



PRÉPASIMPLE

prepasimple.fr

Ecricome

Mathématiques

2021

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

EXERCICE 1

Partie 1 : Étude de trois matrices

1. Il faut calculer A^3 , puis utiliser le fait que toute valeur propre d'une matrice est racine de tout polynôme annulateur de cette matrice.

Calcul de A^2 . Pour $i \in \{1, 2, 3\}$, le coefficient diagonal $(A^2)_{i,i}$ est le produit de la ligne i de A par la colonne i de A :

$$(A^2)_{1,1} = 0 \cdot 0 + 1 \cdot (-1) + (-1) \cdot 1 = -2, \quad (A^2)_{2,2} = (-1) \cdot 1 + 0 \cdot 0 + 1 \cdot (-1) = -2,$$

$$(A^2)_{3,3} = 1 \cdot (-1) + (-1) \cdot 1 + 0 \cdot 0 = -2.$$

Pour les coefficients non diagonaux :

$$(A^2)_{1,2} = 0 \cdot 1 + 1 \cdot 0 + (-1) \cdot (-1) = 1, \quad (A^2)_{1,3} = 0 \cdot (-1) + 1 \cdot 1 + (-1) \cdot 0 = 1,$$

$$(A^2)_{2,1} = (-1) \cdot 0 + 0 \cdot (-1) + 1 \cdot 1 = 1, \quad (A^2)_{2,3} = (-1) \cdot (-1) + 0 \cdot 1 + 1 \cdot 0 = 1,$$

$$(A^2)_{3,1} = 1 \cdot 0 + (-1) \cdot (-1) + 0 \cdot 1 = 1, \quad (A^2)_{3,2} = 1 \cdot 1 + (-1) \cdot 0 + 0 \cdot (-1) = 1.$$

Ainsi

$$A^2 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = J - 3I_3.$$

Chaque ligne de A a une somme de coefficients nulle, donc chaque coefficient de AJ (produit d'une ligne de A par une colonne de J , qui ne contient que des 1) est nul : $AJ = 0$. Par conséquent :

$$A^3 = A \cdot A^2 = A(J - 3I_3) = AJ - 3A = -3A.$$

Le polynôme $Q = X^3 + 3X = X(X^2 + 3)$ est donc annulateur de A . Toute valeur propre réelle λ de A vérifie $Q(\lambda) = 0$, soit $\lambda(\lambda^2 + 3) = 0$. Comme $\lambda^2 + 3 \geq 3 > 0$, on obtient $\lambda = 0$. Donc $\text{Sp}(A) \subset \{0\}$.

Réciproquement, avec $W_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, on a

$$AW_3 = \begin{pmatrix} 0 + 1 - 1 \\ -1 + 0 + 1 \\ 1 - 1 + 0 \end{pmatrix} = 0 = 0 \cdot W_3$$

et $W_3 \neq 0$, donc 0 est valeur propre de A .

Conclusion

$$A^3 = -3A \text{ et } \text{Sp}(A) = \{0\}.$$

Diagonalisabilité. Raisonnons par l'absurde : si A était diagonalisable, il existerait Q_0 inversible et D diagonale avec $A = Q_0 D Q_0^{-1}$, les coefficients diagonaux de D étant les valeurs propres de A , donc tous nuls. On aurait $D = 0$ puis $A = 0$, ce qui est faux.

Conclusion

A n'est pas diagonalisable.

2. Il s'agit de reconnaître des matrices symétriques réelles et de faire un calcul de produits.

Méthode 1 — Théorème spectral.

On a ${}^t J = J$ et

$${}^t S = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = S :$$

J et S sont des matrices symétriques réelles. D'après le théorème spectral, toute matrice symétrique réelle est diagonalisable (avec une matrice de passage orthogonale). Donc J et S sont diagonalisables.

Méthode 2 — Pour J , étude directe des espaces propres.

Les trois colonnes de J sont égales et non nulles, donc $\text{rg}(J) = 1$ et, par le théorème du

rang, $\dim \text{Ker}(J) = 3 - 1 = 2$: 0 est valeur propre de J et $\dim E_0(J) = 2$. De plus $JW_3 = 3W_3$ avec $W_3 \neq 0$, donc 3 est valeur propre et $\dim E_3(J) \geq 1$. La somme des dimensions des espaces propres vaut au moins $2 + 1 = 3 = \dim \mathcal{M}_{3,1}(\mathbb{R})$ et ne peut dépasser 3 : elle vaut 3, donc J est diagonalisable.

La méthode 1 est immédiate dès que l'on voit la symétrie ; la méthode 2 n'utilise que le rang et fournit en plus les valeurs propres de J , ce qui servira à la question 4. Pour S , seule la méthode 1 est disponible à ce stade, puisque le spectre de S n'est donné qu'à la question 3.

Commutation. Chaque ligne et chaque colonne de S a une somme de coefficients nulle : $1 - 1 + 0 = 0$, $-1 + 0 + 1 = 0$, $0 + 1 - 1 = 0$ pour les lignes, et de même pour les colonnes puisque S est symétrique. Le coefficient (i, j) de SJ est la somme des coefficients de la ligne i de S , donc $SJ = 0$. Le coefficient (i, j) de JS est la somme des coefficients de la colonne j de S , donc $JS = 0$.

Conclusion

J et S sont diagonalisables et $SJ = JS (= 0)$.

3. On veut montrer que si $SX = \lambda X$ avec $X \neq 0$, alors JX est colinéaire à X .

Méthode 1 — Commutation et dimension des espaces propres.

S est une matrice de $\mathcal{M}_3(\mathbb{R})$ possédant trois valeurs propres distinctes $0, \sqrt{3}, -\sqrt{3}$. Chaque espace propre est de dimension au moins 1, et la somme des dimensions des espaces propres est au plus 3 (les espaces propres sont en somme directe). Donc chaque espace propre de S est de dimension exactement 1.

Soit X un vecteur propre de S , associé à la valeur propre λ . Comme $SJ = JS$ (question 2) :

$$S(JX) = J(SX) = J(\lambda X) = \lambda JX.$$

Ainsi $JX \in E_\lambda(S)$. Or $E_\lambda(S)$ est de dimension 1 et contient le vecteur non nul X , donc $E_\lambda(S) = \text{Vect}(X)$. Il existe donc un réel μ tel que $JX = \mu X$; comme $X \neq 0$, X est un vecteur propre de J (associé à μ).

Méthode 2 — Calcul explicite à l'aide des sommes de coefficients.

Pour $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, on a $JX = (x_1 + x_2 + x_3)W_3$.

Cas $\lambda \neq 0$. Comme chaque colonne de S a une somme nulle, ${}^tW_3S = 0$. Si $SX = \lambda X$, alors $0 = {}^tW_3SX = \lambda {}^tW_3X = \lambda(x_1 + x_2 + x_3)$, donc $x_1 + x_2 + x_3 = 0$ et $JX = 0 = 0 \cdot X$: X est vecteur propre de J pour la valeur propre 0.

Cas $\lambda = 0$. On a $SW_3 = 0$ (sommes des lignes nulles), donc $W_3 \in E_0(S)$, et $E_0(S)$ est de dimension 1 (argument de dimension de la méthode 1), donc $E_0(S) = \text{Vect}(W_3)$. Tout vecteur propre X associé à 0 s'écrit $X = tW_3$ avec $t \neq 0$, et $JX = tJW_3 = 3tW_3 = 3X$: X est vecteur propre de J pour la valeur propre 3.

La méthode 1 est générale (deux matrices qui commutent, l'une ayant des espaces propres de dimension 1); la méthode 2 est spécifique mais donne en plus la valeur propre de J associée à chaque vecteur propre de S .

Conclusion

tout vecteur propre de S est vecteur propre de J .

4. Il faut construire une base commune de vecteurs propres.

Soient X_1, X_2, X_3 des vecteurs propres de S associés respectivement à 0, $\sqrt{3}$ et $-\sqrt{3}$. Des vecteurs propres associés à des valeurs propres distinctes forment une famille libre; la famille (X_1, X_2, X_3) est donc une famille libre de trois vecteurs de $\mathcal{M}_{3,1}(\mathbb{R})$, espace de dimension 3 : c'est une base. La matrice P dont les colonnes sont X_1, X_2, X_3 est donc inversible.

On a $SP = P \operatorname{diag}(0, \sqrt{3}, -\sqrt{3})$, donc $P^{-1}SP = \operatorname{diag}(0, \sqrt{3}, -\sqrt{3})$.

D'après la question 3, il existe des réels μ_1, μ_2, μ_3 tels que $JX_k = \mu_k X_k$ pour $k \in \{1, 2, 3\}$, donc $JP = P \operatorname{diag}(\mu_1, \mu_2, \mu_3)$ et $P^{-1}JP = \operatorname{diag}(\mu_1, \mu_2, \mu_3)$. D'après la méthode 2 de la question 3, on a d'ailleurs $\mu_1 = 3$ et $\mu_2 = \mu_3 = 0$.

Conclusion

il existe P inversible telle que $P^{-1}SP = \operatorname{diag}(0, \sqrt{3}, -\sqrt{3})$ et $P^{-1}JP = \operatorname{diag}(3, 0, 0)$ soient diagonales.

Partie 2 : Étude des matrices magiques

5. Il faut vérifier que ℓ_1 est à valeurs réelles et linéaire.

Pour $M = (m_{i,j})$ et $N = (n_{i,j})$ dans $\mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{R}$, le coefficient $(1, j)$ de $\lambda M + N$ est $\lambda m_{1,j} + n_{1,j}$, donc par linéarité de la somme finie :

$$\ell_1(\lambda M + N) = \sum_{j=1}^n (\lambda m_{1,j} + n_{1,j}) = \lambda \sum_{j=1}^n m_{1,j} + \sum_{j=1}^n n_{1,j} = \lambda \ell_1(M) + \ell_1(N).$$

Conclusion

ℓ_1 est une application linéaire de $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$, c'est-à-dire une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$.

6. On veut montrer que $K_n = \{M \in E_n, s(M) = 0\}$ est un sous-espace vectoriel de E_n .

Méthode 1 — Écriture comme intersection de sous-espaces.

Pour $M \in E_n$, on a $s(M) = \ell_1(M)$. Donc $K_n = E_n \cap \operatorname{Ker}(\ell_1)$. Le noyau de la forme linéaire ℓ_1 est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ et E_n en est un (admis) ; leur

intersection est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$, contenu dans E_n , donc un sous-espace vectoriel de E_n .

Méthode 2 — Caractérisation directe.

