Master 2 informatique

Parcours IAAA, UE MIA

Feuille d'exercices numéro 3

Année 2024-25

Algèbre linéaire.

1 Concepts fondamentaux en algèbre linéaire

Exercice 1 Fonction linéaire. ($\star$)

On considère la fonction

$$f : \left| \begin{array}{ccc} \mathbf{R}^2 & \to & \mathbf{R}^2 \\ x, y & \mapsto & x+y, x-y \end{array} \right.$$

Montrez que f est linéaire.

Exercice 2 Espaces vectoriels fonctionnels. ($\star\star$)

1. Proposez une notion d'addition et de produit scalaire naturelle sur l'espace $C^0([a,b], \mathbf{R})$ des fonctions continue d'un segment $[a,b]$ de $\mathbf{R}$ vers $\mathbf{R}$
2. Prouvez que les notions d'addition et de produit scalaire proposées munissent bien de $C^0([a,b], \mathbf{R})$ d'une structure d'espace vectoriel.
3. On considère la fonction e_x qui évalue une fonction continue en un point $x \in \mathbf{R}$:

$$e_x : \left| \begin{array}{ccc} C^0([a,b], \mathbf{R}) & \to & \mathbf{R} \\ f & \mapsto & f(x) \end{array} \right.$$

Prouvez que e_x est linéaire.

Exercice 3 Notions élémentaires. ($\star$)

1. Décrivez en language naturel l'espace vectoriel engendré par les vecteurs $(1,1,0)$ et $(3,3,0)$ de $\mathbf{R}^3$.
2. Quelle est la dimension de cet espace ?
3. Les vecteurs $(3,2)$ et $(6,4)$ de $\mathbf{R}^2$ sont-ils linéairement indépendants ?
4. Qu'en est-il des vecteurs $(3,2,1)$ et $(6,4,0)$ de $\mathbf{R}^3$?
5. On considère les vecteurs de $\mathbf{R}^3$ $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ et $e_3 = (0,0,1)$. La famille (e_1, e_2, e_3) est-elle libre ?

6. Est-ce que (e_1, e_2, e_3) est une base de $\mathbf{R}^3$?

7. Est-ce que de $\mathbf{R}^3$ est de dimension finie ? Si oui, quelle est sa dimension ?

8. Démontrer par l'absurde le théorème sur la dimension d'un espace vectoriel vu en cours.

Exercice 4 Représentation matricielle des applications linéaires. $(\star)$

1. Donnez la matrice de l'application $f : x, y \mapsto y, 3x + y$ dans la base canonique de $\mathbf{R}^2$.

2. Utilisez cette matrice pour trouver $f(6, 17)$.

3. A l'aide d'un produit matriciel, obtenez la matrice de la composée de f avec $g : x, y \mapsto x+y$.

4. Même question pour la composée de f avec $h_y : x \mapsto x + y$.

Exercice 5 Changement de système de coordonnées. $(\star)$

1. Montrer que $B = (3e_1 - 2e_2, e_1 + e_2)$ forme une base de $\mathbf{R}^2$.

Solution : $\mathbb{R}^2$ est un ($\mathbb{R}$-)espace vectoriel de dimension 2 et B est une famille de 2 éléments de $\mathbb{R}^2$, il suffit donc de montrer que B est une famille libre, c'est à dire que pour tout $(\lambda_1, \lambda_2) \in \mathbb{R}^2$:

$$\lambda_1 (3e_1 - 2e_2) + \lambda_2 (e_1 + e_2) = 0 \Rightarrow \lambda_1 = \lambda_2 = 0$$

Soient λ_1, λ_2 deux nombres réels. On a :

$$\begin{aligned}
\lambda_1 (3e_1 - 2e_2) + \lambda_2 (e_1 + e_2) = 0 &\Leftrightarrow (3\lambda_1 + \lambda_2) e_1 + (\lambda_2 - 2\lambda_1) e_2 = 0 && (1)\\
&\Leftrightarrow \begin{cases} 3\lambda_1 + \lambda_2 = 0 \\ \lambda_2 - 2\lambda_1 = 0 \end{cases} && (2)\\
&\Leftrightarrow \begin{cases} \lambda_2 = -3\lambda_1 \\ -5\lambda_1 = 0 \end{cases} && (3)\\
&\Leftrightarrow \lambda_1 = \lambda_2 = 0, && (4)
\end{aligned}$$

où en (1) on a utilisé les propriétés élémentaires des opérations sur un espace vectoriel, en (2) on a utilisé le fait que $(e_1,\ e_2)$, la base canonique de $\mathbb{R}^2$, est une base et en (3) et (4) on a effectué des manipulations élémentaires sur un système d'équations dont les inconnues sont des nombres réels.

2. Trouver la matrice de changement de base P telle que si $v = v_1 e_1 + v_2 e_2$, alors $P \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ donne les coordonnées de v dans la base B.

Solution : Notons $a_1 := 3e_1 - 2e_2$ et $a_2 := e_1 + e_2$. Comme nous l'avons montré en cours, la matrice de changement de base $P = (p_{i,j})_{1\leq i\leq 2, 1\leq j\leq 2}$ retournant les coordonnées d'un vecteur dans la base (a_1, a_2) quand on l'applique aux coordonnées de ce vecteur dans la base (e_1, e_2) est telle que $\begin{pmatrix} p_{1,1} \\ p_{2,1} \end{pmatrix}$ et $\begin{pmatrix} p_{1,2} \\ p_{2,2} \end{pmatrix}$ sont respectivement les coordonnées de e_1 et de e_2 dans la base (a_1, a_2).

Or :

$$\begin{aligned} \begin{cases} a_1 = 3e_1 - 2e_2 \\ a_2 = e_1 + e_2 \end{cases} &\Leftrightarrow \begin{cases} a_1 + 2a_2 = 5e_1 \\ a_2 = e_1 + e_2 \end{cases} \\ &\Leftrightarrow \begin{cases} e_1 = \frac{1}{5}(a_1 + 2a_2) \\ e_2 = a_2 - \frac{1}{5}(a_1 + 2a_2) \end{cases} \\ &\Leftrightarrow \begin{cases} e_1 = \frac{1}{5}(a_1 + 2a_2) \\ e_2 = \frac{1}{5}[-a_1 + 3a_2] \end{cases} \end{aligned} \tag{5}$$

Donc :

$$P = \frac{1}{5}\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}.$$

3. Calculer les coordonnées de $v = -e_1 + e_2$ dans la base B.

Solution : Les coordonnées de v dans la base canonique sont $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$, donc ses coordonnées dans la base B sont $P\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 1/5 \end{pmatrix}$.

Exercice 6 Représentation matricielles d'opérations linéaires sur les polynômes. ($\star\star$)

On considère :

$$f : \left| \begin{array}{ccc} \mathbf{R}_2[X] & \to & \mathbf{R}_3[X] \\ P = a_0 + a_1 X + a_2 X^2 & \mapsto & f(P) = (a_2 + a_0)X^3 + a_1 X^2 + a_0 \end{array} \right.$$

et les bases canoniques $V = (1, X, X^2)$ de $\mathbf{R}_2[X]$ et $W = (1, X, X^2, X^3)$ de $\mathbf{R}_3[X]$.

1. Montrez que f est linéaire.

2. Donnez la matrice $M(f, V, W)$.

3. A l'aide de cette matrice, évaluez $f(X^2 - 3X + 1)$

Exercice 7 Différentes visions des opérations matricielles. ($\star$)

1. Écrire le produit :

$$\begin{pmatrix} 3 & 6 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix},$$

   (a) comme une combinaison linéaire des colonnes de la matrice de gauche ;

   (b) comme une matrice contenant deux produits scalaires.

2. Écrire le produit matriciel :

$$\begin{pmatrix} 3 & 6 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 1 & 2 \\ 1 & -1 \end{pmatrix},$$

   (a) comme deux produits matrice-vecteur ;

   (b) comme deux produits vecteur-matrice ;

   (c) comme une somme de 3 matrices de rang 1 ;

   (d) comme une matrice contenant quatre produits scalaires.

Solution :

1. (a)

$$\begin{pmatrix} 3 & 6 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix} = 4 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \begin{pmatrix} 6 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

   (b)

$$\begin{pmatrix} 3 & 6 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a^T c \\ b^T c \end{pmatrix},$$

   où : $a := \begin{pmatrix} 3 & 6 & 2 \end{pmatrix}^T$, $b := \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}^T$ et $c := \begin{pmatrix} 4 & 1 & 1 \end{pmatrix}^T$.

