Module : L'Analyse de données

Filière : IA | GIN -S_7-

Année : 2025-2026

► Professeur de cours : Labloul Bouchra

Ce polycopié contient le cours et les sujets d'exercices, la correction de ces exercices sera bien détaillée aux séances de travaux dirigés

Table des matières

1 Rappels et compléments d'algèbre linéaire Décompositions de matrices **5**

1.1 Le calcul matriciel 7

1.1.1 Égalité et opérations sur les matrices 7

1.1.2 Types particuliers de matrices 9

1.1.3 Transposée d'une matrice 10

1.1.4 Propriétés fondamentales 10

1.1.5 Matrice identité et matrice inverse 11

1.1.6 Propriétés du produit matriciel 12

1.1.7 Le déterminant d'une matrice 13

1.1.8 Diagonalisation d'une matrice 15

1.1.9 Valeurs propres et vecteurs propres 15

Introduction

L'analyse de données occupe une place centrale dans la recherche scientifique et la prise de décision. Elle consiste à collecter, organiser, traiter et interpréter des données afin d'en extraire des informations pertinentes. Grâce aux méthodes statistiques et aux outils informatiques, elle permet de décrire les phénomènes observés, de détecter des structures cachées et de formuler des conclusions objectives.

Dans un monde où la quantité d'informations disponibles ne cesse de croître, l'analyse de données offre une approche méthodique pour transformer ces données brutes en connaissances exploitables. Elle s'applique dans des domaines variés tels que l'économie, la biologie, la sociologie, ou encore le marketing, où elle aide à comprendre les comportements, à identifier des tendances et à appuyer la prise de décision.

Ce cours a pour objectif de présenter les principales étapes du processus d'analyse de données : depuis la préparation et le nettoyage du jeu de données jusqu'à l'application de méthodes exploratoires comme l'Analyse en Composantes Principales (ACP). L'objectif final est de comprendre les relations entre les variables, de réduire la dimensionnalité et de faciliter l'interprétation des résultats. L'étude mettra en évidence la démarche rigoureuse de l'analyse de données, fondée sur des outils statistiques et géométriques, permettant d'obtenir une représentation simplifiée mais fidèle de la réalité observée.

Chapitre 1

Rappels et compléments d'algèbre linéaire Décompositions de matrices

Introduction

L'algèbre linéaire occupe une place essentielle dans les mathématiques modernes et leurs applications. Elle fournit un cadre théorique et pratique pour l'étude des **espaces vectoriels**, des **matrices** et des **transformations linéaires**, qui constituent la base de nombreuses disciplines scientifiques et techniques.

Dans un contexte où la modélisation et l'analyse de données jouent un rôle central, l'algèbre linéaire permet de *représenter, manipuler et résoudre efficacement des systèmes d'équations linéaires*, d'interpréter des phénomènes géométriques ou physiques, et de formuler des modèles dans des domaines aussi variés que la physique, l'économie, l'informatique, ou encore les statistiques.

L'objectif de ce rapport est de présenter les **principes fondamentaux** de l'algèbre linéaire et de montrer leur **importance dans la compréhension et la résolution de problèmes concrets**. Nous aborderons successivement les notions de base (vecteurs, matrices, espaces vectoriels), les **opérations linéaires** et leurs **propriétés**, ainsi que quelques **applications pratiques** illustrant l'utilité de cette discipline dans l'analyse de données et la modélisation mathématique. Ainsi, ce travail vise à offrir une vision claire et synthétique des outils de l'algèbre linéaire, tout en soulignant leur rôle central dans le développement de nombreuses méthodes quantitatives contemporaines.

1.1 Le calcul matriciel

Définition 1. *Soient $m, n \in \mathbb{N}^*$. On appelle* ***matrice*** *de taille (m, n) à coefficients dans un corps $\mathbb{K}$ tout tableau rectangulaire à m lignes et n colonnes dont les éléments appartiennent à $\mathbb{K}$:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

L'ensemble de toutes les matrices de taille (m, n) à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{m,n}(\mathbb{K})$.

