



PULSARS

CORRECTION INDÉPENDANTE

Centrale-Supélec 2026 — Mathématiques 1 (TSI)

SESSION 2026	NIVEAU CPGE	DURÉE —	FILIÈRES TSI
------------------------	-----------------------	-------------------	------------------------

Correction indépendante rédigée par Pulsars. Elle n'est ni officielle ni affiliée à l'organisateur du concours. Les énoncés ne sont pas reproduits : le sujet officiel reste consultable sur le site du concours.

Partie A · I — Distance à un sous-espace (Q1 à Q3)

$F = \text{Vect}(e_1, e_2, e_3)$ dans $M_{\{4,1\}}(\mathbb{R})$ muni du produit scalaire canonique.

- **Q1** — la famille est une base orthogonale ; en donner une orthonormée.

Q1. Vérifions les trois produits scalaires :

$$\begin{aligned}\langle e_1, e_2 \rangle &= 1 - 2 + 1 + 0 = 0, & \langle e_1, e_3 \rangle &= 1 + 0 - 1 + 0 = 0, \\ \langle e_2, e_3 \rangle &= 1 + 0 - 1 + 0 = 0.\end{aligned}$$

La famille est **orthogonale** et formée de vecteurs non nuls : elle est libre, donc c'est une base de F , qui est de dimension 3.

Les normes valent $\|e_1\| = \sqrt{6}$, $\|e_2\| = \sqrt{3}$, $\|e_3\| = \sqrt{2}$, d'où la base orthonormée

$$(e_1/\sqrt{6}, e_2/\sqrt{3}, e_3/\sqrt{2}).$$

- **Q2** — le projeté orthogonal d'un vecteur quelconque sur F .

Q2. Les trois vecteurs ont leur quatrième coordonnée nulle et sont indépendants : ils engendrent exactement l'hyperplan $\{t = 0\}$, qui est de dimension 3. Donc

$$F = \{(x, y, z, 0)\},$$

et le projeté orthogonal de (x, y, z, t) sur F est simplement $(x, y, z, 0)$.

On peut aussi appliquer la formule $\sum \langle u, e_i \rangle e_i / \|e_i\|^2$, mais elle demande trois calculs pour un résultat que la structure de F donne immédiatement.

- **Q3** — la distance de $v = (1, 1, 1, 1)$ à F .

Q3. La distance de v à F est la norme de la composante orthogonale :

$$v - p(v) = (0, 0, 0, 1), \text{ donc } \min_{u \in F} \|v - u\|_2 = 1.$$

Partie A · II — Séries de Fourier (Q4 à Q12)

- **Q4** — les coefficients de Fourier de $x \mapsto \cos(\alpha x)$, prolongée par 2π -périodicité, pour $\alpha \notin \mathbb{Z}$.
- **Q5, Q6** — deux sommes de séries, puis le développement de la cotangente.
- **Q7 à Q10** — coefficients complexes, et résolution de $f'' + e^{ix}f = 0$.
- **Q11, Q12** — orthogonalité de la famille trigonométrique, et la projection comme meilleure approximation.

Q4. La fonction f est **paire** — $\cos(\alpha \cdot)$ l'est sur $]-\pi, \pi]$, et la 2π -périodicité conserve la parité — donc tous les $b_n(f)$ sont nuls.

$$a_0(f) = (1/2\pi) \int_{-\pi}^{\pi} \cos(\alpha t) dt = (1/2\pi) \cdot [\sin(\alpha t)/\alpha]_{-\pi}^{\pi} = \sin(\alpha\pi)/(\alpha\pi).$$

Pour $n \geq 1$, la linéarisation $\cos(\alpha t)\cos(nt) = \frac{1}{2}[\cos((\alpha+n)t) + \cos((\alpha-n)t)]$ donne

$$a_n(f) = (1/\pi) \cdot [\sin((\alpha+n)\pi)/(\alpha+n) + \sin((\alpha-n)\pi)/(\alpha-n)].$$

Or $\sin((\alpha \pm n)\pi) = (-1)^n \sin(\alpha\pi)$, et $1/(\alpha+n) + 1/(\alpha-n) = 2\alpha/(\alpha^2 - n^2)$:

$$a_n(f) = (-1)^n \cdot 2\alpha \cdot \sin(\alpha\pi)/(\pi(\alpha^2 - n^2)).$$

Q5. Les termes généraux sont équivalents à $1/n^2$ et $1/n^4$: les deux séries **convergent** par comparaison aux séries de Riemann.

