Correction : Devoir libre n°4

MP Clemenceau 2022-23

Pour ... avant les vacances de février

Correction : Remarques : ce qui est en rouge concerne les remarques d'ordre général, en bleu les remarques personnelles. Les notes sur 4 par question sont aussi en bleu en début de correction.

Remarques générales importantes :

- *Toujours tout bien justifier même ce qui pourrait paraitre évident. Ce que vous vous dites pour justifier une égalité, une implication, une équivalence ... il faut l'écrire.*
- *Lorsqu'une matrice A est diagonalisable il est inutile de passer par la diagonalisation (matrice de passage et matrice diagonale) pour dire que le déterminant est le produit des valeurs propres et la trace la somme des valeurs propres. Je rappelle que c'est une conséquence du caractère scindé du polynôme caractéristique.*
- *Toute application φ bilinéaire vérifie $\varphi(0,0) = 0$. Il est donc inutile de l'écrire pour le caractère défini d'un produit scalaire (à condition de bien commencer par la bilinéarité).*
- *Eviter d'utiliser des doubles indices dans les matrices colonnes et les coefficients diagonaux d'une matrice diagonale. Cela alourdit la notation.*
- *Ne pas confondre la norme sur les matrices carrées et la norme usuelle sur les matrices colonnes qui n'a pas besoin de l'application trace.*

Calculs de distances entre une matrice et certaines parties de $\mathscr{M}_n(\mathbb{R})$

Notations

Dans ce sujet, n est un entier naturel non nul et on note :
$\mathscr{M}_n(\mathbb{R})$: la $\mathbb{R}$-algèbre des matrices carrées réelles d'ordre n.
$\mathscr{M}_{n,1}(\mathbb{R})$: le $\mathbb{R}$ -espace vectoriel des matrices à n lignes et à une colonne.
Pour une matrice A de $\mathscr{M}_n(\mathbb{R})$, $A^\top$ est sa matrice transposée, $rang(A)$ son rang et $\operatorname{tr}(A)$ sa trace.
I_n : la matrice unité de $\mathscr{M}_n(\mathbb{R})$.
$\mathscr{S}_n(\mathbb{R})$: le sous-espace vectoriel des matrices symétriques de $\mathscr{M}_n(\mathbb{R})$.
$\mathscr{A}_n(\mathbb{R})$: le sous-espace vectoriel des matrices antisymétriques de $\mathscr{M}_n(\mathbb{R})$.
$\mathscr{S}_n^+(\mathbb{R})$: l'ensemble des matrices positives de $\mathscr{S}_n(\mathbb{R})$ c'est-à-dire des matrices A de $\mathscr{S}_n(\mathbb{R})$ vérifiant : pour toute matrice $X \in \mathscr{M}_{n,1}(\mathbb{R})$, $X^\top AX \geqslant 0$.
$GL_n(\mathbb{R})$: le groupe des matrices inversibles de $\mathscr{M}_n(\mathbb{R})$.
$\mathscr{O}_n(\mathbb{R})$: le groupe des matrices réelles orthogonales c'est-à-dire des matrices M de $\mathscr{M}_n(\mathbb{R})$ vérifiant : $M^\top M = I_n$.
Pour p entier naturel, Δ_p est l'ensemble des matrices de $\mathscr{M}_n(\mathbb{R})$ de rang supérieur ou égal à p et ∇_p est l'ensemble des matrices de $\mathscr{M}_n(\mathbb{R})$ de rang inférieur ou égal à p.

Objectifs

Le but du sujet est de calculer la distance (par la norme de Schur définie à la question II.3.) d'une matrice à :
dans la partie II., $\mathscr{S}_n(\mathbb{R})$ et $\mathscr{A}_n(\mathbb{R})$ par le théorème de projection orthogonale,
dans la partie III., $\mathscr{O}_n(\mathbb{R})$ par le théorème de décomposition polaire,
dans la partie IV., Δ_p par des notions de densité,
dans la partie V., ∇_p par le théorème de Courant et Fischer.

La partie I. traite un exemple qui sera utilisé dans les différentes parties.
Remarque : dans le texte, le mot « positif » signifie « supérieur ou égal à 0 ».

I. Exercice préliminaire

1) Soit la matrice $\Gamma = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}$ de $\mathscr{M}_3(\mathbb{R})$, on pose $H = \Gamma^\top \Gamma$.
Diagonaliser la matrice H et déterminer une matrice P de $\mathscr{O}_3(\mathbb{R})$ et une matrice diagonale D à termes tous positifs telles que $D^2 = P^{-1}HP$.

Correction : Par calcul direct on a $H = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix}$.

Il ne faut pas se précipiter sur le polynôme caractéristique. Il y a souvent des éléments propres qui apparaissent rapidement.

On a alors directement que le vecteur $(1, 1, 1)$ est vecteur propre de H associé à la valeur propre 16. De plus la matrice $H - I_3$ est de rang 1 donc 1 est valeur propre double de H. Les vecteurs $(1, -1, 0)$ et $(0, 1, -1)$ sont directement des vecteurs propres de H associés à la valeur propre 1.

