Devoir surveillé 5 (4h) : CCINP PC 2001

I 1) Pour tout $i \in \{1,2,\ldots,p\}, (A^\top A)_{i,i} = \sum_{k=1}^{n} (a_{k,i})^2$
donc si $A^\top A = 0$, pour tout $k \in \{1,2,\ldots,n\}$ et pour tout $i \in \{1,2,\ldots,p\}, (A)_{k,i} = 0$

Remarque : on peut aussi dire qu'avec les deux formes du produit scalaire canonique dans $\mathcal{M}_{n,p}(\mathbb{R})$ donne ici que $\mathrm{Tr}(A^\top A) = \sum_{i,j} a_{i,j}^2$, donc si $A^\top A = 0$ alors tous les $a_{i,j}$ sont nuls, donc $A = 0$.

Réciproquement si $A = 0$ alors $A^\top A = 0$, donc

$$A^\top A = 0 \text{ si et seulement } A = 0$$

I 2) Par le théorème spectral, il suffit de montrer que ces deux matrices sont symétriques réelles.

Ce sont bien des matrices réelles, et par le fait pour deux matrices M, N tels que le produit $M.N$ soit défini, on a $(M.N)^\top = N^\top M^\top$, on a immédiatement

$$(A^\top.A)^\top = A^\top (A^\top)^\top = A^\top A$$

donc $A^\top A$ est bien symétrique réelle, donc diagonalisable au moyen de matrices orthogonales.

De même pour $AA^\top$.

I3) a) $\langle X \mid Y\rangle_n = \sum_{i=1}^{n} x_i y_i = X^\top Y = Y^\top X$.

b) W est un vecteur propre de $A^\top A$ associé à la valeur propre λ, donc $W \neq 0$ et $A^\top AW = \lambda W$.

Or $\|AW\|_n^2 = (AW)^\top AW = W^\top (A^\top AW)$

Donc $\|AW\|_n^2 = W^\top(\lambda W) = \lambda W^\top W = \lambda \|W\|_p^2$:

$$\|AW\|_n^2 = \lambda \|W\|_p^2$$

c) Comme $\|W\|_p^2 > 0$ et $\|AW\|_n^2 \geq 0$, par l'égalité prouvée à la fin du b), $\lambda \geq 0$

Les valeurs propres de $A^\top A$ sont donc réelles, positives ou nulles.

N.B. On pouvait aussi dire que $A^\top A$ est une matrice symétrique positive avec la propriété $X^\top A^\top AX = \|AX\|^2 \geq 0$ et ensuite c'est un résultat de cours que ses v.p. sont positives ou nulles.

d) Si $AX = 0$ alors $A^\top AX = 0$ ce qui donne une inclusion.

Et réciproquement si $A^\top AX = 0$ alors, comme au b), $X^\top A^\top AX = 0$ et donc $\|AX\|^2 = 0$ donc $AX = 0$.

On a donc bien montré l'égalité $\ker(A) = \ker(A^\top A)$.

Alors par théorème du rang on a $\mathrm{rg}(A^\top A) = p - \dim\ker(A^\top A) = p - \dim\ker A = \mathrm{rg}(A)$.

I. 4 a)

$$\text{a) } \begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} \begin{pmatrix} -I_n & 0_{n,p} \\ A^\top & I_p \end{pmatrix} = \begin{pmatrix} AA^\top - xI_n & A \\ 0_{p,n} & I_p \end{pmatrix}$$

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

b) On sait que $\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \times \det B$ si A et B sont des matrices carrées

