J. E. ROMBALDI

Problèmes corrigés d'Analyse Numérique

(Deuxième cycle universitaire, Agrégation)

Février 1996

TABLE DES MATIERES

CHAPITRE 1

Compléments sur le calcul matriciel

Problème 1 : Normes vectorielles et normes matricielles induites

On rappelle que dans un espace métrique une partie K est compacte si et seulement si elle est séquentiellement compacte, c'est-à-dire que de toute suite de points de K on peut extraire une sous suite qui converge vers un élément de K (propriété de Bolzano-Weierstrass).

Soit $(E, \|\ \|)$ *un espace vectoriel normé sur* $\mathbb{R}$ *ou C. On note :*

$$B^1 = \left\{x \in E,\ \|x\| \le 1\right\}$$

sa boule unité et

$$S^1 = \left\{x \in E,\ \|x\| = 1\right\}$$

sa sphère unité.

Q1. *On suppose que E est de dimension finie et on désigne par* $\{e_1, e_2, \cdots, e_n\}$ *une base de E.*

Pour tout $x = \sum_{i=1}^{n} x_i e_i \in E$*, on note* $\|x\|_\infty = Max\left\{|x_i|;\ 1 \le i \le n\right\}$ *et* B_∞^1 *[Resp.* S_∞^1*] la boule [Resp. sphère] unité de E pour* $\|\ \|_\infty$*.*

(a) Montrer que B_∞^1 *et* S_∞^1 *sont compactes dans* $(E, \|\ \|_\infty)$*.*

(b) Montrer que pour toute norme $\|\ \|$ *sur E, l'application* $x \mapsto \|x\|$ *est continue de* $(E, \|\ \|_\infty)$ *dans* $\mathbb{R}$*.*

(c) Montrer que toutes les normes sur E sont équivalentes.

(d) Montrer que B^1 *et* S^1 *sont compactes dans* $(E, \|\ \|)$*, pour toute norme sur E.*

Q2. *E n'est plus nécessairement de dimension finie.*

(a) Montrer que si F est un sous espace vectoriel fermé de E, distinct de E, alors :

$$\forall \varepsilon > 0,\ \exists x \in S^1;\ d(x, F) = Inf\left\{\|x - y\|;\ y \in F\right\} \ge 1 - \varepsilon$$

(b) En déduire que E est de dimension finie si et seulement si sa boule unité B^1 est compacte dans $(E, \| \|)$ (théorème de Riesz).

Soit $x \mapsto \|x\|$ une norme vectorielle sur C^n avec $n \geq 2$. On note B^1 la boule unité et S^1 la sphère unité pour cette norme.
On pose, pour toute matrice $A \in M_n(C)$:

$$\|A\| = Sup\left\{\|Ax\|;\ x \in S^1\right\}$$

Q3. *Montrer qu'on définit ainsi une norme sur $M_n(C)$.*
On dit que $A \mapsto \|A\|$ est la norme matricielle induite par (ou subordonnée à) la norme vectorielle $x \mapsto \|x\|$.
Q4. *Montrer que :*
(a) $\|A\| = Sup\left\{\|Ax\|;\ x \in B^1\right\} = Sup\left\{\dfrac{\|Ax\|}{\|x\|};\ x \in C^n - \{0\}\right\}$
(b) $\forall\, x \in C^n,\ \forall\, A \in M_n(C),\ \|Ax\| \leq \|A\| \cdot \|x\|$
(c) $\forall\, A, B \in M_n(C),\ \|A \cdot B\| \leq \|A\| \cdot \|B\|$
(d) Si I_d est la matrice identité, alors $\|I_d\| = 1$
(e) $\exists x \in C^n;\ \|x\| = 1$ *et* $\|A\| = \|Ax\|$
(f) $\|A\| = Inf\left\{\alpha \in IR;\ \forall x \in C^n,\ \|\mathrm{Ax}\| \leq \alpha\|x\|\right\}$.
Q5. *La « norme de Schur » est définie par :*

$$N_s(A) = \left(Trace\left({}^t\overline{A} \cdot A\right)\right)^{\frac{1}{2}} = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}}$$

Est-elle induite par une norme vectorielle?
Q6. *Calculer la norme matricielle induite par chacune des normes vectorielles suivantes :*
(a) $x \mapsto \|x\|_\infty = Max\left\{|x_i|;\ 1 \leq i \leq n\right\}$.
(b) $x \mapsto \|x\|_1 = \sum_{i=1}^{n} |x_i|$.
(c) $x \mapsto \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$.
Q7. *Calculer $\|A\|_2$ pour toute matrice normale (i. e. telle que $A^*A = AA^*$ où A^* désigne la matrice adjointe de A).*
Q8. *A toute matrice A d'ordre n à coefficients réels et non nulle, on associe l'ensemble :*

$$C(A) = \left\{x \in IR^n;\ \|Ax\|_2 = \|A\|_2 \|x\|_2\right\}$$

(a) Montrer que si A est symétrique positive, alors $C(A)$ est un espace vectoriel non réduit à 0.
(b) Montrer que :

$$C(A) = \left\{x \in IR^n;\ \left\|{}^t AAx\right\|_2 = \|A\|_2{}^2 \|x\|_2\right\}$$

(c) Montrer que $C(A)$ est un espace vectoriel non réduit à 0.

(d) Quelles sont les matrices vérifiant $C(A) = IR^n$?

***Q9.** Montrer les inégalités suivantes pour toute matrice $A \in M_n(C)$:*

(a) $\|A\|_2 \le \sqrt{\|A\|_\infty \|A\|_1}$.

(b) $\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\|A\|_\infty$.

(c) $\frac{1}{\sqrt{n}}\|A\|_1 \le \|A\|_2 \le \sqrt{n}\|A\|_1$.

Solution

Q1. Soit $(x^k)_{k\in IN}$ une suite de points de B_∞^1 avec $x^k = \sum_{i=1}^n x_i^k e_i$, pour tout entier *k*. Pour tout $i \in \{1,2,\cdots,n\}$, on a $|x_i^k| \le \|x^k\|_\infty \le 1$. D'après le théorème de Borel-Lebesgue, on peut extraire de chacune des suites $(x_i^k)_{k\in IN}$ une suite qui converge vers $x_i \in C$, avec $|x_i| \le 1$.

On peut alors extraire de $(x^k)_{k\in IN}$ une suite qui converge vers $x = \sum_{i=1}^n x_i e_i \in B_\infty^1$, ce qui prouve la compacité de B_∞^1.

La sphère unité étant fermée dans B_∞^1, toujours pour la topologie définie par $\|\ \|_\infty$, on en déduit qu'elle est compacte.

(b) Pour tout $x = \sum_{i=1}^n x_i e_i \in E$, on a $\|x\| \le n \cdot Max\{\|e_i\|;\ 1 \le i \le n\}\|x\|_\infty = \beta\|x\|_\infty$.

Quels que soient le réel $\varepsilon > 0$ et *x*, *x'* dans *E* tels que $\|x - x'\|_\infty < \frac{\varepsilon}{\beta}$, on a $\big|\|x\| - \|x'\|\big| \le \|x - x'\| < \varepsilon$. C'est-à-dire que l'application $x \mapsto \|x\|$ est continue de $(E, \|\ \|_\infty)$ dans *IR*.

(c) Il suffit de montrer que $\|\ \|$ et $\|\ \|_\infty$ sont équivalentes.

On a déjà vu qu'il existe une constante réelle $\beta > 0$ telle $\|x\| \le \beta\|x\|_\infty$ pour tout *x* dans *E*.

L'application $x \mapsto \|x\|$ est continue sur le compact S_∞^1 (pour la topologie définie par $\|\ \|_\infty$), elle est donc bornée et atteint ses bornes. En particulier, on peut poser $\alpha = Inf\{\|x\|;\ x \in S_\infty^1\} = \|x_0\| > 0$, avec $x_0 \in S_\infty^1$ (donc $x_0 \ne 0$).

Pour tout *x* non nul dans *E*, on a alors $\left\|\frac{1}{\|x\|_\infty} x\right\| \ge \alpha$ soit $\alpha \cdot \|x\|_\infty \le \|x\|$. Ce qui achève de prouver que les deux normes sont équivalentes.

(d) Soit $\left(x^k\right)_{k\in IN}$ une suite de points de B^1 avec $x^k = \sum_{i=1}^{n} x_i^k e_i$, pour tout entier k. Pour tout $i \in \{1,2,\cdots,n\}$, on a $\left|x_i^k\right| \leq \left\|x^k\right\|_\infty \leq \frac{1}{\alpha}\left\|x^k\right\| \leq \frac{1}{\alpha}$. Donc les suites $\left(x_i^k\right)_{k\in IN}$ sont bornées et on peut extraire de chacune de ces suites une suite convergente.

On peut alors extraire de $\left(x^k\right)_{k\in IN}$ une suite qui converge dans $\left(E,\|\ \|_\infty\right)$ et la convergence a lieu aussi dans $\left(E,\|\ \|\right)$ à cause de l'équivalence des normes. Ce qui prouve la compacité de B^1.

La sphère unité étant fermée dans B^1, pour la topologie définie par $\|\ \|$ (ou par n'importe quelle autre norme, puisqu'on sait maintenant qu'elles sont toutes équivalentes), on en déduit qu'elle est compacte.

Q2. (a) Soit $v \in E - F$, comme F est fermé, on a $\delta = d(v,F) > 0$ et on peut trouver y dans F tel que $0 < \delta \leq \|v-y\| < \frac{\delta}{1-\varepsilon}$, pour tout réel $\varepsilon \in \left]0,1\right[$.

Si on pose $x = \frac{1}{\|v-y\|}(v-y)$, on a $x \in S^1$ et :

$$\forall z \in F,\ \|x-z\| = \left\|\frac{1}{\|v-y\|}(v-y) - z\right\| = \frac{1}{\|v-y\|}\left\|v - \left(y + \|v-y\|z\right)\right\| \geq \frac{1-\varepsilon}{\delta}\delta = 1-\varepsilon$$

car $y + \|v-y\|z \in F$. D'où le résultat.

(b) Si E est de dimension infinie, on peut alors trouver une suite $\left(E_k\right)_{k\in IN}$ de sous espaces vectoriels de dimension finie telle que $E_{k-1} \underset{\neq}{\subset} E_k$, pour tout $k \geq 1$.

Chaque E_k étant fermé, on peut construire, avec le résultat du (a), une suite $\left(x^k\right)_{k\in IN}$ dans S^1 telle que $d\left(x^k, E_{k-1}\right) \geq \frac{1}{2}$ et $x^k \in E_k$.

On a alors :

$$\forall k > l,\ \|x_k - x_l\| \geq \frac{1}{2}$$

et il est impossible d'extraire de $\left(x^k\right)_{k\in IN}$ une sous suite convergente. On déduit donc que ni S^1 ni B^1 ne sont compactes.

Q3. Pour A dans $M_n(C)$, $\|A\|$ est bien définie comme borne supérieure d'une fonction continue sur un compact (la sphère unité est compacte en dimension finie).

L'égalité $\|A\| = 0$ équivaut à $\|Ax\| = 0$ pour tout x dans S^1. En remarquant que pour tout x dans $C^n - \{0\}$, on a $\frac{1}{\|x\|}x \in S^1$, on déduit que $A = 0$.

La vérification des autres propriétés d'une norme ne pose pas de problèmes particuliers.

Q4. (a) Résulte de $S^1 \subset B^1$ et de la linéarité de A.

(b) Résulte de (a).

(c) De (b), on déduit que :

$$\forall A, B \in M_n(C),\ \forall x \in C^n,\ \|ABx\| \le \|A\| \cdot \|Bx\| \le \|A\| \cdot \|B\| \cdot \|x\|$$

d'où le résultat.

(d) Pour tout x dans C^n tel que $\|x\| = 1$, on a $\|I_d x\| = \|x\| = 1$. On déduit donc que $\|I_d\| = 1$.

(e) Résulte du fait qu'en dimension finie la sphère unité est compacte et qu'une fonction continue sur un compact admet une borne supérieure qui est atteinte.

(f) On pose :

$$D = \left\{ \alpha \in IR^+ ;\ \forall x \in C^n,\ \|\mathrm{Ax}\| \le \alpha \|x\| \right\}$$

On a $D \ne \varnothing$, car $\|A\| \in D$. Donc D admet une borne inférieure comme partie non vide et minorée de IR^+. Et on a $Inf(D) \le \|A\|$.

Soient $\alpha \in D$ et $x \in S^1$ tel que $\|A\| = \|Ax\|$. On a alors $\|A\| = \|Ax\| \le \alpha \|x\| = \alpha$, soit $\|A\| \le \alpha$. On en déduit donc que $\|A\| \le Inf(D)$.

Q5. C'est simplement la norme euclidienne de A considérée comme vecteur de C^{n^2}.

Comme $N_s(I_d) = \sqrt{n}$ cette norme n'est pas induite par une norme vectorielle.

Q6. (a) Pour x dans C^n tel que $\|x\|_\infty = 1$, on a :

$$\|Ax\|_\infty = Max\left\{ \sum_{j=1}^{n} \left| a_{ij} x_j \right| ;\ 1 \le i \le n \right\} \le Max\left\{ \sum_{j=1}^{n} \left| a_{ij} \right| ;\ 1 \le i \le n \right\} = \alpha$$

Pour montrer que $\alpha \le \|A\|_\infty$, il suffit de trouver un vecteur x tel que $\|Ax\|_\infty = \alpha$.

Soit k tel que $\alpha = \sum_{j=1}^{n} \left| a_{kj} \right|$. On pose, pour $j = 1, ..., n$:

$$x_j = \begin{cases} \dfrac{\overline{a_{kj}}}{\left| a_{kj} \right|} & si\ a_{kj} \ne 0 \\ 0 & si\ a_{kj} = 0 \end{cases}$$

On a alors $\|Ax\|_\infty = \alpha$ et :

$$\|A\|_\infty = Max\left\{ \sum_{j=1}^{n} \left| a_{ij} \right| ;\ 1 \le i \le n \right\}$$

(b) Pour x dans C^n tel que $\|x\|_1 = 1$, on a :

$$\|Ax\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \sum_{j=1}^{n} \left| x_j \right| \sum_{i=1}^{n} \left| a_{ij} \right| \le Max\left\{ \sum_{i=1}^{n} \left| a_{ij} \right| ;\ 1 \le j \le n \right\} = \alpha$$

Pour montrer que $\alpha \le \|A\|_1$, il suffit de trouver un vecteur x tel que $\|Ax\|_1 = \alpha$.

Soit k tel que $\alpha = \sum_{i=1}^{n} \left| a_{ik} \right|$. Pour $x = e_k$ ($k^{\text{ème}}$ vecteur de base canonique), on a :

$$\|Ax\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| = \sum_{i=1}^{n} \left| a_{ik} \right| = \alpha.$$

D'où $\alpha \le \|A\|_1 \|x\|_1 = \|A\|_1$ et :

$$\|A\|_1 = Max\left\{ \sum_{i=1}^{n} |a_{ij}| \ ; \ 1 \le j \le n \right\}$$

(c) On a $\|A\|_2^{\,2} = Sup\left\{\|Ax\|_2^{\,2};\ \|x\|_2 = 1\right\} = Sup\left\{\left\langle x \,\middle|\, A^* Ax \right\rangle;\ \|x\|_2 = 1\right\}$, où A^* désigne la transposée de A.

La matrice A^*A étant hermitienne positive se diagonalise dans une base orthonormée $\{e_1, e_2, \cdots, e_n\}$ avec des valeurs propres réelles positives $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_n \ge 0$. On a alors pour tout $x = \sum_{i=1}^{n} x_i e_i \in C^n$:

$$\left\langle x \,\middle|\, A^* Ax \right\rangle = \sum_{i=1}^{n} \mu_i |x_i|^2 \le \mu_1 \|x\|_2^{\,2}$$

l'égalité étant réalisée pour $x = e_1$. On a donc :

$$\|A\|_2 = \sqrt{\mu_1} = \sqrt{\rho\left(A^* A\right)}$$

où $\rho(M) = Max\{|\mu| \ ; \ \mu \text{ valeur propre de } M\}$ désigne le rayon spectral de M (voir le problème 2).

Q7. On rappelle qu'une matrice normale se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe une matrice unitaire Q (i. e. telle que $Q^{-1} = Q^*$) et une matrice diagonale D telles que $Q^* AQ = D$.

On a alors $A^*A = QD^*Q^*QDQ^* = QD^*DQ^*$, c'est-à-dire que A^*A et D^*D sont semblables. Elles admettent donc les mêmes valeurs propres et :

$$\|A\|_2 = \sqrt{\rho\left(A^* A\right)} = \sqrt{\rho\left(D^* D\right)} = \rho(D) = \rho(A)$$

Q8. (a) Si A est symétrique réelle, alors c'est une matrice normale et $\|A\|_2 = \rho(A)$.

Pour toute valeur propre λ de A, on note $E_\lambda = \mathrm{Ker}\left(A - \lambda \cdot I_d\right)$ l'espace propre associé.

On désigne par $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$ l'ensemble des valeurs propres de A rangées dans l'ordre décroissant et par $\{e_1, e_2, \cdots, e_n\}$ une base orthonormale de vecteurs propres associés avec $Ae_i = \lambda_i e_i$ pour tout i = 1, 2, ···, n. On a donc $\rho(A) = \lambda_1$.

Pour tout vecteur $x = \sum_{i=1}^{n} x_i e_i \in C(A)$, on a :

$$\|Ax\|_2^{\,2} = \sum_{i=1}^{n} \lambda_i^2 x_i^{\,2} = \|A\|_2^{\,2} \|x\|_2^{\,2} = \lambda_1^{\,2} \sum_{i=1}^{n} x_i^{\,2}$$

On en déduit alors que $\lambda_i = 0$ pour tout indice i tel que $\lambda_i \ne \lambda_1$. C'est-à-dire que $x \in E_{\lambda_1}$.

Réciproquement, pour tout $x \in E_{\lambda_1}$, on a :

$$\|Ax\|_2 = \|\lambda_1 x\|_2 = \lambda_1 \|x\|_2 = \rho(A)\|x\|_2 = \|A\|_2 \|x\|_2$$

c'est-à-dire que $x \in C(A)$.

On a donc ainsi montré que $C(A)$ est l'espace propre associé à une valeur propre dominante de A.

(b) On pose $D(A)=\left\{x\in IR^n;\ \left\|{}^tAAx\right\|_2=\|A\|_2^2\|x\|_2\right\}$.

Pour tout $x\in C(A)$, on a $\left\|{}^tA\cdot Ax\right\|_2\le\left\|{}^tA\right\|_2\|Ax\|_2=\|A\|_2^2\|x\|_2$.

On peut aussi écrire $\|A\|_2^2\|x\|_2^2=\|Ax\|_2^2=\langle Ax\mid Ax\rangle=\left\langle {}^tA\cdot Ax\mid x\right\rangle\le\left\|{}^tA\cdot Ax\right\|_2\|x\|_2$ et on a $\|A\|_2^2\|x\|_2\le\left\|{}^tAAx\right\|_2$. C'est-à-dire que $x\in D(A)$.

Réciproquement, pour tout $x\in D(A)$, on a :

$$\|A\|_2^2\|x\|_2=\left\|{}^tAAx\right\|_2\le\left\|{}^tA\right\|_2\|Ax\|_2=\|A\|_2\|Ax\|_2$$

c'est-à-dire que $\|A\|_2\|x\|_2\le\|Ax\|_2$. Et avec $\|Ax\|_2\le\|A\|_2\|x\|_2$, on déduit que $x\in C(A)$.

On a donc $C(A)=D(A)$.

(c) En Q6, on a montré que pour toute matrice réelle A, on a $\|A\|_2^2=\left\|{}^tA\right\|_2^2=\rho\left({}^tA\cdot A\right)=\left\|{}^tAA\right\|_2$. Ce qui se traduit par $C(A)=C\left({}^tAA\right)$ et $C(A)$ est un espace vectoriel non réduit à $\{0\}$ (d'après (a)).

(d) Si $C(A)=IR^n$, alors $\dfrac{1}{\|A\|_2}A$ est une isométrie (A est supposée non nulle), c'est-à-dire que A est une similitude. La réciproque est évidente.

Q9. (a) Avec les notations de Q6 (c), on a $\|A\|_2=\sqrt{\mu_1}$, avec :

$$\mu_1\|e_1\|_\infty=\|\mu_1e_1\|_\infty=\left\|A^*Ae_1\right\|_\infty\le\left\|A^*A\right\|_\infty\|e_1\|_\infty\le\left\|A^*\right\|_\infty\|A\|_\infty\|e_1\|_\infty$$

et $\left\|A^*\right\|_\infty=\|A\|_1$. On a donc $\|A\|_2^2\le\|A\|_1\|A\|_\infty$.

(b) Pour tout $x\in C^n$, on a :

$$\|Ax\|_2^2=\sum_{i=1}^n\left|\sum_{j=1}^n a_{ij}x_j\right|^2\le\sum_{i=1}^n\left(\sum_{j=1}^n|a_{ij}||x_j|\right)^2\le\sum_{i=1}^n\left(\sum_{j=1}^n|a_{ij}|^2\right)\left(\sum_{j=1}^n|x_j|^2\right)$$

en utilisant l'inégalité de Cauchy-Schwarz. Ce qui donne :

$$\|Ax\|_2^2\le\|x\|_2^2\sum_{i=1}^n\left(\sum_{j=1}^n|a_{ij}|^2\right)\le\|x\|_2^2\,n\,\underset{1\le i\le n}{Max}\left(\sum_{j=1}^n|a_{ij}|^2\right)$$

On en déduit alors que :

$$\|A\|_2^2\le n\,\underset{1\le i\le n}{Max}\left(\sum_{j=1}^n|a_{ij}|^2\right)\le n\left(\underset{1\le i\le n}{Max}\left(\sum_{j=1}^n|a_{ij}|\right)\right)^2=n\|A\|_\infty^2$$

Soit $\|A\|_2\le\sqrt{n}\|A\|_\infty$.

Soit $i\in\{1,2,\cdots,n\}$ tel que $\|A\|_\infty=\sum_{j=1}^n|a_{ij}|$. Avec l'inégalité de Cauchy-Schwarz, on peut écrire que :

$$\|A\|_\infty=\sum_{j=1}^n|a_{ij}|\cdot1\le\sqrt{\sum_{j=1}^n|a_{ij}|^2}\sqrt{n}=\left\|A^*e_i\right\|_2\sqrt{n}\le\left\|A^*\right\|_2\sqrt{n}=\|A\|_2\sqrt{n}$$

C'est-à-dire que $\|A\|_\infty\le\|A\|_2\sqrt{n}$.

(c) L'inégalité résulte de $\|A\|_1 = \|A^*\|_\infty$ et $\|A\|_2 = \|A^*\|_2$.

Problème 2 : Valeurs propres et rayon spectral d'une matrice

On désigne par $M_n(C)$ l'ensemble des matrices d'ordre n à coefficients complexes.

L'ensemble des valeurs propres d'une matrice A dans $M_n(C)$ est appelé le spectre de A et noté Sp(A). C'est une partie finie de C ayant au plus n éléments.

Le rayon spectral de A est alors défini par :

$$\rho(A) = Max\{ |\mu| \; ; \; \mu \in Sp(A) \}$$

Q1. *Montrer que pour toute norme matricielle induite par une norme vectorielle, on a :*

$$\rho(A) \leq \|A\|$$

l'inégalité pouvant être stricte.

Q2. *Pour tout réel $\delta > 0$, on désigne par $D_\delta = ((d_{ij}))_{1 \leq i,j \leq n}$ la matrice diagonale définie par :*

$$\forall i \in \{1,2,\cdots,n\},\; d_{ii} = \delta^{\,i-1}$$

(a) Soit $M = ((m_{ij}))_{1 \leq i,j \leq n}$ dans $M_n(C)$. Calculer $M_\delta = D_\delta^{-1} M D_\delta = ((m_{ij}^{(\delta)}))_{1 \leq i,j \leq n}$.

(b) Montrer que pour tout $\varepsilon > 0$, on peut choisir δ tel que :

$$\sum_{j=i+1}^{n} \left| m_{ij}^{(\delta)} \right| < \varepsilon \;\; (1 \leq i \leq n-1)$$

Q3. *(a) En écrivant que $A' = Q^{-1} A Q$, où Q est la matrice de passage de la base canonique de C^n à une base $\{e_1', e_2', \cdots, e_n'\}$, montrer qu'on définit une norme vectorielle sur C^n en posant :*

$$\forall x = \sum_{i=1}^{n} x_i' e_i' \in C^n,\; \|x\| = Max\{|x_i'|;\; 1 \leq i \leq n\}$$

(b) Montrer que la norme matricielle induite est définie par :

$$\forall B \in M_n(C),\; \|B\| = Max\left\{ \sum_{j=1}^{n} |b_{ij}'| \; ; \; 1 \leq i \leq n \right\}$$

où $B' = Q^{-1} B Q$.

Q4. *En utilisant le fait que A est semblable à une matrice triangulaire, montrer que pour tout $\varepsilon > 0$, il existe une norme matricielle induite par une norme vectorielle telle que :*

$$\|A\| \leq \rho(A) + \varepsilon$$

Q5. *Montrer que* $\rho(A) = Inf\{\|A\|;\; \|\cdot\| \text{ est une norme matricielle induite}\}$

Q6. *Montrer que les quatre conditions suivantes sont équivalentes :*

(i) $\lim\limits_{k \to +\infty} (A^k) = 0$;

(ii) pour toute valeur initiale x_0, la suite définie par $x_{k+1} = A x_k$, pour $k \geq 0$, converge vers le vecteur nul ;

(iii) $\rho(A) < 1$;

(iv) il existe au moins une norme matricielle induite telle que $\|A\| < 1$.

***Q7.** Montrer que pour toute norme matricielle induite par une norme vectorielle, on a :*

$$\rho(A) = \underset{k\to+\infty}{Lim}\left(\left\|A^k\right\|^{\frac{1}{k}}\right) = \underset{k\geq 1}{Inf}\left(\left\|A^k\right\|^{\frac{1}{k}}\right)$$

Solution

Q1. Soit λ une valeur propre de A qui vérifie $\rho(A) = |\lambda|$ et x un vecteur propre associé dans C^n de norme 1. On a alors :

$$\|Ax\| = \|\lambda \cdot x\| = |\lambda| \leq \|A\|$$

d'où $\rho(A) \leq \|A\|$.

En prenant $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, on a $\rho(A) = 0$ et $\|A\| > 0$.

Q2. (a) $m_{ij}^{(\delta)} = \delta^{j-i} \cdot m_{ij}$ $(1 \leq i, j \leq n)$

(la multiplication à droite par D_δ a pour effet de multiplier la colonne numéro j par δ^{j-1} et la multiplication à gauche par D_δ^{-1} a pour effet de diviser la ligne numéro i par δ^{i-1}).

(b) Comme $\underset{\delta\to 0}{Lim}\, m_{ij}^{(\delta)} = 0$ pour $1 \leq i < j \leq n$, on peut choisir δ tel que :

$$\sum_{j=i+1}^{n} \left|m_{ij}^{(\delta)}\right| < \varepsilon \ \ (1 \leq i \leq n-1)$$

Q3. (a) Résulte de $\|x\| = \left\|Q^{-1}x\right\|_\infty$.

(b) On a :

$$\|B\| = Sup\left\{ \frac{\|Bx\|}{\|x\|};\ x \in C^n,\ x \neq 0\right\}$$

$$= Sup\left\{ \frac{\left\|Q^{-1}Bx\right\|_\infty}{\left\|Q^{-1}x\right\|_\infty};\ x \in C^n,\ x \neq 0\right\}$$

Soit en écrivant tout vecteur x sous la forme $x = Qy$:

$$\|B\| = Sup\left\{ \frac{\left\|Q^{-1}BQy\right\|_\infty}{\|y\|_\infty};\ y \in C^n,\ y \neq 0\right\} = \left\|Q^{-1}BQ\right\|_\infty$$

Et il suffit alors d'utiliser le résultat de la question 6 (a) du problème 1.

Q4. Soit P une matrice inversible à coefficients complexes telle que :

$$T = P^{-1}AP$$

avec T triangulaire supérieure.

En prenant, dans la question précédente, $Q = P \cdot D_\delta$ où δ est choisi tel que les coefficients de $T' = Q^{-1}AQ = D_\delta^{-1}TD_\delta$ vérifie $\sum_{j=i+1}^{n}\left|t'_{ij}\right| < \varepsilon$ $(1 \leq i \leq n-1)$, on a :

$$\|A\| = \|T'\|_\infty = Max\left\{ |t_{ii}| + \sum_{j=i+1}^{n} |t'_{ij}| \ ; \ 1 \leq i \leq n \right\} \leq \varepsilon + Max\{|t_{ii}| \ ; \ 1 \leq i \leq n\}$$

soit $\|A\| \leq \rho(A) + \varepsilon$ (les t_{ii} sont les valeurs propres de A).

Q5. Résulte de ce qui précède.

Q6. (i) $\Rightarrow$ (ii). Résulte de : $\|x_k\| = \|A^k x_0\| \leq \|A^k\| \cdot \|x_0\|$.

(ii) $\Rightarrow$ (iii). Soit x_0 est un vecteur propre non nul associé à une valeur propre μ, en écrivant que $x_k = A^k x_0 = \mu^k x_0 \underset{k \to +\infty}{\longrightarrow} 0$, on déduit que $\underset{k \to +\infty}{Lim}\, \mu^k = 0$ et nécessairement $|\mu| < 1$.

(iii) $\Rightarrow$ (iv). Soit $\varepsilon > 0$ tel que $\rho(A) + \varepsilon < 1$. Il suffit de prendre une norme matricielle induite telle que $\|A\| < \rho(A) + \varepsilon$.

(iv) $\Rightarrow$ (i). En prenant une norme matricielle induite qui vérifie $\|A\| < 1$ et en écrivant que $\|A^k\| \leq \|A\|^k$, on déduit que $\underset{k \to +\infty}{Lim}\left(A^k\right) = 0$.

Q7. Soit $\varepsilon > 0$ et $A_\varepsilon = \dfrac{1}{\rho(A) + \varepsilon} A$. On a $\rho(A_\varepsilon) = \dfrac{\rho(A)}{\rho(A) + \varepsilon} < 1$ donc $\underset{k \to +\infty}{Lim}\left(A_\varepsilon{}^k\right) = 0$ et :

$$\exists k_\varepsilon \in IN; \ \forall k \geq k_\varepsilon, \ \left\|A_\varepsilon{}^k\right\| < 1$$

On a alors :

$$\forall k \geq k_\varepsilon, \ \left\|A^k\right\| < \left(\rho(A) + \varepsilon\right)^k$$

Et avec $\rho(A) = \left(\rho\left(A^k\right)\right)^{\frac{1}{k}} \leq \left\|A^k\right\|^{\frac{1}{k}}$, on déduit que

$$\forall k \geq k_\varepsilon, \ \rho(A) \leq \left\|A^k\right\|^{\frac{1}{k}} \leq \rho(A) + \varepsilon$$

c'est-à-dire le résultat.

Problème 3 : Localisation des valeurs propres

Soit A une matrice d'ordre n à coefficients complexes. On note :

$$L_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \ (1 \leq i \leq n); \ C_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \ (1 \leq j \leq n)$$

$$L = \underset{1 \leq i \leq n}{Max}\left\{L_i + |a_{ii}|\right\}; \ C = \underset{1 \leq j \leq n}{Max}\left\{C_j + |a_{jj}|\right\}$$

Q1. *Montrer que pour toute valeur propre de A, $\lambda \in C$, il existe $i \in \{1, 2, \cdots, n\}$ et $j \in \{1, 2, \cdots, n\}$ tels que :*

$$|\lambda - a_{ii}| \leq L_i; \ |\lambda - a_{jj}| \leq C_j$$

Ce résultat est connu sous le nom de « premier théorème de Gerschgörin et Hadamard ».

Q2. *Montrer que pour toute valeur propre de A, $\lambda \in C$, on a :*

$$|\lambda| \leq Min\{L, C\}$$

Q3. *(a) Montrer que s'il existe* $\alpha \in [0,1]$ *tel que :*

$$\forall i \in \{1,2,\cdots,n\},\ |a_{ii}| > L_i^{\alpha} C_i^{1-\alpha}$$

alors A est inversible. On utilisera l'inégalité de Hölder :

$$\left|\sum_{k=1}^{n} a_k \overline{b}_k\right| \leq \left(\sum_{k=1}^{n} |a_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |b_k|^q\right)^{\frac{1}{q}} \quad \left(p>0,\ q>0,\ \frac{1}{p}+\frac{1}{q}=1\right)$$

(b) En déduire que pour tout réel $\alpha \in [0,1]$ *et toute valeur propre de A,* $\lambda \in C$*, il existe* $i \in \{1,2,\cdots,n\}$ *tel que :*

$$|\lambda - a_{ii}| \leq L_i^{\alpha} C_i^{1-\alpha}$$

Ce résultat est le théorème d'Ostrowsky. Pour $\alpha = 1$*, on retrouve le théorème de Gerschgörin et Hadamard.*

Q4. *Montrer que pour toute valeur propre de A,* $\lambda \in C$*, il existe* $i \in \{1,2,\cdots,n\}$ *tel que :*

$$|\lambda|^2 \leq \left(L_i + |a_{ii}|\right)\left(C_i + |a_{ii}|\right)$$

Q5. *Soient a, b dans IR. Calculer les valeurs propres et les vecteurs propres associés de la matrice* $A(a,b) = \begin{pmatrix} a & b & 0 & 0 & \cdots & 0 \\ b & a & b & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b & a & b \\ 0 & \cdots & 0 & 0 & b & a \end{pmatrix}$.

Q6. *Une matrice réelle ou complexe d'ordre* $n \geq 2$*,* $A = \left(\left(a_{ij}\right)\right)_{1\leq i,j\leq n}$*, est à diagonale strictement dominante, si :*

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \quad (i = 1,\cdots,n)$$

(a) Montrer qu'une matrice à diagonale strictement dominante est inversible.
(b) Soit A une matrice symétrique réelle à diagonale strictement dominante. Montrer que A est définie positive si, et seulement si $a_{ii} > 0$*, pour tout* $i \in \{1,\cdots,n\}$.

Solution

Q1. Soit $\lambda \in C$ une valeur propre de A et x un vecteur propre associé avec $\|x\|_\infty = 1$.

Pour $i \in \{1,2,\cdots,n\}$ tel que $|x_i| = \|x\|_\infty$, on a :

$$\sum_{j\neq i} a_{ij} x_j = (\lambda - a_{ii}) x_i$$

et $|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}| = L_i$.

En appliquant ce résultat à la transposée de A qui admet les mêmes valeurs propres que A, on déduit que toute valeur propre de A est dans l'un des disques $\left|\lambda - a_{jj}\right| \leq C_j$.

Q2. Avec les hypothèses et notations de Q1, on peut écrire que $|\lambda| \leq |\lambda - a_{ii}| + |a_{ii}| \leq L_i + |a_{ii}| \leq L$. De manière analogue, on a $|\lambda| \leq C$. On a donc bien $|\lambda| \leq Min\{L, C\}$, pour toute valeur propre de A.

Q3. (a) Si A n'est pas inversible alors 0 est valeur propre et avec le théorème de Gerschgörin et Hadamard on déduit qu'il existe $i \in \{1,2,\cdots,n\}$ et $j \in \{1,2,\cdots,n\}$ tels que $|a_{ii}| \leq L_i$; $\left|a_{jj}\right| \leq C_j$. Ce qui démontre le résultat pour $\alpha = 0$ et $\alpha = 1$.

On suppose maintenant que $\alpha \in \left]0,1\right[$ et que A est non inversible. On désigne par x un vecteur propre non nul associé à la valeur propre 0. C'est à dire que x est solution non nulle du système linéaire :

$$\sum_{j=1}^{n} a_{ij} x_j = 0 \quad (1 \leq i \leq n)$$

On a alors :

$$|a_{ii}||x_i| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j| \quad (1 \leq i \leq n)$$

Avec $L_i^{\alpha} C_i^{1-\alpha} < |a_{ii}|$, pour tout $i \in \{1, 2, \cdots, n\}$, on déduit que :

$$L_i^{\alpha} C_i^{1-\alpha} |x_i| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j| \quad (1 \leq i \leq n)$$

l'inégalité étant stricte pour tous les indices i tels que $x_i \neq 0$.

On utilise alors l'inégalité de Hölder avec $p = \dfrac{1}{\alpha}$ et $q = \dfrac{1}{1-\alpha}$, ce qui donne :

$$L_i^{\alpha} C_i^{1-\alpha} |x_i| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|^{\alpha} \left(|a_{ij}|^{1-\alpha} |x_j|\right) \leq \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|\right)^{\alpha} \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j|^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \quad (1 \leq i \leq n)$$

soit :

$$L_i^{\alpha} C_i^{1-\alpha} |x_i| \leq L_i^{\alpha} \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j|^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \quad (1 \leq i \leq n)$$

Si $x_i \neq 0$, alors l'inégalité est stricte et $L_i > 0$. On déduit donc que :

$$C_i |x_i|^{\frac{1}{1-\alpha}} \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j|^{\frac{1}{1-\alpha}} \quad (1 \leq i \leq n)$$

l'inégalité étant stricte pour $x_i \neq 0$ et évidente pour $x_i = 0$.

En additionnant ces inégalités, on aboutit à :

$$S = \sum_{i=1}^{n} C_i |x_i|^{\frac{1}{1-\alpha}} < \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j|^{\frac{1}{1-\alpha}} = \sum_{j=1}^{n} C_j |x_j|^{\frac{1}{1-\alpha}} = S$$

ce qui est impossible. La matrice A est donc nécessairement inversible.
(b) Si λ est valeur propre de A alors $A-\lambda I_d$ est non inversible et avec le (a), on déduit que pour tout réel $\alpha\in[0,1]$, on peut trouver un indice $i\in\{1,2,\cdots,n\}$ tel que $|a_{ii}-\lambda|\leq L_i^{\alpha}C_i^{1-\alpha}$.

Q4. En prenant $\alpha=\dfrac{1}{2}$ dans Q3 (b), on peut trouver $i\in\{1,2,\cdots,n\}$ tel que $|a_{ii}-\lambda|\leq\sqrt{L_iC_i}$. On déduit alors que $|\lambda|\leq|a_{ii}|+\sqrt{L_iC_i}\leq\sqrt{(|a_{ii}|+L_i)(|a_{ii}|+C_i)}$, la dernière inégalité résultant de $2\sqrt{L_iC_i}\leq C_i+L_i$.

Q5. On peut écrire $A(a,b)=a\cdot I_d+b\cdot A(0,1)$, où I_d désigne la matrice identité. Il suffit donc de considérer le cas $(a,b)=(0,1)$.

La matrice $A(0,1)$ est symétrique réelle, donc toutes ses valeurs propres sont réelles. Avec le théorème de Gerschgörin-Hadamard, on déduit que pour toute valeur propre $\lambda\in IR$ on a $|\lambda|\leq 2$. Une telle valeur propre peut donc s'écrire $\lambda=2\operatorname{Cos}(\alpha)$ avec $\alpha\in[0,\pi]$. Si x est un vecteur propre non nul associé ses composantes sont solutions de la récurrence :

$$x_{k-1}-\lambda x_k+x_{k+1}=0\ (1\leq k\leq n)$$

avec les conditions aux limites $x_0=0$ et $x_{n+1}=0$.

Le polynôme caractéristique de cette récurrence est $P(r)=r^2-2Cos(\alpha)r+1$, soit $P(r)=(r-e^{i\alpha})(r-e^{-i\alpha})$. Les racines sont donc $r_1=e^{i\alpha}$ et $r_2=e^{-i\alpha}$. Ce qui donne $x_k=c_1e^{ik\alpha}+c_2e^{-ik\alpha}\ (0\leq k\leq n+1)$.

Avec $x_0=0$, on déduit que $c_2=-c_1$. Avec $x_{n+1}=0$ et $(c_1,c_2)\neq(0,0)$, on déduit que $\operatorname{Sin}((n+1)\alpha)=0$ et $\alpha=\dfrac{j\pi}{n+1}$ avec $1\leq j\leq n$.

Les valeurs propres de $A(a,b)$ sont donc :

$$\lambda_k(a,b)=a+2\cdot b\cdot\operatorname{Cos}\left(\frac{k\pi}{n+1}\right)\ (1\leq k\leq n)$$

L'espace propre associé à $\lambda_k(a,b)$ est la droite engendrée par le vecteur $v^{(k)}$ de composantes :

$$v_j^{(k)}=\operatorname{Sin}\left(j\frac{k\pi}{n+1}\right)\ (1\leq k,j\leq n)$$

Q6. (a) Avec le théorème de Gershgörin–Hadamard, on déduit que si 0 est valeur propre de A, alors il existe $i\in\{1,\cdots,n\}$ tel que $|a_{ii}|\leq\sum\limits_{j\neq i}|a_{ij}|$, ce qui contredit le fait que A est à diagonale strictement dominante. Donc 0 ne peut pas être valeur propre de A, ce qui équivaut à dire que A est inversible.

(b) Supposons que $a_{ii} > 0$ pour tout $i \in \{1, \cdots, n\}$. Le théorème de Gershgörin-Hadamard nous dit que pour toute valeur propre $\lambda \in IR$ de la matrice symétrique réelle A, il existe $i \in \{1, \cdots, n\}$ tel que :

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| < a_{ii}$$

On en déduit donc que $\lambda > 0$. Les valeurs propres de la matrice A sont donc toutes strictement positives, ce qui équivaut à dire qu'elle est définie positive.

La réciproque provient de l'égalité $a_{ii} = \langle Ae_i | e_i \rangle$, où $\{e_1, \cdots, e_n\}$ est la base canonique de IR^n.

Problème 4 : Le quotient de Rayleigh

Soit A une matrice symétrique réelle. Son quotient de Rayleigh est l'application :

$$R_A : x \in IR^n - \{0\} \mapsto R_A(x) = \frac{q(x)}{\|x\|_2^{\;2}}$$

où $q : x \mapsto q(x) = \langle Ax \,|\, x \rangle$ *est la forme quadratique associée à A.*

Les valeurs propres de A sont rangées dans l'ordre croissant, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ *et on note* $\{e_1, e_2, \cdots, e_n\}$ *une base orthonormée de* IR^n *formée de vecteurs propres associés (*$Ae_k = \lambda_k e_k$*).*

Q1. *Montrer que pour tout k = 1, 2, ···, n :*

$$\lambda_k = R_A(e_k) = \mathrm{Sup}\left\{ R_A(x);\; x = \sum_{j=1}^{k} x_j e_j,\; (x_1, x_2, \cdots, x_k) \in IR^k \right\}$$

$$= \mathrm{Inf}\left\{ R_A(x);\; x = \sum_{j=k}^{n} x_j e_j,\; (x_k, x_{k+1}, \cdots, x_n) \in IR^{n-k+1} \right\}$$

Q2. *Calculer* $\mathrm{Sup}\{R_A(x);\; x \in IR^n\}$ *et* $\mathrm{Inf}\{R_A(x);\; x \in IR^n\}$.

Q3. *Montrer que* $R_A(IR^n) = [\lambda_1, \lambda_n]$.

Q4. *La norme matricielle subordonnée à la norme euclidienne de* IR^n *est définie par :*

$$\forall A \in M_n(IR),\; \|A\|_2 = Sup\{\|Ax\|;\; x \in IR^n,\; \|x\|_2 = 1\}$$

Celle induite par la norme euclidienne de C^n *est définie par :*

$$\forall A \in M_n(IR),\; \|A\|_{2,C} = Sup\{\|Az\|_{2,C};\; z \in C^n,\; \|z\|_{2,C} = 1\}$$

Montrer que ces deux normes sont identiques.

Q5. *Montrer que la norme matricielle subordonnée à la norme euclidienne de* IR^n *est définie par :*

$$\forall A \in M_n(IR),\; \|A\|_2 = \sqrt{\rho({}^tA \cdot A)} = \sqrt{\rho(A \cdot {}^tA)} = \|{}^tA\|_2$$

Q6. *Calculer* $\|A\|_2$ *pour A symétrique réelle.*

Q7. *Calculer* $\|A\|_2$, *pour* $A=\begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ -1 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & -1 & 1 & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$.

Solution

Q1. Pour tout $k = 1, 2, \cdots, n$, on a :

$$R_A(e_k) = \frac{\langle Ae_k | e_k \rangle}{\|e_k\|_2^2} = \lambda_k$$

Pour tout $x = \sum_{j=1}^{k} x_j e_j$, on a :

$$R_A(x) = \frac{\sum_{j=1}^{k} \lambda_j x_j^2}{\sum_{j=1}^{k} x_j^2} \leq \lambda_k \frac{\sum_{j=1}^{k} x_j^2}{\sum_{j=1}^{k} x_j^2} = \lambda_k$$

On en déduit alors que :

$$\lambda_k = \operatorname{Sup}\left\{ R_A(x);\ x = \sum_{j=1}^{k} x_j e_j,\ (x_1, x_2, \cdots, x_k) \in IR^k \right\}$$

De la même façon, on a :

$$\lambda_k = \operatorname{Inf}\left\{ R_A(x);\ x = \sum_{j=k}^{n} x_j e_j,\ (x_k, x_{k+1}, \cdots, x_n) \in IR^{n-k+1} \right\}$$

Q2. En particulier, on a :

$$\lambda_n = R_A(e_n) = \operatorname{Sup}\{R_A(x);\ x \in IR^n\}$$

$$\lambda_1 = R_A(e_1) = \operatorname{Inf}\{R_A(x);\ x \in IR^n\}$$

Q3. On a déjà $R_A(IR^n) \subset [\lambda_1, \lambda_n]$. Et en écrivant tout $\lambda \in [\lambda_1, \lambda_n]$ sous la forme $\lambda = (1-t)\lambda_1 + t\lambda_n$ et en posant $x = \sqrt{1-t}e_1 + \sqrt{t}e_n$, on a $\|x\|_2 = 1$ et $R_A(x) = \lambda$.

Une autre façon de voir les choses est de dire que la sphère unité de IR^n est connexe et son image par une fonction continue est connexe.

Q4. On a $\|z\|_{2,C}^2 = \|x\|_2^2 + \|y\|_2^2$, pour tout $z = x + iy \in C^n$.

Pour tout $x \in IR^n$, on a $\|x\|_{2,C} = \|x\|_2$ et $\|Ax\|_2 = \|Ax\|_{2,C}$, donc $\|A\|_2 \leq \|A\|_{2,C}$. Et pour $z = x + iy \in C^n$, on a :

$\|Az\|_{2,C}^2 = \|Ax\|_2^2 + \|Ay\|_2^2 \leq \|A\|_2^2 \left(\|x\|_2^2 + \|y\|_2^2 \right) = \|A\|_2^2 \|z\|_{2,C}^2$, donc $\|A\|_{2,C} \leq \|A\|_2$.

Q5. La matrice ${}^tA \cdot A$ est symétrique positive, donc son rayon spectral est égal à sa plus grande valeur propre. Avec Q2, on peut alors écrire

$\rho({}^tA\cdot A)=\mathrm{Sup}\left\{R_{{}^tA\cdot A}(x);\ x\in IR^n\right\}$, avec $R_{{}^tA\cdot A}(x)=\dfrac{\left\langle {}^tA\cdot Ax\,\middle|\,x\right\rangle}{\|x\|_2^{\;2}}=\dfrac{\|Ax\|_2^{\;2}}{\|x\|_2^{\;2}}$. C'est-à-dire que $\sqrt{\rho({}^tA\cdot A)}=\|A\|_2$.

Il reste à montrer que $\rho({}^tA\cdot A)=\rho(A\cdot{}^tA)$.

Si $\lambda=\rho({}^tA\cdot A)>0$, il existe $x\in IR^n-\{0\}$ tel que ${}^tA\cdot Ax=\lambda x$. Et nécessairement, on a $Ax\neq 0$.

On alors $A\cdot{}^tA(Ax)=A(\lambda x)=\lambda\cdot Ax$, donc λ est valeur propre de $A\cdot{}^tA$ et $\lambda\leq\rho(A\cdot{}^tA)$. C'est-à-dire que $\rho({}^tA\cdot A)\leq\rho(A\cdot{}^tA)$. On démontre de manière analogue que $\rho(A\cdot{}^tA)\leq\rho({}^tA\cdot A)$.

Si $\rho({}^tA\cdot A)=0$, le raisonnement précédent nous montre alors que $\rho(A\cdot{}^tA)=0$.

Q6. Si A est symétrique réelle, alors les valeurs propres de ${}^tA\cdot A=A^2$ sont les carrés de celles de A et $\sqrt{\rho({}^tA\cdot A)}=\sqrt{\rho(A)^2}=\rho(A)$.

Remarque 1 — Le résultat est encore vrai sur C en remplaçant la transposée par l'adjoint. Soit :

$$\forall A\in M_n(C),\ \|A\|_{2,C}=\sqrt{\rho({}^t\overline{A}\cdot A)}=\sqrt{\rho(A\cdot{}^t\overline{A})}=\left\|{}^t\overline{A}\right\|_{2,C}$$

Si A est normale, c'est-à-dire ${}^t\overline{A}\cdot A=A\cdot{}^t\overline{A}$, alors $\|A\|_{2,C}=\rho(A)$ (voir le problème 1, Q7).

Remarque 2 — Les racines carrées positives des valeurs propres de ${}^tA\cdot A$ sont appelées les *valeurs singulières* de A.

Q7. On a $\|A\|_2=\sqrt{\rho({}^tA\cdot A)}$, avec ${}^tA\cdot A=\begin{pmatrix}2&-1&\cdots&0&-1\\-1&2&-1&\cdots&0\\.&.&.&.&.\\0&\cdots&-1&2&-1\\-1&\cdots&0&-1&2\end{pmatrix}$. La matrice ${}^tA\cdot A$ est symétrique réelle, donc toutes ses valeurs propres sont réelles. De plus, avec le théorème de Gerschgörin-Hadamard (problème 3, Q1), on déduit que pour toute valeur propre $\lambda\in IR$ on a $|\lambda-2|\leq 2$. Une telle valeur propre peut donc s'écrire $\lambda=2(1-\mathrm{Cos}(\alpha))=4\,\mathrm{Sin}(\theta)^2$. Si x est un vecteur propre non nul associé ses composantes sont solutions de la récurrence :

$$x_{k-1}+(\lambda-2)x_k+x_{k+1}=0\ (1\leq k\leq n)$$

avec $\lambda-2=-2\,\mathrm{Cos}(\alpha)$ et les conditions de périodicité $x_0=x_n$ et $x_1=x_{n+1}$.

Il s'agit donc de chercher les solutions $(x_k)_{k\in Z}$ périodiques de période n de cette récurrence.

Le polynôme caractéristique de cette récurrence est $P(r)=r^2-2Cos(\alpha)r+1$, soit $P(r)=(r-e^{i\alpha})(r-e^{-i\alpha})$. Les racines sont donc $r_1=e^{i\alpha}$ et $r_2=e^{-i\alpha}$. Ce qui donne $x_k=c_1e^{ik\alpha}+c_2e^{-ik\alpha}\ (0\leq k\leq n+1)$.

Avec $(c_1,c_2)\neq(0,0)$ et la condition de périodicité on a nécessairement $\alpha = j\dfrac{2\pi}{n}$ avec $0\leq j\leq n-1$.

Les valeurs propres de ${}^tA\cdot A$ sont donc :

$$\lambda_k = 2\left(1-\operatorname{Cos}\left(\frac{2k\pi}{n}\right)\right) = 4\operatorname{Sin}\left(\frac{k\pi}{n}\right)^2 \quad (0\leq k\leq n-1)$$

D'où la norme de A induite par la norme euclidienne :

$$\|A\|_2 = \begin{cases} 2 & \text{si } n=2p \\ 2\,Sin\left(p\dfrac{\pi}{2p+1}\right) & \text{si } n=2p+1 \end{cases}$$

Problème 5 : Matrices tridiagonales. Agrégation 1993, extrait

On désigne par m un réel donné et par T_m la matrice tridiagonale d'ordre n dont les éléments t_{ij} sont définis par :

$$\begin{cases} t_{ii} = m,\ pour\ i=1,2,\cdots,n \\ t_{ij} = -1,\ pour\ i,j=1,2,\cdots,n,\ |i-j|=1 \\ t_{ij} = 0\ dans\ les\ autres\ cas \end{cases}$$

On désigne par A_m la matrice carrée d'ordre n^2 tridiagonale par blocs d'ordre n égale à :

$$A_m = \begin{pmatrix} T_m & -I & 0 & 0 & \cdots & 0 \\ -I & T_m & -I & 0 & \ldots & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & -I & T_m & -I \\ 0 & \cdots & 0 & 0 & -I & T_m \end{pmatrix}$$

où I est la matrice identité d'ordre n.

Q1. *(a) Soit $M = ((m_{ij}))_{1\leq i,j\leq n}$ une matrice carrée complexe d'ordre n, telle que l'on ait :*

$$\forall i\in\{1,\cdots,n\},\ r_i = \sum_{\substack{j=1\\ j\neq i}}^{n} |m_{ij}| < |m_{ii}|$$

Montrer que M est inversible.

(b) En déduire que toute valeur propre de M appartient à la réunion des boules fermées de centre d'affixe m_{ii} et de rayon r_i.

Q2. *(a) Montrer que si λ est valeur propre de T_m, il existe un réel $\theta\in[0,\pi]$, tel que $\lambda = m-2\operatorname{Cos}(\theta)$.*

Déterminer alors les valeurs $\theta_k\in[0,\pi]$ pour lesquelles $\lambda_k = m-2\operatorname{Cos}(\theta_k)$ est effectivement valeur propre de T_m.

(b) Déterminer le vecteur propre V_k associé à λ_k dont la première composante est égale à $Sin(\theta_k)$ et montrer que $\|V_k\|_2^2 = \frac{n+1}{2}$.

En déduire une matrice orthogonale symétrique P, indépendante de m, telle que $P \cdot T_m \cdot P$ soit une matrice diagonale D_m.

(c) Soit O la matrice carrée d'ordre n^2 diagonale par blocs d'ordre n tous égaux à P (et donc orthogonale et symétrique). Calculer le produit $O \cdot A_m \cdot O$.

Q3. *(a) On considère n nombres réels $x_1, x_2, \cdots, x_n$, un vecteur propre V de T_m, puis le vecteur W à n^2 composantes défini par blocs de la façon suivante :*

$$w = \begin{pmatrix} x_1 V \\ x_2 V \\ \vdots \\ x_n V \end{pmatrix}$$

A quelles conditions sur $x_1, x_2, \cdots, x_n$ le vecteur W est-il vecteur propre de A_m?

En déduire les valeurs propres et vecteurs propres de A_m et vérifier que, lorsque m = 4, ces valeurs propres sont les réels :

$$\lambda_{pq} = 4Sin^2\left(\frac{p\pi}{2(n+1)}\right) + 4Sin^2\left(\frac{q\pi}{2(n+1)}\right) \; (1 \le p, q \le n)$$

(b) En déduire une matrice orthogonale symétrique Q, indépendante de m, telle que $Q \cdot A_m \cdot Q$ soit une matrice diagonale.

Solution

Q1. (a) Voir le problème 3, Q6 (a).

(b) C'est le théorème de Gerschgörin-Hadamard (problème 3, Q1).

On peut aussi en donner une démonstration qui utilise le résultat du (a).

Soit M une matrice complexe et λ une valeur propre. La matrice $A - \lambda \cdot I_d$ est donc non inversible et ne peut pas être à diagonale strictement dominante, c'est-à-dire qu'il existe $i \in \{1, 2, \cdots, n\}$ tel que $|\lambda - m_{ii}| \le r_i$.

Q2. (a) Dans le problème 3 Q5, on a vu que les valeurs propres de T_m sont les :

$$\lambda_k = m - 2\,\mathrm{Cos}\left(\frac{k\pi}{n+1}\right) = m - 2\,\mathrm{Cos}(\theta_k) \; (1 \le k \le n)$$

(b) Avec le problème 3 Q5, on a vu qu'un vecteur propre associé à la valeur propre simple λ_k est le vecteur :

$$V_k = \left(\mathrm{Sin}(\theta_k), \quad \mathrm{Sin}(2\theta_k), \quad \cdots, \quad \mathrm{Sin}(n\theta_k)\right) \; (1 \le k \le n)$$

Le carré de la norme euclidienne de ce vecteur est :

$$\|V_k\|_2^2 = \sum_{j=1}^{n} \mathrm{Sin}(j\theta_k)^2 = \frac{1}{2}\sum_{i=1}^{n}\left(1 - \mathrm{Cos}(2j\theta_k)\right)$$

avec :

$$\sum_{i=1}^{n} \text{Cos}(2j\theta_k) = Ré\left(\sum_{i=1}^{n}\left(e^{2i\theta_k}\right)^j\right) = Ré\left(\frac{1-e^{2i(n+1)\theta_k}}{1-e^{2i\theta_k}}-1\right) = -1$$

Ce qui donne :

$$\|V_k\|_2^{\ 2} = \frac{n+1}{2}$$

En posant $e_k = \dfrac{1}{\|V_k\|_2}V_k$ pour $k = 1, 2, \cdots, n$, on définit une base orthonormale de vecteurs propres de T_m (matrice symétrique réelle) et si P est la matrice de passage de la base canonique à cette base alors $\Delta_m={}^tPT_mP$ est diagonale avec les λ_k comme termes diagonaux. En remarquant que pour i, j compris entre 1 et n, on a $p_{ij} = \dfrac{\sqrt{2}}{\sqrt{n+1}}\text{Sin}\left(\dfrac{i\cdot j\cdot \pi}{n+1}\right)$, on déduit que P est symétrique et $\Delta_m = PT_mP$.

(c) On a $OA_mO = \begin{pmatrix} \Delta_m & -I & 0 & 0 & \cdots & 0 \\ -I & \Delta_m & -I & 0 & \ldots & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & -I & \Delta_m & -I \\ 0 & \cdots & 0 & 0 & -I & \Delta_m \end{pmatrix}$.

Q3. (a) Si V un vecteur propre de T_m associé à la valeur propre λ alors W est vecteur propre de A_m associé à la valeur propre μ si :

$$A_mW = \begin{pmatrix} (\lambda x_1 - x_2)V \\ (-x_1+\lambda x_2 - x_3)V \\ \vdots \\ (-x_{n-2}+\lambda x_{n-1}-x_n)V \\ (-x_{n-1}+\lambda x_n)V \end{pmatrix} = \mu\begin{pmatrix} x_1V \\ x_2V \\ \vdots \\ x_{n-1}V \\ x_nV \end{pmatrix}$$

Ce qui équivaut à :

$$\begin{cases} \lambda x_1 - x_2 = \mu x_1 \\ -x_{k-1}+\lambda x_k - x_{k+1} = \mu x_k \quad (2 \le k \le n-1) \\ -x_{n-1}+\lambda x_n = \mu x_n \end{cases}$$

C'est à dire que le vecteur $X = (x_1, x_2, \cdots, x_n)$ est vecteur propre de T_λ associé à la valeur propre μ.

Si $\lambda = \lambda_p = m - 2\,\text{Cos}(\theta_p)$, avec $\theta_p = \dfrac{p\pi}{n+1}$ et $1 \le p \le n$, alors on peut prendre :

$$V = V_p = \left(\text{Sin}(\theta_p),\quad \text{Sin}(2\theta_p),\quad \cdots,\quad \text{Sin}(n\theta_p)\right)$$

avec $\theta_q = \dfrac{q\pi}{n+1}$, $1 \le q \le n$, et :

$$\mu = \lambda_{p,q} = \lambda_p - 2\,\text{Cos}(\theta_q) = m - 2\,\text{Cos}(\theta_p) - 2\,\text{Cos}(\theta_q)$$

$$X = X_q = \left(\mathrm{Sin}\left(\theta_q\right),\quad \mathrm{Sin}\left(2\theta_q\right),\quad \cdots,\quad \mathrm{Sin}\left(n\theta_q\right)\right)$$

On a donc montré que les valeurs propres de A_m sont les $\lambda_{p,q}$, avec $1 \le p,q \le n$ et les vecteurs propres associés sont les :

$$V_{pq} = \begin{pmatrix} \mathrm{Sin}\left(\theta_q\right)V_p \\ \mathrm{Sin}\left(2\theta_q\right)V_p \\ \vdots \\ \mathrm{Sin}\left(n\theta_q\right)V_p \end{pmatrix}$$

De plus, on a :

$$\left\|V_{pq}\right\|^2 = \sum_{j=1}^{n} \mathrm{Sin}\left(j\theta_q\right)^2 \left\|V_p\right\|_2^{\,2} = \left\|V_q\right\|_2^{\,2}\left\|V_p\right\|_2^{\,2} = \frac{(n+1)^2}{4}$$

Dans le cas particulier $m = 4$, on a :

$$\lambda_{p,q} = 2\left(1-\mathrm{Cos}\left(\theta_p\right)\right) + 2\left(1-\mathrm{Cos}\left(\theta_q\right)\right) = 4Sin^2\left(\frac{p\pi}{2(n+1)}\right) + 4Sin^2\left(\frac{q\pi}{2(n+1)}\right)$$

(b) En posant $e_{pq} = \dfrac{1}{\left\|V_{pq}\right\|_2} V_{pq}$ pour *p, q* = 1, 2, ···, *n*, on définit une base orthonormale de vecteurs propres de A_m (matrice symétrique réelle) et si *Q* est la matrice de passage de la base canonique à cette base alors $D_m={}^tQA_mQ$ est diagonale avec les λ_{pq} comme termes diagonaux.

En remarquant que la matrice *Q* s'écrit par blocs d'ordre *n*, $Q_{ij} = \dfrac{2}{n+1}\mathrm{Sin}\left(\dfrac{i\cdot j\cdot\pi}{n+1}\right)P$ pour *i*, *j* compris entre 1 et *n*, avec *P* définie en Q2 (b), on en déduit que *Q* est symétrique et $D_m = QA_mQ$.

Problème 6 : Conditionnement d'un système linéaire

Ce problème utilise les résultats du problème 1.

Soit A une matrice réelle ou complexe, inversible et x solution du système $Ax = b$.

On désigne par $\|\cdot\|$ *une norme matricielle induite par une norme vectorielle.*

Q1. *En notant* $x+\delta x$ *la solution du système perturbé* $Ay = b + \delta b$, *montrer que :*

$$\frac{\|\delta x\|}{\|x\|} \le \|A\|\cdot\left\|A^{-1}\right\|\cdot\frac{\|\delta b\|}{\|b\|}$$

Q2. *De même, si* $x+\delta x$ *est solution de* $(A+\delta A)y = b$, *montrer que :*

$$\frac{\|\delta x\|}{\|x+\delta x\|} \le \|A\|\cdot\left\|A^{-1}\right\|\cdot\frac{\|\delta A\|}{\|A\|}$$

Le conditionnement de A relativement à la norme $\|\cdot\|$ *est défini par :*

$$Cond(A) = \frac{1}{\|A\|\cdot\left\|A^{-1}\right\|}$$

Un système est bien conditionné si Cond(A) est voisin de 1 et mal conditionné si ce conditionnement est proche de 0.

On note $Cond_2$ le conditionnement relatif à la norme euclidienne sur $I\!K^n$, avec $I\!K = I\!R$ ou $I\!K = C$.

***Q3.** Montrer que :*

(a) $Cond(A) \in \left]0,1\right]$ pour toute norme matricielle, avec $Cond_2(A) = 1$, pour A unitaire (i. e. ${}^t\overline{A} \cdot A = I_d$).

(b) $Cond(A) = Cond(A^{-1})$.

(c) $\forall \alpha \in IK - \{0\}$, $Cond(\alpha A) = Cond(A)$.

(d) $Cond_2(A) = \sqrt{\dfrac{\mu_{\min}}{\mu_{\max}}}$, où $\mu_{\min}$ [Resp. $\mu_{\max}$] est la plus petite [Resp. plus grande] valeur propre de ${}^t\overline{A} \cdot A$.

(e) Pour A hermitienne complexe (ou symétrique réelle :

$$Cond_2(A) = \frac{Min\{|\lambda_i|;\ i = 1,2,\cdots,n\}}{Max\{|\lambda_i|;\ i = 1,2,\cdots,n\}}$$

où les λ_i sont les valeurs propres de A.

***Q4.** Soit* $A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$ *et* $b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$.

(a) Résoudre les systèmes $Ax = b$, $Ax = b + \delta b$ et $(A + \delta A)x = b$, avec

$$b + \delta b = \begin{pmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{pmatrix} \text{ et } A + \delta A = \begin{pmatrix} 10 & 7 & 8.1 & 7.2 \\ 7.08 & 5.04 & 6 & 5 \\ 8 & 5.98 & 9.81 & 9 \\ 6.99 & 4.99 & 9 & 9.98 \end{pmatrix}$$

(b) Calculer $Cond_2(A)$.

Solution

Q1. On a, $A(x + \delta x) = b + \delta b$ et $Ax = b$. On en déduit donc que $\delta x = A^{-1}\delta b$ et $\|\delta x\| \le \|A^{-1}\| \cdot \|\delta b\|$. De même $b = Ax$ donne $\|b\| \le \|A\| \cdot \|x\|$. D'où le résultat en faisant le produit de ces deux inégalités.

Q2. On a, $(A + \delta A)(x + \delta x) = b$ et $Ax = b$. On en déduit donc que $\delta x = -A^{-1}\delta A(x + \delta x)$ et $\|\delta x\| \le \|A^{-1}\| \cdot \|\delta A\| \cdot \|x + \delta x\|$.

Remarque — Il faut bien noter que le conditionnement n'est défini que pour une matrice inversible et dépend du choix d'une norme matricielle.

Q3. (a) En écrivant que $\|A \cdot A^{-1}\| = 1$, on déduit que $1 \le \|A\| \cdot \|A^{-1}\|$. C'est-à-dire que le conditionnement est un réel non nul compris entre 0 et 1.

Si A est unitaire, en utilisant la norme euclidienne, on a (problème 1, Q6 (c)) $\|A\|_{2,C} = \sqrt{\rho({}^t\overline{A} \cdot A)} = \sqrt{\rho(I_d)} = 1$ et $\|A\|_{2,C} = \|{}^tA\|_{2,C}$. Donc $\mathrm{Cond}_2(A) = 1$.
(b) et (c) sont évidents.
(d) La matrice ${}^t\overline{A} \cdot A$ est symétrique définie positive, donc :

$$\|A\|_{2,C}^{\ 2} = \rho({}^t\overline{A} \cdot A) = \mu_{\max}$$

et

$$\|A^{-1}\|_{2,C}^{\ 2} = \rho({}^t\overline{A}^{-1} \cdot A^{-1}) = \rho(A^{-1} \cdot {}^t\overline{A}^{-1}) = \rho\left(({}^t\overline{A} \cdot A)^{-1}\right) = \frac{1}{\mu_{\min}}$$

(e) Si A est hermitienne, alors ses valeurs propres sont réelles et celles de ${}^t\overline{A} \cdot A = A^2$ sont les $\lambda_i^{\ 2}$.

On peut aussi remarquer que dans ce cas on a $\|A\|_{2,C} = \rho(A)$.

Q4. (a) La solution de $Ax = b$ est $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

Si on modifie le second membre en $b + \delta b$, c'est-à-dire que b est donné avec une erreur relative de 0.3%, alors la solution devient $x = \begin{pmatrix} 9.2 \\ -12.6 \\ 4.5 \\ -1.1 \end{pmatrix}$. On a donc une erreur relative sur x de l'ordre de 1000%.

De même si on perturbe la matrice en prenant $A + \delta A$ et en gardant le second membre initial, la solution devient $x = \begin{pmatrix} -8.1 \\ 137 \\ -34 \\ 22 \end{pmatrix}$. Un tel système est donc mal conditionné.
(b) Les valeurs propres de A sont 0.01, 0.84, 3.86 et 30.3. Le conditionnement, relativement à la norme euclidienne, est alors $Cond_2(A) \cong \dfrac{0.01}{30.3} \cong 3.3 * 10^{-4}$.

Remarque — Le conditionnement d'une matrice n'est pas lié à un déterminant voisin de 0. Pour l'exemple ci-dessus, la matrice est symétrique définie positive de déterminant égal à 1.

Problème 7 : Conditionnement du problème de valeurs propres pour les matrices symétriques réelles

On utilise notations et les résultats du problème 4.

Soit A une matrice symétrique réelle de valeurs propres, rangées dans l'ordre croissant, $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. *On désigne par* R_A *son quotient de Rayleigh.*

Pour tout k = 1, 2, ⋯, n, on désigne par E_k *l'ensemble des sous espaces vectoriels de dimension k de* IR^n *et on note* $E_0 = \{\{0\}\}$.

Q1. *Montrer que, pour tout k = 1, 2, ⋯, n, on a :*

$$\lambda_k = \operatorname{Inf}\{\mu(V);\ V \in E_k\} \text{ avec } \mu(V) = \operatorname{Sup}\{R_A(x);\ x \in V\}$$

(théorème de Courant-Fischer).

Q2. *Soit* $A: t \in [a,b] \mapsto A(t) \in S_n(IR)$ *une application continue où* $S_n(IR)$ *désigne l'ensemble des matrices symétriques réelles.*

Si, pour tout $t \in [a,b]$, *on note* $\lambda_1(t) \le \lambda_2(t) \le \cdots \le \lambda_n(t)$ *les valeurs propres de* $A(t)$ *rangées dans l'ordre croissant, montrer alors que les fonctions* λ_k *sont continues de* $[a,b]$ *dans IR.*

Q3. *L'exemple qui suit nous montre que les vecteurs propres de* $A(t)$ *ne définissent pas nécessairement des fonctions continus.*

Soit A: $[0,1] \to S_2(IR)$ *définie par :*

$$\forall t \in [0,1],\ A(t) = \begin{pmatrix} a(t) & b(t) \\ b(t) & -a(t) \end{pmatrix}$$

avec :

$$a(t) = \begin{cases} 0 \text{ si } t = 0 \\ \mathrm{e}^{-\frac{1}{t^2}} \operatorname{Cos}\left(\frac{1}{t}\right) \text{ si } t \ne 0 \end{cases};\quad b(t) = \begin{cases} 0 \text{ si } t = 0 \\ \mathrm{e}^{-\frac{1}{t^2}} \operatorname{Sin}\left(\frac{1}{t}\right) \text{ si } t \ne 0 \end{cases}$$

(a) Montrer que A est continue et, pour tout t dans $[0,1]$, *calculer les valeurs et vecteurs propres de* $A(t)$.

(b) Montrer que vecteurs propres de $A(t)$ *ne définissent pas des fonctions continues en 0.*

Q4. *On suppose que* $A(t) = A + tB$, *pour tout* $t \in [a,b]$, *où A et B sont des matrices symétriques réelles. Montrer alors que la fonction* λ_1 *(plus petite valeur propre) est concave.*

Solution

Q1. On a vu que $\lambda_k = \operatorname{Sup}\{R_A(x);\ x \in V_k\}$ où V_k est le sous espace vectoriel de dimension k engendré par une famille orthonormale de vecteurs propres associés à $\{\lambda_1, \lambda_2, \cdots, \lambda_k\}$ (problème 4, Q1).

On déduit donc que $\lambda_k \ge \alpha_k = \operatorname{Inf}\{\mu(V);\ V \in E_k\}$.

Soit $V \in E_k$, on a :

$$\mathrm{Dim}\left(V \cap V_{k-1}{}^{\perp}\right) = \mathrm{Dim}(V) + \mathrm{Dim}\left(V_{k-1}{}^{\perp}\right) - \mathrm{Dim}\left(V + V_{k-1}{}^{\perp}\right)$$
$$= n + 1 - \mathrm{Dim}\left(V + V_{k-1}{}^{\perp}\right) \geq 1$$

Donc $V \cap V_{k-1}{}^{\perp} \neq \{0\}$.

Soit $x \in V \cap V_{k-1}{}^{\perp} - \{0\}$, on a alors :

$$\lambda_k = \mathrm{Inf}\left\{R_A(y);\ y \in V_{k-1}{}^{\perp}\right\} \leq R_A(x) \leq \mathrm{Sup}\left\{R_A(y);\ y \in V\right\}$$

c'est-à-dire que :

$$\forall V \in E_k,\ \lambda_k \leq \mathrm{Sup}\left\{R_A(y);\ y \in V\right\} = \mu(V)$$

soit $\lambda_k \leq \alpha_k$.

Q2. Pour $t \in [a,b]$, soit $\{e_1(t), e_2(t), \cdots, e_n(t)\}$ une base orthonormale de vecteurs propres de $A(t)$, avec $A(t)e_k(t) = \lambda_k(t)e_k(t)$. On note $V_k(t)$ l'espace vectoriel engendré par $\{e_1(t), \cdots, e_k(t)\}$ et E_k l'ensemble des sous espaces vectoriels de dimension k de $I\!R^n$ avec $E_0 = \{\{0\}\}$.

Avec Q1, on déduit que $\lambda_k(t) \leq \mathrm{Sup}\left\{R_{A(t)}(x);\ x \in V_k(t_0)\right\}$.

Avec $R_{A(t)}(x) = R_{A(t_0)}(x) + R_{A(t)-A(t_0)}(x)$, on a alors :

$$\lambda_k(t) \leq \mathrm{Sup}\left\{R_{A(t_0)}(x) + R_{A(t)-A(t_0)}(x);\ x \in V_k(t_0)\right\}$$
$$\leq \mathrm{Sup}\left\{R_{A(t_0)}(x);\ x \in V_k(t_0)\right\} + \mathrm{Sup}\left\{R_{A(t)-A(t_0)}(x);\ x \in V_k(t_0)\right\}$$

Mais on a aussi $\mathrm{Sup}\left\{R_{A(t_0)}(x);\ x \in V_k(t_0)\right\} = \lambda_k(t_0)$, donc :

$$\lambda_k(t) \leq \lambda_k(t_0) + \mathrm{Sup}\left\{R_{A(t)-A(t_0)}(x);\ x \in V_k(t_0)\right\}$$

avec :

$$\mathrm{Sup}\left\{R_{A(t)-A(t_0)}(x);\ x \in V_k(t_0)\right\} \leq \mathrm{Sup}\left\{R_{A(t)-A(t_0)}(x);\ x \in I\!R^n\right\}$$
$$= \mathrm{Sup}\left\{\frac{\left\langle (A(t) - A(t_0))x \middle| x \right\rangle}{\|x\|_2{}^2};\ x \in I\!R^n - \{0\}\right\} \leq \left\|A(t) - A(t_0)\right\|_2$$

On a donc montré que $\lambda_k(t) \leq \lambda_k(t_0) + \left\|A(t) - A(t_0)\right\|_2$.

En permutant les rôles de t et t_0, on déduit que :

$$\left|\lambda_k(t) - \lambda_k(t_0)\right| \leq \left\|A(t) - A(t_0)\right\|_2$$

ce qui suffit à prouver la continuité de λ_k.

Remarque — Ce résultat peut s'interpréter en disant que de petites perturbations sur les coefficients d'une matrice symétrique réelle n'engendreront que de petites perturbations sur les valeurs propres. Ce qui revient à dire que le problème de valeurs propres est bien conditionné dans ce cas.

On a un résultat analogue pour les matrices complexes hermitiennes.

Q3. (a) A est continûment dérivable et, pour tout t dans $[0,1]$, les valeurs propres de $A(t)$ sont :

$$\forall t \in [0,1],\ \lambda_1(t) = \begin{cases} 0 \text{ si } t = 0 \\ \mathrm{e}^{-\frac{1}{t^2}} \text{ si } t \neq 0 \end{cases},\ \lambda_2(t) = -\lambda_1(t)$$

Ce sont des fonctions continûment dérivables sur $[0,1]$.

Les vecteurs propres associés sont :

$$\forall t \in]0,1],\ e_1(t) = \begin{pmatrix} \mathrm{Cos}\left(\dfrac{1}{2t}\right) \\ \mathrm{Sin}\left(\dfrac{1}{2t}\right) \end{pmatrix},\ e_2(t) = \begin{pmatrix} \mathrm{Sin}\left(\dfrac{1}{2t}\right) \\ -\mathrm{Cos}\left(\dfrac{1}{2t}\right) \end{pmatrix}$$

$$e_1(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ e_2(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(b) Il est clair que les fonctions vecteurs propres ne sont pas continues en 0.

Q4. Pour tout $t \in [a,b]$, on désigne par $\{e_1(t), e_2(t), \cdots, e_n(t)\}$ une base orthonormée de vecteurs propres associée aux $\lambda_1(t), \lambda_2(t), \cdots, \lambda_n(t)$ avec $A(t)e_k(t) = \lambda_k(t)e_k(t)$ pour tout k = 1, 2, ···, n.

Pour tout vecteur $x = \sum_{j=1}^{n} x_j e_j(t)$ de norme 1 et tout $t \in [a,b]$, on a :

$$\langle A(t)x \mid x \rangle = \sum_{j=1}^{n} \lambda_j(t) x_j^{\,2} \geq \lambda_1(t) \sum_{j=1}^{n} x_j^{\,2} = \lambda_1(t)$$

C'est-à-dire que :

$$\forall t \in [a,b],\ \langle Ax \mid x \rangle + t \langle Bx \mid x \rangle \geq \lambda_1(t)$$

En prenant $x = e_1(t_0)$ pour $t_0 \in [a,b]$, on a :

$$\begin{cases} \lambda_1(t_0) = \langle Ae_1(t_0) \mid e_1(t_0) \rangle + t_0 \langle Be_1(t_0) \mid e_1(t_0) \rangle \\ \forall t \in [a,b],\ \lambda_1(t) \leq \langle Ae_1(t_0) \mid e_1(t_0) \rangle + t \langle Be_1(t_0) \mid e_1(t_0) \rangle \end{cases}$$

Ce qui peut s'écrire, en posant $b_0 = \langle Be_1(t_0) \mid e_1(t_0) \rangle$:

$$\forall t \in [a,b],\ \lambda_1(t) \leq \lambda_1(t_0) + b_0(t - t_0)$$

Pour u, v dans $[a,b]$ et $\theta \in [0,1]$, on a alors :

$$\begin{cases} \lambda_1(u) \leq \lambda_1(u + \theta(v-u)) - \theta(v-u) b_{\theta,u,v} \\ \lambda_1(v) \leq \lambda_1(u + \theta(v-u)) + (1-\theta)(v-u) b_{\theta,u,v} \end{cases}$$

et :

$$(1-\theta)\lambda_1(u) + \theta\lambda_1(v) \leq \lambda_1(u + \theta(v-u))$$

C'est-à-dire que la fonction λ_1 est concave.

Problème 8 : Dérivabilité des valeurs propres d'une fonction matricielle en dimension 2

Q1. *Soit* φ: $[a,b] \to IR$ *une fonction continûment dérivable et à valeurs positives ou nulles.*

Montrer que :

$$\forall t \in]a,b[, \ (\varphi(\mathrm{t}) = 0 \Rightarrow \varphi'(\mathrm{t}) = 0)$$

Dans ce qui suit, on se donne f, g: $[a,b] \to IR$ *deux fonctions continûment dérivables et on note* $\varphi = f^2 + \mathrm{g}^2$.

(b) Montrer que si $\varphi(t) = 0$ *sur un intervalle* $]\alpha,\beta[\subset]a,b[$*, alors on a :*
$f(t) = \mathrm{g}(t) = f'(t) = \mathrm{g}'(t) = \varphi'(t) = 0$ *sur* $[\alpha,\beta]$.

(c) Montrer que si $\varphi(t) > 0$ *sur un intervalle* $]\alpha,\beta[\subset]a,b[$ *avec* $\varphi(\alpha) = \varphi(\beta) = 0$*, alors* $h = \sqrt{\varphi}$ *est continûment dérivable sur* $[\alpha,\beta]$ *avec :*

$$h'(t) = \begin{cases} \dfrac{f'(t)f(t) + \mathrm{g}'(t)\mathrm{g}(t)}{\sqrt{f(t)^2 + \mathrm{g}(t)^2}} & \text{si } t \neq \alpha,\ t \neq \beta \\ \sqrt{f'(t)^2 + \mathrm{g}'(t)^2} & \text{si } t = \alpha \\ -\sqrt{f'(t)^2 + \mathrm{g}'(t)^2} & \text{si } t = \beta \end{cases}$$

(d) On suppose qu'il existe deux intervalles $]a_0,b_0[$ *et* $]a_1,b_1[$ *dans* $]a,b[$ *avec* $a_1 = b_0$*,* $\varphi(t) > 0$ *sur* $]a_0,b_0[\cup]a_1,b_1[$ *et* $\varphi(a_0) = \varphi(b_0) = \varphi(b_1) = 0$.

Montrer que la fonction h définie sur $[a_0,b_1]$ *par :*

$$h(t) = \begin{cases} \sqrt{\varphi(t)} & \text{si } t \in [a_0,b_0] \\ -\sqrt{\varphi(t)} & \text{si } t \in [a_1,b_1] \end{cases}$$

est continûment dérivable.

(e) On suppose qu'il existe deux intervalles $]a_0,b_0[$ *et* $]a_1,b_1[$ *dans* $]a,b[$ *avec* $a_1 > b_0$*,* $\varphi(t) > 0$ *sur* $]a_0,b_0[\cup]a_1,b_1[$ *et* $\varphi(t) = 0$ *sur* $[b_0,a_1]$.

Montrer que la fonction h définie sur $[a_0,b_1]$ *par :*

$$h(t) = \begin{cases} \sqrt{\varphi(t)} & \text{si } t \in [a_0,b_0] \\ 0 & \text{si } t \in [b_0,a_1] \\ -\sqrt{\varphi(t)} & \text{si } t \in [a_1,b_1] \end{cases}$$

est continûment dérivable.

(f) Déduire de ce qui précède que si f, g: $[a,b] \to IR$ *sont deux fonctions continûment dérivables alors il existe une fonction* h: $[a,b] \to IR$ *continûment dérivable telle que* $h^2 = f^2 + g^2$.

(g) Plus généralement, montrer que si $v\colon [a,b] \to IR^n$ *est une fonction continûment dérivable alors il existe une fonction* $h\colon [a,b] \to IR$ *continûment dérivable telle que* $h(t)^2 = \|v(t)\|^2$ *sur* $[a,b]$ *où* $\|\cdot\|$ *désigne la norme euclidienne de* IR^n.

(h) On note $S_n(IR)$ *l'ensemble des matrices à coefficients réels, d'ordre n et symétriques.*

Soit $A\colon [a,b] \to S_2(IR)$ *continûment dérivable.*

Montrer que, pour tout t dans $[a,b]$*, le polynôme caractéristique de* $A(t)$ *peut s'écrire :*

$$P(\lambda, t) = (\lambda - \lambda_1(t))(\lambda - \lambda_2(t))$$

où λ_1, $\lambda_2\colon [a,b] \to IR$ *sont continûment dérivables.*

Q2. *Soit* $A\colon [0,1] \to M_2(IR)$ *définie par :*

$$\forall t \in [0,1],\ A(t) = \begin{pmatrix} 0 & b(t) \\ t & 0 \end{pmatrix} \text{ avec } b(t) = \begin{cases} 0 \text{ si } t = 0 \\ t^p \operatorname{Sin}\left(\dfrac{1}{t}\right)^2 \text{ si } t \neq 0 \end{cases}$$

où $2 < p \leq 4$.

(a) Montrer que A est continûment dérivable.

(b) Calculer les valeurs propres de $A(t)$ *et montrer que* $A(t)$ *est diagonalisable pour tout t dans* $[0,1]$.

(c) Montrer que le résultat de Q1 (h) est faux pour cette fonction matricielle.

Solution

Q1. (a) Si, pour $t_0 \in]a,b[$, on a $\varphi'(t_0) \neq 0$, alors la fonction φ est strictement monotone, positive ou nulle sur un intervalle centré en t_0 contenu dans $]a,b[$ et nécessairement $\varphi(t_0) > 0$.

(b) Si, pour $t_0 \in]a,b[$, on a $f'(t_0) \neq 0$, alors la fonction *f* est strictement monotone et positive ou nulle sur un intervalle centré en t_0 et contenu dans $]a,b[$. Nécessairement, on a $f(t_0) \neq 0$ et $\varphi(t_0) > 0$. De même pour *g*.

(c) Pour tout $t \in]\alpha, \beta[$, on a:

$$\frac{h(t) - h(\alpha)}{t - \alpha} = \sqrt{\left(\frac{f(t) - f(\alpha)}{t - \alpha}\right)^2 + \left(\frac{g(t) - f(\alpha)}{t - \alpha}\right)^2} \underset{t \to \alpha}{\to} \sqrt{f'(\alpha)^2 + g'(\alpha)^2}$$

c'est-à-dire que *h* est dérivable à droite en α, avec :

$$h'(\alpha) = \sqrt{f'(\alpha)^2 + g'(\alpha)^2}$$

De la même façon, on voit que *h* est dérivable à gauche en β avec :

$$h'(\beta) = -\sqrt{f'(\beta)^2 + g'(\beta)^2}$$

Dans le cas où $\left(f'(\alpha),g'(\alpha)\right)\neq(0,0)$, la continuité à droite en α de h' résulte du théorème des accroissements finis. Il suffit en effet d'écrire, pour tout $t\in\left]\alpha,\beta\right[$: $h'(t)=\dfrac{f'(t)f'(c_t)+g'(t)g'(c_t)}{\sqrt{f'(c_t)^2+g'(c_t)^2}}$ avec $c_t\in\left]\alpha,t\right[$.

Si $\left(f'(\alpha),g'(\alpha)\right)=(0,0)$, en utilisant l'inégalité de Cauchy-Schwarz, on a :

$$|h'(t)|\leq\sqrt{f'(t)^2+g'(t)^2}\underset{t\to\alpha}{\longrightarrow}0$$

On procède de façon analogue en β.

(d) C'est une conséquence immédiate du (c).

(e) C'est encore une conséquence immédiate du (c).

(f) On note $D=\varphi^{-1}(0)$ et $V=[a,b]-D$.

Si $D=\varnothing$, il suffit alors de poser $h=\sqrt{\varphi}$.

Dans le cas contraire, V est un ouvert de $[a,b]$ et il peut alors s'écrire comme une réunion dénombrable d'intervalles ouverts dans $[a,b]$ deux à deux disjoints, soit $V=\bigcup\limits_{i\in I}V_i$ avec $I=\{0,1,\cdots,n\}$ ou $I=\mathbb{N}$.

On peut en outre supposer les extrémités a_i et b_i de V_i vérifient :

$$a\leq a_i<b_i\leq a_{i+1}<b_{i+1}\leq b$$

Les résultats précédents nous permettent alors de poser :

$$\forall t\in[a,b],\ h(t)=\begin{cases}(-1)^i\sqrt{\varphi(t)} & \text{si } t\in V_i\\ 0 & \text{si } t\in D\end{cases}$$

(g) On raisonne par récurrence sur n en écrivant :

$$v_1^{\,2}+\cdots+v_n^{\,2}+v_{n+1}^{\,2}=f^2+v_{n+1}^{\,2}$$

(h) On note :

$$\forall t\in[0,1],\ A(t)=\begin{pmatrix}a(t)&b(t)\\ b(t)&c(t)\end{pmatrix}$$

Le polynôme caractéristique de $A(t)$ s'écrit alors :

$$P(\lambda,t)=\lambda^2-\left(a(t)+c(t)\right)\lambda+a(t)c(t)-b(t)^2$$

Son discriminant est $\Delta(t)=\left(a(t)-c(t)\right)^2+4b(t)^2$.

On peut donc trouver une fonction h de classe C^1 sur $[a,b]$ telle $\Delta=h^2$. Et les valeurs propres de $A(t)$ sont les fonctions de classe C^1:

$$\begin{cases}\lambda_1(t)=\dfrac{a(t)+c(t)+h(t)}{2}\\[2ex] \lambda_2(t)=\dfrac{a(t)+c(t)-h(t)}{2}\end{cases}$$

Q2. (a) Il suffit de montrer que b est de classe C^1. Sur $\mathbb{R}^*$ il n'y a pas de problème. La continuité en 0 résulte de $|b(t)|\leq|t|^p$ avec $p>0$. La Dérivabilité en 0 résulte de

$\left|\dfrac{b(t)}{t}\right| \le |t|^{p-1}$ avec $p > 1$. Enfin la continuité de la dérivée en 0 résulte de

$|b'(t)| = \left|pt^{p-1}\operatorname{Sin}\left(\dfrac{1}{t}\right)^2 - 2t^{p-2}\operatorname{Sin}\left(\dfrac{1}{t}\right)\operatorname{Cos}\left(\dfrac{1}{t}\right)\right| \le p\left|t^{p-1}\right| + 2\left|t^{p-2}\right|$ avec $p > 2$.

(b) Les valeurs propres de $A(t)$ sont :

$$\forall t \in [0,1],\ \begin{cases} \lambda_1(t) = \sqrt{b(t)} \\ \lambda_2(t) = -\sqrt{b(t)} \end{cases}$$

et $A(t)$ est diagonalisable pour tout t dans $[0,1]$.

(c) La fonction $\lambda_1(t) = \begin{cases} 0 \text{ si } t = 0 \\ t^{\frac{p}{2}}\left|\operatorname{Sin}\left(\dfrac{1}{t}\right)\right| \text{ si } t \neq 0 \end{cases}$ n'est pas dérivable sur $[0,1]$ si $p \le 4$.

Problème 9 : Matrices de Hessenberg et tridiagonales

On appelle « matrice de Hessenberg » une matrice A à coefficients complexes qui vérifie :

$$a_{ij} = 0 \text{ pour } j < i - 1$$

On dit que A est irréductible si $a_{i,i-1} \neq 0$ *pour tout i = 2, ⋯, n.*

Q1. *Soit P un polynôme unitaire de degré n. Construire une matrice de Hessenberg irréductible qui admet* $(-1)^n P$ *pour polynôme caractéristique.*

Q2. *Montrer que si A est une matrice de Hessenberg irréductible, alors pour toute valeur propre λ de A, l'espace propre associé est de dimension 1.*

Q3. *En déduire que les valeurs propres d'une matrice de Hessenberg irréductible sont simples si et seulement si la matrice est diagonalisable.*

Q4. *Soit* $A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 0 & b_{n-1} & a_n \end{pmatrix}$ *une matrice à coefficients réel, tridiagonale symétrique et irréductible.*

(a) Montrer que les valeurs propres de A sont simples.

(b) Décrire un algorithme de calcul de l'espace propre associé à une valeur propre de A.

Q5. *Soit* $A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 \\ c_2 & a_2 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & c_{n-1} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 0 & c_n & a_n \end{pmatrix}$ *une matrice tridiagonale à coefficients complexes.*

(a) Donner un algorithme de calcul du polynôme caractéristique de A.

(b) Montrer que A admet les mêmes valeurs propres que la matrice

$$B = \begin{pmatrix} a_1 & c_2b_1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 & c_3b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & a_{n-1} & c_nb_{n-1} \\ 0 & \cdots & 0 & 0 & 1 & a_n \end{pmatrix}$$

(c) Montrer que si $a_i \in IR$ *pour tout i = 1, 2, ···, n et* $c_{i+1}b_i \in IR_+^*$ *pour tout i = 1, 2, ···, n – 1, alors A admet n valeurs propres réelles simples et est diagonalisable.*

(d) Que peut-on dire si $a_i \in IR$ *pour tout i = 1, 2, ···, n et* $c_{i+1}b_i \in IR_-^*$ *pour tout i = 1, 2, ···, n – 1 ?*

Solution

Q1. Soit $P_n(x) = x^n - a_nx^{n-1} - \cdots - a_2x - a_1$, on lui associe la matrice de Hessenberg irréductible $A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & a_1 \\ 1 & 0 & \cdots & 0 & 0 & a_2 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 & a_{n-1} \\ 0 & \cdots & 0 & 0 & 1 & a_n \end{pmatrix}$.

Son polynôme caractéristique est $Q_n(\lambda) = \begin{vmatrix} -\lambda & 0 & 0 & \cdots & 0 & a_1 \\ 1 & -\lambda & 0 & \cdots & 0 & a_2 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & -\lambda & a_{n-1} \\ 0 & \cdots & 0 & 0 & 1 & a_n - \lambda \end{vmatrix}$.

Pour tout $k = 0, 1, \cdots, n-1$, on désigne par $Q_{n-k}(\lambda)$ le polynôme caractéristique associé au polynôme $P_{n-k}(x) = x^{n-k} - a_nx^{n-k-1} - \cdots - a_{k+2}x - a_{k+1}$.

En développant chacun des $Q_{n-k}(\lambda)$ suivant la première ligne, pour $k = 0, \cdots, n-1$, on a :

$$\begin{cases} Q_n(\lambda) = -\lambda \cdot Q_{n-1}(\lambda) + (-1)^{n+1} a_1 \\ Q_{n-1}(\lambda) = -\lambda \cdot Q_{n-2}(\lambda) + (-1)^n a_2 \\ \quad\vdots \\ Q_2(\lambda) = -\lambda \cdot Q_1(\lambda) + (-1)^3 a_{n-1} \\ Q_1(\lambda) = a_n - \lambda \end{cases}$$

Il en résulte $Q_n(\lambda) = (-1)^n P_n(\lambda)$.

Q2. Soit A une matrice de Hessenberg irréductible, λ un réel et

$$A_\lambda = A - \lambda \cdot I_d = \begin{pmatrix} a_{11}-\lambda & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22}-\lambda & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & a_{n-1,n-2} & a_{n-1,n-1}-\lambda & a_{n-1,n} \\ 0 & \cdots & 0 & 0 & a_{n,n-1} & a_{nn}-\lambda \end{pmatrix}$$

Si B_λ est la matrice extraite de A_λ en supprimant la première ligne et la dernière colonne, soit $B_\lambda = \begin{pmatrix} a_{21} & a_{22}-\lambda & a_{23} & \cdots & a_{2,n-2} & a_{2,n-1} \\ 0 & a_{32} & a_{33}-\lambda & \cdots & a_{3,n-2} & a_{3,n-1} \\ . & . & . & . & . & . \\ 0 & 0 & \cdots & 0 & a_{n-1,n-2} & a_{n-1,n-1}-\lambda \\ 0 & 0 & \cdots & 0 & 0 & a_{n,n-1} \end{pmatrix}$

on a alors $\text{Dét}(B_\lambda) = \prod_{i=2}^{n} a_{i,i-1} \neq 0.$

C'est-à-dire que :

$$\forall \lambda \in IR,\ \text{Rang}(A_\lambda) \geq n-1$$

ou encore :

$$\forall \lambda \in IR,\ \text{Dim}(\text{Ker}(A - \lambda \cdot Id)) \leq 1$$

En particulier, si λ est valeur propre de A :

$$\text{Dim}(\text{Ker}(A - \lambda \cdot Id)) = 1$$

Q3. Si A est diagonalisable, alors $C^n = \bigoplus_{k=1}^{p} \text{Ker}(A - \lambda_k Id)$, où $\{\lambda_1, \lambda_2, \cdots, \lambda_p\}$ sont toutes les valeurs propres de A. Avec Q2, on a nécessairement $p = n,$ c'est-à-dire que les valeurs propres de A sont simples.

Q4. (a) Une matrice symétrique réelle a toutes ses valeurs propres réelles et est diagonalisable. Si elle tridiagonale et irréductible alors ses valeurs propres sont nécessairement simples d'après Q3.

(b) Soit λ une valeur propre de A et $x \in IR^n - \{0\}$ un vecteur propre associé. On a alors :

$$\begin{cases} a_1 x_1 + b_1 x_2 = \lambda x_1 \\ b_{k-1} x_{k-1} + a_k x_k + b_k x_{k+1} = \lambda x_k \ \left(2 \le k \le n-1\right) \\ b_{n-1} x_{n-1} + a_n x_n = \lambda x_n \end{cases}$$

Si $x_n = 0$, alors avec l'hypothèse $b_i \neq 0$ pour tout $i = 1, \cdots, n-1$, on déduit que tous les x_i sont nuls. On peut donc prendre $x_n = 1$ et $\left(x_1, x_2, \cdots, x_{n-1}\right)$ est solution du système triangulaire supérieur :

$$\begin{cases} b_{k-1} x_{k-1} + \left(a_k - \lambda\right) x_k + b_k x_{k+1} = \ \left(2 \le k \le n-2\right) \\ b_{n-2} x_{n-2} + \left(a_{n-1} - \lambda\right) x_{n-1} = -b_{n-1} \\ b_{n-1} x_{n-1} = \lambda - a_n \end{cases}$$

La solution de ce système peut se calculer avec l'algorithme :

$$\begin{cases} x_{n-1} = \dfrac{\lambda - a_n}{b_{n-1}} \\ x_{k-1} = -\dfrac{b_k x_{k+1} + \left(a_k - \lambda\right) x_k}{b_{k-1}} \ \left(k = n-1, \cdots, 2\right) \end{cases}$$

Remarque — On retrouve que l'espace propre associé à λ est de dimension 1.

Q5. (a) Si, pour tout $k = 1, 2, \cdots, n$, on désigne par A_k la matrice principale d'ordre k de A, alors la suite $\left(P_k\right)_{1 \le k \le n}$ des polynômes caractéristiques des A_k vérifie la récurrence :

$$\begin{cases} P_0(\lambda) = 1,\ P_1(\lambda) = a_1 - \lambda \\ P_k(\lambda) = \left(a_k - \lambda\right) P_{k-1}(\lambda) - b_{k-1} c_k P_{k-2}(\lambda) \quad \left(2 \le k \le n\right) \end{cases}$$

(b) Le polynôme caractéristique de B s'obtient avec la même récurrence que celui de A en utilisant les mêmes conditions initiales, ces deux polynômes sont donc identiques.

(c) En utilisant deux fois le résultat du (b), on voit que A admet les mêmes valeurs propres que $C = \begin{pmatrix} a_1 & \sqrt{c_2 b_1} & 0 & 0 & \cdots & 0 \\ \sqrt{c_2 b_1} & a_2 & \sqrt{c_3 b_2} & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & \sqrt{c_{n-1} b_{n-2}} & a_{n-1} & \sqrt{c_n b_{n-1}} \\ 0 & \cdots & 0 & 0 & \sqrt{c_n b_{n-1}} & a_n \end{pmatrix}$.

La matrice C est symétrique tridiagonale et irréductible, elle est donc diagonalisable avec n valeurs propres réelles simples.

La matrice A admet donc n valeurs propres réelles simples et est diagonalisable.

(d) Soit $A = \begin{pmatrix} 0 & b_1 & 0 \\ -b_1 & 0 & b_2 \\ 0 & -b_2 & 0 \end{pmatrix}$. Son polynôme caractéristique est $P(\lambda) = -\lambda\left(\lambda^2 + b_2{}^2\lambda + b_1{}^2\right)$ et pour $b_2{}^4 - 4b_1{}^2 = 0$ la matrice A admet une valeur propre double et donc n'est pas diagonalisable si b_1 et b_2 sont non nuls.

CHAPITRE 2

Résolution de systèmes linéaires Déterminants et inverses

Problème 10 : Matrices de Hilbert

Q1. *Pour tout entier* $n \geq 1$, *on pose :* $R_n(x) = \dfrac{(x-1)(x-2)\cdots(x-n)}{x(x+1)\cdots(x+n)}$

Calculer le coefficient de $\dfrac{1}{x+n}$ *dans la décomposition en éléments simples de* R_n. *On note* λ_n *ce coefficient.*

Q2. *Pour tout entier* $n \geq 1$, *la matrice de Hilbert d'ordre n est définie par :*

$$H_n = \left(\left(h_{ij}\right)\right)_{\substack{1\leq i\leq n \\ 1\leq j\leq n}}, \text{ avec : } h_{ij} = \frac{1}{i+j-1} \quad (1 \leq i, j \leq n)$$

On note A_n *la matrice d'ordre n définie par :*

$$A_n = \left(\left(a_{ij}\right)\right)_{\substack{1\leq i\leq n \\ 1\leq j\leq n}}, \text{ avec } a_{ij} = \begin{cases} h_{ij} \ si \ j = 1,2,\cdots,n-1;\ i = 1,2,\ldots,n \\ R_{n-1}(i) \ si \ j = n;\ i = 1,2,\cdots,n \end{cases}$$

(a) Comparer $\text{Dét}(A_n)$ *et* $\text{Dét}(H_n)$.

(b) Comparer $\text{Dét}(A_n)$ *et* $\text{Dét}(H_{n-1})$.

(c) En déduire une récurrence sur les $\Delta_n = \text{Dét}(H_n)$.

Q3. *En posant, pour tout entier* $n \geq 1$, $\Phi_n = \prod_{k=1}^{n} k!$

Montrer que : $\Delta_n = \text{Dét}(H_n) = \dfrac{{\Phi_{n-1}}^4}{\Phi_{2n-1}}$

Q4. *Montrer que pour tout entier* $n \geq 1$, $\Delta_n \leq \dfrac{1}{n^n}$.

Q5. *(a) En posant* $u_n = C_{2n}^n$, *montrer que :* $\Delta_{n+1} = \dfrac{\Delta_n}{(2n+1){u_n}^2}$

(b) En déduire, sous forme de programmation structurée, un algorithme de calcul des Δ_n.

Solution

Q1. La décomposition en éléments simples de R_n s'écrit $R_n(x) = \sum_{k=0}^{n} \frac{\lambda_{k,n}}{(x+k)}$ avec $\lambda_{k,n} = \underset{x \to -k}{Lim}\left((x+k)R_n(x)\right)$. Et en particulier :

$$\lambda_n = \lambda_{n,n} = \frac{(2n)!}{(n!)^2} = C_{2n}^n$$

Q2. (a) On a $A_n = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1,n-1} & R_{n-1}(1) \\ h_{21} & h_{22} & \cdots & h_{2,n-1} & R_{n-1}(2) \\ . & . & . & . & . \\ . & . & . & . & . \\ h_{n1} & h_{n2} & \cdots & h_{n,n-1} & R_{n-1}(n) \end{pmatrix}$ avec :

$$R_{n-1}(i) = \sum_{k=0}^{n-1} \frac{\lambda_{k,n-1}}{(i+k)} = \sum_{k=0}^{n-1} \lambda_{k,n-1} h_{i,k+1}$$

Soit $R_{n-1}(i) = \sum_{j=1}^{n} \lambda_{j-1,n-1} h_{i,j}$ pour tout $i = 1, 2, \cdots, n$.

Si on note C_j la colonne numéro j de H_n, alors la colonne numéro n de A_n est combinaison linéaire des $C_1, C_2, \cdots, C_n$. Le déterminant étant une forme n–linéaire alternée, on en déduit que :

$$\text{Dét}(A_n) = \lambda_{n-1}\,\text{Dét}(H_n) = C_{2n}^n\,\text{Dét}(H_n)$$

(b) D'autre part, on a $R_{n-1}(i) = 0$ pour tout $i = 1, 2, \cdots, n-1$. En développant le déterminant suivant la dernière colonne on en déduit alors :

$$\text{Dét}(A_n) = R_{n-1}(n)\,\text{Dét}(H_{n-1})$$

(c) On déduit de ce qui précède que $\text{Dét}(H_n) = \frac{R_{n-1}(n)}{\lambda_{n-1}}\text{Dét}(H_{n-1})$, avec $\lambda_{n-1} = \frac{(2n-2)!}{((n-1)!)^2}$ et $R_{n-1}(n) = \frac{((n-1)!)^2}{(2n-1)!}$. D'où :

$$\Delta_n = \text{Dét}(H_n) = \frac{((n-1)!)^4}{(2n-1)!(2n-2)!}\Delta_{n-1}$$

Q3. Par récurrence, on déduit que :

$$\Delta_n = \frac{\{(n-1)!(n-2)!\cdots 2!\}^4}{(2n-1)!(2n-2)!\cdots 3!2!} = \frac{{\Phi_{n-1}}^4}{\Phi_{2n-1}}$$

Q4. On a $\Delta_n = \frac{\{(n-1)!(n-2)!\cdots 2!\}^3}{(2n-1)!(2n-2)!\cdots n!}$, soit avec $C_{2k}^k = \frac{(2k)!}{(k!)^2}$:

$$\Delta_n = \frac{(n-1)!}{(2n-1)!}\frac{1}{C_{2(n-1)}^{n-1}}\frac{(n-2)!}{(2n-3)!}\frac{1}{C_{2(n-2)}^{n-2}}\cdots\frac{2!}{5!}\frac{1}{C_4^2}\frac{1}{3!}\frac{1}{C_2^1}.$$

En posant $\psi_k = \dfrac{k!}{(2k+1)!}\dfrac{1}{C_{2k}^k}$, on a $\Delta_n = \prod_{k=1}^{n-1} \psi_k$.

Avec $C_{2k}^k \geq 1$ on déduit que $\psi_k \leq 1$ et $\Delta_n \leq \psi_{n-1}$, avec
$\psi_{n-1} = \dfrac{(n-1)!}{(2n-1)!}\dfrac{1}{C_{2(n-1)}^{n-1}} \leq \dfrac{(n-1)!}{(2n-1)!} = \dfrac{1}{(2n-1)\cdots n} \leq \dfrac{1}{n^n}$. D'où le résultat.

C'est-à-dire que Δ_n tend vers 0 très vite quand n tend vers l'infini.

Pour n grand, on a $\text{Dét}(H_n)_{machine} = 0$ et le système est numériquement dégénéré, mais non dégénéré.

Pour tester la performance d'un algorithme de résolution d'un système linéaire, on pourra l'essayer avec une matrice de Hilbert d'ordre élevé.

Q5. (a) On a $\Phi_n = n!\Phi_{n-1}$, d'où $\Delta_{n+1} = \dfrac{\Phi_n^{\ 4}}{\Phi_{2n+1}} = \dfrac{n!^4\,\Phi_{n-1}^{\ 4}}{(2n+1)!(2n)!\Phi_{2n-1}}$, soit
$\Delta_{n+1} = \dfrac{1}{2n+1}\left(\dfrac{(n!)^2}{(2n)!}\right)^2 \Delta_n$. C'est-à-dire $\Delta_{n+1} = \dfrac{\Delta_n}{(2n+1)u_n^{\ 2}}$.

(b) En remarquant que $u_{n+1} = \dfrac{2(2n+1)}{n+1}u_n$, on déduit l'algorithme de calcul d'un déterminant de Hilbert :

```
Fonction Déterminant_Hilbert(Entrée n : Entier): Réel;
Début
    D = 1; u = 2;
    Pour k allant de 2 à n faire
    Début
        D = D/((2·k – 1)·u²);
        u = 2·(2·k - 1)·u/k;
    Fin;
    Retourner D;
Fin;
```

Problème 11 : Polynômes de Legendre et matrices de Hilbert

Pour tout entier n > 0, on désigne par E_n l'espace vectoriel réel des polynômes de degré au plus n. On munit E_n du produit scalaire défini par :

$$\forall\ P,Q \in E_n,\ \langle P|Q\rangle = \int_0^1 P(x)Q(x)dx$$

Les polynômes de Legendre d'ordre k sont définis par :

$$\begin{cases} L_1(x) = 1 \\ \forall k \geq 2,\ L_k(x) = \left(\dfrac{d}{dx}\right)^{k-1}\left(x^{k-1}(1-x)^{k-1}\right) \end{cases}$$

Q1. *Calculer $\langle P|Q\rangle$, pour $P(x) = x^p$ et $Q(x) = x^q$, avec p et q entiers.*

Q2. *Soit $\{e_1, e_2, \cdots, e_n\}$ la base de E_{n-1} définie par $e_i(x) = x^{i-1}$ pour i = 1, ⋯, n.*

(a) Comparer la matrice du produit scalaire défini ci-dessus avec la matrice de Hilbert H_n du problème 10.
(b) Que dire des valeurs propres de H_n ?
***Q3.** Calculer le degré et le coefficient du terme de plus haut degré de L_k.*
***Q4.** Montrer que pour tout entier $k > 0$ et toute fonction f admettant des dérivées continues jusqu'à l'ordre $k-1$ sur $[0,1]$, on a :*

$$\langle f \mid L_k \rangle = (-1)^{k-1} \int_0^1 f^{(k-1)}(x) x^{k-1} (1-x)^{k-1} dx$$

***Q5.** (a) Montrer que $\{L_1, L_2, \cdots, L_n\}$ est une base orthogonale de E_{n-1}.*
(b) Calculer $\|L_k\|^2$, pour tout entier $k \geq 1$.
***Q6.** Calculer la matrice de passage P_n de $\{e_1, e_2, \cdots, e_n\}$ à $\{L_1, L_2, \cdots, L_n\}$.*
***Q7.** Montrer qu'il existe une matrice D_n diagonale telle que $D_n = {}^tP_n \cdot H_n \cdot P_n$.*
***Q8.** En déduire le déterminant de H_n.*
***Q9.** Calculer l'inverse de H_n. On montrera que ces coefficients sont entiers.*

Solution

Q1. $\langle P | Q \rangle = \int_0^1 x^{p+q} dx = \dfrac{1}{p+q+1}$.

Q2. (a) La matrice du produit scalaire est la matrice de composantes $\langle e_i | e_j \rangle = \dfrac{1}{i+j-1}$, pour $1 \leq i, j \leq n$. C'est la matrice de Hilbert H_n.

(b) La matrice H_n est la matrice d'un produit scalaire, elle est donc symétrique définie positive et toutes ses valeurs propres sont réelles strictement positives.

Q3. Pour tout entier $k \geq 2$, le polynôme $P_k(x) = x^{k-1}(1-x)^{k-1}$ est de degré $2k-2$, donc L_k est de degré $k-1$. On en déduit alors que la famille libre $\{L_1, L_2, \cdots, L_n\}$ est une base de E_{n-1}.

Pour tout entier $k \geq 1$, le coefficient dominant de L_k est α_{k-1} défini par :

$$\alpha_{k-1} x^{k-1} = \left(\frac{d}{dx}\right)^{k-1} \left((-1)^{k-1} x^{2k-2}\right) = (-1)^{k-1} \frac{(2k-2)!}{(k-1)!} x^{k-1}$$

Q4. Pour tout entier $k \geq 2$, on a $\langle f \mid L_k \rangle = \int_0^1 f(x) \left(\dfrac{d}{dx}\right)^{k-1} \left(x^{k-1}(1-x)^{k-1}\right) dx$. En faisant $k-1$ intégrations par parties on a :

$$\langle f \mid L_k \rangle = \left[\sum_{j=0}^{k-2} f^{(j)}(x) \left(\frac{d}{dx}\right)^{k-j-2} \left(x^{k-1}(1-x)^{k-1}\right)\right]_0^1 + (-1)^{k-1} \int_0^1 f^{(k-1)}(x) x^{k-1} (1-x)^{k-1} dx$$

En remarquant que 0 et 1 sont racines d'ordre $k-1$ de P_k on déduit que $\left(\dfrac{d}{dx}\right)^{k-j-2} \left(x^{k-1}(1-x)^{k-1}\right)$ s'annule en 0 et en 1 pour tout $j = 0, 1, \cdots, k-2$ et le résultat. Pour $k = 1$, le résultat est évident.

Q5. (a) En Q3, on a déjà vu que $\{L_1, L_2, \cdots, L_n\}$ est une base de E_{n-1}. De Q4 on déduit que pour tout $k \geq 2$ et tout polynôme P de degré strictement inférieur à

$k-1$ on a $\langle P | L_k \rangle = 0$. Chaque L_k étant de degré $k-1$, il en résulte que $\langle L_p | L_q \rangle = 0$ si $p \neq q$ sont des entiers strictement positifs. Donc $\{L_1, L_2, \cdots, L_n\}$ est une base orthogonale de E_{n-1}.

(b) On a $\|L_k\|^2 = (-1)^{k-1} \int_0^1 L_k^{(k-1)}(x) x^{k-1} (1-x)^{k-1} dx$, avec $L_k^{(k-1)}(x) = \alpha_{k-1}(k-1)!$, soit d'après Q3, $L_k^{(k-1)}(x) = (-1)^{k-1}(2k-2)!$.

Il s'agit alors de calculer $I_k = \int_0^1 x^k (1-x)^k dx$. Ce calcul se fait facilement avec k intégrations par parties qui donnent $I_k = \dfrac{(k!)^2}{(2k+1)!}$.

On a donc en définitive :

$$\|L_k\|^2 = \frac{((k-1)!)^2}{2k-1}$$

Q6. Pour tout entier $k \geq 2$, on a :

$$L_k(x) = \left(\frac{d}{dx}\right)^{k-1} \left(\sum_{i=0}^{k-1} (-1)^i C_{k-1}^i x^{k-1+i} \right) = \sum_{i=0}^{k-1} (-1)^i C_{k-1}^i \frac{(k-1+i)!}{i!} x^i$$

Ce qui s'écrit aussi :

$$\forall k \geq 1,\ L_k(x) = \sum_{i=1}^{k} (-1)^{i-1} C_{k-1}^{i-1} \frac{(k+i-2)!}{(i-1)!} e_i(x)$$

C'est-à-dire que la matrice de passage de $\{e_1, e_2, \cdots, e_n\}$ à $\{L_1, L_2, \cdots, L_n\}$ est la matrice triangulaire supérieure $P_n = ((p_{ij}))_{1 \leq i,j \leq n}$ définie par :

$$p_{ij} = (-1)^{i-1} C_{j-1}^{i-1} \frac{(j+i-2)!}{(i-1)!} \quad (1 \leq i \leq j \leq n)$$

Q7. La matrice du produit scalaire dans la nouvelle base $\{L_1, L_2, \cdots, L_n\}$ est $D_n = {}^tP_n H_n P_n$. Cette dernière base étant orthogonale, la matrice D_n est diagonale de termes diagonaux $\|L_k\|^2$ $(1 \leq k \leq n)$.

Q8. Le déterminant de H_n est alors donné par :

$$\text{Dét}(H_n) = \frac{\text{Dét}(D_n)}{\text{Dét}(P_n)^2} = \frac{\prod_{k=1}^{n} \|L_k\|^2}{\prod_{k=1}^{n} \alpha_{k-1}^{\ 2}} = \prod_{k=1}^{n} \frac{((k-1)!)^2}{((2k-2)!)^2} \frac{((k-1)!)^2}{2k-1}$$

Soit en remarquant que $(2k-1)((2k-2)!)^2 = (2k-1)!(2k-2)!$:

$$\text{Dét}(H_n) = \frac{\prod_{k=1}^{n} ((k-1)!)^4}{\prod_{k=1}^{n} (2k-1)!(2k-2)!} = \frac{\left(\prod_{k=1}^{n-1} k! \right)^4}{\prod_{k=1}^{2n-1} k!}$$

On retrouve bien le résultat du problème 10, Q3.

Q9. On a $H_n = {}^tP_n^{-1} D_n P_n^{-1}$ et $H_n^{-1} = P_n D_n^{-1}\,{}^tP_n$, avec :

$$P_n D_n^{-1} = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1,n} \\ 0 & \lambda_2 p_{22} & \cdots & \lambda_n p_{2,n} \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \cdots & \lambda_n p_{n,n} \end{pmatrix} \text{ où } \lambda_i = \frac{2i-1}{\left((i-1)!\right)^2} \quad (1 \le i \le n)$$

Ce qui donne pour le coefficient d'indices (i,j) de H_n^{-1}, avec $1 \le i \le j \le n$ (la matrice est symétrique) :

$$u_{ij} = \begin{pmatrix} 0 & \cdots & 0 & \lambda_i p_{ii} & \cdots & \lambda_n p_{in} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_{jj} \\ \vdots \\ p_{jn} \end{pmatrix} = \sum_{k=j}^{n} \lambda_k p_{ik} p_{jk}$$

Soit :

$$u_{ij} = (-1)^{i+j} \sum_{k=j}^{n} (2k-1) C_{k-1}^{i-1} C_{k-1}^{j-1} C_{i+k-2}^{k-1} C_{j+k-2}^{k-1} \in Z \quad (1 \le i \le j \le n)$$

Problème 12 : Système tridiagonal. Agrégation 1993, extrait

On reprend les notations du problème 5.

