Correction du DM14

(CCP PC 2001 maths 2)

Partie I

1. Si $A=0$ alors $A^TA=0$; réciproquement, si $A^TA=0$ alors $\|A\|^2=\mathrm{Tr}(A^TA)=0$ donc $A=0$. On a bien $\boxed{A=0\Leftrightarrow A^TA=0}$

2. A^TA et AA^T sont symétriques réelles donc le théorème spectral s'applique.

3. a) On a $\boxed{\langle X|Y\rangle_n=X^TY}$

b) On a $\|AW\|_n^2=W^TA^TAW=W^T(\lambda W)=\lambda\langle W|W\rangle_p$ donc $\boxed{\|AW\|_n^2=\lambda\|W\|_p^2}$

c) Les valeurs propres de A^TA sont réelles car A^TA est symétrique réelle et avec la question précédente, comme $W\neq 0$, on a $\|W\|_p>0$ donc $\lambda\geqslant 0$ et $\boxed{\mathrm{Sp}(A^TA)\subset\mathbb{R}^+}$

4. a) On a $\begin{pmatrix}xI_n & A\\ A^T & I_p\end{pmatrix}\begin{pmatrix}I_n & 0_{n,p}\\ -A^T & -I_p\end{pmatrix}=\begin{pmatrix}xI_n-AA^T & -A\\ 0 & -I_p\end{pmatrix}$ et $\begin{pmatrix}xI_n & A\\ A^T & I_p\end{pmatrix}\begin{pmatrix}I_n & -A\\ 0_{p,n} & xI_p\end{pmatrix}=\begin{pmatrix}xI_n & 0\\ A^T & xI_p-A^TA\end{pmatrix}$

b) En appliquant le déterminant aux deux égalités précédentes, on obtient $(-1)^p\det\begin{pmatrix}xI_n & A\\ A^T & I_p\end{pmatrix}=(-1)^p\mathcal{X}_{AA^T}(x)$ et $x^n\mathcal{X}_{A^TA}(x)=x^p\det\begin{pmatrix}xI_n & A\\ A^T & I_p\end{pmatrix}$. On en déduit $X^n\mathcal{X}_{A^TA}=X^p\mathcal{X}_{AA^T}$ donc les polynômes $\mathcal{X}_{A^TA}$ et $\mathcal{X}_{AA^T}$ ont les mêmes racines non nulles, avec les mêmes ordres de multiplicité. On en déduit $\boxed{\mathrm{Sp}(A^TA)\setminus\{0\}=\mathrm{Sp}(AA^T)\setminus\{0\}}$ et $\boxed{m_\lambda(A^TA)=m_\lambda(AA^T)\text{ si }\lambda\neq 0}$

c) On a aussi $n+m_0(A^TA)=p+m_0(AA^T)$ et comme A^TA et AA^T sont diagonalisables (th spectral), on en déduit $n+\dim(\ker A^TA)=p+\dim(\ker AA^T)$ et enfin (avec $A^TA\in\mathcal{M}_p(\mathbb{R})$ et $AA^T\in\mathcal{M}_n(\mathbb{R})$), avec le théorème du rang $\boxed{\mathrm{rg}(A^TA)=\mathrm{rg}(AA^T)}$

5. Si $n>p$ alors $\mathrm{rg}(AA^T)=\mathrm{rg}(A^TA)\leqslant p<n$ donc $\mathrm{rg}(AA^T)\neq n$ et $AA^T\in\mathcal{M}_n(\mathbb{R})$ donc $AA^T\notin\mathcal{GL}_n(\mathbb{R})$ et $\boxed{0\in\mathrm{Sp}(AA^T)}$ De même, si $n<p$, $\mathrm{rg}(A^TA)=\mathrm{rg}(AA^T)\leqslant n<p$ donc $A^TA\notin\mathcal{GL}_p(\mathbb{R})$ et $\boxed{0\in\mathrm{Sp}(A^TA)}$