Si on désigne par χ_M le polynôme caractéristique d'une matrice $M \in \mathcal{M}_n(\mathbb{R})$, on sait que

$$\chi_M(x) = \det(xI_n - M).$$

Alors, d'une part :

$$\det\left(\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} \begin{pmatrix} -I_n & 0_{n,p} \\ A^\top & I_p \end{pmatrix}\right) = \det \begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} \times \det \begin{pmatrix} -I_n & 0_{n,p} \\ A^\top & I_p \end{pmatrix} \overset{(1)}{=} (-1)^n \det \begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}$$

$$\det \begin{pmatrix} AA^\top - xI_n & A \\ & 0_{p,n} & I_p \end{pmatrix}) \overset{(2)}{=} (-1)^n \chi_{AA^\top}(x)$$

Donc par (1), (2) et le a), on a obtenu :

$$\chi_{AA^\top}(x) \overset{(3)}{=} \det\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}$$

D'autre part :

$$\det\left(\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}\begin{pmatrix} -I_n & A \\ 0_{p,n} & -xI_p \end{pmatrix}\right) = \det\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} \times \det\begin{pmatrix} -I_n & A \\ 0_{p,n} & -xI_p \end{pmatrix}$$
$$\text{donc}\det\left(\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}\begin{pmatrix} -I_n & A \\ 0_{p,n} & -xI_p \end{pmatrix}\right) \overset{(4)}{=} (-1)^n(-x)^p\det\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix}$$

et

$$\det\begin{pmatrix} -xI_n & 0_{n,p} \\ -A^\top & A^\top A - xI_p \end{pmatrix} \overset{(5)}{=} (-x)^n(-1)^p\chi_{A^\top A}(x)$$

Donc par (4), (5) et le a), on obtient :

$$(-1)^n(-x)^p\det\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} = (-x)^n(-1)^p\chi_{A^\top A}(x)$$

donc en simplifiant par $(-1)^{n+p}$ dans les deux membres :

$$x^p\det\begin{pmatrix} xI_n & A \\ A^\top & I_p \end{pmatrix} \overset{(6)}{=} x^n\chi_{A^\top A}(x)$$

En comparant (3) et (6), on a obtenu :

$$\underline{x^n\chi_{A^\top A}(x) \overset{(7)}{=} x^p\chi_{AA^\top}(x)}$$

Par la relation (7) les polynômes $\chi_{A^\top A}$ et $\chi_{AA^\top}$ ont les mêmes racines non nulles avec le même ordre de multiplicité.

Remarque utile pour le c) : ces polynômes sont scindés dans $\mathbb{R}[X]$, puisque les matrices $A^\top A$ et $AA^\top$ sont diagonalisables et mieux les multiplicités algébriques sont aussi les dimensions des s.e.v propres.

Donc ici $\mathrm{Sp}(A.A^\top) \setminus \{0\} = \mathrm{Sp}(A^\top A) \setminus \{0\}$ et pour toute valeur propre λ non nulle dans ce spectre :

$$\dim E_\lambda(A.A^\top) = \dim E_\lambda(A^\top.A).$$

c) Pour toute matrice diagonalisable $M = PDP^{-1}$ avec D diagonale, on a $\mathrm{rg}(M) = \mathrm{rg}(D) = \sum\limits_{\lambda\,\mathrm{Sp}(D)\setminus\{0\}} \dim E_\lambda(D) = \sum\limits_{\lambda\,\mathrm{Sp}(D)\setminus\{0\}} \dim E_\lambda(M)$

Donc ici $\mathrm{rg}(A.A^\top) = \sum\limits_{\lambda\,\mathrm{Sp}(A.A^\top)\setminus\{0\}} \dim E_\lambda(A.A^\top)$ et d'après la remarque faite à la fin de la solution du b), on en déduit que :

$$\mathrm{rg}(A.A^\top) = \mathrm{rg}(A^\top.A).$$

I.5. Si $n > p$, la relation du I.4.b) donne $\chi_{AA^\top}(x) = (x)^{n-p}\chi_{A^\top A}(x)$ donc 0 est racine de $\chi_{AA^\top}$ donc 0 est valeur propre de $AA^\top$.

De même si $p > n$ en échangeant les rôles de A et $A^\top$.

