Calculs de distances par la norme de Schur

Dans ce sujet, n est un entier naturel non nul et on note :

— $M_n(\mathbb{R})$: la $\mathbb{R}$-algèbre des matrices carrées réelles d'ordre n.
$\mathcal{M}_{n,1}(\mathbb{R})$: le $\mathbb{R}$-espace vectoriel des matrices à n lignes et à une colonne.

— Pour une matrice A de $M_n(\mathbb{R})$, ${}^tA$ est sa matrice transposée, $rang(A)$ son rang et $Tr(A)$ sa trace.

— I_n : la matrice unité de $M_n(\mathbb{R})$.

— $\mathcal{S}_n(\mathbb{R})$: le sous-espace vectoriel des matrices symétriques de $M_n(\mathbb{R})$.

— $\mathcal{A}_n(\mathbb{R})$: le sous-espace vectoriel des matrices antisymétriques de $M_n(\mathbb{R})$.

— $\mathcal{S}_n^+(\mathbb{R})$: l'ensemble des matrices positives de $\mathcal{S}_n(\mathbb{R})$ c'est-à-dire des matrices A de $\mathcal{S}_n(\mathbb{R})$ vérifiant : pour toute matrice $X \in \mathcal{M}_{n,1}(\mathbb{R})$, ${}^tXAX \geqslant 0$.

— $GL_n(\mathbb{R})$: le groupe des matrices inversibles de $M_n(\mathbb{R})$.

— $\mathcal{O}_n(\mathbb{R})$: le groupe des matrices réelles orthogonales c'est-à-dire des matrices M de $M_n(\mathbb{R})$ vérifiant : ${}^tMM = I_n$.

— Pour p entier naturel, Δ_p est l'ensemble des matrices de $M_n(\mathbb{R})$ de rang supérieur ou égal à p et ∇_p est l'ensemble des matrices de $M_n(\mathbb{R})$ de rang inférieur ou égal à p.

Partie I: Préliminaire

Soit la matrice $\Gamma = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}$ de $\mathcal{M}_3(\mathbb{R})$, on pose $H = {}^t\Gamma\Gamma$.

1. Diagonaliser la matrice H et déterminer une matrice P de $\mathcal{O}_3(\mathbb{R})$ et une matrice diagonale D à termes tous positifs telles que $D^2 = P^{-1}HP$.
2. On pose $S = PDP^{-1} \in \mathcal{S}_3^+(\mathbb{R})$, montrer que la relation $\Gamma = US$ définit une matrice $U \in \mathcal{O}_3(\mathbb{R})$ et calculer cette matrice.

Partie II: Calcul de la distance de A à $\mathcal{S}_n(\mathbb{R})$ et à $\mathcal{A}_n(\mathbb{R})$

3. Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$, on pose $(A|B) = \operatorname{Tr}\left({}^tAB\right)$.
Montrer que l'on définit ainsi un produit scalaire sur $M_n(\mathbb{R})$.
La norme associée à ce produit scalaire (norme de Schur) est notée : $\|A\| = ((A|A))^{\frac{1}{2}}$. Dans tout le sujet, si Π est une partie non vide de $M_n(\mathbb{R})$, la distance d'une matrice A de $M_n(\mathbb{R})$ à la partie Π est le réel $d(A, \Pi) = \inf\limits_{M\in\Pi} \|A - M\|$.
4. Montrer que $M_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$ et que cette somme directe est orthogonale.
5. Si A est une matrice de $M_n(\mathbb{R})$, montrer que $d(A, \mathcal{S}_n(\mathbb{R})) = \left\| \frac{1}{2}(A - {}^tA) \right\|$ et déterminer de même $d(A, \mathcal{A}_n(\mathbb{R}))$.
6. Calculer $d(\Gamma, \mathcal{A}_3(\mathbb{R}))$ où Γ est la matrice exemple de la partie I.

Partie III: Théorème de la décomposition polaire

7. Montrer qu'une matrice S de $\mathcal{S}_n(\mathbb{R})$ appartient à $\mathcal{S}_n^+(\mathbb{R})$ si et seulement si toutes les valeurs propres de S sont positives ou nulles.
8. Si A est une matrice de $M_n(\mathbb{R})$, montrer que la matrice ${}^tAA \in \mathcal{S}_n^+(\mathbb{R})$.
9. Soit A une matrice de $M_n(\mathbb{R})$, on suppose qu'il existe une matrice diagonale $D = \operatorname{diag}(d_1, d_2, ..d_n)$ à termes positifs telle que ${}^tAA = D^2$.
On note $A_1, A_2, .., A_n$ les matrices de $\mathcal{M}_{n,1}(\mathbb{R})$ qui forment les colonnes de la matrice A.
   (a) Pour tout couple (i, j) d'entiers naturels compris entre 1 et n, évaluer ${}^tA_iA_j$. En particulier, si i est un entier pour lequel $d_i = 0$, que vaut A_i ?
   (b) Montrer que l'on peut trouver une base orthonormée $(E_1, E_2, ..., E_n)$ de $\mathcal{M}_{n,1}(\mathbb{R})$ (par rapport au produit scalaire canonique $\langle X, Y\rangle = {}^tXY$, de $\mathcal{M}_{n,1}(\mathbb{R})$) telle que, pour tout entier naturel i entre 1 et n, $A_i = d_iE_i$.

