Ecole Supérieure d'Informatique et Applications de Lorraine – 1ère année

Rédacteurs : Bruno Pinçon & Tony Bourdier
Date : Vendredi 27 avril 2007
Durée : 2 heures

Corrigé examen final – Mathématiques Numériques

Remarques : La rédaction de ce corrigé est succincte : les étapes intermédiaires des calculs n'y figurent pas et les explications sont sommaires. Pour obtenir la totalité des points il était nécessaire d'expliciter en détail vos résultats.

1. Moindres carrés classiques */* 3,5 points */*

(*a*) *(0.5 point)* La méthode des moindres carrés est utilisée lorsque l'on veut **estimer** un caractère par une **combinaison linéaire** de fonctions d'autres caractères. Elle peut servir à l'identification d'un modèle physique dont on ne connaît pas les paramètres (sous réserve que le modèle soit linéaire en ses paramètres). Elle peut également être utile dans le cadre du lissage d'un nuage de points ...

(*b*) *(1 point)* En passant au logarithme, on obtient :

$$\log(y) = \alpha \log(t) + \beta \log(z) + \gamma$$

Le problème se pose donc sous la forme suivante :

$$\min_u ||\, A\, u - v\, ||^2$$

avec

$$A = \begin{pmatrix} \log(t_1) & \log(z_1) & 1 \\ \vdots & \vdots & \vdots \\ \log(t_n) & \log(z_n) & 1 \end{pmatrix}, \quad u = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}, \quad v = \begin{pmatrix} \log(y_1) \\ \vdots \\ \log(y_n) \end{pmatrix}$$

(*c*) *(1 point)* Les équations normales de ce problème sont (sous forme matricielle) :

$$A^T A u = A^T v$$

soit

$$\begin{pmatrix} SC_t & S_{zt} & S_t \\ Szt & SC_z & S_z \\ S_t & S_z & n \end{pmatrix} . \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} S_{yt} \\ S_{yz} \\ S_y \end{pmatrix}$$

(*d*) *(1 point)*

```
function [alpha beta gamma] = pmc(y,z,t)
  A = [log(t) , log(z) , ones(n,1)];
  v = log(y);
  [LU info] = factorisation_lu(A'*A);
  if info == 0
   disp('Erreur lors de la factorisation LU');
  else
   [alpha ; beta ; gamma] = descente_remontee(LU,A'*v);
  end
```

2. Arithmétique flottante

/ 3 points */*

(a) *(1.5 point)* Pour x (resp. y) supérieur à 2.10^{200}, on a x^2 (resp. y^2) supérieur à M. Or $f(x,y) < M$. On est donc en présence d'un **overflow parasite**. On utilise donc cette autre écriture de f pour résoudre ce problème :

$$f : (x,y) \mapsto \begin{cases} |\,x\,|\sqrt{1+\left(\dfrac{y}{x}\right)^2} - y & \text{si } |\,x\,| > |\,y\,| \\ |\,y\,|\sqrt{1+\left(\dfrac{x}{y}\right)^2} - y & \text{sinon} \end{cases}$$

(b) *(1.5 point)* Lorsque x est proche de 0, $\sqrt{x^2+y^2}$ est proche de y. On est donc en présence d'une **soustraction de deux nombres très proches**, ce qui amplifie les erreurs. Pour remédier au problème, on réécrit f de la manière suivante :

$$f : (x,y) \mapsto \frac{x^2}{\sqrt{x^2+y^2}+y}$$

3. Interpolation de Newton

/ 3 points */*

(a) *(0,25 point)* $f[x_0] = f(x_0) = \dfrac{\det(f(x_0))}{1} = \dfrac{det(f(x_0))}{det(1)}$ et

$$f[x_0,x_1] = \frac{f(x_1)-f(x_0)}{x_1-x_0} = \frac{det\begin{pmatrix} 1 & f(x_0) \\ 1 & f(x_1) \end{pmatrix}}{det\begin{pmatrix} 1 & x_0 \\ 1 & x_1 \end{pmatrix}}$$

(b) *(0,5 point)* $\forall i \in \{0,...,n\},\ f(x_i) = p_n(x_i) = a_0 + a_1.x_i + ... + a_n.x_i^n$

(c) *(0.75 point)* Le système de la question précédente peut également s'écrire :

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$$

(d) *(0.5 point)* On déduit de la remarque précédente une expression de a_n :

$$a_n = \frac{det\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & f(x_0) \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & f(x_1) \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & f(x_n) \end{pmatrix}}{det\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & x_n^n \end{pmatrix}}$$

Or, $a_n = f[x_0,...,x_n]$ (coefficient dominant de p_n), ce qui permet de conclure.

