

POLYTECH INFORMATIQUE

3ÈME ANNÉE

Alexandra Bac

NUMERICAL METHODS

COURS 3

VALEURS SINGULIÈRES, ANALYSE EN COMPOSANTES PRINCIPALE

MATRICES DIAGONALISABLES ...

Matrices $n \times n$

MATRICES
DIAGONALISABLES

MATRICES NON
DIAGONALISABLES

Matrices $n \times m$

Matrices toujours diagonalisables

Symétriques

Cause 1

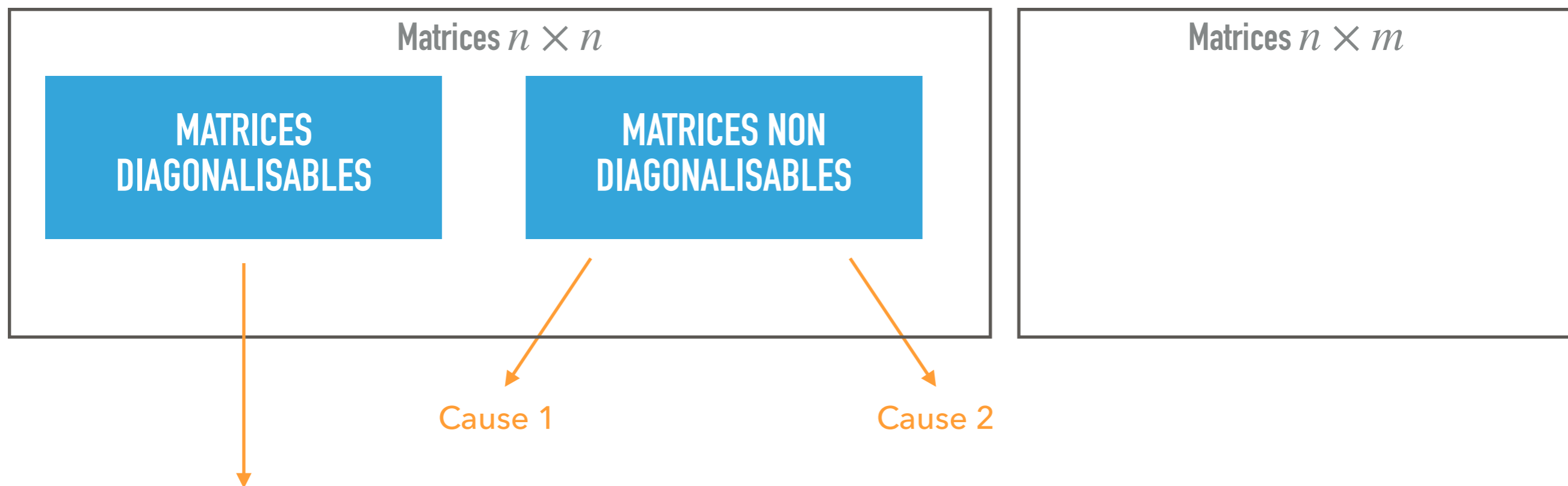
pas carrée

Cause 2

si carrée

↓
pas assez de
vect. propres pour
fabriquer P et C

MATRICES DIAGONALISABLES ...

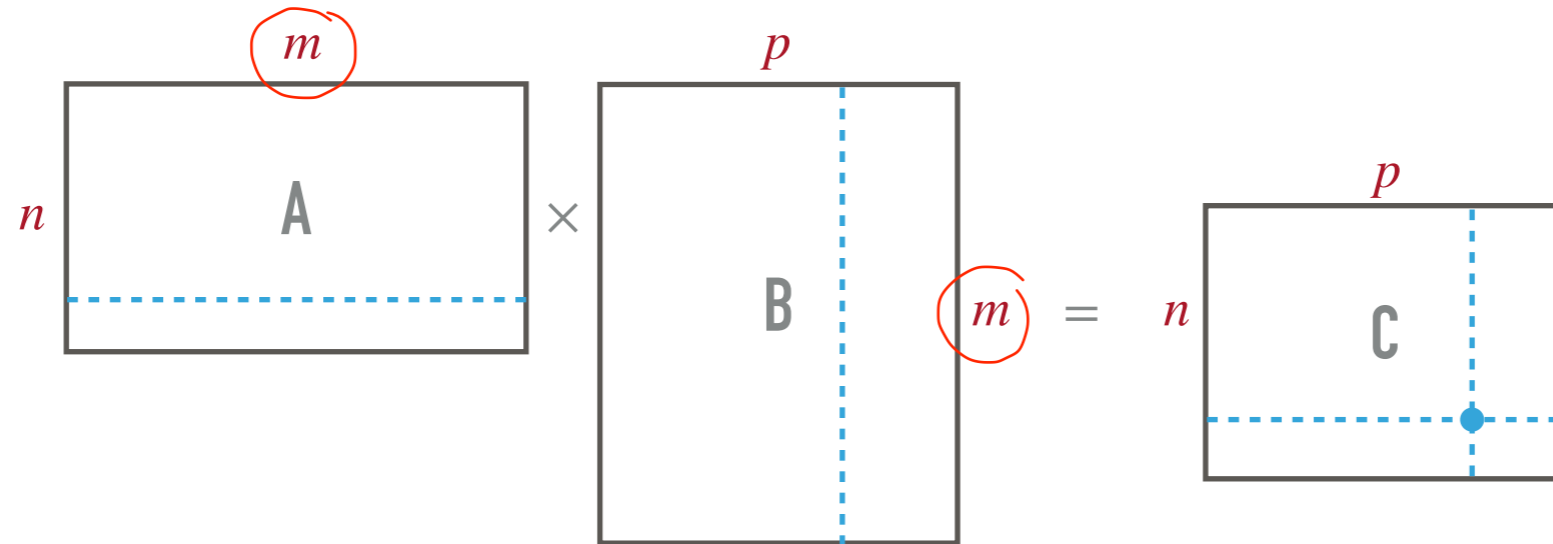


Matrices toujours diagonalisables

Matrices symétriques :
diagonalisables en base
orthonormale.