Calculs de distances par la norme de Schur

(c) En déduire qu'il existe une matrice E de $\mathcal{O}_n(\mathbb{R})$ telle que $A = ED$.

10. Soit A et B deux matrices de $M_n(\mathbb{R})$ vérifiant ${}^tAA = {}^tBB$.
    (a) Montrer qu'il existe une matrice diagonale D à termes positifs et une matrice orthogonale P telles que : $P^{-1\,t}AAP = P^{-1\,t}BBP = D^2$.
    (b) Montrer qu'il existe une matrice U de $\mathcal{O}_n(\mathbb{R})$ telle que $A = UB$.
11. Déduire des questions précédentes le théorème de décomposition polaire : Pour toute matrice A de $M_n(\mathbb{R})$, il existe une matrice U de $\mathcal{O}_n(\mathbb{R})$ et une matrice S de $\mathcal{S}_n^+(\mathbb{R})$ telles que $A = US$. (Remarque : on peut également établir l'unicité de la matrice S de $\mathcal{S}_n^+(\mathbb{R})$ et même l'unicité de la matrice U de $\mathcal{O}_n(\mathbb{R})$ si A est de plus inversible dans cette décomposition mais ce ne sera pas utile pour la suite du problème).

Partie IV: Calcul de $d(A, \mathcal{O}_n(\mathbb{R}))$

12. Montrer que, pour toute matrice M de $M_n(\mathbb{R})$ et pour toute matrice Ω de $\mathcal{O}_n(\mathbb{R})$, $\|M\Omega\| = \|\Omega M\| = \|M\|$.
13. Dans la suite de cette partie, soit A une matrice de $M_n(\mathbb{R})$, soit $U \in \mathcal{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^+(\mathbb{R})$ telles que $A = US$; il existe une matrice diagonale D et une matrice P de $\mathcal{O}_n(\mathbb{R})$ telles que $S = PDP^{-1}$.
    (a) Montrer que, pour toute matrice Ω de $\mathcal{O}_n(\mathbb{R})$, $\|A - \Omega\| = \|S - U^{-1}\Omega\|$ et en déduire que, $d(A, \mathcal{O}_n(\mathbb{R})) = d(S, \mathcal{O}_n(\mathbb{R}))$.
    (b) Montrer que, $d(A, \mathcal{O}_n(\mathbb{R})) = d(D, \mathcal{O}_n(\mathbb{R}))$
14. On note $D = \mathrm{diag}(\lambda_1, \lambda_2, .., \lambda_n)$
    (a) Montrer que pour toute matrice Ω de $\mathcal{O}_n(\mathbb{R})$, $\|D - \Omega\|^2 = \displaystyle\sum_{i=1}^{n} \lambda_i^2 - 2\mathrm{Tr}\,(D\Omega) + n$
    (b) Montrer que pour toute matrice Ω de $\mathcal{O}_n(\mathbb{R})$, $\mathrm{Tr}\,(D\Omega) \leqslant \displaystyle\sum_{i=1}^{n} \lambda_i$.
    (c) Conclure que $d(D, \mathcal{O}_n(\mathbb{R})) = \|D - I_n\|$.
15. Montrer que , $d(A, \mathcal{O}_n(\mathbb{R})) = \|A - U\|$.
16. Calculer $d(\Gamma, \mathcal{O}_3(\mathbb{R}))$ où Γ est la matrice exemple de la partie I.

Partie V: Calcul de la distance de A à Δ_p.

17. (a) Soit M un élément de $M_n(\mathbb{R})$, montrer qu'il existe un réel $\alpha > 0$ tel que pour tout réel λ vérifiant $0 < \lambda < \alpha$, la matrice $M - \lambda I_n$ est inversible.
    (b) En déduire que $GL_n(\mathbb{R})$ est dense dans $M_n(\mathbb{R})$.
18. Soit A un élément de $M_n(\mathbb{R})$, déterminer, pour tout entier naturel $p \leqslant n$, $d(A, \Delta_p)$.

