Méthodes numériques en actuariat avec R

Algèbre linéaire

Vincent Goulet

UNIVERSITÉ LAVAL

Méthodes numériques en actuariat avec R

Algèbre linéaire

Vincent Goulet

Professeur titulaire
École d'actuariat, Université Laval

Avec la collaboration de

Laurent Caron

Édition 2019.05

Code source
Voir sur GitLab

Couverture
Le reptile en couverture est un caméléon tapis (*Furcifer lateralis*) originaire de Madagascar. Adulte, sa taille atteint les 25 cm, queue comprise.

Crédit photo : Michabln Schwarz ; `http://fc-foto.de/2077174`

Introduction

L'algèbre linéaire a le chic de surgir là où on ne l'attend pas nécessairement. On croit qu'en étudiant l'actuariat, les notions de vecteurs et de matrices nous demeurerons étrangères. Et pourtant.

Les matrices et leur algèbre permettent de représenter, de traiter et de résoudre efficacement de grands systèmes d'équations linéaires ou d'équations différentielles. La notion d'erreur quadratique moyenne s'apparente à la projection d'un vecteur dans un espace vectoriel. Les notions d'indépendance stochastique et d'orthogonalité de vecteurs sont liées. On décrit le comportement d'une chaîne de Markov à l'aide d'une matrice de transition. Une classe de lois de probabilités requiert de calculer l'exponentielle d'une matrice. Ce ne sont là que quelques exemples où l'algèbre linéaire joue un rôle en théorie des probabilités, en inférence statistique, en finance ou en théorie du risque.

Le chapitre 7 revient sur les principales notions d'algèbre linéaire normalement étudiées au collège et qu'il est important de maîtriser dans ses études de premier cycle universitaire. Nous nous concentrons sur l'établissement de liens entre des éléments qui peuvent au premier chef sembler disparates. Le chapitre 8 construit sur le précédent pour introduire les concepts de valeurs propres, de vecteurs propres et de diagonalisation d'une matrice. Ceux-ci jouent un rôle, entre autres, en finance mathématique. Pour terminer sur des considérations numériques ayant jusque là traversé le cours, le chapitre 9 expose et compare très succinctement différentes stratégies utilisées pour résoudre des systèmes d'équations linéaires à l'aide d'un ordinateur.

Cette troisième partie de l'ouvrage diffère passablement des deux autres, autant dans sa nature que dans le format des activités d'apprentissage. La matière y est beaucoup plus mathématique et abstraite, il y a de nombreux exercices et les considérations informatiques sont réduites au minimum. Les chapitres 7 et 8 ne comportent que de courtes sections faisant la démonstration de fonctions R dédiées à l'algèbre linéaire. Seul le chapitre chapitre 8 propose un problème à résoudre au fil du texte.

Fonctionnalités interactives

En consultation électronique, ce document se trouve enrichi de plusieurs fonctionnalités interactives.

- Intraliens du texte vers une ligne précise d'une section de code informatique et, en sens inverse, du numéro de la ligne vers le point de la référence dans le texte. Ces intraliens sont marqués par la couleur .
- Intraliens entre le numéro d'un exercice et sa solution, et vice versa. Ces intraliens sont aussi marqués par la couleur .
- Intraliens entre les citations dans le texte et leur entrée dans la bibliographie. Ces intraliens sont marqués par la couleur .
- Hyperliens vers des ressources externes marqués par le symbole et la couleur .
- Table des matières, liste des tableaux, liste des figures et liste des vidéos permettant d'accéder rapidement à des ressources du document.

Blocs signalétiques

Le document est parsemé de divers types de blocs signalétiques inspirés de AsciiDoc qui visent à attirer votre attention sur une notion. Vous pourrez rencontrer l'un ou l'autre des blocs suivants.

Astuce ! Ces blocs contiennent un truc, une astuce, ou tout autre type d'information utile.

Avertissement ! Ces blocs mettent l'accent sur une notion ou fournissent une information importante pour la suite.

Attention ! Vous risquez de vous brûler — métaphoriquement s'entend — si vous ne suivez pas les recommandations de ces blocs.

Important ! Ces blocs contiennent les remarques les plus importantes. Veillez à en tenir compte.

Ces blocs contiennent des remarques additionnelles sur la matière ou des informations amusantes, mais non essentielles.

Ces blocs contiennent des liens vers des vidéos dans ma chaine YouTube dédiée à ce document de référence. Les vidéos sont répertoriées dans la liste des vidéos.

Ces blocs vous invitent à interrompre la lecture du texte pour passer à l'étude du code R des sections d'exemples.

Remarques spécifiques à macOS.

Document libre

Tout comme R et l'ensemble des outils présentés dans ce document, le projet « Méthodes numériques en actuariat avec R » s'inscrit dans le mouvement de l'informatique libre. Vous pouvez accéder à l'ensemble du code source en format LaTeX en suivant le lien dans la page de copyright. Vous trouverez dans le fichier `README.md` toutes les informations utiles pour composer le document.

Votre contribution à l'amélioration du document est également la bienvenue ; consultez le fichier `CONTRIBUTING.md` fourni avec ce document et voyez votre nom ajouté au fichier `COLLABORATEURS`.

Remerciements

Je tiens à souligner la précieuse collaboration de MM. Mathieu Boudreault, de Sébastien Auclair et de Louis-Philippe Pouliot lors de la rédaction des exercices et des solutions.

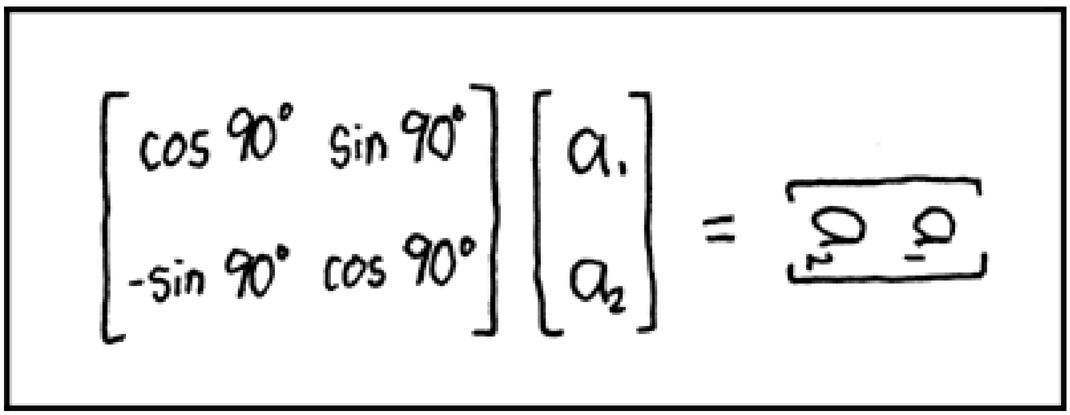
Table des matières

Liste des tableaux

Liste des figures

Liste des vidéos

Le numéro indiqué à gauche est celui de la section dans laquelle se trouve le bloc signalétique.



Tiré de XKCD.com

7 Notions fondamentales d'algèbre linéaire

Objectifs du chapitre

- Utiliser les principaux opérateurs de l'arithmétique matricielle, dont la somme, le produit, la transposée et la trace.
- Exprimer un système d'équations linéaires sous forme matricielle et le résoudre par élimination gaussienne ainsi que par produit de matrices élémentaires.
- Vérifier si des éléments d'un espace vectoriel sont linéairement indépendants.
- Définir les propriétés de l'inverse d'une matrice.
- Définir les propriétés du déterminant d'une matrice.
- Calculer l'inverse d'une matrice à partir de matrices élémentaires ainsi que par la méthode de la matrice adjointe.
- Déterminer le rang d'une matrice.
- Établir les liens entre les concepts ci-dessus.

Ce chapitre propose une révision des éléments fondamentaux d'algèbre linéaire normalement étudiés au collège : matrice, déterminant, inverse, vecteur, espace vectoriel et relations avec les systèmes d'équations linéaires. Nous tâchons d'insister sur l'établissement de liens entre les différentes notions passées en revue. À ce titre, le théorème 7.13 en fin de chapitre est particulièrement important puisqu'il unifie un grand nombre de résultats.

Le texte ne prétend à aucune exhaustivité ; le lecteur devrait donc consulter, au besoin, l'un des nombreux livres d'introduction à l'algèbre linéaire disponibles en librairie ou à la bibliothèque. Les éditions successives de Anton (2000) ont longtemps servi pour les examens professionnels d'actuariat et constituent toujours des références de choix.

Le langage R contient à peu près tous les outils nécessaires pour la manipulation et le traitement des vecteurs et matrices ; voir Goulet (2018). Nous offrons une révision dans le code informatique de la section 7.9.

7.1 Matrices

Une matrice est un tableau rectangulaire de nombres. Les matrices seront généralement représentées par des lettres majuscules et leurs éléments par des minuscules. Ainsi,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

est une matrice $m \times n$ composée des éléments a_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$. De manière plus compacte, nous écrirons parfois

$$\mathbf{A} = [a_{ij}]_{m \times n}$$

ou simplement

$$\mathbf{A} = [a_{ij}]$$

lorsque les dimensions de la matrice sont claires par le contexte.

7.1.1 Matrice carrée

Lorsqu'une matrice compte autant de lignes que de colonnes ($m = n$), la matrice est dite *carrée*. Dans un tel cas, les éléments a_{ii}, $i = 1, \dots, n$ forment la *diagonale* de la matrice.

7.1.2 Matrice symétrique

Si $a_{ij} = a_{ji}$ pour tout i, j, alors la matrice carrée $\mathbf{A}$ est *symétrique*. Par exemple,

$$\begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 1 & 0 & 6 \\ 4 & 0 & 2 & 7 \\ 5 & 6 & 7 & 1 \end{bmatrix}$$

constitue une matrice symétrique.

7.1.3 Matrices triangulaires et diagonales

Une matrice carrée dont tous les éléments sous la diagonale sont nuls est une matrice *triangulaire supérieure*. À l'inverse, une matrice carrée dont tous les éléments au-dessus de la diagonale sont nuls est une matrice *triangulaire inférieure*. Ainsi,

$$\begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

est une matrice triangulaire supérieure, alors que

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 4 & 0 & 2 & 0 \\ 5 & 6 & 7 & 1 \end{bmatrix}$$

est triangulaire inférieure.

Une matrice à la fois triangulaire inférieure et triangulaire supérieure ne comporte que des zéros à l'extérieur de la diagonale et est appelée une matrice *diagonale*, comme par exemple

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

7.1.4 Matrice identité et matrice zéro

La matrice diagonale dont tous les éléments de la diagonale sont égaux à 1 est appelée la *matrice identité*. Celle-ci joue un rôle spécial important en algèbre matricielle. Elle est généralement notée $\mathbf{I}$:

$$\mathbf{I}_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Finalement, la matrice constituée entièrement de 0 est appelée *matrice zéro* ou *matrice nulle*. Elle est généralement notée $\mathbf{0}$:

$$\mathbf{0}_{3\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

La matrice nulle est tout-à-fait différente du scalaire 0.

7.1.5 Vecteur ligne et vecteur colonne

Les matrices comptant une seule ligne ou une seule colonne portent un nom spécial. Une matrice $1 \times n$ est appelée *vecteur ligne* et une matrice $m \times 1$ est appelée *vecteur colonne* ou parfois simplement *vecteur*. Ces deux types de matrices sont généralement notées par des lettres minuscules et leurs éléments sont numérotés par un seul indice :

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

7.2 Arithmétique matricielle

Cette section récapitule l'essentiel des règles d'arithmétique matricielle. Gardez à l'esprit que ces règles sont parfois fort différentes de celles de l'arithmétique des nombres, notamment au chapitre du produit matriciel.

7.2.1 Égalité, somme et différence

Soit $\mathbf{A} = [a_{ij}]$ et $\mathbf{B} = [b_{ij}]$ deux matrices de mêmes dimensions et c, une constante (scalaire). Alors,

1. $\mathbf{A} = \mathbf{B}$ si, et seulement si, $a_{ij} = b_{ij}$ pour tout i et tout j ;
2. $\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}]$;
3. $\mathbf{A} - \mathbf{B} = [a_{ij} - b_{ij}]$;
4. $c\mathbf{A} = [ca_{ij}]$.

7.2.2 Produit

Le produit entre deux matrices, $\mathbf{AB}$, est défini seulement si le nombre de colonnes de la matrice $\mathbf{A}$ est égal au nombre de lignes de la matrice $\mathbf{B}$. Le résultat du produit est une matrice dont l'élément en position (i, j) correspond au produit scalaire entre la ligne i de la matrice $\mathbf{A}$ et la colonne j de la matrice $\mathbf{B}$. Autrement dit, soit $\mathbf{A}_{m \times p}$ et $\mathbf{B}_{p \times n}$ deux matrices. Alors le produit $\mathbf{AB}$ est défini et

$$\mathbf{AB} = \left[\sum_{k=1}^{p} a_{ik} b_{kj} \right]_{m \times n}.$$

Exemple 7.1. Soit les matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}.$$

La matrice **A** compte trois colonnes, soit autant que le nombre de lignes de la matrice **B**. Le produit **AB** est donc défini. Le résultat sera une matrice **C** de deux lignes et quatre colonnes où

$$\begin{aligned} c_{11} &= \begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} \\ &= (1 \cdot 4) + (2 \cdot 0) + (4 \cdot 2) \\ &= 12, \\ c_{12} &= \begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} \\ &= (1 \cdot 1) + (2 \cdot -1) + (4 \cdot 7) \\ &= 27, \end{aligned}$$

etc. Le résultat final est

$$\mathbf{C} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}.$$

Vous remarqueres que le produit **BA** n'est pas défini, ce qui illustre que le produit matriciel n'est pas commutatif. □

Théorème 7.1 (Propriétés de l'arithmétique matricielle)**.** *Soit* **A**, **B** *et* **C** *des matrices de dimensions telles que les opérations ci-dessous sont définies, et soit a et b, des constantes.*

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ *(commutativité de l'addition)*
2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$ *(associativité de l'addition)*
3. $\mathbf{AB} \neq \mathbf{BA}$ *en général*
4. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ *(associativité de la multiplication)*
5. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ *(distributivité par la gauche)*
6. $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$ *(distributivité par la droite)*
7. $\mathbf{A}(\mathbf{B} - \mathbf{C}) = \mathbf{AB} - \mathbf{AC}$

8. $(\mathbf{B}-\mathbf{C})\mathbf{A} = \mathbf{BA} - \mathbf{CA}$

9. $a(\mathbf{B}+\mathbf{C}) = a\mathbf{B} + a\mathbf{C}$

10. $a(\mathbf{B}-\mathbf{C}) = a\mathbf{B} - a\mathbf{C}$

11. $(a+b)\mathbf{C} = a\mathbf{C} + b\mathbf{C}$

12. $(a-b)\mathbf{C} = a\mathbf{C} - b\mathbf{C}$

13. $a(b\mathbf{C}) = (ab)\mathbf{C}$

14. $a(\mathbf{BC}) = (a\mathbf{B})\mathbf{C} = \mathbf{B}(a\mathbf{C})$

Exemple 7.2. Soit $a = 5$, $b = 2$ et

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}.$$

Il est laissé en exercice de vérifier les propriétés du théorème 7.1. Remarquez que $\mathbf{AB}$ est défini, mais pas $\mathbf{BA}$, ce qui vérifie la propriété 3, mais empêche de faire les opérations des propriétés 6 et 8. De plus, $\mathbf{BC} \neq \mathbf{CB}$. □

7.2.3 Rôles de la matrice identité et de la matrice zéro

La matrice identité joue, en arithmétique matricielle, le rôle du nombre 1 en arithmétique usuelle. Ainsi,

$$\mathbf{A}_{m\times n}\mathbf{I}_n = \mathbf{I}_m\mathbf{A}_{m\times n} = \mathbf{A}.$$

De même, la matrice zéro joue un rôle similaire à celui du 0 en arithmétique.

Théorème 7.2 (Propriétés de la matrice zéro)**.** *En supposant que les matrices sont de dimensions appropriées pour que les opérations soient définies, les propriétés ci-dessous sont valides en arithmétique matricielle.*

1. $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$

2. $\mathbf{A} - \mathbf{A} = \mathbf{0}$

3. $\mathbf{0} - \mathbf{A} = -\mathbf{A}$

4. $\mathbf{A0} = \mathbf{0}$ *et* $\mathbf{0A} = \mathbf{0}$

L'exemple suivant illustre toutefois les limites aux similitudes entre les deux systèmes arithmétiques.

Exemple 7.3. Soit les matrices

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}.$$

On vérifie aisément que

$$\mathbf{AB} = \mathbf{AC} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix},$$

d'où $\mathbf{AB} = \mathbf{AC}$ n'implique pas nécessairement $\mathbf{B} = \mathbf{C}$. De même, on vérifie que $\mathbf{AD} = \mathbf{0}$ même si $\mathbf{A} \neq \mathbf{0}$ et $\mathbf{D} \neq \mathbf{0}$. □

7.2.4 Trace

La *trace* d'une matrice carrée $\mathbf{A}_{n\times n}$, notée $\operatorname{tr}(\mathbf{A})$, est égale à la somme des éléments de la diagonale :

$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

7.2.5 Transposée

La *transposée* d'une matrice $\mathbf{A}$ est la matrice obtenue en échangeant les lignes et les colonnes de $\mathbf{A}$. Elle est généralement notée $\mathbf{A}^\mathsf{T}$ ou $\mathbf{A}'$:

$$\mathbf{A} = [a_{ij}] \Rightarrow \mathbf{A}^\mathsf{T} = [a_{ji}].$$

Théorème 7.3 (Propriétés de la transposée)**.** *Soit* $\mathbf{A}$, $\mathbf{B}$ *et* $\mathbf{C}$ *des matrices de dimensions telles que les opérations ci-dessous sont définies, et soit k une constante.*

1. $(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B})^\mathsf{T} = \mathbf{A}^\mathsf{T} + \mathbf{B}^\mathsf{T}$
3. $(k\mathbf{A})^\mathsf{T} = k\mathbf{A}^\mathsf{T}$
4. $(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$
5. $\mathbf{A}^\mathsf{T}\mathbf{A}$ *et* $\mathbf{A}\mathbf{A}^\mathsf{T}$ *sont symétriques.*

Démonstration. Nous ne démontrons que les propriétés 4 et 5. En premier lieu, soit $\mathbf{A}_{m\times p}$ et $\mathbf{B}_{p\times m}$ des matrices, $\mathbf{C} = (\mathbf{AB})^\mathsf{T}$ et $\mathbf{D} = \mathbf{B}^\mathsf{T}_{m\times p}\mathbf{A}^\mathsf{T}_{p\times m}$. L'élément c_{ij} de la matrice $\mathbf{C}$ est égal à l'élément en position (j, i) du produit $\mathbf{AB}$, d'où

$$c_{ij} = \sum_{k=1}^{p} a_{jk} b_{ki}.$$

De plus, puisque $\mathbf{A}^\mathsf{T} = [a_{ji}]$ et $\mathbf{B}^\mathsf{T} = [b_{ji}]$

$$d_{ij} = \sum_{k=1}^{p} b_{ki} a_{jk},$$

d'où $c_{ij} = d_{ij}$ pour tout $i = 1, \dots, m$ et $j = 1, \dots, m$ et, donc, $\mathbf{C} = \mathbf{D}$.

D'autre part, par la propriété 4,

$$\begin{aligned}(\mathbf{A}^\mathsf{T}\mathbf{A})^\mathsf{T} &= \mathbf{A}^\mathsf{T}(\mathbf{A}^\mathsf{T})^\mathsf{T} \\ &= \mathbf{A}^\mathsf{T}\mathbf{A},\end{aligned}$$

d'où $\mathbf{A}^\mathsf{T}\mathbf{A}$ est une matrice symétrique. On procède de même pour le second résultat de la propriété 5. □

Vous rencontrerez certainement ce concept dans la suite de vos études en actuariat : la dérivée d'une matrice. Ce qui pourrait sembler bien intimidant au départ est en fait très simple : si $\mathbf{C} = [c_{ij}(x)]_{m \times n}$ est une matrice formée de fonctions différentiables de x, alors $\mathbf{C}' = [c'_{ij}(x)]_{m \times n}$. En d'autres termes, la dérivée d'une matrice est la matrice des dérivées.

7.3 Lien avec les systèmes d'équations linéaires

L'application de loin la plus fréquente des matrices est la représentation et la résolution de systèmes d'équations linéaires. Soit le système à trois équations et trois inconnues

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}$$

On peut représenter ce système sous la forme d'une équation matricielle avec

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

soit

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

avec les définitions évidentes pour $\mathbf{A}$, $\mathbf{x}$ et $\mathbf{b}$.

Tout système d'équations linéaires comporte ou bien une seule solution, ou bien aucune solution, ou bien une infinité de solutions. La figure 7.1 illustre ces trois possibilités pour un système à deux équations et deux inconnues, de même que pour un système à trois équations et trois inconnues.

> Un système avec plus d'inconnues qu'il n'y a d'équations ($m < n$) ne peut avoir qu'aucune ou une infinité de solutions. (Il est, par exemple, impossible d'avoir un seul point commun entre deux plans.)

7.3.1 Résolution d'un système d'équations linéaires

De manière plus compacte et appropriée pour l'élimination gaussienne (voir plus bas), le système d'équation $\mathbf{A}\mathbf{x} = \mathbf{b}$ peut aussi être représenté sous forme d'une *matrice augmentée*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{bmatrix}.$$

Pour résoudre un système d'équations linéaires, il faut généralement remplacer le système donné par un autre système équivalent, mais plus simple à résoudre. Ce nouveau système est obtenu par diverses opérations :

1. multiplication d'une équation par une constante ;
2. échange de deux équations ;
3. addition d'un multiple d'une équation à une autre.

Puisque les lignes d'une matrice augmentée correspondent aux équations d'un système d'équations, les opérations ci-dessus se traduisent naturellement en opérations sur les lignes de la matrice augmentée :

1. multiplication d'une ligne par une constante ;
2. échange de deux lignes ;
3. addition d'un multiple d'une ligne à une autre.

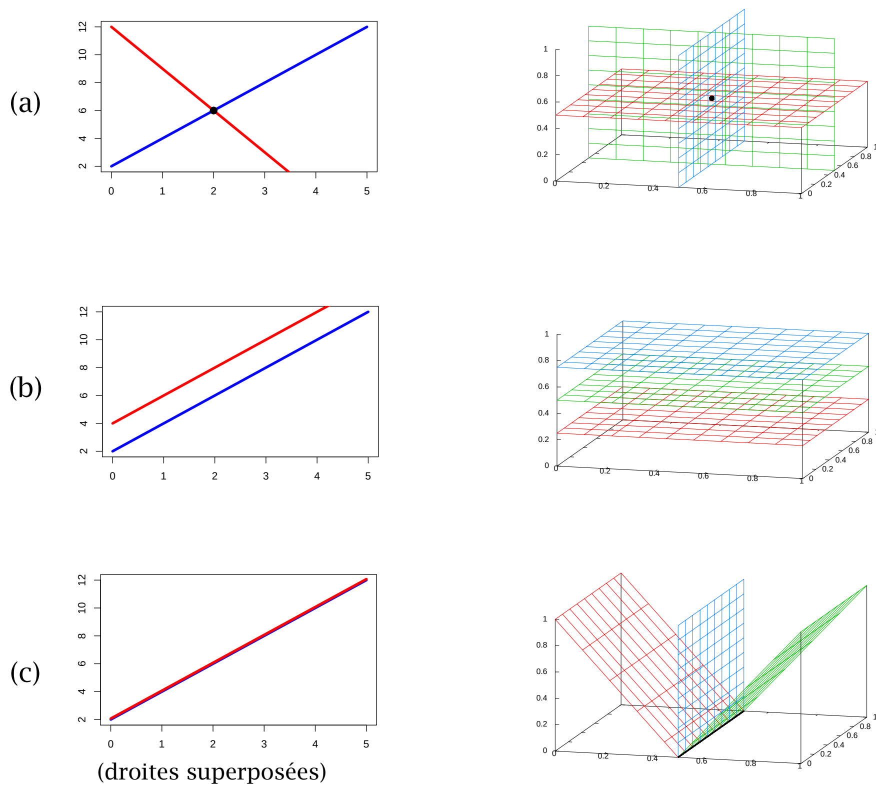


FIG. 7.1 – Solutions possibles de systèmes d'équations linéaires à deux et trois variables : (a) une seule solution ; (b) aucune solution ; (c) infinité de solutions. La ou les solutions sont illustrées par un point ou une droite en noir, lorsque possible.