1.1.1 Égalité et opérations sur les matrices

Définition 2. *Égalité. Deux matrices $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$ sont égales si et seulement si*

$$a_{ij} = b_{ij}, \quad \forall i \in \{1, \dots, m\},\ \forall j \in \{1, \dots, n\}.$$

Somme et multiplication par un scalaire. Si $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$ et $\lambda \in \mathbb{K}$, on définit :

$$A + B = (a_{ij} + b_{ij}) \quad \text{et} \quad \lambda A = (\lambda a_{ij}).$$

Multiplication matricielle. Si $A \in \mathcal{M}_{m,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$, leur produit $AB \in \mathcal{M}_{m,n}(\mathbb{K})$ est défini par :

$$(AB)_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Exemple 1 (Somme de matrices)**.** *Soient deux matrices A et B de même taille $(2,3)$:*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad B = \begin{pmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

La somme $A + B$ est obtenue en additionnant les coefficients correspondants :

$$A + B = \begin{pmatrix} 1+6 & 2+5 & 3+4 \\ 4+3 & 5+2 & 6+1 \end{pmatrix} = \begin{pmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \end{pmatrix}.$$

Exemple 2 (Produit de matrices)**.** *Soient*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \textit{et} \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}.$$

Le produit AB est défini car A est de taille $(2,2)$ et B également. On a :

$$AB = \begin{pmatrix} 1 \times 2 + 2 \times 1 & 1 \times 0 + 2 \times 3 \\ 3 \times 2 + 4 \times 1 & 3 \times 0 + 4 \times 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 10 & 12 \end{pmatrix}.$$

Remarquons que, dans le cas général, la multiplication n'est pas commutative :

$$AB \neq BA.$$

Exemple 3 (Produit non défini)**.** *Si*

$$A = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \quad et \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

alors le produit AB n'est pas défini, car A est de taille $(1,3)$ et B de taille $(2,2)$: le nombre de colonnes de A (3) est différent du nombre de lignes de B (2).

Remarque 1.1.1. *La somme de matrices n'est possible que si les matrices ont la **même taille**. Le produit AB n'est possible que si le nombre de **colonnes de** A est égal au nombre de **lignes de** B.*

1.1.2 Types particuliers de matrices

- **Matrice carrée :** une matrice $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est dite *carrée*.
- **Matrice nulle :** $A = 0$ si $a_{ij} = 0$ pour tout i, j.
- **Matrice identité :**

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

- **Matrice diagonale :** $A = (a_{ij})$ telle que $a_{ij} = 0$ si $i \neq j$.
- **Matrice triangulaire supérieure :** $a_{ij} = 0$ si $i > j$.
- **Matrice symétrique :** $A^T = A$.
- **Matrice antisymétrique :** $A^T = -A$.
- **Matrice idempotente :** $A^2 = A$.

1.1.3 Transposée d'une matrice

Définition 3. *Si* $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$*, sa* ***transposée*** A^T *est la matrice de* $\mathcal{M}_{n,m}(\mathbb{K})$ *définie par :*

$$(A^T)_{ij} = a_{ji}.$$

Exemple 4.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

1.1.4 Propriétés fondamentales

$$(A+B)^T = A^T + B^T, \qquad (\lambda A)^T = \lambda A^T, \qquad (AB)^T = B^T A^T.$$

$$I_n^T = I_n, \qquad (A^T)^T = A.$$

Remarque 1.1.2. *Les matrices sont des outils fondamentaux de l'algèbre linéaire : elles représentent les applications linéaires dans une base donnée. Les propriétés de ces matrices traduisent directement les propriétés des applications correspondantes (inversibilité, symétrie, projection, rotation, etc.).*

1.1.5 Matrice identité et matrice inverse

Définition 4 (Matrice identité)**.** *Pour tout entier $n \geq 1$, on appelle* ***matrice identité*** *d'ordre n la matrice carrée*

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Cette matrice vérifie :

$$AI_n = I_nA = A, \quad \forall A \in \mathcal{M}_{n,n}(\mathbb{K}).$$

Définition 5 (Matrice inverse)**.** *Une matrice carrée $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est dite* ***inversible*** *s'il existe une matrice B telle que :*

$$AB = BA = I_n.$$

Dans ce cas, la matrice B est unique et notée A^{-1}.