(e) *(1 point)* $a_0 = \dfrac{det(f(x_0))}{det(1)} = f(x_0) = 1$

$$a_1 = \frac{\begin{vmatrix} 1 & f(x_0) \\ 1 & f(x_1) \end{vmatrix}}{\begin{vmatrix} 1 & x_0 \\ 1 & x_1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & 1 \\ 1 & -3 \end{vmatrix}}{\begin{vmatrix} 1 & -2 \\ 1 & 0 \end{vmatrix}} = \frac{-4}{2} = 2$$

$$a_2 = \frac{\begin{vmatrix} 1 & -2 & 1 \\ 1 & 0 & -3 \\ 1 & 3 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & -2 & 4 \\ 1 & 0 & 0 \\ 1 & 3 & 9 \end{vmatrix}} = 1$$

Le polynôme d'interpolation dans la base de Newton associé aux points P_0, P_1 et P_2 s'écrit alors :

$$1 - 2(x+2) + (x+2)x$$

4. Système linéaire *`/* 3 points */`*

(a) *(2 points)*

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & -4/7 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 7/3 & 7/3 \\ 0 & 0 & -1 \end{pmatrix}$$

(b) *(1 point)*

$$x = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

5. Moindres carrés dynamiques *`/* 7 points */`*

Partie 1

(a) On a :

$$B^\top = (A + uu^\top)^\top = A^\top + (uu^\top)^\top = A + (u^\top)^\top u^\top = A + uu^\top = B$$

donc B est symétrique.
Soit $x \in \mathbb{R}^n$, $x \neq 0$, il vient :

$$(Bx|x) = (Ax|x) + (uu^\top x|x) = (Ax|x) + x^\top uu^\top x = (Ax|x) + (u|x)^2$$

Il est clair que $(u|x)^2 \geq 0$ [1] et comme $(Ax|x) > 0$ on a bien $(Bx|x) > 0$ pour tout vecteur x non nul. B est donc bien définie positive.

[1] et est nul si x est dans le s.e.v. de dimension $n-1$ des vecteurs orthogonaux au vecteur u.

(b) Il faut vérifier que :

$$B\left(A^{-1} - \frac{A^{-1}uu^\top A^{-\top}}{1+u^\top A^{-\top}u}\right) = I$$

Il suffit juste de développer (en utilisant le fait que $u^\top A^{-\top}u$ est un scalaire et que $A^{-\top} = A^{-1}$) :

$$\begin{aligned}
B\left(A^{-1} - \frac{A^{-1}uu^\top A^{-\top}}{1+u^\top A^{-\top}u}\right) &= (A+uu^\top)(\left(A^{-1} - \frac{A^{-1}uu^\top A^{-\top}}{1+u^\top A^{-\top}u}\right) \\
&= I - \frac{uu^\top A^{-\top}}{1+u^\top A^{-\top}u} + uu^\top A^{-1} - \frac{u(u^\top A^{-1}u)u^\top A^{-\top}}{1+u^\top A^{-\top}u} \\
&= I + \left(-\frac{1}{1+u^\top A^{-\top}u} + 1 - \frac{u^\top A^{-\top}u}{1+u^\top A^{-\top}u}\right) uu^\top A^{-\top} \\
&= I + \left(1 - \frac{1+u^\top A^{-\top}u}{1+u^\top A^{-\top}u}\right) uu^\top A^{-\top} \\
&= I
\end{aligned}$$

(c) avec $y = A^{-1}u$, on a $y^\top = u^\top A^{-\top}$ et le terme de correction s'écrit donc bien :

$$\left(\frac{y}{1+u^\top y}\right) y^\top$$

Donc pour calculer effectivement B^{-1} connaissant A^{-1} et u il faut :
- calculer $y = A^{-1}u$, cad un produit matrice vecteur soit n^2 opérations (n^2 multiplications et $n(n-1)$ additions)
- calculer le produit scalaire $u^\top y$, soit n multiplications et $n-1$ additions
- calculer $1 + u^\top y$ soit une addition
- diviser chaque composante de y par $1 + u^\top y$, soit n divisions (appelons w le vecteur ainsi obtenu) ;
- puis calculer le produit matriciel $wy^\top$ soit n^2 multiplications, appelons *cor* la matrice ainsi obtenue ;
- et enfin $B^{-1} = A^{-1} - cor$ soit n^2 soustractions.

On voit que c'est un calcul en $O(n^2)$ opérations (à peu près $2n^2$ additions et $2n^2$ multiplications) ce qui est beaucoup plus économe que de recalculer l'inverse directement (ce qui coûte environ n^3 opérations).