Cependant on veut une matrice de passage orthogonale, il faut donc normer le premier vecteur et trouver une base orthonormale du sous espace propre associé à 1. On peut remarquer que $2(0, 1, -1) + (1, -1, 0) = (1, 1, -2)$ est orthogonal à $(1, -1, 0)$. On peut donc considérer alors la base orthonormée de vecteurs propres $\left(\frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{2}}(1, -1, 0), \frac{1}{\sqrt{6}}(1, 1, -2)\right)$.

Remarque : on peut partir des deux premiers vecteurs et faire le produit vectoriel pour trouver le dernier. C'est à dire $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \wedge \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}$, ceci en utilisant : $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \wedge \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} bc' - cb' \\ ca' - ac' \\ ab' - ba' \end{pmatrix}$

Conclusion : en posant $D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix}$, on a alors $H = PD^2P^{-1}$.

2) On pose $S = PDP^{-1} \in \mathscr{S}_3^+(\mathbb{R})$, montrer que la relation $\Gamma = US$ définit une matrice $U \in \mathscr{O}_3(\mathbb{R})$ et calculer cette matrice.

Correction : Remarque : le texte peut paraitre confus car il n'est pas justifié que S est un élément de $\mathscr{S}_3^+(\mathbb{R})$. La justification vient plus dans la question **7)**.

S est inversible car ses valeurs propres, qui sont les éléments diagonaux de D, sont non nulles. On pose $U = \Gamma S^{-1}$. On a alors, en utilisant le fait que P est orthogonale et donc $P^{-1} = P^\top$:

$$U^\top U = (S^{-1})^\top \Gamma^\top \Gamma S = (PD^{-1}P^{-1})^\top PD^2P^{-1}PD^{-1}P^{-1} = I_3$$

Donc U est bien une matrice orthogonale.

Par calcul direct on a $U = \Gamma PD^{-1}P^\top = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Remarque : il est vrai que les calculs peuvent être un peu long sans calculatrices (interdites aussi lorsque le sujet est tombé).

II. Calcul de la distance de A à $\mathscr{S}_n(\mathbb{R})$ et à $\mathscr{A}_n(\mathbb{R})$

3) Soit A et B deux matrices de $\mathscr{M}_n(\mathbb{R})$, on pose $(A|B) = \mathrm{tr}(A^\top B)$.
Montrer que l'on définit ainsi un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$.

Correction : *classique donc à savoir faire rapidement*. La linéarité de la trace et la bilinéarité du produit matriciel donne la bilinéarité. La propriété de la trace $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$ permet d'avoir la symétrie. En développant directement $\mathrm{tr}(A^\top A) = \sum_{1 \leqslant i,j \leqslant n} a_{i,j}^2$, on obtient le caractère défini positif.

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 $\mathscr{M}_n(\mathbb{R})$, la distance d'une matrice A de $\mathscr{M}_n(\mathbb{R})$ à la partie Π est le réel $d(A, \Pi) = \inf_{M \in \Pi} \|A - M\|$.

4) Montrer que $\mathscr{M}_n(\mathbb{R}) = \mathscr{S}_n(\mathbb{R}) \oplus \mathscr{A}_n(\mathbb{R})$ et que cette somme directe est orthogonale.

Correction : Soit $(A,B) \in \mathscr{S}_n(\mathbb{R}) \times \mathscr{A}_n(\mathbb{R})$. On a $\left(A \mid B\right) = \operatorname{tr}(A^\top B) = \operatorname{tr}(AB)$, et $\left(B \mid A\right) = \operatorname{tr}(B^\top A) = -\operatorname{tr}(AB)$. Or $\left(A \mid B\right) = \left(B \mid A\right)$, on en déduit que $\left(A \mid B\right) = 0$ est donc les deux espaces sont en somme directe orthogonale.

Soit $A \in \mathscr{M}_n(\mathbb{R})$.

On a $A = \frac{1}{2}\left(A + A^\top\right) + \frac{1}{2}\left(A - A^\top\right)$. Or $\frac{1}{2}\left(A + A^\top\right)$ est une matrice symétrique et $\frac{1}{2}\left(A - A^\top\right)$ une matrice antisymétrique.

On en déduit que $\mathscr{M}_n(\mathbb{R}) = \mathscr{S}_n(\mathbb{R}) \oplus \mathscr{A}_n(\mathbb{R})$.

Remarque : on peut aussi utiliser un argument de dimension pour avoir le dernier résultat.

5) Si A est une matrice de $\mathscr{M}_n(\mathbb{R})$, montrer que $d(A, \mathscr{S}_n(\mathbb{R})) = \|\dfrac{1}{2}(A - A^\top)\|$ et déterminer de même $d(A, \mathscr{A}_n(\mathbb{R}))$.