Ces opérations sont appelées *opérations élémentaires sur les lignes*.

Exemple 7.4. Soit le système d'équations

$$
\begin{aligned}
x + y + 2z &= 9 \\
2x + 4y - 3z &= 1 \\
3x + 6y - 5z &= 0.
\end{aligned}
$$

La procédure de résolution classique du système d'équations se trouve dans la colonne de gauche, alors que la procédure utilisant la matrice augmentée se trouve dans la colonne de droite.

Système d'équations original :

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + 6y - 5z &= 0. \end{aligned}$$

Matrice augmentée originale :

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

Additionner -2 fois la première équation à la seconde et -3 fois la première à la troisième :

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3y - 11z &= -27. \end{aligned}$$

Additionner -2 fois la première ligne à la seconde et -3 fois la première à la troisième :

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

Additionner $-\frac{3}{2}$ fois la seconde équation à la troisième :

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ -\frac{1}{2}z &= -\frac{3}{2}. \end{aligned}$$

Additionner $-\frac{3}{2}$ fois la seconde ligne à la troisième :

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

À cette étape, nous pouvons déjà résoudre facilement le système d'équations par substitution successive : la troisième équation nous donne $z = 3$, la seconde $2y - 7(3) = -17 \Rightarrow y = 2$ et la première, $x + 2 + 2(3) = 9 \Rightarrow x = 1$.

Nous pouvons aussi continuer les opérations élémentaires jusqu'à obtenir un système trivial.

Additionner -14 fois la troisième équation à la seconde et 4 fois la troisième à la première :

$$\begin{aligned} x + y &= 3 \\ 2y &= 4 \\ -\frac{1}{2}z &= -\frac{3}{2}. \end{aligned}$$

Additionner -14 fois la troisième ligne à la seconde et 4 fois la troisième à la première :

$$\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

Additionner $-\frac{1}{2}$ fois la seconde équation à la première :

$$\begin{aligned} x &= 1 \\ 2y &= 4 \\ -\frac{1}{2}z &= -\frac{3}{2}. \end{aligned}$$

Additionner $-\frac{1}{2}$ fois la seconde ligne à la première :

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

Ce dernier système d'équations mène à la même réponse que ci-dessus. □

7.3.2 Élimination gaussienne

La procédure de la matrice augmentée utilisée pour résoudre le système d'équations de l'exemple 7.4 s'apparente à l'*élimination gaussienne*. Plus précisément, l'élimination gaussienne consiste à appliquer des opérations élémentaires à la matrice augmentée correspondant à un système d'équations linéaires jusqu'à ce que la partie gauche de celle-ci se trouve sous forme *échelonnée*, c'est-à-dire :

1. si une ligne n'est pas constituée uniquement de zéros, alors le premier élément non nul de cette ligne est un 1 ;
2. les lignes composées uniquement de zéros sont groupées au bas de la matrice ;
3. pour toutes les lignes adjacentes non composées uniquement de zéros, le premier 1 suivant des zéros apparaît plus à droite dans la ligne du dessous.

Exemple 7.5. Les matrices suivantes sont échelonnées (le symbole $*$ représente un nombre quelconque) :

$$\begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

□

Si nous ajoutons l'exigence que le premier 1 suivant des zéros d'une colonne ne soit lui-même suivi que de zéros, alors la matrice résultante est sous forme *échelonnée réduite*. La procédure de résolution du système d'équations linéaires est alors appelée *élimination de Gauss–Jordan*.

Exemple 7.6. Les matrices suivantes sont sous forme échelonnée réduite :

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\ \begin{bmatrix}1&0&0&*\\0&1&0&*\\0&0&1&*\\0&0&0&0\end{bmatrix},\ \begin{bmatrix}1&0&*&*\\0&1&*&*\\0&0&0&0\\0&0&0&0\end{bmatrix},\ \begin{bmatrix}1&0&*&0&*\\0&1&*&0&*\\0&0&0&1&*\\0&0&0&0&0\end{bmatrix}.$$

□

Exemple 7.7. La matrice échelonnée réduite d'un système d'équations linéaires est la suivante :

$$\begin{bmatrix}1&0&0&4&-1\\0&1&0&2&6\\0&0&1&3&2\end{bmatrix}.$$

Le système d'équations compte plus d'inconnues que d'équations. Il a donc une infinité de solutions. Le système correspondant à la matrice échelonnée réduite est

$$\begin{aligned}x_1 + 4x_4 &= -1\\x_2 + 2x_4 &= 6\\x_3 + 3x_4 &= 2,\end{aligned}$$

soit

$$\begin{aligned}x_1 &= -1 - 4x_4\\x_2 &= 6 - 2x_4\\x_3 &= 2 - 3x_4.\end{aligned}$$

Afin d'exprimer l'ensemble de solutions, nous avons recours à une (ou plusieurs, le cas échéant) *variable libre* égale à une valeur quelconque. Ici, nous choisirons tout naturellement de poser $x_4 = t$. Ainsi, nous obtenons la solution générale

$$x_1 = -1 - 4t, \qquad x_2 = 6 - 2t, \qquad x_3 = 2 - 3t, \qquad x_4 = t.$$

□

L'élimination gaussienne et l'élimination de Gauss–Jordan se prêtent bien à la mise en oeuvre informatique de méthodes de résolution de grands systèmes d'équations linéaires. Un algorithme efficace prendra toutefois soin de minimiser les erreurs d'arrondi ainsi que le nombre d'opérations. Nous étudierons au chapitre 9 la méthode la plus souvent retenue pour résoudre les systèmes d'équations linéaires.

7.3.3 Matrices élémentaires

Les opérations élémentaires réalisées sur une matrice peuvent être représentées comme un produit entre une *matrice élémentaire* et la matrice (dans cet ordre). Une matrice élémentaire $n \times n$ est le résultat d'une — et une seule — opération sur les lignes de la matrice identité $\mathbf{I}_n$.

Par exemple, soit la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Le produit suivant permet d'intervertir les deux lignes de $\mathbf{A}$:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}.$$

Quant au produit suivant, il permet d'additionner à la première ligne 2 fois la seconde :

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 3 & 4 \end{bmatrix}.$$

Les matrices les plus à gauche dans les produits ci-dessus sont des matrices élémentaires.

Les matrices élémentaires sont généralement notées $\mathbf{E}$. Leur principale propriété est formalisée dans le théorème suivant.

Théorème 7.4. *Soit* $\mathbf{E}$ *une matrice obtenue en faisant une opération sur les lignes de la matrice identité. Alors* $\mathbf{EA}$ *est la matrice obtenue en faisant la même opération sur la matrice* $\mathbf{A}$.

Exemple 7.8. Exprimons la solution de l'exemple 7.4 à l'aide de matrices élémentaires. À chaque opération sur les lignes de la matrice augmentée correspond une matrice élémentaire.

1. Additionner -2 fois la première ligne à la seconde :

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. Additionner -3 fois la première ligne à la troisième :

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}.$$

3. Additionner $-\frac{3}{2}$ fois la seconde ligne à la troisième :

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{3}{2} & 1 \end{bmatrix}.$$

Il est laissé en exercice de vérifier que

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1 \begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}.$$

□

Les matrices élémentaires seront utiles lorsque nous étudierons la décomposition *LU* au chapitre 9.

7.3.4 Systèmes d'équations homogènes

Un système d'équations linéaires est dit *homogène* lorsque $\mathbf{b} = \mathbf{0}$, c'est-à-dire de la forme

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

Tout système d'équations homogène compte au moins une solution, la *solution triviale* $x_1 = \cdots = x_n = 0$. Par conséquent, de tels systèmes peuvent n'admettre que la solution triviale ou une infinité de solutions.

Théorème 7.5. *Tout système d'équations homogène composé de plus d'inconnues que d'équations* $(m < n)$ *a une infinité de solutions.*

7.4 Lien avec les espaces vectoriels

De manière très générale, un *espace vectoriel* est un ensemble d'objets pour lequel sont définis un opérateur d'addition et un opérateur de produit par un scalaire, tous deux satisfaisant un certain nombre de conditions.

Définition 7.1 (Espace vectoriel)**.** Soit V un ensemble non vide d'objets. Si les axiomes ci-dessous sont satisfaits pour tous objets $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$ de V et tous scalaires k et l, alors V est un espace vectoriel.

1. Si $\mathbf{u} \in V$ et $\mathbf{v} \in V$, alors $\mathbf{u} + \mathbf{v} \in V$.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
4. Il existe un objet $\mathbf{0} \in V$ tel que $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$ pour tout $\mathbf{u} \in V$.
5. Pour tout $\mathbf{u} \in V$, il existe un objet $-\mathbf{u}$ tel que $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$.
6. Si k est un scalaire et $\mathbf{u} \in V$, alors $k\mathbf{u} \in V$.
7. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
8. $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$
9. $k(l\mathbf{u}) = (kl)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

Définition 7.2 (Produit scalaire)**.** Soit $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$ des éléments d'un espace vectoriel V et k, un scalaire. Un produit scalaire dans V est une fonction $\langle \mathbf{u}, \mathbf{v} \rangle$ associant un nombre réel aux éléments $\mathbf{u}$ et $\mathbf{v}$ de V et respectant les axiomes suivants :

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (symétrie)
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (distributivité)
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ (multiplication par un scalaire)
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ et $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ si, et seulement si, $\mathbf{v} = 0$.

Définition 7.3 (Norme)**.** La norme (longueur) d'un vecteur $\mathbf{u}$, notée $\|\mathbf{u}\|$, est la racine carrée du produit scalaire du vecteur avec lui-même :

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

Définition 7.4 (Espace pré-hilbertien)**.** Un espace pré-hilbertien (*inner product space*) est un espace vectoriel muni d'un produit scalaire.

Définition 7.5 (Orthogonalité)**.** Deux éléments $\mathbf{u}$ et $\mathbf{v}$ d'un espace pré-hilbertien sont orthogonaux si $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Définition 7.6 (Ensemble orthogonal et orthonormal)**.** Un ensemble d'éléments d'un espace pré-hilbertien est dit orthogonal si toutes les paires d'éléments distincts sont orthogonales. Si, de plus, la norme de chaque élément est 1, l'ensemble est dit orthonormal.

Exemple 7.9. L'espace euclidien $\mathbb{R}^n$ est un espace pré-hilbertien pour lequel l'addition et le produit par un scalaire sont respectivement définis par

$$\mathbf{u} + \mathbf{v} = [u_i + v_i]$$

et

$$k\mathbf{v} = [kv_i],$$

où $\mathbf{u} = (u_1, \dots, u_n)$ et $\mathbf{v} = (v_1, \dots, v_n)$ sont deux vecteurs et k est une constante. De plus, le produit scalaire dans $\mathbb{R}^n$ est défini ainsi :

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\mathsf{T}\mathbf{v} = \sum_{i=1}^{n} u_i v_i.$$

Par conséquent, la norme d'un vecteur dans $\mathbb{R}^n$ est la racine carrée de la somme des composantes au carré ou, si l'on préfère :

$$\|\mathbf{u}\|^2 = \sum_{i=1}^{n} u_i^2.$$

□

Exemple 7.10. L'ensemble des variables aléatoires de second moment fini forme un espace pré-hilbertien communément noté $\mathscr{L}^2$. En effet, la somme $X + Y$ de deux variables aléatoires de second moment fini et le produit d'une variable aléatoire par une constante kX sont des variables aléatoires de second moment fini. De plus, le produit scalaire est défini ainsi dans cet espace :

$$\langle X, Y \rangle = E[XY].$$

La notion d'orthogonalité est donc liée à celle d'indépendance stochastique dans cet espace. □

On peut voir la matrice

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

comme la réunion dans un même objet des coordonnées (dans un espace à n dimensions) de m vecteurs $\mathbf{a}_1, \dots, \mathbf{a}_m$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_m \end{bmatrix},$$

où $\mathbf{a}_i = (a_{1i}, \ldots, a_{ni})$. Les règles de l'arithmétique matricielle nous fournissent donc une arithmétique pour les espaces vectoriels.

La principale notion d'algèbre vectorielle qui nous sera utile ici est celle d'indépendance linéaire.

Soit $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un ensemble de vecteurs non nuls. Si la seule solution de l'équation vectorielle

$$k_1\mathbf{v}_1 + \cdots + k_n\mathbf{v}_n = \mathbf{0}$$

est la solution triviale $k_1 = 0, \ldots, k_n = 0$, alors V est un ensemble *linéairement indépendant*. S'il existe une autre solution que la solution triviale à l'équation, alors V est un ensemble *linéairement dépendant*.

Lorsque $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est un ensemble de vecteurs linéairement indépendants, on dit que V forme une *base* pour l'espace vectoriel.

Exemple 7.11. Soit $\mathbf{v}_1 = (2, -1, 0, 3)$, $\mathbf{v}_2 = (1, 2, 5, -1)$ et $\mathbf{v}_3 = (7, -1, 5, 8)$. Ces trois vecteurs sont linéairement dépendants puisque $3\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. Ils ne peuvent donc pas servir de base pour $\mathbb{R}^4$. □

Exemple 7.12. Les vecteurs $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ et $\mathbf{k} = (0, 0, 1)$ sont linéairement indépendants puisque

$$\begin{aligned} k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k} &= (k_1, 0, 0) + (0, k_2, 0) + (0, 0, k_3) \\ &= (k_1, k_2, k_3) \\ &= (0, 0, 0) \end{aligned}$$

seulement si $k_1 = k_2 = k_3 = 0$. Ces vecteurs forment une base pour l'espace $\mathbb{R}^3$. Puisque les vecteurs sont orthogonaux deux à deux et que leur norme est 1, $\mathbf{i}$, $\mathbf{j}$ et $\mathbf{k}$ forment aussi un ensemble orthonormal. □

Pour établir le lien entre l'indépendance linéaire et les matrices, supposons un ensemble de trois vecteurs $V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, où $\mathbf{v}_i = (v_{1i}, v_{2i}, v_{3i})$. Les vecteurs sont linéairement indépendants si

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0},$$

soit

$$k_1(v_{11}, v_{21}, v_{31}) + k_2(v_{12}, v_{22}, v_{32}) + k_3(v_{13}, v_{23}, v_{33}) = (0, 0, 0)$$

ou

$$(k_1v_{11} + k_2v_{12} + k_3v_{13}, k_1v_{21} + k_2v_{22} + k_3v_{23}, k_1v_{31} + k_2v_{32} + k_3v_{33}) = (0, 0, 0)$$

ou encore, exprimé sous forme matricielle,

$$\begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Ainsi, un ensemble de vecteurs est linéairement indépendant si la seule solution de l'équation homogène $\mathbf{A}\mathbf{x} = \mathbf{0}$, où les colonnes de $\mathbf{A}$ sont formées des coordonnées des vecteurs, est la solution triviale.

7.5 Inverse d'une matrice

En arithmétique usuelle, l'inverse d'un nombre x est le nombre y tel que $xy = 1$. Par habitude, nous écrivons simplement $y = 1/x$ ou $y = x^{-1}$. L'inverse d'une matrice est défini de manière similaire, sauf que l'opérateur de division n'existe pas en arithmétique matricielle. Prenez donc garde de ne *jamais* écrire une telle abomination :

$$\mathbf{AB} = \mathbf{C} \Rightarrow \mathbf{A} = \frac{\mathbf{C}}{\mathbf{B}}.$$

Si l'on veut isoler la matrice $\mathbf{A}$ dans l'équation $\mathbf{AB} = \mathbf{C}$, on ne peut que multiplier de part et d'autre par la matrice inverse de $\mathbf{B}$.

Définition 7.7 (Inverse d'une matrice)**.** Soit $\mathbf{A}$ une matrice carrée. S'il existe une matrice $\mathbf{B}$ tel que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I},$$

alors $\mathbf{A}$ est inversible, $\mathbf{B}$ est la matrice inverse — ou simplement inverse — de $\mathbf{A}$ et cette dernière est notée $\mathbf{A}^{-1}$. S'il n'existe pas de matrice $\mathbf{B}$ satisfaisant l'égalité ci-dessus, alors $\mathbf{A}$ est dite singulière ou non-inversible.

Exemple 7.13. Soit l'équation $\mathbf{AB} = \mathbf{C}$. Pour isoler $\mathbf{A}$, il faut multiplier de part et d'autre de l'égalité par $\mathbf{B}^{-1}$. Attention, cependant : le produit matriciel n'étant pas commutatif, le produit doit se faire par la droite :

$$(\mathbf{AB})\mathbf{B}^{-1} = \mathbf{CB}^{-1} \Leftrightarrow \mathbf{A}(\mathbf{BB}^{-1}) = \mathbf{CB}^{-1} \Leftrightarrow \mathbf{A} = \mathbf{CB}^{-1}.$$

Pour isoler $\mathbf{B}$, on devra multiplier de part et d'autre de l'égalité par $\mathbf{A}^{-1}$, mais cette fois par la gauche :

$$\mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{A}^{-1}\mathbf{C} \Leftrightarrow (\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{A}^{-1}\mathbf{C} \Leftrightarrow \mathbf{B} = \mathbf{A}^{-1}\mathbf{C}.$$

□

Théorème 7.6. *Soit* $\mathbf{A}$ *et* $\mathbf{B}$ *des matrices inversibles de mêmes dimensions. Alors on a les résultats suivants.*

1. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
2. $(\mathbf{A}^{-1})^\mathsf{T} = (\mathbf{A}^\mathsf{T})^{-1}$

Démonstration. Pour le premier résultat, on doit vérifier que $(\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{I}$. Or,

$$\begin{aligned}
(\mathbf{AB})(\mathbf{AB})^{-1} &= (\mathbf{AB})\mathbf{B}^{-1}\mathbf{A}^{-1} \\
&= \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} \\
&= \mathbf{AIA}^{-1} \\
&= \mathbf{AA}^{-1} \\
&= \mathbf{I}.
\end{aligned}$$

Pour le second résultat, il faut démontrer que $(\mathbf{A}^{-1})^\mathsf{T}\mathbf{A}^\mathsf{T} = \mathbf{A}^\mathsf{T}(\mathbf{A}^{-1})^\mathsf{T} = \mathbf{I}$. Or, par le résultat 4 du théorème 7.3 (page 7),

$$(\mathbf{A}^{-1})^\mathsf{T}\mathbf{A}^\mathsf{T} = (\mathbf{AA}^{-1})^\mathsf{T} = \mathbf{I}^\mathsf{T} = \mathbf{I}$$

et

$$\mathbf{A}^\mathsf{T}(\mathbf{A}^{-1})^\mathsf{T} = (\mathbf{A}^{-1}\mathbf{A})^\mathsf{T} = \mathbf{I}^\mathsf{T} = \mathbf{I},$$

puisque la matrice identité est symétrique. □

On a également le résultat suivant, utile lorsque l'on travaille avec les matrices élémentaires.

Théorème 7.7. *Toute matrice élémentaire est inversible et son inverse est aussi une matrice élémentaire.*

7.5.1 Calcul de l'inverse

La manière la plus usuelle de calculer l'inverse d'une matrice repose sur l'utilisation du déterminant, qui sera présenté à la section 7.6. D'ici là, nous pouvons calculer l'inverse d'une matrice par une méthode similaire à l'élimination de Gauss–Jordan.

En effet, on peut démontrer que si la matrice $\mathbf{A}$ est inversible, alors on peut trouver des matrices élémentaires $\mathbf{E}_1, \dots, \mathbf{E}_k$ tel que

$$\mathbf{E}_k \cdots \mathbf{E}_1\mathbf{A} = \mathbf{I}.$$

(Ce fait sera confirmé dans le théorème 7.13.) Par conséquent,

$$\mathbf{A}^{-1} = \mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{I},$$

ce qui signifie que l'on peut obtenir l'inverse de **A** en appliquant aux lignes de la matrice identité les opérations élémentaires nécessaires pour convertir **A** en matrice identité.

Exemple 7.14. Soit la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}.$$

L'idée pour calculer l'inverse de cette matrice consiste à convertir **A** en la matrice identité par une suite d'opérations élémentaires et à appliquer simultanément ces opérations à la matrice identité. Pour ce faire, nous joignons côte à côte les matrices **A** et **I** dans une nouvelle matrice

$$[\mathbf{A}|\mathbf{I}] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right].$$

Suite aux opérations élémentaires, on obtiendra

$$[\mathbf{I}|\mathbf{A}^{-1}].$$

La procédure est la suivante :

1. Additionner -2 fois la première ligne à la seconde et -1 fois à la troisième :
$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right]$$

2. Additionner 2 fois la seconde ligne à la troisième, puis multiplier celle-ci par -1 :
$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right].$$

3. Additionner 3 fois la troisième ligne à la seconde et -3 fois à la première :
$$\left[\begin{array}{rrr|rrr} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right].$$

4. Additionner -2 fois la seconde ligne à la première :

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right].$$

Par conséquent,

$$\mathbf{A}^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}.$$

□

Exemple 7.15. Soit la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}.$$

Il est laissé en exercice de vérifier que la procédure d'élimination pour les deuxième et troisième lignes donne

$$\left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right].$$

Comme le côté gauche contient une ligne complète de zéros, la matrice est singulière. □

7.5.2 Résolution de systèmes d'équations par l'inverse

Outre l'élimination gaussienne et l'élimination de Gauss–Jordan, les systèmes d'équations linéaires de la forme $\mathbf{Ax} = \mathbf{b}$ peuvent être résolus avec l'inverse de $\mathbf{A}$, lorsqu'elle existe.

Théorème 7.8. *Si la matrice $\mathbf{A}$ est inversible, alors le système d'équations linéaires $\mathbf{Ax} = \mathbf{b}$ a une solution unique et celle-ci est donnée par*

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Le résultat ci-dessus ne peut évidemment servir que lorsque le système d'équations compte autant d'inconnues que d'équations, c'est-à-dire lorsque la matrice $\mathbf{A}$ est carrée. Dans le cas contraire, il faut utiliser les méthodes d'élimination pour résoudre le système d'équations.

7.5.3 Matrice orthogonale

Pour certaines matrices, le calcul de l'inverse se résume à une simple transposition. Ces matrices jouent un rôle suffisamment important en algèbre linéaire pour leur attribuer un nom spécial.

Définition 7.8 (Matrice orthogonale)**.** Une matrice carrée $\mathbf{A}$ est dite orthogonale lorsque $\mathbf{A}^{-1} = \mathbf{A}^\mathsf{T}$.

De la définition précédente découle qu'une matrice $\mathbf{A}$ est orthogonale si, et seulement si,

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{A}\mathbf{A}^\mathsf{T} = I.$$

7.6 Déterminant

La fonction déterminant, notée $\det(\cdot)$, associe un nombre réel à une matrice carrée $\mathbf{A}$.