La fonction f est continue sur $\mathbb{R}$ — les valeurs en π et $-\pi$ coïncident, toutes deux égales à $\cos(\alpha\pi)$ — et de classe C^1 par morceaux : le théorème de Dirichlet s'applique. En $x = \pi$:

$$\cos(\alpha\pi) = a_0 + \sum_{n \geq 1} (-1)^n a_n \dots = \sin(\alpha\pi)/(\alpha\pi) - (2\alpha \sin(\alpha\pi)/\pi) \cdot \sum_{n \geq 1} 1/(n^2 - \alpha^2).$$

En isolant la somme :

$$\sum_{n \geq 1} 1/(n^2 - \alpha^2) = 1/(2\alpha^2) - \pi/(2\alpha \cdot \tan(\alpha\pi)).$$

Pour la seconde, la formule de Parseval — avec $\|f\|^2 = \frac{1}{2} + \sin(2\alpha\pi)/(4\alpha\pi)$ obtenu par linéarisation — donne

$$\sum_{n \geq 1} 1/(n^2 - \alpha^2)^2 = [\frac{1}{2} + \sin(2\alpha\pi)/(4\alpha\pi) - \sin^2(\alpha\pi)/(\alpha^2\pi^2)] \cdot \pi^2/(2\alpha^2 \sin^2(\alpha\pi)).$$

Q6. Reprenons la première somme et posons $t = \alpha\pi$, soit $\alpha = t/\pi$. La relation de **Q5** se réécrit

$$\pi \cdot \cotan(\alpha\pi) = 1/\alpha + 2\alpha \cdot \sum_{n \geq 1} 1/(\alpha^2 - n^2).$$

En substituant et en divisant par π , les π^2 se compensent :

$$\cos(t)/\sin(t) = 1/t + \sum_{n \geq 1} 2t/(t^2 - (n\pi)^2),$$

valable pour $t \notin \pi\mathbb{Z}$.

C'est le développement en éléments simples de la cotangente, dû à Euler : chaque pôle $n\pi$ apparaît explicitement.

Q7. En développant $e^{-int} = \cos(nt) - i \sin(nt)$ dans la définition de c_n et en comparant aux facteurs $1/\pi$ des définitions de a_n et b_n :

$$c_n(f) = (a_n(f) - i b_n(f))/2 \text{ et } c_{-n}(f) = (a_n(f) + i b_n(f))/2.$$

En additionnant et en soustrayant : $a_n = c_n + c_{-n}$ et $b_n = i(c_n - c_{-n})$.

Q8. Deux intégrations par parties, les termes de bord se compensant par 2π -périodicité de f , f' et e^{-int} :

$$c_n(f'') = (in) \cdot c_n(f') = (in)^2 \cdot c_n(f) = -n^2 c_n(f).$$

Par ailleurs (E) donne $f'' = -e^{ix}f$, donc

$$c_n(f'') = -(1/2\pi) \int f(t) e^{it} e^{-int} dt = -c_{n-1}(f).$$

En rapprochant : $c_{n-1}(f) = n^2 c_n(f)$.

Pour $n = 0$, cela donne $c_{-1}(f) = 0$, puis de proche en proche $c_n(f) = 0$ pour tout $n < 0$. Pour $n \geq 1$, la relation $c_n = c_{n-1}/n^2$ donne par récurrence

$$c_n(f) = c_0(f)/(n!)^2.$$

Q9. Les coefficients vérifient $|c_n(f)| \leq |c_0(f)|/(n!)^2$, terme général d'une série convergente : la série de Fourier converge **normalement**, donc uniformément, et sa somme vaut f puisque f est continue et C^1 par morceaux. Comme les coefficients d'indice négatif sont nuls :

$$f(x) = c_0(f) \cdot \sum_{n \geq 0} e^{inx}/(n!)^2.$$

Q10. (E) est une équation différentielle **linéaire d'ordre 2** à coefficients continus : son espace de solutions est de dimension **2** sur **$\mathbb{C}$** .

