DS9

Exercice

On désigne par id l'endomorphisme identité de $\mathbb{R}^3$, et par I la matrice identité de $\mathscr{M}_3(\mathbb{R})$.

Partie I : Étude d'un endomorphisme de $\mathbb{R}^3$

On note $\mathscr{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et on considère l'endomorphisme f de $\mathbb{R}^3$ dont la matrice dans la base canonique est :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & -2 \\ -1 & 1 & 0 \end{pmatrix}$$

1. ***a)*** Calculer $A^2 - 3\,A$ puis déterminer un polynôme annulateur de A de degré 2.

b) En déduire les deux valeurs propres possibles λ_1 et λ_2 de A (avec $\lambda_1 < \lambda_2$).

c) En **Scilab**, l'instruction `r = rank(M)` permet de stocker dans la variable `r` le rang de la matrice `M`. On considère le programme suivant :

```
A = [1, 0, 0; -2, 3, -2; -1, 1, 0]
r1 = rank(A - eye(3,3))
r2 = rank(A - 2 * eye(3,3))
```

On place ci-dessous le résultat des instructions qui permettent d'obtenir les valeurs `r1` et `r2` dans la console.

```
--> r1
    r1  =
        1.

--> r2
    r2  =
        2.
```

Que peut-on conjecturer quant aux valeurs propres de f et la dimension des sous-espaces propres associés ?

d) Déterminer une base (u_1, v_1) de $\text{Ker}(f - \lambda_1 \,\text{id})$. Le réel λ_1 est-il valeur propre de f ?
On choisira les vecteurs u_1 et v_1 de façon que leurs composantes soient des entiers naturels les plus petits possible, la dernière composante de u_1 et la première de v_1 étant nulles.

e) De même, déterminer une base (v_2) de $\text{Ker}(f - \lambda_2 \,\text{id})$. Le réel λ_2 est-il valeur propre de f ?
On choisira v_2 de façon que ses composantes soient des entiers naturels les plus petits possible.

2. ***a)*** Justifier que la famille (u_1, v_1, v_2) est une base de $\mathbb{R}^3$.
Quelle est la matrice représentative de f dans cette base ?

b) On note $x = (a, b, c)$ un vecteur quelconque de $\mathbb{R}^3$.
Déterminer, en fonction de a, b, c les coordonnées de x dans la base (u_1, v_1, v_2).

3. Utiliser le polynôme annulateur obtenu en question ***1.*** pour déterminer A^{-1} en fonction de A et I.

Partie II : Étude d'une base de $\mathbb{R}^n$

Soient n et p deux entiers naturels tels que $n \geqslant p \geqslant 2$ et soit E un $\mathbb{R}$-espace vectoriel de dimension n. On considère f un endomorphisme de E. On suppose :

× f diagonalisable,

× f possède p valeurs propres $\lambda_1, \ldots, \lambda_p$ deux à deux distinctes.

Dans la suite, on note $\mathscr{B}'$ une base dans laquelle la matrice représentative de f est une matrice diagonale notée D. On note id l'endomorphisme identité de $\mathbb{R}^n$.

4. ***a)*** On note I_n la matrice identité de $\mathscr{M}_n(\mathbb{R})$, montrer :

$$(D - \lambda_1 I_n) \ \ldots \ (D - \lambda_p I_n) \ = \ 0_{\mathscr{M}_n(\mathbb{R})}$$

b) En déduire un polynôme annulateur de f.

Pour tout entier $k \in [\![1, p]\!]$, on définit le polynôme $L_k = \prod\limits_{\substack{j=1 \\ j \neq k}}^{p} \dfrac{X - \lambda_j}{\lambda_k - \lambda_j}$.

5. Soit $k \in [\![1, p]\!]$.

a) En distinguant les cas $i = k$ et $i \neq k$, calculer $L_k(\lambda_i)$.

b) Montrer que $(L_1, \ldots, L_p)$ est une base de $\mathbb{R}_{p-1}[X]$.

c) Établir alors : $\forall P \in \mathbb{R}_{p-1}[X], \ P \ = \ \sum\limits_{k=1}^{p} P(\lambda_k) \ L_k$.

d) En déduire : $\sum\limits_{k=1}^{p} L_k = 1$. Que vaut $\sum\limits_{k=1}^{p} L_k(f)$? Si $x \in E$, que vaut $\sum\limits_{k=1}^{p} \big(L_k(f)\big)(x)$?

