

POLYTECH INFORMATIQUE

3ÈME ANNÉE

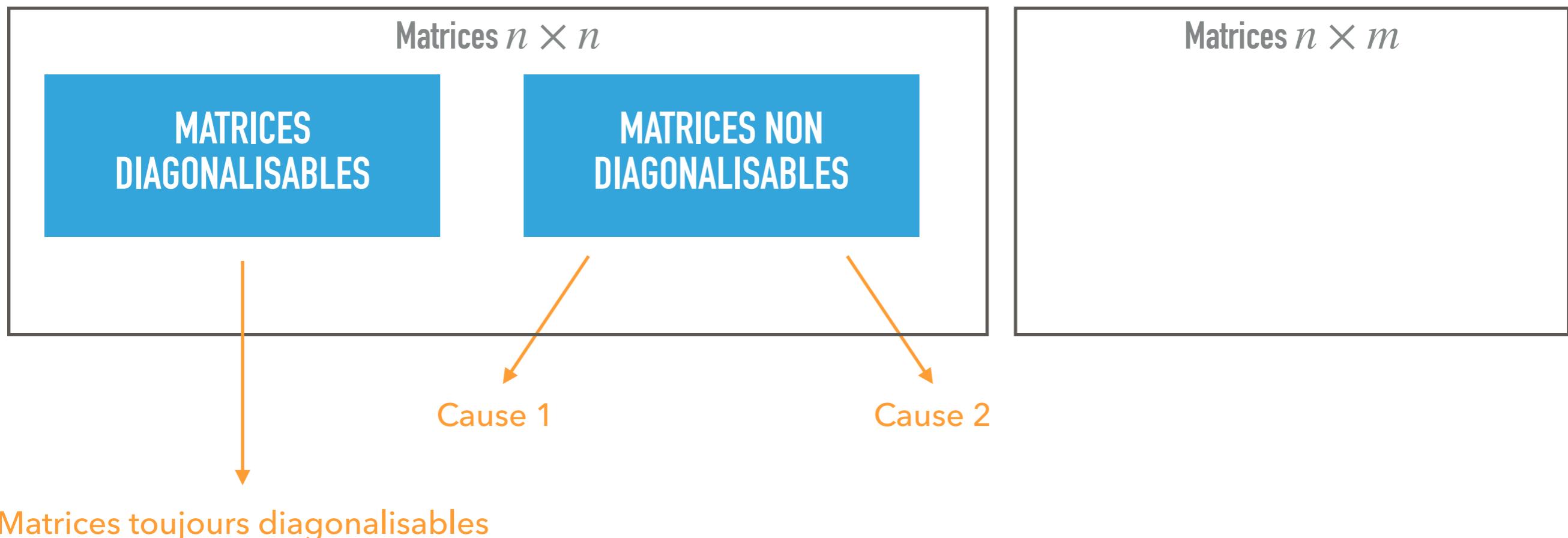
Alexandra Bac

NUMERICAL METHODS

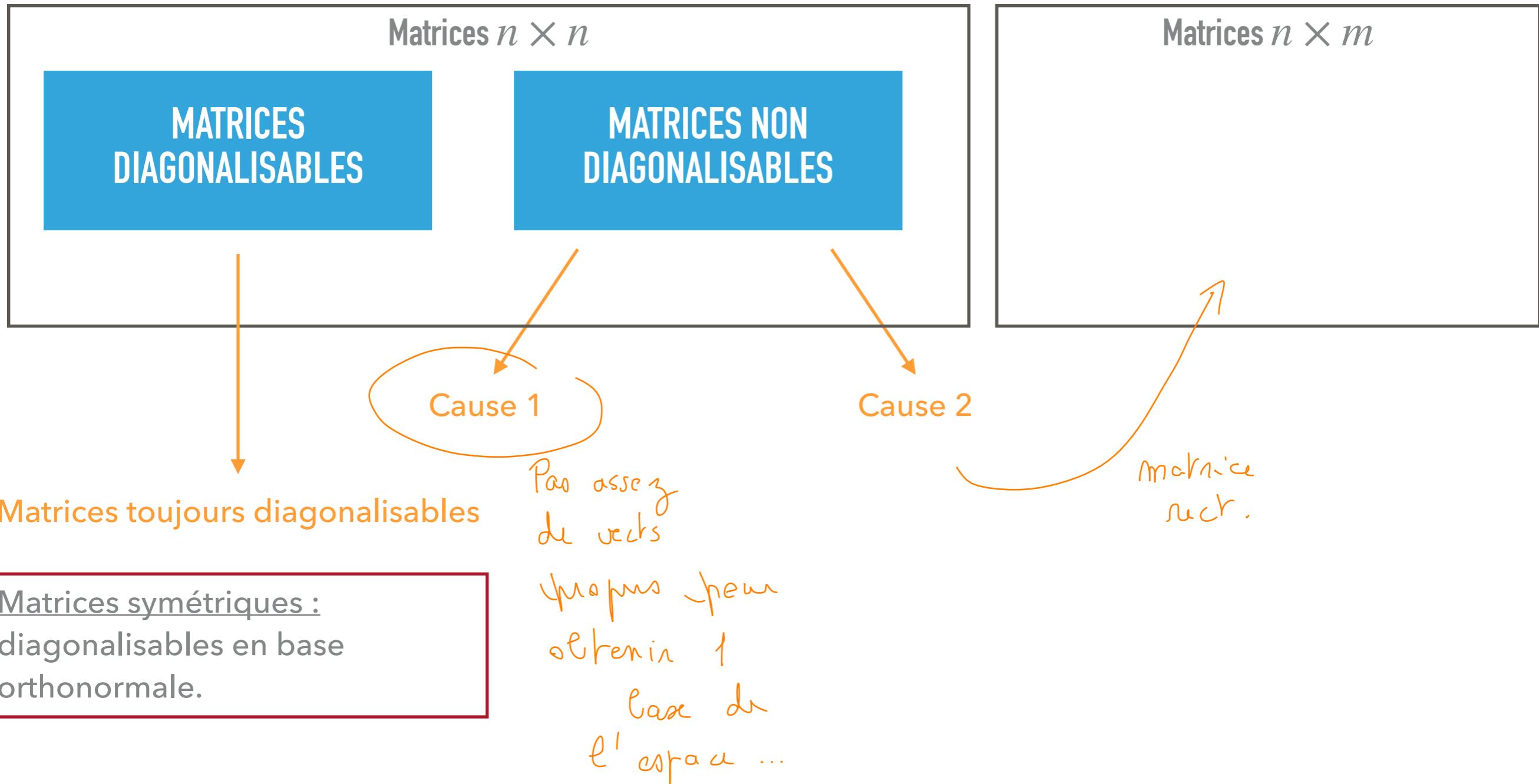
COURS 3

VALEURS SINGULIÈRES, ANALYSE EN COMPOSANTES PRINCIPALE

MATRICES DIAGONALISABLES ...

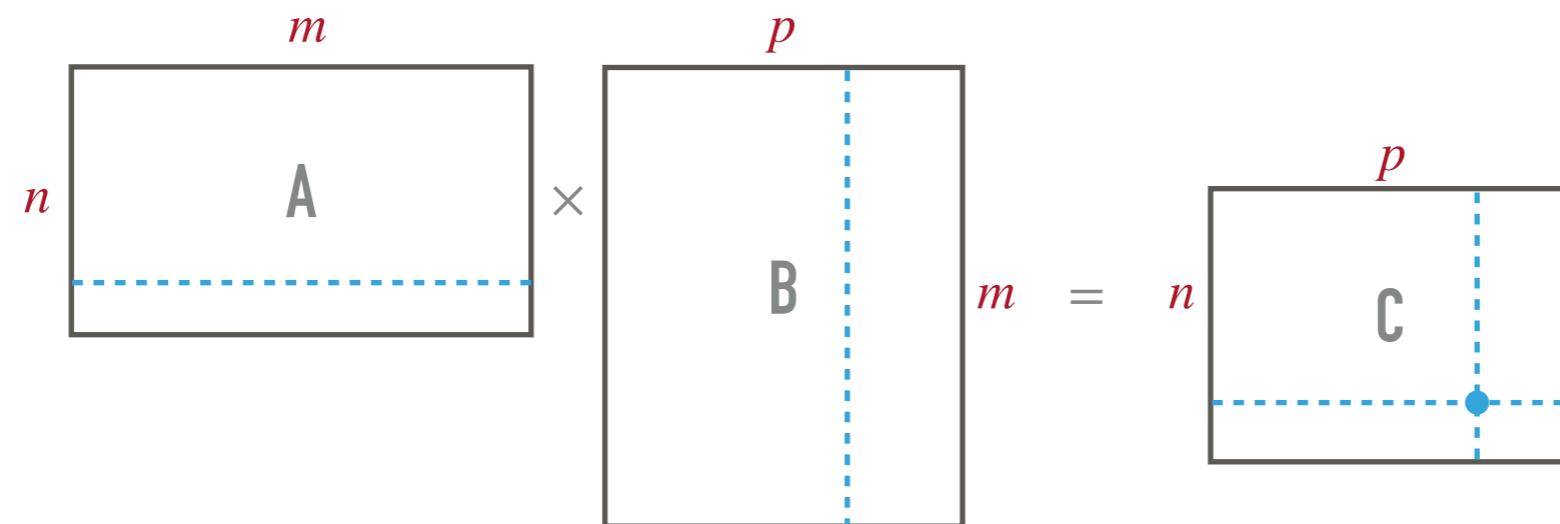


MATRICES DIAGONALISABLES ...



PRODUIT MATRICIEL PAR BLOCS

Produit de matrices :



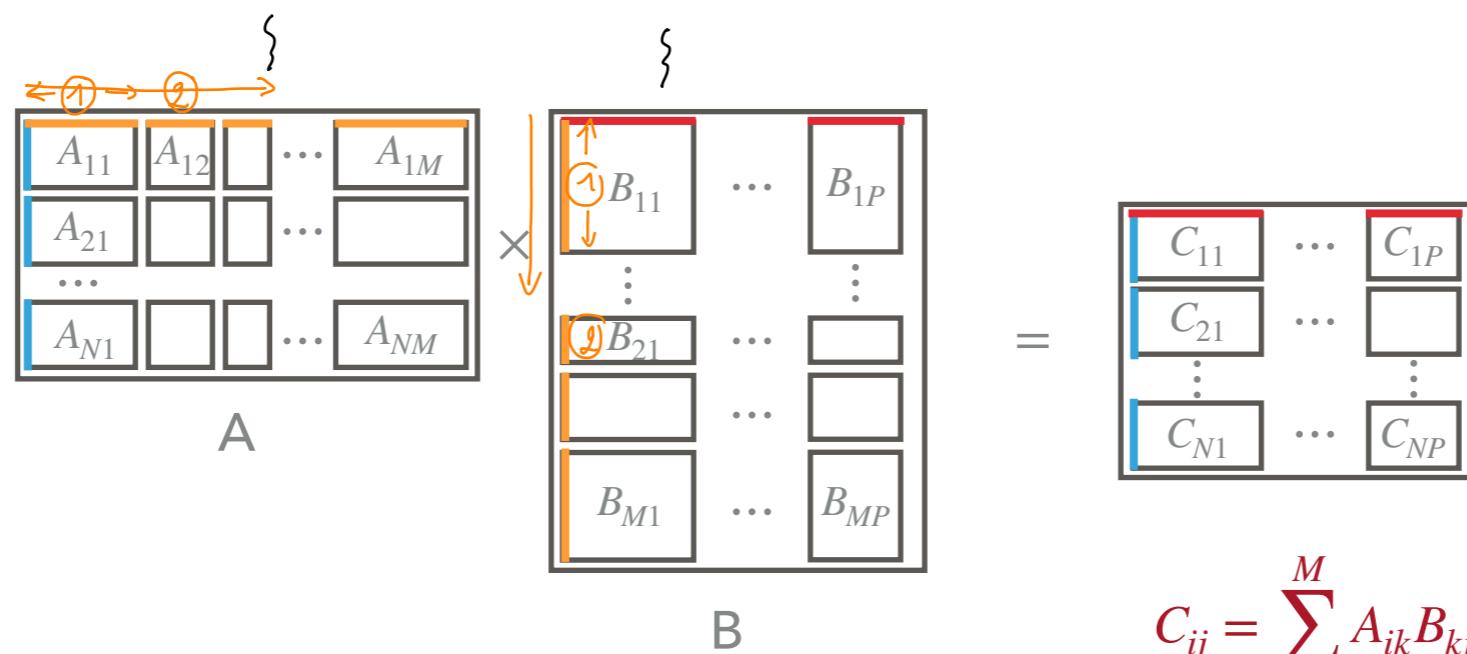
$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$$