$K_n \subset E_n$. La matrice nulle est magique de somme nulle, donc $0 \in K_n$. Si $M, N \in K_n$ et $\lambda \in \mathbb{R}$, alors $\lambda M + N \in E_n$ car E_n est un sous-espace vectoriel, et par linéarité de s (admise) : $s(\lambda M + N) = \lambda s(M) + s(N) = 0$. Donc $\lambda M + N \in K_n$.

La méthode 1 évite toute vérification grâce au noyau ; la méthode 2 est la vérification élémentaire de la définition.

Conclusion

K_n est un sous-espace vectoriel de E_n .

7. On compare les sommes de ${}^tM$ à celles de M .

Notons ${}^tM = (m'_{i,j})$ avec $m'_{i,j} = m_{j,i}$. Pour tout i et tout j de $\{1, \dots, n\}$:

$$\ell_i({}^tM) = \sum_{j=1}^n m_{j,i} = c_i(M) = s(M), \quad c_j({}^tM) = \sum_{i=1}^n m_{j,i} = \ell_j(M) = s(M),$$

$$d_1({}^tM) = \sum_{i=1}^n m_{i,i} = d_1(M) = s(M), \quad d_2({}^tM) = \sum_{i=1}^n m'_{i,n-i+1} = \sum_{i=1}^n m_{n-i+1,i}.$$

Dans la dernière somme, le changement d'indice $k = n - i + 1$ (qui décrit $\{1, \dots, n\}$ quand i décrit $\{1, \dots, n\}$) donne

$$\sum_{k=1}^n m_{k,n-k+1} = d_2(M) = s(M).$$

Toutes les sommes de ${}^tM$ sont égales à $s(M)$.

Conclusion

$${}^tM \in E_n \text{ et } s({}^tM) = s(M).$$

8. On cherche λ tel que $M - \lambda J_n$ soit magique de somme nulle.

Toutes les sommes de lignes, colonnes et diagonales de J_n valent n (somme de n coefficients égaux à 1), donc $J_n \in E_n$ et $s(J_n) = n$.

Pour tout réel λ , $M - \lambda J_n \in E_n$ car E_n est un sous-espace vectoriel, et par linéarité de s :

$$s(M - \lambda J_n) = s(M) - \lambda s(J_n) = s(M) - n\lambda.$$

Ainsi $M - \lambda J_n \in K_n \iff s(M) - n\lambda = 0 \iff \lambda = \frac{s(M)}{n}$, ce qui est licite car $n \neq 0$.

Conclusion

il existe un unique réel λ tel que $M - \lambda J_n \in K_n$, à savoir $\lambda = \frac{s(M)}{n}$.

9. On calcule MW_n .

La i -ème coordonnée de MW_n vaut

$$\sum_{j=1}^n m_{i,j} \cdot 1 = \ell_i(M) = s(M).$$

Donc $MW_n = s(M)W_n$. Comme $W_n \neq 0$, W_n est un vecteur propre de M .

Conclusion

W_n est vecteur propre de M , associé à la valeur propre $s(M)$.

Partie 3 : Étude du cas où $n = 3$

10. On calcule les huit sommes de chaque matrice.

Pour A : lignes $0 + 1 - 1 = 0, -1 + 0 + 1 = 0, 1 - 1 + 0 = 0$; colonnes $0 - 1 + 1 = 0, 1 + 0 - 1 = 0, -1 + 1 + 0 = 0$; $d_1(A) = 0 + 0 + 0 = 0$;
 $d_2(A) = a_{1,3} + a_{2,2} + a_{3,1} = -1 + 0 + 1 = 0$.

Pour J : toutes les sommes valent 3.

Pour S : lignes et colonnes de somme nulle (question 2); $d_1(S) = 1 + 0 - 1 = 0$;
 $d_2(S) = s_{1,3} + s_{2,2} + s_{3,1} = 0 + 0 + 0 = 0$.

Conclusion

A, J, S sont magiques, avec $s(A) = 0, s(J) = 3$ et $s(S) = 0$. En particulier $A \in K_3$ et $S \in K_3$.

11. On procède par analyse-synthèse.

Analyse. Si $M = M_1 + M_2$ avec ${}^tM_1 = -M_1$ et ${}^tM_2 = M_2$, alors ${}^tM = -M_1 + M_2$. En additionnant et en soustrayant : $M_2 = \frac{1}{2}(M + {}^tM)$ et $M_1 = \frac{1}{2}(M - {}^tM)$. Cela prouve l'unicité.

Synthèse. Posons $M_1 = \frac{1}{2}(M - {}^tM)$ et $M_2 = \frac{1}{2}(M + {}^tM)$. Alors $M_1 + M_2 = M$,
 ${}^tM_1 = \frac{1}{2}({}^tM - M) = -M_1$ et ${}^tM_2 = \frac{1}{2}({}^tM + M) = M_2$. Cela prouve l'existence.

Conclusion

il existe un unique couple convenable, donné par $M_1 = \frac{1}{2}(M - {}^tM)$ (antisymétrique) et $M_2 = \frac{1}{2}(M + {}^tM)$ (symétrique).

12. (a) On utilise la question 7 et la structure d'espace vectoriel de K_3 .

Soit $M \in K_3$. D'après la question 7, ${}^tM \in E_3$ et $s({}^tM) = s(M) = 0$, donc ${}^tM \in K_3$.

Comme K_3 est un sous-espace vectoriel (question 6), les combinaisons linéaires

$M_1 = \frac{1}{2}M - \frac{1}{2}{}^tM$ et $M_2 = \frac{1}{2}M + \frac{1}{2}{}^tM$ appartiennent à K_3 .

Conclusion

$M_1 \in K_3$ et $M_2 \in K_3$.

(b) On écrit les conditions « magique de somme nulle » pour une matrice antisymétrique, puis pour une matrice symétrique.

Cas de M_1 . Une matrice antisymétrique a une diagonale nulle; on écrit

$$M_1 = \begin{pmatrix} 0 & p & q \\ -p & 0 & u \\ -q & -u & 0 \end{pmatrix}.$$

Les sommes des lignes sont nulles : $p + q = 0$, $-p + u = 0$, $-q - u = 0$. Donc $q = -p$ et $u = p$, et :

$$M_1 = \begin{pmatrix} 0 & p & -p \\ -p & 0 & p \\ p & -p & 0 \end{pmatrix} = pA.$$

Cas de M_2 . On écrit

$$M_2 = \begin{pmatrix} p & q & r \\ q & u & v \\ r & v & w \end{pmatrix}.$$

Comme $M_2 \in K_3$:

$$p + q + r = 0 \quad (L_1), \quad q + u + v = 0 \quad (L_2), \quad r + v + w = 0 \quad (L_3),$$

$$p + u + w = 0 \quad (D_1), \quad r + u + r = 0 \quad (D_2).$$

(les conditions sur les colonnes sont identiques à celles sur les lignes par symétrie).

En sommant $(L_1), (L_2), (L_3)$: $(p + u + w) + 2(q + r + v) = 0$, donc avec (D_1) : $q + r + v = 0$. Avec (L_2) : $u = r$. Avec (D_2) : $3r = 0$, donc $r = 0$ et $u = 0$. Alors (L_2) donne $v = -q$, (L_1) donne $p = -q$ et (L_3) donne $w = -v = q$. Ainsi :

$$M_2 = \begin{pmatrix} -q & q & 0 \\ q & 0 & -q \\ 0 & -q & q \end{pmatrix} = -q S.$$

Conclusion

$M_1 = \alpha A$ et $M_2 = \beta S$, où α est le coefficient $(1, 2)$ de M_1 et β l'opposé du coefficient $(1, 2)$ de M_2 .

13. On combine les questions 10, 12 et 8.

Base de K_3 . Pour $M \in K_3$, la question 12 donne $M = M_1 + M_2 = \alpha A + \beta S \in \text{Vect}(A, S)$. Réciproquement $A, S \in K_3$ (question 10) et K_3 est un sous-espace vectoriel, donc $\text{Vect}(A, S) \subset K_3$. Ainsi $K_3 = \text{Vect}(A, S)$.

La famille (A, S) est libre : si $\alpha A + \beta S = 0$, en transposant on obtient $-\alpha A + \beta S = 0$; par différence $2\alpha A = 0$, donc $\alpha = 0$ car $A \neq 0$, puis $\beta S = 0$ et $\beta = 0$ car $S \neq 0$.

Donc (A, S) est une base de K_3 et $\dim K_3 = 2$.

Base de E_3 . La famille (A, J, S) est formée d'éléments de E_3 (question 10).

Elle est génératrice de E_3 : pour $M \in E_3$, la question 8 donne

$$M - \frac{s(M)}{3} J \in K_3 = \text{Vect}(A, S), \text{ donc } M \in \text{Vect}(A, J, S).$$

Elle est libre : si $\alpha A + \beta J + \gamma S = 0$, en appliquant la forme linéaire s (sur E_3) : $0 \cdot \alpha + 3\beta + 0 \cdot \gamma = 0$, donc $\beta = 0$, puis $\alpha = \gamma = 0$ par liberté de (A, S) .

Conclusion

(A, S) est une base de K_3 et (A, J, S) est une base de E_3 (donc $\dim E_3 = 3$).

14. On montre la double inclusion, en utilisant la décomposition dans la base (A, J, S) et la non-diagonalisabilité de A .

Inclusion $\text{Vect}(J, S) \subset \Delta$. Soit $M = \beta J + \gamma S$. Alors $M \in E_3$ et

$P^{-1}MP = \beta P^{-1}JP + \gamma P^{-1}SP$ est une combinaison linéaire de matrices diagonales (question 4), donc est diagonale. Ainsi $M \in \Delta$.

Inclusion $\Delta \subset \text{Vect}(J, S)$. Soit $M \in \Delta$. Comme $M \in E_3$, on écrit $M = \alpha A + \beta J + \gamma S$ (question 13). Alors :

$$\alpha P^{-1}AP = P^{-1}MP - \beta P^{-1}JP - \gamma P^{-1}SP$$

est diagonale. Si l'on avait $\alpha \neq 0$, la matrice $P^{-1}AP = \frac{1}{\alpha} (\alpha P^{-1}AP)$ serait diagonale, donc A serait diagonalisable, ce qui contredit la question 1. Donc $\alpha = 0$ et

$$M = \beta J + \gamma S \in \text{Vect}(J, S).$$

Conclusion

$$\Delta = \text{Vect}(J, S).$$

EXERCICE 2

1. On utilise les opérations sur les fonctions de classe $\mathcal{C}^2$.

Les fonctions $(x, y) \mapsto x^2 + y$ et $(x, y) \mapsto -(x^2 + y^2)$ sont polynomiales, donc de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. La fonction $\exp$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$, donc par composition $(x, y) \mapsto e^{-(x^2+y^2)}$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. Par produit, f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

Pour $(x, y) \in \mathbb{R}^2$, en dérivant le produit :

$$\partial_1 f(x, y) = 2x e^{-(x^2+y^2)} + (x^2 + y)(-2x)e^{-(x^2+y^2)} = 2x(1 - x^2 - y) e^{-(x^2+y^2)},$$

$$\partial_2 f(x, y) = e^{-(x^2+y^2)} + (x^2 + y)(-2y)e^{-(x^2+y^2)} = (1 - 2y(x^2 + y)) e^{-(x^2+y^2)}.$$

Conclusion

f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$, avec les dérivées partielles ci-dessus.