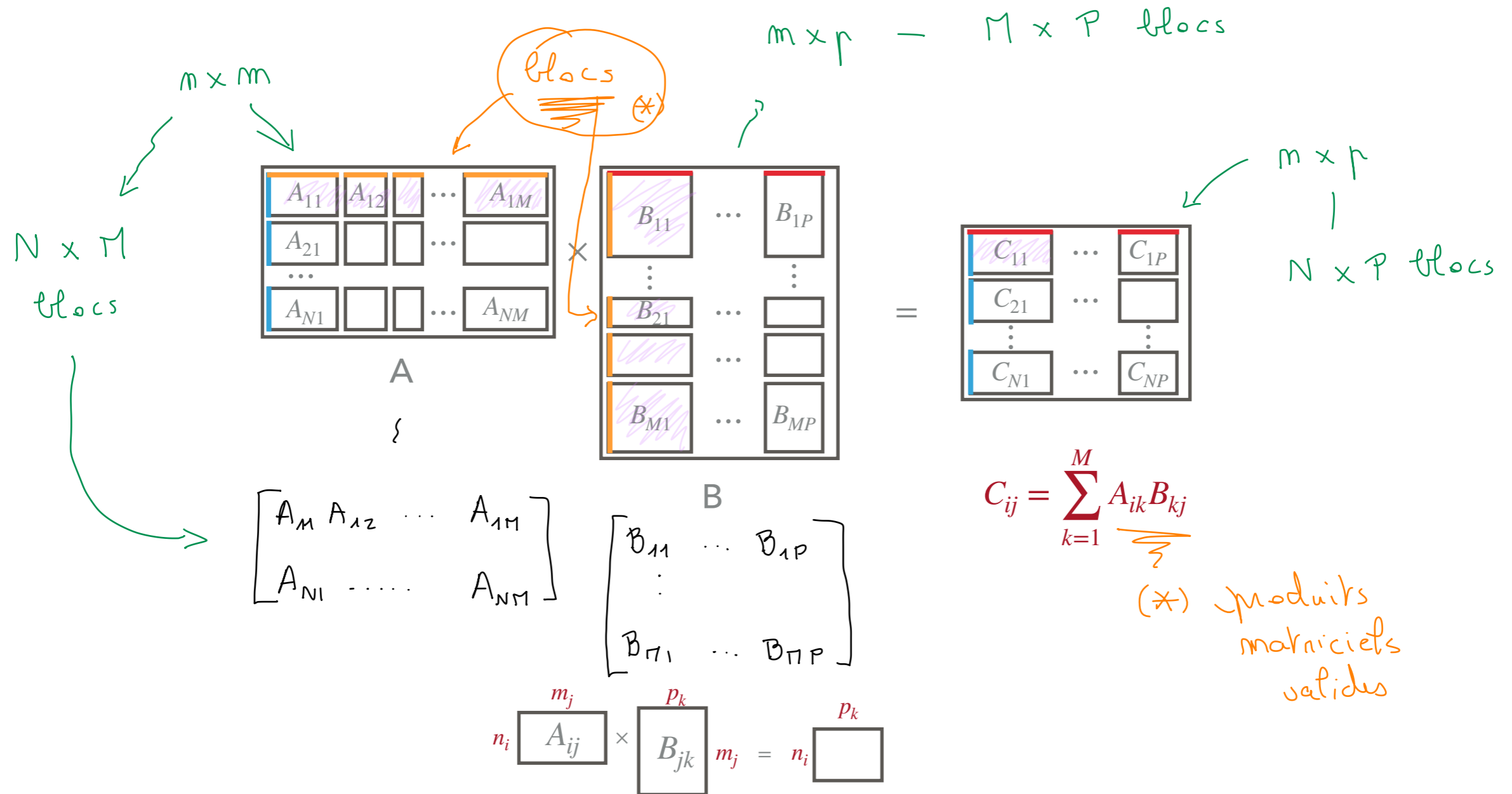
PRODUIT MATRICIEL PAR BLOCS

$$C = A \times B$$

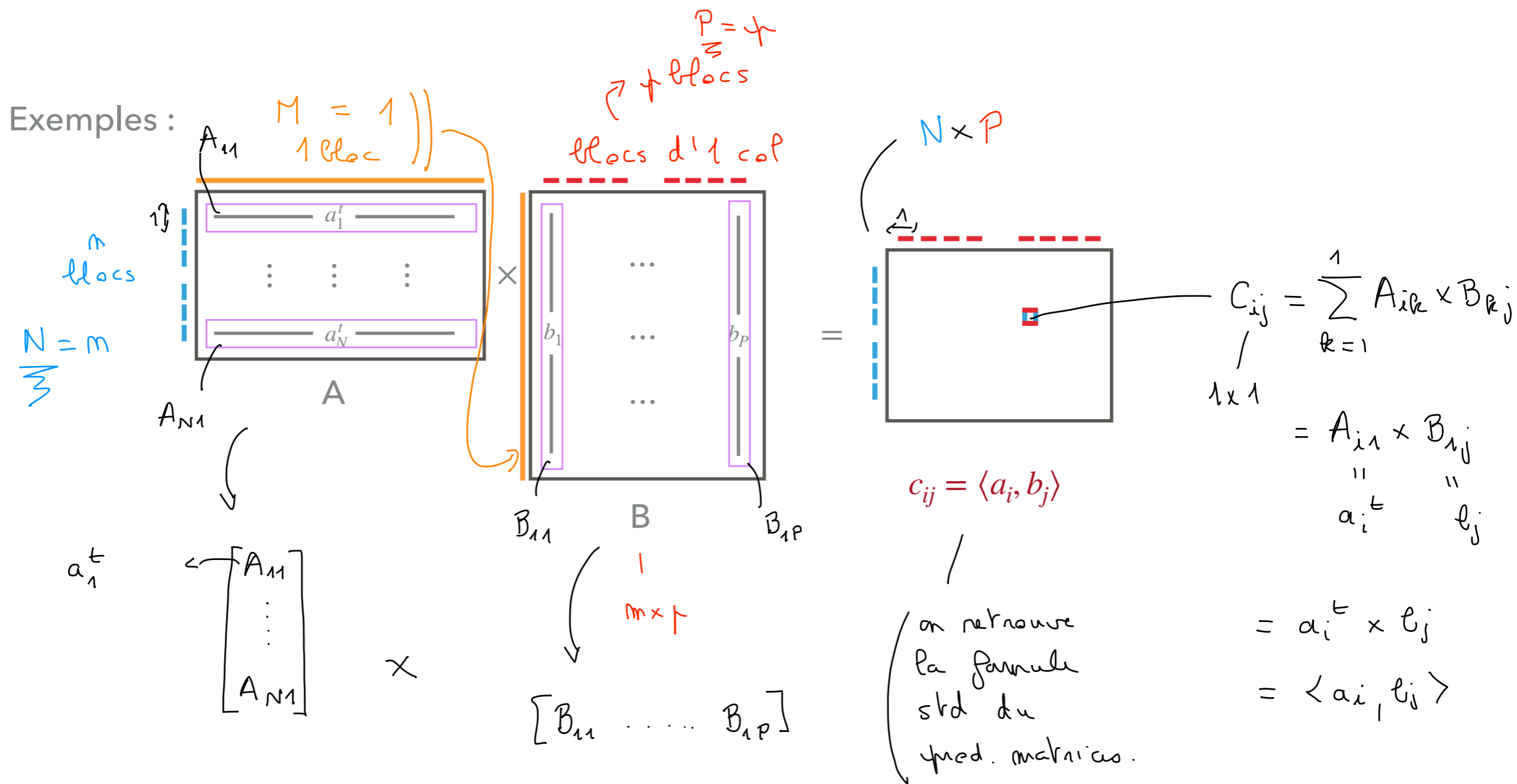


$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$$

PRODUIT MATRICIEL PAR BLOCS

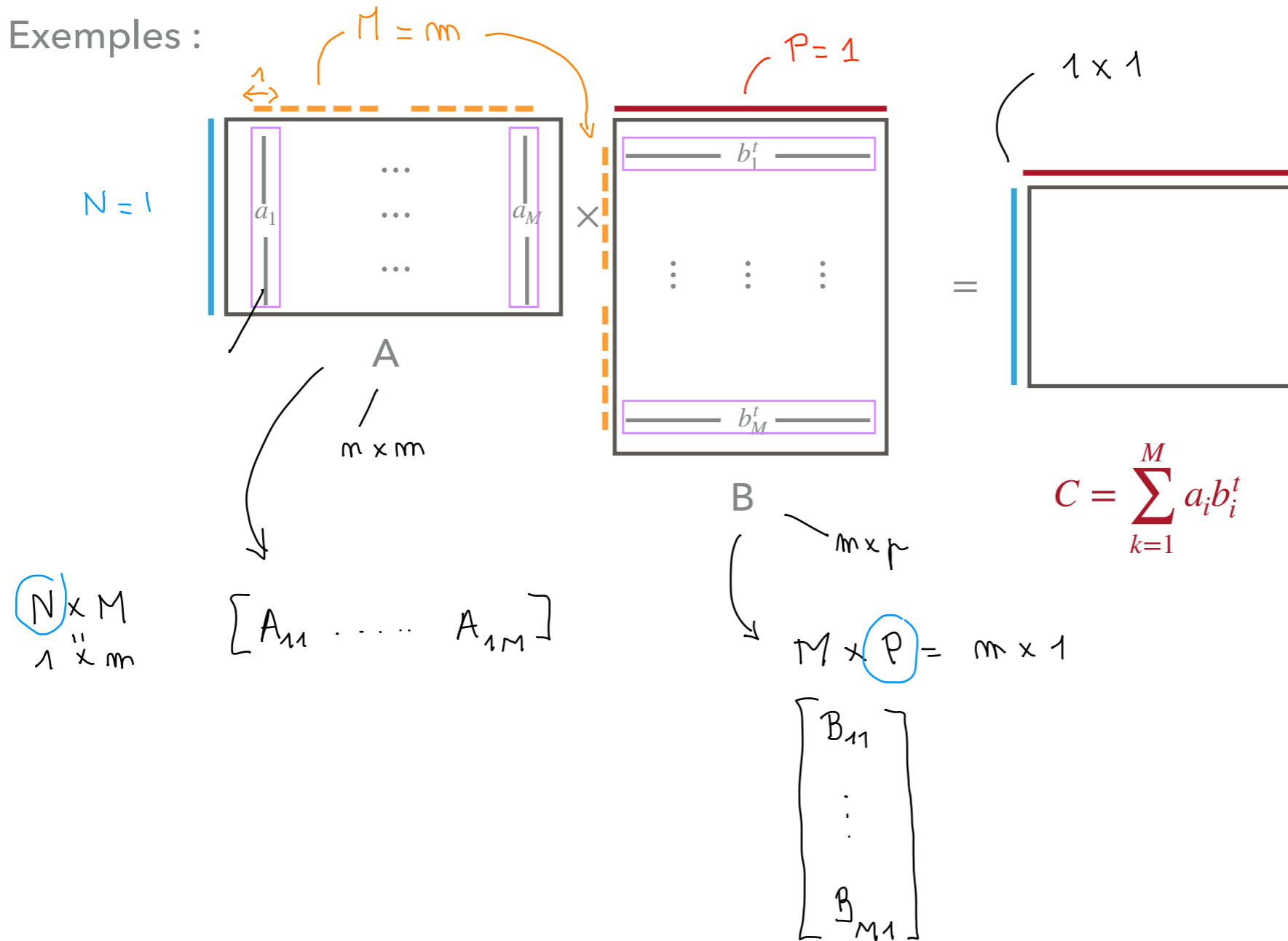


PRODUIT MATRICIEL PAR BLOCS



PRODUIT MATRICIEL PAR BLOCS

Exemples :



$C_{11} = \sum_{k=1}^M A_{1k} \times B_{k1}$
 $C_{11} = \sum_{k=1}^m a_k \times b_k^t$
 $= \sum_{k=1}^m \underbrace{a_k \times b_k^t}_{\text{matrice}}$

k col de A
 a_k
 k ligne de B
 b_k^t

VALEURS SINGULIÈRES

DIAGONALISATION

canée

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

DIAGONALISABLE

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

e_i vecteur propre

$$Ae_i = \lambda_i e_i$$

$$A = PDP^{-1}$$

$$P^{-1}AP = \begin{array}{|c|} \lambda_1 & \\ & \ddots & \\ & & \ddots & \\ & & & \lambda_n \end{array}$$

D

$$\text{où } P^{-1} = \begin{array}{|c|c|c|} e_1 & \dots & e_n \end{array}$$

DIAGONALISATION

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

DIAGONALISABLE

Matrices rectangulaires ?

Matrices non diagonalisables ?

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



e_i vecteur propre

$$Ae_i = \lambda_i e_i$$



$$P^{-1}AP = \begin{array}{|c|} \lambda_1 & \\ & \ddots & \\ & & \ddots & \\ & & & \lambda_n \end{array}$$

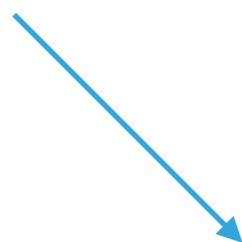
D

$$\text{où } P^{-1} = \begin{array}{|c|c|c|} e_1 & \dots & e_n \end{array}$$

$$A = PDP^{-1}$$

MATRICES RECTANGULAIRES

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



$$B = \begin{matrix} A^T & A \\ \swarrow & \searrow \\ m \times m & n \times n \end{matrix} \begin{matrix} \rightarrow \text{carrée} \\ \rightarrow \text{symétrique} \end{matrix} \quad m \times m$$

D diagonale :

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

diagonalisation ...
 → fabriquer B carrée
 → diagonalisable
 ↳ symétrique ...

⇒ diagonalisable en base o.m.

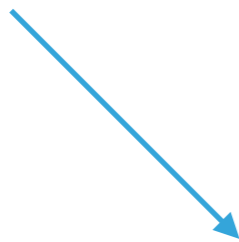
∃ base o.m.

orthogonale $V^{-1} = V^T$ — V la matrice de changement de base

$$B = V \cdot \begin{matrix} \text{diag} \\ | \\ \text{diag} \end{matrix} \cdot \begin{matrix} V^{-1} \\ | \\ V^T \end{matrix}$$

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 orthornormale de \mathbb{R}^m) telle que :

$$B = A^t A = V D V^t \quad \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{pmatrix}$$

↓
caric
symétrique
⊕

$$d_i \geq 0$$

positive $\xrightarrow{\text{donc}}$ ses val. propres $d_i \geq 0$

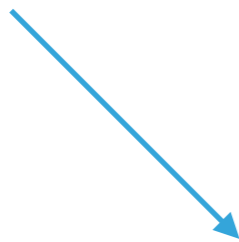
def ←

$$\forall X \neq 0 \quad {}^t X \cdot \text{matrice} \cdot X > 0$$

$$\begin{aligned} {}^t X (A^t \cdot A) \cdot X &= (AX)^t \cdot (AX) \\ &= \langle AX, AX \rangle \\ &= \|AX\|^2 \geq 0 \end{aligned}$$

MATRICES RECTANGULAIRES

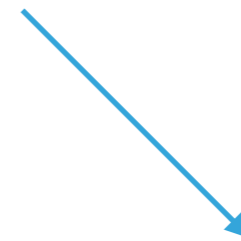
$$A \in \mathcal{M}_{m \times m}(\mathbb{R}) \quad \longleftrightarrow \quad \mathcal{f} : \mathbb{R}^m \longrightarrow \mathbb{R}^m$$



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


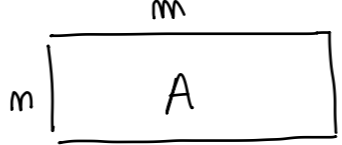
$\exists V$ orthogonale (ie. base orthornormale de \mathbb{R}^m) telle que :

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



.../...

DÉCOMPOSITION EN VALEURS SINGULIÈRES

$A \in \mathcal{M}_{n,m}(\mathbb{R})$  ou  $\longleftrightarrow f : \mathbb{R}^m \rightarrow \mathbb{R}^n$

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 :

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

$\Sigma = \begin{array}{c} \sigma_1 \\ \vdots \\ \sigma_m \\ \hline 0 \end{array}$ si $n \geq m$

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

$\Sigma = \begin{array}{c} \sigma_1 \\ \vdots \\ \sigma_n \\ \hline 0 \end{array}$ si $n \leq m$

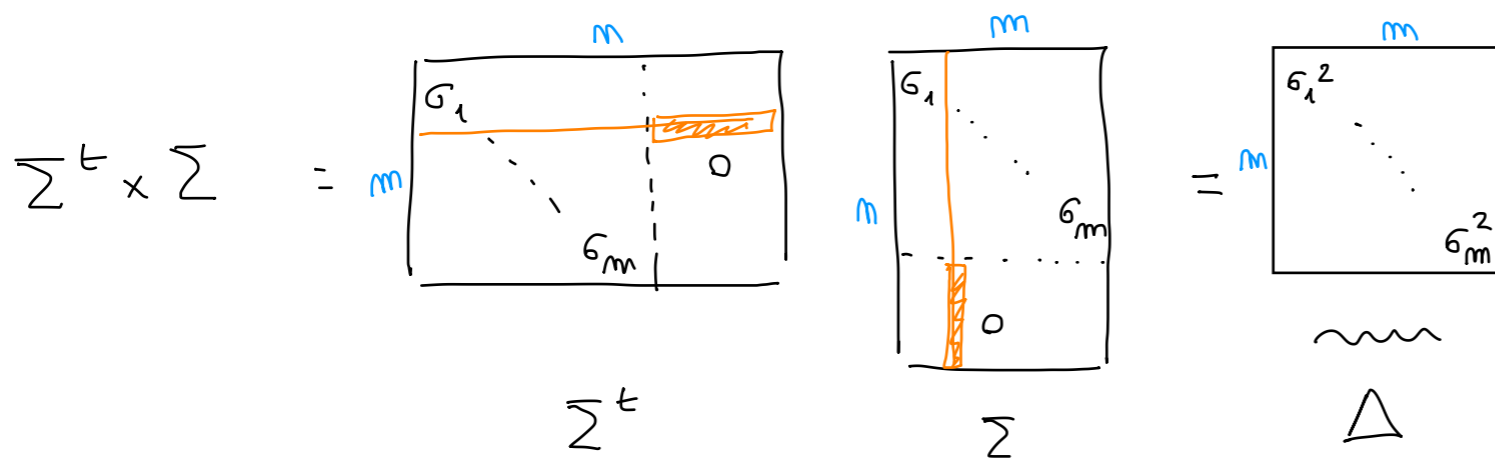
Décomp. valeurs singulières

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

A is $m \times m$.
 U is $m \times m$.
 Σ is $m \times m$ with singular values $\sigma_1, \dots, \sigma_m$ on the diagonal and 0 elsewhere.
 V is $m \times m$.

Diagonalisation ...

$$\begin{aligned}
 B &= A^t \times A \\
 &= (U \Sigma V^t)^t \times (U \Sigma V^t) \\
 &= V \Sigma^t U^t U \Sigma V^t \\
 &= V \Sigma^t \cdot \Sigma \cdot V^t
 \end{aligned}$$



σ_m obtenu

$$B = V \cdot \Delta \cdot V^t$$

Quand on a diagonalisé B , on avait :

$$B = V \cdot D \cdot V^t$$

D is a diagonal matrix with eigenvalues d_i .
 V is the matrix of eigenvectors of B .

