Université Paris-Nord
Institut Galilée
Département de Mathématiques

Année 2009-2010
Licence de Mathématiques - L3

C. Basdevant - F. Cuvelier

Corrigé de l'examen d'Analyse Numérique
du mercredi 1er septembre 2010

Durée : 3h
Aucun document n'est autorisé

Problème I

Soit A une matrice $n \times n$ réelle inversible.

1. Montrez que $C = A^t A$ est symétrique définie positive. En déduire que C admet une décomposition de Cholesky sous la forme $C = B^t B$, avec B triangulaire supérieure et inversible.

   Corrigé : On vérifie d'une part que $C^t = C$, c'est à dire que C est symétrique, et d'autre part, pour $x \neq 0 \in \mathbb{R}^n$, calculons $x^t C x$:

   $$x^t C x = x^t A^t A x = \|Ax\|_2^2$$

   Cette dernière quantité est strictement positive puisque A est inversible. La matrice C est donc symétrique définie positive. D'après un résultat du cours elle admet une décomposition de Cholesky $C = B^t B$ avec B triangulaire supérieure à diagonale strictement positive.

2. Montrez que la matrice $S = AB^{-1}$ est orthogonale.

   Corrigé : $SS^t = AB^{-1}(B^{-1})^t A^t = AC^{-1}A^t = A(A^t A)^{-1}A^t = I$, S est bien une matrice orthogonale.

3. En déduire que la matrice A s'écrit sous la forme $A = QR$, où Q est orthogonale et R triangulaire supérieure à termes diagonaux strictement positifs.

   Corrigé : Comme $S = AB^{-1}$ on en déduit $A = SB$, on a bien mis la matrice A sous la forme QR avec $Q = S$ orthogonale et $R = B$ triangulaire supérieure à termes diagonaux strictement positifs.

4. Montrez que cette décomposition QR est unique.

   Corrigé : Supposons deux décompositions $Q_1 R_1 = Q_2 R_2$, alors $Q_2^{-1} Q_1 = R_2 R_1^{-1}$. Mais la matrice $M = Q_2^{-1} Q_1$ est orthogonale comme produit de matrices orthogonales et la matrice $M = R_1^{-1} R_2$ est triangulaire supérieure à termes diagonaux strictement positifs comme produit de matrices triangulaires supérieures à termes diagonaux strictement positifs. $M^t = M^{-1}$ comme matrice orthogonale, mais M^t est triangulaire inférieure puisque M est triangulaire supérieure et M^{-1} est triangulaire supérieure comme inverse d'une triangulaire supérieure, donc M est diagonale à terme strictement positifs, mais de modules 1 puisque ce sont les valeurs propres d'une matrice orthogonale, M est donc la matrice identité, ce qui prouve que $Q_1 = Q_2$ et $R_1 = R_2$ d'où l'unicité de la décomposition QR d'une matrice inversible.

Problème II

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 :

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

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 III

Soit A une matrice réelle, $n \times n$, symétrique, définie positive.
Pour résoudre le système :

$$Ax = b \;, \qquad b \text{ donné dans } \mathbb{R}^n \tag{1}$$

on considère la méthode itérative suivante pour laquelle x_0 est arbitraire et σ est un paramètre réel :

$$x_{n+1} = x_n - \sigma(Ax_n - b) \tag{2}$$

1. Montrez que la méthode (2) converge vers la solution du système (1) si et seulement si :

$$0 < \sigma < \frac{2}{\rho(A)} \;,$$

   où $\rho(A)$ désigne le rayon spectral de A.

   Corrigé : Si la méthode converge vers une limite y, celle-ci vérifie l'"équation $y = y - \sigma(Ay - b)$, c'est donc bien l'unique solution du problème (1) dès que $\sigma \neq 0$. Pour que la méthode converge il faut et il suffit que le rayon spectral de la matrice d'itération soit inférieur strictement à 1. La matrice d'itération étant $I - \sigma A$ ses valeurs propres sont $1 - \sigma\lambda$ avec λ valeur propre de A ; valeurs propres qui sont par hypothèse réelles strictement positives. On doit donc avoir $-1 < 1 - \sigma\lambda < 1$ soit $0 < \sigma\lambda < 2$ quel que soit λ, d'où le résultat demandé.