2. On résout $\nabla f(x, y) = 0$. Comme $e^{-(x^2+y^2)} > 0$, le système équivaut à :

$$x(1 - x^2 - y) = 0 \quad \text{et} \quad 1 - 2y(x^2 + y) = 0.$$

Cas $x = 0$: la seconde équation devient $1 - 2y^2 = 0$, soit $y = \frac{1}{\sqrt{2}}$ ou $y = -\frac{1}{\sqrt{2}}$.

Cas $x \neq 0$: alors $x^2 + y = 1$, et la seconde équation devient $1 - 2y = 0$, soit $y = \frac{1}{2}$, puis $x^2 = \frac{1}{2}$, soit $x = \frac{1}{\sqrt{2}}$ ou $x = -\frac{1}{\sqrt{2}}$ (qui sont bien non nuls).

Conclusion

f admet exactement quatre points critiques : $\left(0, \frac{1}{\sqrt{2}}\right)$, $\left(0, -\frac{1}{\sqrt{2}}\right)$, $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ et $\left(-\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$.

3. On calcule la hessienne avec les formules admises, puis on étudie le signe de ses valeurs propres.

Au point $\left(0, \frac{1}{\sqrt{2}}\right)$: $x = 0$, $x^2 + y = \frac{1}{\sqrt{2}}$, $x^2 + y^2 = \frac{1}{2}$ et $2y(x^2 + y) = 1$. Donc :

$$\partial_{1,1}^2 f \left(0, \frac{1}{\sqrt{2}}\right) = 2 \left(1 - \frac{1}{\sqrt{2}}\right) e^{-1/2} = (2 - \sqrt{2}) e^{-1/2},$$

$$\partial_{2,2}^2 f \left(0, \frac{1}{\sqrt{2}}\right) = -2 \left(0 + \frac{2}{\sqrt{2}} + \frac{1}{\sqrt{2}}(1 - 1)\right) e^{-1/2} = -2\sqrt{2} e^{-1/2},$$

$$\partial_{1,2}^2 f \left(0, \frac{1}{\sqrt{2}}\right) = 0 \quad (\text{facteur } x = 0).$$

La hessienne

$$\nabla^2 f \left(0, \frac{1}{\sqrt{2}}\right) = e^{-1/2} \begin{pmatrix} 2 - \sqrt{2} & 0 \\ 0 & -2\sqrt{2} \end{pmatrix}$$

est diagonale.

Méthode 1 — Signes des valeurs propres de la hessienne.

Les valeurs propres d'une matrice diagonale sont ses coefficients diagonaux :

$(2 - \sqrt{2})e^{-1/2} > 0$ (car $\sqrt{2} < 2$) et $-2\sqrt{2}e^{-1/2} < 0$. La fonction f est de classe $\mathcal{C}^2$ sur l'ouvert $\mathbb{R}^2$, le point est critique et la hessienne possède une valeur propre strictement positive et une valeur propre strictement négative : c'est un point selle, f n'y admet pas d'extremum local.

Méthode 2 — Restrictions à deux droites.

Sur la droite verticale : $\varphi(y) = f(0, y) = ye^{-y^2}$, avec $\varphi'(y) = (1 - 2y^2)e^{-y^2}$, strictement positive sur $\left] -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right[$ et strictement négative sur $\left] \frac{1}{\sqrt{2}}, +\infty \right[$. Donc pour $y \neq \frac{1}{\sqrt{2}}$ proche de $\frac{1}{\sqrt{2}}$: $f(0, y) < f\left(0, \frac{1}{\sqrt{2}}\right)$.

Sur la droite horizontale : $\psi(x) = f\left(x, \frac{1}{\sqrt{2}}\right) = \left(x^2 + \frac{1}{\sqrt{2}}\right)e^{-x^2 - \frac{1}{2}}$, avec $\psi'(x) = 2x\left(1 - x^2 - \frac{1}{\sqrt{2}}\right)e^{-x^2 - \frac{1}{2}}$. Pour $0 < |x| < \sqrt{1 - \frac{1}{\sqrt{2}}}$, le facteur $1 - x^2 - \frac{1}{\sqrt{2}}$ est strictement positif, donc ψ' est du signe de x : ψ est strictement décroissante puis strictement croissante autour de 0, et $f\left(x, \frac{1}{\sqrt{2}}\right) > f\left(0, \frac{1}{\sqrt{2}}\right)$ pour $x \neq 0$ proche de 0.

Dans tout voisinage du point, f prend des valeurs strictement plus grandes et strictement plus petites que $f\left(0, \frac{1}{\sqrt{2}}\right)$: il n'y a pas d'extremum local.

La méthode 1 applique le critère du cours ; la méthode 2 montre concrètement les deux directions responsables du point selle et reste utilisable même lorsque la hessienne a une valeur propre nulle.

Conclusion

la hessienne en $\left(0, \frac{1}{\sqrt{2}}\right)$ est diagonale et f n'admet pas d'extremum local en ce point (point selle).

4. Même démarche au point $\left(0, -\frac{1}{\sqrt{2}}\right)$: $x = 0$, $x^2 + y = -\frac{1}{\sqrt{2}}$, $x^2 + y^2 = \frac{1}{2}$ et $2y(x^2 + y) = 1$. Donc :

$$\begin{aligned}\partial_{1,1}^2 f &= 2 \left(1 + \frac{1}{\sqrt{2}}\right) e^{-1/2} = (2 + \sqrt{2})e^{-1/2}, & \partial_{2,2}^2 f &= -2 \left(-\frac{2}{\sqrt{2}} + 0\right) e^{-1/2} \\ &= 2\sqrt{2}e^{-1/2}, & \partial_{1,2}^2 f &= 0.\end{aligned}$$

La hessienne

$$e^{-1/2} \begin{pmatrix} 2 + \sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{pmatrix}$$

est diagonale, de valeurs propres $(2 + \sqrt{2})e^{-1/2} > 0$ et $2\sqrt{2}e^{-1/2} > 0$. La fonction f est de classe $\mathcal{C}^2$ sur l'ouvert $\mathbb{R}^2$ et le point est critique ; ses deux valeurs propres sont strictement positives.

Conclusion

f admet en $\left(0, -\frac{1}{\sqrt{2}}\right)$ un minimum local strict, de valeur $f\left(0, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}}e^{-1/2} = -\frac{1}{\sqrt{2}e}$.

5. Au point $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$: $x^2 = \frac{1}{2}$, $x^2 + y = 1$, $x^2 + y^2 = \frac{3}{4}$, $1 - x^2 - y = 0$ et $1 - 2y(x^2 + y) = 0$. Donc :

$$\partial_{1,1}^2 f = 2 \left(0 \cdot 0 - 2 \cdot \frac{1}{2} \right) e^{-3/4} = -2e^{-3/4}, \quad \partial_{2,2}^2 f = -2 \left(\frac{1}{2} + 1 + \frac{1}{2} \cdot 0 \right) e^{-3/4} = -3e^{-3/4},$$

$$\partial_{1,2}^2 f = -2 \cdot \frac{1}{\sqrt{2}} \left(1 + 2 \cdot \frac{1}{2} \cdot 0 \right) e^{-3/4} = -\sqrt{2} e^{-3/4}.$$

On obtient bien $H = e^{-3/4} B$ avec $B = \begin{pmatrix} -2 & -\sqrt{2} \\ -\sqrt{2} & -3 \end{pmatrix}$.

H est une matrice symétrique réelle, donc diagonalisable dans $\mathcal{M}_2(\mathbb{R})$ par le théorème spectral. Ses valeurs propres sont celles de B multipliées par $e^{-3/4} > 0$; il suffit d'étudier le signe des valeurs propres de B .

Méthode 1 — Vecteurs propres explicites.

Avec $U_1 = \begin{pmatrix} \sqrt{2} \\ -1 \end{pmatrix}$ et $U_2 = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$:

$$\begin{aligned} BU_1 &= \begin{pmatrix} -2\sqrt{2} + \sqrt{2} \\ -2 + 3 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} \\ 1 \end{pmatrix} = -U_1, & BU_2 &= \begin{pmatrix} -2 - 2 \\ -\sqrt{2} - 3\sqrt{2} \end{pmatrix} = \begin{pmatrix} -4 \\ -4\sqrt{2} \end{pmatrix} \\ &= -4U_2. \end{aligned}$$

U_1 et U_2 sont non colinéaires, donc forment une base de $\mathcal{M}_{2,1}(\mathbb{R})$ constituée de vecteurs propres : avec $Q = (U_1 \mid U_2)$ inversible, $B = Q \operatorname{diag}(-1, -4) Q^{-1}$. Le spectre de B est donc exactement $\{-1, -4\}$ et celui de H est $\{-e^{-3/4}, -4e^{-3/4}\}$: deux valeurs strictement négatives.

Méthode 2 — Signe de la forme quadratique associée.

Pour $X = \begin{pmatrix} h \\ k \end{pmatrix}$:

$${}^tXBX = -2h^2 - 2\sqrt{2}hk - 3k^2 = -(\sqrt{2}h + k)^2 - 2k^2.$$

Si $X \neq 0$, cette quantité est strictement négative : si $k \neq 0$, le terme $-2k^2$ est strictement négatif ; si $k = 0$, alors $h \neq 0$ et $-(\sqrt{2}h)^2 < 0$. Si λ est une valeur propre de B et X un vecteur propre associé, ${}^tXBX = \lambda {}^tXX = \lambda \|X\|^2$, avec $\|X\|^2 > 0$, donc $\lambda < 0$.

La méthode 1 donne les valeurs propres exactes ; la méthode 2 ne donne que leur signe, mais elle s'applique sans avoir à deviner des vecteurs propres.

Conséquence. f est de classe $\mathcal{C}^2$ sur l'ouvert $\mathbb{R}^2$, le point $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ est critique et toutes les valeurs propres de la hessienne sont strictement négatives.

Conclusion

f admet en $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ un maximum local strict, de valeur $f\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right) = 1 \cdot e^{-3/4} = e^{-3/4}$.

