Algèbre linéaire
Remise à niveau - M1 BEE@Lyon

Pr. Sandrine CHARLES
sandrine.charles@univ-lyon1.fr
Université Claude Bernard Lyon 1 – France

D'après un document de T. Moreau et M. Chavance
http://cesp.vjf.inserm.fr/M2SPR/pdf/polmat02009.pdf

Avec les contributions de I. Amat & C. Bajard

1er octobre 2019

Table des matières

Plan détaillé

Définition d'une matrice

Une *matrice* $\mathbf{A}(n,p)$ est un tableau rectangulaire de nombres réels comprenant n lignes et p colonnes.
Chaque nombre de ce tableau est un *coefficient* (ou *élément*) de $\mathbf{A}$.
Les valeurs de n et p sont les *dimensions* de $\mathbf{A}$.

$$\mathbf{A}\left(n,p\right) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2p} \\ \vdots & \vdots & & & & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{ip} \\ \vdots & \vdots & & & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nj} & \ldots & a_{np} \end{pmatrix} = \left\{a_{ij}\right\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$$

Les p coefficients $(a_{i1}, a_{i2}, \ldots, a_{ip})$ forment la i^{e} ligne de $\mathbf{A}$: $\left\{a_{ij}\right\}_{j\in[1;p]}$
Les n coefficients $(a_{1j}, a_{2j}, \ldots, a_{nj})$ forment la j^{e} colonne de $\mathbf{A}$: $\left\{a_{ij}\right\}_{i\in[1;n]}$

Exemple : tableau floro-faunistique

Situation : n sites sont décrits par l'abondance ou par un indice de présence-absence de p espèces. L'information est contenue dans un tableau $\mathbf{A}(n,p)$. Ici, $n = 10$ et $p = 8$.

	Cogo	Satr	Phph	Neba	Thth	Teso	Chna	Chto
1	0	3	0	0	0	0	0	0
2	0	5	4	3	0	0	0	0
3	0	5	5	5	0	0	0	0
4	0	4	5	5	0	0	0	0
5	0	2	3	2	0	0	0	0
6	0	3	4	5	0	0	0	0
7	0	5	4	5	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	1	3	0	0	0	0
10	0	1	4	4	0	0	0	0

Autres définitions

- Un nombre est une matrice $(1, 1)$.
- La matrice $\mathbf{X} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$ de dimension $(1, 3)$ est appelée *matrice ligne* (ou *vecteur ligne*).
- La matrice $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ de dimension $(3, 1)$ est appelée *matrice colonne* (ou *vecteur colonne*).

Matrice transposée

La matrice transposée de $\mathbf{A}_{(n,p)}$ est la matrice de dimensions (p, n) dont la j$^{\text{e}}$ ligne est la j$^{\text{e}}$ colonne de $\mathbf{A}$ $(j = 1, \ldots, p)$, et dont la i$^{\text{e}}$ colonne est la i$^{\text{e}}$ ligne de $\mathbf{A}$ $(i = 1, \ldots, n)$.

On la note $\mathbf{A}^t$ ou $\mathbf{A}'$ ou $\mathbf{A}^T$.

On a par ailleurs que $\left(\mathbf{A}^T\right)^T = \mathbf{A}$.

Exemple :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \Leftrightarrow \mathbf{A}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Matrices carrées

Une matrice est dite *carrée* si $n = p$.

Exemple de matrice carrée de dimension 3 :

$$\mathbf{A}_{(3,3)} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Remarque : la transposée d'une matrice carrée est aussi une matrice carrée, de même dimension.

Les coefficients $\{a_{ii}\}$ forment la *diagonale* de $\mathbf{A}$.

$$Diag\,(\mathbf{A}) = \{1, 5, 9\}$$

Matrices carrées particulières

Matrice diagonale $\forall i\,,\forall j\,,i \neq j\,,a_{ij} = 0$.

Notation : $Diag\{a_{ii}\}_{1\leqslant i\leqslant n}$.

Exemple : $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$

Matrice symétrique $\forall i\,,\forall j\,,a_{ij} = a_{ji}$. Alors $\mathbf{A}^T = \mathbf{A}$.

Exemple : $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 9 \end{pmatrix}$

Matrice identité (ou unité) matrice diagonale telle que $\forall i\,,a_{ii} = 1$

Exemple : $\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Matrice de variance-covariance

Soient $(X_1, X_2, \ldots, X_n)$ n variables aléatoires telles que $var(X_i) = \sigma_i^2$ et $cov(X_i, X_j) = \sigma_{ij}$.

La matrice de variance-covariance des $\{X_i\}$ est la matrice *carrée symétrique* de dimension n suivante :

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_n^2 \end{pmatrix}$$

Matrice de corrélation

Si on désigne par ρ_{ij} les coefficients de corrélation entres les variables X_i et X_j, $\rho_{ij} = corr(X_i, X_j) = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$, alors la matrice de corrélation des $\{X_i\}$ est la matrice *carrée symétrique* suivante :

$$\Omega = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{pmatrix}$$

Si toutes les variables sont indépendantes deux à deux, alors $\forall i \,, \forall j \,, i \neq j, \sigma_{ij} = 0$. Conséquence : Σ est diagonale et $\Omega = \mathbf{I}_n$.