PRODUIT MATRICIEL PAR BLOCS

$$N \left(\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1P} \\ \vdots & & & \\ A_{N1} & A_{N2} & \dots & A_{NP} \end{bmatrix} \times \begin{bmatrix} B_{11} & \dots & B_{1P} \\ \vdots & & \\ B_{N1} & \dots & B_{NP} \end{bmatrix} \right) M = C$$

P

$$C_{ij} = \sum_{k=1}^M A_{ik} \times B_{kj}$$



$$C_{ij} = \sum_{k=1}^M A_{ik} B_{kj}$$

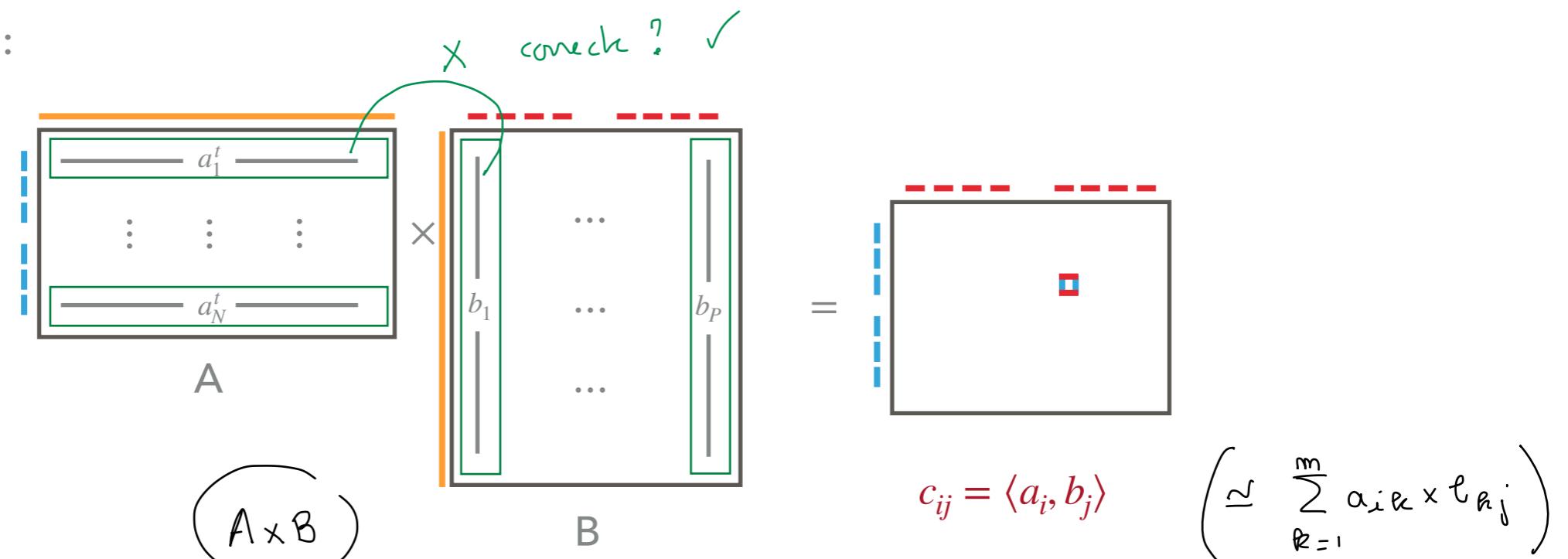
produire du
matrice (blocs)

Si les blocs sont de "tailles compatibles"

$$n_i \begin{matrix} m_j \\ A_{ij} \end{matrix} \times \begin{matrix} p_k \\ B_{jk} \end{matrix} m_j = n_i \begin{matrix} p_k \\ \square \end{matrix}$$

PRODUIT MATRICIEL PAR BLOCS

Exemples :



$$\begin{bmatrix} A_{11} \\ \vdots \\ A_{N1} \end{bmatrix}_{N \times 1 \text{ blocs}} \times A_{i1} = a_i^t$$

$$N \times 1 \quad \times \quad 1 \times P$$

$$B_{11} \dots B_{1P} \quad = \quad \begin{array}{c} \text{1} \times P \text{ blocs} \\ \text{N} \times P \end{array}$$

$$c_{ij} = A_{i1} \times B_{1j} = \langle a_i, b_j \rangle$$

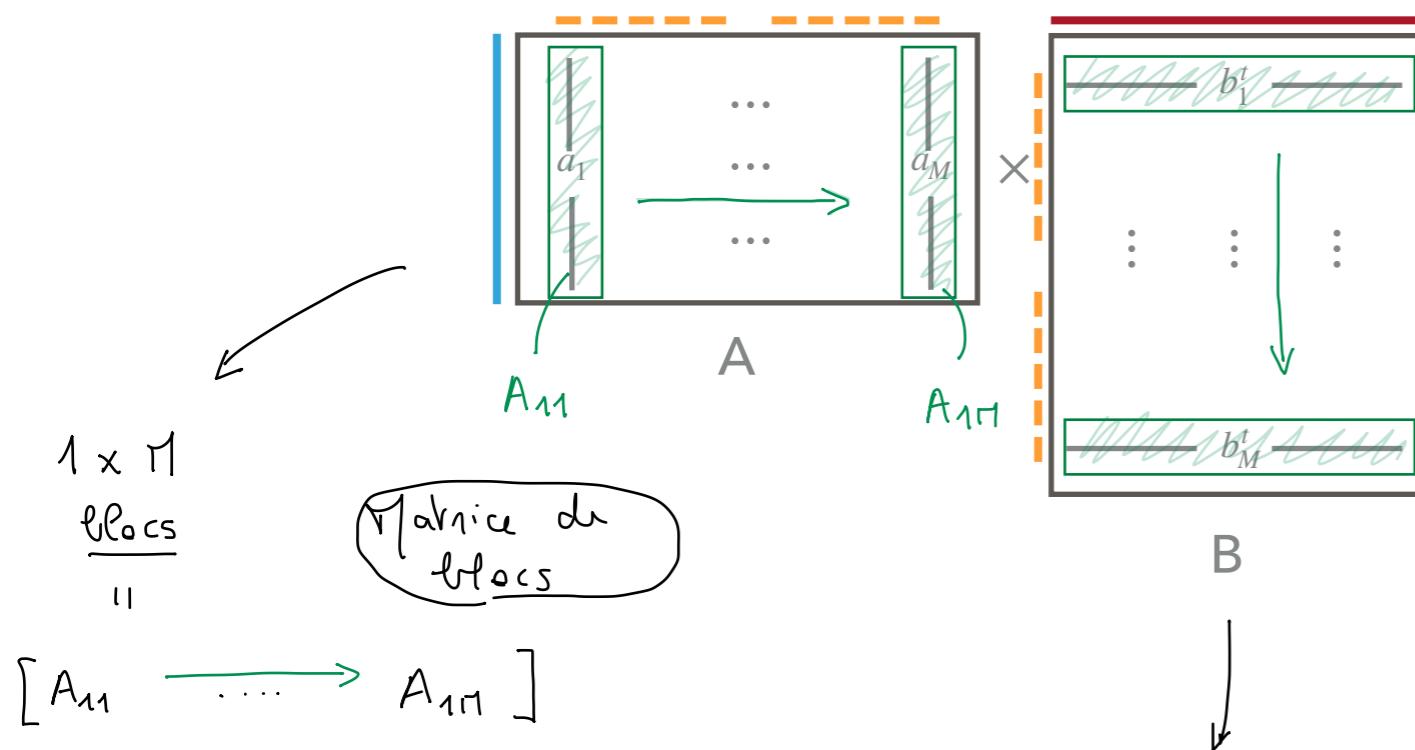
$$a_i^t \times e_{ij}$$

PRODUIT MATRICIEL PAR BLOCS

Vecteur : colonne

$$\text{Vecteur colonne} = (\text{vect. (colonne)})^t$$

Exemples :



$C : 1 \times 1 \text{ blocs}$

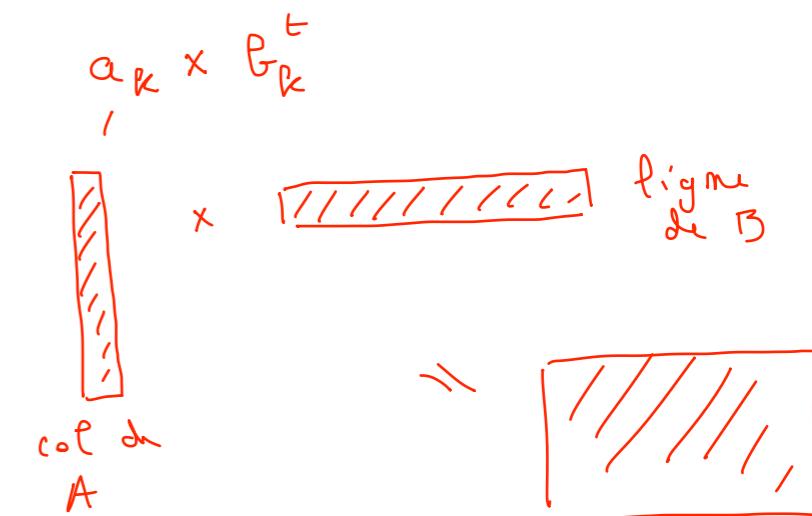
$C_{11} = \sum_{k=1}^M A_{1k} \times B_{k1}$

$C = \sum_{k=1}^M a_k b_k^t$

↓
somme de matrices

$$\begin{matrix} \nearrow n \\ \searrow n \end{matrix} \times \begin{matrix} \nearrow n \\ \searrow n \end{matrix} = \overbrace{\begin{matrix} \nearrow n \\ \searrow n \end{matrix}}^R \begin{matrix} \nearrow 1 \\ \searrow 1 \end{matrix} \begin{matrix} \nearrow 1 \\ \searrow 1 \end{matrix} \text{ blocks}$$

$$\begin{bmatrix} B_{11} \\ \vdots \\ B_{n1} \end{bmatrix}$$



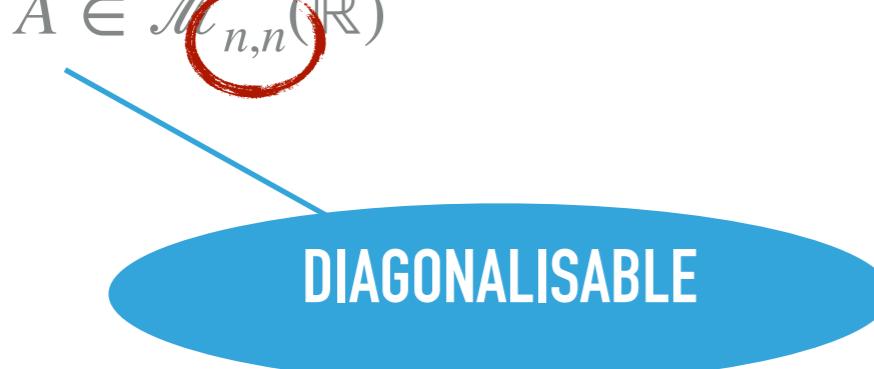
VALEURS SINGULIÈRES

DIAGONALISATION

Matrices rectangulaires ?

Matrices non diagonalisables ?

$$A \in \mathcal{M}_{n,n}(\mathbb{R})$$



$$\exists \mathcal{B} = \{e_1, \dots, e_n\} \text{ où } A \text{ diagonale}$$



e_i vecteur propre

$$Ae_i = \lambda_i e_i$$



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



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

$$\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{matrix}$$

D

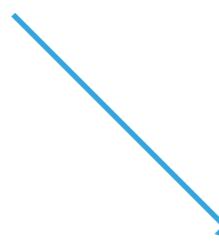
où $P^{-1} =$

$$\begin{bmatrix} | & & | \\ e_1 & \cdots & e_n \\ | & & | \end{bmatrix}$$

MATRICES RECTANGULAIRES

rectangulaire
↑

$$A \in \mathcal{M}_{m,n}(\mathbb{R})$$



$$A^t \cdot A \quad \text{symétrique positive}$$

Diagonalisation garantie
↓

matr. symétriques

diagonalisable en base
orthonormale

matr. changement
de base

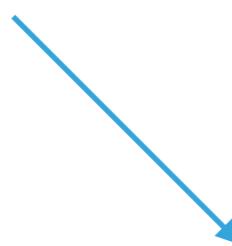
orthogonale

$$\begin{aligned} P^t &= P^{-1} \\ \Leftrightarrow P P^t &= I \end{aligned}$$

$$\begin{aligned} \Leftrightarrow & \text{vect. colonne} \\ & \text{base } \sigma. \text{ no} \end{aligned}$$

MATRICES RECTANGULAIRES

$$A \in \mathcal{M}_{n,m}(\mathbb{R})$$

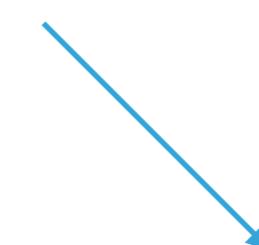


$$A^t A \in \mathcal{M}_{m,m}(\mathbb{R})$$

$\exists V$ orthogonale (ie. base orthonormale de \mathbb{R}^m) telle que :

tq $A^t A$ devient diagonale ds cette base

$$A^t A = V D V^t$$



... / ...

$$A = \dots$$

DÉCOMPOSITION EN VALEURS SINGULIÈRES

$A \in \mathcal{M}_{n,m}(\mathbb{R})$
ndr.

