Département MIDO

Méthodes numériques
année 2023/2024

Corrigé (succinct) de l'examen du 29 mai 2024

Exercice 1 (méthode des directions alternées). Soit n un entier naturel non nul et γ un nombre réel strictement positif. Dans la suite, on note $\|\cdot\|_2$ à la fois la norme euclidienne sur $M_{n,1}(\mathbb{R})$ et la norme matricielle subordonnée sur $M_n(\mathbb{R})$ qui lui est associée.

1. **Questions préliminaires.** Soit C une matrice réelle d'ordre n symétrique positive.

   (a) Montrer que la matrice $\gamma I_n + C$ est inversible et que la matrice $D = (\gamma I_n - C)(\gamma I_n + C)^{-1}$ est symétrique.

   La matrice C étant réelle symétrique positive, ses valeurs propres sont positives. Le paramètre γ étant strictement positif, les valeurs propres de la matrice $\gamma I_n + C$, données par $\gamma + \lambda$ avec λ valeur propre de C, sont donc strictement positives. En particulier, 0 ne peut être valeur propre de $\gamma I_n + C$, qui est donc inversible. Pour montrer que la matrice D est symétrique, on utilise que les matrices $\gamma I_n + C$ et $\gamma I_n - C$ commutent entre elles et sont symétriques (car C est symétrique), ce qui implique que $(\gamma I_n + C)^{-1}$ et $\gamma I_n - C$ commutent.

   (b) Montrer que $\|D\|_2 \leq 1$, puis que $\|D\|_2 = 1$ si et seulement si 1 est valeur propre de D.

   La matrice D étant symétrique, on sait que $\|D\|_2 = \rho(D)$, où $\rho(D)$ est le rayon spectral de la matrice D, d'où

   $$\|D\|_2 = \rho(D) = \max_{\lambda \in \mathrm{Sp}(C)} \left| \frac{\gamma - \lambda}{\gamma + \lambda} \right| \leq 1.$$

   Les valeurs propres de D étant réelles, avoir égalité dans l'inégalité ci-dessus implique que -1 ou 1 sont valeurs propres. Cependant, la valeur propre -1 correspondant au cas $\gamma = 0$, cette possibilité se trouve exclue.

   On suppose à présent que 1 est valeur propre de D.

   (c) Montrer que le sous-espace propre de D associé à 1 est un sous-espace vectoriel du noyau de C.

   Soit x un vecteur propre de la matrice D associé à la valeur propre 1. On a $D x = x$, ce qui implique que $(\gamma I_n - C) x = (\gamma I_n + C) x$, soit encore $2C x = 0$. Le vecteur x appartient donc bien à $\mathrm{Ker}(C)$.

   (d) Montrer enfin qu'une matrice x de $M_{n,1}(\mathbb{R})$ telle que $\|x\|_2 = 1$ et $\|Dx\|_2 = 1$ est un vecteur propre de D associé à 1. On pourra pour cela (en le justifiant) décomposer x dans une base orthonormée formée de vecteur propres de D.

   La matrice D étant réelle symétrique, il existe une base orthonormée de $M_{n,1}(\mathbb{R})$, notée $\{e_i\}_{i=1,\dots,n}$ formée de vecteurs propres de D, dont on note les valeurs propres λ_i, $i = 1, \dots, n$. En décomposant le vecteur x dans cette base, on trouve $x = \sum_{i=1}^n x_i\, e_i$ et $D x = \sum_{i=1}^n x_i \frac{\gamma - \lambda_i}{\gamma + \lambda_i} e_i$. On a alors

   $$0 = {\|x\|_2}^2 - {\|Dx\|_2}^2 = \sum_{i=1}^n |x_i|^2 \left(1 - \left|\frac{\gamma - \lambda_i}{\gamma + \lambda_i}\right|^2\right),$$

   par la relation de Pythagore, ce qui implique que, pour tout entier i de $\{1, \dots, n\}$ pour lequel $x_i \neq 0$, on a $\frac{\gamma - \lambda_i}{\gamma + \lambda_i} = 1$, d'où $D x = x$.

