Université Paris Dauphine
2012 /2013

Analyse Numérique
Groupes 1 et 2

Devoir de Maison 2

Problème 1

Soit f une fonction réelle, continue et dérivable sur $[-1,1]$ et $x_0 \in]0,1[$.

1. Question de cours - Démontrez qu'il existe un unique polynôme Q de degré inférieur ou égal à 5 vérifiant :

$$\begin{array}{lll} Q(-x_0) = f(-x_0) & Q(0) = f(0) & Q(x_0) = f(x_0) \\ Q'(-x_0) = f'(-x_0) & Q'(0) = f'(0) & Q'(x_0) = f'(x_0) \end{array}$$

Corrigé : Soit $\mathcal{P}_5$ l'espace vectoriel des polynômes de degré inférieur ou égal à 5.

Considérons l'application Φ de $\mathcal{P}_5$ dans $\mathbb{R}^6$ qui à un polynôme P associe le sextuplet $(P(-x_0), P(0), P(x_0), P'(-x_0), P'(0), P'(x_0))$. L'existence et l'unicité du polynôme Q est équivalente à la bijectivité de Φ. L'application Φ étant une application linéaire entre deux espace vectoriels de même dimension 6, elle est bijective si et seulement si elle est injective. Pour vérifier l'injectivité de Φ il faut vérifier que son noyau est réduit au polynôme nul, or si $P \in \ker(\Phi)$ alors $-x_0, 0$ et x_0 sont des racines doubles de P, P étant de degré inférieur ou égal à 5 ce n'est possible que si P est le polynôme nul.

2. Question de cours - Montrez que si f est six fois dérivable dans $]-1,1[$ on a la majoration suivante :

$$|f(x) - Q(x)| \leq \frac{M_6}{6!}\varpi(x), \quad \forall x \in [-1,1]$$

où $M_6 = \sup\left\{|f^{(6)}(t)|\,;\, t \in]-1,1[\right\}$ et $\varpi(t) = (x^2 - x_0^2)^2 x^2$.

Corrigé : Si $x = -x_0, 0$ ou x_0 la majoration est vraie car les deux membres sont nuls. Supposons maintenant $x \neq -x_0, 0$ et x_0 et considérons la fonction auxiliaire $\Psi(t)$ définie par :

$$\Psi(t) = f(t) - Q(t) - \frac{f(x) - Q(x)}{\varpi(x)}\varpi(t)$$

Cette fonction s'annule en x, $-x_0$, 0 et x_0, d'après le théorème de Rolle sa dérivée s'annule en 3 points de l'intervalle $]-1,1[$ distincts des points x, , $-x_0$, 0 et x_0. Par ailleurs la dérivée de Φ est nulle par construction en $-x_0$, 0 et x_0. On a donc 6 zéros distincts pour Φ'. Par applications successives du théorème de Rolle on aura 5 zéros distincts pour $\Phi^{(2)}$, 4 pour $\Phi^{(3)}$, ..., et un zéro, soit ξ_x, pour $\Phi^{(6)}$, ce qui s'écrit :

$$\Phi^{(6)}(\xi_x) = f^{(6)}(\xi_x) - \frac{f(x) - Q(x)}{\varpi(x)} 6! = 0$$

Ce qui donne :

$$f(x) - Q(x) = \frac{f^{(6)}(\xi_x)}{6!} \varpi(x)$$

et la majoration demandée s'en déduit.

3. On considère maintenant la formule de quadrature :

$$\int_{-1}^{+1} f(t)dt \approx \alpha f(-x_0) + \beta f(0) + \gamma f(x_0)$$

Déterminez α, β, γ et x_0 pour que la formule soit exacte pour les polynômes de degré inférieur ou égal à 5.