2. Notons C le produit matriciel considéré dans cette question.

   (a) On a :

$$C = \begin{pmatrix} Ab & Ac \end{pmatrix},$$

   où : $A := \begin{pmatrix} 3 & 6 & 2 \\ 2 & 1 & 3 \end{pmatrix}$, $b := \begin{pmatrix} 4 & 1 & 1 \end{pmatrix}^T$ et $c := \begin{pmatrix} 5 & 2 & -1 \end{pmatrix}^T$.

   (b) On a aussi :

$$C = \begin{pmatrix} bA \\ cA \end{pmatrix},$$

   où : $A := \begin{pmatrix} 4 & 1 & 1 \\ 5 & 2 & -1 \end{pmatrix}^T$, $b := \begin{pmatrix} 3 & 6 & 2 \end{pmatrix}$ et $c := \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}$.

Solution :

2. (c) On a également :

$$C = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \begin{pmatrix} 4 & 5 \end{pmatrix} + \begin{pmatrix} 6 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

(d) Et finalement :

$$C = \begin{pmatrix} a^T c & a^T d \\ b^T c & b^T d \end{pmatrix},$$

où : $a := \begin{pmatrix} 3 & 6 & 2 \end{pmatrix}^T$, $b := \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}^T$, $c := \begin{pmatrix} 4 & 1 & 1 \end{pmatrix}^T$ et $d := \begin{pmatrix} 5 & 2 & -1 \end{pmatrix}^T$.

Exercice 8 Matrix product with diagonal matrix. ($\star$)

On note $\text{diag}(x_1, x_2, \dots, x_n)$ la matrice $\begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}$ with all the off-diagonal coefficients equal to zero.

1. Express the matrix product $A\text{diag}(x_1, x_2, \dots, x_n)$ in terms of the columns of A and the diagonal coefficients $x_1, x_2, \dots, x_n$.

Solution : We can write a general matrix product AX in terms of the columns of $X = [\mathbf{c_1}\ \mathbf{c_2}\ \dots\ \mathbf{c_n}]$ as $AX = [A\mathbf{c_1}\ A\mathbf{c_2}\ \dots\ A\mathbf{c_n}]$. Furthermore, we can write a general matrix vector product Av in terms of the columns of $A = [\mathbf{a_1}\ \mathbf{a_2}\ \dots\ \mathbf{a_n}]$ as $Av = v_1\mathbf{a_1} + v_2\mathbf{a_2} + \dots + v_n\mathbf{a_n}$. For $X = \text{diag}(x_1, x_2, \dots, x_n)$, we thus have $A\mathbf{c_i} = c_{1i}\mathbf{a_1} + c_{2i}\mathbf{a_2}\ \dots + c_{ni}\mathbf{a_n} = 0 + 0 + \dots + 0 + x_i\mathbf{a_i} + 0 + \dots + 0 = x_i\mathbf{a_i}$ and therefore

$$A\text{diag}(x_1, x_2, \dots, x_n) = [x_1\mathbf{a_1}\ x_2\mathbf{a_2}\ \dots\ x_n\mathbf{a_n}]$$

2. Express the matrix product $\text{diag}(x_1, x_2, \dots, x_n)B$ in terms of the lines of B and the diagonal coefficients $x_1, x_2, \dots, x_n$.

Solution : Let us write B in terms of its rows as $B = [\mathbf{b_1}\ \mathbf{b_2}\ \dots\ \mathbf{b_n}]^T$. Then $\text{diag}(x_1, x_2, \dots, x_n)B = ((\text{diag}(x_1, x_2, \dots, x_n)B)^T)^T = (B^T\text{diag}(x_1, x_2, \dots, x_n)^T)^T = (B^T\text{diag}(x_1, x_2, \dots, x_n))^T$. From the previous question, we then deduce

$$\text{diag}(x_1, x_2, \dots, x_n)B = [x_1\mathbf{b_1}\ x_2\mathbf{b_2}\ \dots\ x_n\mathbf{b_n}]^T$$

Exercice 9 Matrices nilpotentes. ($\star\star$)

Soit N une matrice nilpotente de taille n par n (c'est à dire qu'il existe un entier p strictement positif, tel que $N^p = 0$). On définit $q := \inf\{r \mid N^r = 0, r \in \mathbf{N}, r > 0\}$ l'indice de nilpotence de N.

1. Montrer que $q \leq n$. Indice : par un raisonnement par l'absurde, montrer qu'il existe une matrice colonne x_0 telle que $(x_0, Nx_0, N^2x_0, ..., N^{q-1}x_0)$ est une famille libre et conclure.

Solution :

Supposons que pour toute matrice colonne x_0 de taille n, $(x_0, Nx_0, N^2x_0, ..., N^{q-1}x_0)$ soit une famille liée. Soit $x_0 \in \mathbb{R}^n$. Il existe alors $(\alpha_i)_{i=0}^{q-1} \neq 0$ (où 0 représente un vecteur de n zéros), tel que :

$$\sum_{i=0}^{q-1} \alpha_i N^i x_0 = 0.$$

Notons $i_0 = \min\{i \mid 0 \leq i \leq q-1,\ \alpha_i \neq 0\}$. On a donc :

$$\sum_{i=i_0}^{q-1} \alpha_i N^i x_0 = \sum_{i=0}^{q-1} \alpha_i N^i x_0 = 0,$$

avec $\alpha_{i_0} \neq 0$.

En multipliant à gauche par N^{q-1-i_0}, on obtient :

$$\sum_{i=i_0}^{q-1} \alpha_i N^{i+q-1-i_0} x_0 = N^{q-1-i_0} 0 = 0.$$

Par définition de q, pour tout $k \geq q$, $N^k x_0 = 0x_0 = 0$. On en déduit que :

$$\sum_{i=i_0}^{q-1} \alpha_i N^{i+q-1-i_0} x_0 = \alpha_{i_0} N^{q-1} x_0 = 0.$$

Solution : (continuée)

Et comme $\alpha_{i_0} \neq 0$, on obtient

$$N^{q-1} x_0 = 0.$$

On a donc obtenu pour tout x_0,

$$N^{q-1} x_0 = 0.$$

On en déduit que $N^{q-1} = 0$, ce qui est une contradiction avec la définition de q.

Il existe donc une matrice colonne x_0 telle que $(x_0, Nx_0, N^2x_0, \ldots, N^{q-1}x_0)$ est une famille libre de q vecteurs de $\mathbb{R}^n$. $\mathbb{R}^n$ étant un espace vectoriel de dimension n, on a $q \leq n$.

2. Soit A et B deux matrices de taille n par n qui commutent (c'est à dire qu'on a $AB = BA$)

et soit k un entier supérieur ou égal à 2. Démontrer l'identité remarquable :

$$A^k - B^k = (A - B)\left(\sum_{i=0}^{k-1} A^{k-1-i}B^i\right).$$

Solution : On peut le montrer par exemple par récurrence.

— Cas de base. Pour $k = 2$, on a :

$$\begin{aligned}(A - B)\left(\sum_{i=0}^{2-1} A^{2-1-i}B^i\right) &= (A - B)(A + B)\\ &= A^2 - BA + AB - B^2\\ &= A^2 - AB + AB - B^2\\ &= A^2 - B^2.\end{aligned}$$

— Propagation. Supposons que :

$$A^k - B^k = (A - B)\left(\sum_{i=0}^{k-1} A^{k-1-i}B^i\right).$$

On a que : $A(A^k - B^k) + (A^k - B^k)B = A^{k+1} - AB^k + A^kB - B^{k+1}$.

Par conséquent : $A^{k+1} - B^{k+1} = A(A^k - B^k) + (A^k - B^k)B + AB^k - A^kB$.

Solution : (continuée)

En utilisant l'hypothèse de récurrence et la commutativité de A et B, qui implique celle de A^i et B^j pour tous entiers naturels i et j, on en déduit que :

$$\begin{aligned}
A^{k+1}-B^{k+1} &= A(A-B)\left(\sum_{i=0}^{k-1} A^{k-1-i}B^i\right) + (A-B)\left(\sum_{i=0}^{k-1} A^{k-1-i}B^i\right)B + AB^k - A^kB \\
&= \left(\sum_{i=0}^{k-1}\left(A^{k+1-i}B^i - A^{k-i}B^{i+1} + A^{k-i}B^{i+1} - A^{k-1-i}B^{i+2}\right)\right) + AB^k - A^kB \\
&= \left(\sum_{i=0}^{k-1}\left(A^{k+1-i}B^i - A^{k-1-i}B^{i+2}\right)\right) + AB^k - A^kB \\
&= A\left(\sum_{i=0}^{k-1} A^{k-i}B^i\right) + AB^k - B\left(\sum_{i=0}^{k-1} A^{k-1-i}B^{i+1}\right) - A^kB \\
&= A\left(\sum_{i=0}^{k-1} A^{k-i}B^i\right) + AB^k - B\left(\sum_{i=1}^{k} A^{k-i}B^i\right) - A^kB \\
&= A\left(\sum_{i=0}^{k} A^{k-i}B^i\right) - B\left(\sum_{i=0}^{k} A^{k-i}B^i\right) \\
&= (A-B)\left(\sum_{i=0}^{k} A^{k-i}B^i\right).
\end{aligned}$$

3. À l'aide d'un choix judicieux de A, B et k, montrer que $I-N$ est inversible et donner son inverse.