Correction : *il est vraiment indispensable de bien connaître son cours afin de pouvoir répondre rapidement à ce genre de question.* La distance d'un vecteur à un sous espace vectoriel est égal à la distance entre ce vecteur et son projeté orthogonal sur cet espace ou encore la norme du projeté orthogonal de ce vecteur sur le supplémentaire orthogonal du sous espace (ceci en dimension finie). D'après la décomposition faite à la question précédente on a directement : $d(A, \mathscr{S}_n(\mathbb{R})) = \|\dfrac{1}{2}(A - A^\top)\|$ et $d(A, \mathscr{A}_n(\mathbb{R})) = \|\dfrac{1}{2}(A + A^\top)\|$.

6) Calculer $d(\Gamma, \mathscr{A}_3(\mathbb{R}))$ où Γ est la matrice exemple de la partie I.

Correction : on a directement que $\frac{1}{2}\left(\Gamma^\top + \Gamma\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix}$ et par suite $d(\Gamma, \mathscr{A}_3(\mathbb{R})) = 2\sqrt{2}$.

III. Calcul de la distance de A à $\mathscr{O}_n(\mathbb{R})$

A. Théorème de la décomposition polaire

7) Montrer qu'une matrice S de $\mathscr{S}_n(\mathbb{R})$ appartient à $\mathscr{S}_n^+(\mathbb{R})$ si et seulement si toutes les valeurs propres de S sont positives ou nulles.

Correction : Soit S une matrice symétrique.
On suppose que S est un élément de $\mathscr{S}_n^+(\mathbb{R})$, on a alors, par définition, que pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$, $X^\top SX \geqslant 0$. Pour X vecteur propre associé à λ on a alors $\lambda X^\top X \geqslant$. Or $X^\top X = \displaystyle\sum_{i=1}^n x_i^2$ et comme X est non nul, ceci est strictement positif, donc λ est aussi positive.

Réciproquement : on suppose que S n'admet que des valeurs propres positives ou nulles. D'après le théorème spectral S est diagonalisable à l'aide d'une matrice orthogonale. Il existe donc $P \in \mathscr{O}_n$ et une matrice diagonale D à coefficients positifs telles que $S = PDP^\top$. On note $D = diag(\lambda_1, \dots, \lambda_n)$. Soit $X \in \mathscr{M}_n(\mathbb{R})$. On a $X^\top SX = ({}^\top P^\top X)D(P^\top X)$. Si on nonte $(y_i)_{i \in [\![1,n]\!]}$ les coefficients de $P^\top X$, alors $X^\top SX = \displaystyle\sum_{i=1}^n \lambda_i y_i^2$.

On en déduit que $X^\top SX \geqslant 0$.

Lors de l'utilisation du théorème spectral l'écriture $A = P^\top DP$ n'est pas fausse mais P n'est pas alors la matrice de passage de la base canonique à la base de vecteurs propres. Dans le cas d'étude d'endomorphismes canoniquement associés cela peut créer des problèmes.

8) Si A est une matrice de $\mathscr{M}_n(\mathbb{R})$, montrer que la matrice $A^\top A \in \mathscr{S}_n^+(\mathbb{R})$.

Correction : Soit $A \in \mathscr{M}_n(\mathbb{R})$. La matrice $A^\top A$ est clairement symétrique. Soit $X \in \mathscr{M}_{n,1}(\mathbb{R})$. On a $X^\top(A^\top A)X = ({}^\top AX)(AX) = \|AX\|^2$. On en déduit que $A^\top A$ est positive.

9) Soit A une matrice de $\mathscr{M}_n(\mathbb{R})$, on suppose qu'il existe une matrice diagonale $D = \operatorname{diag}(d_1, d_2, \dots, d_n)$ à termes positifs telle que $A^\top A = D^2$.
On note $A_1, A_2, \dots, A_n$ les matrices de $\mathscr{M}_{n,1}(\mathbb{R})$ qui forment les colonnes de la matrice A.

a) Pour tout couple (i,j) d'entiers naturels compris entre 1 et n, évaluer $A_i^\top A_j$. En particulier, si i est un entier pour lequel $d_i = 0$, que vaut A_i ?

Correction : $A_i^\top A_j$ est le coefficient d'indices (i,j) de la matrice $A^\top A$. Comme $A^\top A = D^2$ on a donc, pour $i \neq j$ $A_i^\top A_j = 0$ et, pour tout i, $A_i^\top A_i = d_i^2$. En particulier, si $d_i = 0$ alors, comme $A_i^\top A_i = \|A_i\|^2$, A_i est nulle.

b) Montrer que l'on peut trouver une base orthonormée $(E_1, E_2, ..., E_n)$ de $\mathscr{M}_{n,1}(\mathbb{R})$ (par rapport au produit scalaire canonique $\langle X, Y\rangle = X^\top Y$, de $\mathscr{M}_{n,1}(\mathbb{R})$) telle que, pour tout entier naturel i entre 1 et n, $A_i = d_i E_i$.

Correction : pour $i \in [\![1,n]\!]$ tel que $A_i \neq 0$, on pose $E_i = \frac{1}{\|A_i\|}A_i$. On a alors bien $A_i = d_i E_i$. Si on note I l'ensemble des i tel que $A_i \neq 0$ alors la famille $(A_i)_{i\in I}$ est orthonormale. On peut la compléter en une base orthonormale $(E_i)_{i\in[\![1,n]\!]}$.