Or **Q9** montre que les solutions **2π** -périodiques forment une droite vectorielle — elles sont toutes proportionnelles à une même fonction. Comme **$1 < 2$** , **oui**, (E) admet des solutions non **2π** -périodiques.

Q11. Les calculs classiques donnent, pour **$n \neq m$** , **$\varphi(e_n, e_m) = \varphi(f_n, f_m) = 0$** et **$\varphi(e_n, f_m) = 0$** pour tous **$n, m$** : la famille est **orthogonale**.

Elle n'est pas orthonormée : **$\varphi(e_0, e_0) = 1$** , mais

$\varphi(e_n, e_n) = \varphi(f_n, f_n) = 1/2$ pour **$n \geq 1$** .

Q12. Dans une base orthogonale, le projeté s'obtient en divisant par le carré des normes. Avec **$\varphi(g, e_n) = a_n(g)/2$** et **$\|e_n\|^2 = 1/2$** pour **$n \geq 1$** , le quotient vaut exactement **$a_n(g)$** :

$p_{\{V_N\}}(g) = a_0(g)e_0 + \sum_{n=1}^N (a_n(g)e_n + b_n(g)f_n) = S_N(g)$.

La fonction **$h = S_N(g)$** réalise donc le minimum : **la somme partielle de Fourier est la meilleure approximation quadratique** de **g** par un polynôme trigonométrique de degré **N** .

Partie B — Limites de suites de matrices (Q13 à Q16)

- **Q13** — deux exemples de suites (A^k) divergentes.

Q13. Pour $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$: $A^2 = I_2$, donc la suite vaut alternativement A et I_2 . Ces deux matrices étant distinctes, la suite **diverge**.

Pour B , on reconnaît la matrice de la rotation d'angle $2\pi/3$: $B^3 = I_2$ et la suite est périodique de période 3, prenant trois valeurs distinctes. Elle **diverge** également.

- **Q14** — pour A diagonalisable, la convergence équivaut à une condition sur le spectre.

Q14. Écrivons $A = PDP^{-1}$ avec $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$. Alors $A^k = PD^kP^{-1}$, et le résultat admis donne l'équivalence entre la convergence de (A^k) et celle de $(D^k) = (P^{-1}A^kP)$.

Or (D^k) converge si et seulement si chaque suite (λ_i^k) converge, ce qui équivaut à $|\lambda_i| < 1$ ou $\lambda_i = 1$, c'est-à-dire

$$\text{Sp}(A) \subset]-1, 1].$$

- **Q15** — un cas convergent.

Q15. A est **symétrique réelle**, donc diagonalisable dans une base orthonormée (théorème spectral).

En remarquant que $2A = J - I_3$ où J est la matrice dont tous les coefficients valent 1, et que J a pour valeurs propres 3 (simple) et 0 (double), on obtient

$$\text{Sp}(A) = \{1, -1/2\}.$$

Ces valeurs appartiennent à $] -1, 1]$: d'après **Q14**, la suite (A^k) **converge**.

- **Q16** — deux matrices semblables, et la convergence de (A^k) .

Q16. A est la matrice compagnon du polynôme

$$X^3 + X^2 - X - 1 = (X - 1)(X + 1)^2,$$

qui est aussi le polynôme caractéristique de T , triangulaire de diagonale $(1, -1, -1)$.

Les deux matrices ont donc mêmes valeurs propres avec mêmes multiplicités ; de plus

$\text{rg}(A + I_3) = \text{rg}(T + I_3) = 2$, donc le sous-espace propre associé à -1 est de dimension 1 dans les deux cas : les deux matrices ont **même réduite**, elles sont **semblables**.

La suite (A^k) **ne converge pas** : $-1 \notin]-1, 1]$, et le bloc non diagonalisable aggrave encore la divergence.

Partie C — Interpolation polynomiale et moindres carrés (Q17 à Q23)

L_i désigne le i -ième polynôme de Lagrange associé aux points $x_0, \dots, x_n$.

- **Q17 à Q20** — propriété caractéristique des L_i , produit scalaire par évaluation, base orthonormée, et la décomposition d'un polynôme.
- **Q21** — la meilleure approximation aux moindres carrés par un polynôme de degré $\leq m$.
- **Q22, Q23** — l'erreur d'interpolation, via Rolle.