Partie VI: Théorème de Courant et Fischer

Soit A une matrice de $\mathcal{S}_n(\mathbb{R})$. On notera $\lambda_1 \geqslant \lambda_2 \geqslant .. \geqslant \lambda_n$ ses valeurs propres, on notera $D = \mathrm{diag}(\lambda_1, \lambda_2, .., \lambda_n)$, P la matrice de $\mathcal{O}_n(\mathbb{R})$ vérifiant $A = PD^tP$ et $C_1, C_2, .., C_n$ les matrices de $\mathcal{M}_{n,1}(\mathbb{R})$ formant les colonnes de la matrice P.
Si k est un entier entre 1 et n, on note Ψ_k l'ensemble des sous-espaces vectoriels de $\mathcal{M}_{n,1}(\mathbb{R})$ de dimension k. Nous allons montrer que :

$$\lambda_k = \max_{F \in \Psi_k} \min_{X \in F - \{0\}} \frac{{}^tXAX}{{}^tXX} \quad \text{(Théorème de Courant et Fischer).}$$

19. Soit X un vecteur de $\mathcal{M}_{n,1}(\mathbb{R})$ de coordonnées $(x_1, x_2, .., x_n)$ dans la base orthonormée $(C_1, C_2, .., C_n)$ de $\mathcal{M}_{n,1}(\mathbb{R})$. Calculer en fonction des x_i et λ_i. (i compris entre 1 et n) : ${}^tXAX$ et ${}^tXX$ et pour k entier entre 1 et n, $\dfrac{{}^tC_kAC_k}{{}^tC_kC_k}$.

Calculs de distances par la norme de Schur

20. Soit k entier entre 1 et n, on pose $F_k = \text{vect}\{C_1, C_2, .., C_k\}$. Montrer que pour tout X non nul de F_k , $\dfrac{{}^tXAX}{{}^tXX} \geqslant \lambda_k$ et déterminer $\displaystyle\min_{X \in F_k - \{0\}} \dfrac{{}^tXAX}{{}^tXX}$.
21. Soit $F \in \Psi_k$
    (a) Montrer que $\dim(F \cap \text{vect}\,(C_k, C_{k+1}, .., C_n)) \geqslant 1$.
    (b) Si X est un vecteur non nul de $F \cap \text{vect}\{C_k, C_{k+1}, .., C_n\}$, montrer que $\dfrac{{}^tXAX}{{}^tXX} \leqslant \lambda_k$.
22. Conclure.

Partie VII: Calcul de $d(A, \nabla_p)$

Dans toute cette partie : A est une matrice de $M_n(\mathbb{R})$ de rang r et p est un entier naturel, $p < r$.

23. Montrer qu'il existe deux matrices E et P de $\mathcal{O}_n(\mathbb{R})$ et une matrice diagonale D à termes positifs telles que $A = EDP$. En déduire que le rang de la matrice ${}^tAA$ est encore r. (On pourra utiliser les résultats de la question **9.**)
24. Si on note les valeurs propres de la matrice symétrique réelle ${}^tAA$ de rang r : $\mu_1 \geqslant \mu_2 \geqslant .. \geqslant \mu_r > 0$ et $\mu_{r+1} = .. = \mu_n = 0$, si on pose $D = \text{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}, .., \sqrt{\mu_r}, 0..., 0)$, si pour $1 \leqslant \ell \leqslant n$ on note M_ℓ la matrice de $M_n(\mathbb{R})$ dont la l-ième colonne est celle de la matrice $E \in \mathcal{O}_n(\mathbb{R})$ de la question **23.**, tous les autres termes de M_ℓ étant nuls, on a clairement : $ED = \displaystyle\sum_{\ell=1}^{n} \sqrt{\mu_\ell} M_\ell$.
    Montrer alors qu'il existe une famille orthonormale $(R_1, R_2, .., R_n)$ de matrices de $M_n(\mathbb{R})$ (pour le produit scalaire $(A|B) = Tr({}^tAB)$ de $M_n(\mathbb{R})$), toutes de rang un, et telles que $A = \displaystyle\sum_{\ell=1}^{n} \sqrt{\mu_\ell} R_\ell = \sum_{\ell=1}^{r} \sqrt{\mu_\ell} R_\ell$.
25. Avec les notations de la question **24.**, on pose $N = \displaystyle\sum_{\ell=1}^{p} \sqrt{\mu_\ell} R_\ell$.
    Montrer que $\text{rang}(N) \leqslant p$ puis que $d(A, \nabla_p) \leqslant \sqrt{\displaystyle\sum_{\ell=p+1}^{r} \mu_\ell}$.
26. Soit M une matrice de rang p ($p < r$), on note $\alpha_1 \geqslant \alpha_2 \geqslant \alpha_n \geqslant 0$ les valeurs propres de la matrice ${}^t(A-M)(A-M)$ et on pose $G = \text{Ker}M \cap \text{Im}({}^tAA)$.
    Soit k un entier compris entre 1 et $r - p$.
    (a) Montrer que $\dim G \geqslant r - p$.
    (b) Soit F un sous-espace vectoriel de G de dimension k, montrer que : $\alpha_k \geqslant \displaystyle\min_{X \in F - \{0\}} \dfrac{{}^tX{}^tAAX}{{}^tXX}$.
    (c) On note $(V_1, V_2, .., V_n)$ une base de $\mathbb{R}^n$ formée de vecteurs propres de la matrice ${}^tAA$, le vecteur V_i étant associé à la valeur propre μ_i de telle sorte que : $\mu_1 \geqslant \mu_2 \geqslant .. \geqslant \mu_r > 0$ et $\mu_{r+1} = .. = \mu_n = 0$.
    Montrer que $\dim(G \cap \text{vect}\{V_1, V_2, .., V_{k+p}\}) \geqslant k$.
    (d) En déduire que $\alpha_k \geqslant \mu_{k+p}$.
27. En déduire $d(A, \nabla_p)$.
28. Calculer, pour $p \in \{0, 1, 2, 3\}$, $\gamma_p = d(\Gamma, \nabla_p)$ où Γ est la matrice exemple de la partie I.