Pour $i \notin I$ on a $A_i = 0$ et $d_i = 0$ donc $A_i = d_i E_i$.

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

Correction : On pose E la matrice dont les colonnes sont les E_i. Par définition c'est la matrice de passage de la base canonique à la base $(E_i)_{i\in[\![1,n]\!]}$. Comme les deux bases sont orthonormales E est une matrice orthogonale. Comme on a, pour tout $i \in [\![1,n]\!]$, $A_i = d_i E_i$, on en déduit que $A = ED$.

Remarque importante : la multiplication par D se fait à droite car on multiplie les colonnes de E.

10) Soit A et B deux matrices de $\mathscr{M}_n(\mathbb{R})$ vérifiant $A^\top A = B^\top B$.

a) Montrer qu'il existe une matrice diagonale D à termes positifs et une matrice orthogonale P telles que : $P^{-1}A^\top AP = P^{-1}B^\top BP = D^2$.

Correction : comme la matrice $A^\top A$ est symétrique réelle, d'après le théorème spectral il existe $P \in \mathscr{O}_n(\mathbb{R})$ et une matrice diagonale Δ telle que $A^\top A = P\Delta P^{-1}$. D'après les questions **7)** et **8)** les valeurs propres de $A^\top A$ sont positives, il existe donc une matrice diagonale D à coefficients positifs telle que $D^2 = \Delta$. D'où le résultat.

b) Montrer qu'il existe une matrice U de $\mathscr{O}_n(\mathbb{R})$ telle que $A = UB$.

Correction : P étant une matrice orthogonale, $P^{-1} = P^\top$. On a donc, d'après la question précédente $({}^\top AP)AP = ({}^\top BP)BP = D^2$. Il existe alors, d'après la question **9)**, E et F deux matrices orthogonales telles que $AP = ED$ et $BP = FD$. On a alors $A = EDP^\top$ et $F^\top B = DP^\top$, d'où $A = UB$ avec $U = EF^\top$ qui est bien une matrice orthogonale car ${}_n(\mathbb{R})$ est un groupe multiplicatif.

11) Déduire des questions précédentes le théorème de décomposition polaire :
Pour toute matrice A de $\mathscr{M}_n(\mathbb{R})$, il existe une matrice U de $\mathscr{O}_n(\mathbb{R})$ et une matrice S de $\mathscr{S}_n^+(\mathbb{R})$ 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).

Correction : soit $A \in \mathscr{M}_n(\mathbb{R})$, d'après la question **10a)** il existe $P \in \mathscr{O}_(\mathbb{R})$ et D une matrice diagonale à coefficients positifs telles que $A^\top A = PD^2P^\top$. On pose $S = PDP^\top$. On a alors $A^\top A = S^\top S$, et donc, d'après la question il existe $U \in \mathscr{O}_n(\mathbb{R})$ telle que $A = US$. S est bien positive car ses valeurs propres sont positives.

B. Calcul de $d(A, \mathscr{O}_n(\mathbb{R}))$

12) Montrer que, pour toute matrice M de $\mathscr{M}_n(\mathbb{R})$ et pour toute matrice Ω de $\mathscr{O}_n(\mathbb{R})$, $\|M\Omega\| = \|\Omega M\| = \|M\|$.

Correction : soit $M \in \mathscr{M}_n(\mathbb{R})$ et $\Omega \in \mathscr{O}_n(\mathbb{R})$, on a, à l'aide de la propriété de la trace $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, et du fait que $\Omega^\top$ est l'inverse de Ω :

$$\|M\Omega\|^2 = \operatorname{tr}\left(({}^\top M\Omega)M\Omega\right) = \operatorname{tr}\left(\Omega^\top M^\top M\Omega\right) = \operatorname{tr}\left(M^\top M\Omega\Omega^\top\right) = \operatorname{tr}\left(M^\top M\right) = \|M\|^2$$

$$\|\Omega M\|^2 = \operatorname{tr}\left(({}^\top \Omega M)\Omega M\right) = \operatorname{tr}\left(M^\top \Omega^\top \Omega M\right) = \operatorname{tr}\left(M^\top M\right) = \|M\|^2$$

13) Dans la suite de cette partie, soit A une matrice de $\mathscr{M}_n(\mathbb{R})$, soit $U \in \mathscr{O}_n(\mathbb{R})$ et $S \in \mathscr{S}_n^+(\mathbb{R})$ telles que $A = US$; il existe une matrice diagonale D et une matrice P de $\mathscr{O}_n(\mathbb{R})$ telles que $S = PDP^{-1}$.

a) Montrer que, pour toute matrice Ω de $\mathscr{O}_n(\mathbb{R})$, $\|A - \Omega\| = \|S - U^{-1}\Omega\|$ et en déduire que, $d(A, \mathscr{O}_n(\mathbb{R})) = d(S, \mathscr{O}_n(\mathbb{R}))$.

