Gestion des Entreprises et Administrations
S4 — Approfondissement Mathématique

Résumé de cours en mutation — yann.grisel@free.fr

Table des matières

1 Algèbre linéaire (4.5h)

Introduction aux matrices

1.1 L'anneau des matrices réelles

- Une matrice à m lignes et n colonnes est un tableau rectangulaire de $m \times n$ éléments d'un certain ensemble (nombres réels, nombres complexes, ...).
- Les éléments d'une matrice $\mathbb{A}$ sont généralement notés $a_{i,j}$, i étant le numéro de ligne et j étant le numéro de colonne.
- L'ensemble des matrices réelles $m \times n$ est noté $\mathcal{M}_{m \times n}(\mathbb{R})$.
- Le couple (m, n) est appelé taille, ou dimension, de la matrice.
- Si $m = n$, on dira que la matrice est carrée.

Exemple $\begin{pmatrix} 2 & 4 & 3 \\ 3 & 5 & 1 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{R})$ est une matrice 2 par 3.

1.1.1 Multiplication par un réel

Avec $\mathbb{A} \in \mathcal{M}_{m,n}(\mathbb{R})$ $\mathbb{A} = (a_{i,j})$ et $\lambda \in \mathbb{R}$, alors

$$\lambda\mathbb{A} = (\lambda a_{i,j}).$$

Exemple $3 \times \begin{pmatrix} 2 & 4 & 3 \\ 3 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 12 & 9 \\ 9 & 15 & 3 \end{pmatrix}.$

1.1.2 Addition de matrices

Avec $\mathbb{A} \in \mathcal{M}_{m,n}$ $\mathbb{A} = (a_{i,j})$ et $\mathbb{B} \in \mathcal{M}_{m,n}$ $\mathbb{B} = (b_{i,j})$ alors

$$\mathbb{A} + \mathbb{B} = (a_{i,j} + b_{i,j}).$$

Propriétés Avec $\mathbb{A}, \mathbb{B}, \mathbb{C}$ trois matrice de $\mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda \in \mathbb{R}$, on a :

- Commutativité : $\mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}$,
- Associativité : $(\mathbb{A} + \mathbb{B}) + \mathbb{C} = \mathbb{A} + (\mathbb{B} + \mathbb{C}) = \mathbb{A} + \mathbb{B} + \mathbb{C}$,
- Distributivité : $\lambda(\mathbb{A} + \mathbb{B}) = \lambda\mathbb{A} + \lambda\mathbb{B}$
- Elément neutre : matrice avec des 0 partout (notée $\mathbb{0}_{m,n}$),

Exemple $\begin{pmatrix} 3 & 2 & 5 \\ 5 & 6 & 4 \end{pmatrix} + \begin{pmatrix} 4 & 2 & 4 \\ 4 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 7 & 4 & 9 \\ 9 & 9 & 7 \end{pmatrix}$

1.1.3 Multiplication de matrices

Avec $\mathbb{A} \in \mathcal{M}_{m,p}$ et $\mathbb{B} \in \mathcal{M}_{p,n}$ alors

$$\mathbb{A}\mathbb{B} = \mathbb{C} \in \mathcal{M}_{m,n} \text{ avec } c_{i,j} = \sum_{k=1}^{p} a_{i,k} b_{k,j}$$

Exemple $\begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 9 & 11 \\ 7 & 5 & 11 \end{pmatrix}$.

Attention
- Le nombre de colonnes de $\mathbb{A}$ doit être égal au nombre de lignes de $\mathbb{B}$.
- En général, $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$.

Propriétés Soient $\mathbb{A}, \mathbb{B}, \mathbb{C}$ trois matrices de bonnes dimensions, on a :
- Associativité : $(\mathbb{A}\mathbb{B})\mathbb{C} = \mathbb{A}(\mathbb{B}\mathbb{C}) = \mathbb{A}\mathbb{B}\mathbb{C}$,
- Distributivité par rapport à l'addition : $\mathbb{A}(\mathbb{B} + \mathbb{C}) = \mathbb{A}\mathbb{B} + \mathbb{A}\mathbb{C}$.
- Non commutativité : en général, même si les multiplications sont valides, on a malgré tout $\mathbb{A} \times \mathbb{B} \neq \mathbb{B} \times \mathbb{A}$.
- Elément neutre : matrice nulle avec des 1 sur la diagonale (matrice unité, notée $\mathbb{I}_n$).

1.1.4 Transposition de matrice

Soit $\mathbb{A} \in \mathcal{M}_{m,n}$, alors on définit

$$\mathbb{A}^T \in \mathcal{M}_{n,m} \text{ par } (\mathbb{A}^T)_{i,j} = a_{j,i}$$

Exemple $\begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 2 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 3 & 2 \end{pmatrix}$

Propriétés Soient $\mathbb{A}$ et $\mathbb{B}$ des matrices aux dimensions compatibles, alors

$$(\mathbb{A}\mathbb{B})^T = \mathbb{B}^T\mathbb{A}^T.$$

1.1.5 Déterminant d'une matrice carrée

C'est une fonction des matrices *carrées* $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$, qui a une matrice *carrée* $\mathbb{A}$ associe un réel noté $det(\mathbb{A})$ ou $|\mathbb{A}|$ (si $\mathbb{A}$ n'est pas carrée, une convention est de lui affecter un déterminant nul).

Le calcul général du déterminant n'est pas toujours simple et on se limitera aux dimensions 2 et 3.

Déterminant d'une matrice 2×2

$$\left|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = ad - bc$$

Déterminant d'une matrice 3×3

$$\left|\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\right| = (a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3) - (a_3b_2c_1 + b_3c_2a_1 + c_3a_2b_1).$$

Pour s'en souvenir, on peut écrire :

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

et le déterminant est la différence entre la somme des diagonales vers le bas et des diagonales vers le haut.

Propriétés

— $det(\mathbb{A}^T) = det(\mathbb{A})$
— $det(\mathbb{A}\mathbb{B}) = det(\mathbb{A})\, det(\mathbb{B})$

1.1.6 Inversion d'une matrice – Le pivot de Gauss

— Il n'existe pas à proprement parler de division pour les matrices
— On parle plutôt d'inverse pour la multiplication

Définition Une matrice carrée $\mathbb{A}$ de taille (n, n) est dite inversible s'il existe une matrice $\mathbb{B}$ de taille (n, n) telle que

$$\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}_n.$$

Propriétés

— L'inverse de $\mathbb{A}$ existe si et seulement si $det(\mathbb{A}) \neq 0$
— S'il existe, l'inverse est unique et noté $\mathbb{A}^{-1}$
— L'inverse d'une matrice inversible l'est aussi : $\left(\mathbb{A}^{-1}\right)^{-1} = \mathbb{A}$
— Le produit de matrices inversibles l'est aussi : $\left(\mathbb{A}\mathbb{B}\right)^{-1} = \mathbb{B}^{-1}\mathbb{A}^{-1}$
— Calculer l'inverse d'une matrice inversible est difficile, mais vérifier une proposition est facile : il suffit de vérifier si $\mathbb{A}\mathbb{A}^{-1} = \mathbb{A}^{-1}\mathbb{A} = \mathbb{I}_n$
— Une matrice est inversible si et seulement si son déterminant est non nul.

Pour calculer l'inverse de $\mathbb{A}$ avec la méthode du pivot de Gauss, on commence par écrire la matrice $\mathbb{A}$, et à sa droite on écrit la matrice identité : on se retrouve avec une matrice ayant n lignes et $2n$ colonnes.

On va ensuite répéter les 3 étapes suivantes une fois par ligne, en notant i le numéro de la ligne en cours.

1. Vérifier que le terme numéro i de la ligne $\mathbb{l}_i$ est non nul : S'il est égal à zéro, échanger la ligne en cours $\mathbb{l}_i$ avec une ligne *en dessous* telle que son terme numéro i ne soit pas nul.
   Si on doit faire cet échange et qu'on ne peut pas, la matrice n'est pas inversible et on s'arrête ici.
   Sinon, on termine donc cette sous-étape avec une ligne $\mathbb{l}_i$ dont le terme numéro i est non nul.
2. Diviser la ligne $\mathbb{l}_i$ par le terme numéro i.
   On termine donc cette sous-étape avec une ligne $\mathbb{l}_i$ dont le terme numéro i est 1.
3. Modifier toutes les *autres* lignes en leur soustrayant $\mathbb{l}_i$ multipliée par le terme numéro i *de la ligne qu'on modifie*.
   On termine donc cette sous-étape avec toutes les *autres* lignes que $\mathbb{l}_i$ qui auront 0 comme terme numéro i.

A la fin, la partie gauche de cette grosse matrice de travail sera devenue la matrice identité, et la partie droite sera devenue $\mathbb{A}^{-1}$.

Exemple $\begin{pmatrix} 2 & 0 & 3 \\ 3 & -1 & 0 \\ 2 & -4 & 3 \end{pmatrix}^{-1} = \dfrac{1}{36}\begin{pmatrix} 3 & 12 & -3 \\ 9 & 0 & -9 \\ 10 & -8 & 2 \end{pmatrix}$

1.2 Résolution de systèmes

Si $det(\mathbb{A}) \neq 0$, le système linéaire $\mathbb{A}\mathbb{x} = \mathbb{b}$ possède une solution unique.

1.2.1 Ecriture matricielle

Séparation des données et des inconnues :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_1 = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_2 = b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 = b_3 \end{cases} \iff \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Exemple

$$\begin{cases} 3z + 2x = 3 \\ 3x - y = 1 \\ 2x + 3z - 4y = 2 \end{cases} \iff \begin{pmatrix} 2 & 0 & 3 \\ 3 & -1 & 0 \\ 2 & -4 & 3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

1.2.2 Résolution par l'inversion

Si $\mathbb{A}$ est inversible et qu'on connaît son inverse, alors $\mathbb{x} = \mathbb{A}^{-1}\mathbb{b}$.