Il y a différentes façons de présenter le déterminant ; nous ne retiendrons que celle basée sur les cofacteurs. Pour ce faire, définissons tout d'abord le déterminant d'une matrice 2×2, ainsi que son inverse. La justification de ce second résultat viendra plus loin.

Théorème 7.9. *Soit* $\mathbf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. *Alors*

$$\det(\mathbf{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

et

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Soit $\mathbf{A} = [a_{ij}]_{n\times n}$ une matrice carrée. Le *mineur de l'élément* a_{ij}, noté M_{ij}, est le déterminant de la sous-matrice résultante après avoir éliminé la ligne i et la colonne j de la matrice $\mathbf{A}$. De plus, le *cofacteur de l'élément* a_{ij} est $C_{ij} = (-1)^{i+j} M_{ij}$.

Exemple 7.16. Soit

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}.$$

Une fois la première ligne et la première colonne éliminées de la matrice **A**, on a

$$M_{11} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 16,$$

d'où $C_{11} = (-1)^2 M_{11} = M_{11} = 16$. De même, après avoir éliminé la troisième ligne et la deuxième colonne, on obtient

$$M_{32} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

et $C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -26$. □

Le déterminant d'une matrice **A** est le produit scalaire entre les éléments d'une ligne ou d'une colonne quelconque de la matrice et les cofacteurs de cette ligne ou colonne.

Théorème 7.10. *Soit* **A** *une matrice carrée* $n \times n$. *Alors, pour tout* $1 \leq i \leq n$ *et* $1 \leq j \leq n$,

$$\begin{aligned} \det(\mathbf{A}) &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ &= \sum_{j=1}^{n} a_{ij}C_{ij} \end{aligned}$$

ou

$$\begin{aligned} \det(\mathbf{A}) &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \\ &= \sum_{i=1}^{n} a_{ij}C_{ij}. \end{aligned}$$

Exemple 7.17. Calculons le déterminant de la matrice de l'exemple 7.16. En faisant le calcul à partir de la première ligne, nous obtenons

$$\begin{aligned} \det(\mathbf{A}) &= 3(-1)^2 \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} + (-1)^3 \begin{vmatrix} 2 & 6 \\ 1 & 8 \end{vmatrix} + (-4)(-1)^4 \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} \\ &= 3 \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} - \begin{vmatrix} 2 & 6 \\ 1 & 8 \end{vmatrix} - 4 \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} \\ &= 3(16) - 10 - 4(3) \\ &= 26. \end{aligned}$$

La réponse est la même avec un développement selon la deuxième colonne :

$$\begin{aligned}\det(\mathbf{A}) &= -\begin{vmatrix} 2 & 6 \\ 1 & 8 \end{vmatrix} + 5\begin{vmatrix} 3 & -4 \\ 1 & 8 \end{vmatrix} - 4\begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} \\ &= -10 + 5(28) - 4(26) \\ &= 26.\end{aligned}$$

□

7.6.1 Propriétés du déterminant

Du théorème 7.10, il est clair que si une matrice contient une ligne ou une colonne complète de zéros, alors son déterminant est nul.

De plus, si $\mathbf{A}$ est une matrice triangulaire (ou diagonale), alors son déterminant est égal au produit des éléments de la diagonale :

$$\det(\mathbf{A}) = \prod_{i=1}^{n} a_{ii}.$$

Le théorème suivant énonce quelques autres propriétés de la fonction déterminant.

Théorème 7.11. *Soit* $\mathbf{A}$ *une matrice* $n \times n$ *inversible.*

1. $\det(\mathbf{A}^\mathsf{T}) = \det(\mathbf{A})$
2. $\det(k\mathbf{A}) = k^n \det(\mathbf{A})$
3. $\det(\mathbf{A} + \mathbf{B}) \neq \det(\mathbf{A}) + \det(\mathbf{B})$ *en général.*
4. $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$
5. $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$

7.6.2 Calcul de l'inverse avec la matrice adjointe

La manière de loin la plus usuelle de calculer l'inverse d'une matrice — à la main du moins — est intimement liée aux notions de cofacteur et de déterminant.

Avant toute chose, définissons $[C_{ij}]_{n\times n}$, la *matrice des cofacteurs* de la matrice $\mathbf{A}$. La transposée de cette matrice des cofacteurs est appelée la *matrice adjointe* de $\mathbf{A}$ et est notée $\operatorname{adj}(\mathbf{A})$.

Exemple 7.18. Il est laissé en exercice de vérifier que la matrice adjointe de la matrice de l'exemple 7.16 est

$$\begin{bmatrix} 16 & -24 & 26 \\ -10 & 28 & -26 \\ 3 & -11 & 13 \end{bmatrix}.$$

□

Théorème 7.12. *Si* $\mathbf{A}$ *est une matrice inversible, alors*

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\,\mathrm{adj}(\mathbf{A}).$$

7.7 Rang

Le *rang* d'une matrice $\mathbf{A}_{m\times n}$ correspond au nombre de lignes ou de colonnes linéairement indépendantes dans cette matrice. Évidemment,

$$\mathrm{rang}(\mathbf{A}) \leq \min(m, n).$$

Lorsqu'il y a égalité, on dit de la matrice qu'elle est de rang complet.

7.8 Unification des résultats

Le théorème suivant unifie un grand nombre de résultats et de concepts présentés dans les sections précédentes. En ce sens, il fait le sommaire des plus importants résultats à connaître en algèbre linéaire.

Théorème 7.13. *Soit* $\mathbf{A}$ *une matrice* $n \times n$*. Les énoncés suivants sont équivalents.*

1. $\mathbf{A}$ *est inversible.*
2. $\det(\mathbf{A}) \neq 0$
3. $\mathrm{rang}(\mathbf{A}) = n$
4. *La seule solution de* $\mathbf{Ax} = \mathbf{0}$ *est la solution triviale.*
5. *Le système d'équations* $\mathbf{Ax} = \mathbf{b}$ *a une solution pour tout vecteur* $\mathbf{b}$ *et celle-ci est unique.*
6. $\mathbf{A}$ *peut être exprimée comme un produit de matrices élémentaires.*
7. *Les vecteurs lignes de* $\mathbf{A}$ *sont linéairement indépendants.*

8. *Les vecteurs colonnes de* $\mathbf{A}$ *sont linéairement indépendants.*
9. *Les vecteurs lignes de* $\mathbf{A}$ *forment une base de* $\mathbb{R}^n$*.*
10. *Les vecteurs colonnes de* $\mathbf{A}$ *forment une base de* $\mathbb{R}^n$*.*

Le code informatique de la section 7.9 passe en revue les fonctions d'algèbre linéaire offertes dans R.

7.9 Code informatique

Fichier d'accompagnement `notions_fondamentales.R`

```
###
### MATRICES ET VECTEURS
###

## La fonction 'diag' a plusieurs usages. Un premier consiste
## à extraire la diagonale d'une matrice carrée.
(a <- matrix(1:9, nrow = 3))
diag(a)

## Si la matrice passée en argument est rectangulaire 'm x n',
## 'diag' retourne la liste des éléments a[i, i], i = 1, ...,
## min(m, n).
(a <- matrix(1:12, nrow = 4))
diag(a)

## Lorsque l'on passe un vecteur à la fonction 'diag', le
## résultat est une matrice diagonale formée du vecteur.
diag(1:3)

## Enfin, lorsque l'argument de 'diag' est un scalaire 'n', le
## résultat est la matrice identité 'n x n'.
diag(3)

## La fonction 'lower.tri' (respectivement 'upper.tri')
## retourne une matrice de la même dimension que celle passée
## en argument avec des valeurs TRUE (FALSE) aux positions des
## éléments sous la diagonale et FALSE (TRUE) aux positions
## au-dessus de la diagonale. La diagonale sera dans le
## premier ou dans le second groupe selon que l'argument
```

```
## 'diag' est TRUE ou FALSE.
(a <- matrix(1:9, nrow = 3))
lower.tri(a)
upper.tri(a, diag = TRUE)

## On utilise les fonctions 'lower.tri' et 'upper.tri'
## principalement pour extraire les éléments au-dessus ou
## au-dessous de la diagonale d'une matrice carrée.
a[lower.tri(a)]
a[upper.tri(a)]

## Bien que rarement nécessaires en R, on peut créer des
## vecteurs ligne ou colonne avec les fonctions 'rbind' et
## 'cbind', dans l'ordre.
rbind(1:3)
dim(rbind(1:3))
cbind(1:3)

###
### ARITHMÉTIQUE MATRICIELLE
###

## Le produit matriciel s'effectue en R avec l'opérateur
## '%*%', et non avec '*', qui fait le produit élément par
## élément (une opération qui n'est pas définie en
## arithmétique matricielle usuelle, mais qui l'est dans R).
(a1 <- matrix(c(1, 2, 2, 6, 4, 0), nrow = 2))
(a2 <- matrix(c(4, 0, 2, 1, -1, 7, 4, 3, 5, 3, 1, 2), nrow = 3))
a1 %*% a2

## Il n'y a pas d'opérateur pour calculer la trace d'une
## matrice carrée. Il suffit de sommer les éléments de la
## diagonale.
a
sum(diag(a))

## La transposée est obtenue avec la fonction 't'.
t(a)

###
### DÉTERMINANT ET INVERSE
###

## Le déterminant d'une matrice est obtenu avec la fonction
## 'det'.
```

```
A <- matrix(c(1, 2, 1, 2, 5, 0, 3, 3, 8), nrow = 3)
det(A)

## Avec pour seul argument une matrice carrée, la fonction
## 'solve' calcule l'inverse de cette matrice.
solve(A)

## Avec deux arguments, une matrice 'A' et un vecteur 'b',
## 'solve' calcule la solution du système d'équations 'Ax =
## b', c'est-à-dire 'A^{-1} b'.
b <- 1:3
solve(A, b)

## À noter qu'il est *beaucoup* plus rapide de calculer
## directement la solution du système d'équations avec
## 'solve(A, b)' que d'inverser la matrice et calculer la
## solution par la suite avec 'solve(A) %*% b'. Pour en faire
## la démonstration, on doit utiliser de grands objets.
A <- matrix(rnorm(500^2), nrow = 500)
b <- rnorm(500)
system.time(solve(A) %*% b)
system.time(solve(A, b))
```

7.10 Exercices

7.1 Pour chaque cas ci-dessous, supposer que la matrice donnée est une matrice augmentée d'un système d'équations linéaires exprimée sous forme échelonnée. Résoudre le système d'équations.

a) $\begin{bmatrix} 1 & -3 & 4 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & 8 & -5 & 6 \\ 0 & 1 & 4 & -9 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 7 & -2 & 0 & -8 & -3 \\ 0 & 0 & 1 & 1 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -3 & 7 & 1 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

7.2 Résoudre chacun des systèmes d'équations suivants par l'élimination gaussienne et l'élimination de Gauss–Jordan.

a) $$\begin{aligned} x_1 + x_2 + 2x_3 &= 8 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 3x_1 - 7x_2 + 4x_3 &= 10 \end{aligned}$$

b) $$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$$

c) $$\begin{aligned} x - y + 2z - w &= -1 \\ 2x + y - 2z - 2w &= -2 \\ -x + 2y - 4z + w &= 1 \\ 3x \qquad\qquad - 3w &= -3 \end{aligned}$$

d) $$\begin{aligned} - 2b + 3c &= 1 \\ 3a + 6b - 3c &= -2 \\ 6a + 6b + 3c &= 5 \end{aligned}$$

7.3 Résoudre chacun des systèmes d'équations suivants par l'élimination gaussienne et l'élimination de Gauss–Jordan.

a) $$\begin{aligned} 5x_1 - 2x_2 + 6x_3 &= 0 \\ -2x_1 + x_2 + 3x_3 &= 1 \end{aligned}$$

b) $$\begin{aligned} x_1 - 2x_2 + x_3 - 4x_4 &= 1 \\ x_1 + 3x_2 + 7x_3 + 2x_4 &= 2 \\ x_1 - 12x_2 - 11x_3 - 16x_4 &= 5 \end{aligned}$$

c) $$\begin{aligned} w + 2x - y &= 4 \\ x - y &= 3 \\ w + 3x - 2y &= 7 \\ 2u + 4v + w + 7x \qquad &= 7 \end{aligned}$$

7.4 Résoudre les systèmes d'équations homogènes suivants.

a) $$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 2x_2 \qquad &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

b) $$\begin{aligned} 3x_1 + x_2 + x_3 + x_4 &= 0 \\ 5x_1 - x_2 + x_3 - x_4 &= 0 \end{aligned}$$

c) $$\begin{aligned} 2x + 2y + 4z &= 0 \\ w \qquad - y - 3z &= 0 \\ 2w + 3x + y + z &= 0 \\ -2w + x + 3y - 2z &= 0 \end{aligned}$$

7.5 Pour quelle valeur de a le système d'équations

$$\begin{aligned} x + 2y - \qquad 3z &= 4 \\ 3x - y \qquad\quad 5z &= 2 \\ 4x + y + (a^2 - 14)z &= a + 2 \end{aligned}$$

n'aura-t-il aucune solution ? Une seule solution ? Une infinité de solutions ?

7.6 Soit les matrices $\mathbf{A}_{4\times 5}$, $\mathbf{B}_{4\times 5}$, $\mathbf{C}_{5\times 2}$, $\mathbf{D}_{4\times 2}$ et $\mathbf{E}_{5\times 4}$. Lesquelles des expressions matricielles suivantes sont définies ? Pour les expressions définies, donner les dimensions de la matrice résultante.

a) $\mathbf{BA}$ b) $\mathbf{AC}+\mathbf{D}$ c) $\mathbf{AE}+\mathbf{B}$ d) $\mathbf{AB}+\mathbf{B}$
e) $\mathbf{E}(\mathbf{A}+\mathbf{B})$ f) $\mathbf{E}(\mathbf{AC})$ g) $\mathbf{E}^\mathsf{T}\mathbf{A}$ h) $(\mathbf{A}^\mathsf{T}+\mathbf{E})\mathbf{D}$

7.7 Soit les matrices

$$\mathbf{A}=\begin{bmatrix}3 & 0\\ -1 & 2\\ 1 & 1\end{bmatrix}\qquad \mathbf{B}=\begin{bmatrix}4 & -1\\ 0 & 2\end{bmatrix}\qquad \mathbf{C}=\begin{bmatrix}1 & 4 & 2\\ 3 & 1 & 5\end{bmatrix}$$

$$\mathbf{D}=\begin{bmatrix}1 & 5 & 2\\ -1 & 0 & 1\\ 3 & 2 & 4\end{bmatrix}\qquad \mathbf{E}=\begin{bmatrix}6 & 1 & 3\\ -1 & 1 & 2\\ 4 & 1 & 3\end{bmatrix}.$$

Calculer les expressions suivantes (lorsque possible).

a) $\mathbf{D}+\mathbf{E}$
b) $\mathbf{AB}$
c) $(2\mathbf{D}^\mathsf{T}-\mathbf{E})\mathbf{A}$
d) $(4\mathbf{B})\mathbf{C}+2\mathbf{B}$
e) $(-\mathbf{AC})^\mathsf{T}+5\mathbf{D}^\mathsf{T}$
f) $(\mathbf{BA}^\mathsf{T}-2\mathbf{C})^\mathsf{T}$
g) $\mathbf{B}^\mathsf{T}(\mathbf{CC}^\mathsf{T}-\mathbf{A}^\mathsf{T}\mathbf{A})$
h) $\mathbf{D}^\mathsf{T}\mathbf{E}^\mathsf{T}-(\mathbf{ED})^\mathsf{T}$
i) $\mathrm{tr}(\mathbf{DD}^\mathsf{T})$
j) $\mathrm{tr}(4\mathbf{E}^\mathsf{T}-\mathbf{D})$

7.8 Démontrer que si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices $n\times n$, alors $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A})+\mathrm{tr}(\mathbf{B})$, où $\mathrm{tr}(\mathbf{A})=\sum_{i=1}^n a_{ii}$.

7.9 Une matrice $\mathbf{A}$ est dite *symétrique* si $\mathbf{A}^\mathsf{T}=\mathbf{A}$ et *antisymétrique* si $\mathbf{A}^\mathsf{T} = -\mathbf{A}$. Démontrer les énoncés suivants, étant donné une matrice carrée $\mathbf{A}$.

a) $\mathbf{AA}^\mathsf{T}$ et $\mathbf{A}^\mathsf{T}\mathbf{A}$ sont symétriques.

b) $\mathbf{A}-\mathbf{A}^\mathsf{T}$ est antisymétrique.

7.10 Démontrer que si $\mathbf{A}^\mathsf{T}\mathbf{A}=\mathbf{A}$, alors $\mathbf{A}$ est symétrique et $\mathbf{A}=\mathbf{A}^2$. Une matrice qui satisfait cette dernière égalité est dite *idempotente*.

7.11 Démontrer chacun des résultats ci-dessous si les éléments des matrices $\mathbf{A}$ et $\mathbf{B}$ sont des fonctions différentiables de x et que les dimensions des matrices sont appropriées.

a) $\dfrac{d}{dx}(k\mathbf{A}) = k\dfrac{d}{dx}\mathbf{A}$

b) $\dfrac{d}{dx}(\mathbf{A}+\mathbf{B}) = \dfrac{d}{dx}\mathbf{A}+\dfrac{d}{dx}\mathbf{B}$

c) $\dfrac{d}{dx}(\mathbf{AB}) = \dfrac{d\mathbf{A}}{dx}\mathbf{B}+\mathbf{A}\dfrac{d\mathbf{B}}{dx}$

7.12 Lesquels des ensembles de vecteurs dans $\mathbb{R}^3$ suivants sont linéairement dépendants ?

a) $(4,-1,2)$, $(-4,10,2)$

b) $(-3,0,4)$, $(5,-1,2)$, $(1,1,3)$

c) $(8,-1,3)$, $(4,0,1)$

d) $(-2,0,1)$, $(3,2,5)$, $(6,-1,1)$, $(7,0,-2)$

7.13 Soit $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$. Déterminer si les trois vecteurs se trouvent dans un plan.

a) $\mathbf{v}_1 = (2,-2,0)$, $\mathbf{v}_2 = (6,1,4)$, $\mathbf{v}_3 = (2,0,-4)$

b) $\mathbf{v}_1 = (-6,7,2)$, $\mathbf{v}_2 = (3,2,4)$, $\mathbf{v}_3 = (4,-1,2)$

7.14 Soit la matrice

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

où $\prod_{i=1}^{n} a_{ii} \neq 0$. Démontrer que $\mathbf{A}$ est inversible et trouver son inverse.

7.15 Soit les matrices

$$\mathbf{A} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 2 & -7 & 3 \end{bmatrix}.$$

Trouver les matrices élémentaires $\mathbf{E}_1$, $\mathbf{E}_2$ et $\mathbf{E}_3$ satisfaisant les équations suivantes.

a) $\mathbf{E}_1\mathbf{A} = \mathbf{B}$ b) $\mathbf{E}_2\mathbf{B} = \mathbf{A}$ c) $\mathbf{E}_3\mathbf{A} = \mathbf{C}$ d) $\mathbf{E}_4\mathbf{C} = \mathbf{A}$

7.16 Utiliser la méthode des opérations élémentaires sur les lignes pour trouver l'inverse des matrices suivantes, si elle existe. Le cas échéant, vérifier votre réponse par multiplication de la matrice et de son inverse.

a) $\begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$ b) $\begin{bmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \\ 2 & 7 & 7 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

7.17 Trouver l'inverse des matrices suivantes, où $k, k_1, \ldots, k_4$ sont des constantes non nulles.

a) $\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{bmatrix}$

7.18 Soit la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix}.$$

a) Trouver des matrices élémentaires $\mathbf{E}_1$ et $\mathbf{E}_2$ tel que $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$.

b) Écrire $\mathbf{A}^{-1}$ comme le produit de deux matrices élémentaires.

c) Écrire $\mathbf{A}$ comme le produit de deux matrices élémentaires.

7.19 Résoudre le système d'équations suivant par la méthode de la matrice inverse :

$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ x_1 - x_2 + x_3 &= 3 \\ x_1 + x_2 &= 4. \end{aligned}$$

7.20 Soit la matrice et le vecteur

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

a) Vérifier que l'équation $\mathbf{Ax} = \mathbf{x}$ peut s'écrire sous la forme $(\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ et utiliser cette formule pour résoudre $\mathbf{Ax} = \mathbf{x}$ pour $\mathbf{x}$.

b) Résoudre $\mathbf{Ax} = 4\mathbf{x}$.

7.21 Sans faire aucun calcul, déterminer si les systèmes d'équations homogènes ci-dessous admettent une solution autre que la solution triviale, puis déterminer si la matrice des coefficients est inversible.

a) $$\begin{aligned} 2x_1 - x_2 - 3x_3 + x_4 &= 0 \\ 5x_2 + 4x_3 + 3x_4 &= 0 \\ x_3 + 2x_4 &= 0 \\ 3x_4 &= 0 \end{aligned}$$

b) $$\begin{aligned} 5x_1 - x_2 + 4x_3 + x_4 &= 0 \\ 2x_3 - x_4 &= 0 \\ x_3 + x_4 &= 0 \\ 7x_4 &= 0 \end{aligned}$$

7.22 Par simple inspection visuelle de la matrice

$$\mathbf{A} = \begin{bmatrix} -2 & 8 & 1 & 4 \\ 3 & 2 & 5 & 1 \\ 1 & 10 & 6 & 5 \\ 4 & -6 & 4 & -3 \end{bmatrix},$$

expliquer pourquoi $\det(\mathbf{A}) = 0$.