Correction : On a directement, à l'aide de la question précédente :

$$\|A - \Omega\| = \|US - \Omega\| = \left\|U\left(S - U^{-1}\Omega\right)\right\| = \left\|S - U^{-1}\Omega\right\|$$

L'application de $\mathscr{O}_n(\mathbb{R})$ dans lui-même, qui à Ω associe $U^{-1}\Omega$ est une bijection (d'application réciproque $\Omega \mapsto U\Omega$).*cette remarque est importante pour justifier que les bornes inférieures sont identiques*. On en déduit que

$$d\left(A, \mathscr{O}_n(\mathbb{R})\right)=\inf_{\Omega \in \mathscr{O}_n(\mathbb{R})}\|A-\Omega\|=\inf_{\Omega \in \mathscr{O}_n(\mathbb{R})}\left\|S-U^{-1}\Omega\right\|=\inf_{V \in \mathscr{O}_n(\mathbb{R})}\|S-V\|=d(S, \mathscr{O}_n(\mathbb{R}))$$

b) Montrer que, $d(A, \mathscr{O}_n(\mathbb{R}))=d(D, \mathscr{O}_n(\mathbb{R}))$
Correction : Pour $\Omega \in \mathscr{O}_n(\mathbb{R})$, à l'aide de la question **12)**, on a

$$\|S-\Omega\|=\left\|PDP^{-1}-\Omega\right\|=\left\|P\left(D-P^{-1}\Omega P\right)P^{-1}\right\|=\left\|D-P^{-1}\Omega P\right\|$$

L'application de $\mathscr{O}_n(\mathbb{R})$ dans lui-même, qui à Ω associe $P^{-1}\Omega P$ est bijective. On en déduit que

$$d\left(S, \mathscr{O}_n(\mathbb{R})\right)=\inf_{\Omega \in \mathscr{O}_n(\mathbb{R})}\|S-\Omega\|=\inf_{\Omega \in \mathscr{O}_n(\mathbb{R})}\left\|D-P^{-1}\Omega P\right\|=\inf_{V \in \mathscr{O}_n(\mathbb{R})}\|D-V\|=d(D, \mathscr{O}_n(\mathbb{R}))$$

14) On note $D=\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

a) Montrer que pour toute matrice Ω de $\mathscr{O}_n(\mathbb{R})$, $\|D-\Omega\|^2=\sum_{i=1}^{n}\lambda_i^2-2\operatorname{tr}(D\Omega)+n$.

Correction : Soit Ω une matrice orthogonale, par linéarité de la trace :

$$\|D-\Omega\|^2=\operatorname{tr}\left(({}^\top D-\Omega)(D-\Omega)\right)=\operatorname{tr}\left(D^2-\Omega^\top D-D\Omega+\Omega^\top\Omega\right)=\operatorname{tr}(D^2)-\operatorname{tr}(\Omega^\top D)-\operatorname{tr}(D\Omega)+\operatorname{tr}(I_n)$$

En utilisant les propriétés de la trace on a $\operatorname{tr}(\Omega^\top D)=\operatorname{tr}\left(({}^\top D\Omega\right)=\operatorname{tr}(D\Omega)$.

On a donc : $\|D-\Omega\|^2=\sum_{i=1}^{n}\lambda_i^2-2\operatorname{tr}(D\Omega)+n$.

b) Montrer que pour toute matrice Ω de $\mathscr{O}_n(\mathbb{R})$, $\operatorname{tr}(D\Omega)\leqslant\sum_{i=1}^{n}\lambda_i$.

Correction : Soit $\Omega \in \mathscr{O}_n(\mathbb{R})$. On note $\Omega=(\omega_{i,j})$.
On a le coefficient d'indice (i,i) de $D\Omega$ est $\lambda_i\omega_{i,i}$

Comme les colonnes de Ω sont des vecteurs de $\mathscr{M}_{n,1}(\mathbb{R})$ normés on a, pour tout $j \in [\![1,n]\!]$ $\sum_{i=1}^{n}\omega_{i,j}^2=1$.

On en déduit que, pour tout $(i,j)\in[\![1,n]\!]^2$, $|w_{i,j}|\leqslant 1$, et par suite, comme les valeurs propres sont positives, que $\lambda_i\omega_{i,i}\leqslant\lambda_i$.

On a alors $\operatorname{tr}(D\Omega)\leqslant\sum_{i=1}^{n}\lambda_i$.

c) Conclure que $d(D, \mathscr{O}_n(\mathbb{R}))=\|D-I_n\|$.
Correction : d'après les deux questions précédentes on a

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

Par définition de $d\left(D, \mathscr{O}_n(\mathbb{R})\right)$ on a alors $d\left(D, \mathscr{O}_n(\mathbb{R})\right)\geqslant\|D-I_n\|$. Or I_n est un élément de $\mathscr{O}_n(\mathbb{R})$ donc $d(D, \mathscr{O}_n(\mathbb{R}))=\|D-I_n\|$.