I.6. a) A étant non nulle, on a vu au I.1) que $A^\top A$ est non nulle, et elle est diagonalisable, elle a donc au moins une valeur propre non nulle. Toutes ses valeurs propres sont positives, donc la plus grande des valeurs propres est strictement positive :

$$\lambda_1 > 0$$

b) Par déf. de r, r est le rang de la matrice diagonale à laquelle $A^\top A$ est semblable, donc $r = \mathrm{rg}(A^\top A)$.

Mais, encore par ce 4.c), on en déduit que $r = \mathrm{rg}(A.A^\top)$ et comme $A.A^\top \in \mathcal{M}_n(\mathbb{R})$, on a bien :

$$r \leq n$$

Et d'après le théorème du rang ,

$$\dim\left(\operatorname{Ker} AA^\top\right) = n - r$$

Remarque : on peut aussi dire que $r = \mathrm{rg}(A^\top A) = \mathrm{rg}(A)$ par la question I.3 d). On a de même en échangeant les rôles des matrices $\mathrm{rg}(A.A^\top) = \mathrm{rg}(A^\top)$.

Comme par le cours $\mathrm{rg}(A) = \mathrm{rg}(A^\top)$ tous ces rangs sont égaux.

c) Pour tout $i \in \{1, 2, \ldots, r\}, AV_i = \mu_i U_i$ par définition des U_i.

Si $r > p$, alors pour tout $i \in \{r+1, .., p\}, \lambda_i = 0, A^\top AV_i = 0$, donc $V_i \in \ker(A^\top A) = \ker(A)$ donc $AV_i = 0$.

d) Pour tout $i \in \{1, 2, .., r\}, A^\top U_i = \frac{1}{\mu_i}\left(A^\top AV_i\right) = \frac{\lambda_i}{\mu_i} V_i = \mu_i V_i$ par déf. de $\mu_i = \sqrt{\lambda_i}$.

e) Si $n > r$, pour tout $i \in \{r+1, .., n\}, U_i \in \operatorname{Ker} AA^\top = \ker(A^\top)$ par I. 3 d) donc pour tout $i \in \{r+1, .., n\}, A^\top U_i = 0$

f) D'après le d) et le c) pour tout $i \in \{1, 2, .., r\}$, $A^\top U_i = \mu_i V_i$. et donc

$$A.A^\top U_i = \mu_i AV_i = \mu_i^2 U_i = \lambda_i U_i \quad (*)$$

Pour tout $i \in \{1, 2, .., r\}$, pour tout $j \in \{1, 2, .., r\}$,

$$\langle U_i \mid U_j \rangle = \frac{1}{\mu_i \mu_j} V_i^\top \left(A^\top AV_j\right) = \frac{\lambda_j}{\mu_i \mu_j} \langle V_i \mid V_j \rangle = \delta_{ij}$$

puisque $\lambda_j = \mu_j^2$ et que $(V_1, \ldots, V_r)$ est une famille orthonormée.

Ceci montre que la famille $(U_1, \ldots, U_r)$ est orthonormée, en particulier les vecteurs U_i sont non nuls, et par $(*)$ chaque U_i est donc propre pour la v.p. λ_i pour $A.A^\top$.

Si $n > r$, par définition $(U_{r+1}, ..U_n)$ est une famille orthonormale, propre pour la v.p. 0.

Il reste seulement à montrer que si $i \leq r$ et $j > r$ alors $\langle U_i, U_j \rangle = 0 \quad (**)$.

Or, en notant que $U_i = \dfrac{1}{\mu_i} AV_i$ par définition, on a :

$$\langle U_i, U_j \rangle = \frac{1}{\mu_i} V_i^\top A^\top U_j = 0$$

la dernière égalité étant donnée par la question e).

On a donc bien montré $(**)$ ce qui achève de montrer que $(U_1, \ldots, U_n)$ est une famille orthonormale, et vue sa longueur, c'est une base.

