Interrogation n°22 bis. Corrigé

Exercice A

1) u_n est bien définie car $\deg u(f) \leq \deg f$.

On a $u(f)(x) = (x^2-1)f''(x) + 4xf'(x) + 2f(x)$.

On a $u(f)(x^k) = k(k-1)(x^k - x^{k-2}) + 4kx^k + 2x^k = (k^2+3k+2)x^k - k(k-1)x^{k-2}$.

La matrice de u_n dans la base canonique $(1, x, ..., x^n)$ est une matrice triangulaire supérieure dont les coefficients diagonaux sont les $\lambda_k = k^2+3k+2 = (k+1)(k+2)$, avec $0 \leq k \leq n$.

Comme les λ_k sont distincts, u_n est diagonalisable, et les valeurs propres de u_n sont les λ_k, avec $0 \leq k \leq n$.

2) a) Supposons $f(x) = \sum_{n=0}^{+\infty} c_n x^n$ de rayon de convergence $R > 0$.

On a $\forall x \in]-R, R[$, $u(f)(x) = \sum_{n=0}^{+\infty}(n^2+3n+2)c_n x^n - \sum_{n=0}^{+\infty}(n+1)(n+2)c_{n+2}x^{n+2}$.

Donc $u(f) = \lambda f$ ssi $\forall n \in \mathbb{N}$, $(n^2+3n+2-\lambda)c_n = (n+1)(n+2)c_{n+2}$, ssi $\forall n \in \mathbb{N}$, $c_{n+2} = \dfrac{(n+1)(n+2)-\lambda}{(n+1)(n+2)}c_n$.

b) Supposons $c_n \neq 0$ pour tout n. On a alors : $\lim_{n\to+\infty} \dfrac{c_{n+2}x^{n+2}}{c_n x^n} = x^2$.

Donc les séries $\sum c_{2k}x^{2k}$ et $\sum c_{2k+1}x^{2k+1}$ si elles ne sont pas polynomiales sont de rayon de convergence $R = 1$.

Si $\lambda = (n+1)(n+2)$, alors la solution $\sum_{k\leq 2n} c_{n-2k}x^{n-2k}$ est polynomiale.

c) Par a), les solutions sur $]-1,1[$ sont de la forme $c_0\varphi(x) + c_1\psi(x)$, avec φ et ψ solutions non nulles respectivement paire et impaire. donc $F = \mathrm{Vect}(\varphi, \psi)$ est un sev de $C^\infty(]-1,1[, \mathbb{R})$ de dimension 2.

On a $c_{n+2} = \left(1 + O\left(\dfrac{1}{n^2}\right)\right)c_n$. La série $\sum \ln \dfrac{c_{n+2}}{c_n}$ converge.

Donc $(c_{2k})_{k\in\mathbb{N}}$ et $(c_{2k+1})_{k\in\mathbb{N}}$ convergent vers des valeurs non nulles (si c_0 et c_1 non nuls).

Donc φ et ψ divergent en 1 et -1. Comme elles sont respectivement paire et impaire, il en est de même de toute combinaison linéaire non nulle de φ et ψ.

Donc si λ n'est pas de la forme $(n+1)(n+2)$, avec $n \in \mathbb{N}$, la seule solution sur $[-1,1]$ est la fonction nulle.

Exercice B

1) La fonction ln est concave sur $]0, +\infty[$, car $\ln'' < 0$.

Donc pour x et $y > 0$, $\ln\left(\dfrac{1}{p}x^p + \dfrac{1}{q}y^q\right) \geq \dfrac{1}{p}\ln(x^p) + \dfrac{1}{q}\ln(y^q) = \ln(xy)$.

Comme exp est croissante, on obtient $xy \leq \dfrac{1}{p}x^p + \dfrac{1}{q}y^q$. L'inégalité est immédiate si $x = 0$ ou $y = 0$.

2) Si X et Y est presque sûrement nulle, il en est de même de XY et on a $E(XY) = 0$.

Supposons donc que ce n'est pas le cas. Alors $E(X^p) > 0$ et $E(Y^q) > 0$.

Par 1), on a $\dfrac{X}{E(X^p)^{1/p}}\,\dfrac{Y}{E(Y^q)^{1/q}} \leq \dfrac{1}{p}\dfrac{X^p}{E(X^p)} + \dfrac{1}{q}\dfrac{Y^q}{E(Y^q)}$.