15) Montrer que , $d(A, \mathscr{O}_n(\mathbb{R}))=\|A-U\|$.

Correction : d'après les questions précédentes on a $d(A, \mathscr{O}_n(\mathbb{R})=\|D-I_n\|$. On utilise alors le « chemin inverse »des questions précédentes :

$$\|D-I_n\|=\left\|P(D-I_n)P^{-1}\right\|=\|S-I_n\|=\|U(S-I_n)\|=\|A-U\|$$

d'où le résultat.

16) Calculer $d(\Gamma, \mathscr{O}_3(\mathbb{R}))$ où Γ est la matrice exemple de la partie I.

Correction : on a dans l'exemple, $\lambda_1=4$ et $\lambda_2=\lambda_3=1$, d'où $d(\Gamma, \mathscr{O}_3(\mathbb{R}))=3$

IV. Calcul de la distance de A à Δ_p.

17) Un résultat de densité.

a) Soit M un élément de $\mathscr{M}_n(\mathrm{IR})$, 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.

Correction : une matrice élément de $\mathscr{M}_n(\mathrm{IR})$ admet un nombre fini de valeurs propres. Si on note α la plus petite valeur propre strictement positive, avec $\alpha = +\infty$ si M ,n'a pas de valeurs propres strictement positive, alors pour tout $\lambda \in]0, \alpha[$ $\chi_M(\lambda) \neq 0$. Or $\chi_M(\lambda) = \det(M - \lambda I_n)$. On en déduit que $M - \lambda I_n$ est inversible.

b) En déduire que $GL_n(\mathrm{IR})$ est dense dans $\mathscr{M}_n(\mathrm{IR})$.

Correction : Soit $M \in \mathscr{M}_n(\mathrm{IR})$, on pose, pour $k \in \mathrm{IN}$, $M_k = M + \frac{1}{k+1} I_n$. D'après la question précédente les éléments de cette suite $(M_k)_{k\in\mathrm{IN}}$ sont, à partir d'un certain rang, des matrices inversibles. De plus, pour $k \in \mathrm{IN}$, $\|M - M_k\| = \left\|\frac{1}{k+1} I_n\right\| = \frac{\sqrt{n}}{k+1}$. Cette suite tend bien vers 0 lorsque k tend vers $+\infty$, donc $(M_k)_{k\in\mathrm{IN}}$ converge vers M.
Conclusion : $GL_n(\mathrm{IR})$ est dense dans $\mathscr{M}_n(\mathrm{IR})$.

18) Soit A un élément de $\mathscr{M}_n(\mathrm{IR})$, déterminer, pour tout entier naturel $p \leqslant n$, $d(A, \Delta_p)$.

Correction : Soit A un élément de $\mathscr{M}_n(\mathrm{IR})$. Soit p un entier inférieur ou égal à n. L'ensemble $GL_n(\mathrm{IR})$ est inclus dans Δ_p, or d'après la question précédente $d\left(A, GL_n(\mathrm{IR})\right) = 0$, on en déduit que $d\left(A, \Delta_p\right) = 0$.

V. Calcul de la distance de A à ∇_p.

A. Théorème de Courant et Fischer

Soit A une matrice de $\mathscr{S}_n(\mathrm{IR})$. On notera $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ ses valeurs propres, on notera $D = \mathrm{diag}(\lambda_1, \lambda_2, .., \lambda_n)$, P la matrice de $\mathscr{O}_n(\mathrm{IR})$ vérifiant $A = PDP^\top$ et $C_1, C_2, \ldots, C_n$ les matrices de $\mathscr{M}_{n,1}(\mathrm{IR})$ 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 $\mathscr{M}_{n,1}(\mathrm{IR})$ de dimension k. Nous allons montrer que :
$\lambda_k = \max\limits_{F\in\Psi_k} \min\limits_{X\in F\setminus\{0\}} \dfrac{X^\top AX}{X^\top X}$. (théorème de Courant et Fischer).

19) Soit X un vecteur de $\mathscr{M}_{n,1}(\mathrm{IR})$ de coordonnées $(x_1, x_2, \ldots, x_n)$ dans la base orthonormée $(C_1, C_2, .., C_n)$ de $\mathscr{M}_{n,1}(\mathrm{IR})$. Calculer en fonction des x_i et λ_i. (i compris entre 1 et n) : $X^\top AX$ et $X^\top X$ et, pour k entier entre 1 et n, $\dfrac{C_k^\top AC_k}{C_k^\top C_k}$.

Correction : par définition des (x_i) on a $X = \sum\limits_{i=1}^{n} x_i C_i$. Par définition de la base $(C_i)_{i[\![1,n]\!]}$, C_i est un vecteur propre de A associé à la valeur propre λ_i. On a donc $X^\top AX = X^\top \sum\limits_{i=1}^{n} x_i AC_i = \sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} \lambda_i x_j^\top x_i C_j C_i$. Comme la famille $(C_i)_{i\in[\![1,n]\!]}$ est orthonormale on a $X^\top AX = \sum\limits_{i=1}^{n} \lambda_i x_i^2$.

