Introduction aux mathématiques pour les biologistes

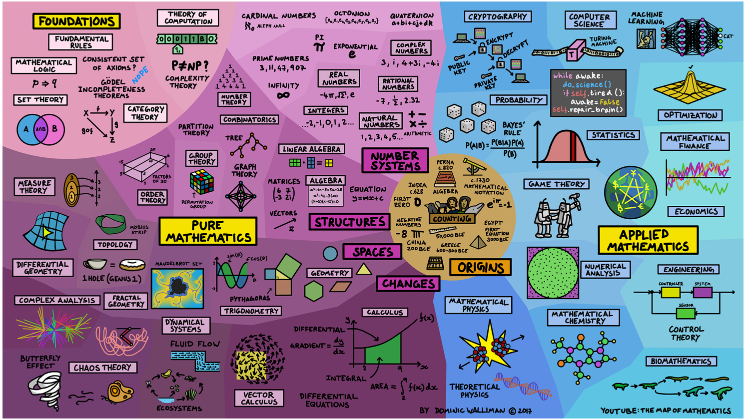
Algèbre linéaire

Pierre Veron

Doctorant
Institut de biologie de l'École normale supérieure – PSL
Écologie Systématique Évolution, Université Paris-Saclay
pierre-veron.github.io / pveron@bio.ens.psl.eu

14 novembre 2024

Qu’allons-nous faire ?



- Algèbre linéaire
- Matrices
- Résolution de systèmes linéaires
- Analyse en composantes principales
- Chaînes de Markov

Algèbre

L'algèbre est une branche mathématique qui s'intéresse aux propriétés d'objets sur lesquels sont définies des opérations.

Par exemple

- Addition sur deux nombres entiers : $1 + 2 = 3$
- Addition sur deux nombres relatifs : $1 + (-2) = -1$
- Addition sur deux nombres rationnels : $\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$
- Addition sur deux nombres réels, complexes...
- Peut-on additionner d'autres choses ?

Qu'est-ce qui définit une *addition* de manière générale ?

Algèbre > Structures algébriques

Une **structure algébrique** est un ensemble défini par plusieurs axiomes.

Par exemple : la loi « addition » notée « + » doit être

- commutative : $x + y = y + x$,
- associative : $(x + y) + z = x + (y + z)$
- admettre un élément neutre : il existe e tel que $x + e = x$ pour tout x,
- ...

En étudiant les propriétés des structures algébriques, on peut en déduire des propriétés générales, même sans préciser le type d'objets avec lequel on travail (nombres, fonctions, vecteurs, matrices...).

Algèbre > Structures algébriques > Algèbre linéaire

L'algèbre linéaire s'intéresse plus spécifiquement aux espaces vectoriels et aux applications linéaires. Les espaces suivants peut être considérés comme des espaces vectoriels : la droite, le plan, l'espace...

Les espaces vectoriels peuvent être de dimension finie (par exemple 3 pour notre espace) ou de dimension infinie (l'espace des fonctions réelles).

Les applications linéaires sont des transformations géométriques de l'espace qui respectent certaines règles. Sous cette hypothèse, les applications linéaires dans des espaces de dimension finie peuvent être représentés par des **matrices**. C'est l'objet de ce cours.

Pourquoi en a-t-on besoin ? (1)

Prenons une séquence de nucléotides :

CCAGTACCTAGGCCATAGGAGAGGGAGAGGCTCCTACTATTAAGAGTTAA

que l'on peut résumer par le nombre de ses nucléotides 16 A,14 G, 10 C, 10 T.

Supposons que la probabilité de chaque transition (A$\leftrightarrow$ G ou C$\leftrightarrow$ T) est de 5% et de chaque transversion (les autres) est de 1% pour chaque nucléotide. Quel est la répartition attendue après 1 génération ? 10 générations ? À l'infini ?

Pourquoi en a-t-on besoin ? (2)

```
Generation 0 : AAAACTAAGCCGTTATTCATATGGACCGTGGCGTATGAGTGCGGGAATGC
Generation 1 : AAAACTAAGCCGTTATACATATGGACAGTGGCTTAGGAGTGAGGGAATGC
Generation 2 : AAACCTAAGCCGTTATAAATATGGACAGTGGCTTATGAGTGAGTGAATTC
Generation 3 : AAACCTAAGCCGTTATAAATATGGACAGTGGCCGATGAGTGAGTGACTTC
Generation 4 : ACACCTAAGCCGTTATAAATATGGACAGGGGCCGTTTAGTGAGTGCGGTC
Generation 5 : ACCCCTAAGCCGTCATAAATCTTGACAGGGCTCGGTTAGGGAGTGCGGTC
Generation 6 : ACCCCTAAGCAGTCATAACTCTTGACAGGGCTCGGTTAGGGAATGCGTTC
Generation 7 : ACCCCTAAGGAGTCATAACACTTTAGAGGGCTCGGTTAGGGAATTCTTTC
Generation 8 : CCCCCTCAGGAGTCATAACACTTTAGAGGGCTCGGTTAGGGAATTCTTTC
Generation 9 : CCCACTCAGGAGGCATAACACTTTAGAGGGCTCGGGTAGGGAATTCTTTC
Generation 10: CCCACGCAGGAGGCATAACACTTTAGAGGGCTCGGGTCGGTAATTCGTTC
```

Pourquoi en a-t-on besoin ? (3)

Réponse : (après le cours) en notant $\pi_0 = (n_A, n_G, n_C, n_T)$ la répartition initiale et M la matrice de transition :

$$M = \begin{pmatrix} 0.93 & 0.05 & 0.01 & 0.01 \\ 0.05 & 0.93 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.93 & 0.05 \\ 0.01 & 0.01 & 0.05 & 0.95 \end{pmatrix} \tag{1}$$

la répartition après n étapes est une transformation linéaire de π_0 notée par $\pi_n = \pi_0 M^n$. Avec les propriétés des chaînes de Markov, la répartition à l'équilibre est donnée par le vecteur propre principal de M, et la vitesse de convergence est donné par la seconde valeur propre de M.

ESE
Écologie Systématique Évolution

Qu’est-ce qu’un espace vectoriel ?

Un espace vectoriel est un ensemble d’objets mathématiques :

- que l’on peut additionner entre eux
- que l’on peut multiplier par un nombre
- qui contient un élément nul, noté $\vec{0}$.

Les éléments d’un espace vectoriel sont appelé des **vecteurs**, alors que les nombres sont appelés des **scalaires**.

