DS9

Correction du devoir surveillé

Problème 1 (EM Lyon 2022)

Partie A : Mise en place d'un exemple

1. (a) La matrice A n'est pas inversible puisque sa deuxième ligne est nulle. Plus précisément, en notant L_i la i-ème ligne de A, on a :

$$\mathrm{rg}(A) = \dim(\mathrm{Vect}(L_1, L_2, L_3)) = \dim(\mathrm{Vect}(L_1, L_3)) \boxed{= 2}$$

car les lignes L_1 et L_3 forment une famille libre (puisqu'elles sont non colinéaires).

(b) On obtient $\boxed{A^2 = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix} \text{ et } A^3 = \begin{pmatrix} -1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \end{pmatrix}.}$ On vérifie alors que :

$$\boxed{A^3 - A^2 + A = 0.}$$

(c) Le polynôme $P(x) = x^3 - x^2 + x = x(x^2 - x + 1)$ est annulateur de A. Le discriminant de $x^2 - x + 1$ étant négatif, P admet une seule racine réelle 0. Ainsi 0 est l'unique valeur propre réelle possible pour A. Et puisque A n'est pas inversible, 0 est bien valeur propre de A. On obtient donc $\boxed{\mathrm{Sp}_{\mathbb{R}}(A) = \{0\}.}$

Supposons A est diagonalisable dans $\mathscr{M}_3(\mathbb{R})$. Puisque $\mathrm{Sp}_{\mathbb{R}}(A) = \{0\}$, A serait semblable à la matrice nulle 0_n, et donc égale à 0_n, ce qui n'est pas. Ainsi, $\boxed{A \text{ n'est pas diagonalisable dans } \mathscr{M}_3(\mathbb{R}).}$

2. (a) Calculons :

$$^{\mathrm{t}}B =^{\mathrm{t}} \left(^{\mathrm{t}}AA\right) =^{\mathrm{t}} A^{\mathrm{t}}(^{\mathrm{t}}A) =^{\mathrm{t}} AA.$$

B étant une matrice symétrique réelle, $\boxed{\text{elle est diagonalisable.}}$

(b) i. On calcule :

$$R^{\mathrm{t}}R = \frac{1}{6}\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = I_3.$$

Ainsi R est inversible et $R^{-1} =^{\mathrm{t}} R$. $\boxed{\text{La matrice } R \text{ est bien orthogonale.}}$

ii. On peut effectuer le produit matriciel qui ne pose pas de difficulté. On peut aussi remarquer que $^{\mathrm{t}}RBR$ correspond à une formule de changement de bases pour l'endomorphisme $\varphi_B : X \in \mathscr{M}_{3,1}(\mathbb{R}) \mapsto BX \in \mathscr{M}_{3,1}(\mathbb{R})$. Plus précisément, cela correspond à écrire φ_B dans la base (C_1, C_2, C_3) de $\mathscr{M}_{3,1}(\mathbb{R})$ formée des vecteurs colonnes de R. Or on a :

$$\varphi_B(C_1) = BC_1 = 3C_1, \quad \varphi_B(C_2) = C_1, \quad \varphi_B(C_3) = 0C_3.$$

Par formule de changement de bases, on obtient donc :

$$\boxed{^{\mathrm{t}}RBR = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.}$$

Partie B : Valeurs singulières d'une matrice

3. Pour tout $x, y \in \mathbb{R}^n$, la correspondance matricielle-vectorielle permet d'écrire $M_{\mathscr{B}}(g(y)) = M_{\mathscr{B}}(g)M_{\mathscr{B}}(y) = {}^{t}MY$ en utilisant les notations de l'énoncé, et de même $M_{\mathscr{B}}(f(x)) = MX$. On obtient en substituant :
$$\langle x, g(y)\rangle = {}^{t}X({}^{t}MY) = {}^{t}X{}^{t}MY = {}^{t}(MX)Y \boxed{= \langle f(x), y\rangle.}$$
De même, on a $M_{\mathscr{B}}(h(x)) = M_{\mathscr{B}}(g)M_{\mathscr{B}}(f)M_{\mathscr{B}}(x) = {}^{t}MMX$, et :
$$\langle x, h(x)\rangle = {}^{t}X({}^{t}MMX) = {}^{t}(MX)(MX) = \langle f(x), f(x)\rangle \boxed{= \|f(x)\|^2.}$$

4. (a) Soit x appartenant à $\mathrm{Ker}(h)$. Alors :
$$\|f(x)\|^2 = \langle x, h(x)\rangle = \langle x, 0_{\mathbb{R}^n}\rangle = 0.$$
Ainsi, $f(x) = 0_{\mathbb{R}^n}$, et $\boxed{x \text{ appartient à } \mathrm{Ker}(f).}$

   (b) On a déjà l'inclusion $\mathrm{Ker}(h) \subset \mathrm{Ker}(f)$ par la question précédente. Réciproquement, si $x \in \mathrm{Ker}(f)$, alors $f(x) = 0_{\mathbb{R}^n}$, et donc :
$$h(x) = g \circ f(x) = g(f(x)) = g(0_{R^n}) = 0_{\mathbb{R}^n}$$
puisque g est linéaire. D'où l'inclusion $\mathrm{Ker}(f) \subset \mathrm{Ker}(h)$, et donc l'égalité $\boxed{\mathrm{Ker}(h) = \mathrm{Ker}(f).}$

   Par le théorème du rang et l'égalité qu'on vient d'obtenir, on a :
$$\mathrm{rg}(h) = n - \dim(\mathrm{Ker}(h)) = n - \dim(\mathrm{Ker}(f)) = \mathrm{rg}(f) = \mathrm{rg}(M) \boxed{= r.}$$

5. (a) Puisque $M_{\mathscr{B}}(h) = M_{\mathscr{B}}(g)M_{\mathscr{B}}(f) = {}^{t}MM$, la matrice de h dans la base canonique est symétrique réelle. $\boxed{\text{Elle est donc diagonalisable dans } \mathscr{M}_n(\mathbb{R}), \text{ et } h \text{ l'est également.}}$ Par le théorème spectral, il existe une base orthonormée $\mathscr{B}_1 = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ de $\mathbb{R}^n$ constituée de vecteurs propres de h.

