



## La méthode des moindres carrés

Ce problème comporte 5 parties. Les parties A, B, C et D sont indépendantes entre elles.

L'objet de ce sujet est l'étude d'une méthode très utile en intelligence artificielle : la méthode des moindres carrés. Les parties A et C présentent des domaines des mathématiques que l'on peut reformuler en termes de problème des moindres carrés. La partie D amène à résoudre des systèmes d'équations linéaires au sens des moindres carrés à l'aide de la notion de décomposition en valeurs singulières. La partie B introduit la notion de limites de suites de matrices, que l'on utilise dans la partie E pour résoudre le problème étudié dans la partie D.

## Rappels et notations

- Si  $k$  et  $l$  sont deux entiers tels que  $k \leq l$ , on note  $[[k, l]]$  l'ensemble des entiers  $i$  tels que  $k \leq i \leq l$ .
- L'espace vectoriel des polynômes à coefficients dans  $\mathbb{R}$  est noté  $\mathbb{R}[X]$ . Pour  $n \in \mathbb{N}$ , on note  $\mathbb{R}_n[X]$  l'espace vectoriel des polynômes à coefficients dans  $\mathbb{R}$  de degré inférieur ou égal à  $n$ .
- Pour  $n \in \mathbb{N}^*$  et  $p \in \mathbb{N}^*$ , on note  $M_{n,p}(\mathbb{R})$  l'espace vectoriel des matrices à  $n$  lignes et  $p$  colonnes. On note aussi  $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$  l'espace vectoriel des matrices carrées de taille  $n$ .
- La transposée d'une matrice  $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in M_{n,p}(\mathbb{R})$  est notée  $A^T$ .

Elle est définie par  $A^T = (a_{j,i})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in M_{p,n}(\mathbb{R})$ .

- La trace d'une matrice  $A = (a_{i,j})_{1 \leq i,j \leq n} \in M_n(\mathbb{R})$  est notée  $\text{Tr}(A)$ . Elle est définie par  $\text{Tr}(A) = \sum_{i=1}^n a_{i,i}$ .
- Le produit scalaire canonique sur l'espace vectoriel  $M_{n,1}(\mathbb{R})$  des vecteurs colonnes à  $n$  lignes est noté  $\langle \cdot, \cdot \rangle$  et sa norme associée est notée  $\|\cdot\|_2$ . Ce produit scalaire est défini par :

$$\forall A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \forall B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad \langle A, B \rangle = A^T B = \sum_{i=1}^n a_i b_i \quad \text{et} \quad \|A\|_2 = \sqrt{\sum_{i=1}^n a_i^2}.$$

- Le produit scalaire canonique sur l'espace vectoriel  $M_n(\mathbb{R})$  des matrices carrées est aussi noté  $\langle \cdot, \cdot \rangle$  et sa norme associée est notée  $\|\cdot\|_2$ . Ce produit scalaire est défini par :

$$\forall A = (a_{i,j})_{1 \leq i,j \leq n}, \forall B = (b_{i,j})_{1 \leq i,j \leq n}, \quad \langle A, B \rangle = \text{Tr}(A^T B) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} b_{i,j} \quad \text{et} \quad \|A\|_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2}.$$

- L'ensemble des valeurs propres complexes d'une matrice  $A \in M_n(\mathbb{R})$  s'appelle le spectre de  $A$  et se note  $\text{Sp}(A)$ .

## Partie A – Quelques problèmes de minimisation et de convergence

## I – Distance entre un vecteur et un sous-espace vectoriel

On munit  $M_{4,1}(\mathbb{R})$  de son produit scalaire canonique. Soit  $F = \text{Vect}(e_1, e_2, e_3)$  avec

$$e_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

**Q1.** Justifier que  $(e_1, e_2, e_3)$  est une base orthogonale de  $F$ . Proposer une base orthonormée de  $F$ .

**Q2.** Calculer le projeté orthogonal sur  $F$  d'un vecteur colonne  $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in M_{4,1}(\mathbb{R})$ .