Définition formelle d'un espace vectoriel

Définition – Espace vectoriel

Soit E un ensemble muni d'une loi de composition interne « + » et d'une loi de composition externe par un scalaire « · » vérifiant les propriétés suivantes $\forall \vec{x}, \vec{y}, \vec{z} \in E, \forall a, b \in \mathbb{R}$:

- « + » est commutative : $\vec{x} + \vec{y} = \vec{y} + \vec{x}$,
- « + » est associative : $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$,
- « + » admet un élément neutre noté $\vec{0}$: $\vec{0} + \vec{x} = \vec{x}$,
- tout vecteur $\vec{x}$ a un opposé pour « + » noté $-\vec{x}$: $\vec{x} + (-\vec{x}) = \vec{0}$,
- « · » est distributive par rapport à « + » :

$$(a+b) \cdot \vec{x} = a \cdot \vec{x} + b \cdot \vec{x} \quad \text{et} \quad a \cdot (\vec{x} + \vec{y}) = a \cdot \vec{x} + a \cdot \vec{y}$$

- « · » est associative par rapport à la multiplication usuelle sur $\mathbb{R}$: $(ab) \cdot \vec{x} = a \cdot (b \cdot \vec{x})$,
- 1 est neutre : $1 \cdot \vec{x} = \vec{x}$.

Remarque et propriétés (1)

En pratique on note plutôt la loi externe sans symbole, c'est à dire $a\vec{x}$ au lieu de $a \cdot \vec{x}$. On ne note pas systématiquement $\vec{x}$ mais parfois $\mathbf{x}$ ou simplement x.

Propriétés

Soit E un $\mathbb{R}$-espace vectoriel. Alors, pour tous $a \in \mathbb{R}$ et tous $\vec{x} \in E$

1. $a\vec{x} = \vec{0} \Leftrightarrow a = 0$ ou $\vec{x} = \vec{0}$.
2. $(-a)\vec{x} = -(a\vec{x}) = a(-\vec{x})$.

Remarque et propriétés (2)

Démonstration –

1. (i) Montrons dans un premier temps que $0\vec{x} = \vec{0}$. Pour cela il suffit d'écrire $0 = 0 + 0$ et d'utiliser la propriété de distributivité donc $0\vec{x} = (0+0)\vec{x} = 0\vec{x} + 0\vec{x}$. En ajoutant à droite et à gauche l'opposé du vecteur $0\vec{x}$ on obtient : $0\vec{x} = \vec{0}$.
   (ii) Montrons ensuite que $a\vec{0} = \vec{0}$. Pour cela, écrivons $a\vec{0} = a(\vec{0} + \vec{0}) = a\vec{0} + a\vec{0}$ en utilisant la distributivité. Finalement, il suffit d'ajouter l'opposé du vecteur $a\vec{0}$ pour constater $a\vec{0} = 0$.
   (iii) Enfin, admettons $a\vec{x} = \vec{0}$. Supposons $a \neq 0$. On a alors

$$\vec{0} = \frac{1}{a}\vec{0} = \frac{1}{a}(a\vec{x}) = \left(\frac{1}{a}a\right)\vec{x} = 1\vec{x} = \vec{x}.$$

   (i), (ii) et (iii) prouvent l'équivalence.

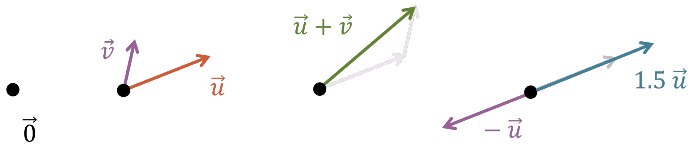
Remarque et propriétés (3)

2. Pour cela il suffit de constater que $(-a)\vec{x} + a\vec{x} = (-a + a)\vec{x} = \vec{0}$ donc $(-a)\vec{x}$ est l'opposé de $a\vec{x}$ donc $(-a)\vec{x} = -(a\vec{x})$.

$\square$

Exemples d'espaces vectoriels

- $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel
- $\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel
- l'espace des vecteurs est un espace vectoriel



Vecteur nul

Addition de vecteurs

Opposé d'un vecteur et multiplication par un scalaire

- l'espace des fonctions réelles
- $\mathbb{R}^n = \{(x_1, x_2, ..., x_n)\}$, l'ensemble des n-uplets (ou n-vecteurs) de réels est un espace vectoriel.

Sous-espace vectoriel (1)

Définition – Sous-espace vectoriel

Soit E un espace vectoriel et $F \subset E$ un sous-ensemble non-vide de F. F est un **sous-espace vectoriel** de E si :

- F est stable par addition, c'est-à-dire pour tous $\vec{x}, \vec{y} \in F$, $\vec{x} + \vec{y} \in F$,
- F est stable par multiplication par un scalaire, c'est-à-dire pour tout $\vec{x} \in F$ et $a \in \mathbb{R}$, $a\vec{x} \in F$.

Sous-espace vectoriel (2)

Propriétés

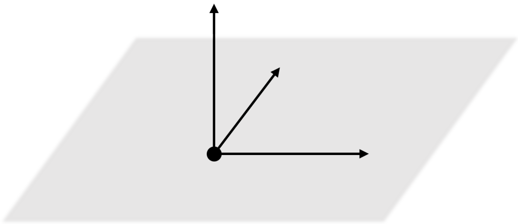
Si F est un sous-espace vectoriel de E alors

- F est lui-même un $\mathbb{R}$-espace vectoriel,
- $\vec{0} \in F$,
- F est stable par combinaison linéaire, c'est-à-dire pour tous $\vec{x}_1, \ldots, \vec{x}_k \in F$ vecteurs et $a_1, \ldots, a_k \in \mathbb{R}$
$$\sum_{i=1}^{k} a_i \vec{x}_i = a_1 \vec{x}_1 + \cdots + a_k \vec{x}_k \in F.$$

Sous-espace vectoriel (3)

Par exemple

- un plan passant par l'origine dans un espace à 3 dimensions est un sous-espace vectoriel



- l'ensemble des nombres réels $\mathbb{R}$ dans $\mathbb{C}$ est un sous-espace vectoriel.

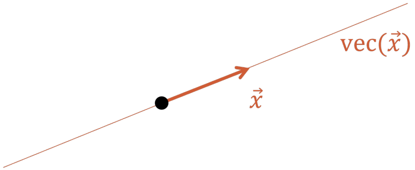
Sous-espace vectoriel (4)

Définition – Sous-espace vectoriel engendré par une famille de vecteurs

Soit $\vec{x}_1, \ldots, \vec{x}_k$ une famille de k vecteurs de E. L'ensemble des vecteurs formés par combinaison linéaire de cette famille, noté $\mathrm{vec}(\vec{x}_1, \ldots, \vec{x}_n)$ est un sous-espace vectoriel de E.

$$\mathrm{vec}(\vec{x}_1, \ldots, \vec{x}_n) = \{a_1\vec{x}_1 + \cdots a_k\vec{x}_k,\ a_1, \ldots, a_k \in \mathbb{R}\}.$$

Exemple 1. $\mathrm{vec}(\vec{x}) = \{a\vec{x},\ a \in \mathbb{R}\}$ est une **droite vectorielle**, c'est-à-dire l'ensemble des vecteurs proportionnels à $\vec{x}$ (si $\vec{x} \neq \vec{0}$),



Sous-espace vectoriel (5)

Exemple 2. Si $\vec{x}$ et $\vec{y}$ ne sont pas proportionnels alors $\mathrm{vec}(\vec{x}, \vec{y})$ est un **plan vectoriel**.