Remarque : comme $f(-x, y) = f(x, y)$ pour tout (x, y) , f admet aussi un maximum local strict en $\left(-\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$, de même valeur $e^{-3/4}$.

6. (a) On majore séparément les deux facteurs de f .

Soit $(x, y) \in \mathbb{R}^2$ et $m = \max(|x|, |y|)$. On a $|x| \leq m$ et $|y| \leq m$, donc par l'inégalité triangulaire :

$$|x^2 + y| \leq x^2 + |y| \leq m^2 + m.$$

D'autre part $x^2 + y^2 \geq \max(x^2, y^2) = m^2$, donc, par croissance de $\exp$,
 $0 < e^{-(x^2+y^2)} \leq e^{-m^2}$. En multipliant ces inégalités entre nombres positifs :

$$0 \leq |f(x, y)| = |x^2 + y| e^{-(x^2+y^2)} \leq (m^2 + m) e^{-m^2}.$$

Conclusion

l'inégalité demandée est établie pour tout $(x, y) \in \mathbb{R}^2$.

(b) Posons $\varphi(u) = (u^2 + u)e^{-u^2}$ pour $u \geq 0$.

Par croissances comparées, avec $v = u^2 \rightarrow +\infty : u^2 e^{-u^2} = v e^{-v} \rightarrow 0$. Pour $u \geq 1$,
 $0 \leq u e^{-u^2} \leq u^2 e^{-u^2}$. Par encadrement, $u e^{-u^2} \rightarrow 0$, puis $\varphi(u) \rightarrow 0$ quand $u \rightarrow +\infty$.

Appliquons la définition de la limite avec le réel strictement positif $\frac{1}{2}e^{-3/4}$: il existe
 $A \geq 0$ tel que pour tout $u \geq A$, $\varphi(u) \leq \frac{1}{2}e^{-3/4}$. Posons $r = \max(A, 1) > 0$.

Si $\max(|x|, |y|) \geq r$, alors avec la question (a) appliquée à $m = \max(|x|, |y|) \geq r \geq A$:

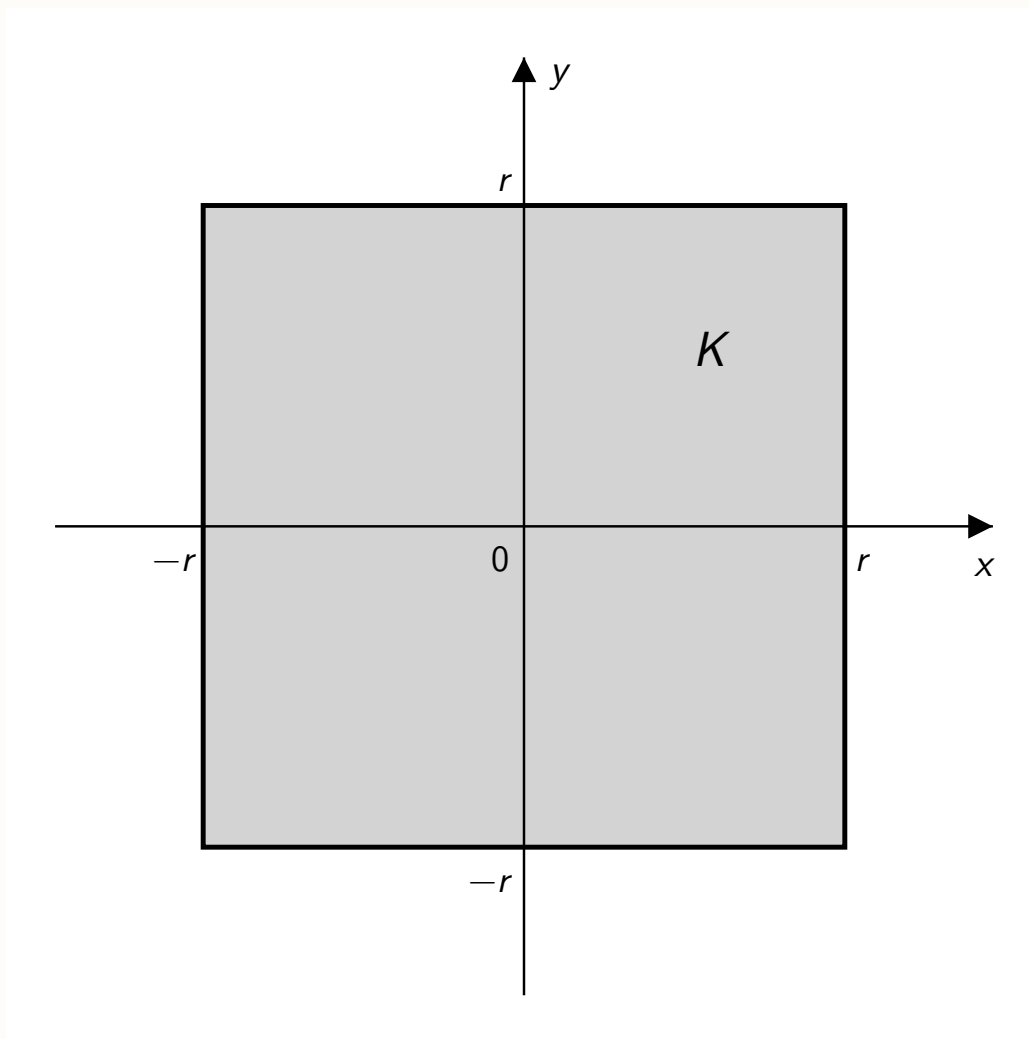
$$0 \leq |f(x, y)| \leq \varphi(m) \leq \frac{1}{2}e^{-3/4}.$$

Conclusion

il existe $r > 0$ tel que $\max(|x|, |y|) \geq r \implies 0 \leq |f(x, y)| \leq \frac{1}{2}e^{-3/4}$.

(c) Représentation. $\max(|x|, |y|) \leq r \iff (|x| \leq r \text{ et } |y| \leq r)$, donc

$\mathcal{K} = [-r, r] \times [-r, r]$: c'est le carré plein de centre l'origine, de côtés parallèles aux axes et de sommets $(\pm r, \pm r)$, bord compris.



Fermé. La fonction $m : (x, y) \mapsto \max(|x|, |y|)$ est continue sur $\mathbb{R}^2$: en effet, pour des réels a et b , $\max(a, b) = \frac{a + b + |a - b|}{2}$, donc $m(x, y) = \frac{|x| + |y| + ||x| - |y||}{2}$ s'obtient par sommes et compositions de fonctions continues (valeur absolue, projections). D'après le résultat admis dans l'énoncé, $\mathcal{K} = \{(x, y) \in \mathbb{R}^2, m(x, y) \leq r\}$ est une partie fermée de $\mathbb{R}^2$. De plus $\mathcal{K}$ est bornée, car pour $(x, y) \in \mathcal{K}$, $\sqrt{x^2 + y^2} \leq \sqrt{2}r$.

Conclusion

$\mathcal{K}$ est le carré plein $[-r, r]^2$; c'est une partie fermée (et bornée) de $\mathbb{R}^2$.

(d) On compare les valeurs de f aux points critiques avec la majoration de (b), puis on utilise le théorème des extrema sur un fermé borné.

Valeurs aux points critiques :

$$f\left(\pm\frac{1}{\sqrt{2}}, \frac{1}{2}\right) = e^{-3/4}, \quad f\left(0, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}e^{-1/2}, \quad f\left(0, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}}e^{-1/2}.$$

On a $e^{-3/4} > \frac{1}{2}e^{-3/4}$, et $\frac{1}{\sqrt{2}}e^{-1/2} > \frac{1}{2}e^{-3/4} \iff \sqrt{2}e^{1/4} > 1$, ce qui est vrai car $\sqrt{2} > 1$ et $e^{1/4} > 1$. Ainsi $|f| > \frac{1}{2}e^{-3/4}$ en chacun des quatre points critiques. Par contraposée de (b), chaque point critique vérifie $\max(|x|, |y|) < r$: tous les points critiques appartiennent à $\mathcal{K}$.

Comparaison utile : $e^{-3/4} > \frac{1}{\sqrt{2}}e^{-1/2} \iff \sqrt{2} > e^{1/4} \iff 4 > e$, ce qui est vrai.

Maximum global. f est continue sur $\mathcal{K}$, partie fermée bornée non vide de $\mathbb{R}^2$: f y admet un maximum $M_{\mathcal{K}}$, atteint en un point de $\mathcal{K}$. Comme $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right) \in \mathcal{K}$, $M_{\mathcal{K}} \geq e^{-3/4}$.

Pour $(x, y) \notin \mathcal{K}$, on a $\max(|x|, |y|) > r$ donc $f(x, y) \leq |f(x, y)| \leq \frac{1}{2}e^{-3/4} < M_{\mathcal{K}}$. Ainsi $M_{\mathcal{K}}$ est le maximum global de f sur $\mathbb{R}^2$, et il n'est atteint qu'en des points de $\mathcal{K}$.

Soit (x_0, y_0) un point où ce maximum global est atteint. f est de classe $\mathcal{C}^1$ sur l'ouvert $\mathbb{R}^2$ et admet en (x_0, y_0) un maximum (global, donc local) : par la condition nécessaire du premier ordre, (x_0, y_0) est un point critique. Donc $M_{\mathcal{K}}$ est l'une des valeurs critiques $e^{-3/4}$, $\frac{1}{\sqrt{2}}e^{-1/2}$, $-\frac{1}{\sqrt{2}}e^{-1/2}$; comme $M_{\mathcal{K}} \geq e^{-3/4}$ et que $e^{-3/4}$ est la plus grande de ces valeurs, $M_{\mathcal{K}} = e^{-3/4}$, atteint exactement aux points critiques de valeur $e^{-3/4}$.

Minimum global. De même, f admet sur $\mathcal{K}$ un minimum

$$m_{\mathcal{K}} \leq f\left(0, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}}e^{-1/2} < -\frac{1}{2}e^{-3/4}. \text{ Pour } (x, y) \notin \mathcal{K},$$

$f(x, y) \geq -|f(x, y)| \geq -\frac{1}{2}e^{-3/4} > m_{\mathcal{K}}$. Donc $m_{\mathcal{K}}$ est le minimum global de f sur $\mathbb{R}^2$; tout point où il est atteint est critique, et $m_{\mathcal{K}} < 0$ est une valeur critique strictement négative : la seule est $-\frac{1}{\sqrt{2}}e^{-1/2}$, atteinte uniquement en $\left(0, -\frac{1}{\sqrt{2}}\right)$.

