Méthodes de calcul et raisonnement — corrigé
Durée : 2 heures

Problème 1 :

I. Résultats préliminaires

1. (a) Soit $j \in \llbracket 0; p \rrbracket$. Alors d'une part $\binom{p}{j} \frac{1}{1+j} = \frac{p!}{j!(p-j)!} \frac{1}{1+j} = \frac{p!}{(j+1)!(p-j)!}$.

Et d'autre part $\binom{p+1}{j+1} \frac{1}{p+1} = \frac{(p+1)!}{(j+1)!(p-j)!} = \frac{p!}{(j+1)!(p-j)!}$.

Donc $\boxed{\binom{p}{j} \frac{1}{1+j} = \binom{p+1}{j+1} \frac{1}{p+1}}$.

(b) Soit $j \in \llbracket 0; p-1 \rrbracket$. Alors

$$\begin{aligned} \binom{p}{j+1} + \binom{p}{j} &= \frac{p!}{(j+1)!(p-j-1)!} + \frac{p!}{j!(p-j)!} \\ &= \frac{p!(p-j)}{(j+1)!(p-j)!} + \frac{p!(j+1)}{(j+1)!(p-j)!} \\ &= \frac{p!(p+1)}{(j+1)!(p-j)!} \\ &= \frac{(p+1)!}{(j+1)!(p-j)!} \\ &= \boxed{\binom{p+1}{j+1}}. \end{aligned}$$

2. Une densité de X : $\boxed{x \mapsto \lambda e^{-\lambda x} 1_{\mathbb{R}^+}(x)}$.

Fonction de répartition : $\boxed{x \mapsto (1 - e^{-\lambda x}) 1_{\mathbb{R}^+}(x)}$.

$\boxed{\mathbf{E}(X) = \frac{1}{\lambda} \text{ et } \mathbf{V}(X) = \frac{1}{\lambda^2}}$.

3. (a) Soit un entier $i \geq 2$. La fonction $t \mapsto \frac{1}{t}$ est décroissante sur $\mathbb{R}_+$, donc : $\forall t \in [i-1, i], \frac{1}{i} \leq \frac{1}{t}$.

Comme $t \mapsto \frac{1}{i}$ et $t \mapsto \frac{1}{t}$ sont continues sur $[i-1, i]$, on peut intégrer ces fonctions. On obtient

par croissance de l'intégrale $\int_{i-1}^i \frac{dt}{i} \leq \int_{i-1}^i \frac{dt}{t}$, c'est-à-dire

$$\boxed{\frac{1}{i} \leq \int_{i-1}^i \frac{dt}{t}}$$

De même, on a : $\forall t \in [i, i+1], \frac{1}{t} \leq \frac{1}{i}$, ce qui donne

$$\boxed{\int_i^{i+1} \frac{dt}{t} \leq \frac{1}{i}}$$

(b) Soit $n \geq 1$. Alors $\ln(n+1) = \ln(n) + \ln\left(1 + \frac{1}{n}\right)$, donc

$$\frac{\ln(n+1)}{\ln(n)} = 1 + \frac{\ln\left(1 + \frac{1}{n}\right)}{\ln(n)} \xrightarrow{n \rightarrow +\infty} 1$$

par opérations sur les limites. Donc $\boxed{\ln(n+1) \underset{n \rightarrow +\infty}{\sim} \ln(n)}$.

- (c) L'encadrement de la question 3.a étant établi pour tout entier $i \geq 2$, on le somme pour i de 2 à n , ce qui donne par relation de Chasles

$$\int_2^{n+1} \frac{dt}{t} \leq \sum_{i=2}^n \frac{1}{i} \leq \int_1^n \frac{dt}{t}.$$

En ajoutant 1 et en divisant par $\ln(n)$ (strictement positif dès que $n \geq 2$) on obtient donc

$$\frac{\ln(n+1)}{\ln(n)} - \frac{\ln(2)}{\ln(n)} + \frac{1}{\ln(n)} \leq \frac{1}{\ln(n)} \sum_{k=1}^n \frac{1}{k} \leq 1 + \frac{1}{\ln(n)}.$$