Plan détaillé

Multiplication par un nombre

Le produit d'une matrice $\mathbf{A}_{(n,p)}$ par un nombre λ est une matrice de *même dimension* $\mathbf{B}_{(n,p)}$ obtenue en multipliant par λ tous les coefficients de $\mathbf{A}$:

$$b_{ij} = \lambda a_{ij}$$

Exemple :

$$2\mathbf{A} = 2\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{pmatrix}$$

Addition

La somme $\mathbf{C} = \mathbf{A} + \mathbf{B}$ de deux matrices *de même dimension* s'obtient en additionnant les coefficients de mêmes indices :

$$c_{ij} = a_{ij} + b_{ij}$$

Exemple :

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \\ 15 & 18 \end{pmatrix}$$

Produit de deux matrices

Pour pouvoir multiplier la matrice $\mathbf{A}$ par la matrice $\mathbf{B}$, il faut que **le nombre de colonnes de $\mathbf{A}$ soit égal au nombre de lignes de $\mathbf{B}$**. Autrement dit, le produit $\mathbf{AB}$ ne sera possible que si $\mathbf{A}_{(n,p)}$ et $\mathbf{B}_{(p,q)}$. Le résultat est alors une matrice $\mathbf{C}_{(n,q)}$ dont les coefficients c_{ij} valent :

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$$

$$\begin{array}{ccccc} \mathbf{C} & = & \mathbf{A} & \times & \mathbf{B} \\ (n,q) & & (n,p) & & (p,q) \end{array}$$

Le coefficient c_{ij} est obtenu en multipliant deux à deux les coefficients de la i^{e} ligne de $\mathbf{A}$ avec ceux de la j^{e} colonne de $\mathbf{B}$.

Produit de deux matrices : exemple

$$\mathbf{B}_{(2,4)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

$$\mathbf{A}_{(3,2)} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} \boxed{11} & 14 & 17 & 20 \\ 23 & 30 & 37 & 44 \\ 35 & 46 & 57 & 68 \end{pmatrix} = \mathbf{C}_{(3,4)}$$

Produit de deux matrices : exemple

$$\mathbf{B}_{(2,4)} = \begin{pmatrix} 1 & \boxed{2} & 3 & 4 \\ 5 & \boxed{6} & 7 & 8 \end{pmatrix}$$

$$\mathbf{A}_{(3,2)} = \begin{pmatrix} \boxed{1 \quad 2} \\ 3 \quad 4 \\ 5 \quad 6 \end{pmatrix} \begin{pmatrix} 11 & \boxed{14} & 17 & 20 \\ 23 & 30 & 37 & 44 \\ 35 & 46 & 57 & 68 \end{pmatrix} = \mathbf{C}_{(3,4)}$$

Produit de deux matrices : exemple

$$\mathbf{B}_{(2,4)} = \begin{pmatrix} 1 & \boxed{2} & 3 & 4 \\ 5 & \boxed{6} & 7 & 8 \end{pmatrix}$$

$$\mathbf{A}_{(3,2)} = \begin{pmatrix} 1 & 2 \\ \boxed{3 \quad 4} \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 11 & 14 & 17 & 20 \\ 23 & \boxed{30} & 37 & 44 \\ 35 & 46 & 57 & 68 \end{pmatrix} = \mathbf{C}_{(3,4)}$$

Produit de deux matrices : exemple

$$\mathbf{B}_{(2,4)} = \begin{pmatrix} 1 & 2 & 3 & \boxed{4} \\ 5 & 6 & 7 & \boxed{8} \end{pmatrix}$$

$$\mathbf{A}_{(3,2)} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ \boxed{5} & \boxed{6} \end{pmatrix} \begin{pmatrix} 11 & 14 & 17 & 20 \\ 23 & 30 & 37 & 44 \\ 35 & 46 & 57 & \boxed{68} \end{pmatrix} = \mathbf{C}_{(3,4)}$$

Propriétés du produit de deux matrices

Non commutativité En général, $\mathbf{A} \times \mathbf{B} \neq \mathbf{B} \times \mathbf{A}$

Associativité $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{A} \times \mathbf{B} \times \mathbf{C}$

Pour une matrice carrée de dimension n, on peut alors définir $\mathbf{A} \times \mathbf{A} \times \mathbf{A} \times \mathbf{A} \times \ldots = \mathbf{A}^k$

Distributivité $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$

Multiplication par un nombre $(\lambda \mathbf{A}) \times \mathbf{B} = \lambda (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \times (\lambda \mathbf{B})$

Produit transposé $(\mathbf{A} \times \mathbf{B})^T = \mathbf{B}^T \times \mathbf{A}^T$

Matrice identité $\mathbf{I}_n \times \mathbf{A} = \mathbf{A} \times \mathbf{I}_n = \mathbf{A}$

Plan détaillé

Trace d'une matrice carrée

La trace d'une matrice carrée $\mathbf{A}_{(n,n)}$, notée $tr(\mathbf{A})$, est la somme des coefficients diagonaux :

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

Exemple :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \Rightarrow tr\left(\mathbf{A}\right) = 1 + 5 + 9 = 15$$

Déterminant d'une matrice carrée

Ordre 1

$$\left| a_{11} \right| = a_{11}$$

Ordre 2

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{12}\,a_{21}$$

Ordre 3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

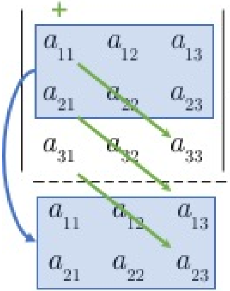
$$= a_{11}\,a_{22}\,a_{33} + a_{21}\,a_{32}\,a_{13} + a_{31}\,a_{12}\,a_{23}$$
$$- a_{13}\,a_{22}\,a_{31} - a_{23}\,a_{32}\,a_{11} - a_{33}\,a_{12}\,a_{21}$$

Déterminant d'ordre 3 : *règle de Sarrus*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}$$