Donc

$$d_i = \sigma_i^2$$

d_i is the eigenvalue of B .
 σ_i is the singular value of A .

et les vect. propres de B

v_i vect. sing à droite de A

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 :

$$Av_i = \sigma_i u_i \quad A^t u_i = \sigma_i v_i$$

"également"

$$\|v_i\| = 1 \quad \|u_i\| = 1$$

Immunément
~ vect. propres
($Ax = \lambda x$)
 $\sigma \quad \delta$

Preuve :

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

$$\text{de } B = A^t A$$

Soient :

- ▶ $\lambda_1, \dots, \lambda_n$ les valeurs propres de $A^t A$
- ▶ $\sigma_1, \dots, \sigma_n$ les valeurs singulières de A

$$\lambda_i = \sigma_i^2$$

Preuve :

cf asant

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 :

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

||
 σ_1

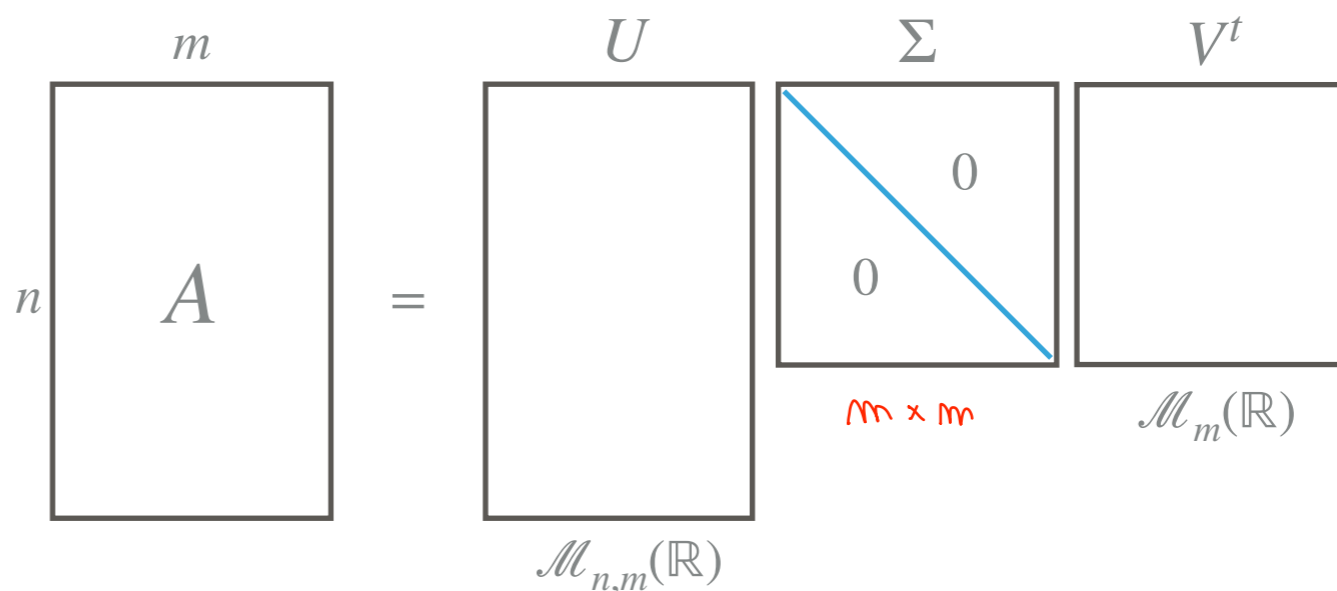
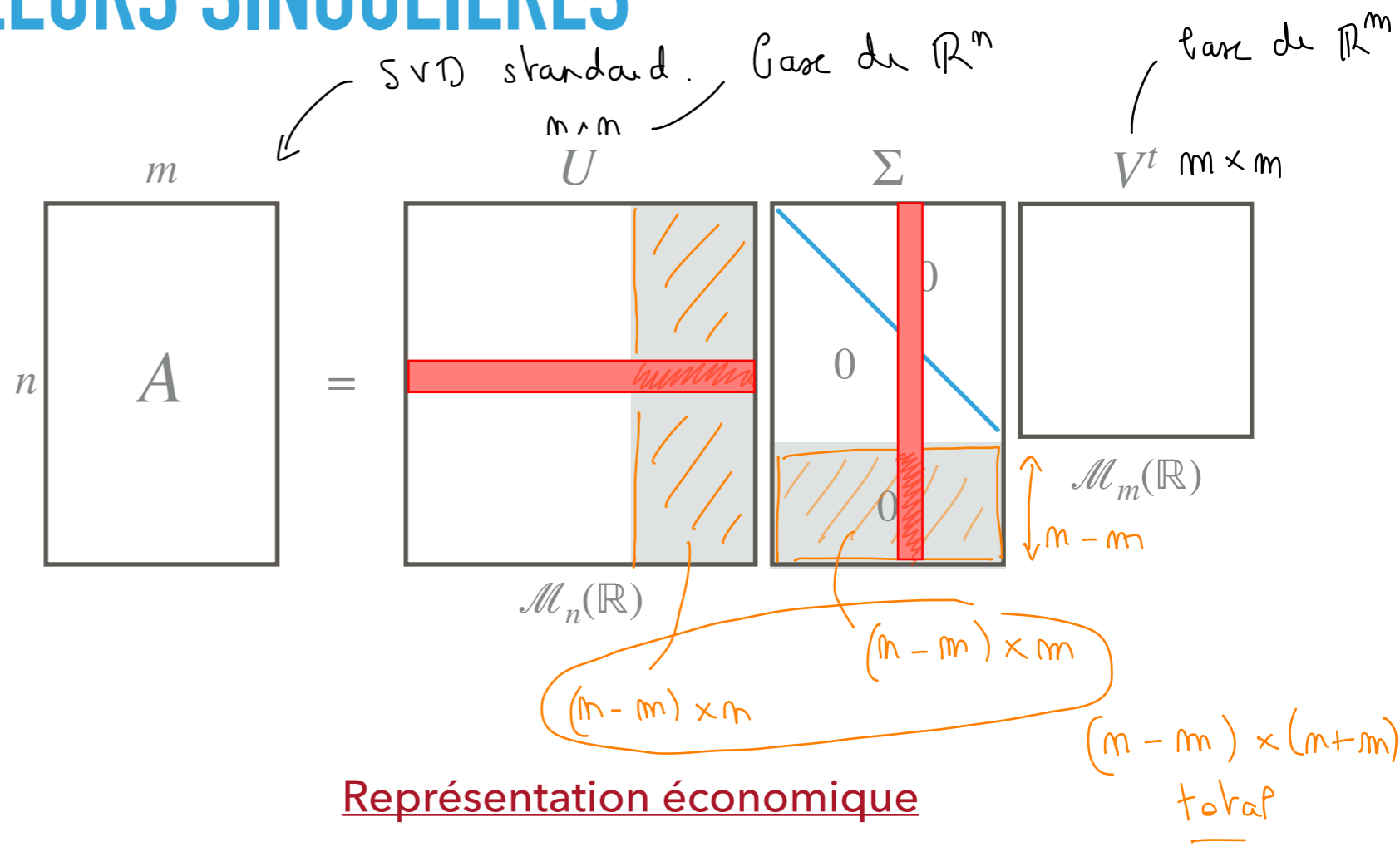
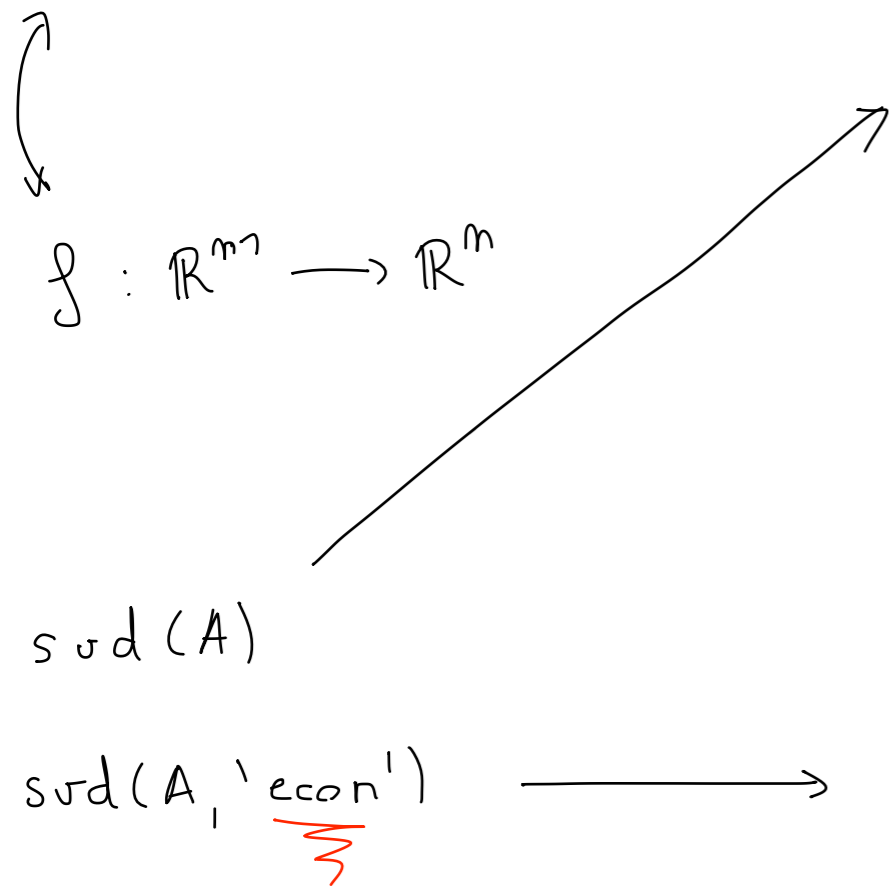
$$A = U \cdot \Sigma \cdot V^t$$

$$\|A\|_2 = \underbrace{\|U\|_2}_{\text{orthogonales}} \|\Sigma\|_2 \underbrace{\|V^t\|_2}_{\text{orthogonales}} = \|\Sigma\|_2 = \sup \frac{\|\Sigma x\|}{\|x\|} \xrightarrow{\text{diag.}} \max\{\sigma_i\}$$

DÉCOMPOSITION EN VALEURS SINGULIÈRES

- $U \in \mathcal{M}_{n,n}(\mathbb{R})$ orthogonale
 - $V \in \mathcal{M}_{m,m}(\mathbb{R})$ orthogonale
- $$A = U \Sigma V^t \quad m \times m$$

On suppose $n \geq m$.



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

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

$$\begin{bmatrix} \sigma_1 v_{11}^t \\ \vdots \\ \sigma_m v_{m1}^t \end{bmatrix}$$

Prod. par blocs

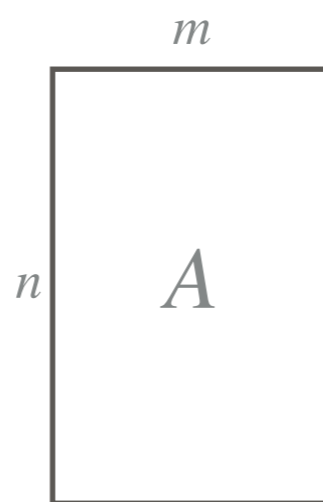
1 bloc

m blocs



$$= \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \sum_{k=1}^m U_{1k} \cdot (\sigma_k V_{k1})$$

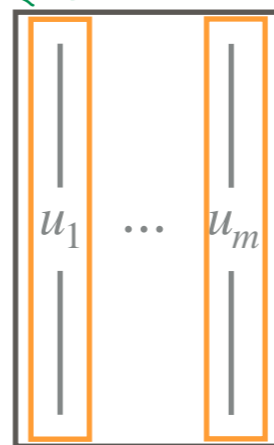
décomp.
économique
 \equiv



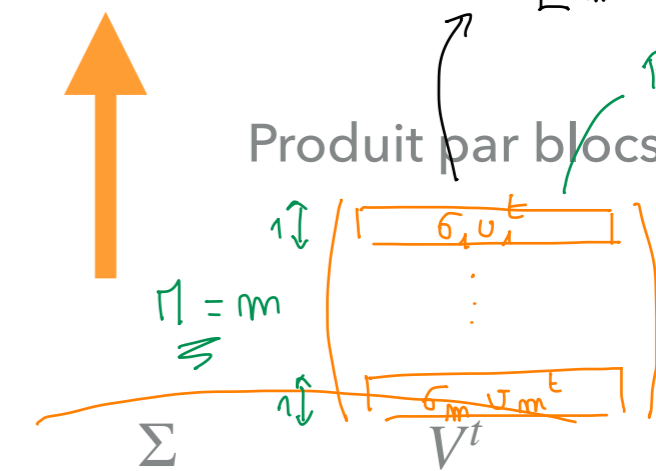
$$[U_{11} \dots U_{1m}]$$

$M = m$
 U

$N = 1$

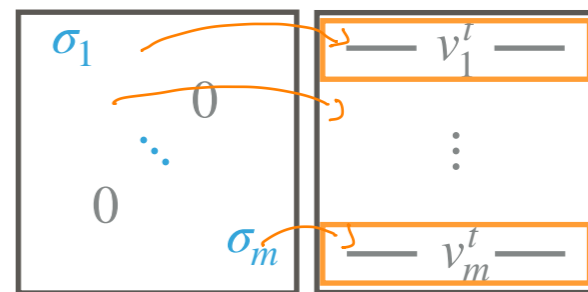


$\mathcal{M}_{n,m}(\mathbb{R})$



Produit par blocs

$P = 1$
 m



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

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

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

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

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 :

