MP 2 **Samedi 2 Mars 2024**

DEVOIR SURVEILLÉ 5 (4H)

Les calculatrices sont autorisées.

Notations

Soit n et p des entiers supérieurs ou égaux à 1. On désigne par $\mathcal{M}_{n,p}(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel des matrices à coefficients réels ayant n lignes et p colonnes. On identifiera $\mathcal{M}_{n,1}(\mathbb{R})$ et $\mathcal{M}_{p,1}(\mathbb{R})$ respectivement à $\mathbb{R}^n$ et $\mathbb{R}^p$ que l'on supposera munis de leurs produits scalaires canoniques notés respectivement $\langle \cdot \mid \cdot \rangle_n$ et $\langle \cdot \mid \cdot \rangle_p$. Les normes associées à ces produits scalaires seront notées respectivement $\| \cdot \|_n$ et $\| \cdot \|_p$.

On notera $(E_i)_{1\leq i\leq p}$ la base canonique de $\mathcal{M}_{p,1}(\mathbb{R})$ et $(F_j)_{1\leq j\leq n}$ celle de $\mathcal{M}_{n,1}(\mathbb{R})$.

Lorsque $p = n, \mathcal{M}_{n,n}(\mathbb{R})$ est noté plus simplement $\mathcal{M}_n(\mathbb{R})$. On note $0_{n,p}$ la matrice nulle de $\mathcal{M}_{n,p}(\mathbb{R})$ et 0_n la matrice nulle de $\mathcal{M}_n(\mathbb{R})$.

Pour A appartenant à $\mathcal{M}_{n,p}(\mathbb{R})$, $A^\top$ désigne la matrice transposée de A : c'est un élément de $\mathcal{M}_{p,n}(\mathbb{R})$.

Ker A est le noyau de A défini par $\operatorname{Ker} A = \{X \in \mathcal{M}_{p,1}(\mathbb{R}) \mid AX = 0\}$

Im A est l'image de A définie par $\operatorname{Im} A = \{AX \mid X \in \mathcal{M}_{p,1}(\mathbb{R})\}$

Enfin, $F^\perp$ désigne l'orthogonal d'un sous-espace vectoriel F d'un espace euclidien.

Partie I : la décomposition en valeurs singulières

Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$.

I.1 Montrer que $A^\top A$ est nulle si et seulement si A est nulle.

Dans toute la suite du problème A sera supposée non nulle.

I.2 Montrer que les matrices $A^\top A$ et $AA^\top$ sont diagonalisables au moyen de matrices orthogonales.

I.3) a) X, Y désignant deux éléments de $\mathcal{M}_{n,1}(\mathbb{R})$, exprimer le produit scalaire $\langle X \mid Y \rangle_n$ sous la forme d'un produit matriciel.

b) Si W est un vecteur propre de $A^\top A$ associé à la valeur propre λ, exprimer $\|AW\|_n^2$ en fonction de λ et $\|W\|_p$.

c) En déduire que les valeurs propres de $A^\top A$ sont réelles, positives ou nulles.

d) Montrer que $\ker(A^\top A) = \ker(A)$ et que $\operatorname{rg}(A^\top A) = \operatorname{rg}(A)$.

I.4 a) Pour x réel, calculer les produits matriciels par bloc suivants :

$$\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}\begin{pmatrix} -I_n & 0_{n,p} \\ A^\top & I_p \end{pmatrix} \text{ et } \begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}\begin{pmatrix} -I_n & A \\ 0_{p,n} & -xI_p \end{pmatrix}$$

b) Déduire du a) une relation entre les polynômes caractéristiques $\chi_{A.A^\top}$ et $\chi_{A^\top A}$. En déduire que les matrices $A^\top A$ et $AA^\top$ ont les mêmes valeurs propres non nulles avec le même ordre de multiplicité.

c) En déduire également que les matrices $A^\top A$ et $AA^\top$ ont même rang.

I.5. Montrer que si $n > p, 0$ est valeur propre de $AA^\top$ et que si $n < p, 0$ est valeur propre de $A^\top A$.

I.6 On note $\lambda_1, \lambda_2, \ldots, \lambda_p$ les valeurs propres de $A^\top A$, chaque valeur propre apparaissant dans cette liste un nombre de fois égal à son ordre de multiplicité et on pose $\mu_i = \sqrt{\lambda_i}$ pour tout i élément de $\{1, 2, \ldots, p\}$. Les réels μ_i sont appelés *valeurs singulières* de A.

On suppose les réels λ_i ordonnés tels que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$.

a) Montrer que λ_1 est non nul.