   (b) On sait déjà que les valeurs propres de h sont réelles (puisque sa matrice dans la base $\mathscr{B}$ est symétrique réelle). Soit λ une valeur propre de h et x un vecteur propre associé (nécessairement non nul donc). En utilisant l'égalité de la question 3., on a d'une part $\langle x, h(x)\rangle = \|f(x)\|^2$. Et d'autre part :
$$\langle x, h(x)\rangle = \langle x, \lambda x\rangle = \lambda\langle x, x\rangle = \lambda\|x\|^2.$$
Il suit (en notant que $\|x\|^2 \neq 0$) :
$$\lambda = \frac{\|f(x)\|^2}{\|x\|^2} \geq 0.$$
Ainsi $\boxed{\text{les valeurs propres de } h \text{ sont bien positives ou nulles.}}$

6. P étant la matrice de passage de la base $\mathscr{B}$ orthonormée pour $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ à la base $\mathscr{B}_1$ également orthonormée pour $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, $\boxed{P \text{ est une matrice orthogonale.}}$

   Par formule de changement de bases, on peut écrire :
$$M_{\mathscr{B}_1}(h) = P^{-1}M_{\mathscr{B}}(h)P = P^{-1}({}^{t}MM)P.$$
$\mathscr{B}_1$ étant une base formée de vecteurs propres de h, la matrice $M_{\mathscr{B}_1}(h)$ est diagonale, de la forme $D = \mathrm{Diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$, où les scalaires $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de h. Par la question 5.(b), ce sont donc des réels positifs ou nuls. Et on obtient bien :
$$\boxed{{}^{t}MM = PDP^{-1} = PD{}^{t}P.}$$

7. Supposons M symétrique. Par le théorème spectral, il existe $Q \in \mathscr{M}_n(\mathbb{R})$ orthogonale telle que :
$$M = Q\mathrm{Diag}(\mu_1, \ldots, \mu_n)Q^{-1}$$
où les scalaires $\mu_1, \ldots, \mu_n$ sont les valeurs propres réelles de M. Calculons :
$$^{\mathrm{t}}MM = M^2 = Q\mathrm{Diag}(\mu_1, \ldots, \mu_n)^2Q^{-1} = Q\mathrm{Diag}(\mu_1^2, \ldots, \mu_n^2)Q^{-1}.$$
Ainsi, $^{\mathrm{t}}MM$ est semblable à $\mathrm{Diag}(\mu_1^2, \ldots, \mu_n^2)$. Ces matrices ont donc les mêmes valeurs propres, à savoir $\mu_1^2, \ldots, \mu_n^2$.
Et puisque $\mathrm{Sp}(^{\mathrm{t}}MM) = \{\lambda_1, \ldots, \lambda_n\}$, on obtient, quitte à renuméroter ces valeurs propres :
$$\forall i \in [\![1, n]\!], \quad \mu_i^2 = \lambda_i \quad \Rightarrow \quad \boxed{\forall i \in [\![1, n]\!], \quad |\mu_i| = \sqrt{\lambda_i} = \sigma_i.}$$
Ainsi lorsque M est symétrique, ses valeurs singulières sont les valeurs absolues de ses valeurs propres.

8. On a établi que $\mathrm{rg}(h) = r$ à la question 4.(b). D'où :
$$\mathrm{rg}(^{\mathrm{t}}MM) = \mathrm{rg}(M_{\mathscr{B}}(h)) = \mathrm{rg}(h) = r$$
Puisque D et $^{\mathrm{t}}MM$ sont semblables, D est également de rang r. Et comme D est une matrice diagonale, $\boxed{\text{elle admet exactement } r \text{ coefficients diagonaux non nuls.}}$

9. (a) Calculons pour tout $i \in [\![1; n]\!]$:
$$\|f(\varepsilon_i)\|^2 \overset{3.}{=} \langle \varepsilon_i, h(\varepsilon_i)\rangle = \langle \varepsilon_i, \lambda_i \varepsilon\rangle = \lambda_i \|\varepsilon_i\|^2 = \lambda_i$$
en utilisant que $(\varepsilon_1, \ldots, \varepsilon_n)$ est une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de h. D'où en prenant la racine carrée :
$$\|f(\varepsilon_i)\| = \sqrt{\lambda_i} \boxed{= \sigma_i.}$$
En particulier pour $1 \leq i \leq r$, $\boxed{f(\varepsilon_i) \text{ est un vecteur non nul}}$ puisque $\sigma_i \neq 0$.

(b) Notons tout d'abord que par définition, les vecteurs u_i sont unitaires (en tant que normalisation des vecteurs non nuls $f(\varepsilon_i)$ pour $1 \leq i \leq r$).
Soit $1 \leq i < j \leq r$. Calculons :
$$\langle f(\varepsilon_i), f(\varepsilon_j)\rangle \overset{3.}{=} \langle \varepsilon_i, g \circ f(\varepsilon_j)\rangle = \langle \varepsilon_i, h(\varepsilon_j)\rangle = \langle \varepsilon_i, \lambda_j \varepsilon_j\rangle = \lambda_j \langle \varepsilon_i, \varepsilon_j\rangle = 0$$
car $\mathscr{B}_1$ est une famille orthogonale. On obtient donc :
$$\langle u_i, u_j\rangle = \frac{1}{\|f(\varepsilon_i)\| \times \|f(\varepsilon_j)\|}\langle f(\varepsilon_i), f(\varepsilon_j)\rangle = 0.$$
$\boxed{\text{La famille } (u_1, \ldots, u_r) \text{ est bien une famille orthonormée.}}$

(c) La famille $(u_1, \ldots, u_r)$ étant orthonormée, on peut la complète en une base $\mathscr{B}_2 = (u_1, \ldots, u_n)$ orthonormée de $\mathbb{R}^n$.