6. a) Si $\lambda_1=0$ alors $\mathrm{Sp}(A^TA)=\{0\}$ donc, comme A^TA est diagonalisable (th spectral), A^TA est semblable à la matrice nulle donc $A^TA=0$ et $A=0$ ce qui est exclu. On a donc $\boxed{\lambda_1>0}$

b) On a $\mathrm{rg}(A^TA)=r$ (utiliser la matrice diagonale semblable à A^TA) donc $\mathrm{rg}(AA^T)=r$ et $AA^T\in\mathcal{M}_n(\mathbb{R})$ donc $\boxed{r\leqslant n\text{ et }\dim(\ker AA^T)=n-r}$ par le théorème du rang.

c) Par définition, on a $\boxed{AV_i=\mu_iU_i\text{ pour }i\leqslant r}$ Si $i>r$, on a $U_i\in\ker(A^TA)$ et $\|AV_i\|_n^2=V_i^TA^TAV_i=0$ donc $\boxed{AV_i=0\text{ pour }i>r}$ *(c'est une partie de la preuve de* $\ker(A)=\ker(A^TA)$*)*

d) Si $i\leqslant r$, on a $A^TU_i=\dfrac{1}{\mu_i}A^TAV_i=\dfrac{1}{\mu_i}\lambda_iV_i$ donc comme $\lambda_i=\mu_i^2$, on a $\boxed{A^TU_i=\mu_iV_i\text{ si }i\leqslant r}$

e) Si $i>r$ alors $U_i\in\ker(AA^T)$ donc $\|A^TU_i\|_p^2=U_i^TAA^TU_i=0$ donc $\boxed{A^TU_i=0\text{ si }i>r}$

f) Si $i\leqslant r$, on a $AA^TU_i=\mu_iAV_i=\mu_i^2U_i=\lambda_iU_i$ et si $i>r$, $AA^TU_i=0=\lambda_iU_i$. Comme U_i est non nul, $\boxed{U_i\text{ est un vecteur propre de }AA^T\text{ associé à }\lambda_i}$

Si $(i,j)\in[\![1,r]\!]^2$, on a $\langle U_i|U_j\rangle_n=\dfrac{1}{\mu_i\mu_j}V_i^TA^TAV_j=\dfrac{\lambda_j}{\mu_i\mu_j}V_i^TV_j=\dfrac{\mu_j}{\mu_i}\langle V_i|V_j\rangle=\dfrac{\mu_j}{\mu_i}\delta_{i,j}=\delta_{i,j}$

On sait déjà que $U_{r+1},\dots,U_n$ est une famille orthonormale ; il reste donc seulement à vérifier que pour $i\leqslant r$ et $j>r$, on a $U_i\perp U_j$: on a $\langle U_i|U_j\rangle=\dfrac{1}{\mu_i}V_i^TA^TU_j=0$ car $A^TU_j=0$. On peut donc conclure $\boxed{(U_1,\dots,U_n)\text{ est une base orthonormale de }\mathbb{R}^n}$

7. a) On a $(U^TAV)_{i,j}=F_i^TU^TAVE_j=U_i^TAV_j=\mu_jU_i^TU_j\mu_j\langle U_i|U_j\rangle_n=\boxed{\mu_j\delta_{i,j}}$ (l'égalité $AV_j=\mu_jU_j$ reste vraie pour $j>r$ car $\mu_j=0$ dans ce cas.

b) On vient de prouver $U^TAV=\Delta$ et comme $(U_i)_{1\leqslant i\leqslant n}$ et $(V_j)_{1\leqslant j\leqslant p}$ sont des bases orthonormales de $\mathbb{R}^n$ et $\mathbb{R}^p$, les matrices U et V sont orthogonales (matrices de passage de la base canonique, orthonormale, à une autre base orthonormale), on a $\boxed{A=U\Delta V^T}$