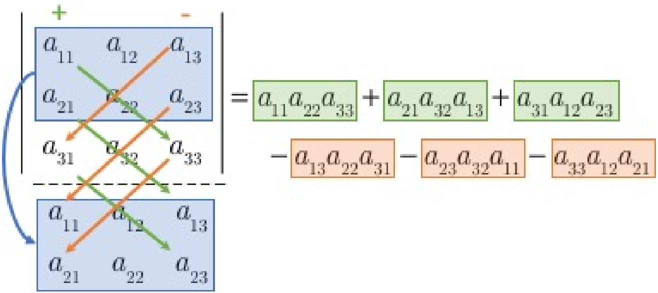
$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}$$

Déterminant d'ordre 3 : *règle de Sarrus*



$$= \boxed{a_{11}a_{22}a_{33}} + \boxed{a_{21}a_{32}a_{13}} + \boxed{a_{31}a_{12}a_{23}}$$
$$- a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}$$

Déterminant d'ordre 3 : *règle de Sarrus*



Déterminant d'ordre 3 : les cofacteurs

On choisit une ligne ou une colonne puis on somme les cofacteurs.

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Par rapport à la première ligne :

$$\Delta = +a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Par rapport à la première colonne :

$$\Delta = +a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Méthode des cofacteurs : généralisation

Le déterminant d'une matrice $\mathbf{A}_{(n,n)}$ s'obtient en choisissant une ligne ou une colonne quelconque (en pratique, avec le plus de coefficients égaux à 0), et en faisant la somme du produit de ses éléments par leur cofacteur.

Le cofacteur de a_{ij} est égal au produit de $(-1)^{i+j}$ et du déterminant extrait de $\mathbf{A}$ en supprimant la ligne i et la colonne j.

Exemple :

$$\Delta = \begin{vmatrix} 1 & 0 & 2 \\ 4 & 1 & 0 \\ 0 & 3 & 3 \end{vmatrix} = +1 \begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0 \begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix} = 3 - 0 + 24 = 27$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} 1 & 0 & 2 \\ 4 & 1 & 0 \\ 0 & 3 & 3 \end{vmatrix} = +1 \begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0 \begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} +1 & -0 & +2 \\ -4 & +1 & -0 \\ +0 & -3 & +3 \end{vmatrix} = +1 \begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0 \begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} +1 & -0 & +2 \\ -4 & +1 & -0 \\ +0 & -3 & +3 \end{vmatrix} = +1 \begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0 \begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} +1 & -0 & +2 \\ -4 & +1 & -0 \\ +0 & -3 & +3 \end{vmatrix} = +1\begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0\begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2\begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} {}^{+}1 & {}^{-}0 & {}^{+}2 \\ {}^{-}4 & {}^{+}1 & {}^{-}0 \\ {}^{+}0 & {}^{-}3 & {}^{+}3 \end{vmatrix} = +1\begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} \boxed{-0\begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix}} +2\begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Méthode des cofacteurs : détails

$$\Delta = \begin{vmatrix} 1 & 0 & 2 \\ 4 & 1 & 0 \\ 0 & 3 & 3 \end{vmatrix} = +1\begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} - 0\begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} + 2\begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix}$$

Propriétés du déterminant (1)

- Si une colonne de $\mathbf{A}$ est nulle, alors $\det(\mathbf{A}) = 0$
- Si une ligne de $\mathbf{A}$ est nulle, alors $\det(\mathbf{A}) = 0$
- Si deux lignes ou deux colonnes de $\mathbf{A}$ sont proportionnelles, alors $\det(\mathbf{A}) = 0$
- Si une ligne (resp. une colonne) de $\mathbf{A}$ est une combinaison linéaire des autres lignes (resp. colonnes), alors $\det(\mathbf{A}) = 0$:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{pmatrix}$$

Propriétés du déterminant (2)

- $\det(\mathbf{A} \times \mathbf{B}) = \det(\mathbf{A}) \times \det(\mathbf{B})$
- $\det(\mathbf{A}^T) = \det(\mathbf{A})$
- $\det\left(Diag\{a_{ii}\}_{1 \leqslant i \leqslant n}\right) = \prod_{i=1}^{n} a_{ii}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \det(\mathbf{A}) = 4$$

Plan détaillé

Définitions

- Une matrice carrée $\mathbf{A}$ de dimension n est dite *régulière* si $\det(\mathbf{A}) \neq 0$.
- L'inverse d'une matrice carrée $\mathbf{A}$ de dimension n existe si $\mathbf{A}$ est régulière.
- Si son inverse existe, alors est dite *inversible*.
- L'inverse est une matrice carrée de dimension n notée $\mathbf{A}^{-1}$ telle que :

$$\mathbf{A} \times \mathbf{A}^{-1} = \mathbf{A}^{-1} \times \mathbf{A} = \mathbf{I}_n$$

Propriétés des inverses

- Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices de dimension n :

$$(\mathbf{A} \times \mathbf{B})^{-1} = \mathbf{B}^{-1} \times \mathbf{A} - 1$$

- Et en général, soient p matrices carrées de dimension n :

$$(\mathbf{A}_1 \times \mathbf{A}_2 \times \ldots \times \mathbf{A}_p)^{-1} = \mathbf{A}_p^{-1} \times \mathbf{A}_{p-1}^{-1} \times \ldots \times \mathbf{A}_1^{-1}$$

- Inverse de la transposée

$$\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$$

- Déterminant de l'inverse

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det\left(\mathbf{A}\right)}$$

Calcul pratique de la matrice inverse

Pour obtenir l'inverse d'une matrice, on peut utiliser la formule suivante :

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \times \mathbf{C}^T$$

avec $\mathbf{C}$ la *comatrice* de $\mathbf{A}$, obtenue en remplaçant chaque coefficient de $\mathbf{A}$ par son cofacteur.