Exemple

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{-36} \begin{pmatrix} -3 & -12 & 3 \\ -9 & 0 & 9 \\ -10 & 8 & -2 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

1.2.3 Résolution par les déterminants

$x_i = \dfrac{det(\mathbb{A}_i)}{det(\mathbb{A})}$, où $\mathbb{A}_i$ est une matrice obtenue à partir de $\mathbb{A}$ en remplaçant la colonne $\mathbb{c}_i$ par la colonne $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$.

Exemple $det(\mathbb{A}) = -36$. Ensuite :

$A_x = \begin{pmatrix} 3 & 0 & 3 \\ 1 & -1 & 0 \\ 2 & -4 & 3 \end{pmatrix}$ et $det(\mathbb{A}_x) = -15$. Ainsi : $x = \frac{-15}{-36}$

$\mathbb{A}_y = \begin{pmatrix} 2 & 3 & 3 \\ 3 & 1 & 0 \\ 2 & 2 & 3 \end{pmatrix}$ et $det(\mathbb{A}_y) = -9$. Ainsi : $y = \frac{-9}{-36}$

$\mathbb{A}_z = \begin{pmatrix} 2 & 0 & 3 \\ 3 & -1 & 1 \\ 2 & -4 & 2 \end{pmatrix}$ et $det(\mathbb{A}_z) = -26$. Ainsi : $z = \frac{-26}{-36}$

1.2.4 Résolution par la méthode de Gauss

Pour calculer la solution de $\mathbb{A}\mathbb{x} = \mathbb{b}$, on peut aussi écrire le vecteur (colonne) $\mathbb{b}$ à droite de $\mathbb{A}$ puis appliquer les opérations décrites en 1.1.6 à la matrice $\mathbb{A}$ en reportant exactement les mêmes transformations au vecteur (colonne) de droite.

Si l'algorithme se termine, la matrice de gauche sera devenue la matrice identité et le vecteur (colonne) de droite sera devenu $(\mathbb{A}^{-1}\mathbb{b})$, c'est à dire $\mathbb{x}$. Sinon, c'est que le système n'a pas une solution unique et sort du cadre de ce résumé.

1.3 Vecteurs propres et Valeurs propres d'une matrice carrée

On peut décomposer de nombreuses manières une matrice en morceaux plus faciles à comprendre.

1.3.1 Polynôme caractéristique

Le polynôme caractéristique d'une matrice $\mathbb{A}$ (carrée) est obtenu en soustrayant x à chaque terme diagonal et en calculant le déterminant de cette nouvelle matrice :

$$P_{\mathbb{A}}(x) := det(\mathbb{A} - x\mathbb{I}_n)$$

Propriétés
- Le degré du polynôme caractéristique est égal à n (le nombre de lignes, ou de colonnes, puisque c'est le même, de $\mathbb{A}$).
- Le polynôme caractéristique est toujours de la forme $P_{\mathbb{A}}(x) = (-1)^n x^n - \text{trace}(\mathbb{A}) \cdot x^{n-1} + ... + det(\mathbb{A})$, la trace de $\mathbb{A}$ étant la somme de sa diagonale.
- Cas particulier avec $n = 2$: $P_{\mathbb{A}}(x) = x^2 - \text{trace}(\mathbb{A}) \cdot x + det(\mathbb{A})$.

Exemple $\mathbb{A} = \begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix}$: $P_{\mathbb{A}}(x) = det\left\{\begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}\right\} = det\begin{pmatrix} 1-x & 4 \\ 3 & 1-x \end{pmatrix} = x^2 - 2x - 11$

1.3.2 Valeurs propres

Les valeurs propres d'une matrice sont les racines du polynôme caractéristique associé :

$$x \text{ v.p. } \Leftrightarrow P_{\mathbb{A}}(x) = 0$$

Exemple $P_{\mathbb{A}}(x) = x^2 - 2x - 11$ s'annule pour $x_1 = 1 - 2\sqrt{3}$ et $x_2 = 1 + 2\sqrt{3}$, qui sont donc les valeurs propres de $\mathbb{A}$.

Remarques
- Trouver les valeurs propres est généralement difficile – on ne le fera que pour des matrices 2×2 – mais vérifier si une proposition convient est toujours facile.
- On peut voir une matrice comme "un chiffre" en plusieurs dimensions : quand on multiplie par une matrice, on multiplie donc dans plusieurs dimensions. Les valeurs propres sont ses coefficients multiplicateurs (mais elles ne disent pas quelles coordonnées elles multiplient).
- Le produit des valeurs propres est égal au déterminant : ceci permet de vérifier qu'on a les bonnes valeurs, et permet aussi d'expliquer vaguement le déterminant comme étant "le" coefficient multiplicateur de la matrice (dont le détail est donné par les valeurs propres et les vecteurs propres associés).

1.3.3 Vecteurs propres

Pour une valeur propre x_i (de $\mathbb{A}$), on dit qu'un vecteur *non nul* $\mathbb{v}_i$ (une colonne) est propre si

$$\mathbb{A}\mathbb{v}_i = x_i \mathbb{v}_i$$

Exemple

— Affirmation : $\mathbb{v}_1 = \begin{pmatrix} 1 \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$ est un vecteur (colonne) propre pour la valeur propre $x_1 = 1 - 2\sqrt{3}$.

— Vérification : on calcule $\mathbb{A} \cdot \mathbb{v}_1 = \begin{pmatrix} 1 - 2\sqrt{3} \\ 3 - \frac{\sqrt{3}}{2} \end{pmatrix}$ et on calcule $x_1 \cdot \mathbb{v}_1 = \begin{pmatrix} 1 - 2\sqrt{3} \\ 3 - \frac{\sqrt{3}}{2} \end{pmatrix}$. On constate que $\mathbb{A}\mathbb{v}_1 = x_1\mathbb{v}_1$; l'affirmation est vérifiée.

— Affirmation : $\mathbb{v}_2 = \begin{pmatrix} 1 \\ 3 - \frac{\sqrt{3}}{2} \end{pmatrix}$ est un vecteur (colonne) propre pour la valeur propre $x_2 = 1 + 2\sqrt{3}$.

— Vérification : On calcule $\mathbb{A} \cdot \mathbb{v}_2 = \begin{pmatrix} -11 + 2\sqrt{3} \\ 6 - \frac{\sqrt{3}}{2} \end{pmatrix}$ et $x_2 \cdot \mathbb{v}_2 = \begin{pmatrix} 1 + 2\sqrt{3} \\ 11\frac{\sqrt{3}}{2} \end{pmatrix}$. On constate que les deux résultats ne sont pas égaux : $\mathbb{v}_2$ n'est pas un vecteur (colonne) propre pour la valeur propre x_2.

Remarques

— Un vecteur propre est associé à une valeur propre (il ne sera pas propre pour une autre)

— Le calcul des vecteurs propres associés aux valeurs propres ne sera pas abordé ici, mais vérifier si une proposition convient est toujours facile.

— Par définition des valeurs propres, la matrice $(\mathbb{A} - x_i\mathbb{I}_n)$ n'est pas inversible : Les méthodes de la section 1.2 ne peuvent être appliquées pour calculer les vecteurs propres.

— Les vecteurs propres associés aux valeurs propres ne sont pas uniques (on peut par exemple multiplier toutes les coordonnées d'un vecteur (colonne) propre par un même chiffre : On obtiendra encore un vecteur (colonne) propre associé à la même valeur propre)

— Les vecteurs propres sont les directions dans lesquelles la matrice multiplie par la valeur propre associée.

1.3.4 Diagonalisation d'une matrice

— Une matrice peut se factoriser de nombreuses manières.

— Si on a n valeurs propres différentes, on peut diagonaliser la matrice ainsi :

$$\mathbb{A} = \mathbb{P}\mathbb{D}\mathbb{P}^{-1},$$

où $\mathbb{D}$ est la matrice ayant les valeurs propres sur la diagonale (et des 0 ailleurs) et $\mathbb{P}$ est la matrice obtenue en juxtaposant les vecteurs propres (dans le même ordre que pour $\mathbb{D}$)

Exemple $\begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \times \begin{pmatrix} 1 - 2\sqrt{3} & 0 \\ 0 & 1 + 2\sqrt{3} \end{pmatrix} \times \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{3}} \\ \frac{1}{2} & \frac{1}{\sqrt{3}} \end{pmatrix}$

Remarque Si on n'a pas n valeurs propres distinctes il peut y avoir des complications, voire pas de diagonalisation du tout ; mais nous garderons ces subtilités pour une autre fois.

Algèbre linéaire – Exercices

$$\mathbb{M}_2 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \end{pmatrix} \qquad \mathbb{M}_3 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\mathbb{M}_4 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{pmatrix} \qquad \mathbb{M}_5 = \begin{pmatrix} 0 & 3 & -1 \\ 2 & -1 & 1 \\ 4 & 0 & 2 \end{pmatrix}$$

Exercice 1 Calcul

1. Effectuer toutes les additions qu'il est possible de faire avec ces matrices.
2. Effectuer toutes les multiplications qu'il est possible de faire avec ces matrices.
3. Calculer les déterminants de ces matrices.
4. Vérification : les matrices $\frac{1}{6}\begin{pmatrix} -5 & -2 & 10 \\ 7 & 4 & -14 \\ 3 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 & 2 \\ 7 & 5 & 0 \\ 2 & 1 & 1 \end{pmatrix}$ sont-elles inverses l'une de l'autre ?

