EML 2020

Exercice 1

On considère la fonction f définie sur $]0,1[$ par :

$$f : x \mapsto \frac{\ln(1-x)}{\ln(x)}$$

Partie A : Étude de la fonction f

1. Montrer que f est dérivable sur $]0,1[$ et que l'on a :

$$\forall x \in]0,1[,\ f'(x) = \frac{1}{x\,(1-x)\,\big(\ln(x)\big)^2}\,\big(-x\,\ln(x) - (1-x)\,\ln(1-x)\big)$$

Démonstration.

- La fonction f est dérivable sur $]0,1[$ car elle est le quotient $f = \dfrac{f_1}{f_2}$ où :
  - $f_1 : x \mapsto \ln(1-x)$ est dérivable sur $]0,1[$ car elle est la composée $f_1 = g_2 \circ g_1$ où :
    - $g_1 : x \mapsto 1-x$ est :
      - dérivable sur $]0,1[$,
      - telle que : $g_1\big(]0,1[\big) \subseteq]0,1[$.
    - $g_2 : x \mapsto \ln(x)$ est dérivable sur $]0,1[$.
  - $f_2 : x \mapsto \ln(x)$:
    - est dérivable sur $]0,1[$,
    - `NE S'ANNULE PAS` sur $]0,1[$.

  La fonction f est dérivable sur $]0,1[$.

- Soit $x \in]0,1[$.

$$\begin{aligned} f'(x) &= \frac{\frac{-1}{1-x}\,\ln(x) - \ln(1-x)\,\frac{1}{x}}{\big(\ln(x)\big)^2} \\ &= \frac{\frac{-x\,(1-x)}{1-x}\,\ln(x) - \ln(1-x)\,\frac{x\,(1-x)}{x}}{x\,(1-x)\,\big(\ln(x)\big)^2} \\ &= \frac{-x\,\ln(x) - (1-x)\,\ln(1-x)}{x\,(1-x)\,\big(\ln(x)\big)^2} \end{aligned}$$

On a bien : $\forall x \in]0,1[,\ f'(x) = \dfrac{1}{x\,(1-x)\,\big(\ln(x)\big)^2}\,\big(-x\,\ln(x) - (1-x)\,\ln(1-x)\big)$.

Commentaire

- Afin de permettre une bonne compréhension de ce point, on a rédigé en détails la dérivabilité de la fonction f_1 qui est obtenue comme composée ($f_2 = g_2 \circ g_1$). Mais un tel niveau de détails n'est certainement pas attendu par les correcteurs.

Commentaire

- De manière générale, il n'est pas nécessaire de rédiger aussi précisément les questions portant sur la régularité de fonctions. Il est conseillé :
  - × de rédiger très proprement la régularité d'une fonction pour les questions que l'on traite en premier. On démontre ainsi au correcteur sa capacité à rédiger ce type de questions et on pourra alors réduire le niveau de détail pour les questions suivantes.
  - × de rédiger très proprement la régularité lorsqu'il s'agit du cœur de la question (« Démontrer que la fonction est continue / de classe $\mathcal{C}^1$ sur ... »

  Dans les autres cas, on pourra se contenter d'écrire que la fonction f est dérivable sur J (intervalle à déterminer) car elle est la composée de fonctions dérivables sur les intervalles adéquats. Évidemment, cela n'apportera pas de point si l'intervalle J n'est pas le bon.
- Les précisions apportées dans cette question permettent de rappeler la formule de dérivation d'une composée. On a :

$$\forall x \in]0,1[,\ (g_2 \circ g_1)'(x) = g_2'\big(g_1(x)\big) \times g_1'(x)$$

  Ici on a : $g_1'(x) = -1$. Il est évidemment primordial de ne pas oublier ce terme dans l'obtention de la dérivée. L'énoncé donne ici la solution ce qui permet de vérifier que l'on n'a pas commis ce type d'erreur.

□

2. a) Justifier : $\forall t \in]0,1[,\ t\ln(t) < 0$.

Démonstration.
Soit $t \in]0,1[$.

Alors $\ln(t) < 0$ *(par définition de la fonction* ln*)*

donc $t\ln(t) < 0$ *(par multiplication par* $t > 0$*)*

$\boxed{\forall t \in]0,1[,\ t\ln(t) < 0}$

□

b) En déduire que la fonction f est strictement croissante sur $]0,1[$.

Démonstration.
Soit $x \in]0,1[$.

- D'après la question ***1.***, on a :

$$f'(x) = \frac{1}{x(1-x)\big(\ln(x)\big)^2}\ \big(-x\ln(x) - (1-x)\ln(1-x)\big)$$

  Or : $x > 0$, $1-x > 0$ et $\big(\ln(x)\big)^2 > 0$. Ainsi : $x(1-x)\big(\ln(x)\big)^2 > 0$.
  On en déduit que le signe de $f'(x)$ est celui de la quantité $-x\ln(x) - (1-x)\ln(1-x)$.
- En utilisant la propriété de la question précédente pour $t = x \in]0,1[$, on obtient : $x\ln(x) < 0$ et donc $-x\ln(x) > 0$.
- En utilisant la propriété de la question précédente pour $t = 1-x \in]0,1[$, on obtient : $(1-x)\ln(1-x) < 0$ et donc $-(1-x)\ln(1-x) > 0$.

Ainsi : $-x\ln(x) - (1-x)\ln(1-x) > 0$.

On en déduit : $\forall x \in]0,1[,\ f'(x) > 0$.
La fonction f est donc strictement croissante sur $]0,1[$.

Commentaire

La propriété démontrée dans la question précédente a été démontrée pour tout $t \in]0,1[$. On ne peut donc l'utiliser que pour un réel $t \in]0,1[$. C'est une évidence qu'il convient toutefois de rappeler car elle est trop régulièrement ignorée par les candidats. Lorsque l'on souhaite utiliser un résultat précédemment démontré ou admis, il faut scrupuleusement vérifier que l'on est dans les conditions d'application de ce résultat.

□

3. a) Montrer que la fonction f est prolongeable par continuité en 0.
On note encore f la fonction ainsi prolongée en 0. Préciser $f(0)$.

Démonstration.

- Comme $\lim\limits_{x\to 0} \ln(x) = -\infty$ alors $\lim\limits_{x\to 0} \dfrac{1}{\ln(x)} = 0$.
- Comme $\lim\limits_{x\to 0} \ln(1-x) = \ln(1) = 0$, on obtient :

$$\lim_{x\to 0} \frac{\ln(1-x)}{\ln(x)} = 0 \times 0 = 0$$

La fonction f est prolongeable par continuité en posant $f(0) = 0$. □

b) Montrer que f est dérivable en 0 et préciser $f'(0)$.

Démonstration.
Soit $x \in]0,1[$.

$$\tau_0(f)(x) = \frac{f(x)-f(0)}{x-0} = \frac{\frac{\ln(1-x)}{\ln(x)} - 0}{x} = \frac{\ln(1-x)}{x\ln(x)} \underset{x\to 0}{\sim} \frac{-x}{x\ln(x)} = \frac{-1}{\ln(x)}$$

Or : $\lim\limits_{x\to 0} \dfrac{-1}{\ln(x)} = 0$.
La fonction taux d'accroissement $\tau_0(f)$ admet donc une limite finie lorsque $x \to 0$.

On en conclut que la fonction f est dérivable en 0, de dérivée $f'(0) = 0$. □

4. Calculer la limite de f en 1. Que peut-on en déduire pour la courbe représentative de f ?

Démonstration.

- Tout d'abord, comme $\lim\limits_{x\to 1} \ln(x) = 0$, on a : $\lim\limits_{x\to 1} \dfrac{1}{\ln(x)} = -\infty$.
- En posant $t = 1-x$, on obtient : $\lim\limits_{x\to 1} \ln(1-x) = \lim\limits_{t\to 0} \ln(t) = -\infty$.

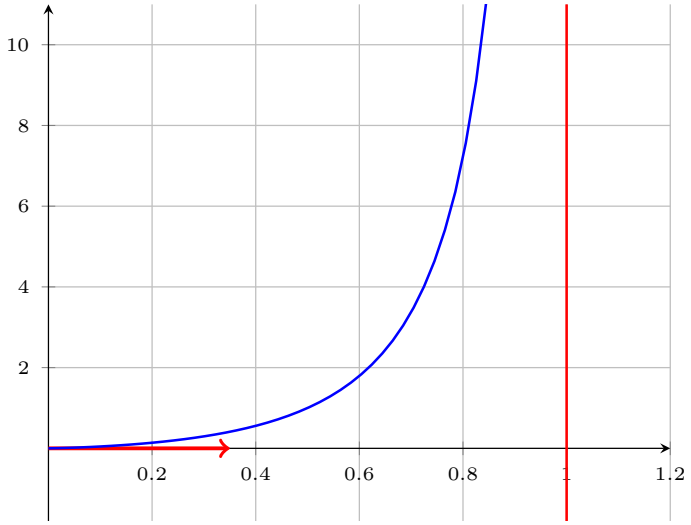
On en déduit, par produit de limites :

$$\lim_{x\to 1} \frac{\ln(1-x)}{\ln(x)} = +\infty$$

Comme $\lim\limits_{x\to 1} f(x) = +\infty$, la droite $x = 1$ est une asymptote verticale de la courbe représentative de f. □

5. Tracer l'allure de la courbe représentative de f dans un repère orthonormé, en faisant figurer la tangente en 0 et les branches infinies éventuelles.

Démonstration.



□

Partie B : Étude d'une suite

On note, pour tout n de $\mathbb{N}^*$, (E_n) l'équation : $x^n + x - 1 = 0$.

6. Soit $n \in \mathbb{N}^*$. Étudier les variations sur $\mathbb{R}_+$ de la fonction $x \mapsto x^n + x - 1$.
En déduire que l'équation (E_n) admet une unique solution sur $\mathbb{R}_+$ que l'on note u_n.

Démonstration.
Soit $n \in \mathbb{N}^*$. Dans la suite, on note $h_n : x \mapsto x^n + x - 1$.

- La fonction h_n est une fonction polynomiale (de degré n).
  Elle est donc dérivable sur $\mathbb{R}_+$. De plus, pour tout $x \in \mathbb{R}_+$:

$$h_n'(x) \;=\; n\,x^{n-1} + 1$$

  Comme $x \geqslant 0$, alors : $x^{n-1} \geqslant 0$. Ainsi : $n\,x^{n-1} + 1 \geqslant 1 > 0$.
  On en déduit le tableau de variation suivant.

x	0 $\qquad\qquad$ $+\infty$
Signe de $h_n'(x)$	$+$
h_n	-1 ↗ $+\infty$

- La fonction h_n est :
  × continue sur $[0, +\infty[$,
  × strictement croissante sur $[0, +\infty[$.
  Elle réalise donc une bijection de $[0, +\infty[$ sur $h_n([0, +\infty[)$. Or :

$$h_n([0, +\infty[) \;=\; [h_n(0), \lim_{x \to +\infty} h_n(x)[\;=\; [-1, +\infty[$$

  Comme $0 \in [-1, +\infty[$, l'équation $h_n(x) = 0$ admet une unique solution $u_n \in [0, +\infty[$.

L'équation (E_n) admet une unique solution sur $\mathbb{R}_+$ notée u_n.

□

7. Montrer que, pour tout n de $\mathbb{N}^*$, u_n appartient à l'intervalle $]0,1[$.

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

- Remarquons :
  - $h_n(0) = -1$.
  - $h_n(u_n) = 0$, par définition.
  - $h_n(1) = 1$.

  Ainsi :
$$h_n(0) \ < \ h_n(u_n) \ < \ h_n(1)$$
- Or, d'après le théorème de la bijection, $h_n^{-1} : [-1,+\infty[\to [0,+\infty[$ est strictement croissante sur $[-1,+\infty[$. En appliquant h_n^{-1}, on obtient alors :
$$\begin{array}{ccccc} h_n^{-1}\big(h_n(0)\big) & < & h_n^{-1}\big(h_n(u_n)\big) & < & h_n^{-1}\big(h_n(1)\big) \\ \shortparallel & & \shortparallel & & \shortparallel \\ 0 & & u_n & & 1 \end{array}$$

On a bien : $\forall n \in \mathbb{N}^*,\ u_n \in]0,1[$. □

8. Déterminer u_1 et u_2.

Démonstration.

- Par définition, u_1 est l'unique solution positive de l'équation : $h_1(x) = 0$. Or :
$$h_1(x) = 0 \ \Leftrightarrow \ x^1 + x - 1 = 0 \ \Leftrightarrow \ 2\,x = 1 \ \Leftrightarrow \ x = \frac{1}{2}$$

  On en déduit : $u_1 = \dfrac{1}{2}$.
- Par définition, u_2 est l'unique solution positive de l'équation : $h_2(x) = 0$. Or :
$$h_2(x) = 0 \ \Leftrightarrow \ x^2 + x - 1 = 0$$

  Notons $P(X) = X^2 + X - 1$ le polynôme de degré 2 associé à la fonction h_2.
  Ce polynôme admet pour discriminant : $\Delta = 1^2 - 4 \times (-1) = 1 + 4 = 5 > 0$.
  Ainsi, P admet deux racines :
$$x_- = \frac{-1-\sqrt{5}}{2} < 0 \qquad \text{et} \qquad x_+ = \frac{-1+\sqrt{5}}{2}$$
  Comme $5 \geqslant 1$, alors : $\sqrt{5} \geqslant \sqrt{1} = 1$ et ainsi : $\sqrt{5} - 1 \geqslant 0$.

  On en déduit : $u_2 = \dfrac{-1+\sqrt{5}}{2}$.

□

9. a) Recopier et compléter la fonction **Scilab** suivante afin que, prenant en argument un entier n de $\mathbb{N}^*$, elle renvoie une valeur approchée de u_n à 10^{-3} près, obtenue à l'aide de la méthode par dichotomie.

```
1  function u = valeur_approchee(n)
2      a = 0
3      b = 1
4      while ...
5          c = (a + b) / 2
6          if (c^n + c - 1) > 0 then
7              ...
8          else
9              ...
10         end
11         u = ...
12     end
13 endfunction
```

Démonstration.

- Afin de bien comprendre tous les mécanismes en jeu, on se permet d'apporter une réponse très détaillée à cette question, accompagnée d'un aparté sur la méthode de recherche par dichotomie. Il faut toutefois garder en tête qu'un tel niveau de détail n'est pas du tout attendu lors des concours. Fournir la fonction **Scilab** démontre la bonne compréhension et permet d'obtenir la totalité des points alloués à cette question.
  Commençons par rappeler le cadre de la recherche par dichotomie.

> *Calcul approché d'un zéro d'une fonction par dichotomie*
> **Données** :
> × une fonction $f : \mathbb{R} \to \mathbb{R}$,
> × un intervalle de recherche $[a, b]$,
> × une précision de recherche ε.
>
> **Résultat** : une valeur approchée à ε près d'un zéro (sur l'intervalle $[a, b]$) de la fonction f.
> Autrement dit, une valeur approchée (à ε près) d'un réel $x \in [a, b]$ tel que : $f(x) = 0$.

- La dichotomie est une méthode itérative dont le principe, comme son nom l'indique, est de découper à chaque itération l'intervalle de recherche en deux nouveaux intervalles. L'intervalle de recherche est découpé en son milieu. On obtient deux nouveaux intervalles :
  × un intervalle dans lequel on sait que l'on va trouver un zéro de f.
  Cet intervalle est conservé pour l'itération suivante.
  × un intervalle dans lequel ne se trouve pas forcément un zéro de f.
  Cet intervalle n'est pas conservé dans la suite de l'algorithme.

  La largeur de l'intervalle de recherche est ainsi divisée par 2 à chaque itération.
  On itère jusqu'à obtenir un intervalle I contenant un zéro de f et de largeur plus faible que ε.
  Les points de cet intervalle I sont tous de bonnes approximations du zéro de f contenu dans I.

- C'est le **théorème des valeurs intermédiaires** qui permet de choisir l'intervalle qu'il faut garder à chaque étape. Rappelons son énoncé et précisons maintenant l'algorithme :

Théorème des Valeurs Intermédiaires

Soit $f : [a, b] \to \mathbb{R}$ continue sur l'**intervalle** $[a, b]$.

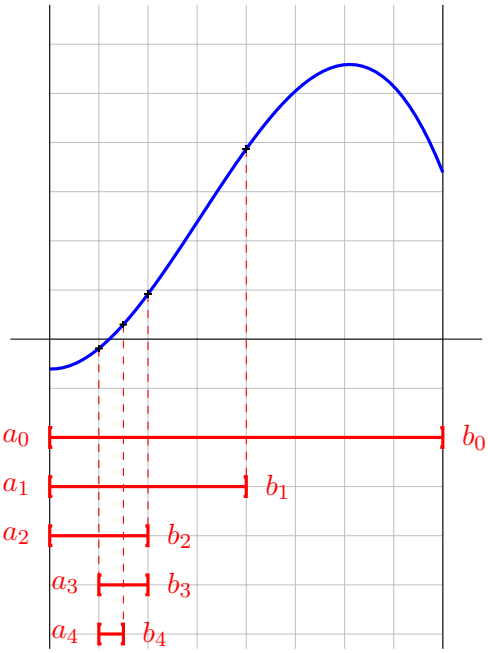
Supposons : $f(a)\ f(b) \leqslant 0$.

Alors il existe $c \in [a, b]$ tel que $f(c) = 0$.