Proposition 1.1.1 (Propriétés)**.**

- *Si A et B sont deux matrices inversibles d'ordre n, alors :*

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (A^{-1})^{-1} = A, \quad (A^T)^{-1} = (A^{-1})^T.$$

- *Cas particulier : matrice 2×2*

  Soit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Si le déterminant de A, noté $\det(A) = ad-bc$*, est non nul, alors A est inversible et :*

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

Exemple 5. *Soit*

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

On calcule :

$$\det(A) = 2 \times 3 - 5 \times 1 = 6 - 5 = 1.$$

Ainsi A est inversible et

$$A^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

On vérifie :

$$AA^{-1} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

1.1.6 Propriétés du produit matriciel

- **Associativité :** $(AB)C = A(BC)$ lorsque les produits sont définis.
- **Distributivité :** $A(B+C) = AB + AC$ et $(A+B)C = AC + BC$.
- **Non commutativité :** en général, $AB \neq BA$.
- **Compatibilité avec la transposée :** $(AB)^T = B^T A^T$.

Exemple 6 (Illustration de la non commutativité). *Soient*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad et \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Alors

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = BA.$$

1.1.7 Le déterminant d'une matrice

Définition 6. *Le* ***déterminant*** *est une application qui associe à toute matrice carrée $A \in \mathcal{M}_{n,n}(\mathbb{K})$ un nombre réel (ou complexe) noté $\det(A)$, appelé déterminant de A. Le déterminant mesure le* ***facteur d'échelle et d'orientation*** *de la transformation linéaire représentée par A.*

Définition 7 (Cas particulier 2×2). *Si*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

alors

$$\det(A) = ad - bc.$$

Exemple 7.

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \quad \Rightarrow \quad \det(A) = 2 \times 4 - 3 \times 1 = 8 - 3 = 5.$$

Le déterminant est non nul, donc A est inversible.

Définition 8 (Cas 3×3). *Si*

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix},$$

alors

$$\det(A) = aei + bfg + cdh - ceg - bdi - afh.$$

Exemple 8.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix}$$

On calcule :

$$\det(A) = 1(1 \cdot 0 - 4 \cdot 6) - 2(0 \cdot 0 - 4 \cdot 5) + 3(0 \cdot 6 - 1 \cdot 5) = 1(-24) - 2(-20) + 3(-5) = -24 + 40 - 15 = 1.$$

Ainsi A est inversible et $\det(A^{-1}) = \frac{1}{\det(A)} = 1$.

Propriétés fondamentales du déterminant

— $\det(I_n) = 1$.

— $\det(A^T) = \det(A)$.

— $\det(AB) = \det(A)\det(B)$.

— Si une ligne ou une colonne est nulle, alors $\det(A) = 0$.

— Si deux lignes (ou colonnes) sont égales, alors $\det(A) = 0$.

— Si on multiplie une ligne par un scalaire λ, le déterminant est multiplié par λ.

— Une matrice A est **inversible** si et seulement si $\det(A) \neq 0$.

1.1.8 Diagonalisation d'une matrice

Définition 9. *Une matrice carrée $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est dite **diagonalisable** s'il existe une matrice inversible P telle que :*

$$P^{-1}AP = D,$$

*où D est une matrice diagonale. Dans ce cas, les colonnes de P sont des **vecteurs propres** de A, et les coefficients diagonaux de D sont les **valeurs propres** associées.*

Proposition 1.1.2 (Caractérisation de la diagonalisabilité)**.** *Une matrice A est diagonalisable si et seulement si il existe une base de $\mathbb{K}^n$ formée de vecteurs propres de A, c'est-à-dire si la somme des dimensions des sous-espaces propres est égale à n.*

1.1.9 Valeurs propres et vecteurs propres

Définition 10. *Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. Un **vecteur propre** de A est un vecteur non nul x tel qu'il existe un scalaire λ vérifiant :*

$$Ax = \lambda x.$$

*Le scalaire λ est alors appelé **valeur propre** de A.*

Proposition 1.1.3. *Les valeurs propres de A sont les racines du* ***polynôme caractéristique*** *:*

$$p_A(\lambda) = \det(A - \lambda I_n).$$

Exemple 9 (Exemple numérique de diagonalisation)**.** *Soit*

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

Le polynôme caractéristique est :

$$p_A(\lambda) = \det(A - \lambda I_2) = (4 - \lambda)(3 - \lambda) - 2 = \lambda^2 - 7\lambda + 10.$$

Les racines sont $\lambda_1 = 5$ et $\lambda_2 = 2$.