Exercice 2 Inversion / Résolution de système

1. Inverser $\mathbb{A} = \begin{pmatrix} 6 & -4 \\ 0 & 2 \end{pmatrix}$ [A préparer pour la séance suivante]
2. Inverser $\mathbb{M}_4$
3. Résoudre les systèmes suivants par la méthode des déterminants, puis par la méthode de Gauss.

$$\begin{cases} 3y + x = 1 \\ 3x - y = 2 \end{cases}$$

$$\begin{cases} x + 3y + 2z = 1 \\ x + z + 4y = 1 \\ x + 2y + 2z = 1 \end{cases}$$

Exercice 3 Valeurs et vecteurs propres

1. Calculer le polynôme caractéristique de chacune des matrices $\mathbb{M}_2$, $\mathbb{M}_3$, $\mathbb{M}_4$ et $\mathbb{M}_5$ [A préparer pour la séance suivante]
2. Calculer les valeurs propres de $\mathbb{M}_4$ et $\mathbb{M}_5$, en utilisant pour $\mathbb{M}_5$ la factorisation donnée en solution de la question 3.1 : $P_5(x) = (x-1)(-x^2+4)$.
3. Montrer que $\mathbb{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \mathbb{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ et $\mathbb{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ sont des vecteurs propres respectivement associés aux valeurs propres $x_1 = 1, x_2 = 2$ et $x_3 = 3$ de $\mathbb{M}_4$.

4. Ecrire la diagonalisation de $\mathbb{M}_4$ ($\mathbb{M}_4 = PDP^{-1}$, D et P à écrire à partir des valeurs et vecteurs propres)
5. Les vecteurs $\mathbb{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, $\mathbb{v}_4 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbb{v}_5 = \begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix}$ sont-ils propres pour $\mathbb{M}_5$?

Exercice 4 Pour s'entraîner On rajoute les matrices suivantes :

$$\mathbb{M}_1 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \end{pmatrix}, \quad \mathbb{M}_6 = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix}, \quad \mathbb{M}_7 = \begin{pmatrix} 0 & 0 & 2 \\ 7 & 5 & 0 \\ 2 & 1 & 1 \end{pmatrix}$$

1. En réutilisant les résultats précédents, calculer la matrice $\mathbb{V} = (\mathbb{M}_1\mathbb{M}_3^T)^T\mathbb{M}_4^T$
2. Vérification : les matrices $\frac{1}{12}\begin{pmatrix} 2 & 4 \\ 0 & 6 \end{pmatrix}$ et $\begin{pmatrix} 6 & -4 \\ 0 & 2 \end{pmatrix}$ sont-elles inverses l'une de l'autre ?
3. Calculer les déterminants de $\mathbb{M}_1$, $\mathbb{M}_6$, $\mathbb{M}_7$
4. Inverser $\mathbb{M}_1$, $\mathbb{M}_6$, $\mathbb{M}_7$
5. Calculer le polynôme caractéristique de $\mathbb{M}_6$ et celui de $\mathbb{M}_7$
6. Vérification : $0.5, 1$ et 2 sont-ils valeurs propres de $\mathbb{M}_6$?
7. Vérification : Pour la matrice $\mathbb{M}_6$, $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est-il un vecteur propre pour la valeur propre 0.5 ? Et pour la valeur propre 1 ?

Réponses

Réponse 1 Calcul

1.

Additions	$\mathbb{M}_2$	$\mathbb{M}_3$	$\mathbb{M}_4$	$\mathbb{M}_5$
$\mathbb{M}_2$	$\mathbb{B}$	$\mathbb{C}$	–	–
$\mathbb{M}_3$		$\mathbb{D}$	–	–
$\mathbb{M}_4$			$\mathbb{E}$	$\mathbb{F}$
$\mathbb{M}_5$				$\mathbb{I}$

Les cases blanches résultent de la la commutativité de l'addition, $\mathbb{M}_i + \mathbb{M}_j = \mathbb{M}_j + \mathbb{M}_i$.

$$\mathbb{B} = \begin{pmatrix} 0 & 2 & 4 & 6 \\ 8 & 10 & 12 & 14 \end{pmatrix} \quad \mathbb{D} = \begin{pmatrix} 0 & 0 & 2 & 2 \\ 0 & 2 & 0 & 2 \end{pmatrix} \quad \mathbb{E} = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 6 & 0 \\ 4 & -8 & 4 \end{pmatrix}$$

$$\mathbb{I} = \begin{pmatrix} 0 & 6 & -2 \\ 4 & -2 & 2 \\ 8 & 0 & 4 \end{pmatrix} \quad \mathbb{C} = \begin{pmatrix} 0 & 1 & 3 & 4 \\ 4 & 6 & 6 & 8 \end{pmatrix} \quad \mathbb{F} = \begin{pmatrix} 1 & 5 & -1 \\ 2 & 2 & 1 \\ 6 & -4 & 4 \end{pmatrix}$$

2.

Multiplications	$\mathbb{M}_2$	$\mathbb{M}_3$	$\mathbb{M}_4$	$\mathbb{M}_5$
$\mathbb{M}_2$	–	–	–	–
$\mathbb{M}_3$	–	–	–	–
$\mathbb{M}_4$	–	–	$\mathbb{A}$	$\mathbb{E}$
$\mathbb{M}_5$	–	–	$\mathbb{B}$	$\mathbb{F}$

$$\mathbb{A} = \begin{pmatrix} 1 & 8 & 0 \\ 0 & 9 & 0 \\ 6 & -16 & 4 \end{pmatrix} \quad \mathbb{E} = \begin{pmatrix} 4 & 1 & 1 \\ 6 & -3 & 3 \\ 0 & 10 & -2 \end{pmatrix} \quad \mathbb{B} = \begin{pmatrix} -2 & 13 & -2 \\ 4 & -3 & 2 \\ 8 & 0 & 4 \end{pmatrix} \quad \mathbb{F} = \begin{pmatrix} 2 & -3 & 1 \\ 2 & 7 & -1 \\ 8 & 12 & 0 \end{pmatrix}$$

3. Seules les matrices $\mathbb{M}_4$ et $\mathbb{M}_5$ sont concernées.
   a) $|\mathbb{M}_4| = 6$
   b) $|\mathbb{M}_5| = -4$
4. Oui

Réponse 2 Inversion / Résolution de système

1. $\mathbb{A}^{-1} = \frac{1}{12}\begin{pmatrix} 2 & 4 \\ 0 & 6 \end{pmatrix}$
2. $\mathbb{M}_4^{-1} = \frac{1}{6}\begin{pmatrix} 6 & -4 & 0 \\ 0 & 2 & 0 \\ -6 & 8 & 3 \end{pmatrix}$
3. a) $det(\mathbb{A}) = -10$, $det(\mathbb{A}_x) = -7$, $det(\mathbb{A}_y) = -1$: $x = 7/10$ et $y = 1/10$.
   b) $det(\mathbb{A}) = -1$, $det(\mathbb{A}_x) = -1$, $det(\mathbb{A}_y) = 0$, $det(\mathbb{A}_z) = 0$: $x = 1$, $y = 0$ et $z = 0$.

Réponse 3 Valeurs et vecteurs propres

1. a) $P_4(x) = -x^3 + 6x^2 - 11x + 6$
   b) $P_5(x) = -x^3 + x^2 + 4x - 4 = (x-1)(-x^2+4)$
2. a) Valeurs propres pour $\mathbb{M}_4$: $x_1 = 1, x_2 = 2$ et $x_3 = 3$.
   b) Valeurs propres pour $\mathbb{M}_5$: $x_1 = 1, x_2 = 2$ et $x_3 = -2$.
3. On voit par le calcul que $\mathbb{M}_4\mathbb{v}_1 = \mathbb{v}_1$, $\mathbb{M}_4\mathbb{v}_2 = 2\mathbb{v}_2$ et $\mathbb{M}_4\mathbb{v}_3 = 3\mathbb{v}_3$.
4. On en déduit
$$\mathbb{M}_4 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ -2 & 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
5. — $\mathbb{v}_4$: oui pour la valeur propre 2.
   — $\mathbb{v}_5$: oui pour la valeur propre -2.
   — $\mathbb{v}_6$: non.

Réponse 4 Pour s'entraîner

1. $\mathbb{V} = (\mathbb{M}_1\mathbb{M}_3^T)^T\mathbb{M}_4^T = \mathbb{M}_3\mathbb{M}_1^T\mathbb{M}_4^T = \mathbb{M}_3(\mathbb{M}_4\mathbb{M}_1)^T = \mathbb{M}_3\mathbb{R}^T$. Ainsi,

$$\begin{aligned}
\mathbb{V} &= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 8 & 11 & 14 & 17 \\ 12 & 15 & 18 & 21 \\ 0 & 0 & 0 & 0 \end{pmatrix}^T \\
&= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 8 & 12 & 0 \\ 11 & 15 & 0 \\ 14 & 18 & 0 \\ 17 & 21 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 31 & 39 & 0 \\ 28 & 36 & 0 \end{pmatrix}
\end{aligned}$$

2. Oui
3. $\mathbb{M}_1$ n'est pas carrée, $|\mathbb{M}_6| = \frac{-1}{8}$ et $|\mathbb{M}_7| = -6$.
4. $\mathbb{M}_1$ n'est pas carrée, $\mathbb{M}_6^{-1} = \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.375 & 0.25 & 0.375 \\ -0.25 & 0.5 & 0.75 \end{pmatrix}$ et $\mathbb{M}_7^{-1} = \frac{1}{6}\begin{pmatrix} -5 & -2 & 10 \\ 7 & 4 & -14 \\ 3 & 0 & 0 \end{pmatrix}$
5. $P_6(x) = -x^3 + \frac{5x^2}{4} - \frac{1}{8}x - \frac{1}{8}$, $P_7(x) = -x^3 + 6x^2 - x - 6$
6. $P_6(0.5) = P_6(1) = 0$, mais $P_6(2) = \frac{-27}{8}$: 0.5 et 1 sont des valeurs propres de $\mathbb{M}_6$, mais pas 2.
7. Non, et oui.