Conclusion

f admet sur $\mathbb{R}^2$ un maximum global égal à $e^{-3/4}$, atteint exactement en $\left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ et $\left(-\frac{1}{\sqrt{2}}, \frac{1}{2}\right)$, et un minimum global égal à $-\frac{1}{\sqrt{2}e}$, atteint uniquement en $\left(0, -\frac{1}{\sqrt{2}}\right)$. Ce sont les seuls extrema globaux.

7. Il s'agit de lire la figure, puis de confirmer la lecture par un calcul.

Lecture de la figure. Le gradient $\nabla f(x, y)$ indique la direction dans laquelle f croît le plus vite à partir de (x, y) . Au point $(1, 0)$, la flèche est verticale, dirigée vers le haut, et non réduite à un point; elle est donc tangente au cercle $\mathcal{C}$ (dont la tangente en $(1, 0)$ est verticale). En se déplaçant sur $\mathcal{C}$ vers le haut à partir de $(1, 0)$, f semble augmenter, et en se déplaçant vers le bas, diminuer : f ne semble donc pas admettre d'extremum sous la contrainte en $(1, 0)$.

Confirmation par le calcul.

$$\nabla f(1, 0) = (2 \cdot 1 \cdot (1 - 1 - 0)e^{-1}, (1 - 0)e^{-1}) = (0, e^{-1}) :$$

c'est bien un vecteur vertical non nul. Pour $y \in]-1, 1[$, le point $(\sqrt{1 - y^2}, y)$ appartient à $\mathcal{C}$ et tend vers $(1, 0)$ quand $y \rightarrow 0$. On a :

$$\theta(y) = f(\sqrt{1 - y^2}, y) = (1 - y^2 + y)e^{-1}, \quad \theta'(y) = (1 - 2y)e^{-1}.$$

Pour $y \in]-1, \frac{1}{2}[$, $\theta'(y) > 0$: θ est strictement croissante sur cet intervalle qui contient

0. Donc, pour $y \in]0, \frac{1}{2}[$, $f(\sqrt{1 - y^2}, y) > f(1, 0)$, et pour $y \in]-1, 0[$, $f(\sqrt{1 - y^2}, y) < f(1, 0)$, avec des points de $\mathcal{C}$ aussi proches que l'on veut de $(1, 0)$.

Conclusion

la figure suggère, et le calcul confirme, que f n'admet pas d'extremum sous la contrainte $x^2 + y^2 = 1$ au point $(1, 0)$.

8. On étudie les variations de g sur $[-1, 1]$.

g est dérivable et $g'(y) = 1 - 2y$: $g' > 0$ sur $[-1, \frac{1}{2}[$ et $g' < 0$ sur $]\frac{1}{2}, 1]$. Donc g est strictement croissante sur $[-1, \frac{1}{2}]$ et strictement décroissante sur $[\frac{1}{2}, 1]$, avec :

$$g(-1) = 1 - 1 - 1 = -1, \quad g\left(\frac{1}{2}\right) = 1 + \frac{1}{2} - \frac{1}{4} = \frac{5}{4}, \quad g(1) = 1.$$

y	-1		$\frac{1}{2}$		1
$g'(y)$		+	0	—	
$g(y)$	-1	croît	$\frac{5}{4}$	décroît	1

Conclusion

sur $[-1, 1]$, g admet un maximum global $\frac{5}{4}$ atteint uniquement en $\frac{1}{2}$, un minimum global -1 atteint uniquement en -1 , et un minimum local (non global) en 1 , de valeur 1 : pour $y \in [\frac{1}{2}, 1]$, $g(y) \geq g(1)$ par décroissance. Il n'y a pas d'autre extremum local, car g est strictement monotone au voisinage de tout autre point de $[-1, 1]$.

9. On paramètre le cercle par l'ordonnée, ce qui ramène l'étude à celle de g .

Pour $(x, y) \in \mathcal{C}$, on a $x^2 = 1 - y^2$, donc :

$$f(x, y) = (1 - y^2 + y) e^{-1} = \frac{g(y)}{e}.$$

De plus, quand (x, y) décrit $\mathcal{C}$, y décrit exactement $[-1, 1]$, car pour tout $y \in [-1, 1]$, le point $(\sqrt{1-y^2}, y)$ est sur $\mathcal{C}$.

On dit que f admet en $(x_0, y_0) \in \mathcal{C}$ un maximum (resp. minimum) local sous la contrainte s'il existe $\delta > 0$ tel que $f(x, y) \leq f(x_0, y_0)$ (resp. $\geq$) pour tout $(x, y) \in \mathcal{C}$ à distance au plus δ de (x_0, y_0) .

Extrema globaux sous contrainte. D'après la question 8, sur $\mathcal{C}$, f a pour maximum $\frac{5}{4e}$, atteint exactement aux points de $\mathcal{C}$ d'ordonnée $\frac{1}{2}$, c'est-à-dire $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ et $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$, et pour minimum $-\frac{1}{e}$, atteint exactement au point d'ordonnée -1 , c'est-à-dire $(0, -1)$.

Le point $(0, 1)$. Si $(x, y) \in \mathcal{C}$ est à distance au plus $\frac{1}{2}$ de $(0, 1)$, alors $|y - 1| \leq \frac{1}{2}$, donc $y \in [\frac{1}{2}, 1]$ et $g(y) \geq g(1)$ (question 8). Ainsi $f(x, y) \geq f(0, 1) = \frac{1}{e}$: minimum local sous contrainte, non global puisque $-\frac{1}{e} < \frac{1}{e}$.

Les autres points. Soit $(x_0, y_0) \in \mathcal{C}$ avec $y_0 \notin \{-1, \frac{1}{2}, 1\}$. Alors $x_0 \neq 0$ (sinon $y_0 = \pm 1$) ; notons $\sigma \in \{-1, 1\}$ le signe de x_0 . Les points $(\sigma\sqrt{1-y^2}, y)$, pour $y \in]-1, 1[$, sont sur $\mathcal{C}$ et tendent vers (x_0, y_0) quand $y \rightarrow y_0$, par continuité de la racine carrée. Comme $g'(y_0) = 1 - 2y_0 \neq 0$ et que g' est continue, g est strictement monotone sur un intervalle ouvert contenant y_0 : dans tout voisinage de (x_0, y_0) , il existe des points de $\mathcal{C}$ où f est strictement supérieure, et d'autres où elle est strictement inférieure, à $f(x_0, y_0)$. Il n'y a pas d'extremum en ces points.

Conclusion

sous la contrainte $x^2 + y^2 = 1$, f admet :

- un maximum global $\frac{5}{4e}$ en $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ et $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$;
- un minimum global $-\frac{1}{e}$ en $(0, -1)$;

- un minimum local non global $\frac{1}{e}$ en $(0, 1)$;

et aucun autre extremum sous cette contrainte. En particulier, ce résultat est cohérent avec la question 7 pour le point $(1, 0)$.

Commentaire au vu de la figure 1. Aux quatre points obtenus, les flèches tracées sur la figure sont dirigées selon le rayon du cercle, c'est-à-dire orthogonalement à $\mathcal{C}$: par exemple $\nabla f(0, 1) = (0, -e^{-1})$ et $\nabla f(0, -1) = (0, -e^{-1})$ sont verticaux, et $\nabla f\left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right) = -\frac{1}{2e} \left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ est colinéaire au rayon. Le gradient n'a alors aucune composante le long du cercle : on ne peut pas faire croître ou décroître f « au premier ordre » en se déplaçant sur $\mathcal{C}$. Ailleurs sur $\mathcal{C}$, comme en $(1, 0)$, les flèches traversent le cercle en biais ou lui sont tangentes. On peut le vérifier par le calcul : avec le vecteur tangent $T = (-y, x)$, pour $(x, y) \in \mathcal{C}$, en utilisant $1 - x^2 - y = y^2 - y$ et $x^2 + y = 1 - y^2 + y$:

$$\nabla f(x, y) \cdot T = e^{-1} \left(-2xy(y^2 - y) + x(1 - 2y(1 - y^2 + y)) \right) = e^{-1}x(1 - 2y),$$

qui s'annule exactement aux quatre points $(0, \pm 1)$ et $\left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right)$. Aux maxima, les flèches pointent vers l'intérieur du disque : f y augmente en quittant le cercle vers l'intérieur, mais pas le long du cercle.

PROBLÈME

Partie 1 : Estimateur du maximum de vraisemblance

1. (a) Question adaptée au programme actuel. Code adapté en Python.

On simule n variables indépendantes de loi $\mathcal{U}([0, a])$ par `np.random.uniform(0, a, n)`, puis on renvoie le maximum des coefficients.

Python

```
import numpy as np

def sim_V(n, a):
    X = np.random.uniform(0, a, n)    # tableau de n réalisations indépendantes de U([0, a
    ])
    V = np.max(X)                      # réalisation de  $V_n = \max(X_1, \dots, X_n)$ 
    return V
```

Variables : X contient une réalisation de $(X_1, \dots, X_n)$; V est le plus grand de ses coefficients, donc une réalisation de V_n . Ce code a été exécuté : sur 20000 appels avec $n = 10$ et $a = 3$, la moyenne empirique obtenue vaut environ 2,728, en accord avec la valeur $\frac{na}{n+1} \approx 2,727$ établie à la question 3 (observation numérique, qui ne constitue pas une preuve).

(b) On lit sur la figure 2 le comportement des cinq trajectoires.

Chaque trajectoire est croissante (ce qui est normal : $V_{n+1} = \max(V_n, X_{n+1}) \geq V_n$), reste toujours en dessous de la droite d'ordonnée $1 = a$, et se rapproche de 1 : dès $n \approx 30$, toutes les valeurs sont supérieures à 0,83, et vers $n = 100$ elles sont comprises entre 0,97 et 1.

Conclusion (conjecture) : V_n semble converger vers a (estimateur convergent), en sous-estimant systématiquement a ; on peut donc conjecturer que V_n est biaisé, avec $E(V_n) < a$.

2. (a) X_1 suit la loi $\mathcal{U}([0, a])$, de densité $t \mapsto \frac{1}{a}$ sur $[0, a]$ et nulle ailleurs. Sa fonction de répartition est :

$$F_{X_1}(t) = \begin{cases} 0 & \text{si } t < 0, \\ \frac{t}{a} & \text{si } 0 \leq t \leq a, \\ 1 & \text{si } t > a. \end{cases}$$

(b) On utilise l'égalité d'événements $(V_n \leq t) = \bigcap_{i=1}^n (X_i \leq t)$, valable car un maximum

est inférieur ou égal à t si et seulement si chacun des termes l'est.