Solution : En prenant $A := I$, $B := N$ et $k := q$, A et B commutent car $IN = N = NI$, et on obtient donc que :

$$I^q - N^q = (I-N)\left(\sum_{i=0}^{q-1} I^{q-1-i}N^i\right).$$

D'où :

$$I = (I-N)\left(\sum_{i=0}^{q-1} N^i\right).$$

De plus :

$$\begin{aligned}
(I-N)\left(\sum_{i=0}^{q-1} N^i\right) &= \sum_{i=0}^{q-1} N^i - \sum_{i=0}^{q-1} N^{i+1} \\
&= (\sum_{i=0}^{q-1} N^i)I - (\sum_{i=0}^{q-1} N^i)N \\
&= \left(\sum_{i=0}^{q-1} N^i\right)(I-N).
\end{aligned}$$

Solution : (continuée)

Au final :

$$I = (I - N)\left(\sum_{i=0}^{q-1} N^i\right) = \left(\sum_{i=0}^{q-1} N^i\right)(I - N),$$

et $I - N$ est donc inversible, d'inverse $\left(\sum_{i=0}^{q-1} N^i\right)$.

4. Soit a un nombre réel. Montrer l'inversibilité et calculer l'inverse de la matrice de taille n par n :

$$M_a := \begin{pmatrix} 1 & -a & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -a \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Solution :

Si $a = 0$, $M_a = I$ et M_a est donc inversible d'inverse I. Considérons maintenant le cas $a \neq 0$.

On a $M_a = I = N_a$, où

$$N_a := \begin{pmatrix} 0 & a & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & a \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}.$$

Calculons les puissances N_a^r de N_a pour r entier positif et montrons que N_a est nilpotente.

On peut décrire N_a de manière équivalente comme la matrice $(n_{a,i,j})$ avec pour tout $1 \leq i \leq n$, $1 \leq j \leq n$, $n_{a,i,j} = a\delta_{i+1,j}$, où pour tous entiers s, t, $d_{st} = 1$ si et seulement si $s = t$ et $d_{st} = 0$ sinon.

En utilisant cette caractérisation et la formule terme à terme pour le produit matriciel ($c_{ij} = \sum_k a_{ik}b_{kj}$ pour un produit matriciel $C = AB$), on montre facilement par récurrence que pour tout entier positif r, $N_a^r = (n_{r,a,i,j})$ avec $n_{r,a,i,j} = a^r\delta_{i+r,j}$.

On en déduit que N_a est nilpotente d'ordre n.

Solution : (continuée)

En appliquant le résultat de la question précédente, on obtient finalement que M_a est inversible d'inverse :

$$M_a^{-1} = \sum_{i=0}^{n-1} N_a^i,$$

c'est à dire que :

$$M_a^{-1} = \begin{pmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a^2 \\ \vdots & & \ddots & \ddots & a \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

2 Algèbre linéaire : orthogonalité

Exercice 10 Produit scalaire usuel dans $\mathbf{R}^n$. ($\star$)

Prouvez que la fonction

$$\left| \begin{array}{l} \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R} \\ x, y \mapsto \sum_{i=1}^n x_i y_i \end{array} \right.$$

définit un produit scalaire sur $\mathbf{R}^n$.

Exercice 11 Matrices orthogonales. ($\star$)

Soit $(v_1, v_2, \dots, v_n) \in (\mathbf{R}^d)^n$ les coordonnées de n vecteurs de dimensions d dans la base canonique de $\mathbf{R}^d$.

1. On suppose que ces vecteurs sont deux à deux orthogonaux, calculez les produits matriciels $[v_1\ v_2\ \dots\ v_n][v_1\ v_2\ \dots\ v_n]^T$ et $[v_1\ v_2\ \dots\ v_n]^T[v_1\ v_2\ \dots\ v_n]$, où on a considéré les v_i comme des matrices colonnes.
2. Même question, si on suppose à présent que les vecteurs sont deux à deux orthonormaux.
3. On suppose à présent que les vecteurs sont deux à deux orthonormaux et que $n = d$. Montrez que la matrice $[v_1\ v_2\ \dots\ v_n]$ est inversible et donnez son inverse.

Exercice 12 Procédé d'orthonormalisation de Gram-Schmidt et décomposition QR. ($\star$)

On considère le vecteur de v_1 de $\mathbf{R}^2$ dont la décomposition dans la base canonique est $(3, 1)$ et le vecteur v_2 de $\mathbf{R}^2$ dont la décomposition dans la base canonique est $(2, 2)$.

1. Appliquer le procédé d'orthonormalisation de Gram-Schmidt à la famille (v_1, v_2).

2. En déduire une décomposition de la matrice

$$\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

commme un produit d'une matrice orthogonale et d'une matrice triangulaire supérieure.

Exercice 13 Validité du processus de Gram-Schmidt. ($\star\star$)

Etant donnés un espace vectoriel E et un produit scalaire $(.|.)$ quelconques (possiblement en dimension infinie), prouvez par récurrence que la famille de vecteurs produite par le processus de Gram-Schmidt est orthonormale.

Exercice 14 Polynômes orthogonaux. ($\star\star\star$)

On considère l'ensemble $\mathbf{R}[X]$ des polynômes à coefficients réels. Pour rappel, pour tout $P \in \mathbf{R}[X]$, il existe un entier positif d et des nombres réels a_0, a_1, $\ldots$, a_d, tels que

$$P = \sum_{i=0}^{d} a_i X^i.$$

1. Définir des notions naturelles d'addition et de multiplication par un scalaire pour les polynômes et montrer que ces opérations permettent de munir $\mathbf{R}[X]$ d'une structure d'espace vectoriel.

Solution : Nous définissons une opération d'addition

$$+ : \left| \begin{array}{lcl} \mathbf{R}[X] \times \mathbf{R}[X] & \to & \mathbf{R}[X] \\ P, Q & \mapsto & P + Q \end{array} \right.$$

sur **R**[X] en prenant pour tout $P = \sum_{i=1}^{d_1} a_i$ et $Q = \sum_{i=1}^{d_2} b_i$:

$$P + Q := \sum_{i=0}^{\max(d1,d2)} (\tilde{a}_i + \tilde{b}_i) X^i,$$

où $\tilde{a}_i = a_i$ si $i \leq d_1$ et 0 sinon et $\tilde{b}_i = b_i$ si $i \leq d_2$ et 0 sinon.

Pour la multiplication par un scalaire :

$$. : \left| \begin{array}{lcl} \mathbf{R} \times \mathbf{R}[X] & \to & \mathbf{R}[X] \\ \alpha, P & \mapsto & \alpha P \end{array} \right.$$

sur **R**[X] nous prenons pour tout $\alpha \in \mathbf{R}$ et $P = \sum_{i=1}^{d_1} a_i$:

$$\alpha P := \sum_{i=0}^{d} (\alpha a_i) X^i.$$

Les propriétés élémentaires de l'addition et de la multiplication sur **R** permettent de vérifier l'associativité, la commutativité, l'existence d'un élément neutre et d'inverses pour $+$ sur $\mathbf{R}[X]$, ainsi que la compatibilité de . avec la multiplication sur **R**, la présence d'un élement neutre pour . et la distributivité de . par rapport à l'addition sur **R** et sur $\mathbf{R}[X]$.

On en conclut que $(\mathbf{R}[X], +, .)$ est un **R**-espace vectoriel.

2. Montrer que :

$$(. \mid .) : \left| \begin{array}{lcl} \mathbf{R}[X] \times \mathbf{R}[X] & \to & \mathbf{R} \\ P, Q & \mapsto & \int_0^1 P(x) Q(x) dx \end{array} \right.$$

définit un produit scalaire sur $\mathbf{R}[X]$.

Solution : Un produit scalaire sur un $\mathbf{R}$-espace vectoriel est une forme bilinéaire symétrique définie positive.