On a vu que pour tout $1 \leq i \leq n$, $\|f(\varepsilon_i)\| = \sigma_i$. En particulier pour $r + 1 \leq i \leq n$, $\|f(\varepsilon_i)\| = 0$ et $f(\varepsilon_i) = 0_{\mathbb{R}^n}$. Pour $1 \leq i \leq r$ à présent, on a :
$$f(\varepsilon_i) = \|f(\varepsilon_i)\|u_i = \sigma_i u_i.$$

La matrice de f dans la base $\mathscr{B}_1$ (au départ) et la base $\mathscr{B}_2$ (à l'arrivée) est donc :

$$M_{\mathscr{B}_1,\mathscr{B}_2}(f) = \begin{array}{c} \begin{array}{cccccc} f(\varepsilon_1) & \dots & f(\varepsilon_r) & f(\varepsilon_{r+1}) & \dots & f(\varepsilon_n) \end{array} \\ \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \sigma_r & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \\ u_1 \\ \vdots \\ u_r \\ u_{r+1} \\ \vdots \\ u_n \end{array} \boxed{= \mathrm{Diag}\left(\sigma_1, \dots, \sigma_r, 0, \dots, 0\right).}$$

10. Q étant une matrice de passage entre deux bases orthonormées de $\mathbb{R}^n$, $\boxed{\text{elle est orthogonale.}}$
Calculons, comme suggéré, de deux façons différentes la matrice de f dans la base $\mathscr{B}$ (au départ) et la base $\mathscr{B}_2$ (à l'arrivée). On a d'une part :

$$M_{\mathscr{B},\mathscr{B}_2}(f) = M_{\mathscr{B},\mathscr{B}_2}(f \circ \mathrm{id}) = M_{\mathscr{B}_1,\mathscr{B}_2}(f) \times M_{\mathscr{B},\mathscr{B}_1}(\mathrm{id}) = \Delta \times P_{\mathscr{B}_1,\mathscr{B}} = \Delta \times P^{-1}.$$

D'autre part, on a :

$$M_{\mathscr{B},\mathscr{B}_2}(f) = M_{\mathscr{B},\mathscr{B}_2}(\mathrm{id} \circ f) = M_{\mathscr{B},\mathscr{B}_2}(\mathrm{id}) \times M_{\mathscr{B},\mathscr{B}}(f) = P_{\mathscr{B}_2,\mathscr{B}}(\mathrm{id}) \times M = Q^{-1} \times M.$$

On obtient donc :

$$Q^{-1}M = \Delta P^{-1} \quad \Rightarrow \quad M = Q\Delta P^{-1} \boxed{= Q\Delta\,{}^{\mathrm{t}}P.}$$

11. On a déjà montré dans la partie A que R est orthogonale et que :

$${}^{\mathrm{t}}RBR = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

avec $B = {}^{\mathrm{t}} AA$. Ainsi, les valeurs singulières de A sont $\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$ et $\sigma_3 = 0$, et $\boxed{\Delta_1 = \mathrm{Diag}(\sqrt{3}, 1, 0)}$. De plus, en reprenant les notations de la partie B, $\boxed{P_1 = R}$ correspond à la matrice de passage P de la base canonique $\mathscr{B}$ à la base de diagonalisation $\mathscr{B}_1$ de B, de sorte que (par lecture matricielle des colonnes de R) :

$$\mathscr{B}_1 = (\varepsilon_1 = \frac{1}{\sqrt{6}}(-1, 1, 2), \varepsilon_2 = \frac{1}{\sqrt{6}}(\sqrt{3}, \sqrt{3}, 0), \varepsilon_3 = \frac{1}{\sqrt{6}}(\sqrt{2}, -\sqrt{2}, \sqrt{2})).$$

Reste à définir la matrice Q_1, et donc à calculer les vecteurs $u_i = \frac{1}{\|f(\varepsilon_i)\|} f(\varepsilon_i)$ (où f est l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à A). Reprenons la construction effectuée dans la partie B, et calculons donc :

$$A \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} = \begin{pmatrix} 3/\sqrt{6} \\ 0 \\ 3/\sqrt{6} \end{pmatrix} = \sqrt{\frac{3}{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad A \begin{pmatrix} \sqrt{3}/\sqrt{6} \\ \sqrt{3}/\sqrt{6} \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{3}/\sqrt{6} \\ 0 \\ -\sqrt{3}/\sqrt{6} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Ainsi, $u_1 = \frac{1}{\sqrt{2}}(1, 0, 1)$ et $u_2 = \frac{1}{\sqrt{2}}(1, 0, -1)$. Comme attendu, (u_1, u_2) est une famille orthonormale, qu'on complète en une famille orthonormale en lui ajoutant le vecteur

$u_3 = (0,1,0)$ (qui est bien unitaire et orthogonal à u_1 et u_2). Posons alors Q_1 la matrice de passage de la base canonique $\mathscr{B}$ à la base $\mathscr{B}_2 = (u_1, u_2, u_3)$:

$$\boxed{Q_1 = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}.}$$

Par l'étude effectuée dans la partie B, on peut alors affirmer que :

$$A = Q_1 \Delta_1 {}^{\mathrm{t}}P_1.$$

Partie C : Pseudo-inverse d'une matrice et application

12. Si M est inversible, alors $r = n$ et $\Delta = \mathrm{Diag}(\sigma_1, \ldots, \sigma_n)$ est inversible, d'inverse $\Delta^{-1} = \mathrm{Diag}\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_n}\right)$. Effectuons le produit :

$$\begin{aligned} M \times M^+ &= (Q\Delta {}^{\mathrm{t}}P) \times (P\Delta^{-1}{}^{\mathrm{t}}Q) = Q\Delta P^{-1} P \Delta^{-1} Q^{-1} \\ &= Q\Delta\Delta^{-1}Q^{-1} = QQ^{-1} = I_n. \end{aligned}$$