2 Régression multi-linéaire (4.5h)

— Description : Un échantillon de m individus, où on mesure n critères différents par individu
— Estimation : Une V.A. expliquée sur la population : Y
— Plusieurs (n) variables explicatives : X_1, X_2, ..., X_n

Le modèle linéaire pour Y

$$Y = a_1 X_1 + a_2 X_2 + ... + a_n X_n + b$$

Exemple On souhaite expliquer le chiffre d'affaires (CA) par le budget publicitaire (PUBS), le nombre de promos en cours (PROMOS) et le taux de change de l'euro en dollars (DOLS). Le modèle linéaire est alors CA = a_1·PUBS + a_2·PROMOS + a_3·DOLS + b.

Remarque C'est un modèle : on propose une formule pour calculer Y qui tente de reproduire la réalité, mais la réalité ne se conformera jamais au modèle.

2.1 Estimation des paramètres du modèle

Les p paramètres $a_1, a_2 ... a_n, b$ (oui, $p = n + 1$) sont calculés tous d'un coup (sous forme d'une matrice $\mathbb{a}$) en se basant sur m mesures de la V.A. expliquée :

Variables	X_1	X_2	X_3	...	X_n	Y	
Données	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	...	$x_{1,n}$	y_1	Mesures
	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$		$x_{2,n}$	y_2	
	$\vdots$				$\vdots$	$\vdots$	
	$x_{m,1}$	$x_{m,2}$	$x_{m,3}$	...	$x_{m,n}$	y_m	

TABLE 2.1 – ensemble des informations disponibles

On commence par nommer les blocs suivants :

— On appelle ensemble des mesures la matrice $\mathbb{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}_{m \times 1}$: ce sont les valeurs de la variable qu'on cherche à expliquer

— On appelle ensemble des paramètres la matrice $\mathbb{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ b \end{pmatrix}_{p\times 1}$: ce sont les coefficients du modèle linéaire qu'on propose pour expliquer Y

— On appelle ensemble des données la matrice $\mathbb{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,n} & 1 \\ x_{2,1} & x_{2,2} & x_{2,3} & & x_{2,n} & 1 \\ \vdots & & & & \vdots & \vdots \\ x_{m,1} & x_{m,2} & x_{m,3} & \dots & x_{m,n} & 1 \end{pmatrix}_{m\times p}$: ce sont les valeurs des variables explicatives pour le modèle proposé

Remarque Le dernier paramètre du modèle (b) est une valeur constante, sous-entendu toujours multipliée par 1. C'est pourquoi $\mathbb{X}$ est systématiquement complétée par une colonne de "1"s.

Avec ces matrices, le résultat du modèle – les prévisions – est la matrice $(\mathbb{X} \cdot \mathbb{a})$. On voudrait que toutes les prévisions soient égales aux mesures réelles correspondantes ($\mathbb{X}\mathbb{a} = \mathbb{y}$) mais c'est impossible (et en fait, ce n'est même pas forcément une bonne idée).

En admettant que les données $x_{i,j}$ soient connues avec exactitude, on va alors chercher les paramètres ($\mathbb{a}$) qui minimisent l'erreur entre les prévisions du modèle ($\mathbb{X}\mathbb{a}$) et les mesures réelles ($\mathbb{y}$) : $\sum_i \left((\mathbb{X}\mathbb{a})_i - \mathbb{y}_i\right)^2$ (le "carré" dans le calcul de l'erreur est essentiellement là pour simplifier les formules). La différence entre prévision et réalité est la plus faible quand $(\mathbb{X}^T\mathbb{X})\mathbb{a} - (\mathbb{X}^T\mathbb{y}) = 0_{m\times 1}$. C'est à dire

$$(\mathbb{X}^T\mathbb{X})\mathbb{a} = (\mathbb{X}^T\mathbb{y}).$$

Si la matrice $(\mathbb{X}^T\mathbb{X})$ est inversible, on calcule $\mathbb{a}$ par une des trois manières présentées en Section 1.2 ; sinon, c'est hors programme.

Exemple On mesure l'influence des PUBS (budget en k€), PROMOS (idem) et DOLS (valeurs de 10€ en $) sur le CA (en k€) pendant 6 mois :

PUBS	PROMOS	DOLS	CA
19	13	10	649
18	14	15	689
10	10	13	492
18	18	17	780
17	14	13	567

La matrice des données est alors $\mathbb{X} = \begin{pmatrix} 19 & 13 & 10 & 1 \\ 18 & 14 & 15 & 1 \\ 10 & 10 & 13 & 1 \\ 18 & 18 & 17 & 1 \\ 17 & 14 & 13 & 1 \end{pmatrix}$ et la matrice de mesures est $\mathbb{y} = \begin{pmatrix} 649 \\ 689 \\ 492 \\ 780 \\ 567 \end{pmatrix}$. On calcule alors les produits $(\mathbb{X}^T\mathbb{X}) = \begin{pmatrix} 1398 & 1161 & 1117 & 82 \\ 1161 & 985 & 958 & 69 \\ 1117 & 958 & 952 & 68 \\ 82 & 69 & 68 & 5 \end{pmatrix}$ et $(\mathbb{X}^T\mathbb{y}) = \begin{pmatrix} 53332 \\ 44981 \\ 43852 \\ 3177 \end{pmatrix}$ pour résoudre le système $(\mathbb{X}^T\mathbb{X})\mathbb{a} = (\mathbb{X}^T\mathbb{y})$; ce qui donne $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b \end{pmatrix} = \begin{pmatrix} 16 \\ 12 \\ 14 \\ 17 \end{pmatrix}$. Au vu de ces mesures, on propose donc le modèle suivant : CA = 16·PUBS +12·PROMOS +14·DOLS +17.

Remarques

— Les paramètres sont obtenus en minimisant la différence entre prédiction et mesures : ce sont les "mieux possibles" mais en général $(\mathbb{X}\mathbb{a})_i$ sera assez différent de $\mathbb{y}_i$.

— La matrice $(\mathbb{X}^T\mathbb{X})$ est toujours symétrique : inutile de calculer la partie basse (idem pour l'inverse).

— Si les V.A. explicatives sont sur des échelles de grandeur trop différentes, l'influence des V.A. à grandes valeurs va masquer l'influence des V.A. à petites valeurs. Il peut alors être nécessaire de toutes les mettre sur un pied d'égalité en centrant et réduisant respectivement chaque colonne des données : à chaque colonne, on lui soustrait sa moyenne et la on divise par l'écart-type estimé de la V.A. correspondante.

— Si on utilise des données centrées réduites, $(\mathbb{X}^T\mathbb{X})$ est encore symétrique et de plus tous les termes diagonaux valent $m-1$ (si on a réduit en utilisant l'écart-type estimé – sinon ils valent m)

— Si les réalisations de X ne sont pas connues avec exactitude, il existe des formules différentes légèrement plus pertinentes.

2.2 Validation statistique du modèle

On peut toujours proposer un modèle linéaire et calculer $a_1, a_2 \dots a_n$ et b, mais est-ce pertinent ?

— Décomposition de la quantité de variation :

$$\underbrace{\sum_i (\mathbb{y}_i - \overline{\mathbb{y}})^2}_{Q_T(\text{variation totale})} = \underbrace{\sum_i (\mathbb{y}_i - (\mathbb{X}\mathbb{a})_i)^2}_{Q_E(\text{quantité d'erreur})} + \underbrace{\sum_i ((\mathbb{X}\mathbb{a})_i - \overline{\mathbb{y}})^2}_{Q_M(\text{variation du modèle})}$$

— Décomposition des degrés de liberté : $\underbrace{m-1}_{Totale} = \underbrace{m-p}_{Erreur} + \underbrace{p-1}_{Modele}$

Estimations des variances

— Estimation de la variance (totale) des mesures : $V_T = \dfrac{Q_T}{m-1}$ (à calculer en premier pour en déduire $Q_T = (m-1)V_T = (m-1)\text{var}(\mathbb{y})$)

— Estimation de la variance de l'erreur : $V_E = \dfrac{Q_E}{m-p}$ (à calculer en second, effort incompressible : on calcule les différences $\mathbb{D} = \mathbb{X} \cdot \mathbb{a} - \mathbb{y}$, puis la quantité d'erreur $Q_E = \mathbb{D}^T \cdot \mathbb{D}$)

— Estimation de la variance du modèle : $V_M = \dfrac{Q_M}{p-1}$ (à calculer en dernier : $Q_M = Q_T - Q_E$)

Coefficient de détermination

— $R^2 = \dfrac{Q_M}{Q_T}$

— $0 \leqslant R^2 \leqslant 1$

— La valeur R^2 est la quantité de variation du modèle, relativement à la variation totale. On l'interprète comme un pourcentage d'explication de la V.A. Y par le modèle $a_1X_1 + a_2X_2 + ... + a_nX_n + b$
— La valeur R^2 augmente mécaniquement avec le nombre de variables et ne tient pas compte du nombre de mesures : il faudra encore l'ajuster pour faire des conclusions précises

Le test

— $\mathcal{H}_0$: La régression linéaire multiple n'est pas significative
— $\mathcal{H}_1$: La régression linéaire multiple est significative
— Statistique de test : Sous certaines hypothèses dont on ne se préoccupera pas ici, on a

$$F = \frac{V_M}{V_E} \sim Fisher(p-1, m-p)$$

— Si la réalisation de F se trouve en dehors de l'intervalle de confiance *unilatéral* pour la loi de Fisher, on en conclut, au risque α, que la variance de l'erreur est (suffisamment) négligeable par rapport à la variance du modèle dans la variance totale.