6. Soit $k \in [\![1, p]\!]$.

a) Démontrer : $\forall i \in [\![1, p]\!], \ (f - \lambda_k \,\mathrm{id}) \circ (f - \lambda_i \,\mathrm{id}) \ = \ (f - \lambda_i \,\mathrm{id}) \circ (f - \lambda_k \,\mathrm{id})$.

b) Écrire $(f - \lambda_k \,\mathrm{id}) \circ L_k(f)$ sous forme d'une composée de p applications linéaires.
En déduire que pour tout $x \in E$, $\big(L_k(f)\big)(x)$ est un élément de $\mathrm{Ker}(f - \lambda_k \,\mathrm{id})$.

c) En combinant le résultat de la question précédente et celui de la question ***5.***, que démontre-t-on ?

Problème

- Toutes les variables aléatoires qui interviennent dans ce problème sont supposées définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.
- Sous réserve d'existence, on note $\mathbb{E}(X)$ et $\mathbb{V}(X)$ respectivement l'espérance et la variance d'une variable aléatoire X, et $\mathrm{Cov}(X, Y)$ la covariance de deux v.a.r. X et Y.
- Dans les parties I et III, la fonction de répartition et une densité d'une variable aléatoire X à densité sont notées respectivement F_X et f_X.
- **On admet** que les formules donnant l'espérance et la variance d'une somme de variables aléatoires discrètes, ainsi que la définition et les propriétés de la covariance et du coefficient de corrélation linéaire de deux variables aléatoires discrètes, s'appliquent au cas de variables aléatoires à densité.
- Pour n entier supérieur ou égal à 2, on dit que les variables aléatoires à densité $X_1, X_2, \ldots, X_n$ sont indépendantes si pour tout n-uplet $(x_1, x_2, \ldots, x_n)$ de réels, les événements $[X_1 \leqslant x_1]$, $[X_2 \leqslant x_2]$, $\ldots$, $[X_n \leqslant x_n]$ sont indépendants.
- L'objet du problème est double. D'une part, montrer certaines analogies entre les lois géométriques et exponentielles, d'autre part mettre en évidence quelques propriétés asymptotiques de variables aléatoires issues de la loi exponentielle. La partie II est indépendante de la partie I. La partie III est indépendante de la partie II et largement indépendante de la partie I.

Partie I. Loi exponentielle

1. ***a)*** Rappeler la valeur de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\, dt$.

Établir pour tout n de $\mathbb{N}^*$ la convergence de l'intégrale $\displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\, dt$.

On pose alors $I_0 = \displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\, dt$ et pour tout n de $\mathbb{N}^*$ $I_n = \displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\, dt$.

b) Soit n un entier de $\mathbb{N}^*$. À l'aide d'une intégration par parties, établir une relation de récurrence entre I_n et I_{n-1}. En déduire la valeur de I_n en fonction de n.

Soit λ un réel strictement positif. Soient X_1 et X_2 deux variables indépendantes de même loi exponentielle de paramètre λ (d'espérance $\frac{1}{\lambda}$).

On pose : $Y = X_1 - X_2$, $T = \max(X_1, X_2)$ et $Z = \min(X_1, X_2)$.

2. Justifier les relations $T + Z = X_1 + X_2$ et $T - Z = |X_1 - X_2| = |Y|$.

3. ***a)*** Rappeler sans démonstration les valeurs de $\mathbb{V}(X_1)$ et de $\mathbb{P}\big(\,[X_1 \leqslant x]\,\big)$, pour tout réel x.

b) Calculer $\mathbb{E}(X_1 + X_2)$, $\mathbb{V}(X_1 + X_2)$, $\mathbb{E}(Y)$, $\mathbb{V}(Y)$.

4. Déterminer pour tout réel z, $F_Z(z)$ et $f_Z(z)$. Reconnaître la loi de Z et en déduire $\mathbb{E}(Z)$ et $\mathbb{V}(Z)$.