Pour $\lambda_1 = 5$, on résout $(A - 5I)x = 0$, soit :

$$\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad \Rightarrow \quad x_1 = x_2.$$

On peut prendre $v_1 = (1, 1)$.

Pour $\lambda_2 = 2$, on résout $(A - 2I)x = 0$, soit :

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad \Rightarrow \quad x_2 = -2x_1.$$

On peut prendre $v_2 = (1, -2)$.

Ainsi,

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \qquad D = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}.$$

On vérifie que :

$$A = PDP^{-1}.$$

La matrice A est donc ***diagonalisable****.*

Interprétation géométrique

- Chaque vecteur propre définit une **direction invariante** par la transformation linéaire associée à A.
- La valeur propre λ indique le **facteur d'allongement** (ou de contraction) sur cette direction.
- La diagonalisation permet donc de **simplifier** les calculs de puissances :

$$A^k = PD^kP^{-1}.$$

Exemple 10. *Soit la matrice*

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

1. *Trouver les valeurs propres de A.*
2. *Pour chaque valeur propre, donner un vecteur propre non nul.*

Solution 1.

Trouver les valeurs propres de A.

▷ *On calcule* $\det(A - \lambda I)$ *:*

$$A - \lambda I = \begin{pmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{pmatrix}.$$

En développant par la première colonne (ou en utilisant la forme tridiagonale) on obtient :

$$\begin{aligned} \det(A - \lambda I) &= (2-\lambda)\det\Big(2-\lambda \quad 1[2pt]1 \quad 2-\lambda\Big) - 1\cdot\det\Big(1 \quad 1[2pt]0 \quad 2-\lambda\Big) \\ &= (2-\lambda)\Big((2-\lambda)^2 - 1\Big) - (2-\lambda) \\ &= (2-\lambda)\Big((2-\lambda)^2 - 2\Big). \end{aligned}$$

L'équation caractéristique est donc

$$(2-\lambda)\Big((2-\lambda)^2 - 2\Big) = 0.$$

On en déduit les racines :

$$\begin{aligned} 2-\lambda = 0 \quad &\Rightarrow \quad \lambda = 2, \\ (2-\lambda)^2 = 2 \quad &\Rightarrow \quad 2-\lambda = \pm\sqrt{2} \end{aligned}$$

Ainsi

$$\boxed{\lambda_1 = 2+\sqrt{2}, \qquad \lambda_2 = 2, \qquad \lambda_3 = 2-\sqrt{2}.}$$

Pour chaque valeur propre, donner un vecteur propre non nul.

▷ *Pour chaque valeur propre λ, on résout le système $(A - \lambda I)x = 0$.*

Pour $\lambda = 2$ *On a*

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

Le système s'écrit

$$\begin{cases} x_2 = 0, \\ x_1 + x_3 = 0. \end{cases}$$

Donc $x_2 = 0$ *et* $x_3 = -x_1$. *Un vecteur propre non nul associé est, par exemple,*

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Pour $\lambda = 2 + \sqrt{2}$ *Posons* $t = 2 - (2 + \sqrt{2}) = -\sqrt{2}$. *Alors*

$$A - (2 + \sqrt{2})I = \begin{pmatrix} t & 1 & 0 \\ 1 & t & 1 \\ 0 & 1 & t \end{pmatrix}.$$

Le premier équation donne $x_2 = -tx_1$. *En remplaçant dans la deuxième on obtient*

$$x_3 = -x_1 - tx_2 = -x_1 - t(-tx_1) = (-1 + t^2)x_1.$$

Comme $t^2 = 2$, *on a* $x_3 = x_1$. *En choisissant* $x_1 = 1$ *on obtient le vecteur propre*

$$v_+ = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

Pour $\lambda = 2 - \sqrt{2}$ *Ici* $t = 2 - (2 - \sqrt{2}) = +\sqrt{2}$*. De la même façon on obtient*

$$x_2 = -\sqrt{2}, x_1, \qquad x_3 = x_1.$$

En choisissant $x_1 = 1$ *on obtient*

$$v_- = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

— *Valeurs propres :* $2 + \sqrt{2}; 2; 2 - \sqrt{2}$.

— *Vecteurs propres correspondants :*

$$v_+ = (1, \sqrt{2}, 1)^\top, \qquad v_2 = (1, 0, -1)^\top, \qquad v_- = (1, -\sqrt{2}, 1)^\top.$$