Avant tout, remarquons que les fonctions polynômiales d'une variable réelle sont continues sur $[0,1]$ intervalle compact de $\mathbf{R}$, il n'y donc pas de problèmes de définition de l'intégrale ci-dessus.

En appliquant notre définition d'addition sur $\mathbf{R}[X]$ et la linéarité de l'intégration, on obtient que $(. \mid .)$ est une forme bilinéaire. Elle est symétrique par commutativité de la multiplication sur $\mathbf{R}$. Elle est positive car pour tout $P \in \mathbf{R}[X]$, $(P \mid P) = \int_0^1 P^2(x)dx$ et l'intégrale d'une fonction positive est toujours positive. Elle est définie car : 1) les fonctions polynomiales sont continues sur $]0,1[$, intervalle ouvert non vide, 2) l'intégrale d'une fonction continue et positive sur un intervalle ouvert non vide ne peut être égale à zéro que si la fonction est égale à zéro sur cet intervalle 3) la fonction polynomiale associé à un polynôme ne peut être égale à zéro sur un intervalle ouvert non vide que si tous les coefficients du polynôme sont égaux à zéro.

3. Montrer qu'il existe une unique base $(P_0, P_1, \dots)$ de $\mathbf{R}[X]$, orthonormale pour $(. \mid .)$ et telle que pour tout entier positif n, le degré de P_n est n et le coefficient d'ordre n de P_n est strictement positif. Indice : adapter le processus d'orthonormalisation de Gram-Schmidt décrit en cours à ce cas de figure.

Solution :

La décomposition canonique des polynômes comme combinaison linéaire de $(1, X, X^2, \dots)$ établit que $\mathcal{B} := (X^i)_{i=0}^{+\infty}$ est une base de $\mathbf{R}[X]$. On applique le procédé d'orthonormalisation de Gram-Schmidt à cette base en définissant par récurrence :

$$P_0 = 1,$$

$$\tilde{P}_n = X^n - \sum_{j=0}^{n-1} (X^n \mid P_j) P_j,$$

et enfin :

$$P_n = \frac{\tilde{P}_n}{\sqrt{(\tilde{P}_n \mid \tilde{P}_n)}},$$

pour tout entier strictement positif n.

Solution : (continuée)

La preuve par récurrence que la famille $\mathcal{P} := (P_i)_{i=0}^{+\infty}$ résultante est est une base orthonormale de $\mathbf{R}[X]$ est identique au cas vu en cours. Une récurrence immédiate établit que pour tout entier positif n, $\tilde{P}_n$ et P_n sont de degré n, que le coefficient dominant de $\tilde{P}_n$ est 1 et que le coefficient dominant de P est $\frac{1}{\sqrt{(\tilde{P}_n|\tilde{P}_n)}} > 0$.

Il reste à prouver l'unicité d'une base possédant ces propriétés. Supposons qu'on ait trouvé une base $\mathcal{Q} := (Q_i)_{i=0}^{+\infty}$ avec les mêmes propriétés qui soit différente de $\mathcal{P}$. Notons k le plus petit indice tel que $P_k \neq Q_k$. Par hypothèse, le degré de Q_k est k, on peut donc écrire $Q_k = \sum_{i=0}^{k}(Q_k \mid P_i)P_i$. Pour tout entier i, $0 \leq i < k$, on a que $Q_i = P_i$ et donc $(Q_k \mid P_i) = (Q_k \mid Q_i) = 0$ par orthonormalité de $\mathcal{Q}$. Donc $Q_k = (Q_k \mid P_k)P_k$. On en déduit que $\|Q_k\| = |(Q_k \mid P_k)|^2\|P_k\|$. Comme $\mathcal{Q}$ et $\mathcal{P}$ sont orthonormale, on a $\|Q_k\| = \|P_k\| = 1$. ON en déduit que $|(Q_k \mid P_k)|^2 = 1$ et donc que $(Q_k \mid P_k)$ est égal à -1 ou 1. Or $(Q_k \mid P_k)$ est le coefficient dominant de Q_k et est donc par hypothèse positif. On en conclut que $(Q_k \mid P_k) = 1$ et donc que $Q_k = P_k$, ce qui contredit notre hypothèse que $\mathcal{Q} \neq \mathcal{P}$.

4. Quelle est la dimension de $\mathbf{R}[X]$?

Solution : Les bases de $\mathbf{R}[X]$ que nous avons trouvé contiennent un nombre d'élément infini (plus précisémment, un infini dénombrable). $\mathbf{R}[X]$ est donc un espace vectoriel de dimension infinie (plus précisémment de dimension égale au cardinal de $\mathbf{N}$).

5. Montrer que pour tout entier $n \geq 2$:

$$P_n(X) = (a_n X + b_n)P_{n-1}(X) + c_n P_{n-2}(X),$$

pour des suites de nombres réels (a_n), (b_n) et (c_n) qu'on explicitera.

Solution : Un élément central de la preuve est la propriété suivante : pour tous polynômes P et Q :

$$(XP \mid Q) = \int_0^1 xP(x)Q(x)dx = \int_0^1 P(x)xQ(x)dx = (P \mid XQ).$$

Solution : (continuée)

Cette propriété permet notamment d'établir que tout polynôme R de la forme :

$$R = (aX + b)P_{n-1} + cP_{n-2},$$

où a, b et c sont des réels, est orthogonal à P_0, à P_1, $\ldots$, et à P_{n-3}. En effet, prenons un entier k dans $\{0, 1, \ldots, n-3\}$. On a :

$$(R \mid P_k) = a(XP_{n-1} \mid P_k) + b(P_{n-1} \mid P_k) + c(P_{n-2} \mid P_k) = a(P_{n-1} \mid XP_k) + 0 + 0.$$

Comme le degré de P_k est $k < n-2$, le degré de XP_k est strictement inférieur à $n-1$ et donc $XP_k \in \mathrm{Vect}\{P_0, P_1, \ldots, P_{n-2}\}$. Par orthonormalité de $\mathcal{P}$, on en déduit que XP_k est orthonormal à P_{n-1} et donc que $(R \mid P_k) = 0$.

Une conséquence de ce résultat est que si pour tout $n \geq 2$ on pouvait trouver des réels a_n, b_n et c_n tels que la projection de $R_n = a_nXP_{n-1} + b_nP_{n-1} + c_nP_{n-2}$ sur, respectivement, P_n, P_{n-1} et P_{n-2} soit, respectivement, 1, 0 et 0, alors on aurait forcément pour tout $n \geq 2$, $R_n = P_n$ (en utilisant le résultat d'unicité de la question 3), ce qui résoudrait la question.

Il ne reste donc plus qu'à montrer qu'on peut trouver de tels réels a_n, b_n et c_n pour tout $n \geq 2$. Soit $n \geq 2$, on a :

$$(R_n \mid P_n) = a_n(XP_{n-1} \mid P_n) + 0 + 0.$$

Calculons plus explicitement $(XP_{n-1} \mid P_n)$ pour vérifier qu'il est non-nul, ce qui nous permettra de garantir $(R_n \mid P_n) = 1$ en prenant $a_n := \frac{1}{(XP_{n-1} \mid P_n)}$.

Le degré de XP_{n-1} est n, on peut donc l'écrire :

$$XP_{n-1} = \sum_{i=0}^{n} (XP_{n-1} \mid P_i)P_i.$$

Les coefficients dominants des polynômes à gauche et à droite du signe d'égalité sont respectivement λ_{n-1} et $\lambda_n(XP_{n-1} \mid P_n)$, où pour tout entier k, λ_k est le coefficient dominant de P_k (donc $\lambda_k \neq 0$). On a donc : $\lambda_{n-1} = \lambda_n(XP_{n-1} \mid P_n)$ avec $\lambda_{n-1} \neq 0$ et $\lambda_n \neq 0$, d'où :

$$(XP_{n-1} \mid P_n) = \frac{\lambda_{n-1}}{\lambda_n} \neq 0.$$

Solution : (continuée)

En prenant :

$$a_n := \frac{\lambda_n}{\lambda_{n-1}},$$

on a donc $(R_n \mid P_n) = 1$.

Considérons maintenant la projection sur P_{n-1} :

$$(R_n \mid P_{n-1}) = a_n(XP_{n-1} \mid P_{n-1}) + b_n.$$

En prenant :

$$b_n := -\frac{\lambda_n}{\lambda_{n-1}}(XP_{n-1} \mid P_{n-1})$$

on a donc $(R_n \mid P_{n-1}) = 0$.

Finalement, considérons la projection sur P_{n-2} :

$$(R_n \mid P_{n-2}) = a_n(XP_{n-1} \mid P_{n-2}) + c_n.$$

En prenant :

$$c_n := -\frac{\lambda_n}{\lambda_{n-1}}(XP_{n-1} \mid P_{n-2})$$

on a donc $(R_n \mid P_{n-2}) = 0$.