Calculs de distances par la norme de Schur

Partie I: Préliminaire

1. On obtient directement :
$$H = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix} = I_3 + 5\,J \text{ avec } J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$
J est clairement de rang 1, donc 0 est valeur propre double de J, la troisième valeur propre étant égale à 3 puisque $\mathrm{Tr}\,(J) = 3$. Comme $(1,1,1)$ est un vecteur propre évident associé à la valeur propre 3, posons
$$e_1 = \frac{\sqrt{3}}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$
Le noyau de J est alors l'orthogonal de e_1. Nous posons donc $e_2 = \dfrac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $e_3 = e_1 \wedge e_2 = \dfrac{\sqrt{6}}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$,
pour obtenir $P^{-1}HP = I_3 + 5\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = D^2$ avec $P = \begin{pmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3} \end{pmatrix}$ et $D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

2. Comme S est inversible (les valeurs propres de S sont égales à celles de D), on peut poser $U = \Gamma S^{-1}$. Nous avons ensuite ${}^tUU = {}^tS^{-1}\,{}^t\Gamma\Gamma S^{-1} = S^{-1}HS^{-1} = PD^{-1}P^{-1}PD^2P^{-1}PDP^{-1} = I_3$ et U est bien orthogonale. Il reste à calculer U :
$$U = \Gamma P D^{-1}\,{}^tP = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Partie II: Calcul de la distance de A à $\mathcal{S}_n(\mathbb{R})$ et à $\mathcal{A}_n(\mathbb{R})$

3. On a, pour $A = (a_{i,j})$ et $B = (b_{i,j})$ matrices de $M_n\,(\mathbb{R})$:
$$(A\,|\,B) = \sum_{1 \leqslant i,j \leqslant n} a_{i,j} b_{i,j}.$$
L'application $(\ |\ )$ est donc le produit scalaire canonique de $M_n\,(\mathbb{R})$, pour lequel la base canonique est une base orthonormale.

4. Si $M \in M_n\,(\mathbb{R})$, on a $M = \dfrac{M + {}^tM}{2} + \dfrac{M - {}^tM}{2}$ avec $\dfrac{M + {}^tM}{2} \in S_n\,(\mathbb{R})$ et $\dfrac{M - {}^tM}{2} \in A_n\,(\mathbb{R})$. Comme $A_n\,(\mathbb{R}) \cap S_n\,(\mathbb{R}) = \{0\}$, les deux espaces sont supplémentaires. Ils sont également orthogonaux car, pour $A \in A_n\,(\mathbb{R})$ et $S \in S_n\,(\mathbb{R})$,
$$(A\,|\,S) = \mathrm{Tr}\left({}^tAS\right) = -\mathrm{Tr}\,(AS) = -\mathrm{Tr}\,(SA) = -\mathrm{Tr}\left({}^tSA\right) = -(S\,|\,A),$$
et donc $(A\,|\,S) = 0$.

5. Si A est une matrice quelconque, la distance de A à $S_n\,(\mathbb{R})$ est égale à la distance de A au projeté orthogonal de A sur $S_n\,(\mathbb{R})$, soit encore à la norme du projeté orthogonal de A sur $A_n\,(\mathbb{R})$, ce qui est exactement le résultat demandé. Par symétrie, on a $d(A, A_n\,(\mathbb{R})) = \left|\left|\frac{1}{2}(A + {}^tA)\right|\right|$.