**Q3.** On note  $v = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ . Calculer  $\min_{u \in F} \|v - u\|_2$ .

## II – Séries de Fourier

Soit  $E = \mathcal{C}_{2\pi}^0(\mathbb{R}, \mathbb{R})$  l'espace vectoriel des fonctions continues  $2\pi$ -périodiques. On munit  $E$  du produit scalaire défini par

$$\forall (f, g) \in E^2, \quad \varphi(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f(t)g(t) dt.$$

On note  $\|\cdot\|$  la norme associée. Les coefficients de Fourier d'une fonction  $f \in \mathcal{C}_{2\pi}^0(\mathbb{R}, \mathbb{C})$  sont définis (pour  $n \geq 1$ ) par

$$a_0(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt \quad a_n(f) = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt) dt \quad b_n(f) = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt) dt$$

Pour tout  $n \in \mathbb{N}$ , nous introduisons les fonctions  $e_n : x \mapsto \cos(nx)$  et  $f_n : x \mapsto \sin(nx)$ .

Les sommes partielles de la série de Fourier d'une fonction  $f \in E$  sont les fonctions définies par

$$\forall N \in \mathbb{N}^*, \quad S_N(f) = a_0(f) + \sum_{n=1}^N (a_n(f)e_n + b_n(f)f_n).$$

**Q4.** Soient  $\alpha \in \mathbb{R} \setminus \mathbb{Z}$  et  $f$  la fonction  $2\pi$ -périodique telle que  $f(x) = \cos(\alpha x)$  pour tout  $x \in ]-\pi, \pi]$ .

Montrer que pour tout  $n \geq 1$ , on a  $b_n(f) = 0$  et  $a_0(f) = \frac{\sin(\alpha\pi)}{\alpha\pi}$  et  $a_n(f) = \frac{(-1)^n 2\alpha \sin(\alpha\pi)}{\pi(\alpha^2 - n^2)}$ .

**Q5.** Montrer que les séries  $\sum_{n \geq 1} \frac{1}{n^2 - \alpha^2}$  et  $\sum_{n \geq 1} \frac{1}{(n^2 - \alpha^2)^2}$  convergent, et calculer  $\sum_{n=1}^{+\infty} \frac{1}{n^2 - \alpha^2}$  et  $\sum_{n=1}^{+\infty} \frac{1}{(n^2 - \alpha^2)^2}$ .

**Q6.** Montrer que pour tout  $t \in \mathbb{R} \setminus \pi\mathbb{Z}$ , on a  $\frac{\cos(t)}{\sin(t)} = \frac{1}{t} + \sum_{n=1}^{+\infty} \frac{2t}{t^2 - (n\pi)^2}$ .

Soit  $f : \mathbb{R} \rightarrow \mathbb{C}$  une fonction  $2\pi$ -périodique de classe  $\mathcal{C}^2$  qui est solution de l'équation différentielle

$$f'' + e^{ix}f = 0. \quad (\mathcal{E})$$

Pour tout  $n \in \mathbb{Z}$ , on pose  $c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t)e^{-int} dt$  et  $c_n(f'') = \frac{1}{2\pi} \int_0^{2\pi} f''(t)e^{-int} dt$ .

**Q7.** Montrer que pour tout  $n \in \mathbb{N}^*$  on a  $c_n(f) = \frac{a_n(f) - ib_n(f)}{2}$  et  $c_{-n}(f) = \frac{a_n(f) + ib_n(f)}{2}$ .

Montrer que pour tout  $n \in \mathbb{N}^*$  on a  $a_n(f) = c_n(f) + c_{-n}(f)$  et  $b_n(f) = i(c_n(f) - c_{-n}(f))$ .

**Q8.** Soit  $n \in \mathbb{Z}$ .

Montrer que  $c_n(f'') = -n^2 c_n(f)$ , puis que  $c_{n-1}(f) = n^2 c_n(f)$ .

En déduire que  $c_n(f) = \begin{cases} \frac{c_0(f)}{(n!)^2} & \text{si } n \geq 0 \\ 0 & \text{si } n < 0 \end{cases}$ .

**Q9.** Montrer que pour tout  $x \in \mathbb{R}$ , on a  $f(x) = c_0(f) + \sum_{n=1}^{+\infty} (c_n(f)e^{inx} + c_{-n}(f)e^{-inx}) = c_0(f) \sum_{n=0}^{+\infty} \frac{e^{inx}}{(n!)^2}$ .