Exemple

1. On commence par calculer la variance estimée des mesures $\text{var}_{est} = 12302.3$ et on en déduit la quantité de variation totale $Q_T = 49209.2$.

   On calcule ensuite les prévisions du modèle pour les données utilisées : $\mathbb{X}\cdot\mathbb{a} = \begin{pmatrix} 617 \\ 683 \\ 479 \\ 759 \\ 639 \end{pmatrix}$. L'adéquation entre modèle ($\mathbb{X}\mathbb{a}$) et mesures ($\mathbb{y}$) peut sembler assez correcte. En effet, le vecteur (colonne) des différences vaut alors $\mathbb{D} = (\mathbb{X}\cdot\mathbb{a} - \mathbb{y}) = \begin{pmatrix} -32 \\ -6 \\ -13 \\ -21 \\ 72 \end{pmatrix}$. On en déduit ainsi la quantité d'erreur $Q_E = 6854$.

   Enfin, on trouve la quantité de variation du modèle par soustraction : $Q_M = Q_T - Q_E = 42355.2$. On s'aperçoit que la variation du modèle donne presque toute la variation des mesures : $R^2 = 86\%$.
2. Avec ceci, on calcule la réalisation de la variable de test $F = \frac{14118.4}{6854} = 2.06$.
3. Sur le tableau de la loi de Fisher(3,1) on lit $F_{5\%} = 215.7$: on peut dire au risque de 5% que le modèle linéaire est (largement) pertinent.

Régression multi-linéaire – Exercices

Quatre caractères – X_1 (PIB par habitant en milliers de dollars), X_2 (taux de croissance de la population en 1/1000), Y_1 (taux de mortalité infantile en 1/100 000), Y_2 (taux d'analphabétisme en 1/100 000) – ont été mesurés dans 6 pays choisis au hasard (voir Table 2.2).

Pays	1	2	3	4	5	6	Moyenne	Variance estimée
PIB (X_1)	2	4	2	4	3	2	2.83	0.97
Croissance (X_2)	2	2	2	5	7	7	4.17	6.17
Mortalité (Y_1)	1646	3031	4177	5970	7039	5952	4635.83	4202525.37
Analphabétisme (Y_2)	3452	5547	9953	13211	14911	10407	9580.17	19248850.57

Table 2.2 – Mesures effectuées dans 6 pays choisis au hasard

Les deux variables X_1 et X_2 sont les variables explicatives, et les variables Y_1 et Y_2 sont les variables à expliquer. Enfin, on considère que la pondération des pays et des variables est uniforme (on ne tient pas compte de la population).

Avec la matrice des données

$$\mathbb{X} = \begin{pmatrix} 2 & 2 & 1 \\ 4 & 2 & 1 \\ 2 & 2 & 1 \\ 4 & 5 & 1 \\ 3 & 7 & 1 \\ 2 & 7 & 1 \end{pmatrix}$$

on donne les produits suivants :

$$\mathbb{X}^T\mathbb{X} = \begin{pmatrix} 53 & 71 & 17 \\ 71 & 135 & 25 \\ 17 & 25 & 6 \end{pmatrix}, \quad \mathbb{X}^T \mathrm{y}_1 = \begin{pmatrix} 80671 \\ 138495 \\ 27815 \end{pmatrix}, \quad \mathbb{X}^T \mathrm{y}_2 = \begin{pmatrix} 167389 \\ 281185 \\ 57481 \end{pmatrix}.$$

Exercice 1 Ecriture d'un modèle linéaire pour le taux de mortalité infantile en fonction du PIB et du taux de Croissance

1. Ecrivez le modèle linéaire pour la mortalité infantile (Y_1) en fonction (linéaire) du PIB (X_1) et de la Croissance (X_2) : proposez une fonction (linéaire) qui permet de calculer Y_1 à partir de X_1 et X_2, avec des coefficients pour l'instant inconnus.
2. Estimez les paramètres de ce modèle : les coefficients de l'équation proposée ci-dessus.
3. Pour la Bordurie, on a mesuré : PIB = 5 et Croissance = 3. Donnez alors une estimation du taux de mortalité en Bordurie à partir du modèle linéaire calculé.
4. Avec le modèle proposé, quelles sont les valeurs estimées pour les pays 4 et 3 ?

Exercice 2 Validation statistique du modèle précédent On donne les valeurs de $(\mathbb{X}\mathbb{a})$:

$$\begin{pmatrix} 2752 \\ 3472 \\ 2752 \\ 5665 \\ 6767 \\ 6407 \end{pmatrix}$$

1. Avec ces valeurs, calculez alors la variance de l'erreur, et déduisez-en la variance du modèle.
2. Avec un test de Fisher, peut-on dire que la variance de l'erreur est négligeable (par rapport à la variance totale) pour la régression (multi-)linéaire de la mortalité infantile ?

Exercice 3 Pour s'entraîner – Explication du taux d'analphabétisme en fonction des variables explicatives PIB et Croissance

1. Ecrivez le modèle linéaire pour l'analphabétisme (Y_2) en fonction (linéaire) du PIB (X_1) et de la Croissance (X_2).
2. Estimez les paramètres de ce modèle.
3. Pour la Bordurie, on a mesuré : PIB = 5 et Croissance = 3. Donnez alors une estimation du taux d'analphabétisme en Bordurie, à partir du modèle linéaire calculé.
4. Avec le modèle proposé, quelles sont les valeurs estimées pour les pays 4 et 3 ?
5. On donne les valeurs arrondies de $(\mathbb{X}\mathbb{a})$:
   $$\begin{pmatrix} 5920 \\ 7700 \\ 5920 \\ 11741 \\ 13545 \\ 12655 \end{pmatrix}$$
   Avec ces valeurs, calculez alors la variance de l'erreur, et déduisez-en la variance du modèle.
6. Avec un test de Fisher, peut-on dire que la variance de l'erreur est négligeable (par rapport à la variance totale) pour la régression (multi-)linéaire de l'Analphabétisme ?

Réponses

Réponse 1

1. $MOR. = a_1 PIB + a_2 CRO + b$
2. $\begin{pmatrix} a_1 \\ a_2 \\ b \end{pmatrix} = \begin{pmatrix} 360 \\ 731 \\ 570 \end{pmatrix}$
3. Pour la Bordurie, le modèle prévoit une Mortalité de 4563 pour 100 000.
4. Pour le pays 4, le modèle nous donne l'estimation MOR= 5665, ce qui est très proche des valeurs mesurées. Par contre, pour le pays 3, le modèle nous donne l'estimation MOR= 2752, ce qui est nettement moins bien. Il faut être conscient des limites du modèle.

Réponse 2

1. $Q_E = 3822376$. De la variance des $\mathbb{y}_1$ on déduit que $Q_T = 21012626.83$, puis que $Q_M = 17190250.83$. Ceci donne $V_E = 1274125.33$ et $V_M = 8595125.42$. On en déduit un R^2 de 0.82, ce qui semble plutôt bon.
2. $F = 6.75$ et $F_{5\%}(2, 3) = 9.55$. On peut donc conclure que la régression linéaire proposée pour la mortalité n'est pas significative au risque de 5% : on ne peut pas suffisamment faire confiance à ce modèle.

Réponse 3

1. $ANA = a_1PIB + a_2CRO + b$ (mais a_1, a_2 et b différents de l'exercice précédent bien sur)
2. $\begin{pmatrix} a_1 \\ a_2 \\ b \end{pmatrix} = \begin{pmatrix} 890 \\ 1347 \\ 1446 \end{pmatrix}$
3. Pour la Bordurie, le modèle prévoit un Analphabétisme de 9937.
4. Pour le pays 4, le modèle nous donne l'estimation ANA= 11741, et pour le pays 3, le modèle nous donne l'estimation ANA= 5920.
5. $Q_E = 36071882$. De la variance des $\mathbb{y}_2$ on déduit que $Q_T = 96244252.83$, puis que $Q_M = 60172370.83$. Ceci donne $V_E = 12023960.67$ et $V_M = 30086185.42$.
6. $F = 2.5$ et $F_{5\%} = 9.55$. On peut donc conclure que la régression linéaire proposée pour l'analphabétisme n'est pas significative au risque de 5% non plus : les données semblent trop erratiques pour pouvoir en déduire un modèle satisfaisant.

3 Optimisation (4.5h)

Calcul du maximum/minimum d'une fonction objectif de plusieurs variables, avec éventuellement des contraintes

3.1 Dérivées de fonctions composées

Rappels du lycée.

3.1.1 Dérivées usuelles

Pour toute fonction dérivable u et tout entier relatif n
- $(cos \circ u)' = -u'(sin \circ u)$
- $(sin \circ u)' = u'(cos \circ u)$
- $(e^u)' = u'e^u$
- $(ln\, u)' = \frac{u'}{u}$
- $(u^n)' = u'nu^{n-1}$
- $(\sqrt{u})' = \frac{u'}{2\sqrt{u}}$

3.1.2 Opérations sur les dérivées

Soient α un nombre réel et u, v et f des fonctions
- $(\alpha u)' = \alpha u'$
- $(u + v)' = u' + v'$
- $(uv)' = u'v + uv'$
- $\left(\frac{u}{v}\right)' = \frac{u'v-uv'}{v^2}$
- $(f \circ u)' = u'(f' \circ u)$

Exemple : $f = sin \circ ln$, c'est à dire $f(x) = sin\,(\,ln(x))$, donne $f'(x) = \frac{1}{x}\,cos\,(\,ln(x))$.