A matrice de $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$

Il existe :

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale $\rightarrow U = (u_1, \dots, u_n)$ vecteurs singuliers à gauche
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale $\rightarrow V = (v_1, \dots, v_m)$ vecteurs singuliers à droite
- $\Sigma \in \mathcal{M}_{n,m}(\mathbb{R})$ diagonale

telles que :

$$A = U \Sigma V^t$$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k$ valeurs singulières

Diagonalisation

$\exists P$ inv tg

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

$$D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{bmatrix} \quad P = \begin{bmatrix} | & & | \\ e_1 & \cdots & e_m \\ | & & | \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & & 0 \end{bmatrix} \quad \text{si } n \geq m$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_n & & 0 \\ & & & \vdots & \\ & & & & 0 \end{bmatrix} \quad \text{si } n \leq m$$

Décomp. SVD de A :

$$A = U \sum V^t$$

diag. $m \times m$

$\swarrow \quad \searrow$
orthogonales

SVD

Diagonalisation

$$B = A^t A$$

$$= (U \sum V^t)^t (U \sum V^t)$$

$$= V \sum^t \underbrace{U^t U}_{I} \sum V^t$$

$$= \sqrt{(\sum^t \sum)} V^t V^{-1}$$

$$B = \boxed{V \underbrace{(\sum^t \sum)}_{\text{diag. de } B} V^{-1}}$$

vec. propres
de B

vec. propres
de B

$$\sum^t \sum = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

vec. propres
de $B = A^t \times A$

$$\sum^t \times \sum = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix} \times \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{pmatrix}$$

$m \times m$

SVD

A rect $m \times m$

$$A = U \cdot \Sigma \cdot V^t$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & & 0 \end{bmatrix} \quad \sigma_1 > \sigma_2 \dots > \sigma_m$$

→ val. singulières
 σ_i

$$\boxed{\sigma_i^2 = d_i}$$

→ U, V orthogonales

$$U = \begin{bmatrix} | & | \\ u_1 & \dots & u_m \\ | & | \end{bmatrix} \quad \text{vecteurs singuliers à gauche}$$

$$V = \begin{bmatrix} | & | \\ v_1 & \dots & v_m \\ | & | \end{bmatrix} \quad \text{vecteurs singuliers à droite}$$

Diagonalisation

$$B = A^t \times A \quad \text{sym}$$



→ valeurs propres de B
 d_i
→ vecteurs propres de B

v_i

DÉCOMPOSITION EN VALEURS SINGULIÈRES

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Preuve :

exercice

Diagonalisation
 $B \rightarrow$ val. propres λ_i
 \downarrow vect. propres e_i
 $B e_i = \lambda_i \cdot e_i$

Propriété :

$$\boxed{Av_i = \sigma_i u_i}$$

$$\boxed{A^t u_i = \sigma_i v_i}$$

$$u_i = \frac{1}{\sigma_i} (A v_i)$$

DÉCOMPOSITION EN VALEURS SINGULIÈRES

- ▶ $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- ▶ $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Propriété : relation avec les valeurs propres

Soient :

- ▶ $\lambda_1, \dots, \lambda_n$ les valeurs propres de $A^t A = \mathcal{B}$
- ▶ $\sigma_1, \dots, \sigma_n$ les valeurs singulières de A

$$\lambda_i = \sigma_i^2$$

Preuve :

Je m'arrête

DÉCOMPOSITION EN VALEURS SINGULIÈRES

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

On suppose $n \geq m$.

Propriété :

$$\|A\|_2 = \|\Sigma\|_2 = \max\{\sigma_i\}$$

car U, V orthogonales
 ne modifient
 pas $\|\cdot\|_2$

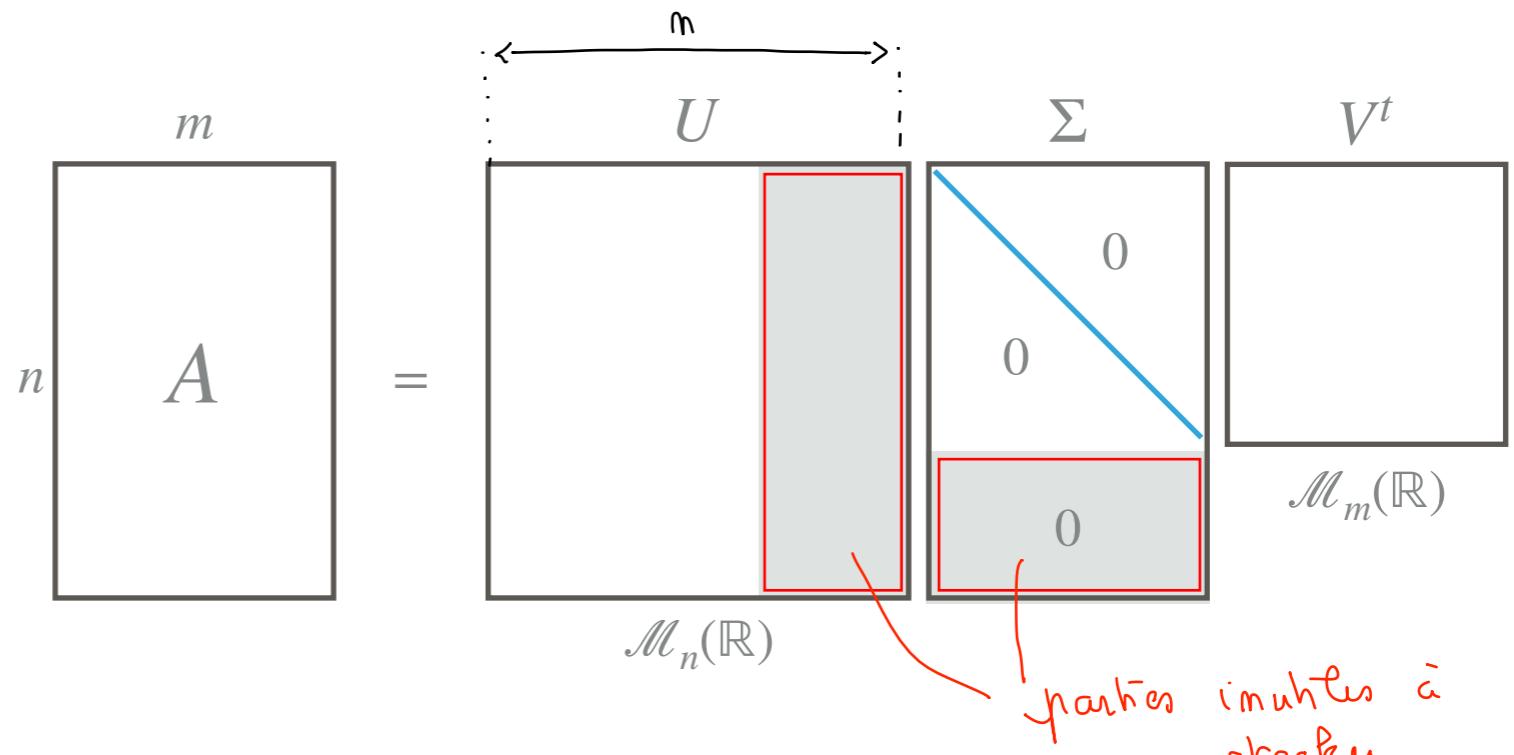
$$\|A\|_2 = \max(\sigma_i)$$

DÉCOMPOSITION EN VALEURS SINGULIÈRES

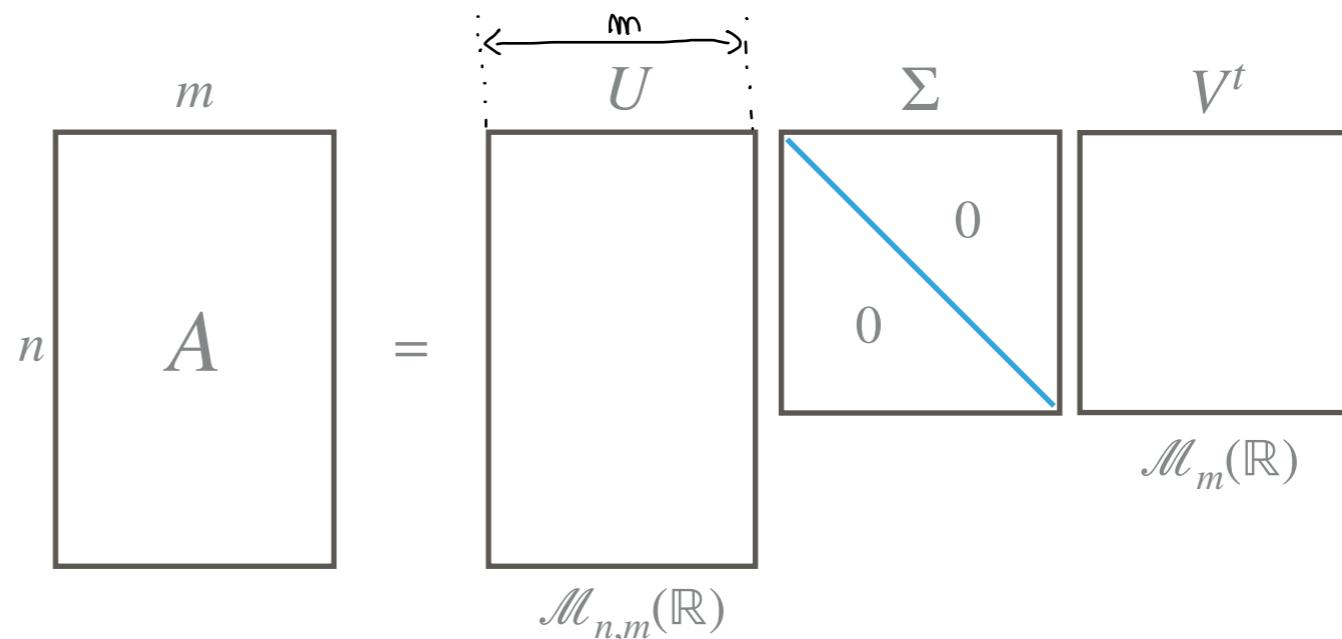
- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

On suppose $n \geq m$.



Représentation économique



SVD de A : (économique)

$$A = \underbrace{\begin{bmatrix} u_1 & \dots & u_m \end{bmatrix}}_U \times \underbrace{\Sigma}_{\sum} \times \underbrace{\begin{bmatrix} v_1^T \\ \vdots \\ v_m^T \end{bmatrix}}_{V^T}$$

$$\begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_m v_m^T \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} u_1 & \dots & u_m \end{bmatrix}}_U \times \underbrace{\begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_m v_m^T \end{bmatrix}}_{V^T}$$

$$\begin{array}{c} \left(\begin{array}{c} 1 \times m \\ \text{blocks} \end{array} \right) \times \left(\begin{array}{c} m \times 1 \\ \text{blocks} \end{array} \right) \\ \left[u_1 \dots u_m \right] \times \begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_m v_m^T \end{bmatrix} = \end{array}$$

prime
block
 $\xrightarrow{*}$

prime
block
 $\xrightarrow{*}$

$$C_{11} = \sum_{k=1}^m \frac{m \text{at-block-1}_{1k} \times m \text{at-block-2}_{k1}}{\sigma_k \sigma_k^T}$$

(*)

$$A = \sum_{k=1}^m \underbrace{m_k \sigma_k}_{\text{IR}} v_k^T u_k$$

$$= \boxed{\begin{bmatrix} m \\ \sum_{k=1}^m \sigma_k u_k v_k^T \end{bmatrix}}$$

DÉCOMPOSITION EN VALEURS SINGULIÈRES

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

On suppose $n \geq m$.