2. Montrez que la vitesse de convergence de la méthode (2) est maximale pour une valeur de σ que l'on exprimera en fonction de $\rho(A)$ et $\rho(A^{-1})$.

   Corrigé : La vitesse de convergence de la méthode est d'autant plus grande que le rayon spectral de la matrice d'itération est petit, il faut donc choisir σ qui minimise la quantité $\max_i |1 - \sigma\lambda_i|$ avec λ_i les valeurs propres de A. Cette fonction s'écrit $f(\sigma) = \max\{1 - \sigma\lambda_1, \sigma\lambda_n - 1\}$ avec λ_1 et λ_n respectivement la plus petite et la plus grande valeur propre de A. Son graphe est tracé sur la Figure **??** et elle est est minimale pour $\sigma = \frac{2}{\lambda_1 + \lambda_n}$, soit :

$$\sigma = \frac{2}{\frac{1}{\rho(A^{-1})} + \rho(A)}$$

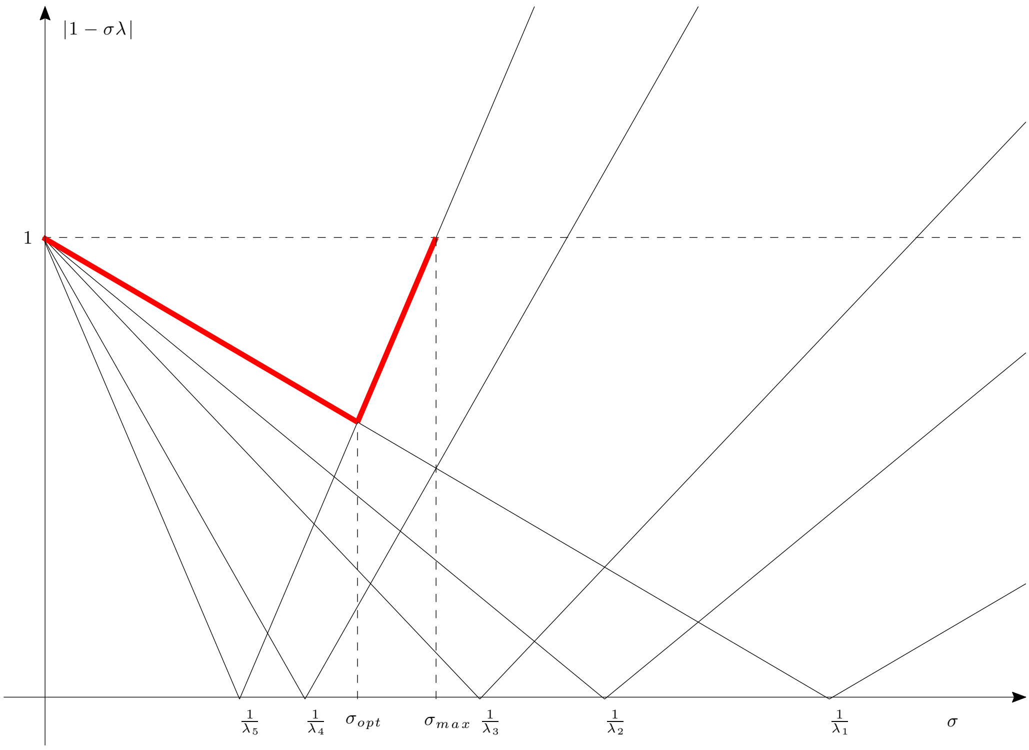

FIG. 1 – $\rho(I - \sigma A)$ en fonction de σ