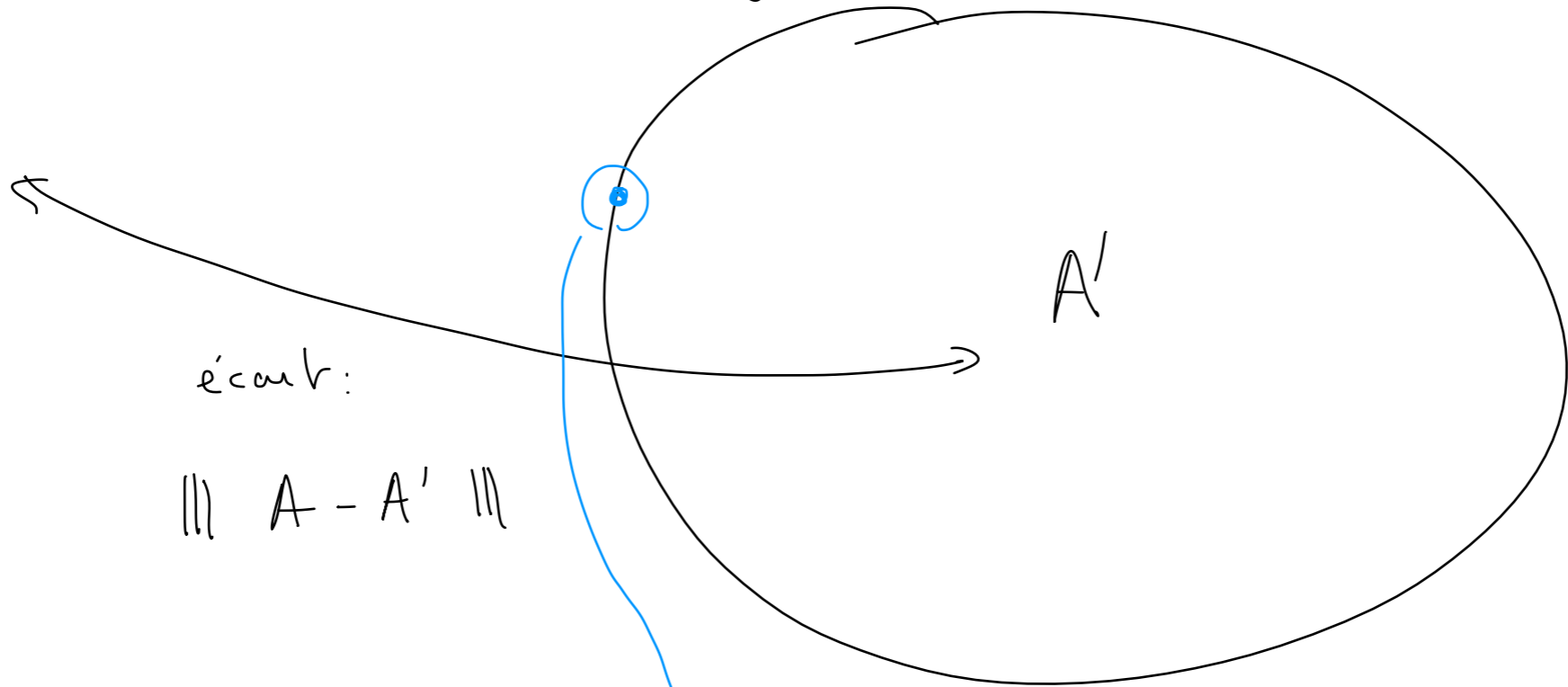
$$A = \sum_{i=1}^r \sigma_i u_i v_i^t$$

— somme seulement jusqu'à $r \leq m$

matrice de rang r

matrices de
rang $r < m$

A
↘



écart:

$$\|A - A'\|$$

la plus
proche
de A

$$\sum_{k=1}^r \sigma_k u_k v_k^T$$

)
meilleure approx de rang k 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)

Théorème d'Eckart-Young

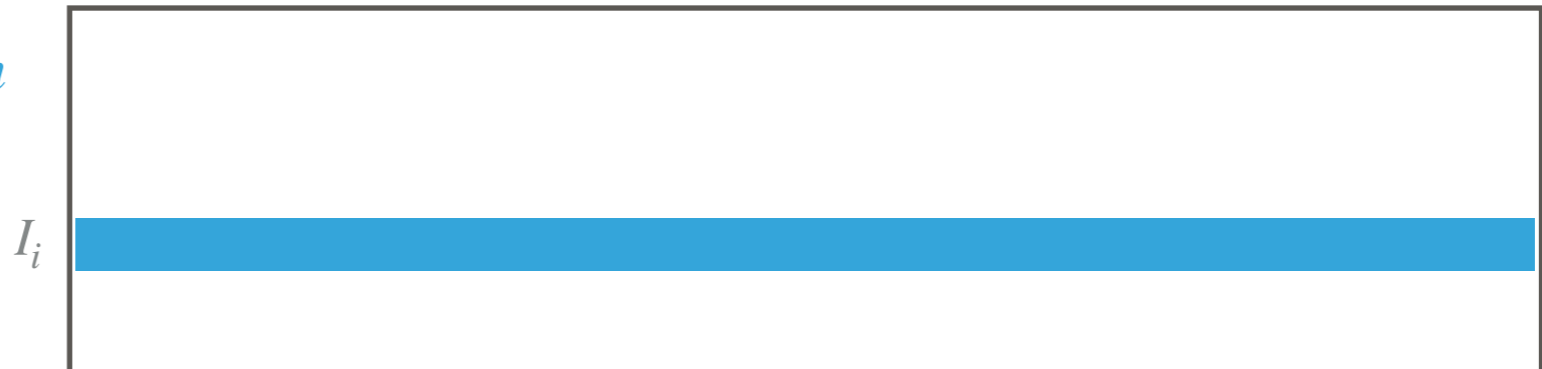
La matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ de rang r la plus proche de A pour les normes $\|\cdot\|_2$ et $\|\cdot\|_{\text{Fro}}$ est :

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

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

Vidéo
↕
Matrice $k \times nm$

Vectorisation en un vecteur nm



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)

Théorème d'Eckart-Young

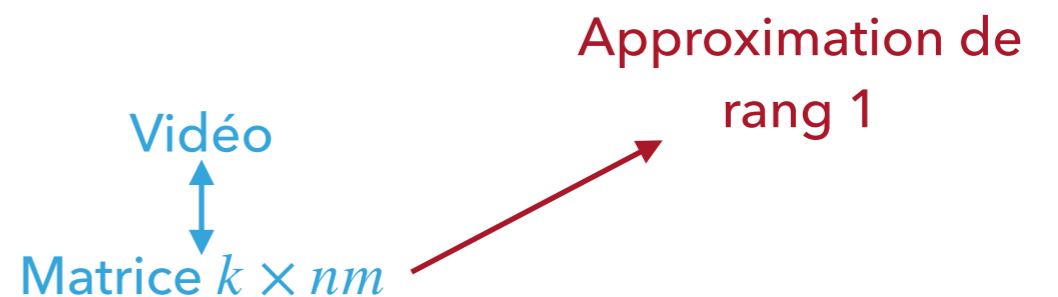
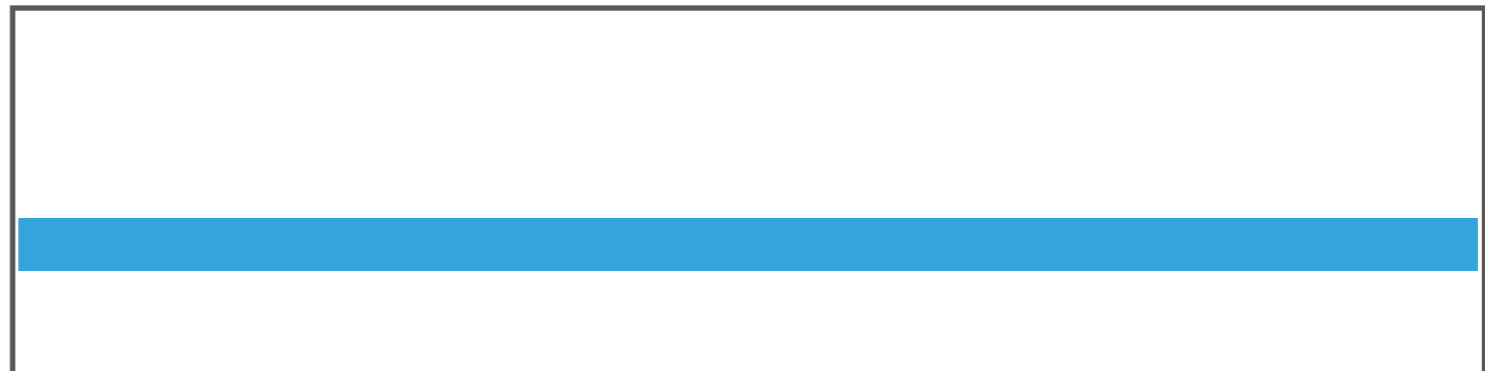
La matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ de rang r la plus proche de A pour les normes $\|\cdot\|_2$ et $\|\cdot\|_{\text{Fro}}$ est :

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

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

Vectorisation en un vecteur nm

I_i



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 **systems 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)

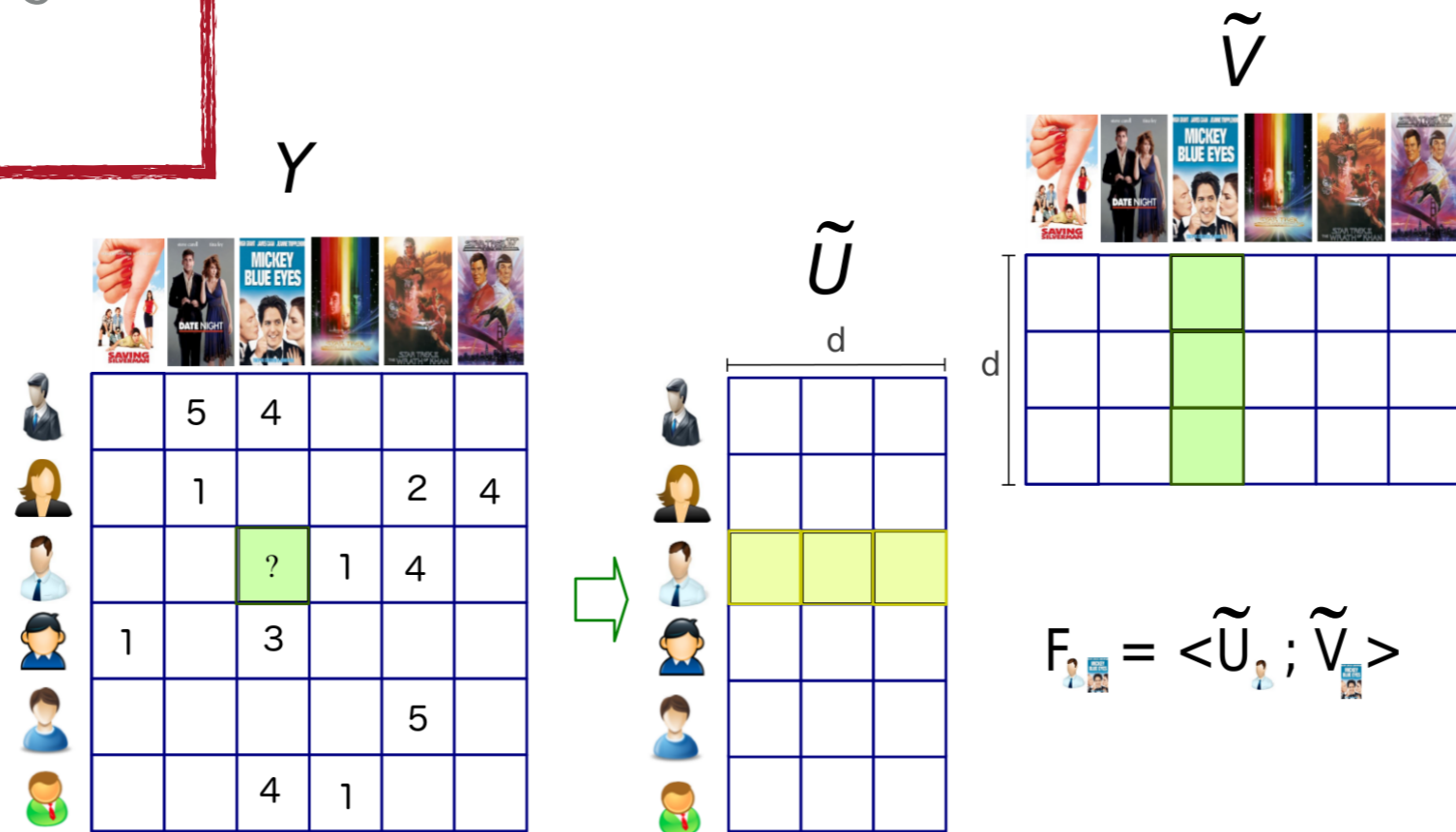
		5	4			
		1			2	4
			?	1	4	
	1		3			
					5	
			4	1		