Application aux matrices de dimension 2

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

On a $\det(\mathbf{A}) = ad - bc$, que l'on va supposer $\neq 0$.

La matrice des cofacteurs s'écrit :

$$\mathbf{C} = \begin{pmatrix} +d & -c \\ -b & +a \end{pmatrix}$$

Et on obtient donc l'inverse :

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

En pratique, il existe de nombreux logiciels permettant de calculer les déterminants et les inverses de matrices, on ne fait donc généralement pas ces calculs à la main !

Application à la résolution d'un système linéaire (1)

Soit à résoudre

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a2_{1n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}$$

où les x_i sont les inconnues.
Ce système s'écrit sous forme matricielle :

$$\mathbf{A} \times \mathbf{X} = \mathbf{B}$$

avec $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$, $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $\mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

Si $\det(\mathbf{A}) \neq 0$, alors il existe une *unique* solution $\mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B}$.

Application à la résolution d'un système linéaire (21)

Soit à résoudre :

$$\mathbf{A} \times \mathbf{X} = \mathbf{B}$$

Si $\det(\mathbf{A}) \neq 0$, alors il existe une *unique* solution :

$$\mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B}$$

$$\begin{aligned}
&\mathbf{AX} = \mathbf{B} \Leftrightarrow \mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{B} \quad \text{si} \quad \det(\mathbf{A}) \neq 0 \\
&\Leftrightarrow \mathbf{IX} = \mathbf{A}^{-1}\mathbf{B} \quad \text{car} \quad \mathbf{A^{-1}A} = \mathbf{I} \\
&\Leftrightarrow \mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \quad \text{l'unique solution}
\end{aligned}$$

Table des matières

Plan détaillé

L'espace vectoriel $\mathbb{R}^2$

Soit l'espace à deux dimensions $\mathbb{R}^2$, muni d'un repère *orthonormé* défini par les deux vecteurs unité $\vec{e}_1$ (communément noté $\vec{i}$) et $\vec{e}_2$ (ou $\vec{j}$).

Chaque point $M \in \mathbb{R}^2$ est caractérisé par ses coordonnées (x_m, y_m) qui sont aussi les composantes du vecteur $\overrightarrow{OM}$, joignant l'origine au point M.

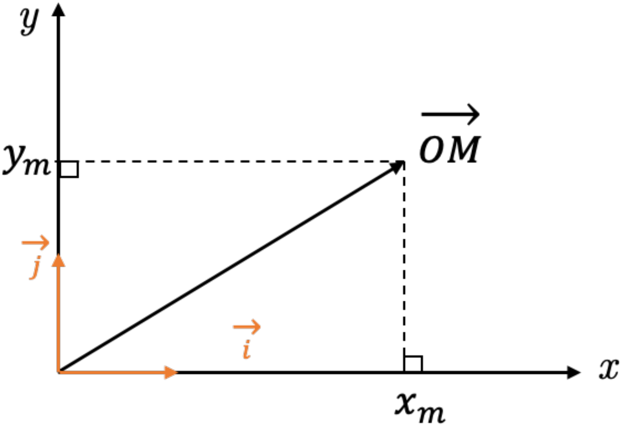
On peut écrire $\overrightarrow{OM} = x_m \vec{e}_1 + y_m \vec{e}_2$.
Les vecteurs $\vec{e}_1$ et $\vec{e}_2$ sont souvent notés $\vec{i}$ et $\vec{j}$.

Sous forme matricielle, on écrira :

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \mathbf{M} = \begin{pmatrix} x_m \\ y_m \end{pmatrix} = x_m \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y_m \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

L'espace vectoriel $\mathbb{R}^2$

Visualisation de la base canonique $(\vec{e}_1, \vec{e}_2) = (\vec{i}, \vec{j})$



Généralisation à $\mathbb{R}^n$

Les vecteurs unité suivants :

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

sont tels qu'on ne peut en exprimer aucun en fonction des autres : on dit qu'ils sont *linéairement indépendants* et qu'ils forment la **base canonique** de $\mathbb{R}^n$.
Ainsi, tout point de $\mathbb{R}^n$ vérifie :

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n = \sum_{i=1}^{n} x_i\mathbf{e}_i$$

Plus généralement, tout système de n vecteurs linéairement indépendants forme une **base** de $\mathbb{R}^n$ et tout $X \in \mathbb{R}^n$ peut s'écrire comme une **combinaison linéaire** des vecteurs de cette base.

Plan détaillé

Définition

Soient $\mathbf{X}_{(p,1)}$ un vecteur de $\mathbb{R}^p$ et $\mathbf{A}_{(n,p)}$.

Le produit $\mathbf{AX}$ est un vecteur de $\mathbb{R}^n$ dont chacune des coordonnées est une *combinaison linéaire* des coordonnées de $\mathbf{X}$.

Par exemple, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix}$

Ainsi, on appelle **application linéaire** la transformation f qui à $\mathbf{X}$ fait correspondre $\mathbf{Y} = \mathbf{AX}$:

$$\begin{array}{rccc} f: & \mathbb{R}^p & \to & \mathbb{R}^n \\ & \mathbf{X} & \mapsto & f(\mathbf{X}) = \mathbf{AX} \end{array}$$

Propriétés des applications linéaires

- $f\left(\mathbf{X}_1 + \mathbf{X}_2\right) = f\left(\mathbf{X}_1\right) + f\left(\mathbf{X}_2\right) \Leftrightarrow \mathbf{A}\left(\mathbf{X}_1 + \mathbf{X}_2\right) = \mathbf{A}\mathbf{X}_1 + \mathbf{A}\mathbf{X}_2$

- $f\left(\lambda \mathbf{X}_1\right) = \lambda f\left(\mathbf{X}_1\right) \Leftrightarrow \mathbf{A}\left(\lambda \mathbf{X}_1\right) = \lambda \mathbf{A}\mathbf{X}_1$

Plan détaillé

Définition dans $\mathbb{R}^2$

On appelle *produit scalaire canonique* de deux vecteurs $\vec{x}$ et $\vec{y}$ de $\mathbb{R}^2$, le **nombre réel** noté $(\vec{x}|\vec{y})$ ou $\langle \vec{x} \mid \vec{y} \rangle$ ou $\vec{x}.\vec{y}$ et défini par :

$$\vec{x}.\vec{y} = x_1 y_1 + x_2 y_2$$

avec $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ les matrices de coordonnées des vecteurs $\vec{x}$ et $\vec{y}$.