Q1. *Soit $(p_k)_{k\ge 1}$ la suite définie par :*

$$\begin{cases} p_1 = m \\ p_{k+1} = m - \dfrac{1}{p_k} \end{cases}$$

Etablir que cette suite est définie pour tout entier $k \ge 1$ lorsque $m > 2$.

Q2. *On considère une matrice colonne B d'ordre n et le système $T_m X = B$ avec $m > 2$.*

Donner un équivalent du nombre d'opérations réelles nécessitées par la résolution de ce système par la méthode du pivot de Gauss sans échange de lignes.

Solution

Q1. Soit $f : x \mapsto m - \frac{1}{x}$ pour $x > 0$. Cette fonction étant strictement croissante, on a $f([1,m]) = \left[m-1, m-\frac{1}{m}\right] \subset [1,m]$, pour $m > 2$ et la suite $(p_k)_{k\geq 1}$ des approximations successives est bien définie avec $1 \leq p_k \leq m$ pour tout $k \geq 1$.

Q2. Si, pour tout k = 1, 2, ···, n, on désigne par L_k la ligne numéro k du système $T_m X = B$, alors l'algorithme du pivot de Gauss sans échange de ligne consiste à effectuer les opérations :

$$L_{k+1} = L_{k+1} + \frac{1}{p_k} L_k \quad (1 \leq k \leq n-1)$$

Soit en notant respectivement $T_m^{(k)} = \left(\left(t_{ij}^{(k)}\right)\right)_{1\leq i,j\leq n}$ et $B^{(k)} = \left(\left(b_i^{(k)}\right)\right)_{1\leq i\leq n}$ la matrice et le second membre à l'étape k, avec $T_m^{(1)} = T_m$:

$$\begin{cases} t_{k+1,k}^{(k+1)} = 0 \\ t_{k+1,k+1}^{(k+1)} = m - \dfrac{1}{t_{k,k}^{(k)}} = m - \dfrac{1}{p_k} = p_{k+1} \ (1 \leq k \leq n-1) \\ b_{k+1}^{(k+1)} = b_{k+1}^{(k)} + \dfrac{1}{p_k} b_k^{(k)} \end{cases}$$

On aboutit alors au bout de n – 1 étapes au système triangulaire $T_m^{(n)} X = B^{(n)}$, avec $T_m^{(n)} = \begin{pmatrix} p_1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & p_2 & -1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 0 & p_{n-1} & -1 \\ 0 & \cdots & 0 & 0 & 0 & p_n \end{pmatrix}$. La résolution de ce système se fait « en remontée » :

$$\begin{cases} x_n = \dfrac{b_n^{(n)}}{p_n} \\ x_k = \dfrac{1}{p_k}\left(x_{k+1} + b_k^{(n)}\right) \ (k = n-1, \cdots, 1) \end{cases}$$

Dans la phase de réduction à la forme triangulaire on a, pour chaque k = 1, ···, n – 1, une division et une soustraction pour le calcul de p_{k+1} et une division et une addition pour le calcul de $b_{k+1}^{(k+1)}$. Ce qui donne un total de $2n-2$ divisions et $2n-2$ additions.

La résolution du système triangulaire nécessite $n-1$ additions et n divisions.

On a donc un total de $3n-2$ divisions et $3n-3$ additions. Soit un nombre d'opérations de l'ordre de $6n$.

Problème 13 : Système tridiagonal

Soit $A_n = \begin{pmatrix} a_1 & c_1 & 0 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & 0 & b_n & a_n \end{pmatrix}$ *une matrice tridiagonale à coefficients réels.*

Q1. *Donner un algorithme de calcul de* $D_n = \text{Dét}(A_n)$.

On suppose que A_n est inversible et que la méthode du pivot de Gauss peut s'effectuer sans échange de ligne. On sait alors que A_n peut s'écrire de manière unique sous la forme $A_n = L \cdot R$ avec L triangulaire inférieure à diagonale unité et R triangulaire supérieure.

Q2. *Donner un algorithme de calcul des matrices L et R.*

Ecrire la programmation structurée correspondante.

Q3. *(a) Décrire un algorithme de résolution de $A_n x = e$ qui utilise la décomposition LR. Ecrire la programmation structurée correspondante.*

(b) Décrire un algorithme de résolution de $A_n x = e$ qui utilise la méthode de Gauss sans échange de ligne. Ecrire la programmation structurée correspondante.

(c) Décrire un algorithme de calcul de l'inverse de A_n qui utilise la décomposition LR. Ecrire la programmation structurée correspondante.

Q4. *On note A_n la matrice symétrique d'ordre n définie par :*

$$A_n = \left(\left(a_{ij}\right)\right)_{\substack{1\le i\le n \\ 1\le j\le n}}, \text{ avec } \begin{cases} a_{ij} = n - j + 1 \text{ pour } 1 \le i \le j \le n \\ a_{ij} = a_{ji} \text{ pour } 1 \le j < i \le n \end{cases}$$

(a) Calculer $\text{Dét}(A_n)$.

(b) Calculer l'inverse de A_n. On notera B_n cette matrice inverse.

(c) Calculer $\text{Dét}(B_n)$ en utilisant l'algorithme de Q1 et retrouver la valeur de $\text{Dét}(A_n)$ calculée en (a).

(d) Donner la décomposition LR de A_n.

Solution

Q1. En développant D_n suivant la dernière ligne, on a :

$$D_n = a_n D_{n-1} - b_n c_{n-1} D_{n-2}$$

Ce qui donne, avec les valeurs initiales $D_0 = 1$ et $D_1 = a_1$, un algorithme très simple.

Q2. On cherche les matrices L et R sous la forme bidiagonale, soit :

$$L=\begin{pmatrix}1&0&0&0&\cdots&0\\L_2&1&0&0&\cdots&0\\.&.&.&.&.&.\\0&\cdots&0&L_{n-1}&1&0\\0&\cdots&0&0&L_n&1\end{pmatrix}\quad\text{et}\quad R=\begin{pmatrix}d_1&r_1&0&0&\cdots&0\\0&d_2&r_2&0&\cdots&0\\.&.&.&.&.&.\\0&\cdots&0&0&d_{n-1}&r_{n-1}\\0&\cdots&0&0&0&d_n\end{pmatrix}$$

On effectuant le produit $A_n = L\cdot R$ et en identifiant avec les coefficients de A_n, on obtient alors :

$$\begin{cases}d_1 = a_1\\ r_j = c_j \ (j=1,\cdots,n-1)\end{cases}$$

puis :

$$\begin{cases}L_{j+1} = \dfrac{b_{j+1}}{d_j}\\ d_{j+1} = a_{j+1} - r_j L_{j+1}\end{cases}\quad (j=1,\cdots,n-1)$$

La programmation structurée est alors évidente.

Q3. (a) Tout d'abord, on résout $Ly=e$, ce qui donne :

$$\begin{cases}y_1 = e_1\\ y_j = e_j - L_j y_{j-1} \ (j=2,\cdots,n)\end{cases}$$

Puis on résout $Rx=y$, ce qui donne :

$$\begin{cases}x_n = \dfrac{y_n}{d_n}\\ x_j = \dfrac{(y_j - r_j x_{j+1})}{d_j} \ (j=n-1,\cdots,1)\end{cases}$$

Là encore la programmation structurée est évidente.

(b) En supposant qu'il n'y a pas de permutations dans la méthode de Gauss (une telle permutation ferait perdre le caractère tridiagonal), on n'aura à chaque étape qu'une ligne à traiter et pour chaque ligne seulement deux opérations.

A l'étape k de la méthode, la ligne L_k du système devient :

$$L_k - m\cdot L_{k-1} \ (k=2,\cdots,n)$$

avec $m = \dfrac{b_k}{a_{k-1}}$.

Ce qui donne les formules de transformations :

$$\begin{aligned}&b_k = 0\\ &a_k = a_k - m\cdot c_{k-1}\\ &c_k \ \text{inchangé}\\ &e_k = e_k - m\cdot e_{k-1}\end{aligned}$$

Le système triangulaire supérieur obtenu sera bidiagonal et aura pour solution :

$$\begin{cases} x_n = \dfrac{e_n}{a_n} \\ x_i = \dfrac{\left(e_i - c_i x_{i+1}\right)}{a_i} \ \left(i = n-1, \cdots, 1\right) \end{cases}$$

Remarque 1 — Le produit des a_i donnera, en fin d'opération, le déterminant de A.
Remarque 2 — La méthode obtenue est appelée *« méthode du double balayage de Cholesky »*.

Pour la programmation structurée, en vue d'économiser la place mémoire, on stocke une matrice tridiagonale sous forme de trois vecteurs. Une telle matrice sera donc notée :

$$T = \left(T_{ij}\right)_{\substack{1 \le i \le n \\ 1 \le j \le 3}}$$

où i est le numéro de la ligne et j celui de la diagonale, c'est-à-dire que $\left(T_{i1}\right)_{1 \le i \le n}$ représente la diagonale inférieure, $\left(T_{i2}\right)_{1 \le i \le n}$ la diagonale et $\left(T_{i3}\right)_{1 \le i \le n}$ la diagonale supérieure.

Ce qui donne la procédure :

PROCEDURE SystèmeTridiagonal(Entrée n : Entier ;
Entrée_Sortie T : MatriceTridiagonale ; b : Vecteur);
Début
Pour i Allant de 2 à n Faire
Début
Si ($T_{i-1,2}$ *= 0) Alors Stop(« Déterminant nul ») ;*
$m = T_{i,1}/T_{i-1,2}$ *;*
$T_{i,2} = T_{i,2} - mT_{i-1,3}$ *;*
$b_i = b_i - mb_{i-1}$ *;*
Fin ;
Si ($T_{n,2}$ *= 0)*
Alors Stop(« Déterminant nul ») ;
$b_n = b_n/T_{n,2}$ *;*
Pour i Allant de n – 1 à 1 Faire $b_i = (b_i - T_{i,3}b_{i+1})/T_{i,2}$ *;*
Fin ;

Remarque 3 — En fin d'opération les valeurs initiales de la matrice T et du second membre b sont perdues, le vecteur b contenant la solution du système.
(c) Avec $A_n = L \cdot R$, on a $A_n^{-1} = R^{-1}L^{-1}$. De plus, si A est une matrice inversible, alors les colonnes de A^{-1} sont les solutions de $Ax = e_j$, où $\{e_1, e_2, \cdots, e_n\}$ est la base canonique de $\mathbb{R}^n$. Ce qui donne, pour les coefficients de $U = R^{-1}$ et $V = L^{-1}$:

$$\forall j \in \{1,2,\cdots,n\} \begin{cases} u_{jj} = \dfrac{1}{d_j} \\ u_{ij} = -\dfrac{r_i u_{i+1,j}}{d_i} \ (i = j-1,\cdots,1) \\ u_{ij} = 0 \ (i = j+1,\cdots,n) \end{cases} \begin{cases} v_{ij} = 0 \ (i = 1,\cdots,j-1) \\ v_{jj} = 1 \\ v_{ij} = -L_i u_{i-1,j} \ (i = j+1,\cdots,n) \end{cases}$$

Les coefficients de $W = A_n^{-1}$ sont alors donnés par :

$$w_{ij} = \sum_{k=Max(i,j)}^{n} u_{ik} v_{kj} \ (1 \le i, j \le n)$$

Ce qui donne la programmation structurée :

PROCEDURE InverseTriangulaireInférieure(Entrée n : Entier ; A : Matrice ;
Sortie W : Matrice) ;

Début
Pour j allant de 1 à n faire
Début
Si (d_j = 0) Alors Stop(« Matrice non inversible ») ;
u_{jj} = 1/d_j ;
v_{jj} = 1;
Pour i allant de j – 1 à 1 faire
Début
$u_{ij} = -\dfrac{r_i u_{i+1,j}}{d_i}$;
v_{ij} = 0 ;
Fin ;
Pour i allant de j + 1 à n faire
Début
u_{ij} = 0 ;
$v_{ij} = -L_i u_{i-1,j}$;
Fin ;
Fin ;
Pour i allant de 1 à n faire
Début
Pour j allant de 1 à i faire
Début
w_{ij} = 0 ;
Pour k allant de 1 à i faire $w_{ij} = w_{ij} + u_{ik}v_{kj}$;
Fn ;
Pour j allant de i + 1 à n faire
Début
w_{ij} = 0 ;
Pour k allant de j à n faire $w_{ij} = w_{ij} + u_{ik}v_{kj}$;
Fn ;
Fin ;

Fin ;

Q4. (a) On a $A_n = \begin{pmatrix} n & n-1 & \cdots & 2 & 1 \\ n-1 & n-1 & \cdots & 2 & 1 \\ . & . & . & . & . \\ 2 & 2 & \cdots & 2 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$. En retranchant la deuxième ligne à la première, on déduit que :

$$\forall n \geq 2,\ Dét(A_n) = Dét(A_{n-1})$$

Et par récurrence :

$$\forall n \geq 2,\ Dét(A_n) = 1$$

(b) D'après (a), la matrice A_n est inversible dans $M_n(Z)$, c'est-à-dire que les coefficients de A_n^{-1} sont entiers.

Après quelques expériences numériques on fait l'hypothèse que :

$$A_n^{-1} = B_n = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -1 & 2 \end{pmatrix}$$

Pour vérifier cette hypothèse, il suffit de vérifier que la colonne numéro j de B_n est solution du système linéaire $A_n x = e_j$ où e_j est le $j^{\text{ème}}$ vecteur de base canonique.

(c) D'après Q1, l'algorithme de calcul du déterminant de B_n est donné par :

$$\begin{cases} \Delta_1 = 1, \quad \Delta_2 = 1 \\ \Delta_n = 2\Delta_{n-1} - \Delta_{n-2} \quad (n \geq 3) \end{cases}$$

On vérifie alors facilement que $\Delta_n = 1$, pour tout $n \geq 1$. On retrouve bien le résultat du (a) : $D_n = \dfrac{1}{\Delta_n} = 1$.

(d) Si $A_n = L \cdot R$, alors $B_n = A_n^{-1} = R'L'$, avec $L' = L^{-1}$ triangulaire inférieure à diagonale unité et $R' = R^{-1}$ triangulaire supérieure.

La matrice B_n étant tridiagonale, on cherche les matrices L' et R' sous la forme :

$$L' = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ L'_2 & 1 & 0 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & L'_{n-1} & 1 & 0 \\ 0 & \cdots & 0 & 0 & L'_n & 1 \end{pmatrix} \text{ et } R' = \begin{pmatrix} d'_1 & r'_1 & 0 & 0 & \cdots & 0 \\ 0 & d'_2 & r'_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 0 & d'_{n-1} & r'_{n-1} \\ 0 & \cdots & 0 & 0 & 0 & d'_n \end{pmatrix}$$

Par identification des coefficients dans le produit $B_n = R'L'$, on obtient :

$$r_i' = -1 \ (i = 1,\cdots,n-1), \ \begin{cases} d_n' = 2 \\ L_i' = -\dfrac{1}{d_i'}, \ d_{i-1}' = 2 + L_i' \quad (i = n,\cdots,3) \end{cases}, \ \begin{cases} L_2' = -\dfrac{1}{d_2'} \\ d_1' = 1 + L_2' \end{cases}$$

Ce qui donne par récurrence :

$$\begin{cases} r_i' = -1 \ (i = 1,\cdots,n-1) \\ d_i' = -\dfrac{1}{L_i'} = \dfrac{n+2-i}{n+1-i} \ (i = 2,\cdots,n) \\ d_1' = \dfrac{1}{n} \end{cases}$$

On vérifie alors que l'inverse de la matrice R' est la matrice :

$$R = \begin{pmatrix} n & n-1 & \cdots & 2 & 1 \\ 0 & \dfrac{n-1}{n} & \cdots & \dfrac{2}{n} & \dfrac{1}{n} \\ . & . & . & . & . \\ 0 & 0 & \cdots & \dfrac{2}{3} & \dfrac{1}{3} \\ 0 & 0 & \cdots & 0 & \dfrac{1}{2} \end{pmatrix}$$

et que l'inverse de la matrice L' est la matrice :

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \dfrac{n-1}{n} & 1 & \cdots & 0 & 0 \\ . & . & . & . & . \\ \dfrac{2}{n} & \dfrac{2}{n-1} & \cdots & 1 & 0 \\ \dfrac{1}{n} & \dfrac{1}{n-1} & \cdots & \dfrac{1}{2} & 1 \end{pmatrix}$$

Problème 14 : Méthodes itératives de résolution des systèmes linéaires

Ce problème utilise les résultats des problèmes 1 et 2.

On désigne par A une matrice d'ordre $n \geq 2$ inversible à coefficients réels.

L'idée des méthodes itératives, pour résoudre le système linéaire $Ax = b$, consiste à construire une suite $\left(x^{(k)}\right)_{k \in IN}$ d'éléments de IR^n qui va converger vers la solution de ce système. Le calcul de chaque $x^{(k)}$ doit être plus simple que la résolution directe du système.

On écrit la matrice A sous la forme $A = M - N$, où M est « facilement inversible » et la résolution de $Ax = b$ est ramenée au « problème de point fixe » : trouver x dans IR^n solution de :

$$x = M^{-1}Nx + M^{-1}b.$$

Pour résoudre ce problème, on utilise la « méthode des approximations successives », c'est-à-dire qu'on considère la suite $\left(x^{(k)}\right)_{k\in IN}$ *définie par :*

$$(1)\quad \begin{cases} x^{(0)} \text{ donné dans } IR^n \\ x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b \end{cases}$$

Q1. *Montrer que si cette suite converge, c'est nécessairement vers la solution de* $Ax = b$.

Q2. *Montrer que la suite* $\left(x^{(k)}\right)_{k\in IN}$ *définie par (1) est convergente vers la solution de* $Ax = b$, *quelle que soit la valeur initiale* $x^{(0)}$ *si et seulement si* $\rho(B) < 1$ *où* $\rho(B)$ *désigne le rayon spectral de* $B = M^{-1}N$ *si* $A = M - N$ *avec M inversible.*

Q3. *Montrer que s'il existe une norme matricielle induite telle que* $\|B\| < 1$, *alors la méthode itérative est convergente.*

Q4. *On suppose que tous les coefficients diagonaux de A sont non nuls et on considère la méthode itérative définie par le choix de la matrice diagonale* $M = D$ *avec* $d_{ii} = a_{ii}$ *pour tout i = 1, ⋯, n. La méthode obtenue est la « méthode de Jacobi ».*

(a) Décrire l'algorithme de construction des $x^{(k)}$.

(b) Ecrire la programmation structurée correspondante.

(c) Montrer que si A est à diagonale strictement dominante, alors la méthode de Jacobi converge.

Q5. *On suppose que tous les coefficients diagonaux de A sont non nuls et on considère la méthode itérative définie par le choix de la matrice M définie par :*

$$\begin{cases} m_{ij} = 0 \ , \ pour\ 1 \le i < j \le n \\ m_{ij} = a_{ij} \ , \ pour\ 1 \le j \le i \le n \end{cases}$$

La méthode obtenue est la « méthode de Gauss–Seidel ».

(a) Décrire l'algorithme de construction des $x^{(k)}$.

(b) Ecrire la programmation structurée correspondante.

(c) Montrer que si A est à diagonale strictement dominante alors la méthode de Gauss–Seidel converge.

(d) Montrer que si A est symétrique définie positive alors la méthode de Gauss-Seidel converge.

(e) Soit $A = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$. *Que peut-on dire de la convergence de la méthode de Jacobi et de celle de Gauss-Seidel ?*

Q6. *On se donne une matrice tridiagonale inversible :*

$$A = \begin{pmatrix} a_1 & c_1 & 0 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & 0 & b_n & a_n \end{pmatrix}$$

avec $n \geq 3$ et tous les a_i non nuls.

Si $A = D + E + F$, où D est la diagonale de A, E le triangle inférieur strict et F le triangle supérieur strict, on pose $J = -D^{-1}(E+F)$ pour la méthode de Jacobi et $G = -(D+E)^{-1}F$ pour la méthode de Gauss-Seidel.

(a) Pour tout matrice $M = \left((m_{ij})\right)_{1 \leq i,j \leq n}$, on définit, pour tout réel non nul t, la matrice $M(t) = \left((m_{ij}(t))\right)_{1 \leq i,j \leq n}$ par $m_{ij}(t) = t^{i-j} m_{ij}$ $(1 \leq i, j \leq n)$.

Montrer que :

$$\forall t \in IR^*, \ \text{Dét}(M(t)) = \text{Dét}(M)$$

(b) Montrer que si P_J est le polynôme caractéristique de J et P_G celui de G, alors on a $P_G(\lambda^2) = \lambda^n P_J(\lambda)$.

En déduire que $\rho(G) = (\rho(J))^2$ et conclure.

Solution

Q1. C'est clair.

Q2. En écrivant que $Ax = b$ équivaut à $x = Bx + M^{-1}b$, on déduit que :

$$\forall k \in IN, \ x^{(k+1)} - x = B\left(x^{(k)} - x\right) = B^{k+1}\left(x^{(0)} - x\right)$$

et donc $\left(x^{(k)}\right)_{k \in IN}$ va converger vers *x,* quelle que soit la valeur initiale si, et seulement si $\underset{k \to +\infty}{Lim} B^k = 0$, ce qui équivaut à $\rho(B) < 1$ (problème 2, Q6).

Q3. Résulte de l'inégalité $\rho(B) < \|B\|$, pour toute norme matricielle induite (problème 2, Q1).

Remarque — On a $\left\|B^k\right\| = Sup\left\{\dfrac{\left\|x^{(k)} - x\right\|}{\left\|x^{(0)} - x\right\|};\ x^{(0)} \in IR^n,\ x^{(0)} \neq x\right\}$, pour toute norme matricielle subordonnée. On peut donc utiliser $\left\|B^k\right\|$ pour mesurer la vitesse de convergence d'une méthode itérative.

Q4. (a) L'algorithme de construction des $x^{(k)}$ est le suivant :

$$a_{ii}x_i^{(k+1)} = -\left(a_{i1}x_1^{(k)} + \cdots + a_{i,i-1}x_{1-1}^{(k)}\right) - \left(a_{i,i+1}x_{i+1}^{(k)} + \cdots + a_{i,n}x_n^{(k)}\right) + b_i \ (i = 1, \cdots, n)$$

Remarque 1 — Le calcul des composantes de $x^{(k+1)}$ nécessite de garder en mémoire le vecteur $x^{(k)}$. Une itération va donc immobiliser 2*n* cases mémoires.

Remarque 2 — On peut décider d'arrêter les itérations à un rang *MaxIter* donné ou lorsque $\left\|x^{(k+1)} - x^{(k)}\right\| < \varepsilon \cdot \left\|x^{(k+1)}\right\|$ (ε une précision donnée).

Remarque 3 — Comme valeur initiale, on peut prendre $x^{(0)} = \left(\left(\dfrac{b_i}{a_{ii}}\right)\right)_{1 \leq i \leq n}$.

(b) Ce qui donne la programmation structurée qui suit.

PROCEDURE Jacobi(Entrée n : Entier ; A : Matrice ; b : Vecteur ; Sortie x : Vecteur) ;

```
Début
    Pour i allant de 1 à n faire
    Début
        Si a_ii = 0 Alors Stop
        Sinon x_i = b_i/a_ii ;
    Fin ;
    p = 0 ; Continuer = Vrai ;
    Tant que (p < MaxIter) et (Continuer) faire
    Début
        p = p + 1 ;
        Pour i allant de 1 à n faire
        Début
            S = b_i ;
            Pour j allant de 1 à n faire
            Début
                Si j ≠ i Alors S = S – a_ij x_j ;
            Fin ;
            y_i = S/a_ii ;
        Fin ;
        Continuer = ‖y – x‖ > ε·‖y‖ ;
        x = y ;
    Fin ;
    Si Non Continuer
    Alors afficher(« Solution approchée = « ,x)
    Sinon Stop(« convergence trop lente ou divergence ») ;
Fin ;
```

(c) Si $B = D^{-1}N = \left(\left(b_{ij}\right)\right)_{1\le i,j\le n}$, avec $A = D - N$ à diagonale strictement dominante, on a (problème 1, Q6 (a)) $\|B\|_\infty = Max\left\{\sum_{j\ne i}\left|\frac{a_{ij}}{a_{ii}}\right|,\ i = 1,\cdots,n\right\} < 1$ et la méthode de Jacobi est convergente.

Q5. (a) Pour tout $k \ge 0$, $x^{(k+1)}$ est solution du système triangulaire inférieur $Mx^{(k+1)} = Nx^{(k)} + b$. D'où l'algorithme de Gauss-Seidel :

$$a_{ii}x_i^{(k+1)} = -\left(a_{i1}x_1^{(k+1)} + \cdots + a_{i,i-1}x_{1-1}^{(k+1)}\right) - \left(a_{i,i+1}x_{i+1}^{(k)} + \cdots + a_{i,n}x_n^{(k)}\right) + b_i\ \ (i = 1,\cdots,n)$$

Remarque 1 — Cet algorithme est en fait une amélioration de l'algorithme de Jacobi. En effet, dans le calcul de $x_i^{(k+1)}$, on utilise les composantes 1 à $i-1$ de $x^{(k+1)}$ (alors que dans la méthode de Jacobi ce sont celles de $x^{(k)}$ qui sont utilisées) et les composantes $i+1$ à n de $x^{(k)}$ (comme dans la méthode de Jacobi). Cet algorithme sera donc en général plus performant que celui de Jacobi.

Remarque 2 — Pour une itération on ne garde donc que *n* termes en mémoire.

(b) Ce qui donne la programmation structurée suivante.

```
PROCEDURE Gauss-Seidel(Entrée n : Entier ; A : Matrice ; b : Vecteur ;
                         Sortie x : Vecteur) ;
```

```
Début
    Continuer = Vrai ; p = 0 ;
    Pour i allant de 1 à n faire
    Début
        Si a_ii = 0 Alors Stop Sinon x_i = b_i/a_ii ;
    Fin ;
    Tant que (p < MaxIter) et (Continuer) faire
    Début
      p = p + 1 ;
      DeltaMax = 0 ;
      Pour i allant de 1 à n faire
      Début
         S = b_i ;
         Pour j allant de 1 à n faire
         Début
            Si j ≠ i Alors S = S – a_ij x_j ;
         Fin ;
         S = S/a_ii ;
         Si |x_i – S| > DeltaMax
         Alors DeltaMax = |x_i – S| ;
         x_i = S ;
     Fin ;
     Continuer = DeltaMax > Epsilon ;
  Fin ;
  Si Non Continuer
  Alors Afficher(« Solution approchée »,x)
  Sinon Stop(« convergence trop lente ou divergence ») ;
Fin ;
```

(c) Soit A une matrice à diagonale strictement dominante, $B = M^{-1}N$, μ une valeur propre de B et x un vecteur propre non nul associé. On a alors : $Nx = \mu \cdot Mx$, c'est-à-dire :

$$\begin{cases} -\sum_{j=i+1}^{n} a_{ij}x_j = \mu \sum_{j=1}^{i} a_{ij}x_j \ (i = 1, \cdots, n-1) \\ 0 = \sum_{j=1}^{n} a_{nj}x_j \end{cases}$$

Si on suppose que $|\mu| \geq 1$, en prenant i tel que $\|x\|_\infty = |x_i|$, on déduit que : $|\mu| \cdot |a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \cdot |\mu|$, ce qui contredit le fait que A soit à diagonale strictement dominante. Le rayon spectral de B est donc strictement inférieur à 1 et la méthode converge.

(d) Soit A symétrique définie positive, μ une valeur propre de $B = M^{-1}N$ et x un vecteur propre non nul associé. En écrivant $A = D + E + F$, où D est la diagonale de A, E le triangle inférieur strict et F le triangle supérieur strict, on a :

$$\langle x \mid Ax \rangle = \langle x \mid Dx \rangle + \langle x \mid Ex \rangle + \langle x \mid Fx \rangle \quad (1)$$

Puis avec $Fx = -\mu(Dx + Ex)$ ($M = D + E$, $N = -F$), on déduit que $\langle x | Ax \rangle = (1-\overline{\mu})\langle x | Dx \rangle + (1-\overline{\mu})\langle x | Ex \rangle$, donc $\mu \neq 1$ $(\langle Ax | x \rangle > 0)$ et :

$$\frac{1}{1-\overline{\mu}}\langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Ex \rangle \quad (2)$$

Par conjugaison complexe de (2), en considérant que A et D sont hermitiennes et que ${}^t\overline{E} = F$ on déduit que :

$$\frac{1}{1-\mu}\langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Fx \rangle \quad (3)$$

En faisant (2) + (3) – (1), on en conclut que :

$$\frac{1-|\mu|^2}{|1-\mu|^2}\langle x | Ax \rangle = \langle x | Dx \rangle$$

Avec $\langle x | Ax \rangle > 0$ et $\langle x | Dx \rangle > 0$ (A et D sont définies positives) on déduit que $|\mu| < 1$. Le rayon spectral de B est donc strictement inférieur à 1 et la méthode de Gauss-Seidel converge.

Remarque — Quand les méthodes de Jacobi et de Gauss-Seidel sont convergentes il vaut mieux choisir celle de Gauss-Seidel. Mais il se peut que la méthode de Gauss-Seidel diverge alors que celle de Jacobi converge comme le montre l'exemple qui suit.

(e) Pour la méthode Jacobi toutes les valeurs propres de B sont nulles, donc la méthode de Jacobi converge.

Et pour la méthode de Gauss-Seidel, on a $B = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & -2 & 4 \end{pmatrix}$ avec les valeurs propres 0, 5.64 et 0.35, donc la méthode de Gauss-Seidel diverge.

$$\Delta(a,b,c)=\frac{c(b-a)^n-a(b-c)^n}{c-a}$$

Q8. (a) Le polynôme caractéristique de G est :

$$P_G(\lambda)=Dét\left(M^{-1}N-\lambda\cdot I_d\right)=Dét\left(M^{-1}\right)(-1)^n Dét\left(M\lambda-N\right)$$

Soit :

$$P_G(\lambda)=\frac{(-1)^n}{\beta^n}\Delta(\lambda\cdot\alpha,\lambda\cdot\beta,\alpha)=\frac{(-1)^n}{\beta^n}\frac{(\beta-\alpha)^n\lambda^n-\lambda(\lambda\cdot\beta-\alpha)^n}{1-\lambda}$$

Ce qui peut aussi s'écrire :

$$P_G(\lambda)=(-1)^n\frac{(\lambda-1)\left(\lambda-\dfrac{\alpha}{\beta}\right)^n+\left(\lambda-\dfrac{\alpha}{\beta}\right)^n-\left(1-\dfrac{\alpha}{\beta}\right)^n\lambda^n}{\lambda-1}$$

Soit :

$$P_G(\lambda)=(-1)^n\left\{\left(\lambda-\frac{\alpha}{\beta}\right)^n+\frac{\alpha}{\beta}\sum_{k=1}^{n}\left(\lambda-\frac{\alpha}{\beta}\right)^{n-k}\left(1-\frac{\alpha}{\beta}\right)^{k-1}\lambda^{k-1}\right\}$$

(b) On en déduit que 0 est valeur propre de G et le produit des modules valeurs propres non nulles est :

$$\begin{aligned}\left|\prod_{i=2}^{n}\lambda_i\right|&=\left|\text{Coefficient de }\lambda\text{ dans }P_G(\lambda)\right|\\&=\left|n\left(-\frac{\alpha}{\beta}\right)^{n-1}+\frac{\alpha}{\beta}\left\{(n-1)\left(-\frac{\alpha}{\beta}\right)^{n-2}+\left(-\frac{\alpha}{\beta}\right)^{n-2}\left(1-\frac{\alpha}{\beta}\right)\right\}\right|\end{aligned}$$

Soit :

$$\left|\prod_{i=2}^{n}\lambda_i\right|=\left|\frac{\alpha}{\beta}\right|^n$$

(c) Si $|\alpha|\geq|\beta|$ alors $\left|\prod_{i=2}^{n}\lambda_i\right|\geq 1$ et $\rho(G)\geq 1$, donc la méthode de Gauss-Seidel ne converge pas.

Remarque — Si $|\beta|>(n-1)|\alpha|$ on a vu alors la méthode de Gauss-Seidel converge.

Il reste à étudier le cas où $\dfrac{1}{n-1}|\beta|\leq|\alpha|<|\beta|$.

Problème 16 : Méthode de relaxation

Ce problème utilise les résultats du problème 14.

A désigne une matrice réelle inversible dont tous les coefficients diagonaux de A sont non nuls.

On écrit la matrice sous la forme $A=D+E+F$*, où D est la diagonale de A, E le triangle inférieur strict et F le triangle supérieur strict.*

En vue d'accélérer la convergence de la méthode de Gauss-Seidel (problème 14, Q5), on introduit un paramètre dans la matrice $G = -(D+E)^{-1}F$. *C'est-à-dire qu'on considère la méthode itérative définie par le choix de* $M_\omega = \frac{1}{\omega}D + E$ *où* ω *est non nul à préciser.*

La méthode obtenue est appelée « méthode de relaxation ».

Q1. *Décrire l'algorithme de construction des* $x^{(k)}$.

Q2. *Ecrire la programmation structurée correspondante.*

Q3. *Montrer que la méthode de relaxation ne peut converger que si* $\omega \in \left]0,2\right[$.

Q4. *On suppose que la matrice A est symétrique définie positive.*

(a) Soit $\lambda \in C$ *une valeur propre de* $L_\omega = M_\omega^{-1}N_\omega$ *et* $x \in C^n - \{0\}$ *un vecteur propre associé.*

Montrer que :

$$\frac{\left(1-|\lambda|^2\right)}{|1-\lambda|^2}\langle x \mid Ax\rangle = \left(\frac{2}{\omega}-1\right)\langle x \mid Dx\rangle$$

(b) En déduire qu'une condition nécessaire et suffisante de convergence de la méthode de relaxation, pour une matrice symétrique définie positive, est $\omega \in \left]0,2\right[$.

Solution

Q1. On note $L_\omega = M_\omega^{-1}N_\omega$ avec $A = M_\omega - N_\omega$ et on a alors :

$$L_\omega = (D+\omega\cdot E)^{-1}\left((1-\omega)D - \omega\cdot F\right)$$

Le vecteur $x^{(k+1)}$ est alors solution du système triangulaire :

$$(D+\omega\cdot E)x^{(k+1)} = \left((1-\omega)D-\omega\cdot F\right)x^{(k)} + \omega\cdot b$$

Ce qui donne l'algorithme de calcul :

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega\cdot \hat{x}_i^{(k+1)} \quad (i = 1,\cdots,n)$$

où $\hat{x}_i^{(k+1)}$ est donné par les formules de Gauss–seidel en fonction des composantes 1 à $i-1$ de $x^{(k+1)}$ et des composantes $i+1$ à n de $x^{(k)}$, soit :

$$a_{ii}\hat{x}_i^{(k+1)} = -\left(a_{i1}x_1^{(k+1)} + \cdots + a_{i,i-1}x_{1-1}^{(k+1)}\right) - \left(a_{i,i+1}x_{i+1}^{(k)} + \cdots + a_{i,n}x_n^{(k)}\right) + b_i \quad (i = 1,\cdots,n)$$

Ce qui peut encore s'écrire :

$$x^{(k+1)} = x^{(k)} - \omega\cdot \xi^{(k+1)} \quad (k \geq 0)$$

où $\xi^{(k+1)}$ est le *« vecteur résidu »* défini par :

$$\xi_i^{(k+1)} = \frac{\left(\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i}^{n} a_{ij}x_j^{(k)} - b_i\right)}{a_{ii}} \quad (i = 1,\cdots,n)$$

Un test de convergence sera alors $\|\xi\| < \varepsilon$, où ε est une précision donnée.

Q2. La programmation structurée est alors la suivante :

PROCEDURE Relaxation(Entrée n : Entier ; A : Matrice ; b : Vecteur ; ω *: Réel ;*

```
                    Sortie : x : Vecteur) ;
Début
   p = 0 ;
   Continuer = Vrai ;
   Pour i Allant de 1 à n Faire
   Début
      Si a_ii = 0 Alors Stop Sinon x_i = b_i/a_ii ;
   Fin ;
   Tant que (p < MaxIter) et (Continuer) Faire
   Début
      p = p + 1 ;
      NormeRésidu = 0 ;
      Pour i Allant de 1 à n Faire
      Début
         Résidu = –b_i ;
         Pour j Allant de 1 à n Faire
           Résidu = Résidu + a_ij x_j ;
         Résidu = Résidu/a_ii ;
         x_i = x_i – ω·Résidu ;
         Si |Résidu| > NormeRésidu
         Alors NormeRésidu = |Résidu| ;
      Fin ;
      Continuer = (NormeRésidu > ε) ;
   Fin ;
   Si (p = MaxIter)
   Alors Stop(« convergence trop lente ou divergence ») ;
Fin ;
```

Q3. Si $\lambda_1, \cdots, \lambda_n$ sont les valeurs propres de L_ω, on a alors :

$$\prod_{i=1}^{n} \lambda_i = Dét(L_\omega) = \frac{Dét((1-\omega)D - \omega \cdot F)}{Dét(D + \omega \cdot E)} = (1-\omega)^n$$

Et pour $\omega \notin]0,2[$ on a $\rho(L_\omega) \geq |\omega - 1| > 1$, ce qui entraîne la divergence de la méthode itérative.

Q4. La matrice A symétrique définie positive est inversible avec $a_{ii} = \langle Ae_i | e_i \rangle > 0$ pour tout $i = 1, 2, \cdots, n$. On peut donc définir la suite $\left(x^{(k)}\right)_{k \in IN}$ de la méthode de relaxation.

(a) Soit $\lambda \in C$ une valeur propre de $L_\omega = M_\omega^{-1} N_\omega$ et $x \in C^n - \{0\}$ un vecteur propre associé.

On a alors $N_\omega x = \lambda \cdot M_\omega x$ et :

$$\langle x | Ax \rangle = \langle x | M_\omega x \rangle - \langle x | N_\omega x \rangle = (1 - \overline{\lambda}) \langle x | M_\omega x \rangle$$

La matrice A étant symétrique définie positive, on en déduit que $\lambda \neq 1$ et :

Q6. (a) On a $M(t)=P(t)MP(t)^{-1}$, avec $P(t)=\begin{pmatrix} t & 0 & 0 & 0 & \cdots & 0 \\ 0 & t^2 & 0 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 0 & t^{n-1} & 0 \\ 0 & \cdots & 0 & 0 & 0 & t^n \end{pmatrix}$.

On peut aussi utiliser le développement du déterminant :

$Dét(M(t))=\sum_{\sigma\in S_n}(-1)^{Sign(\sigma)}\prod_{i=1}^{n}m_{i,\sigma(i)}(t)$ avec $\prod_{i=1}^{n}m_{i,\sigma(i)}(t)=\prod_{i=1}^{n}t^{i-\sigma(i)}\prod_{i=1}^{n}m_{i,\sigma(i)}$ et $\prod_{i=1}^{n}t^{\sigma(i)}=\prod_{i=1}^{n}t^{i}=t^{\frac{n(n+1)}{2}}$. D'où le résultat.

(b) Le polynôme caractéristique de J est :

$$P_J(\lambda)=Dét\left(-D^{-1}(E+F)-\lambda\cdot I_d\right)=(-1)^n Dét\left(D^{-1}\right)Dét(E+F+\lambda\cdot D)$$

et celui de G :

$$P_G(\lambda)=Dét\left(-(D+E)^{-1}F-\lambda\cdot I_d\right)=(-1)^n Dét\left((D+E)^{-1}\right)Dét\left(F+\lambda(D+E)\right)$$

avec $M=F+\lambda(D+E)$ tridiagonale. Avec les notations de (a), on a pour $\lambda>0$:

$$M\left(\frac{1}{\sqrt{\lambda}}\right)=\sqrt{\lambda}(E+F)+\lambda\cdot D=\sqrt{\lambda}\left(E+F+\sqrt{\lambda}D\right)$$

et $\text{Dét}(M)=\text{Dét}\left(M\left(\frac{1}{\sqrt{\lambda}}\right)\right)=\left(\sqrt{\lambda}\right)^n\text{Dét}\left(E+F+\sqrt{\lambda}D\right)$. On a donc :

$$P_G\left(\lambda^2\right)=\frac{(-1)^n}{Dét(D+E)}\lambda^n Dét(E+F+\lambda\cdot D)=\frac{\lambda^n}{Dét(D+E)}P_J(\lambda)Dét(D)=\lambda^n P_J(\lambda).$$

Il en résulte immédiatement que $\rho(G)=\rho(J)^2$.

Remarque — On en déduit que $P_J(-\lambda)=(-1)^n P_J(\lambda)$ et $P_J(\lambda)=P\left(\lambda^2\right)\lambda^q$ où P est un polynôme de degré p avec $P(0)\neq 0$ et $2p+q=n$.

Il résulte de ce qui précède que les deux méthodes convergent ou divergent simultanément, et dans le cas de la convergence, c'est la méthode de Gauss-Seidel qui est la plus rapide car, dans ce cas, on a $\rho(G)<\rho(J)<1$.

Problème 15 : Méthodes itératives sur une matrice particulière

Pour tout le problème, n désigne un entier supérieur ou égal à 3, α *et* β *sont des réels et* $A(\alpha,\beta)=\left((a_{ij})\right)_{1\le i,j\le n}$ *est la matrice réelle d'ordre n définie par :*

$$\forall i\in\{1,2,\ldots,n\},\ \begin{cases} a_{ii}=\beta \\ a_{ij}=\alpha\,\text{si}\,j\in\{1,2,\cdots,n\}-\{i\} \end{cases}$$

On note I_d *la matrice identité d'ordre n.*

Q1. *Calculer* $\Delta(\alpha,\beta)=\text{Dét}(A(\alpha,\beta))$.

Q2. *Calculer le polynôme caractéristique de* $A(\alpha,\beta)$ *:*

$$P_{(\alpha,\beta)}(\lambda) = \text{Dét}\left(A(\alpha,\beta) - \lambda \cdot I_d\right)$$

Q3. *Calculer le rayon spectral* $\rho(\alpha,\beta)$ *de* $A(\alpha,\beta)$.

Q4. *Pour cette question seulement on suppose que* $\alpha \cdot \beta \geq 0$.

Comparer $\|A(\alpha,\beta)\|_\infty$, $\|A(\alpha,\beta)\|_1$ *et* $\rho(\alpha,\beta)$.

Q5. *On suppose que* $\beta \neq 0$ *et on pose* $D = \beta \cdot I_d$ *et* $N = D - A(\alpha,\beta)$.

(a) Calculer la matrice $J = D^{-1}N$ *intervenant dans la méthode de Jacobi.*

(b) Calculer le rayon spectral de J.

(c) Donner une condition nécessaire est suffisante sur la matrice $A(\alpha,\beta)$ *pour que la méthode de Jacobi converge.*

(d) Montrer que si la méthode de Jacobi converge alors celle de Gauss-Seidel converge aussi.

Q6. *On suppose que* $A(\alpha,\beta)$ *est inversible. Résoudre le système* $A(\alpha,\beta)x = e$ *dans les deux cas suivants :*

(a) Toutes les composantes de e valent 1.

(b) e est un vecteur quelconque.

Q7. *Pour tous réels a, b et c, on désigne par* $M(a,b,c) = \left((m_{ij})\right)_{1\leq i,j\leq n}$ *la matrice réelle d'ordre n définie par :*

$$\forall i \in \{1,2,\cdots,n\},\ \begin{cases} m_{ii} = b \\ m_{ij} = c \text{ si } j \in \{i+1,\cdots,n\} \\ m_{ij} = a \text{ si } j \in \{1,\cdots,i-1\} \end{cases}$$

Montrer que, pour $a \neq c$, *le déterminant de* $M(a,b,c)$ *est :*

$$\Delta(a,b,c) = \frac{c(b-a)^n - a(b-c)^n}{c-a}$$

Q8. *On suppose que* $\beta \neq 0$ *et on pose par* $M = M(\alpha,\beta,0)$, $N = M - A(\alpha,\beta)$ *et* $G = M^{-1}N$ *désigne la matrice qui intervient dans la méthode de Gauss-Seidel.*

(a) Calculer le polynôme caractéristique de G.

(b) Montrer que 0 est valeur propre de G et calculer le produit des modules valeurs propres non nulles.

(c) Montrer que si $|\alpha| \geq |\beta|$ *alors la méthode de Gauss-Seidel ne converge pas.*

Solution

La matrice $A(\alpha,\beta)$ est de la forme :

$$A(\alpha,\beta)=\begin{pmatrix}\beta & \alpha & \alpha & \cdots & \alpha\\ \alpha & \beta & \alpha & \cdots & \alpha\\ . & . & . & . & .\\ \alpha & \cdots & \alpha & \beta & \alpha\\ \alpha & \cdots & \alpha & \alpha & \beta\end{pmatrix}$$

Q1. En ajoutant les lignes 2 à n à la première ligne on a :

$$\Delta(\alpha,\beta)=\beta+(n-1)\alpha\begin{vmatrix}1 & 1 & 1 & \cdots & 1\\ \alpha & \beta & \alpha & \cdots & \alpha\\ . & . & . & . & .\\ \alpha & \cdots & \alpha & \beta & \alpha\\ \alpha & \cdots & \alpha & \alpha & \beta\end{vmatrix}$$

Puis en retranchant la première colonne aux colonnes 2 à n on obtient :

$$\Delta(\alpha,\beta)=(\beta+(n-1)\alpha)(\beta-\alpha)^{n-1}$$

Q2. Le polynôme caractéristique de $A(\alpha,\beta)$ est donné par :

$$P_{(\alpha,\beta)}(\lambda)=\Delta(\alpha,\beta-\lambda)=(-1)^n\big(\lambda-(\beta+(n-1)\alpha)\big)\big(\lambda-(\beta-\alpha)\big)^{n-1}$$

Q3. Le rayon spectral de $A(\alpha,\beta)$ est donc :

$$\rho(\alpha,\beta)=Max\{|\beta+(n-1)\alpha|,\ |\beta-\alpha|\}$$

Ce qui donne quatre possibilités.

Premier cas — $\alpha\leq\beta,\ \beta>-(n-1)\alpha$

Alors

$$\rho(\alpha,\beta)=Max\{\beta+(n-1)\alpha,\ \beta-\alpha\}=\begin{cases}\beta+(n-1)\alpha \text{ si } \alpha>0\\ \beta-\alpha \text{ si } \alpha\leq 0\end{cases}$$

Deuxième cas — $\alpha\leq\beta,\ \beta\leq-(n-1)\alpha$

Alors

$$\rho(\alpha,\beta)=Max\{-\beta-(n-1)\alpha,\ \beta-\alpha\}=\begin{cases}-\beta-(n-1)\alpha \text{ si } \beta<-\dfrac{n-2}{2}\alpha\\ \beta-\alpha \text{ si } \beta\geq-\dfrac{n-2}{2}\alpha\end{cases}$$

Troisième cas — $\alpha>\beta,\ \beta>-(n-1)\alpha$

Alors

$$\rho(\alpha,\beta)=Max\{\beta+(n-1)\alpha,\ \alpha-\beta\}=\begin{cases}\beta+(n-1)\alpha \text{ si } \beta>-\dfrac{n-2}{2}\alpha\\ \alpha-\beta \text{ si } \beta\leq-\dfrac{n-2}{2}\alpha\end{cases}$$

Quatrième cas — $\alpha>\beta,\ \beta\leq-(n-1)\alpha$

Alors

$$\rho(\alpha,\beta) = Max\{-\beta-(n-1)\alpha,\ \alpha-\beta\} = \begin{cases} -\beta-(n-1)\alpha \text{ si } \alpha<0 \\ \alpha-\beta \text{ si } \alpha\geq 0 \end{cases}$$

Ce qui peut se traduire par le graphique de la figure 2.1.

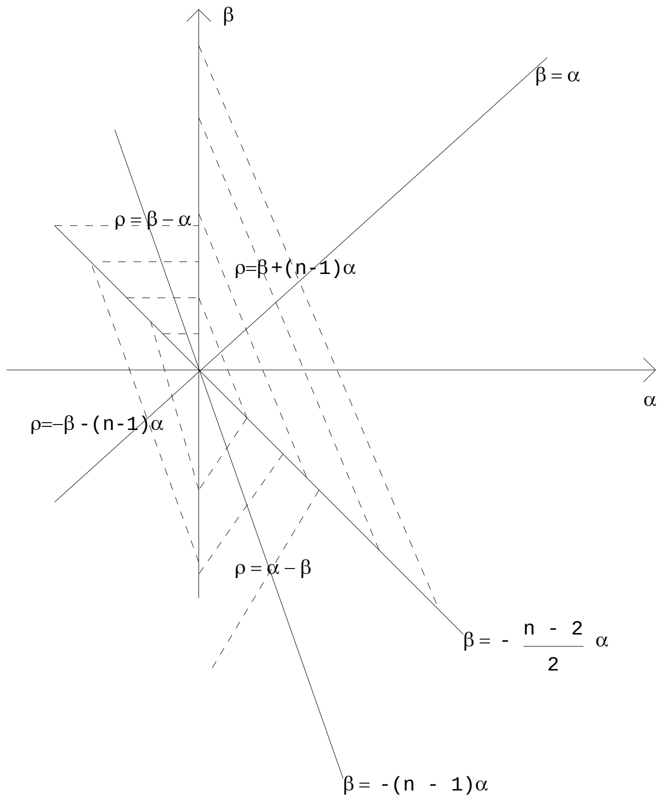

Figure 2.1

Q4. Si $\alpha\cdot\beta\geq 0$, on a :

$$\rho(\alpha,\beta) = |\beta| + (n-1)|\alpha| = \|A(\alpha,\beta)\|_\infty = \|A(\alpha,\beta)\|_1$$

Q5. (a) $J = A\left(-\dfrac{\alpha}{\beta},0\right)$

(b) $\rho(J) = \rho\left(A\left(-\dfrac{\alpha}{\beta},0\right)\right) = (n-1)\dfrac{|\alpha|}{|\beta|}$

(c) La méthode de Jacobi est convergente si et seulement si le rayon spectral de J est strictement inférieur à 1, ce qui équivaut à $|\beta| > (n-1)|\alpha|$ encore équivalent à dire que la matrice $A(\alpha,\beta)$ est à diagonale strictement dominante.

(d) Si la méthode de Jacobi converge alors $A(\alpha,\beta)$ est à diagonale strictement dominante et la méthode de Gauss-Seidel converge aussi.

Q6. (a) Le système est de la forme :

$$\begin{cases} \beta x_1 + \alpha x_2 + \alpha x_3 + \cdots + \alpha x_n = 1 \\ \alpha x_1 + \beta x_2 + \alpha x_3 + \cdots + \alpha x_n = 1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha x_1 + \alpha x_2 + \cdots + \beta x_{n-1} + \alpha x_n = 1 \\ \alpha x_1 + \alpha x_2 + \cdots + \alpha x_{n-1} + \beta x_n = 1 \end{cases}$$

En ajoutant les lignes 2 à n à la première équation, on obtient :

$$\alpha \sum_{j=1}^{n} x_j = \frac{n \cdot \alpha}{\beta + (n-1)\alpha}$$

Puis en retranchant cette équation aux équations 1 à n, on obtient :

$$\forall i \in \{1, 2, \cdots, n\},\ x_i = \frac{1}{\beta + (n-1)\alpha}$$

(b) Avec les mêmes opérations qu'en (a), on aboutit à :

$$\alpha \sum_{j=1}^{n} x_j = \frac{\alpha \cdot S(e)}{\beta + (n-1)\alpha}$$

où $S(e) = \sum_{i=1}^{n} e_i$. Puis :

$$\forall i \in \{1, 2, \cdots, n\},\ x_i = \frac{1}{\beta - \alpha}\left\{ e_i - \frac{\alpha \cdot S(e)}{\beta + (n-1)\alpha} \right\}$$

Q7. La matrice $M(a,b,c)$ est de la forme :

$$M(a,b,c) = \begin{pmatrix} b & c & c & \cdots & c \\ a & b & a & \cdots & c \\ . & . & . & . & . \\ a & \cdots & a & b & c \\ a & \cdots & a & a & b \end{pmatrix}$$

(a) Soit $A = ((a_{ij}))_{1 \le i,j \le n}$ et pour tout réel t, $A(t) = ((a_{ij} + t))_{1 \le i,j \le n}$. En notant C_j la colonne numéro j de A, T le vecteur de composantes toutes égales à t et en utilisant la n–linéarité du déterminant on déduit que :

$$Dét(A(t)) = Dét(C_1 + T, \cdots, C_n + T) = Dét(A) + t \cdot S(A)$$

où on a noté :

$$S(A) = \sum_{j=1}^{n} Dét(C_1, \cdots, C_{j-1}, E, C_j, \cdots, C_n)$$

E désignant le vecteur de composantes toutes égales à 1.

En prenant $A = M(a,b,c)$, on a :

$$\text{Dét}(A(-a)) = (b-a)^n = \text{Dét}(A) - a \cdot S(A)$$

$$\text{Dét}(A(-c)) = (b-c)^n = \text{Dét}(A) - c \cdot S(A)$$

Et on en déduit que :

$$\frac{1}{1-\overline{\lambda}}\langle x\,|\,Ax\rangle=\frac{1}{\omega}\langle x\,|\,Dx\rangle+\langle x\,|\,Ex\rangle\quad(1)$$

Par conjugaison complexe, en considérant que A et D sont hermitiennes et que ${}^t\overline{E}=F$ on déduit que :

$$\frac{1}{1-\lambda}\langle x\,|\,Ax\rangle=\frac{1}{\omega}\langle x\,|\,Dx\rangle+\langle x\,|\,Fx\rangle\quad(2)$$

Enfin, avec $A=D+E+F$, on peut aussi écrire :

$$\langle x\,|\,Ax\rangle=\langle x\,|\,Dx\rangle+\langle x\,|\,Ex\rangle+\langle x\,|\,Fx\rangle\quad(3)$$

En faisant (1) + (2) – (3), on en conclut que :

$$\frac{\left(1-|\lambda|^2\right)}{|1-\lambda|^2}\langle x\,|\,Ax\,\rangle=\left(\frac{2}{\omega}-1\right)\langle x\,|\,Dx\rangle$$

(b) Avec Q3, il nous suffit de montrer que la méthode converge si $\omega\in\left]0,2\right[$.

En tenant compte de la positivité de A et de D, on déduit du (a) que si $\omega\in\left]0,2\right[$ alors toute valeur propre λ de L_ω vérifie $|\lambda|<1$. C'est-à-dire que $\rho(L_\omega)<1$ et que la méthode converge.

Remarque — Pour $\omega=1$, on retrouve le fait que la méthode de Gauss-Seidel converge pour les matrices symétriques définies positives.

Problème 17 : Méthode de relaxation pour les matrices tridiagonales symétriques définies positives

Ce problème utilise les notations et résultats des problèmes 14 et 16.

On suppose dans ce problème que A est tridiagonale symétrique définie positive et on note $A=D+E+F$ où D est la diagonale de A, E le triangle inférieur strict et F le triangle supérieur strict.

On pose $J=-D^{-1}(E+F)$ (méthode de Jacobi, problème 14, Q4) et $L_\omega=\left(\frac{1}{\omega}D+E\right)^{-1}\left(\left(\frac{1}{\omega}-1\right)D-F\right)$ (méthode de relaxation, problème 16).

On désigne par P_J le polynôme caractéristique de J et par P_{L_ω} celui de L_ω.

Q1. *Montrer que pour ω et λ réels non nuls, on a :*

$$P_{L_\omega}\left(\lambda^2\right)=\omega^n\lambda^n P_J\left(\frac{\lambda^2+\omega-1}{\omega\cdot\lambda}\right)$$

On utilisera le résultat du problème 14, Q6 (a).

Q2. *Montrer que :*

$$P_J(\lambda)=(-1)^n\lambda^q\prod_{k=1}^{p}\left(\lambda^2-{\mu_k}^2\right)$$

avec $0<\mu_1\le\mu_2\le\cdots\le\mu_p<1$ dans $I\!R$ et $2p+q=n$.

Q3. *(a) Calculer le polynôme caractéristique de L_ω.*

(b) En déduire les valeurs propres de L_ω.

Q4. *Pour $\mu\in\left]0,1\right[$ et $\omega\in\left]0,2\right[$, on considère l'équation en λ :*

$$(\lambda+\omega-1)^2-\omega^2\mu^2\lambda=0 \quad (E)$$

Montrer qu'il existe un réel $\omega_0(\mu)\in\left]1,2\right[$ *tel que pour* $\omega\in\left]0,\omega_0(\mu)\right[$ *l'équation (E) admet deux racines réelles distinctes, pour* $\omega=\omega_0(\mu)$ *elle admet une racine double et pour* $\omega\in\left]\omega_0(\mu),2\right[$ *elle admet deux racines complexes conjuguées.*

On notera $\lambda_1(\mu,\omega)$ *et* $\lambda_2(\mu,\omega)$ *ces deux racines.*

Calculer $|\lambda_1(\mu,\omega)|$ *et* $|\lambda_2(\mu,\omega)|$ *pour* $\omega\in\left]\omega_0(\mu),2\right[$.

Q5. *(a) Montrer que la méthode de relaxation est convergente si et seulement si* $\omega\in\left]0,2\right[$.

(b) Tracer le graphe de la fonction $\omega\mapsto\rho(L_\omega)$*, pour* $\omega\in\left]0,2\right[$.

(c) Quelle est la valeur optimale du paramètre de relaxation $\omega\in\left]0,2\right[$ *?*

Solution

Q1. On a :

$$P_{L_\omega}(\lambda)=\frac{1}{Dét\left(\left(\frac{1}{\omega}D+E\right)\right)}Dét\left(\left(\frac{1}{\omega}-1\right)D-F-\lambda\left(\frac{1}{\omega}D+E\right)\right)$$

Soit $P_{L_\omega}(\lambda)=\dfrac{\omega^n}{Dét(D)}Dét(M)$, où M est la matrice tridiagonale :

$$M=\frac{1-\omega-\lambda}{\omega}D-\lambda\cdot E-F$$

Avec les notations du problème 14, Q6 (a), on peut écrire, pour $\lambda>0$, que :

$$Dét(M)=Dét\left(M\left(\frac{1}{\sqrt{\lambda}}\right)\right)$$

avec :

$$M\left(\frac{1}{\sqrt{\lambda}}\right)=\sqrt{\lambda}\left(\frac{1-\omega-\lambda}{\omega\sqrt{\lambda}}D-E-F\right)$$

Ce qui donne, en tenant compte de $P_J(\lambda)=\dfrac{(-1)^n}{\text{Dét}(D)}Dét(E+F+\lambda\cdot D)$:

$$P_{L_\omega}(\lambda)=\omega^n\sqrt{\lambda}^nP_J\left(\frac{\lambda+\omega-1}{\omega\sqrt{\lambda}}\right)$$

Q2. La forme du polynôme caractéristique de J résulte du problème 14, Q6 (b). Il reste à montrer que toutes les valeurs propres de J sont réelles. Si $\mu\in C$ est une valeur propre de J et $x\in C^n-\{0\}$ est un vecteur propre associé, on a alors $(1-\mu)Dx=Ax$ et $(1-\mu)\langle Dx|x\rangle=\langle Ax|x\rangle$. Les matrices A et D étant symétriques définies positives, on déduit que μ est réelle et $\mu<1$. Comme $-\mu$ est aussi valeur propre de J on retrouve le fait que $|\mu|<1$.

Q3. (a) De Q1 et Q2, on déduit que le polynôme caractéristique de L_ω s'écrit :

$$P_{L_\omega}(\lambda) = (-1)^n (\lambda + \omega - 1)^q \prod_{k=1}^{p} \left((\lambda + \omega - 1)^2 - \omega^2 \mu_k{}^2 \lambda \right)$$

(b) On en déduit que les valeurs propres de L_ω sont $1-\omega$ d'ordre q et $\lambda_1(\mu_k, \omega)$, $\lambda_2(\mu_k, \omega)$ pour $k = 1, \cdots, p$, où les $\lambda_i(\mu_k, \omega)$ ($i = 1, 2$) sont les racines de :

$$(\lambda + \omega - 1)^2 - \omega^2 \mu_k{}^2 \lambda = 0 \quad (E)$$

Q4. A $\mu \in]0,1[$ fixé, le discriminant de (E) $\Delta(\mu, \omega) = \omega^2 \mu^2 (\omega^2 \mu^2 - 4\omega + 4)$ a une seule racine dans $]0,2[$, $\omega_0(\mu) = \dfrac{2}{1 + \sqrt{1-\mu^2}}$.

Pour $\omega \in]0, \omega_0(\mu)[$ l'équation (E) admet deux racines réelles distinctes :

$$\lambda_1(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) - \omega\mu\sqrt{\omega^2 \mu^2 - 4\omega + 4}}{2}$$

$$\lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) + \omega\mu\sqrt{\omega^2 \mu^2 - 4\omega + 4}}{2} = \frac{(\omega - 1)^2}{\lambda_1(\mu, \omega)}$$

pour $\omega = \omega_0(\mu)$ elle admet une racine double :

$$\lambda_1(\mu, \omega) = \lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2}{2} - (\omega - 1) = \omega_0(\mu) - 1$$

et pour $\omega \in]\omega_0(\mu), 2[$ elle admet deux racines complexes conjuguées :

$$\lambda_1(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) - i\omega\mu\sqrt{-\omega^2 \mu^2 + 4\omega - 4}}{2}$$

$$\lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) + i\omega\mu\sqrt{-\omega^2 \mu^2 + 4\omega - 4}}{2} = \frac{(\omega - 1)^2}{\lambda_1(\mu, \omega)}$$

avec $|\lambda_1(\mu, \omega)| = |\lambda_2(\mu, \omega)| = \omega - 1$.

Q5. (a) C'est un cas particulier du résultat du problème 16, Q4 (b).

(b) On pose $r = \rho(J) \in [0,1[$.

Si $r = 0$, alors toutes les valeurs propres de J sont nulles et $1-\omega$ est la seule valeur propre de L_ω d'après Q3 (a). On a donc $\rho(L_\omega) = |1-\omega|$ (figure 2.2).

Si $r \in]0,1[$, d'après Q4, on a pour $\omega \in]\omega_0(\mu), 2[\subset]1,2[$, $\rho(L_\omega) = |1-\omega| = 1-\omega$. Si $\omega \in]0, \omega_0(\mu)[$, en remarquant que la fonction $\mu \mapsto \lambda_1(\mu, \omega)$ est décroissante et $\mu \mapsto \lambda_2(\mu, \omega)$ croissante sur $I\!R^+$, avec Q4 on déduit que les valeurs propres de L_ω sont toutes positives ou nulles ($\lambda_1(\mu, \omega)\lambda_2(\mu, \omega) = (\omega - 1)^2 > 0$) et vérifient :

$$0 < \lambda_1(\mu_p, \omega) \le \cdots \le \lambda_1(\mu_1, \omega) < 1 - \omega < \lambda_2(\mu_1, \omega) \le \cdots \le \lambda_2(\mu_p, \omega)$$

et $\rho(L_\omega) = \lambda_2(\mu_p, \omega) = \lambda_2(r, \omega)$.

Ce qui donne, en posant $\omega_0(r) = \dfrac{2}{1 + \sqrt{1 - r^2}}$:

$$\rho(L_\omega) = \begin{cases} \dfrac{\omega^2 r^2 - 2(\omega - 1) + \omega r \sqrt{\omega^2 r^2 - 4\omega + 4}}{2} & \text{si } 0 < \omega < \omega_0(r) \\ \omega - 1 \text{ si } \omega_0(r) \le \omega < 2 \end{cases}$$

C'est-à-dire que la fonction $\omega \mapsto \rho(L_\omega)$ a l'allure indiquée par la figure 2.3.

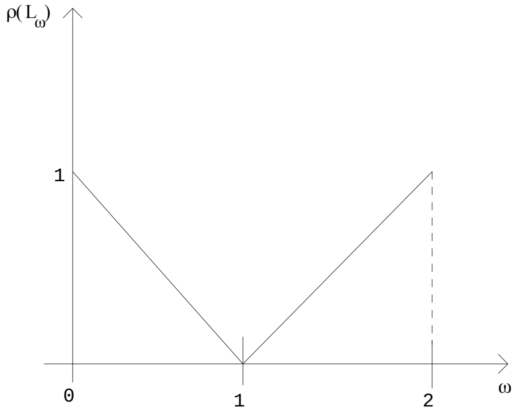

Figure 2.2

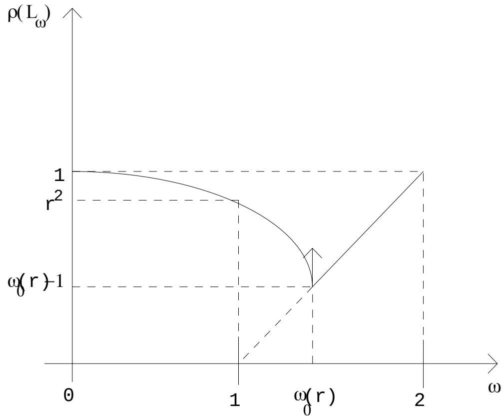

Figure 2.3

(c) La valeur optimale du paramètre de relaxation est $\omega_0(r)$ avec $\rho\left(L_{\omega_0(r)}\right) = \omega_0(r) - 1$.

CHAPITRE 3

Valeurs et vecteurs propres

Problème 18 : Méthode de la puissance itérée et de déflation

Q1. *Soit* $\alpha \in C^n$ *tel que* $\left|\sum_{i=1}^{n} \alpha_i\right| = \sum_{i=1}^{n} |\alpha_i|$. *Montrer qu'il existe* $u \in C$ *et* $\beta \in IR^n$ *tels que :*

$$\begin{cases} |u| = 1 \\ \forall i \in \{1,2,\cdots,n\},\ \beta_i \geq 0 \\ \alpha = u \cdot \beta \end{cases}$$

Q2. *Soit A une matrice d'ordre* $n \geq 2$ *à coefficients réels strictement positifs et* λ *une valeur propre telle que* $\rho(A) = |\lambda|$ *où* $\rho(A)$ *désigne le rayon spectral de A.*

(a) Montrer que $\rho(A) > 0$.

(b) Soit $y \in C^n - \{0\}$ *un vecteur propre de A associé à* λ. *Montrer que le vecteur x de composantes* $x_i = |y_i|$ *est aussi vecteur propre de A associé à* λ *avec* $x_i > 0$ *pour tout i = 1, 2, ⋯, n.*

(c) Montrer que λ *est un réel strictement positif et que la valeur propre dominante de A est unique.*

(d) Montrer que l'espace propre E_λ *associé à* λ *est de dimension 1.*

(e) Si λ *est valeur propre de A de multiplicité supérieure ou égale à 2 et si* $x \in (IR_+^*)^n \cap E_\lambda$, *montrer alors qu'on peut trouver* $y \in (IR_+^*)^n$ *tel que* $Ay = x + \lambda \cdot y$

(f) Montrer que λ *est valeur propre simple de A.*

Q3. *Méthode de la puissance itérée — On se donne une matrice A d'ordre* $n \geq 2$ *à coefficients réels telle que la valeur propre dominante soit unique, c'est-à-dire que si* $\lambda_1, \cdots, \lambda_n$ *sont les valeurs propres de A dans C, alors :*

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

(a) Montrer que λ_1 *est réelle et racine simple du polynôme caractéristique de A.*

(b) On note $E_1 = Ker(A - \lambda_1 I_d)$ *l'espace propre associé à* λ_1 *et* $F_1 = \mathrm{Im}(A - \lambda_1 I_d)$. *Montrer que* $Dim(E_1) = 1$, $IR^n = E_1 \oplus F_1$ *et que les espaces vectoriels* E_1 *et* F_1 *sont stables par A.*

(c) On munit IR^n *d'une norme quelconque et on définit la suite* $\left(x^{(k)}\right)_{k\in IN}$ *de vecteurs de* IR^n *par :*

$$\begin{cases} x^{(0)} = e_1 + f_1 \text{ avec } e_1 \in E_1 - \{0\},\ f_1 \in F_1 \\ x^{(k+1)} = \dfrac{1}{\left\| Ax^{(k)} \right\|} Ax^{(k)},\ (k \geq 0) \end{cases}$$

On note, dans la base canonique de IR^n, $e_{1,j}$ *les composantes du vecteur* e_1, $x_j^{(k)}$ *celles de* $x^{(k)}$ *et* $\left(Ax^{(k)}\right)_j$ *celles de* $Ax^{(k)}$.

Montrer que :

(i) $\underset{k\to+\infty}{Lim} \left\| Ax^{(k)} \right\| = |\lambda_1| = \rho(A)$.

(ii) $\underset{k\to+\infty}{Lim}\, x^{(2k)} = v_1$, $\underset{k\to+\infty}{Lim}\, x^{(2k+1)} = v_2 = \mathrm{Signe}(\lambda_1) v_1$ *où* v_1 *est un vecteur propre non nul associé à* λ_1.

(iii) Pour tout $j \in \{1, 2, \cdots, n\}$ *tel que* $e_{1,j} \neq 0$, *on a* $\underset{k\to+\infty}{Lim} \dfrac{\left(Ax^{(k)}\right)_j}{x_j^{(k)}} = \lambda_1$.

Q4. *Méthode de la puissance inverse — On suppose que A est inversible et que ses valeurs propres vérifient :*

$$|\lambda_n| < |\lambda_{n-1}| \leq |\lambda_3| \leq \cdots \leq |\lambda_1|$$

On note $E_n = Ker(A - \lambda_n I_d)$ *l'espace propre associé à* λ_n *et* $F_n = \mathrm{Im}(A - \lambda_n I_d)$ *et on définit la suite* $\left(x^{(k)}\right)_{k\in IN}$ *de vecteurs de* IR^n *par :*

$$\begin{cases} x^{(0)} = e_n + f_n \text{ avec } e_n \in Ker(A - \lambda_n I_d) - \{0\},\ f_n \in \mathrm{Im}(A - \lambda_n I_d) \\ x^{(k+1)} = \dfrac{1}{\left\| u^{(k+1)} \right\|} u^{(k+1)},\ \text{avec } Au^{(k+1)} = x^{(k)}\ (k \geq 0) \end{cases}$$

Calculer $\underset{k\to+\infty}{Lim}\, x^{(2k)}$, $\underset{k\to+\infty}{Lim}\, x^{(2k+1)}$ *et* $\underset{k\to+\infty}{Lim} \dfrac{\left(Ax^{(k)}\right)_j}{x_j^{(k)}}$ *pour tout* $j \in \{1, 2, \cdots, n\}$ *tel que* $e_{n,j} \neq 0$.

Q5. *Méthode de déflation — On suppose maintenant que les valeurs propres de A vérifient :*

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots > |\lambda_n|.$$

La matrice A est alors diagonalisable avec des valeurs propres réelles et simples.

Soit e_1 *un vecteur propre associé à* λ_1 *de norme euclidienne égale à 1. On lui associe la matrice* $B = A - \lambda_1 e_1 {}^t e_1$.

(a) Montrer que les valeurs propres de B sont 0, $\lambda_2, \cdots, \lambda_n$.

(b) En déduire une procédure de calcul de toutes les valeurs propres et tous les vecteurs propres de A.

Solution

Q1. On note, pour tout $k = 1, 2, \cdots, n$, $\alpha_k = \beta_k e^{i\theta_k}$, avec $\beta_k \geq 0$ et $\theta_k \in]-\pi, \pi]$.

On a alors $\left|\sum_{i=1}^{n} \alpha_i\right|^2 = \left(\sum_{i=1}^{n} |\alpha_i|\right)^2$, soit $\sum_{1\leq j<k\leq n} \beta_j \beta_k \operatorname{Cos}(\theta_j - \theta_k) = \sum_{1\leq j<k\leq n} \beta_j \beta_k$. Ce qui s'écrit aussi $\sum_{1\leq j<k\leq n} \beta_j \beta_k (1 - \operatorname{Cos}(\theta_j - \theta_k))$. Les coefficients de cette somme étant tous positifs ou nuls, on en déduit que $\operatorname{Cos}(\theta_j - \theta_k) = 1$, c'est-à-dire que $e^{i\theta_k} = u$ pour tout $k = 1, 2, \cdots, n$.

Q2. (a) Si $\rho(A) = 0$, alors 0 est valeur propre d'ordre n de A et avec le théorème de Cayley–Hamilton on déduit que $A^n = 0$, ce qui est contradictoire avec le fait que tous les coefficients de A sont strictement positifs. On a donc $\rho(A) > 0$.

(b) On a $\sum_{i=1}^{n} a_{ij} y_j = \lambda \cdot y_i$ pour tout $i = 1, 2, \cdots, n$. Si $\sum_{i=1}^{n} a_{ij} |y_j| > |\lambda| \cdot |y_i|$, pour tout $i = 1, 2, \cdots, n$, alors on peut alors trouver $\varepsilon > 0$ tel que $\sum_{i=1}^{n} a_{ij} |y_j| > (|\lambda| + \varepsilon)|y_i|$, pour tout $i = 1, 2, \cdots, n$. Si x est le vecteur de composantes $x_i = |y_i| \in IR^+$, on a alors $\|x\|_1 = \|y\|_1$ et $\|Ax\|_1 > (|\lambda| + \varepsilon)\|x\|_1$.

Par récurrence sur $k \geq 1$, on déduit alors que $\left\|A^k x\right\|_1 > (|\lambda| + \varepsilon)^k \|x\|_1$ et $\left\|A^k\right\|_1 > (|\lambda| + \varepsilon)^k$, donc $\left\|A^k\right\|_1^{\frac{1}{k}} > (|\lambda| + \varepsilon)$, ce qui contredit $\left\|A^k\right\|_1^{\frac{1}{k}} \underset{k\to+\infty}{\to} \rho(A) = |\lambda|$.

On a donc montré qu'il existe un indice $i \in \{1, 2, \cdots, n\}$ tel que $\sum_{i=1}^{n} a_{ij} |y_j| = |\lambda||y_i| = \left|\sum_{i=1}^{n} a_{ij} y_j\right|$.

Avec Q1, on déduit que x est proportionnel à y, c'est à dire qu'on peut trouver un vecteur propre x associé à λ de composantes positives. S'il existe $i \in \{1, 2, \cdots, n\}$ tel que $x_i = 0$, avec $\sum_{i=1}^{n} a_{ij} x_j = \lambda \cdot x_i = 0$, on déduit alors que $x = 0$. On a donc $x_i > 0$ pour tout $i \in \{1, 2, \cdots, n\}$.

(c) Avec $\sum_{i=1}^{n} a_{ij} x_j = \lambda \cdot x_i$, on déduit que λ est un réel strictement positif.

(d) Soient x et y deux éléments non nuls de E_λ et $\alpha \in C - \{0\}$ tel que $x_1 - \alpha \cdot y_1 = 0$. D'après ce qui précède, on a nécessairement $x - \alpha \cdot y = 0$. En effet $x - \alpha \cdot y$ est un vecteur propre associé à λ et s'il est non nul toutes ses composantes sont non nulles. On en déduit donc que E_λ est de dimension 1.

(e) Si λ est valeur propre de A de multiplicité supérieure ou égale à 2, en tenant compte de $\mathrm{Dim}(E_\lambda)=1$, on déduit que pour tout $x \in \left(IR_+^*\right)^n \cap E_\lambda$, on peut trouver $y \in C^n$ non colinéaire à x tel que $Ay = x + \lambda \cdot y$ (réduction de Jordan). Comme A, λ et x sont réels, on a aussi $A\overline{y} = x + \lambda \cdot \overline{y}$ et $A\dfrac{y+\overline{y}}{2} = x + \lambda \dfrac{y+\overline{y}}{2}$ avec $y+\overline{y} \in IR^n - \{0\}$ ($y+\overline{y}=0 \Rightarrow x=0$). On peut donc trouver $y \in IR^n - \{0\}$ tel que $Ay = x + \lambda \cdot y$.

Pour $\alpha > 0$ assez grand, on aura alors $z = y + \alpha x \in \left(IR_+^*\right)^n$ et $Az = x + \lambda \cdot z$. Soit, pour tout $i = 1, 2, \cdots, n$, $\sum_{j=1}^{n} a_{ij} z_j = x_i + \lambda \cdot z_i$.

(f) Si λ est valeur propre de A de multiplicité supérieure ou égale à 2, avec les notations du (e) on déduit alors que pour $\varepsilon > 0$ tel que $x_i > \varepsilon \cdot z_i$ pour tout $i = 1, 2, \cdots, n$, on a $\sum_{j=1}^{n} a_{ij} z_j > (\lambda+\varepsilon) z_i$, soit $\|Az\|_1 > (\lambda+\varepsilon)\|z\|_1$ et par récurrence $\left\|A^k z\right\|_1 > (\lambda+\varepsilon)^k \|z\|_1$, donc $\left\|A^k\right\|_1^{\frac{1}{k}} > (\lambda+\varepsilon)$, ce qui contredit $\left\|A^k\right\|_1^{\frac{1}{k}} \underset{k\to+\infty}{\to} \rho(A) = |\lambda|$.

Donc λ est valeur propre simple de A.

Q3. (a) On a $\lambda_1 \neq \lambda_j$ pour tout $j = 2, ..., n$, donc λ_1 est valeur propre simple de A. Si λ_1 est complexe non réelle alors $\overline{\lambda_1}$ est une autre valeur propre de A de même module que λ_1, ce qui contredit $|\lambda_1| > |\lambda_j|$ pour tout $j = 2, \cdots, n$.

(b) λ_1 étant simple, l'espace propre associé est de dimension 1.

Les valeurs propres de $A_1 = A - \lambda_1 I_d$ sont $0, \lambda_2 - \lambda_1, \cdots, \lambda_2 - \lambda_n$, la valeur propre nulle étant simple. La matrice A_1^2 a pour valeurs propres $0, (\lambda_2-\lambda_1)^2, \cdots, (\lambda_2-\lambda_n)^2$, la valeur propre nulle étant simple. $Ker\left(A_1^2\right)$ est alors de dimension 1 et avec $Ker(A_1) \subset Ker\left(A_1^2\right)$, on déduit que $Ker(A_1) = Ker\left(A_1^2\right)$. Il en résulte alors que $Ker(A_1) \cap \mathrm{Im}(A_1) = \{0\}$ et avec le théorème du rang on conclut que $IR^n = E_1 \oplus F_1$.

Enfin il est clair que E_1 et F_1 sont stables par A.

Remarque — De manière plus générale, on peut montrer que si λ est une valeur propre de A d'ordre $p \geq 1$, alors $IR^n = E_p \oplus F_p$, où $E_p = Ker(A - \lambda_1 I_d)^p$ et $F_p = \mathrm{Im}(A - \lambda_1 I_d)^p$ avec $Dim(E_p) = p$ et E_p, F_p stables par A.

(c) On identifie une matrice à l'application linéaire qu'elle définit et on note B la restriction de A à F_1. B est un endomorphisme de F_1 de valeurs propres $\lambda_2, \cdots, \lambda_n$.

On a $Ax^{(0)} = \lambda_1 e_1 + Bf_1 \neq 0$, donc $x^{(1)}$ est bien défini. Par récurrence, on voit que $x^{(k)}$ est bien défini avec une projection non nulle sur E_1. De manière plus précise, on a $x^{(k)} = \dfrac{1}{\left\|A^k x^{(0)}\right\|} A^k x^{(0)}$. En effet le résultat est vrai pour $k = 1$ et en le supposant vrai pour $k \geq 1$, on a $Ax^{(k)} = \dfrac{1}{\left\| A^k x^{(0)} \right\|} A^{k+1} x^{(0)}$ et $x^{(k+1)} = \dfrac{\left\| A^k x^{(0)} \right\|}{\left\| A^{k+1} x^{(0)} \right\|} \dfrac{1}{\left\| A^k x^{(0)} \right\|} A^{k+1} x^{(0)}$, d'où le résultat.

Pour tout $k \geq 1$, on a $A^k x^{(0)} = \lambda_1{}^k e_1 + B^k f = \lambda_1{}^k (e_1 + f_k)$ avec $\underset{k \to +\infty}{Lim} f_k = \underset{k \to +\infty}{Lim} \left(\left(\dfrac{1}{\lambda_1} B \right)^k f_1 \right) = 0$ car $\rho\left(\dfrac{1}{\lambda_1} B \right) < 1$.

(i) On peut écrire $x^{(k)} = \left(\dfrac{\lambda_1}{|\lambda_1|} \right)^k \dfrac{1}{\|e_1 + f_k\|} (e_1 + f_k)$ et on a alors $\left\| Ax^{(k)} \right\| = \dfrac{1}{\|e_1 + f_k\|} \|\lambda_1 e_1 + Af_k\|$. Il en résulte que $\underset{k \to +\infty}{Lim} \left\| Ax^{(k)} \right\| = |\lambda_1|$.

(ii) On déduit aussi que $\underset{k \to +\infty}{Lim} x^{(2k)} = v_1 = \dfrac{1}{\|e_1\|} e_1$ et $\underset{k \to +\infty}{Lim} x^{(2k+1)} = \dfrac{\text{Signe}(\lambda_1)}{\|e_1\|} e_1$.

(iii) Avec $Ax^{(k)} - \lambda_1 x^{(k)} = \left(\dfrac{\lambda_1}{|\lambda_1|} \right)^k \dfrac{1}{\|e_1 + f_k\|} (Af_k - \lambda_1 f_k)$, on déduit que $\underset{k \to +\infty}{Lim} \left(Ax^{(k)} - \lambda_1 x^{(k)} \right) = 0$.

Comme $e_1 \neq 0$, il existe $j \in \{1, 2, \cdots, n\}$ tel que $e_{1,j} \neq 0$ et pour k assez grand on a $x_j^{(j)} \neq 0$. Avec $\underset{k \to +\infty}{Lim} \left(\left(Ax^{(k)} \right)_j - \lambda_1 x_j^{(k)} \right) = 0$, on déduit alors que $\underset{k \to +\infty}{Lim} \dfrac{\left(Ax^{(k)} \right)_j}{x_j^{(k)}} = \lambda_1$.

Q4. La matrice A^{-1} vérifie les hypothèses de Q3 avec $\rho\left(A^{-1} \right) = \dfrac{1}{|\lambda_n|}$ et $\dfrac{1}{\lambda_n}$ valeur propre réelle et simple de A^{-1}.

D'autre part, on a $Ker(A - \lambda_n I_d) = Ker\left(A^{-1} - \dfrac{1}{\lambda_n} I_d \right)$ et $\operatorname{Im}(A - \lambda_n I_d) = \operatorname{Im}\left(A^{-1} - \dfrac{1}{\lambda_n} I_d \right)$. En effet $Ax = \lambda_n x$ équivaut à $A^{-1} x = \dfrac{1}{\lambda_n} x$ et $y = Ax - \lambda_n x$ équivaut à $y = A^{-1} x' - \dfrac{1}{\lambda_n} x'$ où $x' = -\lambda_n Ax$.

Enfin $x^{(k+1)} = \dfrac{1}{\left\| A^{-1} x^{(k)} \right\|} A^{-1} x^{(k)}$.

C'est-à-dire qu'on applique la méthode de la puissance itérée à l'inverse de A.

On déduit donc que $\underset{k \to +\infty}{Lim} x^{(2k)} = \dfrac{1}{\|e_n\|} e_n$ et $\underset{k \to +\infty}{Lim} x^{(2k+1)} = \dfrac{\text{Signe}(\lambda_n)}{\|e_n\|} e_n$.

On a $\underset{k\to+\infty}{Lim}\left(A^{-1}x^{(k)}-\dfrac{1}{\lambda_n}x^{(k)}\right)=0$, ce qui équivaut à $\underset{k\to+\infty}{Lim}\left(Ax^{(k)}-\lambda_n x^{(k)}\right)=0$ et $\underset{k\to+\infty}{Lim}\dfrac{\left(Ax^{(k)}\right)_j}{x_j^{(k)}}=\lambda_n$.

Q5. (a) Pour tout j = 1, 2, ···, n, λ_j est aussi valeur propre de ${}^tA$. On désigne alors par f_j un vecteur propre de ${}^tA$ associé à λ_j.

En écrivant que $\left\langle e_1 \middle| f_j\right\rangle=\dfrac{1}{\lambda_1}\left\langle Ae_1 \middle| f_j\right\rangle=\dfrac{1}{\lambda_1}\left\langle e_1 \middle| {}^tAf_j\right\rangle=\dfrac{\lambda_j}{\lambda_1}\left\langle e_1 \middle| f_j\right\rangle$, on déduit que $\left\langle e_1 \middle| f_j\right\rangle=0$ pour tout j = 2, ···, n.

On a alors pour tout j = 2, ···, n :

$${}^tBf_j={}^tAf_j-\lambda_1 e_1{}^te_1 f_j=\lambda_j f_j-\lambda_1\left\langle e_1 \middle| f_j\right\rangle e_1=\lambda_j f_j$$

c'est-à-dire que λ_j est valeur propre de ${}^tB$. Ce sont donc aussi des valeurs propres de B.

Enfin $Be_1=Ae_1-\lambda_1 e_1{}^te_1 e_1=\lambda_1 e_1-\lambda_1\left\langle e_1 \middle| e_1\right\rangle e_1=0$, c'est à dire que 0 est valeur propre de B.

(b) On applique la méthode de la puissance itérée à la matrice B pour avoir λ_2 et un vecteur propre associé, puis on continue ainsi de suite pour obtenir les autres valeurs propres et vecteurs propres.

Pour la programmation, les valeurs propres peuvent être stockées dans le vecteur *Valeurs_Propres* et les vecteurs propres dans la matrice *Vecteurs_Propres*.

La procédure ci-dessous calcule les p premières valeurs propres et les vecteurs propres associés pour p = 1, ···, n. Les valeurs propres sont calculées dans l'ordre des valeurs absolues décroissantes, ce qui suppose qu'elles sont toutes distinctes.

```
PROCEDURE Déflation(Entrée n, p : Entier ; A : Matrice ;
                    Sortie Valeurs_Propres : Vecteur ; Vecteurs_Propres : Matrice) ;
Début
  Pour k allant de 1 à p faire
  Début
    Se donner vk aléatoire ;
    Iter = 0 ;
    Répéter
      Iter = Iter + 1 ; i0 = 1 ;
      v = A.vk ;
      Pour i allant de 2 à n faire
      Début
        Si |vk_i| > |vk_i0|
        Alors i0 = i ;
      Fin ;
      Valeurs_Propres(k) = v_i0/vk_i0;
      Aux = NormeEuclidienne(v) ;
      v = (1/Aux)·v ;
      uAncien = vk ;
```

vk = v ;
Jusqu'à ($\|vk - uAncien\| < \varepsilon$) ou (Iter = MaxIter) ;
Pour i allant de 1 à n faire Vecteurs_Propres$_{ik}$ = v_i ;
Si Iter = MaxIter
Alors Arrêter(« Convergence trop lente ou divergence ») ;
Pour i allant de 1 à n faire
Début
Pour j allant de 1 à n faire
Début
$a_{ij} = a_{ij}$ – Valeurs_Propres(k)v_i*v_j ;*
Fin ;
Fin ;
Fin ;
Fin ;

Problème 19 : Méthode de Jacobi pour les matrices symétriques

Q1. *Pour toute matrice $M = \left(\left(m_{ij}\right)\right)_{1 \le i,j \le n}$ d'ordre $n \ge 2$ et à coefficients réels on désigne par $\|M\|_s = Tr\left({}^tM \cdot M\right)^{\frac{1}{2}}$ la « norme de Schur » de M (problème 1, Q5).*

Montrer qu'on définit bien ainsi une norme sur l'espace vectoriel des matrices réelles d'ordre n et qu'on a $\|M\|_2 \le \|M\|_s$, où $\| \ \|_2$ désigne la norme matricielle induite par la norme euclidienne et $\|M \cdot R\|_s = \|R \cdot M\|_s = \|M\|_s$ pour toute matrice orthogonale R.

Pour tout réel $\theta \in [-\pi, \pi[$ et tout couple d'entiers (p,q) tels que $1 \le p < q \le n$, on note $R_{p,q}(\theta)$ la matrice de rotation d'angle θ dans le plan défini par les vecteurs e_p et e_q de la base canonique de $\mathbb{R}^n$.

Pour toute matrice symétrique réelle d'ordre $n \ge 2$, $M = \left(\left(m_{ij}\right)\right)_{1 \le i,j \le n}$, on note $M_{p,q}(\theta) = R_{p,q}(\theta)^{-1} M R_{p,q}(\theta)$.

Q2. *Calculer les coefficients m'_{ij} de $M_{p,q}(\theta)$ pour tous i, j = 1, 2, ⋯, n ainsi que $\left\|M_{p,q}(\theta)\right\|_s$.*

Q3. *Soit $(p,q) \in IN^2$ tels que $1 \le p < q \le n$ et $m_{pq} \ne 0$.*

(a) Montrer qu'on peut trouver un unique réel $\theta \in \left]-\frac{\pi}{4}, \frac{\pi}{4}\right] - \{0\}$ tel que $m'_{pq} = 0$.

(b) Montrer que $t = \mathrm{Tg}(\theta)$ est l'unique racine dans $]-1,1] - \{0\}$ d'une équation du second degré.

(c) En déduire un calcul algébrique de $c = \mathrm{Cos}(\theta)$ et $s = \mathrm{Sin}(\theta)$.

(d) Déduire de ce qui précède des expressions simplifiées des coefficients de $M_{p,q}(\theta)$ en fonction des coefficients de M, de $c = \mathrm{Cos}(\theta)$, $s = \mathrm{Sin}(\theta)$, $t = \mathrm{Tg}(\theta)$ et $\tau = \mathrm{Tg}\left(\frac{\theta}{2}\right)$.

Dans ce qui suit, A désigne une matrice symétrique réelle d'ordre $n \geq 3$ *de valeurs propres (réelles)* $\lambda_1, \lambda_2, \cdots, \lambda_n$. *On lui associe la suite de matrices symétriques* $(A_k)_{k \in IN}$ *définie par* $A_k = \left(\left(a_{ij}^k\right)\right)_{1 \leq i,j \leq n}$ *et :*

$$\begin{cases} A_0 = A \\ A_{k+1} = R(\theta_k)^{-1} A_k R(\theta_k) \ (k \geq 0) \end{cases}$$

où $R(\theta_k)$ *est la matrice unité si* $a_{pq}^k = 0$ *et une matrice orthogonale* $R_{p,q}(\theta_k)$ *avec* $1 \leq p < q \leq n$ *tels que* $\left|a_{pq}^k\right| = Max\left\{\left|a_{ij}^k\right|;\ 1 \leq i < j \leq n\right\}$ *et* $\theta_k \in \left]-\frac{\pi}{4}, \frac{\pi}{4}\right] - \{0\}$ *tel que* $a_{pq}^{k+1} = 0$.

On décompose chaque matrice A_k *sous la forme* $A_k = D_k + E_k$ *où* D_k *est la partie diagonale de* A_k.

Q4. *(a) Calculer* $\|D_{k+1}\|_s$ *en fonction de* $\|D_k\|_s$.

(b) Calculer $\|E_{k+1}\|_s$ *en fonction de* $\|E_k\|_s$.

(c) Montrer que $\underset{k \to +\infty}{Lim}\|E_k\|_s = 0$ *et* $\underset{k \to +\infty}{Lim} E_k = 0$.

Q5. *Soit E un espace vectoriel normé de dimension finie et* $(x_k)_{k \in IN}$ *une suite d'éléments de E qui vérifie les hypothèses suivantes :*

(i) $(x_k)_{k \in IN}$ *est bornée ;*

(ii) $(x_k)_{k \in IN}$ *n'admet qu'un nombre fini de valeurs d'adhérence ;*

(iii) $\underset{k \to +\infty}{Lim}\|x_{k+1} - x_k\| = 0$.

Montrer alors que $(x_k)_{k \in IN}$ *est convergente dans E.*

Q6. *En utilisant Q5, montrer que la suite* $(D_k)_{k \in IN}$ *converge vers une matrice diagonale* D_σ *de termes diagonaux* $d_i^\sigma = \lambda_{\sigma(i)}$ *pour i = 1, 2, ⋯, n, où* σ *est une permutation de* $\{1, 2, \cdots, n\}$.

Q7. *Avec les notations de Q3, on désigne par* $(P_k)_{k \in IN}$ *la suite de matrices orthogonales définie par* $P_k = R(\theta_0) R(\theta_1) \cdots R(\theta_k)$ *pour tout* $k \geq 0$.

On suppose que toutes les valeurs propres de A sont distinctes. On désigne par D la matrice diagonale de termes diagonaux $d_i = \lambda_i$ *pour i = 1, 2, ..., n et par* $P = (C_1, C_2, \cdots, C_n)$ *une matrice orthogonale telle que* $D = P^{-1} A P$ *, où* C_j *désigne la colonne numéro j de P.*

(a) Montrer que $\underset{k \to +\infty}{Lim} \theta_k = 0$ *et* $\underset{k \to +\infty}{Lim} R_k = I_d$.

(b) En utilisant Q5, montrer que la suite $(P_k)_{k \in IN}$ *converge vers une matrice* $P_\sigma = \left(\pm C_{\sigma(1)}, \pm C_{\sigma(2)}, \cdots, \pm C_{\sigma(n)}\right)$ *où* σ *est une permutation de* $\{1, 2, \cdots, n\}$. *Le vecteur* $C_{\sigma(i)}$ *étant un vecteur propre associé à la valeur propre* $\lambda_{\sigma(i)}$.

Q8. *Ecrire une procédure de calcul des valeurs propres et des vecteurs propres associés d'une matrice symétrique réelle ayant toutes ses valeurs propres distinctes.*

Solution

Q1. On a $\|M\|_s = \left(\sum_{i,j=1}^{n} |m_{ij}|^2 \right)^{\frac{1}{2}}$, c'est-à-dire que la norme de Schur de M est la norme euclidienne de M considérée comme vecteur de IR^{n^2}.

Soient $\lambda_1, \lambda_2, \cdots, \lambda_n$ les valeurs propres, réelles positives de ${}^tMM$ (symétrique réelle et positive), avec $\rho({}^tMM) = \lambda_1$. On sait alors que $\|M\|_2^{\ 2} = \lambda_1$ (problème 1, Q6, (c)). Et avec $\lambda_1 \le Tr({}^tMM) = \|M\|_s^{\ 2}$ on déduit que $\|M\|_2 \le \|M\|_s$.

Si R est une matrice orthogonale, on a :

$$\|MR\|_s = Tr({}^tR \cdot {}^tM \cdot M \cdot R)^{\frac{1}{2}} = Tr(R^{-1} \cdot {}^tM \cdot M \cdot R)^{\frac{1}{2}} = Tr({}^tM \cdot M)^{\frac{1}{2}} = \|M\|_s.$$

On démontre de la même façon que $\|RM\|_s = \|M\|_s$.

Q2. La matrice de rotation $R_{p,q}(\theta)$ est définie par :

$$R_{p,q}(\theta) = \begin{array}{cc} \begin{pmatrix} 1 & 0 & . & . & . & . & 0 \\ . & . & . & . & . & . & \\ . & . & c. & . & -s & & \\ . & . & . & 1 & . & & \\ . & . & s & . & c & & \\ . & . & . & . & . & . & \\ 0 & . & . & . & . & 0 & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow p \\ \\ \leftarrow q \\ \\ \\ \end{array} \\ \begin{array}{ccccccc} & & \uparrow & & \uparrow & & \\ & & \mathrm{p} & & \mathrm{q} & & \end{array} & \end{array}$$

où on a posé $c = \text{Cos}(\theta)$ et $s = \text{Sin}(\theta)$.

La multiplication à gauche par $R_{p,q}(\theta)^{-1} = {}^t R_{p,q}(\theta)$ modifie seulement les lignes p et q de M et la multiplication à droite par $R_{p,q}(\theta)$ change seulement les colonnes p et q de M. Ce qui donne pour les coefficients de $M_{p,q}(\theta) = \left((m'_{ij}) \right)_{1 \le i,j \le n}$:

$$\begin{cases} m'_{ij} = m_{ij} \ (i \ne p,\ i \ne q,\ j \ne p,\ j \ne q) \\ m'_{ip} = c \cdot m_{ip} + s \cdot m_{iq} \ (i \ne p,\ i \ne q) \\ m'_{pp} = c^2 m_{pp} + s^2 m_{qq} + 2s \cdot c \cdot m_{pq} \\ m'_{iq} = c \cdot m_{iq} - s \cdot m_{ip} \ (i \ne p,\ i \ne q) \\ m'_{qq} = s^2 m_{pp} + c^2 m_{qq} - 2s \cdot c \cdot m_{pq} \\ m'_{pq} = (c^2 - s^2) m_{pq} - s \cdot c (m_{pp} - m_{qq}) \end{cases}$$

la matrice $M_{p,q}(\theta)$ étant symétrique.

Avec Q1, on déduit que $\|M_{p,q}(\theta)\|_s = \|M\|_s$.

Q3. (a) La condition $m'_{pq} = 0$ équivaut à :

$$\left(\mathrm{Cos}(\theta)^2 - \mathrm{Sin}(\theta)^2\right)m_{pq} - \mathrm{Sin}(\theta)\cdot\mathrm{Cos}(\theta)\left(m_{pp} - m_{qq}\right) = 0$$

Si $m_{pq} \neq 0$, alors $\theta \neq 0$. Si $m_{pp} = m_{qq}$, on prend alors $\theta = \dfrac{\pi}{4}$ et si $m_{pp} \neq m_{qq}$ la condition ci-dessus s'écrit :

$$\mathrm{Tg}(2\theta) = \frac{2m_{pq}}{m_{pp} - m_{qq}}$$

et cette équation admet une unique solution dans $\theta \in \left]-\dfrac{\pi}{4}, \dfrac{\pi}{4}\right[- \{0\}$.

(b) $t = \mathrm{Tg}(\theta) \in]-1,1] - \{0\}$ est solution de l'équation du second degré $t^2 + 2b_{pq}t - 1 = 0$ où on a posé $b_{pq} = \dfrac{m_{pp} - m_{qq}}{2m_{pq}}$.

Cette équation admet deux racines réelles :

$$t_1 = -b_{pq} + \sqrt{1 + b_{pq}{}^2} = \frac{1}{b_{pq} + \sqrt{1 + b_{pq}{}^2}} \text{ et } t_2 = -\frac{1}{t_1} = \frac{-1}{\sqrt{1 + b_{pq}{}^2} - b_{pq}}$$

La racine de valeur absolue inférieure à 1 est :

$$t = \frac{\mathrm{Signe}(b_{pq})}{\sqrt{b_{pq}{}^2 + 1} + |b_{pq}|}$$

avec la convention $\mathrm{Signe}(0) = 1$.

(c) Avec $\mathrm{Cos}(\theta) = \dfrac{1}{\sqrt{1 + \mathrm{Tg}^2(\theta)}}$ et $\mathrm{Sin}(\theta) = \mathrm{Cos}(\theta)\mathrm{Tg}(\theta)$, on déduit que $c = \dfrac{1}{\sqrt{1 + t^2}}$ et $s = t \cdot c$.

(d) En remarquant que $\mathrm{Tg}\left(\dfrac{\theta}{2}\right) = \dfrac{s}{1 + c}$, les formules donnant les coefficients de $M_{p,q}(\theta)$ se simplifient alors pour donner finalement :

$$\begin{cases} m'_{ij} = m_{ij} \ (i \neq p,\ i \neq q,\ j \neq p,\ j \neq q) \\ m'_{ip} = m_{ip} + s\left(m_{iq} - \tau \cdot m_{ip}\right) \ (i \neq p,\ i \neq q) \\ m'_{pp} = m_{pp} + t \cdot m_{pq} \\ m'_{iq} = m_{iq} - s\left(m_{ip} + \tau \cdot m_{iq}\right) \ (i \neq p,\ i \neq q) \\ m'_{qq} = m_{qq} - t \cdot m_{pq} \\ m'_{pq} = 0 \end{cases}$$

Q4. (a) En utilisant les formules simplifiées de Q3 (b), on a :

$$\|D_{k+1}\|_s^2 = \sum_{i=1}^{n} \left|a_{ii}^{k+1}\right|^2 = \sum_{\substack{i=1 \\ i \neq p, i \neq q}}^{n} \left|a_{ii}^{k}\right|^2 + \left|a_{pp}^{k+1}\right|^2 + \left|a_{qq}^{k+1}\right|^2$$

avec :

$$\left|a_{pp}^{k+1}\right|^2+\left|a_{qq}^{k+1}\right|^2=\left|a_{pp}^{k}+t_k a_{pq}^{k}\right|^2+\left|a_{qq}^{k}-t_k a_{pq}^{k}\right|^2$$
$$=\left|a_{pp}^{k}\right|^2+\left|a_{qq}^{k}\right|^2+2t_k a_{pq}^{k}\left(a_{pp}^{k}-a_{qq}^{k}+t_k a_{pq}^{k}\right)$$

et avec $t_k^{\ 2}+2b_{pq}t_k-1=0$ on a $a_{pp}^{k}-a_{qq}^{k}=\dfrac{1-t_k^{\ 2}}{t_k}a_{pq}^{k}$ et :

$$\left|a_{pp}^{k+1}\right|^2+\left|a_{qq}^{k+1}\right|^2=\left|a_{pp}^{k}\right|^2+\left|a_{qq}^{k}\right|^2+2\left|a_{pq}^{k}\right|^2$$

ce qui donne :

$$\left\|D_{k+1}\right\|_s^{\ 2}=\left\|D_k\right\|_s^{\ 2}+2\left|a_{pq}^{k}\right|^2$$

Remarque — On peut aussi constater que :

$$\begin{pmatrix} a_{pp}^{k+1} & a_{pq}^{k+1} \\ a_{qp}^{k+1} & a_{qq}^{k+1} \end{pmatrix}=\begin{pmatrix} \mathrm{Cos}(\theta_k) & \mathrm{Sin}(\theta_k) \\ -\mathrm{Sin}(\theta_k) & \mathrm{Cos}(\theta_k) \end{pmatrix}\begin{pmatrix} a_{pp}^{k} & a_{pq}^{k} \\ a_{qp}^{k} & a_{qq}^{k} \end{pmatrix}\begin{pmatrix} \mathrm{Cos}(\theta_k) & -\mathrm{Sin}(\theta_k) \\ \mathrm{Sin}(\theta_k) & \mathrm{Cos}(\theta_k) \end{pmatrix}$$

et par conservation des normes on a immédiatement :

$$\left|a_{pp}^{k+1}\right|^2+\left|a_{qq}^{k+1}\right|^2+2\left|a_{pq}^{k+1}\right|^2=\left|a_{pp}^{k}\right|^2+\left|a_{qq}^{k}\right|^2+2\left|a_{pq}^{k}\right|^2$$

On conclut alors avec $a_{pq}^{k+1}=0$.

(b) On a $\left\|A_k\right\|_s^{\ 2}=\left\|D_k\right\|_s^{\ 2}+\left\|E_k\right\|_s^{\ 2}$ avec $\left\|A_k\right\|_s=\left\|A\right\|_s$ pour tout $k\geq 0$. Avec le (a), on déduit alors que :

$$\left\|E_{k+1}\right\|_s^{\ 2}=\left\|E_k\right\|_s^{\ 2}-2\left|a_{pq}^{k}\right|^2$$

(c) En utilisant le fait que a_{pq}^{k} est de module maximal dans E_k, on peut écrire que $\left\|E_k\right\|_s^{\ 2}\leq\left(n^2-n\right)\left|a_{pq}^{k}\right|^2$ et de (b) on déduit que :

$$\left\|E_{k+1}\right\|_s^{\ 2}\leq\left(1-\frac{2}{n^2-n}\right)\left\|E_k\right\|_s^{\ 2}$$

Par récurrence on a alors :

$$\left\|E_k\right\|_s^{\ 2}\leq\left(1-\frac{2}{n^2-n}\right)^k\left\|E_0\right\|_s^{\ 2}$$

Il en résulte que $\underset{k\to+\infty}{Lim}\left\|E_k\right\|_s=0\ \left(n\geq 3\right)$ et $\underset{k\to+\infty}{Lim}E_k=0$.

Q5. Soit $\{a_1,a_2,\cdots,a_n\}$ la suite finie des valeurs d'adhérence de la suite $\left(x_k\right)_{k\in IN}$. Pour $\varepsilon=\dfrac{1}{4}Min\left\{\left\|a_i-a_j\right\|;\ 1\leq i,j\leq p,\ i\neq j\right\}>0$, on peut trouver un entier k_ε tel que :

$$\forall k\geq k_\varepsilon,\ \begin{cases}\left\|x_{k+1}-x_k\right\|<\varepsilon \\ x_k\in\displaystyle\bigcup_{i=1}^{p}B\left(a_i,\varepsilon\right)\end{cases}$$

où $B\left(a_i,\varepsilon\right)=\left\{x\in E;\ \left\|x-a_i\right\|<\varepsilon\right\}$.

Si j est un entier tel que $x_{k_\varepsilon}\in B\left(a_j,\varepsilon\right)$, on a alors :

$$\forall k\geq k_\varepsilon,\ x_k\in B\left(a_j,\varepsilon\right)$$

En effet le résultat est vrai pour k_ε. Supposons le vrai pour $k \geq k_\varepsilon$ et soit i tel que $x_k \in B(a_i, \varepsilon)$. Si $i \neq j$, on a :

$$\|x_{k+1} - x_k\| \geq \|a_i - a_j\| - \|x_{k+1} - a_i\| - \|x_k - a_j\| \geq 4\varepsilon - \varepsilon - \varepsilon = 2\varepsilon$$

ce qui est faux. On a donc $i = j$. On a donc ainsi montré que $\underset{k\to+\infty}{Lim} = a_j$.

Q6. Avec $\|D_k\|_s \leq \|A_k\|_s = \|A\|_s$, on déduit que la suite $(D_k)_{k\in IN}$ est bornée.

Avec $a_{ii}^{k+1} - a_{ii}^k = \begin{cases} 0 \text{ si } i \neq p,\ i \neq q \\ \mathrm{Tg}(\theta_k) a_{pq}^k, \text{ si } i = p \\ -\mathrm{Tg}(\theta_k) a_{pq}^k, \text{ si } i = q \end{cases}$, $|\mathrm{Tg}(\theta_k)| \leq 1$ et $|a_{pq}^k| \leq \|E_k\|_s \underset{k\to+\infty}{\longrightarrow} 0$, on déduit que $\underset{k\to+\infty}{Lim}(D_{k+1} - D_k) = 0$.

Pour montrer que la suite $(D_k)_{k\in IN}$ est convergente il nous reste donc à montrer qu'elle n'a qu'un nombre fini de valeurs d'adhérence.

Soit $\left(D_{\varphi(k)}\right)_{k\in IN}$ une suite extraite de $(D_k)_{k\in IN}$ qui converge vers une matrice D diagonale. Avec $\underset{k\to+\infty}{Lim} E_{\varphi(k)} = 0$, on déduit que $\underset{k\to+\infty}{Lim} A_{\varphi(k)} = D$. Si $P_k(\lambda)$ désigne le polynôme caractéristique de $A_{\varphi(k)}$, on a $\underset{k\to+\infty}{Lim} P_k(\lambda) = Dét(D - \lambda \cdot I_d)$ avec $P_k(\lambda) = Dét(A - \lambda \cdot I_d)$ puisque A et $A_{\varphi(k)}$. On a donc montré que D et A ont le même polynôme caractéristique, donc les mêmes valeurs propres. La matrice D étant diagonale les termes diagonaux de D sont donnés par $d_i^\sigma = \lambda_{\sigma(i)}$ pour tout $i = 1,\ 2,\ \cdots$, n, où σ est une permutation de $\{1, 2, \cdots, n\}$. Le nombre de ces permutations étant fini, on déduit qu'il n'y a qu'un nombre fini de valeurs d'adhérences.

Q7. (a) Avec $\underset{k\to+\infty}{Lim} D_k = D_\sigma$ et l'hypothèse que les valeurs propres de A sont deux à deux distinctes, on déduit qu'il existe un entier k_0 tel que :

$$\forall k \geq k_0,\ \left|a_{pq}^k - a_{qq}^k\right| > \frac{1}{2} Min\left\{\left|\lambda_i - \lambda_j\right|;\ 1 \leq i, j \leq n,\ i \neq j\right\} > 0$$

Avec $|a_{pq}^k| \leq \|E_k\|_s \underset{k\to+\infty}{\longrightarrow} 0$, on déduit alors que $|b_k| = \left|\dfrac{a_{pp}^k - a_{qq}^k}{2a_{pq}^k}\right| \underset{k\to+\infty}{\longrightarrow} +\infty$ et $\mathrm{Tg}(\theta_k) = \dfrac{\mathrm{Signe}(b_k)}{\sqrt{b_k^{\ 2} + 1} + |b_k|} \underset{k\to+\infty}{\longrightarrow} 0$. On a donc $\underset{k\to+\infty}{Lim}\ \theta_k = 0$ et $\underset{k\to+\infty}{Lim} R_k = I_d$.

(b) On a $\|P_k\|_2 = 1$ (P_k est orthogonale) et avec $P_{k+1} - P_k = P_k\left(R(\theta_{k+1}) - I_d\right)$ on déduit que $\underset{k\to+\infty}{Lim}(P_{k+1} - P_k) = 0$.

Il nous reste donc à montrer que la suite $(P_k)_{k\in IN}$ n'a qu'un nombre fini de valeurs d'adhérence.

Soit $\left(P_{\varphi(k)}\right)_{k\in IN}$ une suite extraite de $(P_k)_{k\in IN}$ qui converge vers une matrice orthogonale Q. Avec $A_{\varphi(k)} = {}^tP_{\varphi(k)} A P_{\varphi(k)}$, on déduit que

${}^tP_\sigma AP_\sigma = D_\sigma = \lim\limits_{k\to+\infty} A_{\varphi(k)} = {}^tQAQ$. Les valeurs propres de A étant deux à deux distinctes, on a nécessairement $Q = \left(\pm C_{\sigma(1)}, \pm C_{\sigma(2)}, \cdots, \pm C_{\sigma(n)}\right)$. D'où le résultat.

Q8. *Remarque 1* — D'un point de vue numérique, le calcul du maximum des termes non diagonaux de A_k à chaque étape n'est pas intéressant car il augmente le temps de calcul. On préfère procéder de la manière suivante : à l'étape k du calcul, la matrice A_{k-1} étant construite on construit la matrice A_k en effectuant $\dfrac{n(n-1)}{2}$ transformations de Jacobi en prenant pour valeurs successives de (p,q), les valeurs $(1,2)$, $(1,3)$, $\cdots$, $(1,n)$; $(2,3)$, $\cdots$, $(2,n)$; $\cdots$; $(n-1,n)$. Un tel calcul est appelé un « balayage ».

En notant $S_k = \sum\limits_{1\le i<j\le n} \left|a_{ij}^k\right|$, on arrête les itérations quand $S_k < \varepsilon$, où ε est une précision donnée.

On peut alors montrer que la convergence devient très rapidement quadratique.

Remarque 2 — En pratique on n'effectuera pas de transformation de Jacobi si $\left|a_{pq}^k\right| < \sigma$, où σ est un seuil donné. Usuellement, on prend la valeur $\sigma = \dfrac{S_k}{5n^2}$ seulement pour les trois premiers balayages, puis à partir du quatrième, on décide que $a_{pq}^k = 0$ si $\left|a_{pq}^k\right| < \varepsilon'\left|a_{pp}^k\right|$ et $\left|a_{pq}^k\right| < \varepsilon'\left|a_{qq}^k\right|$ avec ε' assez petit.

Remarque 3 — Le calcul de la matrice P_σ, donnant les vecteurs propres de A, se fait avec les relations :

$$\begin{cases} p_{ij}^{k+1} = p_{ij}^k \text{ si } j \ne p,\ j \ne q \\ p_{ip}^{k+1} = p_{ip}^k + s\left(p_{iq}^k - \tau\cdot p_{ip}^k\right)\ (i = 1,2,\ldots,n) \\ p_{iq}^{k+1} = p_{iq}^k - s\left(p_{ip}^k + \tau\cdot p_{iq}^k\right) \end{cases}$$

où $\tau = \mathrm{Tg}\left(\dfrac{\theta_k}{2}\right) = \dfrac{s}{1+c}$.

Ce qui nous donne en définitive la programmation structurée suivante, où *Valeurs_Propres* est un vecteur qui contiendra les valeurs propres et *Vecteurs_Propres* une matrice dont les colonnes seront des vecteurs propres associés.

```
PROCEDURE Jacobi(Entrée n : Entier ; A : Matrice ;
                 Sortie Valeurs_Propres : Vecteur ; Vecteurs_Propres : Matrice) ;
Début
  Vecteurs_Propres = Id ; Correct = Faux ; Iteration = 0 ; Valeurs_Propres = 0 ;
  Pour i Allant de 1 à n Faire Valeurs_Propres_i = a_ii ;
  Répéter
    Iteration = Iteration + 1 ;
    Sk = 0 ;
    Pour i allant de 1 à n – 1 Faire
    Début
      Pour j allant de i + 1 à n Faire
```

Début
$Sk = Sk + \left|a_{ij}\right|$;
Fin ;
Fin ;
Correct = (Sk < ε) ou (Iteration = MaxIteration) ;
Si Iteration < 4
Alors Seuil = $Sk/(5 \cdot n^2)$
Sinon Seuil = ε^2 ;
Pour p allant de 1 à n – 1 Faire
Début
Pour q allant de p + 1 à n Faire
Début
Si (Iteration ≥ 4) Et $(10^8 \cdot \left|a_{pq}\right| < \left|Valeurs\_Propres_p\right|)$
Et $(10^8 \cdot \left|a_{pq}\right| < \left|Valeurs\_Propres_q\right|)$
Alors $a_{pq} = O$
Sinon Si $\left|a_{pq}\right| > Seuil$
Alors Début
$b = 0.5 \cdot (Valeurs\_Propres_p - Valeurs\_Propres_q)/a_{pq}$;
$t = \dfrac{1}{\sqrt{b^2 + 1} + \left|b\right|}$;
Si b < 0 Alors t = –t ;
$c = \dfrac{1}{\sqrt{1 + t^2}}$;
$s = t \cdot c$;
$\tau = s/(1 + c)$;
$Aux = t \cdot a_{pq}$;
$Valeurs\_Propres_p = Valeurs\_Propres_p + Aux$;

Valeurs_Propres$_q$ = *Valeurs_Propres*$_q$ – *Aux* ;
$a_{pq} = 0$;
Pour j Allant de 1 à p – 1 Faire
Début
Aux = a_{jp} ;
$a_{jp} = a_{jp} + s\cdot(a_{jq} - \tau\cdot a_{jp})$;
$a_{jq} = a_{jq} - s\cdot(Aux + \tau\cdot a_{jq})$;
Fin ;
Pour j Allant de p + 1 à q – 1 Faire
Début
Aux = a_{pj} ;
$a_{pj} = a_{pj} + s\cdot(a_{jq} - \tau\cdot a_{pj})$;
$a_{jq} = a_{jq} - s\cdot(Aux + \tau\cdot a_{jq})$;
Fin ;
Pour j Allant de q + 1 à n Faire
Début
Aux = a_{pj} ;
$a_{pj} = a_{pj} + s\cdot(a_{qj} - \tau\cdot a_{pj})$;
$a_{qj} = a_{qj} - s\cdot(Aux + \tau\cdot a_{qj})$;
Fin ;
Pour i Allant de 1 à n Faire
Début
Aux = *Vecteurs_Propres*$_{ip}$;
Vecteurs_Propres$_{ip}$ =
Vecteurs_Propres$_{ip}$ + $s\cdot$(*Vecteurs_Propres*$_{iq}$ – $\tau\cdot$ *Vecteurs_Propres*$_{ip}$) ;
Vecteurs_Propres$_{iq}$ =
Vecteurs_Propres$_{iq}$ – $s\cdot$(*Aux* + $\tau\cdot$ *Vecteurs_Propres*$_{iq}$) ;
Fin ;
Fin ;
Fin ;
Fin ;
Jusqu'à Correct ;
Si Itération = MaxItération
Alors Arrêter(« Nombre d'itérations trop grand ») ;
Fin ;

Problème 20 : Tridiagonalisation d'une matrice symétrique. Méthode de Householder

$I\!R^n$ est muni du produit scalaire euclidien $\langle\cdot|\cdot\rangle$ et de la norme euclidienne associée $\|\cdot\|$.