I.7. a) Soit $(i, j) \in \{1, 2, .., n\} \times \{1, 2, .., p\}$ alors par définition du produit de matrices :

$$\left(U^\top AV\right)_{ij} = L_i(U^\top) C_j(AV)$$

où $L_i(U^\top)$ est la i-ième ligne de $U^\top$ et $C_j(AV)$ la j-ième colonne de AV.

Mais par déf. de la transposée $L_i(U^\top) = U_i^\top$ et encore par déf. du produit de matrices ; $C_j(AV) = AV_j$.

Ainsi :

$$\left(U^\top AV\right)_{ij} = U_i^\top AV_j = \langle U_i \mid AV_j \rangle$$

Pour tout $j \in \{1, 2, .., r\}, AV_j = \mu_j U_j$ donc $\langle U_i \mid AV_j \rangle = \mu_j \langle U_i \mid U_j \rangle = \mu_j \delta_{ij}$.

En outre si $r < p$, pour tout $j \in \{r+1, \ldots, p\}, \mu_j = 0$ et $AV_j = 0$ donc $\langle U_i \mid AV_j \rangle = \mu_j \delta_{ij} = 0$ dans ce cas.

On a bien montré (avec les deux cas sur j) que :

$$\forall (i, j) \in \{1, 2, .., n\} \times \{1, 2, .., p\}, \left(U^\top AV\right)_{ij} = \mu_j \delta_{ij}$$

b) Par le a), on a montré que $U^\top AV = \Delta$

Les vecteurs colonnes de U constituent une base orthonormale de $\mathbb{R}^n$, donc U est une matrice orthogonale, $U \in O(n)$, U est inversible et $U^{-1} = U^\top$

De même $V \in O(p)$ et V est inversible avec $V^{-1} = V^\top$
Donc, en multipliant l'égalité $U^\top AV = \Delta$ à droite par $V^\top$ et à gauche par U on obtient :

$$A = U\Delta V^\top$$

c) (i) On commence pour A_0, par chercher les vecteurs V_i qui forment une b.o.n. de vecteurs propres de $A_0^\top A_0$.

Or $A_0^\top A_0 = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}$ est déjà diagonale donc de valeurs propres (ordonnées dans le sens décroissant) $\lambda_1 = 6, \lambda_2 = 2$ et vecteurs propres $V_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, V_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Ici $r = 2$, on calcule pour $i = 1, 2$ les $U_i = \dfrac{1}{\mu_i} A_0 V_i$ donc ici ;

$U_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, U_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

On calcule ensuite $A_0.A_0^\top = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 4 \end{pmatrix}$ de noyau dirigé part le vecteur : $U_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$,

Ainsi la décomposition en valeurs singulières s'écrit :

$$A_0 = U\Delta V^\top \text{ avec } U = \begin{pmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}, \Delta = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

De même pour B_0, on calcule ;

$$B_0^\top B_0 = (2), \lambda_1 = 2,\ V_1 = (1),$$
$$B_0.B_0^\top = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},\ U_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, U_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$B_0 = U\Delta V^\top \text{ avec } U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \Delta = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}, V^\top = (1)$$

I.8. a) Par définition du produit de matrice , $V_j E_j^\top$ est la matrice de $\mathcal{M}_p(\mathbb{R})$ dont toutes les colonnes sont nulles sauf la j-ième égale à V_j, donc d'après la définition de V :

$$V = \sum_{j=1}^{p} V_j E_j^\top$$

b) On a de même $U = \sum_{i=1}^n U_i F_i^\top$ donc $A = U\Delta V^\top = \sum_{i=1}^n \left(\sum_{j=1}^p U_i F_i^\top \Delta E_j V_j^\top \right) \quad (*)$
Or pour tout i, j, $F_i^\top \Delta E_j = \Delta_{i,j}$ et $\Delta_{i,j} = \delta_{i,j}\mu_i$ pour $i \leq r$ et 0 sinon.
Donc, en ne gardant dans $(*)$ que les termes tels que $i = j \leq r$, on obtient :