On peut obtenir une expression plus explicite pour c_n en remarquant que :

$$(XP_{n-1} \mid P_{n-2}) = (P_{n-1} \mid XP_{n-2}) = (XP_{n-2} \mid P_{n-1}) = \frac{\lambda_{n-2}}{\lambda_{n-1}},$$

par un raisonnement analogue au cas de $(XP_{n-1} \mid P_n)$. On obtient alors :

$$c_n = -\frac{\lambda_n \lambda_{n-2}}{\lambda_{n-1}^2}.$$

3 Algèbre linéaire : structure des applications linéaires

Exercice 15 Propriétés élémentaires de la décomposition en valeurs singulières. ($\star$)

Indice : dans le cadre de cet exercice, pensez aux différentes interprétations des produits matrice-matrice et matrice-vecteurs que nous avons vu en cours.

Soit A une matrice à coefficients réels de taille m par n. Le théorème sur la décomposition en valeurs singulières indique qu'il existe une matrice orthogonale $U = [u_1\ u_2\ \dots\ u_m]$ de taille m par m, une matrice orthogonale $V = [v_1\ v_2\ \dots\ v_n]$ de taille n par n, un entier naturel $r \leq \min(m,n)$ et une matrice diagonale $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0)$ de taille m par n, tels que toutes les matrices sont à coefficients réels, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ et $AV = U\Sigma$.

1. Montrez que $A = U\Sigma V^T$ et que $\Sigma = U^T AV$.

   Solution : $AV = U\Sigma$ donc $AVV^T = U\Sigma V^T$ or $VV^T = I$, donc $A = U\Sigma V^T$. $AV = U\Sigma$ donc $U^T AV = U^T U\Sigma$ or $U^T U = I$, donc $\Sigma = U^T AV$.

2. Montrez que pour tout i dans $\{1, 2, \dots, r\}$, $Av_i = \sigma_i u_i$.

   Solution : $AV = [Av_1\ Av_2\ \dots\ Av_n]$ et (cf. exercice 8), $U\Sigma = [\sigma_1 u_1\ \sigma_2 u_2\ \dots\ \sigma_r u_r\ 0\ 0\ \dots\ 0]$. Donc $Av_i = \sigma_i u_i$ pour tout i dans $\{1, 2, \dots, r\}$.

3. Montrez que pour tout i dans $\{r+1, r+2, \dots, n\}$, $Av_i = 0$, i.e. $v_i \in \mathrm{Ker}A$.

   Solution : Nous l'avons établi à la question précédente.

4. Montrez que $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$.

   Solution : D'après l'exercice 8, $U\Sigma = [\sigma_1 u_1\ \sigma_2 u_2\ \dots\ \sigma_r u_r\ 0\ 0\ \dots\ 0]$. Donc

   $$A = U\Sigma V^T = [\sigma_1 u_1\ \sigma_2 u_2\ \dots\ \sigma_r u_r\ 0\ 0\ \dots\ 0]V^T.$$

   Nous pouvons écrire ce produit matrice-matrice comme une somme de produits de matrices de rang 1 pour obtenir :

   $$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T + \sum_{i=r+1}^{n} 0 v_i^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T.$$

Exercice 16 Dérivation des théorèmes fondamentaux de l'algèbre linéaire à partir de la décomposition en valeurs singulières. ($\star\star$)

On considère le même point de départ que dans l'exercice précédent (décomposition en valeur singulière d'une matrice A).

1. Montrez que les r premières colonnes de U forment une base orthonormale de l'image de A.

2. Montrez que les $m - r$ dernières colonnes de U forment une base orthonormale du co-noyau de A (i.e. le noyau de A^T).

3. Montrez que les r premières colonnes de V forment une base orthonormale de la co-image de A (i.e. l'image de A^T).

4. Montrez que les $n-r$ dernières colonnes de V forment une base orthonormale du noyau de A.

5. Montrez que le noyau et la co-image de A sont des sous-espaces orthogonaux supplémentaires de $\mathbf{R}^n$, c'est à dire qu'on peut décomposer tout vecteur de $\mathbf{R}^n$ de manière unique en une somme de deux vecteurs orthogonaux, l'un appartenant au noyau de A et l'autre appartenant à la co-image de A.

6. Montrez que le co-noyau et l'image de A sont des sous-espaces orthogonaux supplémentaires de $\mathbf{R}^m$.

7. Montrez qu'il existe un unique entier naturel $r \leq \min(m,n)$, le range de A, tel que l'image et la co-image de A sont de même dimension r.

8. Montrez que l'entier r de la question précédente est aussi la dimension de l'image de toute application linéaire $f : E \to F$ telle qu'il existe des bases B_E et B_F de E et F telles que $A = M(f, B_E, B_F)$.

9. Montrez que la matrice A est inversible—c'est à dire qu'il existe une matrice A^{-1} telle que $AA^{-1} = A^{-1}A = I$—si et seulement si $r = m = n$ et donnez une expression pour A^{-1} en fonction de U, Σ et V.

Exercice 17 Décomposition en valeurs singulières des matrices orthogonales. ($\star$)

Déterminez les valeurs singulières d'une matrice orthogonale.

Exercice 18 Propriétés élémentaires de la décomposition spectrale. ($\star$)

Soit S une matrice à coefficients réels de taille m par m symmétrique. Le théorème spectral indique qu'il existe une matrice orthogonale réelle $Q = [q_1 \; q_2 \; \dots q_m]$ de taille m par m et une matrice diagonale réelle $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ de taille m par m, telles que $S = Q\lambda Q^T$.

1. Montrez que $\Lambda = Q^T SQ$ et que $SQ = Q\Lambda$.

2. Montrez que pour tout i dans $\{1, 2, \dots, n\}$, $Sq_i = \lambda_i q_i$.

3. Montrez que $S = \sum_{i=1}^{n} \lambda_i q_i q_i^T$.

4. Proposez une interprétation intuitive du résultat du produit matrice vecteur Sv en vous appuyant sur la réponse à la question précédente.

5. On considère un polynôme P. Calculer $P(S)$ en fonction de Q et Λ.

Exercice 19 Éléments propres de $A^T A$ et AA^T. ($\star$)

1. Soit A une matrice réelle de taille m par n. Décrire les valeurs propres de $A^T A$ et de AA^T en fonction des valeurs singulières de A.

Solution : On sait qu'il existe une décomposition en valeurs singulières de A, c'est à dire qu'il existe une matrice orthogonale U de taille $m \times m$, une matrice orthogonale V de taille $n \times n$ et une matrice rectangulaire diagonale Σ de taille $m \times n$ dont les $\min(m, n)$ valeurs diagonales sont rangées en ordre décroissant et positives, telles que $A = U\Sigma V^T$. De plus, on a vu que la matrice Σ apparaissant dans une telle décomposition est unique est que les valeurs $\sigma_1, \sigma_2, \dots, \sigma_{\min(m,n)}$ apparaissant sur sa diagonale sont appelées valeurs singulières de A.

Soit donc une telle décomposition en valeurs singulières (Σ, U, V) de A. On a :

$$\begin{aligned} A^T A &= (U\Sigma V^T)^T (U\Sigma V^T) \\ &= (V^T)^T \Sigma^T U^T U\Sigma V^T \\ &= V\Sigma^T (U^T U)\Sigma V^T \\ A^T A &= V\Sigma^T \Sigma V^T \end{aligned}$$

où on a utilisé l'associativité du produit matriciel, la règle donnant la transposée d'un produit matriciel en fonction des transposées des facteurs du produit et les propriétés définissant le caractère orthogonal d'une matrice.

$\Sigma^T \Sigma$ est une matrice diagonale de taille n par n dont les valeurs diagonales sont :

$$\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2, \quad \underbrace{0, 0, \dots, 0}_{\max(0, n-m) \text{ fois}} \quad ,$$

De plus, V est une matrice orthogonale, donc inversible, avec $V^{-1} = V^T$. On a donc $A^T A = VDV^{-1}$ avec $D = \Sigma^T \Sigma$ diagonale. On en déduit que les valeurs propres de $A^T A$ sont :

$$\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2, \quad \underbrace{0, 0, \dots, 0}_{\max(0, n-m) \text{ fois}} \quad .$$

Par un raisonnement directement analogue on obtient que $AA^T = U\Sigma\Sigma^T U^T$ et on en déduit que les valeurs propres de AA^T sont :

$$\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2, \quad \underbrace{0, 0, \dots, 0}_{\max(0, m-n) \text{ fois}} \quad .$$

2. Donner une base orthonormale de vecteurs propres pour chacune de ces deux matrices en

fonction des matrices orthogonales U et V apparaissant dans la décomposition en valeur singulières de A.