On désigne par $\{e_1, e_2, \cdots, e_n\}$ la base canonique de $I\!R^n$.

Pour tout réel x, le signe de x est défini par : $Signe(x)=\begin{cases}1 \text{ si } x>0\\ 0 \text{ si } x=0\\ -1 \text{ si } x<0\end{cases}$.

Pour tout vecteur unitaire $u=\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}$ *dans* $\mathbb{R}^n$, *on désigne par* P_u *la matrice de Householder définie par* $P_u = I_d - 2u\cdot{}^t u$.

Q1. *Interpréter géométriquement l'endomorphisme associé à* P_u *et calculer son inverse.*

Q2. *Soit M une matrice symétrique réelle d'ordre* $n \geq 1$. *Montrer que :*

$$P_u^{-1}MP_u = M - 2\left(v\cdot{}^t u + u\cdot{}^t v\right)$$

où $v = Mu - \langle u \mid Mu\rangle u$.

Dans ce qui suit, on se donne une matrice symétrique réelle $A=\left(\left(a_{ij}\right)\right)_{1\leq i,j\leq n}$ *d'ordre* $n \geq 3$.

Q3. *Soit x la première colonne de A.*

On note $x^{(2)} = x - a_{11}e_1$, $y^{(2)} = x^{(2)} + Signe(a_{21})\left\|x^{(2)}\right\|e_2$ *et* $P_1=\begin{cases}I_d \text{ si } y^{(2)}=0\\ P_{u^{(2)}} \text{ si } y^{(2)}\neq 0\end{cases}$,

où $u^{(2)} = \dfrac{1}{\left\|y^{(2)}\right\|}y^{(2)}$ *pour* $y^{(2)} \neq 0$. *Enfin on note* $A_1 = P_1^{-1}AP_1 = \left(\left(a_{ij}^1\right)\right)_{1\leq i,j\leq n}$.

Montrer que A_1 *est symétrique avec* $a_{i1}^1 = 0$ *pour i = 3, 4, ⋯, n.*

Q4. *Montrer qu'il existe n – 2 matrices orthogonales* $P_1, P_2, \cdots, P_{n-2}$ *telles que*

$$T = P_{n-2}{}^{-1}P_{n-1}{}^{-1}\cdots P_1^{-1}AP_1P_2\ldots P_{n-2}$$

soit tridiagonale.

Q5. *Ecrire une procédure de réduction d'une matrice symétrique réelle à la forme tridiagonale.*

Remarque — Le résultat de cet exercice peut se montrer en utilisant le fait que toute matrice symétrique réelle est orthogonalement semblable à une matrice diagonale, donc tridiagonale. L'avantage de cette méthode est qu'elle fournit un procédé algorithmique simple de réduction à la forme tridiagonale sans avoir à calculer les valeurs propres. Cette réduction sera utilisée pour calculer les valeurs propres d'une matrice symétrique réelle (problèmes 21 et 22).

Solution

Q1. Soit H_u l'hyperplan orthogonal à u dans $\mathbb{R}^n$. Pour tout x dans H_u, on a $P_u x = x$. Et avec $P_u u = -u$, on conclut que P_u est la symétrie orthogonale par rapport à H_u. On a donc $P_u^{-1} = {}^tP_u = P_u$.

Q2. On a :

$$P_u^{-1}MP_u = P_uMP_u = M - 2u\cdot{}^t uM - 2\,Mu\cdot{}^t u + 4u\cdot{}^t uMu\cdot{}^t u$$

avec :

$$u\cdot{}^t uM - u\left({}^t uMu\right){}^t u = u\cdot{}^t\left(Mu\right) - u\langle u \mid Mu\rangle{}^t u = u\cdot{}^t v$$

et :

$$Mu\cdot{}^t u - u\left({}^t uMu\right){}^t u = \left(Mu - \langle u \mid Mu\rangle u\right){}^t u = v\cdot{}^t u$$

d'où le résultat.

Q3. Si $y^{(2)} = 0$, alors $a_{i1} = 0$ pour tout i = 3, 4, $\cdots$,n et $A_1 = A$ vérifie bien la condition voulue. On suppose donc que $y^{(2)} \neq 0$.

On a $A_1 = P_1AP_1$ et A_1 est symétrique puisque P_1 et A le sont.

La première colonne de A_1 est donnée par $A_1e_1 = P_1AP_1e_1$.

Avec $e_1 \in \left(IRy^{(2)}\right)^{\perp}$, on a $P_1e_1 = e_1$ et $A_1e_1 = P_1Ae_1 = P_1x$.

Si on note $y^{(2)} = x^{(2)} + z^{(2)}$, on a :

$$\left\langle y^{(2)} \middle| x^{(2)} - z^{(2)} \right\rangle = \left\langle x^{(2)} + z^{(2)} \middle| x^{(2)} - z^{(2)} \right\rangle = \left\| x^{(2)} \right\|^2 - \left\| x^{(2)} \right\|^2 = 0$$

donc :

$$P_1\left(x^{(2)} + z^{(2)}\right) = -x^{(2)} - z^{(2)}$$

$$P_1\left(x^{(2)} - z^{(2)}\right) = x^{(2)} - z^{(2)}$$

et $P_1x^{(2)} = -z^{(2)} = -Signe\left(a_{21}\right)\left\| x^{(2)} \right\| e_2$. Ce qui donne en définitive :

$$P_1x = P_1\left(a_{11}e_1 + x^{(2)}\right) = a_{11}e_1 - Signe\left(a_{21}\right)\left\| x^{(2)} \right\| e_2$$

C'est à dire que $a_{i1}^1 = \begin{cases} a_{11} \text{ si } i = 1 \\ -\,Signe\left(a_{21}\right)\left\| x^{(2)} \right\| \text{ si } i = 2 \\ 0 \text{ si } i = 3, \cdots, n \end{cases}$

Q4. On raisonne par récurrence sur $n \geq 2$. Pour $n = 2$, la matrice A est déjà tridiagonale. On suppose donc le résultat vrai pour les matrices symétriques d'ordre n – 1.

Avec les notations de Q2, on a $A_1 = \begin{pmatrix} a_{11} & a_{12}^1 & 0 & \cdots & 0 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 & \cdots & a_{2n}^1 \\ 0 & a_{32}^1 & a_{33}^1 & \cdots & a_{3n}^1 \\ . & & & & \\ & . & . & . & . \\ . & & & & \\ 0 & a_{n2}^1 & a_{n3}^1 & \cdots & a_{nn}^1 \end{pmatrix}$, la matrice étant symétrique ainsi que la matrice d'ordre $n-1$, $B = \begin{pmatrix} a_{22}^1 & \cdots & a_{2n}^1 \\ . & . & . \\ a_{n2}^1 & \cdots & a_{nn}^1 \end{pmatrix}$. Avec l'hypothèse de récurrence, on peut trouver des matrices orthogonales Q_2, Q_3, $\cdots$, Q_{n-2} telles que $T_1 = Q_{n-2}^{-1} \cdots Q_2^{-1}BQ_2 \cdots Q_{n-2}$ soit tridiagonale. On pose alors

$P_j = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q_j & \\ 0 & & & \end{pmatrix}$, pour $j = 2, 3, \cdots, n-2$ et $T = \begin{pmatrix} a_{11} & a_{12}^1 & 0 & \cdots & 0 \\ a_{12}^1 & & & & \\ 0 & & & & \\ \vdots & & & T_1 & \\ 0 & & & & \end{pmatrix}$. On a alors $T = P_{n-2}^{-1} \cdots P_2^{-1} A_1 P_2 \cdots P_{n-2} = P_{n-2}^{-1} \cdots P_2^{-1} P_1^{-1} A P_1 P_2 \cdots P_{n-2}$.

Q5. La matrice $T = A_{n-2}$ est définie par la récurrence :

$$\begin{cases} A_0 = A \\ A_k = P_k A_{k-1} P_k \ (1 \le k \le n-2) \end{cases}$$

où P_1 est la matrice définie en Q2 et pour $k = 2, \cdots, n-2$, $P_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & Q_k \end{pmatrix}$, en désignant par I_{k-1} la matrice identité d'ordre $k-1$ et par Q_k la matrice orthogonale associée par le procédé de Q2 à la matrice extraite $\left(\left(a_{ij}^{k-1}\right)\right)_{k \le i, j \le n}$.

Le passage de A_{k-1} à A_k est décrit par la procédure suivante.

```
procédure Transformation_Householder(Entrée n, k : Entier;
                                    Entrée_Sortie A : Matrice) ;
Début
    y = 0 ; v = 0 ;
    pour i allant de k + 1 à n faire
    Début
        y_i := a_ik ;
        Fin ;
    s_1 := ‖y‖ ;
    Si a_{k+1,k} > 0 Alors
        s = 1;
    Sinon Si a_{k+1,k} = 0 Alors
            s = 0;
        Sinon
            s = –1 ;
        Fin si;
    Fin si ;
    y_{k+1} = a_{k+1,k} + s*s_1 ;
    s_2 = 1/‖y‖ ;
    y = s_2*y ;
    a_{k+1,k} = –s*s_1 ;
    Pour i allant de k + 2 à n faire
    Début
        a_ik = 0 ;
    Fin ;
    Pour i allant de k + 1 à n faire
    Début
```

Pour j allant de k + 1 à n faire
Début
$v_i = v_i + a_{ij} * y_j$;
Fin ;
Fin ;
$s_3 = \langle y | v \rangle$;
$v = v - s_3 * y$;
Pour i allant de k + 1 à n faire
Début
Pour j allant de k + 1 à n faire
Début
$a_{ij} = a_{ij} - 2*(v_i * y_j + v_j * y_i)$;
Fin ;
Fin ;
Fin ;

Et la procédure de tridiagonalisation par la méthode de Householder s'écrit comme indiqué ci-dessous, une matrice symétrique et tridiagonale étant stockée sous la forme d'un vecteur de dimension n, D_1 représentant la diagonale principale et un vecteur de dimension $n-1$, D_2 représentant la diagonale supérieure.

procédure Householder(Entrée A : Matrice; Sortie D1, D2 :Vecteur) ;
Début
Pour k allant de 1 à n – 2 faire
Début
Transformation_Householder(n, k, A) ;
Fin ;
Pour i allant de 1 à n – 1 faire
Début
$D1_i = a_{ii}$;
$D2_i = a_{i+1,i}$;
Fin ;
$D1_n = a_{nn}$;
Fin ;

Problème 21 : Méthode de Givens pour les matrices tridiagonales

Ce problème utilise les résultats du problème 9.

Si A est une matrice symétrique réelle d'ordre $n \geq 3$*, alors elle est semblable à une matrice tridiagonale T. L'algorithme de Householder (problème 20) permet de calculer effectivement la matrice T. Le calcul des valeurs propres de T nous donne donc celles de A. La méthode décrite ci-dessous est la méthode de Givens ou de « bissection ».*

Soit $A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 0 & b_{n-1} & a_n \end{pmatrix}$ *une matrice, à coefficients réels, tridiagonale symétrique et supposée irréductible, c'est-à-dire telle que :*

$$\forall i \in \{1,2,\cdots,n-1\}\ b_i \neq 0$$

Q1. *Montrer que A admet n valeurs propres réelles simples.*

Pour tout k = 1, 2, ···, n, on note $A_k = \left(\left(a_{ij}\right)\right)_{1\le i,j\le k}$ *la* $k^{ème}$ *sous–matrice principale de A et* $P_k(\lambda) = Dét\left(A_k - \lambda \cdot I_d\right)$ *son polynôme caractéristique.* P_k *admet donc k racines réelles distinctes que l'on range comme suit :*

$$\lambda_1^k < \lambda_2^k < \cdots < \lambda_k^k$$

Les valeurs propres de A sont les $\lambda_j = \lambda_j^n$ *pour j = 1, 2, ···, n.*

Q2. *Montrer qu'on a la relation de récurrence :*

$$\begin{cases} P_0(\lambda) = 1;\ P_1(\lambda) = a_1 - \lambda \\ P_k(\lambda) = (a_k - \lambda)P_{k-1}(\lambda) - {b_{k-1}}^2 P_{k-2}(\lambda) \quad (k = 2,\cdots,n) \end{cases}$$

Q3. *Montrer que pour tout k = 1, ..., n, on a :*

$$\underset{\lambda\to-\infty}{Lim}\ P_k(\lambda) = +\infty$$

et que, pour k = 2, ···, n, les racines de P_{k-1} *séparent celles de* P_k *dans le sens où :*

$$\begin{cases} \lambda_1^k < \lambda_1^{k-1} \\ \lambda_j^{k-1} < \lambda_{j+1}^k < \lambda_{j+1}^{k-1} \quad (1 \le j \le k-2) \\ \lambda_{k-1}^{k-1} < \lambda_k^k \end{cases}$$

On note, pour tout réel λ et tout entier $k \geq 0$ *:*

$$s_k(l) = \begin{cases} Signe\left(P_k(l)\right)\ si\ P_k(l) \neq 0 \\ s_{k-1}(l)\ si\ P_k(l) = 0 \end{cases}$$

et $N_k(\lambda)$ *désigne le nombre de changements de signes entre deux termes consécutifs de la suite* $\left(s_0(\lambda), s_1(\lambda), \cdots, s_k(\lambda)\right)$.

Q4. *(i) Montrer que* $s_1(\lambda) = \begin{cases} 1\ si\ \lambda \in \left]-\infty, \lambda_1^1\right] \\ -1\ si\ \lambda \in \left]\lambda_1^1, +\infty\right[\end{cases}$ *et pour tout k = 2, ···, n :*

$$s_k(\lambda) = \begin{cases} 1\ si\ \lambda \in \left]-\infty, \lambda_1^k\right] \\ (-1)^j\ si\ \lambda \in \left]\lambda_j^k, \lambda_{j+1}^k\right] \quad (1 \le j \le k-1) \\ (-1)^k\ si\ \lambda \in \left]\lambda_k^k, +\infty\right[\end{cases}$$

(ii) Montrer que $N_k(\lambda)$ est égal au nombre de racines de P_k qui sont strictement inférieures à λ.
Q5. *Montrer que toutes les valeurs propres de A sont dans l'intervalle $[a_0, b_0] = [-\|A\|_\infty, \|A\|_\infty]$.*

Pour tout réel λ, on désigne par $N(\lambda)$ le nombre de racines de P_n qui sont strictement inférieures à λ.
Q6. *Soient $i \in \{1,2,\cdots,n\}$ fixé. On note $c_0 = \dfrac{a_0 + b_0}{2}$ le milieu de $[a_0, b_0]$.*
(i) Montrer que si $N(c_0) \geq i$ alors $\lambda_i \in [a_0, c_0]$, sinon $\lambda_i \in [c_0, b_0]$.
(ii) Montrer qu'on peut construire une suite d'intervalles emboîtés $([a_k, b_k])_{k\in IN}$ telle que :

$$\forall k \geq 0 \begin{cases} b_k - a_k = \dfrac{b_0 - a_0}{2^k} \\ \lambda_i \in [a_k, b_k] \end{cases}$$

(iii) Ecrire la procédure correspondante de calcul des valeurs propres de A.
Q7. *Calcul des vecteurs propres par la méthode de la puissance inverse — Soient $i \in \{1,2,\cdots,n\}$ fixé et λ un réel qui vérifie :*

$$0 < |\lambda - \lambda_i| < Inf\{|\lambda - \lambda_j|;\ 1 \leq j \leq n,\ j \neq i\}$$

On se donne un vecteur x_0 non orthogonal au sous espace propre associé à λ_i et on définit la suite de vecteurs $(x_k)_{k\in IN}$ par :

$$\forall k \geq 0,\ (A - \lambda \cdot I_d)x_{k+1} = x_k$$

Calculer $\underset{k\to+\infty}{Lim} \dfrac{Signe(\lambda_i - \lambda)^k}{\|x_k\|} x_k$.

Solution

Q1. Voir problème 9, Q5 (c).
Q2. Il suffit de développer $P_k(\lambda)$ suivant la dernière ligne.
Q3. Le coefficient dominant de $P_k(\lambda)$ est $(-1)^k \lambda^k$. On a alors $\underset{\lambda\to-\infty}{Lim} P_k(\lambda) = \underset{\lambda\to-\infty}{Lim} (-1)^k \lambda^k = +\infty$.

Pour $k = 1$, on a $P_1(\lambda) = a_1 - \lambda$ qui admet une racine réelle $\lambda_1^1 = a_1$.

Pour $k = 2$, on a $P_2(\lambda) = (a_1 - \lambda)(a_2 - \lambda) - b_1^2$, avec $\underset{\lambda\to-\infty}{Lim} P_2(\lambda) = +\infty$, $P_2(\lambda_1^1) = P_2(a_1) = -b_1^2 < 0$ et $\underset{\lambda\to+\infty}{Lim} P_2(\lambda) = +\infty$. Avec le théorème des valeurs intermédiaires, on déduit donc que P_2 admet une racine réelle $\lambda_1^2 < \lambda_1^1$ et une racine réelle $\lambda_2^2 > \lambda_1^1$. On a donc montré que la racine λ_1^1 de P_1 sépare les deux racines de P_2, avec $P_0(\lambda_1^1) = 1 > 0$. Ce qui est schématisé par la figure 3.1.

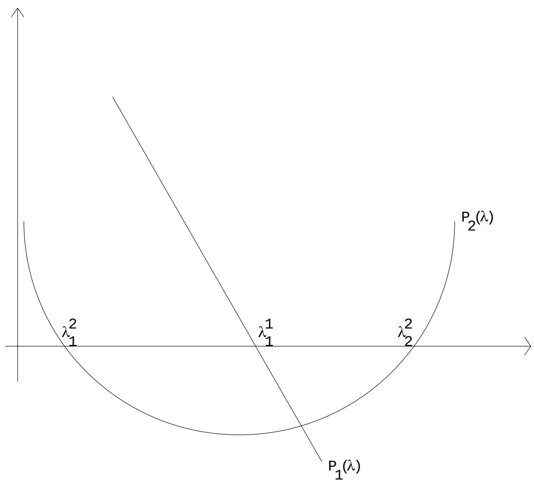

Figure 3.1

Supposons le résultat vrai pour p = 2, ···, $k-1$, avec $P_{k-2}\left(\lambda_j^{k-1}\right)$ du signe de $(-1)^{j-1}$ pour tout j = 1, ···, $k-1$.

Avec $\underset{\lambda\to-\infty}{Lim}\, P_k(\lambda)=+\infty$ et $P_k\left(\lambda_1^{k-1}\right)=-{b_{k-1}}^2 P_{k-2}\left(\lambda_1^{k-1}\right)<0$, on déduit qu'il existe une racine de P_k, $\lambda_1^k<\lambda_1^{k-1}$ telle que $P_{k-1}\left(\lambda_1^k\right)>0$, puisque $P_{k-1}(\lambda)>0$ sur $\left]-\infty,\lambda_1^{k-1}\right[$.

Pour j = 1, 2, ···, $k-2$, on a $P_k\left(\lambda_j^{k-1}\right)=-{b_{k-1}}^2 P_{k-2}\left(\lambda_j^{k-1}\right)$ du signe de $(-1)^j$ et $P_k\left(\lambda_{j+1}^{k-1}\right)=-{b_{k-1}}^2 P_{k-2}\left(\lambda_{j+1}^{k-1}\right)$ du signe de $(-1)^{j+1}$. On en déduit alors que P_k admet une racine $\lambda_{j+1}^k\in\left]\lambda_j^{k-1},\lambda_{j+1}^{k-1}\right[$ (figure 3.2).

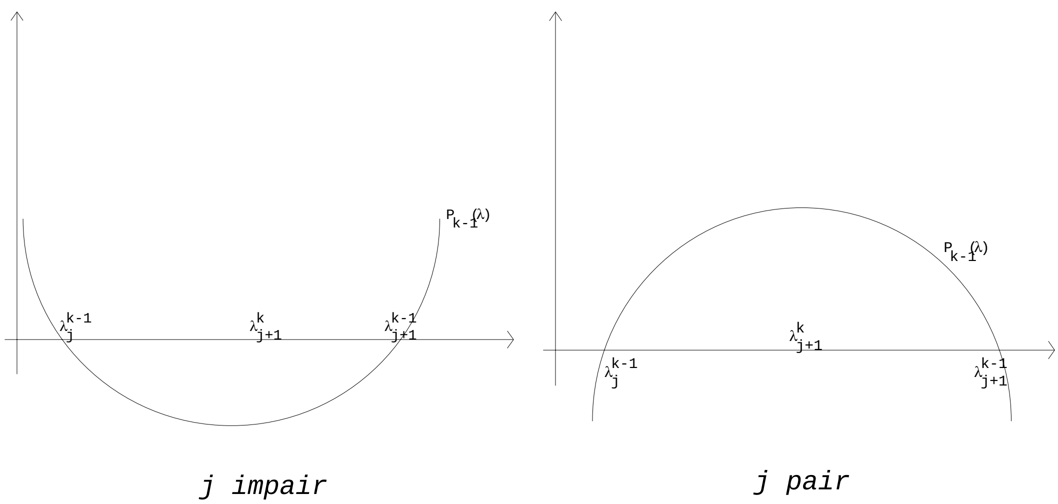

Figure 3.2

Si j est impair, alors $P_{k-1}(\lambda_{j+1}^k) < 0$ et si j est pair alors $P_{k-1}(\lambda_{j+1}^k) > 0$. C'est-à-dire que $P_{k-1}(\lambda_{j+1}^k)$ est du signe de $(-1)^j$.

Enfin $P_k(\lambda_{k-1}^{k-1}) = -b_{k-1}{}^2 P_{k-2}(\lambda_{k-1}^{k-1})$ est du signe de $(-1)^{k-1}$ et $\underset{\lambda\to+\infty}{Lim} P_k(\lambda) = (-1)^k +\infty$. On a donc une racine de P_k, $\lambda_k^k \in \left]\lambda_{k-1}^{k-1}, +\infty\right[$ telle que $P_{k-1}(\lambda_k^k)$ du signe de $(-1)^{k-1}$. D'où le résultat.

Q4. On remarque que si $P_k(\lambda) = 0$ alors $P_{k-1}(\lambda) \neq 0$. Donc, pour tout k = 0, 1, ⋯, n, la fonction s_k est bien définie et à valeurs dans $\{-1,1\}$ avec $s_0(\lambda) = 1$ pour tout réel λ.

(i) Le résultat sur s_1 est évident.

Avec $\underset{\lambda\to-\infty}{Lim} P_k(\lambda) = +\infty$ on déduit que $P_k(\lambda) > 0$ sur $\left]-\infty, \lambda_1^k\right[$ pour tout entier k = 1, 2, ⋯, n. Et avec $\lambda_1^k < \lambda_1^{k-1}$, on conclut que $s_k(\lambda) = 1$ sur $\left]-\infty, \lambda_1^k\right]$.

Pour tout k = 2, 3, ⋯, n et tout j = 1, 2, ⋯, k – 1, P_k est du signe de $(-1)^j$ sur $\left]\lambda_j^k, \lambda_{j+1}^k\right[$. En considérant que $\lambda_{j+1}^k \in \left]\lambda_j^{k-1}, \lambda_{j+1}^{k-1}\right[$, on déduit que $s_k(\lambda) = (-1)^j$ sur $\left]\lambda_j^k, \lambda_{j+1}^k\right]$ pour tout k = 3, ⋯, n et tout j = 1, 2, ⋯, k – 1. Pour k = 2, le résultat est encore valable puisque $s_2(\lambda_2^2) = s_1(\lambda_2^2) = -1$.

Enfin, avec $\underset{\lambda\to+\infty}{Lim} P_k(\lambda) = (-1)^k +\infty$, on déduit que $s_k(\lambda) = (-1)^k$ sur $\left]\lambda_k^k, +\infty\right[$.

(ii) On montre le résultat par récurrence sur $k \geq 1$.

Pour k = 1, on a $(s_0(\lambda), s_1(\lambda)) = \begin{cases} (1,1) \text{ si } \lambda \leq a_1 = \lambda_1^1 \\ (1,-1) \text{ si } \lambda > a_1 = \lambda_1^1 \end{cases}$ et $N_1(\lambda) = \begin{cases} 0 \text{ si } \lambda \leq \lambda_1^1 \\ 1 \text{ si } \lambda > \lambda_1^1 \end{cases}$.

Le résultat est donc vrai pour k = 1.

Supposons le résultat vrai pour $k - 1 \geq 1$ et soit $\lambda \in IR$.

Si $\lambda \leq \lambda_1^k < \lambda_1^j$ (j = 1, 2, ⋯, k – 1), alors $P_j(\lambda) > 0$ pour tout j = 1, 2, ⋯, k et $(s_0(\lambda), s_1(\lambda), \cdots, s_k(\lambda)) = (1,1,\cdots,1)$. C'est-à-dire que $N_k(\lambda) = 0$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Si $\lambda \in \left]\lambda_j^k, \lambda_{j+1}^k\right]$ pour un entier $j \in \{1,2,\cdots,k-1\}$, on distingue alors deux cas en fonction de la position de λ par rapport à $\lambda_j^{k-1} \in \left]\lambda_j^k, \lambda_{j+1}^k\right]$.

Si $\lambda \in \left]\lambda_j^k, \lambda_j^{k-1}\right]$, alors on a j – 1 racines de P_{k-1} strictement inférieures à λ. C'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j - 1$. En tenant compte de $s_{k-1}(\lambda) = (-1)^{j-1}$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Si $\lambda \in \left]\lambda_j^{k-1}, \lambda_{j+1}^k\right]$, alors on a j racines de P_{k-1} strictement inférieures à λ. C'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j$. En tenant compte de $s_{k-1}(\lambda) = (-1)^j$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Enfin si $\lambda > \lambda_k^k > \lambda_{k-1}^{k-1}$, alors, avec l'hypothèse de récurrence $N_{k-1}(\lambda) = k-1$. Avec $s_{k-1}(\lambda) = (-1)^{k-1}$ et $s_k(\lambda) = (-1)^k$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = k$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Q5. Résulte immédiatement de $\rho(A) \le \|A\|_\infty$ où $\rho(A)$ désigne le rayon spectral de A.

Q6. (i) De Q4, on déduit que si $\lambda_i \in [c_0, b_0]$, alors $N(c_0) \le i-1$ et si $\lambda_i \in [a_0, c_0]$, alors $N(c_0) \ge i$. D'où le résultat.

(ii) La construction est immédiate par récurrence en appliquant le raisonnement du (i) à l'intervalle $[a_k, b_k]$.

(iii) Pour la programmation, on stocke une matrice tridiagonale et symétrique dans deux vecteurs a et b. On désigne par c le vecteur de $\mathbb{R}^{n-1}$ défini par $c_i = b_i^2$ pour tout $i = 1, 2, \cdots, n-1$ (on évite de calculer plusieurs fois les b_i^2).

On définit tout d'abord la fonction $\lambda \mapsto N(\lambda)$.

```
fonction Changement_Signe(Entrée n : Entier ; a, c : Vecteur ; λ : Réel) : Réel ;
Début
    Nλ = 0; {Nombre de changements de signes}
    P0λ = 1.0 ;
    P1λ = a1 – λ ;
    sλ = 1;
    Si (Signe(P1λ) = – sλ) alors
        sλ = – sλ;
        Nλ  = Nλ + 1 ;
    Fin Si;
    Pour i allant de 2 à n faire
    Début
        P2λ = (ai – λ)* P1λ – ci–1* P0λ ;
        Si (Signe(P2λ) = – sλ) alors
            sλ = – sλ;
            Nλ = Nλ + 1 ;
        Fin Si ;
        P0λ = P1λ ;
        P1λ = P2λ ;
    Fin ;
    Changement_Signe = Nλ ;
Fin ;
```

On définit ensuite une procédure qui permet de calculer la valeur propre numéro i.

```
procédure Une_Valeur_Propre(Entrée n, i : Entier ; a, c : Vecteur ;
                                Entrée_Sortie ak, bk : Réel ; Sortie λ : Réel) ;
Début
    Répéter
        λ = (ak + bk)*0.5 ;
```

```
        Si Changement_Signe(a, c, λ) ≥ i alors
            bk = λ ;
        Sinon
            ak = λ ;
        Fin Si ;
    Jusqu'à |a_k − b_k| < ε ;
Fin ;
```

Et la procédure de Givens de recherche de toutes les valeurs propres de A s'écrit alors comme suit.

```
procédure Givens(Entrée n : Entier ; a, b : Vecteur ; Sortie λ : Vecteur) ;
Début
    Pour i allant de 1 à n – 1 faire
    Début
        c_i = b_i*b_i ;
    Fin ;
    M = |a_1| + |b_1| ;
    Pour i allant de 2 à n – 1 faire
    Début
        M = Max ( M ,|a_i| + |b_{i−1}| + |b_i|) ;
    Fin ;
    M = Max ( M ,|a_n| + |b_{n−1}|) ;
    ak = –M ;
    bk = M ;
    Pour i allant de 1 à n faire
    Début
        Une_Valeur_Propre(n, i, a, c, ak, bk, λ_i) ;
        ak = bk ;
        bk = M ;
    Fin ;
Fin ;
```

Q7. On remarque que si $x_{k+1} = 0$, alors $x_k = 0$. On déduit donc par récurrence que si $x_0 \neq 0$ alors $x_k \neq 0$ pour tout entier k.

Soit $\{f_1, f_2, \cdots, f_n\}$ une base orthonormée de vecteurs propres avec $Af_i = \lambda_i f_i$ pour tout $i = 1, 2, \cdots, n$.

Pour tout entier k, on note $x_k = \sum_{j=1}^{n} x_j^k f_j$.

Avec $(A - \lambda \cdot I_d)x_{k+1} = x_k$, on déduit que $x_j^{k+1} = \dfrac{1}{\lambda_j - \lambda} x_j^k$ et par récurrence $x_j^k = \dfrac{1}{(\lambda_j - \lambda)^k} x_j^0$.

```
        Pour j allant de k + 1 à n faire
        Début
            v_i = v_i + a_ij*y_j ;
        Fin ;
    Fin ;
    s_3 = ⟨y|v⟩ ;
    v = v – s_3*y ;
    Pour i allant de k + 1 à n faire
    Début
        Pour j allant de k + 1 à n faire
        Début
            a_ij = a_ij – 2*(v_i*y_j + v_j*y_i) ;
        Fin ;
    Fin ;
Fin ;
```

Et la procédure de tridiagonalisation par la méthode de Householder s'écrit comme indiqué ci-dessous, une matrice symétrique et tridiagonale étant stockée sous la forme d'un vecteur de dimension n, D_1 représentant la diagonale principale et un vecteur de dimension $n-1$, D_2 représentant la diagonale supérieure.

```
procédure Householder(Entrée A : Matrice; Sortie D1, D2 :Vecteur) ;
Début
    Pour k allant de 1 à n – 2 faire
    Début
        Transformation_Householder(n, k, A) ;
    Fin ;
    Pour i allant de 1 à n – 1 faire
    Début
        D1_i = a_ii ;
        D2_i = a_i+1,i ;
    Fin ;
    D1_n = a_nn ;
Fin ;
```

Problème 21 : Méthode de Givens pour les matrices tridiagonales

Ce problème utilise les résultats du problème 9.

Si A est une matrice symétrique réelle d'ordre $n \geq 3$, alors elle est semblable à une matrice tridiagonale T. L'algorithme de Householder (problème 20) permet de calculer effectivement la matrice T. Le calcul des valeurs propres de T nous donne donc celles de A. La méthode décrite ci-dessous est la méthode de Givens ou de « bissection ».

Soit $A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 0 & b_{n-1} & a_n \end{pmatrix}$ *une matrice, à coefficients réels, tridiagonale symétrique et supposée irréductible, c'est-à-dire telle que :*

$$\forall i \in \{1,2,\cdots,n-1\}\ b_i \neq 0$$

Q1. *Montrer que A admet n valeurs propres réelles simples.*

Pour tout k = 1, 2, ···, n, on note $A_k = \left((a_{ij})\right)_{1\le i,j\le k}$ *la* $k^{ème}$ *sous–matrice principale de A et* $P_k(\lambda) = Dét\left(A_k - \lambda \cdot I_d\right)$ *son polynôme caractéristique.* P_k *admet donc k racines réelles distinctes que l'on range comme suit :*

$$\lambda_1^k < \lambda_2^k < \cdots < \lambda_k^k$$

Les valeurs propres de A sont les $\lambda_j = \lambda_j^n$ *pour j = 1, 2, ···, n.*

Q2. *Montrer qu'on a la relation de récurrence :*

$$\begin{cases} P_0(\lambda) = 1;\ P_1(\lambda) = a_1 - \lambda \\ P_k(\lambda) = (a_k - \lambda)P_{k-1}(\lambda) - {b_{k-1}}^2 P_{k-2}(\lambda) \quad (k = 2,\cdots,n) \end{cases}$$

Q3. *Montrer que pour tout k = 1, ..., n, on a :*

$$\underset{\lambda\to-\infty}{Lim}\, P_k(\lambda) = +\infty$$

et que, pour k = 2, ···, n, les racines de P_{k-1} *séparent celles de* P_k *dans le sens où :*

$$\begin{cases} \lambda_1^k < \lambda_1^{k-1} \\ \lambda_j^{k-1} < \lambda_{j+1}^k < \lambda_{j+1}^{k-1} \quad (1 \le j \le k-2) \\ \lambda_{k-1}^{k-1} < \lambda_k^k \end{cases}$$

On note, pour tout réel λ et tout entier $k \geq 0$ *:*

$$s_k(l) = \begin{cases} Signe\left(P_k(l)\right) \ si\ P_k(l) \neq 0 \\ s_{k-1}(l)\ si\ P_k(l) = 0 \end{cases}$$

et $N_k(\lambda)$ *désigne le nombre de changements de signes entre deux termes consécutifs de la suite* $\left(s_0(\lambda), s_1(\lambda), \cdots, s_k(\lambda)\right)$.

Q4. *(i) Montrer que* $s_1(\lambda) = \begin{cases} 1\ si\ \lambda \in \left]-\infty, \lambda_1^1\right] \\ -1\ si\ \lambda \in \left]\lambda_1^1, +\infty\right[\end{cases}$ *et pour tout k = 2, ···, n :*

$$s_k(\lambda) = \begin{cases} 1\ si\ \lambda \in \left]-\infty, \lambda_1^k\right] \\ (-1)^j\ si\ \lambda \in \left]\lambda_j^k, \lambda_{j+1}^k\right] \quad (1 \le j \le k-1) \\ (-1)^k\ si\ \lambda \in \left]\lambda_k^k, +\infty\right[\end{cases}$$

(ii) Montrer que $N_k(\lambda)$ *est égal au nombre de racines de* P_k *qui sont strictement inférieures à* λ.

Q5. *Montrer que toutes les valeurs propres de A sont dans l'intervalle* $[a_0, b_0] = [-\|A\|_\infty, \|A\|_\infty]$.

Pour tout réel λ, *on désigne par* $N(\lambda)$ *le nombre de racines de* P_n *qui sont strictement inférieures à* λ.

Q6. *Soient* $i \in \{1,2,\cdots,n\}$ *fixé. On note* $c_0 = \dfrac{a_0+b_0}{2}$ *le milieu de* $[a_0, b_0]$.

(i) Montrer que si $N(c_0) \geq i$ *alors* $\lambda_i \in [a_0, c_0]$, *sinon* $\lambda_i \in [c_0, b_0]$.

(ii) Montrer qu'on peut construire une suite d'intervalles emboîtés $([a_k, b_k])_{k\in IN}$ *telle que :*

$$\forall k \geq 0 \begin{cases} b_k - a_k = \dfrac{b_0 - a_0}{2^k} \\ \lambda_i \in [a_k, b_k] \end{cases}$$

(iii) Ecrire la procédure correspondante de calcul des valeurs propres de A.

Q7. *Calcul des vecteurs propres par la méthode de la puissance inverse — Soient* $i \in \{1,2,\cdots,n\}$ *fixé et* λ *un réel qui vérifie :*

$$0 < |\lambda - \lambda_i| < Inf\left\{|\lambda - \lambda_j|;\ 1 \leq j \leq n,\ j \neq i\right\}$$

On se donne un vecteur x_0 *non orthogonal au sous espace propre associé à* λ_i *et on définit la suite de vecteurs* $(x_k)_{k\in IN}$ *par :*

$$\forall k \geq 0,\ (A - \lambda \cdot I_d)x_{k+1} = x_k$$

Calculer $\displaystyle\lim_{k\to+\infty} \frac{Signe(\lambda_i - \lambda)^k}{\|x_k\|} x_k$.

Solution

Q1. Voir problème 9, Q5 (c).

Q2. Il suffit de développer $P_k(\lambda)$ suivant la dernière ligne.

Q3. Le coefficient dominant de $P_k(\lambda)$ est $(-1)^k \lambda^k$. On a alors $\displaystyle\lim_{\lambda\to-\infty} P_k(\lambda) = \lim_{\lambda\to-\infty} (-1)^k \lambda^k = +\infty$.

Pour $k = 1$, on a $P_1(\lambda) = a_1 - \lambda$ qui admet une racine réelle $\lambda_1^1 = a_1$.

Pour $k = 2$, on a $P_2(\lambda) = (a_1 - \lambda)(a_2 - \lambda) - b_1^{\,2}$, avec $\displaystyle\lim_{\lambda\to-\infty} P_2(\lambda) = +\infty$, $P_2(\lambda_1^1) = P_2(a_1) = -b_1^{\,2} < 0$ et $\displaystyle\lim_{\lambda\to+\infty} P_2(\lambda) = +\infty$. Avec le théorème des valeurs intermédiaires, on déduit donc que P_2 admet une racine réelle $\lambda_1^2 < \lambda_1^1$ et une racine réelle $\lambda_2^2 > \lambda_1^1$. On a donc montré que la racine λ_1^1 de P_1 sépare les deux racines de P_2, avec $P_0(\lambda_1^1) = 1 > 0$. Ce qui est schématisé par la figure 3.1.

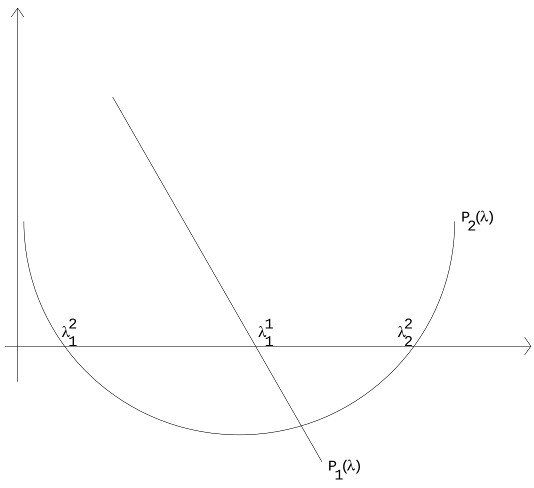

Figure 3.1

Supposons le résultat vrai pour p = 2, ···, $k-1$, avec $P_{k-2}\left(\lambda_j^{k-1}\right)$ du signe de $(-1)^{j-1}$ pour tout j = 1, ···, $k-1$.

Avec $\underset{\lambda\to-\infty}{Lim}\, P_k(\lambda)=+\infty$ et $P_k\left(\lambda_1^{k-1}\right)=-b_{k-1}{}^2 P_{k-2}\left(\lambda_1^{k-1}\right)<0$, on déduit qu'il existe une racine de P_k, $\lambda_1^k<\lambda_1^{k-1}$ telle que $P_{k-1}\left(\lambda_1^k\right)>0$, puisque $P_{k-1}(\lambda)>0$ sur $\left]-\infty,\lambda_1^{k-1}\right[$.

Pour j = 1, 2, ···, $k-2$, on a $P_k\left(\lambda_j^{k-1}\right)=-b_{k-1}{}^2 P_{k-2}\left(\lambda_j^{k-1}\right)$ du signe de $(-1)^j$ et $P_k\left(\lambda_{j+1}^{k-1}\right)=-b_{k-1}{}^2 P_{k-2}\left(\lambda_{j+1}^{k-1}\right)$ du signe de $(-1)^{j+1}$. On en déduit alors que P_k admet une racine $\lambda_{j+1}^k\in\left]\lambda_j^{k-1},\lambda_{j+1}^{k-1}\right[$ (figure 3.2).

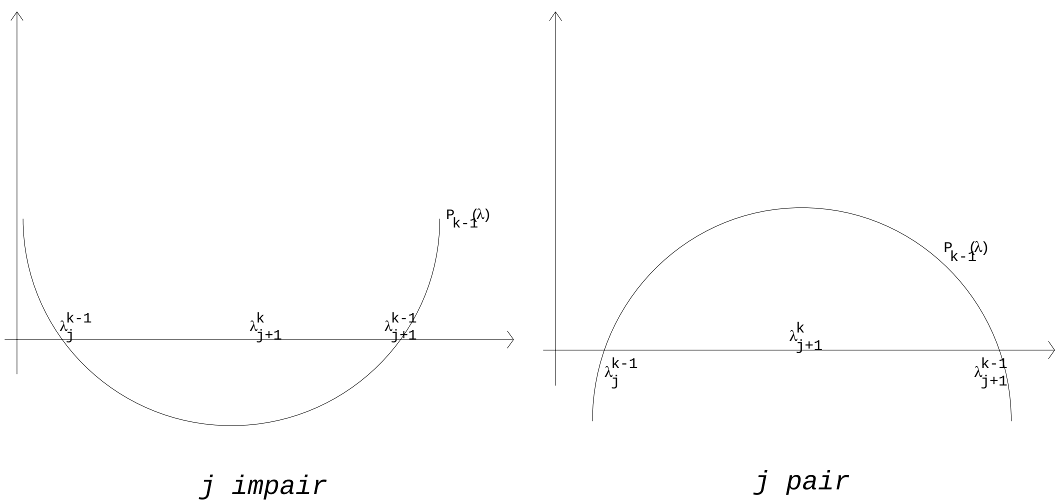

Figure 3.2

Si j est impair, alors $P_{k-1}(\lambda_{j+1}^k) < 0$ et si j est pair alors $P_{k-1}(\lambda_{j+1}^k) > 0$. C'est-à-dire que $P_{k-1}(\lambda_{j+1}^k)$ est du signe de $(-1)^j$.

Enfin $P_k(\lambda_{k-1}^{k-1}) = -b_{k-1}{}^2 P_{k-2}(\lambda_{k-1}^{k-1})$ est du signe de $(-1)^{k-1}$ et $\underset{\lambda\to+\infty}{Lim} P_k(\lambda) = (-1)^k +\infty$. On a donc une racine de P_k, $\lambda_k^k \in \left]\lambda_{k-1}^{k-1}, +\infty\right[$ telle que $P_{k-1}(\lambda_k^k)$ du signe de $(-1)^{k-1}$. D'où le résultat.

Q4. On remarque que si $P_k(\lambda) = 0$ alors $P_{k-1}(\lambda) \neq 0$. Donc, pour tout k = 0, 1, ⋯, n, la fonction s_k est bien définie et à valeurs dans $\{-1,1\}$ avec $s_0(\lambda) = 1$ pour tout réel λ.

(i) Le résultat sur s_1 est évident.

Avec $\underset{\lambda\to-\infty}{Lim} P_k(\lambda) = +\infty$ on déduit que $P_k(\lambda) > 0$ sur $\left]-\infty, \lambda_1^k\right[$ pour tout entier k = 1, 2, ⋯, n. Et avec $\lambda_1^k < \lambda_1^{k-1}$, on conclut que $s_k(\lambda) = 1$ sur $\left]-\infty, \lambda_1^k\right]$.

Pour tout k = 2, 3, ⋯, n et tout j = 1, 2, ⋯, k – 1, P_k est du signe de $(-1)^j$ sur $\left]\lambda_j^k, \lambda_{j+1}^k\right[$. En considérant que $\lambda_{j+1}^k \in \left]\lambda_j^{k-1}, \lambda_{j+1}^{k-1}\right[$, on déduit que $s_k(\lambda) = (-1)^j$ sur $\left]\lambda_j^k, \lambda_{j+1}^k\right]$ pour tout k = 3, ⋯, n et tout j = 1, 2, ⋯, k – 1. Pour k = 2, le résultat est encore valable puisque $s_2(\lambda_2^2) = s_1(\lambda_2^2) = -1$.

Enfin, avec $\underset{\lambda\to+\infty}{Lim} P_k(\lambda) = (-1)^k +\infty$, on déduit que $s_k(\lambda) = (-1)^k$ sur $\left]\lambda_k^k, +\infty\right[$.

(ii) On montre le résultat par récurrence sur $k \geq 1$.

Pour $k = 1$, on a $(s_0(\lambda), s_1(\lambda)) = \begin{cases} (1,1) \text{ si } \lambda \leq a_1 = \lambda_1^1 \\ (1,-1) \text{ si } \lambda > a_1 = \lambda_1^1 \end{cases}$ et $N_1(\lambda) = \begin{cases} 0 \text{ si } \lambda \leq \lambda_1^1 \\ 1 \text{ si } \lambda > \lambda_1^1 \end{cases}$.

Le résultat est donc vrai pour k = 1.

Supposons le résultat vrai pour $k - 1 \geq 1$ et soit $\lambda \in IR$.

Si $\lambda \leq \lambda_1^k < \lambda_1^j$ (j = 1, 2, ⋯, k – 1), alors $P_j(\lambda) > 0$ pour tout j = 1, 2, ⋯, k et $(s_0(\lambda), s_1(\lambda), \cdots, s_k(\lambda)) = (1,1,\cdots,1)$. C'est-à-dire que $N_k(\lambda) = 0$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Si $\lambda \in \left]\lambda_j^k, \lambda_{j+1}^k\right]$ pour un entier $j \in \{1,2,\cdots,k-1\}$, on distingue alors deux cas en fonction de la position de λ par rapport à $\lambda_j^{k-1} \in \left]\lambda_j^k, \lambda_{j+1}^k\right]$.

Si $\lambda \in \left]\lambda_j^k, \lambda_j^{k-1}\right]$, alors on a j – 1 racines de P_{k-1} strictement inférieures à λ. C'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j - 1$. En tenant compte de $s_{k-1}(\lambda) = (-1)^{j-1}$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Si $\lambda \in \left]\lambda_j^{k-1}, \lambda_{j+1}^k\right]$, alors on a j racines de P_{k-1} strictement inférieures à λ. C'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j$. En tenant compte de $s_{k-1}(\lambda) = (-1)^j$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Enfin si $\lambda > \lambda_k^k > \lambda_{k-1}^{k-1}$, alors, avec l'hypothèse de récurrence $N_{k-1}(\lambda) = k-1$. Avec $s_{k-1}(\lambda) = (-1)^{k-1}$ et $s_k(\lambda) = (-1)^k$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = k$ et c'est bien le nombre de racines de P_k strictement inférieures à λ.

Q5. Résulte immédiatement de $\rho(A) \leq \|A\|_\infty$ où $\rho(A)$ désigne le rayon spectral de A.

Q6. (i) De Q4, on déduit que si $\lambda_i \in [c_0, b_0]$, alors $N(c_0) \leq i-1$ et si $\lambda_i \in [a_0, c_0]$, alors $N(c_0) \geq i$. D'où le résultat.

(ii) La construction est immédiate par récurrence en appliquant le raisonnement du (i) à l'intervalle $[a_k, b_k]$.

(iii) Pour la programmation, on stocke une matrice tridiagonale et symétrique dans deux vecteurs a et b. On désigne par c le vecteur de $\mathbb{R}^{n-1}$ défini par $c_i = b_i^2$ pour tout $i = 1, 2, \cdots, n-1$ (on évite de calculer plusieurs fois les b_i^2).

On définit tout d'abord la fonction $\lambda \mapsto N(\lambda)$.

```
fonction Changement_Signe(Entrée n : Entier ; a, c : Vecteur ; λ : Réel) : Réel ;
Début
   Nλ = 0; {Nombre de changements de signes}
   P0λ = 1.0 ;
   P1λ = a1 – λ ;
   sλ = 1;
   Si (Signe(P1λ) = – sλ) alors
      sλ = – sλ;
      Nλ  = Nλ + 1 ;
   Fin Si;
   Pour i allant de 2 à n faire
   Début
      P2λ = (ai – λ)* P1λ – ci–1* P0λ ;
      Si (Signe(P2λ) = – sλ) alors
         sλ = – sλ;
         Nλ = Nλ + 1 ;
      Fin Si ;
      P0λ = P1λ ;
      P1λ = P2λ ;
   Fin ;
   Changement_Signe = Nλ ;
Fin ;
```

On définit ensuite une procédure qui permet de calculer la valeur propre numéro i.

```
procédure Une_Valeur_Propre(Entrée n, i : Entier ; a, c : Vecteur ;
                                        Entrée_Sortie ak, bk : Réel ; Sortie λ : Réel) ;
Début
   Répéter
      λ = (ak + bk)*0.5 ;
```

Si Changement_Signe(a, c, λ) ≥ i alors
bk = λ ;
Sinon
ak = λ ;
Fin Si ;
Jusqu'à $|a_k - b_k| < \varepsilon$ *;*
Fin ;

Et la procédure de Givens de recherche de toutes les valeurs propres de *A* s'écrit alors comme suit.

procédure Givens(Entrée n : Entier ; a, b : Vecteur ; Sortie λ : Vecteur) ;
Début
Pour i allant de 1 à n – 1 faire
Début
$c_i = b_i * b_i$ *;*
Fin ;
$M = |a_1| + |b_1|$ *;*
Pour i allant de 2 à n – 1 faire
Début
$M = Max\left(M, |a_i| + |b_{i-1}| + |b_i|\right)$ *;*
Fin ;
$M = Max\left(M, |a_n| + |b_{n-1}|\right)$ *;*
ak = –M ;
bk = M ;
Pour i allant de 1 à n faire
Début
Une_Valeur_Propre(n, i, a, c, ak, bk, λ_i*) ;*
ak = bk ;
bk = M ;
Fin ;
Fin ;

Q7. On remarque que si $x_{k+1} = 0$, alors $x_k = 0$. On déduit donc par récurrence que si $x_0 \neq 0$ alors $x_k \neq 0$ pour tout entier *k*.

Soit $\{f_1, f_2, \cdots, f_n\}$ une base orthonormée de vecteurs propres avec $Af_i = \lambda_i f_i$ pour tout *i* = 1, 2, ···, *n*.

Pour tout entier *k*, on note $x_k = \sum_{j=1}^{n} x_j^k f_j$.

Avec $(A - \lambda \cdot I_d)x_{k+1} = x_k$, on déduit que $x_j^{k+1} = \dfrac{1}{\lambda_j - \lambda} x_j^k$ et par récurrence $x_j^k = \dfrac{1}{(\lambda_j - \lambda)^k} x_j^0$.

On peut alors écrire que $x_k = \dfrac{1}{(\lambda_i - \lambda)^k} y_k$, où y_k est le vecteur de composantes $y_j^k = \left(\dfrac{\lambda_i - \lambda}{\lambda_j - \lambda}\right)^k x_j^0$.

On a alors $\underset{k \to +\infty}{Lim}\, y_j^k = \begin{cases} 0 \text{ si } j \neq i \\ x_i^0 \text{ si } j = i \end{cases}$, avec $x_i^0 \neq 0$. C'est-à-dire que $\underset{k \to +\infty}{Lim}\, y_k = x_i^0 f_i$.

On déduit alors que $\underset{k \to +\infty}{Lim} \dfrac{Signe(\lambda_i - \lambda)^k}{\|x_k\|} x_k = \underset{k \to +\infty}{Lim} \dfrac{1}{\|y_k\|} y_k = Signe(x_i^0) f_i$.

Remarque 1 — Si la matrice A n'est pas irréductible, on peut alors la découper en blocs de matrices irréductibles et l'algorithme de Givens permet de calculer les valeurs propres.

Remarque 2 — Pour calculer les valeurs propres et vecteurs propres d'une matrice symétrique réelle la méthode de Givens et Householder consiste à réduire la matrice sous tridiagonale dans un premier temps puis à rechercher les valeurs et vecteurs propres de la matrice tridiagonale obtenue par la méthode de Givens et de la puissance inverse (les systèmes linéaires à résoudre sont tridiagonaux).

Problème 22 : Une méthode homotopique de calcul des valeurs propres d'une matrice tridiagonale

Ce problème utilise les notations du problème 21 et les résultats des problèmes 9 et 21.

Q1. *Montrer que la matrice A admet les mêmes valeurs propres que la matrice :*

$$B = \begin{pmatrix} a_1 & b_1^2 & 0 & 0 & \cdots & 0 \\ 1 & a_2 & b_2^2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & a_{n-1} & b_{n-1}^2 \\ 0 & \cdots & 0 & 0 & 1 & a_n \end{pmatrix}$$

Dans tout ce qui suit, on suppose que :

$$A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 0 & 1 & a_n \end{pmatrix}$$

avec :

$$a_i, b_i \in IR,\ b_i > 0$$

On note :

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

On définit alors une « homotopie » de D vers A en posant :

$$\forall t \in [0,1]\ A(t) = (1-t)D + tA$$

Q2. *Calculer les valeurs propres de D.*

Q3. *Montrer que, pour tout t dans* $[0,1]$*,* $A(t)$ *admet n valeurs propres réelles simples.*

Q4. *Montrer qu'il existe une fonction continûment dérivable*

$$f = (f_1, \cdots, f_n) \colon [0,1] \to IR^n$$

telle que le polynôme caractéristique de $A(t)$ *s'écrive :*

$$\forall t \in [0,1],\ P(\lambda, t) = \prod_{j=1}^{n} \left(\lambda - f_j(t)\right)$$

Pour tout t dans $[0,1]$ *et j = 1, 2, ···, n, on pose :*

$$\varphi_j(t) = P\left(f_j(t), t\right)$$

et on a alors $\varphi_j(t) = 0$ *sur* $[0,1]$.

Q5. *Montrer que pour tout j = 1, 2, ···, n,* f_j *est solution sur* $[0,1]$ *d'une équation différentielle avec condition initiale (problème de Cauchy).*

Q6. *En utilisant la méthode Runge–Kutta d'ordre 4, donner un algorithme de calcul de toutes les valeurs propres de la matrice A.*

Q7. *En déduire un algorithme de calcul des valeurs propres et des vecteurs propres d'une matrice symétrique réelle n'ayant que des racines simples.*

Solution

Q1. Le polynôme caractéristique de *B* se calcule en utilisant la même récurrence que celle qui définit le polynôme caractéristique de *A* (problème 9, Q5).

Q2. Dans le problème 3 Q5, on a vu que les valeurs propres de *D* sont les $\lambda_k = 2Cos(k\theta)$ $(1 \le k \le n)$ où $\theta = \dfrac{\pi}{n+1}$.

Q3. Pour tout *t* dans $[0,1]$, on a :

$$A(t) = \begin{pmatrix} ta_1 & 1-t+tb_1 & 0 & 0 & \cdots & 0 \\ 1 & ta_2 & 1-t+tb_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & ta_{n-1} & 1-t+tb_{n-1} \\ 0 & \cdots & 0 & 0 & 1 & ta_n \end{pmatrix}$$

Cette matrice admet *n* valeurs propres réelles simples puisqu'elle est irréductible avec $1-t+tb_i > 0$ pour tout $t \in [0,1]$ et tout *i* = 1, 2, ..., *n* (problème 9, Q5, (c)).

Q4. Soit $t_0 \in [0,1]$ et λ_0 une valeur propre de $A(t_0)$, on a alors :

$$P(\lambda_0,t_0)=0,\ \frac{\partial P}{\partial \lambda}(\lambda_0,t_0)\neq 0$$

Le théorème des fonctions implicites nous dit alors qu'il existe un voisinage ouvert V_0 de t_0 dans $[0,1]$ et une fonction continûment dérivable :

$$f=(f_1,\cdots,f_n):\ V_0 \to IR^n$$

tels que :

$$\forall t \in V_0,\ P(\lambda,t)=\prod_{k=1}^{n}(\lambda-f_k(t))$$

Les valeurs propres de $A(t)$ étant toutes simples, on peut supposer que $f_1(t)<f_2(t)<\cdots<f_n(t)$. Avec cette condition la fonction f est unique.

L'intervalle $[0,1]$ étant compact, on peut trouver une partition du type :

$$[0,1]=[a_0,b_0[\cup[a_1,b_1[\cup\cdots\cup[a_p,b_p]$$

où :

$$0=a_0<b_0=a_1<b_1=a_2<\cdots<b_{p-1}=a_p<b_p=1$$

telle pour tout k = 0, 1, ···, p, il existe une unique fonction continûment dérivable

$$f^k=(f_1^k,\cdots,f_n^k):\ [a_k,b_k]\to IR^n$$

avec :

$$\forall t\in[a_k,b_k],\ \begin{cases} P(\lambda,t)=\prod_{j=1}^{n}(\lambda-f_j^k(t)) \\ f_1^k(t)<f_2^k(t)<\cdots<f_n^k(t)\end{cases}$$

La fonction f définie sur $[0,1]$ par :

$$f(t)=(f_1^k(t),\cdots,f_n^k(t)) \text{ si } t\in[a_k,b_k]$$

est alors continûment dérivable.

Q5. En dérivant la relation $\varphi_j(t)=0$ par rapport à t, pour tout j = 1, 2, ···, n, on déduit alors que f_j est solution du problème de Cauchy :

$$\begin{cases} f_j(0)=a_j \\ f_j'(t)=-\dfrac{\dfrac{\partial P}{\partial t}(f_j(t),t)}{\dfrac{\partial P}{\partial \lambda}(f_j(t),t)} \quad t\in[0,1]\end{cases}$$

La simplicité des valeurs propres de $A(t)$ nous garantit que $\dfrac{\partial P}{\partial \lambda}(f_j(t),t)\neq 0$ sur $[0,1]$.

Q6. Le calcul du polynôme caractéristique $P(\lambda,t)$ se fait en utilisant la récurrence :

$$(1)\ \begin{cases} P_0(\lambda,t)=1;\ P_1(\lambda,t)=ta_1-\lambda \\ P_k(\lambda,t)=(ta_k-\lambda)P_{k-1}(\lambda,t)-(1-t+tb_{k-1})P_{k-2}(\lambda,t)\ (k=2,\cdots,n)\end{cases}$$

En dérivant (2) par rapport à λ et t, on obtient l'algorithme suivant de calcul des dérivées partielles de P :

$$(2)\begin{cases}\dfrac{\partial P_0}{\partial\lambda}(\lambda,t)=0;\ \dfrac{\partial P_0}{\partial t}(\lambda,t)=0;\\ \dfrac{\partial P_1}{\partial\lambda}(\lambda,t)=-1;\ \dfrac{\partial P_1}{\partial t}(\lambda,t)=a_1;\\ \dfrac{\partial P_k}{\partial\lambda}(\lambda,t)=(ta_k-\lambda)\dfrac{\partial P_{k-1}}{\partial\lambda}(\lambda,t)-P_{k-1}(\lambda,\mathrm{t})-(1-t+tb_{k-1})\dfrac{\partial P_{k-2}}{\partial\lambda}(\lambda,t)\\ \dfrac{\partial P_k}{\partial t}(\lambda,t)=(ta_k-\lambda)\dfrac{\partial P_{k-1}}{\partial t}(\lambda,t)+a_kP_{k-1}(\lambda,t)\\ \qquad\qquad -(1-t+tb_{k-1})\dfrac{\partial P_{k-2}}{\partial t}(\lambda,t)-(b_{k-1}-1)P_{k-2}(\lambda,\mathrm{t})\end{cases}$$

Pour la programmation, la matrice tridiagonale symétrique est définie par deux vecteurs a et b. Une fois ces vecteurs entrés on fait les transformations $b_i=b_i^2$ pour tout $i = 1, 2, \cdots, n-1$.

Le calcul des dérivées partielles $P_\lambda=\dfrac{\partial P}{\partial\lambda}(\lambda,t)$ et $P_t=\dfrac{\partial P}{\partial t}(\lambda,t)$ se fait avec la procédure :

Procédure Dérivées_Polynôme_Caractéristique(Entrée n : Entier; a : Vecteur(1..n);
b : Vecteur(1..n – 1); t, λ : Réels;
Sortie P_t, P_λ: Réels) ;

Début
  $P_0 = 1$;
  $P_{\lambda,0} = 0$;
  $P_{t,0} = 0$;
  $P_1 = t\cdot a_1 - \lambda$;
  $P_{\lambda,1} = -1$; $P_{t,1} = a_1$;
  Pour k allant de 2 à n faire
  Début
    $P_2 = (t\cdot a_k - \lambda)\cdot P_1 - (1- t + t\cdot b_{k-1})\cdot P_0$;
    $P_{\lambda,2} = (t\cdot a_k - \lambda)\cdot P_{\lambda,1} - P_1 - (1- t + t\cdot b_{k-1})\cdot P_{\lambda,0}$;
    $P_{t,2} = (t\cdot a_k - \lambda)\cdot P_{t,1} + a_k\cdot P_1 - (b_{k-1} - 1)\cdot P_0 - (1- t + t\cdot b_{k-1})\cdot P_{t,0}$;
    $P_0 = P_1$; $P_1 = P_2$;
    $P_{\lambda,0} = P_{\lambda,1}$; $P_{\lambda,1} = P_{\lambda,2}$;
    $P_{t,0} = P_{t,1}$; $P_{t,1} = P_{t,2}$;
  Fin ;
  $P_\lambda = P_{\lambda,2}$; $P_t = P_{t,2}$;
Fin ;

On peut ensuite définir la fonction F définissant l'équation différentielle à résoudre.

Fonction F(Entrée n : Entier; a : Vecteur(1..n); b : Vecteur(1..n – 1); t, λ : Réels) : Réel ;

```
Début
    Dérivées_Polynôme_Caractéristique(n, a, b, t, λ, P_t, P_λ) ;
    F = –P_λ/P_t ;
Fin ;
```

Pour la résolution des équations différentielles, on se donne un pas de calcul $h = \frac{1}{p}$ sur $[0,1]$ où p est un entier positif. En vue de tracer les graphes des fonctions valeurs propres $t \mapsto f_j(t)$, on peut stocker les valeurs propres des matrices $A(t_k)$, où $t_k = k \cdot h \ (0 \le \mathrm{k} \le \mathrm{p})$, dans un tableau L à p + 1 lignes et n colonnes. C'est-à-dire que $L_{k,j} \approx f_j(t_k) \ (0 \le k \le p,\ 1 \le j \le n)$.

```
Procédure Valeurs_Propres(Entrée n : Entier ; a : Vecteur(1..n) ; b : Vecteur(1..n–1) ;
                          Sortie L : Matrice(0..p,1..n)) ;
Début
    h = 1/p ;
    h_2 = 0.5·h ;
    θ = π/(n + 1) ;
    θ_j = 0 ;
    Pour j allant de 1 à n faire
    Début
        t = 0 ;
        θ_j = θ_j + θ;
        L_0,j = 2Cos(θ_j) ;
        Pour k allant de 1 à p faire
        Début
            K_1 = h·F(n, a, b, t, L_k–1,j) ;
            t = t + h_2 ;
            K_2 = h·F(n, a, b, t, L_k–1,j + 0.5·K_1) ;
            K_3 = h·F(n, a, b, t, L_k–1,j + 0.5·K_2) ;
            t = t + h_2 ;
            K_4 = h·F(n, a, b, t, L_k–1,j + K_3) ;
            L_k,j = L_k–1,j + (K_1 + 2·(K_2 + K_3) + K_4)/6 ;
        Fin ;
    Fin ;
Fin ;
```

Les valeurs propres cherchées sont alors les $L_{p,j}$, pour $j = 1, 2, \cdots, n$.

Q7. Pour calculer les valeurs propres d'une matrice symétrique réelle on commence par réduire la matrice sous forme tridiagonale en utilisant la méthode de Householder, puis on calcule les valeurs propres en utilisant l'algorithme de Q6. Le calcul des vecteurs propres de la matrice tridiagonale obtenue se fait en utilisant la méthode de la puissance inverse (problème 21, Q7). Les systèmes linéaires à résoudre étant tridiagonaux).

Problème 23 : Méthode Q-R. Agrégation 1981, extrait

$C^{n,p}$ (Resp. $IR^{n,p}$) est l'espace vectoriel sur C (Resp. sur IR) des matrices à coefficients complexes (Resp. réels) à n lignes et p colonnes.

$C^{n,p}$ (Resp. $IR^{n,p}$) est muni de la topologie associée à une norme, l'élément de la ligne i et de la colonne j d'une matrice M est noté m_{ij}.

$e_p \in IR^{n,1}$ est la matrice colonne dont tous les éléments sont nuls, excepté celui de la ligne p valant 1.

Une matrice $M \in C^{n,n}$ (Resp. $M \in IR^{n,n}$) est dite tridiagonale si $m_{ij} = 0$ pour tout couple (i, j) tel que $|i - j| \geq 2$.

Une matrice $M \in C^{n,n}$ (Resp. $M \in IR^{n,n}$) est dite triangulaire supérieure si $m_{ij} = 0$ pour tout couple (i, j) tel que $i > j$.

I est la matrice unité d'ordre n.

M^* est la matrice adjointe de la matrice M.

***Q1.** Soit E un espace vectoriel normé de dimension finie, d la distance associée à la norme, $(x_s)_{s\in IN}$ une suite bornée de E.*

Démontrer que si :

(a) la suite $\left(d(x_s, x_{s+1})\right)_{s\in IN}$ a pour limite 0,

(b) la suite $(x_s)_{s\in IN}$ a un nombre fini de valeurs d'adhérence,

alors la suite $(x_s)_{s\in IN}$ a une limite.

***Q2.** Soient $A, B, C \in C^{n,n}$ trois matrices hermitiennes telles que $C = A + B$. On désigne par $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$, $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$, $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$ les valeurs propres respectives de A, B et C.*

Démontrer que :

$$\alpha_i + \beta_n \leq \gamma_i \leq \alpha_i + \beta_1,\ i = 1, 2, \ldots, n$$

On utilisera le fait que si M est une matrice de $C^{n,n}$ hermitienne et de valeurs propres $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, on a :

$$\begin{cases} \lambda_1 = \underset{x\in C^{n,1}-\{0\}}{Max} \dfrac{x^* Ms}{x^* x} \\ \lambda_s = \underset{(p_1,\ldots,p_{s-1})\in\left(C^{n,1}\right)^{s-1}}{Min} \ \underset{\substack{x\in C^{n,1}-\{0\} \\ p_i^* x=0 \\ 1\leq i\leq s-1}}{Max} \dfrac{x^* Ms}{x^* x},\ \text{pour } s = 2, 3, \cdots, n \end{cases}$$

Soit $\theta \in IR$, $(k,l) \in IN^2$ tels que $1 \leq k < l \leq n$ et $S^{k,l}(\theta) = \left(s_{i,j}^{k,l}(\theta)\right)$ la matrice carrée d'ordre n définie par :

$$s_{i,i}^{k,l}(\theta) = 1 \ pour\ i \neq k \text{ et } i \neq l$$

$$s_{k,k}^{k,l}(\theta) = s_{l,l}^{k,l}(\theta) = \mathrm{Cos}(\theta)$$

$$s_{k,l}^{k,l}(\theta) = -s_{l,k}^{k,l}(\theta) = -\mathrm{Sin}(\theta)$$

$$s_{i,j}^{k,l}(\theta) = 0 \ sinon$$

Q3. *Soit* $x=\sum_{i=1}^{n} x_i e_i$ *une matrice colonne réelle, soit p un entier compris entre 1 et n.*

Montrer qu'il existe une matrice Q_p, produit de matrices $S^{k,l}\left(\theta_{k,l}\right)$ et $\alpha_p \in IR$ *tels que :*

$$\begin{cases} Q_p x=\sum_{i=1}^{p-1} x_i e_i+\alpha_p e_p \text{ si } 2\le p\le n \\ Q_1(x)=\alpha_1 e_1 \text{ si } p=1 \end{cases}$$

Q4. *Soit* $A\in IR^{n,n}$ *symétrique. Montrer qu'il existe une matrice U, produit de matrices* $S^{k,l}\left(\theta_{k,l}\right)$, *telle que la matrice* $U^{*}AU$ *soit tridiagonale.*

Q5. *Soit H l'ensemble des matrices* $A\in IR^{n,n}$ *symétriques, tridiagonales et telles que :*

$$a_{i,i-1}\ne 0 \text{ pour } i=2,3,\cdots,n$$

(a) Montrer que pour tout $A\in H$, *il existe une matrice Q, produit de matrices* $S^{k,l}\left(\theta_{k,l}\right)$ *telle que* $Q^{*}A$ *soit une matrice triangulaire supérieure R.*

(b) Etudier le rang de A. Démontrer que $r_{k,k}\ne 0$ *pour k = 1, 2, ···, n – 1 et que $r_{n,n}$ est nul si et seulement si A est singulière.*

(c) Démontrer que :

$$\begin{cases} q_{k,k-1}\ne 0 \text{ pour } k=2,3,\cdots,n \\ q_{k,l}=0 \text{ pour } k\ge l+2 \end{cases}$$

(d) Soit $P\in IR^{n,n}$ *une matrice orthogonale telle que* $P^{*}A$ *soit une matrice triangulaire supérieure S. Démontrer que* $P^{*}Q$ *est une matrice orthogonale diagonale. Comparer les matrices* RQ *et* SP.

Q6. *Soit* $A\in H$, *montrer que l'on peut définir une suite de matrices* $\left(A_s\right)_{s\in IN}$ *par :*

$$\begin{cases} A_0=A \\ si\ A_s=\left(\left(a_{ij}^{(s)}\right)\right)_{1\le i,j\le n};\ A_s-a_{nn}^{(s)}I=Q_sR_s \end{cases}$$

avec Q_s matrice orthogonale et R_s matrice triangulaire supérieure, alors

$$A_{s+1}=R_sQ_s+a_{nn}^{(s)}I$$

Q7. *Démontrer que toutes les matrices A_s sont orthogonalement semblables, symétriques, tridiagonales et que* $a_{i,i-1}^{(s)}$ *est non nul pour* $s\in IN,\ 2\le i\le n-1$.

Q8. *Démontrer que pour tout* $s\in IN$ *:*

$$a_{n,n-1}^{(s+1)}=q_{n,n-1}^{(s)}q_{n-1,n}^{(s)}a_{n,n-1}^{(s)}$$

En déduire que la suite $\left(\left|a_{n,n-1}^{(s)}\right|\right)_{s\in IN}$ *a une limite L.*

Q9. *Dans toute cette question, on suppose L = 0.*

(a) Démontrer que la suite $\left(a_{n,n}^{(s+1)}-a_{n,n}^{(s)}\right)_{s\in IN}$ *a pour limite 0. En déduire que la suite* $\left(a_{n,n}^{(s)}\right)_{s\in IN}$ *a pour limite l.*

(b) On note :

$$A_s=\begin{pmatrix}F_s & G_s\\ K_s & H_s\end{pmatrix}\ avec\ \begin{cases}F_s\in IR^{n-1,n-1},\ K_s\in IR^{1,n-1}\\ G_s\in IR^{n-1,1},\ H_s\in IR^{1,1}\end{cases}$$
$$Q_s=\begin{pmatrix}S_s & T_s\\ U_s & V_s\end{pmatrix}\ avec\ \begin{cases}S_s\in IR^{n-1,n-1},\ U_s\in IR^{1,n-1}\\ T_s\in IR^{n-1,1},\ V_s\in IR^{1,1}\end{cases}$$

Démontrer qu'il existe s_0 tel que $s\geq s_0$, on ait :

(i) $\left|l-a_{nn}^{(s)}\right|\leq\left|a_{n,n-1}^{(s)}\right|$

(ii) $a_{nn}^{(s)}$ *non valeur propre de la matrice F_s.*

En utilisant les notations précédentes et en écrivant que $Q_s^*\left(A_s-a_{nn}^{(s)}I\right)$ est triangulaire supérieure, démontrer que les suites $\left(T_s\right)_{s\in IN}$ et $\left(U_s\right)_{s\in IN}$ sont convergentes.

En déduire qu'il existe une constante $k>0$ telle que pour tout $s\geq s_0$, on ait :

$$\left|a_{n,n-1}^{(s+1)}\right|\leq k\left|a_{n,n-1}^{(s)}\right|^3$$

Solution

Q1. Voir le problème 19, Q5.

Q2. Soit $\{a_1,a_2,\cdots,a_n\}$ une base orthonormée de vecteurs propres de A avec $Aa_k=\alpha_k a_k$ et V_k l'espace vectoriel engendré par $\{a_k,a_{k+1},\cdots,a_n\}$ pour tout $k=1$, $\cdots$, n.

Avec l'indication (théorème de Courant–Fischer, problème 7, Q1), on a :

$$\gamma_k\leq Max\left\{\frac{x^*Cx}{x^*x};\ x\in V_k-\{0\}\right\}$$

soit :

$$\gamma_k\leq Max\left\{\frac{x^*Ax}{x^*x};\ x\in V_k-\{0\}\right\}+Max\left\{\frac{x^*Bx}{x^*x};\ x\in V_k-\{0\}\right\}$$

Pour $x=\sum_{j=k}^{n}x_ja_j\in V_k$, on a $x^*Ax=\sum_{j=k}^{n}\alpha_j\left|x_j\right|^2\leq\alpha_k\sum_{j=k}^{n}\left|x_j\right|^2=\alpha_k x^*x$. Donc :

$$\gamma_k\leq\alpha_k+Max\left\{\frac{x^*Bx}{x^*x};\ x\in V_k-\{0\}\right\}\leq\alpha_k+\beta_1$$

En appliquant ce résultat à $A=C-B$, en tenant compte du fait que les valeurs propres de $-B$ sont $-\beta_n\geq\cdots\geq-\beta_1$, on déduit que $\alpha_k\leq\gamma_k-\beta_n$. On a donc bien :

$$\alpha_k+\beta_n\leq\gamma_k\leq\alpha_k+\beta_1\ \left(1\leq k\leq n\right)$$

Remarque — Ce résultat est le théorème de monotonicité de Weyl.

Pour tout $x=\sum_{i=1}^{n}x_ie_i\in IR^{n,1}$, on a :

$$S^{k,l}(\theta)x=\sum_{\substack{1\leq i\leq n\\ i\neq k,i\neq l}}x_ie_i+\left(\mathrm{Cos}(\theta)x_k-\mathrm{Sin}(\theta)x_l\right)e_k+\left(\mathrm{Sin}(\theta)x_k+\mathrm{Cos}(\theta)x_l\right)e_l$$

Q3. Pour $p = n$, on prend $Q_n = I = S^{1,2}(0)$ et $\alpha_n = x_n$.

On suppose donc que $p \in \{1,2,\cdots,n-1\}$ et on écrit $x \in IR^{n,1}$ sous la forme $x = x' + x''$, où $\begin{cases} x' = \sum_{i=1}^{p-1} x_i e_i,\ x'' = \sum_{i=p}^{n} x_i e_i,\ \text{si } 2 \le p \le n-1 \\ x' = 0,\ x'' = x,\ \text{si } p = 1 \end{cases}$.

Pour $p \le k < l \le n$, on a alors $S^{k,l}(\theta)x' = x'$, pour tout $\theta \in IR$.

En prenant $k = p$ et $l = p + 1$, on a :

$$S^{k,l}(\theta)x'' = \begin{cases} x_p^1 e_p + x_{p+1}^1 e_{p+1} + \sum_{i=p+2}^{n} x_i e_i,\ \text{si } 1 \le p \le n-2 \\ x_{n-1}^1 e_{n-1} + x_n^1 e_n,\ \text{si } p = n-1 \end{cases}$$

avec :

$$\begin{cases} x_p^1 = \mathrm{Cos}(\theta)x_p - \mathrm{Sin}(\theta)x_{p+1} \\ x_{p+1}^1 = \mathrm{Sin}(\theta)x_p + \mathrm{Cos}(\theta)x_{p+1} \end{cases}$$

On pose $r_p^1 = \sqrt{\left(x_p^1\right)^2 + \left(x_{p+1}^1\right)^2}$ et on définit $\theta = \theta_{p,p+1} \in]-\pi,\pi[$ par :

$$\begin{cases} \theta_{p,p+1} = 0,\ \text{si } x_{p+1}^1 = 0 \\ Cos\left(\theta_{p,p+1}\right) = \dfrac{x_p^1}{r_p^1},\ \mathrm{Sin}\left(\theta_{p,p+1}\right) = -\dfrac{x_{p+1}^1}{r_p^1},\ \text{si } x_{p+1}^1 \ne 0 \end{cases}$$

et on a :

$$\begin{cases} x_p^1 = r_p^1,\ \text{si } x_{p+1}^1 \ne 0,\ x_p^1 = x_p,\ \text{si } x_{p+1}^1 = 0 \\ x_{p+1}^1 = 0 \end{cases}$$

Soit, en définitive :

$$S^{p,p+1}\left(\theta_{p,p+1}\right)x = \begin{cases} x' + x_p^1 e_p + \sum_{i=p+2}^{n} x_i e_i,\ \text{si } 1 \le p \le n-2 \\ x' + x_p^1 e_p,\ \text{si } p = n-1 \end{cases}$$

En prenant ainsi de suite $(k,l) = (p,p+1),(p,p+2),\cdots,(p,n)$, on peut trouver des réels $\theta_{p,l} \in]-\pi,\pi[$, pour $p+1 \le l \le n$ tels que :

$$S^{p,l}\left(\theta_{p,l}\right)\cdots S^{p,p+1}\left(\theta_{p,p+1}\right)x = \begin{cases} x' + x_p^{l-p} e_p + \sum_{i=l+1}^{n} x_i e_i,\ \text{si } 1 \le l \le n-1 \\ x' + x_p^{l-p} e_p,\ \text{si } l = n \end{cases}$$

En posant $Q_p = S^{p,n}\left(\theta_{p,n}\right)\cdots S^{p,p+1}\left(\theta_{p,p+1}\right)$, on a $Q_p x = x' + \alpha_p e_p$.

Remarque — On a $Q_p e_j = e_j$, pour $j = 1, 2, \cdots, p-1$, si $2 \le p \le n$. La matrice Q_p étant orthogonale, on a également $Q_p^* e_j = e_j$, pour $j = 1, 2, \cdots, p-1$.

Q4. On a, pour tout $j = 1, 2, \cdots, n$, $Ae_j = \sum_{i=1}^{n} a_{ij} e_i$.

D'après ce qui précède, on peut trouver une matrice orthogonale $Q_2 = S^{2,n}(\theta_{2,n})\cdots S^{2,3}(\theta_{2,3})$, laissant e_1 invariant telle que $Q_2 Ae_1 = a_{11}e_1 + a_{21}^2 e_2$. Et avec $Q_2^* e_1 = e_1$, cela peut s'écrire $Q_2 AQ_2^* e_1 = a_{11}e_1 + a_{21}^2 e_2$.

La matrice $Q_2 AQ_2^*$ étant symétrique (A est symétrique), on en déduit que :

$$\begin{cases} Q_2 AQ_2^* e_2 = \sum_{i=1}^{n} a_{i2}^2 e_i \ \left(a_{12}^2 = a_{21}^2\right) \\ Q_2 AQ_2^* e_j = \sum_{i=2}^{n} a_{ij}^2 e_i \ \left(3 \le j \le n\right) \end{cases}$$

Supposons que, pour tout $k = 2, \cdots, p,$ on ait construit une matrice orthogonale $Q_k = S^{k,n}(\theta_{k,n})\cdots S^{k,k+1}(\theta_{k,k+1})$, laissant $e_1, e_2, \cdots, e_{k-1}$ invariants et telle que si $A_p = Q_p \cdots Q_2 AQ_2^* \cdots Q_p^*$, alors :

$$\begin{cases} A_p e_1 = a_{11} + a_{21}^2 e_2 \\ A_p e_j = a_{j-1,j}^j e_{j-1} + a_{jj}^j e_j + a_{j+1,j}^{j+1} e_{j+1} \quad (2 \le j \le p-1) \\ A_p e_p = a_{p-1,p}^p e_{p-1} + \sum_{i=p}^{n} a_{ip}^p e_i \\ A_p e_j = \sum_{i=p}^{n} a_{ij}^p e_i \ \left(p+1 \le j \le n\right) \end{cases}$$

On peut alors trouver une matrice orthogonale $Q_{p+1} = S^{p+1,n}\left(\theta_{p+1,n}\right)\cdots S^{p+1,p+2}\left(\theta_{p+1,p+2}\right)$, laissant $e_1, e_2, \cdots, e_p$ invariants et telle que $Q_{p+1}A_p e_p = a_{p-1,p}^p e_{p-1} + a_{pp}^p e_p + a_{p+1,p}^{p+1} e_{p+1}$. Avec $Q_{p+1}^* e_p = e_p$, cela peut s'écrire $Q_{p+1}A_p Q_{p+1}^* e_p = a_{p-1,p}^p e_{p-1} + a_{pp}^p e_p + a_{p+1,p}^{p+1} e_{p+1}$.

Q_{p+1} et Q_{p+1}^* laissant invariants $e_1, e_2, \cdots, e_{p-1}$, on a $Q_{p+1}A_p Q_{p+1}^* e_j = A_p e_j$, pour $j = 1, \cdots, p-1$.

Enfin, avec la symétrie de $Q_{p+1}A_p Q_{p+1}^*$, on en déduit que :

$$\begin{cases} Q_{p+1}A_p Q_{p+1}^* e_{p+1} = \sum_{i=p}^{n} a_{i,p+1}^{p+1} e_i \\ Q_{p+1}A_p Q_{p+1}^* e_j = \sum_{i=p+1}^{n} a_{ij}^{p+1} e_i \quad \left(p+2 \le j \le n\right) \end{cases}$$

On a donc ainsi montré par récurrence que la matrice $Q_{n-1}\cdots Q_2 AQ_2^* \cdots Q_{n-1}^*$ est tridiagonale.

Remarque — Cette technique de tridiagonalisation est due à Givens. Elle nécessite deux fois plus d'opérations que celle de Householder (problème 20). On peut accélérer ce procédé (voir Théodor et Lascaux, p. 357).

Q5. (a) En reprenant le raisonnement du début de Q3 avec $x = Ae_1$, on peut trouver $\theta_{12} \in \left]-\pi, \pi\right[$ tel que :

$$\begin{cases} S^{1,2}(\theta_{1,2})Ae_1 = a_{11}^1 e_1 \\ S^{1,2}(\theta_{1,2})Ae_j = \sum_{i=j-1}^{j} a_{ij}^1 e_i + a_{j+1,j}e_{j+1}, \text{ si } 2 \le j \le n-1 \\ S^{1,2}(\theta_{1,2})Ae_n = \sum_{i=n-1}^{n} a_{in}^1 e_i \end{cases}$$

avec $a_{11}^1 = \sqrt{{a_{11}}^2 + {a_{21}}^2} > 0$.

Supposons que, pour tout k = 1, 2, ···, p, on ait trouvé $\theta_{k,k+1} \in \left]-\pi, \pi\right[$ tel que si on pose $A_p = S^{p,p+1}(\theta_{p,p+1})\cdots S^{1,2}(\theta_{1,2})A$, alors :

$$\begin{cases} A_p e_j = \sum_{i=1}^{j} a_{ij}^p e_i, \text{ si } 1 \le j \le p \\ A_p e_j = \sum_{i=1}^{j} a_{ij}^p e_i + a_{j+1,j}e_{j+1}, \text{ si } p+1 \le j \le n-1 \\ A_p e_n = \sum_{i=1}^{n} a_{in}^p e_i \end{cases}$$

avec $a_{jj}^p > 0$.

Toujours avec le raisonnement du début de Q3, on peut trouver $\theta_{p+1,p+2} \in \left]-\pi, \pi\right[$ tel que $S^{p+1,p+2}(\theta_{p+1,p+2})A_p e_{p+1} = \sum_{i=1}^{p} a_{i,p+1}^p e_i + a_{p+1,p+1}^{p+1} e_{p+1}$, avec $a_{p+1,p+1}^{p+1} = \sqrt{(a_{pp}^p)^2 + (a_{p+1,p})^2} > 0$.

La matrice de rotation $S^{p+1,p+2}(\theta_{p+1,p+2})$ laissant invariants les e_j pour $j \ne p+1$, $j \ne p+2$, on déduit que $S^{p+1,p+2}(\theta_{p+1,p+2})A_p e_j = A_p e_j$, pour $j \ne p+1$, $j \ne p+2$.

Enfin pour $j = p + 2$, on a :

$$S^{p+1,p+2}(\theta_{p+1,p+2})A_p e_{p+2} = \sum_{i=1}^{p} a_{i,p+2}^p e_i + a_{p+1,p+2}^{p+1} e_{p+1} + a_{p+2,p+2}^{p+1} e_{p+2} + a_{p+3,p+2} e_{p+3}$$

On a donc ainsi montré qu'il existe $Q = S^{12}(-\theta_{12})\cdots S^{n-1,n}(-\theta_{n-1,n})$ tel que $Q^* A$ soit triangulaire supérieure.

Remarque — Les colonnes de Q forment un système orthonormé. Elles se déduisent de celles de A par le procédé d'orthogonalisation de Gram–Schmidt (voir Lascaux et Théodor, p. 342).

(b) On a $n-1 \le Rang(A) \le n$ (problème 9, Q2).

Avec $Dét(Q) = 1$, on a $Dét(A) = Dét(R) = \prod_{k=1}^{n} r_{kk}$ où $r_{kk} = a_{kk}^k > 0$ pour k = 1, 2, ···, n – 1. On déduit alors que $r_{nn} = 0$ si et seulement si $Dét(A) = 0$.

(c) On a, pour j = 1, 2, ···, n :

$$\begin{aligned}Qe_j &= S^{12}(-\theta_{12})\cdots S^{n-1,n}(-\theta_{n-1,n})e_j = S^{12}(-\theta_{12})\cdots S^{j,j+1}(-\theta_{j,j+1})e_j\\ &= S^{12}(-\theta_{12})\cdots S^{j-1,j}(-\theta_{j-1,j})\big(\mathrm{Cos}(\theta_{j,j+1})e_j - \mathrm{Sin}(\theta_{j,j+1})e_{j+1}\big)\end{aligned}$$

soit :

$$Qe_j = \mathrm{Cos}(\theta_{j,j+1})S^{12}(-\theta_{12})\cdots S^{j-1,j}(-\theta_{j-1,j})e_j - \mathrm{Sin}(\theta_{j,j+1})e_{j+1}$$

Ce qui s'écrit sous la forme :

$$Qe_j = \sum_{i=1}^{j} q_{ij}e_i - \mathrm{Sin}(\theta_{j,j+1})e_{j+1}$$

avec :

$$\mathrm{Sin}(\theta_{j,j+1}) = \frac{a_{j+1,j}}{r_{jj}} \neq 0 \ (1 \leq j \leq n-1)$$

On a donc bien :

$$\begin{cases} q_{k,k-1} = -\mathrm{Sin}(\theta_{k-1,k}) \neq 0 \text{ pour } k = 2,3,\cdots,n \\ q_{k,l} = 0 \text{ pour } k \geq l+2 \end{cases}$$

C'est-à-dire que Q est une matrice orthogonale de Hessenberg supérieure et irréductible (problème 9).

(d) Si $S = P^*A$ et $R = Q^*A$ sont triangulaires supérieures, avec P et Q orthogonales alors $D = P^*Q$ est une matrice orthogonale qui vérifie $DR = S$. Avec $0 = s_{ij} = \sum_{k=1}^{j} d_{ik}r_{kj}$ pour $1 \leq j < i \leq n$, on déduit que :

$$\begin{cases} r_{11}d_{i1} = 0 \ (i = 2,\cdots,n) \\ r_{1j}d_{i1} + r_{2j}d_{i2} + \cdots + r_{jj}d_{ij} = 0 \ (j = 2,\cdots,n-1;\ i = j+1,\cdots,n) \end{cases}$$

Et avec $r_{kk} \neq 0$ pou $k = 1, 2, \cdots, n-1$, on déduit que $d_{ij} = 0$ pour $1 \leq j < i \leq n$. C'est-à-dire que D est triangulaire supérieure et orthogonale. En écrivant que $D^*D = I$, on déduit alors qu'elle est diagonale. En effet on a $\sum_{k=1}^{i} d_{ki}d_{kj} = 0$ pour $1 \leq i < j \leq n$. Et avec $\prod_{i=1}^{n} d_{ii} = D\acute{e}t(D) = \pm 1$, on déduit que $d_{kj} = 0$ pour $1 \leq k < j \leq n$.

La matrice D est donc diagonale et orthogonale. C'est nécessairement une matrice diagonale avec $d_{ii} = \pm 1$ pour tout $i = 1, 2, \cdots, n$.

On a $PS = QR$, $D^* = D^{-1} = D$ et $DSPD = P^*QSPP^*Q = Q^*PSQ = Q^*AQ = RQ$. Soit en définitive $QR = D(SP)D$, c'est-à-dire que les coefficients de QR et SP sont identiques en valeur absolue avec :

$$(QR)_{ij} = d_{ii}d_{jj}(SP)_{ij} = \begin{cases} (SP)_{ij} \text{ si } d_{ii}d_{jj} = 1 \\ -(SP)_{ij} \text{ si } d_{ii}d_{jj} = -1 \end{cases}$$

Remarque — La matrice R est triangulaire supérieure et Q est orthogonale et de Hessenberg supérieure. On en déduit que $B = RQ$ est aussi de Hessenberg supérieure. En effet, on a $b_{ij} = \sum_{k=i}^{n} r_{ik} q_{kj} = 0$ pour $1 \le j \le i-2 \le n$ du fait que $q_{kj} = 0$ pour $k = j+2, \cdots, n$. D'autre part, B est semblable à $A = QR = Q(RQ)Q^*$ avec $B^* = B$. C'est-à-dire que B est de Hessenberg et symétrique, elle est donc tridiagonale. Enfin avec $b_{i,i-1} = r_{ii} q_{i,i-1} \neq 0$, pour tout $i = 2, \cdots, n$, on déduit que B vérifie les mêmes propriétés que A.

Q6. Si on pose $A_0 = A$, alors $A_0 - a_{nn}^{(0)} I$ vérifie les mêmes propriétés que A (tridiagonale, symétrique et irréductible) et on peut la décomposer sous la forme $A_0 - a_{nn}^{(0)} I = Q_0 R_0$, avec R_0 triangulaire supérieure, Q_0 orthogonale et de Hessenberg supérieure irréductible. La matrice $A_1 = R_0 Q_0 + a_{nn}^{(0)} I$ vérifiant les mêmes propriétés que A.

Supposons construite les matrices $A_0, \cdots, A_s$ vérifiant les mêmes propriétés que A. On peut alors écrire $A_s - a_{nn}^{(s)} I = Q_s R_s$ et $A_{s+1} = R_s Q_s + a_{nn}^{(s)} I$ vérifie les mêmes propriétés que A, on peut donc écrire la décomposition $A_{s+1} - a_{nn}^{(s+1)} I = Q_{s+1} R_{s+1}$.

La suite $(A_s)_{s \in IN}$ de matrices tridiagonales symétriques et irréductibles est donc bien définie.

Q7. Montrons le résultat par récurrence sur $s \ge 0$. Pour $s = 0$, c'est clair. Supposons le résultat vrai pour $s \ge 0$.

On a vu que $R_s Q_s$ est tridiagonale symétrique et irréductible (remarque de Q5) et il en est de même de $A_{s+1} = R_s Q_s + a_{nn}^{(s)} I$.

Enfin avec $A_{s+1} = Q_s^* (Q_s R_s) Q_s + a_{nn}^{(s)} I = Q_s^* \left(A_s - a_{nn}^{(s)} I \right) Q_s + a_{nn}^{(s)} I = Q_s^* A_s Q_s$, on déduit que A_{s+1} est orthogonalement semblable à A_s. Donc tous les A_s sont orthogonalement semblables à A.

Q8. Avec $A_{s+1} = R_s Q_s + a_{nn}^{(s)} I$, on déduit que $a_{n,n-1}^{(s+1)} = r_{nn}^{(s)} q_{n,n-1}^{(s)}$. Puis avec $R_s = Q_s^* \left(A_s - a_{nn}^{(s)} I \right)$, on déduit que $r_{nn}^{(s)} = q_{n-1,n}^{(s)} a_{n-1,n}^{(s)}$. Ce qui donne, en tenant compte de la symétrie de A_s :

$$a_{n,n-1}^{(s+1)} = q_{n,n-1}^{(s)} q_{n-1,n}^{(s)} a_{n,n-1}^{(s)}$$

La matrice Q_s est orthogonale, donc ses colonnes forment un système orthonormé et en particulier, on a $\left| q_{ij}^{(s)} \right| \le 1$ pour tous i, j compris entre 1 et n et tout entier s.

On a donc $\left| a_{n,n-1}^{(s+1)} \right| \le \left| a_{n,n-1}^{(s)} \right|$, c'est-à-dire que la suite $\left(\left| a_{n,n-1}^{(s)} \right| \right)_{s \in IN}$ est décroissante et minorée, elle est donc convergente.

Q9. (a) Avec $a_{n,n}^{(s+1)} = r_{n,n}^{(s)} q_{n,n}^{(s)} + a_{n,n}^{(s)} = q_{n-1,n}^{(s)} a_{n,n-1}^{(s)} q_{n,n}^{(s)} + a_{n,n}^{(s)}$, on déduit que $\left| a_{n,n}^{(s+1)} - a_{n,n}^{(s)} \right| \le \left| a_{n,n-1}^{(s)} \right| \underset{s \to +\infty}{\longrightarrow} 0$.

D'autre part, avec $A_s = Q^* A Q$ où Q est orthogonale, on déduit que $\|A_s\|_2 \le \|Q^*\|_2 \|A\|_2 \|Q\|_2 = \|A\|_2$ (si Q est orthogonale alors $\|Q\|_2 = 1$ – problème 1, Q7 –). Il en résulte que la suite $\left(a_{n,n}^{(s)} \right)_{s \in IN}$ est bornée.

Avec le résultat de Q1, il reste à montrer que cette suite n'admet qu'un nombre fini de valeurs d'adhérence.

On remarque que $Dét\left(A_s - a_{nn}^{(s)}I\right) = Dét(R_s) = \prod_{k=1}^{n} r_{kk}^{(s)}$. D'autre part, on peut écrire que $\|R_s\|_2 = \left\|Q_s^*\left(A_s - a_{nn}^{(s)}I\right)\right\|_2 \le \left\|A_s - a_{nn}^{(s)}I\right\|_2 \le \|A_s\|_2 + \left|a_{nn}^{(s)}\right| \le \|A\|_2 + \left|a_{nn}^{(s)}\right|$. On en déduit alors que la suite $\left(\|R_s\|_2\right)_{s\in IN}$ est bornée. Et il en est de même de chaque suite $\left(r_{kk}^{(s)}\right)_{s\in IN}$, pour $k = 1, 2, \cdots, n$.

Enfin avec $\left|r_{nn}^{(s)}\right| = \left|q_{n-1,n}^{(s)}a_{n-1,n}^{(s)}\right| \le \left|a_{n-1,n}^{(s)}\right|$, on déduit que $\underset{s\to+\infty}{Lim}\, r_{nn}^{(s)} = 0$ et $\underset{s\to+\infty}{Lim}\, Dét\left(A_s - a_{nn}^{(s)}I\right) = 0$. En considérant que A_s est semblable à A, on peut écrire que $\underset{s\to+\infty}{Lim}\, Dét\left(A - a_{nn}^{(s)}I\right) = 0$ et si $\left(a_{nn}^{(\varphi(s))}\right)_{s\in IN}$ est une suite extraite de $\left(a_{nn}^{(s)}\right)_{s\in IN}$ qui converge vers λ, alors $Dét(A - \lambda I) = 0$, c'est-à-dire que λ est une valeur propre de A. La suite $\left(a_{nn}^{(s)}\right)_{s\in IN}$ a donc au plus n valeurs d'adhérence.

On a donc ainsi montré que $\left(a_{nn}^{(s)}\right)_{s\in IN}$ converge vers une valeur propre l de A.

(b) On rappelle que A étant tridiagonale symétrique et irréductible, elle est diagonalisable avec toutes ses valeurs propres réelles et simples $\lambda_1 > \lambda_2 > \cdots > \lambda_n$ (problème 9, Q4).

On note $\delta = Min\left\{\left|\lambda_i - \lambda_j\right|;\ 1 \le i, j \le n,\ i \ne j\right\} > 0$.

Comme $\underset{s\to+\infty}{Lim}\, a_{nn}^{(s)} = l$, avec l valeur propre de A, on peut trouver $s_0 \in IN$ tel que :

$$\forall s \ge s_0,\ \left|a_{nn}^{(s)} - l\right| < \frac{\delta}{2}$$

Pour toute valeur propre de A, $\lambda_i \ne l$, on a :

$$\forall s \ge s_0,\ \left|\lambda_i - a_{nn}^{(s)}\right| \ge \left|\lambda_i - l\right| - \left|l - a_{nn}^{(s)}\right| > \delta - \frac{\delta}{2} = \frac{\delta}{2} > \left|l - a_{nn}^{(s)}\right|$$

C'est-à-dire que :

$$\forall s \ge s_0,\ \left|l - a_{nn}^{(s)}\right| = Min\left\{\left|\lambda_i - a_{nn}^{(s)}\right|;\ 1 \le i \le n\right\}$$

On écrit :

$$A_s = \begin{pmatrix} F_s & 0 \\ 0 & H_s \end{pmatrix} + \begin{pmatrix} 0 & G_s \\ K_s & 0 \end{pmatrix} = B_s + C_s, \text{ avec } C_s = \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & \cdots & 0 & 0 & a_{n-1,n}^{(s)} \\ 0 & \cdots & 0 & a_{n,n-1}^{(s)} & 0 \end{pmatrix}$$

La matrice A_s est semblable à A, elle admet donc les mêmes valeurs propres. Les valeurs propres de C_s sont $-\left|a_{n,n-1}^{(s)}\right| < 0 < \left|a_{n,n-1}^{(s)}\right| = \left|a_{n-1,n}^{(s)}\right|$. Enfin, on note $\mu_1(s) > \mu_2(s) > \cdots > \mu_n(s)$ celles de B_s.

Les inégalités de Q2 s'écrivent alors : $-\left|a_{n,n-1}^{(s)}\right| \le \lambda_i - \mu_i(s) \le \left|a_{n,n-1}^{(s)}\right|$. Soit :

$$\forall i \in \{1,2,\cdots,n\},\ \left|\lambda_i - \mu_i(s)\right| \le \left|a_{n,n-1}^{(s)}\right|$$

Comme $a_{nn}^{(s)}$ est valeur propre de B_s, on peut trouver un indice $i \in \{1,2,\cdots,n\}$ tel que $\left|\lambda_i - a_{nn}^{(s)}\right| \le \left|a_{n,n-1}^{(s)}\right|$. On déduit alors que :

$$\forall s \ge s_0,\ \left|l - a_{nn}^{(s)}\right| \le \left|a_{n,n-1}^{(s)}\right| \quad (i)$$

Si $a_{nn}^{(s)}$ est valeur propre de F_s, pour un indice $s \ge s_0$, alors $a_{nn}^{(s)}$ est valeur propre double de B_s et, les valeurs propres de A_s étant toutes simples, il existe deux indices $i \ne j$ dans $\{1,2,\cdots,n\}$ tels que :

$$\begin{cases} \left|\lambda_i - a_{nn}^{(s)}\right| \le \left|a_{n,n-1}^{(s)}\right| \\ \left|\lambda_j - a_{nn}^{(s)}\right| \le \left|a_{n,n-1}^{(s)}\right| \end{cases}$$

On a alors $\left|\lambda_i - \lambda_j\right| \le 2\left|a_{n,n-1}^{(s)}\right|$ pour un indice $s \ge s_0$.

Mais avec $\underset{s\to+\infty}{Lim}\, a_{n,n-1}^{(s)} = 0$, on peut choisir s_0 tel que l'on ait aussi $\left|a_{n,n-1}^{(s)}\right| < \dfrac{\delta}{2}$ pour tout $s \ge s_0$, de sorte que $\left|\lambda_i - \lambda_j\right| < \delta$ avec $i \ne j$, ce qui contredit la définition de δ.

On a donc montré que $a_{nn}^{(s)}$ n'est pas valeur propre de F_s pour tout $s \ge s_0$.

Avec $\left|\lambda_i - \mu_i(s)\right| \le \left|a_{n,n-1}^{(s)}\right|$, pour tout $i \in \{1,2,\cdots,n\}$ et en tenant compte de l'ordre imposé sur les valeurs propres de A_s (ou A) et B_s, on déduit que $\underset{s\to+\infty}{Lim}\, \mu_i(s) = \lambda_i$, pour tout $i \in \{1,2,\cdots,n\}$.

Les valeurs propres de A_s (ou A) sont toutes distinctes, $l = \lambda_{i_0}$ est une valeur propre de A avec $i_0 \in \{1,2,\cdots,n\}$ et $a_{nn}^{(s)}$ est valeur propre de B_s. On peut donc trouver une constante $M > 0$ et un entier $s_1 \ge s_0$ tels que :

$$\forall s \ge s_1,\ \left|\mu_i(s) - a_{nn}^{(s)}\right| > M \ (1 \le i \le n,\ i \ne i_0)$$

Ce qui est résumé par la figure 3.3.

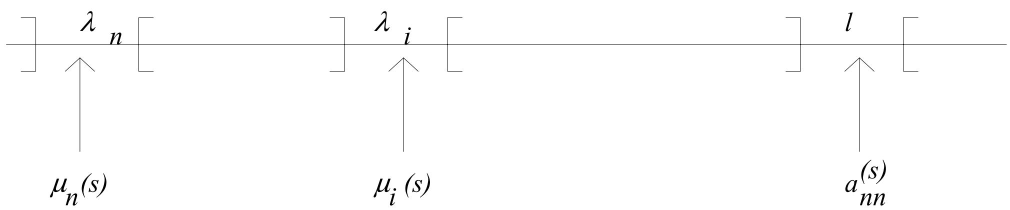

Figure 3.3

D'autre part, avec $Q_s^*\left(A_s - a_{nn}^{(s)}I\right) = R_s$ triangulaire supérieure, on déduit que $T_s^*\left(F_s - a_{nn}^{(s)}I\right) + V_s K_s = 0$ dans $IR^{1,n-1}$. Par transposition, en tenant compte du fait que F_s est symétrique et que $V_s \in IR$, on a $\left(F_s - a_{nn}^{(s)}I\right)T_s = -V_s K_s^*$.

Pour $s \geq s_1$, $F_s - a_{nn}^{(s)}I$ est inversible ($a_{nn}^{(s)}$ n'est pas valeur propre de F_s), on peut donc écrire $T_s = -V_s\left(F_s - a_{nn}^{(s)}I\right)^{-1} K_s^*$ et :

$$\|T_s\|_2 \leq |V_s| \left\|\left(F_s - a_{nn}^{(s)}I\right)^{-1}\right\| \|K_s^*\|_2 \leq \frac{1}{\underset{i\neq i_0}{Min}\left|\mu_i(s) - a_{nn}^{(s)}\right|}\left|a_{n,n-1}^{(s)}\right| = \frac{1}{M}\left|a_{n,n-1}^{(s)}\right|$$

(Q_s est orthogonale, donc $|V_s| \leq 1$).

On déduit donc que $\underset{s\to+\infty}{Lim}\, T_s = 0$.

Avec Q_s orthogonale, on déduit que $S_s^* T_s + V_s U_s^* = 0$, soit $V_s U_s^* = S_s^*\left(F_s - a_{nn}^{(s)}I\right)^{-1} V_s K_s^*$ et $|V_s|\|U_s^*\|_2 \leq \|S_s^*\|_2 \left\|\left(F_s - a_{nn}^{(s)}I\right)^{-1}\right\|_2 |V_s| \|K_s^*\|_2$.

Avec $\underset{s\to+\infty}{Lim}(V_s)^2 = \underset{s\to+\infty}{Lim}\left(1 - T_s^* T_s\right) = 1$, on déduit que $V_s \neq 0$ pour s assez grand et avec $\|S_s^*\|_2 \leq 1$ (Q_s est orthogonale), on conclut que $\|U_s^*\|_2 \leq \frac{1}{M}\left|a_{n,n-1}^{(s)}\right|$ et $\underset{s\to+\infty}{Lim}\, U_s = 0$.

En écrivant que $\left|q_{n,n-1}^{(s)}\right| \leq \|U_s^*\|_2$ et $\left|q_{n-1,n}^{(s)}\right| \leq \|T_s\|_2$, on déduit que $\left|q_{n,n-1}^{(s)}\right| \leq \frac{1}{M}\left|a_{n,n-1}^{(s)}\right|$ et $\left|q_{n-1,n}^{(s)}\right| \leq \frac{1}{M}\left|a_{n,n-1}^{(s)}\right|$.

Enfin avec $a_{n,n-1}^{(s+1)} = q_{n,n-1}^{(s)} q_{n-1,n}^{(s)} a_{n,n-1}^{(s)}$, on déduit qu'il existe une constante $k > 0$ telle que :

$$\forall s \geq s_1,\ \left|a_{n,n-1}^{(s+1)}\right| \leq k\left|a_{n,n-1}^{(s)}\right|^3$$

Problème 24 : Rayon spectral et majorations de normes matricielles. Agrégation 1980, extrait

Soit E un espace vectoriel de dimension finie sur C, muni de la norme notée $\| \ \|$.

Soit r un entier supérieur ou égal à 1 ; on désigne par E^r *l'espace vectoriel des matrices colonnes à r lignes* $U = \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix}$ *avec* $u_i \in E$, $1 \leq i \leq r$. *L'élément U de* E^r *sera noté* $U = (u_1, \cdots, u_r)^T$.

On munit l'espace E^r *de la norme* $\| \ \|_1$ *définie par :*

$$\forall U = (u_1, \cdots, u_r)^T, \quad \|U\|_1 = \sum_{i=1}^{r} \|u_i\|$$

On note $L(E)$ *[Resp.* $L(E^r)$*] l'espace vectoriel sur C des opérateurs linéaires de E [Resp. de* E^r*] dans lui même. On munit ces espaces des normes encore notées* $\| \ \|$ *et* $\| \ \|_1$ *respectivement, définies par :*

$$\forall g \in L(E), \quad \|g\| = \underset{x \in E, x \neq 0}{Max} \frac{\|g(x)\|}{\|x\|}$$

$$\forall B \in L(E^r), \quad \|B\|_1 = \underset{U \in E^r, U \neq 0}{Max} \frac{\|BU\|_1}{\|U\|_1}$$

Soit A une matrice carrée d'ordre r d'élément générique $a_{ij} \in C$, $1 \leq i, j \leq r$ *; on lui associe l'opérateur linéaire, noté encore A, appartenant à* $L(E^r)$*, défini par :*

si $U = (u_1, \cdots, u_r)^T$ *et si* $V = (v_1, \cdots, v_r)^T$ *appartiennent à* E^r *alors* $V = AU$ *équivaut à* $\left(\forall i = 1, \cdots, r, \quad v_i = \sum_{j=1}^{r} a_{ij} u_j \right)$.

Q1. *Montrer que, pour un tel opérateur A,*

$$\|A\|_1 = \underset{1 \leq j \leq r}{Max} \left(\sum_{i=1}^{r} |a_{ij}| \right)$$

Q2. *Etant donnés r nombres réels* $a_1, a_2, \cdots, a_r$ *on leur associe la matrice carrée d'ordre r notée R, d'élément générique* r_{ij} *:*

$$r_{ij} = \delta_{i,j+1} + a_j \delta_{1i}, \quad 1 \leq i \leq r, 1 \leq j \leq r$$

où :

$$\delta_{ij} = \begin{cases} 1 \ si \ i = j \\ 0 \ si \ i \neq j \end{cases}$$

(a) Montrer que si R admet une valeur propre de module 1 et de multiplicité supérieure ou égale à 2, la suite $(R^n)_{n \in IN}$ *des puissances de la matrice R n'est pas bornée ; (on pourra pour cela préciser la dimension de l'espace propre associé à cette valeur propre de module 1).*

(b) Montrer que, pour qu'il existe une matrice carrée H d'ordre r à coefficients dans C, inversible et d'inverse H^{-1}*, telle que* $\|H^{-1}RH\|_1 \leq 1$ *il faut et il suffit que les valeurs propres de R soient toutes de module inférieur ou égal à 1 et que les valeurs propres de module 1 soient simples.*

(c) On suppose maintenant que la matrice R vérifie la condition ;

$$(D) \quad \begin{cases} 1 \ est \ valeur \ propre \ simple \ de \ R \ et \ les \ autres \ valeurs \ propres \\ de \ R \ sont \ de \ module \ strictement \ inférieur \ ou \ égal \ à \ 1, \end{cases}$$

et on considère l'ensemble S des matrices carrées d'ordre r à coefficients complexes telles que :

$$\forall A \in S, \quad A\xi = \xi$$

où ξ *désigne la matrice colonne de* IR^r *dont tous les éléments sont égaux à 1.*

Montrer que si R vérifie la condition (D), il existe une matrice inversible H et un nombre réel positif ε tels que les conditions

$$A \in S \quad et \quad \|A - R\|_1 \leq \varepsilon \text{ entraînent } \left\|H^{-1}AH\right\|_1 \leq 1.$$

Solution

Q1. La démonstration est analogue à celle faite dans le problème 1, Q6 (b).

Q2. La matrice R associée à $(a_0, a_1, \cdots, a_r)$ est définie par :

$$R = \begin{pmatrix} a_1 & a_2 & & \cdots & a_r \\ 1 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

Si $\lambda \in C$ est une valeur propre de R, alors le sous espace propre associé est de dimension 1. En effet ce sous-espace vectoriel est défini par :

$$\begin{cases} \sum_{j=1}^{r} a_j x_j = \lambda \cdot x_1 \\ x_{j-1} = \lambda \cdot x_j \quad (2 \leq j \leq r) \end{cases}$$

C'est-à-dire que c'est le sous-espace vectoriel de C^r de dimension 1 engendré par le vecteur $e_\lambda = \left(\lambda^{r-1}, \lambda^{r-2}2, \cdots, \lambda, 1\right)^T \in C^r$.

On en déduit, en particulier, que la matrice R est diagonalisable si et seulement si toutes ses valeurs propres sont simples.

(a) Si R admet une valeur propre λ de multiplicité supérieure ou égale à 2, alors elle n'est pas diagonalisable et est semblable à une matrice réduite de Jordan de la forme :

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \varepsilon_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \varepsilon_3 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 0 & \lambda_{r-1} & \varepsilon_{r-1} \\ 0 & \cdots & & & 0 & \lambda_r \end{pmatrix} \quad (\varepsilon_i = 0 \text{ ou } 1)$$

C'est-à-dire qu'il existe une base $\{e_1, e_2, \cdots, e_r\}$ de $I\!R^r$ telle que :

$$\begin{cases} Re_1 = \lambda e_1 \\ Re_2 = e_1 + \lambda e_2 \end{cases}$$

Par récurrence, on déduit alors que $R^n e_2 = n\lambda^{n-1}e_1 + \lambda^n e_2$ $(n \geq 1)$ et :

$$\|R^n\| \geq \frac{\|R^n e_2\|}{\|e_2\|} \geq \frac{n|\lambda|^{n-1}\|e_1\| - |\lambda|^n \|e_2\|}{\|e_2\|}$$

où $\| \ \|$ désigne une norme quelconque sur IR^r ainsi que sa norme matricielle induite (sur un espace vectoriel de dimension finie elles sont toutes équivalentes).

Dans le cas où $|\lambda| = 1$, on a :

$$\|R^n\| \geq \frac{n\|e_1\| - \|e_2\|}{\|e_2\|} \underset{n\to+\infty}{\longrightarrow} +\infty$$

et la suite $(R^n)_{n\in IN}$ est non bornée.

(b) La condition $\|H^{-1}RH\|_1 \leq 1$ entraîne que toutes les valeurs propres de $H^{-1}RH$, et donc celles de R, sont de module inférieur ou égal à 1 (problème 1, Q1).

Les majorations :

$$\|R^n\|_1 = \left\|H\left(H^{-1}RH\right)^n H^{-1}\right\|_1 \leq \|H\|_1 \|H^{-1}RH\|_1^{\,n} \|H^{-1}\|_1 \leq \|H\|_1 \|H^{-1}\|_1$$

entraînent que la suite $(R^n)_{n\in IN}$ est bornée et donc que les valeurs propres de module 1 sont simples (d'après (a)).

Réciproquement, supposons que toutes les valeurs propres de R soient de module inférieur ou égal à 1 et que celles de module 1 sont simples. Dans ce cas R est semblable à une matrice réduite de Jordan de la forme :

$$J = H^{-1}RH = \begin{pmatrix} e^{i\theta_1} & 0 & . & . & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & \cdots & e^{i\theta_p} & 0 & . & . & 0 \\ 0 & \cdots & 0 & J_{p+1} & 0 & \cdots & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & 0 & J_q \end{pmatrix}$$

avec $\theta_j \in IR \quad (1 \leq j \leq p), \quad J_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & 0 & \lambda_k & 1 \\ 0 & & \cdots & 0 & \lambda_k \end{pmatrix}$ d'ordre $r_k \geq 1$ $(p+1 \leq k \leq q)$ et $|\lambda_k| < 1$.

Si tous les r_k sont égaux à 1, alors $\|H^{-1}RH\|_1 = \|J\|_1 \leq 1$. Sinon, pour les indices k tels que $r_k \geq 2$, on pose $P_\varepsilon = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \varepsilon & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & 0 & \cdots & 0 & \varepsilon^{r_k} \end{pmatrix}$ avec $|\lambda_k| + \varepsilon < 1$ et la matrice J_k est semblable à la matrice $J_{k,\varepsilon} = P_\varepsilon^{-1} J_k P_\varepsilon = \begin{pmatrix} \lambda_k & \varepsilon & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & 0 & \lambda_k & \varepsilon \\ 0 & & \cdots & 0 & \lambda_k \end{pmatrix}$ qui vérifie

$\|J_{k,\varepsilon}\|_1 < 1$. En notant $J_{k,\varepsilon} = J_k$ pour les indices k tels que $r_k = 1$, on voit que la matrice R est semblable à la matrice :

$$J_\varepsilon = H_\varepsilon^{-1} R H_\varepsilon = \begin{pmatrix} e^{i\theta_1} & 0 & . & . & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & \cdots & e^{i\theta_p} & 0 & . & . & 0 \\ 0 & \cdots & 0 & J_{p+1,\varepsilon} & 0 & \cdots & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & 0 & J_{q,\varepsilon} \end{pmatrix}$$

qui vérifie $\|J_\varepsilon\|_1 \leq 1$.

(c) Si $\lambda = 1$ est valeur propre de R, alors l'espace propre associé est de dimension 1 engendré par le vecteur $\xi = \begin{pmatrix} 1, & \cdots, & 1 \end{pmatrix}^T$. Si, de plus, 1 est valeur propre simple de R et les autres valeurs propres de R sont de module strictement inférieur ou égal à 1, alors le raisonnement du (b) permet de trouver une matrice inversible $H = \begin{pmatrix} 1 & h_{12} & \cdots & h_{1r} \\ . & . & . & . \\ . & . & . & . \\ 1 & h_{r2} & \cdots & h_{rr} \end{pmatrix}$ telle que $J = H^{-1} R H = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R_1 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} e_1 & R' \end{pmatrix}$, en notant $e_1 = \begin{pmatrix} 1, & 0, & \cdots, & 0 \end{pmatrix}^T$.