Par linéarité et croissance de l'espérance, on obtient bien $\dfrac{E(XY)}{E(X^p)^{1/p}E(Y^q)^{1/q}} \leq \dfrac{1}{p} + \dfrac{1}{q} = 1$.

Exercice C

1) On a $Y_v = \langle v, X\rangle = \sum_{i=1}^n v_i X_i$, donc par linéarité, $E(Y_v) = 0$.

Remarque : Autrement dit, par linéarité de la forme linéaire E, on a : $E(\langle v, X\rangle) = \langle v, E(X)\rangle$.

On a $(Y_v)^2 = \sum_{i=1}^n \sum_{i=1}^n v_i v_j X_i X_j$, donc $E((Y_v)^2) = \sum_{i=1}^n \sum_{i=1}^n v_i v_j E(X_i X_j)$, et $E(X_i X_j) = \mathrm{Cov}(X_i, X_j)$.

2) On a $\langle v, Mv\rangle = \sum_{i=1}^n \sum_{i=1}^n v_i v_j \operatorname{Cov}(X_i, X_j) = E((Y_v)^2) \geq 0$, d'où le résultat.

3) En décomposant v dans une base orthonormée de vecteurs propres, on a (cf Hausdorffien) :

$\max(\operatorname{Sp}(M)) = \sup_{\|v\|=1} \ \langle v, Mv\rangle$, d'où le résultat par 2).

4) On a par un calcul analogue $E(Y_v Y_w) = \sum_{i=1}^n \sum_{i=1}^n v_i w_j E(X_i X_j) = \langle v, Mw\rangle$.

Ainsi, $E(Y_v Y_w) = 0$ lorsque $\langle v, Mw\rangle = 0$.

Ce qui est réalisé si on choisit une BON $(v_1, ..., v_n)$ de vecteurs propres de M : $\langle v_i, Mv_j\rangle = \lambda_j \delta_{ij}$.

Exercice D. Minimisation de distance quadratique

1. a) Lorsque X décrit $\mathbb{R}^p$, AX décrit $F = \operatorname{Im} A$. Donc $\inf f$ est le carré de la distance de B à $\operatorname{Im} A$.

Notons Y le projeté orthogonal de B sur $\operatorname{Im} A$.

Alors $f(X_0) = \inf f$ ssi $AX_0 = Y$, et par définition de $\operatorname{Im} A$, il existe un tel X_0.

La solution X_0 est unique ssi A est injective, c'est-à-dire $\operatorname{Ker} A = \{0\}$.

b) Soit $X \in \mathbb{R}^p$. Alors $f(X) = \inf f$ ssi $AX_0 = Y$ donc ssi $\forall V \in (\operatorname{Im} A)$, $\langle V, B - AX\rangle = 0$.

Comme $\operatorname{Im} A = \{AZ,\ Z \in \mathbb{R}^p\}$, on obtient la CNS : $\forall V \in (\operatorname{Im} A)$, $\langle AZ, B - AX\rangle = 0$.

Mais $\langle AZ, Y - AX\rangle = \left\langle Z, A^T B - A^T AX\right\rangle$. D'où la CNS : $\forall Z \in \mathbb{R}^p$, $\left\langle Z, A^T B - A^T AX\right\rangle = 0$.

Ce qui équivaut à $A^T B - A^T AX = 0$, car le vecteur nul est le seul vecteur orthogonal à tout vecteur.

2. a) *Première méthode* : on calcule le gradient par les dérivées partielles selon les coordonnées de X :

$f(X) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j - b_i\right)^2$, donc $\dfrac{\partial f}{\partial x_j}(X) = \sum_{i=1}^n 2a_{ij} \left(\sum_{j=1}^n a_{ij} x_j - b_i\right)$.

On reconnaît dans $\sum_{i=1}^n 2a_{ij} \left(\sum_{j=1}^n a_{ij} x_j - b_i\right)$ le j-ième coefficient de $2A^T(AX - B)$.

Donc $\operatorname{grad} f(X) = 2A^T(AX - B)$.