$$A = U\Delta V^\top = \sum_{i=1}^{r} U_i \mu_i V_i^\top = \sum_{i=1}^{r} \mu_i U_i V_i^\top$$

comme demandé par le sujet. Le reste des calculs est alors immédiat :

$$A^\top A = A^\top \left(\sum_{i=1}^{r} \mu_i U_i V_i^\top \right) = \sum_{i=1}^{r} \mu_i \left(A^\top U_i \right) V_i^\top, \text{ or } A^\top U_i = \mu_i V_i \text{par I.6. d) et } \lambda_i = \mu_i^2, \text{ donc}$$

$$A^\top A = \sum_{i=1}^{r} \lambda_i V_i V_i^\top.$$

Enfin $A^\top = \sum_{i=1}^r \mu_i V_i U_i^\top$, donc $A.A^\top = \sum_{i=1}^r \mu_i \left(AV_i \right) U_i^\top = \sum_{i=1}^r \mu_i \left(\mu_i U_i \right) U_i^\top$, donc

$$AA^\top = \sum_{i=1}^{r} \lambda_i U_i U_i^\top .$$

c) (i) **(M1)** On sait par I4 d) que $\ker(A) = \ker(A^\top A)$ et par définition $(V_1, \dots, V_p)$ est une base de diagonalisation de $A^\top A$ avec $V_{r+1}, \dots, V_p$ les vecteurs propres associés à la valeur propre 0, d'où la conclusion.

(M2) Soit $X \in \mathbb{R}^p, V_i^\top X = \langle V_i \mid X \rangle$, donc $AX = \left(\sum_{i=1}^r \mu_i U_i V_i^\top\right) X = \sum_{i=1}^r \mu_i \langle V_i \mid X \rangle U_i$

$(U_1, U_2, .., U_r)$ est une famille libre de $\mathbb{R}^n$,et pour tout $i \in \{1, 2, .., r\}, \mu_i \neq 0$

donc $AX = 0$ si et seulement si pour tout $i \in \{1, 2, .., r\}, \langle V_i \mid X \rangle = 0$

$(V_1, V_2, \dots, V_p)$ étant une base orthormale de $\mathbb{R}^p$,

(ii) On peut adapter au choix la (M1) ou la (M2) du (i) :

Avec la (M1), on a vu que $(U_1, \dots, U_n)$ est une b.o.n. de dz de $A.A^\top$. et que $U_{r+1}, \dots, U_n$ sont les vecteurs propres associés à la valeur propre 0.

Donc $\boxed{\ker(A^\top) = \ker(A.A^\top) = \mathrm{Vect}(U_{r+1}, \dots, U_n)}$.

Ensuite, par le cours $\boxed{\mathrm{Im}(A^\top) = (\ker A)^\perp = \mathrm{Vect}(V_1, \dots, V_r)}$

et $\boxed{\mathrm{Im}(A) = (\ker(A^\top))^\perp = \mathrm{Vect}(U_1, \dots, U_r).}$

I.9) Le résultat du I.7. b) donne l'existence. La non-unicité vient du fait que si on a une valeur propre multiple, on a plusieurs choix possibles d'une base orthonormale $V_{i_1}, \dots, V_{i_k}$ de vecteurs propres associées, ce qui change la matrice V etc.

Partie II :

II.1. On déduit de I.7.b que

$$A_0^+ = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

On en déduit :

$$A_0 A_0^+ = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad A_0^+ A_0 = I_2$$

puis

$$A_0 A_0^+ A_0 = A_0 \quad A_0^+ A_0 A_0^+ = A_0^+$$

II.2.Notons $U_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, V_0 = \begin{pmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$ et $\Delta_0^+ = \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$, Comme U_0 et V_0 sont orthogonales et que Δ_0^+ est du type voulu, $A_0^+ = U_0 \Delta_0 V_0^\top$ est une décomposition de A_0^+en valeurs singulières, d'où l'on déduit que $(A_0^+)^+ = V_0 \left(\Delta_0^+\right)^+ U_0^\top = V_0 \Delta_0 U_0^\top = A_0$.