Familles libres et génératrices (1)

Soit $\mathcal{F} = (x_1, ..., x_k)$ une famille de k vecteurs de E.

Définition – Famille libre

$\mathcal{F}$ est une **famille libre** si les vecteurs qui la composent sont *linéairement indépendants*, c'est-à-dire qu'aucun de ses vecteurs ne peut s'écrire comme une combinaison linéaire des autres.

Dans le cas contraire, on dit que la famille est **liée**.
Exemples dans $\mathbb{R}^3$:

- la famille de vecteurs $((1,0,1),(1,1,-1),(0,2,0))$ est libre car on ne peut pas écrire un de ces vecteurs comme combinaison linéaire des autres,

Familles libres et génératrices (2)

- la famille de vecteurs $((1,0,1),(0,1,-1),(2,2,0))$ est liée car
$$(2,2,0) = 2 \times (1,0,1) + 2 \times (0,1,-1).$$
- toute famille qui contient le vecteur nul est liée.

Propriété

Retirer des vecteurs à une famille libre, elle restera libre.

Définition – Famille génératrice

$\mathcal{F}$ est une famille génératrice si
$$E = \mathrm{vec}(\vec{x}_1, \ldots, \vec{x}_k)$$
c'est-à-dire que tout vecteur de E peut s'écrire comme une combinaison linéaire des vecteurs de $\mathcal{F}$.

Familles libres et génératrices (3)

Exemples dans $\mathbb{R}^3$:

- la famille $((1,0,0),(0,1,0),(0,0,1))$ est une famille génératrice
- la famille $((1,1,0),(0,1,0),(1,-1,0)),(2,-1,0))$ n'est pas une famille génératrice.

Propriété

Ajouter des vecteurs à une famille génératrice, elle restera génératrice.

Qu'est-ce qu'une base ?

Définition – Base d'un espace vectoriel

Une base d'un espace vectoriel est une famille libre et génératrice.

Remarque

Tous les espaces vectoriels n'ont pas de base.

Exemples dans $\mathbb{R}^3$ la famille $((1,0,0),(0,1,0),(0,0,1))$ est une base. La famille $((1,0,1),(1,1,-1),(0,2,0))$ est aussi une base.

L'espace vectoriel des fonctions réelles n'admet pas de base.

Dimension d'un espace vectoriel

Théorème – Dimension d'un espace vectoriel

Soit E un espace vectoriel. On suppose que $\mathcal{B} = (\vec{x}_1, \ldots, \vec{x}_n)$ est une base de cet espace vectoriel. Alors toute autre base de cet espace vectoriel admet exactement le même nombre d'éléments. Ce nombre n est la **dimension** de E, noté $\dim(E)$.
Si E admet une base, on dit qu'il est de dimension finie.

Exemples : $\mathbb{R}^n$ est un espace vectoriel de dimension n. Une base de cet espace est $\mathcal{B} = ((1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 1))$ appelé la **base canonique** de $\mathbb{R}^n$.

$\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel de dimension 2, une base est $(1, i)$.

Caractérisation des bases

Théorème –

Soit E un espace vectoriel de dimension n.

- Une famille libre avec n éléments est une base.
- Si une famille est libre, elle a au plus n éléments.
- Une famille génératrice avec n éléments est une base.
- Si une famille est génératrice, elle a au moins n éléments.

Calcul dans une base (1)

Mais à quoi sert une base ?

Travailler avec une base permet de se ramener à travailler avec n réels, grâce à ce résultat :

Théorème – Décomposition dans une base

Soit E un espace vectoriel de dimension n et $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_n)$ une base de E. Alors tout vecteur $\vec{x} \in E$ admet une **unique décomposition** dans la base $\mathcal{B}$ c'est-à-dire qu'il existe n uniques réels $a_1, ..., a_n$ tels que

$$\vec{x} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n.$$

Calcul dans une base (2)

Ceci se note de façon compacte

$$\vec{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}}$$

Puisque la décomposition est unique, une fois que l'on a fixé une base, on peut assimiler $\vec{x}$ et le n-uplet $(a_1, \ldots, a_n)$, sa représentation (ou décomposition) dans la base $\mathcal{B}$.

Si $\vec{x}$ et $\vec{y}$ sont des vecteurs de E et $\lambda \in \mathbb{R}$ tels que

$$\vec{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}} \quad \text{et} \quad \vec{y} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{\mathcal{B}}$$

Calcul dans une base (3)

alors

$$\lambda\vec{x} + \vec{y} = \begin{pmatrix} \lambda a_1 + b_1 \\ \lambda a_2 + b_2 \\ \vdots \\ \lambda a_n + b_n \end{pmatrix}_{\mathcal{B}}$$

Application linéaire

Définition – Application linéaire

Soit E et F deux espaces vectoriels. Soit $f : E \to F$ une application (= une fonction). f est une application linéaire si $\forall \vec{x}, \vec{y} \in E, \forall \lambda \in \mathbb{R}$

- $f(\lambda \vec{x}) = \lambda f(\vec{x})$
- $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$.

Conséquence. Si f est une application linéaire alors nécessairement

- $f(\vec{0}) = \vec{0}$ (si $E \neq F$ alors ce n'est pas le même zéro),
- $f(a_1 \vec{x}_1 + \cdots + a_k \vec{x}_k) = a_1 f(\vec{x}_1) + \cdots + a_k f(\vec{x}_k)$.

Exemples d'applications linéaires

- homothétie

$$f : \begin{pmatrix} E & \longrightarrow & E \\ \vec{x} & \longmapsto & c\vec{x} \end{pmatrix}$$

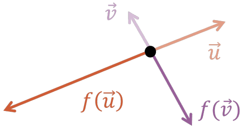
- identité (homothétie avec $c = 1$)

$$\mathrm{id} : \begin{pmatrix} E & \longrightarrow & E \\ \vec{x} & \longmapsto & \vec{x} \end{pmatrix}$$

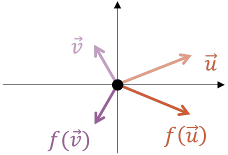
- symétrie, par exemple sur $\mathbb{R}^2$, la symétrie par rapport à l'axe des abscisses :

$$f : \begin{pmatrix} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x, y) & \longmapsto & (x, -y) \end{pmatrix}$$

Homothétie avec $c = -2$



Symétrie par rapport à l'axe des x



Exemples d'applications **non-linéaires**

- fonction affine, avec $\vec{b} \neq \vec{0}$ et $\lambda \in \mathbb{R}$:

$$f : \begin{pmatrix} E & \longrightarrow & E \\ \vec{x} & \longmapsto & c\vec{x} + \vec{b} \end{pmatrix}$$

  en effet $f(\vec{0}) = \vec{b} \neq \vec{0}$.