Calcul des suites (a_m), (b_m), (c_m)

Cas $f(a) \leqslant 0$ et $f(b) \geqslant 0$

- Initialement, $a_0 = a$, $b_0 = b$
- À chaque tour de boucle (tant que $b_m - a_m > \varepsilon$) :
  - $c_m = \dfrac{a_m + b_m}{2}$ (point milieu de $[a_m, b_m]$)
  - si $f(c_m) < 0$ alors :
    - $a_{m+1} = c_m$
    - $b_{m+1} = b_m$
  - si $f(c_m) \geqslant 0$ alors :
    - $a_{m+1} = a_m$
    - $b_{m+1} = c_m$



- On construit ainsi une suite $\big([a_m, b_m]\big)_{m\in\mathbb{N}}$ de segments emboîtés :
  - contenant tous un zéro de f,
  - dont la largeur est divisée par deux d'un rang au suivant.
- Il reste enfin à adapter cet algorithme à l'énoncé.
  Soit $n \in \mathbb{N}^*$. On cherche une valeur de x telle que : $h_n(x) = 0$.
  On se fixe initialement l'intervalle de recherche $[0, 1]$ de sorte que l'équation $h_n(x) = 0$ ne possède qu'une solution, à savoir la valeur u_n qu'on cherche à approcher. D'un point de vue informatique, on crée des variables `a` et `b` destinées à contenir les valeurs succesives de a_m et b_m. Ces variables sont initialisées respectivement à 0 et 1.

```
2    a = 0
3    b = 1
```

On procède alors de manière itérative, tant que l'intervalle de recherche n'est pas de largeur plus faible que la précision 10^{-3} escomptée.

```
4    while (b-a) > 10^(-3)
```

On commence par définir le point milieu du segment de recherche.

```
5        c = (a+b) / 2
```

Puis on teste si $h_n(\texttt{c}) > 0$.
Si c'est le cas, la recherche s'effectue dans le demi-segment de gauche.

```
6        if (c^n + c - 1) > 0 then
7          b = c
```

Sinon, elle s'effectue dans le demi-segment de droite.

```
8       else
9         a = c
10      end
```

En sortie de boucle, on est assuré que le segment de recherche, mis à jour au fur et à mesure de l'algorithme, est de largeur plus faible que 10^{-3} et contient un zéro de h_n. Tout point de cet intervalle est donc une valeur approchée à 10^{-3} près de ce zéro.
On peut alors choisir de renvoyer le point le plus à gauche du segment.

```
12      u = a
```

On peut tout aussi bien choisir le point le plus à droite :

```
12      u = b
```

Ou encore le point milieu :

```
12      u = (a + b) / 2
```

Ce dernier choix présente un avantage : tout point (dont le zéro recherché) du dernier intervalle de recherche se situe à une distance d'au plus $\dfrac{10^{-3}}{2}$ de ce point milieu.
On obtient ainsi une valeur approchée à $\dfrac{10^{-3}}{2}$ du zéro recherché.

Commentaire

- On peut se demander combien de tours de boucle sont nécessaires pour obtenir le résultat. Pour le déterminer, il suffit d'avoir en tête les éléments suivants :
  - × l'intervalle de recherche initial $[0, 1]$ est de largeur 1.
  - × la largeur de l'intervalle de recherche est divisée par 2 à chaque tour de boucle. À la fin du $m^{\text{ème}}$ tour de boucle, l'intervalle de recherche est donc de largeur $\dfrac{1}{2^m}$.
  - × l'algorithme s'arrête lorsque l'intervalle devient de largeur plus faible que 10^{-3}.

  On obtient le nombre d'itérations nécessaires en procédant par équivalence :

$$\begin{aligned}
\frac{1}{2^m} \leqslant 10^{-3} &\Leftrightarrow 2^m \geqslant 10^3 && \textit{(par stricte croissance de la fonction inverse sur } \mathbb{R}_+^*\textit{)}\\
&\Leftrightarrow \ln\left(2^m\right) \geqslant \ln\left(10^3\right) && \textit{(par stricte croissance de la fonction } \ln \textit{ sur } \mathbb{R}_+^*\textit{)}\\
&\Leftrightarrow m\ \ln(2) \geqslant 3\ \ln(10)\\
&\Leftrightarrow m \geqslant 3\ \frac{\ln(10)}{\ln(2)} && \textit{(car } \ln(2) > 0\textit{)}
\end{aligned}$$

  Ainsi : $\left\lceil 3\ \dfrac{\ln(10)}{\ln(2)} \right\rceil$ tours de boucle suffisent.

  On retiendra que si l'on souhaite obtenir une précision de 3 chiffres après la virgule, il suffit d'effectuer de l'ordre de 3 tours de boucle. Cet algorithme est donc extrêmement rapide.

- On obtient le programme complet suivant.

```
1  function u = valeur_approchee(n)
2    a = 0
3    b = 1
4    while (b-a) > 10^(-3)
5      c = (a+b) / 2
6      if (c^n + c - 1) > 0 then
7        b = c
8      else
9        a = c
10     end
11   end
12   u = a
13 endfunction
```

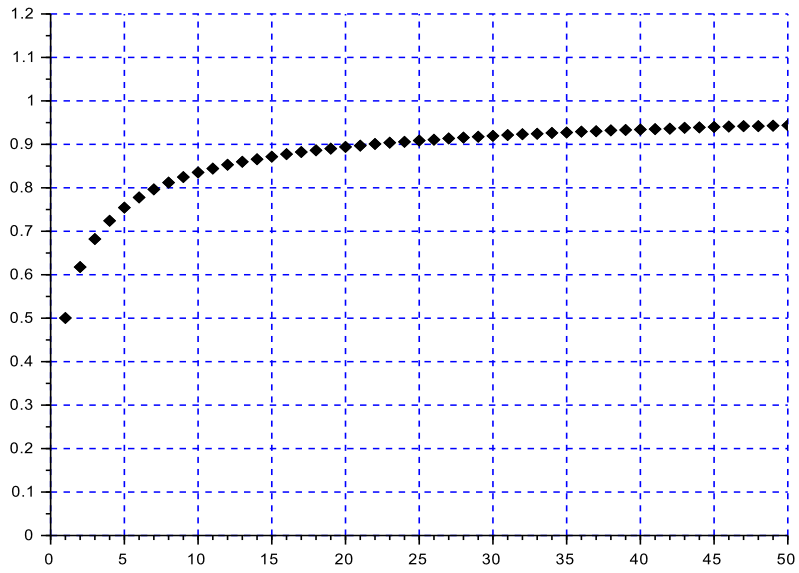
Commentaire

Dans le programme à trous donné par l'énoncé, l'instruction :

```
11     u = ...
```

apparaît en ligne 11, comme dernière instruction de la boucle. Dans ce cas, l'affectation `u = a` va être effectuée à chaque tour de boucle. Ainsi, la valeur de la variable **u** est écrasée à chaque tour de boucle. Il en résulte que la variable **u** contient en fin de boucle la dernière valeur qui lui a été affectée. De ce fait, on peut s'interroger sur la pertinence d'une telle présentation. Comme seule la dernière affectation `u = a` permet de définir la valeur de la variable **u**, il apparaît bien plus raisonnable d'effectuer cette instruction une seule fois en sortie de boucle. C'est le choix qui est fait dans ce corrigé. □

b) On représente alors les premiers termes de la suite $(u_n)_{n\in\mathbb{N}^*}$ et on obtient le graphe suivant. Quelles conjectures peut-on faire sur la suite $(u_n)_{n\in\mathbb{N}^*}$ concernant sa monotonie, sa convergence et son éventuelle limite ?



Démonstration.
Le graphe permet d'effectuer les conjectures suivantes :

× la suite $(u_n)_{n\in\mathbb{N}^*}$ est strictement croissante,

× la suite $(u_n)_{n\in\mathbb{N}^*}$ est minorée par $\frac{1}{2}$,

× la suite $(u_n)_{n\in\mathbb{N}^*}$ est majorée par 1,

× la suite $(u_n)_{n\in\mathbb{N}^*}$ est convergente de limite 1.

Commentaire

On lit sur ce graphe les valeurs :

$$u_1 = 0.5 \qquad \text{et} \qquad u_2 \simeq 0.61$$

Cela permet de vérifier les résultats donnés en question ***8***.

En particulier, 0.61 est bien une valeur approchée de la quantité $u_2 = \dfrac{-1+\sqrt{5}}{2}$.

□

10. a) Montrer, pour tout n de $\mathbb{N}^*$: $f(u_n) = n$.

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

$$\begin{aligned} f(u_n) &= \frac{\ln(1-u_n)}{\ln(u_n)} \\ &= \frac{\ln\left({u_n}^n\right)}{\ln(u_n)} \qquad \textit{(car } {u_n}^n = 1-u_n \textit{ par définition de } u_n\textit{)} \\ &= \frac{n\ln(u_n)}{\ln(u_n)} \end{aligned}$$

On a bien : $\forall n \in \mathbb{N}^*,\ f(u_n) = n$.

□

b) En déduire que la suite $(u_n)_{n\in\mathbb{N}^*}$ est croissante.

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

- D'après la question précédente :

$$f(u_n) = n \leqslant n+1 = f(u_{n+1})$$

- Or, d'après l'étude en ***Partie A***, la fonction f réalise une bijection de $]0,1[$ sur $]0,+\infty[$. D'après le théorème de la bijection, $f^{-1} :]0,+\infty[\to]0,1[$ est strictement croissante sur $]0,+\infty[$.
En appliquant f^{-1}, on obtient alors :

$$\begin{array}{ccc} f^{-1}\big(f(u_n)\big) & \leqslant & f^{-1}\big(f(u_{n+1})\big) \\ \| & & \| \\ u_n & & u_{n+1} \end{array}$$

Ainsi : $\forall n \in \mathbb{N}^*,\ u_n \leqslant u_{n+1}$. La suite (u_n) est donc croissante.

Commentaire

- La **Partie B** consiste en l'étude de la suite (u_n). On parle ici de « suite implicite » car on n'a pas accès à la définition explicite de la suite (u_n) mais simplement à la propriété qui permet de définir chacun de ses termes, à savoir :

  Pour tout $n \in \mathbb{N}^*$, u_n est l'unique solution positive de l'équation $h_n(x) = 0$

  On comprend alors que l'étude de (u_n) va passer par l'étude des propriétés de la fonction h_n.
- De cette définition, on tire la propriété : $\boxed{\forall m \in \mathbb{N}^*,\ h_m(u_m) = 0}$.
  Cette propriété est au cœur de l'étude de la suite implicite (u_n).
  On l'utilise en question ***7.***.
- La question ***10.a)*** établit une propriété équivalente à cette propriété de définition de la suite (u_n), à savoir : $\boxed{\forall m \in \mathbb{N}^*,\ f(u_m) = m}$.
  C'est de cette propriété dont on se sert ici (en $m = n$ et $m = n+1$) pour démontrer la monotonie de la suite (u_n). Comme la suite (u_n) est définie de manière implicite, cette étude ne se réalise pas directement en étudiant la différence $u_{n+1} - u_n$. Il est par contre très classique de passer par l'inégalité :

  $$f(u_n) \leqslant f(u_{n+1})$$

  et de conclure : $u_n \leqslant u_{n+1}$ à l'aide d'une propriété de f. □

c) Montrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ converge et préciser sa limite.

Démonstration.

- D'après ce qui précède, la suite (u_n) est :
  - × croissante,
  - × majorée par 1 (on démontre en question ***7.*** : $\forall n \in \mathbb{N}^*,\ u_n \in]0,1[$).

$\boxed{\text{On en conclut que la suite } (u_n) \text{ est convergente de limite } \ell \in [0,1].}$

Commentaire

- On rappelle que, par passage à la limite, les inégalités strictes deviennent **larges**. Plus précisément :

$$\left.\begin{array}{l} \forall n \in \mathbb{N},\ u_n > a \\ u_n \underset{n\to+\infty}{\longrightarrow} \ell \end{array}\right\} \quad \Rightarrow \quad \ell \geqslant a$$

- On pourra retenir l'exemple classique de la suite $\left(\dfrac{1}{n}\right)$:
  - × d'une part : $\forall n \in \mathbb{N}^*,\ \dfrac{1}{n} > 0$
  - × d'autre part : $\lim\limits_{n\to+\infty} \dfrac{1}{n} = 0 \not> 0$

- Par ailleurs, comme la suite (u_n) est croissante, on a :

$$\forall n \in \mathbb{N}^*,\ u_n \geqslant u_1 = \frac{1}{2}$$

On en conclut, par passage à la limite : $\ell \geqslant \dfrac{1}{2}$.

- Démontrons alors $\ell = 1$. Pour ce faire, on procède par l'absurde.
On suppose : $\ell \neq 1$. D'après ce qui précède, on a donc : $\ell \in [\frac{1}{2}, 1[$.
La fonction f est continue en $\ell \in [\frac{1}{2}, 1[\subseteq]0, 1[$. Ainsi, la suite $\big(f(u_n)\big)$ est convergente de limite $f(\ell) \in \mathbb{R}$. Par passage à la limite dans l'égalité définie en ***10.a)***, on obtient :

$$\begin{array}{ccc} f(u_n) & = & n \\ \overset{u}{\underset{+\infty}{\downarrow}} & & \overset{u}{\underset{+\infty}{\downarrow}} \\ f(\ell) & & +\infty \end{array}$$

Absurde !

On en conclut que la suite (u_n) est convergente de limite $\ell = 1$.

□

Partie C : Étude d'une fonction de deux variables

On considère la fonction F de classe $\mathcal{C}^2$ sur l'ouvert $]0, +\infty[^2$ définie par :

$$F : (x, y) \mapsto x^2\, y + x^2 - \frac{y^2}{2} - 2x$$

11. a) Calculer les dérivées partielles d'ordre 1 de F en tout point (x, y) de $]0, +\infty[^2$.

Démonstration.

- La fonction f est de classe $\mathcal{C}^1$ sur $]0, +\infty[\times]0, +\infty[$ car elle est polynomiale.
Elle admet donc des dérivées partielles premières sur cet ensemble.
- Soit $(x, y) \in]0, +\infty[\times]0, +\infty[$. Tout d'abord :

$$\partial_1(F)(x, y) \ = \ 2y\, x + 2x - 2$$

Par ailleurs :

$$\partial_2(F)(x, y) \ = \ x^2 - y$$

$$\boxed{\partial_1(F)(x, y) \ = \ 2xy + 2x - 2 \quad \text{et} \quad \partial_2(F)(x, y) \ = \ x^2 - y}$$

□

b) Montrer que la fonction F admet $(u_3, {u_3}^2)$ comme unique point critique, où le réel u_3 est l'unique solution sur $\mathbb{R}_+$ de l'équation (E_3) définie sans la partie **B**.

Démonstration.
Soit $(x, y) \in]0, +\infty[^2$.

- Rappelons tout d'abord :

$$(x, y) \text{ est un point critique de } F \ \Leftrightarrow \ \nabla(F)(x, y) = 0_{\mathscr{M}_{2,1}(\mathbb{R})} \ \Leftrightarrow \ \left\{ \begin{array}{rcl} \partial_1(F)(x, y) & = & 0 \\ \partial_2(F)(x, y) & = & 0 \end{array} \right.$$

Ainsi : (x,y) est un point critique de F $\Leftrightarrow \begin{cases} 2xy+2x-2 &= 0 \\ x^2-y &= 0 \end{cases}$

$\Leftrightarrow \begin{cases} 2\,(xy+x-1) &= 0 \\ y &= x^2 \end{cases}$

$\Leftrightarrow \begin{cases} x^3+x-1 &= 0 \\ y &= x^2 \end{cases}$ *(en remplaçant y par x^2 dans la première équation)*

$\Leftrightarrow \begin{cases} x &= u_3 \\ y &= x^2 \end{cases}$ *(car l'unique solution positive de l'équation (E_3) est u_3)*

On en déduit que la fonction F admet comme unique point critique le point de coordonnées $(u_3, {u_3}^2)$.

Commentaire

- La difficulté de cette question réside dans le fait qu'il n'existe pas de méthode générale pour résoudre l'équation $\nabla(f)(x,y) = 0_{\mathscr{M}_{2,1}(\mathbb{R})}$.
  On est donc confronté à une question bien plus complexe qu'une résolution de système d'équations linéaires (que l'on résout aisément à l'aide de la méthode du pivot de Gauss).
- Lors de la recherche de points critiques, on doit faire appel à des méthodes ad hoc. Ici, on fait apparaître une équation du type :

  $$y = \psi(x)$$

  En injectant cette égalité dans la première équation, on obtient une nouvelle équation qui ne dépend plus que d'une variable et qu'il est donc plus simple de résoudre. Plus précisément, on fait apparaître l'équation $h_3(x) = 0$ qu'on a déjà résolue en question ***6***. □

12. a) Écrire la matrice hessienne, notée H, de la fonction F au point (u_3, u_3^2).

Démonstration.

- La fonction F est de classe $\mathcal{C}^2$ sur $]0,+\infty[\times]0,+\infty[$ car elle est polynomiale.
  Elle admet donc des dérivées partielles d'ordre 2 sur $]0,+\infty[\times]0,+\infty[$.
- Soit $(x,y) \in]0,+\infty[\times]0,+\infty[$. Tout d'abord :

  $$\partial^2_{11}(F)(x,y) \;=\; 2y+2$$

- Ensuite :

  $$\partial^2_{12}(F)(x,y) \;=\; 2x \;=\; \partial^2_{21}(F)(x,y)$$

  La dernière égalité est obtenue en vertu du théorème de Schwarz puisque la fonction F est $\mathcal{C}^2$ sur l'**ouvert** $]0,+\infty[\times]0,+\infty[$.

- Enfin :
$$\partial^2_{22}(F)(x,y) \;=\; -1$$

Pour tout $(x,y) \in]0,+\infty[\times]0,+\infty[$, $\partial^2_{11}(F)(x,y) \;=\; 2y+2$,
$\partial^2_{12}(F)(x,y) \;=\; 2x \;=\; \partial^2_{21}(F)(x,y)$ et $\partial^2_{22}(F)(x,y) \;=\; -1$.

Commentaire

- Il faut penser à utiliser le théorème de Schwarz dès que la fonction à deux variables considérée est $\mathcal{C}^2$ sur un ouvert $U \subset \mathbb{R}^2$.
- Ici, le calcul de $\partial^2_{12}(F)(x,y)$ et $\partial^2_{21}(F)(x,y)$ est aisé. Il faut alors concevoir le résultat du théorème de Schwarz comme une mesure de vérification : en dérivant par rapport à la $1^{\text{ère}}$ variable puis par rapport à la $2^{\text{ème}}$, on doit obtenir le même résultat que dans l'ordre inverse.

- On rappelle que la matrice hessienne de F en un point $(x,y) \in]0,+\infty[\times]0,+\infty[$ est :
$$\nabla^2(F)(x,y) = \begin{pmatrix} \partial^2_{11}(F)(x,y) & \partial^2_{12}(F)(x,y) \\ \partial^2_{21}(F)(x,y) & \partial^2_{22}(F)(x,y) \end{pmatrix} = \begin{pmatrix} 2y+2 & 2x \\ 2x & -1 \end{pmatrix}$$
- On en déduit :
$$H = \nabla^2(F)(u_3, {u_3}^2) = \begin{pmatrix} 2{u_3}^2+2 & 2u_3 \\ 2u_3 & -1 \end{pmatrix}$$

$$\boxed{H = \begin{pmatrix} 2{u_3}^2+2 & 2u_3 \\ 2u_3 & -1 \end{pmatrix}}$$

□

b) Montrer que la matrice H admet deux valeurs propres distinctes, notées λ_1 et λ_2, vérifiant :
$$\lambda_1 \, \lambda_2 \;=\; -6u_3^2 - 2$$

Démonstration.