On admet que la fonction définie par  $x \mapsto \sum_{n=0}^{+\infty} \frac{e^{inx}}{(n!)^2}$  est bien une solution de l'équation  $(\mathcal{E})$ .

**Q10.** On rappelle que l'ensemble des solutions de l'équation  $(\mathcal{E})$  est un sous-espace vectoriel de  $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ .  
Quelle est sa dimension ? Justifier. L'équation  $(\mathcal{E})$  admet-elle une solution qui n'est pas  $2\pi$ -périodique ?

**Q11.** Montrer que, pour tout  $N \in \mathbb{N}^*$ , la famille  $(e_0, \dots, e_N, f_1, \dots, f_N)$  est orthogonale pour  $\varphi$ . Est-elle orthonormale ?

**Q12.** Soient  $g \in E$  et  $N \in \mathbb{N}^*$ . On note  $V_N = \text{Vect}(e_0, \dots, e_N, f_1, \dots, f_N)$ .  
Quelle est la projection orthogonale de  $g$  sur  $V_N$  ? En déduire une fonction  $h \in E$  telle que  $\|g-h\| = \min_{v \in V_N} \|g-v\|$ .

## Partie B – Limites de suites de Matrices

Soit  $(M_k)_{k \in \mathbb{N}}$  une suite de matrices de  $M_{n,p}(\mathbb{R})$ , avec  $M_k = \begin{pmatrix} m_{k,1,1} & \dots & m_{k,1,p} \\ \vdots & & \vdots \\ m_{k,n,1} & \dots & m_{k,n,p} \end{pmatrix}$  (pour tout  $k \in \mathbb{N}$ ).

On dit que la suite  $(M_k)_{k \in \mathbb{N}}$  converge si pour tout  $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$  la suite  $(m_{k,i,j})_{k \in \mathbb{N}}$  converge. On note alors

$$\lim_{k \rightarrow +\infty} M_k = \begin{pmatrix} \lim_{k \rightarrow +\infty} (m_{k,1,1}) & \dots & \lim_{k \rightarrow +\infty} (m_{k,1,p}) \\ \vdots & & \vdots \\ \lim_{k \rightarrow +\infty} (m_{k,n,1}) & \dots & \lim_{k \rightarrow +\infty} (m_{k,n,p}) \end{pmatrix}.$$

Lorsque la suite  $(M_k)_{k \in \mathbb{N}}$  ne converge pas, on dit que la suite  $(M_k)_{k \in \mathbb{N}}$  diverge.

On admettra sans preuve que, si  $(M_k)_{k \in \mathbb{N}}$  converge, alors, pour tout  $(q, r) \in \mathbb{N}^* \times \mathbb{N}^*$ , toute matrice  $P \in M_{q,n}(\mathbb{R})$  et toute matrice  $Q \in M_{p,r}(\mathbb{R})$  la suite  $(PM_kQ)_{k \in \mathbb{N}}$  converge et  $\lim_{k \rightarrow +\infty} (PM_kQ) = P \left( \lim_{k \rightarrow +\infty} M_k \right) Q$ .

**Q13.** Soient  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  et  $B = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$ . Montrer que les suites  $(A^k)_{k \in \mathbb{N}}$  et  $(B^k)_{k \in \mathbb{N}}$  divergent.

**Q14.** Soit  $A$  une matrice diagonalisable sur  $\mathbb{R}$ . Montrer que la suite  $(A^k)_{k \in \mathbb{N}}$  converge si et seulement si  $\text{Sp}(A) \subset ]-1, 1]$ .

**Q15.** Soit  $A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ . Montrer que  $A$  est diagonalisable sur  $\mathbb{R}$ , et que la suite de matrices  $(A^k)_{k \in \mathbb{N}}$  converge.

**Q16.** Montrer que les matrices  $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}$  et  $T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$  sont semblables.

La suite de matrices  $(A^k)_{k \in \mathbb{N}}$  converge-t-elle ?