D'après la question précédente, $\frac{\ln(n+1)}{\ln(n)} \xrightarrow{n \rightarrow +\infty} 1$, donc les membres de gauche et de droite de l'encadrement précédent tendent vers 1. Par théorème des gendarmes, il s'ensuit

$$\boxed{\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln(n)}.$$

II. Quelques résultats autour de la loi exponentielle

4. Soit un réel x . Alors $[X_{(1)} > x] = \bigcap_{i=1}^n [X_i > x]$, donc l'indépendance de $X_1, \dots, X_n$ donne

$$\mathbf{P}(X_{(1)} > x) = \prod_{i=1}^n \mathbf{P}(X_i > x).$$

Cependant, toutes les variables X_i suivent la loi $\mathcal{E}(\lambda)$, par conséquent

$$\mathbf{P}(X_{(1)} > x) = \begin{cases} (e^{-\lambda x})^n = e^{-n\lambda x} & \text{si } x \geq 0 \\ 1 & \text{si } x < 0 \end{cases}.$$

Ainsi la fonction de répartition de $X_{(1)}$ est donnée par $F_{X_{(1)}}(x) = (1 - e^{-n\lambda x})1_{\mathbb{R}^+}(x)$. On reconnaît la fonction de répartition de la loi exponentielle $\mathcal{E}(n\lambda)$. Or la fonction de répartition caractérise la loi, donc $\boxed{X_{(1)} \sim \mathcal{E}(n\lambda)}$.

5. Pour tout réel x , on a $[X_{(n)} \leq x] = \bigcap_{i=1}^n [X_i \leq x]$ ce qui donne par indépendance des X_i la fonction de répartition de $X_{(n)}$:

$$F_{X_{(n)}}(x) = \prod_{i=1}^n F_{X_i}(x) = F_{X_1}(x)^n = (1 - e^{-\lambda x})^n 1_{\mathbb{R}^+}(x).$$

On constate :

- par composition de fonctions $\mathcal{C}^1$, $F_{X_{(n)}}$ est de classe $\mathcal{C}^1$ sauf peut-être en 0 ;
- $\lim_{x \rightarrow 0} (1 - e^{-\lambda x})^n = 0$ donc $F_{X_{(n)}}$ est continue en 0.

Donc $\boxed{X_{(n)}$ est une variable à densité.

En outre, $\frac{d}{dx} [(1 - e^{-\lambda x})^n] = n\lambda e^{-\lambda x} (1 - e^{-\lambda x})^{n-1}$, donc une densité de $X_{(n)}$ est bien donnée par

$$\boxed{x \mapsto n\lambda e^{-\lambda x} (1 - e^{-\lambda x})^{n-1} 1_{\mathbb{R}^+}(x)}.$$

6. Préalablement, observons que pour tout réel $\mu > 0$, l'intégrale $\int_0^{+\infty} x e^{-\mu x} dx$ est convergente et vaut $\frac{1}{\mu^2}$. Pour le prouver, il suffit d'observer que $\int_0^{+\infty} \mu x e^{-\mu x} dx$ est l'espérance d'une variable aléatoire de loi $\mathcal{E}(\mu)$, donc est convergente et égale à $\frac{1}{\mu}$.

Ensuite, l'espérance de $X_{(n)}$ est donnée par l'intégrale suivante

$$I = \int_0^{+\infty} n\lambda x e^{-\lambda x} (1 - e^{-\lambda x})^{n-1} dx$$

sous réserve de convergence. Or par binôme de Newton :

$$\begin{aligned} e^{-\lambda x} (1 - e^{-\lambda x})^{n-1} &= e^{-\lambda x} \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j (e^{-\lambda x})^j \\ &= \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j e^{-\lambda(1+j)x} \end{aligned}$$

Par linéarité de l'intégrale, l'intégrale I est donc convergente, l'espérance de $X_{(n)}$ existe et est égale à

$$\begin{aligned} \mathbf{E}(X_{(n)}) &= \sum_{j=0}^{n-1} n\lambda \binom{n-1}{j} (-1)^j \int_0^{+\infty} x e^{-\lambda(1+j)x} dx \\ &= \sum_{j=0}^{n-1} n\lambda \binom{n-1}{j} (-1)^j \frac{1}{(1+j)^2 \lambda^2} \\ &= \frac{1}{\lambda} \sum_{j=0}^{n-1} \binom{n-1}{j+1} \frac{(-1)^j}{1+j} \end{aligned}$$

en utilisant $\binom{n-1}{j} \frac{1}{j+1} = \binom{n}{j+1} \frac{1}{n}$ (d'après la question 1). On a donc bien

$$\mathbf{E}(X_{(n)}) = \frac{1}{\lambda} \sum_{j=0}^{n-1} \binom{n}{j+1} \frac{(-1)^j}{j+1}.$$