- Soit $\lambda \in \mathbb{R}$.
$$\begin{aligned} \det(H - \lambda \, I_2) &= \det \begin{pmatrix} 2{u_3}^2+2-\lambda & 2u_3 \\ 2u_3 & -1-\lambda \end{pmatrix} = (2{u_3}^2+2-\lambda)(-1-\lambda) - 4\,{u_3}^2 \\ &= (-2{u_3}^2-2) + \lambda - (2{u_3}^2+2)\,\lambda + \lambda^2 - 4\,{u_3}^2 \\ &= \lambda^2 - (2{u_3}^2+1)\,\lambda + (-6{u_3}^2-2) \end{aligned}$$

On en déduit :
$$\begin{aligned} \lambda \text{ est valeur propre de } H &\Leftrightarrow H - \lambda \, I_2 \text{ non inversible} \\ &\Leftrightarrow \det(H - \lambda \, I_2) = 0 \\ &\Leftrightarrow \lambda^2 - (2{u_3}^2+1)\,\lambda + (-6{u_3}^2-2) \\ &\Leftrightarrow \lambda \text{ est racine de } Q \end{aligned}$$

où Q est le polynôme de degré 2 défini par :
$$Q(X) \;=\; X^2 - (2{u_3}^2+1)\,X + (-6{u_3}^2-2)$$

Commentaire

À ce stade de l'étude de la nature d'un point critique, le calcul de $\det(H - \lambda\, I_2)$ nous fournit toujours un polynôme de degré 2 en λ, noté ici Q. Deux cas se présentent alors :

× l'expression de Q est « simple » (c'est par exemple le cas lorsque les coefficients de Q sont numériques). Dans ce cas :

1) on détermine explicitement les racines de Q par factorisation ou calcul de discriminant.

2) les racines de Q sont les valeurs propres de H d'après les équivalences ci-dessus.

3) on en déduit le signe des valeurs propres de H et ainsi la nature du point critique étudié.

× l'expression de Q est « compliquée » (c'est par exemple le cas lorsque l'expression de Q dépend de plusieurs paramètres, comme ici). Dans ce cas, **on ne cherchera pas** à déterminer les racines de Q explicitement. On procèdera de la manière suivante :

1) on justifie l'existence de valeurs propres λ_1 et λ_2 de H (la matrice H est symétrique).

2) les valeurs propres de H sont racines de Q d'après les équivalences ci-dessus.
On en déduit la factorisation de Q suivante : $Q(X) = (X - \lambda_1)\,(X - \lambda_2)$.

3) on identifie les coefficients des deux expressions de Q pour en déduire des relations sur λ_1 et λ_2 (elles sont appelées *relations coefficients / racines*).

4) on détermine le signe de λ_1 et λ_2 (valeurs propres de H) grâce à ces relations, et on obtient ainsi la nature du point critique étudié.

- La matrice H est une matrice symétrique (réelle). Elle est donc diagonalisable. On note λ_1 et λ_2 ses valeurs propres **éventuellement égales**.
- D'après les équivalences précédentes, on en déduit que λ_1 et λ_2 sont racines de Q. Ainsi :

$$\begin{aligned} Q(X) &= (X - \lambda_1)\,(X - \lambda_2) \\ &= X^2 - (\lambda_1 + \lambda_2)\,X + \lambda_1\,\lambda_2 \end{aligned}$$

D'où, par définition de Q :

$$X^2 - (2{u_3}^2 + 1)\,X + (-6{u_3}^2 - 2) \;=\; X^2 - (\lambda_1 + \lambda_2)\,X + \lambda_1\,\lambda_2$$

Par identification des coefficients de ces polynômes de degré 2, on en déduit le système d'équations suivant :

$$\begin{cases} \lambda_1 + \lambda_2 &= 2{u_3}^2 + 1 \\ \lambda_1\,\lambda_2 &= -6{u_3}^2 - 2 \end{cases}$$

- Montrons maintenant que λ_1 et λ_2 sont distinctes.
Raisonnons par l'absurde. Supposons alors : $\lambda_1 = \lambda_2$.
D'après le système précédent, on obtient en particulier :

$$\lambda_1^2 \;=\; \lambda_1\,\lambda_2 \;=\; -6u_3^2 - 2 \leqslant -2 < 0$$

Absurde !

Ainsi, H admet deux valeurs propres distinctes λ_1 et λ_2 qui vérifient : $\lambda_1\,\lambda_2 \;=\; -6u_3^2 - 2$. □

13. La fonction F présente-t-elle des extrema locaux sur $]0,+\infty[^2$?

Démonstration.

- Rappelons tout d'abord que les extremaux locaux sur $]0,+\infty[\times]0,+\infty[$ sont parmi les points critiques de la fonction F.
  Or, d'après la question ***11.b)***, la fonction F admet $(u_3, {u_3}^2)$ comme unique point critique.

  Le point $(u_3, {u_3}^2)$ est donc le seul extremum local possible.

- On a montré dans la question précédente : $\lambda_1 \lambda_2 = -6u_3^2 - 2 < 0$.
  Ainsi, les valeurs propres de H sont de signes opposés.

Le point $(u_3, {u_3}^2)$ n'est donc pas un extremum local de F mais un point selle.
Ainsi, la fonction F ne présente pas d'extrema locaux sur $]0,+\infty[\times]0,+\infty[$

□

Exercice 2

On définit, pour tous réels a et b, $M(a,b)$ la matrice carrée d'ordre 4 par :

$$M(a,b) = \begin{pmatrix} a & 0 & 0 & a \\ a & 0 & 0 & a \\ a & 0 & 0 & a \\ b & b & b & b \end{pmatrix}$$

et on note : $E = \left\{M(a,b) \mid (a,b) \in \mathbb{R}^2\right\}$.
L'objectif de cet exercice est de déterminer les matrices de E qui sont diagonalisables.

1. a) Montrer que E est un sous-espace vectoriel de $\mathscr{M}_4(\mathbb{R})$.
Déterminer une base de E et sa dimension.

Démonstration.

- Tout d'abord :

$$\begin{aligned} E &= \{M(a,b) \mid (a,b) \in \mathbb{R}^2\} \\ &= \left\{ \begin{pmatrix} a & 0 & 0 & a \\ a & 0 & 0 & a \\ a & 0 & 0 & a \\ b & b & b & b \end{pmatrix} \mid (a,b) \in \mathbb{R}^2 \right\} \\ &= \left\{ a \cdot \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \mid (a,b) \in \mathbb{R}^2 \right\} \\ &= \left\{a \cdot M(1,0) + b \cdot M(0,1) \mid (a,b) \in \mathbb{R}^2\right\} \\ &= \text{Vect}\left(M(1,0), M(0,1)\right) \end{aligned}$$

On en déduit que E est un sous-espace vectoriel de $\mathscr{M}_4(\mathbb{R})$.

Commentaire

- Ici, E est un ensemble dont les éléments sont des matrices écrites à l'aide de paramètres. Il y a tout lieu de penser que cet ensemble va pouvoir facilement s'écrire comme espace vectoriel engendré par une partie.
- Cette méthode présente un double avantage. En effet, en plus de démontrer que E est un sous-espace vectoriel de $\mathscr{M}_4(\mathbb{R})$, on obtient de plus que famille $\big(M(1,0), M(0,1)\big)$ est génératrice de E.

- La famille $\big(M(1,0), M(0,1)\big)$ est :
  - × génératrice de E d'après le point précédent,
  - × libre car uniquement constituée de deux vecteurs (de $\mathscr{M}_4(\mathbb{R})$) non colinéaires.

On en déduit que $\big(M(1,0), M(0,1)\big)$ est une base de E.

- De plus :

$$\dim(E) = \text{Card}((M(1,0), M(0,1))) = 2$$

$$\boxed{\dim(E) = 2}$$

Commentaire

Si la rédaction précédente est celle attendue, il arrive parfois qu'on ne puisse pas l'utiliser (dans certains cas, l'ensemble étudié ne s'écrit pas naturellement comme espace vectoriel engendré par une partie). Il est donc important de savoir utiliser la méthode consistant à revenir à la définition de sous-espace vectoriel de $\mathscr{M}_4(\mathbb{R})$. Rappelons ci-dessous la rédaction.

(i) $E \subset \mathscr{M}_4(\mathbb{R})$.

(ii) $E \neq \varnothing$ car $0_{\mathscr{M}_4(\mathbb{R})} = M(0,0) \in E$.

(iii) Démontrons que E est stable par combinaison linéaire.

Soit $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ et soit $(U_1, U_2) \in E^2$.

× Comme $U_1 \in E$, il existe $(a_1, b_1) \in \mathbb{R}^2$ tel que $U_1 = M(a_1, b_1)$.

× Comme $U_2 \in E$, il existe $(a_2, b_2) \in \mathbb{R}^2$ tel que $U_2 = M(a_2, b_2)$.

Démontrons que $\lambda_1 \cdot U_1 + \lambda_2 \cdot U_2 \in E$. On a :

$$\begin{aligned}
\lambda_1 \cdot U_1 + \lambda_2 \cdot U_2 &= \lambda_1 \cdot M(a_1, b_1) + \lambda_2 \cdot M(a_2, b_2) \\
&= \lambda_1 \cdot \begin{pmatrix} a_1 & 0 & 0 & a_1 \\ a_1 & 0 & 0 & a_1 \\ a_1 & 0 & 0 & a_1 \\ b_1 & b_1 & b_1 & b_1 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} a_2 & 0 & 0 & a_2 \\ a_2 & 0 & 0 & a_2 \\ a_2 & 0 & 0 & a_2 \\ b_2 & b_2 & b_2 & b_2 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 a_1 & 0 & 0 & \lambda_1 a_1 \\ \lambda_1 a_1 & 0 & 0 & \lambda_1 a_1 \\ \lambda_1 a_1 & 0 & 0 & \lambda_1 a_1 \\ \lambda_1 b_1 & \lambda_1 b_1 & \lambda_1 b_1 & \lambda_1 b_1 \end{pmatrix} + \begin{pmatrix} \lambda_2 a_2 & 0 & 0 & \lambda_2 a_2 \\ \lambda_2 a_2 & 0 & 0 & \lambda_2 a_2 \\ \lambda_2 a_2 & 0 & 0 & \lambda_2 a_2 \\ \lambda_2 b_2 & \lambda_2 b_2 & \lambda_2 b_2 & \lambda_2 b_2 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 a_1 + \lambda_2 a_2 & 0 & 0 & \lambda_1 a_1 + \lambda_2 b_2 \\ \lambda_1 a_1 + \lambda_2 a_2 & 0 & 0 & \lambda_1 a_1 + \lambda_2 b_2 \\ \lambda_1 a_1 + \lambda_2 a_2 & 0 & 0 & \lambda_1 a_1 + \lambda_2 b_2 \\ \lambda_1 b_1 + \lambda_2 b_2 & \lambda_1 b_1 + \lambda_2 b_2 & \lambda_1 b_1 + \lambda_2 b_2 & \lambda_1 b_1 + \lambda_2 b_2 \end{pmatrix}
\end{aligned}$$

D'où : $\lambda_1 \cdot U_1 + \lambda_2 \cdot U_2 = M(\lambda_1 a_1 + \lambda_2 a_2, \lambda_1 b_1 + \lambda_2 b_2) \in E$.

E est un sous-espace vectoriel de $\mathscr{M}_4(\mathbb{R})$.

□

b) Le produit de deux matrices quelconques de E appartient-il encore à E ?

Démonstration.

- On remarque :

$$M(1,0) \times M(0,1) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Ainsi :

× $M(1,0) \in E$ et $M(0,1) \in E$,

× $M(1,0) \times M(0,1) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \notin E$.

L'ensemble E n'est donc pas stable par produit.

□

2. Étude du cas $a = 0$ et $b = 0$.
Justifier que la matrice $M(0,0)$ est diagonalisable.

Démonstration.
La matrice $M(0,0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ est diagonale.

On en déduit que $M(0,0)$ est diagonalisable. □

3. Étude du cas $a \neq 0$ et $b = 0$.
Soit a un réel non nul. On note A la matrice $M(a,0)$.

a) Calculer A^2 et déterminer un polynôme annulateur de A.

Démonstration.

- On remarque tout d'abord : $A = a \cdot M(1,0)$. Ainsi :

$$A^2 = \big(a \cdot M(1,0)\big) \times \big(a \cdot M(1,0)\big) = a^2 \cdot \big(M(1,0)\big)^2$$

- Or :

$$\big(M(1,0)\big)^2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = M(1,0)$$

Finalement : $A^2 = a \cdot A$.

- On en déduit : $A^2 - a \cdot A = 0_{\mathscr{M}_4(\mathbb{R})}$.

Ainsi, le polynôme $Q(X) = X^2 - aX$ est un polynôme annulateur de A. □

b) En déduire les valeurs propres de la matrice A et préciser une base de chacun des sous-espaces propres associés.

Démonstration.

- D'après la question précédente, $Q(X) = X^2 - aX = X(X-a)$ est un polynôme annulateur de A. D'où : $\mathrm{Sp}(A) \subset \{\text{racines de } Q\} = \{0, a\}$.

Ainsi : $\mathrm{Sp}(A) \subset \{0, a\}$ et 0 et a sont les deux valeurs propres possibles de A.

- La matrice A possède une ligne de 0 donc A n'est pas inversible.

On en déduit que 0 est bien une valeur propre de A.

- Démontrons que a est valeur propre de A.

$$A - a \cdot I_4 = \begin{pmatrix} 0 & 0 & 0 & a \\ a & -a & 0 & a \\ a & 0 & -a & a \\ 0 & 0 & 0 & -a \end{pmatrix}$$

On remarque : $C_1 + C_2 + C_3 = 0$ (où on a noté C_i la $i^{\text{ème}}$ colonne de A).
La famille des vecteurs colonnes de $A - a \cdot I_4$ est donc liée.

Ainsi, $A - a \cdot I_4$ est non inversible. On en déduit que a est bien une valeur propre de A.

Commentaire

- On démontre que la matrice $A - a \cdot I_4$ est non inversible en exhibant une relation de dépendance linéaire non triviale entre les colonnes de cette matrice. Il est aussi possible d'exhiber une relation de dépendance linéaire non triviale entre les lignes. En l'occurrence, on a : $L_4 = -L_1$.
- Il est aussi possible d'effectuer un calcul du rang. En procédant par opérations élémentaires successives, on obtient une réduite triangulaire de même rang que la matrice initiale. En particulier, matrice initiale et réduite sont toutes les deux inversibles ou toutes les deux non inversibles. La réduite étant triangulaire, elle est non inversible si et seulement si elle possède (au moins) un coefficient diagonal nul, ce qui permet de conclure quant au caractère inversible (ou non) de la matrice initiale.

$$\begin{aligned}
\operatorname{rg}(A - a \cdot I_4) &= \operatorname{rg}\left(\begin{pmatrix} 0 & 0 & 0 & a \\ a & -a & 0 & a \\ a & 0 & -a & a \\ 0 & 0 & 0 & -a \end{pmatrix}\right) \\
&\overset{L_1 \leftrightarrow L_3}{=} \operatorname{rg}\left(\begin{pmatrix} a & 0 & -a & a \\ a & -a & 0 & a \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & -a \end{pmatrix}\right) \overset{L_2 \leftarrow L_2 - L_1}{=} \operatorname{rg}\left(\begin{pmatrix} a & 0 & -a & a \\ 0 & -a & a & 0 \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & -a \end{pmatrix}\right)
\end{aligned}$$

La réduite obtenue, triangulaire (supérieure), possède un coefficient diagonal nul.
Elle (et la matrice initiale $A - a \cdot I_4$) est donc non inversible.

- Déterminons $E_0(A)$, le sous-espace propre de A associé à la valeur propre 0.

Soit $X \in \mathscr{M}_{4,1}(\mathbb{R})$. Il existe donc $(x, y, z, t) \in \mathbb{R}^4$ tel que $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$.

$$\begin{aligned}
X \in E_0(A) &\iff A\,X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff a \cdot M(1,0)\,X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff M(1,0)\,X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \qquad \textit{(car } a \neq 0\textit{)} \\
&\iff \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{cases} x & & + & t & = & 0 \\ x & & + & t & = & 0 \\ x & & + & t & = & 0 \\ & & & 0 & = & 0 \end{cases} \\
&\overset{\substack{L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1}}{\iff} \begin{cases} x & & + & t & = & 0 \\ & & & 0 & = & 0 \\ & & & 0 & = & 0 \end{cases} \\
&\iff \{\ x = -t
\end{aligned}$$

On en déduit :

$$\begin{aligned} E_0(A) &= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid x = -t \right\} = \left\{ \begin{pmatrix} -t \\ y \\ z \\ t \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid (y,z,t) \in \mathbb{R}^3 \right\} \\ &= \left\{ t \cdot \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid (y,z,t) \in \mathbb{R}^3 \right\} \\ &= \text{Vect}\left(\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right) \end{aligned}$$

La famille $\mathcal{F}_0 = \left(\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$ est :

× génératrice de $E_0(A)$ (d'après ce qui précède).

× libre.

Ainsi, la famille $\left(\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$ est une base de $E_0(A)$.

Commentaire

- On ne développe pas ici la démonstration de la liberté de la famille $\mathcal{F}_0$ car il ne s'agit pas du cœur de la question. De plus, à la lecture des vecteurs de cette famille, il apparaît assez évident qu'il n'existe pas de relation de dépendance linéaire, autre que la triviale, entre ces vecteurs. Il est toutefois primordial de connaître la rédaction formelle permettant de démontrer la liberté d'une famille. C'est pourquoi on la rappelle ci-dessous.
- Démontrons que la famille $\mathcal{F}_0$ est libre.
  Soit $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$.
  Supposons : $\lambda_1 \cdot \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ $(*)$.

  $$\begin{aligned} \text{Or :} \quad (*) &\iff \begin{pmatrix} -\lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff \left\{ \lambda_1 = \lambda_2 = \lambda_3 = 0 \right. \end{aligned}$$

  Ainsi, la famille $\mathcal{F}_0$ est libre.

- Déterminons $E_a(A)$ le sous-espace propre de A associé à la valeur propre a.