Dans la suite de l'énoncé, on pourra utiliser sans preuve le résultat suivant :

si  $M \in M_{n,p}(\mathbb{R})$  est une matrice et si  $(M_k)_{k \in \mathbb{N}}$  est une suite de matrices de  $M_{n,p}(\mathbb{R})$ , alors

$$\lim_{k \rightarrow +\infty} M_k = M \quad \text{si et seulement si} \quad \lim_{k \rightarrow +\infty} \|M_k - M\|_2 = 0.$$

## Partie C – Interpolation polynômiale et moindres carrés

Soient  $n \in \mathbb{N}^*$  et  $(x_0, \dots, x_n) \in \mathbb{R}^{n+1}$  une famille de nombres réels deux à deux distincts. Pour chaque  $i \in \llbracket 0, n \rrbracket$ , on note

$$L_i(X) = \prod_{0 \leq j \leq n, j \neq i} \frac{X - x_j}{x_i - x_j}.$$

**Q17.** Montrer que pour tout  $(i, k) \in \llbracket 0, n \rrbracket^2$ , on a  $L_i(x_k) = \begin{cases} 1 & \text{si } k = i \\ 0 & \text{sinon} \end{cases}$

**Q18.** Montrer que l'on peut munir  $\mathbb{R}_n[X]$  d'un produit scalaire  $\langle \cdot, \cdot \rangle$  en posant  $\langle P, Q \rangle = \sum_{k=0}^n P(x_k)Q(x_k)$ .

**Q19.** Montrer que  $(L_0, \dots, L_n)$  est une base orthonormée de  $\mathbb{R}_n[X]$  muni du produit scalaire  $\langle \cdot, \cdot \rangle$ .

**Q20.** Montrer que pour tout  $P \in \mathbb{R}_n[X]$  et tout  $i \in \llbracket 0, n \rrbracket$ , on a  $\langle L_i, P \rangle = P(x_i)$ , puis que  $P = \sum_{i=0}^n P(x_i)L_i$ .

Soit  $(a, b) \in \mathbb{R}^2$  tel que pour tout  $i \in \llbracket 0, n \rrbracket$  on ait  $a < x_i < b$ . Soit  $f \in \mathcal{C}^{n+1}([a, b], \mathbb{R})$ . On note  $P_f = \sum_{i=0}^n f(x_i)L_i$ .

**Q21.** Soient  $m \leq n$ . En calculant  $\langle P_f - Q, P_f - Q \rangle$ , montrer l'existence et l'unicité d'un polynôme  $R \in \mathbb{R}_m[X]$  tel que

$$\sum_{i=0}^n (f(x_i) - R(x_i))^2 = \min_{Q \in \mathbb{R}_m[X]} \left( \sum_{i=0}^n (f(x_i) - Q(x_i))^2 \right).$$

**Q22.** Soit  $g \in \mathcal{C}^{n+1}([a, b], \mathbb{R})$ . On suppose que  $g$  possède au moins  $n+2$  zéros deux à deux distincts dans  $[a, b]$ . Montrer par récurrence sur  $n$  que  $g^{(n+1)}$  a au moins un zéro dans  $[a, b]$ . On pourra utiliser le théorème de Rolle.

**Q23.** Soit  $x \in [a, b]$  avec  $x \neq x_i$  pour tout  $i \in \llbracket 0, n \rrbracket$ . Soit  $K = \frac{f(x) - P_f(x)}{\prod_{i=0}^n (x - x_i)}$ .

Montrer que  $W : t \mapsto f(t) - P_f(t) - K \prod_{i=0}^n (t - x_i)$  s'annule en  $n+2$  points distincts de  $[a, b]$ .

En déduire l'existence de  $\xi_x \in [a, b]$  tel que  $f(x) - P_f(x) = \frac{\prod_{i=0}^n (x - x_i)}{(n+1)!} f^{(n+1)}(\xi_x)$ .

## Partie D – Régression linéaire et moindres carrés

### I – Le cas particulier de la dimension 2

Soit  $n \geq 2$  un entier. Considérons  $n$  points du plan réel  $(x_1, y_1), \dots, (x_n, y_n)$  avec  $x_1, \dots, x_n$  deux à deux distincts tels que  $(x_1, \dots, x_n)$  ne soit pas colinéaire au vecteur  $(1, \dots, 1)$ .

En dimension 2, la régression linéaire consiste à trouver une relation affine entre deux grandeurs physiques, c'est-à-dire à chercher une droite affine s'approchant le plus possible des points  $(x_1, y_1), \dots, (x_n, y_n)$  (au sens des moindres carrés).