- fonction carré sur $\mathbb{R}$:

$$f : \begin{pmatrix} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{pmatrix}$$

- fonction norme :

$$||\cdot|| : \begin{pmatrix} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (x_1, \ldots, x_n) & \longmapsto & ||\vec{x}|| = \sqrt{x_1^2 + \cdots + x_n^2} \end{pmatrix}$$

  en effet on voit que $||-\vec{x}|| = ||\vec{x}||$ et non $-||\vec{x}||$ pour $\vec{x}$ non nul.

Représentation matricielle d'une application linéaire (1)

E and F are finite dimensional vector speces such that $\dim(E) = n$ and $\dim(F) = m$. Let $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_n)$ be a basis of E and $\mathcal{G} = (\vec{g}_1, ..., \vec{g}_n)$ be a basis of F.

Let $f : E \to F$ a linear map. Each image of the base vector of E by f has a unique decomposition in the base $\mathcal{G}$ of F can be decomposed in a unique way :

$$\left\{\begin{array}{rcl} f(\vec{e}_1) & = & y_{11}\vec{g}_1 + y_{21}\vec{g}_2 + \cdots + y_{m1}\vec{g}_m \\ f(\vec{e}_2) & = & y_{12}\vec{g}_1 + y_{22}\vec{g}_2 + \cdots + y_{m2}\vec{g}_m \\ \vdots & & \\ f(\vec{e}_n) & = & y_{1n}\vec{g}_1 + y_{2n}\vec{g}_2 + \cdots + y_{mn}\vec{g}_m \end{array}\right.$$

Représentation matricielle d'une application linéaire (2)

Définition – Représentation matricielle d'une application linéaire

La matrice représentante de l'application linéaire f de la base $\mathcal{B}$ vers la base $\mathcal{G}$, notée $\mathrm{Mat}_{\mathcal{B},\mathcal{G}}(f)$ est la matrice avec m ligne et n colonnes :

$$\mathrm{Mat}_{\mathcal{B},\mathcal{G}}(f) = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{pmatrix}$$

La i-ème colonne de la matrice est la représentation du vecteur $f(\vec{e}_i)$ dans la base $\mathcal{G}$.

Représentation matricielle d'une application linéaire (3)

Un exemple. Soit l'application linéaire définie par

$$f : \left(\begin{array}{rcl} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x, y, z) & \longmapsto & (2x + z, -x + 3y + 5z) \end{array} \right).$$

Quelle est la représentation matricielle de f dans les bases canoniques ?

Les bases canoniques de $\mathbb{R}^3$ et de $\mathbb{R}^2$ sont respectivement

- $(1, 0, 0), (0, 1, 0), (0, 0, 1)$
- $(1, 0), (0, 1)$.

Représentation matricielle d'une application linéaire (4)

$$
\begin{aligned}
f(1,0,0) &= (2,-1) = 2 \times (1,0) - 1 \times (0,1) \\
f(0,1,0) &= (0,3) = 0 \times (1,0) + 3 \times (0,1) \\
f(0,0,1) &= (1,5) = 1 \times (1,0) + 5 \times (0,1).
\end{aligned}
$$

La représentation matricielle de l'application f est la matrice

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

Calcul matriciel (1)

- $\mathcal{M}_{n,m}(\mathbb{R})$ est l'ensemble des matrices à coefficients réels à n lignes et m colonnes et on note plus simplement $\mathcal{M}_n(\mathbb{R})$ les matrices carrées de taille n
- $\mathcal{M}_{n,m}(\mathbb{R})$ est un $\mathbb{R}$-espace vectoriel de dimension $n \times m$, muni de l'addition coefficient par coefficient
- si M est une matrice, on note $(M)_{i,j}$ le coefficient sur la i-ème ligne et la j-ème colonne
- on peut multiplier des matrices si leurs tailles sont compatibles : si $A \in \mathcal{M}_{n,m}$ et $B \in \mathcal{M}_{m,p}$ alors on peut définir le produit $AB \in \mathcal{M}_{n,p}$ par la matrice dont les coefficients sont :

$$(AB)_{i,j} = \sum_{k=1}^{m} A_{i,k} B_{k,j}.$$

- le produit matriciel n'est pas commutatif, en général $AB \neq BA$ (d'ailleurs BA peut ne même pas être défini)

Calcul matriciel (2)

- la matrice identité de taille n est la matrice carrée définie par

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

  et pour toute matrice de taille compatible $AI_n = A$ et $I_nB = B$. I_n est la matrice représentative de l'application identité d'un espace à dimension n dans n'importe quelle base
- puissance d'une matrice carrée de taille n :

$$A^k = \underbrace{A \times A \times \cdots \times A}_{k \text{ fois}} \quad \text{et} \quad A^0 := I_n$$

Calcul matriciel (3)

- transposée d'une matrice : si $A \in \mathcal{M}_{n,p}$ alors A^T est la matrice dans $\mathcal{M}_{p,n}$ définie par $(A^T)_{i,j} = A_{j,i}$
- multiplication par un vecteur : vecteur colonne à droite, vecteur ligne à gauche

$$A\vec{x} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad \vec{x}^T A = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} A.$$

Utilisation du calcul matriciel pour les applications linéaires

Soit $\vec{x} \in E$ et $X = (x_1, ..., x_n)$ sa représentation dans une base $\mathcal{B} = (\vec{e}_1, ..., \vec{e}_n)$. Soit $f : E \to E$ une application linéaire dont la matrice dans la base $\mathcal{B}$ est M. Alors le vecteur MX est la représentation du vecteur $f(\vec{x})$ dans la même base.

En effet :

$$\begin{aligned} f(\vec{x}) &= f(x_1\vec{e}_1 + \cdots x_n\vec{e}_n) \\ &= x_1 f(\vec{e}_1) + \cdots + x_n f(\vec{e}_n) \end{aligned}$$

or $f(\vec{e}_j)$ s'écrit $\sum_{i=1}^n M_{ij}\vec{e}_i$

$$= \left(\sum_{k=1}^n M_{1k}x_k\right)\vec{e}_1 + \left(\sum_{k=1}^n M_{2k}x_k\right)\vec{e}_2 + \cdots + \left(\sum_{k=1}^n M_{nk}x_k\right)\vec{e}_n.$$

On reconnaît devant chaque élément $\vec{e}_i$ la i-ème ligne du vecteur colonne MX.