On a $X^\top X = \|X\|^2$, or la famille $(C_i)_{i\in[\![1,n]\!]}$ est une base orthonormale et $X = \sum\limits_{i=1}^{n} x_i C_i$, on en déduit que $X^\top X = \sum\limits_{i=1}^{n} x_i^2$.

On applique directement les égalités trouvées pour $X = C_k$, dans ce cas $x_j = 0$ pour $j \neq k$ et $x_k = 1$, d'où $\dfrac{C_k^\top AC_k}{C_k^\top C_k} = \lambda_k$.

20) Soit k entier entre 1et n, on pose $F_k = \mathrm{vect}\{C_1, C_2, \ldots, C_k\}$.
Montrer que pour tout X non nul de F_k , $\dfrac{X^\top AX}{X^\top X} \geqslant \lambda_k$ et déterminer $\min\limits_{X\in F_k\setminus\{0\}} \dfrac{X^\top AX}{X^\top X}$.

Correction : Soit $X \in F_k$. Par définition de F_k, il existe $(x_i)_{i\in[\![1,k]\!]}$ tel que $X = \sum_{i=1}^{k} x_i C_i$. En utilisant les mêmes arguments que la question précédente on a $X^\top AX = \sum_{i=1}^{k} \lambda_i x_i^2$ et $X^\top X = \sum_{i=1}^{k} x_i^2$. La suite des valeurs propres est décroissante, donc $X^\top AX \geqslant \lambda_k \sum_{i=1}^{k} x_i^2$. On en déduit que $\dfrac{X^\top AX}{X^\top X} \geqslant \lambda_k$.

C_k est un élément de F_k donc le minimum est atteint $\min_{X\in F_k\setminus\{0\}} \dfrac{X^\top AX}{X^\top X}$ en C_k et par suite, d'après la question précédente, $\min_{X\in F_k\setminus\{0\}} \dfrac{X^\top AX}{X^\top X} = \lambda_k$

21) Soit $F \in \Psi_k$

a) montrer que $\dim(F \cap \text{vect}\{C_k, C_{k+1}, .., C_n\}) \geqslant 1$.
Correction : on a :

$$\dim(F\cap\text{vect}\{C_k, C_{k+1}, .., C_n\}) = \dim(F)+\dim(\text{vect}\{C_k, C_{k+1}, .., C_n\})-\dim(F + \text{vect}\{C_k, C_{k+1}, .., C_n\})$$

On a $\dim(F + \text{vect}\{C_k, C_{k+1}, .., C_n\}) \leqslant n$, $\dim(F) = k$ par hypothèse et $\dim(\text{vect}\{C_k, C_{k+1}, .., C_n\}) = n - k + 1$. D'où le résultat.

b) Si X est un vecteur non nul de $F \cap \text{vect}\{C_k, C_{k+1}, .., C_n\}$, montrer que $\dfrac{X^\top AX}{X^\top X} \leqslant \lambda_k$.
Correction : soit X est un vecteur non nul de $F \cap \text{vect}\{C_k, C_{k+1}, .., C_n\}$, il existe $(x_i)_{i\in[\![k,n]\!]}$ tel que $X = \sum_{i=k}^{n} x_i C_i$. En utilisant les mêmes arguments que la question précédente on a $X^\top AX = \sum_{i=k}^{n} \lambda_i x_i^2$ et $X^\top X = \sum_{i=k}^{n} x_i^2$. La suite des valeurs propres est décroissante, donc $X^\top AX \leqslant \lambda_k \sum_{i=k}^{n} x_i^2$. On en déduit que $\dfrac{X^\top AX}{X^\top X} \leqslant \lambda_k$.

22) Conclure.
Correction : Les questions **20)**, pour la minoration, et **21)** pour la majoration, donnent l'égalité du théorème de Courant Fischer.

B. Calcul de $d(A, \nabla_p)$

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

23) Montrer qu'il existe deux matrices E et P de $\mathscr{O}_n(\mathbb{R})$ et une matrice diagonale D à termes positifs telles que $A = EDP$. En déduire que le rang de la matrice $A^\top A$ est encore r. *(On pourra utiliser les résultats de la question **9**))*

Correction : Soit P une matrice orthogonale et D une matrice diagonale à diagonale positive telles que $A^\top A = P^\top D^2 P$. On a alors $(AP^\top)^\top(AP^\top) = D^2$. On en déduit, d'après la question **9)**, qu'il existe E une matrice orthogonale telle que $AP^\top = ED$. On a alors $A = EDP$.

Comme E et P sont inversibles on a $rg(A) = rg(D) = rg(D^2) = rg(A^\top A)$.

24) Si on note les valeurs propres de la matrice symétrique réelle $A^\top A$ de rang r : $\mu_1 \geqslant \mu_2 \geqslant .. \geqslant \mu_r > 0$ et $\mu_{r+1} = \cdots = \mu_n = 0$, si on pose $D = \text{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}, \ldots, \sqrt{\mu_r}, 0, \ldots, 0)$, si pour $1 \leqslant l \leqslant n$ on note M_l la matrice de $\mathscr{M}_n(\mathbb{R})$ dont la l-ième colonne est celle de la matrice $E \in \mathscr{O}_n(\mathbb{R})$ de la question **23)**, tous les autres termes de M_l étant nuls, on a clairement : $ED = \sum_{l=1}^{n} \sqrt{\mu_l} M_l$.