Matrices de rang 1

$$A = \sum_{i=1}^m \sigma_i u_i v_i^t$$



$$\begin{array}{c}
 m \\
 | \\
 n \quad A \\
 | \\
 \end{array}
 =
 \begin{array}{c}
 U \\
 | \\
 u_1 \quad \dots \quad u_m \\
 | \\
 \end{array}
 \Sigma
 \begin{array}{c}
 \sigma_1 & & & 0 \\
 & \ddots & & \\
 0 & & \sigma_m & \\
 \end{array}
 \begin{array}{c}
 V^t \\
 | \\
 v_1^t \quad \dots \quad v_m^t \\
 | \\
 \end{array}$$

$\mathcal{M}_{n,m}(\mathbb{R})$

 $\mathcal{M}_m(\mathbb{R})$

THÉORÈME D'ECKART-YOUNG (APPROXIMATION DE RANG FAIBLE)

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

$$\boxed{u_i v_i^t : \text{rang } 1}$$

$$A = \sum_{i=1}^m \sigma_i u_i v_i^t$$

$$A = \sum \text{matrices}$$

$$u_i : \text{vect } m$$

$$v_i^t : \text{vect } m$$

$$u_i \times v_i^t : \underbrace{\begin{array}{|c|} \hline u_i \\ \hline \end{array}}_m \times \underbrace{\begin{array}{|c|} \hline v_i^t \\ \hline \end{array}}_m = m$$

Théorème d'Eckart-Young

La matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ de rang r la plus proche de A pour les normes $\|\cdot\|_2$ et $\|\cdot\|_{\text{Fro}}$ est :

$$A_r = \sum_{i=1}^{r \leq m} \sigma_i u_i v_i^t$$

$r \leq m$

au moins à r (\otimes)

$r \leq m$

rang r

Meilleure approximation de rang r de A

THÉORÈME D'ECKART-YOUNG (APPLICATIONS)

- ▶ $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- ▶ $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Application aux **images** et vidéos niveaux de gris (cf. TP)

matrice d'entiers



Théorème d'Eckart-Young

La matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ de rang r la plus proche de A pour les normes $\|\cdot\|_2$ et $\|\cdot\|_{\text{Fro}}$ est :

$$A = \sum_{i=1}^r \sigma_i u_i v_i^t$$



Initiale - rang 533



Rang 20

THÉORÈME D'ECKART-YOUNG (APPLICATIONS)

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Application aux images et vidéos niveaux de gris (cf. TP)

Suite de k images I_i de taille $n \times m$

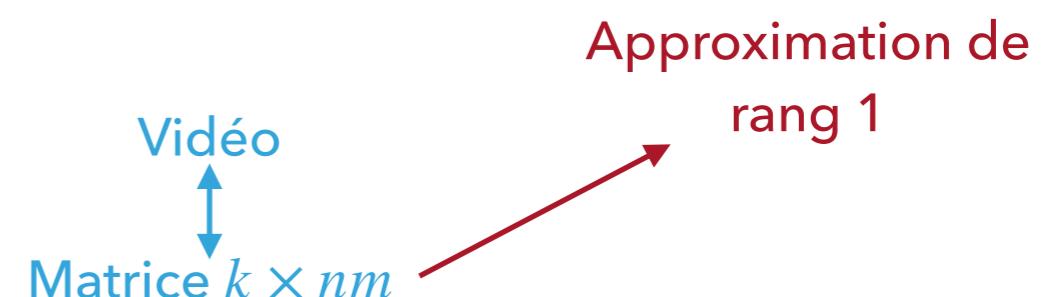
Vectorisation en un vecteur nm

I_i

Théorème d'Eckart-Young

La matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ de rang r la plus proche de A pour les normes $\|\cdot\|_2$ et $\|\cdot\|_{\text{Fro}}$ est :

$$A = \sum_{i=1}^r \sigma_i u_i v_i^t$$



THÉORÈME D'ECKART-YOUNG (APPLICATIONS)

- ▶ $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- ▶ $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Application aux systèmes de recommandation
(d'après V. Emiya)



Netflix Challenge (2007-2009)

Principe

- ▶ Tâches : recommandation de films
- ▶ CineMatch : moteur de reco Netflix
- ▶ But : amélioration des performances de 10%
- ▶ Critère : erreur quadratique
- ▶ Récompense : 1 million de dollars



Données

- ▶ 480 000 utilisateurs
- ▶ 17 000 films
- ▶ 100 millions de scores (de 1 à 5)
- ▶ taux de données connues : 1,3 %
- ▶ test sur 3 millions (les plus récents)
- ▶ 48 000 téléchargement

THÉORÈME D'ECKART-YOUNG (APPLICATIONS)

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Application aux systèmes de recommandation
(d'après V. Emiya)



救人	DATE NIGHT	MICKEY BLUE EYES	THE LOVE GURU	STAR TREK	THE WRATH OF KHAN
1				2	4
	?		1	4	
1		3			
				5	
		4	1		

Recommandation

↓
Complémentation de matrice

THÉORÈME D'ECKART-YOUNG (APPLICATIONS)

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
- $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale

$$A = U\Sigma V^t$$

Y



	5	4			
	1			2	4
		?	1	4	
	1	3			
				5	
		4	1		

Approximation de rang d

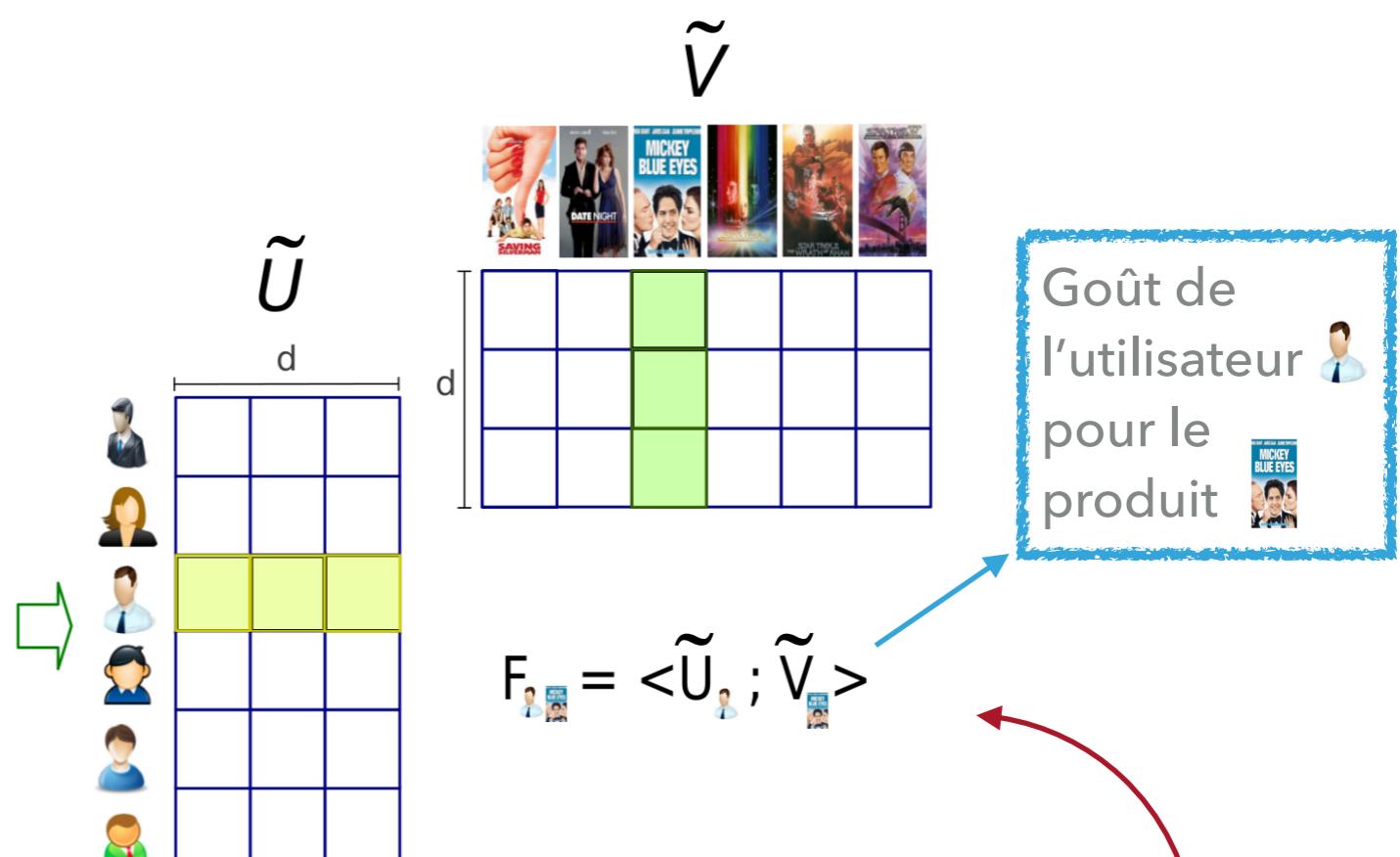
Eckart-Young

Approximation

de rang d

$$Y = \sum_{i=1}^d \sigma_i u_i v_i^t$$

Application aux systèmes de recommandation
(d'après V. Emiya)

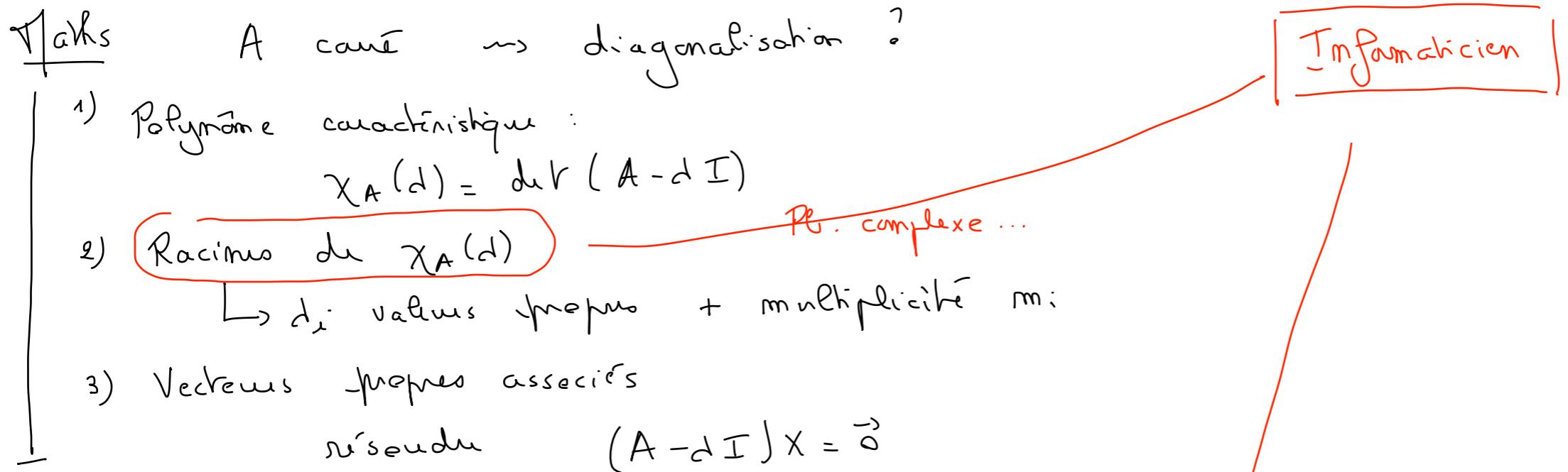


$$F_{\text{user}} = \langle \tilde{U}_{\text{user}}, \tilde{V} \rangle$$

$$\tilde{U} = \begin{bmatrix} | & \dots & | \\ u_1 & \dots & u_d \\ | & \dots & | \end{bmatrix}$$

$$\tilde{V} = \begin{bmatrix} \sigma_1 v_1^t \\ \vdots \\ \sigma_d v_d^t \end{bmatrix}$$