On définit alors un unique entier naturel r appartenant à $\{1, 2, \ldots, p\}$ comme suit : si toutes les valeurs propres de $A^\top A$ sont non nulles, $r = p$, sinon r est tel que pour tout $i \leq r$, $\lambda_i > 0$ et pour tout $i > r, \lambda_i = 0$.

Soit $(V_1, V_2, \ldots, V_p)$ une base orthonormale de vecteurs propres de $A^\top A$ respectivement associés aux valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_p$; $V_1, V_2, \ldots, V_r$ désignent les vecteurs propres associés aux valeurs propres non nulles et lorsque r est strictement inférieur à $p, V_{r+1}, \ldots, V_p$ désignent les vecteurs propres associés à la valeur propre 0 .

b) Montrer que $r \leq n$ et que la dimension de $\operatorname{Ker} AA^\top$ est égale à $n - r$.

Pour tout $i \in \{1, 2, \ldots, r\}$, on pose $U_i = \frac{1}{\mu_i} AV_i$ et si $n > r$, on désigne par $(U_{r+1}, \ldots, U_n)$ une base orthonormale de $\operatorname{Ker} AA^\top$.

c) Montrer que pour tout $i \in \{1, 2, \ldots, r\}, AV_i = \mu_i U_i$ et que si r est strictement inférieur à p, pour tout $i \in \{r+1, \ldots, p\}, AV_i = 0$.

d) Montrer que pour tout $i \in \{1, 2, \ldots, r\}$, $A^\top U_i = \mu_i V_i$.

e) Montrer que si $n > r$, pour tout $i \in \{r+1, \ldots, n\}$, $A^\top U_i = 0$.

f) En déduire que le système de vecteurs $(U_1, U_2, \ldots, U_n)$ constitue une base orthonormale de vecteurs propres de $AA^\top$ et préciser la valeur propre associée à chaque vecteur U_i.

I.7 . On note V la matrice carrée réelle d'ordre p dont le ième vecteur colonne est le vecteur V_i et U la matrice carrée réelle d'ordre n dont le jème vecteur colonne est le vecteur U_j et $(U^\top AV)_{i,j}$ l'élément de la i $^{\text{ème}}$ ligne, jème colonne de la matrice $U^\top AV$.

a) Montrer que :

$$\forall (i,j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, p\}, \quad \left(U^\top AV\right)_{i,j} = \mu_j \delta_{i,j} \text{ où } \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

b) On note Δ la matrice appartenant à $\mathcal{M}_{n,p}(\mathbb{R})$ dont tous les éléments $\Delta_{i,j}$ sont nuls sauf $\Delta_{11}, \Delta_{22}, \ldots, \Delta_{rr}$ respectivement égaux à $\mu_1, \mu_2, \ldots, \mu_r$. Montrer que $A = U\Delta V^\top$.
La factorisation de A ainsi obtenue est dite *décomposition de A en valeurs singulières*.

c) Trouver une décomposition en valeurs singulières de chacune des matrices :

$$A_0 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad B_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

I.8 a) Montrer que $V = \sum_{j=1}^{p} V_j E_j^\top$.

b) En déduire :

$$A = \sum_{i=1}^{r} \mu_i U_i V_i^\top \quad , \quad A^\top A = \sum_{i=1}^{r} \lambda_i V_i V_i^\top \quad , \quad AA^\top = \sum_{i=1}^{r} \lambda_i U_i U_i^\top$$

c) (i) Montrer que si $r < p, \operatorname{Ker} A = \operatorname{Vect}(V_{r+1}, .., V_p)$ et si $r = p, \operatorname{Ker} A = \{0\}$

(ii) Déterminer de même à l'aide des U_i et V_j des bases orthonormales de $\operatorname{Ker} A^\top, \operatorname{Im} A, \operatorname{Im} A^\top$.

I.9 Epilogue de cette partie :

Définition générale : on appelle *décomposition en valeurs singulières* d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$ toute écriture : $A = U\Delta V^\top$ avec $U \in O_n(\mathbb{R})$, $V \in O_p(\mathbb{R})$, Δ comme ci-dessus.

Justifier qu'on vient bien de montrer l'existence d'une décomposition en valeurs singulières pour toute matrice A. Expliquer pourquoi une telle décomposition n'est pas forcément unique.

