

Modèles Linéaires Appliqués

Arthur Charpentier

Automne 2020

Rappels #6 (algèbre linéaire)

Matrices

Soient $m, n \geq 1$. Une matrice \mathbf{A} de taille (m, n) à coefficients réels est un tableau de nombres réels ayant m lignes et n colonnes. On note également par $(\mathbf{A})_{ij}$ ou plus simplement A_{ij} l'élément sur la ligne i et sur la colonne j de \mathbf{A} .

Example:

$$\mathbf{A} = \begin{pmatrix} 1.5 & 2 & 3.1 & 8 \\ -1 & 4 & 5 & 6.5 \end{pmatrix}$$

\mathbf{A} est de taille (2×4) et par exemple $A_{13} = 3.1$.

Une matrice ne contenant qu'une colonne est appelée un vecteur et une matrice ne contenant qu'une ligne est un vecteur ligne. Par exemple $\mathbf{x} = \begin{pmatrix} 1.5 \\ -1 \end{pmatrix}$ et $\mathbf{y} = (1.5 \ 2 \ 3.1 \ 8)$ sont respectivement de taille $(2, 1)$ et $(1, 4)$.

Transposée

Soit \mathbf{A} une matrice réelle de taille (m, n) . La matrice transposée notée \mathbf{A}^\top de taille (n, m) est définie par $(\mathbf{A}^\top)_{ij} = A_{ji}$ pour $i = 1, \dots, n$ et $j = 1, \dots, m$.
Et $(\mathbf{A}^\top)^\top = \mathbf{A}$.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

$$\|\mathbf{x}\|^2 = \mathbf{x}^\top \mathbf{x} = \sum_{i=1}^n x_i^2$$

$$\mathbf{X}^\top \mathbf{X} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top$$

```
1 > t(1:4) %*% rep(1,4)
2      [,1]
3 [1,]    10
4 > (1:4) %*% t(rep(1,4))
5      [,1] [,2] [,3] [,4]
6 [1,]     1     1     1     1
7 [2,]     2     2     2     2
8 [3,]     3     3     3     3
9 [4,]     4     4     4     4
10 > t(1:4) %*% (1:4)
11      [,1]
12 [1,]    30
13 > (1:4) %*% t(1:4)
14      [,1] [,2] [,3] [,4]
15 [1,]     1     2     3     4
16 [2,]     2     4     6     8
17 [3,]     3     6     9    12
18 [4,]     4     8    12    16
```

Products

If **A** and **B** are (respectively) $k \times m$ and $m \times n$ matrices,

$$C_{ij} = \mathbf{A}_{i\cdot}^\top \mathbf{B}_{\cdot j} = A_{i1}B_{1j} + \cdots + A_{im}B_{mj} = \sum_{k=1}^m A_{ik}B_{kj},$$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1p} \\ B_{21} & B_{22} & \cdots & B_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mp} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{np} \end{pmatrix}$$

```
1 > A = matrix(1:6,2,3)
2 > B = matrix(1:12,3,4)
3 > A %*% B
4      [,1] [,2] [,3] [,4]
5 [1,]    22    49    76   103
6 [2,]    28    64   100   136
```

Le produit matriciel n'est pas commutatif pour deux matrices quelconque de même taille: **AB** \neq **BA**

Produit & Trace

Soit \mathbb{I}_n la matrice de taille (n, n) composée de 1 sur la diagonale et de 0 ailleurs. Alors, pour \mathbf{A} de taille (n, n) , \mathbb{I}_n est l'élément neutre tel que $\mathbf{A}\mathbb{I}_n = \mathbb{I}_n\mathbf{A} = \mathbf{A}$.

Soient \mathbf{A} , \mathbf{B} et \mathbf{C} trois matrices réelles de dimension concordante, alors

- ▶ $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ (associativité du produit)
- ▶ $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ (distributivité du produit)
- ▶ $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$.