§ Calcul algorithmique diagonalisation / SVD



Informaticien

Algo du calcul des
valeurs propres

autres (complexes)

simple

Algo des
puissances
itérées

→ calcule la
plus grande val. propre (d_1)
+ vect. propre

⚠
→ Hyp: d_1 est de mult. 1

cf. livre de Ciocan

ex: méthode QR
 $O\left(\frac{m^3}{3}\right)$

$A \xrightarrow[\text{QR}]{\text{decouplage}}$ $A = Q_1 \times R_1$

$A_2 = R_1 \times Q_1 \xrightarrow{\text{QR}} A_2 = Q_2 \times R_2$

$A_3 = R_2 \times Q_2 \dots$

Algèbre des puissances itérées

$$A_i \xrightarrow{i \rightarrow \infty} \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{pmatrix}$$

$A : m \times m$ diagonalisable $d_1 : \oplus$ grande val propre (mult. 1)

Hyp.

$$\left. \begin{array}{l} X_0 \text{ "quelconque" (vect de taille } m) \\ X_{i+1} = A \times X_i \end{array} \right\} X_m = A^m \times X_0$$

A diagonalisable $\Rightarrow \exists$ base de vect. propres e_1, \dots, e_m e_i : vect. propre $A \cdot e_i = d_i \cdot e_i$

Donc $X_0 = x_1 \cdot e_1 + x_2 \cdot e_2 + \dots + x_m \cdot e_m$

$$X_1 = A \times X_0 = A \times (\underbrace{x_1 e_1 + \dots + x_m e_m}_{\text{a des coords ds la base } e_1, \dots, e_m}) = x_1 \cdot \underbrace{Ae_1}_{d_1 e_1} + x_2 \cdot \underbrace{Ae_2}_{d_2 e_2} + \dots + x_m \cdot \underbrace{Ae_m}_{d_m e_m}$$

$$\boxed{X_1 = x_1 d_1 e_1 + \dots + x_m d_m e_m}$$

$$X_2 = A \times X_1 = A \left(\underbrace{x_1 d_1 e_1}_{d_1 e_1} + \underbrace{x_2 d_2 e_2}_{d_2 e_2} + \dots + \underbrace{x_m d_m e_m}_{d_m e_m} \right) = x_1 d_1 \cdot \underbrace{Ae_1}_{d_1 e_1} + \dots + x_m d_m \cdot \underbrace{Ae_m}_{d_m e_m}$$

$$\boxed{X_2 = x_1 d_1^2 e_1 + \dots + x_m d_m^2 e_m}$$

$$x_k = \alpha_1 d_1^k e_1 + \dots + \alpha_m d_m^k e_m$$

$$d_1 > d_i$$

$$\frac{x_k}{d_1^k} = \alpha_1 e_1 + \alpha_2 \left(\frac{d_2}{d_1} \right)^k e_2 + \dots + \alpha_m \left(\frac{d_m}{d_1} \right)^k e_m$$

$\downarrow k \rightarrow \infty$

0

$\downarrow k \rightarrow \infty$

0

$$x_k \xrightarrow{k \rightarrow \infty} \alpha_1 e_1 d_1^k$$

||

$$x_k \xrightarrow{\substack{k \rightarrow \infty \\ R}} \underbrace{\alpha_1 d_1^k}_{R} \cdot \vec{e}_1$$

algo des puissances itérées

cercle: définition du vect. propre

$\Rightarrow d_1 / e_1$

Algorithme 2 Algorithme des puissances itérées (avec normalisation).

Entrées: Matrice $A \in \mathbb{R}^{d \times d}$, nombre d'itérations N

Sorties: Plus grande valeur propre λ de A et vecteur propre u associé.

x nombre ...

1: Initialiser $y_0 \in \mathbb{R}^d$ aléatoirement

2: $x_0 \leftarrow \frac{y_0}{\|y_0\|_2}$

3: pour n de 1 à N faire

4: $y_n \leftarrow Ax_{n-1}$

5: $x_n \leftarrow \frac{y_n}{\|y_n\|_2}$

6: fin pour

7: $u \leftarrow x_N$ Approx du vect. propre

8: $\lambda \leftarrow \langle Au, u \rangle$ Approx. de d_1

9: renvoyer λ, u

normalisation du vect. propre construit

stabilité numérique

$x_k // e_1$

$$A \rightsquigarrow \begin{cases} d_1 \\ e_1 \end{cases} \text{ (approx)}$$

α_m maintenu de manne 1

$\downarrow n \rightarrow \infty$

$$\dots \times e_1$$

1

$$\Rightarrow \alpha_m \xrightarrow{n \rightarrow \infty} e_1$$

PB: calculer les autres d_i / e_i) complexe.

SVD

?

$A : n \times m$

$$\left(\begin{array}{c} \sigma_i \\ u_i \\ v_i \end{array} \right) ?$$

{ SVD1 }

$$\begin{cases} \sigma_1 = \sqrt{\lambda_1} \\ v_1 = e_1 \\ u_1 = \frac{1}{\sigma_1} \cdot A v_1 \end{cases}$$

$$| A \cdot v_1 = \sigma_1 u_1 |$$

$$B = A^T \times A$$

plus. iténés

$$\left(\begin{array}{c} \lambda_1 \\ e_1 \end{array} \right)$$



et les autres σ_i, u_i, v_i ?

Th. Eckart- Young

$$A = \sum_{i=1}^m \sigma_i u_i v_i^T = \sigma_1 u_1 v_1^T + \sum_{i=2}^m \sigma_i u_i v_i^T$$

$$\Rightarrow \underbrace{A - \sigma_1 u_1 v_1^T}_{A'} = \sum_{i=2}^m \sigma_i u_i v_i^T$$

A' $\xrightarrow{\text{SVD}}$

$$\sigma_i u_i v_i^T \quad i > 2$$

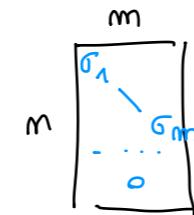
SVD1

plus grande val. sing σ_2
vect. sing u_2, v_2

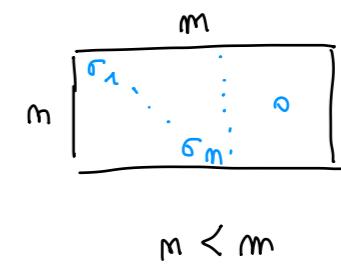
SVD - algorithme

l'itération i du 1 à $\min(\underline{m}, \underline{n})$

$(\sigma_i, u_i, v_i) \leftarrow \text{SVD1}(A)$
 Stépper $(\sigma_i, u_i, v_i) \leftarrow (\sigma_i, u_i, v_i)$
 $A \leftarrow A - \sigma_i u_i v_i^T$



$m > n$



$m < n$

l'itération

A n'est pas fini le calcul de la SVD
 que certains des valeurs / vecteurs propres

- plus simple
- plus efficace

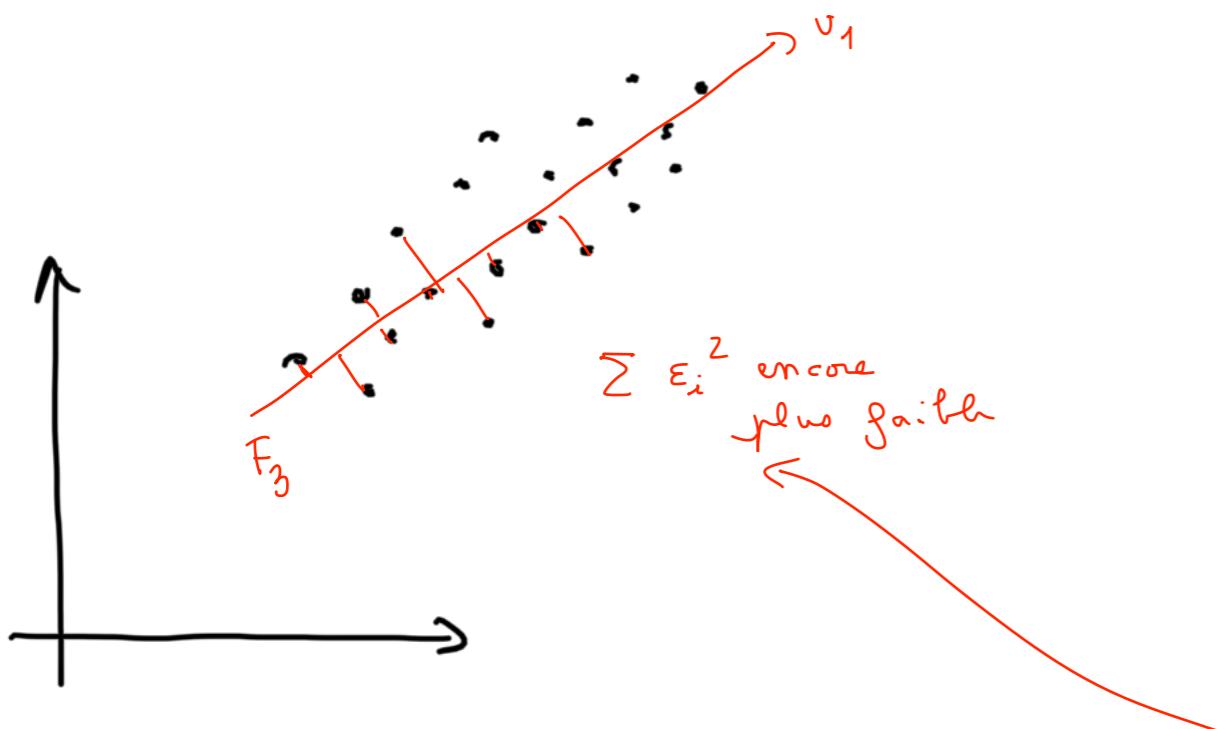
ANALYSE EN COMPOSANTES PRINCIPALES

PRÉSENTATION DU PROBLÈME

Application emblématique de la SVD

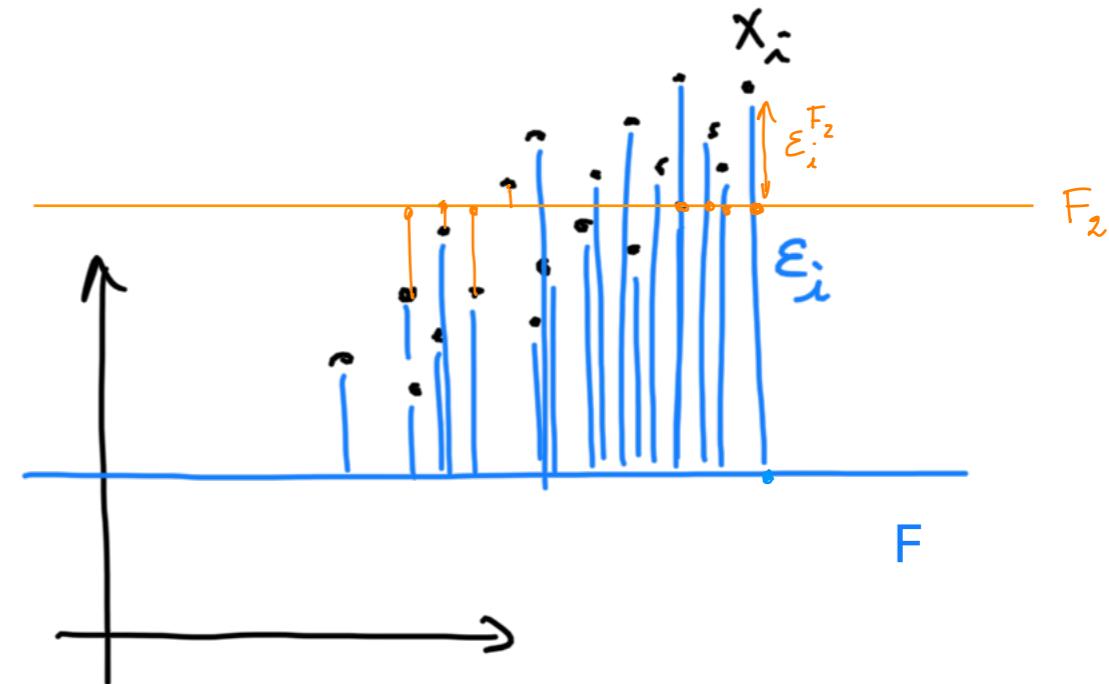
Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- Trouver le sous-espace affine F de dimension k minimisant l'erreur de projection