6. On a facilement $\dfrac{\Gamma + {}^t\Gamma}{2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix}$ puis $d(\Gamma, A_n\,(\mathbb{R})) = 2\sqrt{2}$.

Calculs de distances par la norme de Schur

Partie III: Théorème de la décomposition polaire

7. Soit $S \in S_n(\mathbb{R})$. Le théorème de réduction des matrices symétriques permet d'affirmer qu'il existe une matrice orthogonale P telle que $D = PS^tP$ soit diagonale. Pour $X \in \mathbb{R}^n$, nous obtenons donc :
$$^tXSX = {}^t(PX)D(PX) = \sum_{i=1}^{n} \lambda_i y_i^2$$
en notant λ_i les termes diagonaux de D (i.e. les valeurs propres de S) et y_i les coefficients de la matrice colonne PX. Ainsi, il faut et il suffit que les λ_i soit tous positif pour que S soit positive puisque PX décrit $\mathbb{R}^n$ quand X décrit $\mathbb{R}^n$.
8. La matrice ^tAA est clairement symérique et $^tX(^tAA)X = {}^t(AX)(AX) = ||AX||^2 \geqslant 0$ pour tout X (en notant $||\ ||$ la norme euclidienne canonique de $\mathbb{R}^n$). Pour tout $A \in M_n(\mathbb{R})$, ^tAA est donc symétrique et positive.
9. (a) $^tA_iA_j$ est le coefficient d'indice (i,j) de la matrice ^tAA : il est donc nul si $i \neq j$ et égal à d_i^2 si $i = j$. En particulier, si $d_i = 0$, $||A_i||^2 = {}^tA_iA_i = d_i^2 = 0$ et la colonne A_i est nulle.
   (b) Notons I l'ensemble des i tels que $e_i \neq 0$ et, pour chaque $i \in I$, posons $E_i = \frac{A_i}{||A_i||} = \frac{A_i}{d_i}$. La famille $(E_i)_{i\in I}$ est alors une famille orthonormale, que nous pouvons complétée en une base orthonormale $(E_i)_{1\leqslant i\leqslant n}$. Comme $d_i = 0$ et $A_i = 0$ pour $i \notin I$, l'égalité $A_i = d_iE_i$ est vraie pour tout i.
   (c) Soit E la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base $(E_i)_{1\leqslant i\leqslant n}$. Cette base étant orthonormale, E est une matrice orthogonale et $A_i = d_iE_i$ pour tout i se traduit par $A = ED$.
10. (a) ^tAA est symétrique réelle, donc il existe P orthogonale telle que $P^{-1\,t}AAP$ soit diagonale. D'autre part, ^tAA est positive donc ses valeurs propres sont positives (questions 7 et 8). On en déduit que $D = P^{-1\,t}AAP = P^{-1\,t}BBP$ est une matrice diagonale à termes positifs.
   (b) On déduit de la question 9c qu'il existe deux matrices orthogonales E et F telles que $A = ED$ et $B = FD$, ce qui donne $A = UB$ avec $U = EF^{-1}$, qui est bien orthogonale.
11. Soit $A \in M_n(\mathbb{R})$. Comme ^tAA est symétrique positive, il existe $P \in \mathcal{O}_n(\mathbb{R})$ et D diagonale positive telle que $^tAA = {}^tPD^2P$, que l'on peut écrire $^tAA = {}^tSS$ où $S = {}^tPDP$ est symétrique positive. On déduit de la question précédente qu'il existe U orthogonale telle que $A = US$, ce qui est le résultat demandé.

Partie IV: Calcul de $d(A, \mathcal{O}_n(\mathbb{R}))$

12. Nous avons :
$$||\Omega M||^2 = \mathrm{Tr}\left(^tM\Omega^t\Omega M\right) = \mathrm{Tr}\left(^tMM\right) = ||M||^2$$
et en utilisant la propriété classique $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$:
$$||M\Omega||^2 = \mathrm{Tr}\left(^t\Omega^tMM\Omega\right) = \mathrm{Tr}\left(^tMM\Omega^t\Omega\right) = \mathrm{Tr}\left(^tMM\right) = ||M||^2,$$
ce qui donne bien $||\Omega M|| = ||M\Omega|| = ||M||$.
13. (a) On a $||A - \Omega|| = ||US - \Omega|| = ||U(S - U^{-1}\Omega)|| = ||S - U^{-1}\Omega||$ d'après la question 12.