c) On suit patiemment les calculs de **I.6** et **I.7** : $A_0^TA_0 = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}$ donc $\lambda_1 = 6$, $\mu_1 = \sqrt{6}$ et $V_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ puis $\lambda_2 = 2$, $\mu_2 = \sqrt{2}$ et $V_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. On calcule alors $U_1 = \dfrac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$ et $U_2 = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Pour finir, $A_0A_0^T = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 4 \end{pmatrix}$ donc $\ker(A_0A_0^T) = \mathrm{Vect}\{(1,-1,1)\}$ donc $U_3 = \dfrac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. On en déduit

$$\boxed{A_0 = \underbrace{\frac{1}{\sqrt{6}}\begin{pmatrix} -1 & \sqrt{3} & \sqrt{2} \\ 1 & \sqrt{3} & -\sqrt{2} \\ 2 & 0 & \sqrt{2} \end{pmatrix}}_{U}\underbrace{\begin{pmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}}_{\Delta}\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{V^T}}$$

Idem pour B_0 et on trouve $\boxed{B_0 = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}(1)}$

8. Comme U et V sont inversibles, on a $\mathrm{rg}(A) = \mathrm{rg}(\Delta)$ donc $\boxed{\mathrm{rg}(A) = r}$

9. a) La matrice $V_iE_i^T$ est la matrice $p \times p$ dont toutes les colonnes sont nulles sauf celle d'indice i qui vaut V_i. Par construction de V, on a $\boxed{V = \sum_{i=1}^{p} V_iE_i^T}$

b) On a $A = \sum_{i=1}^{p} U\Delta E_iV_i^T$, de plus $\Delta E_i = \mu_iF_i$ ($= 0$ si $i > r$) donc $A = \sum_{i=1}^{r} \mu_iUF_iV_i^T$; enfin, $UF_i = U_i$ donc $\boxed{A = \sum_{i=1}^{r} \mu_iU_iV_i^T}$ Comme $A^TU_i = \mu_iV_i$ (même si $i > r$), on a $A^TA = \sum_{i=1}^{r} \mu_iA^TU_iV_i^T = \sum_{i=1}^{r} \mu_i^2V_iV_i^T$ donc $\boxed{A^TA = \sum_{i=1}^{r} \lambda_iV_iV_i^T}$ On a de même $AV_i = \mu_iU_i$ donc $AA^T = \sum_{i=1}^{r} \mu_iU_i(AV_i)^T$ donc $\boxed{AA^T = \sum_{i=1}^{r} \lambda_iU_iU_i^T}$

c) $\boxed{\ker(A) = \ker(A^TA) = \mathrm{Vect}\{V_{r+1},\dots,V_p\} \text{ et } \ker(A^T) = \ker(AA^T) = \mathrm{Vect}\{U_{r+1},\dots,U_n\}}$ D'après la question précédente, les colonnes de A sont des multiples des U_i ($i \leqslant r$) donc $\mathrm{Im}(A) \subset \mathrm{Vect}\{U_1,\dots,U_r\}$ et comme $\mathrm{rg}(A) = r$, on a $\boxed{\mathrm{Im}(A) = \mathrm{Vect}\{U_1,\dots,U_r\}}$ De la même façon, on a $\boxed{\mathrm{Im}(A^T) = \mathrm{Vect}\{V_1,\dots,V_r\}}$

Partie II

1. On a $\Delta_0^+ = \begin{pmatrix} 1/\sqrt{6} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 0 \end{pmatrix}$ donc $\boxed{A_0^+ = \dfrac{1}{6}\begin{pmatrix} 3 & 3 & 0 \\ -1 & 1 & 2 \end{pmatrix},\ A_0A_0^+ = \dfrac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix},\ A_0^+A_0 = I_2}$ puis on vérifie $\boxed{A_0A_0^+A_0 = A_0 \text{ et } A_0^+A_0A_0^+ = A_0^2}$

2. Inutile de reprendre tous les calculs de **I.7.c** avec A_0^+, $A_0^+ = V\Delta^+U^T$ est la décomposition en valeur singulières de A_0^+ (en admettant l'unicité qui va être prouvée après, comme le suggère le texte) donc $(\Delta_0^+)^+ = \Delta$ et $\boxed{(A_0^+)^+ = A_0}$