Solution :

On déduit de la diagonalisation $A^T A = V(\Sigma^T \Sigma)V^T$ que les colonnes de V forment une base de vecteurs propres de $A^T A$. Comme V est orthogonale, cette base est orthonormale. Par une raisonnement analogue on déduit de la diagonalisation $AA^T = U(\Sigma\Sigma^T)V$ que les colonnes de U forment une base orthonormale de vecteurs propres de AA^T.

Exercice 20 Positivité des matrices symmétriques et décomposition spectrale. ($\star$)

On considère une matrice symétrique réelle S. Donnez des conditions nécessaire et suffisante sur les éléments de la décomposition spectrale de S pour que S soit semi-définie positive ? Pour que S soit définie positive ?

Exercice 21 Relation d'ordre sur les matrices. ($\star$)

La relation d'ordre sur les matrices symétriques réelles proposée en cours, définit-elle un ordre total ? (C'est à dire, étant donnée deux matrices symmétriques réelles S_1 et S_2, a-t-on toujours soit $S_1 \preceq S_2$, soit $S_2 \preceq S_1$?)

Exercice 22 Vecteurs propres et valeurs propres. ($\star$)

Soit A une matrice à coefficients réels de taille m par m.

1. Montrez que si v est un vecteur propre de A tout produit de v par un scalaire est aussi un vecteur propre de A.
2. Montrer que si $\lambda \in \mathbf{C}$ est une valeur propre de A, le conjugué $\bar{\lambda}$ de λ est aussi une valeur propre de A.

Exercice 23 Matrices non-diagonalisables. ($\star$)

Démontrez que la matrice

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

n'est pas diagonalisable.

Exercice 24 Déterminant et trace. ($\star$)

1. Calculez le déterminant et la trace de

$$\begin{pmatrix} 1 & -1/2 \\ 2 & -1 \end{pmatrix}.$$

2. Calculez le déterminant de la matrice identité I_n de taille n par n en partant de la définitions formelle du déterminant.
3. Calculez le déterminant de la matrice $\operatorname{diag}(x_1, x_2, \ldots, x_n)$ en partant de la définitions formelle du déterminant.
4. Soit P une matrice inversible de taille $m \times m$, Q une matrice orthogonale de taille $m \times m$ et A une matrice de taille $m \times m$. Montrez que $\det(P^{-1}AP) = \det(Q^T AQ) = \det(A)$ et que $\operatorname{Tr}(P^{-1}AP) = \operatorname{Tr}(Q^T AQ) = \operatorname{Tr}(A)$.
5. Soit A une matrice de taille $m \times m$. En utilisant la décomposition de Jordan, calculez la trace et le déterminant de A en fonction de ses valeurs propres.

Exercice 25 Polynôme caractéristique. ($\star$)

1. Calculez le polynôme caractéristique de la matrice
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$
2. Vérifiez que le théorème de Cayley-Hamilton est vérifié dans le cadre de cet exemple.
3. Donnez le polynôme caractéristique d'une matrice de taille 2×2 quelconque en fonction de la trace et du déterminant de la matrice.
4. Montrez que le polynôme caractéristique est invariant au changement de système de coordonnées.
5. Soit A une matrice de taille $m \times m$. Montrez que les racines du polynôme caractéristique de A sont les valeurs propres de A avec leur multiplicité. Commencez par considérer le cas où A est diagonale, puis le cas où A est diagonalisable, puis le cas général.

Exercice 26 Ensembles de matrices. ($\star$)

1. Montrez que l'ensemble $\mathcal{M}_{m,n}(\mathbf{R})$ des matrices à coefficients réels de taille m par n muni de l'addition et du produit scalaire usuel est un espace vectoriel.
2. Montrez que la dimension de $\mathcal{M}_{m,n}(\mathbf{R})$ est mn, par exemple en proposant une base canonique.

Exercice 27 Diagonalisation. ($\star$)

On considère la matrice

$$A := \begin{pmatrix} -1 & 1 \\ 3 & -1 \end{pmatrix}.$$

1. Calculer le polynôme caractéristique de A.

Solution :

$$\chi_A(X) := \det(A - XI) = \begin{vmatrix} -1-X & 1 \\ 3 & -1-X \end{vmatrix} = (X+1)^2 - 3 = X^2 + 2X - 2$$

2. Donner les valeurs propres de A.

Solution : Les valeurs propres de A sont $\sqrt{3}-1$ et $-\sqrt{3}-1$, les racines de son polynôme caractéristique.

3. Trouver un vecteur propre associé à chaque valeur propre de A.

Solution : On cherche $v_1 := \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$ et $v_2 := \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}$ tels que v_1 et v_2 sont non nuls, $Av_1 = (\sqrt{3}-1)v_1$ et $Av_2 = (-\sqrt{3}-1)v_2$.

On a :

$$\begin{aligned} Av_1 = (\sqrt{3}-1)v_1 &\Leftrightarrow \begin{cases} -v_{11} + v_{12} = (\sqrt{3}-1)v_{11} \\ 3v_{11} - v_{12} = (\sqrt{3}-1)v_{12} \end{cases} \\ &\Leftrightarrow \begin{cases} v_{12} = \sqrt{3}v_{11} \\ v_{11} \in \mathbb{R} \end{cases} \end{aligned}$$

On en déduit que les vecteurs propres associés à la valeur propre $\sqrt{3}-1$ sont les élements de $S_1 = \{(a, \sqrt{3}a)^T \mid a \in \mathbb{R} \setminus \{0\}\}$. Par exemple, $(1, \sqrt{3})^T$ est un vecteur propres associé à $\sqrt{3}-1$.

On a aussi :

$$\begin{aligned} Av_2 = (-\sqrt{3}-1)v_2 &\Leftrightarrow \begin{cases} -v_{21} + v_{22} = (-\sqrt{3}-1)v_{21} \\ 3v_{21} - v_{22} = (-\sqrt{3}-1)v_{22} \end{cases} \\ &\Leftrightarrow \begin{cases} v_{22} = -\sqrt{3}v_{21} \\ v_{21} \in \mathbb{R} \end{cases} \end{aligned}$$

On en déduit que les vecteurs propres associés à la valeur propre $-\sqrt{3}-1$ sont les élements de $S_2 = \{(a, -\sqrt{3}a)^T \mid a \in \mathbb{R} \setminus \{0\}\}$. Par exemple, $(1, -\sqrt{3})^T$ est un vecteur propres associé à $\sqrt{3}-1$

4. Donner une diagonalisation de A.

Solution : A est une matrice de taille 2 par 2 possédant 2 valeurs propres distinctes $\sqrt{3}-1$ et $-\sqrt{3}-1$, elle est donc diagonalisable. De plus, $(1,\sqrt{3})^T$ et $(1,-\sqrt{3})^T$ sont respectivement des vecteurs propres associés à $\sqrt{3}-1$ et $-\sqrt{3}-1$. On en déduit que $A = PDP^{-1}$, où : $P := \begin{pmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{pmatrix}$ et $D := \begin{pmatrix} \sqrt{3}-1 & 0 \\ 0 & -\sqrt{3}-1 \end{pmatrix}$.

Par définition, $PP^{-1} = I$, ce qu'on peut exploiter pour trouver une expression explicite pour $P^{-1} := \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$.

On a :

$$PP^{-1} = I \Leftrightarrow \begin{cases} q_{11} + q_{21} = 1 \\ \sqrt{3}q_{11} - \sqrt{3}q_{21} = 0 \\ q_{12} + q_{22} = 0 \\ \sqrt{3}q_{12} - \sqrt{3}q_{22} = 1 \end{cases}.$$

$$\Leftrightarrow \begin{cases} q_{11} = q_{21} = 1/2 \\ q_{12} = \sqrt{3}/6 \\ q_{22} = -\sqrt{3}/6 \end{cases}$$

D'où : $P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{6} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} \end{pmatrix}$.

5. Calculer l'inverse de A.

Solution : A est inversible car toutes ses valeurs propres sont différentes de 0 et $A^{-1} = PD^{-1}P^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ 3/2 & 1/2 \end{pmatrix}$.