Soit $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R})$.

$$
\begin{aligned}
X \in E_a(A) &\iff (A - a \cdot I_4)\, X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff (a \cdot M(1,0) - a \cdot I_4)\, X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff a \cdot (M(1,0) - I_4)\, X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff (M(1,0) - I_4)\, X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \qquad \textit{(car } a \neq 0\textit{)} \\
&\iff \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{cases} t = 0 \\ x - y + t = 0 \\ x - z + t = 0 \\ -t = 0 \end{cases} \\
&\overset{L1 \leftrightarrow L_3}{\iff} \begin{cases} x - z + t = 0 \\ x - y + t = 0 \\ t = 0 \\ -t = 0 \end{cases} \\
&\overset{L_2 \leftarrow L_2 - L_1}{\iff} \begin{cases} x - z + t = 0 \\ -y + z = 0 \\ t = 0 \\ -t = 0 \end{cases} \\
&\iff \begin{cases} x + t = z \\ -y = -z \\ t = 0 \\ -t = 0 \end{cases} \\
&\iff \begin{cases} x + t = z \\ -y = -z \\ t = 0 \end{cases} \\
&\overset{L_1 \leftarrow L_1 - L_3}{\iff} \begin{cases} x = z \\ -y = -z \\ t = 0 \end{cases}
\end{aligned}
$$

On en déduit :

$$
\begin{aligned}
E_a(A) &= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid x = z \text{ et } y = z \text{ et } t = 0 \right\} = \left\{ \begin{pmatrix} z \\ z \\ z \\ 0 \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid z \in \mathbb{R} \right\} \\
&= \left\{ z \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \mid z \in \mathbb{R} \right\} = \operatorname{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)
\end{aligned}
$$

La famille $\mathcal{F}_a = \left(\begin{pmatrix}1\\1\\1\\0\end{pmatrix}\right)$ est :

× génératrice de $E_a(A)$ (d'après ce qui précède).

× libre car constituée uniquement d'un vecteur non nul.

Ainsi, la famille $\left(\begin{pmatrix}1\\1\\1\\0\end{pmatrix}\right)$ est une base de $E_a(A)$.

Commentaire

- Il faut s'habituer à déterminer les ensembles $E_\lambda(A)$ par lecture de la matrice $A - \lambda I$.
- Illustrons la méthode avec la matrice de l'exercice et $\lambda = a$.

  On cherche les vecteurs $X = \begin{pmatrix}x\\y\\z\\t\end{pmatrix}$ de $E_a(A)$ c'est-à-dire les vecteurs tels que :

  $(M(1,0) - I_4)\,X = 0_{\mathscr{M}_{4,1}(\mathbb{R})}$ (d'après ce qui précède). Or :

$$\begin{pmatrix}0&0&0&1\\1&-1&0&1\\1&0&1&1\\0&0&0&-1\end{pmatrix}\begin{pmatrix}x\\y\\z\\t\end{pmatrix} = x\cdot C_1 + y\cdot C_2 + z\cdot C_3 + t\cdot C_4$$

$$= x\cdot\begin{pmatrix}0\\1\\1\\0\end{pmatrix} + y\cdot\begin{pmatrix}0\\-1\\0\\0\end{pmatrix} + z\cdot\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + t\cdot\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}$$

  Pour obtenir le vecteur $\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$ à l'aide de cette combinaison linéaire, on doit forcément choisir $t = 0$. En effet, si t est non nul, les coefficients en $4^{\text{ème}}$ position de la combinaison linéaire sont non nuls. On prend alors $t = 0$.

  La combinaison linéaire restante est nulle si $x - y = 0$ (obligatoire pour éliminer le $2^{\text{ème}}$ coefficient) et $x - z = 0$ (obligatoire pour éliminer le $3^{\text{ème}}$ coefficient). Finalement, il faut nécessairement vérifier $x = y = z$ pour obtenir le vecteur nul.

  En prenant par exemple $x = y = z = 1$, on obtient : $E_a(A) \supset \text{Vect}\left(\begin{pmatrix}1\\1\\1\\0\end{pmatrix}\right)$.

  Et l'égalité est vérifiée car ces deux espaces vectoriels sont de même dimension. □

c) En déduire que la matrice A est diagonalisable. Déterminer une matrice P de $\mathscr{M}_4(\mathbb{R})$ inversible et une matrice D de $\mathscr{M}_4(\mathbb{R})$ diagonale telles que : $A = PDP^{-1}$.

Démonstration.

- D'après la question précédente :

  × la famille $\mathcal{F}_0$ est une base de $E_0(A)$. Ainsi :

$$\dim\left(E_0(A)\right) = \text{Card}(\mathcal{F}_0) = 3$$

  × la famille $\mathcal{F}_a$ est une base de $E_a(A)$. Ainsi :

$$\dim\left(E_a(A)\right) = \text{Card}(\mathcal{F}_a) = 1$$

- On en déduit :
$$\dim\big(E_0(A)\big) + \dim\big(E_a(A)\big) \;=\; 4$$
Or 0 et a sont les seules valeurs propres de A et $A \in \mathscr{M}_4(\mathbb{R})$.

On en déduit que A est diagonalisable.

- Il existe donc une matrice P inversible et une matrice D diagonale telles que $A = P\,D\,P^{-1}$. Plus précisément :
  - × la matrice P est obtenue par concaténation de bases des sous-espaces propres de A,
  - × la matrice D est la matrice diagonale dont les coefficients diagonaux sont les valeurs propres de A (dans le même ordre d'apparition que les vecteurs propres).

En posant $P = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ et $D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$, on a donc : $A = PDP^{-1}$.

□

4. **Étude du cas $a = 0$ et $b \neq 0$.**
Soit b un réel non nul. On note B la matrice $M(0, b)$.

a) Déterminer le rang des matrices B et $B - b\,I_4$, I_4 désignant la matrice identité d'ordre 4.

Démonstration.

- Tout d'abord :
$$\begin{aligned}
\mathrm{rg}(B) &= \mathrm{rg}\left(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ b & b & b & b \end{pmatrix}\right) \\
&= \mathrm{rg}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ b \end{pmatrix}\right) \\
&= \mathrm{rg}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ b \end{pmatrix}\right) \\
&= \mathrm{rg}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right) && \textit{(car } b \neq 0\textit{)} \\
&= 1 && \textit{(car la famille } \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right) \textit{ est libre)}
\end{aligned}$$

Finalement : $\mathrm{rg}(B) = 1$.

- Ensuite :

$$\begin{aligned}
\operatorname{rg}(B - b\cdot I_4) &= \operatorname{rg}\left(\begin{pmatrix} -b & 0 & 0 & 0 \\ 0 & -b & 0 & 0 \\ 0 & 0 & -b & 0 \\ b & b & b & 0 \end{pmatrix}\right) = \operatorname{rg}\left(\begin{pmatrix} -b \\ 0 \\ 0 \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ -b \\ 0 \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -b \\ b \end{pmatrix}\right) \\
&= \operatorname{rg}\left(\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}\right) \qquad \textit{(car } b \neq 0\textit{)} \\
&= 3 \qquad \textit{(car la famille précédente est libre)}
\end{aligned}$$

$$\boxed{\operatorname{rg}(B - b\cdot I_4) = 3}$$

□

b) En déduire l'ensemble des valeurs propres de B en précisant la dimension des sous-espaces propres associés.

Démonstration.

- D'après la question précédente, $\operatorname{rg}(B) = 1 < 4$.
  Ainsi, B est non inversible et donc $0 \in \operatorname{Sp}(B)$.
  De plus, d'après le théorème du rang :

$$\begin{array}{ccccc} \dim\big(E_0(B)\big) & + & \operatorname{rg}(B) & = & \dim\big(\mathscr{M}_{4,1}(\mathbb{R})\big) \\ & & \shortparallel & & \shortparallel \\ & & 1 & & 4 \end{array}$$

  On en conclut que 0 est valeur propre de B et : $\dim\big(E_0(B)\big) = 3$.

- D'après la question précédente, $\operatorname{rg}(B - b\cdot I_4) = 3 < 4$.
  Ainsi, $B - b\cdot I_4$ est non inversible et donc $b \in \operatorname{Sp}(B)$.
  De plus, d'après le théorème du rang :

$$\begin{array}{ccccc} \dim\big(E_b(B)\big) & + & \operatorname{rg}(B - b\cdot I_4) & = & \dim\big(\mathscr{M}_{4,1}(\mathbb{R})\big) \\ & & \shortparallel & & \shortparallel \\ & & 3 & & 4 \end{array}$$

  On en conclut que b est valeur propre de B et : $\dim\big(E_b(B)\big) = 1$.

Commentaire

On peut aussi remarquer que la matrice B est triangulaire (inférieure). Ainsi, ses valeurs propres sont ses coefficients diagonaux. Cela permet de conclure directement :

$$\operatorname{Sp}(B) = \{0, b\}$$

- Comme B est une matrice carrée d'ordre 4, on sait d'après le cours que la somme des dimensions de toutes ses sous-espaces propres est majorée par 4. Or :

$$\dim\big(E_0(B)\big) + \dim\big(E_b(B)\big) = 3 + 1 = 4$$

On en déduit que la matrice B n'admet pas de valeur propre autre que 0 et b. □

c) La matrice B est-elle diagonalisable ?

Démonstration.
D'après la question précédente :

$$\dim(E_0(B)) + \dim(E_b(B)) \;=\; 4$$

Or 0 et b sont les seules valeurs propres de B et $B \in \mathscr{M}_4(\mathbb{R})$.

On en déduit que B est diagonalisable.

□

5. **Étude du cas $a \neq 0$ et $b \neq 0$.**
Soient a et b deux réels non nuls. On note f l'endomorphisme de $\mathbb{R}^4$ dont la matrice dans la base canonique de $\mathbb{R}^4$ est $M(a,b)$.
On pose :

$$v_1 \;=\; (1,1,1,0), \qquad v_2 \;=\; (0,0,0,1) \qquad \text{et} \qquad T \;=\; \begin{pmatrix} a & a \\ 3b & b \end{pmatrix}$$

a) Montrer que $\mathrm{Ker}(f)$ est de dimension 2 et préciser une base (v_3, v_4) de $\mathrm{Ker}(f)$.

Démonstration.
On note $\mathscr{B}$ la base canonique de $\mathbb{R}^4$.

- Soit $u = (x,y,z,t) \in \mathbb{R}^4$. On note : $U = \mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$.

$$\begin{aligned}
u \in \mathrm{Ker}(f) \quad &\iff \quad f(u) = 0_{\mathbb{R}^4} \\
&\iff \quad M(a,b)\,U = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\
&\iff \quad \begin{pmatrix} a & 0 & 0 & a \\ a & 0 & 0 & a \\ a & 0 & 0 & a \\ b & b & b & b \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \quad \left\{ \begin{array}{rcl} a\,x + a\,t &=& 0 \\ a\,x + a\,t &=& 0 \\ a\,x + a\,t &=& 0 \\ b\,x + b\,y + b\,z + b\,t &=& 0 \end{array} \right. \\
&\overset{\substack{L_1 \leftarrow \frac{1}{a} L_1 \\ L_2 \leftarrow \frac{1}{a} L_2 \\ L_3 \leftarrow \frac{1}{a} L_3 \\ L_4 \leftarrow \frac{1}{b} L_4}}{\iff} \quad \left\{ \begin{array}{rcl} x + t &=& 0 \\ x + t &=& 0 \\ x + t &=& 0 \\ x + y + z + t &=& 0 \end{array} \right. \qquad \textit{(avec } a \neq 0 \textit{ et } b \neq 0\textit{)} \\
&\overset{\substack{L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1 \\ L_4 \leftarrow L_4 - L_1}}{\iff} \quad \left\{ \begin{array}{rcl} x + t &=& 0 \\ 0 &=& 0 \\ 0 &=& 0 \\ y + z &=& 0 \end{array} \right. \\
&\iff \quad \left\{ \begin{array}{rcl} x &=& -t \\ y &=& -z \end{array} \right.
\end{aligned}$$

On en déduit :

$$\begin{aligned}\mathrm{Ker}(f) &= \{(x,y,z,t)\in\mathbb{R}^4 \mid x=-t \text{ et } y=-z\}\\ &= \{(-t,-z,z,t)\in\mathbb{R}^4 \mid (z,t)\in\mathbb{R}^2\}\\ &= \{t\cdot(-1,0,0,1)+z\cdot(0,-1,1,0) \mid (z,t)\in\mathbb{R}^2\}\\ &= \mathrm{Vect}\left((-1,0,0,1),(0,-1,1,0)\right)\end{aligned}$$

- Ainsi, la famille $((-1,0,0,1),(0,-1,1,0))$ est :
  - × génératrice de $\mathrm{Ker}(f)$ (d'après ce qui précède).
  - × libre car constituée uniquement de deux vecteurs non colinéaires.

Ainsi, en posant $v_3=(-1,0,0,1)$ et $v_4=(0,-1,1,0)$, on obtient que la famille (v_3,v_4) est une base de $\mathrm{Ker}(f)$.

Finalement : $\dim\big(\mathrm{Ker}(f)\big)=\mathrm{Card}\left((v_3,v_4)\right)=2$

□

b) Montrer que la famille $\mathscr{B}'=(v_1,v_2,v_3,v_4)$ est une base de $\mathbb{R}^4$.

Démonstration.

- Démontrons que la famille $\mathscr{B}'$ est libre.
  Soit $(\lambda_1,\lambda_2,\lambda_3,\lambda_4)\in\mathbb{R}^4$. Supposons : $\lambda_1\cdot v_1+\lambda_2\cdot v_2+\lambda_3\cdot v_3+\lambda_4\cdot v_4=0_{\mathbb{R}^4}$ $(*)$.

Or : $(*) \iff \lambda_1\cdot v_1+\lambda_2\cdot v_2+\lambda_3\cdot v_3+\lambda_4\cdot v_4=0_{\mathbb{R}^4}$

$\iff (\lambda_1-\lambda_3,\lambda_1-\lambda_4,\lambda_1+\lambda_4,\lambda_2+\lambda_3)=(0,0,0,0)$

$$\iff \begin{cases} \lambda_1 - \lambda_3 = 0\\ \lambda_1 - \lambda_4 = 0\\ \lambda_1 + \lambda_4 = 0\\ \lambda_2 + \lambda_3 = 0\end{cases}$$

$$\overset{\substack{L_2\leftarrow L_2-L_1\\ L_3\leftarrow L_3-L_1}}{\iff} \begin{cases} \lambda_1 - \lambda_3 = 0\\ \lambda_3 - \lambda_4 = 0\\ \lambda_3 + \lambda_4 = 0\\ \lambda_2 + \lambda_3 = 0\end{cases}$$

$$\overset{L_2\leftrightarrow L_4}{\iff} \begin{cases} \lambda_1 - \lambda_3 = 0\\ \lambda_2 + \lambda_3 = 0\\ \lambda_3 + \lambda_4 = 0\\ \lambda_3 - \lambda_4 = 0\end{cases}$$

$$\overset{L_4\leftarrow L_4-L_3}{\iff} \begin{cases} \lambda_1 - \lambda_3 = 0\\ \lambda_2 + \lambda_3 = 0\\ \lambda_3 + \lambda_4 = 0\\ -2\lambda_4 = 0\end{cases}$$

$\iff \{\lambda_1=\lambda_2=\lambda_3=\lambda_4=0$

(par remontées successives)

Ainsi, $\mathscr{B}'$ est une famille libre de $\mathbb{R}^4$.

- On a alors :
  - × la famille $\mathscr{B}' = (v_1, v_2, v_3, v_4)$ est une famille libre,
  - × $\text{Card}(\mathscr{B}') = 4 = \dim(\mathbb{R}^4)$.

On en déduit que $\mathscr{B}'$ est une base de $\mathbb{R}^4$. □

c) Déterminer la matrice notée N de l'endomorphisme f dans la base $\mathscr{B}'$.

Démonstration.
Dans la suite, on note :

$$V_1 = \text{Mat}_{\mathscr{B}}(v_1) = \begin{pmatrix}1\\1\\1\\0\end{pmatrix},\ V_2 = \text{Mat}_{\mathscr{B}}(v_2) = \begin{pmatrix}0\\0\\0\\1\end{pmatrix},\ V_3 = \text{Mat}_{\mathscr{B}}(v_3) = \begin{pmatrix}-1\\0\\0\\1\end{pmatrix},\ V_4 = \text{Mat}_{\mathscr{B}}(v_4) = \begin{pmatrix}0\\-1\\1\\0\end{pmatrix}$$

- Tout d'abord :

$$\begin{aligned}
\text{Mat}_{\mathscr{B}}\big(f(v_1)\big) &= \text{Mat}_{\mathscr{B}}(f) \times \text{Mat}_{\mathscr{B}}(v_1)\\
&= M(a,b) \times V_1\\
&= \begin{pmatrix}a&0&0&a\\a&0&0&a\\a&0&0&a\\b&b&b&b\end{pmatrix}\begin{pmatrix}1\\1\\1\\0\end{pmatrix}\\
&= \begin{pmatrix}a\\a\\a\\3b\end{pmatrix}\\
&= a\cdot\begin{pmatrix}1\\1\\1\\0\end{pmatrix} + 3b\cdot\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\\
&= a\cdot V_1 + 3b\cdot V_2\\
&= a\cdot\text{Mat}_{\mathscr{B}}(v_1) + 3b\cdot\text{Mat}_{\mathscr{B}}(v_2)\\
&= \text{Mat}_{\mathscr{B}}(a\cdot v_1 + 3b\cdot v_2)
\end{aligned}$$

Finalement : $\text{Mat}_{\mathscr{B}}\big(f(v_1)\big) = \text{Mat}_{\mathscr{B}}(a\cdot v_1 + 3b\cdot v_2)$.
L'application $\text{Mat}_{\mathscr{B}}(\cdot)$ étant bijective, on en déduit : $f(v_1) = a\cdot v_1 + 3b\cdot v_2$.

Comme $f(v_1) = a\cdot v_1 + 3b\cdot v_2 + 0\cdot v_3 + 0\cdot v_4$.

Ainsi : $\text{Mat}_{(v_1,v_2,v_3,v_4)}\big(f(v_1)\big) = \begin{pmatrix}a\\3b\\0\\0\end{pmatrix}$.