Inversion de matrices carrées

Soit A une matrice carrée de taille n.

Définition – Matrice inversible

A est inversible s'il existe une autre matrice B carrée de taille n telle que $AB = BA = I_n$. Dans ce cas, B est unique, est est appelée l'inverse de A, notée A^{-1}.

On remarque que I_n est inversible et est son propre inverse.

Matrice inversible (1)

Propriétés de l'inverse d'une matrice

A est supposée inversible.

- A^{-1} est inversible et $(A^{-1})^{-1} = A$.
- Pour tout $k \geq 0$, A^k est inversible et $(A^k)^{-1} = (A^{-1})^{-1}$.
- Si B est une matrice inversible de même taille, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$
- A^T est inversible et $(A^T)^{-1} = (A^{-1})^T$.
- Pour tout réel $c \neq 0$, cA est inversible et $(cA)^{-1} = \frac{1}{c}A^{-1}$.

Matrice inversible (2)

Caractérisation d'une matrice inversible

Soit A une matrice carrée de taille n. Les propriétés suivantes sont équivalentes :

- A est inversible.
- $\det A \neq 0$.
- Pour tout $Y \in \mathbb{R}^n$, le système $AX = Y$ avec X inconnue dans $\mathbb{R}^n$ admet une unique solution (dans ce cas cette solution est $X = A^{-1}Y$).
- Le système $AX = 0$ avec X inconnue dans $\mathbb{R}^n$ admet pour seule solution $X = 0$.
- 0 n'est pas valeur propre de A.
- Les colonnes de A sont linéairement indépendantes dans $\mathbb{R}^n$ *ie* elles forment une base.
- Les lignes de A sont linéairement indépendantes dans $\mathbb{R}^n$ (elles forment une base).
- L'application linéaire associée à la matrice A dans une base quelconque est bijective (alors A^{-1} est la représentation canonique de l'application réciproque f^{-1} qui est linéaire).

Calculer l'inverse d'une matrice – Cas particulier d'une matrice 2×2

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

Le déterminant de A vaut $\det A = ad - bc$. A est inversible si et seulement si $\det A \neq 0$ et dans ce cas

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

ESE Écologie Systématique Évolution

Calculer l'inverse d'une matrice – Cas général (1)

The general way to calculate the inverse of a square matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

of size n is to solve the linear system :

$$AX = Y \quad \text{with} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Calculer l'inverse d'une matrice – Cas général (2)

The explicit expression of this linear system is :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = y_2 \\ & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = y_n \end{cases}$$

For instance, the Gaussian elimination can be used to solve it. If the solution is unique, it means that this matrix is invertible. In this case the values x_i are linear compositions of the y_j :

$$\begin{cases} x_1 & = b_{11}y_1 + b_{12}y_2 + \cdots + b_{1n}y_n \\ x_2 & = b_{21}y_1 + b_{22}y_2 + \cdots + b_{2n}y_n \\ \vdots & \\ x_n & = b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nn}y_n \end{cases}$$

Calculer l'inverse d'une matrice – Cas général (3)

Then the inverse of A is the matrix formed by the coefficients of these expressions :

$$A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}.$$

ESE
Écologie Systématique Évolution

Valeurs propres et vecteurs propres

Définition – Valeur propre et vecteur propre d'une application linéaire

Soit $f : E \to E$ une application linéaire. Un **vecteur propre** est un vecteur non nul $\vec{u} \neq \vec{0}$ de E tel qu'il existe un scalaire $\lambda \in \mathbb{R}$ pour lequel :

$$f(\vec{u}) = \lambda \vec{u}. \tag{2}$$

Le nombre λ est appelé **valeur propre** associée à $\vec{u}$.

Pour une application linéaire, l'ensemble des valeurs propres de f est appelé **spectre** de cette application linéaire.

Un exemple (1)

On considère l'application linéaire « symétrie par rapport à l'axe des x » sur $\mathbb{R}^2$:

$$f : \left(\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x, y) & \longmapsto & (x, -y) \end{array} \right).$$

f a-t-elle des valeurs propres ? Si oui quels sont les vecteurs propres associés ?

Pour cela résolvons l'équation $f((x, y)) = \lambda(x, y)$ avec λ, x et y inconnues, tel que $(x, y) \neq (0, 0)$:

$$\begin{aligned} f((x, y)) = \lambda(x, y) \quad & \Leftrightarrow \quad (x, -y) = \lambda(x, y) \\ & \Leftrightarrow \quad \begin{cases} x = \lambda x & (*) \\ -y = \lambda y & (**) \end{cases} \end{aligned}$$

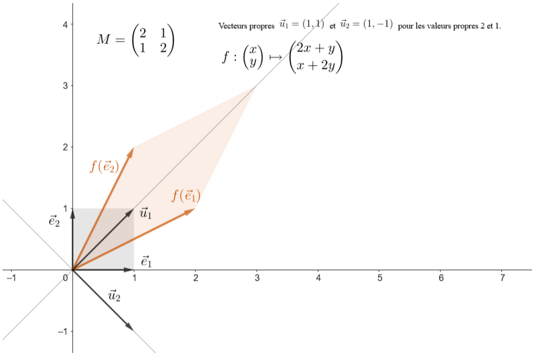
Un exemple (2)

Puisque $(x, y) \neq (0, 0)$ on a au moins une des deux conditions vérifiées : $x \neq 0$ ou $y \neq 0$. Supposons dans un premier cas que $x \neq 0$.

Cas 1. $x \neq 0$.
L'équation (*) se simplifie alors en $\lambda = 1$. Donc l'équation (**) devient $-y = y$ ce qui implique $y = 0$. On voit alors que $\lambda = 1$ est une valeur propre et que les vecteurs propres sont tous les vecteurs $(x, 0)$ avec $x \neq 0$. Par exemple, le vecteur $(1, 0)$ est un vecteur propre pour la valeur propre $\lambda = 1$.

Cas 2. $y \neq 0$.
L'équation (**) se simplifie alors en $\lambda = -1$. Donc l'équation (*) devient $x = -x$ ce qui implique $x = 0$. Donc $\lambda = -1$ est une valeur propre et ses vecteurs propres sont les vecteurs $(0, y)$ avec $y \neq 0$, par exemple le vecteur $(0, 1)$.

Interprétation géométrique



Version interactive https://www.geogebra.org/calculator/fpshctz3

Propriétés des valeurs propres et des vecteurs propres (1)

Propriétés

Soit $f : E \to E$ une application linéaire.