3. $\boxed{\Delta^+\Delta = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}) \text{ alors que } \Delta\Delta^+ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})}$ *(elles ne sont pas égales car elles n'ont pas la même taille)*

4. Si $A \in \mathcal{GL}_n(\mathbb{R})$ alors $n = p = r$ donc $\Delta^+ = \Delta^{-1}$ puis $A^+A = V\Delta^{-1}U^TU\Delta V^T = I_n$ donc $\boxed{A^+ = A^{-1} \text{ si } A \in \mathcal{GL}_n(\mathbb{R})}$

5. On a $U = \sum_{i=1}^{n} U_iF_i^T$ donc $\Delta^+U^T = \sum_{i=1}^{n} \Delta^+F_iU_i^T = \sum_{i=1}^{r} \dfrac{1}{\mu_i}F_iU_i^T$ et $A^+ = \sum_{i=1}^{r} \dfrac{1}{\mu_i}VF_iU_i^T = \boxed{\sum_{i=1}^{r} \dfrac{1}{\mu_i}V_iU_i^T}$

On vérifie ensuite les expressions de AA^+ et A^+A comme en **I.9.b**

6. a) $AA^+U_j = \sum_{i=1}^{r} U_i\underbrace{U_i^TU_j}_{=\delta_{ij}} = \begin{cases} U_j & \text{si } j \leqslant r \\ 0 & \text{sinon} \end{cases}$ donc $\boxed{AA^+ = \mathrm{Mat}_{\mathcal{B}_c}(\pi_{\mathrm{Im}(A)})}$ puisque $(U_1,\dots,U_r)$ est une base de $\mathrm{Im}\,A$ et $(U_{r+1},\dots,U_n)$ une base de $(\mathrm{Im}\,A)^\perp = \ker A^T$

b) On fait de même avec $A^+AV_j = \begin{cases} V_j & \text{si } i \leqslant r \\ 0 & \text{sinon} \end{cases}$

7. AA^+ et A^+A sont les matrices dans des bases orthonormales de projections orthogonales donc $\boxed{AA^+ \text{ et } A^+A \text{ sont symétriques}}$

On a $\operatorname{Im} AA^+ = \operatorname{Im} A$ et, comme AA^+ est une matrice de projecteur, $\operatorname{Im} AA^+ = \ker(AA^+ - I_n)$; on a donc $\operatorname{Im} A \subset \ker(AA^+ - I_n)$ donc $(AA^+ - I_n)A = 0$, ie $\boxed{AA^+A = A}$

On a $A^+(AA^+) = \displaystyle\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{1}{\mu_i}V_i\underbrace{U_i^TU_j}_{=\delta_{i,j}}U_j^T = \sum_{i=1}^{r}\frac{1}{\mu_i}V_iU_i^T$ donc $\boxed{A^+AA^+ = A^+}$ *(on pouvait difficilement faire comme pour AA^+A car on n'a pas encore déterminé $\ker A^+$ et $\operatorname{Im} A^+$)*

8. i) on a $\operatorname{Im} AA^+ \subset \operatorname{Im} A$ puis $\operatorname{Im} A = \operatorname{Im}(AA^+)A \subset \operatorname{Im} AA^+$ donc $\boxed{\operatorname{Im} A = \operatorname{Im} AA^+}$

$\ker A^+ \subset \ker AA^+$ et $\ker AA^+ \subset \ker A^+(AA^+) = \ker A^+$ donc $\boxed{\ker A^+ = \ker AA^+}$

$\operatorname{Im} A^+A \subset \operatorname{Im} A^+$ et $\operatorname{Im} A^+ = \operatorname{Im}(A^+A)A^+ \subset \operatorname{Im} A^+A$ donc $\boxed{\operatorname{Im} A^+ = \operatorname{Im} A^+A}$