Donc $\boxed{(A_0^+)^+ = A_0}$

II.3. Soit $C = \Delta^+ \Delta$. Alors $C \in \mathcal{M}_p(\mathbb{R})$ et pour tout $(i, j) \in [\![1, p]\!]$,

$$c_{i,j} = \sum_{k=1}^{n} \Delta_{i,k}^+ \Delta_{k,j}$$

Comme pour tout $k \in [\![1, n]\!], \Delta_{i,k}^+ = 0$ si $k \neq i$ ou si $k = i$ et $i \geq r + 1$ et que $\Delta_{k,j} = 0$ si $k \neq j$ ou si $k = j$ et $j \geq r + 1$, on en déduit que $c_{i,j}$ est nul sauf si $i = j \leq r$ auquel cas $c_{i,i} = \frac{1}{\mu_i} \mu_i = 1$.

$$\Delta^+ \Delta = J_r(p) = \left(\begin{array}{c|c} \mathrm{I}_r & 0 \\ \hline 0 & 0 \end{array}\right) \in \mathcal{M}_p(\mathbb{R})$$

De même

$$\Delta\Delta^+ = J_r(n) = \left(\begin{array}{c|c} \mathrm{I}_r & 0 \\ \hline 0 & 0 \end{array}\right) \in \mathcal{M}_n(\mathbb{R})$$

II.4. Si $n = p = r$, ce qui précède prouve que $\Delta^+ = \Delta^{-1}$, comme de plus $V = \left({}^tV\right)^{-1}$ et ${}^tU = U^{-1}$, car U et V sont orthogonales,

$$A^+ = A^{-1}$$

II.5.Comme au I.8.a), $U = \sum_{i=1}^n U_i F_i^\top$ car $(F_1, \dots, F_n)$ forment la base canonique de $M_{n,1}(\mathbb{R})$.

Les calculs suivants sont analogues à ceux de I.8.b : $U^\top = \sum_{i=1}^n F_i U_i^\top$ donc

$$\Delta^+ U^\top = \sum_{i=1}^n (\Delta^+ F_i) U_i^\top = \sum_{i=1}^r \frac{1}{\mu_i} F_i U_i^\top$$

car $\Delta^+.F_i$ est simplement la i-ième colonne de Δ^+. Pour $A^+ = V\Delta^+ U^\top$, on en déduit que :

$$A^+ = \sum_{i=1}^r \frac{1}{\mu_i} (VF_i) U_i^\top,$$

et comme $VF_i = V_i$ (la i-ième colonne de V), on en déduit que :

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

Pour les deux autres égalités dont on ne demandait pas la démonstration : en utilisant l'égalité $A = \sum_{j=1}^r \mu_j U_j V_j^\top$ trouvée en I.8, et l'égalité qu'on vient d'obtenir, on a :

$$AA^+ = \sum_{1 \le i,j \le r} \frac{\mu_j}{\mu_i} U_j V_j^\top V_i U_i^\top.$$

Or pour tous i et j dans $[\![1, p]\!]$,

$$V_j^\top V_i = \langle V_j \mid V_i \rangle = \delta_{i,j}$$

car $(V_1, \dots, V_p)$ est une base orthonormée de $\mathbb{R}^p$. D'où

$$AA^+ = \sum_{i=1}^r U_i U_i^\top$$

De la même façon,en échangeant les rôles de U et V, comme $(U_1, \dots, U_n)$ est une base orthonormée de $\mathbb{R}^n$,

$$A^+ A = \sum_{i=1}^r V_i {}^t V_i$$

II.6.a. Pour tout $j \in [\![1, n]\!]$, d'après l'expression de $A.A^+$ donnée au II. 5) :

$$AA^+ U_j = \sum_{i=1}^r U_i \left(U_i^\top U_j\right) = \sum_{i=1}^r \langle U_i \mid U_j \rangle U_i = \sum_{i=1}^r \delta_{i,j} U_i.$$