Pour $A \in S$, on a $A\xi = \xi$ et $H^{-1} A H e_1 = H^{-1} A \xi = H^{-1} \xi = e_1$, de sorte que :

$$H^{-1} A H = \begin{pmatrix} 1 & a'_{12} & \cdots & a'_{1r} \\ 0 & . & . & . \\ \vdots & . & . & . \\ 0 & a'_{r2} & \cdots & a'_{rr} \end{pmatrix} = \begin{pmatrix} e_1 & A' \end{pmatrix}$$

On a alors $\|H^{-1} A H\|_1 = Max\{1, \|A'\|_1\}$, avec :

$$\|A'\|_1 \leq \|A' - R'\|_1 + \|R'\|_1 = \|H^{-1}(A - R)H\|_1 + \|R'\|_1 \leq \|H^{-1}\|_1 \|H\|_1 \|A - R\|_1 + \|R'\|_1$$

Et pour $0 < \varepsilon < \dfrac{1 - \|R'\|_1}{\|H^{-1}\|_1 \|H\|_1}$, $\|A - R\|_1 \leq \varepsilon$, on a $\|A'\|_1 < 1$, ce qui entraîne $\|H^{-1} A H\|_1 \leq 1$.

CHAPITRE 4

Equations et systèmes d'équations non linéaires

Exercice 25 : Polynômes orthogonaux et matrices tridiagonales

On désigne par $IR[x]$ l'espace vectoriel des polynômes à coefficients réels et pour tout entier $n \geq 0$, $IR_n[x]$ est le sous-espace vectoriel de $IR[x]$ constitué des polynômes de degré n au plus.

On se donne un intervalle réel $]a,b[$, avec $-\infty \leq a < b \leq +\infty$ et une fonction poids intégrable $\pi :]a,b[\to IR_+^*$ telle que $0 < \int_a^b |x^k| \pi(x)dx < +\infty$ pour tout entier $k \geq 0$.

On définit un produit scalaire sur $IR[x]$ en posant :

$$\forall P, Q \in IR[x],\ \langle P | Q \rangle = \int_a^b P(x)Q(x)\pi(x)dx$$

On désigne par $\{P_k;\ k \in IN\}$ une famille orthonormée de polynômes pour ce produit scalaire avec $Degré(P_k) = k$ pour tout entier $k \geq 0$ (une telle famille peut être obtenue par le procédé d'orthogonalisation de Gram–Schmidt).

Enfin, on note pour tout entier $k \geq 0$:

$$P_k(x) = \sum_{j=0}^{k} \alpha_j^{(k)} x^j$$

avec $\alpha_k^{(k)} \neq 0$.

Q1. *Montrer que la suite $(P_k)_{k \in IN}$ vérifie la relation de récurrence :*

$$\begin{cases} P_{-1}(x) = 0,\ P_0(x) = \alpha_0^{(0)} = \dfrac{1}{\sqrt{\int_a^b \pi(x)dx}} \\ b_{k+1}P_{k+1}(x) + a_k P_k(x) + b_k P_{k-1}(x) = xP_k(x)\ (k \geq 0) \end{cases}$$

avec :

$$\begin{cases} a_0 = -\dfrac{\alpha_0^{(1)}}{\alpha_1^{(1)}},\ b_0 = 0 \\ a_k = \dfrac{\alpha_{k-1}^{(k)}}{\alpha_k^{(k)}} - \dfrac{\alpha_k^{(k+1)}}{\alpha_{k+1}^{(k+1)}}\ (k \geq 1) \\ b_k = \dfrac{\alpha_{k-1}^{(k-1)}}{\alpha_k^{(k)}}\ (k \geq 1) \end{cases}$$

Q2. *Montrer que, pour tout entier $k \geq 1$, le polynôme P_k admet k racines réelles simples dans l'intervalle $]a,b[$.*

Q3. *On désigne, pour tout entier $n \geq 1$, par T_n la matrice réelle, d'ordre n, tridiagonale et symétrique définie par :*

$$T_n = \begin{pmatrix} a_0 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_1 & b_2 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & b_{n-2} & a_{n-2} & b_{n-1} \\ 0 & \cdots & 0 & 0 & b_{n-1} & a_{n-1} \end{pmatrix}$$

et pour tout réel x, $u_n(x)$ désigne le vecteur de $\mathbb{R}^n$ défini par :

$$u_n(x) = \begin{pmatrix} P_0(x) \\ P_1(x) \\ \vdots \\ P_{n-1}(x) \end{pmatrix}$$

(a) Calculer $T_n u_n(x)$, pour tout entier $n \geq 1$ et tout réel x.

(b) Montrer que les racines de P_n sont les valeurs propres de T_n.

Q4. *On rappelle que la fonction gamma d'Euler est définie par :*

$$\forall x > 0,\ \Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} dt$$

et qu'elle vérifie la relation fonctionnelle $\Gamma(x+1) = x\Gamma(x)$. En particulier, on a $\Gamma(n+1) = n!$ et $\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{2^{2n} n!}\sqrt{\pi}$, pour tout entier $n \geq 0$.

La fonction béta est définie par :

$$\forall p,q > 0,\ B(p,q) = \int_0^1 t^{p-1}(1-t)^{p-1} dt$$

et on a $B(p,q) = \dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

Enfin, on pose, pour tous réels x et y tels que $-1 < y < x$:

$$\mathrm{C}_x^y = \frac{\Gamma(x+1)}{\Gamma(y+1)\Gamma(x-y+1)}$$

Montrer que pour tous réels $\alpha > -1$, $\beta > -1$ et tout entier $n \geq 1$, on a :

$$(i)\ \sum_{k=0}^{n} C_{n+\alpha}^{k} C_{n+\beta}^{n-k} = C_{2n+\alpha+\beta}^{n}$$

$$(ii)\ \sum_{k=1}^{n} k C_{n+\alpha}^{k} C_{n+\beta}^{n-k} = (n+\alpha) C_{2n+\alpha+\beta-1}^{n-1}$$

Q5. *Pour cette question, on prend* $]a,b[\,=\,]-1,1[$ *et* $\pi(x) = (1-x)^{\alpha}(1+x)^{\beta}$ *où* $\alpha > -1$ *et* $\beta > -1$ *sont des réels donnés.*

On définit alors les polynômes de Jacobi par :

$$\forall n \geq 0,\ \forall x \in\,]-1,1[,\ Q_n(x) = \frac{(-1)^n}{2^n n!} \frac{1}{(1-x)^{\alpha}(1+x)^{\beta}} \left(\frac{d}{dx}\right)^n \left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\}$$

(a) Montrer que Q_n *est un polynôme de degré n et que ses coefficients de degrés n et n – 1 respectivement sont donnés par :*

$$\begin{cases} \beta_0^{(0)} = 1,\ \beta_n^{(n)} = \dfrac{1}{2^n} C_{2n+\alpha+\beta}^{n},\ (n \geq 1) \\ \beta_{n-1}^{(n)} = \dfrac{n(\alpha-\beta)}{2n+\alpha+\beta} \beta_n^{(n)},\ (n \geq 1) \end{cases}$$

(b) Calculer $\left\langle Q_n \middle| x^k \right\rangle$ *pour* $0 \leq k < n$. *Que peut-on conclure du résultat ?*

(c) Calculer $\|Q_n\|^2 = \langle Q_n | Q_n \rangle$ *pour tout entier n.*

On définit les polynômes de Jacobi normalisés par :

$$\forall n \in IN,\ P_n = \frac{1}{\|Q_n\|} Q_n$$

(d) Ecrire la récurrence vérifiée par les P_n.

Q6. *On reprend les notations de Q5 avec* $\alpha = \beta$. *Les polynômes orthogonaux obtenus sont appelés les « polynômes ultrasphériques ».*

Ecrire la récurrence vérifiée par les P_n *ainsi que la matrice* T_n *associée.*

Q7. *On reprend les notations de Q6 avec* $\alpha = \beta = -\frac{1}{2}$. *Les polynômes orthogonaux obtenus sont appelés les « polynômes de Tchébychev de première espèce ».*

(a) Ecrire la récurrence vérifiée par les P_n *ainsi que la matrice* T_n *associée.*

(b) Calculer les valeurs propres et vecteurs propres de T_n.

(c) Calculer $P_n(\mathrm{Cos}(t))$ *pour tout réel* $t \in [0,\pi]$ *et en déduire une autre expression des polynômes de Tchébychev.*

Retrouver de manière plus simple les valeurs propres et vecteurs propres de T_n.

Q8. *On reprend les notations de Q6 avec* $\alpha = \beta = \frac{1}{2}$. *Les polynômes orthogonaux obtenus sont appelés les « polynômes de Tchébychev de deuxième espèce ».*

Reprendre les questions de Q7.

Solution

Q1. Avec $\|P_0\| = 1$, on déduit que $P_0(x) = \alpha_0^{(0)} = \dfrac{1}{\sqrt{\int_a^b \pi(x)dx}}$.

On a $xP_k \in IR_{k+1}[x]$ et $\{P_k;\ k = 0,1,\cdots,k+1\}$ est une base de $IR_{k+1}[x]$ puisque $Degré(P_j) = j$ pour tout entier *j*. On peut donc écrire $xP_k = \sum_{j=0}^{k+1} \alpha_j P_j$ avec $\alpha_j = \langle xP_k \mid P_j \rangle = \langle P_k \mid xP_j \rangle = 0$ pour $j+1<k$ puisque xP_j est combinaison linéaire des P_0, P_1, ..., P_{j+1} qui sont orthogonaux à P_k pour $j+1<k$. Il reste donc :

$$xP_k = \alpha_{k+1}P_{k+1} + \alpha_k P_k + \alpha_{k-1}P_{k-1}$$

pour $k \geq 1$.

En identifiant les coefficients de x^{k+1} et x^k respectivement, il vient :

$$\begin{cases} \alpha_{k+1} = \dfrac{\alpha_k^{(k)}}{\alpha_{k+1}^{(k+1)}} = b_{k+1} \neq 0 \\ \alpha_k = \dfrac{\alpha_{k-1}^{(k)}}{\alpha_k^{(k)}} - \dfrac{\alpha_k^{(k+1)}}{\alpha_{k+1}^{(k+1)}} = a_k \end{cases}$$

En écrivant que $\alpha_{k-1} = \langle P_k \mid xP_{k-1} \rangle$ avec $xP_{k-1} = \dfrac{\alpha_{k-1}^{(k-1)}}{\alpha_k^{(k)}} P_k + \sum_{j=0}^{k-1} \gamma_j P_j$, on déduit que $\alpha_{k-1} = \dfrac{\alpha_{k-1}^{(k-1)}}{\alpha_k^{(k)}} = b_k \neq 0$.

Pour $k = 0$, on a :

$$xP_0(x) = \alpha_1 P_1(x) + \alpha_0 P_0(x) = \alpha_1\left(\alpha_0^{(1)} + \alpha_1^{(1)}x\right) + \alpha_0\alpha_0^{(0)}$$

et $\alpha_1 = \dfrac{\alpha_0^{(0)}}{\alpha_1^{(1)}} = b_1 \neq 0$, $\alpha_0 = -\dfrac{\alpha_0^{(1)}}{\alpha_1^{(1)}} = a_0$.

Q2. Soit $k \geq 1$. Si P_k garde un signe constant sur l'intervalle $]a,b[$, alors $\langle P_k \mid P_0 \rangle = \int_a^b P_k(x)P_0(x)\pi(x)dx \neq 0$, ce qui contredit l'orthogonalité de P_k et P_0 pour $k \geq 1$.

Il existe donc une racine x_1 de P_k dans $]a,b[$.

Si x_1 est de multiplicité $p \geq 2$, on peut alors écrire $P_k(x) = (x - x_1)^2 Q_{k-2}(x)$ avec $Q_{k-2} \in IR_{k-2}[x]$.

Mais on a alors $0 = \langle P_k \mid Q_{k-2} \rangle = \int_a^b (x-x_1)^2 Q_{k-2}(x)^2 \pi(x)dx > 0$, ce qui est impossible.

On a donc montré que toutes les racines de P_k qui sont dans $]a,b[$ sont simples. Notons $x_1, x_2, \cdots, x_p$ ces racines.

Si $p < n$, on peut alors écrire $P_k(x) = \prod_{i=1}^{p}(x - x_i)Q_{k-p}(x)$, avec $Q_{k-p} \in IR_{k-p}[x]$ de signe constant dans $]a,b[$.

On a alors $0 = \left\langle P_k \middle| \prod_{i=1}^{p}(x - x_i) \right\rangle = \int_a^b \prod_{i=1}^{p}(x - x_i)^2 Q_{k-p}(x)\pi(x)dx \neq 0$, ce qui est impossible.

On a donc nécessairement $p = n$, c'est-à-dire que toutes les racines de P_k sont dans $]a,b[$ et sont simples.

Q3. (a) Pour n et x fixés, on note $v = T_n u_n(x)$. On a alors :

$$\begin{cases} v_0 = a_0 P_0(x) + b_1 P_1(x) \\ v_k = b_k P_{k-1}(x) + a_k P_k(x) + b_{k+1} P_{k+1}(x) \ (1 \leq k \leq n-2) \\ v_{n-1} = b_{n-1} P_{n-2}(x) + a_{n-1} P_{n-1}(x) \end{cases}$$

Soit en utilisant la récurrence vérifiée par les P_k :

$$\begin{cases} v_k = xP_k(x) \quad (0 \leq k \leq n-2) \\ v_{n-1} = xP_{n-1}(x) - b_n P_n(x) \end{cases}$$

C'est-à-dire que :

$$T_n u_n(x) = xu_n(x) - b_n P_n(x) e_n$$

où on a noté $\{e_1, e_2, \cdots, e_n\}$ la base canonique de IR^n.

(b) Si $P_n(x) = 0$, alors $T_n u_n(x) = xu_n(x)$, c'est-à-dire que x est valeur propre de P_n et $u_n(x)$ est un vecteur propre non nul associé ($P_0(x) \neq 0$).

Le polynôme P_n admettant n racines simples, on a ainsi toutes les valeurs propres de T_n.

Remarque — On retrouve le fait que les valeurs propres de la matrice tridiagonale symétrique et irréductible T_n sont simples et réelles (problème 9, Q5).

Q4. Pour $|x| < 1$, on a :

$$(1+x)^{n+\alpha} = 1 + \sum_{k \geq 1} \frac{(n+\alpha)(n+\alpha-1)\cdots(n+\alpha-k+1)}{k!} x^k$$
$$= \sum_{k=0}^{n} \frac{\Gamma(n+\alpha+1)}{\Gamma(k+1)\Gamma(n+\alpha+1-k)} x^k + R_n(x)$$

On a donc :

$$\begin{cases} (1+x)^{n+\alpha} = \sum_{k=0}^{n} C_{n+\alpha}^k x^k + R_n(x) \\ (1+x)^{n+\beta} = \sum_{k=0}^{n} C_{n+\beta}^k x^k + S_n(x) \end{cases}$$

et en faisant le produit des deux séries entières :

$$(1+x)^{2n+\alpha+\beta} = \sum_{p=0}^{n} \left(\sum_{k=0}^{p} C_{n+\alpha}^k C_{n+\beta}^{p-k} \right) x^p + \Sigma_n(x)$$

ce qui donne la formule (i) en identifiant les coefficients de x^n.

Par dérivation, on a $(n+\alpha)(1+x)^{n+\alpha-1} = \sum_{k=1}^{n} C_{n+\alpha}^{k} k x^{k-1} + R_n'(x)$ et :

$$(n+\alpha)x(1+x)^{n+\alpha-1} = \sum_{k=0}^{n} kC_{n+\alpha}^{k} x^{k} + xR_n'(x)$$

On en déduit alors que :

$$(n+\alpha)x(1+x)^{2n+\alpha+\beta-1} = \sum_{p=0}^{n}\left(\sum_{k=0}^{p} kC_{n+\alpha}^{k} C_{n+\beta}^{p-k}\right)x^{p} + \Xi_n(x)$$

ce qui donne la formule (ii) en identifiant les coefficients de x^n.

Q5. (a) On a $Q_0 = 1$, donc $\beta_0^{(0)} = 1$.

Pour $n \geq 1$, on pose $R_n(x) = \left(\frac{d}{dx}\right)^n \left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\}$.

Avec la formule de Leibnitz, on a :

$$R_n(x) = \sum_{k=0}^{n} (-1)^k C_n^k \frac{\Gamma(n+\alpha+1)}{\Gamma(n+\alpha-k+1)} \frac{\Gamma(n+\beta+1)}{\Gamma(\beta+k+1)} (1-x)^{n+\alpha-k}(1+x)^{\beta+k}$$

ce qui donne :

$$R_n(x) = n!(1-x)^{\alpha}(1+x)^{\beta} \sum_{k=0}^{n} (-1)^k C_{n+\alpha}^{k} C_{n+\beta}^{n-k} (1-x)^{n-k}(1+x)^{k}$$

On en déduit donc que Q_n est un polynôme de degré n avec pour coefficient de x^n :

$$\beta_n^{(n)} = \frac{1}{2^n}\sum_{k=0}^{n} C_{n+\alpha}^{k} C_{n+\beta}^{n-k} = \frac{1}{2^n} C_{2n+\alpha+\beta}^{n} \neq 0$$

En remarquant que :

$$(1-x)^{n-k}(1+x)^{k} = (-1)^{n-k}(x-1)^{n-k}(1+x)^{k} = (-1)^{n-k}\left\{x^n + (2k-n)x^{n-1} + \sum_{j=0}^{n-2}\gamma_j x^j\right\}$$

on déduit que le coefficient de x^{n-1} dans Q_n est :

$$\beta_{n-1}^{(n)} = \frac{1}{2^n}\sum_{k=0}^{n}(2k-n) C_{n+\alpha}^{k} C_{n+\beta}^{n-k} = \frac{1}{2^n}\left(2\sum_{k=0}^{n} kC_{n+\alpha}^{k} C_{n+\beta}^{n-k} - n\sum_{k=0}^{n} C_{n+\alpha}^{k} C_{n+\beta}^{n-k}\right)$$

soit :

$$\beta_{n-1}^{(n)} = \frac{1}{2^n}\left(2(n+\alpha)C_{2n+\alpha+\beta-1}^{n-1} - nC_{2n+\alpha+\beta}^{n}\right)$$

Et avec $C_{2n+\alpha+\beta}^{n} = \frac{2n+\alpha+\beta}{n} C_{2n+\alpha+\beta-1}^{n-1}$, on déduit que $\beta_{n-1}^{(n)} = \frac{n(\alpha-\beta)}{2n+\alpha+\beta}\beta_n^{(n)}$.

(b) On rappelle la formule d'intégration par parties :

$$\int_a^b f^{(n)}(x)g(x)dx = \left[\sum_{j=1}^{p}(-1)^{j+1} f^{(n-j)}(x) g^{(j-1)}(x)\right]_a^b + (-1)^p \int_a^b f^{(n-p)}(x) g^{(p)}(x)dx$$

valable pour $1 \leq p \leq n$.

On remarque, d'autre part, que $\left(\frac{d}{dx}\right)^{n-k}\left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\}$ s'annule en $x=-1$ et $x=1$ pour $1 \leq k \leq n$.

Pour $0 \leq k < n$, on a :

$$\left\langle \frac{2^n n!}{(-1)^n} Q_n \middle| x^k \right\rangle = \int_{-1}^{1} x^k \left(\frac{d}{dx}\right)^n \left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\} dx$$

Et après k intégrations par parties, il reste :

$$\left\langle \frac{2^n n!}{(-1)^n} Q_n \middle| x^k \right\rangle = (-1)^k \int_{-1}^{1} k! \left(\frac{d}{dx}\right)^{n-k} \left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\} dx$$

$$= (-1)^k k! \left[\left(\frac{d}{dx}\right)^{n-k-1} \left\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\right\}\right]_{-1}^{1} = 0$$

On en déduit donc que les polynômes de Jacobi forment un système orthogonal.

(c) Pour $n=0$, on a $\|Q_0\|^2 = \int_{-1}^{+1}(1-x)^{\alpha}(1+x)^{\beta}dx$. En faisant le changement de variable $t = \frac{1}{2}(1+x)$, on a :

$$\|Q_0\|^2 = 2^{\alpha+\beta+1} B(\alpha+1,\beta+1) = 2^{\alpha+\beta+1}\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}$$

Pour $n \geq 1$, on écrit $Q_n(x) = \beta_n^{(n)} x^n + \sum_{j=0}^{n-1} \gamma_j Q_j(x)$ et on a $\|Q_n\|^2 = \langle Q_n | Q_n \rangle = \beta_n^{(n)} \langle x^n | Q_n \rangle$. Puis avec n intégrations par parties :

$$\|Q_n\|^2 = \frac{\beta_n^{(n)}}{2^n} \int_{-1}^{+1} (1-x)^{n+\alpha}(1+x)^{n+\beta} dx$$

Ce qui s'écrit aussi :

$$\|Q_n\|^2 = \frac{2^{\alpha+\beta+1}}{n!(2n+\alpha+\beta+1)} \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+\alpha+\beta+1)} \quad (n \geq 1)$$

(d) Avec les notations de Q1, les coefficients dominants des P_n sont donnés par :

$$\begin{cases} \alpha_0^{(0)} = \dfrac{\beta_0^{(0)}}{\|Q_0\|} = \dfrac{1}{2^{\frac{\alpha+\beta+1}{2}}} \sqrt{\dfrac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)}} \\ \alpha_n^{(n)} = \dfrac{\beta_n^{(n)}}{\|Q_n\|} = \dfrac{1}{2^{n+\frac{\alpha+\beta+1}{2}}} \sqrt{\dfrac{2n+\alpha+\beta+1}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)\Gamma(n+\alpha+\beta+1)n!}}\,\Gamma(2n+\alpha+\beta+1) \\ \alpha_{n-1}^{(n)} = \dfrac{\beta_{n-1}^{(n)}}{\|Q_n\|} = \dfrac{n(\alpha-\beta)}{2n+\alpha+\beta} \alpha_n^{(n)} \end{cases}$$

On en déduit alors que la suite $(P_k)_{k \in IN}$ des polynômes de Jacobi vérifie la relation de récurrence :

$$\begin{cases} P_{-1}(x) = 0,\ P_0(x) = \alpha_0^{(0)} = \dfrac{1}{2^{\frac{\alpha+\beta+1}{2}}} \sqrt{\dfrac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)}} \\ b_{k+1}P_{k+1}(x) + a_k P_k(x) + b_k P_{k-1}(x) = xP_k(x) \ (k \geq 0) \end{cases}$$

avec :

$$\begin{cases} a_0 = -\dfrac{\alpha_0^{(1)}}{\alpha_1^{(1)}} = \dfrac{\beta - \alpha}{2+\alpha+\beta},\ b_0 = 0 \\ a_k = \dfrac{\alpha_{k-1}^{(k)}}{\alpha_k^{(k)}} - \dfrac{\alpha_k^{(k+1)}}{\alpha_{k+1}^{(k+1)}} = \dfrac{\beta^2 - \alpha^2}{(2k+\alpha+\beta)(2k+\alpha+\beta+2)}\ (k \geq 1) \\ b_1 = \dfrac{\alpha_0^{(0)}}{\alpha_1^{(1)}} = \dfrac{2}{2+\alpha+\beta}\sqrt{\dfrac{(1+\alpha)(1+\beta)}{3+\alpha+\beta}} \\ b_k = \dfrac{\alpha_{k-1}^{(k-1)}}{\alpha_k^{(k)}} = \dfrac{2}{2k+\alpha+\beta}\sqrt{\dfrac{k(k+\alpha)(k+\beta)(k+\alpha+\beta)}{(2k+\alpha+\beta-1)(2k+\alpha+\beta+1)}}\ (k \geq 2) \end{cases}$$

Remarque — La formule donnant les b_k pour $k \geq 2$ n'est pas valable pour $k = 1$ si $\alpha+\beta+1=0$.

Q6. Dans le cas particulier où $\alpha = \beta$, tous les coefficients a_k sont nuls et les coefficients b_k sont donnés par :

$$\begin{cases} b_1 = \dfrac{1}{\sqrt{3+2\alpha}} \\ b_k = \sqrt{\dfrac{k(k+2\alpha)}{4(k+\alpha)^2 - 1}}\ (k \geq 2) \end{cases}$$

La matrice T_n s'en déduit immédiatement.

Q7. (a) Dans le cas particulier où $\alpha = \beta = -\dfrac{1}{2}$, on a :

$$\begin{cases} a_k = 0\ (k \geq 0) \\ b_1 = \dfrac{1}{\sqrt{2}} \\ b_k = \dfrac{1}{2}\ (k \geq 2) \end{cases}$$

Ce qui donne la récurrence :

$$\begin{cases} P_0(x) = \dfrac{1}{\|Q_0\|} = \dfrac{1}{\sqrt{\pi}},\ P_1(x) = \sqrt{\dfrac{2}{\pi}}x \\ \dfrac{1}{2}P_2(x) + \dfrac{1}{\sqrt{2}}P_0(x) = xP_1(x) \\ \dfrac{1}{2}P_{k+1}(x) + \dfrac{1}{2}P_{k-1}(x) = xP_k(x)\ (k \geq 2) \end{cases}$$

La matrice T_n est alors donnée par :

$$T_n = \frac{1}{2}\begin{pmatrix} 0 & \sqrt{2} & 0 & 0 & \cdots & 0 \\ \sqrt{2} & 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix} = \frac{1}{2}A_n$$

(b) On sait que les valeurs propres de T_n sont comprises entre –1 et 1, donc celles de A_n sont comprises entre –2 et 2. De plus A_n admet les mêmes valeurs propres que la matrice $B_n = \begin{pmatrix} 0 & 2 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$ (problème 9, Q5 (b)).

Les valeurs propres de B_n sont de la forme $\lambda = 2\,\mathrm{Cos}(\alpha)$ avec $\alpha \in]0,\pi[$. Si x est un vecteur propre non nul associé ses composantes sont solutions de la récurrence :

$$\begin{cases} x_1 = 1 \\ x_2 = \mathrm{Cos}(\alpha)x_1 \\ x_{k-1} - 2\,\mathrm{Cos}(\alpha)x_k + x_{k+1} = 0 \ (2 \le k \le n) \\ x_{n+1} = 0 \end{cases}$$

Le polynôme caractéristique de cette récurrence est $P(r) = r^2 - 2Cos(\alpha)r + 1$, soit $P(r) = (r - e^{i\alpha})(r - e^{-i\alpha})$. Les racines sont donc $r_1 = e^{i\alpha}$ et $r_2 = e^{-i\alpha}$. Ce qui donne $x_k = c_1 e^{ik\alpha} + c_2 e^{-ik\alpha}$ $(1 \le k \le n+1)$.

Avec $x_1 = 1$ et $x_2 = \mathrm{Cos}(\alpha)$, on déduit que $c_1 = \frac{e^{-i\theta}}{2}$ et $c_2 = \frac{e^{i\theta}}{2}$. C'est-à-dire que $x_k = \mathrm{Cos}((k-1)\alpha)$ et avec $x_{n+1} = 0$ on déduit que $\alpha = \frac{(2j+1)}{2n}\pi$ avec $0 \le j \le n-1$.

Les valeurs propres de T_n sont donc :

$$\lambda_j = \mathrm{Cos}\left(\frac{2j+1}{2n}\pi\right) \ (0 \le j \le n-1)$$

L'espace propre associé à λ_j est la droite engendrée par le vecteur $v^{(j)}$ de composantes :

$$v_k^{(j)} = \mathrm{Cos}\left((k-1)\frac{2j+1}{2n}\pi\right) \ (1 \le k \le n)$$

(c) On vérifie par récurrence que $P_n(\mathrm{Cos}(t)) = \sqrt{\frac{2}{\pi}}\mathrm{Cos}(nt)$ pour $n \ge 1$.

On en déduit alors que $P_n(x) = \sqrt{\frac{2}{\pi}}\mathrm{Cos}(n\,\mathrm{ArcCos}(x))$ pour $|x| \le 1$.

Les résultats du (b) se retrouvent alors facilement.

Q8. (a) Pour $\alpha=\beta=\dfrac{1}{2}$, on a :

$$\begin{cases} a_k = 0 \ (k \geq 0) \\ b_k = \dfrac{1}{2} \ (k \geq 1) \end{cases}$$

Ce qui donne la récurrence :

$$\begin{cases} P_{-1}(x)=0,\ P_0(x)=\dfrac{1}{\|Q_0\|}=\sqrt{\dfrac{2}{\pi}} \\ \dfrac{1}{2}P_{k+1}(x)+\dfrac{1}{2}P_{k-1}(x)=xP_k(x)\ (k\geq 0) \end{cases}$$

La matrice T_n est alors donnée par :

$$T_n = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

(b) Dans le problème 3 Q5 on a vu que les valeurs propres de T_n sont :

$$\lambda_j = \mathrm{Cos}\left(\frac{j\pi}{n+1}\right)\ (1 \leq j \leq n)$$

L'espace propre associé à λ_j est la droite engendrée par le vecteur $v^{(j)}$ de composantes :

$$v_k^{(j)} = \mathrm{Sin}\left(k\frac{j\pi}{n+1}\right)\ (1 \leq k \leq n)$$

(c) On vérifie par récurrence que $P_n(\mathrm{Cos}(t)) = \sqrt{\dfrac{2}{\pi}}\dfrac{\mathrm{Sin}((n+1)t)}{\mathrm{Sin}(t)}$ pour $n \geq 1$.

On en déduit alors que $P_n(x) = \sqrt{\dfrac{2}{\pi}}\dfrac{\mathrm{Sin}((n+1)\mathrm{ArcCos}(x))}{\mathrm{Sin}(\mathrm{ArcCos}(x))}$ pour $|x| \leq 1$.

Les résultats du (b) se retrouvent alors facilement.

Problème 26 : Inégalités de Newton et racines réelles d'un polynôme

Pour tout entier $n \geq 1$, on désigne par $C_n[x]$ [Resp. $IR_n[x]$] l'espace vectoriel des polynômes à coefficients complexes [Resp. réels] de degré au plus n.

Les fonctions symétriques élémentaires sont définies par :

$$s_k : \begin{array}{ccc} C^n & \to & C \\ (x_1, x_2, \cdots, x_n) & \mapsto & \displaystyle\sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \ldots x_{i_k} \end{array} \quad (1 \leq k \leq n)$$

On a alors les « formules de Viete » :

$$P(x) = \prod_{i=1}^{n} (x - x_i) = \sum_{k=0}^{n} (-1)^k s_k(x_1, \cdots, x_n) x^{n-k}$$

où on a posé $s_0(x) = 1$.

Les fonctions symétriques élémentaires normalisées sont définies par :

$$\forall x \in C^n,\ S_k(\mathrm{x}) = \frac{1}{\mathrm{C}_\mathrm{n}^\mathrm{k}} s_k(\mathrm{x}) \quad (0 \le k \le n)$$

Q1. *Calculer* $S_k(x)$ *pour* $x = (\alpha, \cdots, \alpha) \in C^n$ *et* $k \in \{0, 1, \cdots, n\}$.

Q2. *Soit* $x = (x_1, \cdots, x_n) \in C^n$, *avec* $x_i \neq 0$ *pour tout i = 1, 2, ···, n.*

Exprimer $S_k\left(\frac{1}{x_1}, \frac{1}{x_2}, \cdots, \frac{1}{x_n}\right)$ *à l'aide de fonctions symétriques de x.*

Q3. *Soit* $P(x) = x^n + \sum_{k=0}^{n-1} a_k x^k \in C_n[x]$ *de racines complexes* $x_1, x_2, \cdots, x_n$ *et* $Q = \frac{1}{n} P'$ *son polynôme dérivé normalisé de racines complexes* $y_1, y_2, \cdots, y_{n-1}$.

Montrer que $S_k(x_1, x_2, \cdots, x_n) = S_k(y_1, y_2, \cdots, y_{n-1})$ *pour* $0 \le k \le n-1$.

Q4. *Montrer que pour tout* $x \in IR^n$ *et tout* $n \ge 2$, *on a :*

$$S_1(x)^2 \ge S_2(x)$$

l'égalité étant réalisée si et seulement si $S_1(x) = S_2(x) = 0$ *ou tous les* x_i *sont identiques.*

Q5. *On suppose que* $n \ge 2$ *et que* $P(x) = x^n + \sum_{k=0}^{n-1} a_k x^k \in IR_n[x]$ *admet n racines réelles* $x_1 \le x_2 \le \cdots \le x_n$.

(a) Montrer que :

$$S_k(x_1, x_2, \cdots, x_n)^2 \ge S_{k-1}(x_1, x_2, \cdots, x_n) S_{k+1}(x_1, x_2, \cdots, x_n) \quad (1 \le k \le n-1)$$

l'égalité étant réalisée si et seulement si les deux membres sont nuls ou tous les x_i *sont identiques.*

(b) Montrer les « inégalités de Newton » :

$$a_k^{\ 2} \ge \frac{n-k+1}{n-k} \frac{k+1}{k} a_{k-1} a_{k+1} \quad (1 \le k \le n-1)$$

l'égalité étant réalisée si et seulement si les deux membres sont nuls ou tous les x_i *sont identiques (on a posé* $a_n = 1$*).*

Q6. *On suppose que* $n \ge 2$ *et que* $P(x) = x^n + \sum_{k=0}^{n-1} a_k x^k \in IR_n[x]$ *admet n racines complexes* $x_1, x_2, \cdots, x_n$ *telles que* $Ré(x_i) < 0$ *pour tout i = 1, 2, ···, n.*

(a) Montrer que $a_k \ge 0$ *pour tout k = 0, 1, ···, n.*

(b) On suppose que P admet une racine réelle négative d'ordre $p \ge 2$ *et on veut montrer qu'il existe alors un indice* $k \in \{1, 2, \cdots, n-1\}$ *tel que :*

$$a_k^{\ 2} - 4 a_{k-1} a_{k+1} \le 0$$

(i) Montrer le résultat pour n = 2 et n = 3.

(ii) Montrer le résultat pour $n > 3$.

Q7. *Soit* $P(x)=\sum_{k=0}^{n} a_k x^k \in C[x]$ *avec* $a_n \neq 0$. *Montrer que pour toute racine* λ *de P, on a :*

$$|\lambda| \leq Max\left\{\frac{|a_0|}{|a_n|},\ 1+\frac{|a_i|}{|a_n|};\ 1 \leq i \leq n-1\right\}$$

Q8. *Soit* $(P_k)_{k\in IN}$ *une suite de polynômes unitaires à coefficients réels de même degré* $n \geq 1$ *et scindés sur IR. On note* $P_k(x)=\sum_{k=0}^{n} a_j^{(k)} x^{n-j} \in C[x]$ *avec* $a_0^{(k)}=1$ *et on suppose que pour tout j = 0, 1, ⋯, n – 1, la suite* $\left(a_j^{(k)}\right)_{k\in IN}$ *converge vers* $a_j \in IR$.

Montrer alors que le polynôme $P(x)=\sum_{j=0}^{n} a_j x^{n-j}$ *(*a_0 *= 1) est aussi scindé sur IR.*

Q9. *On suppose que* $n \geq 2$ *et on se donne* $P(x)=\sum_{k=0}^{n} a_k x^k \in IR_n[x]$ *à coefficients strictement positifs avec* $a_n = 1$.

On se propose de montrer que si $a_k^{\ 2} - 4a_{k-1}a_{k+1} > 0$ *pour tout* $k \in \{1,2,\cdots,n-1\}$*, alors toutes les racines de P sont simples et réelles.*

(a) Montrer le résultat pour n = 2.

On suppose que le résultat est vrai pour $n-1 \geq 2$ *et pour P vérifiant les conditions ci-dessus, on note* $Q(x)=P(x)-a_0$.

(b) Que peut-on dire des racines de Q?

Pour tout réel $t \geq 0$*, on pose* $Q(t,x)=Q(x)+t$ *et on désigne par* $N(t)$ *le nombre de racines réelles distinctes de* $x \mapsto Q(t,x)$.

On pose $S=\{t \in [0,a_0];\ N(t)=n\}$.

(c) Montrer que S admet une borne supérieure $\alpha > 0$ *avec* $N(\alpha)=n$.

(d) Montrer que $a_0 = \alpha$ *et conclure.*

(e) Donner un contre-exemple dans le cas où les coefficients de P ne sont pas tous de même signe.

Solution

Q1. Si tous les coefficients de x sont identiques, alors on a :

$$s_k(x) = Card\{(i_1,i_2,\cdots,i_k)\in IN^k; 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}\alpha^k = C_n^k \alpha^k$$

On a donc $S_k(x)=\alpha^k$, pour tout k = 0, 1, ⋯, n.

Q2. Soit $P(x)=\sum_{k=0}^{n} a_k x^k \in C_n[x]$ admettant $x_1,x_2,\cdots,x_n$ pour racines. On a alors $\frac{a_k}{a_n} = (-1)^{n-k} C_n^k S_{n-k}(x)$.

Pour $x \neq 0$, on a $P\left(\frac{1}{x}\right) = \sum_{k=0}^{n} a_k \frac{1}{x^k} = \frac{1}{x^n} \sum_{k=0}^{n} a_{n-k} x^k$ et $x^n P\left(\frac{1}{x}\right)$ admet les $\frac{1}{x_i}$ pour racines. On en déduit donc que $\frac{a_{n-k}}{a_0} = (-1)^{n-k} C_n^k S_{n-k}\left(\frac{1}{x_1}, \cdots, \frac{1}{x_n}\right)$, ce qui donne :

$$S_k\left(\frac{1}{x_1}, \cdots, \frac{1}{x_n}\right) = \frac{S_{n-k}(x)}{S_n(x)}$$

Q3. On a $P(x) = \prod_{i=1}^{n} (x - x_i) = \sum_{k=0}^{n} (-1)^k C_n^k S_k(x_1, \cdots, x_n) x^{n-k}$, donc :

$$Q(x) = \frac{1}{n} P'(x) = \sum_{k=0}^{n-1} (-1)^k \frac{n-k}{n} C_n^k S_k(x_1, \cdots, x_n) x^{n-k-1}$$
$$= \sum_{k=0}^{n-1} (-1)^k C_{n-1}^k S_k(x_1, \cdots, x_n) x^{n-1-k}$$

Mais on peut aussi écrire que :

$$Q(x) = \prod_{i=1}^{n} (x - y_i) = \sum_{k=0}^{n-1} (-1)^k C_{n-1}^k S_k(y_1, \cdots, y_{n-1}) x^{n-1-k}$$

Et en identifiant les coefficients de x^{n-1-k}, pour tout k = 1, 2, ⋯, n – 1, on déduit que $S_k(x_1, x_2, \cdots, x_n) = S_k(y_1, y_2, \cdots, y_{n-1})$ pour $0 \leq k \leq n-1$.

Q4. On a :

$$S_1(x)^2 - S_2(x) = \frac{1}{n^2(n-1)} \left\{ (n-1)\left(\sum_{i=1}^{n} x_i\right)^2 - 2n \sum_{1 \leq i < j \leq n} x_i x_j \right\}$$
$$= \frac{1}{n^2(n-1)} \left\{ n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \right\}$$

L'inégalité de Cauchy–Schwarz nous dit que $\left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_{i=1}^{n} x_i^2$, l 'égalité étant réalisée si et seulement si tous les x_i sont identiques.

Q5. (a) Montrons le résultat par récurrence sur $n \geq 2$.

Pour $n = 2$, il s'agit de montrer que $\frac{1}{4}(x_1 + x_2)^2 \geq x_1 x_2$, ce qui équivaut à $(x_1 - x_2)^2 \geq 0$, l'égalité étant réalisée si et seulement si $x_1 = x_2$.

Supposons le résultat vrai pour $n - 1 \geq 2$.

Le polynôme P ayant toutes ces racines réelles, le théorème de Rolle nous dit que $Q = \frac{1}{n} P'$ admet n – 1 racines réelles $y_1 \leq y_2 \leq \cdots \leq y_{n-1}$ telles que $x_i \leq y_i \leq x_{i+1}$ pour tout i = 1, 2, ⋯, n – 1.

En écrivant que $S_k(x_1, x_2, \cdots, x_n) = S_k(y_1, y_2, \cdots, y_{n-1})$ pour $0 \leq k \leq n-1$, on déduit les inégalités souhaitées pour k = 1, 2, ⋯, n – 2.

Si les x_i ne sont pas tous identiques, il en est de même des y_i et les inégalités sont strictes ou réduites à 0 = 0.

Il reste à prouver le résultat pour $k = n - 1$. Il s'agit de montrer que :

$$S_{n-1}(x_1, x_2, \cdots, x_n)^2 \geq S_{n-2}(x_1, x_2, \cdots, x_n) \prod_{i=1}^{n} x_i \quad (1)$$

avec égalité si et seulement si tous les x_i sont identiques ou si les deux membres sont nuls.

Si l'un des x_i est nul alors l'inégalité est évidente ainsi que le cas d'égalité. On suppose donc que tous les x_i sont non nuls et on pose $x_i' = \dfrac{1}{x_i}$ pour tout $i = 1, \cdots, n$.

En Q2, on a vu que $\dfrac{S_{n-k}(x_1, x_2, \cdots, x_n)}{S_n(x_1, x_2, \cdots, x_n)} = S_k(x_1', x_2', \cdots, x_n')$ pour tout $k = 1, \cdots, n$.

Et l'inégalité (1) équivaut à $S_1(x_1', x_2', \cdots, x_n')^2 \geq S_2(x_1', x_2', \cdots, x_n')$ qui a été montrée en Q4.

(b) On a $a_{n-k} = (-1)^n C_n^k S_k(x)$ et :

$$\begin{aligned} S_k(x)^2 - S_{k-1}(x) S_{k+1}(x) &= \frac{1}{\left(C_n^k\right)^2} a_{n-k}{}^2 - \frac{1}{C_n^{k-1} C_n^{k+1}} a_{n-k+1} a_{n-k-1} \\ &= \frac{1}{\left(C_n^k\right)^2} \left\{ a_{n-k}{}^2 - \frac{n-k+1}{k} \frac{k+1}{n-k} a_{n-k+1} a_{n-k-1} \right\} \end{aligned}$$

On déduit alors que $a_{n-k}{}^2 - \dfrac{k+1}{k} \dfrac{n-k+1}{n-k} a_{n-k+1} a_{n-k-1} \geq 0$, ce qui équivaut à :

$$a_k{}^2 - \frac{n-k+1}{n-k} \frac{k+1}{k} a_{k-1} a_{k+1} \geq 0 \quad (1 \leq k \leq n-1)$$

Remarque — Pour $n = 2$, on retrouve la condition nécessaire et suffisante sur le discriminant $\Delta = a_1{}^2 - 4a_0 a_2 > 0$, pour avoir deux racines réelles distinctes. Dans les questions qui suivent on se pose la question de savoir si les inégalités de Newton sont aussi des conditions suffisantes pour que le polynôme soit scindé sur *IR*. La réponse est non, mais on va donner une autre condition suffisante.

Q6. (a) En regroupant les racines complexes non réelles avec leurs conjugués (P est à coefficients réels) et en notant $-t_1, -t_2, \cdots, -t_p$ les racines réelles avec $t_i > 0$, on a :

$$P(x) = \prod_{i=1}^{q} \left(x^2 - 2R\acute{e}(x_i) x + |x_i|^2 \right) \prod_{i=1}^{p} (x + t_i)$$

avec $-2R\acute{e}(x_i) > 0$ pour tout $i = 1, \cdots, q$. Il en résulte que tous les coefficients de P sont positifs.

(b) Pour $n = 2$, le résultat provient du fait que si P admet une racine double alors son discriminant $\Delta = a_1{}^2 - 4a_0 a_2$ est nul.

Pour $n = 3$, on a $P(x) = a_0 + a_1 x + a_2 x^2 + x^3 = (x + x_1)^2 (x + x_2)$ avec $x_1 > 0$ et $x_2 > 0$.

Supposons que $a_1{}^2 - 4a_0 a_2 > 0$ et $a_2{}^2 - 4a_1 a_3 > 0$. On a alors $x_1{}^3(x_1 - 4x_2) > 0$ et $x_2(x_2 - 4x_1) > 0$, soit en tenant compte de la positivité de tous les coefficients

$x_1 - 4x_2 > 0$ et $x_2 - 4x_1 > 0$. Ce qui donne $x_2 > 4x_1 > 16x_2 > 0$, c'est-à-dire une impossibilité.

(c) On a $P(x) = \sum_{k=0}^{n} a_k x^k = (x + x_1)^2 \sum_{k=0}^{n-2} b_k x^k$, avec $a_k \geq 0$, $b_k \geq 0$ et $x_1 > 0$.

Par identification, on a alors :

$$a_k = b_{k-2} + 2x_1 b_{k-1} + x_1^2 b_k \quad (0 \leq k \leq n)$$

où on a posé $b_i = 0$ pour $i < 0$ et $i > n - 2$.

Supposons que $a_k^2 - 4a_{k-1}a_{k+1} > 0$ pour tout $k = 1, 2, \cdots, n - 1$. On a alors :

$$x_1^4 b_k^2 + b_{k-2}^2 - 4x_1^3 b_k b_{k-1} - 4x_1 b_{k-1} b_{k-2} > r_k$$

où on a posé :

$$r_k = 14x_1^2 b_k b_{k-2} + 8x_1^3 b_{k+1} b_{k-2} + 4x_1^2 b_{k+1} b_{k-3} + 8x_1 b_k b_{k-3} + 4b_{k-3} b_{k-1} + 4x_1^4 b_{k-1} b_{k+1} \geq 0$$

En notant $p_k = x_1 b_k - 4b_{k-1}$ et $q_k = b_{k-1} - 4x_1 b_k$, on a :

$$p_k x_1^3 b_k + q_{k-1} b_{k-2} > 0 \quad (1 \leq k \leq n-1)$$

Pour $k = 1$, on a $p_1 x_1^3 b_1 > 0$ et $p_1 = x_1 b_1 - 4b_0 > 0$, ce qui entraîne $q_1 = b_0 - 4x_1 b_1 < 0$.

Pour $k = n - 1$, on a $q_{n-2} b_{n-3} > 0$ et $q_{n-2} > 0$.

On désigne par j le plus petit entier compris entre 2 et $n - 2$ qui vérifie $q_j > 0$.

On a alors $p_j x_1^3 b_j + q_{j-1} b_{j-2} > 0$ avec $q_{j-1} \leq 0$, ce qui entraîne $p_j > 0$, soit $x_1 b_j - 4b_{j-1} > 0$ et nécessairement $q_j = b_{j-1} - 4x_1 b_j \leq 0$. C'est-à-dire qu'on aboutit à une contradiction.

Q7. $\dfrac{(-1)^n}{a_n} P$ est le polynôme caractéristique de la matrice de Frobénius définie par $A = \begin{pmatrix} 0 & 0 & \cdots & 0 & b_0 \\ 1 & 0 & \cdots & 0 & b_1 \\ . & . & . & . & . \\ 0 & \cdots & 1 & 0 & b_{n-2} \\ 0 & \cdots & 0 & 1 & b_{n-1} \end{pmatrix}$, où on a noté $b_i = -\dfrac{a_i}{a_n}$ pour $0 \leq i \leq n-1$ (problème 9, Q1). On sait alors que le rayon spectral de A, $\rho(A)$, est majoré par $\|A\|$ pour toute norme matricielle induite par une norme vectorielle (problème 2, Q1). En particulier on a, pour toute racine λ de P :

$$|\lambda| \leq \rho(A) \leq \|A\|_\infty = Max\{|b_0|,\ 1 + |b_i|;\ 1 \leq i \leq n-1\}$$

(problème 1, Q6 (a)).

Q8. Les suites $\left(a_j^{(k)}\right)_{k \in IN}$ étant convergentes sont bornées par une constante $M > 0$. En notant $x^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}\right)$ la suite des zéros de P_k, on a alors avec Q7 :

$$\left|x_j^{(k)}\right| \leq Max\left\{\left|a_0^{(k)}\right|,\ 1 + \left|a_i^{(k)}\right|;\ 1 \leq i \leq n-1\right\} \leq 1 + M$$

C'est-à-dire que la suite $\left(x^{(k)}\right)_{k \in IN}$ est bornée dans IR^n et on peut en extraire une sous suite $\left(x^{(\varphi(k))}\right)_{k \in IN}$ qui converge vers $x \in IR^n$.

Avec $a_j^{(k)} = (-1)^j s_k\left(x^{(k)}\right)$ et la continuité des fonctions symétriques élémentaires, on déduit que $a_j = (-1)^j s_k(x)$ et $P(x) = \prod_{i=1}^{n}(x - x_i)$.

Q9. (a) On a $P(x) = a_0 + a_1 x + x^2$ avec $\Delta = a_1^2 - 4a_0 > 0$, donc P admet deux racines réelles distinctes.

(b) On a $Q(x) = xR(x)$ où $R(x) = x^{n-1} + \sum_{k=0}^{n-2} b_k x^k$ avec $b_k = a_{k+1}$ et :

$$b_k^2 - 4b_{k-1}b_{k+1} = a_{k+1}^2 - 4a_k a_{k+2} > 0 \quad (1 \leq k \leq n-2)$$

Avec l'hypothèse de récurrence, on déduit que R admet $n - 1$ racines réelles simples.

De plus, on a $R(x) \geq a_1 > 0$ pour tout $x \geq 0$. On en conclut alors que Q admet n racines réelles simples, la plus grande étant 0.

(c) S est majoré par a_0 et non vide ($0 \in S$), il admet donc une borne supérieure $\alpha \in [0, a_0]$.

On a $Q(0,x) = Q(x)$ qui admet n racines réelles $x_1(0) < x_2(0) < \cdots < x_n(0) = 0$.

Pour chacune de ces racines, on a $Q(0, x_j(0)) = 0$ avec $\frac{\partial}{\partial x} Q(0, x_j(0)) \neq 0$. Le théorème des fonctions implicites nous dit alors qu'il existe $\varepsilon_j > 0$ et une fonction continue $x_j : [0, \varepsilon_j] \to IR$ telle que $Q(t, x_j(t)) = 0$ pour tout $t \in [0, \varepsilon_j]$.

On en déduit alors qu'on peut trouver $\varepsilon > 0$ tel que pour tout $t \in [0, \varepsilon]$ le polynôme $x \mapsto Q(t,x)$ admet n racines réelles distinctes, ce qui entraîne que $\alpha > 0$.

On peut écrire que $\alpha = \operatorname{Lim}_{k \to +\infty} t_k$ avec $t_k \in S$ pour tout entier k. Avec Q8, on déduit alors que $x \mapsto Q(\alpha, x)$ est scindé sur IR. En tenant compte du fait que $Q(\alpha, x) \geq \alpha > 0$ pour tout $x \geq 0$, on déduit que toutes les racines de $x \mapsto Q(\alpha, x)$ sont réelles et négatives.

Si $x \mapsto Q(\alpha, x) = \sum_{k=0}^{n} c_k x^k$ admet une racine multiple alors il existe $k \in \{1, \cdots, n-1\}$ tel que $c_k^2 - 4c_{k-1}c_{k+1} \leq 0$.

Si $k = 1$, alors on a $c_1^2 \leq 4\alpha \cdot a_2 \leq 4a_0 a_2$ et si $k > 1$, on a $a_k^2 - 4a_{k-1}a_{k+1} \leq 0$. C'est-à-dire que dans tous les cas on aboutit à une impossibilité.

On a donc $N(\alpha) = n$.

(d) Si $\alpha < a_0$, on déduit alors, en reprenant le raisonnement du (b) (continuité des racines), qu'il existe $\varepsilon > 0$ tel que $N(t) = n$ pour $|t - \alpha| < \varepsilon$ et $t \in [0, a_0]$, ce qui contredit le fait que α est la borne supérieure de S.

On a donc $\alpha = a_0$ et P admet n racines réelles simples.

(e) Le polynôme $P(x)=x^3-5x^2+6x+1$ admet deux racines complexes non réelles et les inégalités sur les coefficients sont vérifiées.

Problème 27 : La méthode Newton–Maehly

Pour tout entier $n\geq 1$, on désigne par $IR_n[x]$ l'espace vectoriel des polynômes à coefficients réels de degré au plus n.

On suppose dans ce problème que toutes les racines du polynôme $P\in IR_n[x]$ sont réelles. On note $\lambda_1>\lambda_2>\cdots>\lambda_p$ les racines distinctes de P dans IR. La racine λ_j étant de multiplicité $m_j\geq 1$ pour tout j = 1, 2, ⋯, p, avec $\sum_{j=1}^{p} m_j = n$.

Q1. *Montrer que le polynôme dérivé P' admet les $\lambda_1,\lambda_2,\cdots,\lambda_p$ pour racines de multiplicités respectives $m_1-1,\ m_2-1,\ \cdots,\ m_p-1$ (une multiplicité nulle signifie que λ_j n'est pas racine de P') et des racines simples $\mu_j\in\left]\lambda_j,\lambda_{j+1}\right[$ avec $1\leq j\leq p-1$.*

Q2. *Montrer que pour tout $x>\lambda_1$ et tout k = 0, 1, ⋯, n, on a $a_nP^{(k)}(x)>0$.*

Q3. *Pour tout $x_0>\lambda_1$, on définit la suite $(x_k)_{k\in IN}$ par :*

$$\forall k\geq 0,\ x_{k+1}=x_k-\frac{P(x_k)}{P'(x_k)}$$

Montrer que cette suite converge vers λ_1 en décroissant et donner une majoration de l'erreur.

Q4. *Soit $P(x)=\sum_{k=0}^{n} a_k x^k\in C[x]$ avec $a_n\neq 0$. Montrer que pour toute racine λ de P, on a :*

$$|\lambda|\leq Max\left\{\frac{|a_0|}{|a_n|},\ 1+\frac{|a_i|}{|a_n|};\ 1\leq i\leq n-1\right\}$$

$$|\lambda|\leq Max\left\{1,\ \sum_{i=0}^{n-1}\frac{|a_i|}{|a_n|}\right\}$$

Q5. *Comment utiliser ce qui précède pour calculer toutes les racines de P.*

Solution

Q1. Si, pour $1\leq j\leq p$, on a $m_j\geq 2$, alors λ_j est racine d'ordre m_j-1 de *P*'. On a donc ainsi $\sum_{j=1}^{p}(m_j-1)=n-p$ racines réelles pour *P*'.

D'autre part, le théorème de Rolle nous dit que pour $1\leq j\leq p-1$ il existe $\mu_j\in\left]\lambda_j,\lambda_{j+1}\right[$ tel que $P'(\mu_j)=0$, ce qui donne $p-1$ racines réelles supplémentaires pour *P*'. On a donc un total de $n-1$ racines réelles pour *P*' et les μ_j sont nécessairement simples.

En particulier les racines de P' sont réelles et contenues dans l'intervalle $[\lambda_p, \lambda_1]$.

Q2. On peut supposer, sans restreindre la généralité, que $a_n > 0$.

En Q1, on a vu que pour $0 \leq k \leq n-1$ les racines de $P^{(k)}$ sont dans l'intervalle $[\lambda_p, \lambda_1]$. On en déduit alors que $P^{(k)}$ est de signe constant sur $]\lambda_1, +\infty[$. Soit avec $a_n > 0$, $P^{(k)}(x) > 0$ pour $x > \lambda_1$ et $k = 0, 1, \cdots, n-1$.

Pour $k = n$, on a $P^{(n)}(x) = n! a_n > 0$.

Q3. On pose :

$$\forall x \geq \lambda_1,\ g(x) = \begin{cases} \lambda_1 \text{ si } x = \lambda_1 \\ x - \dfrac{P(x)}{P'(x)} \text{ si } x > \lambda_1 \end{cases}$$

et on définit ainsi une fonction indéfiniment dérivable sur $[\lambda_1, +\infty[$.

Pour $x > \lambda_1$, on a $\dfrac{P(x)}{P'(x)} > 0$ (d'après Q2) et $g(x) < x$.

Avec la formule de Taylor à l'ordre 2, on a :

$$\forall x > \lambda_1,\ g(x) - \lambda_1 = (x - \lambda_1)^2 \frac{P''(c)}{2P'(x)}$$

où c est un réel compris entre x et λ_1. Et avec $\dfrac{P''(c)}{P'(x)} > 0$ (Q2), on déduit que $g(x) > \lambda_1$ pour tout $x > \lambda_1$, c'est-à-dire que $g([\lambda_1, +\infty[) \subset [\lambda_1, +\infty[$ et la suite $(x_k)_{k \in IN}$ est bien définie.

On a donc :

$$\forall k \geq 0,\ \lambda_1 < x_{k+1} < x_k$$

c'est-à-dire que la suite $(x_k)_{k \in IN}$ est strictement décroissante et minorée. Elle converge donc nécessairement vers λ_1.

Pour tout $x > \lambda_1$, on a $1 - g'(x) = \dfrac{P'(x)^2 - P(x)P''(x)}{P'(x)^2} = \dfrac{-\left(\dfrac{P'}{P}\right)'(x)}{\left(\dfrac{P'}{P}\right)^2(x)}$, c'est-à-dire avec $P(x) = a_n \prod_{j=1}^{p} (x - \lambda_j)^{m_j}$ et $P'(x) = a_n \sum_{j=1}^{p} m_j \dfrac{P(x)}{x - \lambda_j}$:

$$1 - g'(x) = \frac{\displaystyle\sum_{j=1}^{p} \frac{m_j}{(x - \lambda_j)^2}}{\left(\displaystyle\sum_{j=1}^{p} \frac{m_j}{x - \lambda_j}\right)^2}$$

Avec l'inégalité de Cauchy–Schwarz, on peut écrire :

$$\left(\sum_{j=1}^{p}\frac{m_j}{x-\lambda_j}\right)^2=\left(\sum_{j=1}^{p}\frac{\sqrt{m_j}}{x-\lambda_j}\sqrt{m_j}\right)^2\leq\sum_{j=1}^{p}\frac{m_j}{\left(x-\lambda_j\right)^2}\sum_{j=1}^{p}m_j=n\sum_{j=1}^{p}\frac{m_j}{\left(x-\lambda_j\right)^2}$$

Il en résulte que :

$$\forall x>\lambda_1, 0<g'(x)<1-\frac{1}{n}$$

Et avec le théorème des accroissements finis, on a :

$$\forall k\geq 0,\ 0\leq x_{k+1}-\lambda_1=g(x_k)-g(\lambda_1)\leq\left(1-\frac{1}{n}\right)(x_k-\lambda_1)$$

ce qui entraîne par récurrence :

$$\forall k\geq 0,\ 0\leq x_k-\lambda_1\leq\left(1-\frac{1}{n}\right)^k(x_0-\lambda_1)$$

Remarque 1 — Avec le théorème des accroissements finis, on a :

$$\forall k\geq 0,\ 0\leq x_k-\lambda_1\leq M_1^{\,k}(x_0-\lambda_1)$$

où $M_1=Sup\left\{\left|g'(x)\right|;\ x\in[\lambda_1,x_0]\right\}$.

En tenant compte de $g'(x)=\dfrac{P(x)P''(x)}{P'(x)^2}$, on déduit que dans le cas où λ_1 est une racine simple, on a $g'(\lambda_1)=0$ et la convergence est d'autant plus rapide que x_0 est proche de λ_1.

De manière plus générale, on a $\underset{x\to\lambda_1}{Lim}\,g'(x)=1-\dfrac{1}{m_1}$.

Q4. Voir le problème 26, Q7.

Q5. Le calcul des autres racines peut se faire en appliquant ce qui précède à $P_1(x)=\dfrac{P(x)}{(x-\lambda_1)}$ et en remarquant que :

$$\frac{P_1(x)}{P_1'(x)}=\frac{P(x)}{P'(x)-\dfrac{P(x)}{x-\lambda_1}}$$

La limite de la suite $(x_k)_{k\in IN}$ correspondante sera encore λ_1 si $m_1>1$ et λ_2 si $m_1=1$.

De manière plus générale si $\xi_1\geq\xi_2\geq\cdots\geq\xi_n$ sont les *n* racines réelles de *P* et si on dispose de valeurs approchées des $m-1$ premières racines $\xi_1\geq\cdots\geq\xi_{m-1}$ alors le calcul de ξ_m se fait en remplaçant *P* par $P_{m-1}(x)=\dfrac{P(x)}{(x-\xi_1)\cdots(x-\xi_{m-1})}$ avec :

$$\frac{P_{m-1}(x)}{P_{m-1}'(x)}=\frac{P(x)}{P'(x)-P(x)\displaystyle\sum_{j=1}^{m-1}\frac{1}{x-\xi_j}}$$

C'est-à-dire qu'on considère la suite $\left(x_k^{(m)}\right)_{k\in IN}$ définie par :

$$\begin{cases} x_0^{(m)} = x_0 \\ x_{k+1}^{(m)} = x_k^{(m)} - \dfrac{P_{m-1}\left(x_k^{(m)}\right)}{P'_{m-1}\left(x_k^{(m)}\right)} \ (k \geq 0) \end{cases}$$

qui converge en décroissant strictement vers ξ_m.

Remarque — Pour accélérer la convergence, on considère tout d'abord la suite $\left(y_k^{(m)}\right)_{k\in IN}$ définie par :

$$\begin{cases} y_0^{(m)} = x_0 \\ y_{k+1}^{(m)} = y_k^{(m)} - 2\dfrac{P_{m-1}\left(y_k^{(m)}\right)}{P'_{m-1}\left(y_k^{(m)}\right)} \ (k \geq 0) \end{cases}$$

On utilise cette suite tant que $y_{k+1}^{(m)} < y_k^{(m)}$ (la suite est théoriquement croissante, mais dès qu'on est très proche de la solution cela peut ne plus être vrai à cause des erreurs d'arrondis), puis dès que $y_{k+1}^{(m)} \geq y_k^{(m)}$ pour un indice k_0, on utilise la suite $\left(x_k^{(m)}\right)_{k\in IN}$ avec $y_{k_0}^{(m)}$ comme point de départ.

Problème 28 : Perturbation d'un polynôme. Influence sur les racines. Agrégation 1995, extrait

Q1. *Soit D un disque fermé du plan complexe, Γ son bord, f une fonction holomorphe sur un voisinage de D, qui ne s'annule pas sur Γ. Montrer que le nombre de racines de f contenues dans D est égal à* $\dfrac{1}{2i\pi}\displaystyle\int_\Gamma \dfrac{f'(z)}{f(z)}dz$.

Q2. *Soit D un disque fermé du plan complexe, Γ son bord, f et g deux fonctions holomorphes sur un voisinage de D, telles que* $\forall z \in \Gamma,\ |f(z)-g(z)| < |f(z)|$. *Montrer que f et g ont le même nombre de zéros dans D. (Ce résultat est connu sous le nom de théorème de Rouché).*

Q3. *On suppose que la racine $z_0 \neq 0$ du polynôme $P(z) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_0$ est simple. Soit $\varepsilon \in \left]0, |z_0|\right[$ tel que le disque fermé de centre z_0 et de rayon ε ne contienne aucune autre racine de P. Soit $k \leq n$ un entier fixé. Montrer qu'il existe $\rho > 0$ tel que pour tout $h \in C$ tel que $|h| < \rho$ le polynôme $Q(z) = P(z) + hz^k$ ait une racine et une seule ζ dans le disque de centre z_0 et de rayon ε, et que l'application $h \mapsto \zeta$ est holomorphe. Calculer $\dfrac{d\zeta}{dh}$ au point $h = 0$.*

Application au polynôme $(z+1)(z+2)(z+3)\cdots(z+19)(z+20)$ pour k = 19. Estimer numériquement $\dfrac{d\zeta}{dh}$ pour la racine $z_0 = -20$. Que conclure de ce résultat ?

Q4. *Soient $z_0, z_1, \cdots, z_{n-1}$ n complexes distincts. On pose :*

$$P(z) = (z - z_0)(z - z_1)\cdots(z - z_{n-1}) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_0$$

Montrer qu'il existe $\rho > 0$ *et* $\varepsilon \in \left]0, \frac{1}{2} \underset{i \neq j}{Min} |z_i - z_j|\right[$ *tels que pour tout élément* $(\alpha_i)_{0 \le i \le n-1}$ *de* C^n *vérifiant* $\forall i \in \{0, 1, \cdots, n-1\}$, $|a_i - \alpha_i| < \rho$, *le polynôme* $Q(z) = P(z) + \sum_{k=0}^{n-1} \alpha_k z^k$ *ait pour tout i une racine* ζ_i *et une seule dans le disque de centre* z_i *et de rayon* ε*. Montrer que l'application de la boule de centre* (a_i) *et de rayon* ρ *de* C^n *dans* C^n *définie par* $(\alpha_i) \mapsto (\zeta_i)$ *est de classe* C^∞.

Solution

Q1. On rappelle que si Ω est un ouvert connexe de C, f est holomorphe non nulle sur Ω et K un compact de Ω, alors f n'a qu'un nombre fini de zéros dans K (voir Cartan, p. 41).

En notant $z_1, z_2, \cdots, z_k$ les zéros distincts de f dans D, de multiplicités respectives $m_1, m_2, \cdots, m_k$ et en utilisant le théorème des résidus, on a :

$$\frac{1}{2i\pi}\int_\Gamma \frac{f'(z)}{f(z)}dz = \sum_{j=1}^{k} Rés\left(\frac{f'(z)}{f(z)}, z_j\right)$$

Le cercle Γ étant parcouru une seule fois dans le sens direct.

En écrivant que $f(z) = (z - z_j)^{m_j} g(z)$ et $f'(z) = m_j (z - z_j)^{m_j - 1} h(z)$ avec g et h holomorphes au voisinage de z_j et telles que $g(z_j) \neq 0$, $h(z_j) \neq 0$, on déduit que :

$$\frac{f'(z)}{f(z)} = \frac{m_j}{z - z_j}\varphi(z)$$

la fonction φ étant holomorphe au voisinage de z_j.

Il en résulte alors que $Rés\left(\frac{f'(z)}{f(z)}, z_j\right) = m_j$ et :

$$\frac{1}{2i\pi}\int_\Gamma \frac{f'(z)}{f(z)}dz = \sum_{j=1}^{k} m_j = n$$

où n est le nombre de zéros de f comptés avec leur multiplicité, dans l'intérieur de D.

Q2. La condition $\forall z \in \Gamma$, $|f(z)| > |f(z) - g(z)| \ge 0$, nous dit que f ne s'annule pas sur Γ.

En écrivant que, pour tout $z \in \Gamma$, on a :

$$|g(z)| \ge |f(z)| - |f(z) - g(z)| > 0$$

on déduit que g ne s'annule pas sur Γ.

En utilisant Q1, il s'agit alors de montrer que $\int_\Gamma \frac{f'(z)}{f(z)}dz = \int_\Gamma \frac{g'(z)}{g(z)}dz$.

On pose $h=\dfrac{g}{f}$ et on définit ainsi une fonction méromorphe sur un voisinage de D et continue sur Γ.

Pour tout $z\in\Gamma$, on a $|h(z)-1|=\left|\dfrac{g(z)-f(z)}{f(z)}\right|<1$, c'est-à-dire que le lacet $t\in[0,2\pi]\mapsto h\circ\Gamma(t)$ a son image contenue dans le disque ouvert de centre 1 et de rayon 1 et en particulier 0 n'est pas à l'intérieur de ce lacet. On déduit alors, avec le théorème des résidus que $\displaystyle\int_{h\circ\Gamma}\frac{du}{u}=0$, soit $\displaystyle\int_0^{2\pi}\frac{h'(\Gamma(t))\Gamma'(t)}{h(\Gamma(t))}dt=\int_\Gamma\frac{h'(z)}{h(z)}dz=0$ (cette intégrale est l'indice de $h\circ\Gamma$ par rapport à 0).

En remarquant que $\dfrac{h'}{h}=\dfrac{g'}{g}-\dfrac{f'}{f}$, on déduit que $\displaystyle\int_\Gamma\frac{f'(z)}{f(z)}dz=\int_\Gamma\frac{g'(z)}{g(z)}dz$, c'est-à-dire que f et g ont le même nombre de zéros à l'intérieur de Γ.

Q3. (i) z_0 étant une racine simple non nulle de P, on peut trouver $\varepsilon\in\left]0,|z_0|\right[$ assez petit tel que $P(z)\neq 0$ pour $|z-z_0|\leq\varepsilon$ et $z\neq z_0$.

On note Γ_ε le cercle de centre z_0 et de rayon ε.

Pour tout $z\in\Gamma_\varepsilon$, on a $0<M_0\leq|z|^k\leq M_1$, où $M_0=(|z_0|-\varepsilon)^k$ et $M_1=(|z_0|+\varepsilon)^k$.

La fonction P est continue et ne s'annule jamais sur le compact Γ_ε, il existe donc une constante $m_0>0$ telle que :

$$\forall z\in\Gamma_\varepsilon,\ |P(z)|\geq m_0>0$$

En choisissant $\rho>0$ tel que $\rho\cdot M_1<\dfrac{m_0}{2}$, on a pour $|h|<\rho$ et $z\in\Gamma_\varepsilon$:

$$|Q(z)|=\left|P(z)+hz^k\right|\geq|P(z)|-\left|hz^k\right|\geq\frac{m_0}{2}$$

On a donc, pour $|h|<\rho$:

$$\forall z\in\Gamma_\varepsilon,\ |Q(z)-P(z)|=\left|hz^k\right|<\frac{m_0}{2}\leq|Q(z)|$$

On déduit alors de Q2, que les polynômes P et Q ont le même nombre de racines à l'intérieur de Γ_ε, c'est-à-dire que le polynôme Q admet une unique racine simple dans le disque de centre z_0 et de rayon ε.

(ii) On note $D(0,\rho)$ le disque ouvert de centre 0 et de rayon ρ dans C.

Pour $|h|<\rho$, la fonction $z\mapsto\dfrac{zQ'(z)}{Q(z)}$ est méromorphe dans un voisinage du disque fermé de centre z_0 et de rayon ε avec un pôle simple en $\zeta=\zeta(h)$ (ce disque fermé ne contient pas 0 si $z_0\neq 0$).

Avec le théorème des résidus, on a alors :

$$\frac{1}{2i\pi}\int_{\Gamma_\varepsilon}\frac{zQ'(z)}{Q(z)}dz=Rés\left(\frac{zQ'(z)}{Q(z)},\zeta\right)=\frac{\zeta Q'(\zeta)}{Q(\zeta)}=\zeta$$

C'est-à-dire que pour tout h tel que $|h|<\rho$, on a :

$$\zeta(h) = \frac{1}{2i\pi}\int_{\Gamma_\varepsilon} \frac{z\frac{\partial}{\partial z}Q(h,z)}{Q(h,z)}\,dz$$

où on a noté $Q(h,z) = P(z) + hz^k$.

Remarque — Si $\zeta(h) = 0$, alors $\frac{zQ'(z)}{Q(z)} = \frac{zQ'(z)}{zR(z)} = \frac{Q'(z)}{R(z)}$ est holomorphe dans un voisinage du disque fermé de centre z_0 et de rayon ε et avec la formule de Cauchy, on a : $\frac{1}{2i\pi}\int_{\Gamma_\varepsilon} \frac{z\frac{\partial}{\partial z}Q(h,z)}{Q(h,z)}\,dz = 0 = \zeta(h)$.

La fonction $(h,z) \mapsto Q(h,z)$ est continue sur $D(0,\rho)\times\Gamma_\varepsilon$ et pour tout $z \in \Gamma_\varepsilon$, la fonction $h \mapsto Q(h,z)$ est holomorphe sur $D(0,\rho)$. L'intégration se faisant sur un compact de *C*, on déduit que l'application $h \mapsto \zeta(h)$ est holomorphe sur $D(0,\rho)$.

Enfin en dérivant par rapport à h l'identité $P(\zeta(h)) + h\zeta(h)^k = 0$, on déduit que :

$$\forall h \in D(0,\rho),\ \zeta'(h) = -\frac{\zeta(h)^k}{P'(\zeta(h)) + h\cdot k\cdot \zeta(h)^{k-1}}$$

(iii) Soit $P(z) = \prod_{j=1}^{n}(z+j)$. Pour $z_0 = -n$ et $k = n-1$, on a :

$$\zeta'(0) = -\frac{n^n}{n!} \cong -\frac{e^n}{\sqrt{2\pi n}}$$

En particulier, pour $n = 20$, $k = 19$, on a $|\zeta'(0)| \cong \frac{e^{20}}{\sqrt{40\pi}} \cong 4.32*10^7$.

En écrivant, pour h voisin de 0, que $\zeta(h) \cong \zeta(0) + h\zeta'(0) \cong -\left(20 + h\frac{20^{20}}{20!}\right)$, on voit qu'une petite perturbation sur un coefficient de P va entraîner une grosse incertitude sur une racine.

Q4. (i) Pour $\varepsilon \in \left]0, \frac{1}{2}\underset{i\neq j}{Min}|z_i - z_j|\right[$, les disques fermés $D(z_i,\varepsilon)$ de centre z_i et de rayon ε sont deux à deux disjoints et P admet z_i comme unique racine dans $D(z_i,\varepsilon)$ pour chaque $i = 0, 1, \cdots, n-1$.

P est continue et ne s'annule jamais sur le compact Γ_{ε_i} bord de $D(z_i,\varepsilon)$, il existe donc une constante $m_i > 0$ telle que :

$$\forall z \in \Gamma_{\varepsilon_i},\ |P(z)| \geq m_i \ (0 \leq i \leq n-1)$$

Soit $M_i > 0$ tel que $|z| \leq M_i$ sur Γ_{ε_i} et $\rho > 0$ tel que $\rho\sum_{k=0}^{n-1} M_i^k < \frac{m_i}{2}$.

Si $Q(z) = P(z) + \sum_{k=0}^{n-1} \alpha_k z^k$ est tel que $|\alpha_k - a_k| < \rho$ pour tout k = 0, 1, ⋯, n – 1, on a alors :

$$\forall z \in \Gamma_{\varepsilon_i},\ |Q(z)| = \left| P(z) + \sum_{k=0}^{n-1} (\alpha_k - a_k) z^k \right| \geq |P(z)| - \sum_{k=0}^{n-1} |\alpha_k - a_k| |z|^k > \frac{m_i}{2}$$

et :

$$\forall z \in \Gamma_{\varepsilon_i},\ |Q(z) - P(z)| = \left| \sum_{k=0}^{n-1} (\alpha_k - a_k) z^k \right| < \frac{m_i}{2} \leq |Q(z)|$$

On déduit alors du théorème de Rouché que Q admet une unique racine $\zeta_i \in D(z_i, \varepsilon)$.

(ii) Cette racine peut s'exprimer à l'aide du théorème des résidus :

$$\zeta_i = \zeta_i(\alpha_0, \alpha_1, \cdots, \alpha_{n-1}) = \frac{1}{2i\pi} \int_{\Gamma_{\varepsilon_i}} \frac{z \frac{\partial}{\partial z} Q(\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, z)}{Q(\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, z)} dz$$

Ce qui définit une fonction C^∞ et même holomorphe de la boule ouverte $B(a, \rho) = \{\alpha \in C^n ;\ Max|\alpha_k - a_k| < \rho\}$ dans C.

On en déduit alors que l'application $\alpha \mapsto \zeta(\alpha)$ est C^∞ de $B(a, \rho)$ dans C^n.

Problème 29 : Transformation de Schur et localisation des zéros d'un polynôme complexe. Agrégation 1995, extrait

Ce problème utilise le théorème de Rouché du problème 28, Q2.

On appelle polynôme de « degré formel » d, un polynôme du type $Q(z) = \sum_{j=0}^{d} a_j z^j$, *où* a_d *peut être nul. Si* $d' < d$ *est le degré effectif de Q, on dit alors que* ∞ *est racine de Q avec la multiplicité* $d - d'$. *On pose* $\frac{1}{0} = \infty$ *et* $\frac{1}{\infty} = 0$.

Pour tout polynôme Q, on notera $Z(Q)$ *le nombre de zéros de Q de module strictement inférieur à 1.*

Soit d un entier positif et $Q(z) = \sum_{j=0}^{d} a_j z^j$ *un polynôme de degré formel d. On appelle polynôme d–réciproque de Q le polynôme* $Q^*(z) = \sum_{j=0}^{d} \bar{a}_j z^{d-j}$ *et on appelle transformée de Schur d'ordre d de Q le polynôme de degré formel d – 1 :* $T_d Q = \bar{a}_0 Q - a_d Q^*$. *On utilisera aussi les itérés de la transformation définis par récurrence par* $T_d^{k+1} Q = T_{d-k} T_d^k Q$, *où k est un entier positif.*

Dans ce problème, n est un entier strictement positif fixé et P un polynôme de degré n.

Q1. *Montrer que si* $w_1, w_2, \cdots, w_n$ *sont les zéros d'un polynôme de degré formel n, les zéros de* Q^* *sont les* $\dfrac{1}{\overline{w}_k}$, *les ordres de multiplicité étant conservés.*

Q2. *Montrer que* $\forall z,\ |z| = 1 \Rightarrow \left|P^*(z)\right| = \left|P(z)\right|$.

Q3. *Vérifier que pour tout k > 0,* $T_n^k P(0)$ *est réel. Montrer que si P n'a pas de racines dans le disque* $D = \{z;\ |z| \le 1\}$, *alors si* $1 \le k \le n$, $T_n^k P(0)$ *est strictement positif.*

On pourra montrer d'abord que $T_n P(0) > 0$ *puis montrer à l'aide du théorème de Rouché que* $T_n P$ *ne s'annule pas sur D.*

Q4. *Réciproquement, montrer que si pour tout* $k \in \{1, 2, \cdots, n\}$, $T_n^k P(0) > 0$, *P ne s'annule pas sur D. On pourra commencer par étudier* $T_n^n P$.

Q5. *Montrer que si* $T_n^n P(0) \ne 0$, *aucun des polynômes* $T_n^k P$, $0 \le k \le n$ *ne s'annule sur le cercle* Γ: $|z| = 1$.

Q6. *Montrer que :*

$$\begin{cases} T_n^k P(0) > 0 \Rightarrow Z\left(T_n^k P\right) = Z\left(T_n^{k-1} P\right) \\ T_n^k P(0) < 0 \Rightarrow Z\left(T_n^k P\right) = Z\left(\left(T_n^{k-1} P\right)^*\right) \end{cases}$$

Q7. *On note* k_j, *j = 1, 2, ⋯, m, les indices k pour lesquels* $T_n P(0) < 0$. *Montrer que si tous les* $T_n^k P(0)$ *sont non nuls :*

$$Z(P) = \sum_{j=1}^{m} (-1)^{j-1}\left(n + 1 - k_j\right)$$

Q8. *Soit* $P(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n$ *un polynôme unitaire de degré n à coefficients complexes. Montrer qu'il existe un unique* $R > 0$ *tel que* $R^n = |a_0| + |a_1| R + \cdots + |a_{n-1}| R^{n-1}$. *Montrer que tous les zéros de P sont contenus dans* $D(0, R)$.

Déterminer un majorant de R s'exprimant de façon simple en fonction des a_i.

Solution

Si $Q \ne 0$ est de degré formel n, alors :

$$Q(z) = \sum_{j=0}^{n} a_j z^j = \sum_{j=p}^{q} a_j z^j$$

où $0 \le p \le q \le n$ et $a_p \ne 0,\ a_q \ne 0$.

p est la « *valuation* » et q est le « *degré effectif* » de Q.

Avec ces notations, 0 est racine d'ordre p de Q (une valuation nulle signifie que 0 n'est pas racine) et ∞ est racine d'ordre $n - q$ de Q (un degré effectif égal au degré formel signifie que ∞ n'est pas racine).

Le polynôme n–réciproque de Q est :

$$Q^*(z)=\sum_{j=0}^{n}\overline{a}_j z^{n-j}=\overline{a}_q z^{n-q}+\cdots+\overline{a}_p z^{n-p}$$

C'est un polynôme de degré formel n, degré effectif $n-p$, de valuation $n-q$ et on a l'égalité :

$$\forall z\neq 0,\ Q^*(z)=z^n\overline{Q\left(\frac{1}{\overline{z}}\right)}$$

Q1. Si 0 est racine d'ordre $p\geq 0$ de Q, alors Q^* est de degré effectif $n-p$ et ∞ est racine d'ordre $n-(n-p)=p$ de Q.

De même si ∞ est racine d'ordre $n-q$ de Q, alors Q^* est de valuation $n-q$ et 0 est racine d'ordre $n-q$ de Q^*.

D'autre part, en notant $Q(z)=a_q z^p\prod_{i=1}^{m}(z-z_i)^{n_i}$ où les z_i sont non nuls, deux à deux distincts et $\sum_{i=1}^{m}n_i+p=q$, on a pour $z\neq 0$:

$$Q^*(z)=\overline{a}_q z^{n-p}\prod_{i=1}^{m}\left(\frac{1}{z}-\overline{z}_i\right)^{n_i}=\overline{a}_q z^{n-q}\prod_{i=1}^{m}\overline{z}_i\left(\frac{1}{\overline{z}_i}-z\right)^{n_i}$$

et $\dfrac{1}{\overline{z}_i}$ est racine d'ordre n_i de Q^*, pour tout $i=1,\cdots,m$.

Q2. La condition $|z|=1$ équivaut à $\dfrac{1}{\overline{z}}=z$. On en déduit alors que :

$$\left|P^*(z)\right|=\left|z^n\overline{P\left(\frac{1}{\overline{z}}\right)}\right|=\left|P(z)\right|$$

Il en résulte que tout zéro de module 1 de P est aussi zéro de P^* (cas particulier de Q1).

Q3. (i) Pour $k=0$, on pose $T_n^0P=P$.

Pour $k=1$, on a :

$$T_n^1P(z)=T_nP(z)=\overline{a}_0P(z)-a_nP^*(z)=\sum_{j=0}^{n-1}\left(\overline{a}_0a_j-a_n\overline{a}_{n-j}\right)z^j=\sum_{j=0}^{n-1}a_j^1z^j$$

en posant $T_n^1P=0$ si $n=0$.

De manière plus générale, on a pour tout $k\in\{0,1,\ldots,n\}$:

$$T_n^kP(z)=\sum_{j=0}^{n-k}a_j^kz^j$$

avec :

$$\begin{cases}a_j^0=a_j\ (0\leq j\leq n)\\ a_j^k=\overline{a}_0^{k-1}a_j^{k-1}-a_{n-(k-1)}^{k-1}\overline{a}_{n-(k-1)-j}^{k-1}\ (1\leq k\leq n,\ 0\leq j\leq n-k)\end{cases}$$

Pour $k>n$, on a $T_n^kP=0$.

En particulier, on a $T_n^kP(0)=\left|a_0^{k-1}\right|^2-\left|a_{n-(k-1)}^{k-1}\right|^2\in IR$, pour tout $k\in\{1,2,\cdots,n\}$.

(ii) Si P n'a pas de racines dans le disque fermé D de centre 0 et de rayon 1, alors $a_0\neq 0$.

Le polynôme P étant de degré effectif n, on a $a_n \neq 0$ et P admet n racines complexes de module strictement supérieur à 1. Le produit de ces racines étant égal à $(-1)^n \dfrac{a_0}{a_n}$, on en déduit que $\left|\dfrac{a_0}{a_n}\right| > 1$ et $T_n P(0) = |a_0|^2 - |a_n|^2 > 0$.

Remarque — Si le degré effectif de P est strictement inférieur à n, alors $a_n = 0$ et on a encore $T_n P(0) = |a_0|^2 > 0$.

D'autre part, on a :

$$|z| = 1 \Rightarrow \left|T_n P(z) - \bar{a}_0 P(z)\right| = \left|a_n P^*(z)\right| = |a_n|\left|P(z)\right| < \left|\bar{a}_0 P(z)\right|$$

On déduit alors du théorème de Rouché (problème 28, Q2) que $T_n P$ a le même nombre de zéros que P dans le disque D. C'est-à-dire que $T_n P$ ne s'annule pas sur D.

Par récurrence, on déduit alors que $T_n^k P(0) > 0$ et $T_n^k P$ ne s'annule pas sur D pour tout $k = 1, 2, \cdots, n$.

Q4. Avec les notations introduites en Q3 (i), les conditions $T_n^k P(0) > 0$ pour tout $k \in \{1, 2, \cdots, n\}$ sont équivalentes à $\left|a_0^k\right|^2 > \left|a_{n-k}^k\right|^2$ pour tout $k \in \{0, 2, \cdots, n-1\}$.

On en déduit alors que :

$$|z| = 1 \Rightarrow \left|T_n^{k+1} P(z) - \bar{a}_0^k T_n^k P(z)\right| = \left|a_{n-k}^k \left(T_n^k P\right)^*(z)\right| = \left|a_{n-k}^k\right|\left|T_n^k P(z)\right| < \left|\bar{a}_0^k T_n^k P(z)\right|$$

Et avec le théorème de Rouché, on conclut que $T_n^{k+1} P$ et $T_n^k P$ ont le même nombre de zéros dans D.

Par récurrence, on en déduit que les $T_n^k P$ ont le même nombre de zéros que $T_n^n P$ dans D, pour tout $k = 0, 1, \cdots, n$. Comme $T_n^n P(z) = T_n^n P(0) > 0$, on déduit que les $T_n^k P$ ne s'annule pas dans D, pour tout $k = 0, 1, \cdots, n$.

On a donc ainsi montré que :

$$(P \text{ n'a pas de zéros dans } D) \Leftrightarrow (\forall k \in \{1, 2, \cdots, n\},\ T_n^k P(0) > 0)$$

Q5. Si $z \in \Gamma = \{z \in C;\ |z| = 1\}$ est racine de $T_n^k P$ pour un indice $k \in \{0, 1, \cdots, n-1\}$, alors z est aussi racine de $\left(T_n^k P\right)^*$ (d'après Q2) et avec l'égalité $T_n^{k+1} P(z) = \bar{a}_0^k T_n^k P(z) - a_{n-k}^k \left(T_n^k P\right)^*(z)$, on déduit que z est aussi racine de $T_n^{k+1} P$. Par récurrence on aboutit alors a $T_n^n P(z) = 0$, ce qui contredit l'hypothèse $T_n^n P(z) = T_n^n P(0) \neq 0$.

On a donc ainsi montré que :

$$\left(T_n^n P(0) \neq 0\right) \Rightarrow \left(\forall k \in \{0, 1, \cdots, n\},\ \forall z \in \Gamma,\ T_n^k P(z) \neq 0\right)$$

Q6. (i) La condition $T_n^k P(0) > 0$ équivaut à $\left|a_0^{k-1}\right|^2 > \left|a_{n-(k-1)}^{k-1}\right|^2$, ce qui entraîne :

$$|z| = 1 \Rightarrow \left|T_n^k P(z) - \bar{a}_0^{k-1} T_n^{k-1} P(z)\right| = \left|a_{n-(k-1)}^{k-1} \left(T_n^{k-1} P\right)^*(z)\right| < \left|\bar{a}_0^{k-1} T_n^{k-1} P(z)\right|$$

Et avec le théorème de Rouché, on conclut que $T_n^k P$ et $T_n^{k-1} P$ ont le même nombre de zéros dans D, c'est-à-dire que $Z\left(T_n^k P\right) = Z\left(T_n^{k-1} P\right)$.

(ii) La condition $T_n^k P(0) < 0$ équivaut à $\left|a_0^{k-1}\right|^2 < \left|a_{n-(k-1)}^{k-1}\right|^2$, ce qui entraîne :

$$|z| = 1 \Rightarrow \left| T_n^k P(z) - a_{n-(k-1)}^{k-1} \left(T_n^{k-1} P \right)^* (z) \right| = \left| \overline{a}_0^{k-1} T_n^{k-1} P(z) \right| < \left| a_{n-(k-1)}^{k-1} \left(T_n^{k-1} P \right)^* (z) \right|$$

Et avec le théorème de Rouché, on conclut que $T_n^k P$ et $\left(T_n^{k-1} P\right)^*$ ont le même nombre de zéros dans D, c'est-à-dire que $Z\left(T_n^k P\right) = Z\left(\left(T_n^{k-1} P\right)^*\right)$.

Q7. Avec Q1, on déduit que $Z\left(\left(T_n^{k-1} P\right)^*\right)$ est égal au nombre de zéros de $T_n^{k-1} P$ en dehors du disque unité fermé. C'est-à-dire que :

$$Z\left(\left(T_n^{k-1} P\right)^*\right) = \deg\left(T_n^{k-1} P\right) - Z\left(T_n^{k-1} P\right) = n - (k-1) - Z\left(T_n^{k-1} P\right)$$

On déduit alors de Q6 que :

$$\forall k \in \{1, 2, \cdots, n\},\ Z\left(T_n^{k-1} P\right) = \begin{cases} Z\left(T_n^k P\right) \text{ si } T_n^k P(0) > 0 \\ n + 1 - k - Z\left(T_n^k P\right) \text{ si } T_n^k P(0) < 0 \end{cases}$$

On range les k_j dans l'ordre croissant, c'est-à-dire qu'on suppose que $1 \leq k_1 < k_2 < \cdots < k_m$.

Par définition de k_1, on a :

$$Z(P) = Z\left(T_n^0 P\right) = \cdots = Z\left(T_n^{k_1 - 1} P\right) = n + 1 - k_1 - Z\left(T_n^{k_1} P\right)$$

avec $Z\left(T_n^{k_1} P\right) = n + 1 - k_2 - Z\left(T_n^{k_2} P\right)$. Ce qui donne par récurrence :

$$Z(P) = \sum_{j=1}^{m} (-1)^{j-1} \left(n + 1 - k_j\right)$$

puisque $Z\left(T_n^n P\right) = 0$ ($T_n^n P(z) = T_n^n P(0) \neq 0$).

Remarque — Avec :

$$\sum_{j=1}^{m} (-1)^{j-1} (n+1) = \begin{cases} 0 \text{ si } m \text{ est pair} \\ n + 1 \text{ si } m \text{ est impair} \end{cases}$$

on déduit que :

$$Z(P) = \begin{cases} \displaystyle\sum_{j=1}^{m} (-1)^j k_j \text{ si } m \text{ est pair} \\ n + 1 + \displaystyle\sum_{j=1}^{m} (-1)^j k_j \text{ si } m \text{ est impair} \end{cases}$$

Q8. (i) Si $a_0 = 0$, alors $P(z) = z^p Q(z)$, avec $p \geq 1$ et $Q(0) \neq 0$. Et on raisonne alors avec Q. On suppose donc que $a_0 \neq 0$.

Soit :

$$\begin{array}{llll} f : & IR^+ & \rightarrow & IR^+ \\ & t & \mapsto & |a_0| t^n + |a_1| t^{n-1} + \cdots + |a_{n-1}| t \end{array}$$

La fonction f est continue, strictement croissante avec $f(0)=0$ et $\underset{t\to+\infty}{Lim} f(t)=+\infty$. Elle réalise donc une bijection de $I\!R^+$ dans $I\!R^+$. En particulier, il existe un unique réel $t>0$ tel que $f(t)=1$. Ce qui équivaut à dire que $R=\frac{1}{t}>0$ est l'unique solution de $R^n=|a_0|+|a_1|R+\cdots+|a_{n-1}|R^{n-1}$.

(ii) Pour tout $\varepsilon>0$ et tout $z\in C$ tel que $|z|=R+\varepsilon$, on a :

$$\left|P(z)-z^n\right|\leq|a_0|+|a_1|(R+\varepsilon)+\cdots+|a_{n-1}|(R+\varepsilon)^{n-1}$$

$$<\left(|a_0|+|a_1|R+\cdots+|a_{n-1}|R^{n-1}\right)\left(\frac{R+\varepsilon}{R}\right)^n=(R+\varepsilon)^n=|z^n|$$

Avec le théorème de Rouché, on déduit alors que P a le même nombre de zéros que z^n dans le disque ouvert $|z|<R+\varepsilon$. Ce résultat étant valable quel que soit $\varepsilon>0$, on en déduit que P a tous ses zéros dans le disque fermé de centre 0 et de rayon R.

(iii) Si $R\geq 1$, on peut écrire que :

$$R^n=|a_0|+|a_1|R+\cdots+|a_{n-1}|R^{n-1}\leq\left(|a_0|+|a_1|+\cdots+|a_{n-1}|\right)R^{n-1}$$

et $R\leq|a_0|+|a_1|+\cdots+|a_{n-1}|$. On déduit donc la majoration :

$$R\leq Max\left\{1,|a_0|+|a_1|+\cdots+|a_{n-1}|\right\}$$

On peut aussi montrer que :

$$R\leq Max\left\{|a_0|,1+|a_1|,\cdots,1+|a_{n-1}|\right\}$$

(problème 27, Q4).

Problème 30 : Méthode de Newton–Kantorovich

Dans la méthode de Newton–Raphson classique, le calcul à chaque étape de la matrice jacobienne $df\left(x^{(k)}\right)$ va ralentir les calculs. L'idée de Kantorovitch consiste à utiliser la même matrice jacobienne $df\left(x^{(0)}\right)$ à chaque étape. La méthode obtenue est convergente pour $x^{(0)}$ convenablement choisi.

$I\!R^n$ est muni d'une norme quelconque qui induit une norme matricielle (problème 1).

On se donne un ouvert non vide $\Omega\subset I\!R^n$, une application différentiable $f:\Omega\to I\!R^n$, un convexe fermé $C\subset\Omega$ et un point $x^{(0)}$ intérieur à C.

On suppose que la différentielle df vérifie dans le convexe C une condition de Lipschitz de paramètre $\lambda>0$ et que $df\left(x^{(0)}\right)$ est inversible.

On note :

$$A_0=df\left(x^{(0)}\right)^{-1},\ \alpha=\|A_0\|,\ \beta=\left\|A_0 f\left(x^{(0)}\right)\right\|\ \text{et}\ \gamma=\alpha\cdot\beta\cdot\lambda$$

On suppose que $\beta\neq 0$ et que $\gamma<\frac{1}{4}$.

On désigne par $t_0 < t_1$ *les racines de l'équation* $\gamma t^2 - t + 1 = 0$ *et on suppose que la boule fermée B_0 de centre* $x^{(0)}$ *et de rayon* $\beta \cdot t_0$ *est contenue dans C.*

Enfin on désigne par g l'application définie sur Ω *par :*

$$\forall x \in \Omega,\ g(x) = x - A_0 f(x)$$

Q1. *Montrer que :*

$$\forall x, y \in C,\ \left\| f(x) - f(y) - df(y)(x-y) \right\| \le \frac{\lambda}{2} \|x-y\|^2$$

Q2. *Montrer que* $g(B_0) \subset B_0$.

Q3. *Montrer que :*

$$\forall x, y \in B_0,\ \|g(x) - g(y)\| \le \frac{1-\sqrt{1-4\gamma}}{2} \|x-y\|$$

Q4. *On définit la suite* $\left(x^{(k)}\right)_{k \in IN}$ *de points de B_0 par :*

$$\forall k \ge 0,\ x^{(k+1)} = x^{(k)} - A_0 f\left(x^{(k)}\right)$$

(a) Montrer que la suite $\left(x^{(k)}\right)_{k \in IN}$ *converge vers l'unique solution* $x \in B_0$ *de l'équation* $f(x) = 0$.

(b) Montrer qu'on a la majoration de l'erreur :

$$\forall k \ge 0,\ \left\| x^{(k)} - x \right\| \le \frac{\beta t_0}{1-\delta} \delta^k$$

où on a posé $\delta = \dfrac{1-\sqrt{1-4\gamma}}{2}$.

Q5. *On suppose que le convexe fermé C contient la boule fermée B_1 de centre* $x^{(0)}$ *et de rayon t_1.*

(a) Montrer que :

$$\beta t_0 \le \left\| x - x^{(0)} \right\| \le \beta t_1 \Rightarrow \left\| g(x) - x^{(0)} \right\| < \left\| x - x^{(0)} \right\|$$

(b) En déduire que l'équation $f(x) = 0$ *admet une unique solution dans B_1.*

Solution

Q1. Soit $\varphi : [0,1] \to IR^n$ définie par $\varphi(t) = f(y + t(x-y))$. Cette fonction est différentiable avec $\varphi'(t) = df(y + t(x-y))(x-y)$ et l'hypothèse *df* lipschitzienne sur *C* entraîne que :

$$\forall t \in [0,1], \left| \varphi'(t) - \varphi'(0) \right| \le \lambda \cdot t \|x-y\|^2$$

Et en écrivant que :

$$f(x) - f(y) - df(y)(x-y) = \varphi(1) - \varphi(0) - \varphi'(0) = \int_0^1 \left(\varphi'(t) - \varphi'(0)\right) dt$$

on déduit l'inégalité souhaitée.

Q2. Pour $x \in B_0$, on a :

$$g(x) - x^{(0)} = x - x^{(0)} - A_0 f(x) = A_0 \left\{ df\left(x^{(0)}\right)\left(x - x^{(0)}\right) - f(x) + f\left(x^{(0)}\right) \right\} - A_0 f\left(x^{(0)}\right)$$

Avec $B_0 \subset C$ et Q1, on déduit que :

$$\left\|g(x)-x^{(0)}\right\| \leq \alpha\frac{\lambda}{2}\left\|x-x^{(0)}\right\|^2+\beta \leq \alpha\frac{\lambda}{2}(\beta t_0)^2+\beta$$

soit, avec la définition de t_0 :

$$\left\|g(x)-x^{(0)}\right\| \leq \beta\left(\frac{\gamma}{2}t_0^{\ 2}+1\right) < \beta\left(\gamma t_0^{\ 2}+1\right)=\beta t_0$$

C'est-à-dire que $g(x)\in B_0$. On a donc $g(B_0)\subset B_0$.

Q3. La fonction g est différentiable avec :

$$\forall x\in\Omega,\ dg(x)=I_d-A_0 df(x)=A_0\left(df\left(x^{(0)}\right)-df(x)\right)$$

Et avec la condition de Lipschitz sur df, on déduit que :

$$\forall x\in C,\ \left\|dg(x)\right\| \leq \alpha\lambda\left\|x-x^{(0)}\right\|$$

En particulier, pour $x\in B_0$, on a $\left\|dg(x)\right\|\leq \alpha\lambda\beta t_0=\gamma t_0$ avec $t_0=\dfrac{1-\sqrt{1-4\gamma}}{2\gamma}$.

Avec le théorème des accroissements finis, on déduit alors que :

$$\forall x,y\in B_0,\ \left\|g(x)-g(y)\right\| \leq \frac{1-\sqrt{1-4\gamma}}{2}\left\|x-y\right\|$$

En remarquant que $\delta=\dfrac{1-\sqrt{1-4\gamma}}{2}<\dfrac{1}{2}$, on déduit que g est une contraction sur B_0.

Q4. La suite $\left(x^{(k)}\right)_{k\in IN}$ est bien définie dans B_0, puisque $x^{(k+1)}=g\left(x^{(k)}\right)$ et $g(B_0)\subset B_0$.

(a) Avec Q3, on déduit que pour tout $k\geq 1$, on a :

$$\left\|x^{(k+1)}-x^{(k)}\right\|=\left\|g\left(x^{(k)}\right)-g\left(x^{(k-1)}\right)\right\| \leq \delta\left\|x^{(k)}-x^{(k-1)}\right\|$$

et par récurrence :

$$\forall k\geq 0,\ \left\|x^{(k+1)}-x^{(k)}\right\| \leq \delta^k\left\|x^{(1)}-x^{(0)}\right\| \leq \beta t_0\delta^k$$

Pour $q>p$ dans IN, on a alors :

$$\left\|x^{(q)}-x^{(p)}\right\| \leq \left\|x^{(q)}-x^{(q-1)}\right\|+\cdots+\left\|x^{(p+1)}-x^{(p)}\right\| \leq \beta t_0\left\{\delta^{q-1}+\cdots+\delta^p\right\}=\beta t_0\delta^p\frac{1-\delta^{q-p}}{1-\delta}$$

C'est-à-dire :

$$\forall q>p,\ \left\|x^{(q)}-x^{(p)}\right\| \leq \frac{\beta t_0}{1-\delta}\delta^p \quad (1)$$

avec $\lim\limits_{p\to+\infty}\delta^p=0$.

On a donc ainsi montré que la suite $\left(x^{(k)}\right)_{k\in IN}$ est de Cauchy dans IR^n, elle est donc convergente vers $x\in B_0$.

Des égalités $x^{(k+1)}=g\left(x^{(k)}\right)$, on déduit que $x=g(x)$, ce qui équivaut à $f(x)=0$.

Si $x'\in B_0-\{x\}$ est une autre solution de $f(x)=0$, alors en écrivant que :

$$\left\|x-x'\right\|=\left\|g(x)-g(x')\right\| \leq \delta\left\|x-x'\right\| < \left\|x-x'\right\|$$

on aboutit à une contradiction.

Donc l'équation $f(x)=0$ admet une unique solution dans B_0.

(b) En faisant tendre q vers l'infini dans les inégalités (1) du (a), on déduit que :

$$\forall p>0,\ \left\|x-x^{(p)}\right\|\leq\frac{\beta t_0}{1-\delta}\delta^p$$

Remarque — La convergence est moins rapide qu'avec la méthode de Newton–Raphson classique.

Q5. (a) En Q2, on a vu que pour tout $x\in C$, on a :

$$\left\|g(x)-x^{(0)}\right\|\leq\alpha\frac{\lambda}{2}\left\|x-x^{(0)}\right\|^2+\beta<\alpha\lambda\left\|x-x^{(0)}\right\|^2+\beta$$

Si $\beta t_0\leq\left\|x-x^{(0)}\right\|\leq\beta t_1$, alors $t=\dfrac{\left\|x-x^{(0)}\right\|}{\beta}\in[t_0,t_1]$ et $\gamma t^2-t+1\leq 0$, c'est-à-dire :

$$\gamma\frac{\left\|x-x^{(0)}\right\|^2}{\beta^2}-\frac{\left\|x-x^{(0)}\right\|}{\beta}+1\leq 0$$

ou encore $\alpha\lambda\left\|x-x^{(0)}\right\|^2+\beta\leq\left\|x-x^{(0)}\right\|$.

On a donc, en définitive :

$$\beta t_0\leq\left\|x-x^{(0)}\right\|\leq\beta t_1\Rightarrow\left\|g(x)-x^{(0)}\right\|<\left\|x-x^{(0)}\right\|$$

(b) On a déjà vu qu'il y a une seule racine de $f(x)=0$ dans B_0.

Si $x'\in B_1$ est une autre racine avec $\beta t_0\leq\left\|x-x^{(0)}\right\|\leq\beta t_1$, alors :

$$\left\|x'-x^{(0)}\right\|=\left\|g(x')-x^{(0)}\right\|<\left\|x'-x^{(0)}\right\|$$

ce qui est impossible. On a donc une seule racine de $f(x)=0$ dans B_1.

Problème 31 : Calcul de l'inverse d'une matrice. Méthode de Schulz

On désigne par $M_n(C)$ l'espace vectoriel des matrices complexes d'ordre $n\geq 1$. Il est muni d'une norme matricielle induite par une norme vectorielle (problème 1). Pour toute matrice $M\in M_n(C)$, on désigne par $\rho(M)$ le rayon spectral de M (problème 2).

Pour toute matrice inversible A, on désigne par $(A_k)_{k\in IN}$ la suite de matrices définie par la récurrence :

$$\begin{cases}A_0\in M_n(C)\\ A_{k+1}=A_k(I_d+E_k)\quad(k\geq 0)\end{cases}$$

où I_d désigne la matrice identité et $E_k=I_d-A\cdot A_k$.

Q1. *Montrer que :*

$$\forall k\geq 0,\ E_k=E_0^{2^k}$$

Q2. *(a) Montrer que la suite $(A_k)_{k\in IN}$ converge vers l'inverse de A si et seulement si $\rho(E_0)<1$.*

(b) Donner une majoration de l'erreur $\left\|A_k - A^{-1}\right\|$.

Q3. *Montrer que le choix de* $A_0 = \dfrac{1}{\|A\|\left\|{}^t\overline{A}\right\|}\,{}^t\overline{A}$ *assure la convergence de la suite* $(A_k)_{k\in IN}$.

Q4. *On écrit la matrice A sous la forme* $A = I_d - B$ *et on suppose que* $\rho(B) < 1$.

En prenant $A_0 = I_d$*, montrer que :*

$$\forall k \geq 0,\ A_k = \sum_{j=0}^{2^k-1} B^j$$

et que la suite $(A_k)_{k\in IN}$ *converge vers* A^{-1}.

Solution

Q1. On a :

$$E_{k+1} = I_d - A\cdot A_k\left(I_d + E_k\right) = E_k - A\cdot A_k E_k = E_k^{\,2}$$

et par récurrence $E_k = E_0^{\,2^k}$ pour tout $\forall k \geq 0$.

Q2. (a) Avec les résultats du problème 2, Q6, on a les équivalences :

$$\underset{k\to+\infty}{Lim} A_k = A^{-1} \Leftrightarrow \underset{k\to+\infty}{Lim} AA_k = I_d \Leftrightarrow \underset{k\to+\infty}{Lim} E_k = 0 \Leftrightarrow \rho(E_0) < 1$$

(b) On a $A^{-1} - A_k = A^{-1}\left(I_d - A\cdot A_k\right) = A^{-1}E_k$ et :

$$\left\|A^{-1} - A_k\right\| \leq \left\|A^{-1}\right\|\left\|E_k\right\| \leq \left\|A^{-1}\right\|\left\|E_0\right\|^{2^k}$$

Q3. La matrice $A\cdot A_0 = \dfrac{1}{\|A\|\left\|{}^t\overline{A}\right\|}\,A\cdot{}^t\overline{A}$ est symétrique définie positive puisque A est inversible, avec $\rho(A\cdot A_0) \leq 1$ (problème 2, Q1).

On en déduit donc que toutes ses valeurs propres sont réelles contenues dans l'intervalle $]0,1]$. Il en résulte que celles de $E_0 = I_d - A\cdot A_0$ sont réelles contenues dans l'intervalle $[0,1[$. Ce qui entraîne $\rho(E_0) < 1$ et la convergence de la suite $(A_k)_{k\in IN}$ vers A^{-1}.

Q4. Pour $k = 1$, on a $A_1 = I + B$. En supposant que pour $k \geq 1$, on a $A_k = \sum_{j=0}^{2^k-1} B^j$, on déduit que $A_{k+1} = \left(\sum_{j=0}^{2^k-1} B^j\right)\left(I + B^{2^k}\right) = \sum_{j=0}^{2^{k+1}-1} B^j$, ce qui démontre le résultat par récurrence.

Problème 32 : Racine carrée d'une matrice complexe

IR^- désigne l'ensemble des réels négatifs ou nuls.

C^+ désigne l'ensemble des nombres complexes de partie réelle strictement positive.

L'espace vectoriel $M_n(C)$ des matrices complexes d'ordre $n \geq 1$ est muni d'une norme non nécessairement induite par une norme vectorielle.

Pour toute matrice $M \in M_n(C)$, on désigne par $Sp(M)$ l'ensemble des valeurs propres de M.

On dit que $M \in M_n(C)$ admet une racine carrée s'il existe $X \in M_n(C)$ telle que $X^2 = M$.

Dans tout le problème, A désigne une matrice complexe d'ordre n telle que :

$$Sp(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_p\} \subset C - IR^- \quad (\lambda_i \neq \lambda_j \text{ si } i \neq j)$$

A cette matrice, on associe la fonction $f : M_n(C) \to M_n(C)$ définie par :

$$\forall X \in M_n(C), f(X) = X^2 - A$$

Q1. *Montrer que pour tout $a \in C - IR^-$, il existe un unique complexe $z \in C^+$ tel que $z^2 = a$. On note $z = \sqrt{a}$.*

Q2. *Soit $a \in C - IR^-$ et g: $C^* \to C$ définie par :*

$$\forall z \in C^*,\ g(z) = \frac{1}{2}\left(z + \frac{a}{z}\right)$$

On prolonge g à $C \cup \{\infty\}$ en posant $f(0) = f(\infty) = \infty$ et on définit les suites $(x_k)_{k \in IN}$ et $(y_k)_{k \in IN}$ par :

$$\begin{cases} x_0 \in C \\ \forall k \geq 0,\ x_{k+1} = g(x_k),\ y_k = \dfrac{x_k - \sqrt{a}}{x_k + \sqrt{a}} \end{cases}$$

(a) Exprimer y_k en fonction de y_0, pour tout $k \geq 0$.

(b) Montrer l'équivalence :

$$\left(\underset{k \to +\infty}{Lim}\, x_k = \sqrt{a}\right) \Leftrightarrow \left(x_0 \in IP^+\right)$$

où IP^+ est un demi-plan du plan complexe stable par f à préciser.

(c) Montrer que pour tout $x_0 > 0$ dans IR, la suite $(x_k)_{k \in IN}$ est bien définie dans C et converge vers $\sqrt{a}$.

Q3. *(a) Donner un exemple de matrice qui n'admet pas de racine carrée.*

(b) Donner un exemple de matrice qui admet une infinité de racines carrées.

Q4. *Soit $T \in M_n(C)$ triangulaire supérieure telle que $Sp(T) \subset C - IR^-$.*

Montrer qu'il existe une unique matrice triangulaire supérieure $X \in M_n(C)$ telle que $Sp(X) = \{\sqrt{\lambda};\ \lambda \in Sp(T)\} \subset C^+$ et $X^2 = T$. On note $X = \sqrt{T}$.

Q5. *Comme indiqué en préambule, $A \in M_n(C)$ est telle que $Sp(A) \subset C - IR^-$.*

Montrer qu'il existe une unique matrice $X \in M_n(C)$ telle que $Sp(X) = \left\{\sqrt{\lambda};\ \lambda \in Sp(A)\right\} \subset C^+$ et $X^2 = A$. On note $X = \sqrt{A}$.

Q6. *Dans cette question, on suppose de plus que A est diagonalisable*

(a) Montrer que :

$$\sqrt{A} = \sum_{i=1}^{p} \frac{\sqrt{\lambda_i}}{L_i(\lambda_i)} L_i(A)$$

où $L_i(t) = \prod_{\substack{j=1 \\ j \neq i}}^{p} (t - \lambda_j)$, pour tout i = 1, 2, ⋯, p.

(b) Pour n = p = 2, donner une expression simple de $\sqrt{A}$, moyennant certaines hypothèses sur les valeurs propres de A.

Q7. *Montrer que la fonction f définie en préambule est indéfiniment dérivable et calculer sa différentielle $df(X)$ en tout point $X \in M_n(C)$.*

Q8. *Montrer que pour tout $M \in M_n(C)$, on a :*

$$Sp(df(M)) = \{\lambda + \mu;\ \lambda, \mu \in Sp(M)\}$$

Q9. *Montrer l'équivalence :*

($df(X)$ est inversible) ⇔ (X n'admet pas deux valeurs propres opposées)

Q10. *On note $\alpha(A) = \dfrac{1}{\left\| df\left(\sqrt{A}\right)^{-1} \right\|}$ où $\left\| df\left(\sqrt{A}\right)^{-1} \right\|$ est la norme induite par la norme choisie sur $M_n(C)$.*

Pour tout $X \in M_n(C)$, on note $e(X) = X - \sqrt{A}$ et si de plus $df(X)$ est inversible, on note $g(X) = X - df(X)^{-1} f(X)$.

(a) Pour tout $X \in M_n(C)$ telle que $df(X)$ soit inversible, montrer que :

(i) $\left(g(X) - X\right)^2 = g(X)^2 - A$.

(ii) $\left(g(X) - X\right)^2 = e(g(X))^2 + e(X)^2 - e(g(X))e(X) - e(X)e(g(X))$.

(iii) $\alpha(A)\left\|e(g(X))\right\| \le \left\|df\left(\sqrt{A}\right)e(g(X))\right\| \le \left\|e(X)\right\|^2 + 2\left\|e(X)\right\|\left\|e(g(X))\right\|$.

(b) Montrer que si X_0 est dans la boule ouverte de centre $\sqrt{A}$ et de rayon $\dfrac{\alpha(A)}{3}$, alors on peut définir la suite $(X_k)_{k \in IN}$ de l'algorithme de Newton par $X_{k+1} = g(X_k)$, pour tout $k \ge 0$, et cette suite converge vers $\sqrt{A}$.

Q11. *Dans cette question, on suppose de plus que A est diagonale.*

Montrer que si dans l'algorithme de Newton associé à la résolution de $f(X) = 0$ on prend $X_0 = I_d$, alors la suite $(X_k)_{k \in IN}$ définie par $X_{k+1} = g(X_k)$ est aussi définie par :

$$\forall k \ge 0,\ X_{k+1} = \frac{1}{2}\left(X_k + A \cdot X_k^{-1}\right)$$

et qu'elle converge vers $\sqrt{A}$.

Q12. *Montrer que le résultat de Q11 est encore vrai pour A diagonalisable.*

Solution

Q1. $a \in C - IR^-$ s'écrit de manière unique $a = \rho e^{i\theta}$, avec $\rho > 0$ et $\theta \in]-\pi, \pi[$. En posant $z = \sqrt{\rho} e^{i\frac{\theta}{2}}$, on a $z \in C^+$ ($\mathrm{Re}(z) = \rho \mathrm{Cos}\left(\frac{\theta}{2}\right) > 0$) et $z^2 = a$.

Si $y \in C^+$ vérifie aussi $y^2 = a$, on a alors $y^2 - z^2 = 0$ et $y = \pm z$. La solution $y = -z$ donne $\mathrm{Re}(y) < 0$, ce qui est exclut, donc $y = z$.

On a donc ainsi montré que a admet une unique racine carrée dans C^+.

Remarque — L'application $a \mapsto \sqrt{a}$ définie sur le plan coupé $C - IR^-$ est la « *détermination principale* » de la racine carrée complexe. Elle définit une fonction holomorphe.

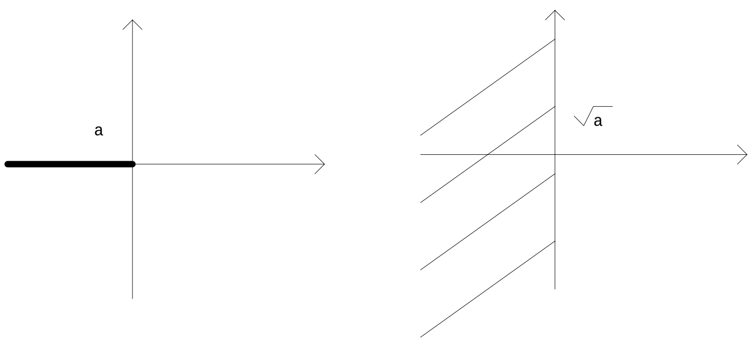

Figure 4.1

Q2. (a) Si $x_k \neq 0$ et $x_k \neq \infty$, on a alors :

$$y_{k+1} = \frac{x_k^{\ 2} - 2\sqrt{a}x_k + a}{x_k^{\ 2} + 2\sqrt{a}x_k + a} = \frac{\left(x_k - \sqrt{a}\right)^2}{\left(x_k + \sqrt{a}\right)^2} = y_k^{\ 2}$$

Si $x_k = 0$ ou $x_k = \infty$, alors $y_k = y_{k+1} = 1$.

Donc, dans tous les cas, on a :

$$\forall k \geq 0,\ y_{k+1} = y_k^{\ 2}$$

et par récurrence :

$$\forall k \geq 0,\ y_k = y_0^{\ 2^k}$$

(b) Si l'un des x_k est nul alors la suite diverge.

Si $\lim\limits_{k\to+\infty} x_k = \sqrt{a}$, alors $\lim\limits_{k\to+\infty} y_k = 0$, ce qui équivaut à $|y_0| < 1$, soit $\left|x_0 - \sqrt{a}\right|^2 < \left|x_0 + \sqrt{a}\right|^2$, c'est-à-dire que x_0 est dans le demi-plan :

$$IP^+ = \left\{z \in C;\ R\acute{e}\left(\sqrt{a}\overline{z}\right) > 0\right\} = \left\{z = x + iy \in C;\ \alpha x - \beta y > 0\right\}$$

où on a noté $\sqrt{a} = \alpha + i\beta$ avec $\alpha > 0$ (figure 4.2).

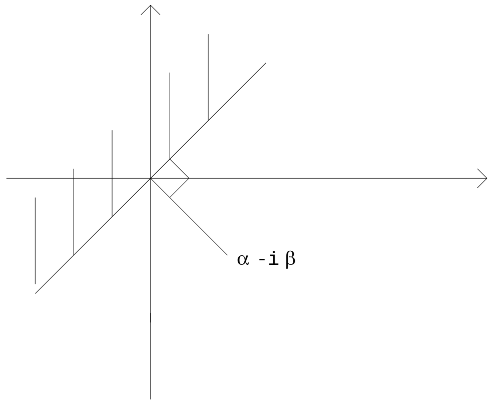

Figure 4.2

Pour $z \in IP^+$, on a $z \neq 0$ et $\dfrac{\left|z-\sqrt{a}\right|}{\left|z+\sqrt{a}\right|} < 1$. Puis avec $\dfrac{\left|f(z)-\sqrt{a}\right|}{\left|f(z)+\sqrt{a}\right|} = \left|\dfrac{z-\sqrt{a}}{z+\sqrt{a}}\right|^2 < 1$, on déduit que $f(z) \in IP^+$. C'est-à-dire que IP^+ est stable par *f*.