Recommandation
↓
Complétion 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
- $$A = U\Sigma V^t$$

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



Approximation
de rang d

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

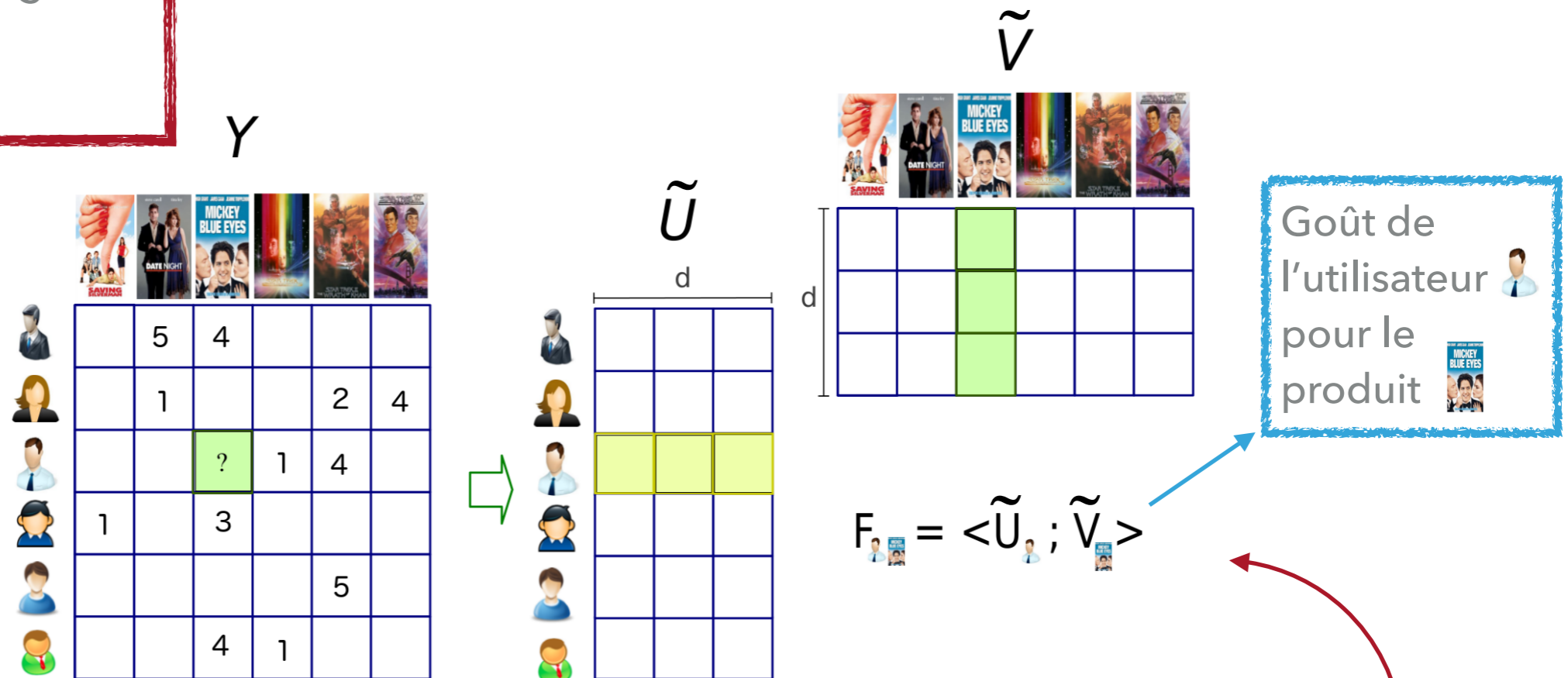
$$\tilde{U} = \begin{bmatrix} | & & | \\ u_1 & \dots & u_d \\ | & & | \end{bmatrix}$$

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

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)



Approximation
de rang d

$$Y = \sum_{i=1}^d \sigma_i u_i v_i^t \longrightarrow \tilde{U} = \begin{bmatrix} | & & | \\ u_1 & \dots & u_d \\ | & & | \end{bmatrix} \quad \tilde{V} = \begin{bmatrix} \text{---} & \sigma_1 v_1^t & \text{---} \\ & \vdots & \\ \text{---} & \sigma_d v_d^t & \text{---} \end{bmatrix}$$

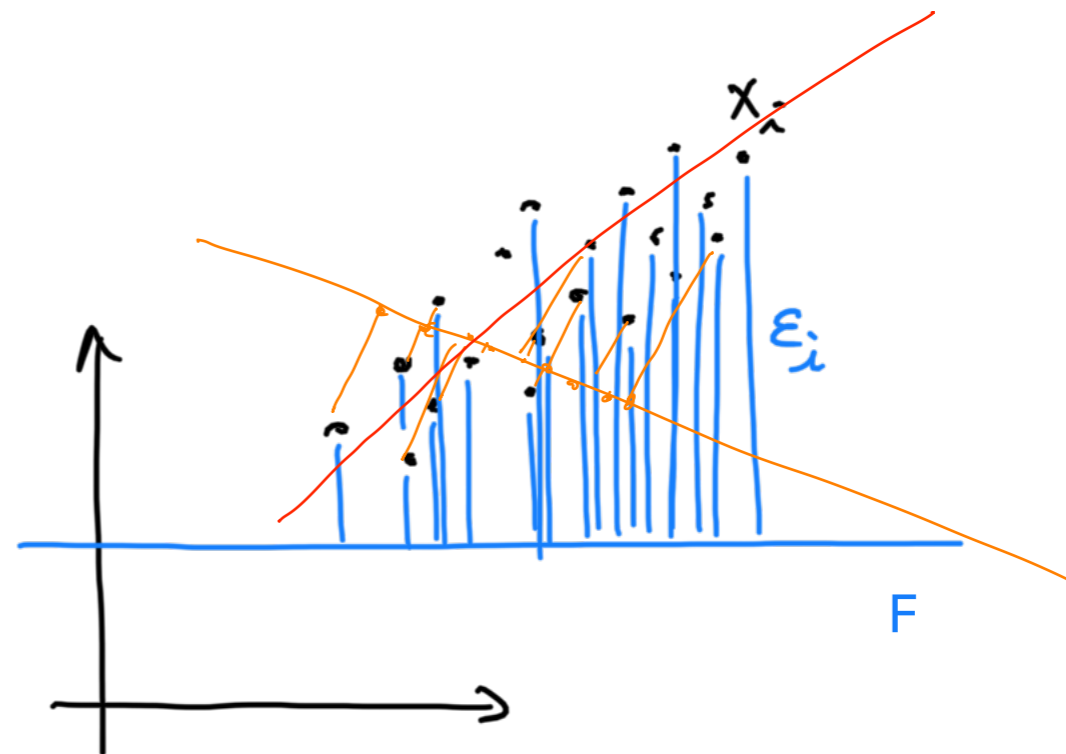
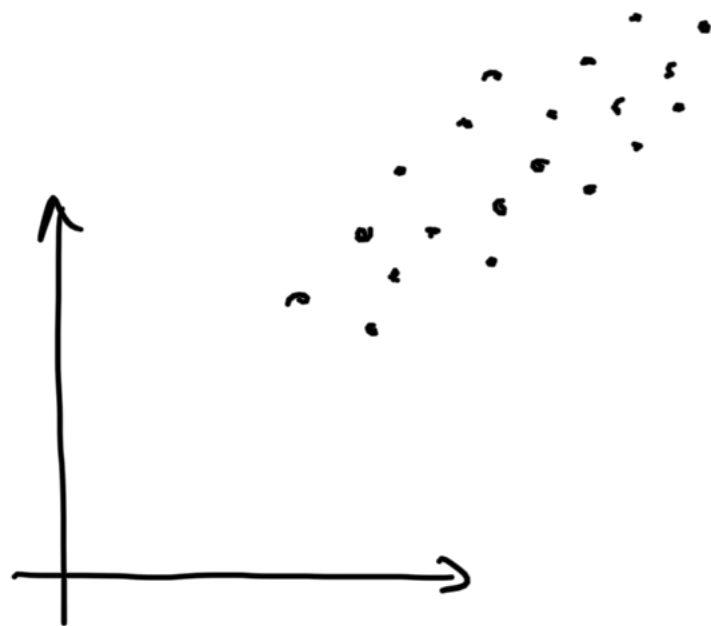
ANALYSE EN COMPOSANTES PRINCIPALES