Partie II : application à la pseudo-inverse d'une matrice

Avec les notations de la partie **I**, pour $A \in \mathcal{M}_{n,p}(\mathbb{R})$ qu'on écrit avec une décomposition en valeurs singulières $A = U\Delta V^\top$, on appelle Δ^+la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ dont tous les éléments $\Delta^+_{i,j}$sont nuls sauf $\Delta^+_{11}, \Delta^+_{22}, \ldots, \Delta^+_{rr}$respectivement égaux à $\frac{1}{\mu_1}, \frac{1}{\mu_2}, \ldots, \frac{1}{\mu_r}$ et on pose

$$A^+ = V\left(\Delta^+\right) U^\top.$$

La matrice Δ^+(resp. A^+) est appelée pseudo-inverse de Δ (resp. de A). A priori, la matrice A^+ainsi définie dépend de la décomposition en valeurs singulières choisie pour la matrice A, mais il sera montré au II. 9 qu'il n'en est rien et que A^+est uniquement déterminée à partir de A.

II.1 Pour la matrice A_0 du I.7 c), calculer les matrices $A_0^+, A_0 A_0^+, A_0^+ A_0, A_0 A_0^+ A_0$ et $A_0^+ A_0 A_0^+$.

II.2 Déterminer aussi $(A_0^+)^+$.

II.3 Evaluer $\Delta^+\Delta$ et $\Delta\Delta^+$ pour toute matrice Δ comme au I. 7. b).

II.4 Montrer que si A est une matrice carrée inversible ($n = p = r$), alors $A^+ = A^{-1}$.

II.5 Montrer de manière analogue au I.8) que :

$$A^+ = \sum_{i=1}^{r} \frac{1}{\mu_i} V_i U_i^\top$$

On obtient aussi facilement (démonstration non demandée) que

$$AA^+ = \sum_{i=1}^{r} U_i U_i^\top \quad , \quad A^+ A = \sum_{i=1}^{r} V_i V_i^\top$$

II.6 Evaluer AA^+U_j pour tout $j \in \{1, 2, \ldots, n\}$ et en déduire que AA^+ est la matrice dans la base canonique de $\mathbb{R}^n$ de la projection orthogonale de $\mathbb{R}^n$ sur $\operatorname{Im} A$.

On montre de même (démonstration non demandée) que A^+A est la matrice dans la base canonique de $\mathbb{R}^p$ de la projection orthogonale de $\mathbb{R}^p$ sur $(\operatorname{Ker} A)^\perp$.

II.7 Etablir les identités suivantes :

$$AA^+ = (AA^+)^\top \quad , \quad A^+A = (A^+A)^\top \quad , \quad AA^+A = A \quad , \quad A^+AA^+ = A^+ \tag{1}$$

II.8 Soit B une matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ vérifiant :

$$AB \overset{(2.1)}{=} (AB)^\top \quad , \quad BA \overset{(2.2)}{=} (BA)^\top \quad , \quad ABA \overset{(2.3)}{=} A \quad , \quad BAB \overset{(2.4)}{=} B \tag{2}$$

a) Montrer que AB est la matrice dans la base canonique de $\mathbb{R}^n$ de la projection orthogonale de $\mathbb{R}^n$ sur $\operatorname{Im} A$.

b) Montrer que BA est la matrice dans la base canonique de $\mathbb{R}^p$ de la projection orthogonale de $\mathbb{R}^p$ sur $(\ker A)^\perp$.

c) On note a (resp. b) l'endomorphisme canonique associé à A (resp. à B) et a' la restriction de a à $(\ker a)^\perp$, qui est un isomorphisme entre $(\ker a)^\perp$ et $\operatorname{Im} a$.

Montrer que $b_{|\operatorname{Im} a} = (a')^{-1}$ et $b_{|(\operatorname{Im} a)^\perp} = 0$, ce qui montre *l'unicité d'une matrice B vérifiant* (2)

II.9 Justifier qu'on vient de démontrer le théorème suivant :

> **Théorème :** *Pour toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$, il existe une unique matrice B vérifiant* (2). *Cette matrice s'appelle* pseudo-inverse *de A et elle coïncide avec la matrice A^+ précédente, ce qui donne un moyen pratique de la calculer.*

L'interprétation géométrique du II.8 c) dit que ce calcul est utile pour les calculs de distance entre un vecteur de $\mathbb{R}^n$ et $\operatorname{Im}(A)$. Précisément :

II.10 a) Montrer que pour tout $Y \in \mathcal{M}_{n,1}(\mathbb{R})$, la distance de Y à $\operatorname{Im} A$ est donnée par :

$$d(Y, \operatorname{Im} A) = \|Y - A.A^+Y\|$$

b) Soit $Y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Calculer $\min_{X \in \mathbb{R}^2} \|A_0 X - Y\|$.

II.11 a) *L'interprétation géométrique du II. 8 c) permet aussi facilement de comprendre l'involutivité de $A \mapsto A^+$:* montrer que $(A^+)^+ = A$.

b) Enfin calculer $(A_0 B_0)^+$et $B_0^+ A_0^+$: a-t-on l'égalité ?