   Quand Ω décrit $\mathcal{O}_n(\mathbb{R})$, $U^{-1}\Omega$ décrit également $\mathcal{O}_n(\mathbb{R})$, donc $d(A, \mathcal{O}_n(\mathbb{R})) = d(S, \mathcal{O}_n(\mathbb{R}))$.
   (b) Pour tout $\Omega \in \mathcal{O}_n(\mathbb{R})$, nous avons $||S - \Omega|| = ||PDP^{-1} - \Omega|| = ||P(D - P^{-1}\Omega P)P^{-1}|| = ||D - P^{-1}\Omega P||$ car $P, P^{-1} \in \mathcal{O}_n(\mathbb{R})$. Une nouvelle fois, $P^{-1}\Omega P$ décrit $\mathcal{O}_n(\mathbb{R})$ quand Ω décrit $\mathcal{O}_n(\mathbb{R})$ donc
$$d(A, \mathcal{O}_n(\mathbb{R})) = d(S, \mathcal{O}_n(\mathbb{R})) = d(D, \mathcal{O}_n(\mathbb{R})).$$
14. (a) $||D - \Omega||^2 = \mathrm{Tr}\left(^t(D - \Omega)(D - \Omega)\right) = \mathrm{Tr}\left(D^2 - {}^t\Omega D - D^t\Omega + I_n\right) = \sum_{i=1}^{n} \lambda_i^2 - 2\mathrm{Tr}(D\Omega) + n.$
   (b) En notant $d_{i,j}$ et $\omega_{i,j}$ les termes génériques de D et de Ω, nous obtenons :
$$\mathrm{Tr}(D\Omega) = \sum_{i=1}^{n}\sum_{k=1}^{n} d_{i,k}\omega_{k,i} = \sum_{i=1}^{n} \lambda_i\omega_{i,i} \leqslant \sum_{i=1}^{n} \lambda_i$$
   car Ω étant orthogonale, les $\omega_{i,j}$ sont éléments de $[-1, 1]$

CALCULS DE DISTANCES PAR LA NORME DE SCHUR

(c) On en déduit que pour tout $\Omega \in \mathcal{O}_n(\mathbb{R})$:

$$||D-\Omega||^2 \geqslant \sum_{i=1}^{n} \lambda_i^2 - 2\sum_{i=1}^{n} \lambda_i + n = \sum_{i=1}^{n} (\lambda_i - 1)^2 = ||D - I_n||^2.$$

Comme $I_n \in \mathcal{O}_n(\mathbb{R})$, ceci prouve que la distance de D à $\Omega \in \mathcal{O}_n(\mathbb{R})$ est minimale pour $\Omega = I_n$.

15. Nous venons de démontrer que $d(A, \mathcal{O}_n(\mathbb{R})) = d(D, \mathcal{O}_n(\mathbb{R})) = d(D, I_n) = \sqrt{\sum_{i=1}^{n} (\lambda_i - 1)^2}$, où les λ_i sont les racines carrées des valeurs propres de ${}^t AA$, appelées *valeurs singulières* de A.
16. Nous avons ici $n = 3$, $\lambda_1 = 4$ et $\lambda_2 = \lambda_3 = 1$, donc $d(\Gamma, \mathcal{O}_n(\mathbb{R})) = 3$.

Partie V: Calcul de la distance de A à Δ_p.

17. (a) Soit α le minimum de l'ensemble des $|\lambda|$ pour λ valeur propre (réelle) non nulle de M (si M n'a aucune valeur propre réelle non nulle, on choisit $\alpha > 0$ quelconque). Pour tout $\lambda \in]0, \alpha[$, $M - \lambda I_n$ est inversible car λ n'est pas valeur propre de M.
    (b) Pour M quelconque et α comme au a, la suite $\left(M - \frac{\alpha}{k+2} I_n\right)_{k \geqslant 0}$ est une suite de matrices inversibles qui converge vers M : $\mathrm{GL}_n(\mathbb{R})$ est donc dense dans $M_n(\mathbb{R})$.
18. On en déduit que pour tout $A \in M_n(\mathbb{R})$, $d(A, \mathrm{GL}_n(\mathbb{R})) = 0$. Comme $\mathrm{GL}_n(\mathbb{R})$ est contenu dans tous les Δ_p pour $p \leqslant n$, on a à plus forte raison $d(A, \Delta_p) = 0$ pour tout $A \in M_n(\mathbb{R})$ et pour tout $p \leqslant n$.

Partie VI: Théorème de Courant et Fischer

19. La base $(C_1, C_2, \dots, C_n)$ est une BON de diagonalisation de A avec $AC_k = \lambda_k C_k$. Nous en déduisons que ${}^t XAX = q(x) = \sum_{i=1}^{n} \lambda_i x_i^2$.