1. Si λ est une valeur propre et $\vec{x}$ un vecteur propre associé, alors $c\vec{x}$ avec $c \neq 0$ est aussi vecteur propre pour la même valeur propre.
2. Les vecteurs propres associés à des valeurs propres différentes sont **linéairement indépendants** (= forment une famille libre).
3. Si $\dim E = n$ alors f admet au plus n valeurs propres différentes.
4. Si λ est une valeur propre, l'ensemble des vecteurs $\vec{x}$ qui vérifient
$$f(\vec{x}) = \lambda x$$
est un sous-espace vectoriel de E appelé **espace propre** associé à la valeur propre λ.

Propriétés des valeurs propres et des vecteurs propres (2)

Démonstration –

1. $\vec{x}$ est un vecteur propre associé à la valeur propre λ et $c \neq 0$.

$$\begin{aligned} f(c\vec{x}) &= cf(\vec{x}) \quad \text{car } f \text{ est linéaire} \\ &= c\lambda\vec{x} \quad \text{car } \vec{x} \text{ est vecteur propre} \\ &= \lambda(c\vec{x}). \end{aligned}$$

   Puisque $c\vec{x} \neq \vec{0}$, $c\vec{x}$ est bien un vecteur propre pour la valeur propre λ.

2. Preuve pour 2 valeurs propres, supposons $\lambda \neq \mu$ deux valeurs propres de f avec les vecteurs propres associés $\vec{x}$ et $\vec{y}$. Supposons que $\vec{x}$ et $\vec{y}$ ne sont pas linéairement indépendants. Il existe donc a et b tels que $a\vec{x} + b\vec{y} = \vec{0}$, avec a et b non tous nuls. Puisque $\vec{x}$ et $\vec{y}$ sont non nuls, cela implique a et b non nuls. On a alors $\vec{x} = -\frac{b}{a}\vec{y}$. Appliquons f :

$$f(\vec{x}) = \lambda\vec{x}$$

Propriétés des valeurs propres et des vecteurs propres (3)

d'une part (car $\vec{x}$ est une valeur propre pour λ). D'autre part :

$$\begin{aligned} f(\vec{x}) &= f\left(-\frac{b}{a}\vec{y}\right) \\ &= -\frac{b}{a}f(\vec{y}) \quad \text{par linéarité de } f \\ &= -\frac{b}{a}\mu\vec{y} \quad \text{car } \vec{y} \text{ vecteur propre pour } \mu. \end{aligned}$$

Donc $\lambda\vec{x} = -\frac{b}{a}\mu\vec{y}$ En remplaçant $\vec{x}$ par $-b/a\vec{y}$ et simplifiant on obtient :

$$\lambda\vec{y} = \mu\vec{y} \quad \Leftrightarrow (\lambda - \mu)\vec{y} = \vec{0}.$$

Puisque $\vec{y} \neq \vec{0}$ on a alors $\lambda = \mu$ ce qui contredit l'hypothèse de départ. Donc $\vec{x}$ et $\vec{y}$ sont linéairement indépendants.
On applique le même type de raisonnement pour plus que deux vecteurs.

Propriétés des valeurs propres et des vecteurs propres (4)

3. Si on avait plus que n valeurs propres, en prenant leurs vecteurs propres on arriverait à construire une famille libre de plus que n vecteurs grâce à la propriété précédente, or dans un espace de dimension finie, la dimension est le nombre maximal de vecteurs qu'une famille libre peut compter.
4. Appelons F l'ensemble des vecteurs $\vec{x}$ tels que $f(\vec{x}) = \lambda\vec{x}$. F est non vide car F contient $\vec{0}$. La stabilité par multiplication par un scalaire non nul est vérifiée par la propriété 1. Pour le scalaire nul, on a bien $0\vec{x} = \vec{0} \in F$. La stabilité par addition utilise la linéarité de l'application f : si $\vec{x}$ et $\vec{y}$ sont dans F alors :

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) = \lambda\vec{x} + \lambda\vec{y} = \lambda(\vec{+}\vec{y})$$

donc $\vec{x} + \vec{y} \in F$. F est bien un sous espace vectoriel de E.

Valeurs propres et vecteurs propres pour les matrices (1)

Les définitions et propriétés sont exactement les mêmes que pour une application linéaire dans un espace de dimension finie, car une matrice n'est autre que la représentation d'une application linéaire dans une certaine base.

Définition – Valeur propre et vecteur propre d'une matrice

Soit M une matrice *carrée* réelle de taille n. $\lambda \in \mathbb{R}$ est une **valeur** propre de cette matrice s'il existe un vecteur colonne $X \in \mathbb{R}^n$ non nul tel que

$$MX = \lambda X.$$

X est appelé alors **vecteur propre** de la matrice M pour la valeur propre λ.

Valeurs propres et vecteurs propres pour les matrices (2)

Propriétés

1. $X \neq 0$ est un vecteur propre de M pour la valeur propre $\lambda \Leftrightarrow (M - \lambda I_n)X = 0$.
2. λ est une valeur propre de $M \Leftrightarrow \det(M - \lambda I_n) = 0$.
3. M admet au plus n valeurs propres différentes.
4. L'espace propre associé à la valeur propre λ, noté $\ker(M - \lambda I_n)$ est un espace vectoriel. Les espaces propres associés à des valeurs propres différents sont indépendants.

La fonction définie sur $\mathbb{R}$ par $\chi_M(\lambda) = \det(M - \lambda I_n)$ est un polynôme dont les racines sont les valeurs propres de M. On appelle cette fonction le **polynôme caractéristique** de M.

Diagonalisation (1)

Définition – Matrice diagonalisable

Soit M une matrice. On dit que M est diagonalisable si il existe une base de $\mathbb{R}^n$ composé de vecteurs propres de la matrice M.

Diagonalisation (2)

Théorème – Définition équivalente

M est diagonalisable si et seulement si il existe une matrice P inversible de taille n et n réels $\lambda_1, \dots, \lambda_n$ tels que

$$M = P \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} P^{-1}.$$

Les λ_i sont les valeurs propres de M et les colonnes de P forment une base de vecteurs propres dans le même ordre.

Diagonalisation (3)

Propriété

M est est diagonalisable si et seulement si la somme des dimensions de ses espaces propres vaut n.

Corollaire

Si M a n valeurs propres différentes, alors M est diagonalisable.

Démonstration –
Si M a n valeurs propres différentes, alors n vecteurs propres associés à ces valeurs propres forment une famille libre, donc une base. Donc M est diagonalisable. □

Diagonalisation (4)

Propriétés

Si M est diagonalisable alors les puissances de M se calculent :

$$M^k = P \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix} P^{-1}$$

et si tous les λ_i sont non nuls, alors M est inversible et

$$M^{-1} = P \begin{pmatrix} 1/\lambda_1 & 0 & \dots & 0 \\ 0 & 1/\lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\lambda_n \end{pmatrix} P^{-1}$$

Exemple de diagonalisation (1)

Diagonaliser une matrice, c'est trouver une base de directions de l'espace qui est stable par l'application de la matrice.