Soit A une matrice réelle d'ordre n symétrique définie positive, C_1 et C_2 deux matrices réelles d'ordre n symétriques positives telles que que $A = C_1 + C_2$. On cherche à résoudre le système linéaire $Ax = b$, avec b une matrice de $M_{n,1}(\mathbb{R})$, par la méthode itérative définie par

$$\forall k \in \mathbb{N},\ (\gamma I_n + C_1)\, x^{(k+\frac{1}{2})} = (\gamma I_n - C_2)\, x^{(k)} + b,\ (\gamma I_n + C_2)\, x^{(k+1)} = (\gamma I_n - C_1)\, x^{(k+\frac{1}{2})} + b.$$

l'approximation initiale $x^{(0)}$ dans $M_{n,1}(\mathbb{R})$ étant donnée.

2. Déterminer des matrices B de $M_n(\mathbb{R})$ et c de $M_{n,1}(\mathbb{R})$ telles que

$$\forall k \in \mathbb{N},\ x^{(k+1)} = B\, x^{(k)} + c.$$

Les matrices C_1 et C_2 étant réelles symétriques positives, on sait d'après la première question que les matrices $(\gamma I_n + C_1)$ et $(\gamma I_n + C_2)$ sont inversibles. En multipliant à gauche la première égalité par $(\gamma I_n + C_1)^{-1}$, en substituant l'expression de $x^{(k+\frac{1}{2})}$ ainsi obtenue dans la seconde égalité et en multipliant à gauche l'égalité résultante par $(\gamma I_n + C_2)^{-1}$, on trouve, après identification,

$$B = (\gamma I_n + C_2)^{-1}(\gamma I_n - C_1)(\gamma I_n + C_1)^{-1}(\gamma I_n - C_2) \text{ et } c = (\gamma I_n + C_2)^{-1}(I_n + (\gamma I_n - C_1)(\gamma I_n + C_1)^{-1})b.$$

3. On pose $D_1 = (\gamma I_n - C_1)(\gamma I_n + C_1)^{-1}$ et $D_2 = (\gamma I_n - C_2)(\gamma I_n + C_2)^{-1}$. Montrer que

$$\rho(B) \leq \|D_1 D_2\|_2 \leq 1,$$

où $\rho(B)$ désigne le rayon spectral de la matrice B.

On remarque qu'on a $(\gamma I_n + C_2)B(\gamma I_n + C_2)^{-1} = D_1 D_2$. Les matrices B et $D_1 D_2$ sont donc semblables, elles ont le même spectre et donc le même rayon spectral. Par conséquent, on a

$$\rho(B) = \rho(D_1 D_2) \leq \|D_1 D_2\|_2 \leq \|D_1\|_2 \|D_2\|_2 \leq 1,$$

la dernière inégalité découlant de ce qui a été établi dans les questions préliminaires.

4. On souhaite à présent montrer que $\rho(B) < 1$. On raisonne pour cela par l'absurde en supposant que $\rho(B) = 1$.
   (a) Montrer qu'il existe une matrice y de $M_{n,1}(\mathbb{R})$ telle que $\|D_1 D_2\|_2 = \|D_2 y\|_2 = \|y\|_2 = 1$.

   Si $\rho(B) = 1$, alors $\|D_1 D_2\|_2 = 1$ d'après les précédentes inégalités. Par définition d'une norme subordonnée, cela implique qu'il existe un vecteur y tel que $\|y\|_2 = 1$ et $\|D_1 D_2 y\|_2 = 1$ On remarque alors que cela entraîne que $\|D_2 y\| = 1$, car $1 = \|D_1 D_2 y\|_2 \leq \|D_1\|_2 \|D_2 y\|_2 \leq \|D_2 y\|_2 \leq \|D_2\|_2 \|y\|_2 \leq 1$.

   (b) Que dire alors de $\|D_1 y\|_2$?

   Puisque $\|D_2 y\|_2 = \|y\|_2 = 1$, on déduit des questions préliminaires que y est un vecteur propre de la matrice D_2 associé à la valeur propre 1. Ainsi, on a $\|D_1 D_2 y\|_2 = \|D_1 y\|_2 = 1$ et y est aussi un vecteur propre de la matrice D_1 associé à la valeur propre 1.

   (c) En déduire que y appartient à $\mathrm{Ker}(C_1) \cap \mathrm{Ker}(C_2)$ et obtenir une contradiction.