Pour formuler ce problème de manière précise et rigoureuse, nous notons

$$E(a, b) = \sum_{i=1}^n (y_i - a - bx_i)^2 \quad \text{et} \quad A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Le problème est alors de trouver les valeurs de  $a$  et  $b$  qui minimisent la fonction  $E$ .

**Q24.** Montrer que  $E$  admet un minimum  $m$  sur  $\mathbb{R}^2$  (qu'on ne demande pas de calculer) et que  $m = \min_{(a,b) \in \mathbb{R}^2} \left\| A \begin{pmatrix} a \\ b \end{pmatrix} - Y \right\|_2^2$ .

**Q25.** Montrer que  $\det(A^T A) = n \left( \sum_{k=1}^n x_k^2 \right) - \left( \sum_{k=1}^n x_k \right)^2$ , puis justifier l'affirmation  $\det(A^T A) > 0$ .

Indication : On pourra utiliser l'inégalité de Cauchy-Schwarz

**Q26.** Calculer le gradient de  $E : (a, b) \mapsto E(a, b)$ .

Soit  $(a_0, b_0)$  un point critique de  $E$ . Montrer que l'on a  $A^T A \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = A^T Y$ .

En déduire qu'il existe un unique couple  $(c, d) \in \mathbb{R}^2$  tel que  $E(c, d) = \min_{(a,b) \in \mathbb{R}^2} E(a, b)$ .

## II – Les moindres carrés : le cas général

Nous allons maintenant généraliser les idées de la sous-section précédente au cas de la dimension quelconque.

Soit  $A \in M_{n,p}(\mathbb{R})$ . Notons  $\text{Col}(A) = \text{Vect}(C_1, \dots, C_p)$  avec  $C_1, \dots, C_p$  les colonnes de  $A$ . Soit  $y \in M_{n,1}(\mathbb{R})$ .

**Q27.** Soit  $x \in M_{p,1}(\mathbb{R})$  tel que  $A^T Ax = 0$ . Montrer que  $\|Ax\|_2 = 0$ .

**Q28.** Montrer que  $\text{Ker}(A^T A) = \text{Ker}(A)$ , puis que  $\text{rg}(A^T A) = \text{rg}(A)$ .

**Q29.** Montrer l'existence d'un unique couple  $(u, d) \in M_{p,1}(\mathbb{R}) \times (\text{Col}(A))^\perp$  tel que  $y = Au + d$ .

**Q30.** Montrer que pour tout  $x \in M_{p,1}(\mathbb{R})$ , on a  $\|Ax - y\|_2 \geq \|d\|_2$  avec égalité si et seulement si  $x - u \in \text{Ker}(A)$ .

**Q31.** Montrer que  $A^T y = A^T Au$ .

**Q32.** Soit  $x_0 \in M_{p,1}(\mathbb{R})$ . Montrer que :  $\|Ax_0 - y\|_2 = \min_{x \in M_{p,1}(\mathbb{R})} \|Ax - y\|_2 \iff A^T Ax_0 = A^T y$ .

**Q33.** Supposons  $n > p$  et que la matrice  $A$  est de rang maximal, c'est-à-dire que le rang de  $A$  est égal à  $p$ .

Montrer que  $A^T A$  est inversible.

En déduire que l'équation  $\|Ab - y\|_2 = \min_{x \in M_{p,1}(\mathbb{R})} \|Ax - y\|_2$  a une unique solution donnée par  $b = (A^T A)^{-1} A^T y$ .

## III – La décomposition en valeurs singulières

Soient  $A \in M_{n,p}(\mathbb{R})$  non nulle et  $y \in M_{n,1}(\mathbb{R})$ . Nous allons généraliser le résultat de la question **Q33**.

Dans la suite, nous notons  $\text{Diag}_{k,l}(\alpha_1, \dots, \alpha_{\min(k,l)})$  la matrice de taille  $k \times l$  dont le coefficient en ligne  $i$  et colonne  $i$  est  $\alpha_i$  (pour tout indice  $i \leq \min(k, l)$ ), et dont tous les autres coefficients sont nuls.

**Q34.** Montrer que pour tout vecteur colonne  $X \in M_{p,1}(\mathbb{R})$ , les nombres  $X^T A^T A X$  et  $X^T X$  sont positifs ou nuls.