7.23 Déterminer si les matrices suivantes sont inversibles.

a) $\begin{bmatrix} 1 & 0 & -1 \\ 9 & -1 & 4 \\ 9 & 9 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 4 & 2 & 8 \\ -2 & 1 & -4 \\ 3 & 1 & 6 \end{bmatrix}$

c) $\begin{bmatrix} \sqrt{2} & -\sqrt{7} & 0 \\ 3\sqrt{2} & -3\sqrt{7} & 0 \\ 5 & -9 & 0 \end{bmatrix}$ d) $\begin{bmatrix} -3 & 0 & 1 \\ 5 & 0 & 6 \\ 8 & 0 & 3 \end{bmatrix}$

7.24 Soit $\mathbf{A}$ une matrice 3×3 dont $\det(\mathbf{A}) = -7$. Trouver les déterminants suivants.

a) $\det(3\mathbf{A})$ b) $\det(\mathbf{A}^{-1})$ c) $\det(2\mathbf{A}^{-1})$ d) $\det((2\mathbf{A})^{-1})$

7.25 Pour quelle(s) valeur(s) de k les matrices ci-dessous ne sont pas inversibles ?

a) $\begin{bmatrix} k-3 & -2 \\ -2 & k-2 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 1 & 6 \\ k & 3 & 2 \end{bmatrix}$

7.26 Soit

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 1 & 6 \\ 0 & 0 & -3 & 3 \\ 4 & 1 & 0 & 14 \\ 4 & 1 & 3 & 2 \end{bmatrix}.$$

Trouver les valeurs suivantes.

a) M_{13} et C_{13}
b) M_{23} et C_{23}
c) M_{22} et C_{22}
d) M_{21} et C_{21}
e) $\det(\mathbf{A})$

Réponses

7.1 a) $x_1 = -37, x_2 = -8, x_3 = 5$
b) $x_1 = -10 + 13t, x_2 = -5 + 13t, x_3 = 2 - t, x_4 = t$
c) $x_1 = -11 - 7s + 2t, x_2 = s, x_3 = -4 - 3t, x_4 = 9 - 3t, x_5 = t$
d) pas de solution

7.2 a) $x_1 = 3, x_2 = 1, x_3 = 2$
b) $x_1 = -\frac{1}{7} - \frac{3}{7}t, x_2 = \frac{1}{7} - \frac{4}{7}t, x_3 = t$
c) $x_1 = t - 1, x_2 = 2s, x_3 = s, x_4 = t$
d) pas de solution

7.3 a) $x_1 = 2 - 12t, x_2 = 5 - 27t, x_3 = t$
b) pas de solution
c) $u = -2s - 3t - 6, v = s, w = -t - 2, x = t + 3, y = t$

7.4 a) solution triviale
b) $x_1 = -s, x_2 = -t - s, x_3 = 4s, x_4 = t$
c) $w = t, x = -t, y = t, z = 0$

7.5 Aucune solution : $a = -4$. Une seule solution : $a \neq \pm 4$. Infinité de solutions : $a = 4$.

7.6 a) non défini b) 4×2 c) non défini d) non défini e) 5×5 f) 5×2 g) non défini h) 5×2

7.7 a) $\begin{bmatrix} 7 & 6 & 5 \\ -2 & 1 & 3 \\ 7 & 3 & 7 \end{bmatrix}$ b) $\begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$ c) $\begin{bmatrix} -6 & -3 \\ 36 & 0 \\ 4 & 7 \end{bmatrix}$ d) non défini

e) $\begin{bmatrix} 2 & -10 & 11 \\ 13 & 2 & 5 \\ 4 & -3 & 13 \end{bmatrix}$ f) $\begin{bmatrix} 10 & -6 \\ -14 & 2 \\ -1 & -8 \end{bmatrix}$ g) $\begin{bmatrix} 40 & 72 \\ 26 & 42 \end{bmatrix}$ h) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ i) 61 j) 35

7.12 d) seulement

7.13 a) non b) oui

7.15 a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$

7.16 a) $\begin{bmatrix} \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ -1 & 1 & 1 \\ -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{bmatrix}$ b) n'existe pas c) $\dfrac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

d) $\begin{bmatrix} \frac{7}{2} & 0 & -3 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ e) $\dfrac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 2 \\ 1 & 1 & -1 \end{bmatrix}$

7.17 a) $\begin{bmatrix} k_1^{-1} & 0 & 0 & 0 \\ 0 & k_2^{-1} & 0 & 0 \\ 0 & 0 & k_3^{-1} & 0 \\ 0 & 0 & 0 & k_4^{-1} \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 0 & k_1^{-1} \\ 0 & 0 & k_2^{-1} & 0 \\ 0 & k_3^{-1} & 0 & 0 \\ k_4^{-1} & 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} k^{-1} & 0 & 0 & 0 \\ -k^{-2} & k^{-1} & 0 & 0 \\ k^{-3} & -k^{-2} & k^{-1} & 0 \\ -k^{-4} & k^{-3} & -k^{-2} & k^{-1} \end{bmatrix}$

7.18 a) $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$, $\mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$

7.19 $x_1 = \frac{16}{3}$, $x_2 = -\frac{4}{3}$, $x_3 = -\frac{11}{3}$

7.21 a) solution triviale seulement, matrice inversible b) infinité de solutions, matrice singulière

7.23 a) oui b) non c) non d) non

7.24 a) -189 b) $-\frac{1}{7}$ c) $-\frac{8}{7}$ d) $-\frac{1}{56}$

7.25 a) $(5 \pm \sqrt{17})/2$ b) -1

7.26 a) $M_{13} = 0$, $C_{13} = 0$

b) $M_{23} = -96$, $C_{23} = 96$

c) $M_{22} = -48$, $C_{22} = -48$

d) $M_{21} = 72$, $C_{21} = -72$

e) -216

8 Valeurs propres, vecteurs propres et diagonalisation

Objectifs du chapitre

- Énoncer les définitions de valeur propre et de vecteur propre d'une matrice.
- Calculer les valeurs propres d'une matrice et les vecteurs propres correspondants.
- Déterminer si une matrice est diagonalisable ou non.
- Diagonaliser une matrice carrée.

Les valeurs et vecteurs propres sont des caractéristiques liées à une matrice carrée. Ces notions admettent une grande variété d'interprétations physiques différentes, chacune ancrée dans son domaine d'application : géométrie, résolution d'équations différentielles, mécanique quantique, finance, moteurs de recherche, etc.

Dans le cadre de ce cours, l'étude des valeurs et vecteurs propres demeurera plutôt collée aux définitions de base. Outre l'exemple résolu du chapitre, nous ne verrons qu'une seule application, celle que vous êtes la plus susceptible de rencontrer dans vos études en actuariat, soit la diagonalisation d'une matrice.

Visionnez la vidéo qui explique comment Google utilise les vecteurs propres pour établir son classement des résultats de recherche.

8.1 Énoncé du problème

Depuis le milieu des années 1960, les systèmes postaux de tri acheminent une énorme proportion du courrier vers la bonne adresse tout à fait automatiquement. Pour y arriver, ils ont recours à la reconnaissance optique de caractères (*optimal character recognition*, OCR), un procédé informatique par lequel une image est convertie en texte.

La reconnaissance optique de caractères repose généralement, après divers pré-traitements, sur le découpage de chaque caractère en un certain nombre de « pixels », chacun de ceux-ci pouvant être « blanc » ou « noir ». À titre d'exemple, la base de données d'apprentissage machine Optical Recognition of Handwritten Digits (Dua et Graff, 2019) contient la représentation matricielle de taille 32×32 de 1 797 chiffres entre 0 et 9 écrits à la main. Le tout premier caractère de cette base de données est représenté à la figure 8.1.

Afin de réduire la dimension des données, l'information est enregistrée dans la base de données sous forme du nombre de pixels noirs dans des carrés disjoints de taille 4×4. Ces carrés forment la grille visible dans la figure 8.1. L'objet `optd` du paquetage R **RSKC** (Kondo et collab., 2016) est une matrice de 1 797 lignes et de 64 colonnes qui contient les représentations compactes des chiffres de la base de données. Par exemple, celle du chiffre de la figure 8.1 est la suivante (présentée sous forme de matrice pour faciliter la comparaison avec l'image) :

```
> matrix(optd[1, ], nrow = 8, byrow = TRUE)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,]    0    0    5   13    9    1    0    0
[2,]    0    0   13   15   10   15    5    0
[3,]    0    3   15    2    0   11    8    0
[4,]    0    4   12    0    0    8    8    0
[5,]    0    5    8    0    0    9    8    0
[6,]    0    4   11    0    1   12    7    0
[7,]    0    2   14    5   10   12    0    0
[8,]    0    0    6   13   10    0    0    0
```

Pour effectuer la reconnaissance optique de caractères, nous souhaitons classer, à l'aide d'un nombre restreint de caractéristiques, les différentes images de la base de données en fonction du chiffre auquel chacune est le plus susceptible de correspondre. Présenté autrement, ce problème consiste à réduire la dimension des données — actuellement de 64 pour chaque image — tout en conservant le plus d'information possible pour réaliser la classi-

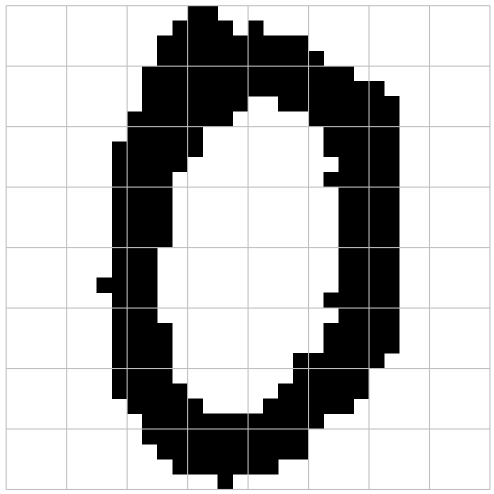


FIG. 8.1 – Représentation matricielle de 32 pixels de haut et 32 pixels de large d'un chiffre écrit à la main. La grille représente le regroupement des pixels en groupes de 4×4.

fication. Cela peut sembler simple avec le caractère de la figure 8.1, mais sauriez-vous identifier facilement le caractère de la figure 8.2 tiré de la même base de données ?

Le contexte du présent problème est celui de l'analyse en composantes principales, une procédure statistique qui, comme nous le verrons, est intimement liée aux notions de valeur et de vecteur propre.

8.2 Définitions

En algèbre vectorielle, le produit par une matrice $\mathbf{A}$ peut être vu comme un opérateur permettant d'appliquer une transformation à un vecteur $\mathbf{x}$. Par exemple :

- La multiplication par un scalaire :

$$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

- La rotation d'un angle θ :

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix} = \begin{bmatrix} w \\ z \end{bmatrix}.$$

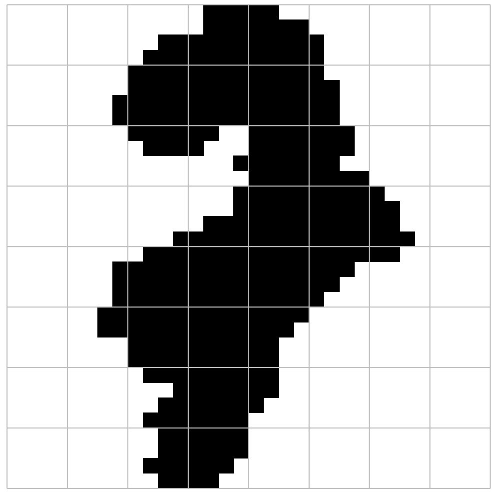
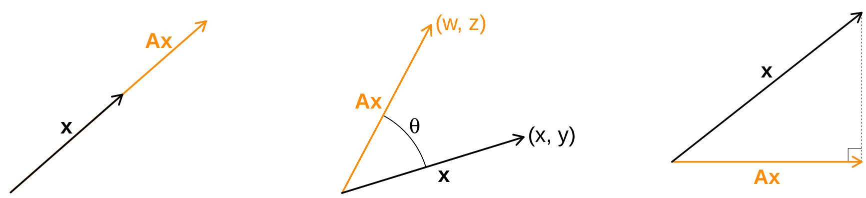


FIG. 8.2 – Représentation matricielle peu lisible d'un chiffre écrit à la main



(a) produit par un scalaire (b) rotation (c) projection

FIG. 8.3 – Illustrations de trois transformations d'un vecteur $\mathbf{x}$ correspondant à un produit matriciel $\mathbf{Ax}$.

- La projection sur un autre vecteur :

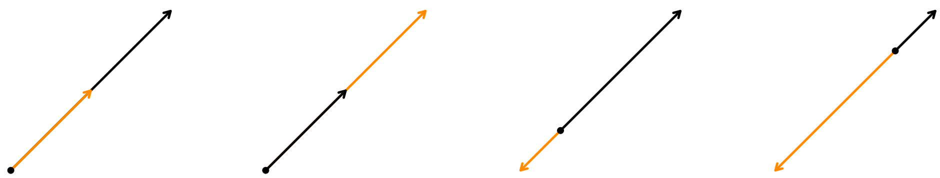
$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Voir la figure 8.3 pour des illustrations des transformations ci-dessus.

Les définitions de valeur propre et de vecteur propre procèdent de cette interprétation du produit matriciel.

Définition 8.1. Soit $\mathbf{A}$ une matrice $n \times n$. Alors le vecteur non nul $\mathbf{x} \in \mathbb{R}^n$ est un *vecteur propre* de $\mathbf{A}$ si

$$\mathbf{Ax} = \lambda\mathbf{x}$$



(a) $0 \leq \lambda \leq 1$ (b) $\lambda \geq 1$ (c) $-1 \leq \lambda \leq 0$ (d) $\lambda \leq -1$

FIG. 8.4 – Illustration de la relation entre une matrice, une valeur propre de celle-ci et le vecteur propre associé. Un vecteur propre $\mathbf{x}$ d'une matrice $\mathbf{A}$ (en noir) est transformé en un multiple λ de lui-même lorsque multiplié par $\mathbf{A}$ (en orange).

pour une valeur de λ appelée *valeur propre* de $\mathbf{A}$. On dit que $\mathbf{x}$ est le vecteur propre *correspondant* à λ.

Ainsi, par définition, lorsqu'un vecteur propre $\mathbf{x}$ est multiplié par la matrice $\mathbf{A}$, il est transformé en un multiple de lui-même ; voir la figure 8.4.

Exemple 8.1. Le vecteur $\mathbf{x} = (1, 2)$ est un vecteur propre de

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

correspondant à la valeur propre $\lambda = 3$ car $\mathbf{Ax} = 3\mathbf{x}$. En effet,

$$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

□

Le qualificatif « propre » doit être pris dans le sens de « qui caractérise ». En anglais, les termes les plus souvent utilisés pour « valeur propre » et « vecteur propre » sont *eigenvalue* et *eigenvector*, dans l'ordre. Ils sont formés à partir du préfixe allemand *eigen* plutôt qu'à partir de son équivalent anglais *own*, probablement pour des raisons générales de viabilité phonétique.

8.3 Indice pour le problème

Nous aurons besoin d'un peu de notation pour résoudre notre problème. Soit $\mathbf{X} = (X_1, \dots, X_p)^\mathsf{T}$ un vecteur $p \times 1$ de variables aléatoires. (La coutume de représenter les variables aléatoires par des lettres majuscules va, ici, l'emporter sur celle suivie jusqu'à présent de représenter les vecteurs par des lettres minuscules.) Chaque *ligne* de la matrice `optd` correspond à un vecteur $\mathbf{X}$. Dans notre exemple, nous avons donc $p = 64$ et 1 797 réalisations de $\mathbf{X}$.

Pour diverses considérations sur lesquelles nous ferons l'impasse ici, et sans perte de généralité, les observations de $\mathbf{X}$ sont généralement normalisées de telle sorte que leur moyenne vaut 0 et leur variance, 1. La matrice de variance-covariance[1] de $\mathbf{X}$ est donc une matrice $\mathrm{Var}[\mathbf{X}] = \boldsymbol{\Sigma}$ avec des 1 sur la diagonale.

Nous avons établi précédemment vouloir réduire la dimension des données afin de faciliter leur classement. Nous considérons qu'une telle réduction est possible parce que les données comportent une certaine redondance, ou corrélation. Nous chercherons donc d'abord à éliminer cette corrélation. Une manière de « décorréler » les éléments d'un vecteur aléatoire consiste à les combiner linéairement. Nous voulons donc créer un nouveau vecteur aléatoire $\mathbf{Y}$ à l'aide de la transformation

$$\mathbf{Y} = \mathbf{W}\mathbf{X},$$

où $\mathbf{W}$ est une matrice $p \times p$ déterminée de telle sorte que

$$\begin{aligned}
\mathrm{Var}[\mathbf{Y}] &= \mathrm{Var}[\mathbf{W}\mathbf{X}] \\
&= \mathbf{W}\,\mathrm{Var}[\mathbf{X}]\mathbf{W}^\mathsf{T} \\
&= \mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^\mathsf{T}
\end{aligned}$$

est une matrice diagonale.

Tel qu'expliqué dans la section précédente, la combinaison linéaire $\mathbf{Y} = \mathbf{W}\mathbf{X}$ correspond à un changement de base ou, de manière équivalente, à une rotation d'axes. Quant au rôle des valeurs propres dans la solution, nous ne le connaissons pas encore.

1. Comme son nom le suggère, une matrice de variance-covariance contient sur la diagonale les variances des éléments de $\mathbf{X}$ et, hors de la diagonale, les covariances des éléments pris deux à deux. Par essence, une matrice de variance est symétrique.

8.4 Calcul des valeurs propres

Pour trouver les valeurs propres d'une matrice $\mathbf{A}$, nous devons résoudre

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

ou, de manière équivalente,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}.$$

pour $\mathbf{x} \neq \mathbf{0}$. Par conséquent, λ est une valeur propre de $\mathbf{A}$ si, et seulement si, il existe une solution autre que la solution triviale à ce système d'équations homogène.

Or, par le théorème 7.13, il existe une solution de $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ autre que la solution triviale si, et seulement si, $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$. Par conséquent :

- les valeurs propres de $\mathbf{A}$ sont les solutions de
$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0;$$
- cette équation est l'*équation caractéristique* de $\mathbf{A}$;
- $\det(\lambda\mathbf{I} - \mathbf{A})$ est le *polynôme caractéristique* (de degré n) de $\mathbf{A}$.

Exemple 8.2. Soit

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}.$$

Le polynôme caractéristique de $\mathbf{A}$ est

$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{vmatrix} \\ &= \lambda[\lambda(\lambda - 8) + 17] - 4 \\ &= \lambda^3 - 8\lambda^2 + 17\lambda - 4. \end{aligned}$$

Les valeurs propres sont les racines de ce polynôme, c'est-à-dire les solutions de

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0.$$

Or,

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = (\lambda - 4)(\lambda^2 - 4\lambda + 1),$$

d'où les valeurs propres sont

$$\begin{aligned}\lambda_1 &= 4\\ \lambda_2 &= 2+\sqrt{3}\\ \lambda_3 &= 2-\sqrt{3}.\end{aligned}$$

□

Exemple 8.3. Soit la matrice triangulaire

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

On a

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ 0 & \lambda - a_{22} & -a_{23} \\ 0 & 0 & \lambda - a_{33} \end{vmatrix}\\ &= (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33}),\end{aligned}$$

d'où $\lambda_1 = a_{11}$, $\lambda_2 = a_{22}$ et $\lambda_3 = a_{33}$. □

Le résultat de l'exemple précédent justifie un théorème.

Théorème 8.1. *Si $\mathbf{A}_{n\times n}$ est une matrice triangulaire (supérieure ou inférieure) ou diagonale, alors ses valeurs propres sont les éléments de la diagonale.*

De ce théorème, nous pouvons établir les faits suivants pour une matrice triangulaire :

1. la trace est égale à la somme des valeurs propres ;
2. le déterminant est égal au produit des valeurs propres.

Le théorème suivant confirme que ces résultats tiennent pour toute matrice carrée.

Théorème 8.2. *Soit $\mathbf{A}_{n\times n}$ une matrice dont les valeurs propres sont $\lambda_1, \dots, \lambda_n$ (égalités et valeurs complexes comprises). Alors*

$$\begin{aligned}\mathrm{tr}(\mathbf{A}) &= \sum_{i=1}^{n} \lambda_i\\ \det(\mathbf{A}) &= \prod_{i=1}^{n} \lambda_i.\end{aligned}$$

Démonstration. On peut démontrer que le polynôme caractéristique d'une matrice carrée **A** est toujours de la forme

$$\begin{aligned}\det(\lambda\mathbf{I} - \mathbf{A}) &= \lambda^n - c_1\lambda^{n-1} + \cdots + c_n \\ &= (\lambda - \lambda_1)\cdots(\lambda - \lambda_n),\end{aligned}$$

où

$$\begin{aligned}c_1 &= \lambda_1 + \cdots + \lambda_n \\ c_n &= (-1)^n\lambda_1\cdots\lambda_n.\end{aligned}$$

(Pensez au cas $n = 2$, soit un polynôme de second degré ; vous savez que le coefficient du terme de degré un est la somme des racines et que le terme constant est le produit des racines.) Il faut maintenant établir que $c_1 = \operatorname{tr}(\mathbf{A})$ et que $c_n = (-1)^n \det(\mathbf{A})$ pour compléter la démonstration. Or, d'une part, il est facile de vérifier que le coefficient de λ^{n-1} dans le polynôme caractéristique provient du produit

$$(\lambda - a_{11})\cdots(\lambda - a_{nn}) = \lambda^n - (a_{11} + \cdots + a_{nn})\lambda^{n-1} + \ldots$$

dans le calcul du déterminant de $\lambda\mathbf{I} - \mathbf{A}$. Par conséquent,

$$\begin{aligned}\operatorname{tr}(\mathbf{A}) &= \sum_{i=1}^{n} a_{ii} \\ &= c_1 \\ &= \sum_{i=1}^{n} \lambda_i.\end{aligned}$$

D'autre part, en posant $\lambda = 0$ dans le polynôme caractéristique, nous obtenons

$$\begin{aligned}\det(-\mathbf{A}) &= (-1)^n \det(\mathbf{A}) \\ &= c_n \\ &= (-1)^n\lambda_1\cdots\lambda_n,\end{aligned}$$

d'où $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$. □

Du théorème précédent, il découle que si au moins une valeur propre d'une matrice est nulle, alors son déterminant est nul et la matrice est singulière. Cette observation nous permet d'ajouter un énoncé au théorème 7.13.

Théorème 8.3. *Soit* $\mathbf{A}$ *une matrice* $n \times n$*. Les énoncés suivants sont équivalents.*

1. $\mathbf{A}$ *est inversible.*
2. $\det(\mathbf{A}) \neq 0$
3. $\text{rang}(\mathbf{A}) = n$
4. *La seule solution de* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *est la solution triviale.*
5. *Le système d'équations* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *a une solution pour tout vecteur* $\mathbf{b}$ *et celle-ci est unique.*
6. $\mathbf{A}$ *peut être exprimée comme un produit de matrices élémentaires.*
7. *Les vecteurs lignes de* $\mathbf{A}$ *sont linéairement indépendants.*
8. *Les vecteurs colonnes de* $\mathbf{A}$ *sont linéairement indépendants.*
9. *Les vecteurs lignes de* $\mathbf{A}$ *forment une base de* $\mathbb{R}^n$*.*
10. *Les vecteurs colonnes de* $\mathbf{A}$ *forment une base de* R^n*.*
11. $\lambda = 0$ *n'est pas une valeur propre de* $\mathbf{A}$*.*

Un dernier résultat : nous observons que si λ est une valeur propre de $\mathbf{A}$ et $\mathbf{x}$ est un vecteur propre correspondant, alors

$$\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x},$$

d'où λ^2 est une valeur propre de $\mathbf{A}^2$ et $\mathbf{x}$ est un vecteur propre correspondant.

Théorème 8.4. *Si* λ *est une valeur propre de la matrice* $\mathbf{A}$ *et* $\mathbf{x}$ *est un vecteur propre correspondant, alors*

- λ^k *est une valeur propre de* $\mathbf{A}^k$*,* $k \in \mathbb{N}$*;*
- $\mathbf{x}$ *est un vecteur propre correspondant.*

Démonstration. Laissée en exercice. □

Vous savez que pour une fonction $y = f(x)$ quelconque, $y = f(x - c)$ correspond à une translation de la fonction de c unités vers la gauche. Par analogie, l'équation caractéristique $\det(\lambda\mathbf{I} - \mathbf{A})$ fournit cette interprétation de la valeur propre : c'est le nombre de fois dont il faut « déplacer » la matrice $\mathbf{A}$ de $\mathbf{I}$ jusqu'à ce qu'elle devienne singulière.

8.5 Calcul des vecteurs propres

Les vecteurs propres de la matrice $\mathbf{A}$ correspondant à la valeur propre λ sont les vecteurs non nuls satisfaisant

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

ou, de manière équivalente,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}.$$

Or, puisque $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$:

- le rang de $\lambda\mathbf{I} - \mathbf{A}$ est inférieur à n ;
- il existe une solution $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ autre que la solution triviale ;
- le système d'équations homogène comporte une infinité de solutions.

Par conséquent, pour chaque valeur propre λ, il faut trouver une base de $\mathbb{R}^k$, où

$$k = n - \mathrm{rang}(\lambda\mathbf{I} - \mathbf{A}).$$

Exemple 8.4. Soit

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

L'équation caractéristique de $\mathbf{A}$ est

$$(\lambda - 2)^2(\lambda - 1) = 0,$$

d'où les valeurs propres sont $\lambda_1 = 1$ et $\lambda_2 = \lambda_3 = 2$.