Ainsi $\boxed{M \text{ est inversible et } M^{-1} = M^+.}$

13. (a) On a :

$$\begin{aligned} M \times M^+ &= \Big(Q\Delta {}^{\mathrm{t}}P\Big) \times \left(P\,\mathrm{Diag}\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0\right) {}^{\mathrm{t}}Q\right) \\ &= Q\mathrm{Diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)\underbrace{P^{-1}P}_{=I_n}\,\mathrm{Diag}\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0\right) Q^{-1} \\ &\boxed{= Q\,\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)\,Q^{-1}} \end{aligned}$$

(b) Notons tout d'abord que $p \in \mathscr{L}(\mathbb{R}^n)$ et que $M_{\mathscr{B}}(p) = M_{\mathscr{B}}(f \circ f^+) = M_{\mathscr{B}}(f) \times M_{\mathscr{B}}(f^+) = M \times M^+$.

D'une part, on a :

$$\begin{aligned} (M \times M^+)^2 &= (Q\,\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)\,Q^{-1})^2 = Q\,\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)^2\,Q^{-1} \\ &= Q\,\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)\,Q^{-1} = M \times M^+. \end{aligned}$$

Ce qui montre que $p \circ p = p$. Donc p est un projecteur.

D'autre part, $\mathscr{B}$ est une base orthonormée de $\mathbb{R}^n$ et la matrice $M_{\mathscr{B}}(p) = M \times M^+$ est symétrique puisque :

$$\begin{aligned} {}^{\mathrm{t}}(M \times M^+) &= {}^{\mathrm{t}}(Q\,\mathrm{Diag}(1, \ldots, 1, 0 \ldots, 0){}^{\mathrm{t}}Q) = {}^{\mathrm{t}}({}^{\mathrm{t}}Q)\,{}^{\mathrm{t}}\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)\,{}^{\mathrm{t}}Q \\ &= Q\,{}^{\mathrm{t}}\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0)\,{}^{\mathrm{t}}Q = M \times M^+. \end{aligned}$$

Donc p est un endomorphisme symétrique.

On peut donc conclure que $\boxed{p \text{ est un projecteur orthogonal.}}$

(c) MM^+ est semblable à $\mathrm{Diag}(\underbrace{1, \ldots 1}_{r \text{ termes}}, 0, \ldots, 0)$ qui est de rang r. Donc $\boxed{\mathrm{rg}\,(MM^+) = r}$.

Montrons que $\mathrm{Im}(p) = \mathrm{Im}(f)$. Tout d'abord, pour tout $y \in \mathrm{Im}(p)$, il existe $x \in \mathbb{R}^n$ tel que :

$$y = p(x) = f \circ f^+(x) = f(f^+(x)) \in \mathrm{Im}(f).$$

Ainsi, $\mathrm{Im}(p) \subset \mathrm{Im}(f)$, et ces sous-espaces sont de même dimension puisque $\mathrm{rg}(p) = \mathrm{rg}(MM^+) = r = \mathrm{rg}(M) = \mathrm{rg}(f)$. D'où l'égalité $\boxed{\mathrm{Im}(p) = \mathrm{Im}(f).}$

14. (a) Nous venons de montrer que p est le projecteur orthogonal sur $F = \mathrm{Im}(p) = \mathrm{Im}(f)$. On sait par le cours que l'ensemble $\{\|y - z\|,\ z \in F\}$ admet un minimum, atteint en un unique vecteur qui est $p(y) \in F$. Ainsi, pour tout $z \in F$, on dispose de l'inégalité :

$$\|y - p(y)\| \leq \|y - z\|.$$

En particulier, pour tout $x \in \mathbb{R}^n$, $z = f(x)$ appartient à $F = \mathrm{Im}(f)$, de sorte que :

$$\boxed{\|y - p(y)\| \leqslant \|y - f(x)\|.}$$

(b) Comme $p(y) = f \circ f^+(y) = f(f^+(y))$, le vecteur $\boxed{x^* = f^+(y) \text{ convient}}$. En effet, on a bien que pour tout $x \in \mathbb{R}^n$,

$$\|y - f(x^*)\| \leq \|y - f(x)\|.$$

Ainsi : $\quad \|y - f(x^*)\| = \min\limits_{x \in \mathbb{R}^n} \|y - f(x)\|.$

Supposons $r < n$. Dans ce cas, $\mathrm{Ker}(f) \neq \{0_{\mathbb{R}^n}\}$ et il existe $x_0 \in \mathrm{Ker}(f) \setminus \{0\}$. Le vecteur $x^* + x_0$ ($\neq x^*$) répond également au problème posé puisque :

$$f(x^* + x_0) = f(x^*) + f(x_0) = f(x^*) + 0_{\mathbb{R}^n} = f(x^*).$$

D'où le résultat lorsque $r < n$.

15. (a) Par définition de la matrice d'une application linéaire dans une base donnée :

$$f_1(e_1) = (0, 0, -1),\ f_1(e_2) = (1, 0, 0),\ f_1(e_3) = (1, 0, 1) = -f_1(e_1) + f_1(e_2).$$

On obtient :

$$\boxed{\mathrm{Im}\,(f_1) = \mathrm{Vect}(f_1(e_1), f_1(e_2), f_1(e_3)) = \mathrm{Vect}((0, 0, -1), (1, 0, 0)).}$$

Le vecteur y appartient à $\mathrm{Im}(f_1)$ si, et seulement si, existe $a, b \in \mathbb{R}$ tels que :

$$(1, 1, 1) = a(0, 0, -1) + b(1, 0, 0) = (b, 0, -a) \quad \Leftrightarrow \quad \begin{cases} 1 = b \\ 1 = 0 \\ 1 = -a \end{cases}.$$

Ce système n'ayant pas de solution, $\boxed{y \text{ n'appartient pas à } \mathrm{Im}\,(f_1).}$

(b) Soit $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. On a :

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 + x_3 \\ 0 \\ -x_1 + x_3 \end{pmatrix},$$

de sorte que $f(x_1, x_2, x_3) = (x_2 + x_3, 0, -x_1 + x_3)$. Calculons alors :

$$\|y - f(x)\|^2 = \|(1 - x_2 - x_3, 1, 1 + x_1 - x_3)\|^2 \boxed{= (x_2 + x_3 - 1)^2 + 1^2 + (x_1 - x_3 + 1)^2.}$$

D'où le résultat voulu.