    D'autre part, la base $(C_1, C_2, \dots, C_n)$ est orthonormale pour le produit scalaire usuel, donc ${}^t XX = ||X||^2 = \sum_{i=1}^{n} x_i^2$.

    En particulier, pour $X = C_k$, nous obtenons :

$$\frac{{}^t C_k A C_k}{{}^t C_k C_k} = \lambda_k.$$

20. Soit X élément non nul de F_k. Avec les notations de la question 19, nous avons $x_i = 0$ pour $i > k$, ce qui donne :

$$\frac{{}^t XAX}{{}^t XX} = \frac{\sum_{i=1}^{k} \lambda_i x_i^2}{\sum_{i=1}^{k} x_i^2} \geqslant \frac{\sum_{i=1}^{k} \lambda_k x_i^2}{\sum_{i=1}^{k} x_i^2} = \lambda_k$$

car les λ_i décroissent. Comme le minorant λ_k est atteint pour $X = C_k$, on en déduit :

$$\min_{X \in F_k \setminus \{O\}} \frac{{}^t XAX}{{}^t XX} = \lambda_k.$$

21. (a) On sait que $\dim F \cap G = \dim F + \dim G - \dim F \cup G$ pour F et G s.e.v. de E, donc

$$\dim (F \cap \mathbf{Vect}(C_k, C_{k+1}, \dots, C_n)) = k + (n - k + 1) - \dim (F \cup \mathbf{Vect}(C_k, C_{k+1}, \dots, C_n)) \geqslant 1$$

car $\dim F \cup \mathbf{Vect} C_k, C_{k+1}, \dots, C_n \leqslant n$.

(b) En reprenant encore les notations de la question 19, nous avons :

$$\frac{{}^tXAX}{{}^tXX} = \frac{\sum_{i=k}^{n} \lambda_i x_i^2}{\sum_{i=k}^{n} x_i^2} \leqslant \frac{\sum_{i=k}^{n} \lambda_k x_i^2}{\sum_{i=1}^{k} x_i^2} = \lambda_k$$

car les λ_i décroissent.

22. En utilisant la question 20, nous obtenons :

$$\max_{F\in\Psi_k} \min_{X\in F\setminus\{O\}} \frac{{}^tXAX}{{}^tXX} \geqslant \min_{X\in F_k\setminus\{O\}} \frac{{}^tXAX}{{}^tXX} = \lambda_k.$$

D'autre part, pour $F \in \Psi_k$, on peut choisir X_0 non nul dans $F \cap \mathbf{Vect} C_k, C_{k+1}, \dots, C_n$ puisque cet espace vectoriel est de dimension non nulle. On en déduit :

$$\min_{X\in F_k\setminus\{O\}} \frac{{}^tXAX}{{}^tXX} \leqslant \frac{{}^tX_0AX_0}{{}^tX_0X_0} \leqslant \lambda_k.$$

Ceci achève la preuve du théorème de Courant et Fischer.

Partie VII: Calcul de $d(A, \nabla_p)$

23. Soit P orthogonale et D diagonale positive telles que ${}^tAA = {}^tPD^2P$. On a alors ${}^t(A{}^tP)(A{}^tP) = D^2$, donc (question 9) il existe $E \in \mathcal{O}_n(\mathbb{R})$ telle que $A{}^tP = ED$, ce qui donne bien $A = EDP$ avec $E, P \in \mathcal{O}_n(\mathbb{R})$ et D diagonale positive.

On en déduit que $\mathbf{rg}A = \mathbf{rg}D = \mathbf{rg}D^2 = \mathbf{rg}{}^tAA$ puisque A est équivalente à D, D est diagonale et D^2 est semblable à ${}^tAA$.

Remarque : il est plus rapide de montrer (classiquement) que A et ${}^tAA$ ont même noyau.

24. Posons $R_l = M_lP$ pour l compris entre 1 et n. On a ainsi $A = EDP = \sum_{i=1}^{n} \sqrt{\mu_l}\ R_l = \sum_{i=1}^{r} \sqrt{\mu_l}\ R_l$ et on vérifie facilement que (R_l) est orthonormale :

$$(R_l \mid R_k) = \mathrm{Tr}\left({}^tP{}^tM_lM_kP\right) = \mathrm{Tr}\left({}^tM_lM_kP{}^tP\right) = \mathrm{Tr}\left({}^tM_lM_k\right)$$

or ${}^tM_lM_k$ a tous ses termes nuls, sauf peut-être celui d'indice (l, k) qui est égal au produit scalaire des l-ième et k-ième colonnes de E. Comme E est orthogonale, on obtient bien $(R_l, R_k) = 0$ si $l \neq k$ et $(R_l, R_k) = 1$ si $l = k$.