1. Cas $\lambda_1 = 1$. Nous devons trouver $\mathbf{x} = (x_1, x_2, x_3)$ tel que

$$(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

soit

$$\begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & -1 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

ou, de manière équivalente,

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Par conséquent :

- la solution du système d'équations est $x_1 = -2x_3$, $x_2 = x_3$ et x_3 est une variable libre ;
- $(-2s, s, s)$ est un vecteur propre correspondant à $\lambda = 1$;
- en fait, tout vecteur de la forme $(-2, 1, 1)$ est un vecteur propre correspondant à $\lambda = 1$.

Ainsi, $(-2, 1, 1)$ est une base de vecteurs propres correspondant à $\lambda = 1$.

2. Cas $\lambda_2 = \lambda_3 = 2$. Nous avons

$$(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

soit

$$\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

ou, de manière équivalente,

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Par conséquent :

- la solution est $x_1 = -x_3$, x_2 et x_3 sont des variables libres ;
- $(s, t, -s)$ est un vecteur propre correspondant à $\lambda = 2$;
- on a

$$\begin{bmatrix} s \\ t \\ -s \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Ainsi $(1, 0, -1)$ et $(0, 1, 0)$ forment une base de vecteurs propres correspondant à $\lambda = 2$.

□

Le nombre de vecteurs propres dans la base est *au plus* la multiplicité de λ. Par conséquent, si toutes les valeurs propres d'une matrice sont différentes, alors elle compte n vecteurs propres différents.

8.6 Diagonalisation

La diagonalisation d'une matrice est la principale application des valeurs et vecteurs propres qu'un étudiant en actuariat est susceptible de rencontrer dans ses études. Elle joue un rôle dans diverses procédures statistiques, dont l'analyse en composantes principales, dans la résolution de systèmes d'équations différentielles ordinaires, ainsi que dans le calcul de l'exponentielle d'une matrice, tel qu'expliqué dans l'exemple suivant.

Exemple 8.5. L'exponentielle d'un nombre réel x est définie comme

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!}. \end{aligned}$$

Par analogie, on définit l'exponentielle d'une matrice $\mathbf{A}$ ainsi :

$$\begin{aligned} e^{\mathbf{A}} &= I + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^3}{3!} + \dots \\ &= \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!}, \end{aligned}$$

avec $\mathbf{A}^0 = \mathbf{I}$ par convention.

L'évaluation numérique de l'exponentielle d'une matrice de manière efficace, fiable et précise est chose difficile. Dans un article célèbre, Moler et Van Loan (1978) passent en revue pas moins de 19 manières différentes de le faire, pour conclure… qu'aucune méthode n'est uniformément meilleure que les autres !

Cela dit, le calcul de l'exponentielle d'une matrice demeure simple si celle-ci est diagonale. Pour les fins de la démonstration et sans perte de généralité, nous utiliserons une matrice diagonale $\mathbf{D}$ de dimension 2 :

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

Ainsi

$$\mathbf{D}^n = \begin{bmatrix} d_1^n & 0 \\ 0 & d_2^n \end{bmatrix}$$

et, par conséquent,

$$\begin{aligned}
e^{\mathbf{D}} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} d_1^2 & 0 \\ 0 & d_2^2 \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} d_1^3 & 0 \\ 0 & d_2^3 \end{bmatrix} + \dots \\
&= \begin{bmatrix} 1 + d_1 + \frac{d_1^2}{2!} + \frac{d_1^3}{3!} + \dots & 0 \\ 0 & 1 + d_2 + \frac{d_2^2}{2!} + \frac{d_2^3}{3!} + \dots \end{bmatrix} \\
&= \begin{bmatrix} e^{d_1} & 0 \\ 0 & e^{d_2} \end{bmatrix}.
\end{aligned}$$

Le calcul de l'exponentielle d'une matrice diagonale se résume donc à calculer l'exponentielle des éléments de la diagonale.

Poussons l'idée une étape plus loin. Supposons que $\mathbf{A}$ est une matrice que nous pouvons écrire sous la forme

$$\mathbf{A} = \mathbf{PDP}^{-1},$$

où $\mathbf{D}$ est diagonale. Puisque

$$\begin{aligned}
\mathbf{A}^2 &= (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1}) \\
&= \mathbf{PD}^2\mathbf{P}^{-1}, \\
\mathbf{A}^3 &= (\mathbf{PDP}^{-1})(\mathbf{PD}^2\mathbf{P}^{-1}) \\
&= \mathbf{PD}^3\mathbf{P}^{-1} \\
&\vdots \\
\mathbf{A}^n &= \mathbf{PD}^n\mathbf{P}^{-1},
\end{aligned}$$

alors

$$\begin{aligned}
e^{\mathbf{A}} &= \sum_{n=0}^{\infty} \frac{\mathbf{PD}^n\mathbf{P}^{-1}}{n!} \\
&= \mathbf{P}\left(\sum_{n=0}^{\infty} \frac{\mathbf{D}^n}{n!}\right)\mathbf{P}^{-1} \\
&= \mathbf{P}e^{\mathbf{D}}\mathbf{P}^{-1}.
\end{aligned}$$

Par conséquent, si nous pouvons écrire une matrice $\mathbf{A}$ sous la forme $\mathbf{A} = \mathbf{PDP}^{-1}$ (ou, inversement, si $\mathbf{P}^{-1}\mathbf{AP}$ est diagonale), alors le calcul de son exponentielle se résumera essentiellement au calcul de l'exponentielle de la matrice diagonale $\mathbf{D}$, calcul qui est simple, comme nous l'avons vu. □

Définition 8.2. Une matrice carrée $\mathbf{A}$ est *diagonalisable* s'il existe une matrice inversible $\mathbf{P}$ tel que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

est une matrice diagonale.

Nous établirons qu'une matrice est diagonalisable si elle possède n vecteurs propres linéairement indépendants.

Exemple 8.6. À l'exemple 8.4, nous avons obtenu que les vecteurs propres de la matrice $\mathbf{A}$ sont

$$\mathbf{p}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{p}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad \mathbf{p}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Ces vecteurs propres sont linéairement indépendants. Vérification rapide :

$$\begin{vmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{vmatrix} = 1 \neq 0.$$

□

Supposons qu'une matrice $\mathbf{A}_{3\times 3}$ possède trois vecteurs propres linéairement indépendants $\mathbf{p}_1, \dots, \mathbf{p}_3$ et soit la matrice $\mathbf{P}$ formée de ces trois vecteurs dans les colonnes :

$$\begin{aligned} \mathbf{P} &= [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3] \\ &= \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}. \end{aligned}$$

Alors,

$$\mathbf{AP} = [\mathbf{Ap}_1\ \mathbf{Ap}_2\ \mathbf{Ap}_3]$$

et, par définition du vecteur propre,

$$\begin{aligned} \mathbf{AP} &= [\lambda_1\mathbf{p}_1\ \lambda_2\mathbf{p}_2\ \lambda_3\mathbf{p}_3] \\ &= [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \\ &= \mathbf{PD}, \end{aligned}$$

où $\mathbf{D}$ est une matrice diagonale. Or, puisque les colonnes de $\mathbf{P}$ sont linéairement indépendantes, $\mathbf{P}^{-1}$ existe et, par conséquent,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

d'où $\mathbf{A}$ est diagonalisable.

Nous pourrions faire le cheminement inverse, ce qui justifie le théorème suivant.

Théorème 8.5. *Soit* $\mathbf{A}$ *une matrice* $n \times n$*. Les énoncés suivants sont équivalents.*

1. $\mathbf{A}$ *est diagonalisable.*
2. $\mathbf{A}$ *possède* n *vecteurs propres linéairement indépendants.*

De plus, si $\mathbf{P}$ *est la matrice qui diagonalise* $\mathbf{A}$*, alors*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

La remarque qui suit l'exemple 8.4 établit que si toutes les valeurs propres sont distinctes (multiplicité 1), alors la matrice compte n vecteurs propres. Or, on peut démontrer que si les valeurs propres sont distinctes, alors les vecteurs propres sont linéairement indépendants. Il en découle le théorème suivant.

Théorème 8.6. *Si la matrice* $\mathbf{A}_{n \times n}$ *possède* n *valeurs propres distinctes, alors elle est diagonalisable.*

> La contraposée du théorème précédent indique que si une matrice n'est pas diagonalisable, alors ses valeurs propres ne sont pas distinctes. En revanche, la matrice peut être diagonalisable même si ses valeurs propres ne sont pas distinctes.

Exemple 8.7. Les valeurs propres de la matrice

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

de l'exemple 8.4 sont $\lambda_1 = 1$ et $\lambda_2 = \lambda_3 = 2$. Néanmoins, les trois vecteurs propres forment une base de $\mathbb{R}^3$. Par conséquent,

$$\begin{aligned}\mathbf{P}^{-1}\mathbf{AP} &= \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.\end{aligned}$$

□

Exemple 8.8. Soit

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}.$$

Le polynôme caractéristique de $\mathbf{A}$ est

$$\det(\lambda\mathbf{I} - \mathbf{A}) = (\lambda - 1)(\lambda - 2)^2,$$

d'où les valeurs propres sont $\lambda_1 = 1$ et $\lambda_2 = \lambda_3 = 2$. Il est laissé en exercice de vérifier que

- le vecteur propre correspondant à $\lambda = 1$ est $\mathbf{p}_1 = (\frac{1}{8}, -\frac{1}{8}, 1)$;
- le vecteur propre correspondant à $\lambda = 2$ est $\mathbf{p}_2 = (0, 0, 1)$.

Puisque nous avons seulement deux vecteurs propres linéairement indépendants, la matrice $\mathbf{A}$ n'est pas diagonalisable. □

Exemple 8.9. Soit

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 4 & 0 \\ 0 & 3 & 1 & 7 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

Nous pouvons facilement établir que :

- la matrice $\mathbf{A}$ est triangulaire ;
- ses valeurs propres sont $\lambda_1 = -1$, $\lambda_2 = 3$, $\lambda_4 = 5$ et $\lambda_4 = -2$;
- les valeurs propres sont distinctes ;
- la matrice est diagonalisable.

□

Terminons par un résultat qui joue un rôle important en statistique. Certaines matrices sont diagonalisables à l'aide d'une matrice **P** orthogonale (définition 7.8), c'est-à-dire que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P}$ est diagonale. On dit d'une telle matrice qu'elle est *orthogonalement diagonalisable*. Le théorème suivant établit que toute matrice orthogonalement diagonalisable est symétrique, et vice versa.

Théorème 8.7. *Soit* **A** *une matrice carrée $n \times n$. Les énoncés suivants sont équivalents.*

1. **A** *est orthogonalement diagonalisable.*
2. **A** *possède un ensemble orthonormal de n vecteurs propres.*
3. **A** *est symétrique.*

Exemple 8.10. Soit la matrice symétrique

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

Il est laissé en exercice de vérifier que les valeurs propres de la matrice sont $\lambda_1 = \lambda_2 = 2$ et $\lambda_3 = 8$. Vous pouvez ensuite vérifier que, d'une part,

$$\mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \quad \text{et} \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

forment une base orthonormale correspondant à $\lambda = 2$ et que, d'autre part,

$$\mathbf{v}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

forme une base orthonormale correspondant à $\lambda = 8$. (La conversion d'une base de vecteur propre « standard » en une base orthonormale se fait par la procédure de Gram–Schmidt, qui est hors de la portée du présent ouvrage.)

Enfin, il est simple de vérifier que la matrice

$$\mathbf{P} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix},$$

formée à partir des vecteurs propres ci-dessus, diagonalise orthogonalement la matrice **A**. □

La fonction `eigen` calcule les valeurs et vecteurs propres d'une matrice dans R. Étudiez attentivement les lignes 15-43 du fichier de script `valeurs_propres.R` reproduit à la section 8.10 pour en connaitre davantage.

8.7 Indice pour le problème

Le théorème 8.7 nous permet d'avancer de manière importante dans la résolution de notre problème. En effet, la matrice $\mathbf{\Sigma}$ étant symétrique, nous savons désormais que :

- la matrice diagonale $\mathrm{Var}[\mathbf{Y}] = \mathbf{W}\mathbf{\Sigma}\mathbf{W}^\top$ existe et elle est formée des valeurs propres de $\mathbf{\Sigma}$;
- la matrice $\mathbf{W}$ est formée des vecteurs propres de $\mathbf{\Sigma}$ sur les lignes ;
- les vecteurs propres en question sont orthonormaux.

Supposons que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ sont les valeurs propres de $\mathbf{\Sigma}$ en ordre décroissant et que $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p$ sont les vecteurs propres correspondants. Pour des raisons qui deviendront claires sous peu, nous allons former les matrices $\mathrm{Var}[\mathbf{Y}]$ et $\mathbf{W}$ avec les valeurs propres dans cet ordre :

$$\mathrm{Var}[\mathbf{Y}] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad \text{et} \quad \mathbf{W} = \begin{bmatrix} \mathbf{w}_1^\top \\ \mathbf{w}_2^\top \\ \vdots \\ \mathbf{w}_p^\top \end{bmatrix}.$$

Le nouveau vecteur aléatoire

$$\mathbf{Y} = \mathbf{W}\mathbf{X} = \begin{bmatrix} \mathbf{w}_1^\top \mathbf{X} \\ \mathbf{w}_2^\top \mathbf{X} \\ \vdots \\ \mathbf{w}_p^\top \mathbf{X} \end{bmatrix}$$

est appelé le vecteur des *composantes principales* de $\mathbf{X}$.

8.8 Formes quadratiques

Cette section offre une courte introduction aux formes quadratiques afin de présenter un théorème important pour l'analyse en composantes principales et, par conséquent, pour résoudre le problème de ce chapitre. Nous n'irons pas plus loin que le cadre de ce théorème, mais sachez que l'étude des formes quadratiques se révèle en soi fort intéressante. Par exemple, les propriétés des formes quadratiques permettent de justifier les conditions — que vous connaissez — selon lesquelles une fonction dans $\mathbb{R}^3$ atteint un maximum, un minimum ou un point de selle en un point donné.

Une forme quadratique est une fonction de n variables $x_1, \ldots, x_n$ de la forme

$$\begin{aligned}\sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i}^{n} a_{ij}x_ix_j &= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 \\ &\quad + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{n-1,n}x_{n-1}x_n.\end{aligned}$$

Par exemple, une forme quadratique de deux variables x_1 et x_2 est

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2,$$

alors qu'une forme quadratique de trois variables x_1, x_2 et x_3 est

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3.$$

Nous pouvons écrire les deux expressions ci-dessus sous forme de produits matriciels ainsi, dans l'ordre :

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

et

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Vous remarquerez que ces deux produits sont de la forme $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$, où $\mathbf{x}$ est le vecteur colonne des variables et $\mathbf{A}$ est la matrice symétrique des coefficients. En fait, il s'agit là de la définition même d'une forme quadratique exprimée de manière matricielle : une fonction de la forme $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$.

Certains auteurs n'incluent pas le facteur 2 devant les termes croisés dans la définition d'une forme quadratique, mais divisent alors par 2 les coefficients a_{ij}, $i \neq j$, dans la matrice $\mathbf{A}$.

Exemple 8.11. Voici quelques formes quadratiques et leurs représentations matricielles respectives.

$$2x^2 + 6xy - 7y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$4x^2 - 5y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$xy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$x^2 + 7y^2 - 3z^2 + 4xy - 2xz + 6yz = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 3 \\ -1 & 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

□

Le théorème suivant contient deux résultats importants liant les formes quadratiques aux valeurs et vecteurs propres.

Théorème 8.8. *Soit* $\mathbf{A}$ *une matrice symétrique* $n \times n$ *dont les valeurs propres sont, en ordre décroissant,* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$*. Si le vecteur* $\mathbf{x}$ *est tel que* $\|\mathbf{x}\| = 1$ *selon le produit scalaire dans* $\mathbb{R}^n$*, alors :*

a) $\lambda_1 \geq \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq \lambda_n$ *;*

b) $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \lambda_n$ *si* $\mathbf{x}$ *est le vecteur propre de* $\mathbf{A}$ *correspondant à* λ_n *et* $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \lambda_1$ *si* $\mathbf{x}$ *est le vecteur propre de* $\mathbf{A}$ *correspondant à* λ_1*.*

Démonstration. Débutons par la partie a). Du théorème 8.7, la matrice symétrique $\mathbf{A}$ compte n vecteurs propres qui forment une base orthonormale de $\mathbb{R}^n$. Supposons que l'ensemble $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ constitue cette base où, par définition, $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ et $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ quand $i \neq j$. Le vecteur $\mathbf{x}$ peut donc s'exprimer comme une combinaison linéaire des vecteurs propres :

$$\mathbf{x} = k_1\mathbf{v}_1 + \cdots + k_n\mathbf{v}_n.$$

De plus, les vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_n$ étant des vecteurs propres de $\mathbf{A}$, nous avons que

$$\begin{aligned} \mathbf{A}\mathbf{x} &= k_1\mathbf{A}\mathbf{v}_1 + \cdots + k_n\mathbf{A}\mathbf{v}_n \\ &= \lambda_1 k_1\mathbf{v}_1 + \cdots + \lambda_n k_n\mathbf{v}_n. \end{aligned}$$

Or, $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = k_1^2 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + \cdots + k_n^2 \langle \mathbf{v}_n, \mathbf{v}_n \rangle = k_1^2 + \cdots + k_n^2 = 1$ et, par conséquent,

$$\begin{aligned}
\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} &= \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle \\
&= \lambda_1 \langle k_1\mathbf{v}_1, k_1\mathbf{v}_1 \rangle + \cdots + \lambda_n \langle k_n\mathbf{v}_n, k_n\mathbf{v}_n \rangle \\
&= \lambda_1 k_1^2 + \cdots + \lambda_n k_n^2 \\
&\leq \lambda_1 k_1^2 + \cdots + \lambda_1 k_n^2 \\
&= \lambda_1 (k_1^2 + \cdots + k_n^2) \\
&= \lambda_1.
\end{aligned}$$

La démonstration que $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq \lambda_n$ est similaire.

Pour la partie b), si $\mathbf{x}$ est le vecteur propre de $\mathbf{A}$ correspondant à λ_1 et que $\|\mathbf{x}\| = 1$, alors

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \lambda_1 \mathbf{x}^\mathsf{T}\mathbf{x} = \lambda_1.$$

De la même manière, $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \lambda_n$ si $\mathbf{x}$ est le vecteur propre de $\mathbf{A}$ correspondant à λ_n. □

Le théorème 8.8 permet d'établir que, sous la contrainte

$$\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} = 1,$$

la valeur maximale de la forme quadratique $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ est λ_1 (la plus grande valeur propre de $\mathbf{A}$) et la valeur minimale est λ_n (la plus petite valeur propre).

8.9 Solution du problème

Par le théorème 8.8, nous savons maintenant que la composante principale $Y_1 = \mathbf{w}_1^\mathsf{T}\mathbf{X}$ est la combinaison linéaire de $\mathbf{X}$ avec la plus grande variance, $Y_2 = \mathbf{w}_2^\mathsf{T}\mathbf{X}$ celle avec la seconde plus grande variance, et ainsi de suite. Les premières composantes principales s'avèrent donc les plus intéressantes puisque ce sont celles qui conservent le plus d'information.

Pour réduire la dimension de nos données, il suffit de choisir les $m \ll p$ premières composantes principales. En fait, la k^e composante principale explique une proportion

$$\frac{\lambda_k}{\lambda_1 + \cdots + \lambda_p}$$

de la variabilité totale qui se trouve dans le vecteur d'origine $\mathbf{X}$ et l'ensemble des m premières composantes principales en explique une proportion

$$\frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i}.$$

La figure 8.5 présente les 1 797 points de la base de données *Optical Recognition of Handwritten Digits* en tenant compte des deux premières composantes principales seulement. Vous constaterez qu'il serait déjà possible d'établir un système de classification relativement fiable. Par exemple, l'immense majorité des points dans l'ensemble $\{(Y_1, Y_2); -8 \leq Y_1 \leq 10, -30 \leq Y_2 \leq -16\}$ représentent le chiffre 0 et ce chiffre est peu présent en dehors de cet ensemble. C'est pourquoi il était facile de reconnaitre le chiffre 0 dans la figure 8.1. Par contre, vous remarquerez que le chiffre 7 s'avère beaucoup plus difficile à distinguer des autres avec seulement deux composantes principales. Eh oui, c'était le chiffre représenté à la figure 8.2 !

Vous pouvez recréer le graphique de la figure 8.5 avec le code des lignes 46-67 du fichier de script `valeurs_propres.R`.

8.10 Code informatique

Fichier d'accompagnement `valeurs_propres.R`

```
14 ###
15 ### VALEURS ET VECTEURS PROPRES
16 ###
17 
18 ## Nous allons illustrer le calcul des valeurs et vecteurs
19 ## propres dans R avec une matrice utilisée dans les exemples
20 ## du chapitre.
21 (A <- matrix(c(0, 1, 1, 0, 2, 0, -2, 1, 3), nrow = 3))
22 
23 ## La fonction 'eigen' calcule les valeurs propres et vecteurs
24 ## propres d'une matrice.
25 (e <- eigen(A))
26 
27 ## Les vecteurs propres sont normalisés de sorte que leur
28 ## norme (longueur) soit toujours égale à 1. Pour vérifier les
29 ## résultats calculés algébriquement, comparer simplement les
30 ## valeurs relatives des coordonnées des vecteurs.
31 e$values[c(1, 2)]          # deux premières valeurs propres...
32 e$vectors[, c(1, 2)]       # ... et vecteurs correspondants
33 e$vectors[, 2]             # équivalent à (1, 0, -1)
34 e$vectors[, 2] * sqrt(2)   # avec norme de 2 plutôt que 1
35 
36 e$values[3]                # troisième valeur propre
```

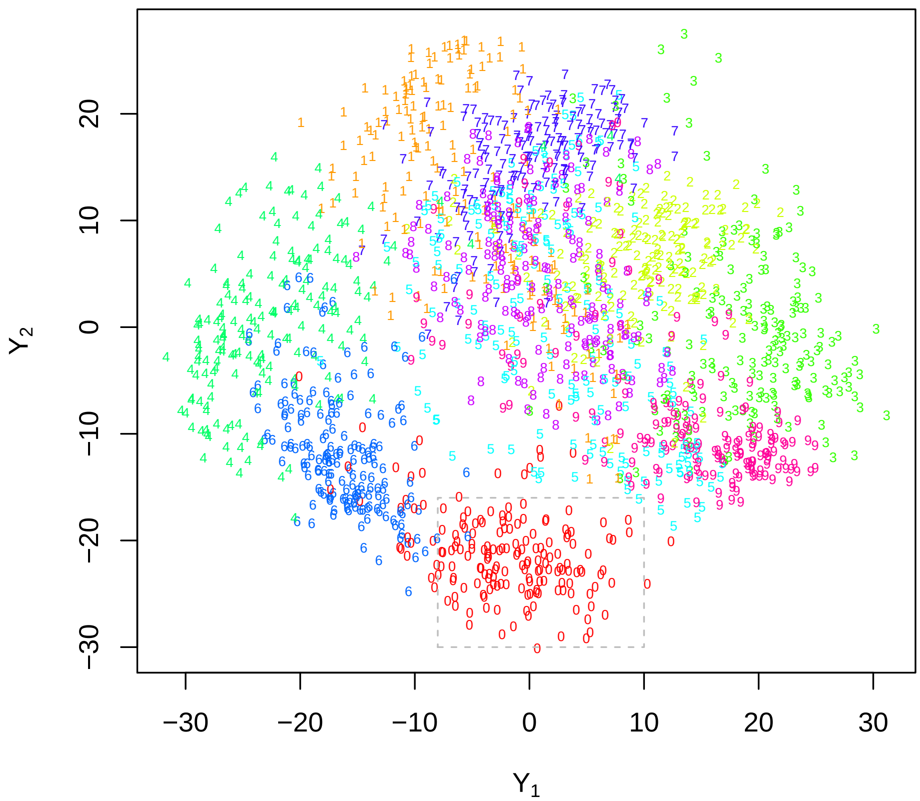


FIG. 8.5 – Représentation des deux premières composantes principales des données *Optical Recognition of Handwritten Digits*. Le rectangle pointillé représente l'ensemble $\{(Y_1, Y_2); -8 \leq Y_1 \leq 10, -30 \leq Y_2 \leq -16\}$

```
e$vectors[, 3]              # vecteur équivalent à (-2, 1, 1)
e$vectors[, 3] * sqrt(6)    # avec norme de 6 plutôt que 1

## Si la matrice est symétrique, les vecteurs propres sont
## orthonormaux.
(A <- matrix(c(4, 2, 2, 2, 4, 2, 2, 2, 4), nrow = 3))
crossprod(eigen(A)$vectors) # matrice identité

###
### ANALYSE EN COMPOSANTES PRINCIPALES
###

## Installation du paquetage RSKC pour obtenir les données
## "Optical Recognition of Handwritten Digits". Exécuter une
## seule fois!
#install.packages("RSKC")

## Chargement du jeu de données 'optd' sans charger tout le
## paquetage dans la session de travail.
data(optd, package = "RSKC")

## Création du graphique des deux composantes principales.
optd.acp <- prcomp(optd, center = TRUE, scale. = FALSE)
optd.coord <- predict(optd.acp, newdata = optd)
par(mar = c(5.1, 4.1, 2.1, 2.1))
plot(x = optd.coord[, 1], y = optd.coord[, 2], type = "n",
     xlab = expression(Y[1]), ylab = expression(Y[2]))
text(x = optd.coord[, 1], y = optd.coord[, 2],
     labels = rownames(optd), cex = 0.5,
     col = rainbow(10)[as.numeric(rownames(optd)) + 1])
rect(-8, -30, 10, -16, lty = 2, border = "gray")
```

8.11 Exercices

8.1 Trouver l'équation caractéristique, les valeurs propres et les bases de vecteurs propres des matrices suivantes. Vérifier les réponses obtenues à l'aide de la fonction `eigen` de R.

a) $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$

c) $\begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ d) $\begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}$

e) $\begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

8.2 Démontrer par induction que si λ est une valeur propre de la matrice $\mathbf{A}$, alors λ^k est une valeur propre de $\mathbf{A}^k$.