Corrigé : Ecrivons que la formule est exacte pour les monômes $f(x) = x^p$ pour p de 0 à 5 :

$$\begin{array}{ll} f(x) = 1 & 2 = \alpha + \beta + \gamma \\ f(x) = x & 0 = -\alpha\, x_0 + \gamma\, x_0 \\ f(x) = x^2 & \frac{2}{3} = \alpha\, x_0^2 + \gamma\, x_0^2 \\ f(x) = x^3 & 0 = -\alpha\, x_0^3 + \gamma\, x_0^3 \\ f(x) = x^4 & \frac{2}{5} = \alpha\, x_0^4 + \gamma\, x_0^4 \\ f(x) = x^5 & 0 = -\alpha\, x_0^5 + \gamma\, x_0^5 \end{array}$$

On déduit de ce système : $x_0 = \sqrt{\frac{3}{5}}$, $\alpha = \gamma = \frac{5}{9}$ et $\beta = \frac{8}{9}$.

4. Montrez que l'erreur de quadrature

$$E(f) = \int_{-1}^{+1} f(t)dt - (\alpha f(-x_0) + \beta f(0) + \gamma f(x_0))$$

vérifie :

$$E(f) = \int_{-1}^{+1} (f(t) - Q(t))\, dt$$

avec Q le polynôme d'interpolation de f défini à la question 1.

Corrigé : Le polynôme Q étant de degré 5 et la formule étant exacte pour les polynômes de degré 5 on a :

$$\int_{-1}^{+1} Q(t)dt = \alpha\, Q(-x_0) + \beta\, Q(0) + \gamma\, Q(x_0)$$

Le résultat demandé se déduit des propriétés d'interpolation de Q.

5. En déduire, pour f six fois dérivable, la majoration de l'erreur :

$$|E(f)| \leq \frac{M_6}{15750}$$

Corrigé : La formule de la question 4 et la majoration de la question 2 permettent d'écrire :

$$|E(f)| \leq \int_{-1}^{+1} \frac{M_6}{6!}\varpi(t)dt = \frac{M_6}{6!}\frac{8}{175}$$

6. Déduisez des résultats précédents une formule de quadrature élémentaire sur un intervalle fini quelconque $]a, b[$

$$\int_a^b f(t)dt \approx \mu_1 f(c_1) + \mu_2 f(c_2) + \mu_3 f(c_3)$$

qui soit exacte pour les polynômes de degré 5.

Corrigé : La transformation affine $t = \frac{a+b}{2} + \frac{b-a}{2}y$ envoie le segment $[-1, 1]$ sur le segment $[a, b]$ et donc :

$$\begin{aligned}\int_a^b f(t)dt &= \frac{b-a}{2}\int_{-1}^{+1} f(\frac{a+b}{2} + \frac{b-a}{2}y)\,dy \\ &\approx \frac{b-a}{2}\left(\alpha f(\frac{a+b}{2} - \frac{b-a}{2}x_0) + \beta f(\frac{a+b}{2}) + \gamma f(\frac{a+b}{2} + \frac{b-a}{2}x_0)\right)\end{aligned}$$

La formule est exacte pour les polynômes de degré 5 puisque la transformation affine préserve le sous-espace vectoriel des polynômes de degré inférieur ou égal p.

7. Donnez une majoration de l'erreur de quadrature pour cette formule en fonction de $(b-a)$ et de $\sup\{|f^{(6)}(x)|\ ;\ x \in]a, b[\}$.

Corrigé : Comte-tenu du résultat de la question 5 l'erreur de quadrature E est majorée par :

$$E \leq \frac{b-a}{2}\frac{1}{15750}\sup\{|g^{(6)}(y)|\ ;\ y \in]-1, 1[\}$$

avec $g(y) = f(\frac{a+b}{2} + \frac{b-a}{2}y)$. Comme $g^{(6)}(y) = (\frac{b-a}{2})^6 f^{(6)}(\frac{a+b}{2} + \frac{b-a}{2}y)$ on en déduit la majoration de l'erreur de quadrature sur $[a, b]$:

$$E \leq (\frac{b-a}{2})^7 \frac{1}{15750}\sup\{|f^{(6)}(x)|\ ;\ x \in]a, b[\}$$

8. Avec ces résultats on peut construire une formule de quadrature composée pour un intervalle $[A, B]$ en le découpant en N sous-intervalles de longueur $h = \frac{B-A}{N}$. Donnez alors un majoration de l'erreur de quadrature en fonction de h, $(B-A)$ et

de $\sup\{|f^{(6)}(x)|\ ;\ x \in]A, B[\}$.