- De même :

$$\mathrm{Mat}_{\mathscr{B}}\big(f(v_2)\big) = M(a,b)\,V_2 = \begin{pmatrix} a & 0 & 0 & a \\ a & 0 & 0 & a \\ a & 0 & 0 & a \\ b & b & b & b \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ a \\ a \\ b \end{pmatrix} = a\cdot V_1 + b\cdot V_2 = \mathrm{Mat}_{\mathscr{B}}\big(a\cdot v_1 + b\cdot v_2\big)$$

Finalement : $\mathrm{Mat}_{\mathscr{B}}\big(f(v_2)\big) = \mathrm{Mat}_{\mathscr{B}}(a\cdot v_1 + b\cdot v_2)$.
L'application $\mathrm{Mat}_{\mathscr{B}}(\cdot)$ étant bijective, on en déduit : $f(v_2) = a\cdot v_1 + b\cdot v_2$.

Comme $f(v_2) = a\cdot v_1 + b\cdot v_2 + 0\cdot v_3 + 0\cdot v_4$.

$$\text{Ainsi : } \mathrm{Mat}_{(v_1,v_2,v_3,v_4)}\big(f(v_2)\big) = \begin{pmatrix} a \\ b \\ 0 \\ 0 \end{pmatrix}.$$

- D'après la question ***5.a)*** : $v_3 \in \mathrm{Ker}(f)$.
Ainsi, par définition : $f(v_3) = 0_{\mathbb{R}^4} = 0\cdot v_1 + 0\cdot v_2 + 0\cdot v_3 + 0\cdot v_4$.

$$\text{Ainsi : } \mathrm{Mat}_{(v_1,v_2,v_3,v_4)}\big(f(v_3)\big) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

- De même : $v_4 \in \mathrm{Ker}(f)$.

$$\text{Ainsi : } \mathrm{Mat}_{(v_1,v_2,v_3,v_4)}\big(f(v_4)\big) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

$$\text{On en conclut : } N = \mathrm{Mat}_{\mathscr{B}'}(f) = \begin{pmatrix} a & a & 0 & 0 \\ 3b & b & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Commentaire

- Rappelons tout d'abord que déterminer la matrice représentative de f dans la base $\mathscr{B}'$ consiste à exprimer l'image par f des vecteurs v_1, v_2, v_3, v_4, suivant la base (v_1, v_2, v_3, v_4).
- L'énoncé ne donne pas directement accès à f mais à $M(a,b)$, sa matrice représentative dans la base $\mathscr{B}$. La base $\mathscr{B}$ étant fixée, l'application $\mathrm{Mat}_{\mathscr{B}}(.)$, appelée parfois isomorphisme de représentation, permet de traduire les propriétés énoncées dans le monde des espaces vectoriels en des propriétés énoncées dans le monde matriciel.
Voici quelques correspondances dans le cas général :

$$\begin{aligned} E \text{ espace vectoriel de dimension } n &\longleftrightarrow \mathscr{M}_{n,1}(\mathbb{R}) \\ f : E \to E \text{ endomorphisme} &\longleftrightarrow \mathrm{Mat}_{\mathscr{B}}(f) \in \mathscr{M}_n(\mathbb{R}) \\ f \text{ bijectif} &\longleftrightarrow \mathrm{Mat}_{\mathscr{B}}(f) \text{ inversible} \end{aligned}$$

Ou encore, dans le cas précis de l'exercice :

$$\text{expression de } f(v_1) \text{ dans } (v_1,v_2,v_3,v_4) \longleftrightarrow \text{expression de } M(a,b)\ V_1 \text{ dans } (V_1,V_2,V_3,V_4)$$

Il est très fréquent que les énoncés de concours requièrent de savoir traduire une propriété d'un monde à l'autre. Il est donc indispensable d'être à l'aise sur ce mécanisme. □

d) Soient λ un réel non nul et $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ une matrice colonne de $\mathscr{M}_{4,1}(\mathbb{R})$.

Montrer :

$$X \text{ est un vecteur propre de } N \text{ associé à la valeur propre } \lambda \quad \Leftrightarrow \quad \left\{ \begin{array}{l} \begin{pmatrix} x \\ y \end{pmatrix} \text{ est un vecteur propre de } T \text{ associé à} \\ \text{la valeur propre } \lambda \text{ et } z = t = 0 \end{array} \right.$$

Démonstration.

- Tout d'abord :

$$\begin{aligned} X \in E_\lambda(N) &\iff \big(N - \lambda I_4\big) X = 0_{\mathscr{M}_{4,1}(\mathbb{R})} \\ &\iff \begin{pmatrix} a-\lambda & a & 0 & 0 \\ 3b & b-\lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff \left\{ \begin{array}{rcrcrcrcl} (a-\lambda)\,x & + & a\,y & & & & & = & 0 \\ 3b\,x & + & (b-\lambda)\,y & & & & & = & 0 \\ & & & - & \lambda\,z & & & = & 0 \\ & & & & & - & \lambda\,t & = & 0 \end{array} \right. \\ &\overset{\substack{L_3 \leftarrow \frac{1}{\lambda} L_3 \\ L_4 \leftarrow \frac{1}{\lambda} L_4}}{\iff} \left\{ \begin{array}{rcrcrcrcl} (a-\lambda)\,x & + & a\,y & & & & & = & 0 \\ 3b\,x & + & (b-\lambda)\,y & & & & & = & 0 \\ & & & - & z & & & = & 0 \\ & & & & & - & t & = & 0 \end{array} \right. \end{aligned} \qquad \textit{(avec } \lambda \neq 0\textit{)}$$

- D'autre part :

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} \in E_\lambda(T) &\iff \big(T - \lambda I_2\big) \begin{pmatrix} x \\ y \end{pmatrix} = 0_{\mathscr{M}_{2,1}(\mathbb{R})} \\ &\iff \begin{pmatrix} a-\lambda & a \\ 3b & b-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\iff \left\{ \begin{array}{rcrcrcl} (a-\lambda)\,x & + & a\,y & = & 0 \\ 3b\,x & + & (b-\lambda)\,y & = & 0 \end{array} \right. \end{aligned}$$

On reconnaît les deux premières lignes du système précédent.

- Finalement :

$$\begin{aligned} X \text{ est un vecteur propre de } N \text{ associé à la valeur propre } \lambda \quad &\Leftrightarrow \quad X \neq 0_{\mathscr{M}_{4,1}(\mathbb{R})} \quad \text{et} \quad X \in E_\lambda(N) \\ &\Leftrightarrow \quad X \neq 0_{\mathscr{M}_{4,1}(\mathbb{R})} \quad \text{et} \quad \begin{pmatrix} x \\ y \end{pmatrix} \in E_\lambda(T) \quad \text{et} \quad z = t = 0 \\ &\Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} \neq 0_{\mathscr{M}_{2,1}(\mathbb{R})} \quad \text{et} \quad \begin{pmatrix} x \\ y \end{pmatrix} \in E_\lambda(T) \quad \text{et} \quad z = t = 0 \\ &\Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} \text{ est un vecteur propre de } T \text{ associé à la valeur propre } \lambda \text{ et } z = t = 0 \end{aligned}$$

On obtient bien la caractérisation souhaitée.

Commentaire

Profitons de cette question pour rappeler que, par définition, le vecteur nul n'est `JAMAIS` vecteur propre. On veillera en particulier à ne pas confondre :

- × sous-espace propre associé à la valeur propre λ (c'est bien un espace vectoriel et il contient en particulier le vecteur nul),
- × et ensemble des vecteurs propres associés à la valeur propre λ (qui n'est pas un espace vectoriel car ne contient pas le vecteur nul).

Plus précisément, avec les notations de l'énoncé, on a :

$$E_\lambda(N) = \left\{ \begin{array}{c} \text{vecteurs propres de } N \text{ associés} \\ \text{à la valeur propre } \lambda \end{array} \right\} \cup \{0_{\mathscr{M}_{4,1}(\mathbb{R})}\}$$

□

e) On suppose dans cette question **uniquement** que $(a,b) = (1,1)$.
Déterminer les valeurs propres de T. En déduire que la matrice $M(1,1)$ est diagonalisable.

Démonstration.

- Par définition de T, lorsque $(a,b) = (1,1)$, on a :

$$T = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}$$

- Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned} \lambda \text{ est valeur propre de } T &\Leftrightarrow T - \lambda I_2 \text{ est non inversible} \\ &\Leftrightarrow \det(T - \lambda I_2) = 0 \\ &\Leftrightarrow \det\left(\begin{pmatrix} 1-\lambda & 1 \\ 3 & 1-\lambda \end{pmatrix}\right) = 0 \\ &\Leftrightarrow (1-\lambda)^2 - 3 = 0 \\ &\Leftrightarrow 1-\lambda = \sqrt{3} \quad \texttt{OU} \quad 1-\lambda = -\sqrt{3} \\ &\Leftrightarrow \lambda = 1-\sqrt{3} \quad \texttt{OU} \quad \lambda = 1+\sqrt{3} \end{aligned}$$

On en déduit que T admet deux valeurs propres distinctes : $\lambda_1 = 1-\sqrt{3}$ et $\lambda_2 = 1+\sqrt{3}$.

- Soit $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ un vecteur propre de T associé à la valeur propre λ_1.

Comme $\lambda_1 \neq 0$, alors, d'après la question ***5.d)***, $X_1 = \begin{pmatrix} x_1 \\ y_1 \\ 0 \\ 0 \end{pmatrix}$ est un vecteur propre de N associé à la valeur propre λ_1. On en déduit :

$$E_{\lambda_1}(N) \supset \text{Vect}\left(\begin{pmatrix} x_1 \\ y_1 \\ 0 \\ 0 \end{pmatrix}\right)$$

On en déduit : $\dim\left(E_{\lambda_1}(N)\right) \geqslant \dim\left(\text{Vect}\left(\begin{pmatrix} x_1 \\ y_1 \\ 0 \\ 0 \end{pmatrix}\right)\right) = 1.$

- De même, en notant $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ un vecteur propre de T associé à la valeur propre $\lambda_2 \neq 0$, on obtient d'après la question ***5.d)*** :

$$E_{\lambda_2}(N) \supset \text{Vect}\left(\begin{pmatrix} x_2 \\ y_2 \\ 0 \\ 0 \end{pmatrix}\right)$$

Et ainsi : $\dim\big(E_{\lambda_2}(N)\big) \geqslant 1$.

- D'autre part :

$$\begin{aligned} \dim\big(E_0(N)\big) &= \dim\big(E_0(f)\big) && \textit{(via la passerelle matrice-endomorphisme)} \\ &= \dim\big(\text{Ker}(f)\big) && \textit{(par définition)} \\ &= 2 && \textit{(d'après la question ***5.a)***)} \end{aligned}$$

- Finalement, on obtient :

$$\dim\big(E_{\lambda_1}(N)\big) + \dim\big(E_{\lambda_2}(N)\big) + \dim\big(E_0(N)\big) \geqslant 1 + 1 + 2 = 4$$

Or, comme N est une matrice carrée d'ordre 4, on sait d'après le cours que la somme des dimensions de tous ses sous-espaces propres est majorée par 4. On en déduit que N n'admet pas d'autres valeurs propres et :

$$\dim\big(E_{\lambda_1}(N)\big) + \dim\big(E_{\lambda_2}(N)\big) + \dim\big(E_0(N)\big) = 4$$

On en conclut que la matrice N est diagonalisable.

- La matrice $N = \text{Mat}_{\mathscr{B}'}(f)$ est diagonalisable.
On en déduit que l'endomorphisme f est diagonalisable.

On en conclut alors que $M(1,1) = \text{Mat}_{\mathscr{B}}(f)$ est diagonalisable. □

f) On suppose dans cette question **uniquement** que $(a,b) = (1,-1)$.
Justifier que T n'admet aucune valeur propre. La matrice $M(1,-1)$ est-elle diagonalisable ?

Démonstration.
On raisonne comme en question précédente.

- Par définition de T, lorsque $(a,b) = (1,-1)$, on a :

$$T = \begin{pmatrix} 1 & 1 \\ -3 & -1 \end{pmatrix}$$

- Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned} \lambda \text{ est valeur propre de } T &\Leftrightarrow T - \lambda I_2 \text{ est non inversible} \\ &\Leftrightarrow \det(T - \lambda I_2) = 0 \\ &\Leftrightarrow \det\left(\begin{pmatrix} 1-\lambda & 1 \\ -3 & -1-\lambda \end{pmatrix}\right) = 0 \\ &\Leftrightarrow (1-\lambda)(-1-\lambda) + 3 = 0 \\ &\Leftrightarrow -(1-\lambda)(1+\lambda) + 3 = 0 \\ &\Leftrightarrow -(1-\lambda^2) + 3 = 0 \\ &\Leftrightarrow \lambda^2 + 2 = 0 \\ &\Leftrightarrow \lambda^2 = -2 \end{aligned}$$

On en déduit que T n'admet pas de valeurs propres réelles.

- On démontre alors que N n'admet pas de valeur propre non nulle.
  Pour ce faire, on procède par l'absurde.
  Supposons que N admet une valeur propre λ non nulle.
  Il existe donc $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ vecteur propre de N associé à cette valeur propre.
  D'après la question ***5.d)***, $\begin{pmatrix} x \\ y \end{pmatrix}$ est alors un vecteur propre de T associé à cette valeur propre λ.
  C'est impossible puisque T n'admet pas de valeur propre.

  On en conclut que la matrice N n'admet pas de valeur propre non nulle.

Commentaire

La propriété de la question ***5.d)*** est axée sur les vecteurs propres. Elle affirme en particulier que l'existence d'un vecteur propre de N assure l'existence d'un vecteur propre de T (associé à la même valeur propre) et inversement. De cet énoncé on peut tirer une propriété sur les valeurs propres. En effet, si $\lambda \neq 0$, on a :

$$\begin{aligned} \lambda \text{ est valeur propre de } N &\Leftrightarrow \text{il existe } X \neq 0_{\mathscr{M}_{4,1}(\mathbb{R})} \text{ vecteur propre de } N \text{ associé à la valeur propre } \lambda \\ &\Leftrightarrow \text{il existe } \begin{pmatrix} x \\ y \end{pmatrix} \neq 0_{\mathscr{M}_{2,1}(\mathbb{R})} \text{ vecteur propre de } T \text{ associé à la valeur propre } \lambda \\ &\Leftrightarrow \lambda \text{ est valeur propre de } T \end{aligned}$$

Ainsi, les valeurs propres non nulles de T sont exactement les valeurs propres non nulles de N. C'est d'ailleurs toute l'idée de la question : en déterminant les valeurs propres de $T \in \mathscr{M}_2(\mathbb{R})$, on récupère les valeurs propres de $N \in \mathscr{M}_2(\mathbb{R})$.

- On en déduit que 0 est l'unique valeur propre de N. Or :

$$\dim\big(E_0(N)\big) = \dim\big(\operatorname{Ker}(f)\big) = 2 \neq 4$$

  On en déduit que N n'est pas diagonalisable.

- Comme $N = \operatorname{Mat}_{\mathscr{B}'}(f)$ n'est pas diagonalisable, l'endomorphisme f n'est pas diagonalisable non plus.

  On en conclut alors que $M(1,-1) = \operatorname{Mat}_{\mathscr{B}}(f)$ n'est pas diagonalisable. □

g) Montrer l'équivalence :

$$M(a,b) \text{ est diagonalisable } \Leftrightarrow\ a^2 + 10\,ab + b^2 > 0$$

Démonstration.
Soit $(a,b) \in \mathbb{R}^2$ tel que : $a \neq 0$ et $b \neq 0$.

- Tout d'abord :

$$\begin{aligned} M(a,b) \text{ diagonalisable} &\Leftrightarrow f \text{ diagonalisable} \quad (\textit{car } M(a,b) = \operatorname{Mat}_{\mathscr{B}}(f)) \\ &\Leftrightarrow N \text{ diagonalisable} \quad (\textit{car } N = \operatorname{Mat}_{\mathscr{B}'}(f)) \end{aligned}$$

- L'idée est alors de relier la diagonalisabilité de N à celle de T, comme cela est fait dans les questions ***5.e)*** et ***5.f)*** qui précèdent.

  Remarquons tout d'abord que d'après la question ***5.d)***, la propriété suivante est vérifiée :

Les matrices N et T ont exactement les mêmes valeurs propres non nulles.	$(*)$

  On peut alors démontrer les proprités suivantes :

  × en adaptant ***5.e*** : si l'une des deux matrices possède deux valeurs propres non nulles, il en est de même de l'autre (d'après $(*)$). On démontre alors que les matrices N et T sont toutes les deux diagonalisables.

  × en adaptant ***5.f*** : si l'une des deux matrices ne possède aucune valeur propre non nulle, il en est de même de l'autre (d'après $(*)$). On démontre alors que les matrices N et T sont toutes les deux non diagonalisables.

  Il est à noter que T ne peut avoir plus de deux valeurs propres non nulles en tant que matrice carrée d'ordre 2. Il en est donc de même de N en vertu de la propriété $(*)$. Il reste donc à traiter le cas où N et T possèdent une seule valeur propre non nulle.

- Supposons que N et T possède une seule valeur propre λ non nulle.
  Étudions alors leur diagonalisabilité.

  × Étude de T

  Remarquons tout d'abord :

  $$\det(T) = \det\left(\begin{pmatrix} a & a \\ 3b & b \end{pmatrix}\right) = ab - 3ab = -2ab \neq 0 \qquad \textit{(car } a \neq 0 \textit{ et } b \neq 0\textit{)}$$

  On en déduit que T est inversible. Ainsi, 0 n'est pas valeur propre de T.
  La matrice T possède donc λ pour seule valeur propre.

  Démontrons alors que T n'est pas diagonalisable. On procède par l'absurde.
  Supposons que T est diagonalisable. Il existe donc :

  ▸ une matrice inversible $P \in \mathscr{M}_2(\mathbb{R})$,

  ▸ une matrice diagonale $D \in \mathscr{M}_2(\mathbb{R})$ dont les coefficients diagonaux sont les valeurs propres de T,

  telles que $T = PDP^{-1}$.

  Or λ est la seule valeur propre de T. Ainsi $D = \lambda\, I_2$ et :

  $$T \;=\; P\,(\lambda\, I_2)\,P^{-1} \;=\; P\,(\lambda\, I_2)\,P^{-1} \;=\; \lambda\, I_2$$

  Absurde !

Si T admet une seule valeur propre non nulle, alors T n'est pas diagonalisable.