8.3 Trouver les valeurs et vecteurs propres de $\mathbf{A}^{25}$ si

$$\mathbf{A} = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}.$$

8.4 Soit

$$(x - x_1)(x - x_2) \cdots (x - x_n) = x^n + c_1 x^{n-1} + \cdots + c_n.$$

Démontrer par induction les identités suivantes.

a) $c_1 = -\sum_{i=1}^{n} x_i$

b) $c_n = (-1)^n \prod_{i=1}^{n} x_i$

8.5 Démontrer que l'équation caractéristique d'une matrice $\mathbf{A}_{2\times 2}$ est

$$\lambda^2 - \operatorname{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0.$$

8.6 Démontrer que si λ est une valeur propre de la matrice inversible $\mathbf{A}$ et que $\mathbf{x}$ est un vecteur propre correspondant, alors λ^{-1} est une valeur propre de $\mathbf{A}^{-1}$ et $\mathbf{x}$ est un vecteur propre correspondant.

8.7 Trouver les valeurs propres et les bases de vecteurs propres de $\mathbf{A}^{-1}$, où

$$\mathbf{A} = \begin{bmatrix} -2 & 2 & 3 \\ -2 & 3 & 2 \\ -4 & 2 & 5 \end{bmatrix}$$

8.8 Démontrer que tout vecteur est un vecteur propre de la matrice identité correspondant à la valeur propre $\lambda = 1$.

8.9 Pour chacune des matrices $\mathbf{A}$ ci-dessous :

i) trouver les valeurs propres de la matrice ;

ii) trouver le rang de la matrice $\lambda\mathbf{I} - \mathbf{A}$ pour chaque valeur propre λ ;

iii) déterminer si la matrice est diagonalisable ;

iv) si la matrice est diagonalisable, trouver la matrice $\mathbf{P}$ qui diagonalise $\mathbf{A}$ et $\mathbf{P}^{-1}\mathbf{AP}$;

v) vérifier les réponses en iv) avec la fonction `eigen` de R.

a) $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$

c) $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

d) $\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$

e) $\begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 5 & -5 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

Réponses

8.1 a) $\lambda = 3$ avec base $(\frac{1}{2}, 1)$, $\lambda = -1$ avec base $(0, 1)$

b) $\lambda = 4$ avec base $(\frac{3}{2}, 1)$

c) $\lambda = 1$ avec base $(0, 1, 0)$, $\lambda = 2$ avec base $(-\frac{1}{2}, 1, 1)$, $\lambda = 3$ avec base $(-1, 1, 1)$

d) $\lambda = -4$ avec base $(-2, \frac{8}{3}, 1)$, $\lambda = 3$ avec base $(5, -2, 1)$

e) $\lambda = 1$ avec base $(0, 0, 0, 1)$ et $(2, 3, 1, 0)$, $\lambda = -2$ avec base $(-1, 0, 1, 0)$, $\lambda = -1$ avec base $(-2, 1, 1, 0)$

8.3 $\lambda = 1$ avec base $(-1, 1, 0)$ et $(-1, 0, 1)$, $\lambda = -1$ avec base $(2, -1, 1)$

8.7 $\lambda = 1$ avec base $(1, 0, 1)$, $\lambda = \frac{1}{2}$ avec base $(\frac{1}{2}, 1, 0)$, $\lambda = \frac{1}{3}$ avec base $(1, 1, 1)$

8.9 a) pas diagonalisable

b) $\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$, $\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

c) pas diagonalisable

d) $\mathbf{P} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}$, $\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

e) $\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

9 Méthodes de résolution de systèmes d'équations linéaires

Objectifs du chapitre

- Comparer la rapidité des méthodes usuelles de résolutions des systèmes d'équations linéaires.
- Calculer la solution d'un système d'équations linéaires par la décomposition LU.

L'utilisation des matrices de loin la plus fréquente en actuariat consiste à résoudre des systèmes d'équations linéaires du type

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

Nous avons jusqu'à maintenant étudié diverses façons — toutes équivalentes d'un point de vue mathématique — d'obtenir la solution

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Nous pouvons maintenant nous demander laquelle est la plus efficace d'un point de vue informatique, surtout lorsque le système compte un grand nombre d'équations.

9.1 Comparaison du nombre d'opérations

Les principales méthodes de résolution d'un système d'équations linéaires sont :

1. l'élimination gaussienne avec substitution successive ;

Méthode	Nombre d'opérations
Élimination gaussienne	$n^3/3$
Élimination de Gauss–Jordan	$n^3/3$
Transformation de $[\mathbf{A}\vert\mathbf{I}]$ en $[\mathbf{I}\vert\mathbf{A}^{-1}]$	n^3
Calcul de $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$	n^3

TAB. 9.1 – Nombre approximatif de multiplications et divisions pour résoudre le système d'équations à n équations et n inconnues $\mathbf{Ax} = \mathbf{b}$

2. l'élimination de Gauss–Jordan ;
3. le calcul de $\mathbf{A}^{-1}$, puis de $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Le calcul de $\mathbf{A}^{-1}$ peut quant à lui s'effectuer par transformation de $[\mathbf{A}|\mathbf{I}]$ en $[\mathbf{I}|\mathbf{A}^{-1}]$, ou par la méthode des cofacteurs.

L'élément décisif dans la comparaison des temps de calcul de ces diverses méthodes de calcul est le nombre d'opérations arithmétiques requis. Le tableau 9.1 présente l'ordre de grandeur (et non le nombre exact) du nombre de multiplications et de divisions nécessaires pour obtenir une réponse avec chacune des méthodes ci-dessus. Concentrons-nous sur les multiplications et divisions, sachant que ces opérations coûtent plus cher en temps de calcul que les additions et les soustractions. Nous constatons que l'élimination gaussienne avec substitution successive et l'élimination de Gauss–Jordan sont les méthodes les plus rapides, leur avantage augmentant rapidement avec la taille du système d'équations.

9.2 Décomposition LU

Le principal inconvénient des méthodes d'élimination réside dans le fait qu'il faut connaître le vecteur des coefficients $\mathbf{b}$ au moment d'effectuer les calculs. Si l'on souhaite résoudre le système $\mathbf{Ax} = \mathbf{b}$ pour un nouveau vecteur $\mathbf{b}$, il faut répéter la procédure depuis le début.

Les principales routines de résolution de systèmes d'équations linéaires disponibles dans les divers outils informatiques (R, Maple, Matlab, Mathematica, etc.) reposent donc plutôt sur la technique de la décomposition *LU*, une variante de l'élimination gaussienne ne nécessitant pas de connaître d'avance le vecteur $\mathbf{b}$.

L'idée est très simple : si la matrice $\mathbf{A}$ peut être factorisée en un produit de matrices $n \times n$

$$\mathbf{A} = \mathbf{LU},$$

où $\mathbf{L}$ est une matrice triangulaire inférieure et $\mathbf{U}$ est une matrice triangulaire supérieure, alors nous obtenons le système d'équations linéaires

$$\mathbf{LUx} = \mathbf{b}. \tag{9.1}$$

Or, en posant

$$\mathbf{Ux} = \mathbf{y} \tag{9.2}$$

il est possible de réécrire (9.1) comme

$$\mathbf{Ly} = \mathbf{b}. \tag{9.3}$$

La solution $\mathbf{y}$ de ce dernier système d'équations est simple à obtenir par substitutions successives puisque $\mathbf{L}$ est triangulaire. De même, une fois $\mathbf{y}$ connu, le vecteur $\mathbf{x}$ est obtenu en résolvant (9.2), toujours par simples substitutions.

Exemple 9.1. Soit le système d'équations linéaires

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

On peut démontrer que

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix},$$

soit $\mathbf{A} = \mathbf{LU}$ avec

$$\mathbf{L} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix}$$

et

$$\mathbf{U} = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Nous avons donc $\mathbf{Ax} = \mathbf{LUx} = \mathbf{b}$. En posant $\mathbf{Ux} = \mathbf{y}$, le système d'équations peut se réécrire sous la forme $\mathbf{Ly} = \mathbf{b}$, soit

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Par substitution successive,

$$\begin{aligned} y_1 &= 1 \\ y_2 &= 2 + 3y_1 = 5 \\ y_3 &= \frac{3 - 4y_1 + 3y_2}{7} = 2. \end{aligned}$$

Pour trouver la solution du système d'équations original, il suffit maintenant de résoudre $\mathbf{Ux} = \mathbf{y}$, soit

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}.$$

Nous obtenons alors

$$\begin{aligned} x_3 &= 2 \\ x_2 &= 5 - 3x_3 = -1 \\ x_1 &= 1 - 3x_2 - x_3 = 2. \end{aligned}$$

□

L'essentiel des calculs dans la décomposition *LU* se trouve dans la factorisation de la matrice **A** en un produit de matrices triangulaires. Pour justifier la technique, supposons que l'on réduit la matrice **A** sous forme échelonnée par une série d'opérations élémentaires sur les lignes. Cela signifie que l'on peut trouver des matrices élémentaires $\mathbf{E}_1, \dots, \mathbf{E}_k$ tel que

$$\mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{A} = \mathbf{U}.$$

Par le théorème 7.7, l'inverse d'une matrice élémentaire existe et est également une matrice élémentaire, d'où

$$\mathbf{A} = \mathbf{E}_1^{-1} \cdots \mathbf{E}_k^{-1} \mathbf{U}$$

et donc

$$\mathbf{L} = \mathbf{E}_1^{-1} \cdots \mathbf{E}_k^{-1}.$$

Cette dernière matrice est triangulaire inférieure *à condition de ne pas échanger des lignes* lors de la réduction de **A** vers **U**. Il existe un algorithme simple pour construire la matrice **L** sans devoir effectuer le produit des matrices élémentaires inverses ; consultez Anton (2000, section 9.9).

Exemple 9.2. Reprenons la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix}$$

de l'exemple 9.1. Il est laissé en exercice de vérifier que les opérations élémentaires suivantes permettent de réduire la matrice $\mathbf{A}$ sous la forme

$$\mathbf{U} = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

1. Multiplier la ligne 1 par $\frac{1}{2}$.
2. Additionner 3 fois la ligne 1 à la ligne 2 et -4 fois la ligne 1 à la ligne 3.
3. Additionner 3 fois la ligne 2 à la ligne 3.
4. Multiplier la ligne 3 par $\frac{1}{7}$.

Les matrices élémentaires correspondant à ces opérations sont, dans l'ordre :

$$\mathbf{E}_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$$

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \qquad \mathbf{E}_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}.$$

Vous pouvez vérifier que $\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{U}$. Nous avons donc $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{E}_4^{-1}\mathbf{U} = \mathbf{LU}$ avec $\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{E}_4^{-1}$. Or, les matrices élémentaires inverses sont :

$$\mathbf{E}_1^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

$$\mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} \qquad \mathbf{E}_4^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix},$$

d'où

$$\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{E}_4^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix}.$$

□

Le nombre d'opérations de la décomposition LU est du même ordre que les méthodes d'élimination. Par contre, on remarquera que la factorisation est tout à fait indépendante du vecteur $\mathbf{b}$. Une fois la factorisation connue, on peut donc résoudre plusieurs systèmes d'équations différents utilisant tous la même matrice de coefficients $\mathbf{A}$ sans devoir répéter une grande partie des calculs.

Plusieurs progiciels mathématiques recourent à LAPACK[1] pour les calculs d'algèbre linéaire numérique. Cette bibliothèque est elle-même basée sur LINPACK, dont les origines remontent aux années 1970. La bibliothèque est écrite en Fortran, un autre très ancien langage de programmation, toujours utilisé pour le calcul scientifique. D'ailleurs, au moins une particularité de ce langage demeure visible en R. En effet, c'est du Fortran que vient cette habitude de remplir les matrices par colonne.

9.3 Exercices

9.1 Résoudre le système d'équations

$$\begin{aligned} 3x_1 - 6x_2 - 3x_3 &= -3 \\ 2x_1 \phantom{{}- 6x_2} + 6x_3 &= -22 \\ -4x_1 + 7x_2 + 4x_3 &= 3 \end{aligned}$$

par la décomposition LU sachant que

$$\begin{bmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

9.2 Soit

$$\begin{aligned} \mathbf{E}_1 &= \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \\ \mathbf{E}_2 &= \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \end{aligned}$$

1. *Linear Algebra PACKage*; https://www.netlib.org/lapack.

et

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Résoudre le système d'équations

$$\mathbf{Ax} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

par la décomposition LU.

Réponses

9.1 $x_1 = -2$, $x_2 = 1$, $x_3 = -3$

9.2 $\mathbf{x} = (2, 1)$

A Solutions des exercices

Chapitre 7

7.1 a) On a une solution unique. Le système d'équations correspondant est

$$\begin{aligned} x_1 - 3x_2 + 4x_3 &= 7 \\ x_2 + 2x_3 &= 2 \\ x_3 &= 5. \end{aligned}$$

Par substitution successive, on obtient $x_3 = 5$, $x_2 = 2 - 2(5) = -8$ et $x_1 = 7 + (3)(-8) - 4(5) = -37$.

b) Le système est sous-déterminé puisqu'il compte plus d'inconnues que d'équations. Il y a donc une infinité de solutions. Le système d'équations correspondant est

$$\begin{aligned} x_1 + 8x_3 - 5x_4 &= 6 \\ x_2 + 4x_3 - 9x_4 &= 3 \\ x_3 + x_4 &= 2. \end{aligned}$$

Posons la variable libre $x_4 = t$. Ainsi, la solution générale du système d'équations est $x_3 = 2 - t$, $x_2 = 3 - 4(2 - t) + 9t = -5 + 13t$ et $x_1 = 6 - 8(2 - t) + 5t = -10 + 13t$.

c) Nous avions à l'origine un système d'équations à quatre équations et cinq inconnues. La matrice échelonnée contient une ligne complète de zéros, ce qui indique que la quatrième équation était une combinaison linéaire des trois autres. Ne reste donc qu'un système sous-déterminé à trois équations et cinq inconnues :

$$\begin{aligned} x_1 + x_2 - 2x_3 - 8x_5 &= -3 \\ x_3 + x_4 + 6x_5 &= 5 \\ x_4 + 3x_5 &= 9. \end{aligned}$$

Ce système a une infinité de solutions et il faudra, pour les exprimer, poser deux variables libres. Les candidates sont les variables correspondant aux colonnes sans un 1 sur la diagonale. Posons $x_2 = s$ et $x_5 = t$. Nous obtenons alors $x_4 = 9 - 3t$, $x_3 = 5 - (9 - 3t) - 6t = -4 - 3t$ et $x_1 = -3 - 7s + 2(-4 - 3t) + 8t = -11 - 7s + 2t$.

d) La dernière ligne de la matrice échelonnée correspond à l'équation impossible à satisfaire

$$0x_1 + 0x_2 + 0x_3 = 1,$$

d'où le système n'a pas de solution.

7.2 a) Solution pour l'élimination gaussienne seulement. La matrice augmentée est

$$\begin{bmatrix} 1 & 1 & 2 & 8 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & 10 \end{bmatrix}.$$

Les opérations élémentaires à effectuer sur cette matrice pour l'exprimer sous forme échelonnée sont, dans l'ordre :

1. additionner la première ligne à la seconde ;
2. additionner -3 fois la première ligne à la troisième ;
3. additionner -3 fois la première ligne à la troisième ;
4. multiplier la deuxième ligne par -1 ;
5. additionner 10 fois la deuxième ligne à la troisième ;
6. diviser la troisième ligne par -52.

On obtient alors la matrice échelonnée

$$\begin{bmatrix} 1 & 1 & 2 & 8 \\ 0 & 1 & -5 & -9 \\ 0 & 0 & 1 & 2 \end{bmatrix},$$

d'où $x_3 = 2$, $x_2 = -9 + 5(2) = 1$ et $x_1 = 8 - 1 - 2(2) = 3$.

b) Solution pour l'élimination de Gauss-Jordan seulement. La matrice augmentée est

$$\begin{bmatrix} 2 & 2 & 2 & 0 \\ -2 & 5 & 2 & 1 \\ 8 & 1 & 4 & -1 \end{bmatrix}.$$

Les opérations élémentaires à effectuer sur cette matrice pour l'exprimer sous forme échelonnée réduite sont les suivantes :

1. multiplier la première ligne par $\frac{1}{2}$;
2. additionner 2 fois la première ligne à la deuxième ;
3. additionner -8 fois la première ligne à la troisième ;
4. additionner la deuxième ligne à la troisième ;
5. multiplier la deuxième ligne par $\frac{1}{7}$;
6. additionner -1 fois la deuxième ligne à la première.

On obtient alors la matrice échelonnée réduite

$$\begin{bmatrix} 1 & 0 & 3/7 & -1/7 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

En posant $x_3 = t$, on a la solution générale $x_2 = \frac{1}{7} - \frac{4}{7}t$ et $x_1 = -\frac{1}{7} - \frac{3}{7}t$.

c) Solution pour l'élimination gaussienne seulement. La matrice augmentée est

$$\begin{bmatrix} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{bmatrix}.$$

Les opérations élémentaires à effectuer sur cette matrice pour l'exprimer sous forme échelonnée sont, dans l'ordre :

1. additionner -2 fois la première ligne à la seconde ;
2. additionner la première ligne à la troisième ;
3. additionner -3 fois la première ligne à la quatrième ;
4. additionner -1 fois la deuxième ligne à la quatrième ;
5. multiplier la deuxième ligne par $\frac{1}{3}$;
6. additionner -1 fois la deuxième ligne à la troisième.

On obtient alors la matrice échelonnée

$$\begin{bmatrix} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

d'où $z = s$, $w = t$, $y = 2s$ et $x = -1 + 2s - 2s + t = -1 + t$.

d) En échangeant les première et troisième équations, on a la matrice augmentée

$$\begin{bmatrix} 6 & 6 & 3 & 5 \\ 3 & 6 & -3 & -2 \\ 0 & -2 & 3 & 1 \end{bmatrix}.$$

Après les opérations élémentaires

1. multiplier la première ligne par $\frac{1}{6}$;
2. additionner -3 fois la première ligne à la deuxième ;
3. multiplier la deuxième ligne par $\frac{1}{3}$;
4. additionner 2 fois la deuxième ligne à la troisième,

on obtient la matrice échelonnée

$$\begin{bmatrix} 1 & 1 & 1/2 & 5/6 \\ 0 & 1 & -3/2 & -3/2 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

La troisième équation étant impossible à satisfaire, le système n'a pas de solution.

7.3 La procédure de réduction des matrices augmentées sous forme échelonnée est similaire à celle utilisée dans les solutions de l'exercice 7.2. On ne donne, ici que les résultats de cette procédure.

a) La matrice échelonnée est

$$\begin{bmatrix} 1 & -2/5 & 6/5 & 0 \\ 0 & 1 & 27 & 5 \end{bmatrix}$$

et la matrice échelonnée réduite est

$$\begin{bmatrix} 1 & 0 & 12 & 2 \\ 0 & 1 & 27 & 5 \end{bmatrix}.$$

Par conséquent, la solution générale est $x_3 = t$, $x_2 = 5 - 27t$ et $x_1 = 2 - 12t$.

b) La matrice échelonnée réduite est

$$\begin{bmatrix} 1 & -2 & 1 & -4 & 1 \\ 0 & 1 & 6/5 & 6/5 & 1/5 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix}.$$

Ce système n'a donc pas de solution.

c) La matrice échelonnée est

$$\begin{bmatrix} 1 & 2 & 1/2 & 7/2 & 0 & 7/2 \\ 0 & 0 & 1 & 2 & -1 & 4 \\ 0 & 0 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

et la matrice échelonnée réduite est

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 3 & -6 \\ 0 & 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Par conséquent, la solution générale est $v = s$, $y = t$, $x = 3 + t$, $w = -2 - t$ et $u = -6 - 3t - 2s$.

7.4 a) Après quelques opérations élémentaires sur les lignes de la matrice augmentée correspondant à ce système d'équations, on obtient

$$\begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 6 & 0 \end{bmatrix},$$

d'où $x_3 = x_2 = x_1 = 0$, soit la solution triviale.

b) La somme des deux équations donne une nouvelle équation $4x_1 + x_3 = 0$, d'où $x_3 = -4x_1$. De la première équation, on a également $x_2 = -3x_1 - x_3 - x_4 = x_1 - x_4$. En posant $x_1 = -s$ et $x_4 = t$, on a la solution générale $x_3 = 4s$ et $x_2 = -s - t$. Il y a évidemment de multiples autres façons d'exprimer la solution générale en utilisant des variables libres différentes.

c) La matrice échelonnée de ce système d'équations est

$$\begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 1 & 7/3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

On a donc $z = 0$ et, en posant $y = t$, $x = -t$, $w = t$.

7.5 Après des opérations élémentaires sur la matrice augmentée du système d'équations, on obtient

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & a^2 - 16 & a - 4 \end{bmatrix}.$$

On constate alors que le système d'équations n'a pas de solution si $a^2 - 16 = 0$ et $a - 4 \neq 0$, donc lorsque $a = -4$. En revanche, si $a^2 - 16 = 0$ et $a - 4 = 0$, soit lorsque $a = 4$, le système a une infinité de solutions. Finalement, si $a^2 - 16 \neq 0 \Leftrightarrow |a| \neq 4$, le système a une solution unique.