7. (a) Par binôme de Newton, $\sum_{k=0}^{p+1} \binom{p+1}{k} (-1)^k = (1-1)^{p+1} = 0.$

(b) On établit le résultat par récurrence sur p .

— Initialisation : pour $p = 1$, le résultat à prouver est

$$\binom{1}{1} \frac{1}{1} = \frac{1}{1}$$

ce qui est vrai.

— Hérité : supposons le résultat acquis pour un rang $p \in \mathbb{N}^*$ donné. Alors

$$\begin{aligned} \sum_{j=0}^p \binom{p+1}{j+1} \frac{(-1)^j}{j+1} &= \sum_{j=0}^{p-1} \binom{p+1}{j+1} \frac{(-1)^j}{j+1} + \binom{p+1}{p+1} \frac{(-1)^p}{p+1} \\ &= \sum_{j=0}^{p-1} \left[\binom{p}{j+1} + \binom{p}{j} \right] \frac{(-1)^j}{j+1} + \frac{(-1)^p}{p+1} && \text{par la question 2} \\ &= \sum_{j=0}^{p-1} \binom{p}{j+1} \frac{(-1)^j}{j+1} + \sum_{j=0}^{p-1} \binom{p}{j} \frac{(-1)^j}{j+1} + \binom{p}{p} \frac{(-1)^p}{p+1} \\ &= \underbrace{\sum_{j=0}^{p-1} \binom{p}{j+1} \frac{(-1)^j}{j+1}}_A + \underbrace{\sum_{j=0}^p \binom{p}{j} \frac{(-1)^j}{j+1}}_B \end{aligned}$$

D'une part, $A = \sum_{k=1}^p \frac{1}{k}$ par hypothèse de récurrence.

D'autre part, en utilisant la première question :

$$\begin{aligned}
 B &= \frac{1}{p+1} \sum_{j=0}^p \binom{p+1}{j+1} (-1)^j \\
 &= -\frac{1}{p+1} \sum_{k=1}^{p+1} \binom{p+1}{k} (-1)^k \\
 &= -\frac{1}{p+1} \left(\sum_{k=0}^{p+1} \binom{p+1}{k} (-1)^k - \binom{p+1}{0} (-1)^0 \right) \\
 &= -\frac{1}{p+1} (0 - 1) && \text{par la question précédente} \\
 &= \frac{1}{p+1}
 \end{aligned}$$

Finalement, on trouve bien $\sum_{j=0}^p \binom{p+1}{j+1} \frac{(-1)^j}{j+1} = \sum_{k=1}^p \frac{1}{k} + \frac{1}{p+1} = \sum_{k=1}^{p+1} \frac{1}{k}$, ce qui conclut la preuve de l'hérédité.

— Conclusion : $\forall p \in \mathbb{N}^*, \sum_{j=0}^{p-1} \binom{p}{j+1} \frac{(-1)^j}{j+1} = \sum_{k=1}^p \frac{1}{k}$.

8. D'après les question 6 et 7, $\mathbf{E}(X_{(n)}) = \frac{1}{\lambda} \sum_{k=1}^n \frac{1}{k}$.

Mais d'après la question 3.c, $\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln(n)$.

Donc $\mathbf{E}(X_{(n)}) \underset{n \rightarrow +\infty}{\sim} \frac{\ln(n)}{\lambda}$.