Soient \mathbf{A} , \mathbf{B} et \mathbf{M} trois matrices réelles $n \times n$ (carré),

- ▶ $\text{trace}(\mathbf{M}) = \sum_{i=1}^n M_{i,i}$
- ▶ $\text{trace}(\mathbf{A} + \mathbf{B}) = \text{trace}(\mathbf{A}) + \text{trace}(\mathbf{B})$
- ▶ $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$

Eigenvalues & Eigenvectors

$\vec{u} \neq \vec{0}$ is an eigenvector of squared matrix \mathbf{M} if $\mathbf{M}\vec{u} = \lambda \vec{u}$ for some λ (called eigenvalue).

Example:

$$\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{M}\vec{u}_1 = 2 \vec{u}_1$$

$$\mathbf{M}\vec{u}_2 = 0 \vec{u}_2$$

Note: $\text{trace}(\mathbf{M}) = 2 = 2 + 0$

```
1 > M = matrix(c(1,1,1,1),2,2)
2 > u = eigen(M)$vector[,1]
3 > u
4 [1] 0.7071068 0.7071068
5 > M**%u
6           [,1]
7 [1,] 1.414214
8 [2,] 1.414214
9 > M**%c(1,1)
10          [,1]
11 [1,]      2
12 [2,]      2
13 > M**%c(1,-1)
14          [,1]
15 [1,]      0
16 [2,]      0
```

Eigenvalues & Eigenvectors

Example:

$$\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \vec{\mathbf{u}}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{\mathbf{u}}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{M}\vec{\mathbf{u}}_1 = 3 \vec{\mathbf{u}}_1$$

$$\mathbf{M}\vec{\mathbf{u}}_2 = -1 \vec{\mathbf{u}}_2$$

Note: $\text{trace}(\mathbf{M}) = 2 = 3 - 1$

for further interpretation of eigenvectors, see [ACT6100](#)

```
1 > eigen(matrix(c(1,1,1,1),2,2))
2 $values
3 [1] 2 0
4
5 $vectors
6           [,1]      [,2]
7 [1,] 0.7071068 -0.7071068
8 [2,] 0.7071068  0.7071068
9 > eigen(matrix(c(1,2,2,1),2,2))
10 $values
11 [1] 3 -1
12
13 $vectors
14           [,1]      [,2]
15 [1,] 0.7071068 -0.7071068
16 [2,] 0.7071068  0.7071068
```

Rank

A set of vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ are linearly independent if

$$\forall (a_1, \dots, a_k) \neq \mathbf{0}, \sum_{i=1}^k a_i \mathbf{x}_i \neq \mathbf{0}$$

The rank of a matrix \mathbf{X} , $\text{rank}(\mathbf{X})$, is the maximum number of linearly independent columns of \mathbf{X} .

A $n \times k$ matrix of rank k is said to be of full rank.

If \mathbf{A} is $p \times q$, $\text{rank}(\mathbf{A}) \leq \min\{p, q\}$

If \mathbf{A} is a $n \times n$ matrix, with $\text{rank}(\mathbf{A}) = n$ then \mathbf{A} is invertible.

If \mathbf{A} is a $n \times n$ matrix, $\text{rank}(\mathbf{A})$ is the number of non-null eigenvalues.

```
1 > (M = matrix(c
      (1,1,1,1,2,1,2,2,2)
      ,3,3))
2      [,1] [,2] [,3]
3 [1,]    1    1    2
4 [2,]    1    2    2
5 [3,]    1    1    2
6 > eigen(M)
7 $values
8 [1] 4.303 0.697 0.000
9
10 $vectors
11      [,1] [,2] [,3]
12 [1,] -0.520 0.370 -0.894
13 [2,] -0.677 -0.852 0.000
14 [3,] -0.520 0.370 0.447
```

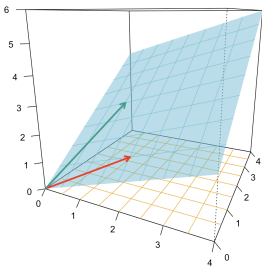
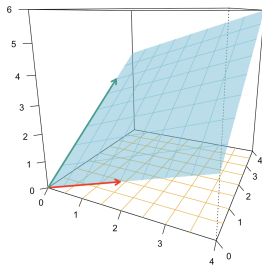

Espace vectoriel engendré

Soient $\mathbf{x}_1, \dots, \mathbf{x}_p \in \mathbb{R}^n$, on définit $\mathcal{V}(\mathbf{x}_1, \dots, \mathbf{x}_p)$ comme

$$\left\{ \mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \sum_{i=1}^p a_i \mathbf{x}_i = \mathbf{X}\mathbf{a}, \mathbf{a} \in \mathbb{R}^p \right\}$$

où $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_p]$ est une matrice $n \times p$.