Q21. Comme $P_f(x_i) = f(x_i)$ par construction,

$$(P_f - Q, P_f - Q) = \sum_{i=0}^n (f(x_i) - Q(x_i))^2.$$

Minimiser cette somme sur $Q \in \mathbb{R}_m[X]$ revient donc à minimiser $\|P_f - Q\|$: le minimum est atteint en l'unique **projeté orthogonal** R de P_f sur $\mathbb{R}_m[X]$, sous-espace de dimension finie. D'où existence et unicité.

Q17. Pour $k = i$, chaque facteur du produit vaut $(x_i - x_j)/(x_i - x_j) = 1$, donc $L_i(x_i) = 1$.

Pour $k \neq i$, le facteur d'indice $j = k$ est $(x_k - x_k)/(x_i - x_k) = 0$, donc $L_i(x_k) = 0$. Les x_j étant deux à deux distincts, tous les dénominateurs sont non nuls.

Q18. L'application est clairement bilinéaire, symétrique et positive.

Pour le caractère défini : si $(P, P) = \sum P(x_k)^2 = 0$, alors P s'annule aux $n + 1$ points $x_0, \dots, x_n$, distincts. Or $\deg P \leq n$: un polynôme non nul de degré $\leq n$ a au plus n racines. Donc $P = 0$.

Q19. D'après **Q17**, $(L_i, L_j) = \sum_k L_i(x_k)L_j(x_k) = \sum_k \delta_{ik}\delta_{jk} = \delta_{ij}$.

La famille est **orthonormale**, donc libre, et compte $n + 1$ éléments dans $\mathbb{R}_n[X]$ de dimension $n + 1$: c'est une **base orthonormée**.

Q20. $(L_i, P) = \sum_k L_i(x_k)P(x_k) = P(x_i)$, un seul terme survivant.

Dans une base orthonormée, tout vecteur est la somme de ses projections :

$$P = \sum_{i=0}^n (L_i, P)L_i = \sum_{i=0}^n P(x_i)L_i.$$

Q22. Récurrence sur n . Si g a $n + 2$ zéros distincts, le théorème de Rolle appliqué entre deux zéros consécutifs fournit $n + 1$ zéros distincts de g' , laquelle est de classe C^n . L'hypothèse de récurrence donne alors un zéro de $(g')^{(n)} = g^{(n+1)}$.

Q23. La fonction W s'annule en $x_0, \dots, x_n$ — car $f - P_f$ et le produit s'y annulent — et en x par le choix de K . Comme x est distinct des x_i , cela fait $n + 2$ zéros distincts dans $[a, b]$.

D'après **Q22**, $W^{(n+1)}$ s'annule en un point ξ_x . Or P_f est de degré $\leq n$ — sa dérivée $(n+1)$ -ième est nulle — et le produit est unitaire de degré $n + 1$, de dérivée $(n+1)$ -ième égale à $(n+1)!$:

$$0 = f^{(n+1)}(\xi_x) - K \cdot (n+1)!.$$

En reportant la valeur de K :

$$f(x) - P_f(x) = \left(\prod_{i=0}^n (x - x_i) \right) \cdot f^{(n+1)}(\xi_x) / (n+1)!.$$

C'est la formule d'erreur de l'interpolation de Lagrange : elle explique pourquoi l'interpolation à nœuds équirépartis se dégrade — le produit $\prod (x - x_i)$ oscille violemment près des bords.

Partie D · I — Régression linéaire en dimension 2 (Q24 à Q26)

- **Q24** — la fonction E admet un minimum, qui s'interprète comme une distance.

Q24. En développant coordonnée par coordonnée, le vecteur $A(a, b)^T - Y$ a pour i -ième composante $a + bx_i - y_i$, donc

$$E(a, b) = \|A(a, b)^T - Y\|_2^2.$$

Quand (a, b) décrit $\mathbb{R}^2$, le vecteur $A(a, b)^T$ décrit exactement $\text{Col}(A)$, sous-espace de dimension finie. Le minimum de la distance de Y à un tel sous-espace **existe** : c'est la distance au projeté orthogonal.

- **Q25** — le déterminant de $A^T A$ est strictement positif.