3.2 Dérivées partielles

Pour des fonctions de plusieurs variables, on peut dériver par rapport à toutes les variables d'un coup (nous ne le ferons pas) ou par rapport à une seule à la fois : on appelle ceci une dérivation partielle.

3.2.1 Définition

Soit $f : D \to \mathbb{R}$ définie par $f : (x_1, x_2 \dots x_n) \mapsto f(x_1, x_2 \dots x_n)$ une fonction. On appelle dérivée partielle par rapport à x_i la dérivée obtenue en considérant toutes les autres variables comme des constantes. On note cette fonction

$$\frac{\partial f}{\partial x_i} \text{ ou } \partial_{x_i} f.$$

Exemple : $f(x, y) = yx^2 - 9y$, $\partial_x f(x, y) = 2xy$ et $\partial_y f(x, y) = x^2 - 9$.

3.2.2 Ordre de dérivation (Théorème de Schwarz)

Si $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ et $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ existent et sont continues, alors elles sont égales et on peut noter le tout

$$\frac{\partial^2 f}{\partial x \partial y} \text{ ou } \partial^2_{xy} f.$$

3.3 Gradient et Hessienne

Le gradient est un vecteur qui joue le rôle de la dérivée première et la Hessienne est une matrice que joue le rôle de la dérivée seconde (et nous jetterons un voile pudique sur les subtiles différences).

3.3.1 Gradient

Dans le cadre qui nous intéresse, le gradient, au point $\mathbb{x} := (x_1, x_2 \dots x_n)$, d'une fonction (dérivable) de $\mathbb{R}^n$ dans $\mathbb{R}$ est noté $\nabla f(x_1, x_2 \dots x_n)$ et défini par

$$\nabla f(x_1, x_2 \dots x_n) = \begin{pmatrix} \partial_{x_1} f(\mathbb{x}) \\ \partial x_2 f(\mathbb{x}) \\ \vdots \\ \partial x_n f(\mathbb{x}) \end{pmatrix}.$$

Remarque Ce n'est que l'empilement des dérivées partielles de f avec des parenthèses autour pour en faire un vecteur (une colonne)

Exemple Avec $f(x, y) = x \cdot exp(-\frac{(x-12)^2}{300} - \frac{(y-18)^2}{900}) - 20 \cdot exp(-\frac{(x+10)^2}{70} - \frac{(y+15)^2}{60})$,

$$\nabla f(x, y) = \begin{pmatrix} -\frac{1}{150}(x^2 - 12x - 150) \cdot exp(-\frac{(x-12)^2}{300} - \frac{(y-18)^2}{900}) + \frac{4}{7}(x + 10) \cdot exp(-\frac{(x+10)^2}{70} - \frac{(y+15)^2}{60}) \\ \frac{1}{450}x(18 - y) \cdot exp(-\frac{(x-12)^2}{300} - \frac{(y-18)^2}{900}) + \frac{2}{3}(y + 15) \cdot exp(-\frac{(x+10)^2}{70} - \frac{(y+15)^2}{60}) \end{pmatrix}.$$

Par exemple, au point $\mathbb{x} = (-1, -20)$ on a donc $\nabla f(-1, -20) = \begin{pmatrix} 1.17 \\ -0.701 \end{pmatrix}$.

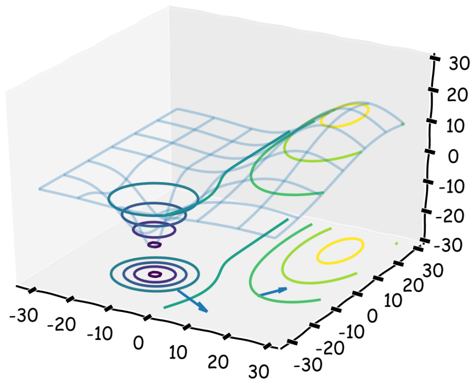


FIGURE 3.1 – Exemple de gradients (fortement rallongés) en (-1,-20) et (16,-14)

3.3.2 Hessienne

Dans le cadre qui nous intéresse, la Hessienne, au point $\mathbb{x} := (x_1, x_2 \dots x_n)$, d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ est notée $H_f(x_1, x_2 \dots x_n)$ et définie par

$$Hf(x_1, x_2 \dots x_n) = \begin{pmatrix} \partial_{x_1 x_1} f(\mathbb{x}) & \partial_{x_1 x_2} f(\mathbb{x}) & \dots & \partial_{x_1 x_n} f(\mathbb{x}) \\ \partial_{x_2 x_1} f(\mathbb{x}) & \partial_{x_2 x_2} f(\mathbb{x}) & \dots & \partial_{x_2 x_n} f(\mathbb{x}) \\ \vdots & & \ddots & \vdots \\ \partial_{x_n x_1} f(\mathbb{x}) & \partial_{x_n x_2} f(\mathbb{x}) & \dots & \partial_{x_n x_n} f(\mathbb{x}) \end{pmatrix}.$$

Remarque : Ce ne sont que les dérivées partielles de ∇f (qui est une colonne) alignées (*ie* en ligne). On obtient donc une matrice. Et cette matrice est symétrique, en vertu du théorème de Schwarz.

Exemple

$$H_f(x,y) = \begin{pmatrix} +\frac{1}{22500}(x^3 + 24x^2 + 306x + 3600) \cdot e^A - \frac{4}{245}(x^2 + 20x + 65) \cdot e^B, \\ \qquad +\frac{1}{67500}(x^2 - 12x - 150) \cdot (y-18) \cdot e^A - \frac{2}{105}(x+10) \cdot (y+15) \cdot e^B \\ +\frac{1}{67500}(x^2 - 12x - 150) \cdot (y-18) \cdot e^A - \frac{2}{105}(x+10) \cdot (y+15) \cdot e^B, \\ \qquad +\frac{1}{202500}x(y^2 - 36y - 126)e^A - \frac{1}{45}(y^2 - 30y + 195) \cdot e^B \end{pmatrix}$$

avec $A = -\frac{(x-12)^2}{300} - \frac{(y-18)^2}{900}$ et $B = -\frac{(x+10)^2}{70} - \frac{(y+15)^2}{60}$. Par exemple, au point $\mathbb{x} = (-1, -20)$ on a ainsi $H_f(-1,-20) = \begin{pmatrix} -0.139 & 0.186 \\ 0.186 & -5.504 \end{pmatrix}$.

3.4 Optimisations sans contraintes

Les extrema de la fonction-objectif sont situés aux points où la dérivée (le gradient) s'annule. Les propriétés de la dérivée seconde (la Hessienne) permettent de dire s'il s'agit d'un minimum, d'un maximum, ou d'aucun des deux.

3.4.1 Points critiques

Si les dérivées partielles de f sont continues et que f admet un min local ou un max local au point $\mathbb{x} := (x_1, x_2 \dots x_n)$, alors

$$\nabla f(\mathbb{x}) = 0.$$

- Les points qui annulent le gradient sont appelés points critiques, ou points stationnaires : ce sont des *candidats* pour trouver le min et le max.
- Ces candidats sont généralement des extrema *locaux*.
- Trouver ces points critiques, même avec un ordinateur, même à peu près, est généralement difficile. Et même quand on en trouve un, on ne sait pas si c'est vraiment *le* meilleur (extremum global).

Exemple Avec $f(x, y) = yx^2 - 9y$, on voit que les deux points critiques sont $(-3, 0)$ et $(3, 0)$.

3.4.2 Valeurs propres

Si les dérivées d'ordre 2 de f sont continues, et que le point $\mathbb{x} := (x_1, x_2, \dots, x_n)$, est un point critique, on calcule les valeurs propres de $H_f(\mathbb{x})$, notées, dans l'ordre, $\lambda_1 \leqslant \lambda_2 \leqslant \dots \leqslant \lambda_n$:

- Si $\lambda_1\lambda_n > 0$, il y a extremum local en $\mathbb{x}$ (maximum si $\lambda_n < 0$, minimum si $\lambda_1 > 0$)
- Si $\lambda_1\lambda_n < 0$, il n'y a pas d'extremum en $\mathbb{x}$ qui est alors un col, ou point-selle.
- Si $\lambda_1\lambda_n = 0$, on ne peut pas conclure directement et les moyens nécessaires pour avancer dépassent le cadre de ce cours.

Exemple Toujours avec $f(x, y) = yx^2 - 9y$:
- Pour $\mathbb{x} = (-3, 0)$, $H_f(-3, 0) = \begin{pmatrix} 0 & -6 \\ -6 & 0 \end{pmatrix}$ et les valeurs propres sont ± 6 : le point (-3,0) est un col.
- Pour $\mathbb{x} = (+3, 0)$, $H_f(+3, 0) = \begin{pmatrix} 0 & 6 \\ 6 & 0 \end{pmatrix}$ et les valeurs propres sont aussi ± 6 : (+3,0) est aussi un col.

Raccourci visuel en 2D : en un point critique $\mathbb{x}$, la fonction ressemble à $\lambda_1 x^2 + \lambda_2 y^2$: si les valeurs propres sont négatives, c'est une approximation concave et on a maximum.

3.5 Programmation linéaire

Ancien nom mal choisi pour un cas particulier d'optimisation avec des contraintes : quand la fonction objectif est affine et que les contraintes sont affines aussi.