Corrigé : L'erreur de quadrature EC de la formule composée est majorée par la somme des majorants des erreurs de quadrature sur chacun des sous-intervalles. Comte-tenu du résultat de la question 7 on a donc :

$$EC \leq \sum_{i=1}^{i=N} (\frac{h}{2})^7 \frac{1}{15750} \sup\{|f^{(6)}(x)|\ ;\ x \in]A, B[\}$$

d'où la majoration, puisque $N = (B - A)/h$:

$$EC \leq (B - A)\, h^6\, C \sup\{|f^{(6)}(x)|\ ;\ x \in]A, B[\}$$

avec $C = 1/(2^7 \times 15750)$.

Problème II

L'objectif du problème est d'étudier une procédure d'accélération de convergence pour une méthode de résolution d'un système linéaire.
On utilisera les polynômes de Chebychev, dont on rappelle la définition et les propriétés :

$$T_n(t) = \cos(n \arccos(t)),\ |t| \leq 1$$

$$T_0(t) = 1, \quad T_1(t) = t, \quad T_{n+1}(t) = 2t\,T_n(t) - T_{n-1}(t), \quad \forall t \in \mathbb{R}$$

Sur $[-1, +1]$ le polynôme T_n, de degré n, atteint $n + 1$ valeurs extrémales $T_n(\xi_i) = (-1)^i$ aux points $\xi_i = \cos(\frac{i\pi}{n})$, $0 \leq i \leq n$.
On désignera dans la suite par $\widetilde{P}_n$ l'ensemble des polynômes p, de degré inférieur ou égal à n, tels que $p(1) = 1$.

1. Soit $0 < a < 1$
   (a) montrez que $\forall n$, $T_n(\frac{1}{a}) \neq 0$.
   On pose :
   $$\widetilde{T}_n(t) = \frac{T_n(\frac{t}{a})}{T_n(\frac{1}{a})}$$
   (b) Montrez que $\widetilde{T}_n \in \widetilde{P}_n$ et déterminez les couples $(t_i, \widetilde{T}_n(t_i))$ avec $t_i \in [-a, a]$ où $|\widetilde{T}_n(t)|$ est maximal.
   (c) Soit $p \in \widetilde{P}_n$ tel que
   $$\max_{-a \leq t \leq a} |p(t)| < \max_{-a \leq t \leq a} |\widetilde{T}_n(t)|$$
   montrez que le polynôme $q = \widetilde{T}_n - p$ s'annule en $n + 1$ points distincts.
   (d) En déduire que :
   $$\inf_{p \in \widetilde{P}_n} \left\{ \max_{-a \leq t \leq a} |p(t)| \right\} = \max_{-a \leq t \leq a} |\widetilde{T}_n(t)|$$

Corrigé :

(a) Comme $1/a > 1$, que les racines des polynôme de Chebychev sont toutes entre -1 et 1, et que $T_n(1) = 1$, alors $T_n(\frac{1}{a}) > 0$.

(b) On vérifie trivialement que $\widetilde{T}_n(1) = 1$ et donc que $\widetilde{T}_n \in \widetilde{P}_n$.
Pour $t \in [-a, a]$, $t/a \in [-1, 1]$, les extrema de $|\widetilde{T}_n(t)|$ dans $[-a, a]$, sont donc les couples $(a\xi_i, (-1)^i/T_n(1/a))$ pour $0 \leq i \leq n$.

(c) Soit $t_i = a\xi_i$, $\widetilde{T}_n(t_i) = (-1)^i \max_{-a \leq t \leq a} |\widetilde{T}_n(t)|$. Compte-tenu de l'hypothèse sur p, $q(t_i)$ est donc du signe de $(-1)^i$. On en déduit que le polynôme $q(t)$ a un zéro entre t_i et t_{i+1} pour $0 \leq i < n$, soit n zéros distincts. De plus par construction $q(1) = 0$ et comme $1 \neq a\xi_i$ (puisque $1/a > 1$ et $|\xi_i| \leq 1$), le polynôme non nul q, de degré n, a donc $n + 1$ zéros distincts ce qui est absurde et contredit l'hypothèse sur le maximum de p.