Q25. Le produit $A^T A$ se calcule directement :

$$A^T A = \begin{bmatrix} n & \sum x_k \\ \sum x_k & \sum x_k^2 \end{bmatrix}, \text{ donc } \det(A^T A) = n \cdot \sum x_k^2 - (\sum x_k)^2.$$

L'inégalité de Cauchy-Schwarz appliquée aux vecteurs $(x_1, \dots, x_n)$ et $(1, \dots, 1)$ donne $(\sum x_k)^2 \leq n \cdot \sum x_k^2$, avec **égalité si et seulement si** les deux vecteurs sont colinéaires. Or l'énoncé l'exclut : l'inégalité est stricte et

$$\det(A^T A) > 0.$$

- **Q26** — le gradient de E , les équations normales, et l'unicité du couple optimal.

Q26. Les dérivées partielles valent

$$\partial E / \partial a = -2 \sum (y_i - a - bx_i) \text{ et } \partial E / \partial b = -2 \sum x_i (y_i - a - bx_i).$$

Annuler le gradient revient à écrire que le vecteur $Y - A(a, b)^T$ est orthogonal aux deux colonnes de A , c'est-à-dire

$$A^T A \cdot (a_0, b_0)^T = A^T Y.$$

$A^T A$ étant inversible (**Q25**), ce système a une **unique** solution $(c, d) = (A^T A)^{-1} A^T Y$. Comme E admet un minimum (**Q24**) et que celui-ci est nécessairement un point critique, (c, d) réalise ce minimum, et c'est le seul.

Partie D · II — Les moindres carrés en dimension quelconque (Q27 à Q33)

- Q27, Q28 — $\text{Ker}(A^T A) = \text{Ker}(A)$, et l'égalité des rangs.
- Q29, Q30 — décomposition de y et caractérisation des minimiseurs.
- Q31, Q32 — les équations normales $A^T A x_0 = A^T y$.
- Q33 — le cas où A est de rang maximal.

Q33. Si $\text{rg}(A) = p$, alors $\text{Ker}(A) = \{0\}$ et Q28 donne $\text{Ker}(A^T A) = \{0\}$: la matrice carrée $A^T A \in M_p(R)$ est inversible.

Les équations normales ont alors une unique solution :

$$b = (A^T A)^{-1} A^T y.$$

Q27. Multiplions à gauche par x^T :

$$0 = x^T A^T A x = (Ax)^T (Ax) = \|Ax\|_2^2, \text{ donc } \|Ax\|_2 = 0.$$

Q28. L'inclusion $\text{Ker}(A) \subset \text{Ker}(A^T A)$ est immédiate, et Q27 donne l'inclusion réciproque : $\text{Ker}(A^T A) = \text{Ker}(A)$.

Le théorème du rang, appliqué aux deux matrices qui ont le même nombre p de colonnes, donne alors $\text{rg}(A^T A) = \text{rg}(A)$.

Q29. $\text{Col}(A) = \text{Im}(A)$ est un sous-espace de dimension finie de $M_{\{n,1\}}(R)$, donc $M_{\{n,1\}}(R) = \text{Im}(A) \oplus \text{Im}(A)^\perp$.

La décomposition de y selon cette somme directe s'écrit $y = p + d$ avec $p \in \text{Im}(A)$ et $d \in \text{Im}(A)^\perp$, et elle est unique. Comme $p \in \text{Im}(A)$, il existe u tel que $p = Au$.

Le couple (Au, d) est unique ; le vecteur u lui-même n'est déterminé que modulo $\text{Ker}(A)$ — c'est exactement ce que la suite exploite.

Q30. En écrivant $Ax - y = A(x - u) - d$, avec $A(x-u) \in \text{Im}(A)$ et $d \in \text{Im}(A)^\perp$, le théorème de Pythagore donne

$$\|Ax - y\|_2^2 = \|A(x-u)\|_2^2 + \|d\|_2^2 \geq \|d\|_2^2,$$

avec égalité si et seulement si $A(x - u) = 0$, c'est-à-dire $x - u \in \text{Ker}(A)$.

Q31. $A^T y = A^T (Au + d) = A^T Au + A^T d$, et $A^T d = 0$: les lignes de A^T sont les colonnes de A , auxquelles d est orthogonal. Donc $A^T y = A^T Au$.

Q32. D'après Q30, x_0 est minimiseur si et seulement si $x_0 - u \in \text{Ker}(A) = \text{Ker}(A^T A)$ (Q28), c'est-à-dire $A^T A x_0 = A^T Au$, ce qui vaut $A^T y$ d'après Q31. D'où l'équivalence

$$x_0 \text{ minimise } \|Ax - y\|_2 \iff A^T A x_0 = A^T y.$$

Ce sont les équations normales.