Exemple $n = 2, k = 1$

comp. - principal



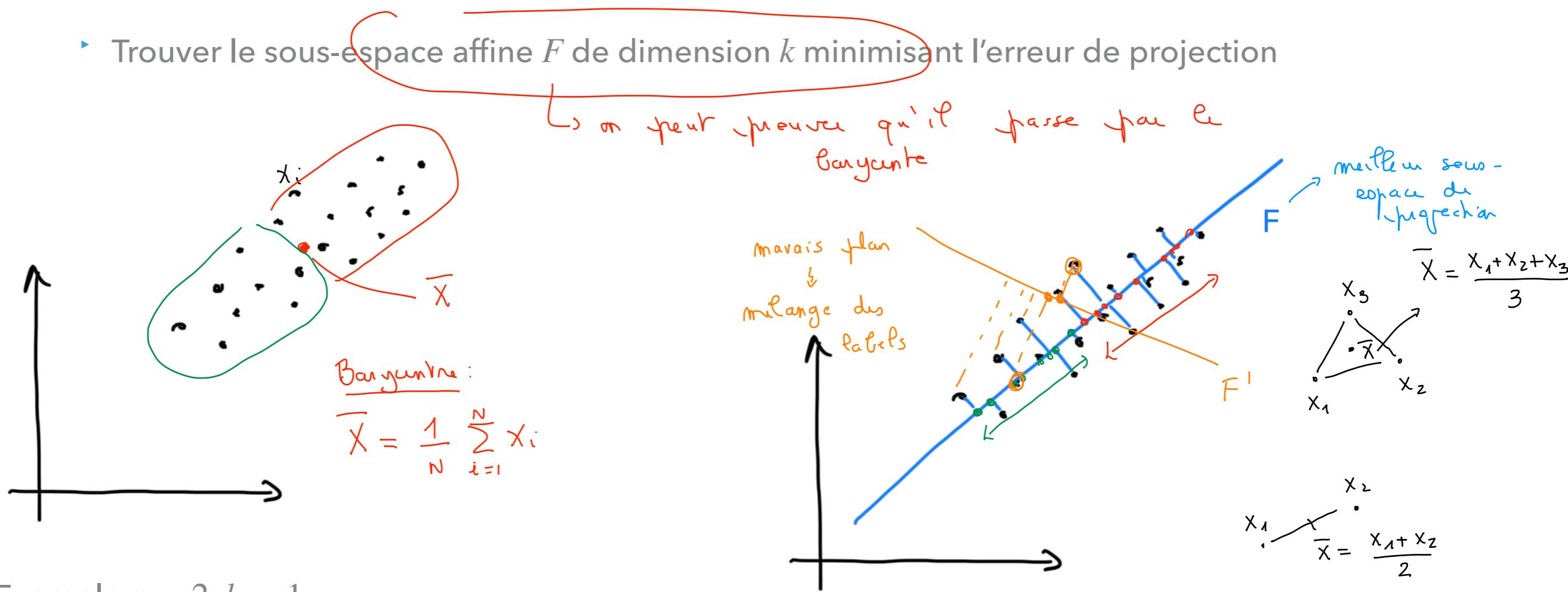
$$\begin{aligned}\sum \varepsilon_i^2 &\text{ minimale} \\ \sum (\varepsilon_i^F)^2 &< \sum (\varepsilon_i^{F^\perp})^2\end{aligned}$$

PRÉSENTATION DU PROBLÈME

Application emblématique de la SVD

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- Trouver le sous-espace affine F de dimension k minimisant l'erreur de projection



Exemple $n = 2, k = 1$

PRÉSENTATION DU PROBLÈME

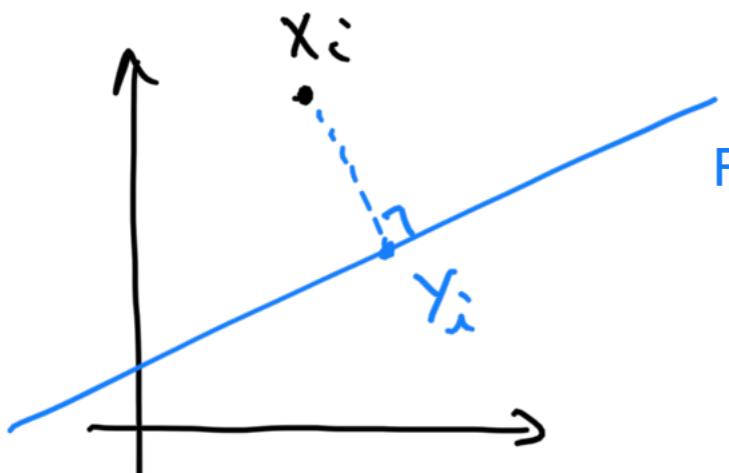
Application emblématique de la SVD

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ▶ Trouver le sous-espace affine F de dimension k minimisant l'erreur de projection

Minimiser :

$$\sum_{i=1}^N \|X_i - Y_i\|_2^2$$



Visualisation
des données

Analyse par
réduction de
dimension

PRÉSENTATION DU PROBLÈME

Application emblématique de la SVD

Etant donnés N points données $\{X_i \in \mathbb{R}^n ; i = 1 \dots N\}$

- ▶ Trouver le sous-espace affine F de dimension k minimisant l'erreur de projection

Exemple : données relatives à des vin

- Alcohol
- Malic acid
- Ash
- Alcalinity of ash
- Magnesium
- Total phenols
- Flavanoids
- Nonflavanoid phenols
- Proanthocyanins
- Color intensity
- Hue
- OD280/OD315 of diluted wines
- Proline

3 classes

14 paramètres / 178 vins

PRÉSENTATION DU PROBLÈME

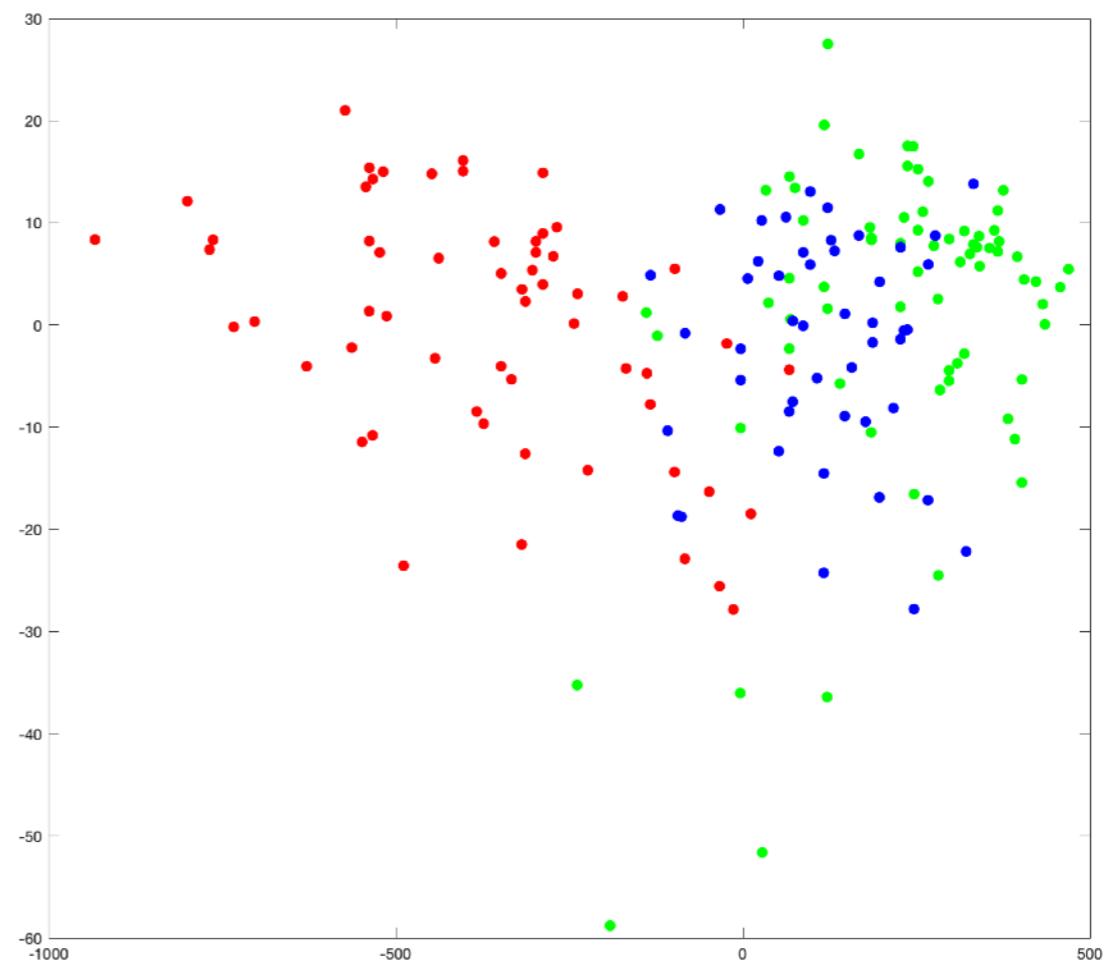
Application emblématique de la SVD

Etant donnés N points données $\{X_i \in \mathbb{R}^n ; i = 1 \dots N\}$

- ▶ Trouver le sous-espace affine F de dimension k minimisant l'erreur de projection

Exemple : données relatives à des vin

Représentation des classes pour $k = 2$

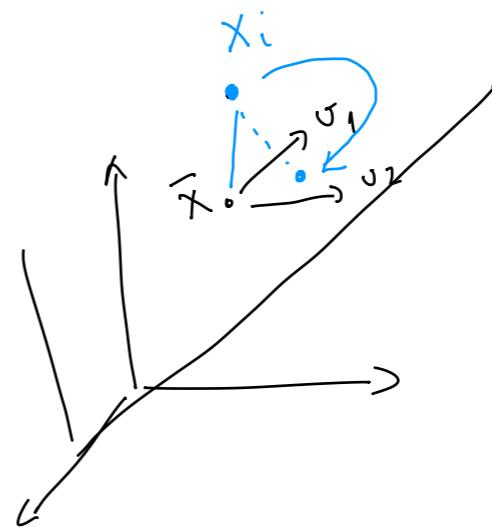


$$A = \begin{pmatrix} \vdots & & \\ -(x_i - \bar{x})^T & \dots \\ \vdots & & \end{pmatrix} \xrightarrow{\text{SVD}} \cancel{\sum_i v^T} \quad \begin{array}{l} \text{Projeter en} \\ \text{dim 2} \\ \downarrow \\ \bar{x}, \vec{v}_1, \vec{v}_2 \end{array}$$

m^{me}
 de
 données

dim
 x
 données

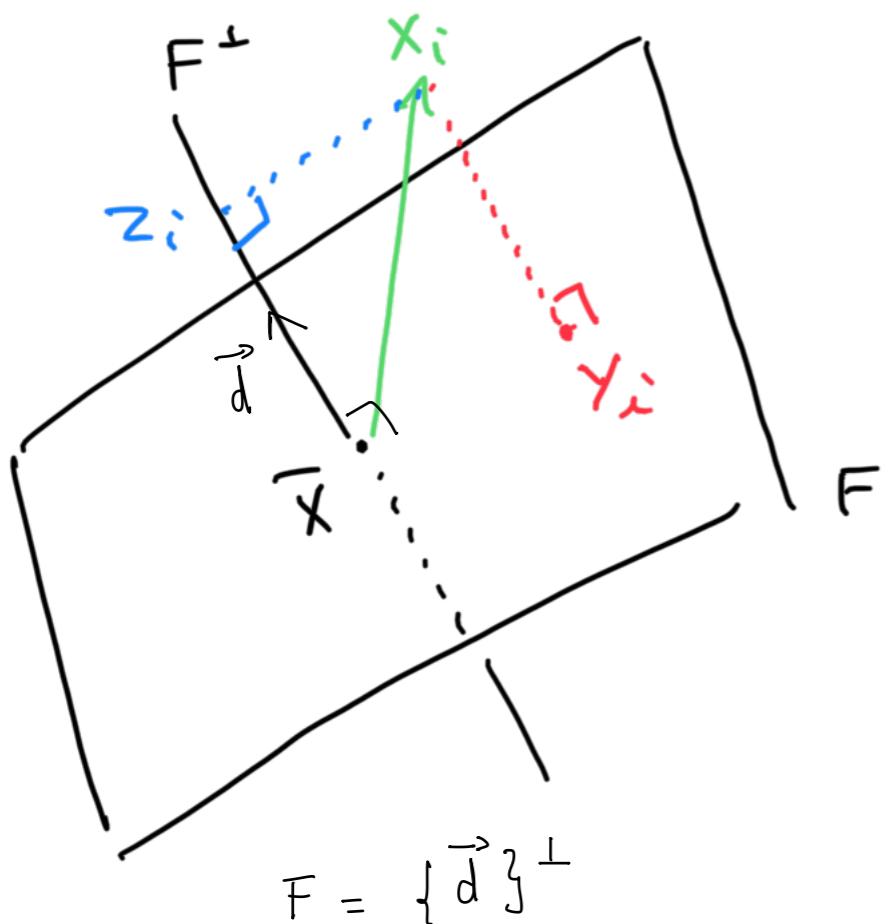
ici: 178 x 14



ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



Minimiser :