En notation matricielle, le produit scalaire s'écrira :

$$\vec{x}.\vec{y} = \mathbf{X}^T \mathbf{Y}$$

Généralisation à $\mathbb{R}^n$

Le produit scalaire de deux vecteurs de coordonnées $\mathbf{X} = \{x_i\}_{1 \leqslant i \leqslant n}$ et $\mathbf{Y} = \{y_i\}_{1 \leqslant i \leqslant n}$ dans $\mathbb{R}^n$ est défini par :

$$\vec{x}.\vec{y} = \sum_{i=1}^{n} x_i y_i = \mathbf{X}^T \mathbf{Y}$$

Exemples :

$$\vec{x}.\vec{y} = \mathbf{X}^T \mathbf{Y} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 1 \times 3 + 2 \times 4 = 11$$

$$\vec{x}.\vec{y} = \mathbf{X}^T \mathbf{Y} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 4 + 10 + 18 = 32$$

Norme (ou longueur) d'un vecteur dans $\mathbb{R}^n$

La norme d'un vecteur $\vec{x}$ de $\mathbb{R}^n$, de coordonnées $\mathbf{X} = \{x_i\}_{1 \leqslant i \leqslant n}$, est définie par :

$$\|\vec{x}\|^2 = \sum_{i=1}^{n} x_i^2 \Leftrightarrow \|\vec{x}\|^2 = \vec{x}.\vec{x} = \mathbf{X}^T\mathbf{X}$$

- $\|\vec{x}\| \geqslant 0$ et $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$
- $\|\lambda\vec{x}\| = |\lambda| \ \|\vec{x}\|$
- $\|\vec{x} + \vec{y}\| \leqslant \ \|\vec{x}\| + \|\vec{y}\|$

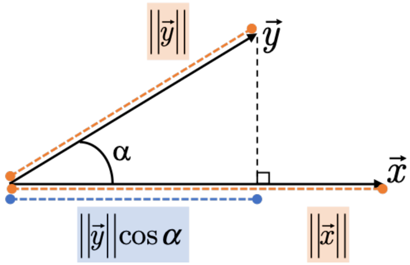
On dira qu'un vecteur est *normé* s'il est de norme égale à 1.
On peut construire un tel vecteur comme suit :

$$\vec{u} = \frac{\vec{x}}{\|\vec{x}\|} \quad \forall \vec{x} \neq \vec{0} \in \mathbb{R}^n$$

Interprétation géométrique

Quels que les soient les vecteurs $\vec{x}$ et $\vec{y}$ de $\mathbb{R}^n$, l'angle qu'ils forment entre eux est l'angle α tel que $0 \leqslant \alpha \leqslant \pi$ et :

$$\cos\alpha = \frac{\vec{x}.\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} \Leftrightarrow \vec{x}.\vec{y} = \|\vec{x}\|\,\|\vec{y}\|\cos\alpha$$



Orthogonalité

Deux vecteurs $\vec{x}$ et $\vec{y}$ de $\mathbb{R}^n$ sont *orthogonaux* si l'angle qu'ils forment vaut $\pm\frac{\pi}{2}$.

Ainsi, deux vecteurs de $\mathbb{R}^n$ seront orthogonaux si $\vec{x}.\vec{y} = 0 \Leftrightarrow \mathbf{X}^T\mathbf{Y} = 0$.

Exemples :

Les vecteurs $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\mathbf{e}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ sont normés et orthogonaux. La base $\{\mathbf{e}_1, \mathbf{e}_2\}$ est dite **orthonormée**.

Les vecteurs $\mathbf{X} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\mathbf{Y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ sont orthogonaux.

Table des matières

Plan détaillé

Exemple

Soit $\vec{x}$ de coordonnées $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

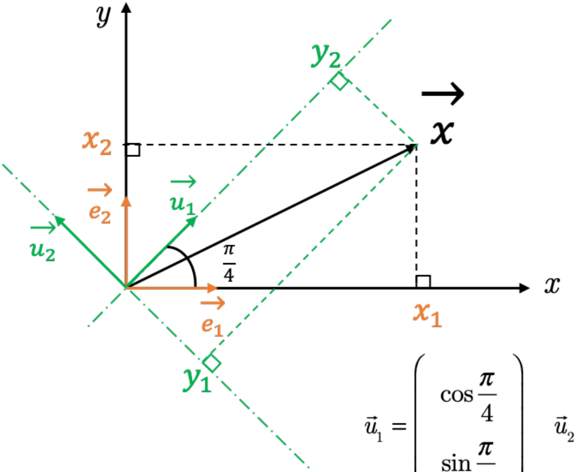
dans la base canonique $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ de $\mathbb{R}^2$:

$$\mathbf{X} = x_1\mathbf{e_1} + x_2\mathbf{e}_2$$

Une rotation des axes d'un angle $\frac{\pi}{4}$ définit une nouvelle base orthonormée dont les vecteurs ont pour coordonnées

$$\mathbf{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Exemple : visualisation



$$\vec{u}_1 = \begin{pmatrix} \cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} \end{pmatrix} \quad \vec{u}_2 = \begin{pmatrix} -\cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} \end{pmatrix}$$