(c) Suivons la méthode exposée à la partie C, et reprenons les résultats obtenus à la question 11. Définissons donc :

$$M_1^+ = P_1 \operatorname{Diag}\left(\frac{1}{\sqrt{3}}, 1, 0\right) {}^{\mathrm{t}}Q_1.$$

Déterminons $x^* = f_1^+(y)$. Pour cela, calculons :

$$X^* = M_1^+ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = P_1 \operatorname{Diag}\left(\frac{1}{\sqrt{3}}, 1, 0\right) \begin{pmatrix} \sqrt{2} \\ 1 \\ 0 \end{pmatrix} = P_1 \begin{pmatrix} \sqrt{\frac{2}{3}} \\ 1 \\ 0 \end{pmatrix}$$

$$\underbrace{=}_{P_1=R} \frac{1}{\sqrt{6}} \begin{pmatrix} -\sqrt{\frac{2}{3}} + \sqrt{3} \\ \sqrt{\frac{2}{3}} + \sqrt{3} \\ 2\sqrt{\frac{2}{3}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} + \frac{1}{\sqrt{2}} \\ \frac{1}{3} + \frac{1}{\sqrt{2}} \\ \frac{2}{3} \end{pmatrix}.$$

Ainsi $\boxed{x^* = \left(-\frac{1}{3} + \frac{1}{\sqrt{2}}, \frac{1}{3} + \frac{1}{\sqrt{2}}, \frac{2}{3}\right) \text{ est solution du problème posé.}}$

Puisque $\operatorname{rg}(f_1) = 2 < 3$, il existe au moins un vecteur distinct de x^* répondant au problème. Pour en exhiber un, il nous faut déterminer $\operatorname{Ker}(f_1)$ ce qui ne pose pas de difficulté. On obtient $\operatorname{Ker}(f_1) = \operatorname{Vect}((1, -1, 1))$. Par exemple,

$$\boxed{z^* = x^* + \frac{1}{3}(1, -1, 1) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 1\right)}$$

répond aussi au problème posé.

Problème 2

Partie I

1. (a) $F_{\mu,a}$ est de classe $\mathscr{C}^2$ sur $\mathbb{R}$ en tant que composée de fonctions usuelles de classe $\mathscr{C}^2$ sur $\mathbb{R}$. Calculons sa dérivée et sa dérivée seconde :

$$\boxed{\forall x \in \mathbb{R}, \quad f_{\mu,a}(x) = \frac{1}{a} \exp\left(\frac{\mu - x}{a}\right) \exp\left(-\exp\left(\frac{\mu - x}{a}\right)\right).}$$

$$\boxed{\forall x \in \mathbb{R}, \quad f'_{\mu,a}(x) = \left[-\frac{1}{a^2} \exp\left(\frac{\mu - x}{a}\right) + \left(\frac{1}{a} \exp\left(\frac{\mu - x}{a}\right)\right)^2\right] \exp\left(-\exp\left(\frac{\mu - x}{a}\right)\right).}$$

(b) Puisque $f_{\mu,a}(x) > 0$ pour tout $x \in \mathbb{R}$, $\boxed{F_{\mu,a} \text{ est strictement croissante sur } \mathbb{R}.}$

Pour étudier la convexité de $F_{\mu,a}$, on s'intéresse au signe de sa dérivée seconde. Pour tout $x \in \mathbb{R}$, $f'_{\mu,a}(x)$ est du signe de :

$$-\frac{1}{a^2} \exp\left(\frac{\mu - x}{a}\right) + \left(\frac{1}{a} \exp\left(\frac{\mu - x}{a}\right)\right)^2 = \frac{1}{a^2} \exp\left(\frac{\mu - x}{a}\right) \left[-1 + \exp\left(\frac{\mu - x}{a}\right)\right].$$

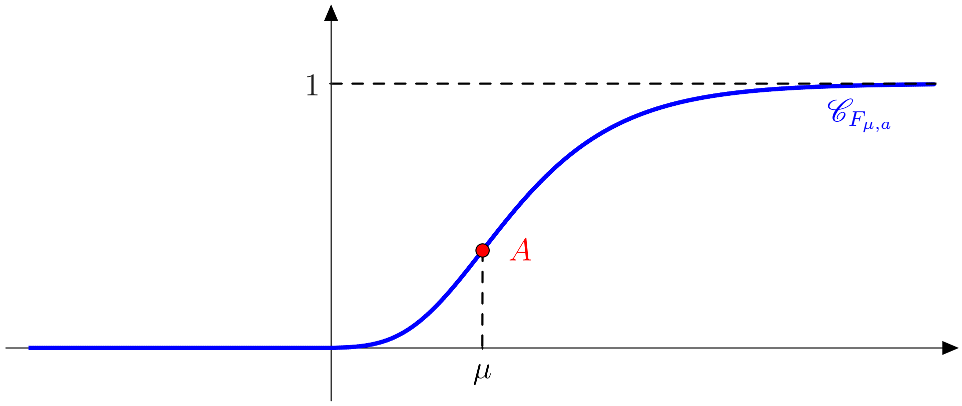
On résout l'inéquation :

$$\exp\left(\frac{\mu - x}{a}\right) - 1 \geq 0 \quad \Leftrightarrow \quad \frac{\mu - x}{a} \geq 0 \quad \Leftrightarrow \quad x \leq \mu.$$