$$\frac{1}{N} \sum_{i=1}^N \|X_i - Y_i\|_2^2$$

$\underbrace{\phantom{\frac{1}{N} \sum_{i=1}^N \|X_i - Y_i\|_2^2}}_{I_F}$

Inertie par rapport au sous-espace F

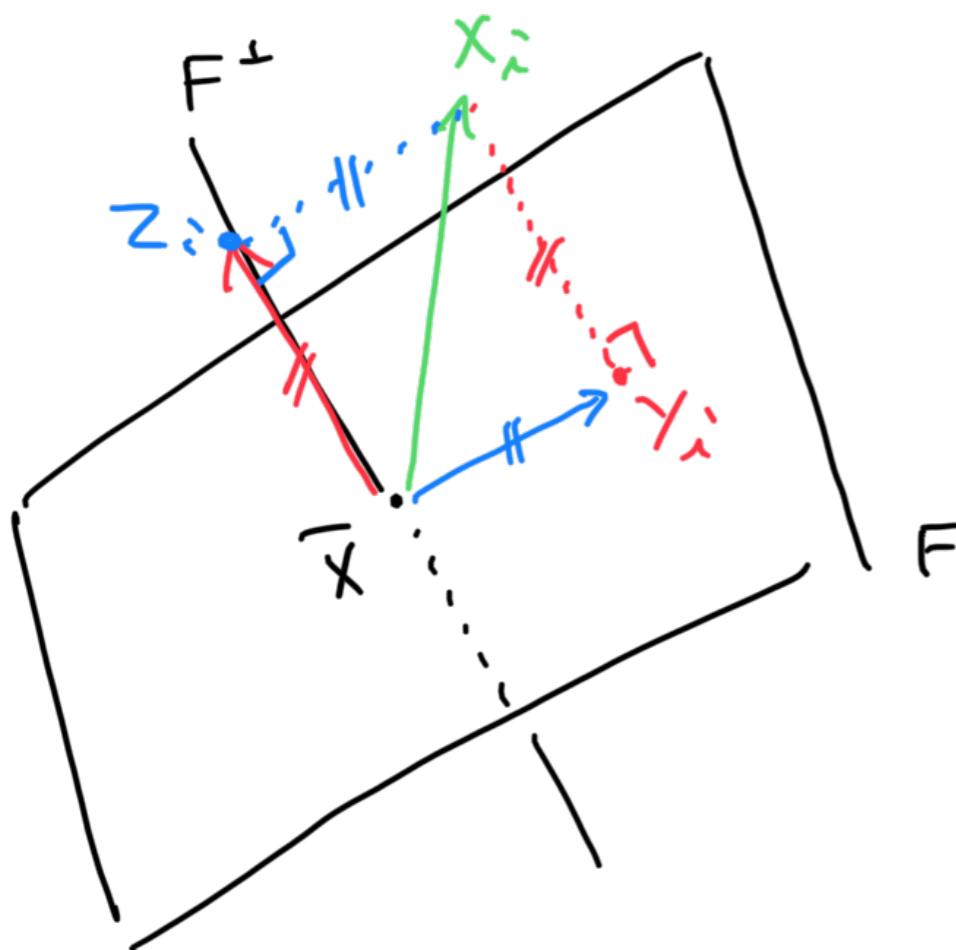
$$I_G = \frac{1}{N} \sum_{i=1}^N \|X_i - \bar{X}\|_2^2$$

Inertie globale

ANALYSE EN COMPOSANTES PRINCIPALES

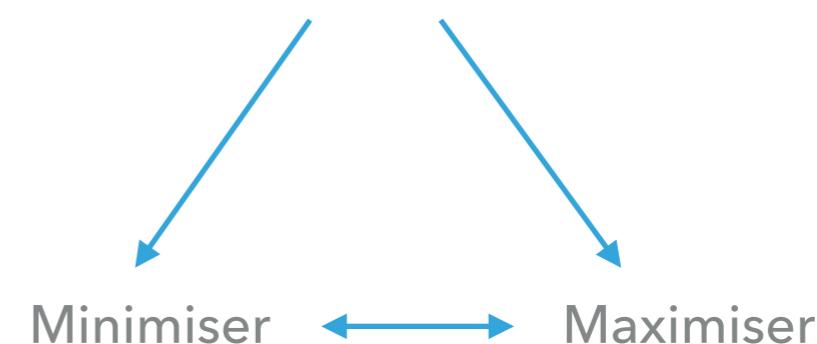
Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



Théorème de Huygens :

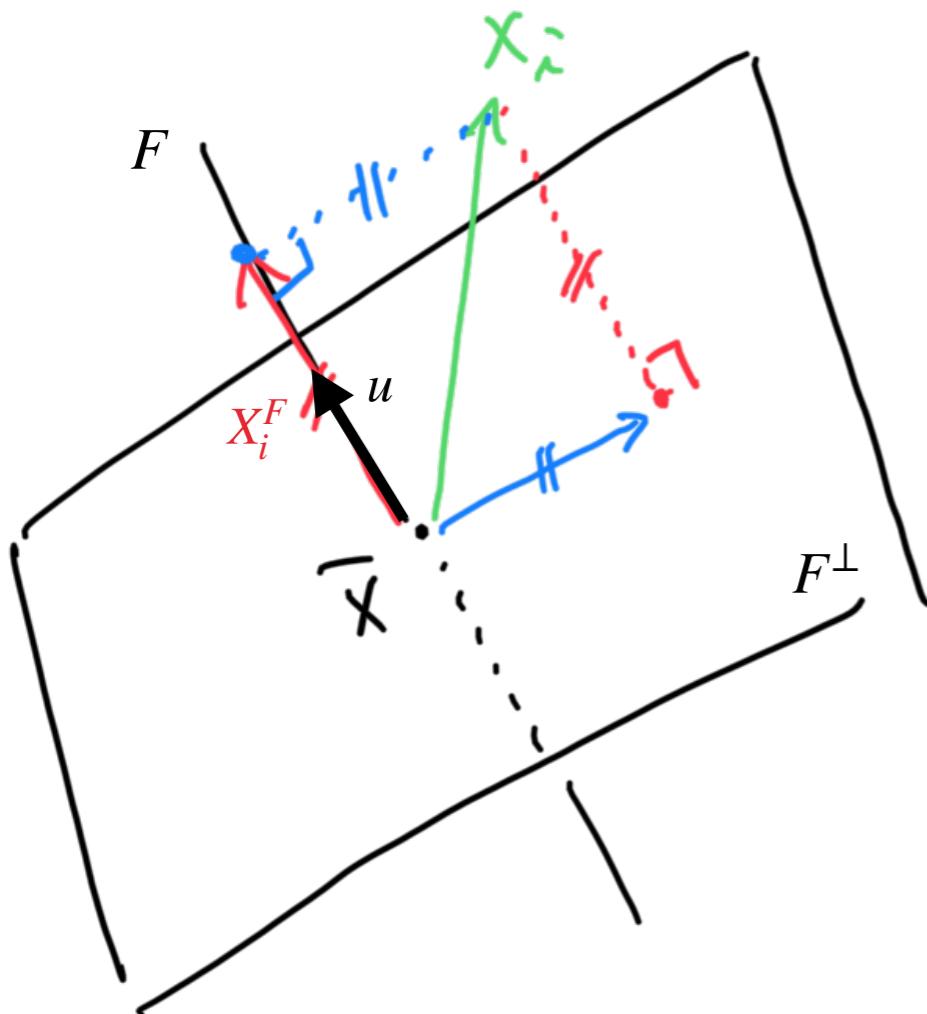
$$I_G = I_F + I_{F^\perp}$$



ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



On cherche $F = \mathcal{D}(\bar{X}, u)$ tel que :

$$I_F \text{ maximum} \leftrightarrow I_{F^\perp} \text{ minimal}$$

Si $\|u\| = 1$, le projeté de X sur F :

$$X_i^F = \langle X_i - \bar{X}, u \rangle u$$

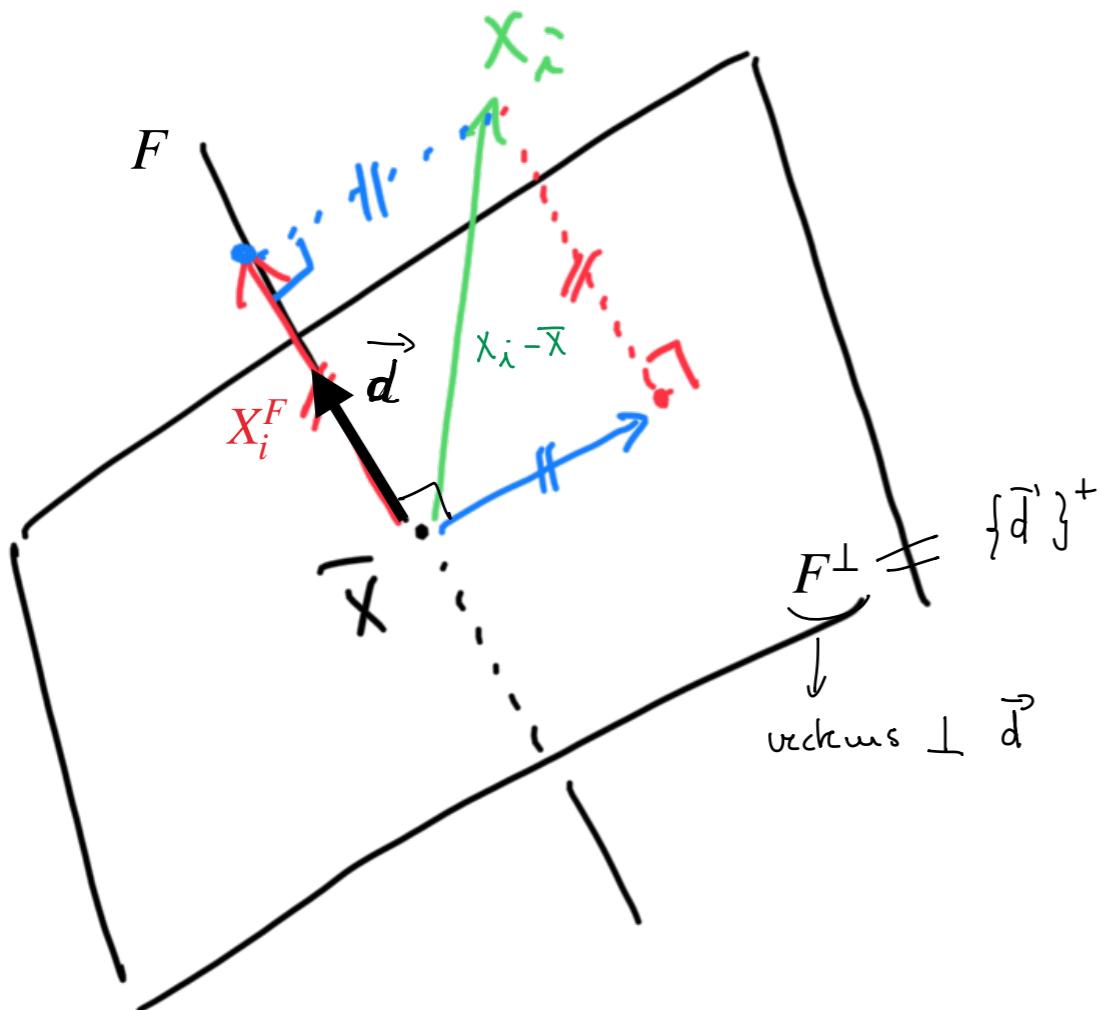
Donc :

$$I_{F^\perp} = \frac{1}{N} \sum_i \langle X_i - \bar{X}, u \rangle^2$$

ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



$$I_{F^\perp} = \frac{1}{N} \sum_i \langle X_i - \bar{X}, u \rangle^2$$

=

$$I_{F^\perp} = d^t \left(\frac{1}{N} \sum_i (X_i - \bar{X})(X_i - \bar{X})^\perp \right) d$$