— Les points qui vérifient toutes les contraintes à la fois sont appelés points admissibles
— On cherche le minimum de la fonction objectif sur l'ensemble des points admissibles, noté C

Les contraintes compliquent le problème : sur le bord du domaine admissible, le gradient de la fonction objectif peut être non-nul. Mais en prenant une fonction et des contraintes affines, on profite de deux simplifications :

— Les contraintes affines forment un ensemble admissible sous forme de polyèdre *sans angle rentrant*
— L'objectif affine fait que les extrema sont forcément situés sur les coins du domaine admissible
— Enfin, les extrema sont *globaux* (il peut y avoir plusieurs points qui minimisent la fonction, mais ils la minimisent tous autant)

Une méthode pour trouver les extrema est alors de partir d'un coin admissible et de voyager de proche en proche tant qu'on peut faire diminuer la fonction objectif (si on cherche un minimum). Si tous les coins en contact avec celui considéré donnent une valeur plus élevée à la fonction-objectif, c'est un min global.

— Les "coins" sont les intersections de n contraintes, où n est le nombre de variables
— Tous les coins ne sont pas admissibles
— La méthode historique pour parcourir les coins à la recherche des extrema est la méthode du simplexe
— En dimension 2, on peut faire un dessin et procéder visuellement

Exemple On prend comme fonction-objectif $f(x, y) := x + 2y$, et on cherche son maximum avec les contraintes définies en Figure 3.3a.

$$\begin{cases} x \geqslant 0 \text{ et } y \geqslant 0 \\ x \leqslant 15 \\ x + y \geqslant 20 \\ 30x + 50y \leqslant 1160 \\ x + 3y \leqslant 60 \end{cases}$$

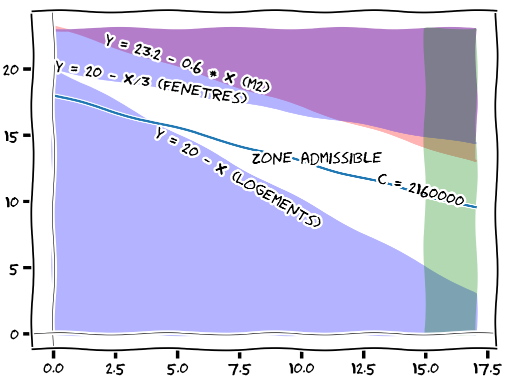


FIGURE 3.2 – Ensemble des contraintes et ensemble admissible

— On voit par exemple que le coin $(0, 0)$, intersection entre les bords $x = 0$ et $y = 0$, n'est pas admissible.
— De même, le coin $(15, 15)$, intersection entre les bords $x = 15$ et $x + 3y = 60$ n'est pas admissible.
— On commence alors, par exemple, avec le coin (0,20), intersection des bords $x + y = 20$ et $x + 3y = 60$. On a $f(0, 20) = 40$.
— On continue par exemple avec le coin (15,5), intersection des bords $x + y = 20$ et $x = 15$. On a $f(15, 5) = 25$: la valeur diminue et on repart dans l'autre sens.
— On continue alors avec le coin (12,16), intersection des bords $x + 3y = 60$ et $30x + 50y = 1160$. On a $f(12, 16) = 44$: la valeur augmente.

— On continue alors avec le coin (15,71/5), intersection des bords $x = 15$ et $30x + 50y = 1160$. On a $f(15, 14.2) = 43.4$: la valeur diminue, on peut donc s'arrêter ici en sachant que le point précédent était le maximum global.

Conclusion : f atteint son maximum sur l'ensemble admissible au point (12,16), avec la valeur 44.

Optimisation - Exercices

Exercice 1 Dérivées partielles - A L'équation d'état des gaz parfaits est donnée par :

$$PV = nRT.$$

1. On a donc $P = \frac{nRT}{V}$. C'est à dire qu'on peut voir P comme une fonction des variables n, R, T et V. Calculer alors $\partial_V P$, $\partial_T P$, $\partial_T(\partial_V P)$ et $\partial_V(\partial_T P)$ et vérifier l'égalité $\frac{\partial^2 P}{\partial V \partial T} = \frac{\partial^2 P}{\partial T \partial V}$.
2. Démontrer la relation suivante :
$$\frac{\partial T}{\partial P} \times \frac{\partial P}{\partial V} \times \frac{\partial V}{\partial T} = -1$$
(sous entendu pour tout P, V, n, R, T).

Exercice 2 Dérivées partielles - B Pour les fonctions suivantes, calculer les cinq dérivées partielles $\frac{\partial f}{\partial x}(x,y)$, $\frac{\partial f}{\partial y}(x,y)$, $\frac{\partial^2 f}{\partial x^2}(x,y)$, $\frac{\partial^2 f}{\partial y^2}(x,y)$ et $\frac{\partial^2 f}{\partial x \partial y}(x,y)$ puis les deux valeurs $\frac{\partial f}{\partial x}(\pi, 2\pi)$ et $\frac{\partial f}{\partial y}(\pi, 2\pi)$.

1. $f(x,y) = 16 - x^2 - e^{3x-y}$
2. $f(x,y) = xy + ln(3x^2 - cos(y))$
3. $f(x,y) = \sqrt{sin(x) + 5y}$

Exercice 3 Optimisation sans contraintes Déterminer, dans $\mathbb{R}^2$, les extrema locaux des fonctions suivantes, puis dire, quand on le peut, s'il s'agit d'un min local, d'un max local, d'un col :

1. $f(x,y) = x^3 + y^3 - 3x - 12y + 20$
2. $f(x,y) = x^4 + y^4 - 2(x-y)^2$. Indication : dans $\mathbb{R}$, $x^3 = -y^3 \Leftrightarrow x = -y$.
3. $f(x,y) = x^3 y^2 (1 - x - y)$

Exercice 4 Optimisation affine - A

1. Dessiner l'ensemble admissible correspondant aux contraintes suivantes :
$$\begin{cases} x \geqslant 0 \text{ et } y \geqslant 0 \\ x \leqslant 24 \\ y \leqslant 16 \\ x + 2y \geqslant 40 \\ x + y \geqslant 26 \end{cases}$$
2. Dans cet ensemble admissible, trouver le minimum de la fonction $f(x,y) := 4x + 6y$.

Exercice 5 Optimisation affine - B Un mignon petit chat à deux plaisirs dans la vie : manger des croquettes et éviscérer des oiseaux. Incluant les temps de déplacement et de sélection, manger une croquette prend (environ) 2 minutes et éventer un oiseau en prend 5. Or, le temps d'activité quotidien du mignon petit chat est inférieur à 25 minutes. Par ailleurs, les temps étants durs, le bol ne contient que 5 croquettes. Enfin, il faut savoir que chasser un oiseau fatigue le chat à hauteur de deux unités de fatigue, alors que gober une croquette le repose à hauteur de -1 unité de fatigue. Encore plein d'énergie dans la prime jeunesse, le chat peut se fatiguer à hauteur de quatre unités de fatigue.

Soit alors c le nombre de croquettes ingurgitées par jour et soit o le nombre d'oiseaux dépecés.

1. Déterminer le système de trois inéquations, portant sur c et o, traduisant les problèmes existentiels du chaton.
2. Représenter graphiquement les droites délimitant chacune de ces contraintes et hachurer la zone valide pour toutes les contraintes à la fois. On pourra par exemple prendre les croquettes en abscisses sur l'intervalle [0, 5[et les oiseaux en ordonnées sur l'intervalle [0, 5[. Un quadrillage est prévu à cet effet sur la page suivante.

Des recherches ont montré que le plaisir du chat est égal au nombre de croquettes plus 5 fois le nombre d'oiseaux attaqués.

3. Ecrire l'équation donnant le plaisir du chat noté P.
4. Déterminer à l'aide du graphique et en expliquant comment, le nombre de croquettes et d'oiseaux à manger par jour pour maximiser le plaisir en respectant les contraintes.

Réponse 1 Dérivées partielles - A

1. — $\partial_V P = -\frac{nRT}{V^2}$
   — $\partial_T P = \frac{nR}{V}$
   — $\partial_T(\partial_V P) = \partial_V(\partial_T P) = -\frac{nR}{V^2}$
2. — $\partial_P T = \frac{V}{nR}$
   — $\partial_V P = -\frac{nRT}{V^2}$
   — $\partial_T V = \frac{nR}{P}$

Réponse 2 Dérivées partielles - B

1. — $\partial_x f(x,y) = -2x - 3e^{3x-y}$
   — $\partial_y f(x,y) = e^{3x-y}$
   — $\partial_{x^2} f(x,y) = -2 - 9e^{3x-y}$
   — $\partial_{y^2} f(x,y) = -e^{3x-y}$
   — $\partial_{xy} f(x,y) = 3e^{3x-y}$
   — $\partial_x f(\pi, 2\pi) = -2\pi - 3e^{\pi}$
   — $\partial_y f(\pi, 2\pi) = e^{\pi}$
2. — $\partial_x f(x,y) = y + \frac{6x}{3x^2 - cos(y)}$
   — $\partial_y f(x,y) = x + \frac{sin(y)}{3x^2 - cos(y)}$
   — $\partial_{x^2} f(x,y) = -\frac{18x^2 + 6\,cos(y)}{\left(3x^2 - cos(y)\right)^2}$
   — $\partial_{y^2} f(x,y) = \frac{cos(y) \times \left(3x^2 - cos(y)\right) - \left(sin(y)\right)^2}{\left(3x^2 - cos(y)\right)^2}$

— $\partial_{xy} f(x,y) = 1 - \frac{6x\,sin(y)}{\left(3x^2 - cos(y)\right)^2}$
— $\partial_x f(\pi, 2\pi) = 2\pi + \frac{6\pi}{3\pi^2 - 1}$
— $\partial_y f(\pi, 2\pi) = \pi$