9. Soit $n \in \mathbb{N}^*$. Alors

$$\begin{aligned}
 u_{n+1} - u_n &= \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) - \sum_{k=1}^n \frac{1}{k} + \ln(n) \\
 &= \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) \\
 &= \frac{1}{n} \cdot \frac{1}{1 + \frac{1}{n}} - \ln\left(1 + \frac{1}{n}\right) \\
 &= g\left(\frac{1}{n}\right)
 \end{aligned}$$

pour $g(x) = \frac{x}{1+x} - \ln(1+x)$. Calculons un développement limité en 0 de g .

$$\begin{aligned}
 g(x) &= x(1-x+o(x)) - \left(x - \frac{x^2}{2} + o(x^2)\right) \\
 &= -\frac{x^2}{2} + o(x^2)
 \end{aligned}$$

Donc $g(x) \underset{x \rightarrow 0}{\sim} -\frac{x^2}{2}$. Comme $\frac{1}{n} \underset{n \rightarrow +\infty}{\rightarrow} 0$, on a donc $u_{n+1} - u_n \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}$.

Par conséquent $u_n - u_{n+1}$ est positif à partir d'un certain rang, et le théorème de comparaison pour les séries à termes positifs indique que les séries $\sum (u_n - u_{n+1})$ et $\sum \frac{1}{n^2}$ sont de même natures. Cette dernière série est une série de référence convergente, donc $\sum (u_n - u_{n+1})$ est convergente.

Par linéarité, $\sum (u_{n+1} - u_n)$ est convergente.

10. La série $\sum (u_{n+1} - u_n)$ est télescopique, et $\sum_{k=1}^{n-1} (u_{k+1} - u_k) = u_n - u_1$. La convergence de la série $\sum (u_{n+1} - u_n)$ est donc équivalente à la convergence de la suite u . Donc u est convergente.

III. Loi de Gumbel

11. Premièrement, la fonction $\exp$ étant positive et continue, f est également positive et continue.

Reste à prouver que $\int_{-\infty}^{+\infty} f$ est convergente et égale à 1.

Vérifions les hypothèses pour effectuer le changement de variable $y = \exp(-x)$.

— $x \mapsto \exp(-x)$ est de classe $\mathcal{C}^1$ et strictement monotone, et on a $dy = -\exp(-x) dx$;

— $\lim_{x \rightarrow -\infty} \exp(-x) = +\infty$ et $\lim_{x \rightarrow +\infty} \exp(-x) = 0$.

Donc $\int_{-\infty}^{+\infty} f$ est de même nature, et en cas de convergence de même valeur, que

$$\int_{+\infty}^0 (-\lambda) \exp(-\lambda y) dy = \int_0^{+\infty} \lambda \exp(-\lambda y) dy.$$

On reconnaît l'intégrale de la densité de la loi $\mathcal{E}(\lambda)$, qui est donc convergente et égale à 1.

Conclusion : f est densité de probabilité.

12. Calculons la fonction de répartition F_Y de Y . Pour tout réel x , on a

$$\begin{aligned} F_Y(x) &= \mathbf{P}(Y \leq x) \\ &= \mathbf{P}(-\ln X_1 \leq x) \\ &= \mathbf{P}(\ln X_1 \geq -x) \\ &= \mathbf{P}(X_1 \geq e^{-x}) && \text{car } \exp \text{ croît strictement ;} \\ &= e^{-\lambda e^{-x}} && \text{car } X_1 \sim \mathcal{E}(\lambda). \end{aligned}$$

Il apparaît que F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc Y admet une densité égale à F'_Y c'est-à-dire

$$x \mapsto \lambda \exp(-x) \exp[-\lambda \exp(-x)]$$

Donc $Y = -\ln X_1$ suit la loi de Gumbel.

13. Soit $x \in \mathbb{R}$.

Pour tout n assez grand, $1 + \frac{x}{n} > 0$, et

$$\ln \left[\left(1 + \frac{x}{n} \right)^n \right] = n \ln \left(1 + \frac{x}{n} \right) \underset{n \rightarrow +\infty}{\sim} n \frac{x}{n} = x$$

donc $\ln \left[\left(1 + \frac{x}{n} \right)^n \right] \xrightarrow{n \rightarrow +\infty} x$. Par composition de limites $\lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n} \right)^n = \exp(x)$.