Réciproquement si $x_0 \in IP^+$, alors la suite $(x_k)_{k\in IN}$ est bien définie puisque $f(IP^+) \subset IP^+$. De plus on a $|y_0| < 1$, ce qui entraîne $\underset{k\to+\infty}{Lim}\, y_k = 0$, soit $\underset{k\to+\infty}{Lim}\, x_k = \sqrt{a}$.

(c) Résulte de $IR^+ \subset IP^+$.

Q3. (a) Soit $A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$. C'est une matrice nilpotente d'ordre *n* ($A^n = 0$ et $A^p \neq 0$ pour $p < n$).

Si $A = X^2$, alors *X* est aussi nilpotente. Soit $p \in \{2, \cdots, n\}$ son ordre. Si $p = 2k$, alors $A^k = X^p = 0$ avec $k < n$, ce qui n'est pas possible. Donc $p = 2k+1$ et $A^{k+1} = X^{p+1} = 0$ avec $k+1 < n$ si $n > 1$, ce qui est encore impossible.

Donc *A* n'admet pas de racine carrée.

(b) Si *X* est une matrice de symétrie, alors $X^2 = I_d$. C'est-à-dire que I_d admet une infinité de racines carrées.

Q4. On procède par récurrence sur $n \geq 1$.

Pour $n = 1$, le résultat est montré en Q1.

Supposons le résultat vrai pour $n-1 \geq 1$.

On décompose *T* par blocs, soit $T = \begin{pmatrix} T' & T_{n-1} \\ 0 & t_{nn} \end{pmatrix}$, où $T' \in M_{n-1}(C)$ est triangulaire supérieure avec $Sp(X') \subset C - IR^-$, $t_{nn} \in C - IR^-$ et $T_{n-1} \in C^{n-1}$.

Il s'agit alors de trouver $X = \begin{pmatrix} X' & X_{n-1} \\ 0 & x_{nn} \end{pmatrix} \in M_n(C)$ telle que $X^2 = T$, c'est-à-dire :

$$\begin{cases} X'^2 = T' \\ x_{nn}{}^2 = t_{nn} \\ (X' + x_{nn} I_{n-1}) X_{n-1} = T_{n-1} \end{cases}$$

Les deux premières équations admettent pour unique solution $X' = \sqrt{T'}$ et $x_{nn} = \sqrt{t_{nn}} \in C^+$ avec $Sp(X') = \left\{\sqrt{\lambda};\ \lambda \in Sp(T')\right\} \subset C^+$.

La matrice $X' + x_{nn} I_{n-1}$ est triangulaire supérieure avec pour termes diagonaux les $\sqrt{t_{ii}} + \sqrt{t_{nn}} \in C^+ \subset C - \{0\}$, elle est donc inversible et la troisième équation a pour unique solution $X_{n-1} = (X' + x_{nn} I_{n-1})^{-1} T_{n-1}$.

Q5. On peut écrire $A = PTP^{-1}$, où T est triangulaire supérieure et P inversible. Comme $Sp(T) = Sp(A)$, la matrice T vérifie les hypothèses de Q4 et admet donc une unique racine carrée $Y = \sqrt{T}$ triangulaire supérieure telle que $Sp(Y) = \left\{\sqrt{\lambda};\ \lambda \in Sp(A)\right\} \subset C^+$. En posant $X = PYP^{-1}$, on a $X^2 = A$ avec $Sp(X) = Sp(Y)$.

Supposons que $X^2 = A = Y^2$, les valeurs propres de X et Y étant dans C^+. On désigne par P_X et P_Y les polynômes minimaux de X et Y respectivement. On a alors $P_X(t) P_Y(t) = \prod_{j=1}^{m} (t - \mu_j)^{m_j}$ avec les μ_j deux à deux distincts tels que les $\mu_j{}^2$ soient aussi deux à deux distincts et non nuls.

On peut alors trouver un polynôme P tel que $P(t^2) - t$ s'annule en chaque μ_j ainsi que ses $m_j - 1$ premières dérivées (polynôme d'interpolation d'Hermite, problème 33). Le polynôme $P(t^2) - t$ est donc un multiple de $P_X P_Y$ et il annule X et Y. On a donc :

$$X = P(X^2) = P(A) = P(Y^2) = Y$$

ce qui prouve l'unicité de la racine carrée de A.

Q6. En reprenant le raisonnement de Q5, on voit qu'on peut écrire $A = PDP^{-1}$ et $\sqrt{A} = P\sqrt{D}P^{-1}$, avec P inversible, D et $\sqrt{D}$ diagonales et $Sp(\sqrt{A}) = \left\{\sqrt{\lambda};\ \lambda \in Sp(A)\right\}$. C'est-à-dire que A et $\sqrt{A}$ se diagonalisent dans une même base, ou encore que :

$$C^n = \bigoplus_{k=1}^{p} Ker(A - \lambda_k I_d), \text{ avec } E_k = Ker(A - \lambda_k I_d) = Ker\left(A - \sqrt{\lambda_k} I_d\right)$$

Pour tout $x \in E_k$, on a alors :

$$\sqrt{A}x = \lambda_k x$$

$$L_i(A)x = \begin{cases} \prod_{\substack{j=1 \\ j\neq i}}^{p} \left(A - \lambda_j I_d\right)x = 0 \text{ si } i \neq k \\ \prod_{\substack{j=1 \\ j\neq k}}^{p} \left(\lambda_k - \lambda_j\right)x = L_k(\lambda_k)x \text{ si } i = k \end{cases}$$

d'où l'égalité cherchée.
(b) Pour $p = 2$ et $n \geq 2$, on a :

$$\sqrt{A} = \frac{1}{\sqrt{\lambda_1} + \sqrt{\lambda_2}}\left(A + \sqrt{\lambda_1}\sqrt{\lambda_2} I_d\right)$$

Dans le cas particulier où $n = 2$, on a $\lambda_1\lambda_2 = Dét(A)$ et $\lambda_1 + \lambda_2 = Trace(A)$. Si $\lambda_1\lambda_2 \in C - IR^-$, on peut alors écrire $\left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)^2 = Trace(A) + 2\sqrt{Dét(A)}$ et en supposant que cette dernière quantité est dans $C - IR^-$, on a :

$$\sqrt{A} = \frac{1}{\sqrt{Trace(A) + 2\sqrt{Dét(A)}}}\left(A + \sqrt{Dét(A)} I_d\right)$$

Remarque — Cette égalité est valable, par exemple, pour A matrice réelle symétrique définie positive.

Q7. En identifiant $M_n(C)$ à C^{n^2}, on voit que l'application $X \mapsto X^2$ est polynomiale homogène de degré 2, elle est donc de classe C^∞.

Pour X, H dans $M_n(C)$, on a :

$$f(X + H) = (X + H)^2 - A = f(X) + XH + HX + H^2$$

c'est-à-dire que la différentielle de f en X est l'application linéaire :

$$\begin{array}{rccc} df(X): & M_n(C) & \to & M_n(C) \\ & H & \mapsto & XH + HX \end{array}$$

Q8. Si $\lambda, \mu \in Sp(M)$, alors $\mu \in Sp({}^tM)$ et il existe x, y dans $C^n - \{0\}$ tels que $Mx = \lambda x$ et ${}^tMy = \mu y$. La deuxième égalité s'écrivant aussi ${}^tyM = \mu{}^ty$. La matrice $H = x{}^ty = \left(\left(x_i y_j\right)\right)_{1\leq i,j\leq n}$ est alors non nulle et on a :

$$df(M)H = x{}^tyM + Mx{}^ty = (\mu + \lambda)H$$

C'est-à-dire que $\lambda + \mu \in Sp(df(M))$.

Réciproquement, soit $\nu \in Sp(df(M))$ et H un vecteur propre associé non nul. On a alors $MH = H(\nu I_d - H)$ et par récurrence sur $p \geq 0$, $M^pH = H(\nu I_d - H)^p$, de sorte que :

$$\forall P \in C[t],\ P(M)H = HP(\nu I_d - H)$$

En prenant pour P le polynôme caractéristique de M, on a $0 = HP(\nu I_d - H)$. Comme $H \neq 0$, la matrice $P(\nu I_d - H)$ ne peut être inversible et elle admet 0 pour valeur propre. En considérant que $Sp(P(B)) = \{P(\nu);\ \nu \in Sp(B)\}$, pour toute matrice B (il suffit de réduire B sous forme triangulaire), on déduit qu'il existe

$\lambda \in Sp(M)$ tel que $P(\nu - \lambda) = 0$, c'est-à-dire qu'il existe $\mu \in Sp(M)$ tel que $\nu - \lambda = \mu$ et $\nu = \lambda + \mu$ avec $\lambda, \mu \in Sp(M)$.

On a donc ainsi prouvé que $Sp(df(M)) = \{\lambda + \mu;\ \lambda, \mu \in Sp(M)\}$.

Q9. $df(X)$ est inversible si et seulement si $0 \notin Sp(df(X))$, ce qui équivaut à dire que pour toutes valeurs propres λ, μ de *X*, on a $\lambda + \mu \neq 0$, c'est-à-dire que *X* n'a pas deux valeurs propres opposées.

Q10. On a vu que $Sp(\sqrt{A}) \subset C^+$ si $Sp(A) \subset C - IR^-$, donc $\sqrt{A}$ ne peut avoir deux valeurs propres opposées et $df(\sqrt{A})$ est inversible, de sorte que la définition de $\alpha(A)$ est légitime.

(a) (i) On a $(g(X) - X)^2 = g(X)^2 + X^2 - Xg(X) - g(X)X$.

D'autre part, avec la définition de la fonction *g*, on a :

$$df(X)g(X) = df(X)X - f(X)$$

soit avec les expressions de $df(X)$ et de $f(X)$:

$$Xg(X) + g(X)X = 2X^2 - f(X) = X^2 + A$$

et :

$$(g(X) - X)^2 = g(X)^2 - A$$

(ii) Il suffit d'écrire que :

$$(g(X) - X)^2 = \left((g(X) - \sqrt{A}) - (X - \sqrt{A})\right)^2 = \left(e(g(X)) - e(X)\right)^2$$

(iii) On a :

$$df(\sqrt{A})e(g(X)) = \sqrt{A}(g(X) - \sqrt{A}) + (g(X) - \sqrt{A})\sqrt{A} = \sqrt{A}g(X) + g(X)\sqrt{A} - 2A$$

Soit :

$$df(\sqrt{A})e(g(X)) = g(X)^2 - A - (g(X) - \sqrt{A})^2$$

Ce qui donne, avec (i) :

$$df(\sqrt{A})e(g(X)) = (g(X) - X)^2 - e(g(X))^2$$

Puis avec (ii) :

$$df(\sqrt{A})e(g(X)) = e(X)^2 - e(g(X))e(X) - e(X)e(g(X))$$

On en déduit alors l'inégalité :

$$\left\|df(\sqrt{A})e(g(X))\right\| \leq \|e(X)\|^2 + 2\|e(X)\|\|e(g(X))\|$$

D'autre part, pour toute matrice *H*, on a :

$$\|H\| = \left\|df(\sqrt{A})^{-1} df(\sqrt{A})H\right\| \leq \left\|df(\sqrt{A})^{-1}\right\|\left\|df(\sqrt{A})H\right\|$$

et :

$$\alpha(A)\|H\| \leq \left\|df(\sqrt{A})H\right\|$$

Cette inégalité est vraie, en particulier, pour $H = e(g(X))$.

(b) On suppose que $\left\|X_0 - \sqrt{A}\right\| < \dfrac{\alpha(A)}{3}$.

Pour tout $H \in M_n(C)$, on a :

$$df\left(\sqrt{A}\right)H = \sqrt{A}H + H\sqrt{A} = \left(\sqrt{A} - X_0\right)H + H\left(\sqrt{A} - X_0\right) + X_0H + HX_0$$

soit :

$$df\left(\sqrt{A}\right)H = df\left(\sqrt{A} - X_0\right)H + df\left(X_0\right)H$$

On en déduit alors que :

$$\left\|df\left(\sqrt{A}\right)H\right\| \leq \left\|df\left(\sqrt{A} - X_0\right)H\right\| + \left\|df\left(X_0\right)H\right\| \leq 2\left\|\sqrt{A} - X_0\right\|\left\|H\right\| + \left\|df\left(X_0\right)H\right\|$$

$$\leq 2\frac{\alpha(A)}{3}\left\|H\right\| + \left\|df\left(X_0\right)H\right\|$$

En (a) (iii), on a vu que, par définition de $\alpha(A)$, on a $\alpha(A)\|H\| \leq \left\|df\left(\sqrt{A}\right)H\right\|$. Il en résulte donc que :

$$\forall H \in M_n(C),\ \alpha(A)\|H\| \leq 2\frac{\alpha(A)}{3}\|H\| + \left\|df\left(X_0\right)H\right\|$$

Ce qui entraîne $\left\|df\left(X_0\right)H\right\| \geq \frac{\alpha(A)}{3}\|H\| > 0$ si $H \neq 0$. On déduit alors que $df\left(X_0\right)$ est inversible et que X_1 est bien défini.

L'inégalité (iii) du (a) s'écrit alors :

$$\alpha(A)\left\|X_1 - \sqrt{A}\right\| \leq \left\|X_0 - \sqrt{A}\right\|\left\{\left\|X_0 - \sqrt{A}\right\| + 2\left\|X_1 - \sqrt{A}\right\|\right\}$$

$$\leq \frac{\alpha(A)}{3}\left\{\left\|X_0 - \sqrt{A}\right\| + 2\left\|X_1 - \sqrt{A}\right\|\right\}$$

Ce qui donne $\left\|X_1 - \sqrt{A}\right\| \leq \left\|X_0 - \sqrt{A}\right\|$.

Par récurrence, en supposant que $X_0, X_1, \cdots, X_k$ sont définis avec $\left\|X_k - \sqrt{A}\right\| \leq \cdots \leq \left\|X_1 - \sqrt{A}\right\| \leq \left\|X_0 - \sqrt{A}\right\|$, le même raisonnement prouve que X_{k+1} est bien défini avec $\left\|X_{k+1} - \sqrt{A}\right\| \leq \left\|X_k - \sqrt{A}\right\| \leq \left\|X_0 - \sqrt{A}\right\|$.

La suite $(e_k)_{k\in IN} = \left(\left\|X_k - \sqrt{A}\right\|\right)_{k\in IN}$ est donc décroissante minorée et elle converge vers une quantité $L \in \left[0, \left\|X_0 - \sqrt{A}\right\|\right] \subset \left[0, \frac{\alpha(A)}{3}\right[$.

L'inégalité (iii) du (a) nous donne $\alpha(A)e_{k+1} \leq e_k{}^2 + 2e_ke_{k+1}$, et en faisant tendre k vers l'infini, on aboutit à $\alpha(A)L \leq 3L^2$.

Si $L > 0$, alors $L \geq \frac{\alpha(A)}{3}$, ce qui contredit $L < \frac{\alpha(A)}{3}$. On a donc nécessairement $L = 0$, c'est-à-dire que $\underset{k\to+\infty}{Lim} X_k = \sqrt{A}$.

Q11. On remarque d'abord que si $X = Diag\left(x_1, x_2, \cdots, x_n\right)$ est diagonale avec $x_i + x_j \neq 0$ pour $i \neq j$, alors $df(X)^{-1}$ est définie par :

$$\forall H \in M_n(C),\ df(X)^{-1} H = \begin{pmatrix} \dfrac{h_{11}}{2x_1} & \cdots & \dfrac{h_{1n}}{x_1 + x_n} \\ \vdots & \cdots & \vdots \\ \dfrac{h_{n1}}{x_n + x_1} & \cdots & \dfrac{h_{nn}}{2x_n} \end{pmatrix}$$

Ce qui donne :

$$g(X) = X - df(X)^{-1} f(X) = Diag(y_1, y_2, \cdots, y_n)$$

avec :

$$y_i = x_i - \frac{x_i^{\,2} - \lambda_i}{2x_i} = \frac{1}{2}\left(x_i + \frac{\lambda_i}{x_i} \right)$$

si $A = Diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

On déduit alors que $(X_k)_{k \in IN}$ est une suite de matrices diagonales définie par $X_k = Diag\left(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}\right)$ où :

$$\begin{cases} x_i^{(0)} = 1 \\ x_i^{(k+1)} = \dfrac{1}{2}\left(x_i^{(k)} + \dfrac{\lambda_i}{x_i^{(k)}} \right) (k \geq 0) \end{cases} \quad (1 \leq i \leq n)$$

En Q2, on a vu que chaque suite $\left(x_i^{(k)}\right)_{k \in IN}$ est bien définie et converge vers $\sqrt{\lambda_i}$ si $x_i^{(0)} > 0$.

C'est-à-dire que la suite $(X_k)_{k \in IN}$ est bien définie et converge vers $\sqrt{A}$.

De plus la récurrence ci-dessus s'écrit :

$$\begin{cases} X_0 = I_d \\ X_{k+1} = \dfrac{1}{2}\left(X_k + A X_k^{\ -1} \right) (k \geq 0) \end{cases}$$

Q12. On a $A = PDP^{-1}$, avec $D = Diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ et P inversible.

On pose $\varphi(Y) = Y^2 - D$.

Pour $X = PYP^{-1}$, on a :

$df(X)H = XH + HX = P\left(Y\left(P^{-1}HP\right) + \left(P^{-1}HP\right)Y\right)P^{-1} = Pd\varphi(Y)\left(P^{-1}HP\right)P^{-1}$

et $df(X)$ est inversible si et seulement si $d\varphi(Y)$ est inversible (ce qui peut aussi se voir en remarquant que X et Y ont les mêmes valeurs propres).

$Z = g(X) = X - df(X)^{-1} f(X)$ est solution de $df(X)Z = df(X)X - f(X)$, soit $Pd\varphi(Y)\left(P^{-1}ZP\right)P^{-1} = Pd\varphi(Y)\left(P^{-1}XP\right)P^{-1} - \left(X^2 - A\right)$. C'est-à-dire que $P^{-1}ZP$ est solution de $d\varphi(Y)\left(P^{-1}ZP\right) = d\varphi(Y)\left(P^{-1}XP\right) - \left(Y^2 - D\right)$.

On a donc $P^{-1}ZP = Y - d\varphi(Y)^{-1} \varphi(Y)$.

On en déduit alors que la suite $(X_k)_{k \in IN}$ est définie par :

$$\forall k \geq 0,\ X_k = PY_k P^{-1}$$

où $(Y_k)_{k\in IN}$ est la suite des itérations de l'algorithme de Newton associé à la résolution de $\varphi(Y) = Y^2 - D = 0$.

Si on prend $X_0 = I_d$, alors $Y_0 = I_d$ et la suite $(Y_k)_{k\in IN}$ est bien définie et converge vers $\sqrt{D}$. Il en résulte alors que la suite $(X_k)_{k\in IN}$ est bien définie et converge vers $\sqrt{A}$.

De plus, on a $X_{k+1} = \dfrac{1}{2}P\left(Y_k + DY_k^{-1}\right)P^{-1} = \dfrac{1}{2}\left(X_k + AX_k^{-1}\right)$.

Remarque — La suite $(X_k)_{k\in IN}$ ne donne un procédé numériquement stable que si le conditionnement de *A,* relatif à une norme matricielle induite (problème 6) est inférieur à 9. Un algorithme stable est donné en considérant les suites $(X_k)_{k\in IN}$ et $(Y_k)_{k\in IN}$ définies par :

$$\begin{cases} X_0 = A;\ Y_0 = I_d \\ X_{k+1} = \dfrac{1}{2}\left(X_k + Y_k^{-1}\right);\ Y_{k+1} = \dfrac{1}{2}\left(Y_k + X_k^{-1}\right)\ (k \geq 0) \end{cases}$$

On a alors $\underset{k\to+\infty}{Lim} X_k = \sqrt{A}$ et $\underset{k\to+\infty}{Lim} Y_k = \sqrt{A}^{-1}$ (Schatzman, p. 243).

CHAPITRE 5

Approximation polynomiale des fonctions numériques

Problème 33 : Polynômes d'interpolation d'Hermite

On désigne par $IR[x]$ l'espace vectoriel des polynômes à coefficients réels et pour tout entier $m \geq 1$, $IR_m[x]$ désigne le sous-espace vectoriel de $IR[x]$ formé des polynômes de degré inférieur ou égal à m.

On se donne un entier $n \geq 1$, une suite de réels deux à deux distincts $X = (x_0, x_1, \cdots, x_n) \in IR^{n+1}$ et une suite d'entiers $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_n) \in IN^{n+1}$.

On note $m = \sum_{i=0}^{n} (\alpha_i + 1)$ et $\pi_m(x) = \prod_{i=0}^{n} (x - x_i)^{\alpha_i + 1}$.

On désigne par $H: IR[x] \to IR^m$ l'application linéaire définie par :

$$\forall P \in IR[x], H(P) = \left(P(x_0), P'(x_0), \cdots, P^{(\alpha_0)}(x_0), \cdots, P(x_n), P'(x_n), \cdots, P^{(\alpha_n)}(x_n)\right)$$

Pour toute fonction continue $\varphi: [a,b] \to IR$, on note :

$$\|\varphi\|_\infty = Sup\{|\varphi(t)|; t \in [a,b]\}$$

Q1. *Calculer le noyau de H.*

Q2. *Montrer que $IR[x] = Ker(H) \oplus IR_{m-1}[x]$.*

Q3. *Montrer que pour toute suite de m réels $Y = \left(y_{0,0}, \cdots, y_{0,\alpha_0}, \cdots, y_{n,0}, \cdots, y_{n,\alpha_n}\right)$ il existe un unique polynôme $P \in IR_{m-1}[x]$ tel que :*

$$P^{(k)}(x_i) = y_{i,k} \quad (0 \leq i \leq n, 0 \leq k \leq \alpha_i)$$

On dit que P est le polynôme d'interpolation d'Hermite associé à α, X et Y.

Dans le cas où $\alpha = 0$ dans IN^m, P est le polynôme d'interpolation de Lagrange associé à X et Y.

Q4. *Pour $i \in \{0,1,\cdots,n\}$ et $k \in \{0,1,\cdots,\alpha_i\}$, on désigne par $H_{i,k}$ le polynôme de $IR_{m-1}[x]$ défini par :*

$$H_{i,k}^{(q)}\left(x_p\right)=\begin{cases}0 \ \ si \ \ (p,q)\neq(i,k)\\ 1 \ \ si \ \ (p,q)=(i,k)\end{cases}$$

(a) Montrer que $B_H=\left\{H_{i,k},\,0\leq i\leq n,\,0\leq k\leq\alpha_i\right\}$ *est une base de* $IR_{m-1}\left[x\right]$.

(b) Donner une expression des $L_i=H_{i,0}$, *pour i = 0, 1, ⋯, n, dans le cas où* $\alpha=0$.

(c) Donner une expression des $H_{i,k}$, *pour i = 0, 1, ⋯, n, k = 0, 1, dans le cas où* $\alpha=(1,1,\cdots,1)$.

(d) Ecrire le polynôme d'interpolation d'Hermite associé à α*, X et Y dans la base* B_H.

Q5. *On se donne un intervalle fermé* $\left[a,b\right]$ *contenant tous les* x_i *pour i = 0, 1, ⋯, n et une fonction* $f:\left[a,b\right]\to IR$ *de classe* C^m.

Le polynôme d'interpolation d'Hermite de f associé à α *et X est défini par :*

$$\begin{cases}d°\left(P_n\right)\leq m-1\\ P_n^{(k)}\left(x_i\right)=f^{(k)}\left(x_i\right) \quad \left(0\leq i\leq n,0\leq k\leq\alpha_i\right)\end{cases}$$

L'erreur d'interpolation est définie par :

$$\forall x\in\left[a,b\right],\,E_n(x)=f(x)-P_n(x)$$

(a) Montrer que si $g:\left[a,b\right]\to IR$ *est une fonction de classe* C^m *qui s'annule en* $m+1$ *points distincts ou confondus, alors il existe* $\xi\in\left[a,b\right]$ *tel que* $g^{(m)}(\xi)=0$.

(b) Soit $x\in\left[a,b\right]-\left\{x_0,x_1,\cdots,x_n\right\}$ *et* Q_{n+1} *le polynôme d'interpolation d'Hermite de f associé à* $\left(\alpha_0,\alpha_1,\cdots,\alpha_n,0\right)$ *et* $\left(x_0,x_1,\cdots,x_n,x\right)$.

(i) Montrer que :

$$Q_{n+1}(t)=P_n(t)+\frac{f(x)-P_n(x)}{\pi_m(x)}\pi_m(t)$$

(ii) Calculer ${Q_{n+1}}^{(m)}$.

(iii) Si on pose $g=f-Q_{n+1}$, *montrer alors qu'on peut trouver* $\xi_x\in\left[a,b\right]$ *tel que* $g^{(m)}\left(\xi_x\right)=0$.

(c) Montrer que pour tout $x\in\left[a,b\right]$, *on peut trouver* $\xi_x\in\left[a,b\right]$ *tel que :*

$$E_n(x)=f(x)-P_n(x)=\frac{1}{m!}\pi_m(x)f^{(m)}\left(\xi_x\right)$$

Q6. *On garde les notations de Q5 en supposant de plus que la fonction f est analytique au voisinage de* $c=\dfrac{a+b}{2}$ *avec un rayon de convergence* $R>\dfrac{b-a}{2}$.

(a) Montrer que pour tout $r\in\left]\dfrac{b-a}{2},R\right[$, *on peut trouver une constante* C_r *telle que :*

$$\forall m\in IN,\left\|f^{(m)}\right\|_\infty\leq\frac{m!\cdot r\cdot C_r}{\left(r-\dfrac{b-a}{2}\right)^{m+1}}$$

(b) En déduire que si $R > \frac{3}{2}(b-a)$, alors la suite $(P_n)_{n\geq 1}$ des polynômes d'interpolation d'Hermite de f converge uniformément vers f.

Solution

Q1. Tout polynôme $P \in Ker(H)$ vérifie :

$$P^{(k)}(x_i) = 0 \quad (0 \leq i \leq n, 0 \leq k \leq \alpha_i)$$

C'est donc un multiple du polynôme π_m. Ce qui montre que :

$$Ker(H) = IR[x] \cdot \pi_m = \{Q \cdot \pi_m ; Q \in R[x]\}$$

Q2. Avec le théorème de division euclidienne, on peut écrire de manière unique tout polynôme $P \in R[x]$ sous la forme :

$$P = Q \cdot \pi_m + R \quad (R = 0 \text{ ou } d°(R) < m)$$

Ce qui donne le résultat.

Q3. Avec Q2, on déduit que $Ker\left(H/_{IR_{m-1}[x]}\right) = \{0\}$, c'est-à-dire que la restriction de H à $IR_{m-1}[x]$ est injective. Les espaces étant de même dimension, on en déduit que c'est un isomorphisme de $IR_{m-1}[x]$ sur IR^m. D'où le résultat.

Q4. (a) La famille B_H est l'image réciproque par l'isomorphisme $H/_{IR_{m-1}[x]}$ de la base canonique de IR^m. C'est donc une base de $IR_{m-1}[x]$.

(b) Si $\alpha = 0$, alors les $L_i = H_{i,0}$ sont définis par :

$$\begin{cases} d°(L_i) \leq m-1 = n \\ L_i(x_j) = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases} \end{cases}$$

On en déduit alors que :

$$L_i(x) = \frac{1}{\prod\limits_{\substack{0\leq j\leq n \\ j\neq i}} (x_i - x_j)} \prod_{\substack{0\leq j\leq n \\ j\neq i}} (x - x_j) \quad (0 \leq i \leq n)$$

En écrivant $\pi_{n+1}(x) = (x - x_i) \prod\limits_{\substack{j=0 \\ j\neq i}}^{n} (x - x_j)$, on déduit que $\pi'_{n+1}(x_i) = \prod\limits_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j)$ et :

$$L_i(x) = \frac{\pi_{n+1}(x)}{\pi'_{n+1}(x_i)(x - x_i)} \quad (0 \leq i \leq n)$$

(c) Si $\alpha = (1, \cdots, 1)$, alors les $H_{i,k}$ sont définis par :

$$\begin{cases} d°\left(H_{i,k}\right) \leq m-1 = 2n+1 \\ H_{i,0}\left(x_j\right) = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases} \\ H'_{i,0}\left(x_j\right) = 0 \\ H_{i,1}\left(x_j\right) = 0 \\ H'_{i,1}\left(x_j\right) = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases} \end{cases}$$

(i) Avec $H_{i,0}\left(x_j\right) = H'_{i,0}\left(x_j\right) = 0$, pour $j \neq i$, on déduit que :

$$H_{i,0}(x) = C(x)\prod_{\substack{j=0 \\ j\neq i}}^{n}\left(x - x_j\right)^2 = Q(x)L_i(x)^2$$

avec Q de degré inférieur ou égal à 1.

Puis avec $H_{i,0}(x_i) = 1 = L_i(x_i)$, on déduit que $Q(x_i) = 1$, c'est-à-dire que $Q(x) = 1 + \beta(x - x_i)$.

Enfin avec $0 = H'_{i,0}(x_i) = Q'(x_i)L_i(x_i)^2 + 2Q(x_i)L_i(x_i)L'_i(x_i) = \beta + 2L'_i(x_i)$, on déduit que :

$$H_{i,0}(x) = \left\{1 - 2L'_i(x_i)(x - x_i)\right\}L_i(x)^2$$

avec :

$$L'_i(x_i) = \sum_{\substack{0\leq j\leq n \\ j\neq i}} \frac{1}{x_i - x_j}$$

(ii) Avec $H_{i,1}\left(x_j\right) = H'_{i,1}\left(x_j\right) = 0$, pour $j \neq i$, on déduit que :

$$H_{i,1}(x) = D(x)\prod_{\substack{j=0 \\ j\neq i}}^{n}\left(x - x_j\right)^2 = R(x)L_i(x)^2$$

avec R de degré inférieur ou égal à 1.

Puis avec $H_{i,1}(x_i) = 0$, on déduit que $R(x) = \gamma(x - x_i)$.

Enfin $1 = H'_{i,1}(x_i) = \gamma \cdot L_i(x_i)^2$ donne $\gamma = 1$. Soit :

$$H_{i,0}(x) = (x - x_i)L_i(x)^2$$

(d) Le polynôme d'interpolation d'Hermite associé à α, X et Y s'écrit dans la base B_H, sous la forme $P(x) = \sum_{i=0}^{n}\sum_{k=0}^{\alpha_i} c_{i,k}H_{i,k}(x)$.

Et avec $P^{(q)}\left(x_p\right) = \sum_{i=0}^{n}\sum_{k=0}^{\alpha_i} c_{i,k}H_{i,k}{}^{(q)}\left(x_p\right) = c_{p,q} = y_{p,q}$, on déduit que :

$$P(x) = \sum_{i=0}^{n}\sum_{k=0}^{\alpha_i} y_{i,k}H_{i,k}(x)$$

Pour $\alpha = 0$, le polynôme d'interpolation de Lagrange est défini par :

$$\begin{cases} d°(P) \leq n \\ P(x_i) = y_i \quad (0 \leq i \leq n) \end{cases}$$

et il s'écrit :

$$P(x) = \sum_{i=0}^{n} y_i L_i(x)$$

Pour $\alpha = (1, \cdots, 1)$, le polynôme d'interpolation d'Hermite est défini par :

$$\begin{cases} d°(P) \leq 2n+1 \\ P(x_i) = y_i \quad (0 \leq i \leq n) \\ P'(x_i) = y_i' \quad (0 \leq i \leq n) \end{cases}$$

et il s'écrit :

$$P(x) = \sum_{i=0}^{n} y_i \{1 - 2L_i'(x_i)(x - x_i)\} L_i(x)^2 + \sum_{i=0}^{n} y_i'(x - x_i) L_i(x)^2$$

Q5. (a) Le théorème de Rolle nous dit qu'entre deux racines de g on peut trouver une racine de g'. La fonction g' admet donc m racines comptées avec leur multiplicité. Par récurrence, on déduit alors que la dérivée d'ordre m, $g^{(m)}$, admet une racine dans $[a,b]$.

(b) (i) En utilisant la somme directe de Q2, on peut écrire que $Q_{n+1} = \varphi + \psi \cdot \pi_m$, avec $\varphi \in IR_{m-1}[t]$ et $\psi \in IR$ ($d°(Q_{n+1}) = m = d°(\pi_m)$).

Pour i = 0, 1, ···, n et k = 0, 1, ···, α_i, on a $f^{(k)}(x_i) = Q^{(k)}(x_i) = \varphi^{(k)}(x_i)$, avec $d°(\varphi) \leq m-1$. C'est-à-dire que φ est le polynôme d'interpolation d'Hermite de f associé à α et X. On a donc $\varphi = P_n$.

En prenant $t = x$, on a alors $f(x) = Q_{n+1}(x) = P(x) + \psi \cdot \pi_m(x)$ et $\psi = \dfrac{f(x) - P(x)}{\pi_m(x)}$.

(ii) Le polynôme Q_{n+1} est de degré m, avec $\dfrac{f(x) - P(x)}{\pi_m(x)} t^m$ pour terme dominant.

On a donc :

$$Q_{n+1}{}^{(m)}(t) = \frac{f(x) - P(x)}{\pi_m(x)} m!$$

(iii) Pour tout i = 0, 1, ···, n, le réel x_i est racine d'ordre supérieur ou égal à α_i de g. Le réel $t = x$ est racine d'ordre supérieur ou égal à 1 de g. La fonction g admet donc au moins m + 1 racines comptées avec leur multiplicité. L'existence de ξ_x découle alors de (a).

(c) Si x est l'un des x_i, alors $\pi_m(x_i) = 0 = E_n(x_i)$ et tout réel ξ_x de $[a,b]$ convient.

On suppose donc que $x \in [a,b] - \{x_0, x_1, \cdots, x_n\}$.

Le résultat de (b) (iii) s'écrit $f^{(m)}(\xi_x) - Q_{n+1}{}^{(m)}(\xi_x) = 0$, soit avec (b) (ii) $f^{(m)}(\xi_x) - \dfrac{f(x) - P(x)}{\pi_m(x)} m! = 0$, c'est-à-dire :

$$f(x) - P_n(x) = \frac{1}{m!}\pi_m(x) f^{(m)}(\xi_x)$$

Q6. (a) Pour tout $x \in]c-R, c+R[$, on peut écrire :

$$f(x) = \sum_{k \geq 0} a_k (x-c)^k$$

En prenant $r \in \left]\frac{b-a}{2}, R\right[$ et $x = c + r$, on obtient la série convergente $\sum_{k \geq 0} a_k r^k$. On a alors, en particulier, $\underset{k \to +\infty}{Lim}(a_k r^k) = 0$ et on peut trouver une constante C_r telle que :

$$k \geq 0, \quad |a_k| \leq \frac{C_r}{r^k}$$

La série ci-dessus peut être dérivée indéfiniment sur $[a,b] \subset]c-R, c+R[$, avec :

$$f^{(m)}(x) = \sum_{k \geq m} a_k \frac{k!}{(k-m)!}(x-c)^{k-m}$$

On a alors, pour tout $x \in [a,b]$, $\left|f^{(m)}(x)\right| \leq C_r \sum_{k \geq m} \frac{1}{r^k}\frac{k!}{(k-m)!}|x-c|^{k-m}$, avec $|x-c| \in \left[0, \frac{b-a}{2}\right]$.

En écrivant que pour $t \in \left[0, \frac{b-a}{2}\right] \subset [0, r[$, on a :

$$\sum_{k \geq m} \frac{1}{r^k}\frac{k!}{(k-m)!}t^{k-m} = \left(\sum_{k \geq 0} \frac{t^k}{r^k}\right)^{(m)} = \left(\frac{1}{1-\frac{t}{r}}\right)^{(m)} = \frac{m!r}{(r-t)^{m+1}} \leq \frac{m!r}{\left(r - \frac{b-a}{2}\right)^{m+1}}$$

on déduit :

$$\left|f^{(m)}(x)\right| \leq \frac{m! \cdot r \cdot C_r}{\left(r - \frac{b-a}{2}\right)^{m+1}}$$

(b) La majoration de l'erreur d'interpolation obtenue en Q5 (c) nous donne :

$$\|f - P_n\|_\infty \leq \frac{1}{m!}\|\pi_m\|_\infty \left\|f^{(m)}\right\|_\infty$$

Avec $|\pi_m(x)| = \left|\prod_{i=0}^{n}(x-x_i)^{\alpha_i+1}\right| \leq (b-a)^m$ et la majoration obtenue en (a), on déduit que :

$$\|f - P_n\|_\infty \leq (b-a)^m \frac{r \cdot C_r}{\left(r - \frac{b-a}{2}\right)^{m+1}} = \frac{r \cdot C_r}{r - \frac{b-a}{2}}\left(\frac{b-a}{r - \frac{b-a}{2}}\right)^m$$

La convergence uniforme est assurée si on peut choisir $r \in \left]\frac{b-a}{2}, R\right[$ tel que $b-a < r - \frac{b-a}{2}$, soit $r > \frac{3}{2}(b-a)$. Ce qui est possible si $R > \frac{3}{2}(b-a)$.

Problème 34 : Polynômes d'interpolation de Fejer–Hermite

Ce problème utilise les notations et résultats du problème 33.

Pour tout entier $n \geq 1$, on désigne par $X = (x_1, \cdots, x_n)$ la suite des racines du polynôme de Tchébychev T_n défini sur $[-1,1]$ par $T_n(x) = Cos(n \cdot ArcCos(x))$.

On note $\pi_n(x) = \prod_{i=1}^{n}(x - x_i)$ et pour $1 \leq i \leq n$, on désigne par $L_i(x) = \frac{1}{\prod_{\substack{1\leq j\leq n \\ j\neq i}}(x_i - x_j)} \prod_{\substack{1\leq j\leq n \\ j\neq i}}(x - x_j)$, le $i^{ème}$ polynôme de base de Lagrange associé à X.

Si $f : [-1,1] \to IR$ est une fonction continue, on définit alors le polynôme d'interpolation de Fejer–Hermite associé à f et X par les conditions :

$$\begin{cases} P_n \in IR_{2n-1}[x] \\ P_n(x_i) = f(x_i),\ P_n'(x_i) = 0 \quad (1 \leq i \leq n) \end{cases}$$

Q1. *Montrer que :*

$$L_i(x) = \frac{\pi_n(x)}{\pi_n'(x_i)(x - x_i)} \quad (1 \leq i \leq n)$$

Q2. *Montrer que :*

$$P_n(x) = \sum_{i=1}^{n} f(x_i) F_i(x)$$

avec :

$$F_i(x) = \left\{1 - \frac{\pi_n''(x_i)}{\pi_n'(x_i)}(x - x_i)\right\} L_i(x)^2 \quad (1 \leq i \leq n)$$

Q3. *Montrer que, pour $1 \leq i \leq n$, on a les identités suivantes :*

(a) $L_i(x) = \frac{(-1)^{i-1}\sqrt{1 - x_i^2}}{n} \frac{T_n(x)}{x - x_i}$.

(b) $F_i(x) = (1 - x \cdot x_i)\left(\frac{T_n(x)}{n(x - x_i)}\right)^2$.

Q4. *Montrer que :*

$$\forall x \in IR, \quad \sum_{i=1}^{n} F_i(x) = 1$$

Q5. *Montrer que :*

$$\forall x \in [-1,1], \quad |f(x) - P_n(x)| \le \sum_{i=1}^{n} |f(x) - f(x_i)| F_i(x)$$

Q6. *Soient* $\varepsilon > 0$ *et* $\eta > 0$ *tels que :*

$$x, x' \in [-1,1] \text{ et } |x - x'| < \eta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon$$

Pour tout $x \in [-1,1]$, *on note* $I_{x,\varepsilon} = \{i \in \{1,2,\cdots,n\};\ |x - x_i| < \eta\}$ *et* $K_{x,\varepsilon} = \{1,2,\cdots,n\} - I_{x,\varepsilon}$.

(a) Montrer que :

$$\sum_{i \in I_{x,\varepsilon}} |f(x) - f(x_i)| F_i(x) < \varepsilon$$

(b) Montrer que :

$$\sum_{i \in K_{x,\varepsilon}} |f(x) - f(x_i)| F_i(x) \le \frac{4\|f\|_\infty}{n \cdot \eta^2}$$

(c) En déduire que la suite $(P_n)_{n\ge1}$ *des polynômes d'interpolation de Fejer–Hermite converge uniformément vers f sur* $[-1,1]$.

Solution

Q1. Cette identité a été établie au problème 33, Q4 (b).

Q2. En utilisant l'expression du polynôme d'interpolation d'Hermite obtenue au problème 33, Q4 (d) et les conditions $P_n'(x_i) = 0$ pour tout i = 1, 2, ···, n, on déduit que :

$$P_n(x) = \sum_{i=1}^{n} f(x_i)\{1 - 2L_i'(x_i)(x - x_i)\} L_i(x)^2$$

D'autre part, avec Q1, on peut écrire :

$$\pi_n(x) = L_i(x)\pi_n'(x_i)(x - x_i)$$
$$\pi_n'(x) = L_i'(x)\pi_n'(x_i)(x - x_i) + L_i(x)\pi_n'(x_i)$$
$$\pi_n''(x) = L_i''(x)\pi_n'(x_i)(x - x_i) + 2L_i'(x)\pi_n'(x_i)$$

et :

$$\pi_n''(x_i) = 2L_i'(x_i)\pi_n'(x_i)$$

C'est-à-dire que le polynôme d'interpolation de Fejer–Hermite peut aussi s'écrire sous la forme :

$$P_n(x) = \sum_{i=1}^{n} f(x_i)\left\{1 - \frac{\pi_n''(x_i)}{\pi_n'(x_i)}(x - x_i)\right\} L_i(x)^2$$

Q3. (a) On a $\pi_n(x) = (x - x_i)\prod_{\substack{j=1 \\ j\ne i}}^{n}(x - x_j)$ et $\pi_n'(x_i) = \prod_{\substack{j=1 \\ j\ne i}}^{n}(x_i - x_j)$ de sorte que :

$$L_i(x) = \frac{\pi_n(x)}{\pi_n'(x_i)(x - x_i)} = \frac{T_n(x)}{T_n'(x_i)(x - x_i)} \quad (1 \le i \le n)$$

D'autre part, pour tout $x \in]-1,1[$, on a $T_n'(x) = \dfrac{nSin(nArcCos(x))}{\sqrt{1-x^2}}$ et pour $x = x_i = Cos\left(\dfrac{2i-1}{2n}\pi\right)$, avec $1 \leq i \leq n$, on a $T_n'(x_i) = \dfrac{nSin\left(\dfrac{2i-1}{2}\pi\right)}{\sqrt{1-x_i^{\,2}}} = \dfrac{n(-1)^{i-1}}{\sqrt{1-x_i^{\,2}}}$, ce qui donne en définitive :

$$L_i(x) = \frac{(-1)^{i-1}\sqrt{1-x_i^{\,2}}}{n}\frac{T_n(x)}{x-x_i}$$

(b) On a $T_n''(x) = n\left\{\dfrac{xSin(nArcCos(x))}{(1-x^2)^{\frac{3}{2}}} - \dfrac{nT_n(x)}{(1-x^2)^{\frac{1}{2}}}\right\}$ et $T_n''(x_i) = \dfrac{n \cdot x_i(-1)^{i-1}}{(1-x_i^{\,2})^{\frac{3}{2}}}$. On en déduit alors que $\dfrac{\pi_n''(x_i)}{\pi_n'(x_i)} = \dfrac{T_n''(x_i)}{T_n'(x_i)} = \dfrac{x_i}{1-x_i^{\,2}}$ et $1 - \dfrac{\pi_n''(x_i)}{\pi_n'(x_i)}(x-x_i) = \dfrac{1-x \cdot x_i}{1-x_i^{\,2}}$.

Ce qui nous donne l'expression suivante des fonctions de base F_i :

$$F_i(x) = \frac{1-x \cdot x_i}{1-x_i^{\,2}} L_i(x)^2 = (1-x \cdot x_i)\left(\frac{T_n(x)}{n(x-x_i)}\right)^2$$

Q4. Les fonctions de base F_i sont définies par les conditions :

$$\begin{cases} F_i \in IR_{2n-1}[x] \\ F_i(x_j) = \begin{cases} 1 \text{ si } j = i \\ 0 \text{ si } j \neq i \end{cases}, \ F_i'(x_j) = 0 \quad (1 \leq j \leq n) \end{cases}$$

Le polynôme $Q(x) = \sum\limits_{i=1}^{n} F_i(x)$ est donc l'unique polynôme de degré inférieur ou égal à $2n-1$ qui vérifie $Q(x_i) = 1$ et $Q'(x_i) = 0$ pour i = 1, 2, $\cdots$, n. Le polynôme constant égal à 1 vérifiant aussi ces conditions, on a $\sum\limits_{i=1}^{n} F_i(x) = 1$, pour tout réel x.

Q5. On peut écrire que $f(x) - P_n(x) = f(x)\sum\limits_{i=1}^{n} F_i(x) - \sum\limits_{i=1}^{n} f(x_i)F_i(x)$ et :

$$|f(x) - P_n(x)| = \left|\sum_{i=1}^{n}(f(x) - f(x_i))F_i(x)\right| \leq \sum_{i=1}^{n}|f(x) - f(x_i)||F_i(x)|$$

Avec l'identité obtenue en Q3, (b), on déduit que $F_i(x) \geq 0$ pour tout $x \in [-1,1]$. On peut donc écrire que :

$$\forall x \in [-1,1], \quad |f(x) - P_n(x)| \leq \sum_{i=1}^{n}|f(x) - f(x_i)|F_i(x)$$

Q6. La fonction f étant uniformément continue sur l'intervalle compact $[-1,1]$, on peut trouver, pour tout réel $\varepsilon > 0$, un réel $\eta > 0$ tel que :

$$x, x' \in [-1,1] \ et \ |x - x'| < \eta \ \Rightarrow \ |f(x) - f(x')| < \varepsilon$$

(a) Pour tout $i \in I_{x,\varepsilon}$, on a $|x - x_i| < \eta$ et $|f(x) - f(x_i)| < \varepsilon$, de sorte que :

$$\sum_{i\in I_{x,\varepsilon}} \left|f(x)-f(x_i)\right| F_i(x) < \varepsilon \sum_{i=1}^{n} F_i(x) = \varepsilon$$

(b) Pour $x\in[-1,1]$ et $i\in K_{x,\varepsilon}$, on a :

$$0<1-x\cdot x_i<2,\quad |T_n(x)|\leq|1,\quad |x-x_i|\geq\eta$$

de sorte que :

$$0\leq F_i(x)=(1-x\cdot x_i)\left(\frac{T_n(x)}{n(x-x_i)}\right)^2\leq\frac{2}{n^2\eta^2}$$

Ce qui donne :

$$\sum_{i\in K_{x,\varepsilon}} \left|f(x)-f(x_i)\right| F_i(x)\leq \sum_{i\in K_{x,\varepsilon}} 2\|f\|_\infty \frac{2}{n^2\cdot\eta^2}\leq\sum_{i=1}^{n}2\|f\|_\infty\frac{2}{n^2\cdot\eta^2}\leq\frac{4\|f\|_\infty}{n\cdot\eta^2}$$

(c) Avec (a) et (b), on déduit que pour tout $\varepsilon>0$, on peut trouver $\eta>0$ tel pour tout $n\geq 1$ et tout $x\in[-1,1]$, on ait :

$$\left|f(x)-P_n(x)\right|\leq\varepsilon+\frac{4\|f\|_\infty}{n\cdot\eta^2}$$

On peut alors trouver un entier n_η tel que :

$$\forall n\geq n_\eta,\quad \frac{4\|f\|_\infty}{n\cdot\eta^2}<\varepsilon$$

On a alors :

$$\forall n\geq n_\eta,\quad \forall x\in[-1,1],\quad \left|f(x)-P_n(x)\right|<2\varepsilon$$

C'est-à-dire que la suite des polynômes d'interpolation de Fejer–Hermite converge uniformément vers *f* sur $[-1,1]$.

Problème 35 : Polynômes de Bernstein

Pour tout entier $n\geq 1$*, on désigne par* $IR_n[x]$ *l'espace vectoriel des polynômes réels de degré inférieur ou égal à n.*

On munit $IR_n[x]$ *de la base canonique* $\{e_i ; 0\leq i\leq n\}$ *définie par :*

$$e_i(x)=x^i\quad(0\leq i\leq n)$$

L'espace vectoriel $E=C^0([0,1];IR)$ *des fonctions continues de* $[0,1]$ *dans IR est muni de la norme de la convergence uniforme définie par :*

$$\forall f\in E,\quad \|f\|_\infty=Sup\{|f(x)|;x\in[0,1]\}$$

Pour tout entier $k\in\{0,1,\cdots,n\}$*, on désigne par* $B_{n,k}$ *le polynôme défini par :*

$$B_{n,k}(x)=C_n^k x^k(1-x)^{n-k}$$

On désigne par B_n *l'application linéaire de E dans* $IR_n[x]$ *définie par :*

$$\forall f\in E,\quad B_n(f)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)B_{n,k}$$

Q1. *Soit* $y\in IR$ *et* $f_y\in E$ *définie par* $f_y(x)=e^{yx}$.

(a) Calculer $B_n(f_y)$.

(b) On désigne par φ_n *la fonction définie sur* IR^2 *par :*

$$\forall (x,y) \in IR^2, \quad \varphi_n(x,y) = B_n(f_y)(x)$$

Montrer que :

$$\forall i \in IN, \quad B_n(e_i)(x) = \frac{\partial^i \varphi_n}{\partial y^i}(x,0)$$

(c) Calculer $B_n(1)$, $B_n(x)$ *et* $B_n(x^2)$.

Q2. *Soient* $f \in E$ $\varepsilon > 0$, $\eta > 0$ *tels que :*

$$x, x' \in [0,1] \text{ et } |x - x'| < \eta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon$$

Pour tout $x \in [0,1]$, *on note* $I_{x,\varepsilon} = \left\{ k \in \{0,1,\cdots,n\};\ \left| x - \frac{k}{n} \right| < \eta \right\}$ *et* $K_{x,\varepsilon} = \{0,1,\cdots,n\} - I_{x,\varepsilon}$.

(a) Montrer que :

$$\sum_{k \in I_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) < \varepsilon$$

(b) Montrer que :

$$\sum_{k \in K_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) \leq \frac{\|f\|_\infty}{2n \cdot \eta^2}$$

(c) En déduire que la suite $(B_n(f))_{n \geq 1}$ *des polynômes de Bernstein converge uniformément vers f sur* $[0,1]$.

Q3. *Montrer que toute fonction continue sur un intervalle fermé borné* $[a,b]$ *peut être approchée uniformément par une suite de polynômes* $(P_n)_{n \geq 1}$ *avec les conditions :*

$$P_n(x_{i,n}) = f(x_{i,n}) \quad (1 \leq i \leq n)$$

où, pour tout entier $n \geq 1$, $X_n = (x_{1,n}, \cdots, x_{n,n})$ *est une suite de n réels distincts dans* $[a,b]$ *(théorème de Walsh).*

Q4. *Soit* $f \in E$ *continûment dérivable.*

(a) Montrer que :

$$B_n(f)'(x) = \sum_{k=0}^{n-1} f'(\xi_{k,n}) B_{n-1,k}(x)$$

où $\xi_{k,n} \in \left] \frac{k}{n}, \frac{k+1}{n} \right[$, *pour tout k = 0, 1, ⋯, n – 1.*

(b) Montrer que la suite $\left(B_n(f)' \right)_{n \geq 1}$ *converge uniformément vers* f' *sur* $[0,1]$.

Q5. *Montrer que si* $f \in E$ *est de classe* C^p, *avec* $p \geq 0$, *alors la suite* $\left(B_n(f)^{(p)} \right)_{n \geq 1}$ *converge uniformément vers* $f^{(p)}$ *sur* $[0,1]$.

Q6. *On suppose, pour cette question, que* $f \in E$ *convexe.*

(a) Soient $x_1 < x_2$ dans $[0,1]$ et $P_1 \in IR_1[x]$ le polynôme d'interpolation de Lagrange de f défini par :

$$P_1(x_1) = f(x_1),\ P_1(x_2) = f(x_2)$$

Montrer que :

$$\forall x \in [x_1, x_2], \quad P_1(x) - f(x) \geq 0$$

(b) Pour tout $x \in [0,1[$, on pose $t = \dfrac{x}{1-x}$. Montrer que :

$$(1-x)^{-n}\left\{B_{n-1}(f)(x) - B_n(f)(x)\right\} = \sum_{k=1}^{n-1} \alpha_k t^k$$

avec $\alpha_k \geq 0$.

(c) En déduire que la suite $(B_n(f))_{n\geq 1}$ est décroissante.

(c) Déduire de (c) que la suite $(B_n(f))_{n\geq 1}$ converge uniformément sur $[0,1]$.

***Q7.** Soit $f \in E$ de classe C^2.*

Montrer que pour tout $x \in [0,1]$, on a :

$$\underset{n\to+\infty}{Lim}\, n\big(B_n(f)(x) - f(x)\big) = \frac{1}{2}\left(x - x^2\right) f''(x)$$

(théorème de Voronovsky).

Solution

Q1. (a) On a :

$$B_n(f_y)(x) = \sum_{k=0}^{n} e^{y\frac{k}{n}} C_n^k x^k (1-x)^{n-k} = \sum_{k=0}^{n} C_n^k \left(xe^{\frac{y}{n}}\right)^k (1-x)^{n-k}$$

C'est-à-dire :

$$B_n(f_y)(x) = \left(xe^{\frac{y}{n}} + 1 - x\right)^n$$

(b) On a :

$$\varphi_n(x,y) = \sum_{k=0}^{n} e^{\frac{k}{n}y} B_{n,k}(x)$$

et pour tout entier $i \geq 0$:

$$\frac{\partial^i \varphi_n}{\partial y^i}(x,y) = \sum_{k=0}^{n} e^{\frac{k}{n}y}\left(\frac{k}{n}\right)^i B_{n,k}(x)$$

Ce qui donne en faisant $y = 0$:

$$\frac{\partial^i \varphi_n}{\partial y^i}(x,0) = \sum_{k=0}^{n}\left(\frac{k}{n}\right)^i B_{n,k}(x) = B_n(e_i)(x)$$

(c) Pour $n \geq 1$, on a :

$$\begin{cases} \varphi_n(x,y) = \left(xe^{\frac{y}{n}} + 1 - x \right)^n \\ \dfrac{\partial \varphi_n}{\partial y}(x,y) = xe^{\frac{y}{n}} \left(xe^{\frac{y}{n}} + 1 - x \right)^{n-1} \\ \dfrac{\partial^2 \varphi_n}{\partial y^2}(x,y) = \dfrac{x}{n} e^{\frac{y}{n}} \left(xe^{\frac{y}{n}} + 1 - x \right)^{n-1} + \dfrac{n-1}{n} x^2 e^{\frac{2y}{n}} \left(xe^{\frac{y}{n}} + 1 - x \right)^{n-2} \end{cases}$$

Et en faisant $y = 0$, on obtient :

$$\begin{cases} B_n(e_0)(x) = \varphi_n(x,0) = 1 \\ B_n(e_1)(x) = \dfrac{\partial \varphi_n}{\partial y}(x,0) = x \\ B_n(e_2)(x) = \dfrac{\partial^2 \varphi_n}{\partial y^2}(x,0) = \dfrac{x}{n} + \dfrac{n-1}{n} x^2 \end{cases}$$

Q2. La fonction *f* étant uniformément continue sur l'intervalle compact $[0,1]$, on peut trouver, pour tout réel $\varepsilon > 0$, un réel $\eta > 0$ tel que :

$$x, x' \in [0,1] \ \ et \ \ |x - x'| < \eta \ \ \Rightarrow \ \ |f(x) - f(x')| < \varepsilon$$

On remarque, d'autre part, que les fonctions $B_{n,k}$ sont positives sur $[0,1]$.

(a) Par définition de $I_{x,\varepsilon}$, on a :

$$\sum_{k \in I_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) < \varepsilon \sum_{k \in I_{x,\varepsilon}} B_{n,k}(x) < \varepsilon \sum_{k=0}^{n} B_{n,k}(x) = \varepsilon B_n(e_0)(x) = \varepsilon$$

(b) Si $\left| x - \dfrac{k}{n} \right| \geq \eta$, alors $1 \leq \dfrac{1}{\eta^2}\left(x - \dfrac{k}{n} \right)^2$ et :

$$\sum_{k \in K_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) \leq 2\|f\|_\infty \sum_{k \in K_{x,\varepsilon}} B_{n,k}(x) \leq 2\|f\|_\infty \sum_{k \in K_{x,\varepsilon}} \frac{1}{\eta^2}\left(x - \frac{k}{n} \right)^2 B_{n,k}(x)$$

Et en écrivant que :

$$\sum_{k=0}^{n} \left(x - \frac{k}{n} \right)^2 B_{n,k}(x) = x^2 B_n(e_0)(x) - 2x B_n(e_1)(x) + B_n(e_1)(x) = \frac{x(1-x)}{n}$$

on déduit que :

$$\sum_{k \in K_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) \leq \frac{2\|f\|_\infty}{n\eta^2} x(1-x)$$

Avec $x(1-x) \leq \dfrac{1}{4}$ sur $[0,1]$ (le maximum est atteint pour $x = \dfrac{1}{2}$), on a :

$$\sum_{k \in K_{x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| B_{n,k}(x) \leq \frac{\|f\|_\infty}{2n\eta^2}$$

(c) Pour tout $x \in [0,1]$, on a :

$$f(x)-B_n(f)(x)=f(x)\sum_{k=0}^{n}B_{n,k}(x)-\sum_{k=0}^{n}f\left(\frac{k}{n}\right)B_{n,k}(x)=\sum_{k=0}^{n}\left(f(x)-f\left(\frac{k}{n}\right)\right)B_{n,k}(x)$$

et avec ce qui précède, on déduit que :

$$\left|f(x)-B_n(f)(x)\right|\leq\varepsilon+\frac{\|f\|_\infty}{2n\eta^2}$$

Puis, avec $\underset{n\to+\infty}{Lim}\frac{\|f\|_\infty}{2n\eta^2}=0$, on conclut qu'il existe $n_{\varepsilon,\eta}\in IN$ tel que :

$$\forall n\geq n_{\varepsilon,\eta},\quad \forall x\in[0,1],\quad \left|f(x)-B_n(f)(x)\right|\leq 2\varepsilon$$

Ce qui prouve la convergence uniforme de la suite $(B_n(f))_{n\geq1}$ vers f sur $[0,1]$ (le réel $\eta>0$ ne dépend que de $\varepsilon>0$ donné).

Q3. Soit $f\in C^0([a,b];IR)$. La fonction g définie sur $[0,1]$ par :

$$\forall t\in[0,1],\quad g(t)=f(a+t(b-a))$$

est continue et limite uniforme de la suite $(B_n(g))_{n\geq1}$ de ses polynômes de Bernstein. En posant, pour tout $n\geq1$ et tout $x\in[a,b]$:

$$Q_n(x)=B_n(g)\left(\frac{x-a}{b-a}\right)$$

on définit une suite de polynômes qui converge uniformément vers f sur $[a,b]$ (c'est le théorème de Weierstrass).

Soit $\varepsilon>0$ et $Q\in IR[x]$ tel que $\|f-Q\|_\infty<\varepsilon$.

Si $X=(x_1,\cdots,x_n)$ est une suite de points distincts de $[a,b]$, on désigne par $P\in IR_{n-1}[x]$ le polynôme d'interpolation de Lagrange défini par :

$$P(x_k)=f(x_k)-Q(x_k)\quad(1\leq k\leq n)$$

On a alors (problème 33, Q4 (d)) :

$$P(x)=\sum_{k=1}^{n}\left(f(x_k)-Q(x_k)\right)L_k(x)$$

avec $L_k(x)=\dfrac{1}{\prod\limits_{\substack{1\leq j\leq n\\ j\neq k}}(x_k-x_j)}\prod\limits_{\substack{1\leq j\leq n\\ j\neq k}}(x-x_j)$, pour $1\leq k\leq n$.

En posant $M=\sum_{k=1}^{n}\|L_k\|_\infty$, on a :

$$\forall x\in[a,b],\quad |P(x)|\leq\sum_{k=1}^{n}\left|f(x_k)-Q(x_k)\right|\left|L_k(x)\right|\leq\varepsilon\sum_{k=1}^{n}\|L_k\|_\infty=M\varepsilon$$

soit $\|P\|_\infty\leq M\varepsilon$.

On pose alors $\varphi=P+Q$ et on a :

$$\varphi(x_k)=P(x_k)+Q(x_k)=f(x_k)\quad(1\leq k\leq n)$$

$$\|f-\varphi\|_\infty\leq\|f-Q\|_\infty+\|P\|_\infty\leq(1+M)\varepsilon$$

D'où le résultat.

Q4. (a) Pour $k=0,1,\cdots,n$, on a :

$$B_{n,k}{}'(x) = C_n^k\left\{kx^{k-1}(1-x)^{n-k} - (n-k)x^k(1-x)^{n-1-k}\right\}$$

soit :

$$B_{n,k}{}'(x) = n\left\{C_{n-1}^{k-1}x^{k-1}(1-x)^{n-k} - C_{n-1}^{k}x^k(1-x)^{n-1-k}\right\}$$

C'est-à-dire :

$$B_{n,k}{}'(x) = n\left\{B_{n-1,k-1}(x) - B_{n-1,k}(x)\right\}$$

avec les conventions $B_{n-1,-1}(x) = B_{n-1,n}(x) = 0$.

Ce qui donne :

$$B_n(f)'(x) = n\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\left\{B_{n-1,k-1}(x) - B_{n-1,k}(x)\right\} = n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right)B_{n-1,k}(x)$$

Et avec le théorème des accroissements finis, on peut écrire :

$$f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) = \frac{1}{n}f'(\xi_{k,n}) \quad \left(\frac{k}{n} < \xi_{k,n} < \frac{k+1}{n}\right)$$

et :

$$B_n(f)'(x) = \sum_{k=0}^{n-1} f'(\xi_{k,n})B_{n-1,k}(x)$$

(b) On a :

$$\left\|B_n(f)' - f'\right\|_\infty \le \left\|B_n(f)' - B_{n-1}(f')\right\|_\infty + \left\|B_{n-1}(f') - f'\right\|_\infty$$

avec $\underset{n\to+\infty}{Lim}\left\|B_{n-1}(f') - f'\right\|_\infty = 0$, pour $f' \in E$.

Il suffit donc de montrer que $\underset{n\to+\infty}{Lim}\left\|B_n(f)' - B_{n-1}(f')\right\|_\infty = 0$.

Pour tout $x \in [0,1]$, on a :

$$\left|B_n(f)'(x) - B_{n-1}(f')(x)\right| = \left|\sum_{k=0}^{n-1} f'(\xi_{k,n})B_{n-1,k}(x) - \sum_{k=0}^{n-1} f'\left(\frac{k}{n}\right)B_{n-1,k}(x)\right|$$

$$\le \sum_{k=0}^{n-1}\left|f'(\xi_{k,n}) - f'\left(\frac{k}{n}\right)\right|B_{n-1,k}(x)$$

La fonction f' étant uniformément continue sur l'intervalle compact $[0,1]$, on peut trouver, pour tout réel $\varepsilon > 0$, un réel $\eta > 0$ tel que :

$$x, y \in [0,1] \;\; et \;\; |x-y| < \eta \;\Rightarrow\; |f'(x) - f'(y)| < \varepsilon$$

De sorte que pour $n > \frac{1}{\eta}$, on a $\left|\xi_{k,n} - \frac{k}{n}\right| < \frac{1}{n} < \eta$ et :

$$\left|B_n(f)'(x) - B_{n-1}(f')(x)\right| \le \varepsilon\sum_{k=0}^{n-1} B_{n-1,k}(x) = \varepsilon B_{n-1}(e_0)(x) = \varepsilon$$

On a donc $\left\|B_n(f)' - B_{n-1}(f')\right\|_\infty \le \varepsilon$, pour $n > \frac{1}{\eta}$. C'est-à-dire que $\underset{n\to+\infty}{Lim}\left\|B_n(f)' - B_{n-1}(f')\right\|_\infty = 0$. On peut donc conclure que la suite $\left(B_n(f)'\right)_{n\ge1}$ converge uniformément vers f' sur $[0,1]$.

Q5. Un simple raisonnement par récurrence sur p prouve le résultat.
Q6. (a)

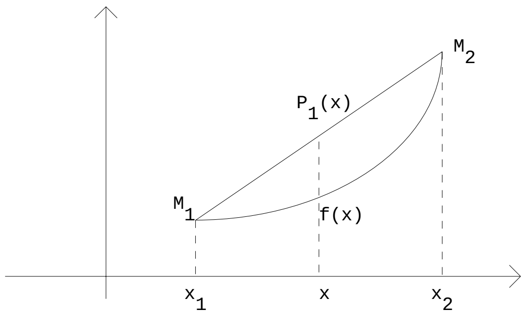

Figure 5.1

Par définition de la convexité, la courbe est sous la corde qui joint les points M_1 et M_2. Ce qui se traduit par :

$$\forall x \in [x_1, x_2], \quad P_1(x) \geq f(x)$$

(b) On a :

$$B_{n-1}(f)(x) - B_n(f)(x) = (1-x)^{n-1}\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)C_{n-1}^k\left(\frac{x}{1-x}\right)^k - (1-x)^n\sum_{k=0}^{n} f\left(\frac{k}{n}\right)C_n^k\left(\frac{x}{1-x}\right)^k$$

Soit :

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) = (1+t)\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)C_{n-1}^k t^k - \sum_{k=0}^{n} f\left(\frac{k}{n}\right)C_n^k t^k$$

$$= \sum_{k=1}^{n-1} f\left(\frac{k}{n-1}\right)C_{n-1}^k t^k + f(0) + \sum_{k=1}^{n-1} f\left(\frac{k-1}{n-1}\right)C_{n-1}^{k-1} t^k + f(1)t^n$$

$$- f(0) - \sum_{k=1}^{n-1} f\left(\frac{k}{n}\right)C_n^k t^k - f(1)t^n$$

Ce qui peut s'écrire aussi :

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) = \sum_{k=1}^{n-1} \alpha_k t^k$$

avec :

$$\alpha_k = C_{n-1}^k f\left(\frac{k}{n-1}\right) + C_{n-1}^{k-1} f\left(\frac{k-1}{n-1}\right) - C_n^k f\left(\frac{k}{n}\right)$$

$$= C_n^k \left\{\frac{n-k}{n} f\left(\frac{k}{n-1}\right) + \frac{k}{n} f\left(\frac{k-1}{n-1}\right) - f\left(\frac{k}{n}\right)\right\}$$

En prenant $x_1 = \dfrac{k-1}{n-1} < x = \dfrac{k}{n} < x_2 = \dfrac{k}{n-1}$ dans (a), on a :

$$P_1(x) = \frac{x-x_2}{x_1-x_2} f(x_1) + \frac{x-x_1}{x_2-x_1} f(x_2) = \frac{n-k}{n} f\left(\frac{k}{n-1}\right) + \frac{k}{n} f\left(\frac{k-1}{n-1}\right) \geq f(x) = f\left(\frac{k}{n}\right)$$

C'est-à-dire que tous les coefficients α_k sont positifs.

(c) De (b), on déduit que :

$$\forall n \geq 2, \quad \forall x \in [0,1[, \quad B_{n-1}(f)(x) \geq B_n(f)(x)$$

Par continuité ces inégalités sont aussi vraies pour $x=1$ (ce qui peut aussi se vérifier directement).

On peut donc conclure que la suite des polynômes de Bernstein de f est décroissante.

(d) Pour tout $x\in[0,1]$, la suite $(B_n(f)(x))_{n\geq 1}$ est décroissante minorée par $\underset{x\in[0,1]}{Inf} f(x)$, elle est donc convergente. Le théorème de Dini nous permet alors de conclure à la convergence uniforme.

Q7. Avec la formule de Taylor à l'ordre 2, on peut écrire, pour tous x, y dans $[0,1]$:

$$f(y)=f(x)+f'(x)(y-x)+\frac{(y-x)^2}{2}f''(x)+\varphi(y)(y-x)^2$$

avec $\underset{y\to x}{Lim}\,\varphi(y)=0$.

En prenant $y=\dfrac{k}{n}$ et en multipliant les deux membres de l'égalité obtenue par $B_{n,k}(x)$, pour tout $k\in\{0,1,\cdots,n\}$, puis en faisant la somme de toutes ces égalités, on déduit que :

$$\begin{aligned}B_n(f)(x)&=f(x)\sum_{k=0}^{n}B_{n,k}(x)+f'(x)\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)B_{n,k}(x)\\&+\frac{f''(x)}{2}\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2B_{n,k}(x)+\sum_{k=0}^{n}\varphi\left(\frac{k}{n}\right)\left(\frac{k}{n}-x\right)^2B_{n,k}(x)\end{aligned}$$

Soit :

$$\begin{aligned}B_n(f)(x)&=f(x)B_n(e_0)(x)+f'(x)\left(B_n(e_1)(x)-xB_n(e_0)(x)\right)\\&+\frac{f''(x)}{2}\left(B_n(e_2)(x)-2xB_n(e_1)(x)+x^2B_n(e_0)(x)\right)+B_n(g)(x)\end{aligned}$$

où on a posé $g(y)=\varphi(y)(y-x)^2$.

Avec les calculs de Q1, (c), cela peut s'écrire :

$$B_n(f)(x)=f(x)+\frac{x(1-x)}{2n}f''(x)+B_n(g)(x)$$

Avec $\underset{y\to x}{Lim}\,\varphi(y)=0$, on déduit que pour tout $\varepsilon>0$, on peut trouver $\eta>0$ tel que :

$$|y-x|<\eta \;\Rightarrow\; |\varphi(y)|<\varepsilon$$

On pose alors :

$$I_{x,\varepsilon}=\left\{k\in\{0,1,\cdots,n\};\ \left|x-\frac{k}{n}\right|<\eta\right\},\quad K_{x,\varepsilon}=\{0,1,\cdots,n\}-I_{x,\varepsilon}$$

et on a :

$$\left|B_n(g)(x)\right| \leq \sum_{k\in I_{x,\varepsilon}} \left|\varphi\left(\frac{k}{n}\right)\right| \left(x-\frac{k}{n}\right)^2 B_{n,k}(x) + \sum_{k\in K_{x,\varepsilon}} \left|\varphi\left(\frac{k}{n}\right)\right| \left(x-\frac{k}{n}\right)^2 B_{n,k}(x)$$

$$\leq \varepsilon \sum_{k=0}^{n} \left(x-\frac{k}{n}\right)^2 B_{n,k}(x) + \|\varphi\|_\infty \sum_{k\in K_{x,\varepsilon}} \left(x-\frac{k}{n}\right)^2 B_{n,k}(x)$$

Soit :

$$\left|B_n(g)(x)\right| \leq \varepsilon \frac{x(1-x)}{n} + \|\varphi\|_\infty \sum_{k\in K_{x,\varepsilon}} \left(x-\frac{k}{n}\right)^2 B_{n,k}(x)$$

$$\leq \frac{\varepsilon}{4n} + \|\varphi\|_\infty \sum_{k\in K_{x,\varepsilon}} \left(x-\frac{k}{n}\right)^2 B_{n,k}(x)$$

Pour $k \in K_{x,\varepsilon}$, on a $\left|x-\frac{k}{n}\right| \geq \eta$ et $1 \leq \frac{1}{\eta^2}\left(x-\frac{k}{n}\right)^2$ de sorte que :

$$\sum_{k\in K_{x,\varepsilon}} \left(x-\frac{k}{n}\right)^2 B_{n,k}(x) \leq \frac{1}{\eta^2} \sum_{k=0}^{n} \left(x-\frac{k}{n}\right)^4 B_{n,k}(x)$$

$$= \frac{1}{\eta^2}\left\{B_n(e_4) - 4xB_n(e_3) + 6x^2 B_n(e_2) - 4x^3 B_n(e_1) + x^4 B_n(e_0)\right\}$$

Et avec Q1, (b), on a :

$$B_n(e_3)(x) = \left(1-\frac{3}{n}+\frac{2}{n^2}\right)x^3 + \left(\frac{3}{n}-\frac{3}{n^2}\right)x^2 + \frac{1}{n^2}x$$

$$B_n(e_4)(x) = \left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)x^4 + \left(\frac{6}{n}-\frac{18}{n^2}+\frac{12}{n^3}\right)x^3 + \left(\frac{7}{n^2}-\frac{7}{n^3}\right)x^2 + \frac{1}{n^3}x$$

Ce qui donne :

$$\sum_{k=0}^{n} \left(x-\frac{k}{n}\right)^4 B_{n,k}(x) = \frac{3}{n^2}x^2(1-x)^2 + \frac{1}{n^3}x(1-x)\left(6x^2-6x+1\right)$$

$$\leq \frac{3}{n^2}\frac{1}{16} + \frac{1}{n^3}\frac{1}{4} \leq \frac{1}{2n^2}$$

sur $[0,1]$.

Soit en définitive :

$$\left|B_n(g)(x)\right| \leq \frac{\varepsilon}{4n} + \|\varphi\|_\infty \frac{1}{2n^2}$$

Et :

$$\left|n\left(B_n(f)(x) - f(x)\right) - \frac{x(1-x)}{2} f''(x)\right| \leq \frac{\varepsilon}{4} + \|\varphi\|_\infty \frac{1}{2n} \leq \varepsilon$$

pour n assez grand. D'où le résultat.

Problème 36 : Théorème de Korovkin

On utilise les notations du problème 35.

Pour f, g dans E, la notation $f \geq g$ *signifie que* $f(x) \geq g(x)$ *pour tout* $x \in [0,1]$.

On désigne par $L(E)$ *l'espace vectoriel des endomorphismes de E.*

On dit que $u \in L(E)$ *est un endomorphisme positif si :*

$$(f \in E, f \geq 0) \Rightarrow (u(f) \geq 0)$$

On rappelle que pour tout entier $j \in IN$, e_j *est la fonction définie sur* $[0,1]$ *par* $e_j(x) = x^j$.

Q1. *Montrer que si* $u \in L(E)$ *est un endomorphisme positif, alors :*

$$\forall g \in E, \quad |u(g)| \leq u(|g|)$$

Q2. *Soient* $u \in L(E)$ *un endomorphisme positif et* $f \in E$.

Montrer que pour tout $x \in [0,1]$, *on a :*

$$|u(f)(x) - f(x)| \leq |u(f - f(x)e_0)(x)| + \|f\|_\infty \|u(e_0) - e_0\|_\infty$$

Q3. *Soit* $f \in E$.

Montrer que pour tout réel $\varepsilon > 0$, *on peut trouver un réel* $\eta > 0$ *tel que :*

$$\forall x, y \in [0,1], \quad |f(y) - f(x)| \leq \varepsilon + 2\frac{\|f\|_\infty}{\eta^2}(y-x)^2$$

Q4. *Soient* $u \in L(E)$ *un endomorphisme positif et* $f \in E$.

Montrer que pour tout réel $\varepsilon > 0$, *on peut trouver un réel* $\eta > 0$ *tel que pour tout* $x \in [0,1]$, *on ait :*

$$|u(f - f(x)e_0)| \leq \varepsilon u(e_0) + 2\frac{\|f\|_\infty}{\eta^2} u\left((e_1 - xe_0)^2\right)$$

Q5. *Soit* $(u_n)_{n\geq 1}$ *une suite d'endomorphismes positifs de E telle que la suite* $(u_n(f))_{n\geq 1}$ *converge uniformément vers f sur* $[0,1]$ *pour toute fonction* $f \in \{e_0, e_1, e_2\}$.

Montrer alors que la suite $(u_n(f))_{n\geq 1}$ *converge uniformément vers f sur* $[0,1]$ *pour toute fonction* $f \in E$ *(théorème de Korovkin).*

Q6. *Montrer que pour toute fonction* $f \in E$, *la suite* $(B_n(f))_{n\geq 1}$ *des polynômes de Bernstein de f converge uniformément vers f sur* $[0,1]$.

Solution

Q1. On a $-|g| \leq g \leq |g|$ et avec la linéarité et la positivité de *u*, on déduit que $-u(|g|) \leq u(g) \leq u(|g|)$, soit $|u(g)| = Max\{-u(g), u(g)\} \leq u(|g|)$.

Q2. Pour $x \in [0,1]$ fixé, on peut écrire pour tout $y \in [0,1]$:

$$u(f)(y)-f(y)=u(f-f(x)e_0)(y)+f(x)u(e_0)(y)-f(y)e_0(y)$$

où $f-f(x)e_0$ est l'application $y\mapsto f(y)-f(x)$.

En particulier, pour $y=x$, on a :

$$\begin{aligned}|u(f)(x)-f(x)| &\le |u(f-f(x)e_0)(x)|+|f(x)||(u(e_0)-e_0)(x)|\\ &\le |u(f-f(x)e_0)(x)|+\|f\|_\infty\|u(e_0)-e_0\|_\infty\end{aligned}$$

Q3. La fonction f étant uniformément continue sur l'intervalle compact $[0,1]$, on peut trouver, pour tout réel $\varepsilon>0$, un réel $\eta>0$ tel que :

$$x,y\in[0,1] \ \ et \ \ |y-x|<\eta \ \Rightarrow \ |f(y)-f(x)|<\varepsilon$$

Pour $x,y\in[0,1]$, on a soit $|f(y)-f(x)|<\varepsilon$, soit $|f(y)-f(x)|\ge\varepsilon$. Dans ce dernier cas, on a $|y-x|\ge\eta$, donc $1\le\dfrac{(y-x)^2}{\eta^2}$ et :

$$0\le|f(y)-f(x)|-\varepsilon<|f(y)-f(x)|\le 2\|f\|_\infty\le 2\frac{\|f\|_\infty}{\eta^2}(y-x)^2$$

D'où le résultat.

Q4. Pour $\varepsilon>0$ donné, on choisit $\eta>0$ comme en Q3. On a alors, avec Q1 :

$$|u(f-f(x)e_0)|\le u(|f-f(x)e_0|)$$

D'autre part, les inégalités de Q3 peuvent s'écrire sous la forme :

$$|f-f(x)e_0|\le\varepsilon\cdot e_0+2\frac{\|f\|_\infty}{\eta^2}(e_1-xe_0)^2$$

Et avec la linéarité et la positivité de u, on déduit que :

$$u(|f-f(x)e_0|)\le\varepsilon u(e_0)+2\frac{\|f\|_\infty}{\eta^2}u((e_1-xe_0)^2)$$

Q5. Pour tout entier $n\ge1$ et tout réel $x\in[0,1]$, on a avec Q2 :

$$|u_n(f)(x)-f(x)|\le|u_n(f-f(x)e_0)(x)|+\|f\|_\infty\|u_n(e_0)-e_0\|_\infty$$

avec $\underset{n\to+\infty}{Lim}\|u_n(e_0)-e_0\|_\infty=0$. Il nous suffit donc de montrer que $\underset{n\to+\infty}{Lim}u_n(f-f(x)e_0)(x)=0$, la convergence étant uniforme sur $[0,1]$.

Pour $\varepsilon>0$, on peut trouver $\eta>0$, ne dépendant que de ε et de f, tel que pour tout entier $n\ge1$ et tout réel $x\in[0,1]$ on ait :

$$|u_n(f-f(x)e_0)|\le\varepsilon u_n(e_0)+2\frac{\|f\|_\infty}{\eta^2}u_n(e_2-2xe_1+x^2e_0)$$

On a alors, en particulier :

$$|u_n(f-f(x)e_0)(x)|\le R_n(x)$$

avec :

$$R_n(x) = \varepsilon u_n(e_0)(x) + 2\frac{\|f\|_\infty}{\eta^2} u_n\left(e_2 - 2xe_1 + x^2 e_0\right)(x)$$
$$= \varepsilon\left(u_n(e_0) - e_0\right)(x) + \varepsilon$$
$$+ 2\frac{\|f\|_\infty}{\eta^2}\left\{\left(u_n(e_2) - e_2\right)(x) - 2x\left(u_n(e_1) - e_1\right)(x) + x^2\left(u_n(e_0) - e_0\right)(x)\right\}$$

puisque $e_2(x) - 2xe_1(x) + x^2 e_0(x) = x^2 - 2x^2 + x^2 = 0$.

Ce qui donne :

$$R_n(x) \le \varepsilon\left\|u_n(e_0) - e_0\right\|_\infty + \varepsilon + 2\frac{\|f\|_\infty}{\eta^2}\left\{\left\|u_n(e_2) - e_2\right\|_\infty + 2\left\|u_n(e_1) - e_1\right\|_\infty + \left\|u_n(e_0) - e_0\right\|_\infty\right\}$$

Et avec $\underset{n \to +\infty}{Lim}\left\|u_n(e_j) - e_j\right\|_\infty = 0$, pour j = 0, 1, 2, on déduit qu'on peut trouver n_0 assez grand tel que :

$$\forall n \ge n_0,\ \forall x \in [0,1],\quad 0 \le R_n(x) \le 3\varepsilon$$

Ce qui achève de prouver la convergence uniforme de la suite $\left(u_n(f)\right)_{n\ge 1}$ vers f sur $[0,1]$.

Q6. Il est clair que pour tout entier $n \ge 1$, l'opérateur B_n est positif.

D'autre part, on a vu (problème 35, Q1 (c)) que :

$$B_n(e_0) = e_0,\quad B_n(e_1) = e_1,\quad B_n(e_2) = e_2 + \frac{1}{n}(e_1 - e_2)$$

On en déduit alors que $\underset{n \to +\infty}{Lim}\left\|B_n(e_j) - e_j\right\|_\infty = 0$, pour j = 0, 1, 2 et on peut utiliser le théorème de Korovkin pour conclure.

Problème 37 : Interpolation spline cubique

Pour tout entier $m \ge 1$, IR^m est muni de la norme définie par :

$$\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in IR^m,\quad \|x\| = Max\left\{|x_i|;\, 1 \le i \le m\right\}$$

Cette norme induit une norme sur l'espace vectoriel des matrices réelles d'ordre m encore notée $\|\ \|$ (voir le problème 1).

On désigne par $IR_m[x]$ l'espace vectoriel des polynômes réels de degré inférieur ou égal à m.

Pour toute fonction continue $g:[a,b] \to IR$, on note :

$$\|g\|_\infty = Sup\left\{|g(x)|;\, x \in [a,b]\right\}$$

Pour tout entier $n \ge 1$ et toute subdivision

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

de l'intervalle $[a,b]$, on appelle « fonction spline cubique » une fonction $s:[a,b] \to IR$ qui vérifie les propriétés suivantes :

$$(1)\quad \begin{cases} (i)\ s \text{ est de classe } C^2 \text{ sur } [a,b] \\ (ii)\ \forall i \in \{0,1,\cdots,n-1\},\ \forall x \in [x_i, x_{i+1}], \\ \qquad s(x) = P_i(x) = \alpha_i (x-x_i)^3 + \beta_i (x-x_i)^2 + \gamma_i (x-x_i) + \delta_i \end{cases}$$

Dans tout le problème, on se donne un intervalle $[a,b]$ *et une fonction* $f:[a,b] \to IR$ *de classe* C^4.

Pour chaque subdivision $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ *de l'intervalle* $[a,b]$, *on note :*

$$h = Max\{h_i = x_{i+1} - x_i\,;\, 0 \le i \le n-1\}$$

Q1. *On se donne une subdivision* $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ *de l'intervalle* $[a,b]$ *et on veut montrer qu'il existe une unique fonction spline cubique s telle que :*

$$(2)\quad \forall i \in \{0,1,\cdots,n\},\ s(x_i) = f(x_i)$$

$$(3)\quad s'(a) = f'(a),\quad s'(b) = f'(b)$$

(a) Montrer que si une fonction spline cubique s vérifie les conditions (2), alors les coefficients $\alpha_i, \beta_i, \gamma_i$ *et* δ_i *qui interviennent dans (1) peuvent s'exprimer en fonction des* $f(x_j)$, h_j *et* $z_j = s''(x_j)$.

(b) Montrer que si une fonction spline cubique s vérifie les conditions (2) et (3), alors le vecteur $z = \begin{pmatrix} z_0 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} s''(x_0) \\ \vdots \\ s''(x_n) \end{pmatrix}$ *est solution d'un système linéaire* $Az = v$,

avec :

$$A = \begin{pmatrix} 2 & \lambda_0 & 0 & 0 & \cdots & 0 \\ \mu_1 & 2 & \lambda_1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & \mu_{n-1} & 2 & \lambda_{n-1} \\ 0 & \cdots & 0 & 0 & \mu_n & 2 \end{pmatrix},\ v = \begin{pmatrix} v_0 \\ \vdots \\ v_n \end{pmatrix}$$

où les coefficients μ_i, λ_i *et* v_i *sont à déterminer.*

(c) Montrer qu'il existe une unique fonction spline cubique s vérifiant (2) et (3).

Q2. *On reprend les notations de Q1 et on pose :*

$$y'' = \begin{pmatrix} y_0'' \\ \vdots \\ y_n'' \end{pmatrix} = \begin{pmatrix} f''(x_0) \\ \vdots \\ f''(x_n) \end{pmatrix},\ r = v - Ay''$$

Montrer que :

(a) $$\|r\| \le \frac{1}{2}\left\|f^{(4)}\right\|_\infty h^2$$

(b) $$\left\|A^{-1}\right\| \le 1$$

(c) $$\|z - y''\| \le \frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2$$

Q3. *Soit* $g:[\alpha,\beta]\to IR$ *de classe* C^2 *et* $h\in IR_1[x]$ *définie par* $h(\alpha)=g(\alpha)$, $h(\beta)=g(\beta)$.

Montrer que :

$$\|g-h\|_\infty \le \frac{(\beta-\alpha)^2}{8}\|g''\|_\infty$$

Q4. *On reprend les notations de Q1 et Q2.*

Pour tout $i\in\{0,1,\cdots,n-1\}$, *on désigne par* $h_i\in IR_1[x]$ *le polynôme de degré 1 défini par* $h_i(x_i)=f''(x_i)$ *et* $h_i(x_{i+1})=f''(x_{i+1})$.

Montrer que :

$$Sup\{|s''(x_i)-h_i(x_i)|;\, x\in[x_i,x_{i+1}]\}\le\frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2 \quad (0\le i\le n-1)$$

Q5. *On reprend les notations de Q1.*

Montrer les inégalités suivantes :

(a) $\|f''-s''\|_\infty \le \dfrac{13}{8}\left\|f^{(4)}\right\|_\infty h^2.$

(b) $\|f'-s'\|_\infty \le \dfrac{13}{8}\left\|f^{(4)}\right\|_\infty h^3.$

(c) $\|f-s\|_\infty \le \dfrac{13}{16}\left\|f^{(4)}\right\|_\infty h^4.$

Q6. *Toujours avec les notations de Q1, monter qu'on a l'inégalité :*

$$\|f'''-s'''\|_\infty \le (3\alpha+1)\left\|f^{(4)}\right\|_\infty h$$

où on a posé $\alpha = Max\left\{\dfrac{h}{h_i};\, 0\le i\le n-1\right\}$.

Solution

Q1. (a) Les conditions $s(x_i)=f(x_i)$ pour tout i = 0, 1, ···, $n-1$ nous donnent immédiatement :

$$\delta_i = f(x_i) \quad (0\le i\le n-1)$$

D'autre part, sur chaque intervalle $[x_i,x_{i+1}]$ on a :

$$s''(x)=P_i''(x)=6\alpha_i(x-x_i)+2\beta_i$$

ce qui donne en prenant respectivement $x=x_i$ et $x=x_{i+1}$:

$$\begin{cases}\beta_i=\dfrac{1}{2}s''(x_i)=\dfrac{1}{2}z_i\\[2mm] \alpha_i=\dfrac{s''(x_{i+1})-s''(x_i)}{6h_i}=\dfrac{z_{i+1}-z_i}{6h_i}\end{cases} \quad (0\le i\le n-1)$$

Enfin les conditions $s(x_{i+1})=P_i(x_{i+1})=f(x_{i+1})$ pour tout i = 0, 1, ···, $n-1$ donnent les égalités $\alpha_i h_i^3+\beta_i h_i^2+\gamma_i h_i+\delta_i=f(x_{i+1})$ qui permettent de déduire les coefficients γ_i. Soit :

$$\gamma_i = \frac{f(x_{i+1}) - f(x_i)}{h_i} - \frac{h_i}{6}\{z_{i+1} + 2z_i\} \quad (0 \le i \le n-1)$$

(b) La fonction s étant continûment dérivable sur l'intervalle $[a,b]$, on doit avoir en particulier les égalités :

$$P_{i-1}'(x_i) = P_i'(x_i) = s'(x_i) \quad (1 \le i \le n-1)$$

Ce qui s'écrit aussi sous la forme :

$$3\alpha_{i-1}h_{i-1}^2 + 2\beta_{i-1}h_{i-1} + \gamma_{i-1} = \gamma_i \quad (1 \le i \le n-1)$$

C'est-à-dire :

$$\frac{z_i - z_{i-1}}{2}h_{i-1} + z_{i-1}h_{i-1} + \frac{f(x_i) - f(x_{i-1})}{h_{i-1}} - \frac{h_{i-1}}{6}(z_i + 2z_{i-1})$$
$$= \frac{f(x_{i+1}) - f(x_i)}{h_i} - \frac{h_i}{6}(z_{i+1} + 2z_i)$$

Ce qui donne le système linéaire de $n-1$ équations à $n+1$ inconnues :

$$\frac{h_{i-1}}{6}z_{i-1} + \frac{h_{i-1} + h_i}{3}z_i + \frac{h_i}{6}z_{i+1} = \frac{f(x_{i+1}) - f(x_i)}{h_i} - \frac{f(x_i) - f(x_{i-1})}{h_{i-1}} \quad (1 \le i \le n-1)$$

En posant, pour $i = 1, 2, \cdots, n-1$:

$$\begin{cases} \mu_i = \dfrac{h_{i-1}}{h_{i-1} + h_i} \\ \lambda_i = \dfrac{h_i}{h_{i-1} + h_i} = 1 - \mu_i \\ v_i = \dfrac{6}{h_{i-1} + h_i}\left\{\dfrac{f(x_{i+1}) - f(x_i)}{h_i} - \dfrac{f(x_i) - f(x_{i-1})}{h_{i-1}}\right\} \end{cases}$$

ce système s'écrit :

$$\mu_i z_{i-1} + 2z_i + \lambda_i z_{i+1} = v_i \quad (1 \le i \le n-1)$$

Les conditions (3) vont nous donner les deux équations manquantes. Ces conditions s'écrivent :

$$\begin{cases} s'(x_0) = P_0'(x_0) = \gamma_0 = f'(x_0) \\ s'(x_n) = P_{n-1}'(x_n) = 3\alpha_{n-1}h_{n-1}^2 + 2\beta_{n-1}h_{n-1} + \gamma_{n-1} = f'(x_n) \end{cases}$$

soit :

$$\begin{cases} \dfrac{f(x_1) - f(x_0)}{h_0} - \dfrac{h_0}{6}(z_1 + 2z_0) = f'(x_0) \\ \dfrac{z_n - z_{n-1}}{2}h_{n-1} + z_{n-1}h_{n-1} + \dfrac{f(x_n) - f(x_{n-1})}{h_{n-1}} - \dfrac{h_{n-1}}{6}(z_n + 2z_{n-1}) = f'(x_n) \end{cases}$$

C'est-à-dire :

$$\begin{cases} 2z_0 + \lambda_0 z_1 = v_0 \\ \mu_n z_{n-1} + 2z_n = v_n \end{cases}$$

où on a posé :

$$\begin{cases} \lambda_0 = 1, \quad v_0 = \dfrac{6}{h_0}\left\{\dfrac{f(x_1)-f(x_0)}{h_0} - f'(x_0)\right\} \\ \mu_n = 1, \quad v_n = \dfrac{6}{h_{n-1}}\left\{f'(x_n) - \dfrac{f(x_n)-f(x_{n-1})}{h_{n-1}}\right\} \end{cases}$$

Avec ces notations, on conclut que le vecteur $z = \begin{pmatrix} z_0 \\ \vdots \\ z_n \end{pmatrix}$ est solution du système linéaire de $n + 1$ équations à $n + 1$ inconnues $Az = v$, avec :

$$A = \begin{pmatrix} 2 & \lambda_0 & 0 & 0 & \cdots & 0 \\ \mu_1 & 2 & \lambda_1 & 0 & \cdots & 0 \\ . & . & . & . & . & . \\ 0 & \cdots & 0 & \mu_{n-1} & 2 & \lambda_{n-1} \\ 0 & \cdots & 0 & 0 & \mu_n & 2 \end{pmatrix}, \; v = \begin{pmatrix} v_0 \\ \vdots \\ v_n \end{pmatrix}$$

(c) Avec les relations $\lambda_i + \mu_i = 1 < 2$ pour tout $i = 0, 1, \cdots, n$ (en posant $\mu_0 = 0 = \lambda_n$), on déduit que la matrice A est à diagonale strictement dominante. Cette matrice est donc inversible (problème 3 Q6 (a)) et le système $Az = v$ admet une unique solution.

On a donc ainsi montré l'existence et l'unicité de la fonction spline s vérifiant les conditions (2) et (3).

Q2. (a) Pour tout $i = 1, 2, \cdots, n-1$, on a :

$$r_i = \frac{1}{h_{i-1} + h_i}\{6u_i - w_i\}$$

où on a posé :

$$\begin{cases} u_i = \dfrac{f(x_{i+1}) - f(x_i)}{h_i} - \dfrac{f(x_i) - f(x_{i-1})}{h_{i-1}} \\ w_i = h_{i-1}f''(x_{i-1}) + 2(h_{i-1} + h_i)f''(x_i) + h_i f''(x_{i+1}) \end{cases}$$

Avec la formule de Taylor à l'ordre 4 en x_i appliquée à f, on peut écrire que :

$$u_i = f'(x_i) + \frac{h_i}{2}f''(x_i) + \frac{h_i^2}{6}f'''(x_i) + \frac{h_i^3}{24}f^{(4)}(\omega_i)$$
$$- f'(x_i) + \frac{h_{i-1}}{2}f''(x_i) - \frac{h_{i-1}^2}{6}f'''(x_i) + \frac{h_{i-1}^3}{24}f^{(4)}(\xi_i)$$

et avec la formule de Taylor à l'ordre 2 en x_i appliquée à f'' :

$$w_i = h_{i-1}\left\{f''(x_i) - h_{i-1}f'''(x_i) + \frac{h_{i-1}^2}{2}f^{(4)}(\psi_i)\right\} + 2(h_{i-1} + h_i)f''(x_i)$$
$$+ h_i\left\{f''(x_i) + h_i f'''(x_i) + \frac{h_i^2}{2}f^{(4)}(\zeta_i)\right\}$$

Ce qui donne :

$$r_i = \frac{1}{h_{i-1}+h_i}\left\{\frac{h_i^3}{4}f^{(4)}(\omega_i)+\frac{h_{i-1}^{\ 3}}{4}f^{(4)}(\xi_i)-\frac{h_{i-1}^{\ 3}}{2}f^{(4)}(\psi_i)-\frac{h_i^3}{2}f^{(4)}(\zeta_i)\right\}$$

On en déduit alors les majorations :

$$|r_i| \le \frac{3}{4}\left\|f^{(4)}\right\|_\infty \frac{h_{i-1}^{\ 3}+h_i^3}{h_{i-1}+h_i} \quad (1\le i\le n-1)$$

Soit :

$$|r_i| \le \frac{3}{4}\left\|f^{(4)}\right\|_\infty \left\{h_{i-1}^{\ 2}-h_{i-1}h_i+h_i^2\right\} \quad (1\le i\le n-1)$$

Et :

$$|r_i| \le \frac{3}{4}\left\|f^{(4)}\right\|_\infty \left\{h_{i-1}^{\ 2}+h_i^2\right\} \le \frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2 \quad (1\le i\le n-1)$$

Pour $i = 0$, on a :

$$r_0 = \frac{1}{h_0}\{6u_0 - w_0\}$$

où on a posé :

$$\begin{cases} u_0 = \dfrac{f(x_1)-f(x_0)}{h_0} - f'(x_0) \\ w_0 = h_0\{2f''(x_0)+f''(x_1)\} \end{cases}$$

Avec la formule de Taylor à l'ordre 4 en x_0 appliquée à f, on a :

$$u_i = \frac{h_0}{2}f''(x_0)+\frac{h_0^{\ 2}}{6}f'''(x_0)+\frac{h_0^{\ 3}}{24}f^{(4)}(\omega_0)$$

et avec la formule de Taylor à l'ordre 2 en x_0 appliquée à f'' :

$$w_0 = h_0\left\{2f''(x_0)+f''(x_0)+h_0f'''(x_0)+\frac{h_0^{\ 2}}{2}f^{(4)}(\psi_0)\right\}$$

Ce qui donne :

$$r_0 = \frac{h_0^{\ 2}}{4}f^{(4)}(\omega_0)-\frac{h_0^{\ 2}}{2}f^{(4)}(\psi_0)$$

Et :

$$|r_0| \le \frac{3}{4}\left\|f^{(4)}\right\|_\infty h^2 < \frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2$$

De manière analogue, pour $i = n$, on a :

$$|r_n| \le \frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2$$

On peut donc conclure que :

$$\|r\| = Max\{|r_i|; 0\le i\le n\} \le \frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2$$

(b) Si $w = Az$, on a alors :

$$w_i = \mu_i z_{i-1} + 2z_i + \lambda_i z_{i+1} \quad (0\le i\le n)$$

en posant $\mu_0 = z_{-1} = \lambda_n = z_{n+1} = 0$.

Pour $i \in \{0,1,\cdots,n\}$ tel que $\|z\| = |z_i|$, on a :

$$\|w\| \ge |w_i| \ge 2|z_i| - \mu_i|z_{i-1}| - \lambda_i|z_{i+1}| \ge (2-\mu_i-\lambda_i)|z_i| = |z_i| = \|z\|$$

puisque $\lambda_i + \mu_i = 1$.

Ce qui peut s'écrire $\|z\| = \|A^{-1}w\| \leq \|w\|$, encore équivalent à $\|A^{-1}\| \leq 1$.

(c) En considérant que $v = Az$, on peut écrire avec (a) et (b) que :

$$\|z - y''\| = \|A^{-1}(v - Ay'')\| = \|A^{-1}r\| \leq \|A^{-1}\|\|r\| \leq \|r\|$$

et $\|z - y''\| \leq \dfrac{3}{2}\|f^{(4)}\|_\infty h^2$.

Q3. On pose $\varphi = g - h$ et on définit ainsi une fonction de classes C^2 sur $[\alpha, \beta]$ avec $\varphi'' = g''$.

La fonction φ est de classes C^1 sur le compact $[\alpha, \beta]$, elle admet donc un minimum et un maximum en des points ξ et η avec $\varphi'(\xi) = 0 = \varphi'(\eta)$. On a donc :

$$\|\varphi\|_\infty = Sup\{|\varphi(x)|; x \in [\alpha, \beta]\} = |\varphi(\zeta)|$$

avec $\zeta \in \{\xi, \eta\} \subset [\alpha, \beta]$ et $\varphi'(\zeta) = 0$.

Si $\zeta \in \left[\alpha, \dfrac{\alpha + \beta}{2}\right]$, on peut alors écrire avec la formule de Taylor à l'ordre 2 :

$$0 = \varphi(\alpha) = \varphi(\zeta) + (\alpha - \zeta)\varphi'(\zeta) + \frac{(\alpha - \zeta)^2}{2}\varphi''(\zeta_1)$$

Soit $\varphi(\zeta) = -\dfrac{(\alpha - \zeta)^2}{2} g''(\zeta_1)$ et :

$$|\varphi(\zeta)| \leq \frac{(\alpha - \zeta)^2}{2}\|g''\|_\infty \leq \frac{1}{2}\left(\frac{\beta - \alpha}{2}\right)^2 \|g''\|_\infty$$

De façon analogue, si $\zeta \in \left[\dfrac{\alpha + \beta}{2}, \beta\right]$, on écrit alors :

$$0 = \varphi(\beta) = \varphi(\zeta) + (\beta - \zeta)\varphi'(\zeta) + \frac{(\beta - \zeta)^2}{2}\varphi''(\zeta_2)$$

et on aboutit encore à $|\varphi(\zeta)| \leq \dfrac{1}{2}\left(\dfrac{\beta - \alpha}{2}\right)^2 \|g''\|_\infty$.

On a donc bien :

$$\|g - h\|_\infty \leq \frac{(\beta - \alpha)^2}{8}\|g''\|_\infty$$

Q4. La fonction $s'' - h_i$ étant affine, on a :

$$Sup\{|s''(x) - h_i(x)|; x \in [x_i, x_{i+1}]\} = Max\{|s''(x_i) - h_i(x_i)|, |s''(x_{i+1}) - h_i(x_{i+1})|\}$$

Ce qui donne, avec les conditions $h_i(x_j) = f''(x_j)$ pour $j = i, i + 1$:

$$Sup\{|s''(x) - h_i(x)|; x \in [x_i, x_{i+1}]\} \leq Max\{|s''(x_j) - f''(x_j)|; 0 \leq i \leq n\} = \|z - y''\|$$

Soit avec la majoration de Q2 (c) :

$$Sup\{|s''(x) - h_i(x)|; x \in [x_i, x_{i+1}]\} \leq \frac{3}{2}\|f^{(4)}\|_\infty h^2$$

Q5. (a) En utilisant les notations de Q4, on peut écrire pour tout $x \in [x_i, x_{i+1}]$:

$$|f''(x) - s''(x)| \leq |f''(x) - h_i(x)| + |h_i(x) - s''(x)|$$

Avec Q3 et Q4, on déduit alors que :

$$\left|f''(x)-s''(x)\right|\le\left\|f^{(4)}\right\|_\infty\frac{h_i^2}{8}+\frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2$$

On peut donc conclure que :

$$\left\|f''-s''\right\|_\infty\le\left(\frac{1}{8}+\frac{3}{2}\right)\left\|f^{(4)}\right\|_\infty h^2=\frac{13}{8}\left\|f^{(4)}\right\|_\infty h^2$$

(b) Avec les conditions $s(x_i)=f(x_i)$ pour $0\le i\le n$ et en utilisant le théorème de Rolle, on déduit qu'on peut trouver des points $\xi_{i+1}\in\left]x_i,x_{i+1}\right[$ tels que $s'(\xi_{i+1})=f'(\xi_{i+1})$ pour $0\le i\le n-1$.

D'autre part, en posant $\xi_0=a$ et $\xi_{n+1}=b$, on peut écrire les égalités $s'(\xi_i)=f'(\xi_i)$ pour $0\le i\le n+1$.

Pour tout réel $x\in[a,b]$, on peut trouver un indice $i\in\{0,1,\cdots,n+1\}$ tel que $|x-x_i|\le h$. On écrit alors que :

$$\left|f'(x)-s'(x)\right|=\left|\int_{\xi_i}^x\left(f''(t)-s''(t)\right)dt\right|\le h\left\|f''-s''\right\|_\infty$$

On conclut donc que :

$$\left\|f'-s'\right\|_\infty\le\frac{13}{8}\left\|f^{(4)}\right\|_\infty h^3$$

(c) Pour $0\le i\le n$, on a $s(x_i)=f(x_i)$ et pour tout réel $x\in[a,b]$, on peut trouver un indice $i\in\{0,1,\cdots,n-1\}$ tel que $|x-x_i|\le\dfrac{h}{2}$. On peut alors écrire que :

$$\left|f(x)-s(x)\right|=\left|\int_{x_i}^x\left(f'(t)-s'(t)\right)dt\right|\le\frac{h}{2}\left\|f'-s'\right\|_\infty$$

de sorte que :

$$\left\|f-s\right\|_\infty\le\frac{13}{16}\left\|f^{(4)}\right\|_\infty h^4$$

Q6. Pour $x\in[x_i,x_{i+1}]$, avec $0\le i\le n-1$, on a :

$$f'''(x)-s'''(x)=f'''(x)-6\alpha_i=f'''(x)-\frac{s''(x_{i+1})-s''(x_i)}{h_i}$$

Ce qui peut s'écrire :

$$f'''(x)-s'''(x)=f'''(x)-\frac{s''(x_{i+1})-f''(x_{i+1})}{h_i}+\frac{s''(x_i)-f''(x_i)}{h_i}-\frac{f''(x_{i+1})-f''(x_i)}{h_i}$$

Avec la formule de Taylor à l'ordre 2 en x appliquée à f'', on a :

$$\frac{f''(x_{i+1})-f''(x_i)}{h_i}=\frac{f''(x_{i+1})-f''(x)}{h_i}-\frac{f''(x_i)-f''(x)}{h_i}$$

$$=\frac{x_{i+1}-x}{h_i}f'''(x)+\frac{(x_{i+1}-x)^2}{2h_i}f^{(4)}(\xi_1)-\frac{x_i-x}{h_i}f'''(x)-\frac{(x_i-x)^2}{2h_i}f^{(4)}(\xi_2)$$

$$=f'''(x)+\frac{(x_{i+1}-x)^2}{2h_i}f^{(4)}(\xi_1)-\frac{(x_i-x)^2}{2h_i}f^{(4)}(\xi_2)$$

Ce qui donne :

$$f'''(x)-s'''(x)=-\frac{s''(x_{i+1})-f''(x_{i+1})}{h_i}+\frac{s''(x_i)-f''(x_i)}{h_i}$$

$$-\frac{(x_{i+1}-x)^2}{2h_i}f^{(4)}(\xi_1)+\frac{(x_i-x)^2}{2h_i}f^{(4)}(\xi_2)$$

Et avec les notations de Q2 :

$$\left|f'''(x)-s'''(x)\right|\le\frac{2}{h_i}\|y''-z\|+h_i\left\|f^{(4)}\right\|_\infty\le\frac{2}{h_i}\frac{3}{2}\left\|f^{(4)}\right\|_\infty h^2+\left\|f^{(4)}\right\|_\infty h$$

Ce qui donne :

$$\|f'''-s'''\|_\infty\le(3\alpha+1)\left\|f^{(4)}\right\|_\infty h$$

Problème 38 : Fonctions splines d'ajustement. Agrégation 1974, extrait

Soient a et b deux réels tels que a < b. On désigne par :
1°) $H^0[a,b]$ (ou simplement H^0) l'ensemble des (classes de) fonctions y à valeurs réelles, définies sur $[a,b]$ et de carré sommable sur $[a,b]$.

Muni des lois usuelles d'addition et de multiplication par un scalaire, H^0 est un espace vectoriel sur IR. Muni, de plus, du produit scalaire habituel :

$$\langle y_1|y_2\rangle_0=\int_a^b y_1(t)y_2(t)dt$$

et de la norme associée $\|y\|_0=\left[\langle y|y\rangle_0\right]^{\frac{1}{2}}$, H^0 est un espace de Hilbert réel.

2°) $H^2[a,b]$ (ou simplement H^2) l'ensemble des fonctions x définies sur $[a,b]$, à valeurs réelles, admettant sur $[a,b]$:

— une dérivée première absolument continue,

— une dérivée seconde définie presque partout sur $[a,b]$ et de carré sommable.

Muni comme en 1°), des lois usuelles d'addition et de multiplication par un scalaire, H^2 est un espace vectoriel sur IR.

On munit H^2 du produit scalaire :

$$\langle x_1|x_2\rangle_2=\int_a^b\left[x_1(t)x_2(t)+x_1'(t)x_2'(t)+x_1''(t)x_2''(t)\right]dt$$

et de la norme associée $\|x\|_2 = \left[\langle x \mid x \rangle_2\right]^{\frac{1}{2}}$.

H^2 *est ainsi un espace de Hilbert réel.*

3°) Soient alors, pour toute la suite, n + 2 valeurs réelles $t_0, t_1, \cdots, t_n, t_{n+1}$ *telles que :*

$$t_0 = a < t_1 < t_2 < \cdots, t_n < b = t_{n+1}$$

(on suppose que $n \geq 2$*).*

On considère l'ensemble S des fonctions s, définies sur $[a,b]$*, à valeurs réelles, telles que :*

(a) Dans tout intervalle $[t_i, t_{i+1}]$ *(i = 1, ···, n – 1), s coïncide avec un polynôme de degré* ≤ 3 *:*

$$\forall t \in [t_i, t_{i+1}], \quad s(t) = \gamma_i^3 t^3 + \gamma_i^2 t^2 + \gamma_i^1 t + \gamma_i^0$$

(b) Dans chacun des deux intervalles $[a, t_1]$ *et* $[t_n, b]$*, s coïncide avec un polynôme de degré* ≤ 1 *:*

$$\forall t \in [a, t_1], \quad s(t) = \gamma_0^1 t + \gamma_0^0$$

$$\forall t \in [t_n, b], \quad s(t) = \gamma_n^1 t + \gamma_n^0$$

(les γ_i^j *sont des réels).*

(c) s, s′, s″ sont définies et continues sur $[a,b]$.

Q1. *Montrer que S est un sous espace vectoriel de* H^2*, de dimension finie.*

Pour tout réel α *on pose* $[\alpha]_+ = Sup(\alpha, 0)$.

Q2. *Montrer qu'un élément s de* H^2 *appartient à S si et seulement si :*

$$\forall t \in [a,b], \quad s(t) = \alpha_0 + \alpha_1 t + \sum_{i=1}^{n} \frac{\delta_i}{3!}\left[(t - t_i)^3\right]_+$$

où $\alpha_0, \alpha_1, \delta_i$ *(i = 1, ···, n) sont des réels tels que :*

$$\sum_{i=1}^{n} \delta_i = 0, \quad \sum_{i=1}^{n} \delta_i t_i = 0$$

En déduire la dimension de S dans H^2.

On désigne par B, l'application de H^2 *dans* H^0 *qui à tout x de* H^2 *fait correspondre sa dérivée seconde :*

$$B(x) = x''$$

On note c l'application de H^2 *dans* $I\!R^n$ *qui à tout x de* H^2 *fait correspondre le vecteur des valeurs de x en* $t_1, \cdots, t_n$ *:*

$$c(x) = \left(x(t_1), x(t_2), \cdots, x(t_n)\right)$$

On note d l'application de S dans $I\!R^n$ *qui à tout s de S fait correspondre le vecteur de* $I\!R^n$ *:*

$$d(s) = (\delta_1, \delta_2, \cdots, \delta_n)$$

où δ_i *(i = 1, ···, n) représente le saut de la dérivée troisième de s en* t_i.

On remarquera à ce propos que, s coïncidant avec un polynôme de degré inférieur ou égal à 3 sur tout intervalle $[t_i, t_{i+1}]$*, la dérivée troisième de s est définie et constante sur* $]t_i, t_{i+1}[$ *(i = 0, 1, ···, n).*

On note $\langle \mid \rangle$ *le produit scalaire usuel sur* $I\!R^n$ *:*

$$\forall u=(u_1,\cdots,u_n)\in IR^n,\ \forall v=(v_1,\cdots,v_n)\in IR^n,\ \langle u|v\rangle=\sum_{i=1}^{n}u_iv_i$$

On note $\| \ \|$ la norme associée sur IR^n :

$$\forall u=(u_1,\cdots,u_n)\in IR^n,\ \|u\|=\left[\sum_{i=1}^{n}{u_i}^2\right]^{\frac{1}{2}}=\langle u|u\rangle^{\frac{1}{2}}$$

Q3. *Montrer que B, c et d sont linéaires et continues.*
Sont-elles surjectives?

Q4. *Montrer que pour tout x dans H^2 et tout s dans S on a :*

$$\langle B(x)|B(s)\rangle_0=\langle c(x)|d(s)\rangle$$

Q5. *Montrer que pour tout $z=(z_1,\cdots,z_n)\in IR^n$, il existe un élément σ unique dans S tel que :*

$$c(\sigma)+d(\sigma)=z$$

Q6. *Montrer que l'élément σ défini en Q5 vérifie :*

$$\|B(\sigma)\|_0^{\ 2}+\|c(\sigma)-z\|^2=\underset{x\in H^2}{Min}\left(\|B(x)\|_0^{\ 2}+\|c(x)-z\|^2\right)$$

Soit Z_0 le sous-espace vectoriel de IR^n, de dimension n – 2 formé des vecteurs $z=(z_1,\cdots,z_n)$ tels que :

$$\sum_{i=1}^{n}z_i=0 \ et \ \sum_{i=1}^{n}z_it_i=0$$

On désigne par $(\gamma_j)_{1\le j\le n-2}$ une base de Z_0.

On pourra par exemple prendre pour γ_j le vecteur de composantes γ_{ji} suivantes :

$$\gamma_{ji}=\begin{cases}\dfrac{1}{(t_j-t_{j+1})(t_j-t_{j+2})} & si\ i=j\\[2mm] \dfrac{1}{(t_{j+1}-t_j)(t_{j+1}-t_{j+2})} & si\ i=j+1\\[2mm] \dfrac{1}{(t_{j+2}-t_j)(t_{j+2}-t_{j+1})} & si\ i=j+2\\[2mm] 0 & si\ i<j\ ou\ si\ i>j+2\end{cases}$$

On rappelle que pour tout $x\in H^2$ on a (développement de Taylor) :

$$x(t)=x(a)+(t-a)x'(a)+\int_a^b[(t-\xi)]_+x''(\xi)d\xi$$

Q7. *Montrer que pour tout j (j = 1, 2, ···, n – 2) il existe une fonction $\beta_j\in H^0$ telle que pour tout $x\in H^2$ on ait :*

$$\langle\gamma_j|c(x)\rangle=\langle\beta_j|B(x)\rangle_0$$

Expliciter les fonctions β_j pour le choix de vecteurs γ_j proposé ci-dessus (montrer en particulier que β_j vérifie alors $\beta_j(t)=0$, lorsque $t\notin[t_j,t_{j+2}]$).

Soit s arbitraire. D'après Q2, il existe des réels μ_j (j = 1, 2, ⋯, n – 2) tels que :

$$d(s) = \sum_{j=1}^{n-2} \mu_j \gamma_j$$

Q8. *Montrer que l'on a alors :*

$$B(s) = \sum_{j=1}^{n-2} \mu_j \beta_j$$

On se propose de déterminer numériquement l'élément $\sigma \in S$ qui vérifie $c(\sigma) + d(\sigma) = z$. Pour cela on calcule d'abord $B(\sigma) = \sigma'' \in H^0$ et $c(\sigma) = (\sigma(t_1), \sigma(t_2), \cdots, \sigma(t_n)) \in IR^n$.

Q9. *Montrer que l'on a :*

$$B(\sigma) = \sum_{j=1}^{n-2} \mu_j \beta_j$$

$$c(\sigma) = z - \sum_{j=1}^{n-2} \mu_j \gamma_j$$

où les μ_j (j = 1, 2, ⋯, n – 2) sont solutions du système d'équations linéaires :

$$\sum_{j=1}^{n-2} \omega_{kj} \mu_j = \eta_k \quad (1 \le k \le n-2)$$

avec

$$\omega_{kj} = \langle \beta_k | \beta_j \rangle_0 + \langle \gamma_k | \gamma_j \rangle$$

et

$$\eta_k = \langle z | \gamma_k \rangle$$

Solution

Rappels — Une fonction $f:[a,b] \to IR$ est dite absolument continue si :

$$\forall \varepsilon > 0, \exists \eta > 0; \quad \sum_{i=1}^{n} (\beta_i - \alpha_i) < \eta \Rightarrow \sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| < \varepsilon$$

pour toute famille de segments disjoints $]\alpha_i, \beta_i[$ dans $[a,b]$.

Une fonction absolument continue est uniformément continue.