Ainsi $f'_{\mu,a}(x)$ est positif sur $]-\infty, \mu[$, s'annule en changeant de signe en μ et est négatif sur $]\mu, +\infty[$. La fonction $\boxed{F_{\mu,a} \text{ est donc convexe sur }]-\infty, \mu[\text{, concave sur }]\mu, +\infty[}$, et sa courbe représentative admet pour $\boxed{\text{point d'inflexion } A = (\mu, F_{\mu,a}(\mu)) = (\mu, e^{-1}).}$

Par composition des limites, on vérifie que $\boxed{\lim_{x\to-\infty} F_{\mu,a}(x) = 0 \text{ et } \lim_{x\to+\infty} F_{\mu,a}(x) = 1.}$

On peut à présent donner l'allure de la courbe représentative de $F_{\mu,a}$.



(c) On a montré que :

- $F_{\mu,a}$ est continue sur $\mathbb{R}$;
- $F_{\mu,a}$ est strictement croissante sur $\mathbb{R}$;
- $\lim_{x\to-\infty} F_{\mu,a}(x) = 0$ et $\lim_{x\to+\infty} F_{\mu,a}(x) = 1$.

Par le théorème de la bijection, $\boxed{F_{\mu,a} \text{ réalise une bijection de } \mathbb{R} \text{ sur } I =]0,1[.}$

Notons G la bijection réciproque de $F_{0,1} : x \in \mathbb{R} \mapsto \exp(-\exp(-x)) \in]0,1[$. Explicitons G. Prenons pour cela $(x, y) \in \mathbb{R} \times I$ et résolvons :

$$y = F_{0,1}(x) \Leftrightarrow y = \exp(-\exp(-x)) \Leftrightarrow \ln(y) = -\exp(-x) \Leftrightarrow -\ln(-\ln(y)) = x$$

On obtient l'expression suivante de G :

$$\boxed{G : y \in]0,1[\mapsto -\ln(-\ln(y)) \in \mathbb{R}.}$$

2. On a montré précédemment que :

- $F_{\mu,a}$ est croissante sur $\mathbb{R}$,
- $F_{\mu,a}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$,
- $\lim_{x\to-\infty} F_{\mu,a}(x) = 0$ et $\lim_{x\to+\infty} F_{\mu,a}(x) = 1$.

On peut donc affirmer que $F_{\mu,a}$ est la fonction de répartition d'une variable aléatoire réelle X à densité, et dont une densité associée est donnée par $F'_{\mu,a} = f_{\mu,a}$. Dit autrement, $\boxed{f_{\mu,a} \text{ est une densité, et que } F_{\mu,a} \text{ est la fonction de répartition associée.}}$

Le saviez-vous ?

La loi de Gumbel a été introduite par le mathématicien allemand Emil Julius Gumbel (1891 - 1966). De confession juive, il est le premier professeur allemand expulsé de l'université de Heidelberg sous la pression des nazis en 1932. Il est déchu de la nationalité allemande en 1933. Il part alors à Paris, puis New York, où il continue son combat contre le nazisme en aidant les services secrets américains.

Quant à la loi de Gumbel, elle peut, par exemple, servir à prévoir le niveau des crues d'un fleuve, si on possède le relevé des débits sur dix ans. Elle peut aussi servir à prédire la probabilité d'un événement critique, comme un tremblement de terre.

3. Étant donné la formulation de la question, il nous est suggéré de redémontrer le résultat de cours sur les transformations affines d'une variable à densité. Faisons le. Pour tout $x \in \mathbb{R}$, on a :

$$F_X(x) = P(X \leq x) = P(aZ+\mu \leq x) \underbrace{=}_{a>0} P\left(Z \leq \frac{x-\mu}{a}\right) = F_Z\left(\frac{x-\mu}{a}\right)$$
$$= \exp\left(-\exp\left(\frac{\mu - x}{a}\right)\right) = F_{\mu,a}(x).$$

Ainsi $\boxed{X = aZ+\mu \text{ est une variable aléatoire à densité qui suit la loi } \mathscr{G}(\mu,a).}$

4. (a) Notons que $Y = -\ln(-\ln(U)) = G(U)$ où $G :]0,1[\to \mathbb{R}$ est la bijection réciproque de $F_{0,1} : \mathbb{R} \to]0,1[$ est continue. Ainsi $Y(\Omega) = \mathbb{R}$ et pour tout $x \in \mathbb{R}$:

$$F_Y(x) = P(Y \leq x) = P(G(U) \leq x) \underbrace{=}_{G \text{ strict. croiss.}} P(U \leq F(x)) = F_U(F(x)) \underbrace{=}_{F(x) \in]0,1[} F(x)$$

Ainsi $\boxed{Y = -\ln(-\ln(U)) \text{ suit la loi de Gumbel de paramètre } (0,1).}$

Déjà vu ?

On a ici un exemple d'application de la méthode d'inversion vue au TP5, permettant de simuler une loi de Gumbel à partir d'une loi uniforme sur $]0,1[$.

(b) On va utiliser pour écrire cette fonction les résultats suivants (obtenus précédemment) :

- Si $U \hookrightarrow \mathscr{U}(]0,1[)$, alors $Y = G(U) = -\ln(-\ln(U))$ suit une loi $\mathscr{G}(0,1)$.
- Si Y suit une loi $\mathscr{G}(0,1)$, alors $X = aY+\mu$ suit une loi $\mathscr{G}(\mu,a)$.