  × Étude de N

  La matrice N possède alors deux valeurs propres : $\lambda \neq 0$ et 0. De plus, on a déjà démontré : $\dim\big(E_0(N)\big) = 2$.
  Démontrons alors que N n'est pas diagonalisable. On procède par l'absurde.
  Supposons que N est diagonalisable. On a alors :

  $$\begin{array}{ccccc} \dim\big(E_\lambda(N)\big) & + & \dim\big(E_0(N)\big) & = & \dim\big(\mathscr{M}_{4,1}(\mathbb{R})\big) \\ & & \shortparallel & & \shortparallel \\ & & 2 & & 4 \end{array}$$

Ainsi : $\dim\big(E_\lambda(N)\big) = 4-2 = 2$. On en déduit qu'il existe :

$$X_1 = \begin{pmatrix} x_1 \\ y_1 \\ 0 \\ 0 \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R}) \quad \text{et} \quad X_2 = \begin{pmatrix} x_2 \\ y_2 \\ 0 \\ 0 \end{pmatrix} \in \mathscr{M}_{4,1}(\mathbb{R})$$

tels que (X_1, X_2) est une base de $E_\lambda(N)$. D'après la question ***5.d)***, on a alors :

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \in E_\lambda(T) \quad , \quad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in E_\lambda(T) \quad \text{et ainsi} \quad E_\lambda(T) \ \supseteq \ \text{Vect}\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)$$

On en déduit :

$$\dim\big(E_\lambda(T)\big) \ \geqslant \ \dim\left(\text{Vect}\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)\right) = 2$$

En effet, la famille (X_1, X_2) étant libre, il en est de même de la famille $\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)$.

On en déduit : $\dim\big(E_\lambda(T)\big) = 2 = \dim\big(\mathscr{M}_{2,1}(\mathbb{R})\big)$. Ainsi, la matrice T est diagonalisable. Impossible ! (d'après l'étude précédente, si T admet une seule valeur propre non nulle elle n'est pas diagonalisable).

Si N admet une seule valeur propre non nulle, alors N n'est pas diagonalisable.

- Finalement, N et T sont toutes deux diagonalisables à la même condition qu'elles possèdent deux valeurs propres non nulles distinctes. Ainsi :

$$\begin{aligned} M(a,b) \text{ diagonalisable} &\Leftrightarrow N \text{ diagonalisable} \\ &\Leftrightarrow N \text{ admet deux valeurs propres non nulles distinctes} \\ &\Leftrightarrow T \text{ admet deux valeurs propres non nulles distinctes} \\ &\Leftrightarrow T \text{ admet deux valeurs propres distinctes} \quad \textit{(car 0 n'est pas valeur propre de } T\textit{)} \end{aligned}$$

Or, on a :

$$\begin{aligned} \lambda \text{ valeur propre de } T &\Leftrightarrow T - \lambda I_2 \text{ non inversible} \\ &\Leftrightarrow \det(T - \lambda I_2) = 0 \\ &\Leftrightarrow \det\left(\begin{pmatrix} a-\lambda & a \\ 3b & b-\lambda \end{pmatrix}\right) = 0 \\ &\Leftrightarrow (a-\lambda)(b-\lambda) - 3ab = 0 \\ &\Leftrightarrow -2ab - (a+b)\lambda + \lambda^2 = 0 \\ &\Leftrightarrow \lambda \text{ est racine du polynôme } Q(X) = -2ab - (a+b)X + X^2 \end{aligned}$$

Le discriminant du polynôme Q est :

$$\Delta = (a+b)^2 - 4(-2ab) = (a^2 + 2ab + b^2) + 8ab = a^2 + 10ab + b^2$$

Finalement :

$$\begin{aligned} M(a,b) \text{ diagonalisable} &\Leftrightarrow T \text{ admet deux valeurs propres distinctes} \\ &\Leftrightarrow \Delta > 0 \end{aligned}$$

$M(a,b)$ diagonalisable $\Leftrightarrow \ a^2 + 10ab + b^2 > 0.$

□

Exercice 3

Dans cet exercice, toutes les variables aléatoires sont supposées définies sur un même espace probabilisé noté $(\Omega, \mathscr{A}, \mathbb{P})$.

Partie A : Loi de Pareto

Soient a et b deux réels strictement positifs. On définit la fonction f sur $\mathbb{R}$ par :

$$f : x \mapsto \begin{cases} 0 & \text{si } x < b \\ a\,\dfrac{b^a}{x^{a+1}} & \text{si } x \geqslant b \end{cases}$$

1. Montrer que f est une densité de probabilité.

Démonstration.

- La fonction f est continue :
  - × sur $]-\infty, b[$ en tant que fonction constante,
  - × sur $]b, +\infty[$ car elle est l'inverse de la fonction $x \mapsto x^{a+1}$ qui :
    - est continue sur $]b, +\infty[$,
    - `NE S'ANNULE PAS` sur $]b, +\infty[$.

La fonction f est continue sur $\mathbb{R}$ sauf éventuellement en b.

- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  - × si $x \in]-\infty, b[$, alors : $f(x) = 0 \geqslant 0$.
  - × si $x \in [b, +\infty[$, alors, comme $a > 0$ et $b > 0$: $f(x) = a\,\dfrac{b^a}{x^{a+1}} \geqslant 0$.

Finalement : $\forall x \in \mathbb{R},\ f(x) \geqslant 0$.

- Démontrons que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx$ converge et vaut 1.
  - × Tout d'abord, comme f est nulle en dehors de $[b, +\infty[$:

$$\int_{-\infty}^{+\infty} f(x)\,dx \;=\; \int_{b}^{+\infty} f(x)\,dx$$

  - × La fonction f est continue par morceaux sur $[b, +\infty[$. L'intégrale $\displaystyle\int_{b}^{+\infty} f(x)\,dx$ est donc seulement impropre en $+\infty$.
  - × Soit $B \in [b, +\infty[$.

$$\begin{aligned} \int_b^B f(x)\,dx &= \int_b^B a\,\frac{b^a}{x^{a+1}}\,dx \;=\; a\,b^a \int_b^B x^{-a-1}\,dx \\ &= a\,b^a \left[\,\frac{x^{-a}}{-a}\,\right]_b^B \;=\; -b^a \left[\,\frac{1}{x^a}\,\right]_b^B \qquad \textit{(car } a \neq 0\textit{)} \\ &= -b^a \left(\frac{1}{B^a} - \frac{1}{b^a}\right) \end{aligned}$$

× Or, comme $a > 0$: $\lim\limits_{B\to+\infty} \dfrac{1}{B^a} = 0$. D'où :

$$\lim_{B\to+\infty} -b^a \left(\frac{1}{B^a} - \frac{1}{b^a}\right) = -b^a \left(0 - \frac{1}{b^a}\right) = 1$$

Ainsi, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt$ est convergente et vaut 1.

Finalement la fonction f est une densité de probabilité.

□

On dit qu'une variable aléatoire suit la loi de Pareto de paramètres a et b lorsqu'elle admet pour densité la fonction f.
Dans toute la suite de l'exercice, on considère une variable aléatoire X suivant la loi de Pareto de paramètres a et b.

2. Déterminer la fonction de répartition de X.

Démonstration.

- Dans la suite, on considère : $X(\Omega) = [b, +\infty[$
- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  × si $x \in]-\infty, b[$, alors : $[X \leqslant x] = \varnothing$ (car $X(\Omega) = [b, +\infty[$). D'où :

$$F_X(x) = \mathbb{P}([X \leqslant x]) = \mathbb{P}(\varnothing) = 0$$

  × si $x \in [b, +\infty[$, alors :

$$\begin{aligned}
F_X(x) &= \mathbb{P}([X \leqslant x]) \\
&= \int_{-\infty}^{x} f(t)\,dt \\
&= \int_{b}^{x} f(t)\,dt && \textit{(car } f \textit{ est nulle en dehors de } [b, +\infty[\textit{)} \\
&= \int_{b}^{x} a\,\frac{b^a}{t^{a+1}}\,dt \\
&= a\,b^a \int_{b}^{x} t^{-a-1}\,dt \\
&= \cancel{a}\,b^a \left[\frac{t^{-a}}{-\cancel{a}}\right]_b^x && \textit{(car } a \neq 0\textit{)} \\
&= -b^a \left(\frac{1}{x^a} - \frac{1}{b^a}\right) = 1 - \left(\frac{b}{x}\right)^a
\end{aligned}$$

Finalement : $F_X : x \mapsto \begin{cases} 0 & \text{si } x < b \\ 1 - \left(\dfrac{b}{x}\right)^a & \text{si } x \geqslant b \end{cases}$

Commentaire

Profitons de cette question pour faire une remarque sur la notation $X(\Omega)$.

- Rappelons qu'une v.a.r. X est une application $X : \Omega \to \mathbb{R}$.
  Comme la notation le suggère, $X(\Omega)$ est l'image de Ω par l'application X.
  Ainsi, $X(\Omega)$ n'est rien d'autre que l'ensemble des valeurs prises par la v.a.r. X :

$$\begin{aligned} X(\Omega) &= \{X(\omega) \mid \omega \in \Omega\} \\ &= \{x \in \mathbb{R} \mid \exists \omega \in \Omega,\ X(\omega) = x\} \end{aligned}$$

  Il faut bien noter que dans cette définition aucune application probabilité $\mathbb{P}$ n'apparaît.

- Il est toujours correct d'écrire : $X(\Omega) \subseteq]-\infty, +\infty[$.
  En effet, cette propriété signifie que toute v.a.r. X est à valeurs dans $\mathbb{R}$, ce qui est toujours le cas par définition de la notion de variable aléatoire réelle.

- Dans le cas des v.a.r. discrètes, il est d'usage relativement courant de confondre :
  - × l'ensemble de valeurs possibles de la v.a.r. X (*i.e.* l'ensemble $X(\Omega)$),
  - × l'ensemble $\{x \in \mathbb{R} \mid \mathbb{P}\big([X = x]\big) \neq 0\}$, ensemble des valeurs que X prend avec probabilité non nulle. Dans le cas qui nous intéresse ici, à savoir X est une v.a.r. discrète, cet ensemble est appelé support de X et est noté $\text{Supp}(X)$.

- Dans le cas des v.a.r. à densité, la détermination de l'ensemble image est plus technique. Dans certains sujets, l'ensemble image des v.a.r. étudiées sera précisé (« On considère une v.a.r. à valeurs strictement positives »). Si ce n'est pas le cas :
  - × si X suit une loi usuelle, on peut se référer à l'ensemble image donné en cours. Par exemple, si $X \hookrightarrow \mathcal{U}([0,1])$, on se permet d'écrire :

    « Comme $X \hookrightarrow \mathcal{U}([0,1])$, on **considère** : $X(\Omega) = [0,1]$. »

  - × si X ne suit pas une loi usuelle, on étudie l'ensemble : $I = \{x \in \mathbb{R} \mid f_X(x) > 0\}$.
    On se permet alors d'écrire :

    « Dans la suite, on **considère** : $X(\Omega) = I$. »

  En **décrétant** la valeur de $X(\Omega)$, on ne commet pas une erreur mais on décide d'ajouter une hypothèse qui ne fait pas partie de l'énoncé. Cette audace permet de travailler avec un ensemble image connu, ce qui permet de structurer certaines démonstrations (l'ensemble image étant connu, on se rappelle que la fonction de répartition, par exemple, s'obtient à l'aide d'une dsijonction de cas). □

3. a) Soit U une variable aléatoire suivant la loi uniforme sur $[0,1[$.
Montrer que la variable aléatoire $bU^{-\frac{1}{a}}$ suit la loi de Pareto de paramètres a et b.

Démonstration.

- Notons $h : x \mapsto b\,x^{-\frac{1}{a}}$, de sorte que $Y = bU^{-\frac{1}{a}} = h(U)$.

  On **considère** : $U(\Omega) =]0,1]$. On en déduit :

$$\begin{aligned} Y(\Omega) &= \big(h(U)\big)(\Omega) = h\big(U(\Omega)\big) \\ &= h(]0,1]) \\ &= [h(1), \lim_{x\to 0} h(x)[\qquad \textit{(car, comme } a>0 \textit{ et } b>0\textit{, la fonction } h \textit{ est continue et strictement décroissante sur }]0,1]\textit{)} \\ &= [b, +\infty[\end{aligned}$$

  Ainsi : $Y(\Omega) = [b, +\infty[$.

Commentaire

Rappelons que la v.a.r. $Y = h(U)$ est par définition l'application :

$$\begin{array}{rccc} Y = h(U) : & \Omega & \to & \mathbb{R} \\ & \omega & \mapsto & h\big(U(\omega)\big) \end{array}$$

Comme la fonction h est définie uniquement sur $]0,+\infty[$, la v.a.r. $Y = h(U)$ est bien définie seulement si :

$$\forall \omega \in \Omega,\ U(\omega) \subset]0,+\infty[$$

Autrement dit, il est primordial, pour la bonne définition de l'objet $Y = h(U)$, de considérer que U est à valeurs dans $]0,1] \subset]0,+\infty[$ (et non $[0,1[\not\subset]0,+\infty[$ comme pouvait le suggérer l'énoncé).

- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  - × si $x \in]-\infty, b[$, alors : $[Y \leqslant x] = \varnothing$ (car $Y(\Omega) = [b,+\infty[$). D'où :

$$F_Y(x) \ = \ \mathbb{P}\big([Y \leqslant x]\big) \ = \ \mathbb{P}(\varnothing) \ = \ 0$$

  - × si $x \in [b,+\infty[$, alors :

$$\begin{array}{rcll} F_Y(x) & = & \mathbb{P}\big([Y \leqslant x]\big) & \\ & = & \mathbb{P}\left(\left[b\,U^{-\frac{1}{a}} \leqslant x\right]\right) & \\ & = & \mathbb{P}\left(\left[U^{-\frac{1}{a}} \leqslant \dfrac{x}{b}\right]\right) & \textit{(car } b > 0\textit{)} \\ & = & \mathbb{P}\left(\left[U \geqslant \left(\dfrac{x}{b}\right)^{-a}\right]\right) & \textit{(par stricte décroissance de la fonction } x \mapsto x^{-a} \textit{ sur }]0,+\infty[\textit{, car } a > 0\textit{)} \\ & = & 1 - F_U\left(\left(\dfrac{b}{x}\right)^{a}\right) & \textit{(car } U \textit{ est une v.a.r. à densité)} \end{array}$$

  - × Or :

$$\begin{array}{lcll} & x \geqslant b & & \\ \text{donc} & \dfrac{1}{x} \leqslant \dfrac{1}{b} & \textit{(par décroissance de la fonction inverse sur }]0,+\infty[\textit{, car } b > 0\textit{)} \\ \text{d'où} & \dfrac{b}{x} \leqslant 1 & \textit{(car } b > 0\textit{)} \\ \text{ainsi} & \left(\dfrac{b}{x}\right)^{a} \leqslant 1 & \textit{(par croissance de la fonction } x \mapsto x^a \textit{ sur } [0,+\infty[\textit{, car } a > 0\textit{)} \end{array}$$

    Comme $a > 0$, $b < 0$ et $x > 0$, on en déduit : $0 < \left(\dfrac{b}{x}\right)^{a} \leqslant 1$.

× De plus, comme $U \hookrightarrow \mathcal{U}(]0,1]\,)$: $F_U : u \mapsto \begin{cases} 0 & \text{si } u \leqslant 0 \\ u & \text{si } 0 < u \leqslant 1 \\ 1 & \text{si } u > 1 \end{cases}$.

Ainsi : $F_U\left(\left(\frac{b}{x}\right)^a\right) = \left(\frac{b}{x}\right)^a$.

Finalement : $F_Y : x \mapsto \begin{cases} 0 & \text{si } x < b \\ 1 - \left(\frac{b}{x}\right)^a & \text{si } x \geqslant b \end{cases}$.

× D'après la question **2.**, on reconnaît la fonction de répartition F_X de X qui suit la loi de Pareto de paramètres a et b.
Or la fonction de répartition caractérise la loi d'une v.a.r. .

On en déduit que Y suit la loi de Pareto de paramètres a et b.

Commentaire

- On a démontré, lors de l'étude de $Y(\Omega)$, que h réalise une bijection de $]0,1]$ sur $[b,+\infty[$. Il est possible de déterminer l'expression de $h^{-1} : [b,+\infty[\to]0,1]$.
Pour ce faire, on remarque que pour tout $x \in]0,1]$ et $y \in [b,+\infty[$, on a :

$$\begin{aligned} y = h(x) &\Leftrightarrow y = b\,x^{-\frac{1}{a}} \\ &\Leftrightarrow x = \left(\frac{b}{y}\right)^a \\ &\Leftrightarrow x = h^{-1}(y) \end{aligned}$$

On démontre ainsi que h^{-1} a pour expression : $h^{-1} : x \mapsto \left(\frac{b}{x}\right)^a$.

- On retrouve ici l'expression de la quantité $\left(\frac{b}{x}\right)^a$ apparaissant à la fin de la résolution de la question. Ce n'est pas surprenant car la méthode utilisée ici consiste justement à faire apparaître, étape par étape, la quantité $h^{-1}(x)$. Plus précisément, on a :

$$F_Y(x) = \mathbb{P}\big([Y \leqslant x]\big) = \mathbb{P}\big([h(U) \leqslant x]\big) = \mathbb{P}\big(\big[U \leqslant h^{-1}(x)\big]\big) = F_U\big(h^{-1}(x)\big)$$

On comprend mieux pourquoi cette manière de procéder est appelée **méthode d'inversion**. □

b) En déduire une fonction **Scilab** d'en-tête `function X = pareto(a,b)` qui prend en arguments deux réels a et b strictement positifs et qui renvoie une simulation de la variable aléatoire X.

Démonstration.
On propose la fonction **Scilab** suivante :

```
1 function X = pareto(a, b)
2     U = rand()
3     X = b * (U.^(-1/a))
4 endfunction
```

Détaillons les éléments de ce script.

- **Début de la fonction**
  On commence par préciser la structure de la fonction :
  - × cette fonction se nomme `pareto`,
  - × elle prend en entrée 2 paramètres `a` et `b`,
  - × elle admet pour variable de sortie la variable `X`.

```
1  function X = pareto(a, b)
```

- **Contenu de la fonction**
  En ligne 2, on stocke dans la variable `U` une simulation de la v.a.r. U de loi $\mathcal{U}(]0,1])$.

```
2        U = rand()
```

  D'après la question précédente, la v.a.r. $Y = bU^{-\frac{1}{a}}$ suit la même loi que la v.a.r. X, dès lors que $U \hookrightarrow \mathcal{U}(]0,1])$.
  Ainsi, en ligne 3, on stocke dans la variable `Y` une simulation de la v.a.r. X.

```
3        X = b * (U.^(-1/a))
```

Commentaire

On rappelle que l'opérateur `.^` est l'opérateur de puissance terme à terme, contrairement à l'opérateur `^` qui est l'opérateur de puissance mathématique classique. Par exemple, en notant : $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, on obtient :

```
--> A = [1, 2 ; 3, 4]
--> A.^2
    ans =
        1.    4.
        9.    16.
--> A^2
    ans =
        7.    10.
        15.    22.
```

□

c) On considère la fonction **Scilab** ci-dessous.
Que contient la liste `L` renvoyée par la fonction `mystere` ?

```
1  function L = mystere(a, b)
2      L = []
3      for p = 2 : 6
4          S = 0
5          for k = 1 : 10^p
6              S = S + pareto(a,b)
7          end
8          L = [L, S / 10^p]
9      end
10 endfunction
```

Démonstration.
Détaillons les éléments de ce script.