Partie 2

Question préliminaire : x vérifie $M^\top Mx = M^\top y$

(a) La nouvelle matrice est :

$$\left[\begin{array}{lll}
\phi_1(t_1) & \dots & \phi_n(t_1) \\
\phi_1(t_2) & \dots & \phi_n(t_2) \\
\vdots & \vdots & \vdots \\
\phi_1(t_m) & \dots & \phi_n(t_m) \\
\hline
\phi_1(t_{m+1}) & \dots & \phi_n(t_{m+1})
\end{array}\right] = \left[\begin{array}{c} M \\ \hline l \end{array}\right]$$

et le nouveau vecteur des mesures est :

$$\left[\begin{array}{c} y_1 \\ \vdots \\ y_m \\ \hline y_{m+1} \end{array}\right] = \left[\begin{array}{c} y \\ \hline z \end{array}\right]$$

(b) Pour ce nouveau problème les équations normales sont :

$$\left[\begin{array}{c} M \\ \hline l \end{array}\right]^\top \left[\begin{array}{c} M \\ \hline l \end{array}\right] x_c = \left[\begin{array}{c} M \\ \hline l \end{array}\right]^\top \left[\begin{array}{c} y \\ \hline z \end{array}\right]$$

On a :

$$\left[\begin{array}{c} M \\ \hline l \end{array}\right]^\top = \left[\begin{array}{c|c} M^\top & l^\top \end{array}\right]$$

Les équations normales se réécrivent donc :

$$\left[\begin{array}{c|c} M^\top & l^\top \end{array}\right] \left[\begin{array}{c} M \\ \hline l \end{array}\right] x_c = \left[\begin{array}{c|c} M^\top & l^\top \end{array}\right] \left[\begin{array}{c} y \\ \hline z \end{array}\right]$$

et la formule du produit par bloc nous donne donc :

$$(M^\top M + l^\top l)x_c = M^\top y + l^\top z$$

Comme M est de rang maximum (ses colonnes sont linéairement indépendantes) la matrice $A = M^\top M$ est symétrique et définie positive. En posant $u = l^\top$ on se retrouve donc avec $B(= M^\top M + l^\top l) = A + uu^\top$, et le résultat (a) de la partie 1 s'applique : B est symétrique et définie positive et donc a fortioti inversible.

(c) $B = M^\top M + l^\top l$ étant inversible et comme (cf question préliminaire) $M^\top y = M^\top Mx$, on obtient :

$$\begin{aligned}
x_c &= B^{-1}\left(M^\top Mx + l^\top z\right) \\
&= B^{-1}\left(M^\top Mx + l^\top lx - l^\top lx + l^\top z\right) \\
&= B^{-1}\left((M^\top M + l^\top l)x + l^\top(z - lx)\right) \\
&= B^{-1}Bx + B^{-1}l^\top(z - lx) \\
&= x + K(z - lx)
\end{aligned}$$

Le terme lx (on a $lx = f(t_{m+1}, [x])$) représente ce que donne le modèle à l'instant t_{m+1} avec les paramètres “actuels”, cad le vecteur x. Si la nouvelle mesure $z = y_{m+1}$ est exactement égale à ce que prédit le modèle avec les paramètres x, il n'y a donc aucune correction pour ces derniers.

(d) Pour obtenir la nouvelle correction rapidement connaissant l'inverse de la matrice $M^\top M$ il suffit d'utiliser la formule rapide vue dans la partie 1 et qui permet de calculer rapidement l'inverse de $M^\top M + l^\top l$ (cet inverse doit être enregistré en mémoire pour calculer rapidement l'apport de la prochaine mesure au temps t_{m+2}). Une fois cet inverse calculé il suffit de faire les opérations suivantes :

$$\begin{aligned}
temp &\leftarrow z - lx \\
vtemp &\leftarrow l^\top temp \\
xc &\leftarrow x + B^{-1}vtemp
\end{aligned}$$