Exemple (suite)

On peut alors étalibr des relations entre les vecteurs de la base canonique $(\vec{e}_1, \vec{e}_2)$ et les vecteurs de la nouvelle base $(\vec{u}_1, \vec{u}_2)$:

$$\begin{cases} \mathbf{u}_1 = \dfrac{\mathbf{e}_1 + \mathbf{e}_2}{\sqrt{2}} \\ \mathbf{u}_2 = \dfrac{\mathbf{e}_2 - \mathbf{e}_1}{\sqrt{2}} \end{cases} \Leftrightarrow \begin{cases} \mathbf{e}_1 = \dfrac{\mathbf{u}_1 - \mathbf{u}_2}{\sqrt{2}} \\ \mathbf{e}_2 = \dfrac{\mathbf{u}_1 + \mathbf{u}_2}{\sqrt{2}} \end{cases}$$

Dans cette nouvelle base $\{\mathbf{u}_1, \mathbf{u}_2\}$, les nouvelles coordonnées de $\vec{x}$ sont alors :

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{x_2-x_1}{\sqrt{2}} \end{pmatrix}$$

Exemple (suite)

Relations entre $(\vec{e}_1, \vec{e}_2)$ et $(\vec{u}_1, \vec{u}_2)$

$$\mathbf{u_1} = \begin{pmatrix} \cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$\mathbf{u_1} = \frac{\mathbf{e_1}+\mathbf{e_2}}{\sqrt{2}} \qquad \|\mathbf{u_1}\| = 1$$

$$\mathbf{u}_2 = \begin{pmatrix} -\cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \left[-\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

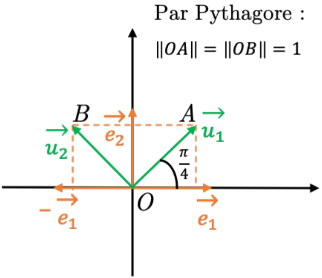
$$\mathbf{u}_2 = \frac{\mathbf{e}_2-\mathbf{e}_1}{\sqrt{2}} \qquad \|\mathbf{u_2}\| = 1$$

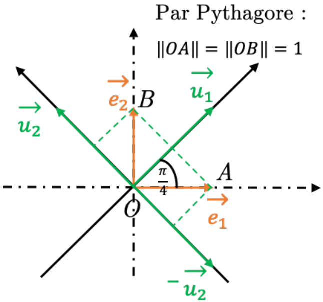
et

$$\begin{cases} \mathbf{e_1} + \mathbf{e_2} = \sqrt{2}\mathbf{u_1} \\ -\mathbf{e_1} + \mathbf{e_2} = \sqrt{2}\mathbf{u}_2 \end{cases} \Leftrightarrow \begin{cases} 2\mathbf{e}_2 = \sqrt{2}\,(\mathbf{u_1} + \mathbf{u}_2) & \text{par somme} \\ 2\mathbf{e}_1 = \sqrt{2}\,(\mathbf{u_1} - \mathbf{u}_2) & \text{par soustraction} \end{cases}$$

$$\Leftrightarrow \begin{cases} \mathbf{e_1} = \frac{\mathbf{u_1}-\mathbf{u}_2}{\sqrt{2}} \\ \mathbf{e}_2 = \frac{\mathbf{u_1}+\mathbf{u}_2}{\sqrt{2}} \end{cases}$$

Exemple : visualisation (suite)





Matrice de passage

Soit $\mathbf{X} = \{x_i\}_{1 \leqslant i \leqslant n}$ un vecteur de $\mathbb{R}^n$.
Ses coordonnées dans une autre base, qui n'est pas nécessairement orthonormée, $\mathbf{Y} = \{y_i\}_{1 \leqslant i \leqslant n}$, vérifient :

$$\mathbf{X} = \mathbf{P}\mathbf{Y}$$

avec $\mathbf{P}$ la *matrice de passage* de l'ancienne vers la nouvelle base, telle que sa i^{e} colonne contient les coordonnées du i^{e} nouveau vecteur de la base dans l'ancienne base.
$\mathbf{P}$ est une matrice carrée dont les colonnes sont linéairement indépendantes, donc $\mathbf{P}$ est inversible et :

$$\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$$

Retour à l'exemple

D'après les coordonnées des vecteurs $(\mathbf{u}_1, \mathbf{u}_2)$ dans la base $(\mathbf{e}_1, \mathbf{e}_2)$, on a :

$$\mathbf{P} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \Leftrightarrow \mathbf{P}^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

Ainsi, les coordonnées $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ deviennent :

$$\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{x_2-x_1}{\sqrt{2}} \end{pmatrix}$$

Plan détaillé

Définitions

Soit $\mathbf{A}$ une matrice carrée de dimension n.

On appelle *vecteur propre* de $\mathbf{A}$ tout vecteur $\mathbf{X} \neq \mathbf{0}$ tel qu'il existe un réel λ qui vérifie :

$$\mathbf{AX} = \lambda\mathbf{X} \Leftrightarrow (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{X} = 0$$

On appelle alors λ la *valeur propre* associée à $\mathbf{X}$.

L'équation $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{X} = 0$ admet des solutions $\mathbf{X} \neq 0$ si et seulement si

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

ce qui définit un polynôme de degré n en λ qu'on appelle *l'équation caractéristique*.

Exemple

Trouver les valeurs et vecteurs propres de $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}$.