La dimension de $\mathcal{V}(\mathbf{x}_1, \dots, \mathbf{x}_p)$ est le rang de \mathbf{X} .



Definite Positive

Let \mathbf{A} be an $n \times n$ matrix, $\mathbf{x} \mapsto \mathbf{x}^\top \mathbf{A} \mathbf{x}$ is called a quadratic form

- ▶ if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$, \mathbf{A} is a positive definite matrix
- ▶ if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$, \mathbf{A} is a positive semidefinite matrix

Toute matrice \mathbf{A} qui peut s'écrire sous la forme $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$ est semi-définie positive.

- ▶ $\text{trace}(\mathbf{B}^\top \mathbf{B}) \geq 0$
- ▶ $\text{trace}(\mathbf{B}^\top \mathbf{B}) = 0 \Leftrightarrow \mathbf{B}^\top \mathbf{B} = \mathbf{0} \Leftrightarrow \mathbf{B} = \mathbf{0}$.

Inverse

Soit \mathbf{A} une matrice carrée de taille (n, n) dont le déterminant est non nul, alors \mathbf{A} est dite non singulière et il existe une matrice inverse (de même taille) notée \mathbf{A}^{-1} vérifiant $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbb{I}_n$
Soient \mathbf{A} et \mathbf{B} deux matrices inversibles de taille (n, n) alors

► $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$

(\mathbf{A}^{-1}) est symétrique ssi \mathbf{A} est symétrique)

► $(\mathbf{A}^{-1})^{-1} = \mathbf{A}.$

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

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

```
1 > A = matrix(  
2   c(3,2,4,3),2,2)  
3 > A  
4           [,1] [,2]  
5 [1,]        3    4  
6 [2,]        2    3  
7 > solve(A)  
8           [,1] [,2]  
9 [1,]        3   -4  
10 [2,]       -2    3
```

Idempotent Matrices

Symmetric matrix \mathbf{P} is idempotent if $\mathbf{P}^2 = \mathbf{P}$.

Note that \mathbf{P} is positive semi-definite.

And $\text{rank}(\mathbf{P}) = \text{trace}(\mathbf{P})$

E.g. $\mathbf{P} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}$ is idempotent

Théorème de Cochran: si \mathbf{A} est une matrice idempotente

- ▶ $\text{rank}(\mathbf{A}) = \text{trace}(\mathbf{A})$.
- ▶ les valeurs propres de \mathbf{A} valent soit 0 soit 1.

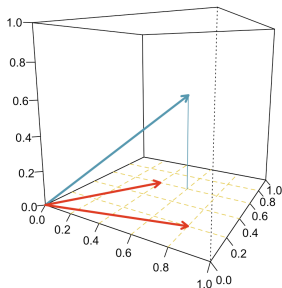
Soit $\mathbf{A} = \mathbf{A}_1 + \cdots + \mathbf{A}_k$. Alors les deux énoncés suivants sont équivalents

- ▶ \mathbf{A} est idempotente et $\text{rank}(\mathbf{A}) = \sum_{i=1}^k \text{rank}(\mathbf{A}_i)$.
- ▶ \mathbf{A}_i est idempotente pour tout i et $\mathbf{A}_i \mathbf{A}_j = 0$ pour tout $i \neq j$.