Enfin, chaque R_l est de rang 1 car $\mathbf{rg}R_l = \mathbf{rg}M_l = 1$ (P est inversible et M_l a une et une seule colonne non nulle).

25. On a clairement $\Im N \subset \Im R_1 + \Im R_2 + \cdots + \Im R_p$, puis $\mathbf{rg}N \leqslant p$ (les $\Im R_i$ sont des droites).

Comme $N \in \nabla_p$, $d(A, \nabla_p) \leqslant d(A, N) = \left\| \sum_{l=p+1}^{r} \sqrt{\mu_i}\ R_i \right\| = \sqrt{\sum_{l=p+1}^{r} \mu_i}$ car (R_i) est une famille orthonormale.

26. (a) $\dim G = \dim \mathrm{Ker}M + \dim \Im{}^tAA - \dim \mathrm{Ker}M \cup \Im{}^tAA \geqslant (n-p) + r - n = r - p$.

(b) En appliquant le théorème de Courant et Fischer (plus exactement en appliquant la question 21) à la matrice $A - M$, nous obtenons :

$$\alpha_k \geqslant \min_{X\in F\setminus\{0\}} \frac{{}^tX{}^t(A-M)(A-M)X}{{}^tXX}$$

mais pour $X \in F$, $MX = 0$ et ${}^tX{}^tM = 0$, donc

$$\alpha_k \geqslant \min_{X\in F\setminus\{0\}} \frac{{}^tX{}^tAAX}{{}^tXX}.$$

(c) On a $G \cap \mathbf{Vect} V_1, \ldots, V_{k+p} = \mathrm{Ker} M \cap \mathbf{Vect} V_1, \ldots, V_{k+p}$ car $\mathbf{Vect} V_1, \ldots, V_{k+p} \subset \mathbf{Vect} V_1, \ldots, V_r = \Im {}^t AA$ (on a $k \leqslant r - p$). On en déduit donc (comme au a) que $G \cap \mathbf{Vect} V_1, \ldots, V_{k+p}$ est de dimension au moins $(k+p) + (n-p) - n = k$.

(d) Comme $G \cap \mathbf{Vect} V_1, \ldots, V_{k+p}$ est de dimension au moins égale à k, on peut choisir un sous-espace F de dimension k contenu dans $G \cap \mathbf{Vect} V_1, \ldots, V_{k+p}$. Nous avons alors :

— $\alpha_k \geqslant \min\limits_{X \in F \setminus \{O\}} \dfrac{{}^t X {}^t AAX}{{}^t XX}$ d'après le b ;

— pour X élément quelconque de F, que l'on peut écrire sous la forme $X = \sum\limits_{i=1}^{k+p} x_i V_i$:

$$\frac{{}^t X {}^t AAX}{{}^t XX} = \frac{\sum\limits_{i=1}^{k+p} \mu_i x_i^2}{\sum\limits_{i=1}^{k+p} x_i^2} \geqslant \mu_{k+p}$$

car les μ_i décroissent.

On en déduit l'inégalité demandée : $\alpha_k \geqslant \mu_{k+p}$.

27. Soit M une matrice de rang $q \leqslant p < r$. En reprenant les notations et les résultats de la question 26, et en remplaçant p par q (l'inégalité obtenue fonctionne aussi quand $q = 0$), nous obtenons :

$$d^2(A, M) = \mathrm{Tr}\left({}^t(A-M)(A+M)\right) = \sum_{i=1}^{n} \alpha_i \geqslant \sum_{i=1}^{r-q} \alpha_i \geqslant \sum_{i=1}^{r-q} \mu_{i+q} = \sum_{i=q+1}^{r} \mu_i \geqslant \sum_{i=p+1}^{r} \mu_i.$$

On en déduit que $d(A, \nabla_p) \geqslant \sum\limits_{i=p+1}^{r} \mu_i$, ce qui donne, avec la question 25 :

$$d(A, \nabla_p) = \sqrt{\sum_{l=p+1}^{r} \mu_i}$$

où les μ_i sont les valeurs propres (décroissantes) de ${}^t AA$.

28. Ici, nous avons $\mu_1 = 16$, $\mu_2 = \mu_3 = 1$ et $r = 3$. Nous en déduisons donc :

$$\begin{aligned}
\gamma_0 &= ||\Gamma|| = 3\sqrt{2} \\
\gamma_1 &= \sqrt{2} \\
\gamma_2 &= 1 \\
\gamma_3 &= d(\Gamma, \Gamma) = 0.
\end{aligned}$$

CCP-MP-2003