$\ker A \subset \ker A^+A$ et $\ker A^+A \subset \ker A(A^+A) = \ker A$ donc $\boxed{\ker A = \ker A^+A}$

ii) AA^+ est un projecteur de $\mathbb{R}^n$ donc $\mathbb{R}^n = \ker AA^+ \oplus \operatorname{Im} AA^+$, ie $\boxed{\mathbb{R}^n = \ker A^+ \oplus \operatorname{Im} A}$ De même A^+A est un projecteur de $\mathbb{R}^p$ donc $\boxed{\mathbb{R}^p = \ker A \oplus \operatorname{Im} A^+}$

9. a) i) $B = BAB = B(AB)^T = BB^TA^T$ et $B = (BA)^TB = A^TB^TB$

ii) $A = ABA = A(BA)^T = AA^TB^T$ et $A = (AB)^TA = A^TA^TA$

iii) il suffit de transposer les relations de ii)

b) Comme A^+ vérifie (1), elle vérifie aussi i), ii) et iii) donc $A^T = A^TAA^+$ d'après iii) ; on reporte dans i) avec B : $B = B(B^TA^TA)A^+ \overset{B\ \text{ii})}{=} BAA^+ \overset{A^+\ \text{i})}{=} BAA^T(A^+)^TA^+ \overset{B\ \text{iii})}{=} A^T(A^+)^TA^+ \overset{A^+\ \text{i})}{=} A^+$

10. A vérifie (1) en prenant A^+ à la place de A donc, par unicité, $\boxed{A = (A^+)^+}$ De même $(A^+)^T$ vérifie (1) avec A^T à la place de A donc $\boxed{(A^+)^T = (A^T)^+}$

11. On vérifie $A_0B_0 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$ puis $(A_0B_0)^+ = \dfrac{1}{4}\begin{pmatrix}1 & 0 & -1\end{pmatrix}$ alors que $B_0^+ = \dfrac{1}{2}\begin{pmatrix}1 & -1\end{pmatrix}$ donc $B_0^+A_0^+ = \dfrac{1}{6}\begin{pmatrix}2 & 1 & -1\end{pmatrix}$ et $\boxed{(A_0B_0)^+ \neq B_0^+A_0^+}$ *(cette égalité serait vérifiée si A et B étaient carrées et inversibles puisque dans ce cas $(AB)^+ = (AB)^{-1} = B^{-1}A^{-1} = B^+A^+$)*

12. a) $\langle AX | H - AA^+H\rangle_n = X^TA^TH - X^T\underbrace{A^TAA^+}_{=A^T}H = 0$ et de même $\langle AA^+H | H - AA^+H\rangle_n = 0$ donc en ajoutant les deux relations (bilinéarité du produit scalaire) $\boxed{AX - AA^+H \perp H - AA^+H}$

L'idée du texte est de redémontrer que la distance de H à $\{AX, X \in \mathbb{R}^p\} = \operatorname{Im} A$ est atteinte au point $A\overline{H}$ avec le théorème de Pythagore (cf cours) mais on peut le voir directement car $A\overline{H} = AA^+H = \pi_{\operatorname{Im}(A)}(H)$ d'après **II.6.a**. On a donc $\boxed{d(H, \operatorname{Im}(A)) = \left\|A\overline{H} - H\right\|_n}$

b) $d(H, \operatorname{Im} A)$ étant atteinte uniquement au point $A\overline{H}$, si $\left\|A\tilde{H} - N\right\|_n = \left\|A\overline{H} - N\right\|_n$ alors $A\tilde{H} = A\overline{H}$ donc $\tilde{H} - \overline{H} \in \ker(A)$; d'autre part, $\overline{H} = A^+H \in \operatorname{Im} A^+ = (\ker A)^\perp$ et $\tilde{H} = \left(\tilde{H} - \overline{H}\right) + \overline{H}$ donc Pythagore donne $\left\|\tilde{H}\right\|_p^2 = \left\|\tilde{H} - \overline{H}\right\|_p^2 + \left\|\overline{H}\right\|_p^2 > \left\|\overline{H}\right\|_p^2$.

c) $\boxed{d(H, \operatorname{Im} A_0) = \left\|A_0\overline{H} - H\right\|_3 = \dfrac{1}{\sqrt{3}}}$