Si $\varphi:[a,b] \to IR$ est Lebesgue–intégrable, alors la fonction *f* définie par

$$\forall t \in [a,b], \quad f(t) = \int_a^t \varphi(u) du$$

est absolument continue. De plus elle admet une dérivée définie presque partout par :

$$pp\, t \in [a,b], \quad f'(t) = \varphi(t)$$

Q1. (a) Montrons tout d'abord que $S \subset H^2$.

Une fonction $s \in S$ est de classe C^2 sur $[a,b]$, en particulier sa dérivée seconde s'' est continue et elle est donc de carré intégrable sur $[a,b]$.

Pour toute suite $\left(\left]\alpha_i,\beta_i\right[\right)_{1\le i\le n}$ d'intervalles ouverts deux à deux disjoints dans $[a,b]$, on a :

$$\sum_{i=1}^{n}\left|s'(\beta_i)-s'(\alpha_i)\right|=\sum_{i=1}^{n}\left|\int_{\alpha_i}^{\beta_i}s''(t)dt\right|\le \underset{t\in[a,b]}{Sup}\left|s''(t)\right|\sum_{i=1}^{n}(\beta_i-\alpha_i)<\varepsilon$$

dés que $\sum_{i=1}^{n}(\beta_i-\alpha_i)<\eta=\dfrac{\varepsilon}{\underset{t\in[a,b]}{Sup}\left|s''(t)\right|}$.

Donc s' est absolument continue sur $[a,b]$.

On a donc bien $S\subset H^2$.

(b) Montrons que *S* est un sous-espace vectoriel de dimension finie de H^2.

L'ensemble des fonctions de classe C^2 et l'ensemble des polynômes réels de degré inférieur ou égal à *p* (donné dans *IN*) étant des espaces vectoriels, on en déduit que S est un espace vectoriel.

S est un sous-espace vectoriel de l'espace vectoriel *P* des fonctions polynomiales par morceaux qui coïncident avec un polynôme de degré inférieur ou égal à 1 sur $[a,t_1]$ et $[t_n,b]$ et avec un polynôme de degré inférieur ou égal à 3 sur chaque intervalle $[t_i,t_{i+1}]$ (i = 1, ···, n – 1) (on ne s'occupe pas des conditions de continuité de s, s', s'' pour les fonctions *s* de P).

Une fonction de *P* étant uniquement déterminée par les coefficients γ_i^j définis en 3°) (a) et (b) de l'énoncé, on déduit que $Dim(P)=2+4(n-1)+2=4n$ et :

$$Dim(S)\le Dim(P)=4n$$

En conclusion, S est un sous espace vectoriel de dimension finie de H^2.

Q2. Soient $(\alpha_0,\alpha_1,\delta_1,\cdots,\delta_n)\in IR^{n+2}$ et $g:[a,b]\to IR$ définie par :

$$\forall t\in[a,b],\quad g(t)=\alpha_0+\alpha_1 t+\sum_{i=1}^{n}\frac{\delta_i}{3!}\left[(t-t_i)^3\right]_+$$

Une telle fonction est définie de manière équivalente par :

$$(*)\quad \forall t\in[a,b],\quad \begin{cases} s(t)=\alpha_0+\alpha_1 t \text{ si } t\in[t_0,t_1] \\ s(t)=\alpha_0+\alpha_1 t+\displaystyle\sum_{k=1}^{i}\frac{\delta_k}{3!}(t-t_k)^3 \text{ si } t\in[t_i,t_{i+1}]\quad (1\le i\le n)\end{cases}$$

(a) Montrons tout d'abord que pour toute fonction spline $s\in S$, il existe un unique vecteur $(\alpha_0,\alpha_1,\delta_1,\cdots,\delta_n)\in IR^{n+2}$ vérifiant les conditions $\sum_{i=1}^{n}\delta_i=0$ et $\sum_{i=1}^{n}\delta_i t_i=0$ et tel que *s* soit définie par (*).

Pour $t\in[t_0,t_1]$, on a $s(t)=\gamma_0^1 t+\gamma_0^0$, ce qui détermine le couple de réels (α_0,α_1) de manière unique. Soit :

$$(\alpha_0,\alpha_1)=(\gamma_0^0,\gamma_0^1)$$

Pour $t\in[t_1,t_2]$, on doit avoir :

$$s(t)=\gamma_1^3 t^3+\gamma_1^2 t^2+\gamma_1^1 t+\gamma_1^0=\alpha_0+\alpha_1 t+\frac{\delta_1}{3!}(t-t_1)^3$$

Ce qui donne en identifiant les coefficients de t^k pour k = 0, 1, 2, 3 :

$$\begin{cases} \alpha_0 - \dfrac{\delta_1}{3!} t_1^{\ 3} = \gamma_1^0 & (i) \\ \alpha_1 + \dfrac{\delta_1}{2} t_1^{\ 2} = \gamma_1^1 & (ii) \\ -\dfrac{\delta_1}{2} t_1 = \gamma_1^2 & (iii) \\ \dfrac{\delta_1}{3!} = \gamma_1^3 & (iv) \end{cases}$$

La dernière équation donne $\delta_1 = 3!\,\gamma_1^3$.

Les autres équations traduisent en fait les conditions de continuité de s, s', s'' en t_1.

En effet, en notant pour tout $i = 0, 1, \cdots, n$ s_i l'expression polynomiale de s sur l'intervalle $[t_i, t_{i+1}]$, on a :

$$s^{(k)}(t_1) = s_0^{(k)}(t_1) = s_1^{(k)}(t_1) \quad (k = 0,1,2)$$

Ce qui donne pour $k = 2$:

$$0 = 3!\,\gamma_1^3 t_1 + 2\gamma_1^2$$

c'est-à-dire $-\dfrac{3!\,\gamma_1^3}{2} t_1 = \gamma_1^2$, ou encore $-\dfrac{\delta_1}{2} t_1 = \gamma_1^2$, soit l'équation ($iii$).

Les conditions (ii) et (i) sont équivalentes respectivement à $s_0'(t_1) = s_1'(t_1)$ et $s_0(t_1) = s_1(t_1)$.

Montrons par récurrence que pour $i \in \{2, \cdots, n+1\}$, il existe des coefficients uniques $\delta_1, \cdots, \delta_{i-1}$ tels que :

$$\forall t \in [t_0, t_i], \quad s(t) = \alpha_0 + \alpha_1 t + \sum_{k=1}^{i-1} \frac{\delta_k}{3!} \left[(t - t_k)^3\right]_+$$

avec :

$$\begin{cases} (\alpha_0, \alpha_1) = (\gamma_0^0, \gamma_0^1) \\ \delta_k = 3!(\gamma_k^3 - \gamma_{k-1}^3) \quad (1 \le k \le i-1) \end{cases}$$

où on a posé $\gamma_0^3 = \gamma_n^3 = 0$.

On vient de vérifier cette propriété pour $i = 2$ et il s'agit de montrer qu'elle est vraie pour $i + 1$ si on suppose qu'elle est vraie pour $i \in \{2, \cdots, n\}$.

Avec l'hypothèse de récurrence, on a :

$$\forall t \in [t_{i-1}, t_i], \quad s(t) = \gamma_{i-1}^3 t^3 + \gamma_{i-1}^2 t^2 + \gamma_{i-1}^1 t + \gamma_{i-1}^0 = \alpha_0 + \alpha_1 t + \sum_{k=1}^{i-1} \frac{\delta_k}{3!} (t - t_k)^3$$

Ce qui donne en identifiant les coefficients de t^k pour $k = 0, 1, 2, 3$:

$$(**)\quad \begin{cases} \alpha_0 - \dfrac{1}{3!}\displaystyle\sum_{k=1}^{i-1} \delta_k t_k{}^3 = \gamma_{i-1}^0 & (i) \\ \alpha_1 + \dfrac{1}{2}\displaystyle\sum_{k=1}^{i-1} \delta_k t_k{}^2 = \gamma_{i-1}^1 & (ii) \\ -\dfrac{1}{2}\displaystyle\sum_{k=1}^{i-1} \delta_k t_k = \gamma_{i-1}^2 & (iii) \\ \dfrac{1}{3!}\displaystyle\sum_{k=1}^{i-1} \delta_k = \gamma_{i-1}^3 & (iv) \end{cases}$$

Si la propriété est vraie pour $i + 1$, alors on a nécessairement :

$$\begin{cases} \alpha_0 - \dfrac{1}{3!}\displaystyle\sum_{k=1}^{i} \delta_k t_k{}^3 = \gamma_i^0 & (i) \\ \alpha_1 + \dfrac{1}{2}\displaystyle\sum_{k=1}^{i} \delta_k t_k{}^2 = \gamma_i^1 & (ii) \\ -\dfrac{1}{2}\displaystyle\sum_{k=1}^{i} \delta_k t_k = \gamma_i^2 & (iii) \\ \dfrac{1}{3!}\displaystyle\sum_{k=1}^{i} \delta_k = \gamma_i^3 & (iv) \end{cases}$$

En utilisant (**) (*iv*), l'équation (*iv*) ci-dessus donne alors :

$$(***)\quad \delta_i = 3!\,\gamma_i^3 - \sum_{k=1}^{i-1} \delta_k = 3!\left(\gamma_i^3 - \gamma_{i-1}^3\right)$$

La condition $s''(t_i) = s_{i-1}''(t_i) = s_i''(t_i)$ se traduit par :

$$3!\,\gamma_{i-1}^3 t_i + 2\gamma_{i-1}^2 = 3!\,\gamma_i^3 t_i + 2\gamma_i^2$$

C'est-à-dire $-\dfrac{3!}{2}\left(\gamma_i^3 - \gamma_{i-1}^3\right)t_i + \gamma_{i-1}^2 = \gamma_i^2$. Ce qui peut s'écrire avec (**) (*ii*) et (***) :

$$-\frac{\delta_i}{2} t_i - \frac{1}{2}\sum_{k=1}^{i-1} \delta_k t_k = \gamma_i^2$$

Ce qui donne bien l'égalité (*iii*).

De façon analogue, on voit que les conditions $s_{i-1}'(t_i) = s_i'(t_i)$ et $s_{i-1}(t_i) = s_i(t_i)$ donnent les égalités (*ii*) et (*i*).

En particulier, pour $i = n$, on a $\gamma_n^3 = \gamma_n^2 = 0$ et les conditions (*iv*) et (*iii*) s'écrivent respectivement $\sum_{k=1}^{n} \delta_k = 0$ et $\sum_{k=1}^{n} \delta_k t_k = 0$.

(b) Réciproquement, on doit montrer que si $s \in H^2$ s'écrit sous la forme :

$$s(t) = \alpha_0 + \alpha_1 t + \sum_{i=1}^{n} \frac{\delta_i}{3!}\left[(t - t_i)^3\right]_+$$

avec $\sum_{k=1}^{n} \delta_k = 0$ et $\sum_{k=1}^{n} \delta_k t_k = 0$, alors $s \in S$.

Sur $[t_0, t_1]$ s est bien une fonction affine.

Sur $[t_i, t_{i+1}]$, pour i = 1, ···, n, la fonction s est polynomiale de degré inférieur ou égal à 3.

En particulier, pour $i = n$, les coefficients de t^3 et t^2 sont donnés respectivement par $\gamma_n^3 = \frac{1}{3!}\sum_{k=1}^{n}\delta_k = 0$ et $\gamma_n^2 = -\frac{1}{2}\sum_{k=1}^{n}\delta_k t_k = 0$. C'est-à-dire que s est bien une fonction affine sur $[t_n, t_{n+1}]$.

Comme s est dans H^2, elle est de classe C^1.

Il est clair que s est de classe C^2 sur chaque intervalle ouvert $]t_i, t_{i+1}[$, pour i = 0, 1, ···, n, avec :

$$\forall t \in]t_i, t_{i+1}[, \quad s''(t) = \sum_{k=1}^{i}\delta_k(t - t_k) = s_i''(t)$$

On a alors $s_{i-1}''(t_i) = \sum_{k=1}^{i-1}\delta_k(t_i - t_k) = s_i''(t_i)$ et s est de classe C^2 sur $[a,b]$.

(c) Il reste enfin à calculer la dimension de S.

Ce qui précède montre que S est isomorphe à l'espace vectoriel :

$$V = \left\{ v = (\alpha_0, \alpha_1, \delta_1, \cdots, \delta_n) \in IR^{n+2}; \sum_{k=1}^{n}\delta_k = 0, \sum_{k=1}^{n}\delta_k t_k = 0 \right\} = Ker(\varphi)$$

où :

$$\begin{array}{llll} \varphi: & IR^{n+2} & \to & IR^2 \\ & v & \mapsto & \left(\sum_{k=1}^{n}\delta_k, \sum_{k=1}^{n}\delta_k t_k \right) \end{array}$$

La matrice de cette application linéaire est $A = \begin{pmatrix} 0 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & t_1 & t_2 & \cdots & t_n \end{pmatrix}$.

Les t_i étant deux à deux distincts, la matrice A est de rang 2. On a donc $Dim(Ker(\varphi)) = n$. On en déduit donc que :

$$Dim(S) = n$$

Q3. (a) L'application $B : x \mapsto x''$ est clairement linéaire de H^2 dans H^0.

Pour tout $x \in H^2$, on a :

$$\|B(x)\|_0^2 = \int_a^b x''(t)^2 dt \le \int_a^b \left(x(t)^2 + x'(t)^2 + x''(t)^2 \right) dt = \|x\|_2^2$$

On a donc :

$$\forall x \in H^2, \quad \|B(x)\|_0 \le \|x\|_2$$

C'est-à-dire que l'application linéaire B est continue de norme inférieure ou égale à 1.

Pour toute fonction $y \in H^0$, on peut écrire en utilisant le fait que la fonction constante égale à 1 est intégrable sur $[a,b]$ et l'inégalité de Cauchy–Schwarz :

$$\int_a^b |y(t)| dt = \langle |y| | 1 \rangle_0 \le \|y\|_0 \|1\|_0 < +\infty$$

C'est-à-dire que y est intégrable sur $[a,b]$. On peut alors définir la fonction $z : [a,b] \to IR$ par :

$$\forall t \in [a,b], \quad z(t) = \int_a^t y(u)du$$

La fonction z est alors absolument continue sur $[a,b]$ et admet une dérivée définie presque partout par $z'(t) = y(t)$.

La fonction $x:[a,b] \to IR$ définie par :

$$\forall t \in [a,b], \quad x(t) = \int_a^t z(u)du$$

est alors de classe C^1 avec une dérivée absolument continue donnée par $x' = z$.

Cette fonction admet aussi une dérivée seconde définie presque partout par $x''(t) = z'(t) = y(t)$ avec $y \in H^0$.

On a donc ainsi montré que pour tout $y \in H^0$, il existe $x \in H^2$ telle que $B(x) = x'' = y$.

C'est-à-dire que B est une surjection de H^2 sur H^0.

(b) Il est clair que $c: x \mapsto c(x) = (x(t_1), x(t_2), \cdots, x(t_n))$ est linéaire de H^2 dans IR^n.

La fonction $x \in H^2$ étant de classe C^1, on peut écrire pour tout $i \in \{1, \cdots, n\}$:

$$x(t_i) = x(t) + \int_t^{t_i} x'(u)du$$

Ce qui donne :

$$x(t_i)^2 = x(t)^2 + 2x(t)\int_t^{t_i} x'(u)du + \left(\int_t^{t_i} x'(u)du\right)^2$$

avec :

$$2x(t)\int_t^{t_i} x'(u)du \leq x(t)^2 + \left(\int_t^{t_i} x'(u)du\right)^2$$

D'où :

$$x(t_i)^2 \leq 2\left(x(t)^2 + \left(\int_t^{t_i} x'(u)du\right)^2\right)$$

D'autre part, avec l'inégalité de Cauchy–Schwarz on peut écrire :

$$\left(\int_t^{t_i} x'(u)\cdot 1du\right)^2 \leq \left(\int_t^{t_i} x'(u)^2 du\right)\left(\int_t^{t_i} 1du\right)$$

soit :

$$\left(\int_t^{t_i} x'(u)\cdot 1du\right)^2 \leq (t - t_i)\int_t^{t_i} x'(u)^2 du \leq (b-a)\int_a^b x'(u)^2 du$$

On a donc :

$$x(t_i)^2 \leq 2\left(x(t)^2 + (b-a)\int_a^b x'(u)^2 du\right)$$

et :

$$\int_a^b x(t_i)^2 du \leq 2\left(\int_a^b x(u)^2 du + \int_a^b \left((b-a)\int_a^b x'(u)^2 du\right)dt\right)$$

Ce qui peut aussi s'écrire :

$$x(t_i)^2 \leq \frac{2}{b-a}\int_a^b x(u)^2 du + 2(b-a)\int_a^b x'(u)^2 du \leq Max\left\{\frac{2}{b-a}, 2(b-a)\right\}\|x\|_2^2$$

Ce qui donne en définitive :

$$\|c(x)\|^2 = \sum_{i=1}^{n} x(t_i)^2 \leq n Max\left\{\frac{2}{b-a}, 2(b-a)\right\}\|x\|_2^{\ 2} = \lambda\|x\|_2^{\ 2}$$

ou encore :

$$\forall x \in H^2, \quad \|c(x)\| \leq \sqrt{\lambda}\|x\|_2$$

Ce qui prouve que c est une application linéaire continue de H^2 dans $I\!R^n$.

Pour tout $y = (y_1, \cdots, y_n) \in IR^n$, on sait qu'il existe un unique polynôme P de degré inférieur ou égal à $n-1$ tel que $P(t_i) = y_i$ (i = 1, ···, n) (polynôme d'interpolation de Lagrange – problème 33 –). Un tel polynôme est dans H^2 et vérifie $c(P) = y$.

On peut donc conclure que c est une surjection de H^2 sur $I\!R^n$.

(c) Pour tout $s \in S$, on a :

$$\forall i \in \{0, \cdots, n\}, \forall t \in \left]t_i, t_{i+1}\right[, \quad s'''(t) = 3!\,\gamma_i^3$$

avec $\gamma_0^3 = \gamma_n^3 = 0$.

Le saut de s''' en t_i est donc donné par :

$$\forall i \in \{1, \cdots, n\}, \quad \delta_i = 3!\left(\gamma_i^3 - \gamma_{i-1}^3\right)$$

On retrouve bien les coefficients δ_i définis en Q2.

Les applications $s \mapsto \gamma_i^3$ étant linéaires, on en déduit que d est une application linéaire de S dans $I\!R^n$.

Les espaces S et $I\!R^n$ étant de dimension finie, l'application d est continue.

On a vu que pour toute fonction $s \in S$, on a $\sum_{k=1}^{n} \delta_k = 0$, donc un vecteur $y = (y_1, \cdots, y_n) \in IR^n$ tel que $\sum_{k=1}^{n} y_k \neq 0$ ne peut pas être dans l'image de d.

En conclusion d n'est pas surjective.

On peut aussi remarquer que

$$Im(d) = \left\{\delta \in IR^n; \quad \sum_{i=1}^{n} \delta_i = \sum_{i=1}^{n} \delta_i t_i = 0\right\}$$

est de dimension $n-2$.

Q4. Pour $x \in H^2$ et $s \in S$, on a :

$$\langle B(x) | B(s)\rangle_0 = \int_a^b x''(t)s''(t)dt = \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} x''(t)s''(t)dt$$

du fait que $s'' = 0$ sur $[t_0, t_1]$ et $[t_n, t_{n+1}]$.

En faisant une intégration par parties (s'' est dérivable sur $\left]t_i, t_{i+1}\right[$), on a :

$$\int_{t_i}^{t_{i+1}} x''(t)s''(t)dt = x'(t_{i+1})s''(t_{i+1}) - x'(t_i)s''(t_i) - \int_{t_i}^{t_{i+1}} x'(t)s'''(t)dt$$

et :

$$\langle B(x) | B(s)\rangle_0 = x'(t_n)s''(t_n) - x'(t_1)s''(t_1) - 3!\sum_{i=1}^{n-1} \gamma_i^3 \int_{t_i}^{t_{i+1}} x'(t)dt$$

$$= -3!\sum_{i=1}^{n-1} \gamma_i^3\left(x(t_{i+1}) - x(t_i)\right)$$

($s''(t_1) = s''(t_n) = 0$).

En écrivant que $\delta_i = 3!(\gamma_i^3 - \gamma_{i-1}^3)$, on déduit que :

$$\langle B(x) | B(s) \rangle_0 = -3! \sum_{i=1}^{n-1} \gamma_i^3 x(t_{i+1}) + 3! \sum_{i=1}^{n-1} \gamma_i^3 x(t_i) = 3! \sum_{i=1}^{n} \gamma_i^3 x(t_i) - 3! \sum_{i=1}^{n} \gamma_{i-1}^3 x(t_i)$$

($\gamma_0^3 = \gamma_n^3 = 0$).

Ce qui peut aussi s'écrire :

$$\langle B(x) | B(s) \rangle_0 = \sum_{i=1}^{n} \delta_i x(t_i) = \langle c(x) | d(s) \rangle$$

Q5. Il s'agit de montrer que l'application linéaire $c + d : S \to IR^n$ est surjective, ce qui équivaut à montrer qu'elle injective puisque les espaces de départ et d'arrivée ont la même dimension.

Soit $s \in Ker(c+d)$, avec Q4 on peut écrire que :

$$0 \leq \|B(s)\|_0^2 = \langle c(s) | d(s) \rangle = -\|c(s)\|^2 \leq 0$$

Donc $\|B(s)\|_0 = \|c(s)\| = 0$ et $c(s) = d(s) = 0$.

Avec Q2, on déduit que la condition $d(s) = 0$ entraîne que $s(t) = \alpha_0 + \alpha_1 t$ pour tout $t \in [a,b]$.

D'autre part la condition $c(s) = 0$ entraîne que s est un polynôme de degré inférieur ou égal à 1 qui s'annule en $n \geq 2$ points distincts et nécessairement s est le polynôme nul.

On a donc $Ker(c+d) = \{0\}$ et $c + d$ est un isomorphisme de S sur IR^n.

Q6. Soit $z \in IR^n$ et $\sigma \in S$ définie par $c(\sigma) + d(\sigma) = z$.

Pour tout $x \in H^2$, on a :

$$\|B(x)\|_0^2 = \|B(x-\sigma) + B(\sigma)\|_0^2 = \|B(x-\sigma)\|_0^2 + 2\langle B(x-\sigma) | B(\sigma) \rangle_0 + \|B(\sigma)\|_0^2$$

En utilisant Q4, on a :

$$\langle B(x-\sigma) | B(\sigma) \rangle_0 = \langle c(x-\sigma) | d(\sigma) \rangle$$

D'autre part, on a :

$$\|c(x) - z\|^2 = \|c(x) - c(\sigma) - d(\sigma)\|^2 = \|c(x-\sigma) - d(\sigma)\|^2$$
$$= \|c(x-\sigma)\|^2 - 2\langle c(x-\sigma) | d(\sigma) \rangle + \|d(\sigma)\|^2$$

Ce qui donne :

$$\|B(x)\|_0^2 + \|c(x) - z\|^2 = \|B(x-\sigma)\|_0^2 + \|B(\sigma)\|_0^2 + \|c(x-\sigma)\|^2 + \|d(\sigma)\|^2$$
$$\geq \|B(\sigma)\|_0^2 + \|d(\sigma)\|^2$$

On donc montré que :

$$\|B(\sigma)\|_0^2 + \|c(\sigma) - z\|^2 = \underset{x \in H^2}{Min}\left(\|B(x)\|_0^2 + \|c(x) - z\|^2\right)$$

Q7. (a) Pour tout $j \in \{1, \cdots, n-2\}$, on a :

$$\langle \gamma_j | c(x) \rangle = \sum_{i=1}^{n} \gamma_{ji} x(t_i)$$

En écrivant, pour tout $i \in \{1, \cdots, n\}$:

$$x(t_i) = x(a) + (t_i - a)x'(a) + \int_a^b \left[(t_i - \xi)\right]_+ x''(\xi)d\xi$$

et en utilisant les égalités $\sum_{i=1}^{n} \gamma_{ji} = \sum_{i=1}^{n} \gamma_{ji} t_i = 0$, on aboutit à :

$$\langle \gamma_j | c(x) \rangle = \sum_{i=1}^{n} \gamma_{ji} \int_a^b \left[(t_i - \xi)\right]_+ x''(\xi)d\xi = \int_a^b \beta_j(t) x''(t)dt = \langle \beta_j | B(x) \rangle_0$$

où β_j est la fonction de H_0 définie par :

$$\forall t \in [a,b], \quad \beta_j(t) = \sum_{i=1}^{n} \gamma_{ji}(t_i - t)_+$$

On a donc ainsi montré l'existence d'une fonction $\beta_j \in H^0$ telle que :

$$\forall x \in H^2, \quad \langle \gamma_j | c(x) \rangle = \langle \beta_j | B(x) \rangle_0$$

(b) Par définition des coefficients γ_{ji}, on a :

$$\forall t \in [a,b], \quad \beta_j(t) = \sum_{i=j}^{j+2} \gamma_{ji}(t_i - t)_+$$

Pour $t > t_{j+2}$, on a $t_i - t < 0$ pour $i = j, j+1, j+2$, de sorte que $\beta_j(t) = 0$.

Pour $t < t_j$, on a $\beta_j(t) = \sum_{i=j}^{j+2} \gamma_{ji}(t_i - t) = \sum_{i=1}^{n} \gamma_{ji} t_i - t \sum_{i=1}^{n} \gamma_{ji} = 0$ par définition de Z_0.

On a donc bien :

$$\forall t \notin [t_j, t_{j+2}], \quad \beta_j(t) = 0$$

Pour $t \in [t_j, t_{j+1}]$, on a :

$$\beta_j(t) = \sum_{i=j+1}^{j+2} \gamma_{ji}(t_i - t) = \sum_{i=j+1}^{j+2} \gamma_{ji} t_i - t \sum_{i=j+1}^{j+2} \gamma_{ji} = -\gamma_{jj} t_j + t\gamma_{jj}$$

du fait que $\sum_{i=j}^{j+2} \gamma_{ji} = \sum_{i=j}^{j+2} \gamma_{ji} t_i = 0$. Ce qui donne :

$$\forall t \in [t_j, t_{j+1}], \quad \beta_j(t) = (t - t_j)\gamma_{jj} = \frac{t - t_j}{(t_j - t_{j+1})(t_j - t_{j+2})}$$

Pour $t \in [t_{j+1}, t_{j+2}]$, on a :

$$\beta_j(t) = \gamma_{j,j+2}(t_{j+2} - t) = \frac{t_{j+2} - t}{(t_{j+2} - t_j)(t_{j+2} - t_{j+1})}$$

Q8. L'égalité $B(s) = \sum_{j=1}^{n-2} \mu_j \beta_j$ est équivalente à :

$$\forall z \in H^0, \quad \left\langle B(s) - \sum_{j=1}^{n-2} \mu_j \beta_j \middle| z \right\rangle_0 = 0$$

L'application B étant une surjection de H^2 sur H^0, il est équivalent de montrer que :

$$\forall x \in H^2, \quad \left\langle B(s) - \sum_{j=1}^{n-2} \mu_j \beta_j \middle| B(x) \right\rangle_0 = 0$$

Pour $s \in S$ et $x \in H^2$, on a avec Q4 :

$$\langle B(s) | B(x) \rangle_0 = \langle c(x) | d(s) \rangle = \sum_{j=1}^{n-2} \mu_j \langle c(x) | \gamma_j \rangle$$

avec $\langle c(x) | \gamma_j \rangle = \langle \beta_j | B(x) \rangle_0$ d'après Q7.

On en déduit alors que :

$$\langle B(s) | B(x) \rangle_0 = \sum_{j=1}^{n-2} \mu_j \langle \beta_j | B(x) \rangle_0$$

Soit :

$$\left\langle B(s) - \sum_{j=1}^{n-2} \mu_j \beta_j \middle| B(x) \right\rangle_0 = 0$$

D'où le résultat.

Q9. Soit $z \in IR^n$ et $\sigma \in S$ définie par $c(\sigma) + d(\sigma) = z$.

Le vecteur $d(\sigma) \in Z_0$ s'écrit de manière unique :

$$d(\sigma) = \sum_{j=1}^{n-2} \mu_j \gamma_j$$

Avec Q8, on a aussi :

$$B(\sigma) = \sum_{j=1}^{n-2} \mu_j \beta_j$$

Avec $c(\sigma) + d(\sigma) = z$, on déduit que :

$$c(\sigma) = z - \sum_{j=1}^{n-2} \mu_j \gamma_j$$

Avec Q7, on a pour tout $k = 1, \cdots, n-2$, $\langle \gamma_k | c(\sigma) \rangle = \langle \beta_k | B(\sigma) \rangle_0$. C'est-à-dire que :

$$\left\langle \gamma_k \middle| z - \sum_{j=1}^{n-2} \mu_j \gamma_j \right\rangle = \left\langle \beta_k \middle| \sum_{j=1}^{n-2} \mu_j \beta_j \right\rangle_0$$

Ou encore :

$$\sum_{j=1}^{n-2} \left(\langle \beta_k | \beta_j \rangle_0 + \langle \gamma_k | \gamma_j \rangle \right) \mu_j = \langle z | \gamma_k \rangle$$

C'est-à-dire que $(\mu_1, \cdots, \mu_{n-2})$ est solution du système linéaire de $n-2$ équations à $n-2$ inconnues :

$$\sum_{j=1}^{n-2} \omega_{kj} \mu_j = \eta_k \quad (1 \leq k \leq n-2)$$

avec :

$$\omega_{kj} = \langle \beta_k | \beta_j \rangle_0 + \langle \gamma_k | \gamma_j \rangle \text{ et } \eta_k = \langle z | \gamma_k \rangle$$

CHAPITRE 6

Calcul approché des intégrales

Problème 39 : Méthodes de Newton–Cotes

Pour tout entier $n \geq 0$, on désigne par $IR_n[x]$ l'espace vectoriel des polynômes à coefficients réels de degré au plus n.

Dans tout le problème on se donne un intervalle réel $[a,b]$ avec $a < b$ et pour tout entier $n \geq 1$ on définit une subdivision de cet intervalle en posant :

$$h = \frac{b-a}{n}, \quad x_i = a + i \cdot h \quad (0 \leq i \leq n)$$

Pour une telle subdivision, on désigne par $\pi_{n+1} \in IR_{n+1}[x]$ le polynôme défini par :

$$\pi_{n+1}(x) = \prod_{i=0}^{n} (x - x_i)$$

On rappelle le théorème de la moyenne qui nous dit que si w est une fonction continue de signe constant sur un intervalle $[\alpha,\beta]$, alors pour toute fonction f continue sur $[\alpha,\beta]$, il existe un réel $c \in]\alpha,\beta[$ tel que :

$$\int_\alpha^\beta f(x)w(x)dx = f(c)\int_\alpha^\beta w(x)dx$$

Q1. *(a) Montrer que pour tout entier $n \geq 1$, il existe des coefficients réels uniques $\lambda_{n,0}, \lambda_{n,1}, \cdots, \lambda_{n,n}$ indépendants de l'intervalle $[a,b]$ tels que :*

$$(1) \quad \forall P \in IR_n[x], \quad \int_a^b P(x)dx = \frac{b-a}{n}\sum_{i=0}^{n} \lambda_{n,i} P(x_i)$$

(b) Montrer que les coefficients $\lambda_{n,i}$ $(0 \leq i \leq n)$ définis en (a) sont des nombres rationnels et qu'ils vérifient :

$$\begin{cases} \displaystyle\sum_{i=0}^{n} \lambda_{n,i} = n \\ \lambda_{n,n-i} = \lambda_{n,i} \quad (0 \leq i \leq n) \end{cases}$$

(c) Montrer que si n est un entier pair, alors la propriété (1) est encore valable pour $P \in IR_{n+1}[x]$.

Q2. *Pour cette question, on suppose que n est un entier pair (*$n = 2p$*).*

On désigne par $\psi_{n+2} \in IR_{n+2}[x]$ *la primitive de* π_{n+1} *qui s'annule en* $x = a$ *et pour tout* $k \in \{0,1,\cdots,n-1\}$*, on pose* $I_k = \int_{x_k}^{x_{k+1}} \pi_{n+1}(x)dx$.

Montrer que :

(a) $\forall k \in \{0,1,\cdots,p-1\}, \quad I_{2k+1} \leq 0 \leq I_{2k}$.

(b) $\forall k \in \{0,1,\cdots,p\}, \quad \psi_{n+2}(x_{n-k}) = \psi_{n+2}(x_k)$.

(c) $\forall k \in \{0,1,\cdots,p-1\}, \quad |I_{k+1}| \leq |I_k|$.

(d) $\forall x \in [a,b], \quad \psi_{n+2}(x) \geq 0$.

Pour tout entier $n \geq 1$ *et toute fonction continue* $f:[a,b] \to IR$*, on désigne par* $P_n \in IR_n[x]$ *le polynôme d'interpolation de Lagrange de f associé aux points* x_i *(*$0 \leq i \leq n$*) (problème 33, Q4).*

On désigne par E_n *l'erreur d'interpolation définie par :*

$$\forall x \in [a,b], \quad E_n(x) = f(x) - P_n(x)$$

et par g_n *la fonction définie par :*

$$\forall x \in [a,b] - \{x_0, x_1, \cdots, x_n\}, \quad g_n(x) = \frac{E_n(x)}{\pi_{n+1}(x)}$$

Q3. *(a) Montrer que si f est de classe* C^2 *sur* $[a,b]$*, alors la fonction* g_n *définie ci-dessus peut être prolongée en une fonction de classe* C^1 *sur* $[a,b]$.

(b) Montrer que si f est de classe C^{n+2} *sur* $[a,b]$*, alors pour tout* $x \in [a,b]$ *il existe* $\xi_x \in [a,b]$ *et* $\eta_x \in [a,b]$ *tels que :*

$$g_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}$$

$$g_n'(x) = \frac{f^{(n+2)}(\eta_x)}{(n+2)!}$$

Q4. *(a) On suppose que n est un entier pair (*$n = 2p$*) et on se donne une fonction f de classe* C^{n+2} *sur* $[a,b]$.

Montrer qu'il existe $\eta \in [a,b]$ *tel que :*

$$\int_a^b f(x)dx = \frac{b-a}{n}\sum_{i=0}^{n} \lambda_{n,i} f(x_i) + \frac{f^{(n+2)}(\eta)}{(n+2)!}\left(\int_0^n t^2(t-1)\cdots(t-n)dt\right)h^{n+3}$$

(b) On suppose que n est un entier impair ($n = 2p+1$*) et on se donne une fonction f de classe* C^{n+1} *sur* $[a,b]$.

Montrer qu'il existe $\eta \in [a,b]$ *tel que :*

$$\int_a^b f(x)dx = \frac{b-a}{n}\sum_{i=0}^{n} \lambda_{n,i} f(x_i) + \frac{f^{(n+1)}(\eta)}{(n+1)!}\left(\int_0^n t(t-1)\cdots(t-n)dt\right)h^{n+2}$$

Q5. *Dans cette question on se fixe un entier* $n \geq 1$.

Pour tout entier $p \geq 1$ *on définit une subdivision de* $[a,b]$ *en posant :*

$$\delta = \frac{b-a}{p}, \quad t_k = a + k\cdot\delta \quad (0 \le k \le p)$$

Pour toute fonction continue $f:[a,b] \to IR$, on pose :

$$(3) \quad T_p(f) = \frac{\delta}{n}\sum_{k=0}^{p-1}\sum_{i=0}^{n}\lambda_{n,i} f(x_{k,i})$$

où on a posé $x_{k,i} = t_k + i\frac{\delta}{n}$ (méthodes de Newton–Cotes composées).

(a) Montrer que $\underset{p\to+\infty}{Lim}\, T_p(f) = \int_a^b f(x)dx$.

(b) Donner une majoration de l'erreur de quadrature $\left|\int_a^b f(x)dx - T_p(f)\right|$ dans les deux cas suivants :

(i) $n = 2p$ et f est une fonction de classe C^{n+2} sur $[a,b]$.

(ii) $n = 2p+1$ et f est une fonction de classe C^{n+1} sur $[a,b]$.

Solution

Q1. (a) On désigne par $\{L_i ; 0 \le i \le n\}$ la base de Lagrange de $IR_n[x]$ (problème 33 Q4, (b)) définie par :

$$\forall i \in \{0,1,\cdots,n\}, \quad \begin{cases} L_i \in IR_n[x] \\ L_i(x_j) = \begin{cases} 1 \text{ si } j = i \\ 0 \text{ si } j \ne i \end{cases} \end{cases}$$

ou encore par :

$$\forall i \in \{0,1,\cdots,n\}, \quad L_i(x) = \prod_{\substack{k=0 \\ k\ne i}}^{n} \frac{x - x_k}{x_i - x_k}$$

On vérifie facilement que $\{L_i ; 0 \le i \le n\}$ forme une base de l'espace vectoriel $IR_n[x]$.

La propriété (1) est vérifiée pour tout $P \in IR_n[x]$ si et seulement si elle est vérifiée pour tous les L_i ($0 \le i \le n$), ce qui équivaut à :

$$\forall i \in \{0,1,\cdots,n\}, \quad \int_a^b L_i(x)dx = \frac{b-a}{n}\lambda_{n,i}$$

Ce qui détermine de manière unique les coefficients $\lambda_{n,i}$ ($0 \le i \le n$), avec :

$$\forall i \in \{0,1,\cdots,n\}, \quad \lambda_{n,i} = \frac{n}{b-a}\int_a^b L_i(x)dx$$

Le changement de variable $x = a + t(b-a)$ donne alors :

$$\forall i \in \{0,1,\cdots,n\}, \quad \lambda_{n,i} = n\int_0^1 L_i(a + t(b-a))dt$$

avec :

$$L_i(a + t(b-a)) = \frac{(-1)^{n-i}}{n!} C_n^i \prod_{\substack{k=0 \\ k\ne i}}^{n} (n\cdot t - k)$$

Ce qui donne en définitive :

$$\forall i \in \{0,1,\cdots,n\},\quad \lambda_{n,i} = \frac{(-1)^{n-i}}{n!} C_n^i n \int_0^1 \prod_{\substack{k=0 \\ k \neq i}}^{n} (n \cdot t - k) dt = \frac{(-1)^{n-i}}{n!} C_n^i \int_0^n \prod_{\substack{k=0 \\ k \neq i}}^{n} (t - k) dt$$

C'est-à-dire que les coefficients $\lambda_{n,i}$ ne dépendent que de n et de $i \in \{0,1,\cdots,n\}$.

(b) Pour tout $i \in \{0,1,\cdots,n\}$, le polynôme $w_{n,i}(t) = \prod_{\substack{k=0 \\ k \neq i}}^{n} (n \cdot t - k)$ est à coefficients entiers, donc $\int_0^1 w_{n,i}(t) dt \in Q$ et $\lambda_{n,i} \in Q$.

En prenant $P = 1$ dans (1), on a $b - a = \int_a^b dx = \frac{b-a}{n} \sum_{i=0}^{n} \lambda_{n,i}$ et $\sum_{i=0}^{n} \lambda_{n,i} = n$.

On pose $w_n(t) = \prod_{k=0}^{n} (n \cdot t - k)$ et pour tout $i \in \{0,1,\cdots,n\}$, on a $w_{n,i}(t) = \frac{w_n(t)}{n \cdot t - i}$.

Le changement de variable $t = 1 - u$ sur $[0,1]$ donne $w_n(t) = (-1)^{n+1} w_n(u)$ et :

$$\int_0^1 w_{n,n-i}(t) dt = \int_0^1 \frac{w_n(t)}{n \cdot t - (n-i)} dt = (-1)^n \int_0^1 \frac{w_n(u)}{n \cdot u - i} dt$$

Ce qui donne $\lambda_{n,n-i} = \lambda_{n,i}$ pour tout $i \in \{0,1,\cdots,n\}$, en tenant compte de $C_n^{n-i} = C_n^i$.

(c) Si $n = 2p$, alors $x_p = \frac{a+b}{2}$, $a - x_p = \frac{a-b}{2} = -(b - x_p)$ et par symétrie :

$$\int_a^b (x - x_p)^{n+1} dx = \left[\frac{(x - x_p)^{2p+2}}{2p+2} \right]_a^b = 0.$$

En considérant que pour tout $i \in \{0,1,\cdots,p\}$, on a $\lambda_{n,n-i} = \lambda_{n,i}$ et $(x_{n-i} - x_p)^{2p+1} = -(x_i - x_p)^{2p+1}$, on déduit qu'on a aussi $\sum_{i=0}^{n} \lambda_{n,i} (x_i - x_p)^{n+1} = 0$.

Donc la propriété (1) est vérifiée sur la base $\{(x - x_p)^i ; 0 \leq i \leq n+1\}$ de $IR_{n+1}[x]$, elle donc vraie sur tout $IR_{n+1}[x]$ par linéarité.

Q2. Pour $n = 2p$, on a le tableau de variations 6.1 pour ψ_{n+2} :

x	x_0	x_1	$\cdots$	x_{n-1} $\quad$ x_n
$\psi_{n+2}' = \pi_{n+1}$	0 $\quad$ +	0 $\quad$ −	$\cdots$	0 $\quad$ − $\quad$ 0
ψ_{n+2}	0 ↗	↘		↘

Tableau 6.1

(a) Du tableau 6.1 on déduit que pour tout $k \in \{0,1,\cdots,p-1\}$, on a :

$$I_{2k} = \psi_{n+2}(x_{2k+1}) - \psi_{n+2}(x_{2k}) \geq 0$$
$$I_{2k+1} = \psi_{n+2}(x_{2k+2}) - \psi_{n+2}(x_{2k+1}) \leq 0$$

C'est-à-dire que la suite $(I_k)_{0\leq k\leq n-1}$ est alternée.

(b) Pour tout $k \in \{0,1,\cdots,p\}$, on a $\psi_{n+2}(x_{n-k}) - \psi_{n+2}(x_k) = \int_{x_k}^{x_{n-k}} \pi_{n+1}(x)dx$.

Le changement de variable $x = x_p + t$ ($x_p = \frac{a+b}{2}$ étant le milieu de $[a,b]$) donne :

$$\int_{x_k}^{x_{n-k}} \pi_{n+1}(x)dx = \int_{-(p-k)h}^{(p-k)h} \prod_{i=0}^{2p}(t+(p-i)h)dt = \int_{-(p-k)h}^{(p-k)h} t\prod_{j=1}^{p}(t^2 - j^2h^2)dt = 0$$

par symétrie. C'est-à-dire que $\psi_{n+2}(x_{n-k}) = \psi_{n+2}(x_k)$.

(c) Pour $k = p-1$, on a $I_{k+1} = I_p = -I_{p-1} = -I_k$, donc $|I_{k+1}| = |I_k|$.

Pour $k \in \{0,1,\cdots,p-2\}$, le changement de variable $x = t-h$ donne :

$$I_k = \int_{x_k}^{x_{k+1}} \prod_{i=0}^{n}(x-x_i)dx = \int_{x_{k+1}}^{x_{k+2}} \prod_{i=0}^{n}(t-h-x_i)dt = \int_{x_{k+1}}^{x_{k+2}} \pi_{n+1}(t)\frac{t-h-x_n}{t-x_0}dt$$

Le polynôme π_{n+1} gardant un signe constant sur $[x_{k+1}, x_{k+2}]$, on peut utiliser le théorème de la moyenne pour écrire que :

$$I_k = \frac{c_k - h - x_n}{c_k - x_0}\int_{x_{k+1}}^{x_{k+2}} \pi_{n+1}(t)dt = \frac{c_k - h - x_n}{c_k - x_0} I_{k+1}$$

avec $c_k \in]x_{k+1}, x_{k+2}[$.

Pour $k \in \{0,1,\cdots,p-2\}$, on a $|c_k - (x_n + h)| > |c_k - x_0|$ et donc $|I_k| > |I_{k+1}|$.

(d) D'après le tableau de variations 6.1, pour montrer que la fonction ψ_{n+2} reste positive sur $[a,b]$, il suffit de montrer que :

$$\forall k \in \{0,1,\cdots,p\}, \quad \psi_{n+2}(x_{2k}) \geq 0$$

D'autre part en considérant que $\psi_{n+2}(x_{n-2k}) = \psi_{n+2}(x_{2k})$, il suffit de montrer que $\psi_{n+2}(x_{2k}) \geq 0$ pour $0 \leq 2k \leq p$.

Dans ce cas, on a $\psi_{n+2}(x_k) = \int_a^{x_{2k}} \pi_{n+1}(x)dx = I_0 + I_1 + \cdots + I_{2k-1}$, avec $I_{2j} \geq 0$, $I_{2j+1} \leq 0$ et $|I_{2j+1}| < |I_{2j}|$ pour $j = 0,1,\cdots,k-1$. Il en résulte alors que $\psi_{n+2}(x_{2k}) \geq 0$.

Q3. (a) Si f est de classe C^1 sur $[a,b]$, alors g_n est de classe C^1 sur $[a,b] - \{x_0, x_1, \cdots, x_n\}$ et pour tout $i \in \{0,1,\cdots,n\}$, on a :

$$\underset{\substack{x\to x_i \\ x\neq x_i}}{Lim}\, g_n(x) = \underset{\substack{x\to x_i \\ x\neq x_i}}{Lim}\left(\frac{E_n(x) - E_n(x_i)}{x - x_i}\,\frac{x - x_i}{\pi_{n+1}(x) - \pi_{n+1}(x_i)}\right) = \frac{f'(x_i) - P_n'(x_i)}{\pi_{n+1}'(x_i)}$$

ce qui permet de prolonger g_n par continuité sur $[a,b]$.

Pour tout $x \in [a,b] - \{x_0, x_1, \cdots, x_n\}$, on a :

$$g_n'(x) = \frac{\pi_{n+1}(x)E_n'(x) - E_n(x)\pi_{n+1}'(x)}{\pi_{n+1}(x)^2}$$

Si de plus f est de classe C^2, alors E_n est aussi de classe C^2 et avec la formule de Taylor le numérateur de $g_n'(x)$ s'écrit :

$$(x-x_i)^2\left\{\pi_{n+1}'(x_i)E_n''(x_i+\beta_x)+E_n'(x_i)\frac{\pi_{n+1}''(x_i+\alpha_x)}{2}-E_n'(x_i)\pi_{n+1}''(x_i+\delta_x)-\pi_{n+1}'(x_i)\frac{E_n''(x_i+\gamma_x)}{2}\right\}$$

$$+\frac{(x-x_i)^3}{2}\left\{\pi_{n+1}''(x_i+\alpha_x)E_n''(x_i+\beta_x)-E_n''(x_i+\gamma_x)\pi_{n+1}''(x_i+\delta_x)\right\}$$

et le dénominateur :

$$(x-x_i)^2\left\{\pi_{n+1}'(x_i)+\frac{x-x_i}{2}\pi_{n+1}''(x_i+\alpha_x)\right\}$$

avec $\alpha_x, \beta_x, \gamma_x, \delta_x$ qui tendent vers 0 quand x tend vers x_i.

Ce qui donne :

$$\lim_{\substack{x\to x_i\\ x\neq x_i}} g_n'(x) = \frac{\pi_{n+1}'(x_i)E_n''(x_i)-\pi_{n+1}''(x_i)E_n'(x_i)}{2\pi_{n+1}'(x_i)^2}$$

En utilisant le théorème des accroissements finis, on en déduit alors que g_n est de classe C^1 sur $[a,b]$.

(b) (i) Pour x donné dans $[a,b]-\{x_0,x_1,\cdots,x_n\}$, on désigne par $Q_{n+1}\in IR_{n+1}[x]$ le polynôme d'interpolation de Lagrange de f associé aux $n+2$ points $x_0,x_1,\cdots,x_n,x$.

Les polynômes $Q_{n+1}-P_n$ et π_{n+1} sont dans $IR_{n+1}[x]$ et s'annulent en $x_0,x_1,\cdots,x_n$, ils sont donc proportionnels. C'est-à-dire qu'il existe une constante réelle c telle que :

$$\forall t\in[a,b],\quad Q_{n+1}(t)-P_n(t)=c\pi_{n+1}(t)$$

D'autre part la fonction $\varphi: t\mapsto f(t)-Q_{n+1}(t)$ est de classe C^{n+1} et s'annule en $n+2$ points distincts $x_0,x_1,\cdots,x_n,x$. Le théorème de Rolle nous dit alors que sa dérivée φ' s'annule en $n+1$ points distincts et par récurrence que la dérivée d'ordre $n+1$ $\varphi^{(n+1)}$ s'annule en un point $\xi_x\in[a,b]$.

En écrivant que $\varphi(t)=f(t)-P_n(t)-c\pi_{n+1}(t)$ et que $P_n^{(n+1)}(t)=0$, $\pi_{n+1}^{(n+1)}(t)=(n+1)!$, on déduit que $c=\frac{1}{(n+1)!}f^{(n+1)}(\xi_x)$.

Puis avec $\varphi(x)=0$, on déduit que :

$$E_n(x)=f(x)-P_n(x)=\frac{1}{(n+1)!}f^{(n+1)}(\xi_x)\pi_{n+1}(x)$$

et :

$$g_n(x)=\frac{f^{(n+1)}(\xi_x)}{(n+1)!}\quad\left(x\in[a,b]-\{x_0,x_1,\cdots,x_n\}\right)$$

Enfin en notant $m_{n+1} = \underset{t\in[a,b]}{Inf} \dfrac{f^{(n+1)}(t)}{(n+1)!}$ et $M_{n+1} = \underset{t\in[a,b]}{Sup} \dfrac{f^{(n+1)}(t)}{(n+1)!}$, on peut en déduire que :

$$\forall x \in [a,b] - \{x_0, x_1, \cdots, x_n\}, \quad m_{n+1} \leq g_n(x) \leq M_{n+1}$$

La fonction g_n étant continue sur $[a,b]$, on déduit que cet encadrement est encore vrai sur tout l'intervalle $[a,b]$.

La fonction $f^{(n+1)}$ étant continue, le théorème des valeurs intermédiaires nous dit que pour tout $i \in \{0,1,\cdots,n\}$ il existe $\xi_{x_i} \in [a,b]$ tel que $g_n(x_i) = \dfrac{f^{(n+1)}(\xi_{x_i})}{(n+1)!}$.

On a donc bien le résultat souhaité pour tout x dans $[a,b]$.

(ii) Supposons maintenant que f est de classe C^{n+2} sur $[a,b]$.

Soient x dans $[a,b] - \{x_0, x_1, \cdots, x_n\}$ et $R_{n+2} \in IR_{n+2}[x]$ le polynôme d'interpolation d'Hermite de f défini par :

$$R_{n+2}(x_i) = f(x_i) \quad (0 \leq i \leq n)$$
$$R_{n+2}(x) = f(x)$$
$$R_{n+2}'(x) = f'(x)$$

(voir le problème 33).

On a alors :

$$\forall t \in [a,b], \quad R_{n+2}(t) - P_n(t) = \alpha(t-\beta)\pi_{n+1}(t)$$

D'autre part la fonction $\psi : t \mapsto f(t) - R_{n+2}(t)$ est de classe C^{n+2} et s'annule en $n+2$ points distincts $x_0, x_1, \cdots, x_n, x$. Le théorème de Rolle nous dit alors que sa dérivée ψ' s'annule en $n+1$ points distincts de $[a,b] - \{x_0, x_1, \cdots, x_n\}$. Elle s'annule également en x, c'est-à-dire qu'elle s'annule en $n+2$ points distincts.

Par récurrence, on déduit alors avec le théorème de Rolle que sa dérivée d'ordre $n+2$ $\varphi^{(n+2)}$ s'annule en un point $\eta_x \in [a,b]$.

En écrivant que $\psi(t) = f(t) - P_n(t) - \alpha(t-\beta)\pi_{n+1}(t)$ et que $P_n^{(n+2)}(t) = 0$, $((t-\beta)\pi_{n+1})^{(n+2)}(t) = (n+2)!$, on déduit que $\alpha = \dfrac{1}{(n+2)!} f^{(n+2)}(\eta_x)$.

Puis avec :

$$\psi(t) = E_n(t) - \alpha(t-\beta)\pi_{n+1}(t), \ \psi'(t) = E_n'(t) - \alpha\pi_{n+1}(t) - \alpha(t-\beta)\pi_{n+1}'(t)$$

et $\psi(x) = 0 = \psi'(x)$, on déduit que

$$E_n(x) = \alpha(x-\beta)\pi_{n+1}(x)$$
$$E_n'(x) = \alpha\pi_{n+1}(x) - \alpha(x-\beta)\pi_{n+1}'(x)$$

ce qui donne :

$$g_n(x) = \alpha(x-\beta) = \frac{f^{(n+2)}(\eta_x)}{(n+2)!}(x-\beta)$$

$$g_n'(x) = \alpha = \frac{f^{(n+2)}(\eta_x)}{(n+2)!}$$

En notant $m_{n+2} = \underset{t\in[a,b]}{Inf} \frac{f^{(n+2)}(t)}{(n+2)!}$ et $M_{n+2} = \underset{t\in[a,b]}{Sup} \frac{f^{(n+2)}(t)}{(n+2)!}$, on peut en déduire que :

$$\forall x \in [a,b] - \{x_0, x_1, \cdots, x_n\}, \quad m_{n+2} \leq g_n'(x) \leq M_{n+2}$$

La fonction g_n' étant continue sur $[a,b]$, on déduit que cet encadrement est encore vrai sur tout l'intervalle $[a,b]$.

Comme en (i), le théorème des valeurs intermédiaires et la continuité de $f^{(n+2)}$ permettent de conclure que pour tout $x \in [a,b]$ il existe $\eta_x \in [a,b]$ tel que $g_n'(x) = \frac{f^{(n+2)}(\eta_x)}{(n+2)!}$.

Q4. Si P_n est le polynôme d'interpolation de Lagrange de f associé aux points x_i ($0 \leq i \leq n$), alors $\frac{b-a}{n}\sum_{i=0}^{n}\lambda_{n,i} f(x_i) = \frac{b-a}{n}\sum_{i=0}^{n}\lambda_{n,i} P(x_i) = \int_a^b P_n(x)dx$ et avec les notations de Q3, on a :

$$R_n(f) = \int_a^b f(x)dx - \frac{b-a}{n}\sum_{i=0}^{n}\lambda_{n,i} f(x_i) = \int_a^b (f(x) - P_n(x))dx$$

$$= \int_a^b E_n(x)dx = \int_a^b g_n(x)\pi_{n+1}(x)dx$$

avec g_n de classe C^1 si f est de classe C^2.

(a) En reprenant les notations de Q2, on désigne par ψ_{n+2} la primitive de π_{n+1} qui s'annule en $x = a$.

En Q2 (b), on a vu que $\psi_{n+2}(b) = \psi_{n+2}(a) = 0$ si $n = 2p$.

Une intégration par parties nous donne alors :

$$R_n(f) = -\int_a^b g_n'(x)\psi_{n+2}(x)dx$$

En Q2 (d), on a vu que la fonction ψ_{n+2} est toujours positive si $n = 2p$, on peut donc utiliser le théorème de la moyenne pour écrire que :

$$R_n(f) = -g_n'(\xi)\int_a^b \psi_{n+2}(x)dx$$

avec $\xi \in [a,b]$.

Une deuxième intégration par parties donne :

$$\int_a^b \psi_{n+2}(x)dx = [(x-a)\psi_{n+2}(x)]_a^b - \int_a^b (x-a)\pi_{n+1}(x)dx = -\int_a^b (x-a)\pi_{n+1}(x)dx$$

et :

$$R_n(f) = g_n'(\xi)\int_a^b (x-a)\pi_{n+1}(x)dx$$

avec $g_n'(\xi) = \dfrac{f^{(n+2)}(\eta)}{(n+2)!}$ où $\eta \in [a,b]$.

Enfin le changement de variable $x = a + t \cdot h$ donne :

$$\int_a^b (x-a)\pi_{n+1}(x)dx = h^{n+3}\int_0^n t^2(t-1)\cdots(t-n)dt$$

D'où le résultat.

(b) On note $\pi_{n+1}(x) = \prod_{i=0}^{2p+1}(x-x_i) = \prod_{i=0}^{2p}(x-x_i)(x-b) = \pi_n(x)(x-b)$ $(n = 2p+1)$ et on désigne par ψ_{n+1} la primitive de π_n qui s'annule en $x = a$.

Comme en Q2, on voit que $\psi_{n+1}(x_{2p}) = \psi_{n+1}(a) = 0$ et $\psi_{n+1}(x) \geq 0$ sur $[a, x_{2p}]$.

L'erreur de quadrature peut s'écrire sous la forme :

$$R_n(f) = \int_a^{x_{2p}} E_n(x)dx + \int_{x_{2p}}^b E_n(x)dx$$

On désigne par $P_{n-1} \in IR_{n-1}[x]$ le polynôme d'interpolation de Lagrange de f aux points $x_0, x_1, \cdots, x_{2p}$.

On a :

$$\int_a^{x_{2p}} E_n(x)dx = \int_a^{x_{2p}} \left(f(x) - P_{n-1}(x)\right)dx + \int_a^{x_{2p}} \left(P_{n-1}(x) - P_n(x)\right)dx$$
$$= R_{n-1}(f) + \int_a^{x_{2p}} \left(P_{n-1}(x) - P_n(x)\right)dx$$

Le polynôme $P_{n-1}(x) - P_n(x)$ est dans $IR_n[x]$ et s'annule en $x_0, x_1, \cdots, x_{2p}$, donc $P_{n-1}(x) - P_n(x) = \alpha\prod_{i=0}^{2p}(x-x_i) = \alpha \cdot \pi_n(x)$, de sorte que :

$$\int_a^{x_{2p}} \left(P_{n-1}(x) - P_n(x)\right)dx = \alpha\int_a^{x_{2p}} \pi_n(x)dx = \alpha\left\{\psi_{n+1}(x_{2p}) - \psi_{n+1}(a)\right\} = 0$$

On a alors, en utilisant le résultat de (a) :

$$\int_a^{x_{2p}} E_n(x)dx = R_{n-1}(f) = -\frac{f^{(n+1)}(\xi_1)}{(n+1)!}\int_a^{x_{2p}} \psi_{n+1}(x)dx$$

Une intégration par parties donne :

$$\int_a^{x_{2p}} \psi_{n+1}(x)dx = \left[(x-b)\,\psi_{n+1}(x)\right]_a^{x_{2p}} - \int_a^{x_{2p}} (x-b)\pi_n(x)dx = -\int_a^{x_{2p}} \pi_{n+1}(x)dx$$

et :

$$\int_a^{x_{2p}} E_n(x)dx = \frac{f^{(n+1)}(\xi_1)}{(n+1)!}\int_a^{x_{2p}} \pi_{n+1}(x)dx$$

D'autre part, la fonction π_{n+1} étant négative sur l'intervalle $[x_{2p}, b]$, on peut utiliser le théorème de la moyenne pour écrire que :

$$\int_{x_{2p}}^b E_n(x)dx = \int_{x_{2p}}^b g_n(x)\pi_{n+1}(x)dx = g_n(\xi_2)\int_{x_{2p}}^b \pi_{n+1}(x)dx = \frac{f^{(n+1)}(\xi_2)}{(n+1)!}\int_{x_{2p}}^b \pi_{n+1}(x)dx$$

Ce qui donne en définitive :

$$R_n(f) = \frac{f^{(n+1)}(\xi_1)}{(n+1)!}\int_a^{x_{2p}} \pi_{n+1}(x)dx + \frac{f^{(n+1)}(\xi_2)}{(n+1)!}\int_{x_{2p}}^{b} \pi_{n+1}(x)dx$$
$$= \alpha_n f^{(n+1)}(\xi_1) + \beta_n f^{(n+1)}(\xi_2)$$

avec α_n et β_n négatifs pour n impair. En effet, avec une intégration par parties, on a :

$$\int_a^{x_{2p}} \pi_{n+1}(x)dx = \int_a^{x_{2p}} \pi_n(x)(x-b)dx$$
$$= \left[(x-b)\psi_{n+1}(x)\right]_a^{x_{2p}} - \int_a^{x_{2p}} \psi_{n+1}(x)dx = -\int_a^{x_{2p}} \psi_{n+1}(x)dx \leq 0$$

et π_{n+1} est négative sur l'intervalle $\left[x_{2p}, b\right]$.

En notant $m_{n+1} = \underset{t\in[a,b]}{Inf} f^{(n+1)}(t)$ et $M_{n+1} = \underset{t\in[a,b]}{Sup} f^{(n+1)}(t)$, on déduit que :

$$m_{n+1} \leq \frac{R_n(f)}{\alpha_n + \beta_n} \leq M_{n+1}$$

La fonction $f^{(n+1)}$ étant continue, on déduit avec le théorème des valeurs intermédiaires qu'il existe $\eta \in [a,b]$ tel que $R_n(f) = (\alpha_n + \beta_n) f^{(n+1)}(\eta)$, soit :

$$R_n(f) = \frac{f^{(n+1)}(\eta)}{(n+1)!}\int_a^b \pi_{n+1}(x)dx$$

Enfin le changement de variable $x = a + t \cdot h$ donne :

$$\int_a^b \pi_{n+1}(x)dx = h^{n+2}\int_0^n t(t-1)\cdots(t-n)dt$$

D'où le résultat.

Q5. (a) On a $T_p(f) = \frac{1}{n}\sum_{i=0}^{n} \lambda_{n,i} S_{i,p}(f)$, où $S_{i,p}(f) = \delta \sum_{k=0}^{p-1} f(x_{k,i})$ est une somme de Riemann de f associée à la subdivision de $[a,b]$ définie par les t_k ($0 \leq k \leq p$).

On a alors $\underset{p\to+\infty}{Lim} S_{i,p}(f) = \int_a^b f(x)dx$ pour tout $i \in \{0,1,\cdots,n\}$. Et avec $\sum_{i=0}^{n} \lambda_{n,i} = n$, on déduit que $\underset{p\to+\infty}{Lim} T_p(f) = \int_a^b f(x)dx$.

(b) (i) Pour $n = 2p$ et f de classe C^{n+2} sur $[a,b]$, on a d'après Q4 (a) :

$$\int_a^b f(x)dx - T_p(f) = \frac{I_n}{(n+2)!}\sum_{k=0}^{p-1} f^{(n+2)}(\eta_k)\left(\frac{t_{k+1}-t_k}{n}\right)^{n+3}$$

avec $I_n = \int_0^n t^2(t-1)\cdots(t-n)dt$, $t_{k+1} - t_k = \delta = \frac{b-a}{p}$ et $\eta_k \in [t_k, t_{k+1}]$.

Ce qui donne en notant $M_{n+2} = \underset{t\in[a,b]}{Sup}\left|f^{(n+2)}(t)\right|$:

$$\left|\int_a^b f(x)dx - T_p(f)\right| \leq \frac{|I_n| M_{n+2}(b-a)}{n^{n+3}(n+2)!}\left(\frac{b-a}{p}\right)^{n+2}$$

(ii) Pour $n = 2p+1$ et f de classe C^{n+1} sur $[a,b]$, on a d'après Q4 (b) :

$$\int_a^b f(x)dx - T_p(f) = \frac{J_n}{(n+1)!}\sum_{k=0}^{p-1} f^{(n+1)}(\xi_k)\left(\frac{t_{k+1}-t_k}{n}\right)^{n+2}$$

avec $J_n = \int_0^n t(t-1)\cdots(t-n)dt$ et $\xi_k \in [t_k, t_{k+1}]$.

Ce qui donne en notant $M_{n+1} = \underset{t\in[a,b]}{Sup}\left|f^{(n+1)}(t)\right|$:

$$\left|\int_a^b f(x)dx - T_p(f)\right| \leq \frac{|J_n| M_{n+1}(b-a)}{n^{n+2}(n+1)!}\left(\frac{b-a}{p}\right)^{n+1}$$

Problème 40 : Méthode de Romberg

Ce problème utilise certains résultats du problème 39.

Dans tout le problème on se donne un intervalle réel $[a,b]$ avec $a<b$ et une fonction continue $f:[a,b]\to IR$.

Pour tout entier $m\geq 1$ on définit une subdivision de $[a,b]$ en posant :

$$h = \frac{b-a}{m}, \quad x_i = a + i\cdot h \quad (0\leq i\leq m)$$

Le calcul approché de $\int_a^b f(x)dx$ par la méthode des trapèzes composée se fait en approchant cette intégrale par le réel $T(h)$ défini par :

$$T(h) = h\left\{\frac{f(a)+f(b)}{2} + \sum_{i=1}^{m-1} f(x_i)\right\}$$

(voir le problème 39, Q5 avec n = 1).

On définit la suite de Romberg $(R_{n,k})_{\substack{k\in IN\\ n\geq k}}$ par la récurrence :

$$\begin{cases} R_{n,0} = T\left(\dfrac{b-a}{2^n}\right) & (n\geq 0) \\[2mm] R_{n,k+1} = \dfrac{1}{4^{k+1}-1}\left(4^{k+1}R_{n,k} - R_{n-1,k}\right) & (k\geq 0, n\geq k+1) \end{cases}$$

Q1. *Montrer que les termes de la suite $(R_{n,0})_{n\geq 0}$ peuvent se calculer par la récurrence :*

$$\begin{cases} R_{0,0} = (b-a)\dfrac{f(a)+f(b)}{2} \\[2mm] R_{n+1,0} = \dfrac{1}{2}\left(R_{n,0} + R_{n,0}{}'\right) & (n\geq 0) \end{cases}$$

où $R_{n,0}{}' = \dfrac{b-a}{2^n}\sum_{i=1}^{2^n} f\left(a + \left(i-\dfrac{1}{2}\right)\dfrac{b-a}{2^n}\right)$.

Q2. *Montrer que, pour tout $k\geq 0$, on a $\underset{\substack{n\to+\infty\\ n\geq k}}{Lim} R_{n,k} = \int_a^b f(x)dx$.*

Q3. *(a) Montrer qu'il existe une suite double $(\alpha_{i,j})_{\substack{j\in IN\\ i\geq j}}$ telle que :*

$$R_{n,k} = \sum_{j=0}^{k} \alpha_{k,j} R_{n-j,0} \quad (0 \le k \le n)$$

(b) Montrer que $\sum_{j=0}^{k} \alpha_{k,j} = 1 \quad (k \ge 0)$.

(c) Montrer que pour tout $k \ge 1$ *et tout* $x \in IR$, *on a :*

$$\sum_{j=0}^{k} \alpha_{k,j} x^j = \prod_{j=1}^{k} \frac{4^j - x}{4^j - 1}$$

$$\sum_{j=0}^{k} \left|\alpha_{k,j}\right| x^j = \prod_{j=1}^{k} \frac{4^j + x}{4^j - 1}$$

(d) Montrer que pour tout entier $j \ge 0$, *on a* $\underset{\substack{n \to +\infty \\ n \ge j}}{Lim}\, \alpha_{n,n-j} = 0$.

Q4. *Pour cette question, on suppose que f est de classe C^2 sur* $[a,b]$.

Donner une majoration de l'erreur de quadrature $\left|R_{n,k} - \int_a^b f(x)dx\right|$, *pour tout entier* $k \ge 0$.

Q5. *Montrer que, pour toute fonction continue f sur* $[a,b]$, *on a* $\underset{n \to +\infty}{Lim} R_{n,n} = \int_a^b f(x)dx$.

Solution

Q1. Pour $n = 0$, on a bien $R_{0,0} = T(b-a) = (b-a)\dfrac{f(a)+f(b)}{2}$.

Pour tout entier $n \ge 0$, on pose $h_n = \dfrac{b-a}{2^n}$.

On a $R_{n+1,0} = T(h_{n+1}) = T\left(\dfrac{h_n}{2}\right) = \dfrac{h_n}{2}\left\{\dfrac{f(a)+f(b)}{2} + \sum_{j=1}^{2\cdot 2^n - 1} f\left(x_j'\right)\right\}$, avec :

$$x_j' = a + j \cdot h_{n+1} = \begin{cases} a + i \cdot h_n \text{ si } j = 2i & \left(1 \le i \le 2^n - 1\right) \\ a + \left(i - \dfrac{1}{2}\right)h_n \text{ si } j = 2i - 1 & \left(1 \le i \le 2^n\right) \end{cases}$$

C'est-à-dire que les x_{2i}' sont tous les points qui interviennent dans le calcul de $R_{n,0}$. Ce qui donne bien $R_{n+1,0} = \dfrac{1}{2}\left\{R_{n,0} + h_n \sum_{i=1}^{2^n} f\left(x_{2i-1}'\right)\right\}$.

C'est-à-dire qu'une fois calculé $R_{n,0}$, le calcul de $R_{n+1,0}$ ne nécessite que le calcul de 2^n nouveaux termes et non pas 2^{n+1}.

Q2. On a $R_{n,0}=\frac{1}{2}\left\{S_{0,n}+S_{1,n}\right\}$, où $S_{0,n}=\frac{b-a}{2^n}\sum_{i=0}^{2^n-1}f(x_i)$ et $S_{1,n}=\frac{b-a}{2^n}\sum_{i=1}^{2^n}f(x_i)$ sont des sommes de Riemann associées à la subdivision $(x_i)_{0\le i\le 2^n}$. On en déduit donc que $\underset{n\to+\infty}{Lim}\,R_{n,0}=\int_a^b f(x)dx$.

Si on suppose maintenant que $\underset{\substack{n\to+\infty\\ n\ge k}}{Lim}\,R_{n,k}=\int_a^b f(x)dx$, pour $k\ge 0$, alors la formule de récurrence $R_{n,k+1}=\frac{1}{4^{k+1}-1}\left(4^{k+1}R_{n,k}-R_{n-1,k}\right)$ entraîne la convergence de la suite $\left(R_{n,k+1}\right)_{n\ge k+1}$ vers $\int_a^b f(x)dx$.

Q3. (a) Pour $k = 0$, on a $\alpha_{0,0}=1$.

Supposons construits, pour $k\ge 0$, les coefficients $\left(\alpha_{k,j}\right)_{0\le j\le k}$ tels que $R_{n,k}=\sum_{j=0}^{k}\alpha_{k,j}R_{n-j,0}$ pour tout $n\ge k$ (ces coefficients ne dépendent pas de n).

Pour tout entier $n\ge k+1$, on a alors :

$$R_{n,k+1}=\frac{1}{4^{k+1}-1}\left(4^{k+1}\sum_{j=0}^{k}\alpha_{k,j}R_{n-j,0}-\sum_{j=0}^{k}\alpha_{k,j}R_{n-1-j,0}\right)=\sum_{j=0}^{k+1}\alpha_{k+1,j}R_{n-j,0}$$

avec :

$$\begin{cases}\alpha_{k+1,0}=\dfrac{4^{k+1}}{4^{k+1}-1}\alpha_{k,0}\\[2mm] \alpha_{k+1,j}=\dfrac{1}{4^{k+1}-1}\left\{4^{k+1}\alpha_{k,j}-\alpha_{k,j-1}\right\}\quad(1\le j\le k-1)\\[2mm] \alpha_{k+1,k+1}=-\dfrac{1}{4^{k+1}-1}\alpha_{k,k}\end{cases}$$

Ce qui détermine de manière unique les coefficients $\alpha_{k,j}$ ($k\ge 0, 0\le j\le k$).

(b) Les coefficients $\alpha_{k,j}$ ne dépendant ni de la fonction f ni de l'intervalle $[a,b]$, on peut prendre $f = 1$ et $[a,b]=[0,1]$, de sorte que $R_{n,k}=1$ pour tous $n\ge k\ge 0$.

Les relations $R_{n,k}=\sum_{j=0}^{k}\alpha_{k,j}R_{n-j,0}$ donnent alors $\sum_{j=0}^{k}\alpha_{k,j}=1$.

(c) Posons, pour tout entier $k\ge 1$ et tout réel x, $\varphi_k(x)=\prod_{j=1}^{k}\frac{4^j-x}{4^j-1}=\sum_{j=0}^{k}\beta_{k,j}x^j$.

En remarquant que $\varphi_{k+1}(x)=\frac{4^{k+1}-x}{4^{k+1}-1}\varphi_k(x)$, on vérifie que les coefficients $\beta_{k,j}$ pour $k\ge 1, 0\le j\le k$ vérifient la même récurrence que les $\alpha_{k,j}$ avec les mêmes conditions initiales $\alpha_{1,0}=\beta_{1,0}=\frac{4}{3}$ et $\alpha_{1,1}=\beta_{1,1}=-\frac{1}{3}$.

On déduit donc que $\alpha_{k,j}=\beta_{k,j}$, pour $k\ge 1, 0\le j\le k$. On a donc :

$$\sum_{j=0}^{k} \alpha_{k,j} x^j = \prod_{j=1}^{k} \frac{4^j - x}{4^j - 1}$$

Remarque — En prenant $x = 1$ dans l'identité ci-dessus, on retrouve $\sum_{j=0}^{k} \alpha_{k,j} = 1$.

D'autre part, en utilisant les formules de Viete, le numérateur de $\varphi_k(x)$ peut s'écrire sous la forme $\prod_{j=1}^{k}(4^j - x) = \sum_{j=0}^{k}(-1)^j S_{k-j} x^j$, où $S_{k-j} = \sum_{1 \le i_1 < \cdots < i_{k-j} \le k} 4^{i_1 + \cdots + i_{k-j}} > 0$.

On en déduit alors que $(-1)^j \alpha_{k,j} > 0$ et :

$$\varphi_k(-x) = \prod_{j=1}^{k} \frac{4^j + x}{4^j - 1} = \sum_{j=0}^{k} \alpha_{k,j} (-1)^j x^j = \sum_{j=0}^{k} \left|\alpha_{k,j}\right| x^j$$

(d) Pour $n \ge j$, $\alpha_{n-j,j}$ est le coefficient de x^{n-j} dans $\varphi_n(x)$. C'est donc :

$$\alpha_{n-j,j} = \frac{1}{\prod_{i=1}^{n}(4^i - 1)}(-1)^{n-j} S_j = \frac{(-1)^{n-j} \sum_{1 \le i_1 < \cdots < i_j \le n} 4^{i_1 + \cdots + i_j}}{\prod_{i=1}^{n}(4^i - 1)}$$

On a alors :

$$\left|\alpha_{n-j,j}\right| \le \frac{C_n^j 4^{(n-j+1)+\cdots+n}}{\prod_{i=1}^{n}(4^i - 1)} \le \frac{C_n^j 4^{n \cdot j - \frac{j(j-1)}{2}}}{\prod_{i=1}^{n}(4^i - 1)}$$

D'autre part, la convergence du produit infini $\prod_{j=1}^{+\infty} \frac{4^j}{4^j - 1}$ (prendre le logarithme) nous garantit l'existence d'une constante $A > 0$ telle que $\prod_{j=1}^{n} \frac{4^j}{4^j - 1} \le A$ pour tout $n \ge 1$. Ce qui donne $\prod_{j=1}^{n} \frac{1}{4^j - 1} \le \frac{A}{\prod_{j=1}^{n} 4^j} = \frac{A}{4^{\frac{n(n+1)}{2}}}$ et :

$$\left|\alpha_{n-j,j}\right| \le \frac{A}{4^{\frac{j(j-1)}{2}}} \frac{C_n^j}{4^{n\left(\frac{n+1}{2} - j\right)}} = u_{n,j}$$

avec $\underset{\substack{n \to +\infty \\ n \ge j}}{Lim} \frac{u_{n+1,j}}{u_{n,j}} = 0$. On en déduit donc que $\underset{\substack{n \to +\infty \\ n \ge j}}{Lim} \alpha_{n-j,j} = 0$.

Q4. Pour $n = 0$, on a $R_{n,0} = T\left(\frac{b-a}{2^n}\right)$. Dans le problème 39, Q5 (b) (ii), on a obtenu la majoration suivante de l'erreur de quadrature dans la méthode des trapèzes composée :

$$\left|\int_a^b f(x)dx - R_{n,0}\right| \le \frac{|J_1| M_2 (b-a)}{2} \left(\frac{b-a}{2^n}\right)^2$$

avec $J_1 = \int_0^1 t(t-1)dt = -\dfrac{1}{6}$ et $M_2 = \underset{t\in[a,b]}{Sup}\left|f''(t)\right|$, ce qui donne :

$$\left|\int_a^b f(x)dx - R_{n,0}\right| \leq \frac{M_2(b-a)^3}{12}4^{-n}$$

Pour $k \geq 1$, on a $R_{n,k} = \sum_{j=0}^{k} \alpha_{k,j} R_{n-j,0}$ et avec $\sum_{j=0}^{k} \alpha_{k,j} = 1$, on peut écrire que :

$$\left|\int_a^b f(x)dx - R_{n,k}\right| = \left|\sum_{j=0}^{k} \alpha_{k,j}\left(\int_a^b f(x)dx - R_{n-j,0}\right)\right| \leq \frac{M_2(b-a)^3}{12}4^{-n}\sum_{j=0}^{k}\left|\alpha_{k,j}\right|4^j$$

avec $\sum_{j=0}^{k}\left|\alpha_{k,j}\right|4^j = \varphi_k(-4) = \prod_{j=1}^{k}\frac{4^j+4}{4^j-1}$. Le produit infini $\prod_{j=1}^{+\infty}\frac{4^j+4}{4^j-1}$ étant convergeant, il existe une constante $B > 0$ telle que $\prod_{j=1}^{k}\frac{4^j+4}{4^j-1} \leq B$ pour tout $k \geq 1$ et :

$$\forall n \geq k,\quad \left|\int_a^b f(x)dx - R_{n,k}\right| \leq \frac{B\cdot M_2(b-a)^3}{12}4^{-n}$$

Remarque — Si on suppose que f est de classe C^∞, on peut alors montrer en utilisant la formule d'Euler Mac–Laurin, que $\left|\int_a^b f(x)dx - R_{n,k}\right|$ est un $O\left(4^{-(k+1)n}\right)$.

Q5. Si on suppose que f est de classe C^2, alors le résultat découle de Q4.

On a $\int_a^b f(x)dx - R_{n,n} = \sum_{j=0}^{n}\alpha_{n,n-j}\left(\int_a^b f(x)dx - R_{j,0}\right)$.

D'après Q4, pour tout $\varepsilon > 0$ donné on peut trouver $j_0 \in IN$ tel que :

$$\forall j \geq j_0,\quad \left|\int_a^b f(x)dx - R_{j,0}\right| \leq \varepsilon$$

Pour $n > j_0$, on a alors :

$$\left|\int_a^b f(x)dx - R_{n,n}\right| \leq \sum_{j=0}^{j_0}\left|\alpha_{n,n-j}\right|\left|\int_a^b f(x)dx - R_{j,0}\right| + \varepsilon\sum_{j=j_0+1}^{n}\left|\alpha_{n,n-j}\right|$$

Le produit infini $\prod_{j=1}^{+\infty}\frac{4^j+1}{4^j-1}$ étant convergeant, il existe une constante $C > 0$ telle que $\sum_{j=0}^{n}\left|\alpha_{n,n-j}\right| = \prod_{j=1}^{n}\frac{4^j+1}{4^j-1} \leq C$ pour tout $n \geq 1$ et :

$$\forall n > j_0,\quad \left|\int_a^b f(x)dx - R_{n,n}\right| \leq D\sum_{j=0}^{j_0}\left|\alpha_{n,n-j}\right| + C\cdot\varepsilon$$

où on a noté $D = Max\left\{\left|\int_a^b f(x)dx - R_{j,0}\right|;\quad 0 \leq j \leq j_0\right\}$.

Et avec $\underset{\substack{n\to+\infty\\ n\geq j}}{Lim}\,\alpha_{n-j,j} = 0$, on en déduit que $\underset{n\to+\infty}{Lim}\,R_{n,n} = \int_a^b f(x)dx$.

Problème 41 : Méthodes de Gauss

Ce problème utilise les notations et résultats du problème 25.

On rappelle qu'on définit un produit scalaire sur $IR[x]$ *en posant :*

$$\forall P, Q \in IR[x],\ \langle P | Q \rangle = \int_a^b P(x)Q(x)\pi(x)dx$$

où $-\infty \leq a < b \leq +\infty$ *et* $\pi :]a,b[\to IR_+^*$ *est une fonction poids intégrable telle que* $0 < \int_a^b |x^k| \pi(x)dx < +\infty$ *pour tout entier* $k \geq 0$. *La norme associée à ce produit scalaire est notée* $P \mapsto \|P\| = \sqrt{\langle P | P \rangle}$.

On désigne par $\{P_n;\ n \in IN\}$ *une base orthonormée pour ce produit scalaire et on note pour tout entier* $n \geq 0$, $P_n(x) = \sum_{j=0}^{n} \alpha_j^{(n)} x^j$ *avec* $\alpha_n^{(n)} \neq 0$.

Dans le problème 25, on a montré que la suite la suite $(P_n)_{n\in IN}$ *vérifie la relation de récurrence :*

$$(1)\quad \begin{cases} P_{-1}(x) = 0,\ P_0(x) = \alpha_0^{(0)} = \dfrac{1}{\sqrt{\int_a^b \pi(x)dx}} \\ b_{n+1}P_{n+1}(x) + a_n P_n(x) + b_n P_{n-1}(x) = xP_n(x) \quad (n \geq 0) \end{cases}$$

avec :

$$\begin{cases} a_0 = -\dfrac{\alpha_0^{(1)}}{\alpha_1^{(1)}},\ b_0 = 0 \\ a_n = \dfrac{\alpha_{n-1}^{(n)}}{\alpha_n^{(n)}} - \dfrac{\alpha_n^{(n+1)}}{\alpha_{n+1}^{(n+1)}} \quad (n \geq 1) \\ b_n = \dfrac{\alpha_{n-1}^{(n-1)}}{\alpha_n^{(n)}} \quad (n \geq 1) \end{cases}$$

et que, pour tout $n \geq 1$, P_n *admet n racines réelles simples dans l'intervalle* $]a,b[$.

Q1. *Montrer la formule de Darboux–Christoffel :*

$$\forall n \in IN, \forall x, y \in IR,\quad (x-y)\sum_{k=0}^{n} P_k(x)P_k(y) = \frac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}}\{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)\}$$

Q2. *Dans cette question, on se fixe un entier* $n \geq 1$ *et on cherche des réels* $x_{n,1} < x_{n,2} < \cdots < x_{n,n}$ *dans l'intervalle* $]a,b[$ *et des coefficients réels* $\lambda_{n,1}, \lambda_{n,2}, \cdots, \lambda_{n,n}$ *tels que :*

$$(2)\quad \forall P \in IR_{2n-1}[x],\quad \int_a^b P(x)\pi(x)dx = \sum_{i=1}^{n} \lambda_{n,i} P(x_{n,i})$$

(a) Montrer qu'il existe des coefficients $\lambda_{n,i}$ $(1 \leq i \leq n)$ *tels que (2) soit vérifié si et seulement si les réels* $x_{n,i}$ $(1 \leq i \leq n)$ *sont les racines de* P_n.

(b) On suppose maintenant que les $x_{n,i}$ $(1 \leq i \leq n)$ *sont les racines de* P_n.

Montrer que les coefficients $\lambda_{n,i}$ $(1 \le i \le n)$ *vérifiant (2) sont donnés par :*

$$\lambda_{n,i} = \frac{\alpha_n^{(n)}}{\alpha_{n-1}^{(n-1)}} \frac{1}{P_n'(x_{n,i}) P_{n-1}(x_{n,i})} \quad (1 \le i \le n)$$

Les coefficients $\lambda_{n,i}$ $(1 \le i \le n)$ *sont appelés coefficients de Christoffel.*

Q3. *Pour cette question, on prend* $]a,b[=]-1,1[$ *et* $\pi(x) = 1$. *La base orthonormée* $\{P_n;\ n \in IN\}$ *est alors formée des polynômes de Legendre normalisés définis par :*

$$\forall n \ge 0, \forall x \in IR, \quad P_n(x) = \sqrt{\frac{2n+1}{2}} \frac{1}{2^n n!} \left(\frac{d}{dx}\right)^n \left\{(x^2-1)^n\right\}$$

Ils sont également définis par la récurrence :

$$\begin{cases} P_{-1}(x) = 0,\ P_0(x) = \dfrac{1}{\sqrt{2}} \\ b_{n+1} P_{n+1}(x) + b_n P_{n-1}(x) = x P_n(x) \quad (n \ge 0) \end{cases}$$

avec :

$$b_0 = 0, \quad b_n = \frac{n}{\sqrt{2n-1}\sqrt{2n+1}} \quad (n \ge 1)$$

(voir le problème 25, Q5).

(a) Calculer $P_n(1)$ *pour tout* $n \ge 1$.

(b) Montrer que :

$$\forall n \ge 0, \quad (x^2-1) P_n'(x) = n \cdot b_{n+1} P_{n+1}(x) - (n+1) b_n P_{n-1}(x)$$

(c) Montrer que les coefficients de Christoffel correspondants sont donnés par :

$$\lambda_{n,i} = \frac{(2n-1)(1 - x_{n,i}{}^2)}{\left(n P_{n-1}(x_{n,i})\right)^2} \quad (1 \le i \le n)$$

Q4. *Pour cette question, on prend* $]a,b[=]-1,1[$ *et* $\pi(x) = \dfrac{1}{\sqrt{1-x^2}}$. *La base orthonormée* $\{P_n;\ n \in IN\}$ *est alors formée des polynômes de Tchébychev normalisés de première espèce définis par :*

$$\begin{cases} P_0(x) = \dfrac{1}{\sqrt{\pi}} \\ \forall n \ge 1, \forall x \in [-1,1], \quad P_n(x) = \sqrt{\dfrac{2}{\pi}} Cos(nArcCos(x)) \end{cases}$$

Ils sont également définis par la récurrence :

$$\begin{cases} P_0(x) = \dfrac{1}{\sqrt{\pi}}, \quad P_1(x) = \sqrt{\dfrac{2}{\pi}} x \\ \dfrac{1}{2} P_2(x) + \dfrac{1}{\sqrt{2}} P_0(x) = x P_1(x) \\ \dfrac{1}{2} P_{n+1}(x) + \dfrac{1}{2} P_{n-1}(x) = x P_n(x) \quad (n \ge 2) \end{cases}$$

(voir le problème 25, Q7).

Calculer, pour tout entier $n \geq 1$, les racines de P_n et les coefficients de Christoffel correspondants.

Q5. *Pour cette question, on prend $]a,b[=]0,+\infty[$ et $\pi(x) = x^{\alpha}e^{-x}$, où $\alpha > -1$ est un réel donné. On définit les polynômes de Laguerre par :*

$$\forall n \geq 0, \forall x \in IR, \quad Q_n(x) = \frac{x^{-\alpha}e^{x}}{n!}\left(\frac{d}{dx}\right)^n \left\{x^{n+\alpha}e^{-x}\right\}$$

(a) Montrer que $\{Q_n;\ n \in IN\}$ est une base orthogonale de $IR[x]$.

(b) Etablir la récurrence vérifiée par les polynômes de Laguerre.

(c) Calculer $\|Q_n\|$ pour tout entier n.

(d) Montrer que :

$$\forall n \geq 0, \quad xQ_n'(x) = nQ_n(x) - (n+\alpha)Q_{n-1}(x)$$

(e) Calculer, pour tout entier $n \geq 1$, les coefficients de Christoffel correspondants.

Q6. *On se place de nouveau dans le cas général avec une fonction poids $\pi :]a,b[\to IR_+^*$ et une base orthonormée $\{P_n;\ n \in IN\}$. Chaque P_n ayant n racines réelles distinctes auxquelles on associe les coefficients de Christoffel.*

On se donne une fonction f de classe C^{2n} sur $[a,b]$ telle que $\int_a^b f(x)\pi(x)dx$ soit convergente.

Montrer qu'il existe $\eta \in]a,b[$ tel que :

$$\int_a^b f(x)\pi(x)dx = \sum_{i=1}^{n} \lambda_{n,i} f(x_{n,i}) + \frac{1}{\left(\alpha_n^{(n)}\right)^2}\frac{f^{(2n)}(\eta)}{(2n)!}$$

Solution

Q1. Pour tout entier $k \geq 0$ et tout réel x, on a :

$$xP_k(x) = b_{k+1}P_{k+1}(x) + a_k P_k(x) + b_k P_{k-1}(x)$$

En multipliant les deux membres de cette égalité par $P_k(y)$, pour $y \in IR$, il vient :

$$xP_k(x)P_k(y) = b_{k+1}P_{k+1}(x)P_k(y) + a_k P_k(x)P_k(y) + b_k P_{k-1}(x)P_k(y)$$

Ce qui peut aussi s'écrire en permutant les rôles de x et y :

$$xP_k(y)P_k(x) = b_{k+1}P_{k+1}(y)P_k(x) + a_k P_k(y)P_k(x) + b_k P_{k-1}(y)P_k(x)$$

En faisant la différence des deux égalités obtenues, on obtient alors :

$$(x-y)P_k(x)P_k(y)$$
$$= b_{k+1}\left(P_{k+1}(x)P_k(y) - P_k(x)P_{k+1}(y)\right) - b_k\left(P_k(x)P_{k-1}(y) - P_{k-1}(x)P_k(y)\right)$$

Et la somme pour k allant de 0 à n donne :

$$(x-y)\sum_{k=0}^{n} P_k(x)P_k(y) = b_{n+1}\left\{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)\right\} - b_0\left\{P_0(x)P_{-1}(y) - P_{-1}(x)P_0(y)\right\}$$

En considérant que $b_0 = 0$, $P_{-1} = 0$ et $b_{n+1} = \dfrac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}}$, on déduit que :

$$(x-y)\sum_{k=0}^{n} P_k(x)P_k(y) = \frac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}}\left\{P_{n+1}(x)P_n(y)-P_n(x)P_{n+1}(y)\right\}$$

Q2. Supposons qu'il existe des coefficients $\lambda_{n,i}$ $(1 \le i \le n)$ tels que la propriété (2) soit vérifiée.

On désigne par ω_n le polynôme unitaire de degré n qui s'annule en chacun des $x_{n,i}$ $(1 \le i \le n)$, soit $\omega_n(x) = \prod_{i=1}^{n}(x - x_{n,i})$.

Pour tout polynôme $Q \in I\!R_{n-1}[x]$, on a $\omega_n Q \in I\!R_{2n-1}[x]$ et avec (2) :

$$\langle \omega_n \,|\, Q\rangle = \int_a^b \omega_n(x)Q(x)\pi(x)dx = \sum_{i=1}^{n}\lambda_{n,i}\omega_n(x_{n,i})Q(x_{n,i}) = 0$$

C'est-à-dire que ω_n est orthogonal à $I\!R_{n-1}[x]$. Ce polynôme est donc proportionnel à P_n et les $x_{n,i}$ $(1 \le i \le n)$ sont racines de P_n.

Réciproquement, on suppose que les $x_{n,i}$ $(1 \le i \le n)$ sont les racines de P_n.

Par division euclidienne, tout polynôme $P \in I\!R_{2n-1}[x]$ s'écrit sous la forme $P = Q \cdot P_n + R$ avec $Q \in I\!R_{n-1}[x]$ et $R \in I\!R_{n-1}[x]$. On a alors :

$$\int_a^b P(x)\pi(x)dx = \langle P_n \,|\, Q\rangle + \int_a^b R(x)\pi(x)dx = \int_a^b R(x)\pi(x)dx$$

puisque $P_n \in (I\!R_{n-1}[x])^{\perp}$.

En remarquant que $P(x_{n,i}) = R(x_{n,i})$ pour tout $i = 1,\cdots,n$, on déduit qu'il suffit de vérifier la propriété (2) pour tous les polynômes $P \in I\!R_{n-1}[x]$, ce qui équivaut à prouver l'existence de coefficients $\lambda_{n,i}$ $(1 \le i \le n)$ solutions du système linéaire de n équations à n inconnues :

$$\sum_{i=1}^{n}(x_{n,i})^{k-1}\lambda_{n,i} = \int_a^b x^{k-1}\pi(x)dx \quad (1 \le k \le n)$$

Le déterminant de ce système est de type Van Der Monde égal à $\prod_{1\le j<i\le n}(x_{n,i} - x_{n,j}) \ne 0$, ce qui assure l'existence et l'unicité d'une solution.

(b) On désigne par $\{L_i ; 1 \le i \le n\}$ la base de Lagrange de $I\!R_{n-1}[x]$ (problème 33 Q4, (b)) définie par :

$$\forall i \in \{1,\cdots,n\}, \quad L_i(x) = \prod_{\substack{k=1\\k\ne i}}^{n}\frac{x - x_{n,k}}{x_{n,i} - x_{n,k}} = \frac{P_n(x)}{x - x_{n,i}}\frac{1}{P_n'(x_{n,i})}$$

On a alors $\int_a^b L_i(x)\pi(x)dx = \sum_{k=1}^{n}\lambda_{n,k}L_i(x_{n,k}) = \lambda_{n,i}$, soit :

$$\lambda_{n,i} = \frac{1}{P_n'(x_{n,i})}\int_a^b \frac{P_n(x)}{x - x_{n,i}}\pi(x)dx \quad (1 \le i \le n)$$

Avec l'identité de Darboux–Christoffel, on a :

$$(x - x_{n,i})\sum_{k=0}^{n}P_k(x)P_k(x_{n,i}) = -\frac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}}P_n(x)P_{n+1}(x_{n,i})$$

Ce qui donne, pour $x \ne x_{n,i}$:

$$\frac{P_n(x)}{x-x_{n,i}} = -\frac{\alpha_{n+1}^{(n+1)}}{\alpha_n^{(n)}} \frac{1}{P_{n+1}(x_{n,i})} \sum_{k=0}^{n} P_k(x) P_k(x_{n,i})$$

et :

$$\lambda_{n,i} = -\frac{\alpha_{n+1}^{(n+1)}}{\alpha_n^{(n)}} \frac{1}{P_{n+1}(x_{n,i})} \frac{1}{P_n'(x_{n,i})} \sum_{k=0}^{n} \frac{P_k(x_{n,i})}{\alpha_0^{(0)}} \langle P_k \mid P_0 \rangle \quad (1 \leq i \leq n)$$

Avec l'orthogonalité de la famille $\{P_k;\quad 0 \leq k \leq n\}$, on déduit alors que :

$$\lambda_{n,i} = -\frac{\alpha_{n+1}^{(n+1)}}{\alpha_n^{(n)}} \frac{1}{P_{n+1}(x_{n,i})} \frac{1}{P_n'(x_{n,i})} \quad (1 \leq i \leq n)$$

En utilisant la récurrence (1), on a $b_{n+1} P_{n+1}(x_{n,i}) + b_n P_{n-1}(x_{n,i}) = 0$, donc :

$$\frac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}} = b_{n+1} = -b_n \frac{P_{n-1}(x_{n,i})}{P_{n+1}(x_{n,i})} = -\frac{\alpha_{n-1}^{(n-1)}}{\alpha_n^{(n)}} \frac{P_{n-1}(x_{n,i})}{P_{n+1}(x_{n,i})}$$

et :

$$\lambda_{n,i} = \frac{\alpha_n^{(n)}}{\alpha_{n-1}^{(n-1)}} \frac{1}{P_n'(x_{n,i}) P_{n-1}(x_{n,i})} \quad (1 \leq i \leq n)$$

Q3. (a) On a $P_0(1) = \frac{1}{\sqrt{2}}$ et par récurrence sur $n \geq 0$, $P_n(1) = \sqrt{\frac{2n+1}{2}}$.

(b) Pour $n = 0$, l'identité se réduit à $0 = 0$. On suppose donc $n \geq 1$.

Le polynôme $(x^2-1)P_n'$ est dans $IR_{n+1}[x]$, il s'écrit donc :

$$(x^2-1)P_n'(x) = \sum_{k=0}^{n+1} \mu_k P_k(x)$$

avec $\mu_k = \left\langle (x^2-1)P_n' \middle| P_k \right\rangle$ pour $0 \leq k \leq n+1$.

Une intégration par parties donne :

$$\mu_k = -\int_{-1}^{1} P_n(x) \frac{d}{dx}\left\{(x^2-1)P_k(x)\right\} dx$$

Pour $0 \leq k \leq n-2$, on a $\frac{d}{dx}\left\{(x^2-1)P_k(x)\right\} \in IR_{n-1}[x]$ et $\mu_k = 0$ puisque P_n est orthogonal à $IR_{n-1}[x]$.

D'autre part, avec la récurrence (1), on déduit que P_n est de même parité que n, donc que la fonction $(x^2-1)P_n'(x)P_n(x)$ est impaire et $\mu_n = \int_{-1}^{1} (x^2-1)P_n'(x)P_n(x)dx = 0$.

Il reste donc :

$$(x^2-1)P_n'(x) = \mu_{n+1} P_{n+1}(x) + \mu_{n-1} P_{n-1}(x)$$

En prenant $x = 1$ dans cette dernière égalité, il vient, compte tenu de (a) :

$$\mu_{n+1} P_{n+1}(1) + \mu_{n-1} P_{n-1}(1) = 0$$

Ce qui donne :

$$\mu_{n-1} = -\frac{P_{n+1}(1)}{P_{n-1}(1)}\mu_{n+1} = -\frac{\sqrt{2n+3}}{\sqrt{2n-1}}\mu_{n+1}$$

Enfin, en identifiant les coefficients de x^{n+1}, on a $n \cdot \alpha_n^{(n)} = \mu_{n+1}\alpha_{n+1}^{(n+1)}$ et :

$$\begin{cases} \mu_{n+1} = n\dfrac{\alpha_n^{(n)}}{\alpha_{n+1}^{(n+1)}} = n \cdot b_{n+1} = \dfrac{n(n+1)}{\sqrt{2n+1}\sqrt{2n+3}} \\ \mu_{n-1} = -\dfrac{n(n+1)}{\sqrt{2n-1}\sqrt{2n+1}} = -(n+1)b_n \end{cases}$$

Ce qui donne bien :

$$\forall n \geq 0, \quad (x^2-1)P_n'(x) = n \cdot b_{n+1}P_{n+1}(x) - (n+1)b_n P_{n-1}(x)$$

(c) Avec (b) et la récurrence (1), on a :

$$\begin{cases} ({x_{n,i}}^2-1)P_n'(x_{n,i}) = n \cdot b_{n+1}P_{n+1}(x_{n,i}) - (n+1)b_n P_{n-1}(x_{n,i}) \\ b_{n+1}P_{n+1}(x_{n,i}) + b_n P_{n-1}(x_{n,i}) = 0 \end{cases}$$

Ce qui donne en éliminant $P_{n+1}(x_{n,i})$, $({x_{n,i}}^2-1)P_n'(x_{n,i}) = -(2n+1)b_n P_{n-1}(x_{n,i})$, de sorte que :

$$\lambda_{n,i} = \frac{1-{x_{n,i}}^2}{(2n+1)\left(b_n P_{n-1}(x_{n,i})\right)^2} = \frac{(2n-1)\left(1-{x_{n,i}}^2\right)}{\left(nP_{n-1}(x_{n,i})\right)^2} \quad (1 \leq i \leq n)$$

Q4. Les racines de P_n sont les :

$$x_{n,i} = Cos\left(\frac{2i-1}{2n}\pi\right) = Cos(\theta_{n,i}) \quad (1 \leq i \leq n)$$

On a alors $P_n'(x_{n,i}) = \dfrac{n(-1)^{i-1}}{Sin(\theta_{n,i})}\sqrt{\dfrac{2}{\pi}}$ et $P_{n-1}(x_{n,i}) = (-1)^{i-1}Sin(\theta_{n,i})\sqrt{\dfrac{2}{\pi}}$, ce qui donne :

$$\lambda_{n,i} = \frac{\pi}{n} \quad (1 \leq i \leq n)$$

Q5. (a) Avec la formule de dérivation de Leibnitz, on a :

$$Q_n(x) = \sum_{k=0}^{n} \frac{(n+\alpha)(n+\alpha-1)\cdots(n+\alpha-(n-k-1))}{k!(n-k)!}(-1)^k x^k$$
$$= \sum_{k=0}^{n} \frac{(-1)^k}{k!}\frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+k+1)(n-k)!}x^k$$

où Γ désigne la fonction gamma d'Euler.

En utilisant les notations du problème 25 Q4, cela peut encore s'écrire :

$$Q_n(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} C_{n+\alpha}^{k+\alpha} x^k$$

On en déduit en particulier que le polynôme Q_n est de degré n, pour tout entier n, et donc que le système $\{Q_n;\ n \in IN\}$ est une base de $IR[x]$.

Pour $n \geq 1$ et $0 \leq k \leq n-1$, on obtient après k intégrations par parties :

$$\left\langle x^k \middle| Q_n \right\rangle = \frac{1}{n!}\int_0^{+\infty} x^k \left(\frac{d}{dx}\right)^n \left\{x^{n+\alpha}e^{-x}\right\}dx = (-1)^k \frac{k!}{n!}\int_0^{+\infty}\left(\frac{d}{dx}\right)^{n-k}\left\{x^{n+\alpha}e^{-x}\right\}dx$$

$$= (-1)^k \frac{k!}{n!}\left[\left(\frac{d}{dx}\right)^{n-k-1}\left\{x^{n+\alpha}e^{-x}\right\}\right]_0^{+\infty} = 0$$

Avec $Q_k \in IR_k[x]$, on déduit alors que $\left\langle Q_k \middle| Q_n \right\rangle = 0$ pour $0 \le k \le n-1$. Ce qui prouve que la famille $\{Q_n;\ n \in IN\}$ est orthogonale.

(b) Le polynôme xQ_n est dans $IR_{n+1}[x]$, il s'écrit donc :

$$xQ_n(x) = \sum_{k=0}^{n+1} \mu_k Q_k(x)$$

avec $\mu_k = \dfrac{\left\langle xQ_n \middle| Q_k \right\rangle}{\|Q_k\|^2}$ pour $0 \le k \le n+1$.

Pour $0 \le k \le n-2$, on a $\left\langle xQ_n \middle| Q_k \right\rangle = \left\langle Q_n \middle| xQ_k \right\rangle = 0$, puisque $xQ_k \in IR_{n-1}[x]$ et Q_n est orthogonal à $IR_{n-1}[x]$. Ce qui donne $\mu_k = 0$ pour $0 \le k \le n-2$.

On a donc :

$$(*) \quad xQ_n(x) = \mu_{n+1}Q_{n+1}(x) + \mu_n Q_n(x) + \mu_{n-1}Q_{n-1}(x)$$

En utilisant le développement de Q_n établi en (a), on voit que les coefficients de x^n et x^{n-1} dans Q_n sont donnés respectivement par $\beta_n^{(n)} = \dfrac{(-1)^n}{n!}$ et $\beta_{n-1}^{(n)} = \dfrac{(-1)^{n-1}}{(n-1)!}(\alpha+n)$. L'identification des coefficients de x^{n+1} et x^n dans (*) donne alors $\mu_{n+1} = \dfrac{\beta_n^{(n)}}{\beta_{n+1}^{(n+1)}} = -(n+1)$, $\mu_n = 2n+\alpha+1$.

Enfin en prenant $x = 0$ dans (*), on a $\mu_{n+1}Q_{n+1}(0) + \mu_n Q_n(0) + \mu_{n-1}Q_{n-1}(0) = 0$, avec $Q_k(0) = C_{k+\alpha}^{\alpha}$ pour tout $k \ge 0$. Ce qui donne $\mu_{n-1} = -(n+\alpha)$.

En définitive, les polynômes Q_n vérifient la relation de récurrence :

$$\begin{cases} Q_{-1}(x) = 0,\ Q_0(x) = 1 \\ -(n+1)Q_{n+1}(x) + (2n+\alpha+1)Q_n(x) - (n+\alpha)Q_{n-1}(x) = xQ_n(x) \quad (n \ge 0) \end{cases}$$

(c) On a, en reprenant les calculs de (a) avec $k = n$:

$$\|Q_n\|^2 = \beta_n^{(n)}\left\langle x^n \middle| Q_n \right\rangle = (-1)^n \beta_n^{(n)} \int_0^{+\infty} x^{n+\alpha}e^{-x}dx = (-1)^n \beta_n^{(n)}\Gamma(n+\alpha+1)$$

C'est-à-dire :

$$\|Q_n\|^2 = \frac{\Gamma(n+\alpha+1)}{n!}$$

(d) Pour $n = 0$, l'identité se réduit à $0 = 0$. On suppose donc $n \ge 1$.

En utilisant le développement de Q_n obtenu en (a), on a :

$$xQ_n'(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} C_{n+\alpha}^{k+\alpha} k \cdot x^k = Q_n(x) - \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} C_{n+\alpha}^{k+\alpha}(n-k)x^k$$

Et avec $C_{n+\alpha}^{k+\alpha}(n-k) = C_{n-1+\alpha}^{k+\alpha}(n+\alpha)$, on déduit que :

$$xQ_n'(x) = Q_n(x) - (n+\alpha)Q_{n-1}(x)$$

(e) Pour $1 \le i \le n$, on a $\lambda_{n,i} = \dfrac{\alpha_n^{(n)}}{\alpha_{n-1}^{(n-1)}} \dfrac{1}{P_n'(x_{n,i})P_{n-1}(x_{n,i})}$, avec $P_n = \dfrac{1}{\|Q_n\|}Q_n$ et $\alpha_n^{(n)} = \dfrac{1}{\|Q_n\|}\beta_n^{(n)}$, ce qui donne $\lambda_{n,i} = \dfrac{\beta_n^{(n)}}{\beta_{n-1}^{(n-1)}} \|Q_{n-1}\|^2 \dfrac{1}{Q_n'(x_{n,i})Q_{n-1}(x_{n,i})}$.

Enfin avec $Q_n'(x_{n,i}) = -\dfrac{n+\alpha}{x_{n,i}}Q_{n-1}(x_{n,i})$, on déduit que :

$$\lambda_{n,i} = \frac{\Gamma(n+\alpha)}{(n+\alpha)n!} \frac{x_{n,i}}{(Q_{n-1}(x_{n,i}))^2}$$

Q6. Soit H_{2n-1} le polynôme d'interpolation d'Hermite de f associé aux $x_{n,i}$ défini par :

$$\begin{cases} H_{2n-1} \in IR_{2n-1}[x] \\ H_{2n-1}{}^{(k)}(x_{n,i}) = f^{(k)}(x_{n,i}) \quad (k = 0,1;\ 1 \le i \le n) \end{cases}$$

Dans le problème 33, Q5 (c), on a vu que l'erreur d'interpolation est donnée par :

$$f(x) - H_{2n-1}(x) = \frac{1}{(2n)!} \frac{(P_n(x))^2}{\left(\alpha_n^{(n)}\right)^2} f^{(2n)}(\xi_x) \quad (\xi_x \in]a,b[)$$

En considérant que $\int_a^b H_{2n-1}(x)\pi(x)dx = \sum_{i=1}^n \lambda_{n,i}H_{2n-1}(x_{n,i}) = \sum_{i=1}^n \lambda_{n,i}f(x_{n,i})$, on déduit que l'erreur de quadrature dans la méthode de Gauss est donnée par :

$$\begin{aligned} \int_a^b f(x)\pi(x)dx - \sum_{i=1}^n \lambda_{n,i}f(x_{n,i}) &= \int_a^b (f(x) - H_{2n-1}(x))\pi(x)dx \\ &= \frac{1}{\left(\alpha_n^{(n)}\right)^2(2n)!} \int_a^b (P_n(x))^2 f^{(2n)}(\xi_x)\pi(x)dx \end{aligned}$$

Si on pose $g_n(x) = f^{(2n)}(\xi_x)$, on a alors :

$$g_n(x) = \left(\alpha_n^{(n)}\right)^2(2n)! \frac{f(x) - H_{2n-1}(x)}{(P_n(x))^2} \quad (x \in]a,b[- \{x_{n,1}, \cdots, x_{n,n}\})$$

Avec $f(x) - H_{2n-1}(x) = (x - x_{n,i})^2 h_n(x)$, où $\underset{x \to x_{n,i}}{Lim} h_n(x) = \dfrac{f''(x_{n,i}) - H_{2n-1}''(x_{n,i})}{2}$, on déduit que :

$$\underset{x \to x_{n,i}}{Lim} \frac{f(x) - H_{2n-1}(x)}{(P_n(x))^2} = \underset{x \to x_{n,i}}{Lim} \left(\frac{x - x_{n,i}}{P_n(x)}\right)^2 h_n(x) = \frac{f''(x_{n,i}) - H_{2n-1}''(x_{n,i})}{2\left(P_n'(x_{n,i})\right)^2}$$

C'est-à-dire que la fonction g_n se prolonge par continuité sur $]a,b[$.

D'autre part, la dérivée $f^{(2n)}$ est continue sur $[a,b]$ (i. e. continue sur $]a,b[$ avec des limites finies en *a* et *b*), donc il en est de même de g_n et on peut utiliser le théorème de la moyenne pour écrire que :

$$\int_a^b (P_n(x))^2 f^{(2n)}(\xi_x)\pi(x)dx = f^{(2n)}(\eta)\int_a^b (P_n(x))^2 \pi(x)dx = f^{(2n)}(\eta)\|P_n\|^2 = f^{(2n)}(\eta)$$

avec $\eta \in]a,b[$.

Ce qui donne bien :

$$\int_a^b f(x)\pi(x)dx = \sum_{i=1}^n \lambda_{n,i} f(x_{n,i}) + \frac{1}{\left(\alpha_n^{(n)}\right)^2}\frac{f^{(2n)}(\eta)}{(2n)!}$$

Problème 42 : Transformation de Fourier discrète. Algorithme de Cooley et Tukey

On désigne par $L^1(IR^+)$ l'espace vectoriel des (classes de) fonctions Lebesgue–intégrables sur $[0,+\infty[$. Cet espace est muni d'une structure d'espace de Banach avec la norme $f \mapsto \|f\|_1 = \int_0^{+\infty} |f(t)|dt$.

On rappelle que l'ensemble des fonctions indéfiniment dérivables et à support compact dans $]0,+\infty[$ est dense dans $\left(L^1(IR^+), \|\ \|_1\right)$.

La transformée de Fourier d'une fonction $f \in L^1(IR^+)$ est la fonction $\hat{f}$ définie par :

$$\forall x \in IR, \quad \hat{f}(x) = \int_0^{+\infty} f(t)e^{-2i\pi xt}dt$$

On approche cette intégrale par la fonction $\hat{f}_T$ définie, pour tout $T > 0$, par :

$$\forall x \in IR, \quad \hat{f}_T(x) = \int_0^{T} f(t)e^{-2i\pi xt}dt$$

Q1. *Montrer que pour toute fonction $f \in L^1(IR^+)$, la transformée de Fourier $\hat{f}$ est continue et nulle à l'infini.*

Q2. *Dans cette question, on se donne une fonction $f \in L^1(IR^+)$ continue et telle que $\underset{t\to+\infty}{Lim} f(t) = 0$.*

(a) Montrer que pour tout $\varepsilon > 0$, on peut trouver un réel $T > 0$ tel que :

$$(1)\quad \begin{cases} \forall t \geq T, & |f(t)| \leq \varepsilon \\ \forall x \in IR, & \left|\hat{f}(x) - \hat{f}_T(x)\right| \leq \varepsilon \end{cases}$$

(b) Pour tous $\varepsilon > 0$, $T > 0$ définis comme en (a) et pour tout entier $n \geq 1$, on note $\delta = \dfrac{T}{n}$ et $X = \dfrac{1}{2\delta}$ (fréquence de Nyquist).

Montrer qu'on peut choisir n assez grand de sorte que :

$$(2)\quad |x| \geq X \Rightarrow \left|\hat{f}(x)\right| \leq \varepsilon$$

(c) Pour ε, T, X et n = 2p entier pair choisis de sorte que (1) et (2) soient vérifiés, on définit des subdivisions de $[0,T]$ *et* $[-X,X]$ *en posant :*

$$\begin{cases} \delta = \dfrac{T}{n}, \quad t_j = j\cdot\delta \quad (0 \le j \le n-1) \\ h = \dfrac{2X}{n}, \quad x_k = k\cdot h \quad \left(-\dfrac{n}{2} \le k \le \dfrac{n}{2}\right) \end{cases}$$

Montrer, en utilisant la méthode des rectangles, que le calcul approché des $\hat{f}(x_k)$ *pour tout* $k \in \left\{-\dfrac{n}{2},\cdots,\dfrac{n}{2}\right\}$ *se ramène au calcul des quantités :*

$$y_k = \sum_{j=0}^{n-1} (\omega_n)^{j\cdot k} f_j \quad (0 \le k \le n-1)$$

où on a posé $\omega_n = e^{-\frac{2i\pi}{n}}$ *(racine* $n^{\text{ème}}$ *de l'unité) et* $f_j = f(t_j)$ *(*$0 \le j \le n-1$*).*

Dans ce qui suit, n désigne un entier positif non nul et $\omega_n = e^{-\frac{2i\pi}{n}}$ *une racine* $n^{\text{ème}}$ *de l'unité.*

La transformation de Fourier discrète est l'application linéaire $F : C^n \to C^n$ *qui associe à tout vecteur* $x = (x_0, x_1, \cdots, x_{n-1}) \in C^n$ *le vecteur* $y = (y_0, y_1, \cdots, y_{n-1}) \in C^n$ *dont les composantes sont définies par :*

$$y_k = \sum_{j=0}^{n-1} (\omega_n)^{j\cdot k} x_j \quad (0 \le k \le n-1)$$

Q3. *(a) Montrer que la transformation de Fourier discrète est un isomorphisme de* C^n *et calculer son inverse.*
(b) Montrer que si $y = F(x)$*, alors :*

$$\sum_{j=0}^{n-1} |x_j|^2 = \frac{1}{n}\sum_{k=0}^{n-1} |y_k|^2$$

(Formule de Plancherel).
(c) Donner un algorithme simple de calcul de la transformée de Fourier discrète ne faisant intervenir que les opérations d'addition et de multiplication de nombres complexes.

On calculera le nombre d'opérations élémentaires (additions et multiplications complexes) que nécessite un tel algorithme.

Dans ce qui suit, on suppose que $n = 2^p$ *avec* $p \ge 1$.

On note, pour tout entier $r \in \{0,1,\cdots,2^p - 1\}$ *:*

$$r = \sum_{i=0}^{p-1} r_i 2^i \quad (r_i \in \{0,1\})$$

son écriture en base 2.

Un nombre complexe indexé par l'entier r sera noté $z_r = z_{r_0, r_1, \cdots, r_{p-1}}$ *et une somme sur r allant de 0 à n – 1 peut alors s'écrire :*

$$\sum_{r=0}^{2^p-1} z_r = \sum_{r_0=0}^{1}\sum_{r_1=0}^{1}\cdots\sum_{r_{p-1}=0}^{1} z_{r_0,r_1,\cdots,r_{p-1}}$$

Q4. *Soient* $x \in C^n$, $y = F(x) \in C^n$ *et* $k \in \{0,1,\cdots,2^p-1\}$ *d'écriture binaire* $k = \sum_{i=0}^{p-1} k_i 2^i$. *Montrer que :*

$$(3) \quad y_k = \sum_{j_0=0}^{1} \omega_n^{2^{p-1}k_{p-1}j_0} \sum_{j_1=0}^{1} \omega_n^{2^{p-2}k_{p-2}(j_0+2j_1)} \cdots \sum_{j_{p-1}=0}^{1} \omega_n^{k_0(j_0+\cdots+2^{p-1}j_{p-1})} x_{j_0,\cdots,j_{p-1}}$$

Q5. *Déduire de (3) un algorithme rapide de calcul de la transformée de Fourier discrète.*

On calculera le nombre d'opérations élémentaires (additions et multiplications complexes) que nécessite un tel algorithme.

Solution

Q1. Pour presque tout $t \in IR^+$, la fonction $t \mapsto f(t)e^{-2i\pi xt}$ est continue sur IR et $\left|f(t)e^{-2i\pi xt}\right| = |f(t)|$ avec $|f| \in L^1(IR^+)$. On en déduit alors que la fonction $\hat{f}$ est continue sur IR (théorème de Lebesgue).