Comment diagonaliser une matrice ?

1. Résoudre l'équation $\det(M - \lambda I_n) = 0$ pour trouver les valeurs propres.
2. Pour chacune des valeurs propres, résoudre le système $MX = \lambda X$. La dimension de l'espace propre est le nombre de vecteurs linéairement indépendants.
3. La matrice M est diagonalisable si on arrive à trouver n vecteurs propres indépendants.
4. Si c'est le cas, la matrice P se construit en accolant ces vecteurs propres en colonne. Calculer son inverse P^{-1}.

Exemple de diagonalisation (2)

Exemple 1. La matrice suivante est-elle diagonalisable ?

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Le polynôme caractéristique est

$$\begin{aligned} \chi_M(\lambda) &= \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \\ &= (2-\lambda)^2 - 1 \\ &= (1-\lambda)(3-\lambda). \end{aligned}$$

Ce polynôme caractéristique admet exactement deux racines distinctes 1 et 3 donc la matrice a 2 valeurs propres. Elle est donc diagonalisable. Trouvons ses vecteurs propres.

Exemple de diagonalisation (3)

- Vecteur propre pour $\lambda = 1$. Résolvons le système $MX = X$ avec $X \in \mathbb{R}^2$. En posant $X = \begin{pmatrix} x \\ y \end{pmatrix}$ on a le système

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} 2x + y & = x \\ x + 2y & = x \end{cases}$$

$$\Leftrightarrow \quad x + y = 0$$

$$\Leftrightarrow \quad x = -y$$

L'espace propre est donc l'ensemble des vecteurs $\begin{pmatrix} x \\ -x \end{pmatrix}$ avec $x \in \mathbb{R}$. Sa dimension est 1. Par exemple le vecteur $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est un vecteur propre pour la valeur propre 1.

Exemple de diagonalisation (4)

- Vecteur propre pour $\lambda = 3$. Résolvons le système $MX = 3X$.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} 2x + y = 3x \\ x + 2y = 3y \end{cases}$$
$$\Leftrightarrow x = y.$$

L'espace propre est donc l'ensemble des vecteurs $\begin{pmatrix} x \\ x \end{pmatrix}$ avec $x \in \mathbb{R}$. Sa dimension est 1. Par exemple le vecteur $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est un vecteur propre pour la valeur propre 3.

Exemple de diagonalisation (5)

On voit que les vecteurs propres sont bien indépendants, on a donc une base de $\mathbb{R}^2$:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

La matrice P est donc :

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

et son inverse peut être calculé aisément :

$$P = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Exemple de diagonalisation (6)

D'où finalement la diagonalisation :

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

Exemple 2. La matrice suivante est-elle diagonalisable ?

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

Le polynôme caractéristique est

Exemple de diagonalisation (7)

$$\begin{aligned}\chi_M(\lambda) &= \det\begin{pmatrix}-\lambda & 1\\ 0 & -\lambda\end{pmatrix}\\ &= \lambda^2\end{aligned}$$

Sa seule racine est 0, donc la seule valeur propre de M est 0. Si M était diagonalisable, alors on aurait

$$M = P\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}P^{-1}$$

et donc M serait la matrice nulle. Donc M n'est pas diagonalisable.

Exemple de diagonalisation (8)

Sinon, on aurait pu également calculer l'espace propre associé à la valeur propre 0. Pour cela, il faut résoudre le système $MX = 0$. En posant $X = \begin{pmatrix} x \\ y \end{pmatrix}$, cela équivaut à l'équation $y = 0$. L'espace propre est donc l'ensemble des vecteurs

$$\begin{pmatrix} x \\ 0 \end{pmatrix}$$

avec $x \in \mathbb{R}$. Cet espace est de dimension $1 < 2$ donc M n'est pas diagonalisable.

Variance et covariance (1)

Définition – Variance d'une variable aléatoire

Soit X une variable aléatoire. On introduit sa variance comme la grandeur :

$$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

où $\mathbb{E}$ désigne l'espérance (= la moyenne).

La variance est la moyenne des carrés des écarts à la moyenne.

La variance d'une variable aléatoire est toujours positive et son écart-type est la racine carrée de la variance : $\sigma_X = \sqrt{V(X)}$.

Statistiquement, la variance peut être estimée à partir de l'estimateur non-biaisé de la variance :

$$V(X) \approx \frac{1}{n-1}\left((x_1 - \mu)^2 + \cdots + (x_n - \mu)^2\right)$$

où $\mu = (x_1 + \cdots + x_n)/n$ est la moyenne empirique.

Variance et covariance (2)

Définition – Covariance de deux variables aléatoires

Soit X et Y deux variables aléatoires. On appelle **covariance de X et de Y** le nombre réel :

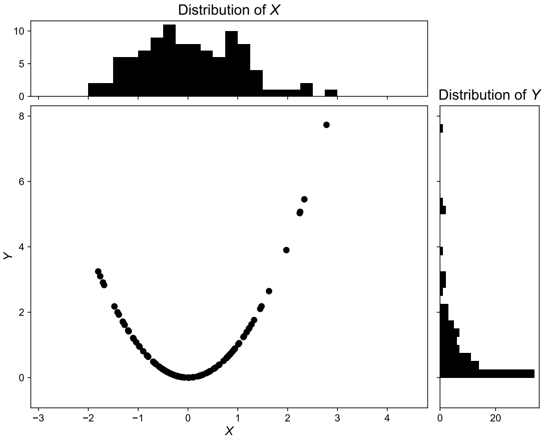
$$\mathrm{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right].$$

Si X et Y sont indépendantes, alors $\mathrm{Cov}(X, Y) = 0$.

Attention.

La réciproque est fausse, deux variables peuvent avoir une covariance nulle et pourtant être dépendantes. La covariance est seulement une mesure de la corrélation linéaire. Par exemple, la variable X tirée selon une loi normale centrée et $Y = X^2$ ont une covariance nulle pourtant elles ne sont pas indépendantes. Elles sont seulement **linéairement décorrélées**.

Variance et covariance (3)



Variance et covariance (4)

Lien avec le reste des statistiques

Le coefficient de corrélation entre deux variables, souvent noté r mesure leur lien linéaire. Il est défini par :

$$r_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

où σ_X et σ_Y sont les écart-types des variables X et Y. Lors d'une régression linéaire, le coefficient de détermination R^2 n'est autre que le carré de $r_{X,Y}$.