5. ***a)*** Montrer que pour tout réel t, on a : $F_T(t) = \begin{cases} \left(1 - \mathrm{e}^{-\lambda t}\right)^2 & \text{si } t \geqslant 0 \\ 0 & \text{si } t < 0 \end{cases}$.

Exprimer pour tout réel t, $f_T(t)$.

b) Justifier l'existence de $\mathbb{E}(T)$ et $\mathbb{V}(T)$. Montrer que $\mathbb{E}(T) = \dfrac{3}{2\lambda}$ et $\mathbb{V}(T) = \dfrac{5}{4\lambda^2}$.
(on pourra utiliser des changements de variables affine)

6. On note r le coefficient de corrélation linéaire de Z et T. Montrer que $r = \dfrac{1}{\sqrt{5}}$.

7. ***a)*** Préciser $Y(\Omega)$ et $|Y|(\Omega)$.

b) Déterminer une densité de la variable aléatoire $-X_2$.

c) Montrer que pour tout réel y, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_{X_1}(t)\, f_{-X_2}(y-t)\, dt$ est convergente et qu'elle vaut $\dfrac{\lambda}{2}\, \mathrm{e}^{-\lambda\,|y|}$ (on distinguera les deux cas : $y \geqslant 0$ et $y < 0$).

d) Établir que la fonction $y \mapsto \dfrac{\lambda}{2}\mathrm{e}^{-\lambda\,|y|}$ est une densité de probabilité sur $\mathbb{R}$; on admet que c'est une densité de la variable aléatoire Y.

e) Déterminer pour tout y réel, $f_{|Y|}(y)$. Reconnaître la loi de $|Y| = T - Z$.

Partie II. Loi géométrique

Non proposée ici.

Partie III. Convergences

Dans les questions 1 à 4, λ désigne un paramètre réel strictement positif, inconnu.
Pour n élément de $\mathbb{N}^*$, on considère un n-échantillon $(X_1, X_2, \ldots, X_n)$ de variables aléatoires à valeurs strictement positives, indépendantes, de même loi exponentielle de paramètre λ.
On pose pour tout n de $\mathbb{N}^*$: $S_n = \sum\limits_{k=1}^{n} X_k$ et $J_n = \lambda S_n$.

1. Calculer pour tout n de $\mathbb{N}^*$, $\mathbb{E}(S_n)$, $\mathbb{V}(S_n)$, $\mathbb{E}(J_n)$ et $\mathbb{V}(J_n)$.

2. On admet qu'une densité f_{J_n} de J_n est donnée par $f_{J_n}(x) = \begin{cases} \dfrac{\mathrm{e}^{-x} x^{n-1}}{(n-1)!} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$.

a) À l'aide du théorème de transfert, établir pour tout n supérieur ou égal à 3, l'existence de $\mathbb{E}\left(\dfrac{1}{J_n}\right)$ et de $\mathbb{E}\left(\dfrac{1}{J_n^2}\right)$, et donner leur valeurs respectives.

b) On pose pour tout n supérieur ou égal à 3 : $\widehat{\lambda_n} = \dfrac{n}{S_n}$. Justifier que $\widehat{\lambda_n}$ est un estimateur de λ. Est-il sans biais ? Calculer la limite, lorsque n tend vers $+\infty$, du risque quadratique associé à $\widehat{\lambda_n}$ en λ.

3. Dans cette question, on veut déterminer un intervalle de confiance du paramètre λ au risque α.
On note Φ la fonction de répartition de la loi normale centrée réduite, et u_α le réel strictement positif tel que $\Phi(u_\alpha) = 1 - \dfrac{\alpha}{2}$.

a) Énoncer le théorème de la limite centrée. En déduire que la variable aléatoire N_n définie par $N_n = \lambda \dfrac{S_n}{\sqrt{n}} - \sqrt{n}$ converge en loi vers la loi normale centrée réduite.

b) En déduire que pour n assez grand, on a approximativement : $\mathbb{P}\big([-u_\alpha \leqslant N_n \leqslant u_\alpha]\big) = 1 - \alpha$.

c) Montrer que pour n assez grand, l'intervalle $\left[\left(1 - \dfrac{u_\alpha}{\sqrt{n}}\right) \widehat{\lambda_n}, \left(1 + \dfrac{u_\alpha}{\sqrt{n}}\right) \widehat{\lambda_n}\right]$ est un intervalle de confiance de λ au risque α. On note λ_0 la réalisation de $\widehat{\lambda_n}$ sur le n-échantillon.