- L'équation caractéristique s'écrit $\lambda^2 - 3\lambda + 2 = 0$. Elle se factorise en $(\lambda - 1)(\lambda - 2) = 0$.
  Les valeurs propres sont donc $\lambda_1 = 1$ et $\lambda_2 = 2$.
- Soit $\mathbf{V} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ le vecteur propre associé à λ_1 ; il vérifie $\mathbf{AV} = \lambda_1 \mathbf{V}$, ce qui conduit à la relation $3v_1 + 2v_2 = 0$ et on peut prendre par exemple $\mathbf{V} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$.
- De même, si on note $\mathbf{W} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ le vecteur propre associé à λ_2 ; il vérifie $\mathbf{AW} = \lambda_2 \mathbf{W}$, ce qui conduit à la relation $w_1 + w_2 = 0$ et on peut prendre par exemple $\mathbf{W} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Relations avec $tr(\mathbf{A})$ et $\det(\mathbf{A})$

Soit $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

À partir de $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$, l'équation caractéristique s'écrit

$$\lambda^2 - (a_{11} + a_{22})\,\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$$
$$\Leftrightarrow \lambda^2 - tr(\mathbf{A})\,\lambda + \det(\mathbf{A}) = 0$$

Supposons que $\mathbf{A}$ admette deux valeurs propres λ_1 et λ_2. Alors, l'équation caractéristique se factorise :

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = 0 \Leftrightarrow \lambda^2 - (\lambda_1 + \lambda_2)\,\lambda + \lambda_1\lambda_2 = 0$$

On en déduit que $\begin{cases} tr(\mathbf{A}) = \lambda_1 + \lambda_2 \\ \det(\mathbf{A}) = \lambda_1\lambda_2 \end{cases}$

Plan détaillé

Définition

Une matrice carrée $\mathbf{A}$ est dite *diagonalisable* s'il existe une matrice carrée $\mathbf{P}$ inversible et une matrice diagonale $\mathbf{D}$ telles que :

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

Les coefficients de $\mathbf{D}$ sont les valeurs propres de $\mathbf{A}$ et les colonnes de $\mathbf{P}$ sont les vecteurs propres associés.

Retour à l'exemple :

$$\mathbf{P} = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix} \quad \mathbf{P}^{-1} = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}$$

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix}$$

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} -2 & 2 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Propriétés

1. $$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \Leftrightarrow \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

   La matrice $\mathbf{P}$ s'interprète comme la *matrice de passage* de la base canonique dans la base formée par les vecteurs propres.

2. $$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \Rightarrow tr\left(\mathbf{D}\right) = tr\left(\mathbf{A}\right)$$

Puissance n^{ieme} d'une matrice carrée

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \Rightarrow \mathbf{D}^n = \mathbf{P}^{-1}\mathbf{A}^n\mathbf{P}$$

On en déduit que :

- les valeurs propres de $\mathbf{A}^n$ sont les puissances n^{ieme} de celles de $\mathbf{A}$;
- les vecteurs propres de $\mathbf{A}^n$ sont les vecteurs propres de $\mathbf{A}$;
- $\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$, ce qui fournit un moyen de calculer $\mathbf{A}^n$.

Pour aller plus loin

1. Calcul de la matrice de variance-covariance
2. Régression linéaire gaussienne et estimateur des moindres carrés

Calcul de la matrice de variance-covariance (1)

Supposons que l'on possède un échantillon de taille n pour deux variables X_1 et X_2.

Soit $\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{pmatrix}$ la matrice des observations.

Alors $\mathbf{X}^T\mathbf{X} = \begin{pmatrix} \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} \\ \sum_{i=1}^{n} x_{i1}x_{i2} & \sum_{i=1}^{n} x_{i2}^2 \end{pmatrix}$.

Posons

$\mathbf{B} = \begin{pmatrix} \sum_{i=1}^{n} x_{i1} & \dots & \sum_{i=1}^{n} x_{i1} \\ \sum_{i=1}^{n} x_{i2} & \dots & \sum_{i=1}^{n} x_{i2} \end{pmatrix}$ et $\mathbf{E} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$.

Calcul de la matrice de variance-covariance (2)

On peut alors montrer que
$\mathbf{B} = \mathbf{X}^T\mathbf{E}$ et que

$$\mathbf{B}\mathbf{X} = \begin{pmatrix} \left(\sum_{i=1}^n x_{i1}\right)^2 & \sum_{i=1}^n x_{i1} \sum_{i=1}^n x_{i2} \\ \sum_{i=1}^n x_{i1} \sum_{i=1}^n x_{i2} & \left(\sum_{i=1}^n x_{i2}\right)^2 \end{pmatrix} = n^2 \overline{\mathbf{X}}\overline{\mathbf{X}}^T$$

avec $\overline{\mathbf{X}} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \end{pmatrix}$.

On montre finalement que la matrice de variance-coriance s'écrit :

$$\Sigma = \frac{1}{n-1}\mathbf{X}^T\mathbf{X} - \frac{n}{n-1}\overline{\mathbf{X}}\overline{\mathbf{X}}^T$$

Matrice de corrélation

Si on définit les variables centrées réduites à partir de X_1 et X_2 comme suit :

$$\mathbf{Y} = \begin{pmatrix} \frac{x_{11}-\overline{X}_1}{s_{X_1}} & \frac{x_{12}-\overline{X}_2}{s_{X_2}} \\ \frac{x_{21}-\overline{X}_1}{s_{X_1}} & \frac{x_{22}-\overline{X}_2}{s_{X_2}} \\ \vdots & \vdots \\ \frac{x_{n1}-\overline{X}_1}{s_{X_1}} & \frac{x_{n2}-\overline{X}_2}{s_{X_2}} \end{pmatrix}$$