Montrer alors qu'il existe une famille orthonormale $(R_1, R_2, \ldots, R_n)$ de matrices de $\mathscr{M}_n(\mathbb{R})$ (pour le produit scalaire $(A|B) = Tr(A^\top B)$ de $\mathscr{M}_n(\mathbb{R})$), toutes de rang un, et telles que $A = \sum_{l=1}^{n} \sqrt{\mu_l} R_l = \sum_{l=1}^{r} \sqrt{\mu_l} R_l$.

Correction : Posons $R_l = M_l P$ 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 que (R_l) est orthonormale :

$$(R_l \mid R_k) = \operatorname{tr}(P^\top M_l^\top M_k P) = \operatorname{tr}(M_l^\top M_k P P^\top) = \operatorname{tr}(M_l^\top M_k)$$

or $M_l^\top M_k$ 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 $\operatorname{rg} R_l = \operatorname{rg} M_l = 1$ (P est inversible et M_l a une et une seule colonne non nulle).

25) Avec les notations de la question **24)**, on pose $N = \sum_{l=1}^{p} \sqrt{\mu_l} R_l$.
Montrer que $\operatorname{rang}(N) \leqslant p$ puis que $d(A, \nabla_p) \leqslant \sqrt{\mu_{p+1} + \cdots + \mu_r}$.
Correction : On a clairement $\operatorname{Im}(N) \subset \operatorname{Im}(R_1) + \operatorname{Im}(R_2) + \cdots + \operatorname{Im}(R_p)$, puis $\operatorname{rg}(N) \leqslant p$ (les $\operatorname{Im}(R_i)$ sont des droites).

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

26) 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 $({}^\top A - M)(A - M)$ et on pose $G = \ker(M) \cap \operatorname{Im}(A^\top A)$.
Soit k un entier compris entre 1 et $r - p$.

a) Montrer que $\dim G \geqslant r - p$.
Correction : $\dim G = \dim(\ker M) + \dim(\operatorname{Im}(A^\top A)) - \dim(\ker(M) \cup \operatorname{Im}(A^\top A)) \geqslant (n-p) + r - n = r - p$.

b) Soit F un sous-espace vectoriel de G de dimension k, montrer que : $\alpha_k \geqslant \min\limits_{X \in F \setminus \{0\}} \dfrac{X^\top A^\top A X}{X^\top X}$.
Correction : En appliquant le théorème de Courant et Fischer (plus exactement en appliquant la question **21**) à la matrice $A - M$, nous obtenons :

$$\alpha_k \geq \min_{X \in F \setminus \{0\}} \frac{X^\top({}^\top A - M)(A - M)X}{X^\top X}$$

mais pour $X \in F$, $MX = 0$ et $X^\top M^\top = 0$, donc

$$\alpha_k \geqslant \min_{X \in F \setminus \{0\}} \frac{X^\top A^\top A X}{X^\top X}$$

c) On note $(V_1, V_2, .., V_n)$ une base de $\mathbb{R}^n$ formée de vecteurs propres de la matrice $A^\top A$, le vecteur V_i étant associé à la valeur propre μ_i de telle sorte que : $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_r > 0$ et $\mu_{r+1} = \cdots = \mu_n = 0$.
Montrer que $\dim(G \cap \operatorname{vect}\{V_1, V_2, .., V_{k+p}\}) \geqslant k$.
Correction : On a $G \cap Vect(V_1, \ldots, V_{k+p}) = \ker(M) \cap Vect(V_1, \ldots, V_{k+p})$ car $Vect(V_1, \ldots, V_{k+p}) \subset Vect(V_1, \ldots, V_r) = \operatorname{Im}(A^\top A)$ (on a $k \leqslant r-p$). On en déduit donc (comme au **a**) que $G \cap Vect(V_1, \ldots, V_{k+p})$ est de dimension au moins $(k+p) + (n-p) - n = k$.

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

- $\alpha_k \geqslant \min\limits_{X \in F \setminus \{O\}} \dfrac{X^\top A^\top A X}{X^\top X}$ d'après le **b** ;
- pour X élément quelconque de F, que l'on peut écrire sous la forme $X = \sum_{i=1}^{k+p} x_i V_i$:

$$\frac{X^\top A^\top A X}{X^\top X} = \frac{\sum_{i=1}^{k+p} \mu_i x_i^2}{\sum_{i=1}^{k+p} x_i^2} \geqslant \mu_{k+p}$$

car les μ_i décroissent.

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

27) En déduire $d(A, \nabla_p)$.

Correction : 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 :

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

On en déduit que $d(A, \nabla_p) \geq \sum_{i=p+1}^{r} \mu_i$, 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 $A^\top A$.

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.

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

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