14. On a calculé à la question 5 la fonction de répartition de $X_{(n)}$, qui était (pour $\lambda = 1$)

$$x \mapsto (1 - e^{-x})^n 1_{\mathbb{R}^+}(x)$$

Mais pour tout réel x , on a $\mathbf{P}(X_{(n)} - \ln(n) \leq x) = \mathbf{P}(X_{(n)} \leq x + \ln(n))$. Donc la fonction de répartition de $X_{(n)} - \ln(n)$, que l'on notera H_n , est égale à

$$H_n : x \mapsto (1 - e^{-(x+\ln n)})^n 1_{[-\ln(n), +\infty[}(x) = \left(1 - \frac{e^{-x}}{n} \right)^n 1_{[-\ln(n), +\infty[}(x).$$

Or, pour un réel x donné :

— d'une part $\left(1 - \frac{e^{-x}}{n} \right)^n \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}}$ par la question précédente,

— d'autre part $1_{[-\ln(n), +\infty[}(x) = 1$ pour tout n assez grand, car $-\ln(n) \xrightarrow{n \rightarrow +\infty} -\infty$.

Donc $H_n(x) \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}}$. Mais $x \mapsto e^{-e^{-x}}$ est (dans le cas où $\lambda = 1$) la fonction de répartition de la loi de Gumbel, qu'on avait trouvée à la question 12. C'est donc la fonction de répartition de Y . Par conséquent $(X_{(n)} - \ln(n))_{n \in \mathbb{N}^*}$ converge en loi vers Y .

Problème 2 :

15. $(B^T, A^T) \in \mathcal{M}_{q,p}(\mathbb{R}) \times \mathcal{M}_{p,n}(\mathbb{R})$ donc les matrices $(AB)^T$ et $B^T A^T$ sont toutes les deux de taille (q, n) . De plus, pour tous $i \in \llbracket 1; q \rrbracket$ et $j \in \llbracket 1; n \rrbracket$, on a :

$$\begin{aligned} [B^T A^T]_{i,j} &= \sum_{k=1}^p [B^T]_{i,k} [A^T]_{k,j} \\ &= \sum_{k=1}^p A_{j,k} B_{k,i} \\ &= [AB]_{j,i} \\ &= [(AB)^T]_{i,j} \end{aligned}$$

donc $\boxed{B^T A^T = (AB)^T}$.

16. Si M est inversible, alors la matrice M^{-1} vérifie $MM^{-1} = M^{-1}M = I_n$. En appliquant la transposée et la question précédente, il vient $(M^{-1})^T M^T = M^T (M^{-1})^T = I_n^T = I_n$. Donc

$\boxed{M^T \text{ est inversible, d'inverse } (M^{-1})^T}$.

17. On va prouver le résultat par double inclusion.

— Soit $X \in \text{Ker}(A)$, c'est-à-dire $X \in \mathcal{M}_{p,1}(\mathbb{R})$ et $AX = 0$.

En multipliant à gauche par A^T , il vient $A^T AX = 0$, donc $X \in \text{Ker}(A^T)$.

Donc $\text{Ker}(A) \subset \text{Ker}(A^T)$.

— Soit $X \in \text{Ker}(A^T A)$, c'est-à-dire $X \in \mathcal{M}_{p,1}(\mathbb{R})$ et $A^T AX = 0$.

En multipliant à gauche par X^T , il vient $0 = X^T A^T AX = (AX)^T AX$.

Mais AX étant une matrice colonne, $(AX)^T AX = \|AX\|^2$.

Donc $\|AX\| = 0$, ce qui prouve $AX = 0$ donc $X \in \text{Ker}(A)$.

Donc $\text{Ker}(A^T A) \subset \text{Ker}(A)$.

Conclusion : $\boxed{\text{Ker}(A) = \text{Ker}(A^T A)}$.

18. $(A^T A)^T = A^T (A^T)^T = A^T A$. Donc A est symétrique, et elle est réelle, donc par théorème spectral

$\boxed{A^T A \text{ est diagonalisable.}}$

19. $A^T A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$. Soit $\lambda \in \mathbb{C}$. Alors $\lambda \in \text{Sp}(A^T A)$ si et seulement si $\text{rg}(A^T A - \lambda I_3) < 3$.