6. Calculer l'exponentielle de A.

Solution :

$$\begin{aligned} e^A &= Pe^DP^{-1} \\ &= P \begin{pmatrix} e^{\sqrt{3}-1} & 0 \\ 0 & e^{-\sqrt{3}-1} \end{pmatrix} P^{-1} \\ &= \begin{pmatrix} \frac{1}{2}\left(e^{\sqrt{3}-1} + e^{-\sqrt{3}-1}\right) & \frac{\sqrt{3}}{6}\left(e^{\sqrt{3}-1} - e^{-\sqrt{3}-1}\right) \\ \frac{\sqrt{3}}{2}\left(e^{\sqrt{3}-1} - e^{-\sqrt{3}-1}\right) & \frac{1}{2}\left(e^{\sqrt{3}-1} + e^{-\sqrt{3}-1}\right) \end{pmatrix} \end{aligned}.$$

4 Algèbre linéaire et optimisation

Exercice 28 Optimisation, normes, valeurs singulières et éléments propres. ($\star$)

1. Prouvez que $\|x\|_2 = \sqrt{x^T x}$ et que $\|Qx\|_2 = \|x\|_2$ pour toute matrice orthogonale Q.
2. Prouvez que $\|A\|_F = \|(\sigma_1, \sigma_2, \ldots, \sigma_r)\|_2$.
3. Prouvez que $\|A\|_2 = \sigma_1$.
4. Soit M une matrice carrée à coefficients réels. Prouvez que $\sigma_r(M) \leq \lambda \leq \sigma_1(M)$ pour toute valeur propre λ de M. Commencez par le cas d'une matrice de rang 1.
5. Montrez que toutes les valeurs propres d'une matrice orthogonale ont un module de 1.

Exercice 29 Moindres carrés linéaires et pseudo-inverse. ($\star\star$)

Etant donné A une matrice à coefficients réels de taille m par n et y une matrice colonne de taille m, on cherche à déterminer si $l : x \mapsto \|Ax - y\|_2$ possède une minimum global et à identifier le ou les $x \in \mathbf{R}^n$ où cet éventuel minimum global est atteint.

1. Montre qu'on peut répondre à notre question en étudiants $f : x \mapsto (Ax - y)^T (Ax - y)$.
2. Justifier que f est de classe C^1 et calculer le gradient de f.
3. En déduire une condition nécessaire d'optimum local pour f.
4. Calculez la Hessienne de f.
5. En déduire que f admet (au moins) un minimum global.
6. Donnez une condition nécessaire d'optimum global pour f.
7. On suppose que $A^T A$ est inversible. Prouvez que l admet un unique minimum global et donnez une expression explicite pour ce minimum global.
8. On définit la pseudo-inverse A^+ d'une matrice A à coefficients réels de taille m par n par le biais de sa décomposition en valeurs singulières $A = U\Sigma V^T$ comme $A^+ := V\Sigma^+ U^T$ où Σ^+ est la matrice diagonale contenant sur la case i de sa diagonale l'inverse de la i-ième valeurs singulière de A si celle-ci est strictement positive et 0 sinon. Montrez que si A est carrée et inversible, $A^+ = A$.
9. Montrez que $(A^+)^+ = A$, que $(A^+)^T = (A^T)^+$, et que $A(A^T)(A^+)^T = A = (A^+)^T(A^T)A$.
10. Montrez que AA^+ est la projection orthogonale sur l'image de A, que A^+A est la projection orthogonale sur l'espace des lignes de A (co-image), que $I - AA^+$ est la projection orthogonale sur le co-noyau de A et que $I - A^+A$ est la projection orthogonale sur sur le noyau de A.

11. Soit X une matrice à coefficients réels de taille n par d. Montrez que $\theta^* = X^+ y$ est la solution de norme minimale du problème des moindres carrés linéaires consistant à minimiser $\|y - X\theta\|_2$. Précisez l'ensemble des autres solutions possibles s'il y en a.

Exercice 30 Régression linéaire et régularisation L_2. ($\star\star$)

On observe $(x_1, y_1), \dots (x_n, y_n) \in (\mathbf{R}^d \times \mathbf{R})^n$. On modélise $x_1, \dots, x_n$ comme les valeurs observées de variables aléatoires $X_1, \dots, X_n$ i.i.d. de loi P et $y_1, \dots, y_n$ comme les valeurs observées de variables aléatoires $Y_1, \dots Y_n$ telles qu'il existe $\theta^* \in \mathbf{R}^d$ tel que pour tout i dans $\{1, \dots n\}$, $Y_i = X_i^T \theta^* + \epsilon_i$, avec $\epsilon_1, \dots, \epsilon_n$ i.i.d. de loi N telle que $\mathbf{E}(N) = 0$ et $\mathbf{Var}(N) = \sigma^2$.

On note X la matrice de taille n par d dont les lignes sont $x_1, \dots, x_n$ et Y la matrice colonne contenant $y_1, \dots, y_n$. On cherche à estimer θ^* et on considère deux estimateurs :

— L'estimateur de la régression « ridge » :

$$\hat{\theta}_\lambda := \arg\min_{\theta \in \mathbf{R}^d} \frac{1}{n} \|Y - X\theta\|_2^2 + \lambda \|\theta\|_2^2,$$

défini pour $\lambda > 0$;

— L'estimateur de norme minimale parmi les estimateurs au moindre carrés :

$$\hat{\theta} := \arg\min_{\theta \in \Theta_{LS}} \|\theta\|_2$$

où $\Theta_{LS} = \arg\min_{\theta \in \mathbf{R}^d} \|Y - X\theta\|_2^2$.

1. Montrez que pour toute matrice M de taille n par m et tout réel $\delta > 0$, $M^T M + \delta I$ est inversible. (Vous pouvez montrer, par exemple, que la matrice est symétrique définie positive et utiliser le théorème spectral pour conclure.)

Solution : $(M^TM + \delta I)^T = (M^TM)^T + \delta I^T = M^T(M^T)^T + \delta I = M^TM + \delta I$, donc $M^TM + \delta I$ est symmétrique.

Soit $v \in \mathbf{R}^m \backslash \{0\}$. On a vu dans le DM 1 que M^TM est positive (car elle est symmétrique et ses valeurs propres sont positives), donc $v^TM^TMv \geq 0$. De plus $v^T\delta Iv = \delta\|v\|_2^2 > 0$ car $v \neq 0$ et $\delta > 0$. Donc $v^T(M^TM + \delta I)v = v^TM^TMv + v^T\delta Iv > 0$. Donc $M^TM + \delta I$ est définie positive.

Comme $M^TM + \delta I$ est symmétrique on peut appliquer le théorème spectral et l'écrire $U\text{diag}(\lambda_1, \dots, \lambda_m)U^T$ pour une matrice orthogonale (réelle) U et des réels $\lambda_1, \dots, \lambda_m$. Ces réels sont strictement positifs car $M^TM + \delta I$ est définie positive. On peut donc définir $A := U^T\text{diag}(1/\lambda_1, \dots, 1/\lambda_m)U$. On vérifie facilement que $(M^TM + \delta I)^TA = A(M^TM + \delta I)^T = I$. $M^TM + \delta I$ est donc inversible.

2. Montrez que le problème d'optimisation dont $\hat{\theta}_\lambda$ est défini comme la solution possède bien une unique solution donnée par $\hat{\theta}_\lambda = \frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^TY$.

Solution : Considérons la fonction $f : \theta \mapsto \frac{1}{n}\|Y - X\theta\|_2^2 + \lambda\|\theta\|_2^2$, définie sur $\mathbf{R}^d$.

f est clairement coercive, elle admet donc un minimum global.

De plus f est de classe C^2. Son gradient est :

$$\begin{aligned}
\nabla f(\theta) &= \frac{1}{n}\nabla((Y - X\theta)^T(Y - X\theta)) + \lambda\nabla(\theta^T\theta) \\
&= \frac{1}{n}(-2\nabla(\theta^TX^TY) + \nabla(\theta^TX^TX\theta)) + \lambda\nabla(\theta^T\theta) \\
&= \frac{1}{n}(-2X^TY + 2X^TX\theta) + 2\lambda\theta.
\end{aligned}$$

Sa Hessienne est :

$$\begin{aligned}
\nabla^2 f(\theta) &= \frac{1}{n}\nabla(2X^TX\theta) + 2\lambda\nabla(\theta) \\
\nabla^2 f(\theta) &= 2\left(\frac{X^TX}{n} + \lambda I\right)
\end{aligned}$$

D'après la question 1, sa Hessienne est toujours définie positive et donc f est strictement convexe. Il existe donc bien une unique solution globale $\hat{\theta}_\lambda$ au problème d'optimisation considéré.