Seconde méthode : on commence par déterminer le développement limité à l'ordre 1 :

On pose $Y = AX - B$. On a $f(X) = \langle Y, Y\rangle$; où $\langle\ ,\ \rangle$ désigne le produit scalaire canonique dans $\mathbb{R}^n$.

On a $f(X + H) = \langle Y + AH, Y + AH\rangle = f(X) + 2\langle AH, AY\rangle + \langle AH, AH\rangle$.

On a donc $\boxed{f(X + H) = f(X) + \left\langle H, 2A^T Y\right\rangle + \langle AH, AH\rangle}$.

On en conclut que $\operatorname{grad} f(X) = 2A^T Y$, c'est-à-dire $\operatorname{grad} f(X) = 2A^T(AX - B)$.

b) Supposons que f admette un minimum en X. Comme $\mathbb{R}^p$ est ouvert, $\operatorname{grad} f(X) = \overrightarrow{0}$.

Donc on a bien $A^T AX = A^T B$.

c) Résulte de la seconde preuve de a).

En effet, avec X tel que $\operatorname{grad} f(X) = \overrightarrow{0}$, on a $\forall H \in \mathbb{R}^p$, $f(X + H) = f(X) + \langle AH, AH\rangle \geq f(X)$.

Exercice E. Décomposition de Cholesky et d'Iwasawa par les matrices de Givens

1) Le coefficient d'indice $(1, 2)$ de $R_\theta A$ est $b\cos\theta - d\sin\theta$.

Si $d \neq 0$, on choisit $\theta = \arctan(\frac{b}{d})$. Ainsi, $\tan(\theta) = \frac{b}{d}$, et $b\cos\theta - d\sin\theta = 0$.

Si $d = 0$, la matrice $R_{\pi/2} = \left(\begin{array}{c|c} 0 & -1 \\ \hline 1 & 0 \end{array}\right)$ convient.

2) Important : Multiplier par $G_{i,j,\theta}$ à gauche revient à opérer sur les lignes L_i et L_j, en les remplaçant respectivement par $(\cos\theta)L_i - (\sin\theta)L_j$ et $(\sin\theta)L_i + (\cos\theta)L_j$.

On montre par récurrence sur i qu'il existe un produit de matrices de Givens R tel que sur les i premières lignes de RA, les coefficients situés à droite des coefficients diagonaux sont nuls.

La propriété est immédiate pour $i = 0$. Supposons la propriété vraie au rang $i - 1$, avec $1 \leq i \leq n$.

Ainsi, il ,existe R telle que $RA = \left(\begin{array}{c|c} L & O \\ \hline * & B \end{array}\right)$, où L triangulaire inférieure et d'ordre $(i - 1)$.

La matrice B est inversible.

On applique alors le même principe qu'au 1) a) pour chaque $j \in \{i + 1, .., n\}$.

On obtient ainsi une nouvelle matrice $R'RA$ dont la i-ième colonne vérifie les propriétés requises, car B est inversible. Par récurrence, on en déduit que la propriété est vraie au rang n, c'est-à-dire RA triangulaire supérieure.

Remarque : Il s'agit d'un principe analogue à celui utilisé dans la méthode du pivot de Gauss : en opérant sur les lignes (c'est-à-dire en multipliant à gauche par des matrices élémentaires), on trasnforme toute matrice en une matrice angulaire supérieure (cf résolution des systèmes linéaires).

3) a) Les matrices de Givens sont des matrices de rotation (car diagonales par blocs avec un bloc égal à I_{n-2} et un bloc égal à une matrice de rotation dans un plan). Un produit de matrices de rotation est a fortiori une matrice orthogonale U. Par 2) appliqué à P, il existe une matrice dont $U \in O_n(\mathbb{R})$ produit de matrices de rotations telle que $UP = L$.

Donc $L^T L = P^T U^T U P = P^T P = S$.

b) Par 2) appliqué à $(A^T)^{-1}$, il existe $U \in O_n(\mathbb{R})$ telle que $U(A^T)^{-1} = L$ triangulaire inférieure inversible (car A et U inversibles). Donc $A = (U^T)^{-1}(L^T)^{-1} = UT$, avec T triangulaire supérieure inversible (car l'inverse d'une matrice triangulaire supérieure inversible est triangulaire supérieure inversible).