Partie D · III — La décomposition en valeurs singulières (Q34 à Q43)

- Q34, Q35 — positivité des formes quadratiques associées, et du spectre de $A^T A$.
- Q36 à Q38 — construction de la décomposition $A = UDV^T$.
- Q39, Q40 — $\tilde{A}\tilde{A}^T$ est un projecteur orthogonal, et $\tilde{b} = \tilde{A}y$ est un minimiseur.
- Q41 à Q43 — $\tilde{b}$ est le minimiseur de **norme minimale**.

Q34. $X^T A^T A X = (A X)^T (A X) = \|A X\|_2^2 \geq 0$, et $X^T X = \|X\|_2^2 \geq 0$.

Q35. Soit λ une valeur propre de $A^T A$, de vecteur propre $X \neq 0$. En multipliant $A^T A X = \lambda X$ par X^T : $\|A X\|_2^2 = \lambda \|X\|_2^2$, donc $\lambda = \|A X\|_2^2 / \|X\|_2^2 \geq 0$.

Q36. $A^T A$ est **symétrique réelle** : le théorème spectral fournit une matrice orthogonale V telle que $V^T A^T A V$ soit diagonale, de coefficients les valeurs propres, toutes positives (**Q35**).

Quitte à réordonner les colonnes de V , on les range par ordre décroissant. Le nombre de valeurs propres non nulles est le rang, soit r (**Q28**), d'où

$$A^T A = V \cdot \text{Diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) \cdot V^T \text{ avec } \lambda_1 \geq \dots \geq \lambda_r > 0.$$

Q37. Pour tous i, j :

$$\langle A v_i, A v_j \rangle = v_i^T A^T A v_j = \lambda_j \cdot v_i^T v_j = \lambda_j \cdot \delta_{ij},$$

les v_i étant orthonormés. La famille $(A v_1, \dots, A v_r)$ est donc **orthogonale**, et $\|A v_i\|_2 = \sqrt{\lambda_i}$ pour $i \leq r$, tandis que $\|A v_i\|_2 = 0$ — donc $A v_i = 0$ — pour $i \geq r + 1$.

Les vecteurs $u_i = A v_i / \sqrt{\lambda_i}$, $i \leq r$, forment ainsi une famille orthonormale de $M_{\{n,1\}}(\mathbb{R})$, que le théorème de la base orthonormée incomplète permet de compléter en une base orthonormée $(u_1, \dots, u_n)$.

Q38. U a ses colonnes orthonormées, donc $U^T U = I_n$: elle est **orthogonale**.

La i -ième colonne de AV est $A v_i$, qui vaut $\sqrt{\lambda_i} \cdot u_i$ pour $i \leq r$ et 0 sinon — c'est exactement la i -ième colonne de UD . Donc $AV = UD$, et comme V est orthogonale, $A = UDV^T$.

Q39. En substituant, les facteurs $V^T V$ et $U^T U$ se simplifient :

$$A\tilde{A} = UDV^T \cdot V\Delta U^T = U(D\Delta)U^T.$$

Or $D\Delta = \text{Diag}_{\{n,n\}}(1, \dots, 1, 0, \dots, 0)$ avec r coefficients égaux à 1 . Donc

$$A\tilde{A} = \sum_{i \leq r} u_i u_i^T, \text{ qui est la matrice de la projection orthogonale sur } \text{Vect}(u_1, \dots, u_r).$$

Et cet espace est exactement $\text{Im}(A)$: d'après **Q37**, $\text{Im}(A)$ est engendré par les $A v_i$, c'est-à-dire par les u_i pour $i \leq r$.

Q40. $Ab = A\tilde{A}y$ est le projeté orthogonal de y sur $\text{Im}(A)$ (**Q39**), donc $\|Ab - y\|_2$ est la distance de y à $\text{Im}(A)$: c'est le minimum, d'après **Q30**.

Q41. $Ab - y$ est l'opposé de la composante de y orthogonale à $\text{Im}(A)$: il appartient à $\text{Im}(A)^\perp$.