$$\begin{aligned} A^T A - \lambda I_3 &= \begin{pmatrix} 2-\lambda & -1 & -1 \\ -1 & 2-\lambda & -1 \\ -1 & -1 & 2-\lambda \end{pmatrix} \\ &\underset{L}{\sim} \begin{pmatrix} -1 & -1 & 2-\lambda \\ -1 & 2-\lambda & -1 \\ 2-\lambda & -1 & -1 \end{pmatrix} && L_1 \leftrightarrow L_3 \\ &\underset{L}{\sim} \begin{pmatrix} -1 & -1 & 2-\lambda \\ 0 & 3-\lambda & -3+\lambda \\ 0 & \lambda-3 & 3-4\lambda+\lambda^2 \end{pmatrix} && \begin{array}{l} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 + (2-\lambda)L_1 \end{array} \\ &\underset{L}{\sim} \begin{pmatrix} -1 & -1 & 2-\lambda \\ 0 & 3-\lambda & -3+\lambda \\ 0 & 0 & -3\lambda+\lambda^2 \end{pmatrix} && L_3 \leftarrow L_3 + L_2 \end{aligned}$$

donc $\lambda \in \text{Sp}(A)$ si et seulement si $\lambda = 3$ ou $\lambda^2 - 3\lambda = 0$, ce qui donne donc $\boxed{\text{Sp}(A) = \{0, 3\}}$.

Les opérations sur les lignes conservent le noyau, donc $E_\lambda(A) = \text{Ker} \begin{pmatrix} -1 & -1 & 2-\lambda \\ 0 & 3-\lambda & -3+\lambda \\ 0 & 0 & -3\lambda+\lambda^2 \end{pmatrix}$.

Pour $\lambda = 3$, on a donc :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_3(A) \iff x + y + z = 0$$

donc $E_3(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right).$

Et pour $\lambda = 0$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_3(A) \iff \begin{cases} -x - y + 2z = 0 \\ y - z = 0 \end{cases}$$

$$\iff x = y = z$$

Donc $E_0(A) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$

Par conséquent $A = PDP^{-1}$ pour les matrices suivantes :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

20. La famille $(C_1, \dots, C_p)$ est libre si et seulement si le seul p -uplet $(\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p$ qui vérifie $\sum_{i=1}^p \alpha_i C_i = 0$ est le p -uplet $(0, \dots, 0)$.

Cependant, pour tout vecteur colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{R})$, on a $AX = \sum_{i=1}^p x_i C_i$.

Par conséquent, la liberté de $(C_1, \dots, C_p)$ signifie aussi que le seul vecteur $X \in \mathcal{M}_{p,1}(\mathbb{R})$ tel que $AX = 0$ est le vecteur nul. Autrement dit, $\text{Ker}(A) = \{0\}$. Mais alors, d'après la question 17, $\text{Ker}(A^T A) = \{0\}$.

Or $A^T A$ est une matrice carrée, donc $\text{Ker}(A^T A) = \{0\}$ implique que $A^T A$ est inversible.

21. (a) D'abord, $H^2 = A \underbrace{(A^T A)^{-1} A^T A}_{I_p} (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T = H$, donc $H^2 = H$.

$$\begin{aligned} \text{Ensuite, } H^T &= [A(A^T A)^{-1} A^T]^T \\ &= (A^T)^T [(A^T A)^{-1}]^T A^T \\ &= A[(A^T A)^T]^{-1} A^T \\ &= A[A^T (A^T)^T]^{-1} A^T \\ &= A(A^T A)^{-1} A^T \\ &= H \end{aligned}$$

donc $H^T = H$.

- (b) Soit $X \in \text{Ker}(H)$ et $Y \in \text{Im}(H)$.

Alors par définition $HX = 0$ et il existe $U \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $Y = HU$. Mais alors :

$$\begin{aligned} \langle Y, X \rangle &= Y^T X \\ &= (HU)^T X \\ &= U^T HX \\ &= 0 \quad \text{car } HX = 0 \end{aligned}$$

Ainsi $\forall (X, Y) \in \text{Ker}(H) \times \text{Im}(H)$, $X \perp Y$. Par conséquent $\text{Ker}(H) \subset (\text{Im}(H))^\perp$.

Cependant, on sait par le cours que $\dim((\text{Im}(H))^\perp) = n - \dim(\text{Im}(H))$; et par théorème du rang on sait aussi $\dim(\text{Ker}(H)) = n - \dim(\text{Im}(H))$. Ainsi

$$\begin{cases} \text{Ker}(H) \subset (\text{Im}(H))^\perp \\ \dim(\text{Ker}(H)) = \dim((\text{Im}(H))^\perp) < +\infty \end{cases}$$

ce qui suffit pour conclure $\boxed{\text{Ker}(H) = (\text{Im}(H))^\perp}$.