vec $\sqrt{\bar{X}X_i}$

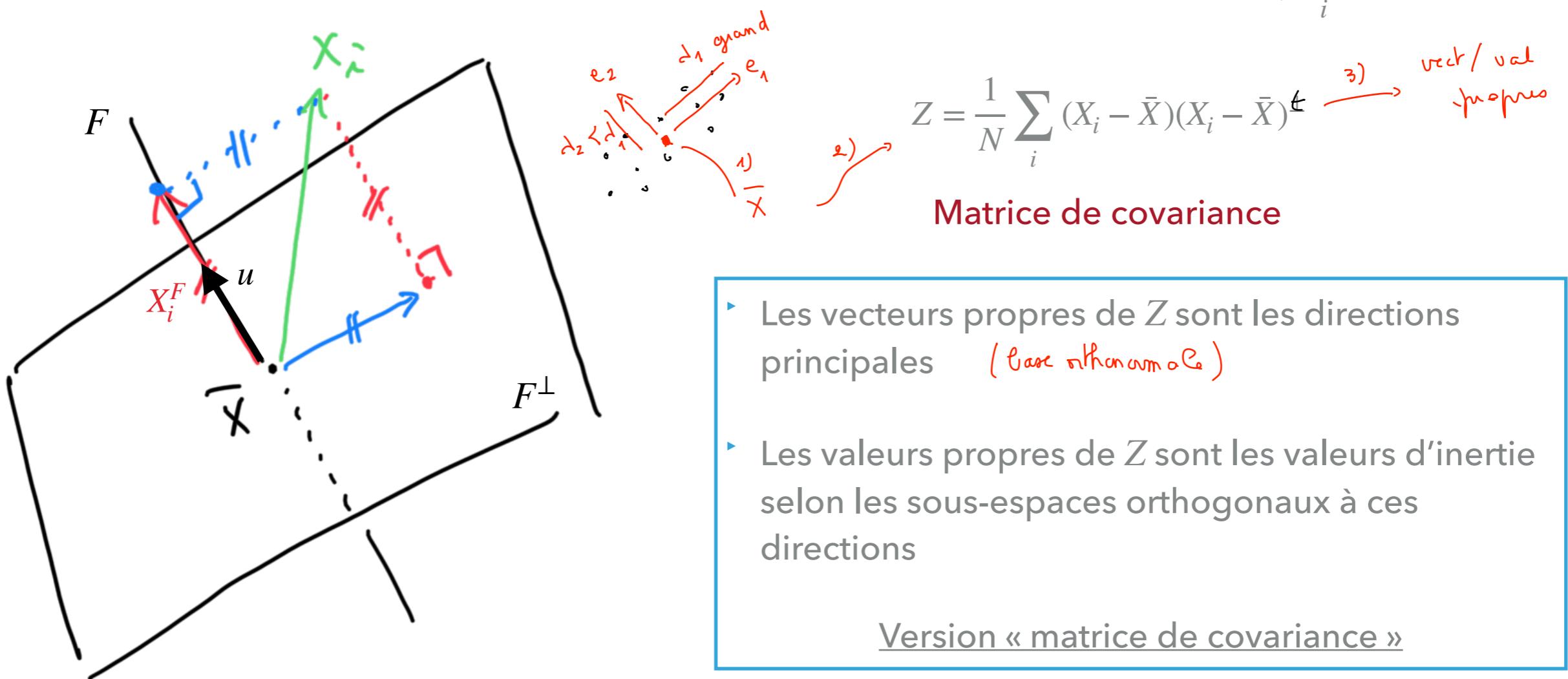
Z

Matrice de covariance

ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

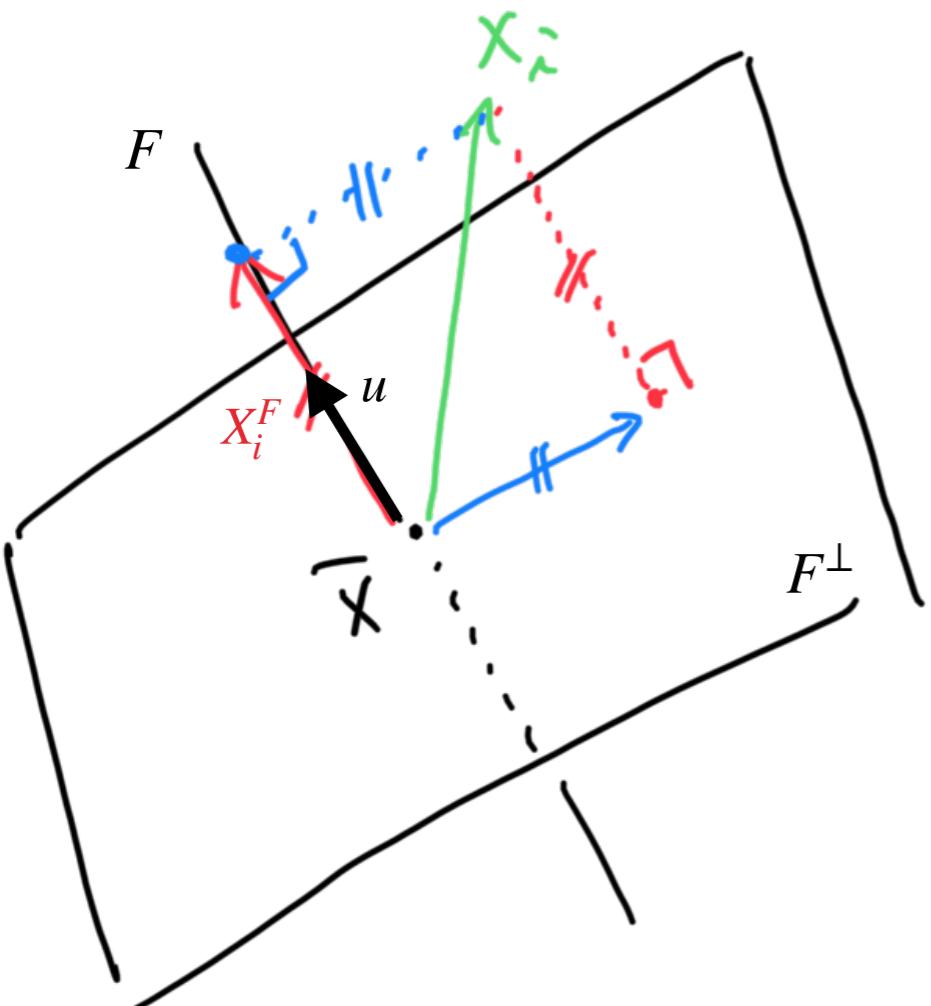
- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



Soit :

$$Z = \frac{1}{N} \sum_i (X_i - \bar{X})(X_i - \bar{X})^t$$

$$A = \begin{bmatrix} -(X_1 - \bar{X})^t \\ \vdots \\ -(X_N - \bar{X})^t \end{bmatrix}$$

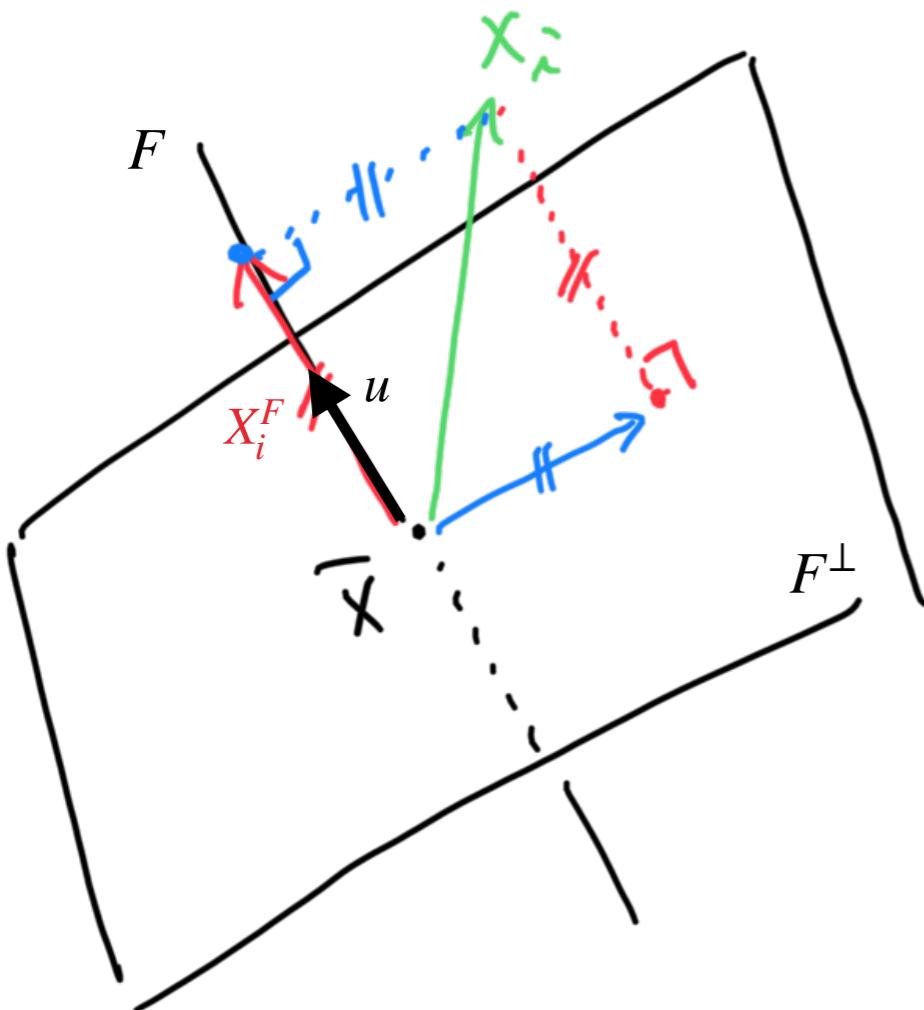
Produit
par
blocs

$$Z = \frac{1}{N} A^t A$$

ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$



$$A = \begin{bmatrix} (X_1 - \bar{X})^t \\ \vdots \\ (X_N - \bar{X})^t \end{bmatrix} \xrightarrow{\text{SVD}} A = U\Sigma V^t$$

- Les vecteurs de V sont les directions principales
- Les valeurs singulières sont les racines des valeurs d'inertie selon les sous-espaces orthogonaux à ces directions

Version « SVD »

ANALYSE EN COMPOSANTES PRINCIPALES

Etant donnés N points données $\{X_i \in \mathbb{R}^n; i = 1 \dots N\}$

- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$

(B)

$$Z = \frac{1}{N} \sum_i (X_i - \bar{X})(X_i - \bar{X})^\perp$$

$Z = \frac{A^t A}{N}$

$$A = \begin{bmatrix} -(X_1 - \bar{X})^t \\ \vdots \\ -(X_N - \bar{X})^t \end{bmatrix}$$

SVD

$$A = U\Sigma V^t$$

Valeurs propres

$$\lambda_1 \geq \dots \geq \lambda_n$$



$$I_{F_i^\perp} = \lambda_i$$

Vecteurs propres associés

$$d_1, \dots, d_n$$



Directions principales

$$F_i = \mathcal{D}(\bar{X}, d_i)$$

Valeurs singulières

$$\sigma_1 \geq \dots \geq \sigma_n$$



$$I_{F_i^\perp} = \sigma_i^2$$

Vecteurs singuliers à droite

$$v_1, \dots, v_n$$



Directions principales

$$F_i = \mathcal{D}(\bar{X}, v_i)$$

ANALYSE EN COMPOSANTES PRINCIPALES

Exemple sur un nuage de points 3D

```
clear
clc
clf
hold on

% Génératation d'un nuage de points aléatoire
a = 1 ;
b = 2 ;
c = 4 ;
N = 40 ;
O = RandOrthMat(3) ;

M = 2*(rand(3,N) - .5) .* [a;b;c] ; % tirage aléatoire / homothétie
M = O*M ; % application de la matrice de transformation orthogonale

% Affichage
plot3(M(1,:), M(2,:), M(3,:), '.b', 'MarkerSize', 20)

% Calcul du barycentre
Xbar = sum(M,2) ;
Xbar = Xbar/N ;

% Affichage
plot3(Xbar(1,:), Xbar(2,:), Xbar(3,:), '.r', 'MarkerSize', 20)

% ACP du nuage
A = (M-Xbar)';
[U,S,V] = svd(A, 'econ') ;

% Affichage des composantes
VV = V*S ;
colors = diag([1.;1.;1.]);
for i=1:3
    tmp = [Xbar,Xbar+VV(:,i)] ;
    plot3(tmp(1,:), tmp(2,:), tmp(3,:), 'Color', colors(:,i)) ;
end

axis equal
```