Finalement, $AA^+ U_j = \begin{cases} U_j & \text{si } 1 \le j \le r \\ 0 & \text{si } j \ge r+1 \end{cases}$

Comme $(U_1, \dots, U_n)$ est une base orthonormée de $\mathbb{R}^n$, l'endomorphisme associé à AA^+dans la base canonique de $\mathbb{R}^n$ est la projection orthogonale de $\mathbb{R}^n$ sur $\mathrm{Vect}\,(U_1, \dots, U_r) = \mathrm{Im}\,A$.

N.B. pour la seconde affirmation : Pour tout $j \in [\![1, p]\!]$,
$A^+ AV_j = \sum_{i=1}^r V_i \left(V_i^\top V_j\right) = \sum_{i=1}^r \langle V_i \mid V_j \rangle V_i = \sum_{i=1}^r \delta_{i,j} V_i$.

Soit $A^+ AV_j = \begin{cases} V_j & \text{si } 1 \le j \le r \\ 0 & \text{si } j \ge r+1 \end{cases}$.

Comme $(V_1, \dots, V_p)$ est une base orthonormée de $\mathbb{R}^p$, l'endomorphisme associé à $A^+ A$ dans la base canonique de $\mathbb{R}^p$ est la projection orthogonale de $\mathbb{R}^p$ sur $\mathrm{Vect}\,(V_1, \dots, V_r) = (\mathrm{Ker}\,A)^\perp$.

II.7. Un projecteur orthogonal est diagonalisable dans une base orthonormée donc autoadjoint. Donc sa matrice dans toute b.o.n. est symétrique.

Donc AA^+ et A^+A sont des matrices symétriques par II. 6. ce qui montre les deux premières égalités demandées.

Pour la troisième : on a $\mathbb{R}^p = \operatorname{Ker} A \oplus (\operatorname{Ker} A)^\perp$.

$\forall X \in \operatorname{Ker} A, AA^+AX = 0 = AX$.

$\forall X \in (\operatorname{Ker} A)^\perp, (A^+A)\,X = X$ car A^+A est la matrice dans la base canonique de $\mathbb{R}^p$ de la projection orthogonale sur $(\operatorname{Ker} A)^\perp$ et donc $A\,(A^+A)\,X = AX$. On en déduit que

$$\forall\, X \in M_{p,1}(\mathbb{R}),\ AA^+AX = AX \qquad \text{donc} \quad AA^+A = A$$

Pour la dernière égalité, on va utiliser plutôt les formules du II. 5) qui font le travail « algébriquement ».

En effet par ces formules :

$$\begin{aligned} A^+A.A^+ = A^+.(A.A^+) = (\sum_{i=1}^{r} \frac{1}{\mu_i} V_i U_i^\top.)(\sum_{j=1}^{r} U_j U_j^\top) &= \sum_{(i,j)\in[\![1,r]\!]^2} \frac{1}{\mu_i} V_i U_i^\top.U_j U_j^\top \\ &= \sum_{i=1}^{r} \frac{1}{\mu_i} V_i.U_i^\top \quad \text{car } U_i^\top U_j = \delta_{i,j} \\ &= A^+ \end{aligned}$$

Bien sûr, on pouvait aussi prouver la troisième égalité avec ces formules. □

II.8. a) Montrons déjà que AB est une matrice de projecteur : $(AB)^2 = A(BAB) = AB$ par (2.4).

