Data Science for Actuaries (ACT6100)

Arthur Charpentier

Rappels # 2 (Matrices & Linear Transformations)

automne 2020

https://github.com/freakonometrics/ACT6100/

Matrices

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

Example:

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

$\boldsymbol{A}$ 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 $\boldsymbol{x} = \begin{pmatrix} 1.5 \\ -1 \end{pmatrix}$ et $\boldsymbol{y} = (1.5\ 2\ 3.1\ 8)$ sont respectivement de taille $(2, 1)$ et $(1, 4)$.

Products

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

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

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

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

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

Products

For vectors $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$, one can define the dot product

$$\boldsymbol{a} \cdot \boldsymbol{b} = \langle \vec{\boldsymbol{a}}, \vec{\boldsymbol{b}} \rangle = \boldsymbol{a}^\top \boldsymbol{b} = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}$$

For matrices $\boldsymbol{A} \in \mathbb{R}^{m,n}$ and $\boldsymbol{B} \in \mathbb{R}^{m,p}$, one can define the cross product

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{A}^\top \boldsymbol{B} \in \mathbb{R}^{n \times p}$$

```
> crossprod(A, B)
```

For matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^{m \times n}$, one can define the Hadamard product - or element-wise product - as

$$\boldsymbol{A} \odot \boldsymbol{B} = \boldsymbol{C} \in \mathbb{R}^{m \times n}, \ \boldsymbol{C}_{ij} = \boldsymbol{A}_{ij} \boldsymbol{B}_{ij}$$

```
> A*B
```

Products

For matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^{n \times p}$, one can define the matrix product as

$$\boldsymbol{AB} = \boldsymbol{C} \in \mathbb{R}^{m \times p},\ \ \boldsymbol{C}_{ij} = \sum_{k=1}^{n} \boldsymbol{A}_{ik}\boldsymbol{B}_{kj} = \boldsymbol{A}_{i\cdot}^{\top} \cdot \boldsymbol{B}_{\cdot j} = \langle \vec{\boldsymbol{A}}_{i\cdot}, \vec{\boldsymbol{B}}_{\cdot j} \rangle$$

```
> A %*% B
```

For matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^{p \times q}$, one can define the Kronecker product as

$$\boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{C} \in \mathbb{R}^{mp \times nq} = \begin{bmatrix} \boldsymbol{A}_{11}\boldsymbol{B} & \cdots & \boldsymbol{A}_{1n}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{m1}\boldsymbol{B} & \cdots & \boldsymbol{A}_{mn}\boldsymbol{B} \end{bmatrix}$$

```
> kronecker(A, B)
```

Rotation

One can use matrices to transform vectors, e.g. $\vec{\boldsymbol{y}} = \boldsymbol{A}\vec{\boldsymbol{x}}$, with $\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}} \in \mathbb{R}^n$, and $\boldsymbol{A}$ is some $n \times n$ matrix.

Example:

$$\boldsymbol{A}\vec{\boldsymbol{x}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}.$$

If $\boldsymbol{A} = R_{\boldsymbol{0}}(\theta)$, $\boldsymbol{A}^\top = R_{\boldsymbol{0}}(-\theta) = \boldsymbol{A}^{-1}$

Example:

$\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = R_{\boldsymbol{0}}(\theta)$ and $\boldsymbol{B} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = R_{\boldsymbol{0}}(\phi)$

then $\boldsymbol{AB} = \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix}$

i.e. $\boldsymbol{AB} = \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix} = R_{\boldsymbol{0}}(\theta+\phi)$

Rotation & Orthogonal Matrices

In higher dimension n, a $n \times n$ matrix $\boldsymbol{A}$ is orthogonal if its columns and rows are orthogonal unit vectors, i.e.

$$\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\top = \mathbb{I}$$

or equivalently, $\boldsymbol{A}^{-1} = \boldsymbol{A}^\top$.
In dimension 2,

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \text{rotation}, \quad \text{and} \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = \text{reflection}$$

Linear vs. affine

Note that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

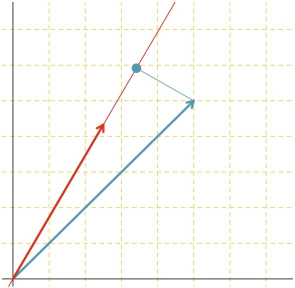
is linear but not affine (there is no constant here). Trick

$$\begin{bmatrix} 1 & 0 & 0 \\ a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ a + bx + cy \\ d + ex + fy \end{bmatrix}$$

Projection

Consider the projection (in $\mathbb{R}^2$) on $\{\vec{\boldsymbol{x}}_1\}$. Let $\boldsymbol{X} = [\boldsymbol{x}_1]$, $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$ is the (orthogonal) projection on $\{\vec{\boldsymbol{x}}_1\}$

```
> theta = pi/3
> u=c(cos(theta),sin(theta))
> X = matrix(u,2,1)
> P = X %*% solve(t(X)%*%X) %*% t(X)
> P
          [,1]      [,2]
[1,] 0.2500000 0.4330127
[2,] 0.4330127 0.7500000
> P %*% c(1,1)
          [,1]
[1,] 0.6830127
[2,] 1.1830127
```

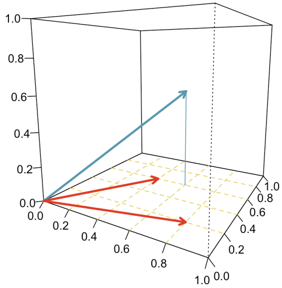


Note: projection on $\{\vec{\boldsymbol{x}}_1\}$ is the projection on the straight line that goes through $\mathbf{0}$, with direction $\vec{\boldsymbol{x}}_1$.

Projection

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

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



Rank

The rank of a matrix $r \times c$ is defined as

- the maximum number of linearly independent column vectors in the matrix
- the maximum number of linearly independent row vectors in the matrix

(the two are equivalent).

Note: $\text{rank}(\boldsymbol{M}) \leq \min\{r, c\}$.

The rank of matrix $\boldsymbol{M}$ is the dimension of the vector space generated (spanned) by its columns. It is equals to the number of non-zero singular values in SVD (or eigenvalues for squared matrices).

Example: $\text{rank}(\boldsymbol{M}^\top\boldsymbol{M}) = \text{rank}(\boldsymbol{M}\boldsymbol{M}^\top) = \text{rank}(\boldsymbol{M}^\top) = \text{rank}(\boldsymbol{M})$