Par indépendance mutuelle des X_i , puis parce qu'elles ont toutes la loi de X_1 : pour tout réel t ,

$$F_n(t) = P(V_n \leq t) = \prod_{i=1}^n P(X_i \leq t) = (F_{X_1}(t))^n.$$

Conclusion

$$F_n(t) = \begin{cases} 0 & \text{si } t < 0, \\ \left(\frac{t}{a}\right)^n & \text{si } 0 \leq t \leq a, \\ 1 & \text{si } t > a. \end{cases}$$

(c) On vérifie les hypothèses de la caractérisation des variables à densité.

F_n est continue sur $\mathbb{R}$: elle l'est sur chacun des intervalles $] -\infty, 0[,]0, a[,]a, +\infty[$, et en 0 et a les raccords donnent $\left(\frac{0}{a}\right)^n = 0$ et $\left(\frac{a}{a}\right)^n = 1$. Elle est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé des deux points 0 et a . Donc V_n est une variable à densité, et une densité s'obtient en dérivant F_n là où elle est dérivable et en choisissant des valeurs arbitraires (positives) aux points restants.

Conclusion

V_n est à densité, de densité $f_n(t) = \frac{n t^{n-1}}{a^n}$ si $t \in [0, a]$ et $f_n(t) = 0$ sinon.

3. f_n est nulle hors du segment $[0, a]$ et continue sur ce segment, donc

$\int_{-\infty}^{+\infty} |t| f_n(t) dt = \int_0^a t f_n(t) dt$ converge : V_n admet une espérance, et :

$$E(V_n) = \int_0^a t \cdot \frac{n t^{n-1}}{a^n} dt = \frac{n}{a^n} \left[\frac{t^{n+1}}{n+1} \right]_0^a = \frac{n}{n+1} a.$$

Le biais vaut $E(V_n) - a = -\frac{a}{n+1} \neq 0$ car $a > 0$.

Conclusion

$E(V_n) = \frac{na}{n+1}$; l'estimateur V_n n'est pas sans biais (il sous-estime a en moyenne), mais son biais tend vers 0.

4. On remarque d'abord que $P(V_n > a) = 1 - F_n(a) = 0$.

L'événement $(|V_n - a| \geq \varepsilon)$ est la réunion disjointe de $(V_n \leq a - \varepsilon)$ et de $(V_n \geq a + \varepsilon)$, ce dernier étant inclus dans $(V_n > a)$, de probabilité nulle. Donc :

$$P(|V_n - a| \geq \varepsilon) = P(V_n \leq a - \varepsilon) = F_n(a - \varepsilon).$$

Si $\varepsilon \geq a$, alors $a - \varepsilon \leq 0$ et $F_n(a - \varepsilon) = 0$. Si $0 < \varepsilon < a$, alors $F_n(a - \varepsilon) = \left(1 - \frac{\varepsilon}{a}\right)^n$, et comme $0 < 1 - \frac{\varepsilon}{a} < 1$, cette quantité tend vers 0 quand $n \rightarrow +\infty$.

Dans tous les cas, $P(|V_n - a| \geq \varepsilon) \rightarrow 0$ pour tout $\varepsilon > 0$.

Conclusion

$P(|V_n - a| \geq \varepsilon) = F_n(a - \varepsilon)$ et (V_n) converge en probabilité vers a : l'estimateur V_n est convergent.

5. On exprime la fonction de répartition de $Y_n = n(a - V_n)$, puis on calcule sa limite.

Pour tout réel t : $(n(a - V_n) \leq t) = \left(V_n \geq a - \frac{t}{n}\right)$. Comme V_n est à densité,

$P\left(V_n = a - \frac{t}{n}\right) = 0$, donc :

$$P(n(a - V_n) \leq t) = 1 - P\left(V_n < a - \frac{t}{n}\right) = 1 - F_n\left(a - \frac{t}{n}\right).$$

Calcul selon t .

Si $t < 0$: $a - \frac{t}{n} > a$, donc $F_n\left(a - \frac{t}{n}\right) = 1$ et $P(Y_n \leq t) = 0$.

Si $t \geq 0$: pour $n \geq \frac{t}{a}$ (ce qui est vrai pour n assez grand), $0 \leq a - \frac{t}{n} \leq a$, donc

$$P(Y_n \leq t) = 1 - \left(1 - \frac{t}{na}\right)^n.$$

Pour $t = 0$, cette quantité vaut 0. Pour $t > 0$ et $n > \frac{t}{a}$, $1 - \frac{t}{na} \in]0, 1[$ et :

$$\left(1 - \frac{t}{na}\right)^n = \exp\left(n \ln\left(1 - \frac{t}{na}\right)\right).$$

Comme $\frac{t}{na} \rightarrow 0$ et $\ln(1 + u) \sim u$ quand $u \rightarrow 0$, on a

$n \ln\left(1 - \frac{t}{na}\right) \sim n \cdot \left(-\frac{t}{na}\right) = -\frac{t}{a}$, donc $n \ln\left(1 - \frac{t}{na}\right) \rightarrow -\frac{t}{a}$, et par continuité de $\exp$: $\left(1 - \frac{t}{na}\right)^n \rightarrow e^{-t/a}$.

Ainsi, pour tout réel t :

$$\lim_{n \rightarrow +\infty} P(n(a - V_n) \leq t) = G(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1 - e^{-t/a} & \text{si } t \geq 0. \end{cases}$$

G est la fonction de répartition de la loi exponentielle de paramètre $\frac{1}{a}$ (d'espérance a).
La convergence a lieu en tout réel, donc en particulier en tout point de continuité de G .

Conclusion

$P(n(a - V_n) \leq t) = 1 - F_n\left(a - \frac{t}{n}\right)$ et $(n(a - V_n))_{n \geq 1}$ converge en loi vers une variable de loi exponentielle $\mathcal{E}\left(\frac{1}{a}\right)$.

6. On cherche un intervalle I_n , dont les bornes ne dépendent que de V_n , n et α , tel que $P(a \in I_n) \rightarrow 1 - \alpha$.

Posons $c = \ln\left(\frac{1}{\alpha}\right) = -\ln \alpha > 0$, de sorte que $1 - e^{-c} = 1 - \alpha$. Pour $n > c$, posons :

$$I_n = \left[V_n, \frac{n V_n}{n - c} \right].$$

Comme $n - c > 0$: $a \leq \frac{n V_n}{n - c} \iff a(n - c) \leq n V_n \iff n(a - V_n) \leq ac$. Donc :

$$(a \in I_n) = (V_n \leq a) \cap (n(a - V_n) \leq ac).$$

Comme $P(V_n > a) = 0$ (question 4), on obtient $P(a \in I_n) = P(n(a - V_n) \leq ac)$. En appliquant la question 5 au réel fixé $t = ac \geq 0$:

$$P(a \in I_n) \xrightarrow{n \rightarrow +\infty} 1 - e^{-ac/a} = 1 - e^{-c} = 1 - \alpha.$$

Conclusion

pour $n > \ln\left(\frac{1}{\alpha}\right)$, $\left[V_n, \frac{n V_n}{n - \ln(1/\alpha)} \right]$ est un intervalle de confiance asymptotique de a au niveau de confiance $1 - \alpha$.

Remarque : la quantité $\frac{n(a - V_n)}{a}$ converge en loi vers la loi $\mathcal{E}(1)$, qui ne dépend pas de a ; c'est ce qui permet d'isoler a sans remplacer a par une estimation dans une convergence en loi. Une simulation de contrôle (exécutée, $a = 3$, $n = 50$, $\alpha = 0,05$, 200000 répétitions) donne une fréquence de recouvrement d'environ 0,955, compatible avec ce résultat.

7. (a) f_n est nulle hors de $[0, a]$ et continue sur $[0, a]$, donc $\int_0^a t^2 f_n(t) dt$ converge (intégrale sur un segment) : par le théorème de transfert, V_n admet un moment d'ordre 2 et :

$$E(V_n^2) = \int_0^a t^2 \frac{n t^{n-1}}{a^n} dt = \frac{n}{a^n} \left[\frac{t^{n+2}}{n+2} \right]_0^a = \frac{n}{n+2} a^2.$$

Conclusion

$$E(V_n^2) = \frac{na^2}{n+2}.$$

(b) Question adaptée au programme actuel.

V_n admet un moment d'ordre 2, donc $(V_n - a)^2 = V_n^2 - 2aV_n + a^2$ admet une espérance, et par linéarité :

$$E((V_n - a)^2) = \frac{na^2}{n+2} - 2a \cdot \frac{na}{n+1} + a^2 = a^2 \frac{n(n+1) - 2n(n+2) + (n+1)(n+2)}{(n+1)(n+2)}.$$

Le numérateur vaut $n^2 + n - 2n^2 - 4n + n^2 + 3n + 2 = 2$. Donc :

$$E((V_n - a)^2) = \frac{2a^2}{(n+1)(n+2)}.$$

Cette quantité tend vers 0 ; elle permet de retrouver la convergence de l'estimateur V_n (question 4).

Méthode 1 — Inégalité de Markov.

Soit $\varepsilon > 0$. La variable $(V_n - a)^2$ est positive et admet une espérance ; par l'inégalité de Markov :

$$P(|V_n - a| \geq \varepsilon) = P((V_n - a)^2 \geq \varepsilon^2) \leq \frac{E((V_n - a)^2)}{\varepsilon^2} = \frac{2a^2}{\varepsilon^2(n+1)(n+2)} \xrightarrow{n \rightarrow +\infty} 0.$$

Méthode 2 — Condition suffisante par l'espérance et la variance.

Par la formule de Koenig-Huygens appliquée à $V_n - a$:

$$V(V_n) = V(V_n - a) = E((V_n - a)^2) - (E(V_n) - a)^2, \text{ donc } 0 \leq V(V_n) \leq E((V_n - a)^2) \rightarrow 0.$$

Comme de plus $E(V_n) = \frac{na}{n+1} \rightarrow a$, le critère du cours (espérance tendant vers le paramètre et variance tendant vers 0) montre que V_n est un estimateur convergent de a .

La méthode 1 utilise directement la quantité calculée ; la méthode 2 la relie au critère usuel des estimateurs asymptotiquement sans biais.

Conclusion

$$E((V_n - a)^2) = \frac{2a^2}{(n+1)(n+2)}, \text{ ce qui redonne la convergence de } V_n \text{ vers } a \text{ en probabilité.}$$

Partie 2 : Méthode des moments

8. Question adaptée au programme actuel. Code adapté en Python.

Python

```
import numpy as np

def sim_M(n, a):
    X = np.random.uniform(0, a, n) # réalisation de (X_1, ..., X_n)
    y = 2 * np.mean(X)             # réalisation de M_n = 2 * moyenne empirique
    return y
```

La moyenne des coefficients de X est une réalisation de $\bar{X}_n$, et y une réalisation de $M_n = 2\bar{X}_n$. Ce code a été exécuté : sur 20000 appels avec $n = 10$ et $a = 3$, la moyenne des valeurs obtenues vaut environ 2,99, et la moyenne des carrés des écarts à a environ

0,297, en accord avec $E(M_n) = a = 3$ et $E((M_n - a)^2) = \frac{a^2}{3n} = 0,3$ (observation numérique, qui ne remplace pas les preuves des questions 9 et 10).