   On a trouve un vecteur y non nul appartenant à la fois au noyau de C_1 et à celui de C_2 d'après les questions préliminaires. Ce vecteur appartient alors au noyau de la matrice $A = C_1 + C_2$, contredisant le fait que cette dernière est définie positive.

5. Montrer enfin que la suite $(x^{(k)})_{k\in\mathbb{N}}$ construite par la méthode est convergente, de limite égale à la solution du système linéaire $Ax = b$.

Le rayon spectral de la matrice d'itération de la méthode étant strictement inférieur à 1, on sait que la suite $(x^{(k)})_{k\in\mathbb{N}}$ qu'elle construit est convergente. Notons x sa limite. Celle-ci satisfait, par passage à la limite dans les relations de récurrence définissant la méthode,

$$(\gamma I_n + C_1)x = (\gamma I_n - C_2)x + b \iff (C_1 + C_2)x = b.$$

Exercice 2 (une caractérisation du polynôme d'interpolation de Lagrange). Soit n un entier naturel. Montrer que le polynôme d'interpolation de Lagrange d'une fonction polynomiale p est le reste de de la division euclidienne de p par $\omega_{n+1}(x) = \prod_{i=0}^{n}(x - x_i)$, où les réels $x_0, \dots, x_n$ sont les nœuds d'interpolation.
La division euclidienne de p par ω_{n+1} s'écrit

$$\forall x \in \mathbb{R},\ p(x) = q(x)\omega_{n+1}(x) + r(x),$$

où q (le quotient) et r (le reste) sont des fonctions polynomiales, de degré strictement inférieur à $n+1$ pour r. Il suffit donc de montrer que r satisfait les contraintes d'interpolation. En remarquant que les nœuds d'interpolation sont des zéros de ω_{n+1}, on a

$$\forall i \in \{0, \dots, n\},\ p(x_i) = q(x_i)\omega_{n+1}(x_i) + r(x_i) = r(x_i),$$

ce qui permet de conclure, par unicité du polynôme d'interpolation de Lagrange.

Exercice 3 (formules de quadrature avec correction aux extrémités). Soit $[a,b]$ un intervalle d'intérieur non vide et borné de $\mathbb{R}$ et f une fonction à valeurs réelles de classe $\mathscr{C}^1$ sur $[a,b]$. Pour l'approximation de l'intégrale $I(f) = \int_a^b f(x)\,\mathrm{d}x$, on considère la *règle du point milieu avec correction aux extrémités* donnée par :

$$I_{1,\mathrm{cor}}(f;a,b) = (b-a)f\left(\frac{a+b}{2}\right) + \frac{(b-a)^2}{24}\left(f'(b) - f'(a)\right).$$

1. Rappeler la valeur du degré d'exactitude de la règle du point milieu et déterminer celle de cette version modifiée.
Indication : on pourra considérer le cas $a=-1$ et $b=1$ pour obtenir un contre-exemple à l'exactitude de la formule.

La règle du point milieu a un degré d'exactitude égal à 1. On remarque ensuite que la formule corrigée coïncide avec la règle du point milieu pour les fonctions polynomiales $x\mapsto 1$ et $x\mapsto x$. On en déduit que son degré d'exactitude est au moins égal à 1 et on la teste avec des fonctions monomiales de degré supérieur. Pour $x\mapsto x^2$, on a

$$I_{1,\text{cor}}(x\mapsto x^2;a,b)=(b-a)\left(\frac{a+b}{2}\right)^2+\frac{(b-a)^2}{24}(2b-2a)=\frac{1}{3}(b^3-a^3),$$

et le degré d'exactitude est donc au moins égal à 2. Pour $x\mapsto x^3$, on a

$$I_{1,\text{cor}}(x\mapsto x^3;a,b)=(b-a)\left(\frac{a+b}{2}\right)^3+\frac{(b-a)^2}{24}(3b^2-3a^2)=\frac{1}{4}(b^4-a^4),$$

et le degré d'exactitude est donc au moins égal à 3. En revanche, pour $x\mapsto x^4$, $a=-1$ et $b=1$, on a

$$I_{1,\text{cor}}(x\mapsto x^4;-1,1)=\frac{1}{6}(4+4)=\frac{2}{3},$$

qui n'est pas égal à la valeur de l'intégrale correspondante, c'est-à-dire $\frac{2}{5}$. Le degré d'exactitude de la formule est donc égal à 3.
La règle de Simpson peut être corrigée de façon similaire pour obtenir une formule dont le degré d'exactitude vaut cinq, moyennant cependant un changement dans les poids de quadrature de la formule originelle. On pose alors :

$$I_{2,\text{cor}}(f;a,b)=\frac{b-a}{2}\left(\alpha f(a)+\beta f\left(\frac{a+b}{2}\right)+\alpha f(b)\right)+\frac{(b-a)^2}{4}\gamma\left(f'(b)-f'(a)\right).$$

où α, β et γ sont les poids de la formule dans le cas $a=-1$ et $b=1$.