(d) On a donc démontré que parmi les polynômes de $\widetilde{P}_n$, $\widetilde{T}_n$ était le polynôme ayant le plus petit maximum sur le segment $[-a, a]$.

2. Soit B une matrice réelle $m \times m$ telle que $\rho(B) < 1$, $\rho(B)$ désignant le rayon spectral de B. Pour résoudre le système linéaire, trouver $x \in \mathbb{R}^m$ tel que

$$(I - B)x = c$$

avec c donné dans $\mathbb{R}^m$, on considère la méthode itérative :

$$x^{(k+1)} = Bx^{(k)} + c, \quad x^{(0)} \text{ arbitraire}$$

Démontrez que la suite $(x^{(k)})$ converge vers x.
Corrigé : Posons $\varepsilon^{(k)} = x^{(k)} - x$, l'erreur à l'itération k. On vérifie facilement que $\varepsilon^{(k+1)} = B\varepsilon^{(k)} = B^{k+1}\varepsilon^{(0)}$. Comme $\rho(B) < 1$ la suite $(\varepsilon^{(k)})$ tend vers zéro quel que soit $\varepsilon^{(0)}$. La suite $(x^{(k)})$ converge vers x.

3. Pour accélérer sa convergence on construit, à partir de la suite $(x^{(k)})$ une suite $(y^{(k)})$ de $\mathbb{R}^m$ par :

$$y^{(k)} = \sum_{i=0}^{i=k} \alpha_{ik}\, x^{(i)}$$

où les coefficients réels α_{ik}, $0 \leq i \leq k$, $k \geq 0$ vérifient $\sum_{i=0}^{i=k} \alpha_{ik} = 1$ et sont à déterminer au mieux.
On définit pour cela le polynôme $p_k \in \widetilde{P}_k$ par :

$$p_k(t) = \sum_{i=0}^{i=k} \alpha_{ik}\, t^i$$

et on pose $\varepsilon^{(k)} = x^{(k)} - x, \quad \eta^{(k)} = y^{(k)} - x$.
Montrez que $\varepsilon^{(k)} = B^k\varepsilon^{(0)}$ et $\eta^{(k)} = p_k(B)\varepsilon^{(0)}$.
Corrigé : On a vu que $\varepsilon^{(k)} = B^k\epsilon^{(0)}$. Par ailleurs comme $p_k(1) = 1$ on a $x = \sum_{i=0}^{i=k} \alpha_{ik}x$, on en déduit que $\eta^{(k)} = \sum_{i=0}^{i=k} \alpha_{ik}(x^{(k)} - x) = \sum_{i=0}^{i=k} \alpha_{ik}B^k\varepsilon^{(0)} = p_k(B)\varepsilon^{(0)}$.

4. On suppose désormais que B est une matrice symétrique de valeurs propres notées $\lambda_i,\ 1 \le i \le m$. Montrez que :

$$\|p_k(B)\|_2 = \max_{1 \le i \le m} |p_k(\lambda_i)|$$

Corrigé : La matrice $p_k(B)$, est symétrique comme B, elle a les mêmes vecteurs propres que B avec pour valeurs propres les $p_k(\lambda_i)$. On sait que pour les matrices normales, et en particulier les matrices symétriques, la norme matricielle associée à la norme vectorielle euclidienne est égale au rayon spectral de la matrice, d'où le résultat demandé.

5. On suppose que l'on a l'estimation $\rho(B) \lesssim r < 1$, justifiez alors le choix suivant pour le polynôme p_k :

$$p_k(t) = \frac{T_k(\frac{t}{r})}{T_k(\frac{1}{r})}$$

On montrera en particulier que la suite $(y^{(k)})$ converge vers x et qu'elle converge plus vite que la suite $(x^{(k)})$.