Lors d'une régression linéaire `Y ~ aX + b`, la pente estimée est :

$$a = \frac{\mathrm{Cov}(X,Y)}{V(X)}.$$

Matrice de covariance

Soit $X_1, \ldots, X_n$, n variables aléatoires qui peuvent être dépendantes. Par exemple, on prend un individu au hasard dans la population et on mesure son poids, sa taille, son âge, sa tension artérielle, ... Ce sont n grandeurs que l'on peut mesurer sur une observation.

Définition – Matrice de covariance

On définit la matrice de covariance comme la matrice carrée de taille n suivante

$$C = \begin{pmatrix} V(X_1) & \mathrm{Cov}(X_1, X_2) & \ldots & \mathrm{Cov}(X_1, X_n) \\ \mathrm{Cov}(X_2, X_1) & V(X_2) & \ldots & \mathrm{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n, X_1) & \mathrm{Cov}(X_n, X_2) & \ldots & V(X_n) \end{pmatrix}$$

Soit Z une variable construite comme combinaison linéaire des X_i : $Z = u_1 X_1 + \cdots + u_n X_n$. En notant $U = (u_1, \ldots, u_n)$ on a alors :

$$V(Z) = U^T C U.$$

Théorème spectral (1)

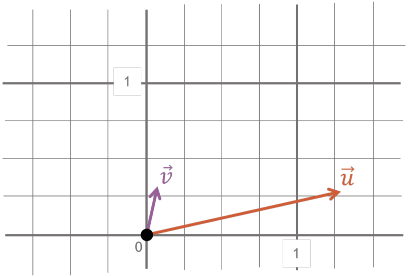
Définition – Matrice symétrique

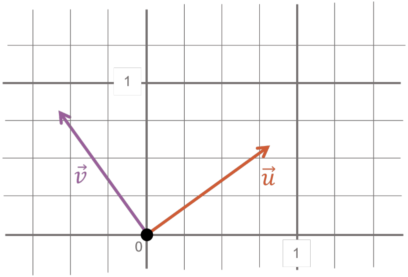
Une matrice carrée de taille n est dite symétrique si elle est égale à sa transposée, c'est-à-dire pour tout i, j, $M_{i,j} = M_{j,i}$.

Définition – Produit scalaire et base orthonormée

- On définit le **produit scalaire** de deux vecteurs $X = (x_1, \ldots, x_n)$ et $Y = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$ comme le nombre $X^T Y$, c'est-à-dire $x_1 y_1 + \cdots + x_n y_n$.
- Deux vecteurs sont **orthogonaux** si leur produit scalaire est nul. Un vecteur est **unitaire** (de longueur 1) si $X^T X = 1$.
- Une base $\mathcal{B}$ de $\mathbb{R}^n$ est dire **orthonormale** si tous ses vecteurs sont unitaires et orthogonaux entre eux.

Théorème spectral (2)





$\vec{u}$ et $\vec{v}$ forment une base
non orthonormée

$\vec{u}$ et $\vec{v}$ forment une base
orthonormée

Théorème spectral (3)

Théorème – Théorème spectral

Soit M une matrice réelle symétrique. Il existe une base $\mathcal{B}$ de vecteurs orthogonaux et de longueur 1 qui diagonalise la matrice M. En pratique, cela implique que si $\mathcal{B} = \{V_1, ..., V_n\}$ alors $V_i^T V_i = 1$ et $V_i^T V_j = 0$ pour tous $i \neq j$.

Analyse en composantes principales (1)

La matrice de covariance étant une matrice symétrique, le théorème spectral s'applique. Il existe donc une base $\mathcal{B} = \{\vec{v}_1, ..., \vec{v}_n\}$ de vecteurs orthonormaux tels que dans cette base, la matrice de covariance est diagonale. On peut montrer (...) que ces valeurs propres sont strictement positives. En notant $\lambda_1, ..., \lambda_n$ les valeurs propres rangées dans l'ordre décroissant : $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$, on a donc

$$C\vec{v}_i = \lambda_i \vec{v}_i.$$

Introduisons maintenant n nouvelles variables aléatoires Y_i définies ainsi : pour $i = 1, ..., n$, on a

$$V_i := \sum_{j=1}^{n} v_{i,j} X_j = \vec{v}_i^{\,T} X_j$$

où $v_{i,j}$ est la j-ème coordonnée du vecteur $\vec{v}_i$.

ESE Écologie Systématique Évolution

Analyse en composantes principales (2)

Grâce au résultat sur la matrice de covariance, on peut aisément calculer la variance de Y_i :

$$\begin{aligned}
V(Y_i) &= \vec{v}_i^{\,T} C \vec{v}_i \\
&= \vec{v}_i^{\,T} \lambda_i \vec{v}_i \\
&= \lambda_i \vec{v}_i^{\,T} \vec{v}_i \\
&= \lambda_i \quad \text{car } \vec{v}_i^{\,T} v_i = 1.
\end{aligned}$$

Analyse en composantes principales (3)

On peut également calculer la covariance de deux variables Y_i, Y_j avec $i \neq j$:

$$\begin{aligned}
\mathrm{Cov}(Y_i, Y_j) &= \vec{v}_i^T C \vec{v}_j \\
&= \vec{v}_i^T \lambda_j \vec{v}_j \\
&= \lambda_j \vec{v}_i^T \vec{v}_j \\
&= 0 \quad \text{car } \vec{v}_i^T \vec{v}_j = 0.
\end{aligned}$$

Les variables Y_i sont donc linéairement décorrélées et de variances décroissantes $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

La transformation que l'on vient de faire $Y_i = \sum_{j=1}^{n} v_{i,j} X_i$ est une transformation linéaire des X_i pour trouver des nouvelles variables selon des directions définies par les $\vec{v}_i$. La direction

Analyse en composantes principales (4)

$\vec{v}_1$ est la direction de Y_1 : c'est la combinaison linéaire des $X_1, ..., X_n$ qui **maximise la variance**. C'est ce qu'on appelle le premier axe de l'analyse en composantes principales. Le pourcentage de variance expliqué est donné par $\lambda_1/(\lambda_1 + \cdots + \lambda_n)$.

Si l'on devait résumer au mieux une observation $(x_1, ..., x_n)$ avec un seul nombre, la position de cette observation le long de ce premier axe est celui qui nous donnerait le plus d'information.

Les axes suivants sont les combinaisons linéaires indépendantes de la première qui expliquent le reste de la variance.

Remarque –
Si certaines des valeurs propres sont nulles (par exemple $\lambda_n = 0$) cela signifie que seules $n-1$ combinaisons linéaires des $X_1, ..., X_n$ suffisent à décrire 100% de la variance de ces

Analyse en composantes principales (5)

variables. Il existe une des mesures X_k qui est totalement déterminée par une combinaison linéaire des autres :

$$X_k = \sum_{j \neq k} \alpha_j X_j.$$