Si x_0 est un autre minimiseur, alors Ax_0 est aussi le projeté orthogonal de y , lequel est **unique** : $Ax_0 = Ab$, donc $A(x_0 - b) = 0$, c'est-à-dire

$$x_0 - b \in \text{Ker}(A).$$

Q42. Les colonnes de Δ d'indice $> r$ sont nulles, donc

$$\text{Im}(\tilde{A}) \subset \text{Vect}(v_1, \dots, v_r),$$

et $b = \tilde{A}y$ appartient à cet espace.

Par ailleurs, $A v_i = 0$ exactement pour $i \geq r+1$ (**Q37**), et les v_i forment une base :

$$\text{Ker}(A) = \text{Vect}(v_{r+1}, \dots, v_p). \text{ La base étant orthonormée,}$$

$$\text{Ker}(A)^\perp = \text{Vect}(v_1, \dots, v_r).$$

Donc $b \in \text{Ker}(A)^\perp$.

Q43. Écrivons $x_0 = b + (x_0 - b)$, avec $b \in \text{Ker}(A)^\perp$ (Q42) et $x_0 - b \in \text{Ker}(A)$ (Q41) : les deux termes sont orthogonaux. Pythagore donne

$$\|x_0\|_2^2 = \|b\|_2^2 + \|x_0 - b\|_2^2 \geq \|b\|_2^2,$$

avec égalité si et seulement si $x_0 = b$.

Parmi tous les minimiseurs, $b = \tilde{A}y$ est donc celui de **norme minimale** : c'est la définition de la solution par pseudo-inverse, et c'est ce que calculent les bibliothèques d'algèbre linéaire.

Partie E — La méthode du gradient (Q44 à Q52)

On suppose ici toutes les valeurs propres de $A^{\top}A$ strictement positives.

- Q44, Q45 — $A^{\top}A$ est inversible, et définit un produit scalaire.
- Q46 à Q48 — l'inégalité de Kantorovich, via une majoration élémentaire.
- Q49 à Q52 — la descente de gradient à pas optimal converge, avec une vitesse gouvernée par le conditionnement.

Q44. Toutes les valeurs propres de $A^{\top}A$ sont non nulles, donc son déterminant — leur produit — l'est aussi : $A^{\top}A$ est **inversible**.

D'après **Q28**, $\text{Ker}(A) = \text{Ker}(A^{\top}A) = \{0\}$.

Q45. La bilinéarité est claire, et la symétrie vient de celle de $A^{\top}A$:

$$\langle A^{\top}Ax, y \rangle = x^{\top}A^{\top}Ay = \langle x, A^{\top}Ay \rangle.$$

Pour le caractère défini positif, $\langle A^{\top}Ax, x \rangle = \|Ax\|_2^2 \geq 0$, avec égalité si et seulement si $Ax = 0$, c'est-à-dire $x = 0$ puisque $\text{Ker}(A) = \{0\}$ (**Q44**).

Q46. L'inégalité $\alpha\beta \leq \frac{1}{4}(\alpha+\beta)^2$ équivaut à $0 \leq (\alpha - \beta)^2$, toujours vraie.

Dans la base propre, $\langle A^{\top}Av, v \rangle = \sum \lambda_i v_i^2$ et $\langle (A^{\top}A)^{-1}v, v \rangle = \sum v_i^2 / \lambda_i$. Appliquons la majoration à

$$\alpha = (\sum \lambda_i v_i^2) / \sqrt{(\lambda_1 \lambda_p)} \text{ et } \beta = \sqrt{(\lambda_1 \lambda_p)} \cdot \sum v_i^2 / \lambda_i,$$

dont le produit est exactement le produit cherché. Leur somme se factorise :

$$\alpha + \beta = \sqrt{(\lambda_1 / \lambda_p)} \cdot \sum \{i\} (\lambda_i / \lambda_1 + \lambda_p / \lambda_i) v_i^2,$$

d'où l'inégalité annoncée.

Q47. $\psi'(t) = 1/\lambda_1 - \lambda_p/t^2$ s'annule en $t = \sqrt{(\lambda_1 \lambda_p)}$, et ψ est convexe sur $]0, +\infty[$: son maximum sur le segment $[\lambda_1, \lambda_p]$ est atteint **en l'une des bornes**. Or

$$\psi(\lambda_1) = 1 + \lambda_p/\lambda_1 = \psi(\lambda_p).$$

Le maximum vaut donc $1 + c(A)$.