PRÉSENTATION DU PROBLÈME

Application emblématique de la SVD

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

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



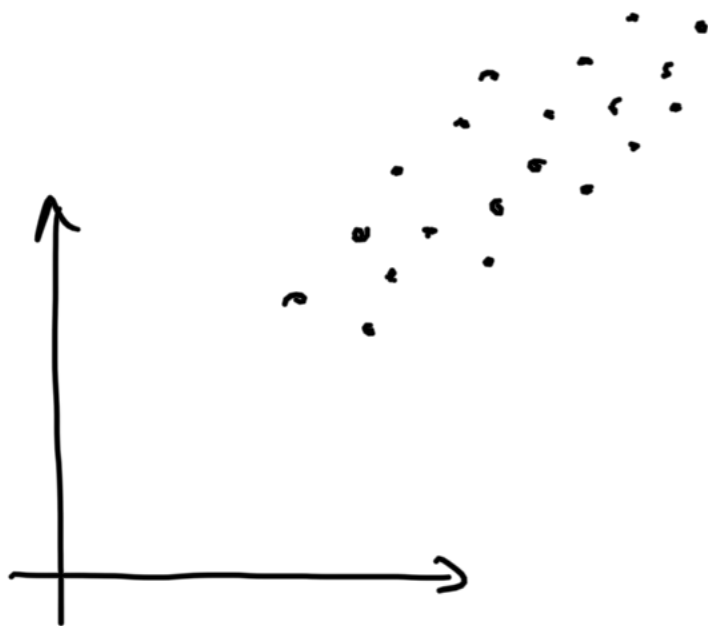
Exemple $n = 2, k = 1$

PRÉSENTATION DU PROBLÈME

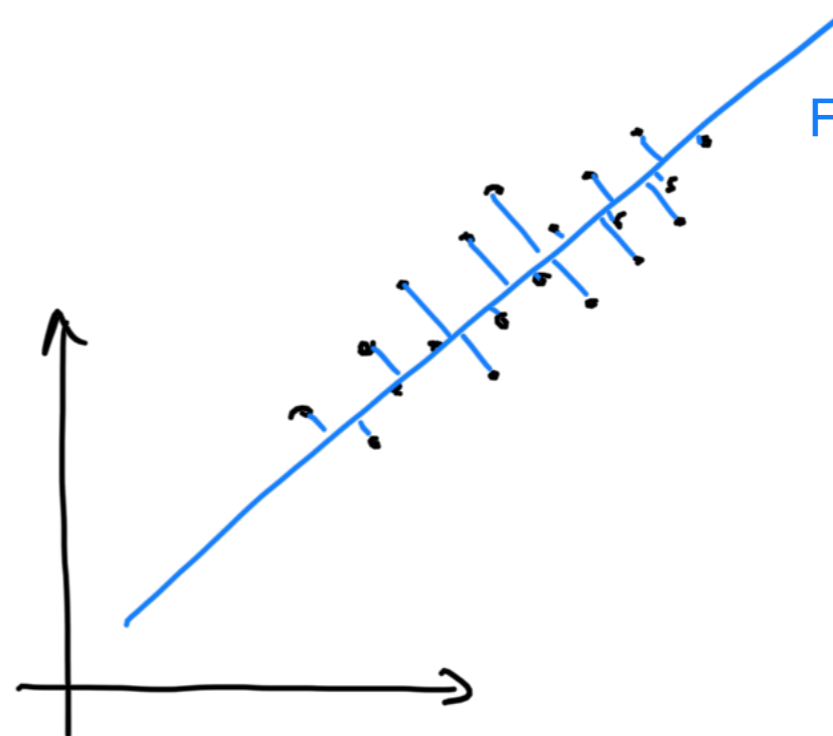
Application emblématique de la SVD

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

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



Exemple $n = 2, k = 1$

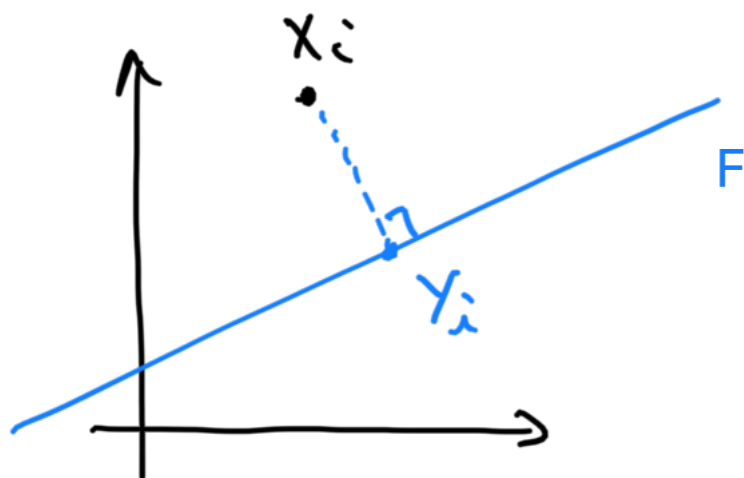


PRÉSENTATION DU PROBLÈME

Application emblématique de la SVD

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

- ▶ 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

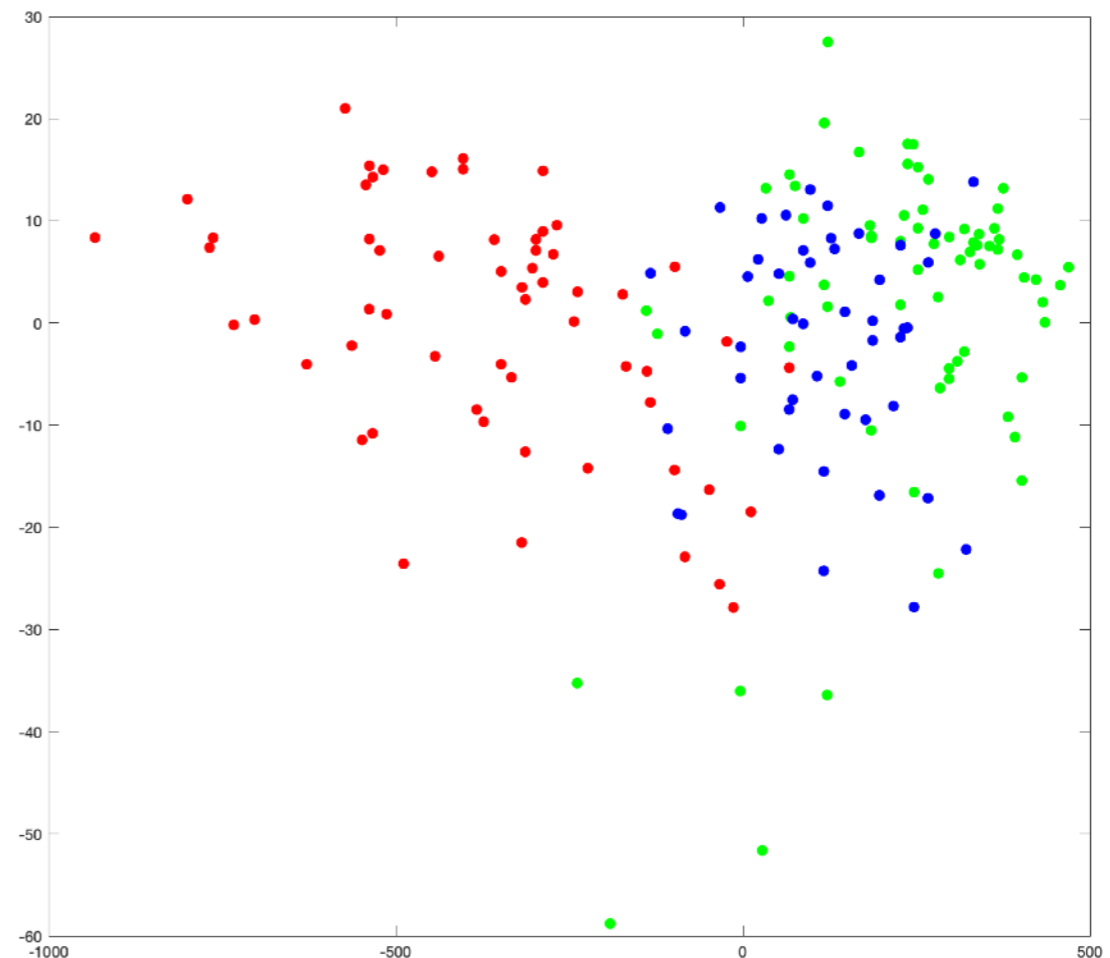
Application emblématique de la SVD

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

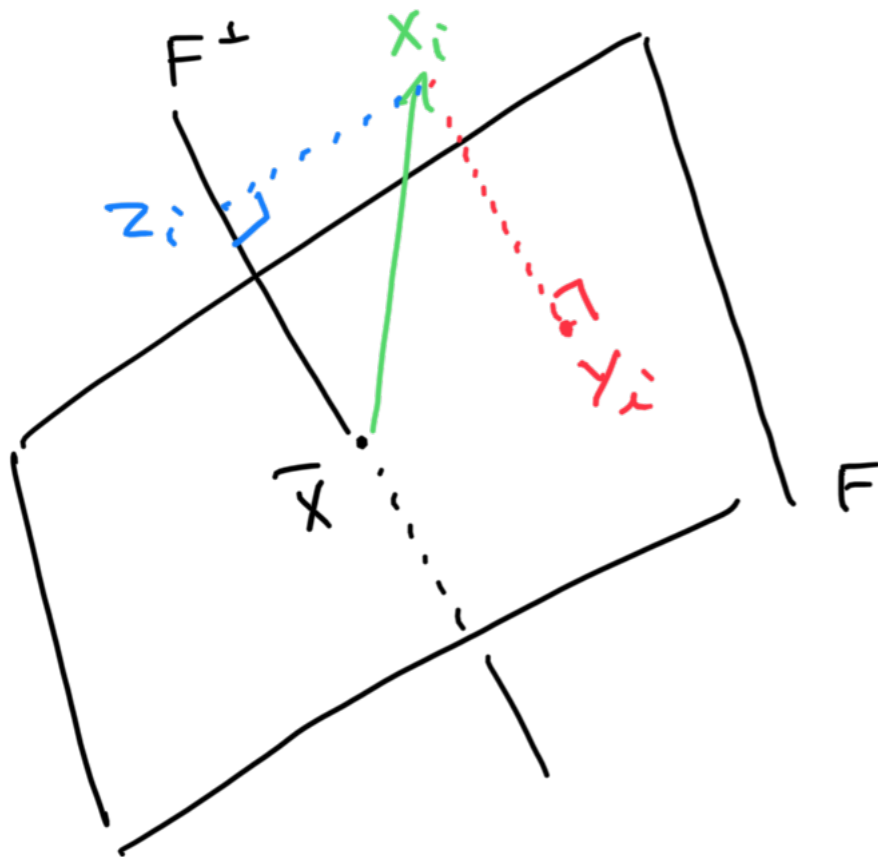
Représentation des
classes pour $k = 2$



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 :

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

Inertie par rapport au sous-espace F

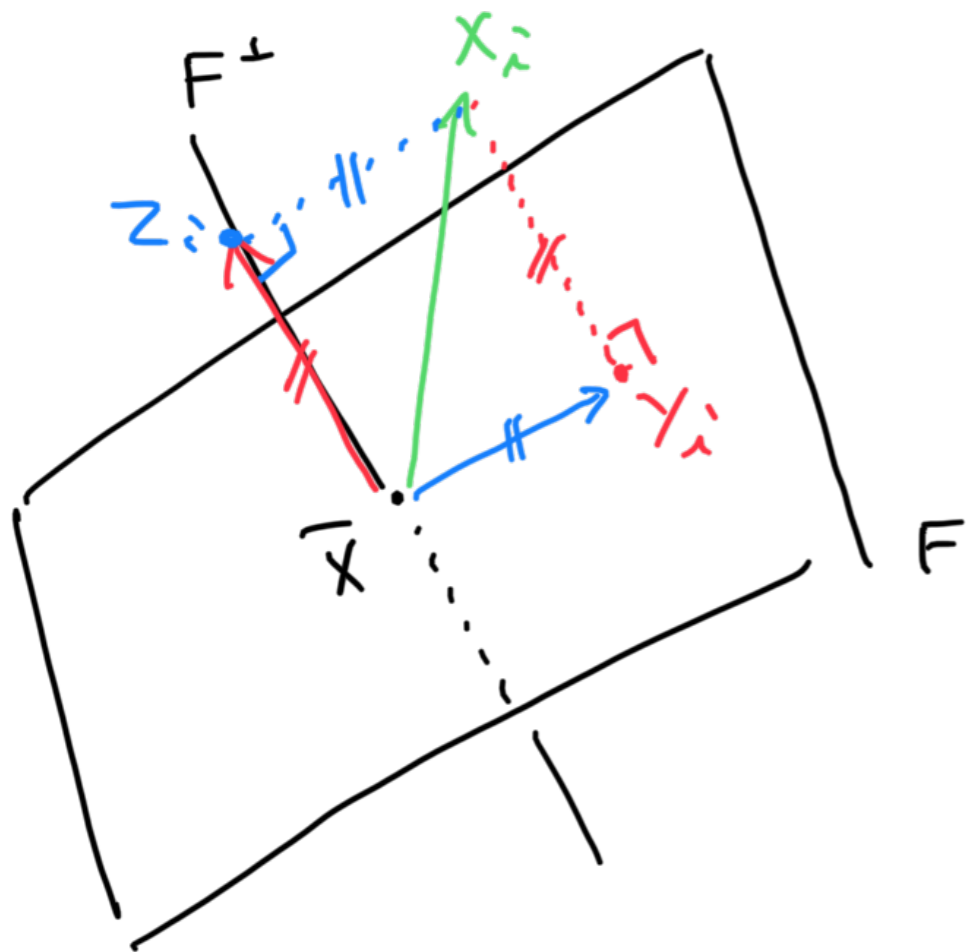
$$I_G = \frac{1}{N} \sum_{i=1}^N \|X_i - \bar{X}\|_2^2$$

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 :

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

Diagram illustrating the decomposition of the inertia tensor I_G into I_F and I_{F^\perp} . The equation is shown at the top. Two blue arrows point downwards from I_F and I_{F^\perp} to the words "Minimiser" and "Maximiser" respectively. A blue double-headed arrow connects "Minimiser" and "Maximiser".