Projection

Consider the projection (in \mathbb{R}^3) on $\{\vec{x}_1, \vec{x}_2\}$. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$.
 $\mathbf{P} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ is the (orthogonal) projection on $\{\vec{x}_1, \vec{x}_2\}$

```
1 > X=cbind(c(.8,.2,0),c(.4,.6,0))
2 > P=X %*% solve(t(X)%*%X) %*% t(X)
3 > P
4      [,1] [,2] [,3]
5 [1,]    1    0    0
6 [2,]    0    1    0
7 [3,]    0    0    0
8 > P %*% c(.6,.6,1)
9      [,1]
10 [1,]  0.6
11 [2,]  0.6
12 [3,]  0.0
```



$$\mathbf{P}^2 = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \cdot \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top = \mathbf{P}$$
$$\mathbf{P}\vec{x}_1 = \vec{x}_1, \mathbf{P}\vec{x}_2 = \vec{x}_2, \text{ and } \text{rank}(\mathbf{P}) = 2.$$

Matrice par blocs

Soient n_1 et n_2 tels que $n = n_1 + n_2$. On note

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

\mathbf{A}_{11} , \mathbf{A}_{12} , \mathbf{A}_{21} , \mathbf{A}_{22} sont de taille (n_1, n_1) , (n_1, n_2) , (n_2, n_1) , (n_2, n_2) .

Soient \mathbf{A} et \mathbf{B} deux matrices par blocs de taille identique, alors la matrice $\mathbf{C} = \mathbf{AB}$ est aussi une matrice par blocs dont les termes sont définis (dans le cas de 4 blocs) par $\mathbf{C}_{ij} = \sum_{k=1}^2 \mathbf{A}_{ik} \mathbf{B}_{kj}$ pour $i, j = 1, 2$.

$\det(\mathbf{A}) = \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12})$ si \mathbf{A}_{11} est inversible

Soit \mathbf{A} une matrice par blocs inversible alors

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22,1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22,1}^{-1} \\ -\mathbf{A}_{22,1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22,1}^{-1} \end{pmatrix}$$

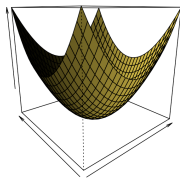
où $\mathbf{A}_{22,1} = \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$.

Quadratic Forms

On peut tracer la surface $S(\mathbf{z}) = \mathbf{z}^\top \mathbf{M} \mathbf{z}$, i.e.

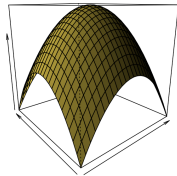
$$S : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & y \end{pmatrix} \mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix}$$

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



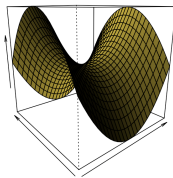
positive
definite

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



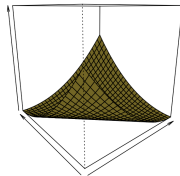
negative
definite

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



indefinite

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



positive
semi-definite

Quadratic Forms

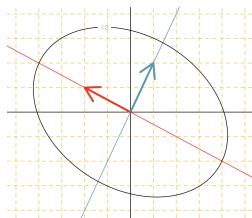
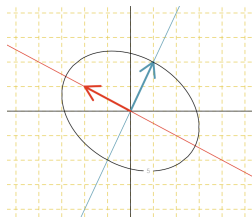
Consider $\mathbf{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$,
and function $\mathbf{z} \mapsto \mathbf{z}^T \mathbf{M} \mathbf{z}$, i.e.

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or $ax^2 + 2bxy + cy^2$ is a quadratic form.

If $\mathbf{M} > 0$, points $\mathbf{z} = (x, y)$ such that $\mathbf{z}^T \mathbf{M} \mathbf{z} = \gamma$, for some $\gamma > 0$, are on an **ellipse** (centered on $\mathbf{0}$)

Let $\lambda_1 \geq \lambda_2 > 0$ denote the eigenvalues of \mathbf{M}
and $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ denote the eigenvectors.



Quadratic Forms

On the picture, $\mathbf{M} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.9 \end{pmatrix}$

```
1 > M=matrix(c(.6,.2,.2,.9),2,2)
2 > eigen(M)
3 eigen() decomposition
4 $values
5 [1] 1.0 0.5
6 $vectors
7           [,1]      [,2]
8 [1,] 0.4472136 -0.8944272
9 [2,] 0.8944272  0.4472136
```

i.e. $\lambda_1 = 1$ and $\lambda_2 = 1/2$, and

$$\vec{v}_1 = \sqrt{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{v}_2 = \sqrt{5} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

Note that $\|\vec{v}_1\| = \|\vec{v}_2\| = 1$ and $\vec{v}_1 \perp \vec{v}_2$