7.6 La somme (ou la différence) entre deux matrices est définie si les dimensions des matrices sont identiques. Les dimensions du résultat seront les mêmes. Quant au produit, il est défini si le nombre de colonnes de la première matrice est égal au nombre de lignes de la seconde ; le résultat est une matrice dont le nombre de lignes est égal à celui de la première matrice et le nombre de colonnes à celui de la seconde matrice.

a) Le produit $\mathbf{BA}$ n'est pas défini.

b) Le résultat de $\mathbf{AC}$ est une matrice 4×2, donc l'opération $\mathbf{AC} + \mathbf{D}$ est définie et le résultat est une matrice 4×2.

c) Le résultat de $\mathbf{AE}$ est une matrice 4×4, donc l'opération $\mathbf{AE} + \mathbf{B}$ n'est pas définie.

d) Le produit $\mathbf{AB}$ n'est pas défini.

e) La somme $\mathbf{A} + \mathbf{B}$ est définie, tout comme le produit $\mathbf{E}(\mathbf{A} + \mathbf{B})$. Le résultat est une matrice 5×5.

f) Le résultat du produit $\mathbf{AC}$ est une matrice 4×2, donc $\mathbf{E}(\mathbf{AC})$ est une matrice 5×2.

g) Puisque $\mathbf{E}^{\mathsf{T}}$ est une matrice 4×5, le produit $\mathbf{E}^{\mathsf{T}}\mathbf{A}$ n'est pas défini.

h) Le résultat de $\mathbf{A}^{\mathsf{T}} + \mathbf{E}$ est une matrice 5×4, donc $(\mathbf{A}^{\mathsf{T}} + \mathbf{E})\mathbf{D}$ est une matrice 5×2.

7.7 a) On a

$$\mathbf{D} + \mathbf{E} = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix} + \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 7 & 6 & 5 \\ -2 & 1 & 3 \\ 7 & 3 & 7 \end{bmatrix}$$

b) On a

$$\mathbf{AB} = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}.$$

c) On a

$$
\begin{aligned}
(2\mathbf{D}^{\mathsf{T}} - \mathbf{E})\mathbf{A} &= \left(2\begin{bmatrix} 1 & -1 & 3 \\ 5 & 0 & 2 \\ 2 & 1 & 4 \end{bmatrix} - \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}\right)\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \left(\begin{bmatrix} 2 & -2 & 6 \\ 10 & 0 & 4 \\ 4 & 2 & 8 \end{bmatrix} - \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}\right)\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -4 & -3 & 3 \\ 11 & -1 & 2 \\ 0 & 1 & 5 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -6 & -3 \\ 36 & 0 \\ 4 & 7 \end{bmatrix}.
\end{aligned}
$$

d) On a $(4\mathbf{B})\mathbf{C} + 2\mathbf{B} = 2\mathbf{B}(2\mathbf{C} + \mathbf{I})$. Or la somme entre parenthèses n'est pas définie puisque la matrice $\mathbf{C}$ n'est pas carrée.

e) On a

$$
\begin{aligned}
(-\mathbf{A}\mathbf{C})^{\mathsf{T}} + 5\mathbf{D}^{\mathsf{T}} &= \left(-\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}\right)^{\mathsf{T}} + 5\begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}^{\mathsf{T}} \\
&= \left(-\begin{bmatrix} 3 & 12 & 6 \\ 5 & -2 & 8 \\ 4 & 5 & 7 \end{bmatrix}\right)^{\mathsf{T}} + \begin{bmatrix} 5 & -5 & 15 \\ 25 & 0 & 10 \\ 10 & 5 & 20 \end{bmatrix} \\
&= \begin{bmatrix} -3 & -5 & -4 \\ -12 & 2 & -5 \\ -6 & -8 & -7 \end{bmatrix} + \begin{bmatrix} 5 & -5 & 15 \\ 25 & 0 & 10 \\ 10 & 5 & 20 \end{bmatrix} \\
&= \begin{bmatrix} 2 & -10 & 11 \\ 13 & 2 & 5 \\ 4 & -3 & 13 \end{bmatrix}.
\end{aligned}
$$

f) On a

$$
\begin{aligned}
(\mathbf{BA}^\mathsf{T} - 2\mathbf{C})^\mathsf{T} &= \left(\begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 3 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 8 & 4 \\ 6 & 2 & 10 \end{bmatrix}\right)^\mathsf{T} \\
&= \left(\begin{bmatrix} 12 & -6 & 3 \\ 0 & 4 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 8 & 4 \\ 6 & 2 & 10 \end{bmatrix}\right)^\mathsf{T} \\
&= \begin{bmatrix} 10 & -6 \\ -14 & 2 \\ -1 & -8 \end{bmatrix}
\end{aligned}
$$

g) On a

$$
\mathbf{CC}^\mathsf{T} = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 4 & 1 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 21 & 17 \\ 17 & 35 \end{bmatrix}
$$

et

$$
\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & -1 \\ -1 & 5 \end{bmatrix},
$$

d'où

$$
\mathbf{B}^\mathsf{T}(\mathbf{CC}^\mathsf{T} - \mathbf{A}^\mathsf{T}\mathbf{A}) = \begin{bmatrix} 4 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 10 & 18 \\ 18 & 30 \end{bmatrix} = \begin{bmatrix} 40 & 72 \\ 26 & 42 \end{bmatrix}
$$

h) On a $\mathbf{D}^\mathsf{T}\mathbf{E}^\mathsf{T} - (\mathbf{ED})^\mathsf{T} = (\mathbf{ED})^\mathsf{T} - (\mathbf{ED})^\mathsf{T} = \mathbf{0}_{3\times 3}$.

i) Pour calculer la trace, il suffit de connaître les éléments de la diagonale. Or l'élément d_{ii} de $\mathbf{DD}^\mathsf{T}$ est la somme des carrés des éléments de la ligne i de la matrice $\mathbf{D}$. Par conséquent, $\operatorname{tr}(\mathbf{DD}^\mathsf{T}) = (1+25+4) + (1+0+1) + (9+4+16) = 61$.

j) Encore une fois, on se concentre seulement sur les éléments de la diagonale de $4\mathbf{E}^\mathsf{T} - \mathbf{D}$. La transposée ne joue donc aucun rôle, ici. Ainsi, $\operatorname{tr}(4\mathbf{E}^\mathsf{T} - \mathbf{D}) = (24-1) + (4-0) + (12-4) = 35$.

7.8 Puisque $\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$, alors

$$
\begin{aligned}
\operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}) &= \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} \\
&= \sum_{i=1}^n (a_{ii} + b_{ii}) \\
&= \operatorname{tr}(\mathbf{A} + \mathbf{B}).
\end{aligned}
$$

7.9 a) On a que $(\mathbf{A}\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} = \mathbf{A}\mathbf{A}^{\mathsf{T}}$, d'où $\mathbf{A}\mathbf{A}^{\mathsf{T}}$ est symétrique. On procède de même pour le second résultat.

b) On a que $(\mathbf{A} - \mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} - \mathbf{A} = -(\mathbf{A} - \mathbf{A}^{\mathsf{T}})$, d'où $\mathbf{A} - \mathbf{A}^{\mathsf{T}}$ est antisymétrique.

7.10 Tout d'abord, $\mathbf{A}^{\mathsf{T}} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}}\mathbf{A} = \mathbf{A}$, d'où $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ est symétrique. De plus, si $\mathbf{A}$ est symétrique, alors $\mathbf{A} = \mathbf{A}^{\mathsf{T}}\mathbf{A} = \mathbf{A}\mathbf{A} = \mathbf{A}^2$. La matrice $\mathbf{A}$ est donc idempotente.

7.11 Les résultats a) et b) découlent directement des propriétés de l'arithmétique matricielle et de celles des dérivées. Quant au résultat c), en supposant que

$$\mathbf{A} = [a_{ij}(x)]_{m\times n} \qquad\qquad \mathbf{B} = [b_{ij}(x)]_{n\times m},$$

nous avons

$$\begin{aligned}
\frac{d}{dx}(\mathbf{A}\mathbf{B}) &= \frac{d}{dx}\left[\sum_{k=1}^{n} a_{ik}(x)b_{kj}(x)\right]_{m\times m} \\
&= \left[\sum_{k=1}^{n}(a'_{ik}(x)b_{kj}(x) + a_{ik}(x)b'_{kj}(x))\right] \\
&= \left[\sum_{k=1}^{n} a'_{ik}(x)b_{kj}(x)\right] + \left[\sum_{k=1}^{n} a_{ik}(x)b'_{kj}(x)\right] \\
&= \frac{d\mathbf{A}}{dx}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dx}.
\end{aligned}$$

7.12 a) Deux vecteurs de $\mathbb{R}^3$ sont nécessairement linéairement indépendants à moins que l'un ne soit un multiple de l'autre. Ce n'est pas le cas ici.

b) Trois vecteurs sont linéairement dépendants si l'un est une combinaison linéaire des deux autres. Si c'est le cas, le déterminant de la matrice formée des coordonnées des vecteurs sera nul. Or, ici,

$$\begin{vmatrix} -3 & 5 & 1 \\ 0 & -1 & 1 \\ 4 & 2 & 3 \end{vmatrix} = 39 \neq 0.$$

Les vecteurs sont donc linéairement indépendants.

c) Le second vecteur n'est pas un multiple du premier, donc les vecteurs sont linéairement indépendants.

d) Un ensemble de quatre vecteurs de $\mathbb{R}^3$ est nécessairement linéairement dépendant.

7.13 Les trois vecteurs se trouvent dans un plan si seulement deux vecteurs sont linéairement indépendants (ceux-ci engendrent un plan dont fait alors partie le troisième).

a) Soit la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 2 \\ -2 & 1 & 0 \\ 0 & 4 & -4 \end{bmatrix}.$$

En additionnant la première et la deuxième ligne, on obtient

$$\begin{bmatrix} 2 & 6 & 2 \\ 0 & 7 & 2 \\ 0 & 4 & -4 \end{bmatrix}.$$

La troisième ligne n'étant pas un multiple de la deuxième, on constate que la seule solution du système d'équations homogène $\mathbf{Ax} = \mathbf{0}$ est la solution triviale. Par conséquent, les trois vecteurs sont linéairement indépendants et ne se trouvent pas dans un plan. On pourrait également vérifier que $\det(\mathbf{A}) \neq 0$.

b) Soit la matrice

$$\mathbf{A} = \begin{bmatrix} -6 & 3 & 4 \\ 7 & 2 & -1 \\ 2 & 4 & 2 \end{bmatrix}.$$

À la suite d'opérations élémentaires sur les lignes, on obtient

$$\begin{bmatrix} -6 & 3 & 4 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Il n'y a donc que deux vecteurs linéairement indépendants formant un plan dans cet ensemble.

7.14 On peut interpréter la condition $\prod_{i=1}^{n} a_{ii} \neq 0$ comme le fait que le déterminant de la matrice est non nul, ou comme le fait qu'il n'y aucun zéro sur la diagonale et, par conséquent, aucune ligne de zéros dans la matrice. Dans un cas comme dans l'autre, cela signifie que l'inverse existe. En utilisant la technique consistant à transformer la matrice $[\mathbf{A}|\mathbf{I}]$ en

$[\mathbf{I}|\mathbf{A}^{-1}]$, on voit immédiatement que

$$\mathbf{A} = \begin{bmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^{-1} \end{bmatrix}.$$

7.15 Une matrice élémentaire est le résultat d'une opération élémentaire sur la matrice identité.

a) On souhaite inverser les première et troisième ligne de la matrice $\mathbf{A}$. Par conséquent,

$$\mathbf{E}_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

b) Même chose que ci-dessus.

c) La matrice $\mathbf{C}$ est obtenue à partir de la matrice $\mathbf{A}$ en additionnant -2 fois la première ligne à la troisième ligne. Par conséquent,

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

d) De façon similaire, on obtient la matrice $\mathbf{A}$ en additionnant 2 fois la première ligne de la matrice $\mathbf{C}$ à la troisième ligne. On a donc

$$\mathbf{E}_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

7.16 Par des opérations élémentaire sur les lignes, on cherche à transformer la matrice $[\mathbf{A}|\mathbf{I}]$ en $[\mathbf{I}|\mathbf{A}^{-1}]$. Ci-dessous, L_i identifie la ligne i d'une matrice, le symbole $\leftarrow$ signifie qu'une ligne est remplacée et le symbole $\leftrightarrow$ que deux lignes sont échangées.

a) Matrice de départ

$$\left[\begin{array}{rrr|rrr} 3 & 4 & -1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 5 & -4 & 0 & 0 & 1 \end{array}\right]$$

$L_1 \leftrightarrow L_2$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 3 & 4 & -1 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 0 & 1 \end{array}\right]$$

$L_2 \leftarrow -3L_1 + L_2$
$L_3 \leftarrow -2L_1 + L_3$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 4 & -10 & 1 & -3 & 0 \\ 0 & 5 & -10 & 0 & -2 & 1 \end{array}\right]$$

$L_3 \leftarrow -L_2 + L_3$
$L_2 \leftrightarrow L_3$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 4 & -10 & 1 & -3 & 0 \end{array}\right]$$

$L_3 \leftarrow -4L_2 + L_3$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & -10 & 5 & -7 & -4 \end{array}\right]$$

$L_3 \leftarrow -L_3/10$
$L_1 \leftarrow -3L_3 + L_1$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{array}\right]$$

b) Matrice de départ
$$\left[\begin{array}{rrr|rrr} -1 & 3 & -4 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 & 1 & 0 \\ -4 & 2 & -9 & 0 & 0 & 1 \end{array}\right]$$

$L_2 \leftarrow 2L_1 + L_2$
$L_3 \leftarrow -4L_1 + L_3$
$$\left[\begin{array}{rrr|rrr} -1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 10 & -7 & 2 & 1 & 0 \\ 0 & -10 & 7 & -4 & 0 & 1 \end{array}\right]$$

$L_3 \leftarrow L_2 + L_3$
$$\left[\begin{array}{rrr|rrr} -1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 10 & -7 & 2 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{array}\right]$$

L'inverse n'existe pas puisque l'on obtient une ligne de zéros du côté gauche de la matrice augmentée.

c) Matrice de départ
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$L_3 \leftarrow -L_1 + L_3$
$$\left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array}\right]$$

$L_3 \leftarrow -L_2 + L_3$
$L_3 \leftarrow -L_3/2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$

$L_2 \leftarrow -L_3 + L_2$
$L_1 \leftarrow -L_3 + L_1$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$

d) Matrice de départ

$$\left[\begin{array}{ccc|ccc} 2 & 6 & 6 & 1 & 0 & 0 \\ 2 & 7 & 6 & 0 & 1 & 0 \\ 2 & 7 & 7 & 0 & 0 & 1 \end{array}\right]$$

$L_2 \leftarrow -L_1 + L_2$
$L_3 \leftarrow -L_1 + L_3$

$$\left[\begin{array}{ccc|ccc} 2 & 6 & 6 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right]$$

$L_1 \leftarrow L_1/2$
$L_3 \leftarrow -L_2 + L_3$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 3 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

$L_1 \leftarrow -3L_3 + L_1$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & \frac{1}{2} & 3 & -3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

$L_1 \leftarrow -3L_2 + L_1$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & \frac{7}{2} & 0 & -3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

e) Matrice de départ

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$L_2 \leftrightarrow L_3$
$L_3 \leftarrow L_1 + L_3$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & 1 & 0 \end{array}\right]$$

$L_3 \leftarrow -L_2 + L_3$
$L_3 \leftarrow L_3/2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$

$L_1 \leftarrow -L_3 + L_1$ $$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$

7.17 a) Tel que démontré à l'exercice 7.14, l'inverse est

$$\begin{bmatrix} k_1^{-1} & 0 & 0 & 0 \\ 0 & k_2^{-1} & 0 & 0 \\ 0 & 0 & k_3^{-1} & 0 \\ 0 & 0 & 0 & k_4^{-1} \end{bmatrix}.$$

b) Pour obtenir l'inverse de cette matrice par la méthode utilisée à l'exercice 7.16, il suffit d'échanger les lignes de bas en haut, puis de diviser celles-ci par k_4, k_3, k_2 et k_1, respectivement. On trouve alors que l'inverse est

$$\begin{bmatrix} 0 & 0 & 0 & k_4^{-1} \\ 0 & 0 & k_3^{-1} & 0 \\ 0 & k_2^{-1} & 0 & 0 \\ k_1^{-1} & 0 & 0 & 0 \end{bmatrix}.$$

c) On utilise la méthode de l'exercice 7.16.

Matrice de départ $$\left[\begin{array}{cccc|cccc} k & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & k & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$$

$L_1 \leftarrow L_1/k$
$L_2 \leftarrow -L_1 + L_2$
$L_2 \leftarrow L_2/k$
$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & k^{-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -k^{-2} & k^{-1} & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$$

$L_3 \leftarrow -L_2 + L_3$
$L_3 \leftarrow L_3/k$
$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & k^{-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -k^{-2} & k^{-1} & 0 & 0 \\ 0 & 0 & 1 & 0 & k^{-3} & -k^{-2} & k^{-1} & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$$

$L_4 \leftarrow -L_3 + L_4$
$L_4 \leftarrow L_4/k$

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & k^{-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -k^{-2} & k^{-1} & 0 & 0 \\ 0 & 0 & 1 & 0 & k^{-3} & -k^{-2} & k^{-1} & 0 \\ 0 & 0 & 0 & 1 & -k^{-4} & k^{-3} & -k^{-2} & k^{-1} \end{array}\right]$$

7.18 a) Les opérations élémentaires à effectuer pour transformer la matrice $\mathbf{A}$ en matrice identité sont, dans l'ordre :

1. additionner 5 fois la première ligne à la deuxième ;
2. multiplier la deuxième ligne par $\frac{1}{2}$.

Les matrices élémentaires correspondantes sont

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

b) De la partie a), on a que

$$\mathbf{A}^{-1} = \mathbf{E}_2\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ \frac{5}{2} & \frac{1}{2} \end{bmatrix}.$$

c) Du théorème 7.7 des notes de cours, une matrice élémentaire est inversible et son inverse est aussi une matrice élémentaire. Par conséquent, on peut isoler $\mathbf{A}$ dans l'équation $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$ pour obtenir $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}$. Or,

$$\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

d'où

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

7.19 On a le système d'équations linéaires $\mathbf{Ax} = \mathbf{b}$ avec

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}.$$

Par la même technique que les exercices précédents, on trouve

$$\mathbf{A}^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 1 & 3 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{bmatrix}.$$

Par conséquent,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{3}\begin{bmatrix} 16 \\ -4 \\ -11 \end{bmatrix}.$$

7.20 a) On a $\mathbf{Ax} = \mathbf{x} \Leftrightarrow \mathbf{x} - \mathbf{A}x = \mathbf{0} \Leftrightarrow (\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$. On veut par la suite résoudre ce système d'équations homogène, où

$$\mathbf{I} - \mathbf{A} = \begin{bmatrix} -1 & -1 & -2 \\ -2 & -1 & 2 \\ -3 & -1 & 0 \end{bmatrix}.$$

Or, on peut vérifier que le déterminant de cette matrice est non nul, ce qui signifie que le système d'équations n'admet que la solution triviale $x_1 = x_2 = x_3 = 0$. On verra au chapitre 8 que cela signifie que 1 n'est pas une valeur propre de la matrice $\mathbf{A}$.

b) On procède comme ci-dessus en résolvant cette fois le système d'équations linéaires homogène $(4\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$. Or,

$$4\mathbf{I} - \mathbf{A} = \begin{bmatrix} 2 & -1 & -2 \\ -2 & 2 & 2 \\ -3 & -1 & 3 \end{bmatrix}.$$

Après quelques opérations élémentaires sur les lignes de cette matrice, on obtient

$$\begin{bmatrix} 2 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Le système d'équations a donc une infinité de solutions. En posant $x_3 = t$, on a $x_2 = 0$ et $x_1 = t$. Autrement dit, tout vecteur de la forme

$$\begin{bmatrix} t & 0 & t \end{bmatrix}$$

est une solution de l'équation $\mathbf{Ax} = 4\mathbf{x}$. Encore une fois, on verra au chapitre 8 que 4 est une valeur propre de la matrice $\mathbf{A}$ et que $(t, 0, t)$ est un vecteur propre correspondant.

7.21 Tel que vu au théorème 7.13 des notes de cours, les concepts d'inversibilité et de solution d'un système d'équations linéaires homogène sont liés : une matrice carrée $\mathbf{A}$ est inversible si et seulement si $\mathbf{Ax} = \mathbf{0}$ n'admet que la solution triviale.

a) Les quatre équations sont linéairement indépendantes, donc la matrice des coefficients est inversible et la seule solution du système est la solution triviale.

b) Puisque que le coefficient de x_2 est 0 dans la seconde équation, les équations sont linéairement dépendantes, d'où la matrice des coefficients est singulière et il existe une infinité de solution pour le système d'équations homogène.

7.22 On voit que la troisième ligne de la matrice est la somme des deux premières. Les lignes n'étant pas linéairement indépendantes, le déterminant est nul.

7.23 a) Les lignes (et les colonnes) sont linéairement indépendantes, donc la matrice est inversible.

b) La troisième colonne est un multiple de la première. Par conséquent, les colonnes ne sont pas linéairement indépendantes et la matrice est singulière.

c) Le déterminant est clairement nul, donc la matrice est singulière.

d) Idem.

7.24 On utilise les propriétés du déterminant énoncées dans le théorème 7.11 des notes de cours.

a) $\det(3\mathbf{A}) = 3^3 \det(\mathbf{A}) = -189$

b) $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A}) = -\frac{1}{7}$

c) $\det(2\mathbf{A}^{-1}) = 2^3/\det(\mathbf{A}) = -\frac{8}{7}$

d) $\det((2\mathbf{A})^{-1}) = 2^{-3}/\det(\mathbf{A}) = -\frac{1}{56}$

7.25 Une matrice n'est pas inversible si son déterminant est nul.

a) Le déterminant de la matrice est $k^2 - 5k + 2$, donc la matrice est singulière lorsque $k = (5 \pm \sqrt{17})/2$.

b) Le déterminant de la matrice est $8k+8$, donc la matrice est singulière lorsque $k = -1$.

7.26 a) On a

$$M_{13} = \begin{vmatrix} 0 & 0 & 3 \\ 4 & 1 & 14 \\ 4 & 1 & 2 \end{vmatrix} = 0,$$

donc $C_{13} = 0$.

b) On a

$$M_{23} = \begin{vmatrix} 4 & -1 & 6 \\ 4 & 1 & 14 \\ 4 & 1 & 2 \end{vmatrix} = 4(-12) - 4(-8) + 4(-20) = -96$$

et $C_{23} = (-1)^5(-96) = 96$.

c) On a

$$M_{22} = \begin{vmatrix} 4 & 1 & 6 \\ 4 & 0 & 14 \\ 4 & 3 & 2 \end{vmatrix} = 4(-42) - 4(-16) + 4(14) = -48$$

et $C_{22} = (-1)^4(-48) = -48$.

d) On a

$$M_{21} = \begin{vmatrix} -1 & 1 & 6 \\ 1 & 0 & 14 \\ 1 & 3 & 2 \end{vmatrix} = -(-12) - 3(-20) = 72$$

et $C_{21} = (-1)^3(72) = -72$.

e) En développant par la deuxième ligne, on a $\det(\mathbf{A}) = -3C_{23} + 3C_{24}$. Or, de ce qui précède, $C_{23} = 96$ et $C_{24} = M_{24} = 24$. Par conséquent, $\det(\mathbf{A}) = -216$.