6. Classification

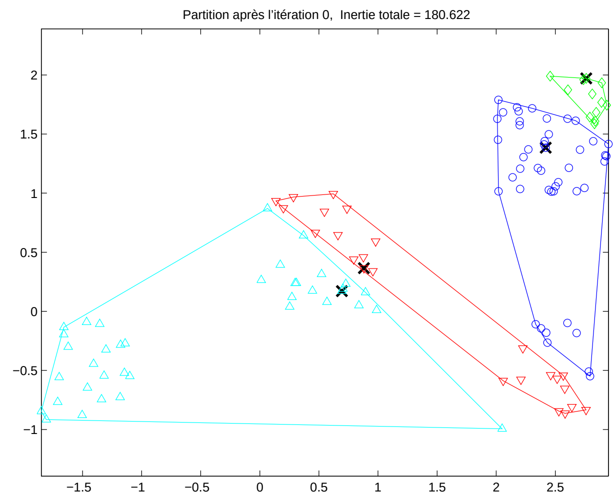
/* 1,5 point */

(*a*) *(0,5 point)* L'intérêt de la classification est de **regrouper** des individus selon des **caractéristiques communes**. Exemples : Sondage (individus : électeurs, caractéristiques : C.S.P., salaires, ...), traitement de l'image (individus : pixels, caractéristiques : niveau de gris compris dans un certain intervalle ou dérivée seconde supérieure à un certain seuil, ...), etc.

(*b*) *(1 point)* Soient n individus à regrouper en K classes. Les H-means procèdent comme suit :

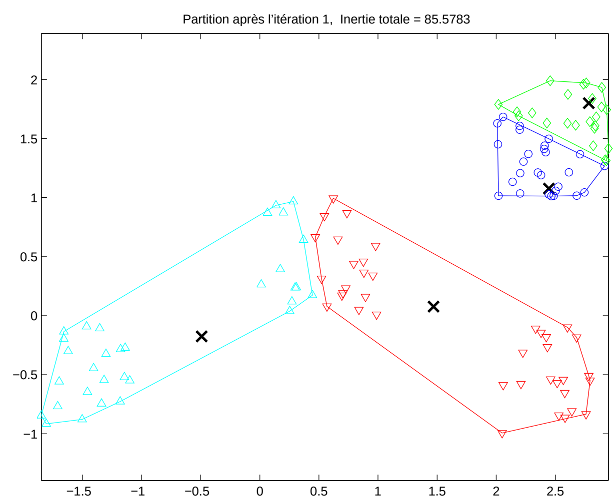
1 $\longrightarrow$ choix (au hasard) de K individus parmi les n qui serviront de **centres aux classes initiales**.

2 $\longrightarrow$ construction de la **partition initiale** : affectation des individus à la classe dont le centre est le plus proche.

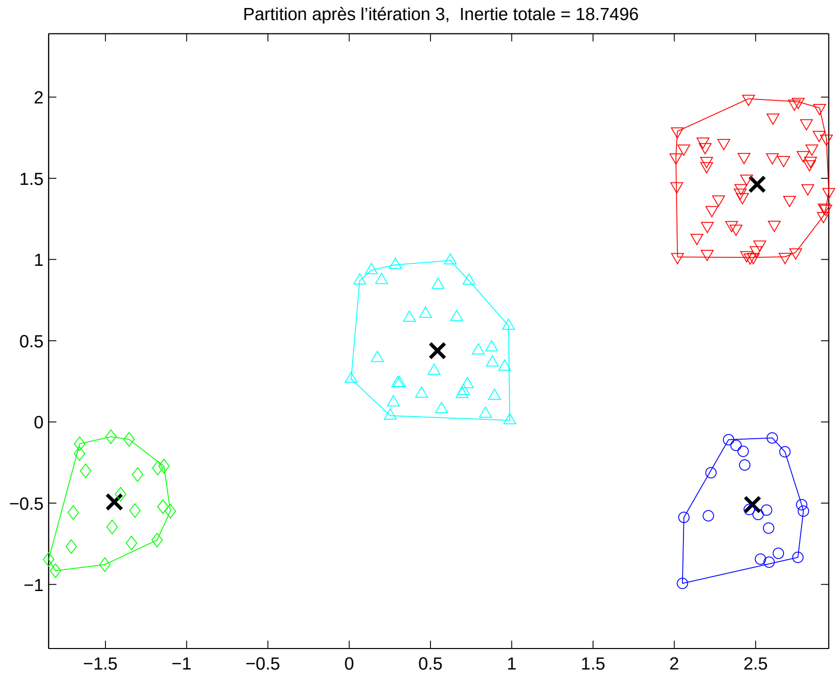


3 $\longrightarrow$ {**Barycentrage**} pour chaque classe, on calcule le nouveau barycentre

4 $\longrightarrow$ {**Affectation**} on réaffecte chaque individu à la classe dont le centre (calculé à l'étape 3) est le plus proche.



5 $\longrightarrow$ On réitère les étapes 3 et 4 jusqu'à **stabilité**, *i.e.* jusqu'à ce que les classes n'évoluent plus.



La partition finale obtenue est optimale **pour la partition initiale**. Il s'agit donc d'un optimum **local**, c'est pourquoi il est souvent nécessaire de réitérer plusieurs fois l'algorithme pour obtenir une partition convenable.