Si la fonction f est indéfiniment dérivable à support compact dans l'intervalle $]0,a[$, alors une intégration par parties donne, pour $x \neq 0$:

$$\hat{f}(x) = \int_0^a f(t)e^{-2i\pi xt}dt = \left[f(t)\frac{e^{-2i\pi xt}}{-2i\pi x}\right]_0^a + \frac{1}{2i\pi x}\int_0^a f'(t)e^{-2i\pi xt}dt$$

Soit :

$$\left|\hat{f}(x)\right| = \frac{1}{2\pi|x|}\left|\int_0^a f'(t)e^{-2i\pi xt}dt\right| \leq \frac{1}{2\pi|x|}\|f'\|_1 \underset{|x|\to+\infty}{\longrightarrow} 0$$

Si $f \in L^1(IR^+)$, alors $f = \underset{n\to+\infty}{Lim} f_n$ dans $\left(L^1(IR^+), \| \|_1\right)$ où $(f_n)_{n\in IN}$ est une suite de fonctions indéfiniment dérivables à support compact. Pour tout réel $x \neq 0$, on a alors $\left|\hat{f}(x)\right| \leq \|f - f_n\|_1 + \left|\hat{f}_n(x)\right|$.

Pour $\varepsilon > 0$, on peut trouver $n_0 \in IN$ tel que $\|f - f_n\|_1 \leq \frac{\varepsilon}{2}$ et $X_0 > 0$ tel que pour $|x| \geq X_0$ on ait $\left|\hat{f}_{n_0}(x)\right| \leq \frac{\varepsilon}{2}$, de sorte que $\left|\hat{f}(x)\right| \leq \varepsilon$ pour $|x| \geq X_0$.

On a donc ainsi prouvé que $\underset{|x|\to+\infty}{Lim} \hat{f}(x) = 0$.

Q2. (a) Comme $\underset{t\to+\infty}{Lim} f(t) = 0$, on peut trouver $T_1 > 0$ tel que $|f(t)| \leq \varepsilon$ pour $t \geq T_1$.

D'autre part, pour tout $T > 0$ et tout $x \in IR$, on a $\left|\hat{f}(x) - \hat{f}_T(x)\right| \leq \int_T^{+\infty}|f(t)|dt \underset{T\to+\infty}{\longrightarrow} 0$, on peut donc trouver $T_2 > 0$ tel que :

$$\forall x \in IR, \quad \forall T \geq T_2, \quad \left|\hat{f}(x) - \hat{f}_T(x)\right| \leq \varepsilon$$

Il suffit donc de prendre $T = Max\{T_1, T_2\}$.

(b) On a vu en Q1 que $\underset{|x|\to+\infty}{Lim}\hat{f}(x) = 0$, on peut donc trouver $A > 0$ tel que $\left|\hat{f}(x)\right| \le \varepsilon$ pour $|x| \ge A$. Il suffit alors de prendre $n \ge 2TA$ pour avoir $X \ge A$ et $\left|\hat{f}(x)\right| \le \varepsilon$ pour $|x| \ge X$.

(c) Chaque $\hat{f}(x_k)$, pour $-\frac{n}{2} \le k \le \frac{n}{2}$, est approché (à ε près) par $\hat{f}_T(x_k)$. L'utilisation la méthode des rectangles à gauche pour le calcul approché de $\hat{f}_T(x_k)$ donne alors :

$$\hat{f}_T(x_k) \cong \frac{T}{n}\sum_{j=0}^{n-1} f(t_j)e^{-2i\pi x_k t_j}$$

avec $x_k t_j = \frac{k \cdot j}{n}$. Ce qui s'écrit aussi :

$$\hat{f}_T(x_k) \cong \delta\sum_{j=0}^{n-1} \omega_n^{\ j\cdot k} f_j = \delta\cdot y_k \quad \left(k = -\frac{n}{2}, \cdots, \frac{n}{2}\right)$$

Avec la 2π–périodicité de la fonction $t \mapsto e^{it}$, on voit que les y_k sont définis pour tout k dans Z et qu'ils forment une suite n–périodique ($y_{k+n} = y_k$ pour tout $k \in Z$).

Il nous suffit donc, pour déterminer complètement cette suite, de calculer les y_k pour $k \in \{0, 1, \cdots, n-1\}$.

Du point de vue pratique, les valeurs approchées des $\hat{f}(x_k)$ sont données par :

$$\begin{cases} k = 0 & \text{pour } x_k = 0 \\ k = 1,\ 2, \cdots,\ \frac{n}{2} - 1 & \text{pour } x_k \in]0, X[\\ k = \frac{n}{2} + 1, \cdots,\ n\ -\ 1 & \text{pour } x_k \in]-X, 0[\\ k = \frac{n}{2} & \text{pour } x_k = X \text{ et } x_k = -X \end{cases}$$

Q3. (a) La matrice de F dans la base canonique de C^n est :

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ . & . & \cdots & . \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$$

C'est une matrice de Van Der Monde de déterminant $Dét(A) = \prod_{0 \le j < i \le n-1}\left(\omega_n^{\ i} - \omega_n^{\ j}\right) \ne 0$. On déduit donc que F est un automorphisme.

Pour tout $j \in \{0, 1, \cdots, n-1\}$, on a :

$$\sum_{k=0}^{n-1} \omega_n^{\ -j\cdot k} y_k = \sum_{k=0}^{n-1} \omega_n^{\ -j\cdot k} \sum_{i=0}^{n-1} \omega_n^{\ i\cdot k} x_i = \sum_{i=0}^{n-1}\left(\sum_{k=0}^{n-1} \omega_n^{\ (i-j)k}\right)x_i$$

Pour $i \neq j$ dans $\{0,1,\cdots,n-1\}$, on a $0 < |i-j| < n$ et $\omega_n^{\ i-j} \neq 1$, de sorte que :

$$\sum_{k=0}^{n-1} \omega_n^{\ (i-j)k} = \frac{1-\omega_n^{\ (i-j)n}}{1-\omega_n^{\ i-j}} = 0$$

Pour $i = j$, cette somme vaut n. On a alors, en définitive :

$$\sum_{k=0}^{n-1} \omega_n^{\ -j\cdot k} y_k = n \cdot x_i$$

C'est-à-dire que la transformation de Fourier discrète est un isomorphisme de C^n, d'inverse F^{-1} définie par :

$$\left(x = F^{-1}(y)\right) \Leftrightarrow \left(x_j = \frac{1}{n}\sum_{k=0}^{n-1} \omega_n^{\ -j\cdot k} y_k \quad (0 \leq j \leq n-1)\right)$$

(b) Pour tout $k \in \{0,1,\cdots,n-1\}$, on a :

$$|y_k|^2 = \sum_{j=0}^{n-1} \omega_n^{\ j\cdot k} x_j \sum_{i=0}^{n-1} \omega_n^{\ -i\cdot k} \overline{x_i} = \sum_{j=0}^{n-1}\left(\sum_{i=0}^{n-1} \omega_n^{\ (j-i)k} \overline{x_i}\right) x_j$$

Et :

$$\sum_{k=0}^{n-1} |y_k|^2 = \sum_{0 \leq i,j \leq n}\left(\sum_{k=0}^{n-1} \omega_n^{\ (j-i)k}\right) \overline{x_i} x_j$$

Avec :

$$\sum_{k=0}^{n-1} \omega_n^{\ (j-i)k} = \begin{cases} \dfrac{1-\omega_n^{\ (i-j)n}}{1-\omega_n^{\ i-j}} = 0 \text{ si } j \neq i \\ 0 \text{ si } j = i \end{cases}$$

On en déduit alors que :

$$\sum_{k=0}^{n-1} |y_k|^2 = n \sum_{j=0}^{n-1} |x_j|^2$$

(c) Tout d'abord, on écrit une procédure de calcul des $\omega_n^{\ r}$, pour $r \in \{0,1,\cdots,n-1\}$, ces valeurs étant stockées dans un vecteur w de C^n.

Ce qui donne, en disposant d'opérations sur les nombres complexes :

PROCEDURE PuissancesOmega(Entrée n : Entier ; Sortie w : VecteurComplexe) ;
Début
$w_0 = 1$; $w_1 = Cos(2\pi/n) - i\cdot Sin(2\pi/n)$;
Pour r allant de 2 à n - 1 Faire $w_r = w_{r-1}\cdot w_1$;
Fin ;

Ce calcul nécessitant $n - 2$ multiplications complexes.

En remarquant que pour $0 \leq j,k \leq n-1$, on a $\omega_n^{\ k\cdot j} = \omega_n^{\ r}$, où r est le reste dans la division euclidienne de $j\cdot k$ par n, on déduit la procédure suivante de calcul du vecteur y :

PROCEDURE FourierDiscret_1(Entrée n : Entier ; w, x : VecteurComplexe ;
Sortie y : VecteurComplexe) ;
Début

Pour k Allant de 0 à n – 1 Faire $y_k = x_0$ *;*
Pour j Allant de 1 à n – 1 Faire $y_0 = y_0 + x_j$ *;*
Pour k Allant de 1 à n – 1 Faire
Début
  Pour j Allant de 1 à n – 1 Faire
  Début
    $r = j \cdot k$ *Modulo n ;*
    $y_k = y_k + w_r \cdot x_j$ *;*
  Fin ;
Fin ;
Fin ;

Le calcul de y_0 nécessite $n - 1$ additions complexes. Pour $k \in \{0,1,\cdots,n-1\}$, le calcul de chaque y_k nécessite $n - 1$ multiplications complexes et $n - 1$ additions complexes. Ce qui donne en ajoutant le nombre d'opérations nécessaires au calcul de w un total de $(n-2)+(n-1)+2(n-1)^2$ opérations complexes.

Remarque — On peut aussi utiliser l'algorithme de Hörner pour évaluer chaque y_k considéré comme un polynôme en $\omega_n{}^k$. Là encore le nombre d'opérations élémentaires est de l'ordre de n^2.

Q4. On a, pour tout $k \in \{0,1,\cdots,n-1\}$:

$$y_k = \sum_{j_0=0}^{1}\sum_{j_1=0}^{1}\cdots\sum_{j_{p-1}=0}^{1}\omega_n{}^{k\cdot j}x_{j_0,\cdots,j_{p-1}}$$

En remarquant que, pour $j = \sum_{i=0}^{p-1} j_i 2^i$ et $k = \sum_{i=0}^{p-1} k_i 2^i$ dans $\{0,1,\cdots,n-1\}$, on a :

$$j \cdot k \equiv k_0\left(j_0 + \cdots + j_{p-1}2^{p-1}\right) + 2k_1\left(j_0 + \cdots + j_{p-2}2^{p-2}\right) + \cdots + 2^{p-1}k_{p-1}j_0 \quad (Mod \quad n)$$

on déduit que :

$$\omega_n{}^{k\cdot j} = \omega_n{}^{2^{p-1}k_{p-1}j_0}\,\omega_n{}^{2^{p-2}k_{p-2}(j_0+2j_1)}\cdots\omega_n{}^{k_0\left(j_0+\cdots+2^{p-1}j_{p-1}\right)}$$

et donc que :

$$y_k = \sum_{j_0=0}^{1}\omega_n{}^{2^{p-1}k_{p-1}j_0}\sum_{j_1=0}^{1}\omega_n{}^{2^{p-2}k_{p-2}(j_0+2\cdot j_1)}\cdots\sum_{j_{p-1}=0}^{1}\omega_n{}^{k_0\left(j_0+\cdots+2^{p-1}j_{p-1}\right)}x_{j_0,\cdots,j_{p-1}}$$

Q5. Avec les formules (3), on déduit le schéma de calcul suivant.

Etape 0 — On pose $x^{(0)} = x$.

Etape r + 1 (r = 0, 1, ..., p – 1) — L'étape précédente a donné le vecteur $x^{(r)}$ de composantes $x^{(r)}_{j_0,\cdots,j_{p-r-1},k_{r-1},\cdots,k_0}$ et on veut calculer celles de $x^{(r+1)}$ notées $x^{(r+1)}_{j_0,\cdots,j_{p-r-2},k_r,\cdots,k_0}$.

Avec (3), on déduit que :

$$x^{(r+1)}_{j_0,\cdots,j_{p-r-2},k_r,\cdots,k_0} = \sum_{j_{p-r-1}=0}^{1}\omega_n{}^{2^r k_r\left(j_0+\cdots+2^{p-r-1}j_{p-r-1}\right)}x^{(r)}_{j_0,\cdots,j_{p-r-1},k_{r-1},\cdots,k_0}$$

On pose alors :

$$m = \frac{n}{2^{r+1}} = 2^{p-r-1}$$
$$L = j_0 + \cdots + j_{p-r-2} 2^{p-r-2} \in \{0, \cdots, m-1\}$$
$$h = k_{r-1} + \cdots + k_0 2^{r-1} \in \left\{0, \cdots, 2^r - 1 = \frac{n}{2m} - 1\right\}$$
$$i = j_0 + \cdots + j_{p-r-2} 2^{p-r-2} + 0 \cdot 2^{p-r-1} + k_{r-1} 2^{p-r} + \cdots + k_0 2^{p-1} = L + 2 \cdot m \cdot h$$
$$j = j_0 + \cdots + j_{p-r-2} 2^{p-r-2} + 1 \cdot 2^{p-r-1} + k_{r-1} 2^{p-r} + \cdots + k_0 2^{p-1} = i + m$$

et les formules précédentes vont s'écrire :

$$\begin{cases} x_i^{(r+1)} = x_i^{(r)} + x_j^{(r)} & (k_r = 0) \\ x_j^{(r+1)} = \omega_n^{2^r L} x_i^{(r)} + \omega_n^{2^r\left(L + 2^{p-r-1}\right)} x_j^{(r)} & (k_r = 1) \end{cases}$$

Avec $\omega_n^{2^{r+p-r-1}} = -1$ et $\omega_n^{2^r} = \omega_{2^{p-r}}$, on déduit alors que :

$$\begin{cases} x_i^{(r+1)} = x_i^{(r)} + x_j^{(r)} \\ x_j^{(r+1)} = \omega_{2^{p-r}}^{L} \left(x_i^{(r)} - x_j^{(r)}\right) \end{cases}$$

Et au bout de p étapes, on a $x_{k_{p-1}, \cdots, k_0}^{(p)} = y_{\sigma(k)}$, où σ est l'inversion binaire définie par $\sigma\left(k_0 + \cdots + k_{p-1} 2^{p-1}\right) = k_{p-1} + \cdots + k_0 2^{p-1}$.

Pour retrouver les composantes du vecteur y dans le bon ordre on devra donc faire une inversion binaire sur celles de $x^{(p)}$.

L'algorithme ainsi obtenu est dû à Cooley et Tukey.

En notant, pour $p \geq 1$, A_p le nombre d'additions complexes et M_p le nombre de multiplications complexes dans cet algorithme, on a la récurrence :

$$\begin{cases} A_1 = 2; \ \ M_1 = 0 \\ A_p = 2 A_{p-1} + 2^p; \ \ M_p = 2 M_{p-1} + 2^{p-1} \end{cases}$$

Ce qui donne :

$$\begin{cases} M_p = A_{p-1} \\ A_p = p 2^p \end{cases}$$

C'est-à-dire qu'on a un nombre d'opérations élémentaires de l'ordre de $n \cdot Log_2(n)$.

Remarque — Cet algorithme peut également être utilisé pour le calcul des coefficients de Fourier d'une fonction périodique.

Si $f : IR \to C$ est continue par morceaux et T–périodique alors ses coefficients de Fourier complexes sont définis par :

$$c_k = \frac{1}{T} \int_0^T f(t) e^{-\frac{2ik\pi}{T} t} dt = \frac{1}{T} \int_0^T f(t) e^{-2i\pi x_k t} dt = \frac{1}{T} \hat{f}(x_k) \quad \left(-\frac{n}{2} \leq k \leq \frac{n}{2}\right)$$

en notant encore f la fonction coïncidant avec la précédente sur $[0, T]$ et nulle ailleurs.

Problème 43 : Transformation de Fourier discrète. Agrégation 1993, extrait

On désigne par n et N des entiers supérieurs à 1, tels que $n = 2^N - 1$.

L'espace C^{n+1} est rapporté à sa base canonique et muni de la norme usuelle définie par $\|x\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}$.

On convient de noter $\omega_{n+1} = e^{\frac{2i\pi}{n+1}}$ et d'appeler :

— Polynôme associé à un élément $a = (a_0, a_1, \cdots, a_n)$ de C^{n+1} le polynôme :

$$A(x) = a_0 + a_1 x + \cdots + a_n x^n$$

— Transformation de Fourier discrète de C^{n+1} l'application ϕ_{n+1} associant à tout élément $a = (a_0, a_1, \cdots, a_n)$ de C^{n+1} l'élément suivant de C^{n+1} :

$$\phi_{n+1}(a) = \left(A(1), A(\omega_{n+1}), \cdots, A\left({\omega_{n+1}}^p\right), \cdots, A\left({\omega_{n+1}}^n\right)\right)$$

***Q1.** (a) Etablir que ϕ_{n+1} est un automorphisme de l'espace vectoriel C^{n+1}.*

(b) Exprimer pour tout $a \in C^{n+1}$ la norme de $\phi_{n+1}(a)$ en fonction de la norme de a.

(c)Expliciter la matrice M_{n+1} de l'automorphisme ϕ_{n+1} dans la base canonique.

En déduire que l'inverse de ϕ_{n+1} est défini par :

$${\phi_{n+1}}^{-1}(a) = \frac{1}{n+1}\left(A(1), A\left({\omega_{n+1}}^{-1}\right), \cdots, A\left({\omega_{n+1}}^{-p}\right), \cdots, A\left({\omega_{n+1}}^{-n}\right)\right)$$

***Q2.** Etant donné un élément $a = (a_0, a_1, \cdots, a_n)$ de C^{n+1}, on lui associe les deux éléments notés $b = (a_0, a_2, a_4, \cdots, a_{n-1})$ et $c = (a_1, a_3, a_5, \cdots, a_n)$ de $C^{\frac{n+1}{2}}$ et l'on considère alors leurs transformées de Fourier discrètes, notées respectivement :*

$$\phi_{n+1}(a) = (\alpha_0, \alpha_1, \cdots, \alpha_n) \text{ et } \begin{cases} \phi_{\frac{n+1}{2}}(b) = \left(\beta_0, \beta_1, \cdots, \beta_{\frac{n-1}{2}}\right) \\ \phi_{\frac{n+1}{2}}(c) = \left(\gamma_0, \gamma_1, \cdots, \gamma_{\frac{n-1}{2}}\right) \end{cases}$$

(a) On suppose connus $\phi_{\frac{n+1}{2}}(b)$ et $\phi_{\frac{n+1}{2}}(c)$ et l'on considère l'algorithme suivant (dans lequel la notation $A \leftarrow B$ indique que l'on affecte à la variable complexe A la valeur complexe B) :

$E \leftarrow 1$;

pour p de 0 à** (n – 1)/2 **faire :

début

$F \leftarrow E \cdot \gamma_p$; $\alpha \leftarrow \beta_p + F$; $\alpha_{p+\frac{n+1}{2}} \leftarrow \beta_p - F$; $E \leftarrow \omega_{n+1} \cdot E$;

fin.

Prouver que l'élément $(\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, \alpha_n)$ *de* C^{n+1} *ainsi obtenu est égal à* $\phi_{n+1}(a)$.

(b) Soit u_N *et* v_N *les nombres des additions et multiplications complexes nécessaires au calcul de* $\phi_{n+1}(a)$ *par l'algorithme récursif décrit ci–dessous (et ne tenant pas compte du calcul des nombres* ω_{n+1}*, supposés connus) :*

— Si $N = 1$*, on calcule* $\phi_2(a_0, a_1) = (a_0 + a_1, a_0 - a_1)$ *et donc* $u_1 = 2$ *et* $v_1 = 0$.

— Sinon, on calcule récursivement $\phi_{\frac{n+1}{2}}(b)$ *et* $\phi_{\frac{n+1}{2}}(c)$ *en effectuant pour chacun d'eux* u_{N-1} *additions complexes et* v_{N-1} *multiplications complexes et l'on obtient* $\phi_{n+1}(a)$ *par l'algorithme développé ci–dessus à la question (a).*

Exprimer u_N *et* v_N *en fonction de* u_{N-1}, v_{N-1} *et N.*

En déduire en fonction de N, puis de n, le nombre d'additions et multiplications complexes (resp. réelles) nécessaires au calcul de $\phi_{n+1}(a)$.

(c) Comparer enfin ces nombre d'opérations à ceux qui résulteraient du calcul de $\phi_{n+1}(a)$ *par utilisation répétée de l'algorithme de Hörner.*

On rappelle que le calcul de $P(x) = a_p x^p + \cdots + a_0$ *par cet algorithme s'effectue de la façon suivante :*

$$P(x) = \left(\cdots\left(\left(a_p x + a_{p-1}\right)x + a_{p-2}\right)\cdots + a_1\right)x + a_0$$

Solution

Q1. (a) Comme dans le problème 42, Q3 (a) on peut remarquer que la matrice de ϕ_{n+1} dans la base canonique de C^{n+1} est une matrice de Van Der Monde de déterminant non nul, ce qui prouve que c'est un automorphisme de C^{n+1}.

On peut aussi calculer le noyau de ϕ_{n+1}. Il est formé des vecteurs $a \in C^{n+1}$ tels que le polynôme associé A s'annule sur les $n+1$ racines $(n+1)^{\text{ème}}$ de l'unité 1, $\omega_{n+1}, \cdots, \omega_{n+1}{}^n$, c'est-à-dire que A est un polynôme de degré au plus n qui s'annule en $n+1$ points distincts, c'est donc le polynôme nul et $a = 0$. Donc $Ker(\phi_{n+1}) = \{0\}$ et ϕ_{n+1} est un automorphisme de C^{n+1}.

(b) On a la formule de Plancherel (problème 42, Q3 (b)) :

$$\forall a \in C^{n+1}, \quad \left\|\phi_{n+1}(a)\right\|_2{}^2 = (n+1)\left\|a\right\|_2{}^2$$

C'est-à-dire que ϕ_{n+1} est une similitude de rapport $\sqrt{n+1}$.

(c) La matrice de ϕ_{n+1} dans la base canonique de C^{n+1} est une matrice de Van Der Monde :

$$M_{n+1} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_{n+1} & \cdots & \omega_{n+1}{}^n \\ . & . & \cdots & . \\ 1 & \omega_{n+1}{}^n & \cdots & \omega_{n+1}{}^{n^2} \end{pmatrix}$$

La formule de Plancherel nous dit que l'automorphisme $\dfrac{1}{\sqrt{n+1}}\phi_{n+1}$ est unitaire, son inverse est donc égal à son adjoint, de matrice $\dfrac{1}{\sqrt{n+1}}{}^t\overline{M}_{n+1} = \dfrac{1}{\sqrt{n+1}}\overline{M}_{n+1}$ (matrice complexe conjuguée).

C'est-à-dire que $\phi_{n+1}{}^{-1}$ a pour matrice $\dfrac{1}{n+1}\overline{M}_{n+1}$ et son expression est :

$$\phi_{n+1}{}^{-1}(a) = \frac{1}{n+1}\left(A(1), A(\overline{\omega}_{n+1}), \cdots, A(\overline{\omega}_{n+1}{}^{n})\right) = \frac{1}{n+1}\left(A(1), A(\omega_{n+1}{}^{-1}), \cdots, A(\omega_{n+1}{}^{-n})\right)$$

Q2. (a) L'algorithme proposé est dû à Cooley et Tukey (problème 42, Q5).

En notant F_p la valeur de la variable F à l'étape $p \in \left\{0, 1, \cdots, \dfrac{n-1}{2}\right\}$ de la boucle « pour », on a :

$$F_p = \omega_{n+1}{}^{p}\gamma_p, \quad \alpha_p = \beta_p + \omega_{n+1}{}^{p}\gamma_p$$

avec :

$$\beta_p = B\left(\omega_{\frac{n+1}{2}}{}^{p}\right) = \sum_{k=0}^{(n-1)/2} a_{2k}\left(\omega_{\frac{n+1}{2}}{}^{p}\right)^k, \quad \gamma_p = C\left(\omega_{\frac{n+1}{2}}{}^{p}\right) = \sum_{k=0}^{(n-1)/2} a_{2k+1}\left(\omega_{\frac{n+1}{2}}{}^{p}\right)^k$$

En remarquant que $\omega_{\frac{n+1}{2}} = \omega_{n+1}{}^{2}$, on déduit que :

$$\alpha_p = \sum_{k=0}^{(n-1)/2} a_{2k}\left(\omega_{n+1}{}^{p}\right)^{2k} + \sum_{k=0}^{(n-1)/2} a_{2k+1}\left(\omega_{n+1}{}^{p}\right)^{2k+1} = \sum_{j=0}^{n} a_j\left(\omega_{n+1}{}^{p}\right)^j = A\left(\omega_{n+1}{}^{p}\right)$$

pour $p = 0, 1, \cdots, \dfrac{n-1}{2}$.

De façon analogue, on voit que :

$$\alpha_{p+\frac{n+1}{2}} = \sum_{j=0}^{n} a_j\left(-\omega_{n+1}{}^{p}\right)^j = A\left(-\omega_{n+1}{}^{p}\right) = A\left(\omega_{n+1}{}^{p+\frac{n+1}{2}}\right)$$

pour $p + \dfrac{n+1}{2} = \dfrac{n-1}{2} + 1, \cdots, n$.

On a donc bien $(\alpha_0, \cdots, \alpha_n) = \phi_{n+1}(a)$ en fin de boucle.

(b) Pour $\phi_{\frac{n+1}{2}}(b)$ et $\phi_{\frac{n+1}{2}}(c)$ connus, le calcul de $\phi_{n+1}(a)$ par l'algorithme du (a) nécessite $n+1$ additions complexes et $n+1$ multiplications complexes.

C'est-à-dire que les suites $(u_N)_{N\geq 1}$ et $(v_N)_{N\geq 1}$ vérifient la même relation de récurrence :

$$x_N = 2x_{N-1} + n + 1 = 2x_{N-1} + 2^N \quad (N \geq 2)$$

La suite $(y_N)_{N\geq 1}$ définie par $y_N = \dfrac{x_N}{2^N}$ est arithmétique de raison 1.

Avec les conditions initiales $u_1 = 2$ et $v_1 = 0$, on déduit alors que les nombres d'additions et de multiplications complexes sont données par :

$$u_N = 2^N N = (n+1)\log_2(n+1)$$
$$v_N = 2^N(N-1) = (n+1)\{\log_2(n+1) - 1\}$$

On a donc un nombre d'opérations complexes de l'ordre de $n\log_2(n)$.

(c) Le calcul direct de $\phi_{n+1}(a) = \left(A(1), A(\omega_{n+1}), \cdots, A\left({\omega_{n+1}}^p\right), \cdots, A\left({\omega_{n+1}}^n\right)\right)$ nécessite $n + 1$ évaluations du polynôme A. Chacune des évaluations $A\left({\omega_{n+1}}^p\right)$ par l'algorithme de Hörner nécessite n additions et n multiplications. Ce qui donne un total de $2n(n+1)$ opérations complexes, soit un nombre d'opérations de l'ordre de n^2.

CHAPITRE 7

Systèmes différentiels

Problème 44 : Problème de Cauchy. Méthodes à un pas

Dans tout le problème, on désigne par $I = [a,b]$ un intervalle fermé borné, p un entier strictement positif, y_a un vecteur de IR^p et $f : I \times IR^p \to IR^p$ une fonction continue.

L'espace vectoriel IR^p est muni d'une norme $y \mapsto \|y\|$.

Le problème de Cauchy associé à I, y_a et f consiste à trouver une fonction y dans l'espace $C^1(I, IR^p)$ des fonctions continûment dérivables de I dans IR^p telle que :

$$(1) \quad \begin{cases} y(a) = y_a \\ \forall t \in I,\ y'(t) = f(t, y(t)) \end{cases}$$

Q1. *Soient (E,d) un espace métrique complet et $\varphi : E \to E$.*

Montrer que s'il existe un entier k > 0 tel que l'itérée $\varphi^k = \varphi \circ \varphi \circ \cdots \circ \varphi$ (k fois) soit contractante, alors φ admet un unique point fixe dans E.

On rappelle qu'une application $\psi : E \to E$ est contractante s 'il existe un réel $\lambda \in [0,1[$ tel que :

$$\forall (y,z) \in E \times E, \quad d(\psi(z), \psi(y)) \le \lambda d(z,y)$$

On suppose de plus, dans tout ce qui suit, que la fonction f est lipschitzienne par rapport à sa deuxième variable, c'est-à-dire qu'il existe un réel $\lambda \ge 0$ tel que :

$$\forall t \in I, \forall (y,z) \in IR^p \times IR^p, \quad \|f(t,z) - f(t,y)\| \le \lambda \|z - y\|$$

Q2. *Montrer que pour tous $t_0 \in I$ et $y_0 \in IR^p$, le problème de Cauchy :*

$$(2) \quad \begin{cases} y(t_0) = y_0 \\ \forall t \in I,\ y'(t) = f(t, y(t)) \end{cases}$$

admet une unique solution $y \in C^1(I, IR^p)$ (théorème de Cauchy–Lipschitz).

A tout entier n > 0, on associe une subdivision de l'intervalle I en posant :

$$h_n = \frac{b-a}{n}, \quad t_k = a + k \cdot h_n \quad (0 \leq k \leq n)$$

Dans tout ce qui suit, on note $J = [0, b-a]$ *et on se donne une fonction continue* $\phi : I \times IR^p \times J \to IR^p$. *On associe à cette fonction la « méthode à un pas » définie par la relation de récurrence :*

$$(3) \quad \begin{cases} y_0 = y_a + \zeta_n \\ y_{k+1} = y_k + h_n \phi(t_k, y_k, h_n) \quad (0 \leq k \leq n-1) \end{cases}$$

où ζ_n *est donné dans* IR^p.

Pour toute solution y du système différentiel $y'(t) = f(t, y(t))$ *et pour tout* $k \in \{0, 1, \cdots, n-1\}$, *on pose :*

$$\varepsilon_k = \frac{1}{h_n}\{y(t_{k+1}) - y(t_k)\} - \phi(t_k, y(t_k), h_n)$$

Cette quantité est appelée « erreur de consistance » de la méthode en t_k.

On dit que la méthode à un pas associée à ϕ *est :*

— consistante avec le problème de Cauchy (1) si :

$$\lim_{n \to +\infty}\left(\max_{0 \leq k \leq n-1} \|\varepsilon_k\|\right) = 0$$

pour toute solution y du système différentiel $y'(t) = f(t, y(t))$.

— stable s'il existe deux constantes M_1 *et* M_2, *indépendantes de n, telles que pour toutes suites* $(y_k)_{0 \leq k \leq n}$, $(z_k)_{0 \leq k \leq n}$ *et* $(\varepsilon_k)_{0 \leq k \leq n-1}$ *vérifiant :*

$$\begin{cases} y_{k+1} = y_k + h_n \phi(t_k, y_k, h_n) \\ z_{k+1} = z_k + h_n \{\phi(t_k, z_k, h_n) + \varepsilon_k\} \end{cases} \quad (0 \leq k \leq n-1)$$

avec y_0 *et* z_0 *donnés dans* IR^p, *on ait :*

$$\max_{0 \leq k \leq n} \|z_k - y_k\| \leq M_1 \|z_0 - y_0\| + M_2 \max_{0 \leq k \leq n-1} \|\varepsilon_k\|$$

On désigne par y la solution du problème de Cauchy (1) et par $(y_k)_{0 \leq k \leq n}$ *une suite définie par (3).*

Q3. *Montrer que si la méthode à un pas associée à* ϕ *est consistante avec le problème de Cauchy (1) et stable, alors :*

$$\left(\lim_{n \to +\infty}(\zeta_n) = 0\right) \Rightarrow \left(\lim_{n \to +\infty}\left(\max_{0 \leq k \leq n} \|y_k - y(t_k)\|\right) = 0\right)$$

On dit alors que la méthode à un pas associée à ϕ *est convergente.*

Q4. *(a) Montrer que :*

$$\lim_{n \to +\infty}\left(\max_{0 \leq k \leq n-1} \|\varepsilon_k\|\right) = \sup_{t \in [a,b]} \|f(t, y(t)) - \phi(t, y(t), 0)\|$$

pour toute solution y du système différentiel $y'(t) = f(t, y(t))$.

(b) Montrer que la méthode à un pas associée à ϕ *est consistante avec le problème de Cauchy (1) si et seulement si :*

$$\forall t \in I, \ \forall z \in IR^p, \quad \phi(t, z, 0) = f(t, z)$$

Q5. *Soit* $(x_k)_{k \in IN}$ *une suite de réels positifs telle qu'il existe deux constantes réelles strictement positives* α *et* β *avec :*

$$\forall k \geq 0, \quad x_{k+1} \leq (1+\alpha) x_k + \beta$$

Montrer alors que :

$$\forall k \geq 0, \quad 0 \leq x_k \leq \left(x_0 + \frac{\beta}{\alpha} \right) e^{k\alpha} - \frac{\beta}{\alpha}$$

Q6. *Montrer que si ϕ est lipschitzienne par rapport à sa deuxième variable, alors la méthode à un pas associée à ϕ est stable.*

On suppose de plus, dans tout ce qui suit, que :

— ϕ est lipschitzienne par rapport à sa deuxième variable ;

— $\forall t \in I, \ \forall z \in IR^p, \quad \phi(t,z,0) = f(t,z)$.

Q7. *Montrer que $\underset{n \to +\infty}{Lim}\left(\underset{0 \leq k \leq n}{Max} \| y_k - y(t_k) \| \right) = 0$ pour toute suite $(y_k)_{0 \leq k \leq n}$ vérifiant (3) avec $\underset{n \to +\infty}{Lim}(\zeta_n) = 0$.*

Q8. *Pour tout entier n > 0, on désigne par φ_n la fonction affine par morceaux sur I définie par :*

$$\begin{cases} \varphi_n(t_k) = y_k & (0 \leq k \leq n) \\ \varphi_n \text{ affine sur } [t_k, t_{k+1}] & (0 \leq k \leq n-1) \end{cases}$$

où la suite $(y_k)_{0 \leq k \leq n}$ est définie par (3) avec $\underset{n \to +\infty}{Lim}(\zeta_n) = 0$.

Montrer que la suite de fonctions $(\varphi_n)_{n \in IN}$ converge uniformément vers y sur I.

Solution

Q1. On montre tout d'abord que la fonction $\psi = \varphi^k$ admet un unique point fixe dans E.

Soit $(x_n)_{n \in IN}$ une suite d'approximations successives définie par :

$$\begin{cases} x_0 \in E \\ x_{n+1} = \psi(x_n) & (n \geq 0) \end{cases}$$

Pour $m > n \geq 1$, on a $d(x_m, x_n) \leq d(x_m, x_{m-1}) + \cdots + d(x_{n+1}, x_n)$, avec :

$$\forall j \geq 0, \quad d(x_{j+1}, x_j) = d(\psi(x_j), \psi(x_{j-1})) \leq \lambda d(x_j, x_{j-1}) \leq \cdots \leq \lambda^j d(x_1, x_0)$$

Ce qui donne :

$$d(x_m, x_n) \leq \sum_{j=n}^{m-1} \lambda^j d(x_1, x_0) \leq \frac{\lambda^n}{1-\lambda} d(x_1, x_0) \underset{n \to +\infty}{\to} 0$$

si $\lambda \in [0,1[$.

On a donc ainsi montré que la suite $(x_n)_{n \in IN}$ est de Cauchy dans l'espace métrique complet *E* et qu'elle converge vers un élément *x* de *E*.

Avec la continuité de ψ (une application contractante est continue), on déduit que $\psi(x) = x$, c'est-à-dire que *x* est un point fixe de ψ.

Si $x' \in E$ est un autre point fixe de ψ, alors :

$$0 \leq d(x', x) = d(\psi(x'), \psi(x)) \leq \lambda d(x', x)$$

et $x \neq x'$ donne $d(x', x) < d(x', x)$, ce qui est impossible, donc $x = x'$ et ψ admet un unique point fixe dans *E*.

Avec $\psi(\varphi(x)) = \varphi^{k+1}(x) = \varphi(\psi(x)) = \varphi(x)$, on déduit que $\varphi(x)$ est aussi point fixe de ψ et $\varphi(x) = x$ du fait de l'unicité de x. Donc x est aussi point fixe de φ.

Réciproquement si x est un point fixe de φ, alors $\varphi^k(x) = x$ et x est aussi point fixe de ψ. D'où l'unicité du point fixe de φ.

Q2. On désigne par E l'espace des applications continues de I dans IR^p.

Muni de la norme de la convergence uniforme définie par $z \mapsto \|z\|_\infty = \underset{t\in[a,b]}{Sup}\|z(t)\|$, c'est un espace de Banach.

On définit l'application $\varphi : E \to E$ par :

$$\forall z \in E, \forall t \in I, \quad \varphi(z)(t) = y_0 + \int_{t_0}^{t} f(x, z(x))dx$$

et on remarque que la fonction $y \in C^1(I, IR^p)$ est solution de (2) si et seulement si elle vérifie $\varphi(y) = y$.

Il s'agit donc de montrer que φ admet un unique point fixe. Pour ce faire, on va montrer que l'une des itérées de φ est contractante.

Pour y_1, y_2 dans E et t dans I, on a :

$$\|\varphi(y_1)(t) - \varphi(y_2)(t)\| \le \left|\int_{t_0}^{t} \|f(x, y_1(x)) - f(x, y_2(x))\|dx\right| \le \lambda|t - t_0|\|y_1 - y_2\|_\infty$$

Et par récurrence sur $k > 0$:

$$\|\varphi(y_1) - \varphi(y_2)\|_\infty \le \frac{\lambda^k|t - t_0|^k}{k!}\|y_1 - y_2\|_\infty \le \frac{(\lambda(b-a))^k}{k!}\|y_1 - y_2\|_\infty$$

Comme $\underset{k\to+\infty}{Lim} \frac{(\lambda(b-a))^k}{k!} = 0$, il existe $k > 0$ tel que $\frac{(\lambda(b-a))^k}{k!} < 1$ et pour un tel k, l'itérée φ^k est contractante. Ce qui prouve que φ admet un unique point fixe dans E et que (2) admet une unique solution.

Q3. En posant $z_k = y(t_k)$, pour tout $k \in \{0, 1, \cdots, n\}$, on a a:

$$z_{k+1} = z_k + h_n\{\phi(t_k, z_k, h_n) + \varepsilon_k\}$$

et avec l'hypothèse de stabilité de la méthode :

$$\underset{0\le k\le n}{Max}\|z_k - y_k\| \le M_1\|z_0 - y_0\| + M_2 \underset{0\le k\le n-1}{Max}\|\varepsilon_k\| = M_1\|\zeta_n\| + M_2 \underset{0\le k\le n-1}{Max}\|\varepsilon_k\|$$

Avec l'hypothèse de consistance, on a $\underset{n\to+\infty}{Lim}\left(\underset{0\le k\le n-1}{Max}\|\varepsilon_k\|\right) = 0$. Et donc :

$$\left(\underset{n\to+\infty}{Lim}(\zeta_n) = 0\right) \Rightarrow \left(\underset{n\to+\infty}{Lim}\left(\underset{0\le k\le n}{Max}\|y_k - y(t_k)\|\right) = \underset{n\to+\infty}{Lim}\left(\underset{0\le k\le n}{Max}\|y_k - z_k\|\right) = 0\right)$$

Q4. (a) En appliquant la formule des accroissements finis à la fonction $y \in C^1(I, IR^p)$ solution de (1), les erreurs de consistance peuvent s'écrire :

$$\varepsilon_k = y'(\tau_k) - \phi(t_k, y(t_k), h_n) = f(\tau_k, y(\tau_k)) - \phi(t_k, y(t_k), h_n) \quad (0 \le k \le n-1)$$

avec $\tau_k \in [t_k, t_{k+1}]$.

On pose :

$$\begin{cases} \alpha_k = f(\tau_k, y(\tau_k)) - \phi(\tau_k, y(\tau_k), 0) \\ \beta_k = \phi(\tau_k, y(\tau_k), 0) - \phi(t_k, y(t_k), h_n) \end{cases} \quad (0 \le k \le n-1)$$

et on a $\varepsilon_k = \alpha_k + \beta_k$.

La fonction $g : t \mapsto \left\| f(t, y(t)) - \phi(t, y(t), 0) \right\|$ est continue sur le compact $[a,b]$, elle est donc bornée et atteint sa borne supérieure en un point $\tau \in [a,b]$. On peut alors poser $\alpha = \underset{t \in [a,b]}{Sup} \left\| f(t, y(t)) - \phi(t, y(t), 0) \right\| = g(\tau)$.

On a, de manière évidente :

$$\forall n \geq 1, \quad \underset{0 \leq k \leq n-1}{Max} \|\alpha_k\| \leq \alpha$$

Pour tout $\varepsilon > 0$, on peut trouver $\eta > 0$ tel que :

$$u, v \in [a,b], |u - v| < \eta \quad \Rightarrow \quad |g(u) - g(v)| < \varepsilon$$

Soit $n_0 \in IN$ tel que $h_n < \eta$ pour tout entier $n \geq n_0$.

Pour $n \geq n_0$, il existe $k \in \{0, 1, \cdots, n-1\}$ tel que $\tau \in [t_k, t_{k+1}]$. On a alors :

$$\forall t \in [t_k, t_{k+1}], \quad -\varepsilon < g(\tau) - g(t) < \varepsilon$$

On déduit alors que $\|\alpha_k\| = g(\tau_k) > g(\tau) - \varepsilon = \alpha - \varepsilon$ et $\underset{0 \leq k \leq n-1}{Max} \|\alpha_k\| > \alpha - \varepsilon$ pour tout entier $n \geq n_0$.

On a donc prouvé que pour tout réel $\varepsilon > 0$ il existe un entier n_0 tel que :

$$\forall n \geq n_0, \quad \alpha - \varepsilon < \underset{0 \leq k \leq n-1}{Max} \|\alpha_k\| \leq \alpha$$

Ce qui prouve que $\underset{n \to +\infty}{Lim} \left(\underset{0 \leq k \leq n-1}{Max} \|\alpha_k\| \right) = \alpha$.

D'autre part, la fonction $\varphi : (t, h) \mapsto \phi(t, y(t), h)$ est uniformément continue sur le compact $I \times J$. Pour tout réel $\varepsilon > 0$, on peut donc trouver un réel $\eta > 0$ tel que :

$$t, t' \in I, |t - t'| < \eta, \ h, h' \in J, |h - h'| < \eta, \quad \Rightarrow \quad \|\varphi(t, h) - \varphi(t', h')\| < \varepsilon$$

Soit $n_0 \in IN$ tel que $h_n < \eta$ pour tout entier $n \geq n_0$. Pour tout entier $n \geq n_0$ et tout $k \in \{0, 1, \cdots, n-1\}$, on a alors $|h_n| < \eta$ et $|\tau_k - t_k| \leq h_n < \eta$, de sorte que :

$$\forall n \geq n_0, \quad \underset{0 \leq k \leq n-1}{Max} \|\beta_k\| < \varepsilon$$

Ce qui prouve que $\underset{n \to +\infty}{Lim} \left(\underset{0 \leq k \leq n-1}{Max} \|\beta_k\| \right) = 0$.

On peut donc conclure que $\underset{n \to +\infty}{Lim} \left(\underset{0 \leq k \leq n-1}{Max} \|\varepsilon_k\| \right) = \alpha$.

(b) Si la méthode à un pas associée à ϕ est consistante avec le problème de Cauchy (1), alors $\underset{n \to +\infty}{Lim} \left(\underset{0 \leq k \leq n-1}{Max} \|\varepsilon_k\| \right) = 0$ et $f(t, y(t)) = \phi(t, y(t), 0)$ pour tout $t \in I$.

Pour $(x, z) \in I \times IR^p$, le problème de Cauchy $y'(t) = f(t, y(t))$ avec la condition initiale $y(x) = z$ admet un unique solution $\tilde{y} \in C^1(I, IR^p)$ (d'après Q2) et :

$$f(x, z) - \phi(x, z, 0) = f(x, \tilde{y}(x)) - \phi(x, \tilde{y}(x), 0) = 0$$

La réciproque est évidente.

Q5. Par récurrence, on déduit que :

$$\forall k \geq 0, \quad 0 \leq x_k \leq (1 + \alpha)^k x_0 + \beta \sum_{i=0}^{k-1} (1 + \alpha)^i = (1 + \alpha)^k x_0 + \beta \frac{(1 + \alpha)^k - 1}{\alpha}$$

Ou encore :

$$\forall k \geq 0, \quad 0 \leq x_k \leq \left(x_0 + \frac{\beta}{\alpha}\right)(1+\alpha)^k - \frac{\beta}{\alpha} \leq \left(x_0 + \frac{\beta}{\alpha}\right)e^{k\cdot\alpha} - \frac{\beta}{\alpha}$$

Ce résultat est le lemme de Gronwall.

Q6. Soit $\lambda \geq 0$ tel que :

$$\forall t \in I, \forall (y,z) \in IR^p \times IR^p, \forall h \in J, \quad \|\phi(t,z,h) - \phi(t,y,h)\| \leq \lambda \|z-y\|$$

On suppose que pour $n \geq 1$ donné, les suites $(y_k)_{0\leq k\leq n}$, $(z_k)_{0\leq k\leq n}$ et $(\varepsilon_k)_{0\leq k\leq n-1}$ vérifient :

$$\begin{cases} y_{k+1} = y_k + h_n \phi(t_k, y_k, h_n) \\ z_{k+1} = z_k + h_n \{\phi(t_k, z_k, h_n) + \varepsilon_k\} \end{cases} \quad (0 \leq k \leq n-1)$$

On définit alors la suite $(e_k)_{0\leq k\leq n}$ par $e_k = \|y_k - z_k\|$ pour tout $k \in \{0,1,\cdots,n\}$ et on a :

$$\begin{aligned} e_{k+1} &= \|y_{k+1} - z_{k+1}\| \\ &= \|y_k - z_k + h_n(\phi(t_k, y_k, h_n) - \phi(t_k, z_k, h_n)) - h_n \varepsilon_k\| \leq (1 + \lambda \cdot h_n) e_k + h_n \|\varepsilon_k\| \end{aligned}$$

Soit $e_{k+1} \leq (1+\alpha)e_k + \beta$, avec $\alpha = \lambda \cdot h_n$ et $\beta = h_n \underset{0\leq k\leq n-1}{Max} \|\varepsilon_k\|$.

Avec Q5, on déduit alors que :

$$\forall k \in \{0,1,\cdots,n\}, \quad 0 \leq e_k \leq e^{k\cdot\alpha} e_0 + \frac{e^{k\cdot\alpha} - 1}{\alpha}\beta$$

Ce qui entraîne :

$$\underset{0\leq k\leq n}{Max} \|z_k - y_k\| \leq M_1 \|z_0 - y_0\| + M_2 \underset{0\leq k\leq n-1}{Max} \|\varepsilon_k\|$$

avec $M_1 = e^{\lambda(b-a)}$ et $M_2 = \dfrac{e^{\lambda(b-a)} - 1}{\lambda}$ ($k\alpha = k\lambda \dfrac{b-a}{n} \leq \lambda(b-a)$ pour $0 \leq k \leq n$).

C'est-à-dire que la méthode à un pas associée à ϕ est stable.

Q7. C'est une conséquence immédiate des résultats de Q3, Q4 (b) et Q6.

Q8. La fonction φ_n est définie par :

$$\forall t \in [t_k, t_{k+1}], \quad \varphi_n(t) = y_k + \frac{t - t_k}{h_n}(y_{k+1} - y_k) \quad (0 \leq k \leq n-1)$$

Pour tout $t \in [t_k, t_{k+1}]$, on a :

$$\begin{aligned} \|\varphi_n(t) - y(t)\| &\leq \|y_k - y(t)\| + \|y_{k+1} - y_k\| \\ &\leq \|y_k - y(t_k)\| + \|y(t_k) - y(t)\| + \|y_{k+1} - y(t_{k+1})\| + \|y(t_{k+1}) - y(t_k)\| + \|y(t_k) - y_k\| \end{aligned}$$

On déduit alors que :

$$\|\varphi_n(t) - y(t)\| \leq 3 \underset{0\leq k\leq n}{Max} \|y_k - y(t_k)\| + 2 \underset{\substack{t,t' \in [a,b] \\ |t-t'| \leq h_n}}{Sup} \|y(t) - y(t')\|$$

La méthode étant convergente, pour tout réel $\varepsilon > 0$, on peut trouver $n_0 \in IN$ tel que $\underset{0\leq k\leq n}{Max} \|y_k - y(t_k)\| < \dfrac{\varepsilon}{6}$ pour tout entier $n \geq n_0$.

La fonction y étant uniformément continue sur le compact I, on aura également pour n_0 choisi assez grand, $\underset{\substack{t,t'\in[a,b] \\ |t-t'|\le h_n}}{Sup}\|y(t)-y(t')\|<\dfrac{\varepsilon}{4}$ dès que $n\ge n_0$.

Ce qui donne en définitive $\underset{t\in[a,b]}{Sup}\|\varphi_n(t)-y(t)\|<\varepsilon$, pour tout $n\ge n_0$. On a donc ainsi prouvé que la suite de fonctions $(\varphi_n)_{n\in IN}$ converge uniformément vers y sur I.

Problème 45 : Ordre d'une méthode à un pas. Méthode RK4

Ce problème utilise les notations et résultats du problème 44.
On rappelle que :
— $I=[a,b]$ est un intervalle réel fermé borné, p un entier strictement positif.
— $f:I\times IR^p\to IR^p$ une fonction continue et lipschitzienne par rapport à sa deuxième variable.
— A tout entier n > 0, on associe une subdivision de l'intervalle I en posant :

$$h_n=\frac{b-a}{n},\quad t_k=a+k\cdot h_n\quad (0\le k\le n)$$

— On note $J=[0,b-a]$ et $\phi: I\times IR^p\times J\to IR^p$ est une fonction continue lipschitzienne par rapport à sa deuxième variable.

On désigne par y la solution du problème de Cauchy (1) du problème 44 et par $(y_k)_{0\le k\le n}$ une suite définie par la relation de récurrence :

$$(*)\quad \begin{cases} y_0=y_a+\zeta_n \\ y_{k+1}=y_k+h_n\phi(t_k,y_k,h_n)\quad (0\le k\le n-1)\end{cases}$$

où ζ_n est donné dans IR^p.

On dit la méthode à un pas associée à ϕ est d'ordre q s'il existe un réel γ tel que :

$$\forall n\ge 1,\quad \underset{0\le k\le n-1}{Max}\|\varepsilon_k\|\le\gamma\cdot h_n^{\,q}$$

où ε_k désigne l'erreur de consistance de la méthode en t_k.

Q1. *On suppose que la méthode à un pas associée à ϕ est d'ordre q.*
Donner alors une majoration de l'erreur de discrétisation définie par :

$$\underset{0\le k\le n}{Max}\|y_k-y(t_k)\|$$

pour toute suite $(y_k)_{0\le k\le n}$ définie par (*).

Dans tout ce qui suit, on fait les hypothèses supplémentaires suivantes :
— p = 1 (pour simplifier les notations) ;
— f est de classe C^q sur I, où q est un entier strictement positif donné ;
— ϕ admet des dérivées partielles continues par rapport à sa troisième variable h jusqu'à l'ordre q.

On définit la suite de fonctions $\left(f^{(j)}\right)_{0\le j\le q}$ par la relation de récurrence :

$$\begin{cases} f^{(0)} = f \\ f^{(j+1)} = \dfrac{\partial f^{(j)}}{\partial t} + f\,\dfrac{\partial f^{(j)}}{\partial y} \quad (0 \le j \le q-1) \end{cases}$$

Q2. *Montrer que la méthode à un pas associée à ϕ est d'ordre $q \ge 1$ si et seulement si :*

$$\forall t \in I,\ \forall z \in IR,\quad \frac{\partial^j \phi}{\partial h^j}(t,z,0) = \frac{1}{j+1} f^{(j)}(t,z) \quad (0 \le j \le q-1)$$

Q3. *On suppose maintenant que f est de classe C^4.*

On peut définir une méthode de résolution approchée du problème de Cauchy (1) du problème 44 en écrivant, pour tout entier $k \in \{0,1,\cdots,n-1\}$:

$$y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(t, y(t))dt$$

et en utilisant une formule d'intégration approchée.

(a) Montrer que l'utilisation de la formule d'intégration approchée de Simpson conduit à la méthode à un pas associée à ϕ définie par :

$$\phi(t,y,h) = \frac{1}{6}\{K_1(t,y,h) + 2K_2(t,y,h) + 2K_3(t,y,h) + K_4(t,y,h)\}$$

où :

$$K_1(t,y,h) = f(t,y), \quad K_2(t,y,h) = f\left(t + \frac{h}{2}, y + \frac{h}{2}K_1(t,y,h)\right)$$

$$K_3(t,y,h) = f\left(t + \frac{h}{2}, y + \frac{h}{2}K_2(t,y,h)\right), \quad K_4(t,y,h) = f(t+h, y + hK_3(t,y,h))$$

Cette méthode est appelée « méthode RK4 ».

(b) Montrer que la méthode RK4 est consistante.

(c) Montrer que la méthode RK4 est stable.

(d) Montrer que la méthode RK4 est convergente.

(e) Montrer que la méthode RK4 est d'ordre 4.

Solution

Q1. La définition de l'erreur de consistance permet de prendre $z_k = y(t_k)$ dans la question Q6 du problème 44. On a alors la majoration :

$$\underset{0\le k\le n}{Max}\|y(t_k) - y_k\| \le M_1\|y(a) - y_0\| + M_2 \underset{0\le k\le n-1}{Max}\|\varepsilon_k\| \le e^{\lambda(b-a)}\|y(a) - y_0\| + \frac{e^{\lambda(b-a)} - 1}{\lambda}\gamma \cdot h_n^{\ q}$$

Dans le cas particulier où la valeur initiale y_0 est exactement $y(a)$, on a une erreur de l'ordre de $h_n^{\ q}$.

Q2. Avec la formule de Taylor en $h = 0$, on peut écrire que :

$$\phi(t_k, y(t_k), h_n) = \sum_{j=0}^{q-1} \frac{h_n^{\ j}}{j!}\frac{\partial^j \phi}{\partial h^j}(t_k, y(t_k), 0) + \frac{h_n^{\ q}}{q!}\frac{\partial^q \phi}{\partial h^q}(t_k, y(t_k), \xi_k)$$

avec $\xi_k \in]0, h_n[$.

Par récurrence, on vérifie facilement que :

$$y^{(j)}(t)=f^{(j-1)}(t,y(t))\quad (1\leq j\leq q+1)$$

En effet, pour $j=1$, on a $y'(t)=f(t,y(t))=f^{(0)}(t,y(t))$ et si on suppose le résultat vrai pour $j\in\{1,\cdots,q\}$, alors :

$$y^{(j+1)}(t)=\frac{\partial}{\partial t}\left(f^{(j-1)}(t,y(t))\right)=\frac{\partial f^{(j-1)}}{\partial t}(t,y(t))+\frac{\partial f^{(j-1)}}{\partial y}(t,y(t))y'(t)$$

$$=\frac{\partial f^{(j-1)}}{\partial t}(t,y(t))+\frac{\partial f^{(j-1)}}{\partial y}(t,y(t))f(t,y(t))=f^{(j)}(t,y(t))$$

Un développement de Taylor en t_k donne alors :

$$y(t_{k+1})=y(t_k+h_n)=y(t_k)+\sum_{j=1}^{q}\frac{h_n^{\,j}}{j!}f^{(j-1)}(t_k,y(t_k))+\frac{h_n^{\,q+1}}{(q+1)!}f^{(q)}(\eta_k,y(\eta_k))$$

avec $\eta_k\in\left]t_k,t_{k+1}\right[$ (y est de classe C^{q+1} puisque f est de classe C^q).

On déduit alors le développement suivant de l'erreur de consistance :

$$\varepsilon_k=\sum_{j=0}^{q-1}\frac{h_n^{\,j}}{j!}\left(\frac{1}{j+1}f^{(j)}(t_k,y(t_k))-\frac{\partial^j\phi}{\partial h^j}(t_k,y(t_k),0)\right)$$

$$+\frac{h_n^{\,q}}{q!}\left(\frac{1}{q+1}f^{(q)}(\eta_k,y(\eta_k))-\frac{\partial^q\phi}{\partial h^q}(t_k,y(t_k),\xi_k)\right)$$

Les conditions :

$$\forall t\in I,\ \forall z\in IR,\quad \frac{\partial^j\phi}{\partial h^j}(t,z,0)=\frac{1}{j+1}f^{(j)}(t,z)\quad (0\leq j\leq q-1)$$

entraînent alors :

$$\varepsilon_k=\frac{h_n^{\,q}}{q!}\left(\frac{1}{q+1}f^{(q)}(\eta_k,y(\eta_k))-\frac{\partial^q\phi}{\partial h^q}(t_k,y(t_k),\xi_k)\right)$$

Avec la continuité de la fonction $t\mapsto f^{(q)}(t,y(t))$ sur le compact I et la continuité de $(t,h)\mapsto\frac{\partial^q\phi}{\partial h^q}(t,y(t),h)$ sur le compact $I\times J$, on déduit qu'il existe une constante réelle γ indépendante de n telle que $\underset{0\leq k\leq n-1}{Max}\|\varepsilon_k\|\leq\gamma\cdot h_n^{\,q}$. C'est-à-dire que la méthode est d'ordre q.

Réciproquement supposons que la méthode soit d'ordre $q\geq 1$ et qu'il existe $j\in\{0,1,\cdots,q-1\}$ tel que la fonction $\frac{\partial^j\phi}{\partial h^j}(t,z,0)-\frac{1}{j+1}f^{(j)}(t,z)$ ne soit pas identiquement nulle sur $I\times IR$. On désigne par $p\in\{0,1,\cdots,q-1\}$ le plus petit de ces entiers.

Comme dans la question Q4 (b) du problème 44, on voit que la fonction $\frac{\partial^p\phi}{\partial h^p}(t,y(t),0)-\frac{1}{p+1}f^{(p)}(t,y(t))$ n'est pas identiquement nulle sur I.

En utilisant des développements de Taylor, on a :

$$\varepsilon_k = \frac{h_n{}^p}{p!}\left(\frac{1}{p+1}f^{(p)}(t_k,y(t_k))-\frac{\partial^p\phi}{\partial h^p}(t_k,y(t_k),0)\right)$$
$$+\frac{h_n{}^{p+1}}{(p+1)!}\left(\frac{1}{p+2}f^{(p+1)}(\eta_k,y(\eta_k))-\frac{\partial^{p+1}\phi}{\partial h^{p+1}}(t_k,y(t_k),\xi_k)\right)$$

Avec la continuité des fonctions considérées, on peut trouver un réel $\delta>0$ et un intervalle non réduit à un point $[\alpha,\beta]\subset[a,b]$ tels que :

$$\forall t\in[\alpha,\beta],\quad \left|\frac{\partial^p\phi}{\partial h^p}(t,y(t),0)-\frac{1}{p+1}f^{(p)}(t,y(t))\right|\geq\delta$$

On peut également trouver un entier $n_0\in IN$ tel que $h_n<\beta-\alpha$ pour tout $n\geq n_0$. Pour tout $n\geq n_0$, il existe $k\in\{0,1,\cdots,n-1\}$ tel que $[t_k,t_{k+1}]\subset[\alpha,\beta]$. On déduit alors que :

$$\delta\leq\left|\frac{1}{p+1}f^{(p)}(t_k,y(t_k))-\frac{\partial^p\phi}{\partial h^p}(t_k,y(t_k),0)\right|\leq\frac{p!}{h_n{}^p}|\varepsilon_k|+\frac{h_n}{(p+1)}M_{p+1}$$

où

$$M_{p+1}=\underset{I\times y(I)}{Sup}\left|\frac{1}{p+2}f^{(p+1)}(\eta_k,y(\eta_k))-\frac{\partial^{p+1}\phi}{\partial h^{p+1}}(t_k,y(t_k),\xi_k)\right|$$

Et avec $\frac{p!}{h_n{}^p}|\varepsilon_k|+\frac{h_n}{(p+1)!}M_{p+1}\leq p!\,\gamma\cdot h_n{}^{q-p}+\frac{h_n}{(p+1)}M_{p+1}\underset{n\to+\infty}{\longrightarrow}0$, on aboutit à une contradiction.

On a donc bien $\frac{\partial^j\phi}{\partial h^j}(t,z,0)-\frac{1}{j+1}f^{(j)}(t,z)=0$ pour tout $j\in\{0,1,\cdots,q-1\}$.

Q3. (a) La formule d'intégration approchée de Simpson peut s'exprimer par :

$$\int_\alpha^\beta g(t)dt\cong\frac{\beta-\alpha}{6}\left\{g(\alpha)+4g\left(\frac{\alpha+\beta}{2}\right)+g(\beta)\right\}$$

On peut alors écrire :

$$y(t_{k+1})\cong y(t_k)+\frac{h_n}{6}\left\{f(t_k,y(t_k))+4f\left(t_k+\frac{h_n}{2},y\left(t_k+\frac{h_n}{2}\right)\right)+f(t_k+h_n,y(t_k+h_n))\right\}$$

Le calcul approché de $f(t_k,y(t_k))$ se fait avec :

$$f(t_k,y(t_k))\cong K_1=f(t_k,y_k)$$

où y_k est une approximation de $y(t_k)$.

Un premier calcul approché de $y\left(t_k+\frac{h_n}{2}\right)$ peut se faire en utilisant la méthode des rectangles à gauche. Ce qui donne :

$$y\left(t_k+\frac{h_n}{2}\right)=y(t_k)+\int_{t_k}^{t_k+\frac{h_n}{2}}f(t,y(t))dt\cong y(t_k)+\frac{h_n}{2}f(t_k,y(t_k))\cong y_k+\frac{h_n}{2}K_1$$

et :

$$f\left(t_k+\frac{h_n}{2},y\left(t_k+\frac{h_n}{2}\right)\right)\cong K_2=f\left(t_k+\frac{h_n}{2},y_k+\frac{h_n}{2}K_1\right)$$

En utilisant la méthode des rectangles à droite, on obtient :

$$y\left(t_k+\frac{h_n}{2}\right)\cong y(t_k)+\frac{h_n}{2}f\left(t_k+\frac{h_n}{2},y\left(t_k+\frac{h_n}{2}\right)\right)\cong y_k+\frac{h_n}{2}f\left(t_k+\frac{h_n}{2},y_k+\frac{h_n}{2}K_1\right)$$

soit $y\left(t_k+\frac{h_n}{2}\right)\cong y_k+\frac{h_n}{2}K_2$ et :

$$f\left(t_k+\frac{h_n}{2},y\left(t_k+\frac{h_n}{2}\right)\right)\cong K_3=f\left(t_k+\frac{h_n}{2},y_k+\frac{h_n}{2}K_2\right)$$

Enfin un calcul approché de $f\left(t_k+h_n,y(t_k+h_n)\right)$ se fait en utilisant la méthode du point milieu, ce qui donne :

$$y(t_k+h_n)=y(t_k)+\int_{t_k}^{t_k+h_n}f(t,y(t))dt\cong y(t_k)+h_nf\left(t_k+\frac{h_n}{2},y\left(t_k+\frac{h_n}{2}\right)\right)$$

$$\cong y_k+h_nf\left(t_k+\frac{h_n}{2},y_k+\frac{h_n}{2}K_2\right)\cong y_k+h_nK_3$$

et :

$$f\left(t_k+h_n,y(t_k+h_n)\right)\cong K_4=f\left(t_k+h_n,y_k+h_nK_3\right)$$

Ce qui donne en définitive la méthode définie par :

$$\phi(t,y,h)=\frac{1}{6}\left\{K_1(t,y,h)+2K_2(t,y,h)+2K_3(t,y,h)+K_4(t,y,h)\right\}$$

avec les coefficients k_j définis dans l'énoncé.

(b) Pour $h=0$, on a :

$$K_1(t,y,0)=K_2(t,y,0)=K_3(t,y,0)=K_4(t,y,0)=f(t,y)$$

Donc $\phi(t,y,0)=f(t,y)$ et la méthode RK4 est consistante avec le problème de Cauchy (problème 44, Q4 (b)).

(c) Pour montrer que la méthode RK4 est stable, il suffit de montrer que l'application ϕ est lipschitzienne par rapport à sa deuxième variable (problème 44, Q6), ce qui résulte immédiatement du caractère lipschitzien de la fonction f.

(d) Avec le résultat de la question Q3 du problème 44, on déduit que la méthode RK4 est convergente.

(e) On note, pour $j=1,2,3,4$:

$$K_j=K_j(t,y,h)=f\left(t+c_jh,y_j\right)$$

avec :

$$c_1=0,\quad c_2=c_3=\frac{1}{2},\quad c_4=1$$

$$K_0=0,\quad y_j=y+c_jhK_{j-1}\quad (j=1,2,3,4)$$

On remarque que :

$$K_j(t,y,0)=f(t,y)\quad (j=1,2,3,4)$$

$$\frac{\partial^iK_1}{\partial h^i}=0\quad (i=1,2,3)$$

Pour $i=1$, on a :

$$\frac{\partial\phi}{\partial h}(t,y,h)=\frac{1}{6}\left\{2\frac{\partial K_2}{\partial h}(t,y,h)+2\frac{\partial K_3}{\partial h}(t,y,h)+\frac{\partial K_4}{\partial h}(t,y,h)\right\}$$

avec, pour $j = 2, 3, 4$:

$$\frac{\partial K_j}{\partial h}(t,y,h)=c_j\frac{\partial f}{\partial t}(t+c_jh,y_j)+\frac{\partial f}{\partial y}(t+c_jh,y_j)\frac{\partial y_j}{\partial h}(t,y,h)$$

$$\frac{\partial y_j}{\partial h}(t,y,h)=c_jK_{j-1}(t,y,h)+hc_j\frac{\partial K_j}{\partial h}(t,y,h)$$

Ce qui donne, pour $h = 0$:

$$\frac{\partial y_j}{\partial h}(t,y,0)=c_jf(t,y)$$

$$\frac{\partial K_j}{\partial h}(t,y,0)=c_j\left(\frac{\partial f}{\partial t}(t,y)+f(t,y)\frac{\partial f}{\partial y}(t,y)\right)$$

et :

$$\frac{\partial\phi}{\partial h}(t,y,0)=\frac{1}{6}(2c_2+2c_3+c_4)\left(\frac{\partial f}{\partial t}(t,y)+f(t,y)\frac{\partial f}{\partial y}(t,y)\right)=\frac{1}{2}f^{(1)}(t,y)$$

Pour $i = 2$, on a :

$$\frac{\partial^2\phi}{\partial h^2}(t,y,h)=\frac{1}{6}\left\{2\frac{\partial^2 K_2}{\partial h^2}(t,y,h)+2\frac{\partial^2 K_3}{\partial h^2}(t,y,h)+\frac{\partial^2 K_4}{\partial h^2}(t,y,h)\right\}$$

avec, pour $j = 2, 3, 4$:

$$\frac{\partial^2 K_j}{\partial h^2}(t,y,h)=c_j^2\frac{\partial^2 f}{\partial t^2}(t+c_jh,y_j)+2c_j\frac{\partial^2 f}{\partial t\partial y}(t+c_jh,y_j)\frac{\partial y_j}{\partial h}(t,y,h)$$
$$+\frac{\partial^2 f}{\partial y^2}(t+c_jh,y_j)\left(\frac{\partial y_j}{\partial h}(t,y,h)\right)^2+\frac{\partial f}{\partial y}(t+c_jh,y_j)\frac{\partial^2 y_j}{\partial h^2}(t,y,h)$$

$$\frac{\partial^2 y_j}{\partial h^2}(t,y,h)=c_j\left(2\frac{\partial K_{j-1}}{\partial h}(t,y,h)+h\frac{\partial^2 K_{j-1}}{\partial h^2}(t,y,h)\right)$$

Ce qui donne, pour $h = 0$:

$$\frac{\partial^2 y_j}{\partial h^2}(t,y,0)=2c_j\frac{\partial K_{j-1}}{\partial h}(t,y,0)=2c_jc_{j-1}\left(\frac{\partial f}{\partial t}(t,y)+f(t,y)\frac{\partial f}{\partial y}(t,y)\right)$$

$$\frac{\partial^2 K_j}{\partial h^2}(t,y,0)=c_j^2\left(\frac{\partial^2 f}{\partial t^2}(t,y)+2\frac{\partial^2 f}{\partial t\partial y}(t,y)f(t,y)+\frac{\partial^2 f}{\partial y^2}(t,y)f(t,y)^2\right)$$
$$+2c_jc_{j-1}\frac{\partial f}{\partial y}(t,y)\left(\frac{\partial f}{\partial t}(t,y)+f(t,y)\frac{\partial f}{\partial y}(t,y)\right)$$

et :

$$\frac{\partial^2\phi}{\partial h^2}(t,y,0)=\frac{1}{6}\left(2c_2^2+2c_3^2+c_4^2\right)\left(\frac{\partial^2 f}{\partial t^2}(t,y)+2\frac{\partial^2 f}{\partial t\partial y}(t,y)f(t,y)+\frac{\partial^2 f}{\partial y^2}(t,y)f(t,y)^2\right)$$
$$+\frac{1}{3}\left(2c_3c_2+c_4c_3\right)\frac{\partial f}{\partial y}(t,y)\left(\frac{\partial f}{\partial t}(t,y)+f(t,y)\frac{\partial f}{\partial y}(t,y)\right)$$

avec :

$$2c_2^2+2c_3^2+c_4^2=2,\quad 2c_3c_2+c_4c_3=1$$

Ce qui donne :

$$\frac{\partial^2\phi}{\partial h^2}(t,y,0)=\frac{1}{3}f^{(2)}(t,y)$$

De manière analogue, on peut aussi vérifier que $\frac{\partial^3\phi}{\partial h^3}(t,y,0)=\frac{1}{4}f^{(3)}(t,y)$ et $\frac{\partial^4\phi}{\partial h^4}(t,y,0)\neq\frac{1}{5}f^{(4)}(t,y)$.

C'est-à-dire que la méthode RK4 est d'ordre exactement 4.

Problème 46 : Méthode de différentiation rétrograde de Gear. Agrégation 1980, extrait

Ce problème utilise les notations et les résultats du problème 24.

Soit T un nombre réel positif ; pour $m\in IN$, *on note* $C^m([0,T];E)$ *l'espace des fonctions m fois continûment différentiables de* $[0,T]$ *à valeurs dans E et* $C^m([0,T]\times E;E)$ *l'espace des fonctions m fois continûment différentiables de* $[0,T]\times E$ *à valeurs dans E.*

On se donne une fonction $f:(t,y)\mapsto f(t,y)$ *appartenant à* $C^1([0,T]\times E;E)$ *et on note* $D_yf(t,y)$ *la dérivée partielle de f par rapport à y au point* (t,y) *; on supposera dans toute la suite qu'il existe un nombre réel positif L tel que :*

$$\forall t\in[0,T],\quad \forall y\in E,\quad \left\|D_yf(t,y)\right\|\leq L.$$

On se propose d'approcher la solution du problème différentiel de condition initiale suivant :

$$(P)\quad\begin{cases}\textit{Déterminer } y\in C^1([0,T];E) \textit{ vérifiant };\\ \forall t\in[0,T],\quad y'(t)=f(t,y(t))\\ y(0)=y_0 \textit{ donné dans E}\end{cases}$$

où $y'(t)$ *désigne la dérivée de la fonction y au point t.*

Q1. *Montrer que pour tout* $t\in[0,T]$, *tout y et z de E, il existe un opérateur* $g(t;y,z)$ *de* $L(E)$ *tel que :*

$$f(t,y)-f(t,z)=g(t;y,z)(y-z)$$

et

$$\left\|g(t;y,z)\right\|\leq L$$

On considère maintenant une subdivision :

$$0 = t_0 < t_1 < t_2 < \cdots < t_n < \cdots < t_N = T$$

de l'intervalle $[0,T]$ *et on note, pour* $n = 1,\cdots,N$, *par* $h_n = t_n - t_{n-1}$ *le* $n^{ième}$ *pas.*

On se propose de déterminer par récurrence des valeurs approchées $y_0, y_1, \cdots, y_n, \cdots, y_N$ *de la solution* $y(\cdot)$ *du problème (P) aux points* $t_0, t_1, \cdots, t_n, \cdots, t_N$. *Pour cela on se donne un entier r avec* $1 \leq r \leq N$ *et on suppose déterminées les valeurs approchées* $y_0, y_1, \cdots, y_{n-1}$ *(*$r \leq n \leq N$*).*

Q2. *Soit* $\pi_{n,r}$ *la fonction « polynomiale » de* $[0,T]$ *à valeurs dans E :*

$$\pi_{n,r}(t) = \sum_{k=0}^{r} t^k d_{k,n,r}$$

où les coefficients $d_{k,n,r}$ *appartiennent à E et sont déterminés par les relations :*

$$\begin{cases} \forall i = 1,2,\cdots,r, \quad \pi_{n,r}(t_{n-i}) = y_{n-i} \\ \qquad\qquad\qquad \pi_{n,r}(t_n) = z \text{ donné dans } E \end{cases}$$

(a) On pose $z^* = \dfrac{d}{dt}\pi_{n,r}(t_n)$. *Montrer que z et* z^* *sont liés par une relation de la forme :*

$$z = \sum_{i=1}^{r} a_{n,i} y_{n-i} + h_n b_n z^*$$

Calculer les coefficients $a_{n,i}$ *et* b_n *en fonction des abscisses* $t_n, t_{n-1}, \cdots, t_{n-r}$.

(b) Montrer que lorsque le pas $h_n = h = \dfrac{T}{N}$ *est constant les coefficients* $a_{n,i}$ *et* b_n *ne dépendent que de i et r ; on les notera respectivement* a_i *et b.*

(c) Montrer que si la subdivision de l'intervalle $[0,T]$ *vérifie la condition :*

$$(1) \textit{il existe } \delta > 0, \textit{ tel que } \forall n = 1,2,\cdots,N-1, \quad 0 < \frac{1}{\delta} \leq \frac{h_{n+1}}{h_n} \leq \delta$$

alors il existe un nombre réel positif $C_1(\delta)$ *tel que, pour tout n vérifiant* $r \leq n \leq N$, *on ait :*

$$\forall i = 1,2,\cdots,r, \quad |a_{n,i}| \leq C_1(\delta), \quad |b_n| \leq C_1(\delta)$$

(d) Etablir les relations :

$$\sum_{i=1}^{r} a_{n,i} = 1, \quad \sum_{i=1}^{r} a_{n,i}(t_n - t_{n-i}) = h_n b_n$$

(e) Montrer que si $h_n b_n L < 1$, *il existe un élément unique* y_n *de E tel que :*

$$y_n = \sum_{i=1}^{r} a_{n,i} y_{n-i} + h_n b_n f(t_n, y_n)$$

Q3. *Soit* $z \in C^{r+2}([0,T];E)$ *; on pose, pour* $r \leq n \leq N$ *:*

$$\varepsilon_n(z) = z(t_n) - \sum_{i=1}^{r} a_{n,i} z(t_{n-i}) - h_n b_n z'(t_n)$$

où z' *désigne la dérivée de la fonction z.*

(a) Montrer que, sous l'hypothèse (1) de Q2 (c), il existe un nombre réel positif $C_2(\delta)$ *indépendant de n, tel que :*

$$\left\|\varepsilon_n(z)\right\| \le C_2(\delta) h^r \int_{t_{n-r}}^{t_n} \left\|z^{(r+1)}(t)\right\| dt$$

où on a posé $h = \underset{1\le n\le N}{Max}\, h_n$ *et où* $z^{(k)}$ *désigne la dérivée* $k^{ième}$ *de la fonction z.*

(b) Montrer que, sous l'hypothèse (1), il existe des nombres réels $C_3(\delta)$ *(indépendant de n) et* $\gamma_{n,r}$*, tels que :*

$$\left\|\varepsilon_n(z) - \gamma_{n,r} b_n h_n^{\,r+1} z^{(r+1)}(t_n)\right\| \le C_3(\delta) h^{r+1} \int_{t_{n-r}}^{t_n} \left\|z^{(r+2)}(t)\right\| dt$$

Donner une expression simple de $\gamma_{n,r}$ *en fonction de* $t_n, t_{n-1}, \cdots, t_{n-r}$ *; préciser la valeur* γ_r *de* $\gamma_{n,r}$ *lorsque le pas* $h_n = h$ *est constant.*

On définit par récurrence des valeurs approchées y_n *de* $y(t_n)$ *par :*

$$(P_h)\quad \begin{cases} y_\mu = \eta_\mu \ \text{ donné dans } E,\ \text{pour } \mu = 0,1,\cdots,r-1 \\[2mm] y_n = \sum_{i=1}^{r} a_{n,i} y_{n-i} + h_n b_n f(t_n, y_n),\ \text{pour } r \le n \le N \end{cases}$$

où les valeurs de démarrage $\eta_0, \eta_1, \cdots, \eta_{r-1}$ *sont calculées par une autre méthode que nous ne préciserons pas ici.*

Nous supposerons dans tout ce qui suit qu'il existe un nombre réel positif $\nu < 1$ *tel que :*

$$(2)\quad \forall n \ge r,\quad h_n b_n L \le \nu$$

de sorte que le problème (P_h) *admet une solution unique.*

On considère maintenant deux suites $(\varepsilon_n)_{r\le n\le N}$ *et* $(z_n)_{0\le n\le N}$ *vérifiant les « équations perturbées » :*

$$(Q_h)\quad \begin{cases} z_0, z_1, \cdots, r_{r-1} \ \text{ donnés dans } E \\[2mm] z_n = \sum_{i=1}^{r} a_{n,i} z_{n-i} + h_n b_n f(t_n, z_n) + \varepsilon_n,\ \text{pour } r \le n \le N \end{cases}$$

On pose :

$$u_n = y_n - z_n \ \text{ et } U_n = (u_n, u_{n-1}, \cdots, u_{n-r+1})^T$$

Q4. *(a) Montrer qu'il existe un opérateur* $G_n \in L(E^r)$ *ne dépendant que de* $g(t_n; y_n, z_n)$*, une matrice carrée* R_n *d'ordre r ne dépendant que des coefficients* $a_{n,i}$ *et un élément* E_n *de* E^r *ne dépendant que de* ε_n*, tels que l'on ait :*

$$(I - h_n b_n G_n) U_n = R_n U_{n-1} + E_n,\ \text{pour } r \le n \le N$$

où I désigne l'opérateur identité.

(b) Lorsque le pas $h_n = h$ *est constant la matrice* R_n *ne dépend pas de n ; on la notera R. Calculer la matrice R pour r = 2, 3 et 4 ; montrer que cette matrice vérifie la condition (D) de Q2 (c) du problème 24.*

On suppose désormais que r = 2, 3 ou 4.

(c) En déduire qu'il existe un nombre réel positif α_r *et une matrice carrée d'ordre r inversible, notée* H_r*, tels que si la subdivision de l'intervalle* $[0, T]$ *vérifie la condition :*

$$(3)\quad \forall n = 2, \cdots, N,\quad \left|1 - \frac{h_n}{h_{n-1}}\right| \le \alpha_r$$

alors on a :

$$\forall n = r, \cdots, N, \quad \left\| H_r^{-1} R_n H_r \right\|_1 \leq 1 \textit{ et la condition (1) de Q2 (c).}$$

(d) Montrer qu'il existe un nombre réel positif C_4 tel que, pour $r \leq n \leq N$:

$$\left\| H_r^{-1} \left(I - h_n b_n G_n \right)^{-1} H_r \right\|_1 \leq 1 + C_4 h_n L$$

(e) En déduire que, sous l'hypothèse (3), il existe un nombre réel positif C_5 tel que :

$$\forall n = r, \cdots, N, \quad \left\| U_n \right\|_1 \leq C_5 \exp\left(C_4 L t_n \right) \left[\left\| U_{r-1} \right\|_1 + \sum_{i=r}^{n} \left\| \varepsilon_i \right\| \right]$$

Q5. *On suppose maintenant que $f \in C^r\left([0, T] \times E; E \right)$. Montrer que, sous l'hypothèse (3), il existe un nombre réel positif C_6 tel que :*

$$\forall n = r, \cdots, N, \quad \left\| y_n - y(t_n) \right\|_1 \leq C_5 \exp\left(C_4 L t_n \right) \left[\sum_{\mu=0}^{r-1} \left\| y_\mu - y(t_\mu) \right\| + C_6 h^r \int_0^{t_n} \left\| y^{(r+1)}(t) \right\| dt \right]$$

Solution

Q1. A $t \in [0, T]$, $y \in E$ et $z \in E$ fixés, on associe la fonction $\varphi . IR \to E$ définie par :

$$\forall x \in IR, \quad \varphi(x) = f\left(t, z + x(y - z) \right)$$

C'est une application continûment dérivable de dérivée définie par :

$$\forall x \in IR, \quad \varphi'(x) = D_y f\left(t, z + x(y - z) \right)(y - z)$$

On a alors :

$$f(t, y) - f(t, z) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(x) dx = g(t; y, z)(y - z)$$

où :

$$g(t; y, z) = \int_0^1 D_y f\left(t, z + x(y - z) \right) dx \in L(E)$$

La condition de Lipschitz $\left\| D_y f(t, y) \right\| \leq L$ sur $[0, T] \times E$ donne alors :

$$\forall (t; y, z)[0, T] \times E^2, \quad \left\| g(t; y, z) \right\| \leq L$$

Q2. Le polynôme $\pi_{n,r}$ est le polynôme d'interpolation de Lagrange associé à $\left(t_n, t_{n-1}, \cdots, t_{n-r} \right)$ et $\left(z, y_{n-1}, \cdots, y_{n-r} \right)$ (voir le problème 33, Q3).

On peut l'écrire sous la forme :

$$\pi_{n,r}(t) = L_0(t) z + \sum_{i=1}^{r} L_i(t) y_{n-i}$$

où on a posé :

$$L_i(t) = \prod_{\substack{j=0 \\ j \neq i}}^{r} \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}} \quad (0 \leq i \leq r)$$

(a) Le polynôme dérivé de $\pi_{n,r}$ est donné par $\pi_{n,r}'(t) = L_0'(t) z + \sum_{i=1}^{r} L_i'(t) y_{n-i}$ et pour $t = t_n$, on a :

$$\begin{cases} L_i'(t_n) = \underset{\substack{t \to t_n \\ t \neq t_n}}{Lim} \dfrac{L_i(t)}{t - t_n} = \dfrac{1}{t_{n-i} - t_n} \displaystyle\prod_{\substack{j=1 \\ j \neq i}}^{r} \dfrac{t_n - t_{n-j}}{t_{n-i} - t_{n-j}} \quad (1 \leq i \leq r) \\ L_0'(t_n) = \displaystyle\sum_{j=1}^{r} \dfrac{1}{t_n - t_{n-j}} \end{cases}$$

Ce qui donne :

$$\begin{aligned} z^* = \pi_{n,r}'(t_n) &= L_0'(t_n) z + \sum_{i=1}^{r} L_i'(t_n) y_{n-i} \\ &= \left(\sum_{j=1}^{r} \frac{1}{t_n - t_{n-j}} \right) z + \sum_{i=1}^{r} \frac{1}{t_{n-i} - t_n} \prod_{\substack{j=1 \\ j \neq i}}^{r} \frac{t_n - t_{n-j}}{t_{n-i} - t_{n-j}} y_{n-i} \end{aligned}$$

et $z = \sum_{i=1}^{r} a_{n,i} y_{n-i} + h_n b_n z^*$, où on a posé :

$$\begin{cases} b_n = \dfrac{1}{h_n L_0'(t_n)} = \dfrac{1}{h_n \displaystyle\sum_{j=1}^{r} \dfrac{1}{t_n - t_{n-j}}} \\ a_{n,i} = -\dfrac{L_i'(t_n)}{L_0'(t_n)} = \dfrac{1}{(t_n - t_{n-i}) \displaystyle\sum_{j=1}^{r} \dfrac{1}{t_n - t_{n-j}}} \displaystyle\prod_{\substack{j=1 \\ j \neq i}}^{r} \dfrac{t_n - t_{n-j}}{t_{n-i} - t_{n-j}} \quad (1 \leq i \leq r) \end{cases}$$

(b) Dans le cas particulier où le pas est constant, on a $h_n = h = \frac{T}{N}$, $t_i = i \cdot h$ ($0 \leq i \leq N$). Ce qui donne $t_n - t_{n-j} = j \cdot h$ pour $1 \leq j \leq r$ et :

$$L_0'(t_n) = \frac{1}{h} \sum_{j=1}^{r} \frac{1}{j}$$

de sorte que :

$$b_n = b = \frac{1}{\displaystyle\sum_{j=1}^{r} \frac{1}{j}}$$

Pour $1 \leq i \leq r$, on a :

$$L_i'(t_n) = -\frac{1}{i \cdot h} \prod_{\substack{j=1 \\ j \neq i}}^{r} \frac{j}{j - i} = -\frac{1}{i^2 h} \frac{(-1)^{i-1} r!}{(i-1)!(r-i)!} = \frac{(-1)^i}{i \cdot h} C_r^i$$

et :

$$a_{n,i} = a_i = \frac{(-1)^{i-1} C_r^i}{i \displaystyle\sum_{j=1}^{r} \frac{1}{j}} \quad (1 \leq i \leq r)$$

(c) L'hypothèse (1) entraîne $\delta \geq 1$ et :

$$\begin{cases} h_{n-i} \leq \delta h_{n-i+1} \leq \cdots \leq \delta^i h_n \\ h_{n-i} \geq \dfrac{1}{\delta} h_{n-i+1} \geq \cdots \geq \dfrac{1}{\delta^i} h_n \end{cases} \quad (0 \leq i \leq n-1)$$

En écrivant, pour $j = 1,2,\cdots,r$, que $t_n - t_{n-j} = h_n + h_{n-1} + \cdots + h_{n-(j-1)}$, on déduit que :

$$\begin{cases} h_n \leq t_n - t_{n-j} \leq h_n\left(1+\delta+\cdots+\delta^{j-1}\right) \quad (1 \leq j \leq r) \\ t_{n-i} - t_{n-j} = h_{n-i} + \cdots + h_{n-(j-1)} > h_{n-i} \geq \dfrac{1}{\delta^i h_n} \quad (1 \leq i < j \leq r) \end{cases}$$

On déduit alors les majorations suivantes :

$$|b_n| = \frac{1}{h_n L_0'(t_n)} \leq \frac{1}{\displaystyle\sum_{j=1}^{r} \frac{1}{1+\delta+\cdots+\delta^{j-1}}} = C(\delta)$$

$$\begin{aligned} |a_{n,i}| &= \frac{\left|L_i'(t_n)\right|}{L_0'(t_n)} = h_n b_n \left|L_i'(t_n)\right| \\ &\leq C(\delta) \prod_{j=1}^{i-1}\left(1+\delta+\cdots+\delta^{j-1}\right)\delta^i \prod_{j=i+1}^{r}\left(1+\delta+\cdots+\delta^{j-1}\right)\delta^j \\ &\leq C(\delta)\left(1+\delta+\cdots+\delta^{r-1}\right)^{r-1}\delta^{\frac{r(r+1)}{2}} \quad (1 \leq i \leq r) \end{aligned}$$

Ce qui donne en définitive les majorations :

$$\forall i = 1,2,\cdots,r, \quad |a_{n,i}| \leq C_1(\delta), \quad |b_n| \leq C_1(\delta)$$

en posant :

$$C_1(\delta) = C(\delta)\left\{\delta^{\frac{r+1}{2}}\left(1+\delta+\cdots+\delta^{r-1}\right)\right\}^r$$

(d) Le polynôme d'interpolation de Lagrange associé à $(t_n, t_{n-1}, \cdots, t_{n-r})$ et $(z, y_{n-1}, \cdots, y_{n-r})$, avec $y_{n-i} = z \neq 0$ pour $1 \leq i \leq r$ est l'unique polynôme de degré inférieur ou égal à r qui vérifie $\pi_{n,r}(t_{n-i}) = z$ pour $0 \leq i \leq r$. Le polynôme constant égal à z convient, donc $z = \pi_{n,r}(t) = \left(\sum_{i=0}^{r} L_i(t)\right)z$ et :

$$\sum_{i=0}^{r} L_i(t) = 1, \quad \sum_{i=0}^{r} L_i'(t) = 0$$

On déduit alors que :

$$\sum_{i=1}^{r} a_{n,i} = -\sum_{i=1}^{r} \frac{L_i'(t_n)}{L_0'(t_n)} = 1$$

D'autre part, on a $\sum_{i=1}^{r}(t_{n-i} - t_n)L_i'(t_n) = \sum_{i=1}^{r} P_i(t_n)$, où $P_i(t)$ est le polynôme de degré $r-1$ défini par :

$$P_i(t)=\prod_{\substack{j=1\\j\neq i}}^{r}\frac{t-t_{n-j}}{t_{n-i}-t_{n-j}}\quad(1\leq i\leq r)$$

Le polynôme $P(t)=\sum_{i=1}^{r}P_i(t)$ est de degré inférieur ou égal à $r-1$ et vérifie $P(t_{n-k})=1$ pour $1\leq k\leq r$. Il est donc identiquement égal à 1 et on a :

$$\sum_{i=1}^{r}(t_{n-i}-t_n)L_i'(t_n)=1$$

Ce qui donne :

$$\sum_{i=1}^{r}(t_n-t_{n-i})a_{n,i}=\sum_{i=1}^{r}(t_{n-i}-t_n)\frac{L_i'(t_n)}{L_0'(t_n)}=\frac{1}{L_0'(t_n)}=h_nb_n$$

(e) Il s'agit de montrer que l'application $\varphi\colon E\to E$ définie par :

$$\forall y\in E,\quad \varphi(y)=\sum_{i=1}^{r}a_{n,i}y_{n-i}+h_nb_nf(t_n,y)$$

admet un unique point fixe dans *I*.

La condition $\|D_yf(t,y)\|\leq L$ sur $[0,T]\times E$ entraîne que *f* est *L*–lipschitzienne par rapport à sa deuxième variable (conséquence immédiate de l'inégalité des accroissements finis) et pour *y*, *z* dans *E*, on a :

$$\|\varphi(y)-\varphi(z)\|=h_nb_n\|f(t_n,y)-f(t_n,z)\|\leq h_nb_nL\|y-z\|$$

La condition $h_nb_nL<1$ nous dit alors que l'application φ est contractante sur l'espace complet *E* et avec le théorème du point fixe (problème 44, Q1) on déduit alors que φ admet un unique point fixe y_n dans *E*.

Q3. (a) La formule de Taylor avec reste intégral à l'ordre *r* permet d'écrire, pour tout $i\in\{1,2,\cdots,r\}$:

$$z(t_{n-i})=\sum_{k=0}^{r}\frac{(t_{n-i}-t_n)^k}{k!}z^{(k)}(t_n)+\int_{t_n}^{t_{n-i}}\frac{(t_{n-i}-t)^r}{r!}z^{(r+1)}(t)dt$$

Ce qui donne :

$$\sum_{i=1}^{r}a_{n,i}z(t_{n-i})=\sum_{k=0}^{r}\frac{1}{k!}\left(\sum_{i=1}^{r}a_{n,i}(t_{n-i}-t_n)^k\right)z^{(k)}(t_n)+\sum_{i=1}^{r}\frac{a_{n,i}}{r!}\int_{t_n}^{t_{n-i}}(t_{n-i}-t)^rz^{(r+1)}(t)dt$$

On connaît déjà $\sum_{i=1}^{r}a_{n,i}(t_{n-i}-t_n)^k$ pour $k=0$ et $k=1$ (Q2 (d)). Le calcul de ces sommes pour $k\in\{2,3,\cdots,r\}$ peut se faire en remarquant que le polynôme d'interpolation de Lagrange associé à $(t_n,t_{n-1},\cdots,t_{n-r})$ et $(z,y_{n-1},\cdots,y_{n-r})$, avec $z\neq 0$ et $y_{n-i}=(t_{n-i}-t_n)^kz$ pour $1\leq i\leq r$ est l'unique polynôme de degré inférieur ou égal à *r* qui vérifie $\pi_{n,r}(t_n)=z$ et $\pi_{n,r}(t_{n-i})=(t_{n-i}-t_n)^kz$ pour $1\leq i\leq r$. Le polynôme $(t-t_n)^kz+\left(\prod_{i=1}^{r}\frac{t-t_{n-i}}{t_n-t_{n-i}}\right)z$ convient, il est donc égal à $\pi_{n,r}(t)$.

Et avec les égalités :

$$z^* = \pi_{n,r}'(t_n) = \left(\sum_{i=1}^{r} \frac{1}{t_n - t_{n-i}}\right) z = L_0'(t_n) z = \frac{1}{h_n b_n} z$$

$$z = \sum_{i=1}^{r} a_{n,i} (t_{n-i} - t_n)^k z + h_n b_n z^*$$

on déduit que :

$$\sum_{i=1}^{r} a_{n,i} (t_{n-i} - t_n)^k = 0 \quad (2 \le k \le r)$$

En définitive, l'erreur de consistance $\varepsilon_n(z)$ peut s'écrire sous la forme :

$$\varepsilon_n(z) = \sum_{i=1}^{r} \frac{a_{n,i}}{r!} \int_{t_{n-i}}^{t_n} (t_{n-i} - t)^r z^{(r+1)}(t) dt$$

Avec l'hypothèse (1), on a les majorations de Q2 (c) pour les coefficients $a_{n,i}$ et :

$$\|\varepsilon_n(z)\| \le \frac{C_1(\delta)}{r!} \sum_{i=1}^{r} \int_{t_{n-i}}^{t_n} (t - t_{n-i})^r \left\| z^{(r+1)}(t) \right\| dt \le \frac{C_1(\delta)}{r!} \sum_{i=1}^{r} \int_{t_{n-i}}^{t_n} (i \cdot h)^r \left\| z^{(r+1)}(t) \right\| dt$$

$$\le \frac{C_1(\delta)}{r!} (r \cdot h)^r r \int_{t_{n-r}}^{t_n} \left\| z^{(r+1)}(t) \right\| dt$$

Soit :

$$\|\varepsilon_n(z)\| \le C_2(\delta) h^r \int_{t_{n-r}}^{t_n} \left\| z^{(r+1)}(t) \right\| dt$$

en posant :

$$C_2(\delta) = \frac{r^r}{(r-1)!} C_1(\delta)$$

(b) Un développement de Taylor avec reste intégral à l'ordre $r + 1$ donne :

$$\sum_{i=1}^{r} a_{n,i} z(t_{n-i}) = z(t_n) - h_n b_n z'(t_n) + \frac{1}{(r+1)!} \left(\sum_{i=1}^{r} a_{n,i} (t_{n-i} - t_n)^{r+1}\right) z^{(r+1)}(t_n)$$

$$+ \sum_{i=1}^{r} \frac{a_{n,i}}{(r+1)!} \int_{t_n}^{t_{n-i}} (t_{n-i} - t)^{r+1} z^{(r+2)}(t) dt$$

Le polynôme d'interpolation de Lagrange associé à $(t_n, t_{n-1}, \cdots, t_{n-r})$ et $(z, y_{n-1}, \cdots, y_{n-r})$, avec $z \ne 0$ et $y_{n-i} = (t_{n-i} - t_n)^{r+1} z$ pour $1 \le i \le r$ est donné par :

$$\pi_{n,r}(t) = (t - t_n)^{r+1} z - \left(\prod_{i=0}^{r} (t - t_{n-i})\right) z + \left(\prod_{i=1}^{r} \frac{t - t_{n-i}}{t_n - t_{n-i}}\right) z$$

Et avec les égalités :

$$z^* = \pi_{n,r}'(t_n) = -\left(\prod_{i=1}^{r} (t_n - t_{n-i})\right) z + L_0'(t_n) z = -\left(\prod_{i=1}^{r} (t_n - t_{n-i})\right) z + \frac{1}{h_n b_n} z$$

$$z = \sum_{i=1}^{r} a_{n,i} (t_{n-i} - t_n)^{r+1} z + h_n b_n z^*$$

on déduit que :

$$\sum_{i=1}^{r} a_{n,i} (t_{n-i} - t_n)^{r+1} = h_n b_n \prod_{i=1}^{r} (t_n - t_{n-i})$$

Ce qui donne en définitive, pour l'erreur de consistance $\varepsilon_n(z)$:

$$\varepsilon_n(z) = -\frac{h_n b_n}{(r+1)!}\left(\prod_{i=1}^{r}(t_n - t_{n-i})\right) z^{(r+1)}(t_n) + \sum_{i=1}^{r}\frac{a_{n,i}}{(r+1)!}\int_{t_{n-i}}^{t_n}(t_{n-i}-t)^{r+1} z^{(r+2)}(t)\,dt$$

Soit :

$$\varepsilon_n(z) - \gamma_{n,r} b_n h_n^{\,r+1} z^{(r+1)}(t_n) = \sum_{i=1}^{r}\frac{a_{n,i}}{(r+1)!}\int_{t_{n-i}}^{t_n}(t_{n-i}-t)^{r+1} z^{(r+2)}(t)\,dt$$

où on a posé :

$$\gamma_{n,r} = -\frac{1}{h_n^{\,r}(r+1)!}\prod_{i=1}^{r}(t_n - t_{n-i}) = -\frac{1}{(r+1)!}\prod_{i=2}^{r}\frac{t_n - t_{n-i}}{t_n - t_{n-1}}$$

Avec l'hypothèse (1), on a alors la majoration :

$$\left\|\varepsilon_n(z) - \gamma_{n,r} b_n h_n^{\,r+1} z^{(r+1)}(t_n)\right\| \le \frac{C_1(\delta)}{(r+1)!}\sum_{i=1}^{r}\int_{t_{n-r}}^{t_n}\left(1+\delta+\cdots+\delta^{r-1}\right)^{r+1} h_n^{\,r+1}\left\|z^{(r+2)}(t)\right\|dt$$

$$\le C_3(\delta) h_n^{\,r+1}\int_{t_{n-r}}^{t_n}\left\|z^{(r+2)}(t)\right\|dt$$

où on a posé :

$$C_3(\delta) = \frac{r}{(r+1)!}C_1(\delta)\left(1+\delta+\cdots+\delta^{r-1}\right)^{r+1}$$

Dans le cas particulier où le pas est constant, on a $h_n = h$ et pour $2 \le i \le r$, $t_n - t_{n-i} = i \cdot h = i(t_n - t_{n-1})$, de sorte que :

$$\gamma_{n,r} = \gamma_r = -\frac{1}{r+1}$$

Q4. (a) On a :

$$u_n = y_n - z_n = \sum_{i=1}^{r} a_{n,i} u_{n-i} + h_n b_n\big(f(t_n, y_n) - f(t_n, z_n)\big) - \varepsilon_n$$

Ce qui peut s'écrire, avec le résultat de Q1 :

$$u_n = \sum_{i=1}^{r} a_{n,i} u_{n-i} + h_n b_n g(t_n; y_n, z_n) u_n - \varepsilon_n$$

Ou encore, sous forme matricielle :

$$U_n = R_n U_{n-1} + h_n b_n G_n U_n + E_n \quad (r \le n \le N)$$

en posant :

$$R_n = \begin{pmatrix} a_{n,1} & a_{n,2} & & \cdots & a_{n,r} \\ 1 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ . & . & . & . & . \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix},\ G_n = \begin{pmatrix} g(t_n; y_n, z_n) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & 0 \end{pmatrix},\ E_n = \begin{pmatrix} -\varepsilon_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(b) Dans le cas d'une subdivision à pas constant, on a vu en Q2 (b) que :

$$a_{n,i} = a_i = \frac{(-1)^{i-1} C_r^i}{i\sum_{j=1}^{r}\frac{1}{j}} \quad (1 \le i \le r)$$

La relation $\sum_{i=1}^{r} a_i = 1$, obtenue en Q2 (d), montre que 1 est valeur propre de R avec $\xi = (1, \ \cdots, \ 1)^T$ pour vecteur propre associé.

Pour $r = 2$, on a $R = \begin{pmatrix} \frac{4}{3} & -\frac{1}{3} \\ 1 & 0 \end{pmatrix}$ avec $\lambda_1 = 1$ et $\lambda_2 = \frac{1}{3}$ pour valeurs propres et la condition (D) est vérifiée.

Pour $r = 3$, on a $R = \begin{pmatrix} \frac{18}{11} & -\frac{9}{11} & \frac{2}{11} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ de valeurs propres $\lambda_1 = 1$, λ_2, λ_3 avec $\lambda_2 + \lambda_3 = Tr(R) - 1 = \frac{7}{11}$ et $\lambda_2 \cdot \lambda_3 = Det(R) = \frac{2}{11}$. Les valeurs propres λ_2 et λ_3 sont donc complexes conjuguées avec $|\lambda_2|^2 = \frac{2}{11} < 1$ et l'hypothèse (D) du problème 24 est vérifiée.

Pour $r = 4$, on a :

$$R = \begin{pmatrix} \frac{48}{25} & -\frac{36}{25} & \frac{16}{25} & -\frac{3}{25} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Le polynôme caractéristique de R est :

$$P_R(x) = (x-1)\left\{-x^3 + \frac{23}{25}x^2 - \frac{13}{25}x + \frac{3}{25}\right\} = (x-1)Q(x)$$

Comme $Q'(x) < 0$ sur $I\!R$, on déduit que Q décroît strictement de $+\infty$ à $-\infty$ et donc qu'elle s'annule une seule fois en $\lambda_2 \in \left]0,1\right[$ ($Q(0)Q(1) < 0$), les deux autres racines λ_3 et λ_4 étant complexes conjuguées. Avec $Q\left(\frac{1}{3}\right)Q(1) < 0$, on déduit que $\lambda_2 \in \left]\frac{1}{3}, 1\right[$ et $|\lambda_3|^2 = |\lambda_4|^2 = \frac{1}{\lambda_1}\frac{3}{25} < \frac{9}{25} < 1$. Donc la condition (D) est vérifiée.

(c) La relation $\sum_{i=1}^{r} a_{n,i} = 1$ entraîne que 1 est valeur propre de R_n avec ξ pour vecteur propre associé, c'est-à-dire que la matrice R_n est dans l'ensemble S défini en Q1 (c) du problème 24.

D'autre part, pour $r \in \{2,3,4\}$, la matrice R vérifie la condition (D). On peut donc trouver, d'après Q2 (c) du problème 24, une matrice inversible H_r et un réel $\varepsilon > 0$ tels que :

$$\|R_n - R\|_1 \leq \varepsilon \quad \Rightarrow \quad \left\|H_r^{-1} R_n H_r\right\|_1 \leq 1$$

Il s'agit alors de montrer qu'il existe un réel α_r tel que la condition (3) entraîne $\|R_n - R\|_1 < \varepsilon$.

Pour ce faire, on remarque que la matrice R_n est fonction des coefficients $a_{n,i}$ $(1 \leq i \leq r)$ qui sont des fonctions rationnelles des $\gamma_n = \dfrac{h_n}{h_{n-1}}$ $(2 \leq n \leq N)$.

Ces fonctions rationnelles étant continues au voisinage de $\gamma = (1, \cdots, 1)$, il en est de même de R_n.

Dans le cas où tous les γ_n sont égaux à 1 (pas constant), on a $R_n = R$ et par continuité il existe un réel $\alpha_r > 0$ tel que :

$$\left(\forall n = 2,\cdots,N, \quad \left|1 - \frac{h_n}{h_{n-1}}\right| \leq \alpha_r\right) \quad \|R_n - R\|_1 \leq \varepsilon \quad \Rightarrow \quad \left\|H_r^{-1} R_n H_r\right\|_1 \leq 1$$

On peut toujours prendre $\alpha_r \in \left]0,1\right[$ et en posant $\dfrac{1}{\delta} = 1 - \alpha_r$, on a :

$$\frac{1}{\delta} = 1 - \alpha_r < \frac{h_n}{h_{n-1}} < 1 + \alpha_r < \delta$$

C'est-à-dire que la condition (1) de Q2 (c) est vérifiée.

(d) On a $\|G_n\|_1 = \|g(t_n; y_n, z_n)\| \leq L$ (voir Q1) et avec la condition $h_n b_n L \leq \nu < 1$ $(n \geq r)$, on déduit que $\|h_n b_n G_n\|_1 < 1$ et que la matrice $I - h_n b_n G_n$ est inversible d'inverse $\sum_{k \geq 0} h_n{}^k b_n{}^k G_n{}^k$.

On a alors :

$$H_r^{-1}(I - h_n b_n G_n)^{-1} H_r = I + H_r^{-1} M H_r$$

avec $M = \sum_{k \geq 1} h_n{}^k b_n{}^k G_n{}^k = h_n b_n G_n \sum_{k \geq 0} h_n{}^k b_n{}^k G_n{}^k = h_n b_n G_n (I - h_n b_n G_n)^{-1}$.

Il reste donc à majorer $\|M\|_1$. Ce qui se fait avec :

$$\|M\|_1 \leq h_n b_n \|G_n\|_1 \sum_{k \geq 0} h_n{}^k b_n{}^k \|G_n\|_1^k \leq h_n b_n L \sum_{k \geq 0} (h_n b_n L)^k \leq h_n b_n L \sum_{k \geq 0} \nu^k = \frac{h_n b_n L}{1 - \nu}$$

Ce qui donne :

$$\left\|H_r^{-1}(I - h_n b_n G_n)^{-1} H_r\right\|_1 \leq 1 + \left\|H_r^{-1}\right\|_1 \|H_r\|_1 \frac{b_n}{1 - \nu} h_n L$$

Et avec la majoration $0 < |b_n| \leq C_1(\delta)$ de Q2 (c), on déduit en définitive que :

$$\left\|H_r^{-1}(I - h_n b_n G_n)^{-1} H_r\right\|_1 \leq 1 + C_4 h_n L$$

où on a posé $C_4 = \left\|H_r^{-1}\right\|_1 \|H_r\|_1 \dfrac{C_1(\delta)}{1 - \nu}$.

(e) En utilisant le résultat de (a), on peut écrire que :

$$H_r^{-1} U_n = H_r^{-1}(I - h_n b_n G_n)^{-1} H_r \left(H_r^{-1} R_n H_r\right) H_r^{-1} U_{n-1} + H_r^{-1}(I - h_n b_n G_n)^{-1} H_r \left(H_r^{-1} E_n\right)$$

Les résultats de (c) et (d) entraînent que, sous l'hypothèse (3), on a :

$$\left\|H_r^{-1} U_n\right\|_1 \leq (1 + C_4 h_n L)\left(\left\|H_r^{-1} U_{n-1}\right\|_1 + \left\|H_r^{-1} E_n\right\|_1\right)$$

Posons, à r fixé :

$$x_n = \left\| H_r^{-1} U_n \right\|_1, \quad \alpha_n = C_4 h_n L, \quad \beta_n = \left\| H_r^{-1} E_n \right\|_1 \quad (r-1 \le n \le N)$$

On a alors :

$$0 \le x_n \le (1+\alpha_n)(x_{n-1} + \beta_n) \quad (r-1 \le n \le N)$$

Par récurrence, on déduit que :

$$0 \le x_n \le (1+\alpha_n)(1+\alpha_{n-1})\cdots(1+\alpha_r)x_{r-1}$$
$$+\beta_n(1+\alpha_n) + \beta_{n-1}(1+\alpha_n)(1+\alpha_{n-1}) + \cdots + \beta_r(1+\alpha_n)\cdots(1+\alpha_r)$$

et :

$$0 \le x_n \le \prod_{i=r}^{n} (1+\alpha_i)\left(x_{r-1} + \sum_{i=r}^{n} \beta_i \right)$$

Et en écrivant que $\prod_{i=r}^{n}(1+\alpha_i) \le \prod_{i=r}^{n} e^{\alpha_i} = e^{\sum_{i=r}^{n}\alpha_i}$, on obtient l'inégalité de Gronwall (voir le problème 44, Q5) :

$$0 \le x_n \le e^{\sum_{i=r}^{n}\alpha_i}\left(x_{r-1} + \sum_{i=r}^{n} \beta_i \right) \quad (r \le n \le N)$$

Avec les inégalités :

$$\sum_{i=r}^{n} \alpha_i = C_4 L \sum_{i=r}^{n} h_i = C_4 L (t_n - t_{r-1}) \le C_4 L t_n$$

$$\sum_{i=r}^{n} \beta_i \le \left\| H_r^{-1} \right\|_1 \sum_{i=r}^{n} \left\| E_i \right\|_1 = \left\| H_r^{-1} \right\|_1 \sum_{i=r}^{n} \left\| \varepsilon_i \right\|$$

on déduit que :

$$\left\| H_r^{-1} U_n \right\|_1 \le e^{C_4 L t_n} \left\{ \left\| H_r^{-1} U_{n-1} \right\|_1 + \left\| H_r^{-1} \right\|_1 \sum_{i=r}^{n} \left\| \varepsilon_i \right\| \right\}$$

Enfin avec $\left\| H_r^{-1} U_{n-1} \right\|_1 \le \left\| H_r^{-1} \right\|_1 \left\| U_{n-1} \right\|_1$ et $\left\| U_n \right\|_1 = \left\| H_r H_r^{-1} U_n \right\|_1 \le \left\| H_r \right\|_1 \left\| H_r^{-1} U_n \right\|_1$, on déduit que :

$$\left\| U_n \right\|_1 \le C_5 e^{C_4 L t_n} \left\{ \left\| U_{r-1} \right\|_1 + \sum_{i=r}^{n} \left\| \varepsilon_i \right\| \right\}$$

où on a posé $C_5 = \left\| H_r^{-1} \right\|_1 \left\| H_r \right\|_1$.

Remarque — Ce résultat montre que la méthode est stable (voir le problème 44).

Q5. Dans cette question on s'intéresse à la convergence de la méthode en donnant une majoration de l'erreur $y_n - y(t_n)$.

Si f est de classe C^r, alors la solution y du problème (P) est de classe C^{r+1} et l'erreur de consistance $\varepsilon_n(y) = y(t_n) - \sum_{i=1}^{r} a_{n,i} y(t_{n-i}) - h_n b_n f(t_n, y(t_n))$ vérifie la majoration de Q3 (a) $\left\| \varepsilon_n(y) \right\| \le C_2(\delta) h^r \int_{t_{n-r}}^{t_n} \left\| y^{(r+1)}(t) \right\| dt$.

En prenant $u_n = y_n - y(t_n)$ dans Q4, on déduit que :

$$\left\| y_n - y(t_n) \right\| = \left\| u_n \right\| \le \left\| U_n \right\|_1 \le C_5 e^{C_4 L t_n} \left\{ \left\| U_{r-1} \right\|_1 + \sum_{i=r}^{n} \left\| \varepsilon_i(y) \right\| \right\}$$

avec $\left\|U_{r-1}\right\|_1 = \sum_{\mu=0}^{r-1}\left\|y_\mu - y(t_\mu)\right\|$ et :

$$\sum_{i=r}^{n}\left\|\varepsilon_i(y)\right\| \le C_2(\delta)h^r\sum_{i=r}^{n}\int_{t_{i-r}}^{t_i}\left\|y^{(r+1)}(t)\right\|dt \le C_2(\delta)h^r(N-r)\int_0^{t_n}\left\|y^{(r+1)}(t)\right\|dt$$

on déduit qu'on a la majoration de l'erreur :

$$\left\|y_n - y(t_n)\right\|_1 \le C_5 e^{C_4 L t_n}\left\{\sum_{\mu=0}^{r-1}\left\|y_\mu - y(t_\mu)\right\| + C_6 h^r\int_0^{t_n}\left\|y^{(r+1)}(t)\right\|dt\right\}$$

où on a posé $C_6 = (N-r)C_2(\delta)$.

CHAPITRE 8

Analyse fonctionnelle

Problème 47 : Théorème de projection sur un convexe fermé dans un espace de Hilbert. Applications

Dans ce problème E désigne un espace de Hilbert réel. On note $\langle \ | \ \rangle$ un produit scalaire sur E et $\| \ \|$ la norme associée.

On désigne par E' le « dual topologique » de E, c'est-à-dire l'espace vectoriel applications linéaires continues de E dans IR.

On rappelle que si E_1 et E_2 sont deux espaces vectoriels normés, alors une application linéaire $u: E_1 \to E_2$ est continue si et seulement si :

$$\|u\| = Sup\left\{ \frac{\|u(x)\|}{\|x\|}; \quad x \in E_1 - \{0\} \right\} < +\infty$$

Une forme bilinéaire $\varphi: E \times E \to IR$ est continue si et seulement si

$$Sup\left\{ \frac{|\varphi(x,y)|}{\|x\|\|y\|}; \quad (x,y) \in E \times E - \{(0,0)\} \right\} < +\infty$$

On dit qu'une forme bilinéaire $\varphi: E \times E \to IR$ est coercive s'il existe une constante $\alpha > 0$ telle que :

$$\forall x \in E, \quad |\varphi(x,x)| \geq \alpha \|x\|^2$$

Pour toute partie A non vide de E, on désigne par $\overline{A}$ l'adhérence de A et par $A^\perp$ l'orthogonal de A dans E, c'est-à-dire :

$$A^\perp = \{ y \in E; \quad \forall x \in A, \langle x | y \rangle = 0 \}$$

Q1. *Soit F une partie convexe, fermée et non vide de E.*
(a) Montrer que pour tout $x \in E$ il existe un unique $y \in F$ tel que :

$$\|x - y\| = d(x,F) = Inf\{\|x - z\|; z \in F\}$$

On note $y = P_F(x)$ et l'application $P_F: E \to F$ ainsi définie est appelée la projection de E sur F (on dit aussi que $y = P_F(x)$ est la meilleure approximation de x dans F).

(b) Soit $x \in E$. *Montrer l'équivalence :*

$$\left(y = P_F(x)\right) \iff \left(y \in F \text{ et } \forall z \in F, \langle x-y \mid z-y\rangle \leq 0\right)$$

(c) Montrer que l'application P_F *est lipschitzienne.*

Q2. *Pour cette question, on suppose que* $E = IR^n$ *est muni du produit scalaire usuel.*

(a) On désigne par F « l'hyperoctant positif » défini par :

$$F = \left\{x \in IR^n;\quad \forall i \in \{1,2,\cdots,n\}, x_i \geq 0\right\}$$

Calculer $P_F(x)$ *pour tout* $x \in E$.

(b) On désigne par F un pavé de E défini par :

$$F = \prod_{i=1}^{n}[a_i, b_i] = \left\{x \in IR^n;\quad \forall i \in \{1,2,\cdots,n\}, a_i \leq x_i \leq b_i\right\}$$

Calculer $P_F(x)$ *pour tout* $x \in E$.

Q3. *Pour cette question, on suppose que F est un sous-espace vectoriel fermé de E.*

(a) Soit $x \in E$. *Montrer l'équivalence :*

$$\left(y = P_F(x)\right) \iff \left(y \in F \text{ et } y - x \in F^{\perp}\right)$$

(b) Montrer que l'application P_F *est linéaire continue de E sur F.*

Calculer sa norme et son noyau.

(c) Montrer que $E = F \oplus F^{\perp}$.

(d) Montrer que $\left(F^{\perp}\right)^{\perp} = F$.

Q4. *Pour cette question, on suppose que F est un sous-espace vectoriel de E, non nécessairement fermé.*

(a) Montrer que $F^{\perp}$ *est un sous-espace vectoriel fermé de E.*

(b) Montrer que $\left(F^{\perp}\right)^{\perp} = \overline{F}$.

Q5. *Pour tout* $x \in E$, *on désigne par* $\varphi(x)$ *la forme linéaire définie par :*

$$\forall y \in E,\quad \varphi(x)(y) = \langle x \mid y\rangle$$

Montrer que φ *définie un isomorphisme de E sur E′ qui conserve la norme (isométrie).*

Ce résultat est le théorème de représentation de Riesz–Fréchet.

Q6. *Montrer que pour tout opérateur linéaire et continu* $u: E \to E$ *il existe un unique opérateur linéaire et continu* ${}^t u: E \to E$ *tel que :*

$$\forall (x,y) \in E \times E,\quad \langle u(x) \mid y\rangle = \langle x \mid {}^t u(y)\rangle$$

On dit que ${}^t u$ *est l'adjoint de u.*

Montrer que $\left\|{}^t u\right\| = \|u\|$.