9. Les X_i suivent la loi $\mathcal{U}([0, a])$: elles admettent une espérance $\frac{a}{2}$ et une variance $\frac{a^2}{12}$.

Par linéarité de l'espérance :

$$E(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{a}{2}.$$

Les X_i étant indépendantes, la variance de leur somme est la somme des variances :

$$V(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \cdot n \cdot \frac{a^2}{12} = \frac{a^2}{12n}.$$

Par linéarité, $E(M_n) = 2E(\bar{X}_n) = a$.

Conclusion

$$E(\bar{X}_n) = \frac{a}{2}, \quad V(\bar{X}_n) = \frac{a^2}{12n}, \text{ et } M_n \text{ est un estimateur sans biais de } a.$$

10. Question adaptée au programme actuel.

Comme $E(M_n) = a$, on a

$$E((M_n - a)^2) = E((M_n - E(M_n))^2) = V(M_n) = 4V(\bar{X}_n) = \frac{a^2}{3n}.$$

Convergence.

Méthode 1 — Inégalité de Bienaymé-Tchebychev.

M_n admet une variance ; pour tout $\varepsilon > 0$:

$$P(|M_n - a| \geq \varepsilon) = P(|M_n - E(M_n)| \geq \varepsilon) \leq \frac{V(M_n)}{\varepsilon^2} = \frac{a^2}{3n\varepsilon^2} \rightarrow 0.$$

Méthode 2 — Loi faible des grands nombres.

Les X_i sont indépendantes, de même loi, admettent une espérance $\frac{a}{2}$ et une variance : par la loi faible des grands nombres, $\bar{X}_n$ converge en probabilité vers $\frac{a}{2}$. La fonction $x \mapsto 2x$ est continue sur $\mathbb{R}$, donc par composition $M_n = 2\bar{X}_n$ converge en probabilité vers $2 \cdot \frac{a}{2} = a$.

La méthode 1 donne en plus une majoration explicite de la probabilité d'erreur ; la méthode 2 est plus rapide mais purement qualitative.

Conclusion

$$E((M_n - a)^2) = \frac{a^2}{3n} \text{ et } M_n \text{ est un estimateur convergent de } a.$$

11. On applique le théorème limite central à la suite (X_n) .

Les X_i sont indépendantes, de même loi, d'espérance $m = \frac{a}{2}$ et de variance $\sigma^2 = \frac{a^2}{12} > 0$ (car $a > 0$) ; on note $\sigma = \frac{a}{2\sqrt{3}}$. D'après le théorème limite central :

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \frac{a}{2})}{\sigma} \text{ converge en loi vers une variable } Z \text{ de loi } \mathcal{N}(0, 1).$$

$$\text{Or } \sqrt{n}(M_n - a) = 2\sqrt{n}\left(\bar{X}_n - \frac{a}{2}\right) = 2\sigma Z_n = \frac{a}{\sqrt{3}}Z_n.$$

La fonction $x \mapsto \frac{a}{\sqrt{3}}x$ est continue sur $\mathbb{R}$; par composition de la convergence en loi par une fonction continue, $\sqrt{n}(M_n - a)$ converge en loi vers $\frac{a}{\sqrt{3}}Z$. Une transformation affine d'une variable normale est normale : $\frac{a}{\sqrt{3}}Z$ suit la loi $\mathcal{N}\left(0, \frac{a^2}{3}\right)$.

On peut aussi le vérifier directement : pour tout réel t ,

$$P\left(\frac{a}{\sqrt{3}}Z_n \leq t\right) = P\left(Z_n \leq \frac{\sqrt{3}t}{a}\right) \rightarrow \Phi\left(\frac{\sqrt{3}t}{a}\right), \text{ qui est la fonction de répartition de la loi } \mathcal{N}\left(0, \frac{a^2}{3}\right).$$

Conclusion

$$(\sqrt{n}(M_n - a))_{n \geq 1} \text{ converge en loi vers la loi normale } \mathcal{N}\left(0, \frac{a^2}{3}\right).$$

12. On utilise une quantité dont la loi limite ne dépend pas de a , puis on isole a par un calcul exact.

D'après la question 11, $T_n = \frac{\sqrt{3n}(M_n - a)}{a} = Z_n$ converge en loi vers $\mathcal{N}(0, 1)$, de fonction de répartition Φ continue sur $\mathbb{R}$.

Soit u_α l'unique réel tel que $\Phi(u_\alpha) = 1 - \frac{\alpha}{2}$ (Φ est une bijection continue strictement croissante de $\mathbb{R}$ sur $]0, 1[$); on a $u_\alpha > 0$. Notons $P_n(t) = P(T_n \leq t)$. Alors :

$$P(-u_\alpha < T_n \leq u_\alpha) = P_n(u_\alpha) - P_n(-u_\alpha) \xrightarrow{n \rightarrow +\infty} \Phi(u_\alpha) - \Phi(-u_\alpha) = 2\Phi(u_\alpha) - 1 = 1 - \alpha.$$

Isolons a . Posons $\delta_n = \frac{u_\alpha}{\sqrt{3n}}$ et supposons $n > \frac{u_\alpha^2}{3}$, de sorte que $0 < \delta_n < 1$. Comme $a > 0$:

$$T_n \leq u_\alpha \iff M_n - a \leq a\delta_n \iff a \geq \frac{M_n}{1 + \delta_n},$$

$$T_n > -u_\alpha \iff M_n - a > -a\delta_n \iff M_n > a(1 - \delta_n) \iff a < \frac{M_n}{1 - \delta_n}.$$

Donc, avec $J_n = \left[\frac{M_n}{1 + \delta_n}, \frac{M_n}{1 - \delta_n} \right]$, on a $P(a \in J_n) = P(-u_\alpha < T_n \leq u_\alpha) \rightarrow 1 - \alpha$.

Remarque : M_n est une variable à densité (somme de variables indépendantes à densité, par produit de convolution), donc $P(T_n = -u_\alpha) = 0$ et l'on peut aussi prendre l'intervalle fermé $\left[\frac{M_n}{1 + \delta_n}, \frac{M_n}{1 - \delta_n} \right]$, de même probabilité de contenir a .

Conclusion

pour $n > \frac{u_\alpha^2}{3}$, $\left[\frac{M_n}{1 + u_\alpha/\sqrt{3n}}, \frac{M_n}{1 - u_\alpha/\sqrt{3n}} \right]$ est un intervalle de confiance asymptotique de a au niveau $1 - \alpha$. Une simulation de contrôle (exécutée, $a = 3$, $n = 50$, $\alpha = 0,05$) donne une fréquence de recouvrement d'environ 0,951.

Comparaison avec la question 6. Les deux intervalles ont le même niveau asymptotique ; on compare leurs longueurs.

Pour I_n (question 6) : $\ell(I_n) = \frac{nV_n}{n-c} - V_n = \frac{cV_n}{n-c} \leq \frac{a \ln(1/\alpha)}{n - \ln(1/\alpha)}$, car $V_n \leq a$ presque sûrement : la longueur est de l'ordre de $\frac{1}{n}$.

Pour J_n : $\ell(J_n) = M_n \left(\frac{1}{1 - \delta_n} - \frac{1}{1 + \delta_n} \right) = \frac{2\delta_n M_n}{1 - \delta_n^2}$, avec $\delta_n = \frac{u_\alpha}{\sqrt{3n}}$ et M_n proche de a pour n grand : la longueur est de l'ordre de $\frac{2u_\alpha a}{\sqrt{3n}}$, donc de l'ordre de $\frac{1}{\sqrt{n}}$.

Conclusion

pour n grand, l'intervalle de la question 6 est nettement plus court à niveau égal ; il est donc meilleur.

13. Question adaptée au programme actuel.

On compare $\frac{2a^2}{(n+1)(n+2)}$ et $\frac{a^2}{3n}$ pour $n \geq 1$. Comme $a^2 > 0$ et les dénominateurs

sont strictement positifs :

$$\begin{aligned} \frac{2a^2}{(n+1)(n+2)} &\leq \frac{a^2}{3n} \iff 6n \leq (n+1)(n+2) \iff n^2 - 3n + 2 \geq 0 \\ &\iff (n-1)(n-2) \geq 0. \end{aligned}$$

Pour un entier $n \geq 1$, $(n-1)(n-2) \geq 0$, avec égalité si et seulement si $n \in \{1, 2\}$.

Conclusion

$E((V_n - a)^2) \leq E((M_n - a)^2)$ pour tout $n \geq 1$, l'inégalité étant stricte dès $n \geq 3$. De plus $E((V_n - a)^2)$ est de l'ordre de $\frac{2a^2}{n^2}$ alors que $E((M_n - a)^2) = \frac{a^2}{3n}$: l'écart quadratique moyen de V_n à a tend vers 0 beaucoup plus vite.

Commentaire de la figure 3. À gauche, les trajectoires de V_n sont croissantes, toujours inférieures à $1 = a$, et se resserrent très vite contre la droite d'ordonnée 1 : à partir de $n \approx 40$, elles sont toutes au-dessus de 0,95 environ. À droite, les trajectoires de M_n oscillent de part et d'autre de 1, avec de fortes fluctuations pour les petites valeurs de n (valeurs entre 0,25 et 1,9 environ au départ) et encore des écarts à 1 pouvant atteindre 0,05 à 0,08 vers $n = 100$. La figure illustre donc le calcul : V_n , bien que biaisé, est beaucoup plus proche de a que M_n , qui est sans biais mais plus dispersé.

Partie 3 : Consistance de ces estimateurs

Dans toute cette partie, comme $P(0 \leq X_1 \leq 2a) = 1$ et $P(0 \leq X_i \leq a) = 1$ pour $i \geq 2$, on suppose, comme il est d'usage, que X_1 prend ses valeurs dans $[0, 2a]$ et les X_i ($i \geq 2$) dans $[0, a]$.

14. (a) Soit $t \in]a, 2a]$. Comme dans la partie 1, $(V_n \leq t) = \bigcap_{i=1}^n (X_i \leq t)$ et, par indépendance mutuelle :

$$P(V_n \leq t) = P(X_1 \leq t) \prod_{i=2}^n P(X_i \leq t) = \frac{t}{2a} \cdot 1^{n-1} = \frac{t}{2a},$$

car $X_1 \sim \mathcal{U}([0, 2a])$ avec $0 \leq t \leq 2a$, et $P(X_i \leq t) = 1$ pour $i \geq 2$ puisque $t > a$. (Pour $n = 1$, le produit vide vaut 1.)