On suppose avoir importé les librairies `numpy` (avec le suffixe `np`) et `numpy.random` (avec le suffixe `rd`). Rappelons qu'on simule une réalisation d'une variable suivant une loi uniforme sur $]0,1[$ à l'aide de la commande `rd.random()`. On a tout ce qu'il faut à présent pour écrire la fonction demandée.

```
def gumbel(mu,a):
        U = rd.random()
        Y = -np.log(-np.log(U))
        X = a*Y+mu
        return X
```

5. (a) La fonction $f : u \mapsto \ln(u)e^{-u}$ est continue sur $]0,+\infty[$. L'intégrale $I = \int_0^{+\infty} \ln(u)e^{-u}\,du$ est donc généralisée en 0 et en $+\infty$.

En 0, on a :

- $f(u) \underset{u\to 0}{\sim} \ln(u)$;

- $\ln(u) \leq 0$ pour tout $u \in]0,1]$;
- $\displaystyle\int_0^1 \ln(u)\,\mathrm{d}u$ converge (c'est une intégrale de référence).

Par théorème de comparaison, $\displaystyle\int_0^1 f(u)\,\mathrm{d}u$ converge.

En $+\infty$ à présent, on a :

- $u^2\ln(u)e^{-u} = \left(\frac{\ln(u)}{u}\right) \times (u^3 \times e^{-u}) \underset{u\to+\infty}{\to} 0$ par croissances comparées, de sorte que $f(u) \underset{u\to+\infty}{=} o\left(\frac{1}{u^2}\right)$;
- $\frac{1}{u^2} \geq 0$ pour tout $u \geq 1$;
- $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}u}{u^2}$ converge en tant qu'intégrale de Riemann en $+\infty$ d'exposant $2 > 1$.

Par théorème de comparaison, $\displaystyle\int_1^{+\infty} f(u)\,\mathrm{d}u$ converge.

Ainsi, $\boxed{I \text{ est une intégrale convergente.}}$

(b) La fonction $\varphi : u \mapsto e^{-u}$ est de classe $\mathscr{C}^1$ et strictement décroissante sur $]0,+\infty[$. Le changement de variable $t = e^{-u}$ est donc licite. De plus, $\mathrm{d}t = -e^{-u}\,\mathrm{d}u$ et $u : +\infty \to 0$ lorsque $t : 0 \to 1$ lorsque . Par théorème de changement de variable dans les intégrales généralisées, $J = \displaystyle\int_0^1 \ln(-\ln(t))\,\mathrm{d}t$ est de même nature que $\displaystyle\int_{+\infty}^0 \ln(-\ln(e^{-u}))(-e^{-u})\,\mathrm{d}u = \int_0^{+\infty} \ln(u)e^{-u}\,\mathrm{d}u$. On reconnaît l'intégrale I de la question précédente, qui est convergente. Donc $\boxed{J \text{ est convergente.}}$

(c) Remarquons tout d'abord que Z suit une loi de Gumbel de paramètre $(0,1)$.

Z admet un espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x f_{(0,1)}(x)\mathrm{d}x$ converge absolument.

La fonction $F_{(0,1)} : x \mapsto \exp(-\exp(-x))$ est de classe $\mathscr{C}^1$ et strictement croissante sur $\mathbb{R}$ et satisfait $\lim\limits_{x\to-\infty} F_{(0,1)}(x) = 0$ et $\lim\limits_{x\to+\infty} F_{(0,1)}(x) = 1$ (question 1). Le changement de variable $u = \exp(-\exp(-x)) = F_{(0,1)}(x)$ est donc licite. De plus $x = G(u)$, $\mathrm{d}u = f_{(0,1)}(x)\mathrm{d}x$ et $u : 0 \to 1$ lorsque $x : -\infty \to +\infty$.

Par théorème de changement de variable dans les intégrales généralisées, $\displaystyle\int_{-\infty}^{+\infty} |x| f_{(0,1)}\mathrm{d}x$ est de même nature que $\displaystyle\int_0^1 |G(u)|\,\mathrm{d}u = \int_0^1 \ln(-\ln(u))\,\mathrm{d}u$. Or cette dernière intégrale converge d'après la question précédente. Ainsi $\boxed{Z \text{ admet une espérance}}$, et on obtient (sans les valeurs absolues cette fois) :

$$\boxed{E(Z) = -\int_0^1 \ln(-\ln(u))\,\mathrm{d}u = \gamma.}$$

(d) Rappelons que si Z admet une espérance, alors $X = aZ + b$ admet également une espérance qui vaut $E(X) = aE(Z) + b$.

Ici, $Z = \frac{X-\mu}{a}$, de sorte que $X = aZ + \mu$. Puisque Z admet une espérance, $E(X)$ existe et vaut :

$$\boxed{E(X) = aE(Z) + \mu = a\gamma + \mu.}$$

6. (a) On pourrait là aussi utiliser le résultat de cours sur les transformations affines de variables à densité. Redémontrons le en posant $U = -Z$. Pour tout $x \in \mathbb{R}$, on a :

$$\begin{aligned} F_U(x) &= P(U \leq x) = P(-Z \leq x) = P(Z \geq -x) = 1 - P(Z < -x) \\ &\underbrace{=}_{Z \text{ cont.}} 1 - F_{(0,1)}(-x). \end{aligned}$$

Or $F_{(0,1)}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$. Par composition, il en est de même de F_U. $\boxed{U = -Z \text{ est donc une variable à densité}}$, et une densité de $-Z$ est :

$$\boxed{g : t \in \mathbb{R} \mapsto F'_{(0,1)}(-t) = e^t e^{-e^t}.}$$

(b) Soit $x \in \mathbb{R}$ fixé. On va se ramener à une intégrale Γ en effectuant le changement de variable $s = (e^{-x} + 1)u$ qui est affine donc licite. Puisque $e^{-x} + 1 > 0$, $s : 0 \to +\infty$ lorsque $u : 0 \to +\infty$.