De plus, d'après la condition nécessaire d'existence d'optimum local pour les fonction de classe C^1, on doit avoir $\nabla f(\hat{\theta}_\lambda) = 0$, c'est à dire $\frac{X^TY}{n} = \left(\frac{X^TX}{n} + \lambda I\right)\hat{\theta}_\lambda$. D'après la question 1, $\frac{X^TX}{n} + \lambda I$ est inversible et donc $\hat{\theta}_\lambda = \left(\frac{X^TX}{n} + \lambda I\right)^{-1} \frac{X^TY}{n}$.

3. Montrez que le problème d'optimisation dont $\hat{\theta}$ est défini comme la solution possède bien une unique solution donnée par $\hat{\theta} = X^\dagger Y$, où $X^\dagger$ est la pseudo-inverse (de Moore-Penrose) de X.

4. Montrez que $\hat{\theta}_\lambda$ tend vers $\hat{\theta}$ quand λ tend vers 0 par valeurs positives. (Indice : commencer par le cas où $d = n = 1$, puis le cas où X est diagonale, puis le cas général.)

5. Montrez qu'une simple descente de gradient à pas fixe sur $f(\theta) := \|Y - X\theta\|^2$ converge vers $\hat{\theta}$. Précisez comment choisir le pas. (Vous pouvez vous inspirer de l'exercice sur la convergence de la descente de gradient dans le cas lisse et fortement convexe.)

6. Calculez le biais de $\hat{\theta}_\lambda$.

Solution : Notons ϵ le vecteur colonne contenant $\epsilon_1, \ldots, \epsilon_n$. On a :

$$\begin{aligned}
\mathbf{E}[\hat{\theta}_\lambda] &= \mathbf{E}\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^TY\right] \\
&= \mathbf{E}\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^T(X\theta^* + \epsilon)\right] \\
&= \mathbf{E}_X\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^TX\theta^*\right] + \mathbf{E}_{X,\epsilon}\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^T\epsilon\right].
\end{aligned}$$

Or :

$$\begin{aligned}
\mathbf{E}_{X,\epsilon}\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^T\epsilon\right] &= \mathbf{E}_X\left[\mathbf{E}_\epsilon\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^T\epsilon\right]\right] \\
&= \mathbf{E}_X\left[\frac{1}{n}\left(\frac{X^TX}{n} + \lambda I\right)^{-1} X^T\mathbf{E}_\epsilon(\epsilon)\right] \\
&= 0.
\end{aligned}$$

Donc :

$$\begin{aligned}
\mathbf{E}[\hat{\theta}_\lambda] &= \mathbf{E}_X\left[\left(\frac{X^TX}{n} + \lambda I\right)^{-1} \frac{X^TX}{n}\theta^*\right] \\
&= \mathbf{E}\left[\left(\hat{\Sigma} + \lambda I\right)^{-1} \hat{\Sigma}\right]\theta^*,
\end{aligned}$$

où on a noté $\hat{\Sigma} := \frac{X^TX}{n}$ la covariance empirique.
Le biais de $\hat{\theta}_\lambda$ est donc $\left(\mathbf{E}\left[\left(\hat{\Sigma} + \lambda I\right)^{-1} \hat{\Sigma}\right] - I\right)\theta^*$.

7. Calculez $v(\hat{\theta}_\lambda) := \mathbf{E}\left[\left\|\hat{\theta}_\lambda - \mathbf{E}[\hat{\theta}_\lambda]\right\|_2^2\right]$, la variance totale de $\hat{\theta}_\lambda$.

Solution :

$$
\begin{aligned}
v(\hat{\theta}_\lambda) &= \mathbf{E}\left[\hat{\theta}_\lambda^T \hat{\theta}_\lambda\right] - \mathbf{E}\left[\hat{\theta}_\lambda\right]^T \mathbf{E}\left[\hat{\theta}_\lambda\right] \\
&= \mathbf{E}\left[\frac{(X\theta^* + \epsilon)^T X}{n}\left(\frac{X^T X}{n} + \lambda I\right)^{-2} \frac{X^T(X\theta^* + \epsilon)}{n}\right] - \left\|\mathbf{E}\left[\left(\hat{\Sigma} + \lambda I\right)^{-1}\hat{\Sigma}\right]\right\|_2^2 \\
&= 0 + 0 + 0 + \mathbf{E}\left[(\theta^*)^T \hat{\Sigma}\left(\hat{\Sigma} + \lambda I\right)^{-2}\hat{\Sigma}\theta^*\right] - \left\|\mathbf{E}\left[\left(\hat{\Sigma} + \lambda I\right)^{-1}\hat{\Sigma}\right]\right\|_2^2 \\
&= \mathbf{E}\left[\left\|\left(\hat{\Sigma} + \lambda I\right)^{-1}\hat{\Sigma}\right\|_2^2\right] - \left\|\mathbf{E}\left[\left(\hat{\Sigma} + \lambda I\right)^{-1}\hat{\Sigma}\right]\right\|_2^2 .
\end{aligned}
$$

8. Calculez le biais et la variance (totale) de $\hat{\theta}$ et commentez.

Exercice 31 Weyl's inequality. ($\star\star$)

Nous allons prouver un résultat qui peut être utile, par exemple, pour analyser les propriétés statistiques de l'estimateur usuel utilisé dans le cadre d'une analyse en composantes principales.

Considérons une matrice symmétrique M à coefficients réels et de taille d par d à laquelle on ajoute une perturbation, représentée par une autre matrice symétrique P à coefficients réels et de taille d par d. On cherche à trouver des conditions sous lesquelles on peut garantir que les valeurs propres de M et celles de $M' = M + P$ sont proches.

Pour toute matrice symmétrique A à coefficients réels et de taille d par d, on note $\lambda_1(A), \lambda_2, \ldots, \lambda_d(A)$ les valeurs propres de A triées par ordre décroissant.

1. Montrer que :

$$|\lambda_1(M') - \lambda_1(M)| \leq \|P\|.$$

   (Indice : nous avons défini et prouvé certaines propriétés de la norme matricielle $\|.\|_2$ en cours.)

2. Soit $i \in \{2, 3, \ldots, d\}$. Notons $\mathcal{E}_i^d$ l'ensemble de tous les sous-espaces de dimension i de $\mathbf{R}^d$. Montrer que pour toute matrice symmétrique A à coefficients réels et de taille d par d, la i-ième valeur propre de A est donnée par :

$$\lambda_i(A) = \min_{E \in \mathcal{E}_i^d} \max_{v \in E^\perp \cap \mathcal{S}^{d-1}} v^T A v,$$

   où $E^\perp$ est l'ensemble des vecteurs de $\mathbf{R}^d$ orthogonaux à tous les éléments de E et $\mathcal{S}^{d-1}$ est l'ensemble des vecteurs de $\mathbf{R}^d$ de norme euclidienne 1.

3. Montrer que :

$$\max_{i\in\{1,2,\ldots,d\}} |\lambda_i(M') - \lambda_i(M)| \leq \|P\|_2.$$

Exercice 32 Preuve de l'existence de la décomposition en valeurs singulières. ($\star\star$)

1. Prouver que toute matrice semi-orthogonale de taille n par r, avec $r \leq n$, peut être complétée en une matrice orthogonale de taille n par n.

2. Soit A une matrice réelle de taille m par n. Montrer qu'il est possible de trouver une matrice colonne x de taille n et une matrice colonne y de taille m, telles que : $Ax = \sigma y$, $\|x\|_2 = 1$, $\|y\|_2 = 1$ et $\sigma = \|A\|_2$. Indice : utiliser l'une des deux définitions équivalentes de $\|A\|_2$.

3. En utilisant le résultat de la première question, x et y, montrer qu'il existe une matrice orthogonale U de taille m par m, une matrice orthogonale V de taille n par n, une matrice colonne w de taille $n-1$ et une matrice B de taille $m-1$ par $n-1$, telles que :

$$A_1 := U^T A V = \begin{pmatrix} \sigma & w^T \\ 0 & B \end{pmatrix}.$$

4. Montrer que $\|A_1\|_2^2 \geq \sigma^2 + w^T w$. Indice : multiplier A_1 par $(\sigma\ w^T)^T$.

5. Montrer que $\|A_1\|_2^2 = \|A\|_2^2$ et en déduire la valeur de w.

6. Utiliser un raisonnement par récurrence pour compléter la preuve de l'existence de la décomposition en valeurs singulières.

5 Algèbre linéaire et probabilité

Exercice 33 Gaussienne multivariée. ($\star\star$)

1. Montrez que la densité d'un gaussienne multivariée est maximale en $x = \mu$.

2. Exprimez la vitesse de décroissance d'une gaussienne multivariée en $x = \mu$ le long d'un vecteur u de norme 1 en fonction des éléments propres de la matrice de covariance de la distribution.