ANALYSE EN COMPOSANTES PRINCIPALES

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

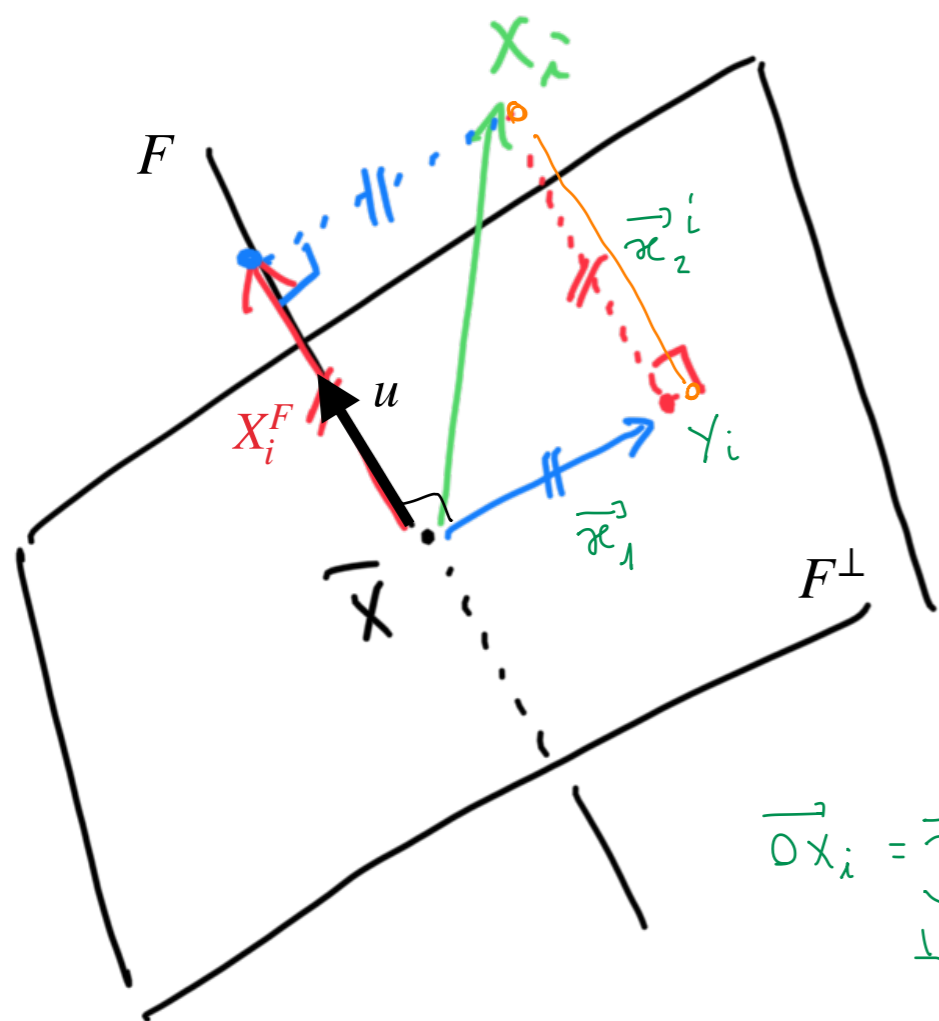
- ACP centrée : on considère les espaces affines passant par le barycentre $\bar{X} = \frac{1}{N} \sum_i X_i$

$$N \cdot I_{F^\perp} = \sum_i \|X_i - Y_i\|_{\mathbb{R}}^2$$

$$= \sum_i \left\| \underbrace{(Ox_i \cdot \vec{u})}_{\mathbb{R}} \vec{u} \right\|^2 \quad \text{avec } \|\vec{u}\| = 1$$

$$= \sum_i (Ox_i \cdot \vec{u})^2$$

origine : \bar{X}
barycentre :



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

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

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

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

$$\vec{Ox}_i \cdot \vec{u} = \underbrace{\vec{x}_2 \cdot \vec{u}}_{\| \vec{x}_2 \| \cdot \| \vec{u} \|} = \| \vec{x}_2 \| \cdot 1$$

$$\| \vec{x}_2 \| = \frac{\vec{Ox}_i \cdot \vec{u}}{\| \vec{u} \|} \Rightarrow \vec{x}_2 = \frac{\vec{Ox}_i \cdot \vec{u}}{\| \vec{u} \|^2} \vec{u} = (\vec{Ox}_i \cdot \vec{u}) \vec{u}$$

$$N \cdot I_{F \perp} = \sum_i \left(\overbrace{(x_i - \bar{x}) \cdot \vec{u}}^{\in \mathbb{R}} \right)^2 \quad \vec{u} \text{ tq le plus max?}$$

$$\left(\underbrace{(x_i - \bar{x}) \cdot \vec{u}} \right) \times \left(\underline{\underline{(x_i - \bar{x}) \cdot \vec{u}}} \right) \quad \left| \begin{array}{l} \text{si } x \in \mathbb{R}: \\ x^t = x \end{array} \right.$$

$$\left(\vec{x}_1 \cdot \vec{x}_2 = \vec{x}_1^t \times \vec{x}_2 \right) \quad \left(\rightarrow \right) \quad \left(\cdot \right)$$

$$\left(\underbrace{(x_i - \bar{x})^t \times \vec{u}} \right) \times \left(\underline{\underline{(x_i - \bar{x})^t \times \vec{u}}} \right)$$

$$N \cdot I_{F \perp} = \sum_i \underbrace{\vec{u}^t \times (x_i - \bar{x})}_{\text{matrice}} \times \underbrace{(x_i - \bar{x})^t \times \vec{u}}_{\text{matrice}}$$

$$= \vec{u}^t \times \left(\sum_i (x_i - \bar{x})(x_i - \bar{x})^t \right) \times \vec{u}$$

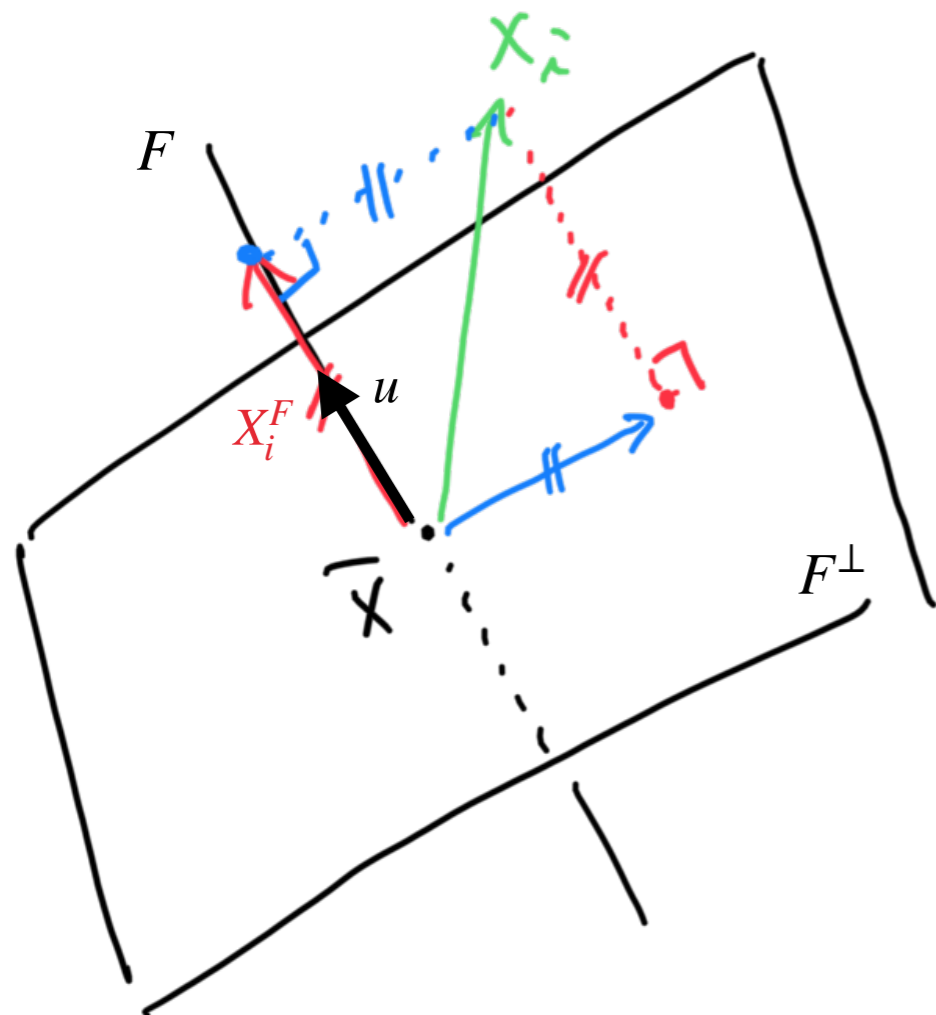
somme de matrices

\Rightarrow matrice de covariance (Z)

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 :

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

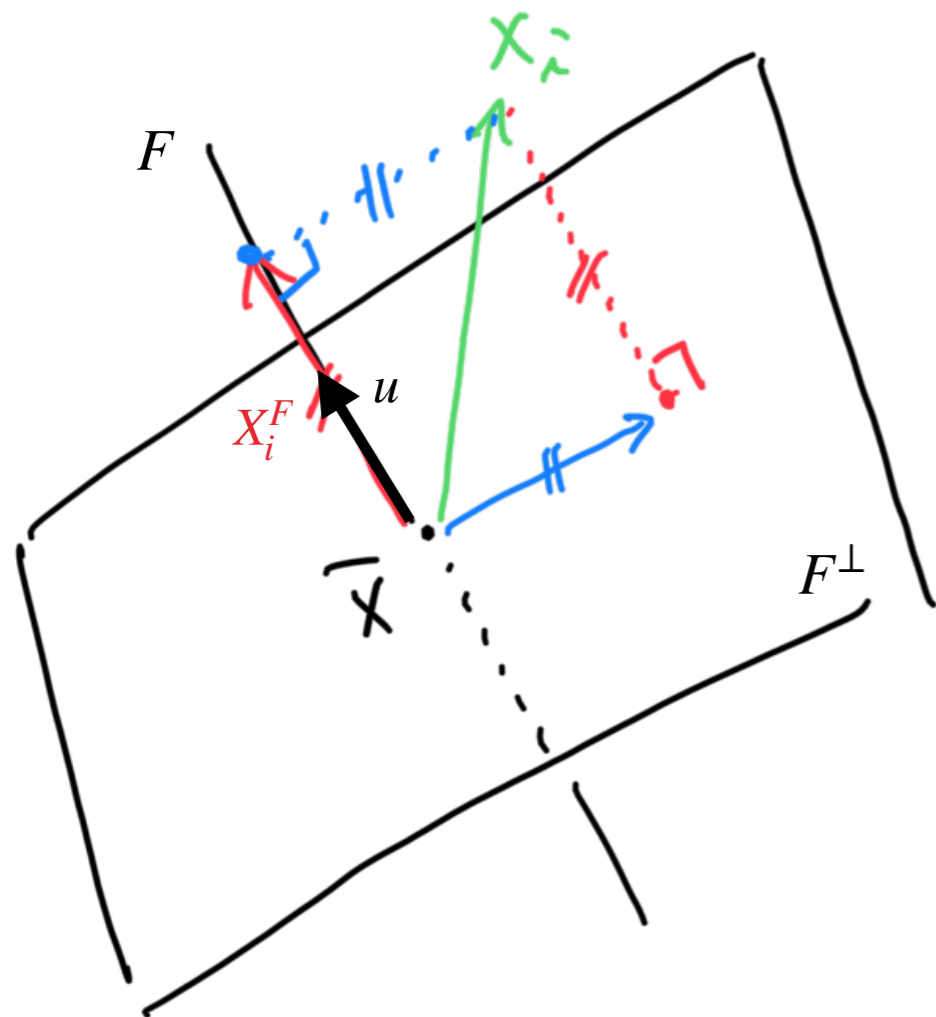
Donc :

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

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 $\bar{X} = \frac{1}{N} \sum_i X_i$



Composantes princip. du nuage
 → vect. propres de Z (\perp)
 diag. en base orthonormale
 ↗ diag. de Z (symétrique)

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

Matrice de covariance

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

sym

val. propre λ_i

vect. propre u_i

$$Z = A^t \cdot A$$

$$\lambda_i = \sigma_i^2$$

$$u_i = v_i$$

A =

$$\begin{pmatrix} \vdots \\ (x_i - \bar{x})^t \\ \vdots \end{pmatrix}$$

→ σ_i

→ u_i

→ v_i

SVD

SVD

A rect
 $m \times m$

U, Σ, V

↓
 σ_i

(u_i)

vect. sing.