Conclusion

$$P(V_n \leq t) = \frac{t}{2a} \text{ pour tout } t \in]a, 2a].$$

(b) Même calcul selon la position de t ; on note G_n la fonction de répartition de V_n .

Si $t < 0$, tous les facteurs sont nuls : $G_n(t) = 0$. Si $0 \leq t \leq a$:

$G_n(t) = \frac{t}{2a} \left(\frac{t}{a}\right)^{n-1} = \frac{t^n}{2a^n}$. Si $a < t \leq 2a$: $G_n(t) = \frac{t}{2a}$ par (a). Si $t > 2a$: tous les facteurs valent 1, $G_n(t) = 1$.

$$G_n(t) = \begin{cases} 0 & \text{si } t < 0, \\ \frac{t^n}{2a^n} & \text{si } 0 \leq t \leq a, \\ \frac{t}{2a} & \text{si } a < t \leq 2a, \\ 1 & \text{si } t > 2a. \end{cases}$$

Limite simple. Pour $0 \leq t < a$, $0 \leq \frac{t}{a} < 1$ donc $\frac{t^n}{2a^n} \rightarrow 0$; pour $t = a$, $G_n(a) = \frac{1}{2}$ pour tout n . Ainsi, pour tout réel t , $G_n(t) \rightarrow G(t)$ avec :

$$G(t) = \begin{cases} 0 & \text{si } t < a, \\ \frac{t}{2a} & \text{si } a \leq t \leq 2a, \\ 1 & \text{si } t > 2a. \end{cases}$$

G est la fonction de répartition d'une variable aléatoire : c'est celle de $W = \max(X_1, a)$.

En effet $P(W \leq t) = 0$ si $t < a$ (car $W \geq a$), et pour $t \geq a$, $(W \leq t) = (X_1 \leq t)$, d'où

$P(W \leq t) = F_{X_1}(t)$, qui vaut $\frac{t}{2a}$ sur $[a, 2a]$ et 1 au-delà. La convergence $G_n(t) \rightarrow G(t)$ a

lieu en tout réel, donc en tout point de continuité de G .

Conclusion

(V_n) converge en loi, vers la loi de $\max(X_1, a)$ (qui vaut a avec probabilité $\frac{1}{2}$ et est uniforme sur $]a, 2a]$ avec probabilité $\frac{1}{2}$); la limite n'est pas la loi certaine égale à a .

(c) Comme $\frac{3}{2}a \in]a, 2a]$, la question (a) donne :

$$P\left(V_n > \frac{3}{2}a\right) = 1 - P\left(V_n \leq \frac{3}{2}a\right) = 1 - \frac{3a/2}{2a} = 1 - \frac{3}{4} = \frac{1}{4}.$$

Avec $\varepsilon = \frac{a}{2} > 0$: $\left(V_n > \frac{3}{2}a\right) \subset \left(|V_n - a| \geq \frac{a}{2}\right)$, donc pour tout n ,

$P\left(|V_n - a| \geq \frac{a}{2}\right) \geq \frac{1}{4}$, qui ne tend pas vers 0.

Conclusion

$P\left(V_n > \frac{3}{2}a\right) = \frac{1}{4}$ pour tout $n \in \mathbb{N}^*$, et (V_n) ne converge pas en probabilité vers a : l'estimateur V_n n'est plus convergent.

15. (a) Par définition, $X_2 + \dots + X_n = \frac{n-1}{2}M'_n$, donc :

$$M_n = \frac{2}{n} \left(X_1 + \frac{n-1}{2}M'_n \right) = \frac{2X_1}{n} + \frac{n-1}{n}M'_n.$$

Conclusion

$$M_n = \frac{2X_1}{n} + \frac{n-1}{n}M'_n.$$

(b) On écrit $M_n - a$ en faisant apparaître $M'_n - a$:

$$M_n - a = \frac{2X_1}{n} + \frac{n-1}{n}(M'_n - a) + \frac{n-1}{n}a - a = \frac{2X_1 - a}{n} + \frac{n-1}{n}(M'_n - a).$$

Comme $X_1 \in [0, 2a]$, on a $-a \leq 2X_1 - a \leq 3a$, donc $|2X_1 - a| \leq 3a$. Par l'inégalité triangulaire et $0 < \frac{n-1}{n} \leq 1$:

$$|M_n - a| \leq \frac{|2X_1 - a|}{n} + \frac{n-1}{n} |M'_n - a| \leq \frac{3a}{n} + |M'_n - a|.$$

Conclusion

$$\text{pour tout } n \geq 2, \quad |M_n - a| \leq \frac{3a}{n} + |M'_n - a|.$$

(c) Soit $n \geq n_0$. Alors $\frac{3a}{n} \leq \frac{3a}{n_0} < \varepsilon$. Si l'événement $(|M'_n - a| < \varepsilon)$ est réalisé, la question (b) donne $|M_n - a| < \varepsilon + \varepsilon = 2\varepsilon$.

Conclusion

$$\text{pour tout } n \geq n_0, \quad (|M'_n - a| < \varepsilon) \subset (|M_n - a| < 2\varepsilon), \text{ et donc } P(|M_n - a| < 2\varepsilon) \geq P(|M'_n - a| < \varepsilon).$$

(d) Montrons que, pour tout $\eta > 0$, $P(|M_n - a| \geq \eta) \rightarrow 0$.

Méthode 1 — Avec la question (c).

Soit $\eta > 0$ et $\varepsilon = \frac{\eta}{2}$. Choisissons un entier $n_0 \geq 2$ tel que $\frac{3a}{n_0} < \varepsilon$ (possible car $\frac{3a}{n} \rightarrow 0$). Par passage aux complémentaires dans (c), pour $n \geq n_0$:

$$0 \leq P(|M_n - a| \geq \eta) = P(|M_n - a| \geq 2\varepsilon) \leq P(|M'_n - a| \geq \varepsilon).$$

Le majorant tend vers 0 car (M'_n) converge en probabilité vers a . Par encadrement, $P(|M_n - a| \geq \eta) \rightarrow 0$.

Méthode 2 — Somme de deux suites convergeant en probabilité.

D'après (a), $M_n = M'_n + R_n$ avec $R_n = \frac{2X_1}{n} - \frac{M'_n}{n}$. Comme $X_1 \in [0, 2a]$ et $M'_n \in [0, 2a]$, on a $|R_n| \leq \frac{4a}{n} + \frac{2a}{n} = \frac{6a}{n}$. Pour $\eta > 0$ et $n > \frac{6a}{\eta}$, l'événement $(|R_n| \geq \eta)$ est impossible, donc $P(|R_n| \geq \eta) = 0$: (R_n) converge en probabilité vers 0. La suite (M'_n) converge en probabilité vers a ; la somme de deux suites convergeant en probabilité converge en probabilité vers la somme des limites, donc (M_n) converge en probabilité vers $a + 0 = a$.

La méthode 1 suit l'enchaînement proposé par l'énoncé ; la méthode 2 met en évidence que l'observation erronée n'intervient dans M_n qu'avec un poids $\frac{2}{n}$, qui tend vers 0.

Conclusion

$(M_n)_{n \geq 2}$ converge en probabilité vers a : M_n reste un estimateur convergent malgré la perturbation.

16. Synthèse.

Dans le modèle non perturbé, V_n est biaisé mais converge très vite vers a : son écart quadratique moyen est de l'ordre de $\frac{1}{n^2}$ et son intervalle de confiance a une longueur de l'ordre de $\frac{1}{n}$, contre $\frac{1}{n}$ et $\frac{1}{\sqrt{n}}$ respectivement pour M_n , qui est sans biais (questions 7, 10, 12 et 13).

En revanche, une seule observation erronée suffit à faire perdre à V_n sa convergence : avec probabilité $\frac{1}{2}$, $X_1 > a$ et alors $V_n = X_1$ pour tout n (pour $n \geq 1$, $V_n \geq X_1 > a \geq X_i$ pour $i \geq 2$), si bien que l'erreur ne se corrige jamais. Le maximum dépend entièrement des valeurs extrêmes : il n'est pas robuste. L'estimateur M_n , qui moyenne les observations, dilue l'erreur avec un poids $\frac{2}{n}$ et reste convergent : il est robuste.

Conclusion

V_n est le plus précis lorsque le modèle est exactement vérifié, M_n est plus fiable lorsque des données aberrantes peuvent apparaître ; le choix de l'estimateur dépend de la confiance que l'on accorde aux mesures.

Relevé des adaptations

- Exercice 2, question 6.(c) : ajout dans l'énoncé du résultat admis « si g est continue sur $\mathbb{R}^2$ et c réel, $\{g \leq c\}$ est fermé ». Justification : les propriétés topologiques utilisées doivent être précisées dans le sujet adapté.
- Exercice 2, questions 3 à 5 : énoncés conservés ; la nature des points critiques est établie uniquement par les signes des valeurs propres de la hessienne (et, en complément, par des restrictions ou par la forme quadratique), sans critère historique.
- Exercice 2, questions 7 et 9 : énoncés conservés ; l'étude sous contrainte non linéaire est menée par paramétrisation élémentaire du cercle, sans théorème général des multiplicateurs de Lagrange.
- Exercice 2, question 7 : coquille « contraine » rectifiée en « contrainte » (sans incidence mathématique).
- Problème, question 1.(a) : Code adapté en Python. Le rappel de la commande Scilab de génération de lois uniformes est remplacé par celui de `np.random.uniform(a, b, (n, m))` et de `np.max`, et l'en-tête de fonction Scilab par `def sim_V(n, a):`.
- Problème, question 8 : Code adapté en Python. L'en-tête de fonction Scilab est remplacé par `def sim_M(n, a):` et la commande `np.mean` est rappelée.
- Problème, questions 7.(b), 10 et 13 : la notion de risque quadratique, hors programme, est remplacée par le calcul explicite de $E((V_n - a)^2)$ et $E((M_n - a)^2)$, avec le même objectif (comparer la précision des deux estimateurs et retrouver leur convergence).

- Problème, questions 6 et 12 : énoncés conservés ; les intervalles de confiance sont construits à partir de quantités dont la loi limite ne dépend pas de a , avec résolution exacte en a , sans remplacer a par un estimateur dans une convergence en loi.
- Figures 1, 2 et 3 : la figure est retranscrite en SVG dans l'énoncé à partir de l'original, avec les valeurs mesurées (positions et formes des flèches, trajectoires et graduations extraites directement des tracés vectoriels du PDF ; cercle tracé exactement).