**Q35.** Montrer que les valeurs propres de la matrice  $A^T A$  sont toutes positives ou nulles.

**Q36.** Notons  $r = \text{rg}(A^T A) = \text{rg}(A)$  (égalité prouvée à la question **Q28**).

Montrer qu'il existe des nombres réels  $\lambda_1 \geq \dots \geq \lambda_r > 0$  et une matrice orthogonale  $V \in O_p(\mathbb{R})$  tels que  $A^T A = V \text{Diag}_{p,p}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) V^T$ .

**Q37.** Notons  $v_i$  la  $i$ -ème colonne de  $V$ .

Montrer que la famille  $(Av_1, \dots, Av_r)$  est orthogonale pour le produit scalaire canonique de  $M_{n,1}(\mathbb{R})$ , et que,

pour tout  $i \in \llbracket 1, p \rrbracket$ , on a  $\|Av_i\|_2 = \begin{cases} \sqrt{\lambda_i} & \text{si } i \leq r \\ 0 & \text{si } i \geq r + 1 \end{cases}$ .

Montrer l'existence d'une base orthonormée  $(u_1, \dots, u_n)$  de  $M_{n,1}(\mathbb{R})$  telle que  $u_i = \frac{1}{\sqrt{\lambda_i}} Av_i$  pour tout  $i \leq r$ .

**Q38.** Notons  $U$  la matrice dont la  $i$ -ème colonne est  $u_i$  et  $D = \text{Diag}_{n,p}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}, 0 : \dots, 0) \in M_{n,p}(\mathbb{R})$ .

Vérifier que  $U \in O_n(\mathbb{R})$  et que  $AV = UD$ .

En particulier, on a  $A = UDV^T$ . Cette écriture s'appelle une décomposition en valeur singulière de  $A$ . Posons

$$\Delta = \text{Diag}_{p,n} \left( \frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_r}}, 0, \dots, 0 \right) \in M_{p,n}(\mathbb{R}) \quad \text{et} \quad \tilde{A} = V \Delta U^T \in M_{p,n}(\mathbb{R})$$

**Q39.** Montrer que  $A\tilde{A}$  est la matrice de la projection orthogonale sur  $\text{Im}(A)$ .

**Q40.** Soit  $y \in M_{n,1}(\mathbb{R})$ . Soit  $b = \tilde{A}y \in M_{p,1}(\mathbb{R})$ . Vérifier que  $\|Ab - y\|_2 = \min_{x \in M_{p,1}(\mathbb{R})} \|Ax - y\|_2$ .

Soit  $x_0 \in M_{p,1}(\mathbb{R})$  un autre vecteur qui vérifie  $\|Ax_0 - y\|_2 = \min_{x \in M_{p,1}(\mathbb{R})} \|Ax - y\|_2$ . Nous allons montrer que  $\|x_0\|_2 \geq \|b\|_2$ .

**Q41.** Montrer que  $Ab - y \in \text{Im}(A)^\perp$ . En déduire que  $x_0 - b \in \text{Ker}(A)$ .

**Q42.** Montrer que  $b \in \text{Ker}(A)^\perp$ .

Indication : on pourra exprimer  $\text{Im}(\tilde{A})$  et  $\text{Ker}(A)$  en fonction des vecteurs colonnes  $v_i$ .

**Q43.** Montrer que  $\|x_0\|_2 \geq \|b\|_2$  avec égalité si et seulement si  $x_0 = b$ .

## Partie E – La méthode du gradient

Soit  $A \in M_{n,p}(\mathbb{R})$  et  $y \in M_{n,1}(\mathbb{R})$ . Nous avons vu à la question **Q35** que les valeurs propres  $\lambda_1 \leq \dots \leq \lambda_p$  de  $A^T A$  sont des nombres réels positifs ou nuls. Dans cette partie, nous supposons en plus que pour tout  $i$ , on a  $\lambda_i > 0$ .

