

# Analyse des données structurées en blocs : généricité des données, polymorphisme des méthodes et questions émergentes

Mohamed Hanafi  
Unité de Sensométrie et  
Chimiométrie (USC)

# Sommaire

- **Partie 1 : introduction et motivation**
  - Préliminaires
  - Définition générique des données structurées en blocs
  - Objectifs méthodologiques
  - Etat de l'art et perspectives
- **Partie 2 : Analyse des tableaux de données d'ordre 3.**
  - Rangs.
  - Modèles.
  - Illustration sur un exemple.
- **Partie 3 : Analyse des Tableaux d'ordre 2 partitionnés.**
  - Démarche
  - Polymorphisme des méthodes à partir d'un exemple (HPCA,ACCPS)
  - Illustration sur un exemple

# Abondance des données

- Accroissement des capacités de stockage des ordinateurs (prix d'un mégaoctet : (\$11.54 en 1988) ; (\$0.003 en 2007)).
- Développement technologiques et diversification des outils d'acquisition des données.



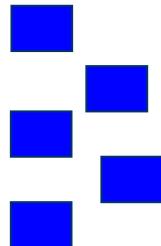
# Entrée et ordre d'un tableau

- Tableau de données.

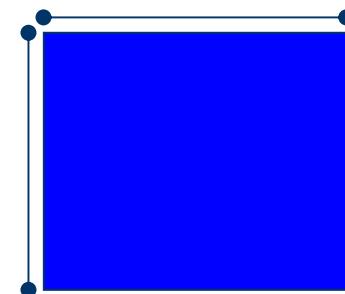
Une application d'un produit cartésien d'ensembles finis à valeurs réelles.

$$T : E_1 \times E_2 \times \dots \times E_K \rightarrow R$$

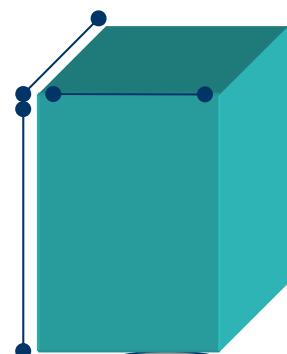
$$T(e_{i_1}^{(1)}, e_{i_2}^{(2)}, \dots, e_{i_K}^{(K)}) = T_{i_1, i_2, \dots, i_p}$$