2. Déterminer les réels α, β et γ.

La règle corrigée étant exacte pour toute fonction polynomiale de degré inférieur ou égal à 5, on obtient des conditions sur les coefficients α, β et γ en considérant successivement les fonctions $x\mapsto 1$, $x\mapsto x$, $x\mapsto x^2$, $x\mapsto x^3$, $x\mapsto x^4$ et $x\mapsto x^5$. Pour des raisons de symétrie dans la formule et de l'intervalle $[-1,1]$, seules les fonctions de degré pair fournissent des conditions utiles. On arrive ainsi au système linéaire

$$\begin{cases}2\alpha+\beta=2\\ 2\alpha+4\gamma=\dfrac{2}{3}\\ 2\alpha+8\gamma=\dfrac{2}{5}\end{cases}$$

d'où on tire que $\alpha=\frac{7}{15}$, $\beta=\frac{16}{15}$ et $\gamma=-\frac{1}{15}$.

3. Montrer qu'une formule de quadrature composée associée à l'une ou l'autre des formules de quadrature corrigées ci-dessus ne nécessite d'évaluer la dérivée de la fonction f qu'aux extrémités de l'intervalle d'intégration.

Soit m un entier naturel non nul. On considère une partition de l'intervalle $[a,b]$ en m sous-intervalles $[x_{j-1},x_j]$, $j=1,\dots,m$, de longueur $\frac{b-a}{m}$. Les formules de quadrature corrigées composées sont alors respectivement données par $\sum_{j=1}^m I_{1,\text{cor}}(f;x_{j-1},x_j)$ et $\sum_{j=1}^m I_{2,\text{cor}}(f;x_{j-1},x_j)$, qui font toutes deux intervenir la somme télescopique $\sum_{j=1}^m (f'(x_j)-f'(x_{j-1}))=f'(x_m)-f'(x_0)=f'(b)-f'(a)$.

Exercice 4 (méthode de la puissance avec décalage). Soit A une matrice d'ordre n diagonalisable dont les valeurs propres λ_i, $i=1,\dots,n$, sont réelles et distinctes, numérotées de manière à avoir

$$\lambda_1<\lambda_2<\dots<\lambda_n.$$

Une variante de la méthode de la puissance consiste à appliquer cette dernière à la matrice $A-\mu I_n$, avec μ un réel, dont le spectre est constitué des valeurs propres de A décalées du facteur $-\mu$. On parle alors de *méthode de la puissance avec décalage*. On cherche ici à utiliser cette technique pour accélérer la convergence de la méthode de la puissance en choisissant judicieusement la valeur du paramètre μ.

1. Rappeler la quantité dont dépend la vitesse de convergence de la méthode de la puissance.

La vitesse de convergence de la méthode de la puissance dépend directement du module du rapport entre la valeur propre sous-dominante et la valeur propre dominante.
On fixe la valeur du paramètre μ.

2. En étudiant les valeurs propres de la matrice $A-\mu I_n$, déterminer quelle(s) valeur(s) propre(s) de la matrice A peu(ven)t être approchée(s) en utilisant la méthode de la puissance avec décalage.

La méthode va permettre d'approcher la valeur propre dominante de la matrice $A-\mu I_n$, c'est-à-dire $\lambda_j-\mu$ où l'entier naturel j est tel que $\left|\lambda_j-\mu\right| = \max\{|\lambda_i-\mu| \mid i\in\{1,\dots,n\}\}$. On voit ainsi qu'on pourra approcher la valeur propre de A la plus éloignée de μ, c'est-à-dire λ_1 si $\mu > \frac{\lambda_1+\lambda_n}{2}$ ou λ_n si $\mu < \frac{\lambda_1+\lambda_n}{2}$.