Q48. Chaque λ_i appartient à $[\lambda_1, \lambda_p]$, donc $\lambda_i/\lambda_1 + \lambda_p/\lambda_i = \psi(\lambda_i) \leq 1 + c(A)$ d'après **Q47**. En reportant dans **Q46** et en utilisant $\lambda_1/\lambda_p = 1/c(A)$:

$$\langle A^{\top}Av, v \rangle \cdot \langle (A^{\top}A)^{-1}v, v \rangle \leq \frac{1}{4} \cdot (1/c(A)) \cdot (1 + c(A))^2 \cdot \|v\|_2^4 = ((c(A)+1)^2 / (4c(A))) \cdot \|v\|_2^4.$$

Q49. Utilisons $A^{\top}A_{\ell} = A^{\top}y$ pour remplacer $A^{\top}y$ dans l'expression de f :

$$f(x) - f(\ell) = \langle A^{\top}Ax, x \rangle - 2\langle x, A^{\top}A_{\ell} \rangle + \langle A^{\top}A_{\ell}, \ell \rangle = \langle A^{\top}A(x - \ell), x - \ell \rangle,$$

en développant le carré pour le produit scalaire de **Q45**.

Q50. Posons $e_k = x_k - \ell$, de sorte que $d_k = A^{\top}A e_k$ et, par **Q49**,

$$f(x_k) - f(\ell) = \langle d_k, e_k \rangle = \langle (A^{\top}A)^{-1}d_k, d_k \rangle.$$

En développant $f(x_{k+1}) - f(\ell) = \langle A^{\top}A(e_k - t_k d_k), e_k - t_k d_k \rangle$:

$$= \langle d_k, e_k \rangle - 2t_k \|d_k\|_2^2 + t_k^2 \langle A^{\top}A d_k, d_k \rangle.$$

Le choix $t_k = \|d_k\|_2^2 / \langle A^{\top}A d_k, d_k \rangle$ — celui qui minimise cette expression du second degré en t_k — laisse

$$f(x_{k+1}) - f(\ell) = (f(x_k) - f(\ell)) - \|d_k\|_2^4 / \langle A^{\top}A d_k, d_k \rangle,$$

que l'on met en facteur en divisant par $f(x_k) - f(\ell) = \langle (A^{\top}A)^{-1}d_k, d_k \rangle$:

$$f(x_{k+1}) - f(\ell) = (f(x_k) - f(\ell)) \cdot (1 - \|d_k\|_2^4 / (\langle A^{\top}A d_k, d_k \rangle \langle (A^{\top}A)^{-1}d_k, d_k \rangle)).$$

Q51. L'inégalité de Kantorovich (**Q48**) appliquée à $v = d_k$ minore le quotient qui apparaît ci-dessus par $4c/(c+1)^2$, donc le facteur est majoré par

$$1 - 4c/(c+1)^2 = (c-1)^2/(c+1)^2.$$

Une récurrence immédiate donne alors

$$f(x_k) - f(\ell) \leq (f(x_0) - f(\ell)) \cdot ((c(A)-1)/(c(A)+1))^{2k}.$$

Enfin, λ_1 étant la plus petite valeur propre de A^TA :

$$\lambda_1 \|x_k - \ell\|_2^2 \leq \langle A^TA(x_k - \ell), x_k - \ell \rangle = f(x_k) - f(\ell).$$

Q52. Comme $c(A) \geq 1$, le rapport $(c-1)/(c+1)$ appartient à $[0, 1[$: la majoration de **Q51** tend vers 0, donc $f(x_k) - f(\ell) \rightarrow 0$, puis

$$\|x_k - \ell\|_2^2 \leq (f(x_k) - f(\ell))/\lambda_1 \rightarrow 0 :$$

la suite (x_k) converge vers ℓ .

Or $\ell = (A^TA)^{-1}A^Ty$ vérifie les équations normales, donc réalise le minimum de $\|Ax - y\|_2$ d'après **Q32**.

La vitesse est *géométrique*, de raison $((c-1)/(c+1))^2$: elle se dégrade quand le conditionnement $c(A)$ grandit. C'est la difficulté centrale des méthodes de gradient, et la raison d'être des techniques de préconditionnement.