Donc AB est bien une matrice de projecteur, et par (2.1) elle est symétrique donc AB est un projecteur autoadjoint, donc un projecteur orthogonal.

Montrons maintenant que $\operatorname{Im} AB = \operatorname{Im} A$.

D'une manière générale, si $M \in \mathcal{M}_{q,m}(\mathbb{R})$ et si $N \in \mathcal{M}_{m,s}(\mathbb{R}), \operatorname{Im}(MN) \subset \operatorname{Im} M$

On a donc toujours ici l'inclusion $\operatorname{Im} AB \subset \operatorname{Im} A$.

Et avec (2.3), on a $\operatorname{Im} A = \operatorname{Im} ABA \subset \operatorname{Im} AB$.

D'où les deux inclusions, et la conclusion.

b) On montre de même que $(BA)^2 = BA$ et BA est symétrique donc BA est matrice d'un projecteur orthogonal.

Reste à montrer que $\operatorname{Im}(BA) = (\ker A)^\perp$. Comme BA est un projecteur orthogonal, $\operatorname{Im}(BA) = \ker(BA)^\perp$.

Il suffit donc de montrer que $\ker(BA) = \ker(A)$.

Comme au a), on a une inclusion évidente, $\ker(A) \subset \ker(BA)$.

Et comme $A = ABA$ on a aussi $\ker(BA) \subset \ker(ABA) = \ker(A)$. D'où la conclusion.

c) Comme $\operatorname{Im}(ba) = (\ker(a))^\perp$, on sait que pour tout $x \in \ker(a)^\perp$, $b(a(x)) = x$.

Pour tout $y \in \operatorname{Im}(a)$, en posant $x = (a')^{-1}(y) \in \ker(a)^\perp$, on a donc :

$$b(y) = b(a(x)) = x = (a')^{-1}(y)$$

ce qui montre bien que $b_{|\operatorname{Im}(a)} = a'$.

D'autre part, si on prend $x \in (\operatorname{Im}(a))^\perp$, comme ab est projecteur orthogonal sur $\operatorname{Im}(a)$, on a $(ab)(x) = 0$ donc $b(x) = b(ab(x)) = 0$.

D'où les deux conditions annoncées.

II.9 Le c) précédent montre l'unicité de la pseudo-inverse, si elle existe, et en donne une description géométrique. Le II. 7 montre que A^+ vérifie les conditions de la déf. donc donne l'existence.

II.10. Par l'unicité démontré ci-dessus, $(A^+)^+$est l'unique matrice $B \in \mathcal{M}_{n,p}(\mathbb{R})$ vérifiant

$$A^+B = (A^+B)^\top \quad BA^+ = (BA^+)^\top \quad A^+BA^+ = A^+ \quad BA^+B = B \tag{2}$$

Comme $A \in \mathcal{M}_{n,p}(\mathbb{R})$ et que A vérifie (2), on conclut bien que :

$$A = (A^+)^+$$

Remarque sur l'interprétation géométrique : je recopie ici l'explication très claire des exercices de Francinou-Gianella-Nicolas *L'application linéaire a est quelconque et a priori ni injective ni surjective. Une équation* $a(x) = y$ *d'inconnue* $x \in \mathrm{E}$ *n'a donc pas forcément de solution et si elle en a, il n'y a pas forcément unicité. Si* $y \in \mathrm{F}$ *est fixé, l'idée est chercher l'élément de* $\mathrm{Im}(a)$ *le plus proche de* y *à savoir le projeté orthogonal de* y *sur* $\mathrm{Im}\, a$*. Notons-le* z*. Parmi tous les antécédents de* z*, il y en a un unique dans l'orthogonal de Ker* a*. C'est cet antécédent qui est l'image de* y *par* b*. Ce choix a le mérite de préserver la symétrie : si* b *est le pseudo-inverse de* a*, alors* a *est le pseudo-inverse de* b*.*

Le tout en un dessin (même source)