Alors la matrice de corrélation s'écrit :

$$\Omega = \frac{1}{n-1}\mathbf{Y}^T\mathbf{Y}$$

Exemple d'application

Considérons un échantillon de mesures de pression systolique et diastolique chez 10 patients [1]. Soit

$$\mathbf{X}^T = \begin{pmatrix} 174 & 178 & 182 & 178 & 182 & 162 & 158 & 162 & 154 & 170 \\ 86 & 94 & 98 & 106 & 90 & 86 & 94 & 74 & 86 & 86 \end{pmatrix}$$

la matrice transposée des observations.
$\overline{X}_1 = 170$ et $\overline{X}_2 = 90$.
Alors

$$\Sigma = \begin{pmatrix} 106.67 & 46.22 \\ 46.22 & 74.67 \end{pmatrix} \quad \text{Matrice des variances-covariances}$$

$$\Omega = \begin{pmatrix} 1 & 0.518 \\ 0.518 & 1 \end{pmatrix} \quad \text{Matrice des corrélations}$$

1. `https://fr.wikipedia.org/wiki/Pression_artérielle`

Régression linéaire gaussienne

Soit un échantillon $\{(x_i, y_i)\}_{1 \leqslant i \leqslant n}$ d'observations telles que les x_i sont contrôlées et les y_i à expliquer à partir des x_i.
Sous l'hypothèse d'un modèle linéaire gaussien simple, on a :

$$M : y_i = a + bx_i + \varepsilon_i \quad \text{avec} \quad \varepsilon_\mathrm{i} \sim \mathcal{N}(0, \sigma^2)$$

En notation matricielle, on écrira :

$$\mathbf{Y} = \mathbf{X}\Theta + \mathrm{E}$$

$$\Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Estimateur des moindres carrés (1)

Des estimations $\hat{a}$ et $\hat{b}$ des paramètres s'obtiennent par minimisation de la somme des carrés des écarts définie comme suit :

$$SCE = \sum_{i=1}^{n} (y_1 - a - bx_i)^2$$

Le minimum s'obtient à partir du système suivant :

$$\begin{cases} \dfrac{\partial SCE}{\partial a} = 0 \\ \dfrac{\partial SCE}{\partial b} = 0 \end{cases} \Leftrightarrow \begin{cases} na + \hat{b} \displaystyle\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i \\ a \displaystyle\sum_{i=1}^{n} x_i + \hat{b} \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

Ce qui conduit aux expressions :

$$\hat{b} = \frac{s_{xy}}{s_x^2} \quad \text{et} \quad \hat{a} = \overline{y} - \hat{b}\overline{x}$$

avec s_{xy} la covariance des (x_i, y_i) et s_x^2 la variance des x_i.

Estimateur des moindres carrés (2)

En notation matricielle, on obtient :

$$\hat{\Theta} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}$$

$$\text{var}\left(\hat{\Theta}\right) = \sigma^2\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\Theta} \Leftrightarrow \hat{\mathbf{Y}} = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y} \Leftrightarrow \hat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}$$

avec $\mathbf{H}$ la “matrice chapeau” ou *matrice de projection*.

$$\hat{\mathrm{E}} = \mathbf{Y} - \mathbf{X}\hat{\Theta} \Leftrightarrow \hat{\mathrm{E}} = \left(\mathbf{I} - \mathbf{H}\right)\hat{\Theta}$$

$$\text{var}\left(\hat{\mathrm{E}}\right) = \sigma^2\left(\mathbf{I} - \mathbf{H}\right)$$

A l’aide du produit scalaire, on peut établir que $\hat{\mathrm{E}}^T\hat{\mathbf{Y}} = 0$, c’est-à-dire que les résidus sont indépendants des valeurs prédites, ce qui justifie l’utilisation du graphe “résidus vs valeurs prédites”.

Décomposition de la variabilité

La variation totale des y_i peut se décomposer en la somme de la variation expliquée par le modèle + la variation résiduelle :

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} \hat{y}_i^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

En notation matricielle, on obtient :

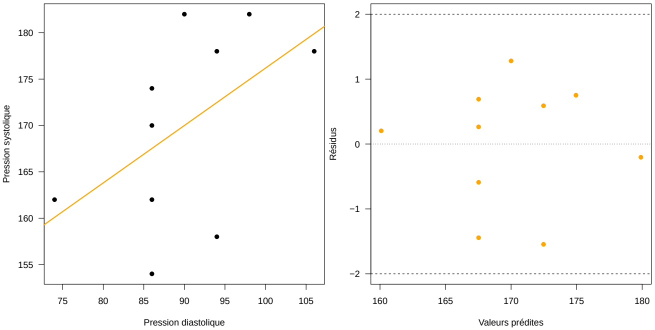
- Variation totale $= \mathbf{Y}^T\mathbf{Y}$
- Variation expliquée $= \hat{\mathbf{Y}}^T\hat{\mathbf{Y}} = \left(\mathbf{X}\hat{\Theta}\right)^T\mathbf{X}\hat{\Theta} = \hat{\Theta}^T\mathbf{X}^T\mathbf{X}\hat{\Theta}$
- Variation résiduelle $=$ Variation totale - Variation expliquée

La quantité $s^2 = \frac{\text{Variation résiduelle}}{n-2}$ est une estmation de σ^2.

Exemple d'application

Retour aux pressions systolique et diastolique (en supposant cette dernière contrôlée).

$$\mathbf{X}^T = \begin{pmatrix} 174 & 178 & 182 & 178 & 182 & 162 & 158 & 162 & 154 & 170 \\ 86 & 94 & 98 & 106 & 90 & 86 & 94 & 74 & 86 & 86 \end{pmatrix}$$



$$\hat{\Theta} = \begin{pmatrix} 114.3 \\ 0.619 \end{pmatrix}$$