Par théorème de changement de variable, l'intégrale $I_x = \displaystyle\int_0^{+\infty} ue^{-\left(e^{-x}+1\right)u}\,\mathrm{d}u$ est de même nature que $\displaystyle\int_0^{+\infty} \frac{s}{e^{-x}+1} e^{-s} \frac{\mathrm{d}s}{e^{-x}+1} = \frac{1}{(e^{-x}+1)^2}\Gamma(1)$. $\boxed{I_x \text{ converge}}$ donc et vaut :

$$\boxed{I_x = \frac{1}{(e^{-x}+1)^2}.}$$

(c) Soit $x \in \mathbb{R}$ fixé. On étudie l'intégrale :

$$\begin{aligned} J_x &= \int_{-\infty}^{+\infty} f_{0,1}(x-t)g(t)\mathrm{d}t = \int_{-\infty}^{+\infty} e^{-(x-t)} e^{-e^{-(x-t)}} e^t e^{-e^t}\,\mathrm{d}t \\ &= e^{-x} \int_{-\infty}^{+\infty} e^{2t} e^{-e^{t-x} - e^t}\,\mathrm{d}t = e^{-x} \int_{-\infty}^{+\infty} e^{2t} e^{-e^t(e^{-x}+1)}\,\mathrm{d}t \end{aligned}$$

La fonction $t \mapsto e^t$ est de classe $\mathscr{C}^1$ et strictement croissante sur $\mathbb{R}$. Le changement de variable $u = e^t$ est donc licite. De plus, on a $\mathrm{d}u = e^t\,\mathrm{d}t$ et $u : 0 \to +\infty$ lorsque $t : -\infty \to +\infty$.

Par théorème de changement de variable, l'intégrale J_x est de même nature que :

$$e^{-x} \int_0^{+\infty} u^2 e^{-u(e^{-x}+1)} \frac{\mathrm{d}u}{u} = e^{-x} I_x.$$

$\boxed{\text{L'intégrale } J_x \text{ converge}}$ donc, et elle vaut :

$$\boxed{J_x = \frac{e^{-x}}{(e^{-x}+1)^2}.}$$

(d) Vérifions les différents points permettant d'effectuer un produit de convolution :

- Les variables Y et $-Z$ sont indépendantes par lemme de coalition, puisque Y et Z le sont.
- D'après l'étude faite à la question 1.(b), $f_{(0,1)}$ est croissante sur $]-\infty, 0[$ et décroissante sur $]0, +\infty[$, de sorte que $0 < f_{(0,1)}(t) \leq f_{(0,1)}(0) = e^{-1}$ pour tout $t \in \mathbb{R}$. La fonction $f_{(0,1)}$ est donc bornée sur $\mathbb{R}$.

Par le cours, on peut donc affirmer que $\boxed{Y + (-Z) = Y - Z \text{ est une variable à densité}}$, de densité (continue sur $\mathbb{R}$) :

$$\boxed{h : x \in \mathbb{R} \mapsto \int_{-\infty}^{+\infty} f_{0,1}(x-t)g(t)\mathrm{d}t \overset{6.(c)}{=} \frac{e^{-x}}{(e^{-x}+1)^2}.}$$

Partie II

7. (a) Les fonctions $x \mapsto \alpha x$ et $x \mapsto \beta x$ étant continues sur $\mathbb{R}$, on en déduit que (αV_n) et (βW_n) convergent en probabilité vers αV et βW respectivement. Par compatibilité de la convergence en probabilités avec la somme, on obtient :

$$\boxed{\alpha V_n + \beta W_n \xrightarrow{P} \alpha V + \beta W.}$$

(b) $(X_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes ayant toutes même espérance $E(X_1)$ et même variance (X_1 admettant un moment d'ordre 4 (résultat admis), elle admet en particulier un moment d'ordre 2). Par la loi faible des grands nombres, $\boxed{M_n \text{ converge en probabilité vers } E(X_1).}$

De même, $(X_n^2)_{n\geq 1}$ est une suite de variables aléatoires indépendantes (par lemme de coalition) ayant toutes même espérance $E(X_1^2)$ et même variance (X_1 admettant un moment d'ordre 4). Toujours par la loi faible des grands nombres, $\boxed{C_n \text{ converge en probabilité vers } E(X_1^2).}$

(c) La fonction $x \mapsto x^2$ étant continue sur $\mathbb{R}$, on sait par le cours que M_n^2 converge en probabilité vers $E(X_1)^2$. À l'aide de la question 7.(a), on en déduit que $C_n - M_n^2$ converge en probabilité vers $E(X_1^2) - E(X_1)^2 = V(X_1) \geq 0$.

Par composition avec la fonction $x \mapsto \frac{1}{\sqrt{c}}\sqrt{x}$ continue sur $\mathbb{R}_+$, A_n converge donc en probabilité vers $\frac{1}{\sqrt{c}}\sqrt{V(X_1)} = \frac{1}{\sqrt{c}}\sqrt{a^2c} = a$.

Ainsi $\boxed{A_n = \frac{1}{\sqrt{c}}\sqrt{C_n - M_n^2} \text{ est un estimateur convergent de } a.}$

(d) On l'a vu, M_n converge en probabilité vers $E(X_1)$ et A_n converge en probabilité vers a. À l'aide de la question 7.(a), on peut donc en déduire que $S_n = M_n - A_n\gamma$ converge en probabilité vers :

$$E(X_1) - \gamma a = a\gamma + \mu - \gamma a = \mu.$$

Ainsi $\boxed{S_n \text{ est un estimateur convergent de } \mu.}$

8. (a) On suppose avoir importé les librairies `numpy` (avec le préfixe `np`) et `numpy.random` (avec le préfixe `rd`). Pour définir cette fonction, il suffit de reprendre les expressions des différents estimateurs C_n, M_n et A_n et de les traduire en langage `Python`.

```python
def estimateur_a(X):
        M = np.mean(X)
        C = np.mean(X**2)
        A = np.sqrt((C-M**2)/c)
        return A
```

(b) On observe sur ce graphique la convergence lorsque $n \to +\infty$ de A_n vers $a = 1$. Ceci confirme le résultat obtenu à la question 7.(c), à savoir la convergence en probabilité de A_n vers a.