Nous reformulons notre problème des moindres carrés en remarquant (par le calcul) que

$$\forall x \in M_{p,1}(\mathbb{R}), \quad \|Ax - y\|_2^2 = \langle A^T Ax, x \rangle - 2\langle x, A^T y \rangle + \langle y, y \rangle.$$

Nous allons étudier une suite de matrices convergente dont la limite sera un minimum de la fonction

$$f : \begin{cases} M_{p,1}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \langle A^T Ax, x \rangle - 2\langle x, A^T y \rangle \end{cases}.$$

Soit  $(e_1, \dots, e_p)$  une base orthonormale de  $M_{p,1}(\mathbb{R})$  telle que pour tout  $i \in \llbracket 1, p \rrbracket$  on ait  $A^T A e_i = \lambda_i e_i$ .

**Q44.** Montrer que la matrice  $A^T A$  est inversible, et justifier que  $\text{Ker}(A) = \{0\}$ .

**Q45.** Montrer que l'application  $\begin{cases} M_{p,1}(\mathbb{R}) \times M_{p,1}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & \langle A^T Ax, y \rangle \end{cases}$  est un produit scalaire.

Soient  $(v_1, \dots, v_p) \in \mathbb{R}^p$  et  $v = \sum_{i=1}^p v_i e_i$ .

**Q46.** Montrer que pour tout  $(\alpha, \beta) \in \mathbb{R}^2$  on a :  $\alpha\beta \leq \frac{1}{4}(\alpha + \beta)^2$ .

$$\text{Montrer que } \langle A^T Av, v \rangle \langle (A^T A)^{-1} v, v \rangle \leq \frac{1}{4} \frac{\lambda_1}{\lambda_p} \left( \sum_{i=1}^p \left( \frac{\lambda_i}{\lambda_1} + \frac{\lambda_p}{\lambda_i} \right) v_i^2 \right)^2.$$

**Q47.** Soit  $\psi : t \mapsto \frac{t}{\lambda_1} + \frac{\lambda_p}{t}$ . Calculer le maximum de  $\psi$  sur  $[\lambda_1, \lambda_p]$ .

**Q48.** On note  $c(A) = \frac{\lambda_p}{\lambda_1}$ . Montrer l'inégalité de Kantorovich :  $\langle A^T Av, v \rangle \langle (A^T A)^{-1} v, v \rangle \leq \frac{(c(A) + 1)^2}{4c(A)} \|v\|_2^4$ .

Nous souhaitons maintenant déterminer une approximation de l'unique solution  $\ell$  à l'équation  $A^T A \ell = A^T y$ .

Nous définissons une suite par récurrence en choisissant un vecteur  $x_0 \in M_{p,1}(\mathbb{R})$  et en posant pour tout  $k \in \mathbb{N}$  :

$$d_k = A^T Ax_k - A^T y \quad t_k = \frac{\|d_k\|_2^2}{\langle A^T A d_k, d_k \rangle} \quad x_{k+1} = x_k - t_k d_k.$$

Pour que cette définition ait un sens, nous supposons que le vecteur  $d_k$  est non nul pour tout  $k \in \mathbb{N}$ . Cette hypothèse n'est pas restrictive : s'il existe  $k$  tel que  $d_k = 0$  alors  $x_k = \ell$  et nous avons ainsi la valeur exacte de  $\ell$ .

Soit  $\ell = (A^T A)^{-1} A^T y \in M_{p,1}(\mathbb{R})$ .

**Q49.** Vérifier que pour tout  $x \in M_{p,1}(\mathbb{R})$ , on a  $f(x) - f(\ell) = \langle A^T A(x - \ell), x - \ell \rangle$ .

**Q50.** Montrer que pour tout  $k \in \mathbb{N}$ , on a  $f(x_{k+1}) - f(\ell) = (f(x_k) - f(\ell)) \left( 1 - \frac{\|d_k\|_2^4}{\langle A^T A d_k, d_k \rangle \langle (A^T A)^{-1} d_k, d_k \rangle} \right)$ .

**Q51.** Montrer que pour tout  $k \in \mathbb{N}$ , on a  $\lambda_1 \|x_k - \ell\|_2^2 \leq f(x_k) - f(\ell) \leq (f(x_0) - f(\ell)) \left( \frac{c(A) - 1}{c(A) + 1} \right)^{2k}$ .

**Q52.** Montrer que la suite  $(x_k)_{k \in \mathbb{N}}$  converge vers  $\ell$ , et que  $\|A\ell - y\|_2 = \min_{x \in M_{p,1}(\mathbb{R})} \|Ax - y\|_2$ .

◇ Fin ◇