Corrigé : On a vu à la question 1 que parmi les polynômes de degré k, vérifiant $p(1) = 1$, $T_k(t/r)/T_k(1/r)$ était celui dont le maximum sur $[-r, r]$ était le plus petit, ne connaissant pas a priori la position des valeurs propres de B dans cet intervalle, ce polynôme nous assure une valeur minimale pour les $p_k(\lambda_i)$ et ainsi pour $\|p_k(B)\|_2$. Par ailleurs, le polynôme t^k appartient à $\widetilde{T}_k$ et donc $r^k = \max_{-r \le t \le r} t^k \ge \max_{-r \le t \le r} |p_k(t)|$, ce qui prouve que $r^k \approx \|B^k\|_2 \ge \|p_k(B)\|_2$, et donc que la suite $(y^{(k)})$converge vers x plus vite que la suite $(x^{(k)})$.

6. Montrez qu'alors :

$$T_{k+1}(\tfrac{1}{r})\, \eta^{(k+1)} = \frac{2}{r}\, T_k(\tfrac{1}{r})\, B\eta^{(k)} - T_{k-1}(\tfrac{1}{r})\, \eta^{(k-1)}$$

Corrigé : De la relation de récurrence $T_{n+1}(t) = 2t\,T_n(t) - T_{n-1}(t)$ et de la définition de p_k on déduit :

$$T_{k+1}(\tfrac{1}{r})\, p_{k+1}(t) = 2\frac{t}{r} T_k(\tfrac{1}{r})\, p_k(t) - T_{k-1}(\tfrac{1}{r})\, p_{k-1}(t)$$

En appliquant cette formule pour une variable matricielle B et un argument vectoriel $\varepsilon^{(0)}$ on obtient l'identité demandée.

7. Vérifiez que :

$$\frac{2\,T_k(\frac{1}{r})}{r\,T_{k+1}(\frac{1}{r})} = 1 + \frac{T_{k-1}(\frac{1}{r})}{T_{k+1}(\frac{1}{r})}$$

.

Corrigé : Cette formule est une conséquence immédiate de la relation de récurrence des T_n.

8. En déduire que :

$$y^{(k+1)} = \omega_{k+1}\left(By^{(k)} + c - y^{(k-1)}\right) + y^{(k-1)}$$

avec

$$\omega_{k+1} = \frac{2\,T_k(\frac{1}{r})}{r\,T_{k+1}(\frac{1}{r})} = 1 + \frac{T_{k-1}(\frac{1}{r})}{T_{k+1}(\frac{1}{r})}$$

Corrigé : En divisant la relation de la question 6 par $T_{k+1}(\frac{1}{r})$ on obtient :

$$\eta^{(k+1)} = \omega_{k+1}\,B\eta^{(k)} - (\omega_{k+1} - 1)\,\eta^{(k-1)}$$

On trouve alors la relation demandée en remplaçant $\eta^{(k)}$ par $y^{(k)} - x$ et en utilisant $x - Bx = c$.

9. Expliquez comment mettre en œuvre en pratique cette méthode, on explicitera en particulier les différentes étapes de calcul.
Corrigé : Initialisation de l'algorithme - Notons que $p_1(t) = t$, d'où, pour $x^{(0)} = y^{(0)}$ arbitraire, $x^{(1)} = y^{(1)} = Bx^{(0)} + c$, on a également $T_0(\frac{1}{r}) = 1$ et $T_1(\frac{1}{r}) = 1/r$.

Pour passer de l'étape k, à l'étape $k+1$, connaissant $y^{(k-1)}$, $y^{(k)}$, $T_{k-1}(\frac{1}{r})$ et $T_k(\frac{1}{r})$: on calcule $T_{k+1}(\frac{1}{r})$ avec la formule de récurrence des polynômes de Chebychev, on en déduit ω_{k+1}, ce qui permet de calculer $y^{(k+1)}$ avec la formule de la question 8 donc au prix d'une seule multiplication matricielle par la matrice B pour chaque itération.
Il faut pour tout cela estimer $r \approx \rho(B)$. On peut à cet effet utiliser la "méthode de la puissance" qui converge si il y a une seule valeur propre de module maximal qui de plus est de multiplicité unité. Cette méthode consiste à construire la suite : v_0 arbitraire, puis $v_{k+1} = Bv_k/\|B(v_k)\|$. On démontre facilement (en se plaçant sur une base orthonormale de vecteurs propres) que la suite v_k converge vers un vecteur propre normé associé à la valeur propre de plus grand module. Et donc pour k assez grand $\|B(v_k)\| \approx \rho(B)$.