On suppose à présent que le paramètre μ est choisi de façon à ce que $\lambda_n-\mu$ soit la plus grande valeur propre en valeur absolue de la matrice $A-\mu I_n$.

3. Montrer que la deuxième plus grande valeur propre en valeur absolue est donnée soit par $\lambda_{n-1}-\mu$, soit par $\lambda_1-\mu$.
Indication : on pourra considérer la position du réel μ par rapport au milieu de l'intervalle $]\lambda_1,\lambda_{n-1}[$, en faisant par exemple un dessin.

Pour que $\lambda_n-\mu$ soit la valeur propre dominante de $A-\mu I_n$, on doit nécessairement avoir que $\mu < \frac{\lambda_1+\lambda_n}{2}$. Par ailleurs, la valeur propre sous-dominante de $A-\mu I_n$ sera $\lambda_1-\mu$ si $\frac{\lambda_1+\lambda_{n-1}}{2} < \mu < \frac{\lambda_1+\lambda_n}{2}$, alors que ce sera $\lambda_{n-1}-\mu$ si $\mu < \frac{\lambda_1+\lambda_{n-1}}{2}$.

4. En déduire que la méthode de la puissance avec décalage permet dans ce cas d'approcher λ_n et qu'elle converge le plus rapidement possible pour le choix
$$\mu = \frac{\lambda_1+\lambda_{n-1}}{2}.$$

La méthode permet d'approcher λ_n si $\lambda_n-\mu$ est la valeur propre dominante, ce qui est le cas par hypothèse. La vitesse de convergence est alors liée à la quantité $\max\left\{\left|\frac{\lambda_1-\mu}{\lambda_n-\mu}\right|,\left|\frac{\lambda_{n-1}-\mu}{\lambda_n-\mu}\right|\right\}$, qui est minimale lorsque $|\lambda_1-\mu| = |\lambda_{n-1}-\mu|$, c'est-à-dire pour $\mu = \dfrac{\lambda_1+\lambda_{n-1}}{2}$.

Exercice 5 (calcul des différences divisées pour la forme de Newton du polynôme d'interpolation de Lagrange). Soit n un entier naturel non nul. Étant donné $n+1$ nœuds d'interpolation $x_0,\dots,x_n$ deux à deux distincts et $n+1$ valeurs $y_0,\dots,y_n$, les différences divisées associées peuvent être calculées de la manière suivante :

$$[x_i]y = y_i,\ i=0,\dots,n,$$
$$[x_{i-k},\dots,x_i]y = \frac{[x_{i-k+1},\dots,x_i]y-[x_{i-k},\dots,x_{i-1}]y}{x_i-x_{i-k}},\ i=0,\dots,n, k=0,\dots,i.$$

1. Rappeler quelles sont les différences divisées intervenant dans la forme de Newton du polynôme d'interpolation de Lagrange associé aux couples (x_i,y_i), $i=0,\dots,n$.

Les différences divisées intervenant dans la forme de Newton du polynôme d'interpolation de Lagrange sont $[x_0]y, [x_0,x_1]y,\dots,[x_0,\dots,x_n]y$.

2. Écrire une fonction `newton_lagrange_coefficients`, ayant pour arguments d'entrée des tableaux contenant les nœuds et valeurs d'interpolation et renvoyant un tableau contenant les coefficients de la forme de Newton du polynôme d'interpolation de Lagrange correspondant, calculés et stockés en exploitant pleinement le procédé de calcul récursif donné ci-dessus.

La formule de récurrence donne lieu à un tableau triangulaire de différences divisées, dans lequel les différences intervenant dans la forme de Newton se trouvent sur la diagonale. Ceci peut-être exploité de manière à n'utiliser qu'un tableau à une seule dimension pour le calcul et le stockage des différences divisées d'intérêt. Pour ce faire, il suffit de calculer successivement les colonnes du tableau traingulaire en commençant par le dernier élément. On obtient le code suivant.

```
def newton_lagrange_coefficients(x,y):
    m=len(x)
    coefficients=np.copy(y)
    for j in range(1,m):
        for i in range(m-1,j-1,-1):
            coefficients[i]=(coefficients[i]-coefficients[i-1])/(x[i]-x[i-j])
    return coefficients
```