(c) Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Prouvons l'équivalence

$$X \in \text{Ker}(H) \iff X \in \text{Ker}(A^\top)$$

par double implication.

— Supposons $X \in \text{Ker}(H)$. Alors $HX = 0$, c'est-à-dire $A(A^\top A)^{-1}A^\top X = 0$. En multipliant à gauche par $A^\top$, il vient $\underbrace{A^\top A(A^\top A)^{-1}}_{I_p} A^\top X = 0$ donc $A^\top X = 0$, ce qui prouve $X \in \text{Ker}(A^\top)$.

— Réciproquement supposons $X \in \text{Ker}(A^\top)$, c'est-à-dire $A^\top X = 0$. En multipliant à gauche par $A(A^\top A)^{-1}$, on trouve $HX = 0$, donc $X \in \text{Ker}(H)$.

Donc $\boxed{\text{Ker}(H) = \text{Ker}(A^\top)}$. Comme H et $A^\top$ ont n colonnes, il s'ensuit par théorème du rang :

$$\text{rg}(H) = n - \dim(\text{Ker } H) = n - \dim(\text{Ker } A^\top) = \text{rg}(A^\top) = \text{rg}(A).$$

Par conséquent, $\text{Im}(H)$ et $\text{Im}(A)$ sont de même dimension.

Prouvons $\text{Im}(H) \subset \text{Im}(A)$. Pour tout vecteur $Y \in \text{Im}(H)$, il existe $X \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $Y = HX$, c'est-à-dire $Y = A(A^\top A)^{-1}A^\top X$. Mais alors $Y = AU$ pour le vecteur colonne $U = (A^\top A)^{-1}A^\top X$, donc $Y \in \text{Im}(A)$. Cela prouve $\text{Im}(H) \subset \text{Im}(A)$, et l'égalité des dimensions permet alors de conclure $\boxed{\text{Im}(H) = \text{Im}(A)}$.

22. Soit $\hat{Y} = A\hat{X} = A(A^\top A)^{-1}A^\top B$. D'après l'énoncé, c'est le projeté orthogonal de B sur $\text{Im}(A)$. Donc c'est le point de $\text{Im}(A)$ le plus proche de B . Plus précisément :

$$\forall Y \in \text{Im}(A), \begin{cases} \|B - Y\| \geq \|B - \hat{Y}\| \\ \|B - Y\| = \|B - \hat{Y}\| \iff Y = \hat{Y}. \end{cases}$$

Comme $\hat{Y} = A\hat{X}$, le résultat précédent se réécrit ainsi :

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \begin{cases} \|B - AX\| \geq \|B - A\hat{X}\| \\ \|B - AX\| = \|B - A\hat{X}\| \iff AX = A\hat{X} \end{cases}$$

et par conséquent

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \begin{cases} g(X) \geq g(\hat{X}) \\ g(X) = g(\hat{X}) \iff X - \hat{X} \in \text{Ker}(A). \end{cases}$$

Mais on a vu à la question 20 que $\text{Ker}(A) = \{0\}$. On trouve donc

$$\boxed{\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \begin{cases} g(X) \geq g(\hat{X}) \\ g(X) = g(\hat{X}) \iff X = \hat{X} \end{cases}}$$

ce qui est exactement la propriété demandée.

23. Quand $n = p$, alors A est carrée donc le résultat $\text{Ker}(A) = \{0\}$ montre que A est inversible, et par conséquent $A^\top$ aussi. Mais alors $(A^\top A)^{-1} = A^{-1}(A^\top)^{-1}$, donc $\boxed{\hat{X} = A^{-1}B}$.

De fait, A étant inversible, AX peut prendre toutes les valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$, donc $g(X)$ est minimal quand AX est égal à B , c'est-à-dire quand $X = A^{-1}B$. Le résultat est cohérent.

24. On trouve $A^\top A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}$ et $(A^\top A)^{-1} = \frac{1}{6} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$, donc $\boxed{\hat{X} = \frac{2}{3} \begin{pmatrix} 1 \\ 2 \end{pmatrix}}$.