3. — $\partial_x f(x,y) = \frac{cos(x)}{2\sqrt{sin(x)+5y}}$
— $\partial_y f(x,y) = \frac{5}{2\sqrt{sin(x)+5y}}$
— $\partial_{x^2} f(x,y) = -\frac{1+(sin(x))^2+10y\,sin(x)}{4(sin(x)+5y)^{3/2}}$
— $\partial_{y^2} f(x,y) = -\frac{25}{4(sin(x)+5y)^{3/2}}$
— $\partial_{xy} f(x,y) = -\frac{5\,cos(x)}{4(sin(x)+5y)^{3/2}}$
— $\partial_x f(\pi, 2\pi) = -\frac{1}{2\sqrt{10\pi}}$
— $\partial_y f(\pi, 2\pi) = \frac{5}{2\sqrt{10\pi}}$

Réponse 3 Optimisation sans contraintes

1. — Points critiques : $\nabla f(x,y) = \begin{pmatrix} 3x^2 - 3 \\ 3y^2 - 12 \end{pmatrix}$ s'annule aux points (1,2), (-1,2), (1,-2), (-1,-2).
— Caractérisation des points critiques : $H_f(x,y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$
— $H_f(1,2) = \begin{pmatrix} 6 & 0 \\ 0 & 12 \end{pmatrix}$: les valeurs propres sont 6 et 12 : (1,2) est un min local
— $H_f(1,-2) = \begin{pmatrix} 6 & 0 \\ 0 & -12 \end{pmatrix}$: les valeurs propres sont 6 et -12 : (1,2) n'est ni un min, ni un max
— $H_f(-1,2) = \begin{pmatrix} -6 & 0 \\ 0 & 12 \end{pmatrix}$: les valeurs propres sont -6 et 12 : (-1,2) n'est ni un min, ni un max
— $H_f(-1,-2) = \begin{pmatrix} -6 & 0 \\ 0 & -12 \end{pmatrix}$: les valeurs propres sont -6 et -12 : (-1,-2) est un max local

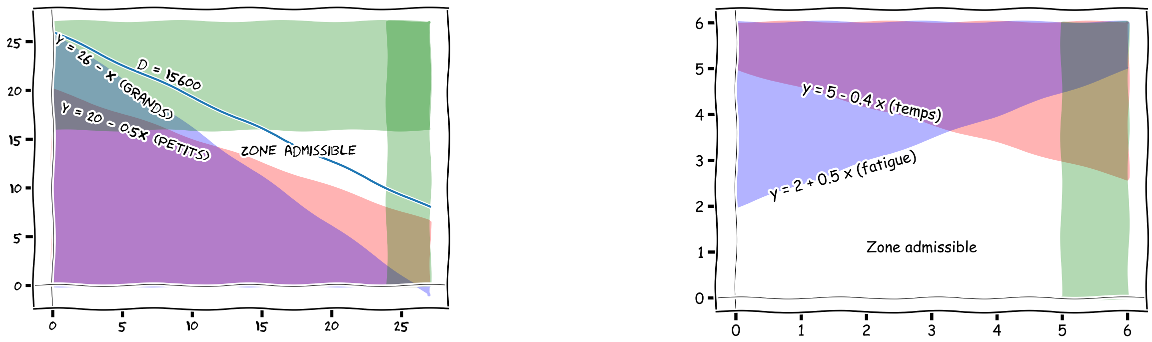
2. — Points critiques : $\nabla f(x,y) = 4 \begin{pmatrix} x^3 - (x-y) \\ y^3 + (x-y) \end{pmatrix}$ s'annule en $(0,0)$, $(\sqrt{2}, -\sqrt{2})$ et $(-\sqrt{2}, \sqrt{2})$
— Caractérisation des points critiques : $H_f(x,y) = 4 \begin{pmatrix} 3x^2 - 1 & 1 \\ 1 & 3y^2 - 1 \end{pmatrix}$
— $H_f(0,0) = 4 \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$: les valeurs propres sont -8 et 0 : on ne peut pas statuer pour $(0,0)$
— $H_f(-\sqrt{2}, \sqrt{2}) = 4 \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$: les valeurs propres sont 16 et 24 : $(-\sqrt{2}, \sqrt{2})$ est un min local
— $H_f(\sqrt{2}, +\sqrt{2}) = 4 \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$: les valeurs propres sont 16 et 24 : $(\sqrt{2}, -\sqrt{2})$ est un min local

3. — Points critiques : $\nabla f(x,y) = \begin{pmatrix} x^2y^2(3 - 4x - 3y) \\ x^3y(2 - 2x - 3y) \end{pmatrix}$ s'annule en $(0,y), \forall y \in \mathbb{R}$, $(x,0), \forall x \in \mathbb{R}$ et $(1/2, 1/3)$
— Caractérisation des points critiques : $H_f(x,y) = \begin{pmatrix} 6xy^2(1-2x-y) & x^2y(6-8x-9y) \\ x^2y(6-8x-9y) & 2x^3(1-x-3y) \end{pmatrix}$
— $H_f(0,y) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$: les valeurs propres sont 0 et 0 : on ne peut pas statuer pour $(0,y)$
— $H_f(x,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2x^3(1-x) \end{pmatrix}$: les valeurs propres sont 0 et $2x^3(1-x)$: on ne pourra pas statuer pour $(x,0)$ non plus

— $H_f(1/2,1/3) = -\begin{pmatrix} 1/9 & 1/12 \\ 1/12 & 1/8 \end{pmatrix}$: les valeurs propres sont environ -7/288 et +75/288 : $(1/2, 1/3)$ est un min local

Réponse 4 Optimisation affine - A

1. Voir Figure 3.3a
2. — En regardant f, on voit que la valeur augmente quand x ou y augmentent, on va donc chercher vers en bas à gauche
   — En regardant le dessin, on voit que seuls 3 coins sont admissibles et intéressants
   — Au coin $(10, 16)$, intersection entre $y = 16$ et $x + y = 26$, $f(10, 16) = 136$
   — Au coin $(24, 8)$, intersection entre $x = 24$ et $x + 2y = 40$, $f(24, 8) = 144$
   — Au coin $(12, 14)$, intersection entre $x + 2y = 40$ et $x + y = 26$, $f(12, 14) = 132$
   Le minimum global de f sur l'ensemble admissible est donc atteint au point $(12, 14)$.



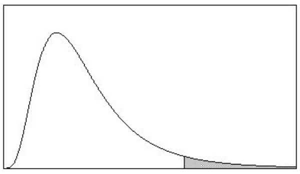
(a) Ensemble admissible A

(b) Ensemble admissible B

FIGURE 3.3 – Ensembles admissibles pour les problèmes d'optimisation affine

Réponse 5 Optimisation affine - B

1. $2 \times c + 5 \times o \leq 25$,
   $c \leq 5$,
   $-1 \times c + 2 \times o \leq 4$.
2. Voir Figure 3.3b.
3. $P = c + 5 \times o$.
4. En traçant la droite précédente pour une valeur de P quelconque on voit qu'il faut chercher vers la droite et vers le haut. Ensuite, on voit clairement sur l'image que le chat ne pourra pas tuer 4 oiseaux et que pour 3 oiseaux, le plus grand nombre de croquettes mangeables (en respectant toutes les contraintes) et 5. C'est donc dans ces conditions que le plaisir est maximal (et P=20).



F_α

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	40	60	80	100	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,7	245,4	245,9	248,0	251,1	252,2	252,7	253,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,45	19,47	19,48	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,66	8,59	8,57	8,56	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,80	5,72	5,69	5,67	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,56	4,46	4,43	4,41	4,41	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,87	3,77	3,74	3,72	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,44	3,34	3,30	3,29	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,15	3,04	3,01	2,99	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,94	2,83	2,79	2,77	2,76	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,77	2,66	2,62	2,60	2,59	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,65	2,53	2,49	2,47	2,46	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,54	2,43	2,38	2,36	2,35	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,46	2,34	2,30	2,27	2,26	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,39	2,27	2,22	2,20	2,19	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,33	2,20	2,16	2,14	2,12	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,28	2,15	2,11	2,08	2,07	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,23	2,10	2,06	2,03	2,02	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,19	2,06	2,02	1,99	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,16	2,03	1,98	1,96	1,94	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,12	1,99	1,95	1,92	1,91	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,10	1,96	1,92	1,89	1,88	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,07	1,94	1,89	1,86	1,85	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,05	1,91	1,86	1,84	1,82	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,03	1,89	1,84	1,82	1,80	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09	2,01	1,87	1,82	1,80	1,78	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,09	2,07	1,99	1,85	1,80	1,78	1,76	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17	2,13	2,10	2,08	2,06	1,97	1,84	1,79	1,76	1,74	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	1,96	1,82	1,77	1,74	1,73	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,08	2,05	2,03	1,94	1,81	1,75	1,73	1,71	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,06	2,04	2,01	1,93	1,79	1,74	1,71	1,70	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,84	1,69	1,64	1,61	1,59	1,51
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95	1,92	1,89	1,87	1,78	1,63	1,58	1,54	1,52	1,44
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84	1,75	1,59	1,53	1,50	1,48	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93	1,89	1,86	1,84	1,81	1,72	1,57	1,50	1,47	1,45	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91	1,88	1,84	1,82	1,79	1,70	1,54	1,48	1,45	1,43	1,32
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,90	1,86	1,83	1,80	1,78	1,69	1,53	1,46	1,43	1,41	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89	1,85	1,82	1,79	1,77	1,68	1,52	1,45	1,41	1,39	1,28
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,72	1,69	1,67	1,57	1,39	1,32	1,27	1,24	1,00

Table 5 : ***Valeurs seuils des lois de Fisher au risque 5% illustré par le graphique.***
En ligne : le degré de liberté du numérateur.
En colonne : le degré de liberté du dénominateur.