- **Début de la fonction**
  On commence par préciser la structure de la fonction :
  - × cette fonction se nomme `mystere`,
  - × elle prend en entrée 2 paramètres `a` et `b`,
  - × elle admet pour variable de sortie la variable `L`.

```
1  function L = mystere(a, b)
```

  La variable de sortie `L` est ensuite initialisée à la matrice vide.

```
2        L = []
```

- **Structure itérative**
  Commençons par décrire ce qu'il se passe à l'intérieur de cette première structure itérative (`for p = 2 : 6`) avant de s'intéresser à l'utilité de la variable `p`.
  - × On commence par initialiser une variable `S` à 0 (choix naturel d'initialisation lorsqu'on souhaite coder une somme puisque 0 est l'élément neutre de l'opérateur de sommation).
  - × Les lignes 5 à 7 permettent de mettre à jour la variable `S` pour qu'elle contienne la somme de 10^p simulations de la v.a.r. X suivant la loi de Pareto de paramètres `a` et `b`. Pour cela, on met de nouveau en place une structure conditionnelle (boucle for) et on utilise la fonction `pareto` définie en question précédente pour obtenir les simulations de la v.a.r. X.

```
5          for k = 1 : 10^p
6              S = S + pareto(a,b)
7          end
```

  - × Enfin, on met à jour la vecteur `L` en lui concaténant à droite la valeur stockée dans la variable `S / 10^p`.

  Remarquons que, comme `S` contient la somme de 10^p simulations de la v.a.r. X, la variable `S / 10^p` contient la moyenne des 10^p simulations de X.
  Or, l'idée naturelle pour obtenir une approximation de l'espérance $\mathbb{E}(X)$ est :
  - × de simuler un grand nombre de fois ($N = 10^p$ est ici ce grand nombre) la v.a.r. X.
    Formellement, on souhaite obtenir un N-uplet $(x_1, \dots, x_N)$ qui correspond à l'observation d'un N-échantillon $(X_1, \dots, X_N)$ de la v.a.r. X.
    (les v.a.r. X_i sont indépendantes et de même loi que X)
  - × de réaliser la moyenne des résultats de cette observation.

  Cette idée est justifiée par la loi faible des grands nombres (LfGN) qui affirme :

$$\text{moyenne de l'observation} \;=\; \frac{1}{N}\sum_{k=1}^{N} x_k \;\simeq\; \mathbb{E}(X)$$

  La variable `S / 10^p` est donc une approximation de $\mathbb{E}(X)$.

- **Sortie de la fonction**
  - × À la fin de la boucle `for p = 2 : 6`, la variable `L` est donc un vecteur à 5 coordonnées (autant que de valeurs prises par la variable `p`) contenant 5 approximations de $\mathbb{E}(X)$.
  - × Plus précisément, la variable `p` prend successivement les valeurs de l'ensemble $[\![2, 6]\!]$.
    Ainsi, la variable `L` contient :

    - en 1ère coordonnée, une approximation de $\mathbb{E}(X)$ obtenue à l'aide de 10^2 simulations de X,

- en $2^{\text{ème}}$ coordonnée, une approximation de $\mathbb{E}(X)$ obtenue à l'aide de 10^3 simulations de X,
- ...
- en $5^{\text{ème}}$ coordonnée, une approximation de $\mathbb{E}(X)$ obtenue à l'aide de 10^6 simulations de X.

La variable `L` est donc un vecteur contenant 5 approximations de plus en plus précises de $\mathbb{E}(X)$ où X suit une loi de Pareto de paramètres `a` et `b`.

Commentaire

- Afin de permettre une bonne compréhension des mécanismes en jeu, on a détaillé avec précision la réponse à cette question. Cependant, fournir la bonne réponse démontre la bonne compréhension de l'algorithme et permet d'obtenir tous les points alloués à cette question. On procèdera de même dans les autres questions **Scilab**.
- Il est étonnant que l'énoncé mentionne la notion de liste. Dans beaucoup de langages informatiques, on trouve la distinction entre liste et tableau. Sans trop entrer dans les détails, ces deux objets répondent à des objectifs différents. En général, la distinction s'établit comme suit :
  - × un tableau est un objet dont la taille est fixé à l'avance et ne peut varier. L'accès et la modification d'un élément du tableau se fait de manière rapide.
  - × une liste est un objet dont on peut faire évoluer facilement la taille. Grossièrement, une liste est implantée suivant la logique d'une pile d'assiettes. L'ajout se fait en début de liste (on pose un nouvelle assiette en haut de la pile). L'accès à un élément de la liste requiert de passer en revue tous les éléments de la liste (on dépile les assiettes jusqu'à l'obtention de celle que l'on souhaite).

  En **Scilab**, il n'y a pas vraiment lieu de parler de liste. Ce langage est pensé avec une faible variété de types. L'idée générale est la suivante : « en **Scilab** tout est matrice ». D'ailleurs, l'objet renvoyé par la fonction `mystere` est bien une matrice ligne à 5 éléments (on parle aussi de vecteur). L'implantation du type matriciel en **Scilab** ne nécessite pas que la taille de la matrice soit fixée en amont. On peut donc agir comme le propose le concepteur et ainsi copier l'idée de la liste consistant à ajouter des éléments au fur et à mesure. Mais cela ne semble pas pertinent car on connaît dès le début de la fonction la taille de l'objet renvoyé. Il est donc **fortement recommandé** de penser l'objet `L` plus comme un tableau dont la taille fixée initialement n'évolue plus. Cela peut se faire en initialisant `L` de la manière suivante :

```
2    L = zeros(1, 5)
```

  Cela permet d'allouer une bonne fois pour toute l'espace mémoire nécessaire à la définition de l'objet `L`. On échappe ainsi au risque qu'une nouvelle allocation de mémoire soit nécessaire si la taille de l'objet grandit au-delà de l'espace mémoire alloué initialement.
- Les considérations détaillées dans le point précédent ne sont pas un attendu du programme. Elles sont présentées ici uniquement car le concepteur mentionne la notion de liste, terme non présent dans le programme officiel.

□

d) On exécute la fonction précédente avec différentes valeurs de a et de b.
Comment interpréter les résultats obtenus ?

```
-->mystere(2,1)
   ans =
       1.9306917   1.9411352   1.9840089   1.9977684   2.0012415
-->mystere(3,2)
   ans =
       3.1050951   3.0142956   2.9849407   2.9931656   2.9991517
-->mystere(1,4)
   ans =
       21.053151   249.58609   51.230522   137.64549   40.243918
```

Démonstration.

- L'instruction `mystere(2,1)` renvoie un vecteur contenant 5 approximations de plus en plus précises de $\mathbb{E}(X)$, où X suit une loi de Pareto de paramètres 2 et 1.
  Les 5 valeurs affichées semblent être de plus en plus proche de 2.

  On peut donc conjecturer que, si X suit une loi de Pareto de paramètres 2 et 1, alors : $\mathbb{E}(X) = 2$.

- L'instruction `mystere(3,2)` renvoie un vecteur contenant 5 approximations de plus en plus précises de $\mathbb{E}(X)$, où X suit une loi de Pareto de paramètres 3 et 2.
  Les 5 valeurs affichées semblent être de plus en plus proche de 3.

  On peut donc conjecturer que, si X suit une loi de Pareto de paramètres 3 et 2, alors : $\mathbb{E}(X) = 3$.

- L'instruction `mystere(1,4)` renvoie un vecteur contenant 5 approximations de plus en plus précises de $\mathbb{E}(X)$, où X suit une loi de Pareto de paramètres 1 et 4.
  Les 5 valeurs affichées ne semblent pas converger vers une valeur en particulier.

  On peut donc conjecturer que, si X suit une loi de Pareto de paramètres 1 et 4, alors elle n'admet pas d'espérance.

□

4. a) Montrer que X admet une espérance si et seulement si $a > 1$ et que, dans ce cas :

$$\mathbb{E}(X) \;=\; \frac{a\,b}{a-1}$$

Démonstration.

- La v.a.r. X admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x\,f(x)\,dx$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour des calculs de moments du type $\displaystyle\int_{-\infty}^{+\infty} x^r\,f(x)\,dx$.
- Comme f est nulle en dehors de $[b, +\infty[$:

$$\int_{-\infty}^{+\infty} x\,f(x)\,dx \;=\; \int_{b}^{+\infty} x\,f(x)\,dx$$

- De plus, pour tout $x \in [b, +\infty[$:

$$x\,f(x) \;=\; x\,\frac{a\,b^a}{x^{a+1}} \;=\; a\,b^a\,\frac{1}{x^a}$$

- Or $\displaystyle\int_b^{+\infty} \frac{1}{x^a}\,dx$ est une intégrale de Riemann, impropre en $+\infty$ ($b>0$), d'exposant a. Elle est donc convergente si et seulement si $a>1$.

  On en déduit que la v.a.r. X admet une espérance si et seulement si $a>1$.

- Supposons alors $a>1$.
  Soit $B\in[b,+\infty[$.

$$\begin{aligned}\int_b^B x\,f(x)\,dx &= a\,b^a\int_b^B x^{-a}\,dx\\ &= a\,b^a\left[\frac{x^{-a+1}}{-a+1}\right]_b^B \qquad (car\ a\neq 1)\\ &= -\frac{a\,b^a}{a-1}\left[\frac{1}{x^{a-1}}\right]_b^B\\ &= -\frac{a\,b^a}{a-1}\left(\frac{1}{B^{a-1}}-\frac{1}{b^{a-1}}\right)\end{aligned}$$

  Comme $a-1>0$: $\displaystyle\lim_{B\to+\infty}\frac{1}{B^{a-1}}=0$. D'où :

$$\mathbb{E}(X) = -\frac{a\,b^a}{a-1}\left(0-\frac{1}{b^{a-1}}\right) = -\frac{a\,b^a}{a-1}\left(-\frac{1}{b^{a-1}}\right) = \frac{a\,\cancel{b^a}}{(a-1)\cancel{b^a}\,b^{-1}} = \frac{a\,b}{a-1}$$

  Ainsi, si $a>1$: $\mathbb{E}(X)=\dfrac{a\,b}{a-1}$.

Commentaire

On remarque qu'on trouve bien des résultats cohérents avec les résultats obtenus avec **Scilab** en question précédente :

× si X suit une loi de Pareto de paramètres 2 et 1, alors X admet une espérance et :

$$\mathbb{E}(X) = \frac{2\times 1}{2-1} = 2$$

× si X suit une loi de Pareto de paramètres 3 et 2, alors X admet une espérance et :

$$\mathbb{E}(X) = \frac{3\times 2}{3-1} = 3$$

× si X suit une loi de Pareto de paramètres 1 et 4, alors X n'admet pas d'espérance. □

b) Montrer que X admet une variance si et seulement si $a>2$ et que, dans ce cas :

$$\mathbb{V}(X) = \frac{a\,b^2}{(a-1)^2\,(a-2)}$$

Démonstration.

- La v.a.r. X admet une variance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^2\,f(x)\,dx$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour des calculs de moments du type $\displaystyle\int_{-\infty}^{+\infty} x^r\,f(x)\,dx$.
- Comme f est nulle en dehors de $[b,+\infty[$:

$$\int_{-\infty}^{+\infty} x^2\,f(x)\,dx = \int_b^{+\infty} x^2\,f(x)\,dx$$

- De plus, pour tout $x \in [b, +\infty[$:

$$x^2\, f(x) \;=\; x^2\, \frac{a\, b^a}{x^{a+1}} \;=\; a\, b^a\, \frac{1}{x^{a-1}}$$

- Or $\displaystyle\int_b^{+\infty} \frac{1}{x^{a-1}}\, dx$ est une intégrale de Riemann, impropre en $+\infty$ ($b > 0$), d'exposant $a-1$. Elle est donc convergente si et seulement si $a - 1 > 1$.

On en déduit que la v.a.r. X admet une variance si et seulement si $a > 2$.

- Supposons alors $a > 2$.
Soit $B \in [b, +\infty[$.

$$\begin{aligned}
\int_b^B x\; f(x)\; dx &= a\, b^a \int_b^B x^{-a+1}\, dx \\
&= a\, b^a \left[\, \frac{x^{-a+2}}{-a+2} \,\right]_b^B \qquad (\textit{car } a \neq 2) \\
&= -\frac{a\, b^a}{a-2} \left[\, \frac{1}{x^{a-2}} \,\right]_b^B \\
&= -\frac{a\, b^a}{a-2} \left(\frac{1}{B^{a-2}} - \frac{1}{b^{a-2}} \right)
\end{aligned}$$

Comme $a - 2 > 0$: $\displaystyle\lim_{B \to +\infty} \frac{1}{B^{a-2}} = 0$. D'où :

$$\mathbb{E}(X^2) \;=\; -\frac{a\; b^a}{a-2} \left(0 - \frac{1}{b^{a-2}} \right) \;=\; -\frac{a\; b^a}{a-2} \left(-\frac{1}{b^{a-2}} \right) \;=\; \frac{a\; \cancel{b^a}}{(a-2)\cancel{b^a}\; b^{-2}} \;=\; \frac{a\, b^2}{a-2}$$

Ainsi, si $a > 2$: $\mathbb{E}(X^2) = \dfrac{a\, b^2}{a-2}$.

- Par la formule de Kœnig-Huygens :

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2 \\
&= \frac{a\, b^2}{a-2} - \left(\frac{a\, b}{a-1} \right)^2 \\
&= ab^2 \left(\frac{1}{a-2} - \frac{a}{(a-1)^2} \right) \\
&= ab^2 \left(\frac{(a-1)^2 - a(a-2)}{(a-2)\,(a-1)^2} \right) \\
&= ab^2 \left(\frac{(\cancel{a^2} - \cancel{2a} + 1) - (\cancel{a^2} - \cancel{2a})}{(a-2)\,(a-1)^2} \right) \\
&= ab^2\, \frac{1}{(a-2)\,(a-1)^2}
\end{aligned}$$

Ainsi, si $a > 2$: $\mathbb{V}(X) = \dfrac{a\, b^2}{(a-2)\,(a-1)^2}$.

□

Partie B : Estimation du paramètre b

On suppose **dans cette partie uniquement** que $a = 3$ et on chercher à déterminer un estimateur performant de b.
Ainsi, la variable aléatoire X admet pour densité la fonction f définie par :

$$f : x \mapsto \begin{cases} 0 & \text{si } x < b \\ \dfrac{3\,b^3}{x^4} & \text{si } x \geqslant b \end{cases}$$

Soient $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires indépendantes, toutes de même loi que X.
On définit :

$$Y_n \;=\; \min(X_1, \ldots, X_n) \qquad \text{et} \qquad Z_n = \frac{X_1 + \cdots + X_n}{n}$$

On admet que Y_n et Z_n sont encore des variables aléatoires définies sur $(\Omega, \mathscr{A}, \mathbb{P})$.

5. a) Calculer, pour tout x de $[b, +\infty[$, $\mathbb{P}(\,[Y_n > x]\,)$.

Démonstration.
Soit $x \in [b, +\infty[$.

$$\begin{aligned}
\mathbb{P}(\,[Y_n > x]\,) &= \mathbb{P}\left(\bigcap_{k=1}^{n} [X_k > x]\right) \\
&= \prod_{k=1}^{n} \mathbb{P}(\,[X_k > x]\,) && \textit{(car les v.a.r. } X_1, \ldots, X_n \textit{ sont indépendantes)} \\
&= \prod_{k=1}^{n} (1 - F_{X_k}(x)) \\
&= (1 - F_X(x))^n && \textit{(car les v.a.r. } X_1, \ldots, X_n \textit{ ont même loi que } X\textit{)} \\
&= \left(\cancel{1} - \left(\cancel{1} - \left(\frac{b}{x}\right)^3\right)\right)^n && \textit{(d'après 2., car } x \geqslant b\textit{)} \\
&= \left(\frac{b}{x}\right)^{3n}
\end{aligned}$$

$$\boxed{\forall x \in [b, +\infty[,\ \mathbb{P}(\,[Y_n > x]\,) = \left(\frac{b}{x}\right)^{3n}}$$

□

b) En déduire que Y_n suit une loi de Pareto dont on précisera les paramètres.

Démonstration.

- Tout d'abord : $\forall k \in [\![1, n]\!]$, $X_k(\Omega) = [b, +\infty[$.

  $$\boxed{\text{Ainsi : } Y_n(\Omega) \subset [b, +\infty[.}$$

- Déterminons F_{Y_n}, la fonction de répartition de Y_n.
  Soit $x \in \mathbb{R}$. Deux cas se présentent :

  × si $x \in]-\infty, b[$, alors $[Y_n \leqslant x] = \varnothing$ (car $Y_n(\Omega) \subset [b, +\infty[$). D'où :

  $$F_{Y_n}(x) \;=\; \mathbb{P}(\,[Y_n \leqslant x]\,) \;=\; \mathbb{P}(\varnothing) \;=\; 0$$

× si $x \in [b, +\infty[$, alors :

$$\begin{aligned} F_{Y_n}(x) &= \mathbb{P}(\,[Y_n \leqslant x]\,) \\ &= 1 - \mathbb{P}(\,[Y_n > x]\,) \\ &= 1 - \left(\frac{b}{x}\right)^{3n} \qquad \textit{(d'après la question précédente)} \end{aligned}$$

$$\text{Finalement : } F_{Y_n} : x \mapsto \begin{cases} 0 & \text{si } x < b \\ 1 - \left(\dfrac{b}{x}\right)^{3n} & \text{si } x \geqslant b \end{cases} .$$

- D'après la question ***2.***, on reconnaît la fonction de répartition d'une v.a.r. de loi de Pareto de paramètres $3n$ et b.
  Or la fonction de répartition caractérise la loi d'une v.a.r. .

  On en déduit que Y_n suit la loi de Pareto de paramètres $3n$ et b.