Chapitre 8

8.1 Dans tous les cas, la matrice mentionnée dans l'énoncé est notée $\mathbf{A}$.

a) On a

$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} \\ &= (\lambda - 3)(\lambda + 1). \end{aligned}$$

Les valeurs propres sont donc $\lambda_1 = 3$ et $\lambda_2 = -1$. D'une part, la forme échelonnée du système d'équations $(3\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ est

$$\begin{bmatrix} 0 & 0 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

soit $x_1 = s/2$ et $x_2 = s$. Une base de vecteurs propres correspondant à $\lambda = 3$ est donc $(\frac{1}{2}, 1)$. D'autre part, le système d'équations $(-\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

soit $x_1 = 0$ et $x_2 = s$. Une base de vecteurs propres correspondant à $\lambda = -1$ est donc $(0, 1)$. Vérification :

```
> m <- matrix(c(3, 8, 0, -1), nrow = 2)
> eigen(m)
eigen() decomposition
$values
[1]  3 -1

$vectors
          [,1] [,2]
[1,] 0.4472136    0
[2,] 0.8944272    1
```

Remarquez que les vecteurs propres obtenus avec `eigen` sont normalisés de sorte que leur norme soit toujours égale à 1.

b) On a

$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda - 10 & 9 \\ -4 & \lambda + 2 \end{vmatrix} \\ &= (\lambda - 10)(\lambda + 2) + 36 \\ &= (\lambda - 4)^2. \end{aligned}$$

La matrice compte donc une seule valeur propre : $\lambda = 4$. Le système d'équations $(4\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & -\frac{3}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

soit $x_1 = 3s/2$ et $x_2 = s$. Une base de vecteurs propres correspondant à $\lambda = 4$ est donc $(\frac{3}{2}, 1)$. Vérification :

```
> m <- matrix(c(10, 4, -9, -2), nrow = 2)
> eigen(m)
```

```
eigen() decomposition
$values
[1] 4 4

$vectors
           [,1]      [,2]
[1,] -0.8320503 0.8320503
[2,] -0.5547002 0.5547002
```

c) On a

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) &= \begin{vmatrix} \lambda-4 & 0 & -1 \\ 2 & \lambda-1 & 0 \\ 2 & 0 & \lambda-1 \end{vmatrix} \\ &= (\lambda-4)(\lambda-1)^2+2(\lambda-1) \\ &= (\lambda-1)(\lambda-2)(\lambda-3).\end{aligned}$$

Les valeurs propres sont donc $\lambda_1 = 1$, $\lambda_2 = 2$ et $\lambda_3 = 3$. En premier lieu, le système d'équations $(\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = x_3 = 0$ et $x_2 = s$. Une base de vecteurs propres correspondant à $\lambda = 1$ est donc $(0,1,0)$. Deuxièmement, le système d'équations $(2\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = -s/2$, $x_2 = s$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = 2$ est donc $(-\frac{1}{2},1,1)$. Finalement, le système d'équations $(3\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = -s$, $x_2 = s$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = 2$ est donc $(-1,1,1)$. Vérification :

```
> m <- matrix(c(4, -2, -2, 0, 1, 0, 1, 0, 1),
+              nrow = 3)
> eigen(m)
eigen() decomposition
$values
[1] 3 2 1

$vectors
           [,1]       [,2] [,3]
[1,]  0.5773503 -0.3333333    0
[2,] -0.5773503  0.6666667    1
[3,] -0.5773503  0.6666667    0
```

d) On a

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) &= \begin{vmatrix}\lambda-5 & -6 & -2\\ 0 & \lambda+1 & 8\\ -1 & 0 & \lambda+2\end{vmatrix}\\ &= (\lambda-5)(\lambda+1)(\lambda+2)-2(\lambda+1)+48\\ &= \lambda^3-2\lambda^2-15\lambda+36\\ &= (\lambda-3)^2(\lambda+4).\end{aligned}$$

Les valeurs propres sont donc $\lambda_1 = 3$ et $\lambda_2 = \lambda_3 = -4$. En premier lieu, le système d'équations $(3\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix}1 & 0 & -5\\ 0 & 1 & 2\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

soit $x_1 = 5s$, $x_2 = -2s$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = 3$ est donc $(5, -2, 1)$. Deuxièmement, le système d'équations $(-4\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix}1 & 0 & 2\\ 0 & 1 & -\frac{8}{3}\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

soit $x_1 = -2s$, $x_2 = 8s/3$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = -4$ est donc $(-2, \frac{8}{3}, 1)$. Vérification :

```
> m <- matrix(c(5, 0, 1, 6, -1, 0, 2, -8, -2),
+              nrow = 3)
> eigen(m)
eigen() decomposition
$values
[1] -4  3  3

$vectors
           [,1]       [,2]       [,3]
[1,]  0.5746958 -0.9128709  0.9128709
[2,] -0.7662610  0.3651484 -0.3651484
[3,] -0.2873479 -0.1825742  0.1825742
```

e) On a

$$\begin{aligned}
\det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda & 0 & -2 & 0 \\ -1 & \lambda & -1 & 0 \\ 0 & -1 & \lambda + 2 & 0 \\ 0 & 0 & 0 & \lambda - 1 \end{vmatrix} \\
&= (\lambda - 1) \begin{vmatrix} \lambda & 0 & -2 \\ -1 & \lambda & -1 \\ 0 & -1 & \lambda + 2 \end{vmatrix} \\
&= (\lambda - 1)(\lambda^3 + 2\lambda^2 - \lambda - 2) \\
&= (\lambda - 1)^2(\lambda + 1)(\lambda + 2).
\end{aligned}$$

Les valeurs propres sont donc $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -1$ et $\lambda_4 = -2$. En premier lieu, le système d'équations $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = 2s$, $x_2 = 3s$, $x_3 = s$ et $x_4 = t$. Or, $(2s, 3s, s, t) = s(2, 3, 1, 0) + t(0, 0, 0, 1)$. Une base de vecteurs propres correspondant à $\lambda = 1$ est donc composée des vecteurs $(2, 3, 1, 0)$ et $(0, 0, 0, 1)$. Deuxièmement, le système d'équations $(-\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = -2s$, $x_2 = s$, $x_3 = s$ et $x_4 = 0$. Une base de vecteurs propres correspondant à $\lambda = -1$ est donc $(-2, 1, 1, 0)$. Finalement, le système d'équations $(-2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = -s$, $x_2 = 0$, $x_3 = s$ et $x_4 = 0$. Une base de vecteurs propres correspondant à $\lambda = -2$ est donc $(-1, 0, 1, 0)$. Vérification :

```
> m <- matrix(c(0, 1, 0, 0, 0, 0, 1, 0,
+               2, 1, -2, 0, 0, 0, 0, 1),
+             nrow = 4)
> eigen(m)
eigen() decomposition
$values
[1] -2 -1  1  1

$vectors
              [,1]       [,2] [,3]       [,4]
[1,] -7.071068e-01  0.8164966    0 -0.5345225
[2,] -2.055734e-16 -0.4082483    0 -0.8017837
[3,]  7.071068e-01 -0.4082483    0 -0.2672612
[4,]  0.000000e+00  0.0000000    1  0.0000000
```

8.2 Premièrement, l'énoncé est clairement vrai pour $k = 1$. Nous supposons par la suite qu'il est vrai pour $k = n$, soit que si $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, alors $\mathbf{A}^n\mathbf{x} = \lambda^n\mathbf{x}$. Ainsi,

$$\mathbf{A}^{n+1}\mathbf{x} = \mathbf{A}(\mathbf{A}^n\mathbf{x}) = \mathbf{A}(\lambda^n\mathbf{x}) = \lambda^n(\mathbf{A}\mathbf{x}) = \lambda^n(\lambda\mathbf{x}) = \lambda^{n+1}\mathbf{x},$$

d'où l'énoncé est vrai pour $k = n + 1$. Ceci complète la preuve.

8.3 Nous allons utiliser les résultats de l'exercice 8.2. Tout d'abord,

$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda + 1 & 2 & 2 \\ -1 & \lambda - 2 & -1 \\ 1 & 1 & \lambda \end{vmatrix} \\ &= (\lambda + 1)(\lambda - 1)^2, \end{aligned}$$

d'où les valeurs propres de $\mathbf{A}$ sont $\lambda_1 = \lambda_2 = 1$ et $\lambda_3 = -1$. Par conséquent, les valeurs propres de $\mathbf{A}^{25}$ sont $\lambda_1 = \lambda_2 = 1^{25} = 1$ et $\lambda_3 = (-1)^{25} = -1$. Toujours par le résultat de l'exercice 8.2, les vecteurs propres de $\mathbf{A}$ correspondant à $\lambda = 1$ et $\lambda = -1$ sont également des vecteurs propres de $\mathbf{A}^{25}$. Or, le système d'équations $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = -s - t$, $x_2 = s$ et $x_3 = t$. Puisque $(-s-t, s, t) = s(-1, 1, 0) + t(-1, 0, 1)$, une base de vecteurs propres correspondant à $\lambda = 1$ est composée de $(-1, 1, 0)$ et $(-1, 0, 1)$. D'autre part, le système d'équations $(-\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = 2s$, $x_2 = -s$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = -1$ est donc $(2, -1, 1)$.

8.4 Nous pouvons démontrer les deux résultats simultanément. Tout d'abord, les résultats sont clairement vrais pour $n = 1$, c'est-à-dire $c_1 = c_n = -x_1$. Supposons ensuite que les résultats sont vrais pour $n = k$, soit

$$\prod_{i=1}^{k}(x - x_i) = x^k + \left(-\sum_{i=1}^{k} x_i\right) x^{k-1} + \cdots + \prod_{i=1}^{k} x_i.$$

Si $n = k + 1$, alors

$$\begin{aligned}
\prod_{i=1}^{k+1}(x - x_i) &= \left[\prod_{i=1}^{k}(x - x_i)\right](x - x_{k+1}) \\
&= \left[x^k + \left(-\sum_{i=1}^{k} x_i\right) x^{k-1} + \cdots + \prod_{i=1}^{k} x_i\right](x - x_{k+1}) \\
&= x^{k+1} + \left(-\sum_{i=1}^{k} x_i - x_{k+1}\right) x^k + \cdots + x_{k+1}\prod_{i=1}^{k} x_i \\
&= x^{k+1} + \left(-\sum_{i=1}^{k+1} x_i\right) x^k + \cdots + \prod_{i=1}^{k+1} x_i.
\end{aligned}$$

Les résultats sont donc vrais pour $n = k + 1$. Ceci complète la preuve.

8.5 Soit
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$
Le polynôme caractéristique de cette matrice est
$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{vmatrix} \\ &= (\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) \\ &= \lambda^2 - \mathrm{tr}(\mathbf{A})\lambda + \det(\mathbf{A}), \end{aligned}$$
d'où l'équation caractéristique est $\lambda^2 - \mathrm{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0$.

8.6 Par définition $\mathbf{Ax} = \lambda\mathbf{x}$. En multipliant de part et d'autre (par la gauche) par $\mathbf{A}^{-1}$, on obtient $\mathbf{x} = \lambda\mathbf{A}^{-1}\mathbf{x}$, soit $\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$. Par conséquent, λ^{-1} est une valeur propre de $\mathbf{A}^{-1}$ et $\mathbf{x}$ est un vecteur propre correspondant.

8.7 Utilisons le résultat de l'exercice 8.6 pour éviter de calculer l'inverse de la matrice. Le polynôme caractéristique de la matrice $\mathbf{A}$ est
$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \begin{vmatrix} \lambda + 2 & -2 & -3 \\ 2 & \lambda - 3 & -2 \\ 4 & -2 & \lambda - 5 \end{vmatrix} \\ &= \lambda^3 - 6\lambda^2 + 11\lambda - 6 \\ &= (\lambda - 1)(\lambda - 2)(\lambda - 3), \end{aligned}$$
d'où les valeurs propres de $\mathbf{A}$ sont $\lambda_1 = 1$, $\lambda_2 = 2$ et $\lambda_3 = 3$. Par conséquent, les valeurs propres de $\mathbf{A}^{-1}$ sont $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2}$ et $\lambda_3 = \frac{1}{3}$. Toujours par le résultat de l'exercice 8.6, les vecteurs propres de $\mathbf{A}$ correspondant à $\lambda = 1$, $\lambda = 2$ et $\lambda = 3$ sont également des vecteurs propres de $\mathbf{A}^{-1}$ correspondant à $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2}$ et $\lambda_3 = \frac{1}{3}$. Or, le système d'équations $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à
$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
soit $x_1 = s$, $x_2 = 0$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = 1$ est donc $(1, 0, 1)$. D'autre part, le système d'équations $(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à
$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = s/2$, $x_2 = s$ et $x_3 = 0$. Une base de vecteurs propres correspondant à $\lambda = 2$ (ou $\lambda = \frac{1}{2}$) est donc $(\frac{1}{2}, 1, 0)$. Finalement, le système d'équations $(3\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ se réduit à

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

soit $x_1 = s$, $x_2 = s$ et $x_3 = s$. Une base de vecteurs propres correspondant à $\lambda = 3$ (ou $\lambda = \frac{1}{3}$) est donc $(1, 1, 1)$.

8.8 Le résultat découle simplement du fait que l'équation $\mathbf{Ix} = \mathbf{x}$ est vraie pour tout vecteur $\mathbf{x}$.

8.9 a) Le polynôme caractéristique est $\lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, donc les valeurs propres sont $\lambda_1 = \lambda_2 = 2$. Par conséquent,

$$2\mathbf{I} - \mathbf{A} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix},$$

d'où $\text{rang}(2\mathbf{I} - \mathbf{A}) = 1$. Puisque la matrice $\mathbf{A}$ ne possède qu'un seul vecteur propre, elle n'est pas diagonalisable. Vérification :

```
> m <- matrix(c(2, 1, 0, 2), nrow = 2)
> eigen(m)
eigen() decomposition
$values
[1] 2 2

$vectors
     [,1]          [,2]
[1,]    0  4.440892e-16
[2,]    1 -1.000000e+00
```

b) Le polynôme caractéristique est $(\lambda-3)(\lambda^2-8\lambda+15) = (\lambda-3)^2(\lambda-5)$, donc les valeurs propres sont $\lambda_1 = \lambda_2 = 3$ et $\lambda_3 = 5$. La forme échelonnée de la matrice $3\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

d'où $\text{rang}(3\mathbf{I} - \mathbf{A}) = 1$. La base de vecteurs propres correspondant à $\lambda = 3$ est composée des vecteurs $(-1, 0, 1)$ et $(0, 1, 0)$. D'autre part, la forme échelonnée de la matrice $5\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

d'où $\text{rang}(5\mathbf{I} - \mathbf{A}) = 2$. La base de vecteurs propres correspondant à $\lambda = 5$ est $(1, 2, 1)$. Bien que les valeurs propres de la matrice $\mathbf{A}$ ne sont pas toutes distinctes, les vecteurs propres $(1, 0, -1)$, $(0, 1, 0)$ et $(1, 2, 1)$ sont linéairement indépendants. Par conséquent, la matrice

$$\mathbf{P} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

diagonalise $\mathbf{A}$ et

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Vérification :

```
> m <- matrix(c(4, 2, 1, 0, 3, 0, 1, 2, 4), nrow = 3)
> eigen(m)
eigen() decomposition
$values
[1] 5 3 3

$vectors
          [,1] [,2]       [,3]
[1,] 0.4082483    0 -0.7071068
[2,] 0.8164966    1  0.0000000
[3,] 0.4082483    0  0.7071068
```

c) La matrice étant triangulaire, nous savons que les valeurs propres sont $\lambda_1 = 3$ et $\lambda_2 = \lambda_3 = 2$. La forme échelonnée de la matrice $3\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

d'où $\mathrm{rang}(3\mathbf{I} - \mathbf{A}) = 2$. La base de vecteurs propres correspondant à $\lambda = 3$ est $(1, 0, 0)$. D'autre part, la forme échelonnée de la matrice $2\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

d'où $\mathrm{rang}(2\mathbf{I} - \mathbf{A}) = 2$. La base de vecteurs propres correspondant à $\lambda = 2$ est $(0, 0, 1)$. Par conséquent, la matrice $\mathbf{A}$ n'a pas trois vecteurs linéairement indépendants, donc elle n'est pas diagonalisable.

```
> m <- matrix(c(3, 0, 0, 0, 2, 1, 0, 0, 2), nrow = 3)
> eigen(m)
eigen() decomposition
$values
[1] 3 2 2

$vectors
     [,1] [,2]          [,3]
[1,]    1    0  0.000000e+00
[2,]    0    0  4.440892e-16
[3,]    0    1 -1.000000e+00
```

d) Le polynôme caractéristique est $\lambda^3 - 6\lambda^2 + 11\lambda - 6 = (\lambda - 1)(\lambda - 2)(\lambda - 3)$, donc les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$ et $\lambda_3 = 3$. Les valeurs propres étant distinctes, la matrice est diagonalisable. Or, la forme échelonnée de la matrice $\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

d'où $\mathrm{rang}(\mathbf{I} - \mathbf{A}) = 2$. La base de vecteurs propres correspondant à $\lambda = 1$ est $(1, 1, 1)$. Deuxièmement, la forme échelonnée de la matrice $2\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & -\frac{2}{3} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

d'où $\mathrm{rang}(2\mathbf{I} - \mathbf{A}) = 2$ et la base de vecteurs propres correspondant à $\lambda = 2$ est $(2, 3, 3)$. Enfin, la forme échelonnée de la matrice $3\mathbf{I} - \mathbf{A}$

est

$$\begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 \end{bmatrix}$$

d'où rang$(3\mathbf{I} - \mathbf{A}) = 2$ et la base de vecteurs propres correspondant à $\lambda = 3$ est $(1, 3, 4)$. Par conséquent,

$$\mathbf{P} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

diagonalise $\mathbf{A}$ et

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Vérification :

```
> m <- matrix(c(-1, -3, -3, 4, 4, 1, -2, 0, 3),
+             nrow = 3)
> eigen(m)
eigen() decomposition
$values
[1] 3 2 1

$vectors
          [,1]      [,2]       [,3]
[1,] 0.1961161 0.4264014 -0.5773503
[2,] 0.5883484 0.6396021 -0.5773503
[3,] 0.7844645 0.6396021 -0.5773503
```

e) La matrice est triangulaire : les valeurs propres sont donc $\lambda_1 = \lambda_2 = -2$ et $\lambda_3 = \lambda_4 = 3$. La forme échelonnée de la matrice $-2\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

d'où rang$(-2\mathbf{I} - \mathbf{A}) = 2$ et la base de vecteurs propres correspondant à $\lambda = -2$ est composée des vecteurs $(1, 0, 0, 0)$ et $(0, 1, 0, 0)$. D'autre part, la forme échelonnée de la matrice $3\mathbf{I} - \mathbf{A}$ est

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

d'où $\text{rang}(2\mathbf{I} - \mathbf{A}) = 2$ et la base de vecteurs propres correspondant à $\lambda = 3$ est composée des vecteurs $(0, 1, 1, 0)$ et $(0, -1, 0, 1)$. Les valeurs propres de la matrice $\mathbf{A}$ ne sont pas toutes distinctes, mais les vecteurs propres $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 1, 1, 0)$ et $(0, -1, 0, 1)$ sont linéairement indépendants. Par conséquent, la matrice

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

diagonalise $\mathbf{A}$ et

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Vérification :

```
> m <- matrix(c(-2, 0, 0, 0, 0, -2, 0, 0,
+               0, 5, 3, 0, 0, -5, 0, 3),
+             nrow = 4)
> eigen(m)
eigen() decomposition
$values
[1]  3  3 -2 -2

$vectors
          [,1]       [,2] [,3] [,4]
[1,] 0.0000000  0.0000000    1    0
[2,] 0.7071068 -0.7071068    0    1
[3,] 0.7071068  0.0000000    0    0
[4,] 0.0000000  0.7071068    0    0
```

Chapitre 9

9.1 La matrice

$$\mathbf{A} = \begin{bmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{bmatrix}$$

est exprimée sous la forme $\mathbf{A} = \mathbf{LU}$, où

$$\mathbf{L} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & -1 & 2 \end{bmatrix}$$

et

$$\mathbf{U} = \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ainsi, le système d'équations $\mathbf{Ax} = \mathbf{b}$ (où $\mathbf{b} = (-3, -22, 3)^\mathsf{T}$) peut être exprimé sous la forme $\mathbf{LUx} = \mathbf{b}$, soit $\mathbf{Ly} = \mathbf{b}$ et $\mathbf{Ux} = \mathbf{y}$. On résoud tout d'abord $\mathbf{Ly} = \mathbf{b}$ par simple substitution successive. On trouve

$$\begin{aligned} y_1 &= -1 \\ y_2 &= \frac{-22 - 2y_1}{4} = -5 \\ y_3 &= \frac{3 + 4y_1 + y_2}{2} = -3. \end{aligned}$$

Par la suite, on résoud de même $\mathbf{Ux} = \mathbf{y}$, ce qui donne

$$\begin{aligned} x_3 &= -3 \\ x_2 &= -5 - 2x_3 = 1 \\ x_1 &= -1 + 2x_2 + x_3 = -2. \end{aligned}$$

9.2 Cherchons tout d'abord des matrices triangulaires inférieure et supérieure $\mathbf{L}$ et $\mathbf{U}$, respectivement, tel que $\mathbf{A} = \mathbf{LU}$. On nous donne dans l'énoncé

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix},$$

d'où

$$\mathbf{U} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix},$$

et $\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}$. Or,

$$\begin{aligned}
\mathbf{E}_1^{-1} &= \frac{1}{3}\begin{bmatrix} 3 & 0 \\ -2 & 1 \end{bmatrix} \\
\mathbf{E}_2^{-1} &= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \\
&= 3\mathbf{I}
\end{aligned}$$

et, par conséquent,

$$\mathbf{L} = \begin{bmatrix} 3 & 0 \\ -2 & 1 \end{bmatrix}.$$

Pour résoudre par décomposition *LU* le système d'équations $\mathbf{Ax} = \mathbf{LUx} = \mathbf{b}$, où $\mathbf{b} = (0, 1)$, on pose $\mathbf{Ux} = \mathbf{y}$ et résout d'abord $\mathbf{Ly} = \mathbf{b}$ par substitution. On a donc le système d'équations

$$\begin{bmatrix} 3 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

dont la solution est $y_1 = 0$ et $y_2 = 1$. Par la suite, on a

$$\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

d'où, finalement, $x_1 = 2$ et $x_2 = 1$.

Bibliographie

Anton, H. 2000, *Elementary Linear Algebra*, 8e éd., Wiley, ISBN 0-4711705-5-0.

Dua, D. et C. Graff. 2019, «UCI machine learning repository», URL http://archive.ics.uci.edu/ml.

Goulet, V. 2018, *Programmer avec R*, Document libre. URL https://vigou3.gitlab.io/programmer-avec-r.

Kondo, Y., M. Salibian-Barrera et R. Zamar. 2016, «RSKC: An R package for a robust and sparse K-means clustering algorithm», *Journal of Statistical Software*, vol. 72, no 5, doi :10.18637/jss.v072.i05, p. 1–26.

Moler, C. et C. Van Loan. 1978, «Nineteen dubious ways to compute the exponential of a matrix», *SIAM Review*, vol. 20, p. 801–836.

Ce document a été produit avec le système de mise en page XƎLaTeX. Le texte principal est composé en Lucida Bright OT 11 points, les mathématiques en Lucida Bright Math OT, le code informatique en Lucida Grande Mono DK et les titres en Fira Sans. Des icônes proviennent de la police Font Awesome. Les graphiques ont été réalisés avec R.