Q7. *L'exemple qui suit montre que le théorème de Riesz–Fréchet est en défaut si on suppose seulement l'espace préhilbertien.*

Soit $E = C^0([0,1], IR)$ *l'espace préhilbertien des fonctions continues de* $[0,1]$ *dans IR, avec le produit scalaire* $(f,g) \mapsto \langle f \mid g\rangle = \int_0^1 f(x)g(x)dx$.

On définit sur E l'application $u: f \mapsto u(f) = \int_0^a x^p f(x)dx$ *avec* $p > 0$ *et* $0 < a < 1$.

(a) Montrer que u est une forme linéaire continue et calculer sa norme.
(b) Montrer qu'il n'existe pas de vecteur $g \in E$ tel que $u(f) = \langle f \mid g \rangle$ pour tout $f \in E$.
Q8. *On désigne par F une partie convexe fermée non vide de E.*
(a) Soient $z \in E$ et $v : E \to E$ une application linéaire continue telle que la forme bilinéaire $\psi : (x, y) \mapsto \langle v(x) \mid y \rangle$ soit coercive.

Montrer qu'il existe un réel $\lambda > 0$ tel que l'application $f : E \to E$ définie par :

$$\forall x \in E, \quad f(x) = p_F(\lambda z + x - \lambda v(x))$$

admette un point fixe.
(b) Soit $\varphi : E \times E \to IR$ une forme bilinéaire continue et coercive.

Montrer que pour toute forme linéaire continue $u \in E'$, il existe un unique $x \in F$ tel que :

$$\forall y \in F, \quad \varphi(x, y - x) \geq u(y - x)$$

(théorème de Stampacchia).
(c) On suppose de plus que la forme bilinéaire φ est symétrique. Pour toute forme linéaire continue $u \in E'$, on définit l'application $J : E \to IR$ par :

$$\forall y \in E, \quad J(y) = \frac{1}{2}\varphi(y, y) - u(y)$$

Montrer l'équivalence :

$$\left(x \in F \text{ et } \forall y \in F, \varphi(x, y - x) \geq u(y - x)\right) \Leftrightarrow \left(x \in F \text{ et } J(x) = \operatorname*{Inf}_{y \in F} J(y)\right)$$

Q9. *Soit $\varphi : E \times E \to IR$ une forme bilinéaire continue et coercive.*
(a) Montrer que pour toute forme linéaire continue $u \in E'$, il existe un unique $x \in E$ tel que :

$$\forall y \in E, \quad u(y) = \varphi(x, y)$$

(théorème de Lax–Milgram).
(b) On suppose de plus que la forme bilinéaire φ est symétrique. Comme en Q8, (c), on associe à toute forme linéaire continue $u \in E'$, l'application $J : E \to IR$

Montrer l'équivalence :

$$\left(\forall y \in E, \quad u(y) = \varphi(x, y)\right) \Leftrightarrow \left(J(x) = \operatorname*{Inf}_{y \in E} J(y)\right)$$

Solution

Q1. (a) La distance de x à F, $\delta = d(x, F)$, est bien définie comme borne inférieure d'une partie non vide et minorée (par 0) de *IR*. Par définition de cette borne inférieure, on peut trouver une suite $(y_n)_{n \geq 1}$ de points de F telle que :

$$\forall n \geq 1, \quad \delta^2 \leq \|x - y_n\|^2 \leq \delta^2 + \frac{1}{n} \quad (1)$$

En utilisant l'identité de la médiane $\left(\|a + b\|^2 + \|a - b\|^2 = 2\left(\|a\|^2 + \|b\|^2\right)\right)$, on peut écrire, pour $q > p \geq 1$:

$$\left\|y_q - y_p\right\|^2 = \left\|\left(y_q - x\right) + \left(x - y_p\right)\right\|^2 = 2\left(\left\|y_q - x\right\|^2 + \left\|x - y_p\right\|^2\right) - \left\|\left(y_q - x\right) - \left(x - y_p\right)\right\|^2$$

avec :

$$\left\|\left(y_q - x\right) - \left(x - y_p\right)\right\|^2 = 4\left\|x - \frac{1}{2}\left(y_p + y_q\right)\right\|^2 \geq \delta^2$$

du fait que $\frac{1}{2}\left(y_p + y_q\right) \in F$ (F est convexe).

On a donc, pour $q > p \geq 1$:

$$\left\|y_q - y_p\right\|^2 \leq 2\left(\delta^2 + \frac{1}{q} + \delta^2 + \frac{1}{p}\right) - 4\delta^2$$

Soit :

$$\forall q > p \geq 1, \quad \left\|y_q - y_p\right\|^2 \leq 2\left(\frac{1}{q} + \frac{1}{p}\right) < \frac{4}{p} \underset{p \to +\infty}{\to} 0$$

On a donc ainsi montré que la suite $\left(y_n\right)_{n \geq 1}$ est de Cauchy dans l'espace complet E, elle est donc convergente et sa limite y est dans F qui est fermé.

Et en faisant tendre n vers l'infini dans (1), on déduit que $\|x - y\| = \delta$.

Il reste à montrer l'unicité d'un tel point réalisant la distance de y à F.

Si $z \in F$ est un autre point de F réalisant la distance de x à F, on peut alors écrire en utilisant l'identité de la médiane :

$$\|y - z\|^2 = \|(y - x) + (x - z)\|^2 = 2\left(\|y - x\|^2 + \|x - z\|^2\right) - 4\left\|x - \frac{1}{2}(y + z)\right\|^2 \leq 4\delta^2 - 4\delta^2 = 0$$

(la convexité de F entraîne que $\frac{1}{2}(y + z)$ est dans F). Et nécessairement $z = y$.

(b) Soit $y = P_F(x)$. On sait déjà que $y \in F$. L'ensemble F étant convexe, on a :

$$\forall z \in F, \quad \forall t \in [0,1], \quad v = (1 - t)y + tz \in F$$

et :

$$\|x - y\|^2 \leq \|x - v\|^2 = \|x - y\|^2 - 2\langle x - y | z - y\rangle t + t^2\|z - y\|^2$$

C'est-à-dire que :

$$\forall z \in F, \quad \forall t \in [0,1], \quad 2\langle x - y | z - y\rangle t \leq t^2\|z - y\|^2$$

En divisant ces inégalités par $t \in]0,1]$, puis en faisant tendre t vers 0 (par valeurs positives, on déduit que (voir figure 8.1) :

$$\forall z \in F, \quad \langle x - y | z - y\rangle \leq 0 \quad (2)$$

Réciproquement supposons que $y \in F$ vérifie (2). Pour tout $z \in F$, on peut écrire :

$$\|x - z\|^2 \leq \|(x - y) - (z - y)\|^2 = \|x - y\|^2 - 2\langle x - y | z - y\rangle + \|z - y\|^2 \geq \|x - y\|^2$$

ce qui équivaut à dire que $y = P_F(x)$.

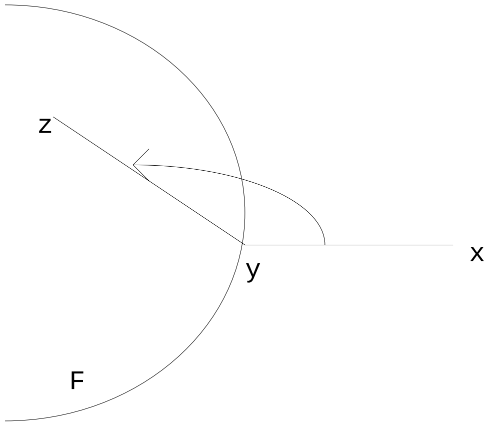

Figure 8.1

(c) Soient x_1, x_2 dans E et $y_1 = P_F(x_1)$, $y_2 = P_F(x_2)$ (dans F).

En utilisant la caractérisation du (b) de la projection, on obtient les inégalités :

$$\langle x_1 - y_1 | y_2 - y_1 \rangle \le 0$$

$$\langle x_2 - y_2 | y_1 - y_2 \rangle \le 0$$

et en additionnant ces deux inégalités :

$$\langle (x_1 - x_2) + (y_2 - y_1) | y_2 - y_1 \rangle \le 0$$

Ce qui donne, avec l'inégalité de Cauchy–Schwarz :

$$\|y_2 - y_1\|^2 \le \langle x_1 - x_2 | y_1 - y_2 \rangle \le \|x_1 - x_2\| \|y_2 - y_1\|$$

C'est-à-dire :

$$\|P_F(x_2) - P_F(x_1)\| \le \|x_2 - x_1\|$$

On a donc ainsi prouvé que P_F est lipschitzienne (et donc uniformément continue).

Q2. (a) Soit $x = (x_1, \cdots, x_n) \in IR^n$ et $y \in F$ défini par :

$$y_i = Max\{0, x_i\} \quad (1 \le i \le n)$$

Pour tout $z \in F$, on a alors :

$$\langle x - y | z - y \rangle = \sum_{i=1}^{n} (x_i - y_i)(z_i - y_i) = \sum_{x_i < 0} x_i z_i \le 0$$

ce qui prouve que $y = P_F(x)$ (figure 8.2).

(b) Soit $x = (x_1, \cdots, x_n) \in IR^n$ et $y \in F$ défini par :

$$y_i = Min\{Max\{x_i, a_i\}, b_i\} = \begin{cases} a_i \text{ si } x_i < a_i \\ x_i \text{ si } a_i \le x_i \le b_i \\ b_i \text{ si } b_i < x_i \end{cases} \quad (1 \le i \le n)$$

Pour tout $z \in F$, on a alors :

$$\langle x - y | z - y \rangle = \sum_{x_i < a_i} (x_i - a_i)(z_i - a_i) + \sum_{x_i > b_i} (x_i - b_i)(z_i - b_i) \le 0$$

ce qui prouve que $y = P_F(x)$ (figure 8.3).

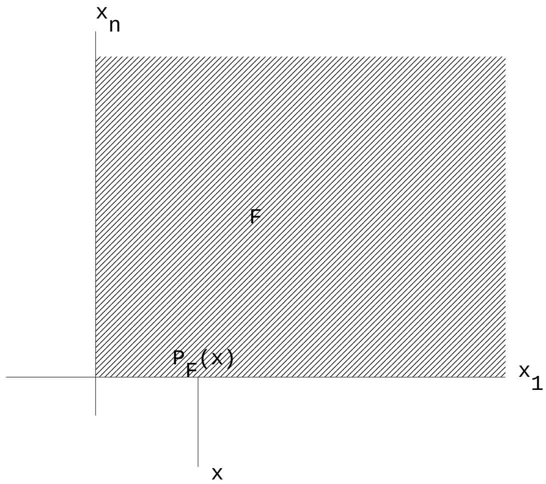

Figure 8.2

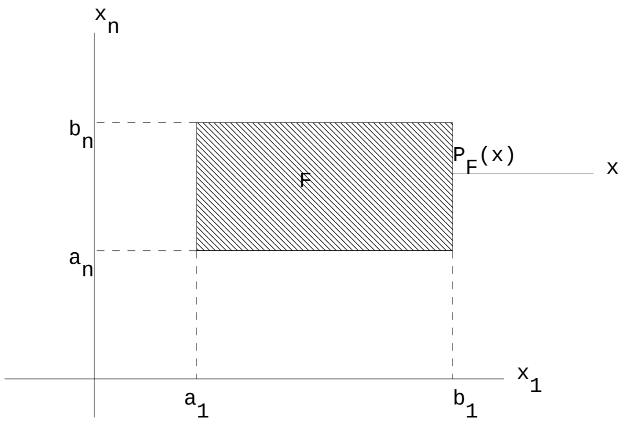

Figure 8.3

Q3. (a) Soit $x \in E$ et $y = P_F(x)$. Pour tout $z \in F$ et tout $t \in IR$, on a $t \cdot z \in F$ (F est un espace vectoriel) et $\langle x - y | t \cdot z - y \rangle \leq 0$. C'est-à-dire que pour tout $z \in F$, on a :

$$\forall t \in IR, \quad \langle x - y | z \rangle t - \langle x - y | y \rangle \leq 0$$

et nécessairement $\langle x - y | z \rangle = 0$.

On a donc $y \in F$ et $x - y \in F^{\perp}$.

Réciproquement soit $y \in F$ tel que $x - y \in F^{\perp}$. On a alors :

$$\forall z \in F, \quad \langle x - y | z - y \rangle = 0$$

et $y = P_F(x)$ d'après Q1 (b).

(b) Soient x_1, x_2 dans E et α_1, α_2 dans IR. F étant un espace vectoriel, on a $\alpha_1 P_F(x_1) + \alpha_2 P_F(x_2) \in F$. De plus pour tout $z \in F$, on a :

$$\langle \alpha_1 x_1 + \alpha_2 x_2 - \alpha_1 P_F(x_1) - \alpha_2 P_F(x_2) | z \rangle = \alpha_1 \langle x_1 - P_F(x_1) | z \rangle + \alpha_2 \langle x_2 - P_F(x_2) | z \rangle = 0$$

ce qui prouve, d'après (a) que :

$$P_F(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 P_F(x_1) + \alpha_2 P_F(x_2)$$

C'est-à-dire que P_F est une application linéaire de E dans F.

On sait déjà que P_F est 1–lipschitzienne (Q1 (c)), c'est donc une application continue avec $\|P_F\| \leq 1$.

En considérant que $p_F(x) = x$ pour tout x dans F, on déduit que $\|P_F\| = 1$.

Le noyau de P_F est formé des vecteurs $x \in E$ qui vérifient $\langle x | z \rangle = 0$ pour tout $z \in F$, c'est donc l'orthogonal de F : $Ker(P_F) = F^{\perp}$.

(c) Tout vecteur $x \in E$ peut s'écrire $x = (x - p_F(x)) + p_F(x)$ avec $p_F(x) \in F$ et $x - p_F(x) \in F^{\perp}$. On a donc $E = F + F^{\perp}$.

Si $x \in F \cap F^{\perp}$; alors $\|x\|^2 = \langle x | x \rangle = 0$ et $x = 0$. On a donc la somme directe $E = F \oplus F^{\perp}$.

(d) Il est clair que $F \subset (F^{\perp})^{\perp}$.

Soit $x \in (F^{\perp})^{\perp}$ et $y = P_F(x) \in F$ (F est fermé). On a $x - p_F(x) \in F^{\perp}$ et pour tout $z \in F^{\perp}$:

$$\langle x - y | z \rangle = \langle x | z \rangle - \langle y | z \rangle = 0$$

C'est-à-dire que $x - y \in F^{\perp} \cap (F^{\perp})^{\perp} = \{0\}$ et $x = y \in F$.

On a donc bien $(F^{\perp})^{\perp} = F$ pour tout sous-espace vectoriel fermé de E.

Q4. (a) Il est clair que $F^{\perp}$ est un sous-espace vectoriel fermé de E. L'inégalité de Cauchy–Schwarz (qui entraîne la continuité du produit scalaire) montre que c'est un fermé.

(b) On a déjà $F \subset (F^{\perp})^{\perp}$, donc $\overline{F} \subset (F^{\perp})^{\perp}$ ($(F^{\perp})^{\perp}$ est un fermé).

L'inclusion $F \subset \overline{F}$ entraîne $\overline{F}^{\perp} \subset F^{\perp}$ et $(F^{\perp})^{\perp} \subset (\overline{F}^{\perp})^{\perp} = \overline{F}$ (d'après Q3 (d)).

On a don bien $(F^{\perp})^{\perp} = \overline{F}$ pour tout sous-espace vectoriel de E.

Q5. Avec la bilinéarité du produit scalaire, on déduit que pour tout $x \in E$ l'application $\varphi(x)$ est une forme linéaire. Avec l'inégalité de Cauchy–Schwarz que $\varphi(x)$ est continue et que $\|\varphi(x)\| \leq \|x\|$.

En écrivant que $\varphi(x)(x) = \|x\|^2$, on déduit l'égalité $\|\varphi(x)\| = \|x\|$.

Il reste à montrer que l'application φ est surjective (l'injectivité découlant de $\|\varphi(x)\| = \|x\|$).

Soit $u \in E'$, il s'agit de trouver $x \in E$ tel que $u = \varphi(x)$, c'est-à-dire tel que :

$$\forall y \in E, \quad u(y) = \langle x | y \rangle$$

Et en particulier :

$$\forall y \in F = Ker(u), \quad \langle x | y \rangle = 0$$

c'est-à-dire que x est nécessairement dans $F^{\perp}$.

Si $u = 0$, il suffit alors de prendre $x = 0$. Pour $u \neq 0$, on a $F = Ker(u) \neq E$ et $F^{\perp} \neq \{0\}$ (puisque $E = F \oplus F^{\perp}$ pour F fermé).

Soit $z \in F^{\perp}$ tel que $\|z\| = 1$ et $x = \alpha \cdot z$ où $\alpha \in IR^*$ est défini par $u(x) = \|x\|^2$ ($\alpha^2 = \|x\|^2 = u(x) = \alpha \cdot u(z) \Rightarrow \alpha = u(z)$).

Pour tout $y \in E$, on pose $y_1 = y - \frac{u(y)}{\|x\|^2} x \in F$ ($u(y_1) = 0$), $y_2 = \frac{u(y)}{\|x\|^2} x \in F^{\perp}$ et on a $\langle x | y \rangle = \langle x | y_1 \rangle + \langle x | y_2 \rangle = \langle x | y_2 \rangle = u(y)$, c'est-à-dire que $u = \varphi(x)$.

Q6. Soit $y \in E$ et $f \in E'$ définie par :

$$\forall x \in E, \quad f(x) = \langle u(x) | y \rangle$$

Le théorème de représentation de Riesz (Q5) nous dit alors qu'il existe un unique vecteur ${}^t u(y) \in E$ tel que :

$$\forall x \in E, \quad f(x) = \langle x | {}^t u(y) \rangle$$

Ce qui détermine de manière unique l'application ${}^t u$.

Pour y_1, y_2 dans E et α_1, α_2 dans IR, on a :

$$\forall x \in E, \langle x | {}^t u(\alpha_1 y_1 + \alpha_2 y_2) \rangle = \langle u(x) | \alpha_1 y_1 + \alpha_2 y_2 \rangle = \alpha_1 \langle u(x) | y_1 \rangle + \alpha_2 \langle u(x) | y_2 \rangle$$
$$= \alpha_1 \langle x | {}^t u(y_1) \rangle + \alpha_2 \langle x | {}^t u(y_2) \rangle = \langle x | \alpha_1 {}^t u(y_1) + \alpha_2 {}^t u(y_2) \rangle$$

Ce qui prouve la linéarité de l'application ${}^t u$.

D'autre part, en utilisant l'inégalité de Cauchy–Schwarz, on obtient l'inégalité :

$$\|{}^t u(y)\|^2 = \langle u({}^t u(y)) | y \rangle \leq \|u\| \|{}^t u(y)\| \|y\|$$

qui entraîne $\|{}^t u(y)\| \leq \|u\| \|y\|$ et $\|{}^t u\| \leq \|u\|$.

En remarquant que ${}^t({}^t u) = u$, on déduit que $\|u\| \leq \|{}^t u\|$ et $\|u\| = \|{}^t u\|$.

Q7. (a) En utilisant l'inégalité de Cauchy–Schwarz, on obtient, pour tout $f \in E$:

$$|u(f)| \leq \sqrt{\frac{a^{2p+1}}{2p+1}} \|f\|$$

ce qui prouve que l'application linéaire u est continue avec $\|u\| \leq \alpha = \sqrt{\frac{a^{2p+1}}{2p+1}}$.

Pour tout entier n tel que $a + \frac{1}{n} < 1$ on définit la fonction $f_n \in E$ par :

$$f_n(x) = \begin{cases} x^p \text{ sur } [0, a] \\ n\left(a + \frac{1}{n} - x\right) a^p \text{ sur } \left[a, a + \frac{1}{n}\right] \\ 0 \text{ sur } \left[a + \frac{1}{n}, 1\right] \end{cases}$$

et on a :

$$u(f_n)=\alpha^2,\quad \|f_n\|^2\leq \alpha^2+\frac{1}{n}$$

Avec les inégalités $\dfrac{u(f_n)}{\|f_n\|}\leq\|u\|\leq\alpha$, on déduit que $\dfrac{\alpha^2}{\sqrt{\alpha^2+\frac{1}{n}}}\leq\|u\|\leq\alpha$ et $\|u\|=\alpha$ en faisant tendre n vers l'infini.

On a donc :

$$\|u\|=\sqrt{\frac{a^{2p+1}}{2p+1}}$$

(b) Supposons qu'il existe $g\in E$ tel que $u(f)=\langle f\,|\,g\rangle$ pour tout $f\in E$. On a alors :

$$\forall f\in E,\quad \int_0^a x^p f(x)dx=\int_0^1 f(x)g(x)dx$$

Et pour $f\in E$, nulle sur $[0,a]$:

$$0=\int_0^a x^p f(x)dx=\int_0^1 f(x)g(x)dx\quad (3)$$

Pour tout entier n tel que $a+\frac{1}{n}<1$ on définit la fonction $f_n\in E$ par :

$$f_n(x)=\begin{cases}0\text{ sur }[0,a]\\ n(x-a)g\left(a+\frac{1}{n}\right)\text{ sur }\left[a,a+\frac{1}{n}\right]\\ g(\mathrm{x})\text{ sur }\left[a+\frac{1}{n},1\right]\end{cases}$$

De (3) on déduit que :

$$\int_a^{a+\frac{1}{n}} f_n(x)g(x)dx=-\int_{a+\frac{1}{n}}^1 g(x)^2\,dx$$

puis avec $\left|\int_a^{a+\frac{1}{n}} f_n(x)g(x)dx\right|\leq\frac{1}{n}\left(\underset{[0,1]}{Sup}|g(x)|\right)^2$, on déduit que :

$$\int_a^1 g(x)^2\,dx=\underset{n\to+\infty}{Lim}\int_{a+\frac{1}{n}}^1 g(x)^2\,dx=0$$

et donc que $g=0$ sur $[a,1]$.

D'autre part, pour $f(x)=x^p-g(x)$, l'égalité (3) nous donne :

$$0=\int_0^a x^p f(x)dx=\int_0^1 f(x)g(x)dx=\int_0^a f(x)g(x)dx$$

soit $\int_0^a\left(x^p-g(x)\right)f(x)dx=0$, c'est-à-dire $\int_0^a\left(x^p-g(x)\right)^2dx=0$ et $g(x)=x^p$ sur $[0,a]$, ce qui est impossible avec $g=0$ sur $[a,1]$ et g continue sur $[0,1]$.

En conclusion il n'existe pas de fonction $g\in E$ tel que $u(f)=\langle f\,|\,g\rangle$ pour tout $f\in E$.

Q8. En utilisant le théorème du point fixe de Banach (problème 44, Q1), il suffit de montrer que l'application f est contractante.

La projection P_F étant 1–lipschitzienne (Q1 (c)), on a :
$$\forall x, y \in E, \quad \|f(x) - f(y)\| \leq \|(x-y) - \lambda(v(x) - v(y))\|$$
soit :
$$\|f(x) - f(y)\|^2 \leq \|x-y\|^2 - 2\lambda\langle v(x-y) | x-y\rangle + \lambda^2 \|v(x-y)\|^2$$
Avec $\langle v(x-y) | x-y\rangle \geq \alpha\|x-y\|^2$ (coercivité de ψ) et $\|v(x-y)\| \leq \|v\|\|x-y\|$ (continuité de v), on déduit alors que :
$$\|f(x) - f(y)\|^2 \leq \left(1 - 2\alpha\lambda + \|v\|^2 \lambda^2\right)\|x-y\|^2 = \rho^2\|x-y\|^2$$
Pour $\lambda \in \left]0, \frac{2\alpha}{\|v\|^2}\right[$, on a $0 \leq \rho < 1$ et l'application f est contractante.

(b) Soit $z \in E$ tel que $u(y) = \langle z | y\rangle$ pour tout $y \in E$ (théorème de Riesz).

Pour tout $x \in E$, l'application $y \mapsto \varphi(x, y)$ est linéaire et continue (φ est continue), on peut donc trouver un vecteur $v(x) \in E$ tel que $\varphi(x, y) = \langle v(x) | y\rangle$ pour tout $y \in E$. On définit ainsi une application linéaire $v : E \to E$.

En écrivant que :
$$\|v(x)\|^2 = \langle v(x) | v(x)\rangle = \varphi(x, v(x)) \leq c\|x\|\|v(x)\|$$
on déduit que $\|v(x)\| \leq c\|x\|$ et que l'application linéaire v est continue.

En utilisant le vecteur z et l'application v, l'inégalité $\varphi(x, y-x) \geq u(y-x)$ est équivalente à :
$$\forall y \in E, \quad \langle v(x) | y-x\rangle \geq \langle z | y-x\rangle$$
ce qui équivaut encore à :
$$\forall y \in E, \quad \langle z - v(x) | y-x\rangle \leq 0$$
En utilisant (a), on peut trouver un réel $\lambda > 0$ tel que l'application :
$$x \mapsto f(x) = p_F(\lambda z + x - \lambda v(x))$$
admette un point fixe x dans F.

On a alors, en utilisant la caractérisation de la projection orthogonale établie en Q1 (b) :
$$\forall y \in F, \quad \langle \lambda z + x - \lambda v(x) - x | y-x\rangle \leq 0$$
c'est-à-dire :
$$\forall y \in F, \quad \langle z - v(x) | y-x\rangle \leq 0$$
D'où le résultat.

L'unicité de x se déduit de l'unicité du point fixe de f.

(c) La forme bilinéaire symétrique et coercive φ définit un produit scalaire sur E ($\alpha\|x\|^2 \leq \varphi(x, x) = 0 \Rightarrow x = 0$). De plus avec la continuité et la coercivité de φ, on a

$\alpha\|x\|^2 \leq \varphi(x, x) \leq c\|x\|^2$, c'est-à-dire que la norme associée à φ est équivalente à $\| \ \|$ et (E, φ) est aussi un espace de Hilbert. On peut donc utiliser le théorème de représentation de Riesz pour écrire que :
$$\forall y \in E, \quad u(y) = \varphi(z, y)$$

où $z \in E$ est uniquement déterminé.

On a alors les équivalences :

$$\left(x \in F \text{ et } \forall y \in F,\ \varphi(x, y-x) \geq u(y-x)\right)$$
$$\Leftrightarrow \left(x \in F \text{ et } \forall y \in F,\ \varphi(z-x, y-x) \leq 0\right) \Leftrightarrow \left(x = P'_F(z)\right)$$

où $P'_F(z)$ désigne la projection de x sur F pour le produit scalaire φ.

Par définition (Q1 (a)), on a :

$$\left(x = P'_F(z)\right) \Leftrightarrow \varphi(z-x, z-x) = \underset{y \in F}{Inf}\ \varphi(z-y, z-y)$$

c'est-à-dire :

$$\left(x = P'_F(z)\right) \Leftrightarrow \frac{1}{2}\varphi(x,x) - \varphi(x,z) = \underset{y \in F}{Inf}\left\{\frac{1}{2}\varphi(y,y) - \varphi(y,z)\right\}$$

Soit avec $\varphi(z, y) = u(y)$:

$$\left(x = P'_F(z)\right) \Leftrightarrow J(x) = \underset{y \in F}{Inf}\ J(y)$$

D'où le résultat.

Q9. (a) Soit $x \in E$ tel que $\varphi(x, y-x) \geq u(y-x)$, pour tout $y \in E$ (Q8 (b)). On a alors :

$$\forall y \in E, \quad \forall t \in IR, \quad \varphi(x, t \cdot y - x) \geq u(t \cdot y - x)$$

Soit, pour $y \in E$ donné :

$$\forall t \in IR, \quad \left(\varphi(x,y) - u(y)\right)t + u(x) - \varphi(x,x) \geq 0$$

ce qui équivaut à $\varphi(x, y) = u(y)$.

(b) Résulte immédiatement de Q8 (c).

Problème 48 : L'espace de Sobolev $H^1(]a,b[)$

Dans tout le problème, on désigne par :

$I =]a,b[$ un intervalle ouvert et borné de IR ;

$C^0(\bar{I})$ l'espace des fonctions continues de $\bar{I} = [a,b]$ dans IR ;

$C^\infty(I)$ l'espace des fonctions indéfiniment dérivables de I dans IR ;

$D(I)$ l'espace des fonctions indéfiniment dérivables et à support compact de I dans IR ;

$L^1(I)$ l'espace des (classes de) fonctions Lebesgue intégrables de I dans IR ;

$L^2(I)$ l'espace des (classes de) fonctions de carré intégrable de I dans IR.

On rappelle que :

On définit une norme sur $C^0(\bar{I})$ avec $f \mapsto \|f\|_\infty = \underset{t \in [a,b]}{Sup}|f(t)|$;

$L^1(I)$ muni de la norme $f \mapsto \|f\|_1 = \int_a^b |f(t)|dt$ est un espace de Banach ;

$L^2(I)$ muni du produit scalaire $(f,g) \mapsto \langle f \mid g\rangle_2 = \int_a^b f(t)g(t)dt$ est un espace de Hilbert séparable (i. e. admettant une partie dense dénombrable). On note $f \mapsto \|f\|_2 = \sqrt{\int_a^b f(t)^2 dt}$ la norme associée ;

Les espaces $C^\infty(I)$ *et* $D(I)$ *sont denses dans* $(L^1(I), \| \|_1)$ *et dans* $(L^2(I), \| \|_2)$.

Q1. *Montrer que* $L^2(I) \subset L^1(I)$ *et que :*

$$\forall f \in L^2(I), \quad \|f\|_1 \leq \sqrt{b-a}\,\|f\|_2$$

Q2. *Soit* $f \in L^2(I)$ *et* $F : \bar{I} \to IR$ *définie par :*

$$\forall x \in \bar{I}, \quad F(x) = \int_a^x f(t)dt$$

Montrer que F est uniformément continue sur $\bar{I}$ *avec* $\|F\|_\infty \leq \sqrt{b-a}\,\|f\|_2$ *et que F = 0 si et seulement si f = 0 presque partout sur I.*

Q3. *Soit* $f \in L^2(I)$.

(a) Montrer que :

$$(f = 0 \quad p.p.) \Leftrightarrow \left(\forall \varphi \in D(I), \quad \int_a^b f(t)\varphi(t)dt = 0\right)$$

(b) Montrer que :

$$(\exists c \in IR; f = c \quad p.p.) \Leftrightarrow \left(\forall \varphi \in D(I), \quad \int_a^b f(t)\varphi'(t)dt = 0\right)$$

Q4. *Soient* $f, g \in L^2(I)$ *et* $F, G : \bar{I} \to IR$ *définies par :*

$$\forall x \in \bar{I}, \quad F(x) = \int_a^x f(t)dt, \quad G(x) = \int_a^x g(t)dt$$

Montrer la « formule d'intégration par parties » :

$$\forall x \in \bar{I}, \quad \int_a^x F(t)g(t)dt = F(x)G(x) - \int_a^x f(t)G(t)dt$$

On désigne par $H^1(I)$ *l'espace des (classes de) fonctions* $f \in L^2(I)$ *pour lesquelles il existe un réel* α *et une fonction* $g \in L^2(I)$ *tels que :*

$$(1) \quad f(x) = \alpha + \int_a^x g(t)dt \quad (p.p.)$$

Q5. *Montrer que pour tout* $f \in H^1(I)$, *le couple* $(\alpha, g) \in IR \times L^2(I)$ *vérifiant (1) est unique.*

On note $g = f'$ *et on dit que g est la dérivée généralisée de f.*

Q6. *Montrer que tout élément* $f \in H^1(I)$ *admet un unique représentant continu sur* $\bar{I}$. *On identifiera f à ce représentant encore noté f.*

Q7. *Montrer que* $f \in H^1(I)$ *si et seulement si* $f \in L^2(I)$ *et il existe une unique fonction* $g \in L^2(I)$ *telle que :*

$$(2) \quad \forall \varphi \in D(I), \quad \int_a^b f(t)\varphi'(t)dt = -\int_a^b g(t)\varphi(t)dt$$

On précisera le lien entre g et la dérivée généralisée de f.

Q8. *Montrer que la fonction* $t \mapsto \frac{1}{2}(|t| + t)$ *est dans* $H^1(]-1,1[)$ *et que* $f' \notin H^1(]-1,1[)$.

Q9. *Montrer* $f \in H^1(I)$ *si et seulement si* $f \in L^2(I)$ *et il existe* $\beta > 0$ *tel que :*

$$(3) \quad \forall \varphi \in D(I), \quad \left|\int_a^b f(t)\varphi'(t)dt\right| \leq \beta\|\varphi\|_2$$

Q10. *(a) Montrer que $H^1(I)$ est un espace de Hilbert séparable pour le produit scalaire défini par :*

$$\forall(f,g)\in H^1(I)\times H^1(I),\quad \langle f\,|\,g\rangle_{H^1}=\langle f\,|\,g\rangle_2+\langle f'\,|\,g'\rangle_2$$

On note $f\mapsto\|f\|_{H^1}$ la norme associée à ce produit scalaire.

(b) Montrer qu'il existe $\gamma>0$ tel que :

$$\forall f\in H^1(I),\quad \|f\|_\infty\le\gamma\|f\|_{H^1}$$

(c) Montrer que $C^\infty(I)$ est dense dans $\left(H^1(I),\|\ \|_{H^1}\right)$.

(d) Est ce que $D(I)$ est dense dans $\left(H^1(I),\|\ \|_{H^1}\right)$?

Q11. *Montrer que :*

(a) Si $f,g\in H^1(I)$ alors $f\cdot g\in H^1(I)$ et :

$$(f\cdot g)'=f'\cdot g+f\cdot g'.$$

(b) Si $f\in H^1(I)$ et ρ est continûment dérivable sur $I\!R$, alors $\rho\circ f\in H^1(I)$ et :

$$(\rho\circ f)'=(\rho'\circ f)f'$$

On désigne par $H_0^1(I)$ l'adhérence de $D(I)$ dans $\left(H^1(I),\|\ \|_{H^1}\right)$.

Q12. *Montrer que :*

(a) L'espace $\left(H_0^1(I),\|\ \|_{H^1}\right)$ est un espace de Hilbert séparable.

(b) Si $f\in H^1(I)$ est à support compact dans I, alors $f\in H_0^1(I)$.

On pourra utiliser le résultat suivant :

Soit $(\rho_n)_{n\in IN}$ une suite régularisante définie par :

$$\rho_n\in D(IR),\quad Support(\rho_n)\subset\left[-\frac{1}{n},\frac{1}{n}\right],\quad \rho_n(x)\ge 0\ sur\ IR,\quad \int_{IR}\rho_n(t)dt=1$$

Pour toute fonction $g\in L^2(I)$, la suite de fonctions $(f_n)_{n\in IN}$ définie par :

$$\forall n\ge1,\quad \forall x\in IR,\quad f_n(x)=(\rho_n*f)(x)=\int_{IR}\rho_n(x-t)f(t)dt$$

(la fonction f étant prolongée par 0 sur $I\!R$) est dans $C^\infty(I)$ avec $(\rho_n*f)^{(k)}=\rho_n^{(k)}*f$ pour tout entier k. De plus cette suite converge vers f dans $L^2(I)$.

(c) $H_0^1(I)=\{f\in H^1(I);\quad f(a)=f(b)=0\}$.

(d) L'application $f\mapsto\|f\|_{H_0^1}=\|f'\|_2$ est une norme sur $H_0^1(I)$ équivalente à la norme $\|\ \|_{H^1}$.

Solution

Q1. L'intervalle I étant borné, la fonction constante égale à 1 est dans $L^2(I)$ et avec l'inégalité de Cauchy–Schwarz, on a pour toute fonction $f\in L^2(I)$:

$$f=f\cdot1\in L^1(I),\quad \|f\|_1=\langle|f|\,|\,1\rangle_2\le\|1\|_2\|f\|_2=\sqrt{b-a}\,\|f\|_2$$

Q2. Les inclusions $L^2(I) \subset L^1(I) \subset L^1(]a,x[)$ valables pour tout $x \in [a,b]$ entraînent que la fonction F est bien définie sur $x \in \bar{I} = [a,b]$.

Pour $x \leq y$ dans $\bar{I}$, l'inégalité de Cauchy–Schwarz donne :

$$\left|F(x) - F(y)\right| = \left|\int_x^y f(t)dt\right| \leq \left(\int_x^y 1dt\right)^{\frac{1}{2}} \left(\int_x^y \left|f(t)\right|^2 dt\right)^{\frac{1}{2}} \leq \sqrt{y-x}\left\|f\right\|_2$$

On en déduit alors que :

$$\forall (x,y) \in \bar{I}^2, \quad \left|F(x) - F(y)\right| \leq \sqrt{|y-x|}\left\|f\right\|_2$$

Ce qui prouve que F est uniformément continue sur $\bar{I}$.

En prenant $y = a$ dans l'inégalité précédente, on a :

$$\forall x \in \bar{I}, \quad \left|F(x)\right| \leq \sqrt{x-a}\left\|f\right\|_2 \leq \sqrt{b-a}\left\|f\right\|_2$$

Ce qui équivaut à :

$$\left\|F\right\|_\infty \leq \sqrt{b-a}\left\|f\right\|_2$$

Supposons que F soit identiquement nulle sur $\bar{I}$. On a alors $\int_x^y f(t)dt = 0$ pour tout $(x,y) \in \bar{I}^2$ et par linéarité de l'intégrale $\int_a^b f(t)\varphi(t)dt = 0$ pour toute fonction φ en escaliers sur $\bar{I}$.

De la densité de l'ensemble des fonctions en escaliers dans $L^2(I)$, on déduit alors que $f = 0$ presque partout.

Q3. (a) Ce résultat se déduit immédiatement de la densité de $D(I)$ dans $L^2(I)$.

(b) Soit $f \in L^2(I)$ telle que $\int_a^b f(t)\varphi'(t)dt = 0$, pour toute fonction $\varphi \in D(I)$.

On se donne une fonction $\rho \in D(I)$ telle que $\int_a^b \rho(t)dt = 1$.

Pour toute fonction $\varphi \in D(I)$, la fonction ψ définie par :

$$\forall x \in \bar{I}, \quad \psi(x) = \varphi(x) - \left(\int_a^b \varphi(t)dt\right)\rho(x)$$

est dans $D(I)$ avec $\int_a^b \psi(t)dt = 0$.

La primitive χ définie par $\chi(x) = \int_a^x \psi(t)dt$ ($x \in \bar{I}$) est alors dans $D(I)$ et :

$$0 = \int_a^b f(t)\chi'(t)dt = \int_a^b f(t)\left(\varphi(t) - \left(\int_a^b \varphi(x)dx\right)\rho(t)\right)dt$$

Ce qui peut aussi s'écrire sous la forme :

$$\int_a^b f(t)\varphi(t)dt - \left(\int_a^b \varphi(x)dx\right)\int_a^b f(t)\rho(t)dt = 0$$

Soit en posant $\alpha = \int_a^b f(t)\rho(t)dt$:

$$\forall \varphi \in D(I), \quad \int_a^b \left(f(t) - \alpha\right)\varphi(t)dt = 0$$

Ce qui équivaut, d'après Q3 (a), à $f = \alpha$ presque partout.

La réciproque est évidente.

Q4. Si on suppose que f et g sont indéfiniment dérivables, alors il en est de même de F et G avec $F' = f$, $G' = g$ et on a la formule d'intégration par parties usuelle.

Pour *f*, *g* dans $L^2(I)$, on peut écrire $f = \underset{n\to+\infty}{Lim} f_n$ et $g = \underset{n\to+\infty}{Lim} g_n$ dans $\left(L^2(I), \| \ \|_2\right)$, avec $f_n, g_n \in C^\infty(I)$ pour tout entier *n*.

On définit alors les suites $(F_n)_{n\in IN}$ et $(G_n)_{n\in IN}$ par :

$$\forall n \in IN,\ \forall x \in \bar{I},\quad F_n(x) = \int_a^x f_n(t)dt,\ G_n(x) = \int_a^x g_n(t)dt$$

En utilisant la majoration de Q2, on déduit que la suite $(F_n)_{n\in IN}$ [Resp. $(G_n)_{n\in IN}$] converge uniformément vers *F* [Resp. *G*] sur $\bar{I}$.

D'autre part, avec la majoration :

$$\left|\int_a^x F(t)g(t)dt - \int_a^x F_n(t)g_n(t)dt\right| \leq \|F - F_n\|_\infty \|g\|_1 + \|F_n\|_\infty \|g - g_n\|_1$$
$$\leq \|F - F_n\|_\infty \|g\|_1 + \sqrt{b-a}\left(\|F\|_\infty + \|F - F_n\|_\infty\right)\|g - g_n\|_2$$

on déduit que la suite $\left(\int_a^x F_n(t)g_n(t)dt\right)_{n\in IN}$ converge uniformément sur $\bar{I}$ vers $\int_a^x F(t)g(t)dt$.

De manière analogue, on voit que la suite $\left(\int_a^x f_n(t)G_n(t)dt\right)_{n\in IN}$ converge uniformément sur $\bar{I}$ vers $\int_a^x f(t)G(t)dt$.

Ce qui donne, en définitive :

$$\int_a^x F(t)g(t)dt = \underset{n\to+\infty}{Lim}\int_a^x F_n(t)g_n(t)dt$$
$$= \underset{n\to+\infty}{Lim}\left\{F_n(x)G_n(x) - \int_a^x f_n(t)G_n(t)dt\right\} = F(x)G(x) - \int_a^x f(t)G(t)dt$$

C'est-à-dire la formule d'intégration par parties.

Q5. Pour montrer l'unicité du couple $(\alpha, g) \in IR \times L^2(I)$ vérifiant (1), il suffit de considérer le cas où *f* = 0.

Dans ce cas, l'égalité $\alpha + \int_a^x g(t)dt = 0$ pour tout *x* (la fonction est continue) entraîne $\alpha = 0$ (il suffit de prendre *x* = *a*) et $\int_a^x g(t)dt = 0$ pour tous *x* dans $\bar{I}$, ce qui équivaut, d'après Q2, à *g* = 0 presque partout.

Q6. Pour toute fonction $f \in H^1(I)$, la fonction $\tilde{f} : x \mapsto \alpha + \int_a^x f'(t)dt$ est définie, continue sur $\bar{I}$, avec $\tilde{f}(a) = \alpha$ et $\tilde{f} = f$ presque partout.

L'unicité d'un tel représentant continu est évidente.

Q7. Soit $f \in H^1(I)$ (continue) de dérivée généralisée $f' \in L^2(I)$.

Pour toute fonction $\varphi \in D(I)$, le théorème d'intégration par parties (Q4) appliqué à f' et φ' donne :

$$\int_a^b (f(t) - \alpha)\varphi'(t)dt = (f(b) - \alpha)\varphi(b) - \int_a^b f'(t)(\varphi(t) - \varphi(a))dt$$

avec $\varphi(a) = \varphi(b) = 0$ et $\int_a^b \varphi'(t)dt = \varphi(b) - \varphi(a) = 0$, donc :

$$\int_a^b f(t)\varphi'(t)dt = -\int_a^b f'(t)\varphi(t)dt$$

Ce qui prouve que la fonction f' vérifie (2).

Si $g \in L^2(I)$ est une autre fonction vérifiant (2), alors :

$$\forall \varphi \in D(I), \quad \int_a^b \left(g(t) - f'(t)\right)\varphi(t)dt = 0$$

ce qui équivaut à $g = f'$ d'après Q3 (a).

Réciproquement soit $f \in L^2(I)$ telle qu'il existe $g \in L^2(I)$ vérifiant (2).

On définit la fonction $h \in C^0(\bar{I})$ par :

$$\forall x \in \bar{I}, \quad h(x) = \int_a^x g(t)dt$$

En utilisant le théorème d'intégration par parties (Q4), on peut écrire pour toute fonction $\varphi \in D(I)$:

$$\int_a^b h(t)\varphi'(t)dt = -\int_a^b g(t)\varphi(t)dt = \int_a^b f(t)\varphi'(t)dt$$

C'est-à-dire que :

$$\forall \varphi \in D(I), \quad \int_a^b \left(h(t) - f(t)\right)\varphi'(t)dt = 0$$

Ce qui équivaut à dire que $h - f$ est égale à une constante presque partout, c'est-à-dire qu'il existe $\alpha \in IR$ telle que $f(x) = \alpha + h(x) = \alpha + \int_a^x g(t)dt$ presque partout. Ce qui prouve que $f \in H^1(I)$ avec $f' = g$.

Q8. La fonction *h* définie par :

$$\forall x \in [-1,1], \quad h(x) = \begin{cases} 1 \text{ si } 0 \le x \le 1 \\ 0 \text{ si } -1 \le x < 0 \end{cases}$$

est dans $L^2([-1,1])$ avec :

$$\int_{-1}^x h(t)dt = \begin{cases} x \text{ si } 0 \le x \le 1 \\ 0 \text{ si } -1 \le x < 0 \end{cases} = \frac{1}{2}\left(|x| + x\right)$$

On a donc $f \in H^1([-1,1])$, avec $f' = h$.

Si $h \in H^1([-1,1])$, alors pour toute fonction $\varphi \in D(]-1,1[)$, on a :

$$\int_{-1}^1 h'(t)\varphi(t)dt = -\int_{-1}^1 h(t)\varphi'(t)dt = -\int_0^1 \varphi'(t)dt = \varphi(0)$$

En prenant $\varphi \in D(]-1,0[)$ [Resp. $\varphi \in D(]0,1[)$], on déduit que $h' = 0$ presque partout sur $]-1,0[$ [Resp. sur $]0,1[$], donc $h' = 0$ presque partout sur $]-1,1[$ et $\varphi(0) = 0$ pour toute fonction $\varphi \in D(]-1,1[)$, ce qui est faux.

On peut donc conclure que $h \notin H^1([-1,1])$.

Q9. Si $f \in H^1(I)$, alors $f \in L^2(I)$ et $\int_a^b f(t)\varphi'(t)dt = -\int_a^b f'(t)\varphi(t)dt$, pour tout $\varphi \in D(I)$. L'inégalité de Cauchy–Schwarz donne alors :

$$\left|\int_a^b f(t)\varphi'(t)dt\right| \le \|f'\|_2 \|\varphi\|_2$$

Réciproquement, supposons que $f \in L^2(I)$ vérifie (3) pour une constante $\beta > 0$. Dans ces conditions la forme linéaire :

$$\begin{array}{llll} u: & D(I) & \to & IR \\ & \varphi & \mapsto & u(\varphi) = \int_a^b f(t)\varphi'(t)dt \end{array}$$

est continue. Avec la densité de $D(I)$ dans $L^2(I)$, on peut la prolonger sur $L^2(I)$ en posant pour tout $h = \underset{n\to+\infty}{Lim}\, \varphi_n \in L^2(I)$ ($\varphi_n \in D(I)$) :

$$u(h) = \underset{n\to+\infty}{Lim}\, u(\varphi_n)$$

Cette définition ne dépend pas du choix d'une suite $(\varphi_n)_{n\in IN}$ d'éléments de $D(I)$ qui converge vers h. En effet si $(\psi_n)_{n\in IN}$ est une autre suite avec les mêmes propriétés, alors :

$$|u(\varphi_n) - u(\psi_n)| \le \beta \|\varphi_n - \psi_n\|_2 \underset{n\to+\infty}{\to} 0$$

L'application u ainsi définie sur $L^2(I)$ est linéaire et :

$$\forall h = \underset{n\to+\infty}{Lim}\, \varphi_n \in L^2(I), \quad |u(h)| = \underset{n\to+\infty}{Lim} |u(\varphi_n)| \le \beta \underset{n\to+\infty}{Lim} \|\varphi_n\|_2 = \beta \|h\|_2$$

C'est-à-dire que u est linéaire et continue sur l'espace de Hilbert $L^2(I)$.

Le théorème de Riesz nous dit alors qu'il existe $g \in L^2(I)$ telle que :

$$\forall h \in L^2(I), \quad u(h) = \langle h | g \rangle_2$$

Ce qui donne, en particulier :

$$\forall \varphi \in D(I), \quad \int_a^b f(t)\varphi'(t)dt = \int_a^b g(t)\varphi(t)dt$$

C'est-à-dire que $f \in H^1(I)$, avec $f' = -g$.

Q10. (a) De l'unicité de la dérivée généralisée et de la linéarité de l'intégrale, on déduit que $H^1(I)$ est un sous-espace vectoriel de $L^2(I)$ et que $(\alpha f + \beta g)' = \alpha f' + \beta g'$ pour tous $\alpha, \beta \in IR$ et $f, g \in H^1(I)$.

Il est clair que l'application $(f,g) \mapsto \langle f | g \rangle_{H^1} = \langle f | g \rangle_2 + \langle f' | g' \rangle_2$ définit un produit scalaire sur $H^1(I)$.

Si $(f_n)_{n\in IN}$ est une suite de Cauchy dans $(H^1(I), \| \ \|_{H^1})$, alors les suites $(f_n)_{n\in IN}$ et $\left(f_n'\right)_{n\in IN}$ sont de Cauchy dans $(L^2(I), \| \ \|_2)$ et elles convergent respectivement vers f et g dans $(L^2(I), \| \ \|_2)$.

De l'inégalité de Cauchy–Schwarz, on déduit que pour tout $\varphi \in D(I)$, on a :

$$\int_a^b f(t)\varphi'(t)dt = \underset{n\to+\infty}{Lim} \int_a^b f_n(t)\varphi'(t)dt, \quad \int_a^b g(t)\varphi(t)dt = \underset{n\to+\infty}{Lim} \int_a^b f_n'(t)\varphi(t)dt$$

Et avec $\int_a^b f_n(t)\varphi'(t)dt = -\int_a^b f_n'(t)\varphi(t)dt$, on déduit que :

$$\int_a^b f(t)\varphi'(t)dt = -\int_a^b g(t)\varphi(t)dt$$

C'est-à-dire que $f \in H^1(I)$, avec $f' = g$.

Enfin avec $\|f - f_n\|_{H^1}^{\ 2} = \|f - f_n\|_2^{\ 2} + \left\|f' - f_n'\right\|_2^{\ 2} = \|f - f_n\|_2^{\ 2} + \left\|g - f_n'\right\|_2^{\ 2}$, on déduit que $f = \underset{n\to+\infty}{Lim}\, f_n$ dans $(H^1(I), \| \ \|_{H^1})$.

On peut donc conclure que $H^1(I)$ est muni d'une structure d'espace de Hilbert.

L'espace de Hilbert $L^2(I)$ étant séparable, il en est de même de $L^2(I)\times L^2(I)$ pour la structure hilbertienne produit.

L'application

$$\chi : \begin{array}{ccc} H^1(I) & \to & L^2(I)\times L^2(I) \\ f & \mapsto & (f, f') \end{array}$$

est une isométrie qui identifie $H^1(I)$ à un sous espace fermé de $L^2(I)\times L^2(I)$. On en déduit alors que $H^1(I)$ est séparable (tout sous espace d'un espace de Hilbert séparable est séparable).

(b) Une fonction $f \in H^1(I)$ peut être définie par $f : x \mapsto f(a) + \int_a^x f'(t)dt$ (voir Q6).

Pour *x*, *y* dans $\bar{I}$, on a alors $f(x) = f(y) + \int_y^x f'(t)dt$ et en intégrant par rapport à *y* sur *I*, on a :

$$(b-a)f(x) = \int_a^b f(y)dy + \int_a^b \left(\int_y^x f'(t)dt\right)dy$$

En utilisant l'inégalité de Cauchy–Schwarz, on déduit que :

$$(b-a)|f(x)| \le \int_a^b |f(y)|dy + \int_a^b \left|\int_y^x f'(t)dt\right|dy \le \sqrt{b-a}\|f\|_2 + (b-a)\sqrt{b-a}\|f'\|_2$$

et :

$$\|f\|_\infty \le \left(\frac{1}{\sqrt{b-a}} + \sqrt{b-a}\right)\|f\|_{H^1}$$

(c) Soit $f \in H^1(I)$. Avec la densité de $C^\infty(I)$ dans $L^2(I)$, on peut trouver une suite $(g_n)_{n\in IN}$ dans $C^\infty(I)$ qui converge vers f' dans $\left(L^2(I), \| \ \|_2\right)$.

On définit alors la suite $(f_n)_{n\in IN}$ par :

$$\forall n \in IN,\ \forall x \in \bar{I},\quad f_n(x) = f(a) + \int_a^x g_n(t)dt$$

et on a $f_n \in C^\infty(I)$, avec $f_n' = g_n$, $f_n(a) = f(a)$, pour tout entier *n*.

L'inégalité de Cauchy–Schwarz donne, pour tout $x \in \bar{I}$ et tout entier *n* :

$$|f(x) - f_n(x)| = \left|\int_a^x (f'(t) - g_n(t))dt\right| \le \sqrt{b-a}\|f' - g_n\|_2$$

c'est-à-dire $\|f - f_n\|_\infty \le \sqrt{b-a}\|f' - g_n\|_2$ et la suite $(f_n)_{n\in IN}$ converge uniformément vers *f* sur $\bar{I}$. Elle converge donc aussi vers *f* dans $\left(L^2(I), \| \ \|_2\right)$.

Enfin, avec $\|f - f_n\|_{H^1}^2 = \|f - f_n\|_2^2 + \left\|f' - f_n'\right\|_2^2 = \|f - f_n\|_2^2 + \|f' - g_n\|_2^2$, on déduit que $f = \underset{n\to+\infty}{Lim} f_n$ dans $\left(H^1(I), \| \ \|_{H^1}\right)$.

Ce qui prouve que $C^\infty(I)$ est dense dans $\left(H^1(I), \| \ \|_{H^1}\right)$.

(d) Si $f \in H^1(I)$ est limite d'une suite $(\varphi_n)_{n\in IN}$ de $D(I)$, alors cette suite converge uniformément vers *f* sur $\bar{I}$ (d'après l'inégalité de (b)) et en particulier, on aura $f(a) = \underset{n\to+\infty}{Lim}\, \varphi_n(a) = 0$ et $f(b) = \underset{n\to+\infty}{Lim}\, \varphi_n(b) = 0$.

Comme $H^1(I)$ contient des fonctions qui ne s'annulent pas en *a* (ou en *b*), on déduit que $D(I)$ n'est pas dense dans $H^1(I)$.

Q11. (a) Si *f*, *g* sont dans $H^1(I)$, alors elles sont égales presque partout à des fonctions continues sur $\bar{I}$. Leur produit est donc continu presque partout et définit un élément de $L^2(I)$.

Soit $(\varphi_n)_{n\in IN}$ [Resp. $(\psi_n)_{n\in IN}$] une suite de $D(I)$ qui converge vers *f* [Resp. *g*] dans $\left(H^1(I), \|\ \|_{H^1}\right)$.

Avec l'inégalité de Q10 (b), on déduit que $(\varphi_n)_{n\in IN}$ [Resp. $(\psi_n)_{n\in IN}$] converge uniformément vers *f* [Resp. *g*] sur $\bar{I}$ et donc la suite $(\varphi_n\psi_n)_{n\in IN}$ converge uniformément et dans $L^2(I)$ vers le produit $f \cdot g$.

On a alors $\underset{n\to+\infty}{Lim}(\varphi_n\psi_n)' = \underset{n\to+\infty}{Lim}\left(\varphi_n'\psi_n + \varphi_n\psi_n'\right) = f'g + fg'$ dans $L^2(I)$ ($\left(\varphi_n'\right)_{n\in IN}$ converge dans $L^2(I)$ vers f' et $\left(\psi_n'\right)_{n\in IN}$ vers g') et pour tout $\varphi \in D(I)$:

$$\int_a^b \left(f'(t)g(t) + f(t)g'(t)\right)\varphi(t)dt = \underset{n\to+\infty}{Lim}\int_a^b \left(\varphi_n(t)\psi_n(t)\right)'\varphi(t)dt$$

$$= -\underset{n\to+\infty}{Lim}\int_a^b \varphi_n(t)\psi_n(t)\varphi'(t)dt = -\int_a^b f(t)g(t)\varphi'(t)dt$$

Ce qui prouve que $f \cdot g \in H^1(I)$, avec $(f \cdot g)' = f'g + fg'$.

(b) Les fonctions $\rho \circ f$ et $\rho' \circ f$ sont continues sur $\bar{I}$, donc $\rho \circ f$ et $(\rho' \circ f)f'$ sont dans $L^2(I)$ (*I* est borné).

Avec la densité de $C^\infty(I)$ dans $\left(H^1(I), \|\ \|_{H^1}\right)$, on peut trouver une suite $(f_n)_{n\in IN}$ de $C^\infty(I)$ qui converge vers *f*. Cette suite converge également de manière uniforme vers *f* sur $\bar{I}$ (Q10 (b)). Avec la continuité de ρ et ρ', on déduit que les suites $(\rho \circ f_n)_{n\in IN}$ et $(\rho' \circ f_n)_{n\in IN}$ convergent uniformément sur $\bar{I}$ vers $\rho \circ f$ et $\rho' \circ f$ respectivement et que la suite $\left((\rho' \circ f_n)f_n'\right)_{n\in IN}$ converge dans $L^2(I)$ vers $(\rho' \circ f)f'$.

Pour tout $x \in I$, la suite $\left((\rho' \circ f_n)f_n'\right)_{n\in IN}$ converge alors dans $\left(L^1(]a,x[), \|\ \|_1\right)$ vers $(\rho' \circ f)f'$ (d'après Q1) et avec :

$$(\rho \circ f_n)(x) = (\rho \circ f_n)(a) + \int_a^x (\rho' \circ f_n)(t)f_n'(t)dt$$

on déduit, en faisant tendre *n* vers l'infini, que :

$$(\rho \circ f)(x) = (\rho \circ f)(a) + \int_a^x (\rho' \circ f)(t)f'(t)dt$$

C'est-à-dire que $\rho\circ f\in H^1(I)$, avec $(\rho\circ f)'=(\rho'\circ f)f'$.

Q12. (a) Le sous-espace vectoriel $H_0^1(I)$ est la fermeture de $D(I)$ dans $H^1(I)$, c'est donc un fermé et il hérite de la structure d'espace de Hilbert de $H^1(I)$.

(b) Soit $f\in H^1(I)$ à support compact dans $[\alpha,\beta]\subset]a,b[$.

On se donne une suite régularisante $(\rho_n)_{n\in IN}$ et on désigne par $(f_n)_{n\in IN}$ la suite définie par $f_n=\rho_n * f$ pour tout entier n.

Si $\delta=Min\{\alpha-a,b-\beta\}$, alors pour $x\in\left[a,a+\dfrac{\delta}{2}\right]\cup\left[b-\dfrac{\delta}{2},b\right]$ et $t\in[\alpha,\beta]$, on a $|x-t|\geq\dfrac{\delta}{2}$, de sorte qu'il existe $n_0\in IN$ tel que $\rho_n(x-t)=0$ pour tout $n\geq n_0$ et donc $f_n(x)=\int_\alpha^\beta\rho_n(x-t)f(t)dt=0$. Pour $n\geq n_0$, la fonction f_n est donc dans $D(I)$.

D'autre part, pour toute fonction $\varphi\in D(I)$, le théorème de Fubini nous permet d'écrire que :

$$\int_a^b(\rho_n * f)(x)\varphi'(x)dx=\int_a^b\left(\int_a^b\rho_n(x-t)f(t)dt\right)\varphi'(x)dx$$
$$=\int_a^b f(t)\left(\int_a^b\rho_n(x-t)\varphi'(x)dx\right)dt=\int_a^b f(t)(\tilde{\rho}_n*\varphi')(t)dt$$

où $\tilde{\rho}_n$ est définie par $\tilde{\rho}_n(y)=\rho_n(-y)$.

En considérant que $\tilde{\rho}_n*\varphi'=(\tilde{\rho}_n*\varphi)'$, on déduit, en utilisant encore le théorème Fubini, que :

$$\int_a^b(\rho_n * f)(x)\varphi'(x)dx=\int_a^b f(t)(\tilde{\rho}_n*\varphi)'(t)dt=-\int_a^b f'(t)(\tilde{\rho}_n*\varphi)(t)dt$$
$$=-\int_a^b(\rho_n * f')(x)\varphi(x)dx$$

Ce qui prouve que $(\rho_n * f)'=\rho_n * f'$.

On peut alors conclure, en utilisant le rappel de cette question, que $\underset{n\to+\infty}{Lim}\rho_n * f=f$ et $\underset{n\to+\infty}{Lim}(\rho_n * f)'=f'$ dans $L^2(I)$, ce qui entraîne que $\underset{n\to+\infty}{Lim}f_n=f$ dans $\left(H^1(I),\|\ \|_{H^1}\right)$ et que $f\in H_0^1(I)$.

(c) Si $f\in H_0^1(I)$, alors $f=\underset{n\to+\infty}{Lim}\varphi_n$ dans $\left(H^1(I),\|\ \|_{H^1}\right)$ avec les φ_n dans $D(I)$. De Q10 (b), on déduit que la suite $(\varphi_n)_{n\in IN}$ converge uniformément vers f sur $\bar{I}$ et en particulier $f(a)=\underset{n\to+\infty}{Lim}\varphi_n(a)=0$ et $f(b)=\underset{n\to+\infty}{Lim}\varphi_n(b)=0$.

Réciproquement soit $f\in H^1(I)$ nulle en a et en b.

Pour tout $n\geq 1$, on désigne par θ_n la fonction affine par morceaux définie par la figure 8.4 et par f_n la fonction définie par $f_n=\theta_n f$.

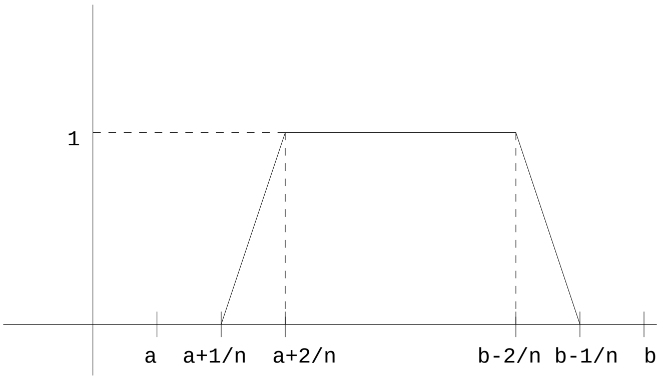

Figure 8.4

La fonction θ_n est dans $H^1(I)$, donc $f_n \in H^1(I)$ (d'après Q11 (a)). De plus f_n est à support compact dans $\left[a+\frac{1}{n}, b-\frac{1}{n}\right] \subset \left]a,b\right[$, donc $f_n \in H_0^1(I)$ (d'après (b)).

Nous allons montrer que $\underset{n\to+\infty}{Lim} f_n = f$ dans $H^1(I)$, ce qui prouvera que $f \in H_0^1(I)$, puisque $H_0^1(I)$ est fermé dans $H^1(I)$.

Avec la continuité de f et les conditions $f(a) = f(b) = 0$, on déduit que pour tout réel $\varepsilon > 0$, on peut trouver un réel $\eta > 0$ tel que :

$$\forall x \in [a, a+\eta] \cup [b-\eta, b], \quad |f(x)| < \varepsilon$$

Pour $n_0 \in IN$ tel que $\eta \geq \frac{2}{n_0}$, on a alors :

$$\forall n \geq n_0, \quad \|f_n - f\|_2^{\ 2} = \int_{I_n} |f_n(t) - f(t)|^2 dt \leq 4\int_{I_n} |f(t)|^2 dt \leq 4(b-a)\varepsilon^2$$

où on a noté $I_n = \left[a+\frac{1}{n}, a+\frac{2}{n}\right] \cup \left[b-\frac{2}{n}, b-\frac{1}{n}\right]$.

Ce qui prouve que $\underset{n\to+\infty}{Lim} f_n = f$ dans $L^2(I)$.

Pour montrer que $\underset{n\to+\infty}{Lim} f_n = f$ dans $H^1(I)$, il reste à montrer que $\underset{n\to+\infty}{Lim} f_n^{\prime} = f'$.

D'après Q11 (a), on a, pour tout entier n, $f_n^{\prime} = \theta_n f' + \theta_n^{\prime} f$.

Montrons tout d'abord que $\underset{n\to+\infty}{Lim} \theta_n^{\prime} f = 0$ dans $L^2(I)$.

On a $\left\|\theta_n^{\prime} f\right\|_2^{\ 2} = \int_{I_n} n^2 |f(t)|^2 dt$. Et en écrivant que $f(t) = \int_a^t f'(x)dx$ [Resp. $f(t) = -\int_t^b f'(x)dx$] sur $\left[a+\frac{1}{n}, a+\frac{2}{n}\right]$ [Resp. $\left[b-\frac{2}{n}, b-\frac{1}{n}\right]$], du fait que $f(a) = 0$ [Resp. $f(b) = 0$], on déduit en utilisant l'inégalité de Cauchy–Schwarz que :

$$\left\|\theta_n' f\right\|_2^2 = n^2 \int_{a+\frac{1}{n}}^{a+\frac{2}{n}} \left|\int_a^t f'(x)dx\right|^2 dt + n^2 \int_{b-\frac{2}{n}}^{b-\frac{1}{n}} \left|\int_t^b f'(x)dx\right|^2 dt$$

$$\leq n^2 \int_{a+\frac{1}{n}}^{a+\frac{2}{n}} (t-a) \int_a^{a+\frac{2}{n}} \left|f'(x)\right|^2 dxdt + n^2 \int_{b-\frac{2}{n}}^{b-\frac{1}{n}} (b-t) \int_{b-\frac{2}{n}}^{b} \left|f'(x)\right|^2 dxdt$$

Soit :

$$\left\|\theta_n' f\right\|_2^2 \leq 2\int_a^{a+\frac{2}{n}} \left|f'(x)\right|^2 dx + 2\int_{b-\frac{2}{n}}^{b} \left|f'(x)\right|^2 dx \underset{n\to+\infty}{\longrightarrow} 0$$

Enfin avec :

$$\left\|\theta_n f' - f'\right\|_2^2 = \int_{\left[a,a+\frac{1}{n}\right]\cup\left[b-\frac{1}{n},b\right]} \left|f'(x)\right|^2 dx + \int_{\left[a+\frac{1}{n},a+\frac{2}{n}\right]\cup\left[b-\frac{2}{n},b-\frac{1}{n}\right]} \left|(\theta_n f')(x) - f'(x)\right|^2 dx$$

$$\leq \int_{\left[a,a+\frac{1}{n}\right]\cup\left[b-\frac{1}{n},b\right]} \left|f'(x)\right|^2 dx + 4\int_{\left[a+\frac{1}{n},a+\frac{2}{n}\right]\cup\left[b-\frac{2}{n},b-\frac{1}{n}\right]} \left|f'(x)\right|^2 dx \underset{n\to+\infty}{\longrightarrow} 0$$

on déduit que $\underset{n\to+\infty}{Lim}\, \theta_n f' = f'$ dans $L^2(I)$. D'où le résultat.

(d) Si $f \in H_0^1(I)$ est telle que $\|f\|_{H_0^1} = 0$, alors $f' = 0$ presque partout et la relation $f(x) = \int_a^x f'(t)dt$ donne $f = 0$. Donc $\|\ \|_{H_0^1}$ est bien une norme sur $H_0^1(I)$.

On a de manière évidente l'inégalité :

$$\forall f \in H_0^1(I), \quad \|f\|_{H_0^1} \leq \|f\|_{H^1}$$

Avec $f(x) = \int_a^x f'(t)dt$ et l'inégalité de Cauchy–Schwarz, on déduit que $\left|f(x)\right|^2 \leq (x-a)\|f'\|_2^2$ et $\|f\|_2^2 \leq \dfrac{(b-a)^2}{2}\|f'\|_2^2$, donc :

$$\|f\|_{H^1} \leq \sqrt{\frac{(b-a)^2}{2} + 1}\,\|f'\|_2$$

Ce qui achève de prouver l'équivalence des deux normes.

Remarque — La dernière inégalité est l'inégalité de Poincaré.

Problème 49 : Problème de Dirichlet dans $H_0^1(]-1,1[)$. Agrégation 1994, extrait

On désigne par Λ l'intervalle ouvert $]-1,1[$. On note $L^2(\Lambda)$ l'espace des (classes de) fonctions à valeurs réelles de carré intégrable sur Λ. On le munit du produit scalaire

$$(u,v) = \int_{-1}^{1} u(x)v(x)dx$$

et de la norme $\|\cdot\|_{L^2(\Lambda)}$ associée à ce produit scalaire. On note $C^\infty(\Lambda)$ l'espace des fonctions indéfiniment dérivables sur Λ, on admettra que $C^\infty(\Lambda)$ est dense dans $L^2(\Lambda)$.

On considère le problème de Dirichlet homogène, en dimension 1, qui consiste à chercher une fonction u continue sur $\overline{\Lambda}$ vérifiant

$$-u'' + a \cdot u = f \text{ presque partout dans } \Lambda \qquad (1)$$

avec les conditions aux limites

$$u(-1) = u(1) = 0 \qquad (2)$$

lorsque a est une fonction réelle continue ≥ 0 sur $\overline{\Lambda}$ et f une fonction donnée dans $L^2(\Lambda)$.

Q1. *Montrer que toute fonction v de $L^2(\Lambda)$ est intégrable sur Λ et vérifie*

$$\int_{-1}^{1} |v(x)| dx \leq \sqrt{2} \|v\|_{L^2(\Lambda)}$$

Q2. *On désigne par $H^1(\Lambda)$ l'espace des (classes de) fonctions v de $L^2(\Lambda)$ pour lesquelles il existe un réel μ et une fonction w de $L^2(\Lambda)$ tels que :*

$$v(x) = \mu + \int_{-1}^{x} w(t) dt \text{ pour presque tout } x \in \Lambda$$

Montrer que le couple (w, μ) associé à la fonction v est unique.

La fonction w est alors appelée dérivée première (au sens généralisé) de v et notée v'.

On munit l'espace $H^1(\Lambda)$ de la norme

$$\|v\|_{H^1(\Lambda)} = \left(\|v\|_{L^2(\Lambda)}{}^2 + \|v'\|_{L^2(\Lambda)}{}^2 \right)^{\frac{1}{2}}$$

Q3. *Montrer que l'espace $H^1(\Lambda)$ est un espace de Hilbert. Vérifier également que l'espace $C^\infty(\Lambda)$ est dense dans $H^1(\Lambda)$. A partir de la formule*

$$v(x) = v(y) + \int_{y}^{x} v'(t) dt$$

prouver que toute fonction v de $H^1(\Lambda)$ admet un représentant continu sur $\overline{\Lambda}$ vérifiant

$$\underset{-1 \leq x \leq x}{Sup} |v(x)| \leq 2 \|v\|_{H^1(\Lambda)}$$

Q4. *On note $H_0^1(\Lambda)$ l'adhérence dans $H^1(\Lambda)$ de l'espace des fonctions de $C^\infty(\Lambda)$ à support compact dans Λ. Montrer que toute fonction de $H_0^1(\Lambda)$ s'annule aux deux extrémités ± 1 de Λ. Vérifier que l'espace $H_0^1(\Lambda)$ est un espace de Hilbert. Montrer que la semi–norme*

$$|v|_{H^1(\Lambda)} = \|v'\|_{L^2(\Lambda)}$$

est une norme sur l'espace $H_0^1(\Lambda)$, équivalente à la norme $\|\cdot\|_{H^1(\Lambda)}$.

Q5. *On définit la forme bilinéaire $\alpha(\cdot,\cdot)$ sur $H^1(\Lambda) \times H^1(\Lambda)$ par la formule*

$$\alpha(u, v) = \int_{-1}^{1} u'(x) v'(x) dx + \int_{-1}^{1} a(x) u(x) v(x) dx$$

Montrer que la forme $\alpha(\cdot,\cdot)$ est continue sur $H^1(\Lambda) \times H^1(\Lambda)$ et que, pour toute fonction v de $H^1(\Lambda)$,

$$\alpha(v, v) \geq |v|_{H^1(\Lambda)}{}^2$$

Q6. *On admettra que* $H_0^1(\Lambda)$ *est exactement l'espace des fonctions de* $H^1(\Lambda)$ *qui s'annulent en* ± 1. *Montrer que toute fonction u deux fois continûment dérivable sur* $\overline{\Lambda}$ *est solution du problème (1) (2) si et seulement si elle est solution du problème : trouver u dans* $H_0^1(\Lambda)$ *tel que*

$$\forall v \in H_0^1(\Lambda), \quad \alpha(u,v) = \int_{-1}^{1} f(x)v(x)dx \quad (4)$$

On considérera désormais uniquement le problème (4).

Q7. *On rappelle le théorème de Riesz (problème 47 Q5) : étant donné un espace de Hilbert H pour le produit scalaire* $(\cdot,\cdot)_H$, *pour toute forme linéaire l continue sur H, il existe un élément u de H tel que*

$$\forall v \in H, \quad l(v) = (u,v)_H$$

Montrer en appliquant ce théorème que le problème (4) a une unique solution u dans $H_0^1(\Lambda)$.

Solution

Q1. Voir le problème 48, Q1.
Q2. Voir le problème 48, Q5.
Q3. Voir le problème 48, Q10 et Q6.
Q4. Voir le problème 48, Q12.
Q5. Avec la continuité de la fonction *a* et l'inégalité de Cauchy–Schwarz, on a pour tous *u*, *v* dans $H^1(\Lambda)$:

$$|\alpha(u,v)| \le \|u'\|_{L^2(\Lambda)}\|v'\|_{L^2(\Lambda)} + \|a\|_\infty \|u\|_{L^2(\Lambda)}\|v\|_{L^2(\Lambda)}$$

où on a noté $\|a\|_\infty = \underset{\overline{\Lambda}}{Sup}|a(t)|$.

On déduit alors que :

$$\forall (u,v) \in H^1(\Lambda)\times H^1(\Lambda), \quad |\alpha(u,v)| \le (1+\|a\|_\infty)\|u\|_{H^1(\Lambda)}\|v\|_{H^1(\Lambda)}$$

ce qui prouve la continuité de la forme bilinéaire α.

Pour toute fonction *v* de $H^1(\Lambda)$, on a :

$$|\alpha(v,v)| = \|v'\|_{L^2(\Lambda)}{}^2 + \int_{-1}^{1} a(x)v(x)^2 dx \ge \|v'\|_{L^2(\Lambda)}{}^2 = |v|_{H^1(\Lambda)}{}^2$$

du fait de la positivité de la fonction *a* sur $\overline{\Lambda}$.

Q6. Soit $u \in C^2(\overline{\Lambda})$ solution au sens classique du problème de Dirichlet homogène (1) et (2).

La condition (2) et la dérivabilité de *u* entraînent que $u \in H_0^1(\Lambda)$.

En multipliant les deux membres de l'égalité (1) par une fonction φ indéfiniment dérivable et à support compact dans Λ, puis en effectuant une intégration par parties, on obtient :

$$\int_{-1}^{1} (u'(x)\varphi'(x) + a(x)u(x)\varphi(x))dx = \int_{-1}^{1} f(x)\varphi(x)dx$$

On a donc $\alpha(u,\varphi)=\beta(\varphi)=\int_{-1}^{1} f(x)\varphi(x)dx$ pour toute fonction φ indéfiniment dérivable et à support compact dans Λ.

Avec la densité, dans $H_0^1(\Lambda)$, de l'ensemble des fonctions indéfiniment dérivables à support compact dans Λ et la continuité des applications α et β (l'inégalité de Cauchy–Schwarz nous dit que la forme linéaire β est bien définie et continue sur $L^2(\Lambda)$ et donc qu'elle est continue sur $H^1(\Lambda)$), on déduit que u est bien solution de (4).

Réciproquement, soit $u \in H_0^1(\Lambda) \cap C^2(\overline{\Lambda})$ solution de (4). La fonction u étant dans $H_0^1(\Lambda)$ s'annule en -1 et en 1. Donc la condition (2) est vérifiée.

En appliquant (4) à une fonction φ indéfiniment dérivable et à support compact dans Λ, puis en effectuant une intégration par parties (la fonction u est de classe C^2), on déduit que $\int_{-1}^{1}(-u''(x)+a(x)u(x)-f(x))\varphi(x)dx=0$.

Il en résulte alors que $-u''(x)+a(x)u(x)-f(x)=0$ presque partout sur Λ (problème 48, Q3 (a)).

Q7. L'inégalité établie en Q5 (coercivité de α sur $\left(H_0^1(\Lambda),|\ |_{H^1(\Lambda)}\right)$) entraîne que la forme bilinéaire symétrique α définit un produit scalaire sur $H_0^1(\Lambda)$. De plus, avec la continuité (et la coercivité) de α, on déduit que la norme associée à ce produit scalaire est équivalente à $|\ |_{H^1(\Lambda)}$ et à $\|\cdot\|_{H^1(\Lambda)}$. Il en résulte alors que $\left(H_0^1(\Lambda),\alpha\right)$ est un espace de Hilbert.

D'autre part, pour toute fonction $f \in L^2(\Lambda)$, la forme linéaire $\beta : v \mapsto \int_{-1}^{1} f(x)v(x)dx$ est continue sur $H_0^1(\Lambda)$ muni de $\|\cdot\|_{H^1(\Lambda)}$, elle est dons aussi continue sur $\left(H_0^1(\Lambda),\alpha\right)$.

On peut donc utiliser le théorème de Riesz pour conclure à l'existence d'une fonction $u \in H_0^1(\Lambda)$ telle que $\alpha(u,v)=\beta(v)$ pour toute fonction $v \in H_0^1(\Lambda)$.

Pour montrer l'unicité d'une telle solution, il suffit de considérer le cas où $f = 0$ dans $L^2(\Lambda)$.

Dans ce cas, on a $\int_{-1}^{1} u'(x)^2 dx + \int_{-1}^{1} a(x)u(x)^2 dx = 0$ et $|u|_{H^1(\Lambda)}{}^2 = \int_{-1}^{1} u'(x)^2 dx = 0$ (la fonction a est positive). On en déduit alors que $u = 0$ dans $H_0^1(\Lambda)$ ($|\ |_{H^1(\Lambda)}$ est une norme sur $H_0^1(\Lambda)$).

On a donc ainsi prouvé que le problème (4) admet une unique solution dans $H_0^1(\Lambda)$.

Problème 50 : Discrétisation par des polynômes du problème de Dirichlet dans $H_0^1(]-1,1[)$. Agrégation 1994, extrait

Ce problème utilise les notations et résultats du problème 49.

Pour tout entier $n \geq 0$, on note $IP_n(\Lambda)$ l'espace des restrictions à Λ des polynômes à une variable de degré $\leq n$. On note $IP_n^0(\Lambda)$ l'espace des polynômes de $IP_n(\Lambda)$ qui s'annulent aux deux extrémités de Λ.

Soit N un entier ≥ 3 fixé. Le premier problème discret consiste à trouver u_N dans $IP_n^0(\Lambda)$ tel que

$$\forall v_N \in IP_N^0(\Lambda), \quad \alpha(u_N, v_N) = \int_{-1}^{1} f(x) v_N(x) dx \quad (1)$$

Q1. *Indiquer la dimension des espaces $IP_n(\Lambda)$ et $IP_n^0(\Lambda)$. Montrer que le problème (1) admet une solution unique u_N dans $IP_n^0(\Lambda)$.*

Q2. *Pour un polynôme v_N quelconque dans $IP_n^0(\Lambda)$, calculer la quantité*

$$\int_{-1}^{1} (u' - u'_N)(x) v'_N(x) dx + \int_{-1}^{1} a(x)(u - u_N)(x) v_N(x) dx$$

où u et u_N sont respectivement les solutions des problèmes (4) de l'énoncé 49 et (1) de cet énoncé. En déduire l'existence d'une constante c positive indépendante de N telle qu'on ait la relation suivante entre la solution u du problème (4) de l'énoncé 49 et la solution u_N du problème (1) de cet énoncé :

$$\|u - u_N\|_{H^1(\Lambda)} \leq c \underset{v_N \in IP_N^0(\Lambda)}{Inf} \|u - v_N\|_{H^1(\Lambda)}$$

On rappelle qu'il existe une unique famille $(L_n)_{n \in IN}$ de polynômes, appelés polynômes de Legendre, qui sont deux à deux orthogonaux dans $L^2(\Lambda)$ pour le produit scalaire $(\cdot,\cdot)$ et tels que chaque polynôme L_n soit de degré n et vérifie $L_n(1) = 1$.

Q3. *Montrer que l'équation différentielle suivante est satisfaite par tout polynôme L_n, $n \in IN$:*

$$\left((1 - x^2) L'_n\right)'(x) + n(n+1) L_n(x) = 0 \quad (2)$$

(on pourra vérifier que $\left((1 - x^2) L'_n\right)'$ est orthogonal dans $L^2(\Lambda)$ à tous les polynômes de degré $\leq n-1$). Puis, pour tous entiers positifs m et n, calculer $\int_{-1}^{1} L'_m(x) L'_n(x)(1 - x^2) dx$ en fonction de $\|L_n\|_{L^2[\Lambda]}^{\ 2}$.

Q4. *Montrer que les polynômes L_n, $n \in IN$, forment une famille totale de $L^2(\Lambda)$. On note π_N l'opérateur de projection orthogonale de $L^2(\Lambda)$ sur $IP_n(\Lambda)$ pour le* *produit scalaire* $(\cdot,\cdot)$. *Etant donnée une fonction v de* $L^2(\Lambda)$, *écrire le développement de* $\pi_N(v)$ *dans la base* $\{L_n,\ 0 \le n \le N\}$ *en fonction de v.*

Q5. *On note A l'opérateur*

$$Av = -\left(\left(1-x^2\right)v'\right)'$$

Montrer que, pour toutes fonctions u et v deux fois continûment dérivables sur $\overline{\Lambda}$, *on a*

$$\int_{-1}^{1}(Au)(x)v(x)dx = \int_{-1}^{1}u(x)(Av)(x)dx \ \ et \ \int_{-1}^{1}(Au)(x)u(x)dx \ge 0$$

Q6. *A partir de l'espace* $H^1(\Lambda)$, *on définit les espaces* $H^m(\Lambda)$ *pour tout entier* $m \ge 2$ *par la formule de récurrence*

$$H^m(\Lambda) = \left\{v \in H^{m-1}(\Lambda);\quad v' \in H^{m-1}\right\}$$

On note aussi $H^0(\Lambda)$ *l'espace* $L^2(\Lambda)$. *On introduit la notation* d^m *pour la dérivée (au sens généralisé) d'ordre m pour toute fonction v de* $H^m(\Lambda)$:

$$d^0v = v \ \ et \ \ d^m v = \left(d^{m-1}v\right)' \ \ si \ \ m \ge 1.$$

On munit l'espace $H^m(\Lambda)$ *de la norme*

$$\|v\|_{H^m(\Lambda)} = \left(\sum_{l=0}^{m}\left\|d^l v\right\|_{L^2(\Lambda)}{}^2\right)^{\frac{1}{2}}$$

Montrer que, pour tout entier $k \ge 0$, *l'opérateur A est continu de* $H^{k+2}(\Lambda)$ *dans* $H^k(\Lambda)$.

On munit également l'espace $H^m(\Lambda)$ *de la norme (*A^0 *désigne l'opérateur identité)*

$$\|v\|_{D^m(\Lambda)} = \begin{cases} \left(\|v\|_{L^2(\Lambda)}{}^2 + \left\|A^k v\right\|_{L^2(\Lambda)}{}^2\right)^{\frac{1}{2}} & \text{si } m \text{ est un entier pair égal à } 2k \\ \left(\|v\|_{L^2(\Lambda)}{}^2 + \displaystyle\int_{-1}^{1}\left(\left(A^k v\right)'\right)^2(x)\left(1-x^2\right)dx\right)^{\frac{1}{2}} & \text{si } m \text{ est un entier impair égal à } 2k+1 \end{cases}$$

Montrer que, pour tout entier $m \ge 1$, *il existe une constante c positive telle que, pour toute fonction v de* $H^m(\Lambda)$,

$$\|v\|_{D^m(\Lambda)} \le c\|v\|_{H^m(\Lambda)}$$

Q7. *Démontrer la majoration suivante : pour tout entier* $m \ge 1$ *et pour toute fonction v de* $H^m(\Lambda)$,

$$\left\|v - \pi_N(v)\right\|_{L^2(\Lambda)} \le N^{-m}\|v\|_{D^m(\Lambda)}$$

On traitera séparément les cas où m est pair et impair :

— dans le cas où m est pair égal à 2k, on calculera les coefficients de $A^k v$ *dans la famille des* L_n, $n \in IN$, *en fonction de ceux de v, à partir de l'équation (2) ;*

— dans le cas où m est impair égal à 2k + 1, on calculera les coefficients de $\left(A^k v\right)'$ *dans la famille des* L'_n, $n \geq 1$, *dont on montrera qu'elle est une base orthogonale de l'espace des (classes de) fonctions de carré intégrable pour la mesure* $\left(1-x^2\right)dx$ *sur Λ.*

Q8. *On définit l'opérateur* π_N^1 *sur les fonctions de* $H_0^1(\Lambda)$ *par la formule*

$$\left(\pi_N^1(v)\right)(x) = \int_{-1}^{x} \left(\pi_{N-1}(v')\right)(t)dt$$

Démontrer que l'opérateur π_N^1 *envoie* $H_0^1(\Lambda)$ *sur* $IP_n^0(\Lambda)$. *Etablir la majoration suivante : pour tout entier* $m \geq 1$, *il existe une constante c positive indépendante de N telle que, pour toute fonction v de* $H_0^1(\Lambda) \cap H^m(\Lambda)$,

$$\left\| v - \pi_N^1(v) \right\|_{H^1(\Lambda)} \leq cN^{1-m} \|v'\|_{D^{m-1}(\Lambda)} \qquad (3)$$

Q9. *Etablir également la majoration suivante : pour tout entier* $m \geq 1$, *il existe une constante c positive indépendante de N telle que, pour toute fonction v de* $H_0^1(\Lambda) \cap H^m(\Lambda)$,

$$\left\| v - \pi_N^1(v) \right\|_{L^2(\Lambda)} \leq cN^{-m} \|v'\|_{D^{m-1}(\Lambda)} \qquad (4)$$

On pourra introduire la fonction w de $H_0^1(\Lambda)$ *dont la dérivée seconde est égale à* $v - \pi_N^1(v)$ *et montrer que*

$$\int_{-1}^{1} \left(v - \pi_N^1(v)\right)^2(x)dx \leq \left\| v' - \left(\pi_N^1(v)\right)' \right\|_{L^2(\Lambda)} \left\| w' - \left(\pi_N^1(w)\right)' \right\|_{L^2(\Lambda)}$$

Q10. *Etant donnée une fonction v de* $H^1(\Lambda)$, *on pose*

$$v_0(x) = v(x) - \frac{1-x}{2}v(-1) - \frac{1+x}{2}v(1)$$

et on étend l'opérateur π_N^1 *à* $H^1(\Lambda)$ *par la formule*

$$\pi_N^1(v) = \pi_N^1\left(v_0\right) + \frac{1-x}{2}v(-1) + \frac{1+x}{2}v(1)$$

Montrer que, pour tout entier $m \geq 1$, *l'application :* $v \mapsto v_0$ *est continue de* $H^m(\Lambda)$ *dans* $H_0^1(\Lambda) \cap H^m(\Lambda)$. *Démontrer un analogue des majorations (3) et (4) qui soit vrai pour toute fonction v de* $H^m(\Lambda)$, $m \geq 1$.

Q11. *On suppose la solution u du problème (1) de l'énoncé 49 dans* $H^m(\Lambda)$, *pour un entier* $m \geq 1$ *donné. Etablir la majoration d'erreur suivante entre la solution u du problème (1) de l'énoncé 49 et la solution* u_N *du problème (1) de cet énoncé : il existe une constante c positive indépendante de N telle que*

$$\left\| u - u_N \right\|_{H^1(\Lambda)} \leq cN^{1-m} \|u'\|_{D^{m-1}(\Lambda)}$$

Solution

Q1. On a $Dim\left(IP_N(\Lambda)\right) = N+1$ et $Dim\left(IP_N^0(\Lambda)\right) = N-1$ du fait que tout polynôme P de $IP_N^0(\Lambda)$ peut s'écrire $P = \left(x^2+1\right)Q$ avec Q de degré inférieur ou égal à $N-2$.

En désignant par $\left\{\varphi_j;\ 1 \le j \le N-1\right\}$ une base de $IP_N^0(\Lambda)$ et en écrivant le polynôme u_N de $IP_N^0(\Lambda)$ sous la forme $u_N = \sum_{j=1}^{N-1} \lambda_j \varphi_j$, on voit que le problème (1) est équivalent à la résolution du système linéaire de $N-1$ équations à $N-1$ inconnues :

$$\sum_{j=1}^{N-1} \alpha\left(\varphi_i, \varphi_j\right)\lambda_j = \int_{-1}^{1} f(x)\varphi_i(x)dx \quad \left(1 \le i \le N-1\right)$$

La matrice de ce système est celle du produit scalaire α dans la base $\left\{\varphi_j;\ 1 \le j \le N-1\right\}$, elle est donc non dégénérée et le système admet une unique solution.

On a donc ainsi prouvé que le problème (1) admet une unique solution dans $IP_N^0(\Lambda)$.

Q2. Les équations (4) du problème 49 et (1) de ce problème appliquées à une fonction $v_N \in IP_N^0(\Lambda) \subset H_0^1(\Lambda)$ donnent :

$$\alpha\left(u, v_N\right) = \int_{-1}^{1} f(x)v_N(x)dx = \alpha\left(u_N, v_N\right)$$

On déduit alors que $\alpha\left(u - u_N, v_N\right) = 0$, c'est-à-dire que :

$$\forall v_N \in IP_N^0(\Lambda), \quad \int_{-1}^{1} \left(u' - u_N'\right)(x)v_N'(x)dx + \int_{-1}^{1} a(x)\left(u - u_N\right)(x)v_N(x)dx = 0$$

Avec l'inégalité triangulaire, on a :

$$\forall v_N \in IP_N^0(\Lambda), \quad \left\|u - u_N\right\|_{H^1(\Lambda)} \le \left\|u - v_N\right\|_{H^1(\Lambda)} + \left\|v_N - u_N\right\|_{H^1(\Lambda)}$$

D'autre part, en utilisant l'équivalence des normes $\left|\ \right|_{H^1(\Lambda)}$ et $\|\cdot\|_{H^1(\Lambda)}$ sur $H_0^1(\Lambda)$ (problème 49 Q4) et la coercivité de α (problème 49 Q5), on peut écrire que :

$$\left\|v_N - u_N\right\|_{H^1(\Lambda)}^2 \le c_1\left|v_N - u_N\right|_{H^1(\Lambda)}^2 \le c_1\alpha\left(v_N - u_N, v_N - u_N\right)$$

avec :

$$\alpha\left(v_N - u_N, v_N - u_N\right) = \alpha\left(v_N - u, v_N - u_N\right) + \alpha\left(u - u_N, v_N - u_N\right) = \alpha\left(v_N - u, v_N - u_N\right)$$

Avec la continuité de α, on a $\alpha\left(v_N - u, v_N - u_N\right) \le c_2\left\|v_N - u\right\|_{H^1(\Lambda)}\left\|v_N - u_N\right\|_{H^1(\Lambda)}$.

Ce qui donne en définitive :

$$\left\|v_N - u_N\right\|_{H^1(\Lambda)} \le c_1c_2\left\|v_N - u\right\|_{H^1(\Lambda)}$$

et :

$$\forall v_N \in IP_N^0(\Lambda), \quad \left\|u - u_N\right\|_{H^1(\Lambda)} \le \left(1 + c_1c_2\right)\left\|u - v_N\right\|_{H^1(\Lambda)} = c\left\|u - v_N\right\|_{H^1(\Lambda)}$$

Ce qui entraîne :

$$\|u-u_N\|_{H^1(\Lambda)} \le c \operatorname*{Inf}_{v_N \in IP_N^0(\Lambda)} \|u-v_N\|_{H^1(\Lambda)}$$

la constante c étant indépendante de N.

Q3. Deux intégrations par parties successives permettent d'écrire, pour tout polynôme P de degré inférieur ou égal à $n-1$:

$$\int_{-1}^{1}\left(\left(1-x^2\right)L_n'\right)'(x)P(x)dx = -\int_{-1}^{1} L_n'(x)\left(1-x^2\right)P'(x)dx$$

$$= \int_{-1}^{1} L_n(x)\left(\left(1-x^2\right)P'\right)'(x)dx = 0$$

(le polynôme $\left(1-x^2\right)P'$ est de degré strictement inférieur à n).

Le polynôme $\left(\left(1-x^2\right)L_n'\right)'$ est donc orthogonal, dans $L^2(\Lambda)$, à tous les polynômes de degré inférieur ou égal à $n-1$. Ce polynôme étant de degré exactement n, on déduit qu'il est proportionnel à L_n (unicité des L_n). C'est-à-dire qu'il existe une constante λ_n telle que $\left(\left(1-x^2\right)L_n'\right)' = \lambda_n L_n$. L'identification des coefficients de x^n donne immédiatement $\lambda_n = -n(n+1)$.

On a donc ainsi prouvé que L_n est solution de l'équation différentielle (2).

En effectuant une intégration par parties puis en utilisant cette équation différentielle, on a :

$$\int_{-1}^{1} L_m'(x)L_n'(x)\left(1-x^2\right)dx = -\int_{-1}^{1} L_m(x)\left(L_n'\left(1-x^2\right)\right)'(x)dx = n(n+1)\int_{-1}^{1} L_m(x)L_n(x)dx$$

Soit :

$$\int_{-1}^{1} L_m'(x)L_n'(x)\left(1-x^2\right)dx = \begin{cases} n(n+1)\|L_n\|^2_{L^2(\Lambda)} & \text{si } n=m \\ 0 \text{ si } n \ne m \end{cases}$$

Q4. (a) On sait déjà que les polynômes de Legendre sont orthogonaux deux à deux. Pour montrer qu'ils forment une famille totale il reste à montrer que l'espace vectoriel qu'ils engendrent est dense dans $L^2(\Lambda)$, ce qui revient à montrer que le sous espace vectoriel $IP(\Lambda)$ de $L^2(\Lambda)$ formé des fonctions polynomiales est dense dans $L^2(\Lambda)$.

On sait que $C^\infty\left(\overline{\Lambda}\right)$ est dense dans $\left(L^2(\Lambda), \|\ \|_{L^2(\Lambda)}\right)$ et que $IP(\Lambda)$ est dense dans $\left(C^\infty\left(\overline{\Lambda}\right), \|\ \|_\infty\right)$ (théorème de Stone–Weierstrass). Donc pour toute fonction $f \in L^2(\Lambda)$ et pour tout réel $\varepsilon > 0$, on peut trouver une fonction $\varphi \in C^\infty\left(\overline{\Lambda}\right)$ et un polynôme $P \in IP(\Lambda)$ tels que :

$$\|f-\varphi\|_2 \le \varepsilon \text{ et } \|P-\varphi\|_\infty \le \varepsilon$$

On a alors :

$$\|f-P\|_2 \le \|f-\varphi\|_2 + \|P-\varphi\|_2 \le \varepsilon + \sqrt{2}\|P-\varphi\|_\infty \le \left(1+\sqrt{2}\right)\varepsilon$$

D'où le résultat.

(b) La famille des polynômes de Legendre étant totale dans $L^2(\Lambda)$, toute fonction $v \in L^2(\Lambda)$ peut se décomposer sous la forme $v = \sum_{n=0}^{+\infty} c_n L_n$, les coefficients (de Fourier) c_n étant donnés par :

$$c_n = \frac{1}{\|L_n\|^2_{L^2(\Lambda)}} \int_{-1}^{1} v(x) L_n(x) dx \quad (n \in IN)$$

La projection orthogonale de v sur $IP_N(\Lambda)$ est alors donné par (voir le problème 47 Q3) :

$$\pi_N(v) = \sum_{n=0}^{N} c_n L_n$$

Q5. Soient u et v deux fonctions dans $C^2(\overline{\Lambda})$. En effectuant deux intégrations par parties, on a :

$$\int_{-1}^{1} (Au)(x) v(x) dx = \int_{-1}^{1} u'(x) v'(x)(1 - x^2) dx = \int_{-1}^{1} u(x)(Av)(x) dx$$

Pour $u = v$, une intégration par parties donne :

$$\int_{-1}^{1} (Au)(x) u(x) dx = \int_{-1}^{1} u'(x)^2 (1 - x^2) dx \geq 0$$

Q6. (a) L'opérateur A est aussi défini sur $H^{k+2}(\Lambda)$, pour tout entier $k \geq 0$, avec :

$$\forall v \in H^{k+2}(\Lambda), \quad Av = -\left((1 - x^2)v'\right)' = 2xv' - (1 - x^2)v'' \in H^k(\Lambda)$$

la formule de Leibnitz étant encore valable pour les dérivées généralisées (voir le problème 48 Q11 (a)).

Par récurrence sur $l \in \{0, \cdots, k\}$, on vérifie que :

$$\forall v \in H^{k+2}(\Lambda), \quad d^l Av = l(l+1) d^l v + 2(l+1) x d^{l+1} v - (1 - x^2) d^{l+2} v \in H^{k-l}(\Lambda)$$

et :

$$\begin{aligned} \|d^l Av\|_{L^2(\Lambda)} &\leq l(l+1)\|d^l v\|_{L^2(\Lambda)} + 2(l+1)\|d^{l+1} v\|_{L^2(\Lambda)} + \|d^{l+2} v\|_{L^2(\Lambda)} \\ &\leq \left((l+2)(l+1) + 1\right)\|v\|_{H^{k+2}(\Lambda)} \end{aligned}$$

La continuité de A en résulte alors immédiatement.

(b) Soit $m = 2k$ un entier pair. L'itérée A^k étant continue de $H^m(\Lambda)$ dans $L^2(\Lambda)$, on peut trouver une constante c_1 telle que :

$$\forall v \in H^m(\Lambda), \quad \|A^k v\|_{L^2(\Lambda)} \leq c_1 \|v\|_{H^m(\Lambda)}$$

et on a alors :

$$\forall v \in H^m(\Lambda), \quad \|v\|_{D^m(\Lambda)} \leq \sqrt{1 + c_1^{\,2}}\, \|v\|_{H^m(\Lambda)}$$

(c) Soit $m = 2k + 1$ un entier impair. L'itérée A^k étant continue de $H^m(\Lambda)$ dans $H^1(\Lambda)$, on peut trouver une constante c_2 telle que :

$$\forall v \in H^m(\Lambda), \quad \|A^k v\|_{H^1(\Lambda)} \leq c_2 \|v\|_{H^m(\Lambda)}$$

Avec :

$$\int_{-1}^{1}\left(\left(A^k v\right)'\right)^2(x)\left(1-x^2\right)dx \le \int_{-1}^{1}\left(\left(A^k v\right)'\right)^2(x)dx = \left\|\left(A^k v\right)'\right\|_{L^2(\Lambda)}^2 \le \left\|A^k v\right\|_{H^1(\Lambda)}^2$$

on déduit alors que :

$$\forall v \in H^m(\Lambda),\quad \|v\|_{D^m(\Lambda)} \le \sqrt{1+c_2{}^2}\,\|v\|_{H^m(\Lambda)}$$

Q7. Pour toute fonction $v \in H^m(\Lambda)$, on a $v = \sum_{n=0}^{+\infty} c_n L_n$ et $A^k v = \sum_{n=0}^{+\infty} d_n L_n$, la convergence de ces séries ayant lieu dans $L^2(\Lambda)$. Les coefficients d_n sont donnés par :

$$d_n = \frac{1}{\|L_n\|_{L^2(\Lambda)}^2}\int_{-1}^{1} A^k v(x)\,L_n(x)dx = \frac{1}{\|L_n\|_{L^2(\Lambda)}^2}\int_{-1}^{1} v(x)\,A^k L_n(x)dx \quad (n \in IN)$$

(les résultats de Q5 sont encore valables dans $H^1(\Lambda)$ du fait de la densité de $C^\infty(\overline{\Lambda})$ dans $H^1(\Lambda)$ et de la continuité de l'opérateur A et du produit scalaire de $L^2(\Lambda)$).

L'équation différentielle vérifiée par les polynômes de Legendre (équation (2) de Q3) peut s'écrire $AL_n = n(n+1)L_n$. On a donc $A^k L_n = n^k(n+1)^k L_n$ et les coefficients d_n sont donnés par :

$$d_n = n^k(n+1)^k c_n \quad (n \in IN)$$

(a) Soit $m = 2k$ un entier pair.

On déduit des calculs qui précédent que :

$$\begin{aligned}\|v-\pi_N(v)\|_{L^2(\Lambda)}^2 &= \sum_{n=N+1}^{+\infty} c_n{}^2\|L_n\|_{L^2(\Lambda)}^2 = \sum_{n=N+1}^{+\infty} \frac{d_n{}^2}{n^{2k}(n+1)^{2k}}\|L_n\|_{L^2(\Lambda)}^2\\ &\le \frac{1}{N^{2k}(N+1)^{2k}}\sum_{n=0}^{+\infty} d_n{}^2\|L_n\|_{L^2(\Lambda)}^2 = \frac{1}{N^{2k}(N+1)^{2k}}\left\|A^k v\right\|_{L^2(\Lambda)}^2\end{aligned}$$

et :

$$\forall v \in H^m(\Lambda),\quad \|v-\pi_N(v)\|_{L^2(\Lambda)} \le N^{-m}\|v\|_{D^m(\Lambda)}$$

(b) Soit $m = 2k+1$ un entier impair.

Pour $n \ge 1$, le polynôme L_n' est de degré $n-1$. La famille des L_n' ($n \ge 1$) forme donc une base de $IP(\Lambda)$ et comme pour les polynômes de Legendre, on déduit que l'espace vectoriel engendré par les L_n' ($n \ge 1$) est dense dans $\left(L^2(\Lambda), \|\ \|_{L^2(\Lambda)}\right)$.

D'autre part, pour n, m dans $IN-\{0\}$, une intégration par parties donne :

$$\int_{-1}^{1} L_n'(x)L_m'(x)\left(1-x^2\right)dx = \int_{-1}^{1} L_n(x)\left(AL_m\right)(x)dx = m(m+1)\left(L_n, L_m\right)$$

On déduit donc que la famille des L_n' ($n \ge 1$) forme une base orthogonale de $L^2(\Lambda)$ pour la mesure $\left(1-x^2\right)dx$.

La fonction $\left(A^k v\right)' \in L^2(\Lambda)$ peut alors se décomposer sous la forme :

$$\left(A^k v\right)' = \sum_{n=1}^{+\infty} e_n L_n'$$

avec :

$$e_n = \frac{1}{n(n+1)\|L_n\|^2_{L^2(\Lambda)}} \int_{-1}^{1} \left(A^k v\right)'(x) L_n'(x)\left(1-x^2\right)dx$$

$$= \frac{1}{n(n+1)\|L_n\|^2_{L^2(\Lambda)}} \int_{-1}^{1} \left(A^k v\right)(x)\left(AL_n\right)(x)dx = \frac{1}{\|L_n\|^2_{L^2(\Lambda)}} \int_{-1}^{1} \left(A^k v\right)(x) L_n(x)dx = d_n$$

(notations du (a)).

On a alors :

$$\int_{-1}^{1} \left(\left(A^k v\right)'\right)^2 (x)\left(1-x^2\right)dx = \sum_{n=1}^{+\infty} d_n{}^2 n(n+1)\|L_n\|^2_{L^2(\Lambda)}$$

et :

$$\|v - \pi_N(v)\|^2_{L^2(\Lambda)} = \sum_{n=N+1}^{+\infty} \frac{d_n{}^2}{n^{2k+1}(n+1)^{2k+1}} n(n+1)\|L_n\|^2_{L^2(\Lambda)}$$

$$\leq \frac{1}{N^{2k+1}(N+1)^{2k+1}} \sum_{n=1}^{+\infty} d_n{}^2 n(n+1)\|L_n\|^2_{L^2(\Lambda)} \leq \frac{1}{N^{2k+1}(N+1)^{2k+1}} \|v\|^2_{D^m(\Lambda)}$$

C'est-à-dire que :

$$\forall v \in H^m(\Lambda), \quad \|v - \pi_N(v)\|_{L^2(\Lambda)} \leq N^{-m}\|v\|_{D^m(\Lambda)}$$

Q8. (a) Pour $v \in H^1(\Lambda)$, on a $\pi_{N-1}(v') = \sum_{n=0}^{N-1} c_n' L_n \in IP_{N-1}(\Lambda)$ et $\pi_N^1(v) \in IP_N(\Lambda)$. De plus il est clair que $\pi_N^1(v)(-1) = 0$.

Avec l'orthogonalité des polynômes de Legendre, on peut écrire que :

$$\int_{-1}^{1} \pi_{N-1}(v')(t)dt = \sum_{n=0}^{N-1} c_n'\left(L_n, L_0\right) = c_0'\|L_0\|^2_{L^2(\Lambda)} = \int_{-1}^{1} v'(t)dt = v(1) - v(-1)$$

et dans le cas particulier où $v \in H_0^1(\Lambda)$, on en déduit que $\pi_N^1(v)(1) = 0$.

On a donc $\pi_N^1\left(H_0^1(\Lambda)\right) \subset IP_N^0(\Lambda)$.

(b) En remarquant que $\pi_N^1(v) = v$ pour tout polynôme $v \in IP_N^0(\Lambda)$, on déduit que π_N^1 est surjective.

(c) Dans le problème 49, on a montré que la semi–norme $|v|_{H^1(\Lambda)} = \|v'\|_{L^2(\Lambda)}$ est une norme sur $H_0^1(\Lambda)$, équivalente à la norme $\|\cdot\|_{H^1(\Lambda)}$. Il existe donc une constante positive c telle que :

$$\forall v \in H_0^1(\Lambda), \quad \|v - \pi_N^1(v)\|_{H^1(\Lambda)} \leq c\left|v - \pi_N^1(v)\right|_{H^1(\Lambda)} = c\|v' - \pi_{N-1}(v')\|_{L^2(\Lambda)}$$

($v - \pi_N^1(v) \in H_0^1(\Lambda)$ d'après (a)).

Si $v \in H_0^1(\Lambda) \cap H^m(\Lambda)$, alors $v' \in H^{m-1}(\Lambda)$ et avec l'inégalité de Q7, on a :

$$\left\|v' - \pi_{N-1}(v')\right\|_{L^2(\Lambda)} \leq N^{-(m-1)} \|v'\|_{D^{m-1}(\Lambda)}$$

et :

$$\left\|v - \pi^1_N(v)\right\|_{H^1(\Lambda)} \leq cN^{1-m} \|v'\|_{D^{m-1}(\Lambda)}$$

Q9. Soit $v \in H_0^1(\Lambda)$. La fonction $v - \pi^1_N(v)$ est dans $H_0^1(\Lambda)$ et la fonction w_1 définie par $w_1(x) = \int_{-1}^{x} \left(v - \pi_N^1(v)\right)(t)dt$ est dans $H^2(\Lambda)$ avec $w_1' = v - \pi_N^1(v)$. La fonction w définie par $w(x) = \int_{-1}^{x} w_1(t)dt + \gamma(x+1)$ où γ est tel que $w(1) = 0$ ($\gamma = -\dfrac{1}{2}\int_{-1}^{1} w_1(t)dt$) est dans $H_0^1(\Lambda) \cap H^3(\Lambda)$ et vérifie $w_1'' = v - \pi_N^1(v)$.

En effectuant une intégration par parties, on a :

$$\begin{aligned}\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} &= \int_{-1}^{1} \left(v - \pi_N^1(v)\right)(x) w''(x)dx \\ &= -\int_{-1}^{1} \left(v - \pi_N^1(v)\right)'(x) w'(x)dx = -\int_{-1}^{1} \left(v' - \pi_{N-1}(v')\right)(x) w'(x)dx\end{aligned}$$

Mais, par définition de la projection orthogonale, $v' - \pi_{N-1}(v') \in IP_{N-1}(\Lambda)^{\perp}$, donc $\int_{-1}^{1} \left(v' - \pi_{N-1}(v')\right)(x)\pi_{N-1}(w')(x)dx = 0$ et on peut écrire que :

$$\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} = \int_{-1}^{1} \left(v' - \pi_{N-1}(v')\right)(x)\left(\pi_{N-1}(w') - w'\right)(x)dx$$

Puis en utilisant l'inégalité de Cauchy–Schwarz, on déduit que :

$$\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} \leq \left\|v' - \pi_{N-1}(v')\right\|_{L^2(\Lambda)} \left\|\pi_{N-1}(w') - w'\right\|_{L^2(\Lambda)}$$

Soit :

$$\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} \leq \left\|v' - \pi^1_N(v)'\right\|_{L^2(\Lambda)} \left\|w' - \pi^1_N(w)'\right\|_{L^2(\Lambda)}$$

Si de plus $v \in H^m(\Lambda)$, on déduit alors de (3) que :

$$\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} \leq c_1 N^{1-m} \|v'\|_{D^{m-1}(\Lambda)} c_1 N^{1-3} \|w'\|_{D^{3-1}(\Lambda)}$$

Avec $\|w'\|_{D^2(\Lambda)} \leq c_2 \|w'\|_{H^2(\Lambda)} \leq c_2 \|w\|_{H^3(\Lambda)} \leq c_2 c_3 \left\|v - \pi_N^1(v)\right\|_{H^1(\Lambda)}$, on déduit que :

$$\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} \leq c_4 N^{-1-m} \|v'\|_{D^{m-1}(\Lambda)} \left\|v - \pi_N^1(v)\right\|_{H^1(\Lambda)}$$

Enfin, en utilisant l'inégalité (3), on aboutit à $\left\|v - \pi_N^1(v)\right\|^2_{L^2(\Lambda)} \leq c_5 N^{-2m} \|v'\|^2_{D^{m-1}(\Lambda)}$, soit :

$$\forall v \in H_0^1(\Lambda) \cap H^m(\Lambda), \quad \left\|v - \pi_N^1(v)\right\|_{L^2(\Lambda)} \leq cN^{-m} \|v'\|_{D^{m-1}(\Lambda)}$$

Q10. (a) Il est clair que $v_0 \in H_0^1(\Lambda) \cap H^m(\Lambda)$, pour toute fonction $v \in H^m(\Lambda)$.

Avec $\|v_0\|_{H^m(\Lambda)} \leq \|v\|_{H^m(\Lambda)} + \frac{\|v\|_\infty}{2}\left\{\|1-x\|_{H^m(\Lambda)} + \|1+x\|_{H^m(\Lambda)}\right\}$ et $\|v\|_\infty \leq 2\|v\|_{H^1(\Lambda)}$ (problème 49, Q3), on déduit que l'application $v \mapsto v_0$ est linéaire continue ($\|v\|_{H^1(\Lambda)} \leq \|v\|_{H^m(\Lambda)}$).

(b) On a $v - \pi_N^1(v) = v_0 - \pi_N^1(v_0)$ et en utilisant les majorations (3) et (4) :

$$\left\|v_0 - \pi_N^1(v_0)\right\|_{H^1(\Lambda)} + N\left\|v_0 - \pi_N^1(v_0)\right\|_{L^2(\Lambda)} \leq c_6 N^{1-m}\|v_0'\|_{D^{m-1}(\Lambda)}$$

avec $\|v_0'\|_{D^{m-1}(\Lambda)} \leq c_7\|v_0'\|_{H^{m-1}(\Lambda)}$ (Q6) et $\|v_0'\|_{H^{m-1}(\Lambda)} \leq \|v_0\|_{H^m(\Lambda)}$. Ce qui donne :

$$\left\|v - \pi_N^1(v)\right\|_{H^1(\Lambda)} + N\left\|v - \pi_N^1(v)\right\|_{L^2(\Lambda)} \leq c_8 N^{1-m}\|v_0\|_{H^m(\Lambda)}$$

Et avec la continuité de $v \mapsto v_0$:

$$\left\|v - \pi_N^1(v)\right\|_{H^1(\Lambda)} + N\left\|v - \pi_N^1(v)\right\|_{L^2(\Lambda)} \leq cN^{1-m}\|v\|_{H^m(\Lambda)}$$

Q11. On a $\|u - u_N\|_{H^1(\Lambda)} \leq c_9 \underset{v_N \in IP_N^0(\Lambda)}{Inf} \|u - v_N\|_{H^1(\Lambda)}$ (Q2), donc :

$$\|u - u_N\|_{H^1(\Lambda)} \leq c_9\left\|u - \pi_N^1(u)\right\|_{H^1(\Lambda)}$$

et avec la majoration (3) :

$$\|u - u_N\|_{H^1(\Lambda)} \leq cN^{1-m}\|u'\|_{D^{m-1}(\Lambda)}$$

BIBLIOGRAPHIE

AHLBERG, NILSON, WALSH — *The theory of Splines and their applications.* Academic Press (1967).

AYANT, BORG — *Fonctions spéciales.* Dunod (1971).

BARANGER J. — *Analyse numérique.* Hermann (1991).

BAYEN F. MARGARIA C. — *Espaces de Hilbert et opérateurs.* Ellipse (1986).

BOURGEOIS G. — *Calcul numérique de la racine carrée d'une matrice symétrique positive.* Revue des Mathématiques Spéciales (Novembre 1993).

BREZIS H. — *Analyse fonctionnelle* . Masson (1993).

BURLISCH R. STOER J. — *Introduction to Numerical Analysis.* Springer Verlag (1980).

CARTAN H. — *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Masson (1961).

CIARLET P. G. — *Introduction à l'analyse numérique matricielle et à l'optimisation.* Masson (1982).

CROUZIEX M. MIGNOT A. — *Analyse numérique des équations différentielles.* Masson (1984).

CROUZIEX M. MIGNOT A. — *Exercices d'analyse numérique des équations différentielles.* Masson (1986).

DAVIS P. — *Interpolation and approximation.* Dover Publications (1975).

DEMAILLY J. P. — *Analyse numérique des équations différentielles.* Presses Universitaires de Grenoble (1991).

DIEUDONNE J. — *Calcul infinitésimal.* Hermann (1968).

HENRICE P. — *Applied and computational complex analysis. Vol. 1.* Wiley (1998).

HOUSEHOLDER A. S. — *The theory of matrices in numerical analysis.* Dover (1975).

KOLMOGOROV A. FOMINE S. — *Eléments de la théorie des fonctions et de l'analyse fonctionnelle.* Mir (1974).

KURTZ D. C. — *A sufficient condition for all the roots of a polynomial to be real.* American Mathematical monthly (1992).

LASCAUX P. THEODOR R. — *Analyse numérique matricielle appliquée à l'art de l'ingénieur, volumes 1 et 2.* Masson (1986, 1987).

LEBEDEV N. — *Special functions and their applications.* Prentice-Hall (1965).

LI T. Y. ZENG Z. — *Homotopy determinant algorithm for solving nonsymetric eigenvalue problems.* Mathematics of computation, volume 59, N° 200 (Octobre 1992).

MONASSE D. — *La méthode de Newton et les racines des polynômes.* Revue des Mathématiques Spéciales (Janvier 1987).

ORTEGA J. M. RHEINBOLDT W. C. — *Iterative solution of nonlinear equations in several variables.* Academic Press 1970).

PARODI M. — *La localisation des valeurs caractéristiques des matrices et ses applications.* Gauthier-Villars 1959).

RELLICH F. — *Perturbation theory of Eigenvalue problems*. Gordon and Breach (1950).

ROMBALDI J. E. — *Algorithmique numérique et Ada.* Masson (1994).

ROSSET S. — *Normalized symetric functions, Newton's inequalities and a new set of stronger inequalities.* American Mathematical monthly (1989).

RUDIN W. — *Analyse réelle et complexe.* Masson (1975).

RUDIN W. — *Functional Analysis.* Mc Graw Hill (1991).
SCHATZMAN M. — *Analyse numérique. Cours et exercices pour la licence.* InterEditions (1991).
SCHMITT J. L. — *Inversion de matrices par la méthode de Newton.* Revue des Mathématiques Spéciales. (Décembre 1992).
WARUSFEL A. — *A propos de la méthode d'intégration de Romberg.* Revue de Mathématiques spéciales (Janvier 1986).
WASSERSTROM E. — *Numerical solutions by the continuation method.* siam review (1973).

INDEX