□

c) Montrer que $Y'_n = \dfrac{3n-1}{3n}\, Y_n$ est un estimateur sans biais de b.

Calculer le risque quadratique de cet estimateur.

Démonstration.

- La v.a.r. $Y'_n = \dfrac{3n-1}{3n}\, Y_n = \dfrac{3n-1}{3n}\, \min(X_1, \ldots, X_n)$ s'exprime :
  × à l'aide d'un n-échantillon $(X_1, \ldots, X_n)$ de la v.a.r. X,
  × sans mention du paramètre b.

  La v.a.r. Y'_n est donc un estimateur de b.

- Comme $3n \geqslant 3 > 1$, d'après la question ***4.a)***, la v.a.r. Y_n admet une espérance.
  Ainsi, la v.a.r. Y'_n admet une espérance (donc un biais) en tant que transformée linéaire d'une v.a.r. qui en admet une.
- De plus :

$$\begin{aligned} \mathbb{E}(Y'_n) &= \mathbb{E}\left(\frac{3n-1}{3n}\, Y_n\right) \\ &= \frac{3n-1}{3n}\, \mathbb{E}(Y_n) \qquad \textit{(par linéarité de l'espérance)} \\ &= \frac{3n-1}{3n}\, \frac{3n\,b}{3n-1} \qquad \textit{(d'après les questions 4.a) et 5.b))} \\ &= b \end{aligned}$$

  La v.a.r. Y'_n est donc un estimateur sans biais de b.

- Comme $3n \geqslant 3 > 2$, d'après la question ***4.b)***, la v.a.r. Y_n admet une variance.
  Ainsi, la v.a.r. Y'_n admet une variance (donc un risque quadratique) en tant que transformée linéaire d'une v.a.r. qui en admet une.

- Par décomposition biais-variance :

$$\begin{aligned} r_b(Y_n') &= \mathbb{V}(Y_n') + \big(b_b(Y_n')\big)^2 \\ &= \mathbb{V}(Y_n') + 0 && \textit{(car } Y_n' \textit{ est un estimateur sans biais de } b\textit{)} \\ &= \mathbb{V}\left(\frac{3n-1}{3n}\, Y_n\right) \\ &= \left(\frac{3n-1}{3n}\right)^2 \mathbb{V}(W_n) \\ &= \frac{\cancel{(3n-1)^2}}{(3n)^2}\, \frac{3n\, b^2}{\cancel{(3n-1)^2}\,(3n-2)} && \textit{(d'après les questions } \textbf{4.a)} \textit{ et } \textbf{5.b)}\textit{)} \end{aligned}$$

Finalement : $r_b(Y_n') = \dfrac{b^2}{3n\,(3n-2)}$ □

6. a) Déterminer l'espérance et la variance de Z_n.

Démonstration.

- La v.a.r. Z_n admet une variance (donc une espérance) en tant que combinaison linéaire de v.a.r. qui en admettent une.
- De plus :

$$\begin{aligned} \mathbb{E}(Z_n) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\ &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_k) && \textit{(par linéarité de l'espérance)} \\ &= \frac{1}{n}\sum_{k=1}^{n} \left(\frac{3\,b}{3-1}\right) && \textit{(d'après } \textbf{4.a)}\textit{)} \\ &= \frac{1}{\cancel{n}} \times \cancel{n}\, \frac{3}{2}\, b \end{aligned}$$

Ainsi : $\mathbb{E}(Z_n) = \dfrac{3}{2}\, b.$

- Enfin :

$$\begin{aligned} \mathbb{V}(V_n) &= \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\ &= \left(\frac{1}{n}\right)^2 \mathbb{V}\left(\sum_{k=1}^{n} X_k\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(X_k) && \textit{(car les v.a.r. } X_1, \ldots, X_n \textit{ sont indépendantes)} \\ &= \frac{1}{n^2}\sum_{k=1}^{n} \frac{3\,b^2}{(3-1)^2\,(3-2)} && \textit{(d'après } \textbf{4.b)}\textit{)} \\ &= \frac{1}{\cancel{n^2}} \times \cancel{n}\, \frac{3\,b^2}{4} \end{aligned}$$

On en déduit : $\mathbb{V}(Z_n) = \dfrac{3\,b^2}{4n}.$ □

b) En déduire un estimateur noté Z'_n sans biais de b de la forme $\alpha\, Z_n$ où α est un réel à préciser. Calculer le risque quadratique de cet estimateur.

Démonstration.

- D'après la question précédente :

$$\mathbb{E}(Z_n) = \frac{3}{2}\, b$$

$$\text{donc} \qquad \frac{2}{3}\, \mathbb{E}(Z_n) = b$$

$$\text{d'où} \qquad \mathbb{E}\left(\frac{2}{3}\, Z_n\right) = b \qquad \textit{(par linéarité de l'espérance)}$$

On pose alors $\alpha = \dfrac{2}{3}$. La v.a.r. $Z'_n = \alpha\, Z_n = \dfrac{2}{3}\, Z_n$ s'exprime :

× à l'aide d'un n-échantillon $(X_1, \ldots, X_n)$ de la v.a.r. X,

× sans mention du paramètre b.

La v.a.r. $Z'_n = \dfrac{2}{3}\, Z_n$ est donc un estimateur de b.

- La v.a.r. Z'_n admet une espérance (donc un biais) en tant que transformée linéaire de la v.a.r. Z_n qui en admet une. De plus :

$$b_b(Z'_n) \;=\; b_b(\alpha\, Z_n) \;=\; \mathbb{E}(\alpha\, Z_n) - b \;=\; 0 \qquad \textit{(d'après un point précédent)}$$

La v.a.r. Z'_n est un estimateur sans biais de b.

- La v.a.r. Z'_n admet une variance (donc un risque quadratique) en tant que transformée linéaire de la v.a.r. Z_n qui en admet une.
- Par décomposition biais-variance :

$$\begin{aligned}
r_b(Z'_n) &= \mathbb{V}(Z'_n) + \big(b_b(Z'_n)\big)^2 \\
&= \mathbb{V}(\alpha\, Z_n) + 0 && \textit{(car } Z'_n \textit{ est un estimateur sans biais de } b\textit{)} \\
&= \alpha^2\, \mathbb{V}(Z_n) \\
&= \frac{4}{9}\, \frac{3\, b^2}{4n} && \textit{(d'après la question précédente)} \\
&= \frac{b^2}{3n}
\end{aligned}$$

$r_b(Z'_n) = \dfrac{b^2}{3n}$

□

7. Entre Y'_n et Z'_n, quel estimateur choisir ? Justifier.

Démonstration.

- Comme les estimateurs Y'_n et Z'_n de b sont tous les deux sans biais (questions ***5.c)*** et ***6.b)***), le meilleur des deux est celui dont le risque quadratique est le plus faible.

- Or, toujours d'après ***5.c)*** et ***6.b)*** :

$$\begin{aligned} r_b(Y_n') \leqslant r_b(Z_n') &\Leftrightarrow \frac{b^2}{3n\,(3n-2)} \leqslant \frac{b^2}{3n} \\ &\Leftrightarrow \frac{1}{3n-2} \leqslant 1 && \textit{(car } \frac{3n}{b^2} > 0\textit{)} \\ &\Leftrightarrow 3n-2 \geqslant 1 && \textit{(par stricte décroissance de la fonction inverse sur }]0,+\infty[\textit{)} \\ &\Leftrightarrow 3n \geqslant 3 \end{aligned}$$

Cette dernière inégalité est vraie car $n \in \mathbb{N}^*$. Ainsi, par équivalence, la 1ère aussi.

On en déduit que Y_n' est un meilleur estimateur de b que Z_n'.

□

Partie C : Estimation du paramètre a

On suppose **dans cette partie uniquement** que $b = 1$ et on cherche à construire un intervalle de confiance pour a.
Ainsi, la variable aléatoire X admet pour densité la fonction f définie par :

$$f : x \mapsto \begin{cases} 0 & \text{si } x < 1 \\ \dfrac{a}{x^{a+1}} & \text{si } x \geqslant 1 \end{cases}$$

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, toutes de même loi que X.

8. Soit $n \in \mathbb{N}^*$. On pose : $W_n = \ln(X_n)$.
Montrer que la variable aléatoire W_n suit une loi exponentielle dont on précisera le paramètre.
En déduire l'espérance et la variance de W_n.

Démonstration.

- Commençons par déterminer $W_n(\Omega)$.
Notons $h : x \mapsto \ln(x)$, de telle sorte que $W_n = h(X_n)$.
On considère ici $X_n(\Omega) = [1,+\infty[$. On en déduit :

$$\begin{aligned} W_n(\Omega) &= \big(h(X_n)\big)(\Omega) \;=\; h\big(X_n(\Omega)\big) \\ &= h([1,+\infty[) \\ &= [h(1), \lim_{x\to+\infty} h(x)[&& \textit{(car la fonction } h \textit{ est continue et strictement croissante sur } [1,+\infty[\textit{)} \\ &= [0,+\infty[\end{aligned}$$

Et ainsi : $W_n(\Omega) = [0,+\infty[$.

- Déterminons la fonction de répartition F_{W_n} de W_n.
Soit $x \in \mathbb{R}$. Deux cas se présentent :

× si $x \leq 0$, alors $[W_n \leqslant x] = \varnothing$ (car $W_n(\Omega) = [0,+\infty[$). D'où :

$$F_{W_n}(x) = \mathbb{P}([W_n \leqslant x]) = \mathbb{P}(\varnothing) = 0$$

× si $x \geqslant 0$ alors :

$$\begin{aligned} F_{W_n}(x) &= \mathbb{P}([W_n \leqslant x]) \\ &= \mathbb{P}([\ln(X_n) \leqslant x]) \\ &= \mathbb{P}([X_n \leqslant \mathrm{e}^x]) && \textit{(par stricte croissance de la fonction } \exp \textit{ sur } \mathbb{R}) \\ &= F_{X_n}(\mathrm{e}^x) \\ &= 1 - \left(\frac{1}{\mathrm{e}^x}\right)^a && \textit{(car } b = 1 \textit{ et, comme } x \geqslant 0, \ e^x \geqslant 1) \\ &= 1 - (\mathrm{e}^{-x})^a \ = \ 1 - \mathrm{e}^{-ax} \end{aligned}$$

On obtient finalement : $F_{W_n} : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathrm{e}^{-ax} & \text{si } x \geqslant 0 \end{cases}$.

- On reconnaît la fonction de répartition d'une v.a.r. qui suit une loi $\mathcal{E}(a)$. Or la fonction de répartition caractérise la loi.

On en déduit : $W_n \hookrightarrow \mathcal{E}(a)$

On en conclut : $\mathbb{E}(W_n) = \dfrac{1}{a}$ et $\mathbb{V}(W_n) = \dfrac{1}{a^2}$.

□

9. On définit, pour tout n de $\mathbb{N}^*$:

$$M_n = \frac{\ln(X_1) + \cdots + \ln(X_n)}{n} \qquad \text{et} \qquad T_n = \sqrt{n}\,(a\,M_n - 1)$$

a) Justifier que la suite de variables aléatoires $(T_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

Démonstration.
Soit $n \in \mathbb{N}^*$.
On cherche ici à appliquer le théorème central limite. On commence donc par déterminer $\overline{W}_n^*$ et on souhaite relier cette v.a.r. à la v.a.r. T_n de l'énoncé.

- Intéressons nous d'abord à la v.a.r. $M_n = \overline{W}_n$.
  - × La v.a.r. M_n admet une variance (et donc une espérance) en tant que combinaison linéaire de v.a.r. qui en admettent une.
  - × De plus :

$$\begin{aligned} \mathbb{E}(M_n) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} W_k\right) \\ &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(W_k) && \textit{(par linéarité de l'espérance)} \\ &= \frac{1}{n}\sum_{k=1}^{n} \frac{1}{a} && \textit{(d'après la question précédente)} \\ &= \frac{1}{n} \times n\,\frac{1}{a} \end{aligned}$$

Ainsi : $\mathbb{E}(M_n) = \dfrac{1}{a}$.

× Enfin :

$$\begin{aligned}
\mathbb{V}(M_n) &= \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} W_k\right) \\
&= \left(\frac{1}{n}\right)^2 \mathbb{V}\left(\sum_{k=1}^{n} W_k\right) \\
&= \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(W_k) && \textit{(car les v.a.r. } W_1, \ldots, W_n \textit{ sont indépendantes par lemme des coalitions)} \\
&= \frac{1}{n^2}\sum_{k=1}^{n} \frac{1}{a^2} && \textit{(d'après la question précédente)} \\
&= \frac{1}{n^2} \times n\, \frac{1}{a^2}
\end{aligned}$$

$$\boxed{\text{D'où : } \mathbb{V}(M_n) = \frac{1}{n\,a^2}.}$$

- Comme la v.a.r. $M_n = \overline{W}_n$ admet une variance non nulle, la v.a.r. $\overline{W}_n^*$ est bien définie. De plus :

$$\begin{aligned}
\overline{W}_n^* &= \frac{\overline{W}_n - \mathbb{E}(\overline{W}_n)}{\sqrt{\mathbb{V}(\overline{W}_n)}} \\
&= \frac{M_n - \frac{1}{a}}{\sqrt{\frac{1}{n\,a^2}}} && \textit{(d'après ce qui précède)} \\
&= \frac{M_n - \frac{1}{a}}{\frac{1}{a\sqrt{n}}} \\
&= a\sqrt{n}\left(M_n - \frac{1}{a}\right) \\
&= \sqrt{n}\,(a\,M_n - 1)
\end{aligned}$$

$$\boxed{\text{On en déduit : } \overline{W}_n^* = T_n.}$$

- La suite $(W_i)_{i\in\mathbb{N}}$ est une suite de v.a.r. :
  - × indépendantes par lemme des coalitions (car la suite $(X_i)_{i\in\mathbb{N}}$ est une suite de v.a.r. indépendantes),
  - × de même loi $\mathcal{E}(a)$,
  - × qui admettent une variance non nulle $\dfrac{1}{a^2}$.

$$\boxed{\text{Ainsi, par théorème central limite : } T_n \xrightarrow[n\to+\infty]{\mathscr{L}} Z, \text{ où } Z \hookrightarrow \mathcal{N}(0,1).}$$

□

b) En déduire que l'intervalle $\left[\dfrac{\sqrt{n}-2}{\sqrt{n}\,M_n}\ ;\ \dfrac{\sqrt{n}+2}{\sqrt{n}\,M_n}\right]$ est un intervalle de confiance asymptotique pour a au niveau de confiance 95%.
On admettra que $\Phi(2) \geqslant 0,975$, où Φ désigne la fonction de répartition d'une variable aléatoire suivant la loi normale centrée réduite.

Commentaire

- L'énoncé considère ici : $M_n(\Omega) \subset]0,+\infty[$. En effet, dans la cas contraire, les v.a.r. $\dfrac{\sqrt{n}-2}{\sqrt{n}\,M_n}$ et $\dfrac{\sqrt{n}+2}{\sqrt{n}\,M_n}$ ne sont pas bien définies.
- Cette hypothèse n'est pas aberrante puisque, d'après la question ***8.*** : $W_n \hookrightarrow \mathcal{E}(a)$. Ainsi, on peut considérer : $W_n(\Omega) \subset]0,+\infty[$.
  Comme $M_n = \dfrac{1}{n}\sum_{k=1}^{n} W_k$, on en déduit : $M_n(\Omega) \subset]0,+\infty[$.
- Pour une remarque plus complète sur la notation $W_n(\Omega)$, on se reportera à la question ***2***.

Démonstration.
On cherche à démontrer :

$$\lim_{n\to+\infty} \mathbb{P}\left(\left[\frac{\sqrt{n}-2}{\sqrt{n}\,M_n} \leqslant a \leqslant \frac{\sqrt{n}+2}{\sqrt{n}\,M_n}\right]\right) \geqslant 95\%$$

- Soit $n \in \mathbb{N}^*$. Étudions d'abord la probabilité.

$$\begin{aligned}
&\mathbb{P}\left(\left[\frac{\sqrt{n}-2}{\sqrt{n}\,M_n} \leqslant a \leqslant \frac{\sqrt{n}+2}{\sqrt{n}\,M_n}\right]\right)\\
&= \mathbb{P}\left([\sqrt{n}-2 \leqslant a\sqrt{n}\,M_n \leqslant \sqrt{n}+2]\right) \qquad \textit{(car } \sqrt{n}>0 \textit{ et } M_n(\Omega)\subset]0,+\infty[\textit{)}\\
&= \mathbb{P}\left([-2 \leqslant a\sqrt{n}\,M_n - \sqrt{n} \leqslant 2]\right)\\
&= \mathbb{P}\left([-2 \leqslant \sqrt{n}\,(a\,M_n - 1) \leqslant 2]\right)\\
&= \mathbb{P}\left([-2 \leqslant T_n \leqslant 2]\right)
\end{aligned}$$

- Or, d'après la question précédente : $T_n \xrightarrow[n\to+\infty]{\mathscr{L}} Z$, où $Z \hookrightarrow \mathcal{N}(0,1)$. D'où :

$$\begin{aligned}
\lim_{n\to+\infty} \mathbb{P}\left([-2 \leqslant T_n \leqslant 2]\right) &= \mathbb{P}\left([-2 \leqslant Z \leqslant 2]\right)\\
&= \Phi(2) - \Phi(-2) \qquad \textit{(car } Z \hookrightarrow \mathcal{N}(0,1)\textit{)}\\
&= \Phi(2) - \big(1 - \Phi(2)\big)\\
&= 2\,\Phi(2) - 1
\end{aligned}$$

- De plus, d'après l'énoncé :

$$\begin{aligned}
&\Phi(2) \geqslant 0,975\\
\text{donc}\quad &2\,\Phi(2) \geqslant 1,95\\
\text{d'où}\quad &2\,\Phi(2) - 1 \geqslant 0,95
\end{aligned}$$

Ainsi :

$$\lim_{n\to+\infty} \mathbb{P}\left(\left[\frac{\sqrt{n}-2}{\sqrt{n}\,M_n} \leqslant a \leqslant \frac{\sqrt{n}+2}{\sqrt{n}\,M_n}\right]\right) \geqslant 95\%$$

On en déduit que l'intervalle $\left[\dfrac{\sqrt{n}-2}{\sqrt{n}\,M_n}\ ;\ \dfrac{\sqrt{n}+2}{\sqrt{n}\,M_n}\right]$ est un intervalle de confiance asymptotique pour a au niveau de confiance 95%.

□