$$A u_i = \sigma_i v_i$$

à droite

$$\sigma_i^2 = \lambda_i$$

val. vect. propre

$$B = A^t \cdot A$$

$m \times m$

↓

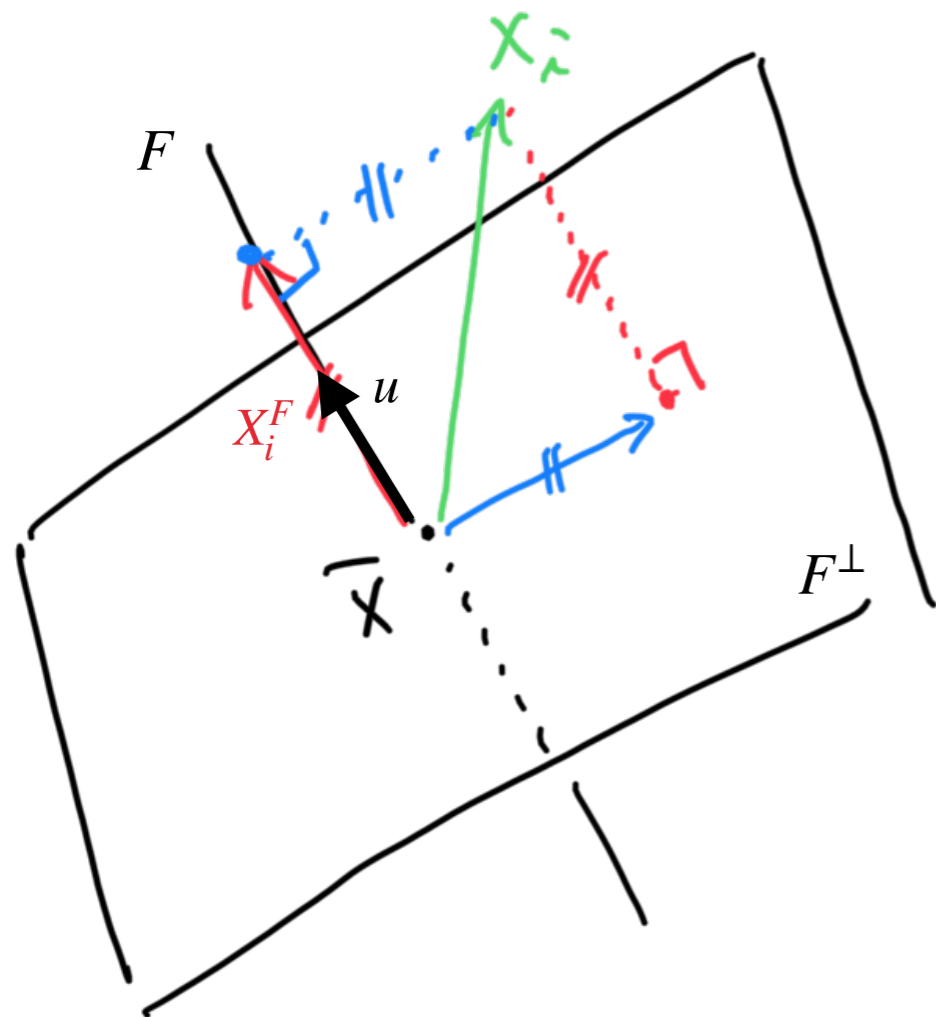
λ_i propre

(v_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 $\bar{X} = \frac{1}{N} \sum_i X_i$



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

$$=$$

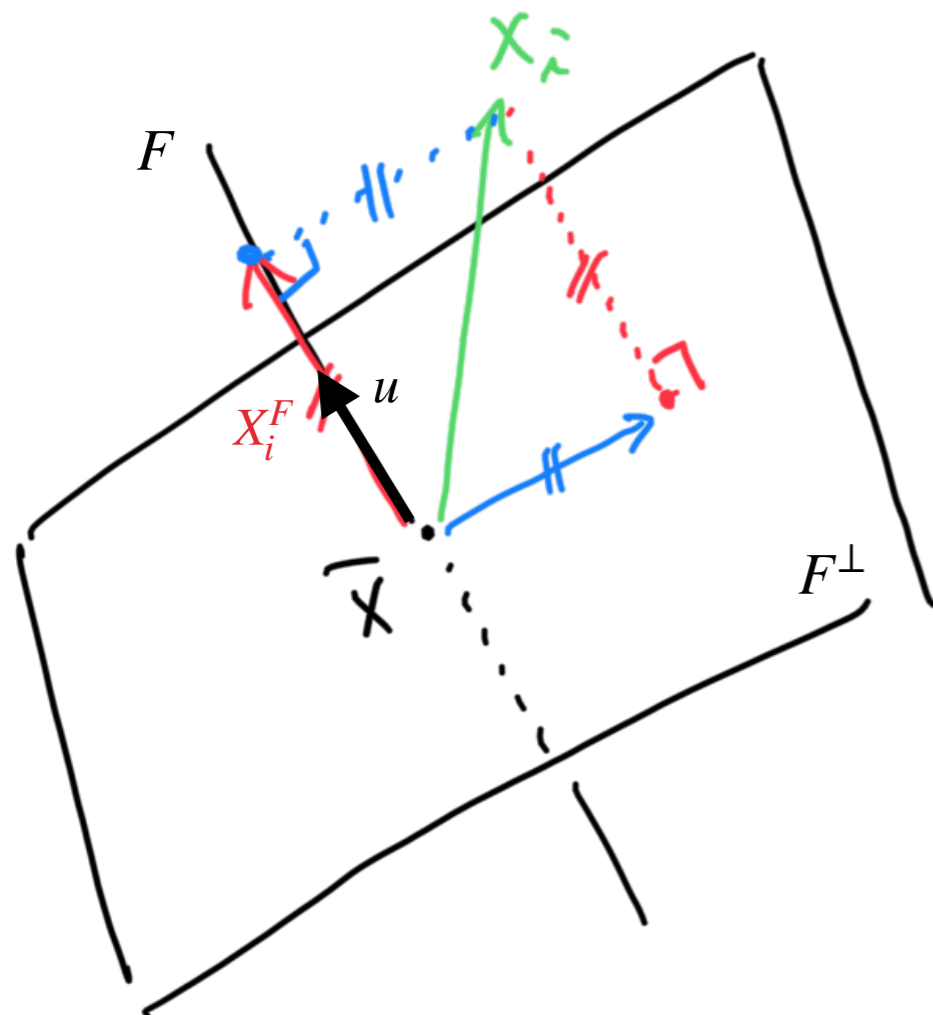
$$I_{F^\perp} = u^t \left(\underbrace{\frac{1}{N} \sum_i (X_i - \bar{X})(X_i - \bar{X})^\perp}_Z \right) u$$

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 $\bar{X} = \frac{1}{N} \sum_i X_i$



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

Matrice de covariance

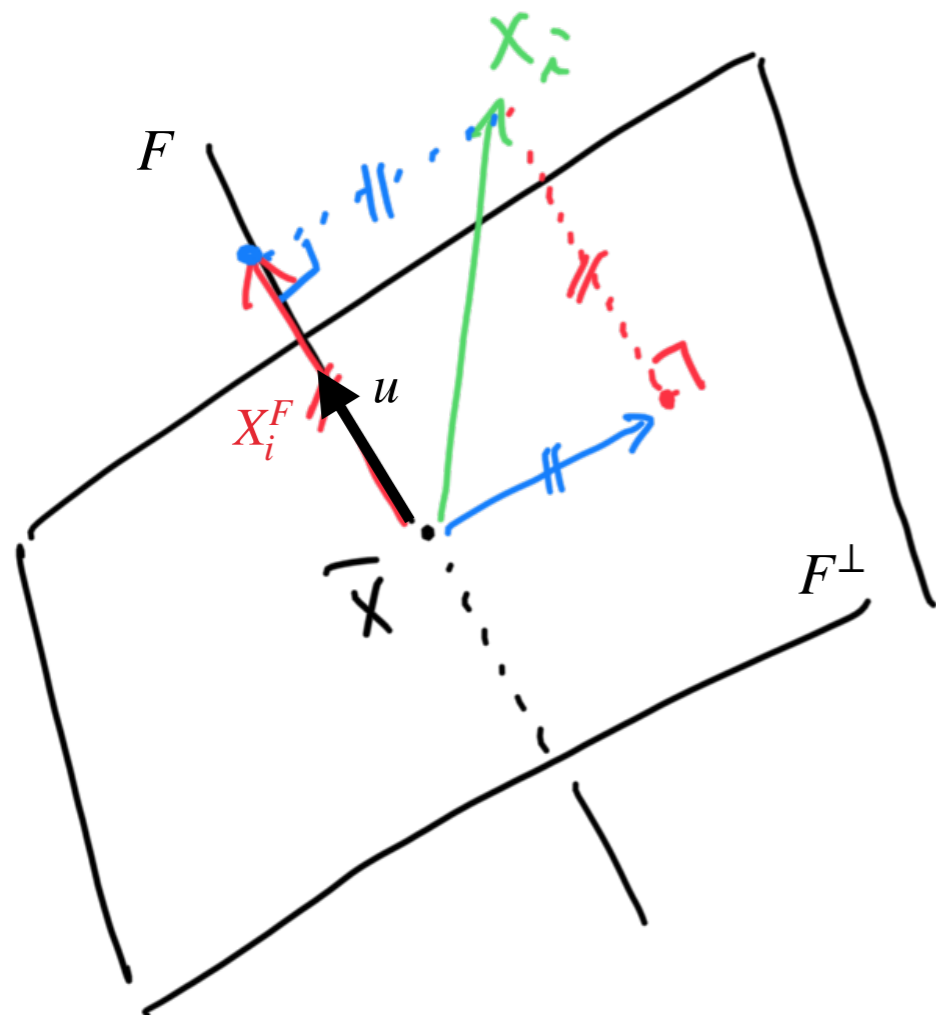
- ▶ Les vecteurs propres de Z sont les directions principales
- ▶ Les valeurs propres de Z sont les valeurs d'inertie selon les sous-espaces orthogonaux à ces directions

Version « 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 $\bar{X} = \frac{1}{N} \sum_i X_i$



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

Soit :

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

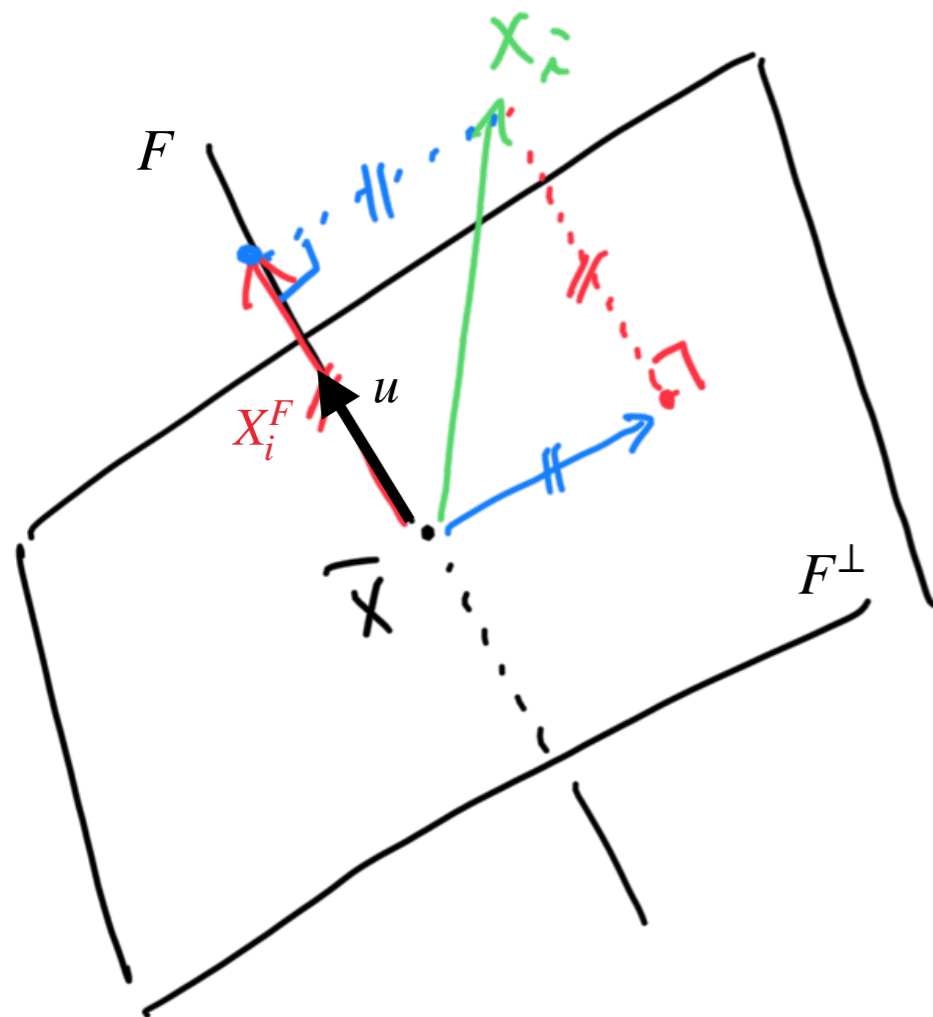
Produit par blocs

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

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 $\bar{X} = \frac{1}{N} \sum_i X_i$



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

- 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 - RESUME

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 $\bar{X} = \frac{1}{N} \sum_i X_i$

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

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

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

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

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

Base orthonormale $\{\vec{u}_1, \dots, \vec{u}_m\}$

$\vec{x} \rightarrow$ coord α_i dans cette base ?
=

$$\alpha_1 \cdot \vec{u}_1 + \dots + \alpha_m \cdot \vec{u}_m$$

$$\vec{x} = \alpha_1 \cdot \vec{u}_1 + \dots + \alpha_m \cdot \vec{u}_m$$

$$\langle \vec{x}, \vec{u}_i \rangle = \langle \alpha_1 \vec{u}_1 + \dots + \alpha_m \vec{u}_m, \vec{u}_i \rangle$$
$$= \alpha_i \underbrace{\|\vec{u}_i\|}_{1}^2$$

$$\langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

ANALYSE EN COMPOSANTES PRINCIPALES

Exemple sur un nuage de points 3D

```
clear
clc
clf
hold on

% Génération 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
```