4. Avec le n-échantillon $(X_1, X_2, \ldots, X_n)$, on construit un nouvel intervalle de confiance de λ au risque β $(\beta \neq \alpha)$, tel que la longueur de cet intervalle soit k $(k > 1)$ fois plus petite que celle obtenue avec le risque α.

a) Justifier l'existence de la fonction réciproque Φ^{-1} de Φ.
Quel est le domaine de définition de Φ^{-1} ?

b) Établir l'égalité $\beta = 2\Phi\left(\dfrac{1}{k}\Phi^{-1}\left(\dfrac{\alpha}{2}\right)\right)$.
En déduire que $\beta > \alpha$. Ce dernier résultat était-il prévisible ?

Dans les questions 5 à 7, on suppose que $\lambda = 1$.

5. On pose pour tout n de $\mathbb{N}^*$: $T_n = \max(X_1, X_2, \ldots, X_n)$.
Pour tout n de $\mathbb{N}^*$, pour tout réel x positif ou nul, on pose :

$$g_n(x) = \int_0^x F_{T_n}(t)\, dt \quad \text{et} \quad h_n(x) = \int_0^x t\, f_{T_n}(t)\, dt$$

a) Exprimer $h_n(x)$ en fonction de $F_{T_n}(x)$ et $g_n(x)$.

b) Déterminer pour tout réel t, l'expression de $F_{T_n}(t)$ en fonction de t.
Établir pour tout n supérieur ou égal à 2, la relation : $g_{n-1}(x) - g_n(x) = \dfrac{1}{n}\, F_{T_n}(x)$.

c) En déduire que pour tout n de $\mathbb{N}^*$, pour tout réel x positif ou nul, l'expression de $g_n(x)$ en fonction de x, $F_{T_1}(x)$, $F_{T_2}(x)$, $\ldots$, $F_{T_n}(x)$.

d) Montrer que $F_{T_n}(x) - 1$ est équivalent à $-n\mathrm{e}^{-x}$, lorsque x tend vers $+\infty$.

e) Déduire des questions ***c)*** et ***d)*** l'existence de $\mathbb{E}(T_n)$ et montrer que $\mathbb{E}(T_n) = \sum\limits_{k=1}^{n} \dfrac{1}{k}$.

6. On veut étudier dans cette question la convergence en loi de la suite de variables aléatoires $(G_n)_{n\geqslant 1}$ définie par : pour tout n de $\mathbb{N}^*$, $G_n = T_n - \mathbb{E}(T_n)$.
On pose pour tout n de $\mathbb{N}^*$: $\gamma_n = -\ln(n) + \mathbb{E}(T_n)$ et on admet sans démonstration que la suite $(\gamma_n)_{n\geqslant 1}$ est convergente ; on note γ sa limite.

a) Montrer que pour tout x réel et n assez grand, on a : $F_{G_n}(x) = \left(1 - \dfrac{1}{n}\, \mathrm{e}^{-(x+\gamma_n)}\right)^n$.

b) En déduire que pour tout x réel, on a : $\lim\limits_{n\to+\infty} F_{G_n}(x) = \mathrm{e}^{-\mathrm{e}^{-(x+\gamma)}}$.

c) Montrer que la fonction $F_G : \mathbb{R} \to \mathbb{R}$ définie par $F_G : x \mapsto \mathrm{e}^{-\mathrm{e}^{-(x+\gamma)}}$ est la fonction de répartition d'une variable aléatoire G à densité. Conclure.

7. ***a)*** Soit X une variable aléatoire à densité de fonction de répartition F_X strictement croissante. Déterminer la loi de la variable aléatoire Y définie par $Y = F_X(X)$.

b) Écrire une fonction **Scilab** d'en-tête `Gumbel` qui permet de simuler la variable aléatoire G.
On supposera que la constante γ est définie en langage **Scilab** par une constante `gamma`.
On rappelle que la fonction **Scilab** `rand()` permet de simuler la loi uniforme sur $]0, 1[$.