ordre=1



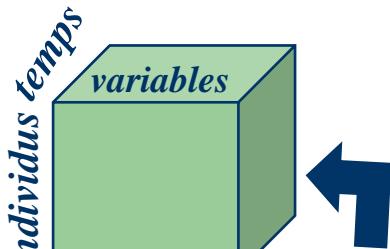
ordre=2



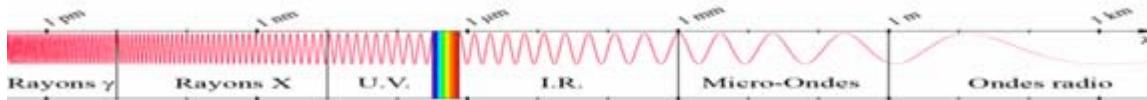
ordre=3

- $K$  : ordre du tableau
- $E_k$  :  $k^{\text{ème}}$  entrée du tableau

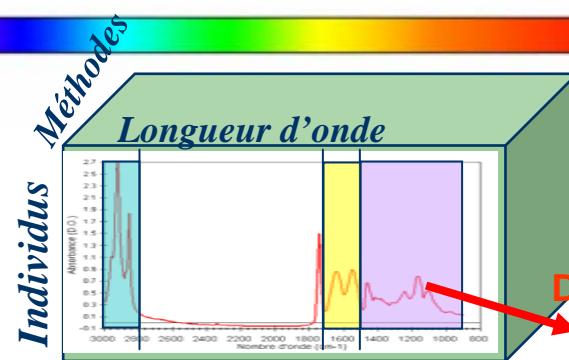
# Exemples



Données à composantes temporelles



Données sensorielles

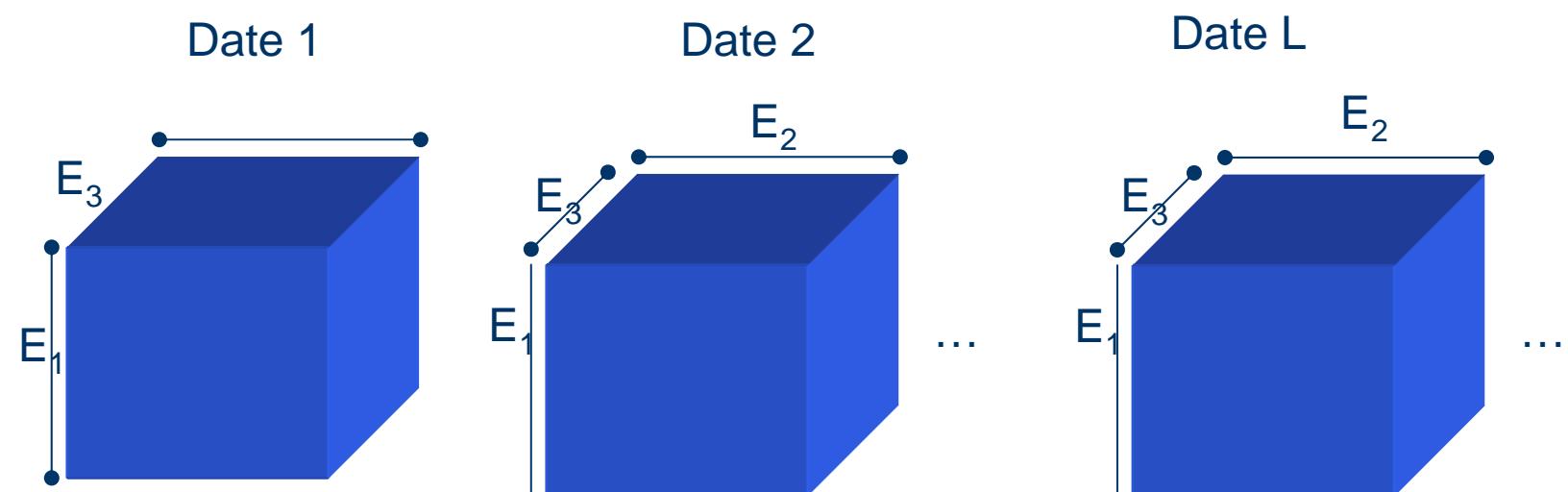


Données spectroscopiques  
spectre Infrarouge moyen



Image couleur

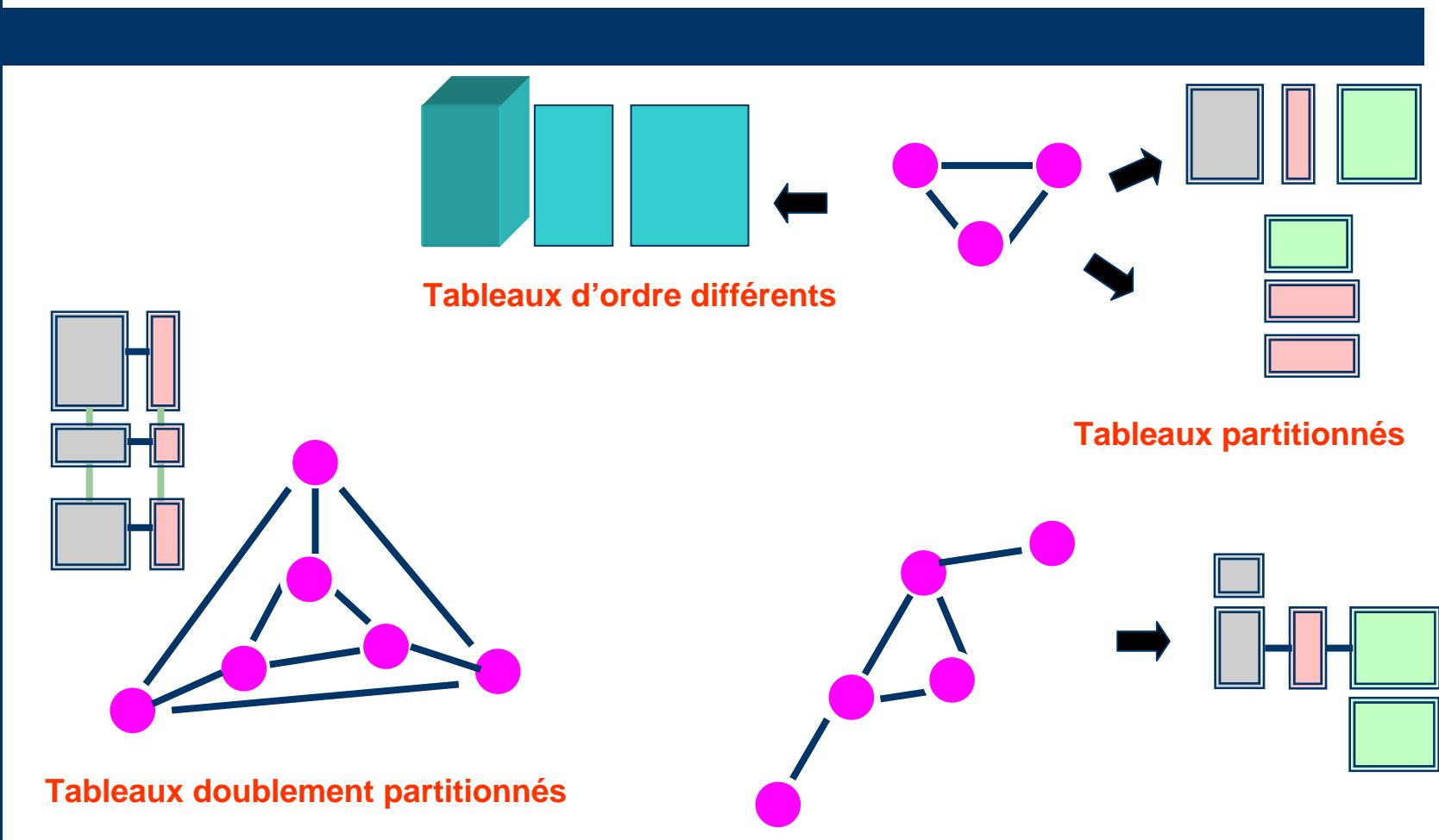
# Tableau d'ordre 4



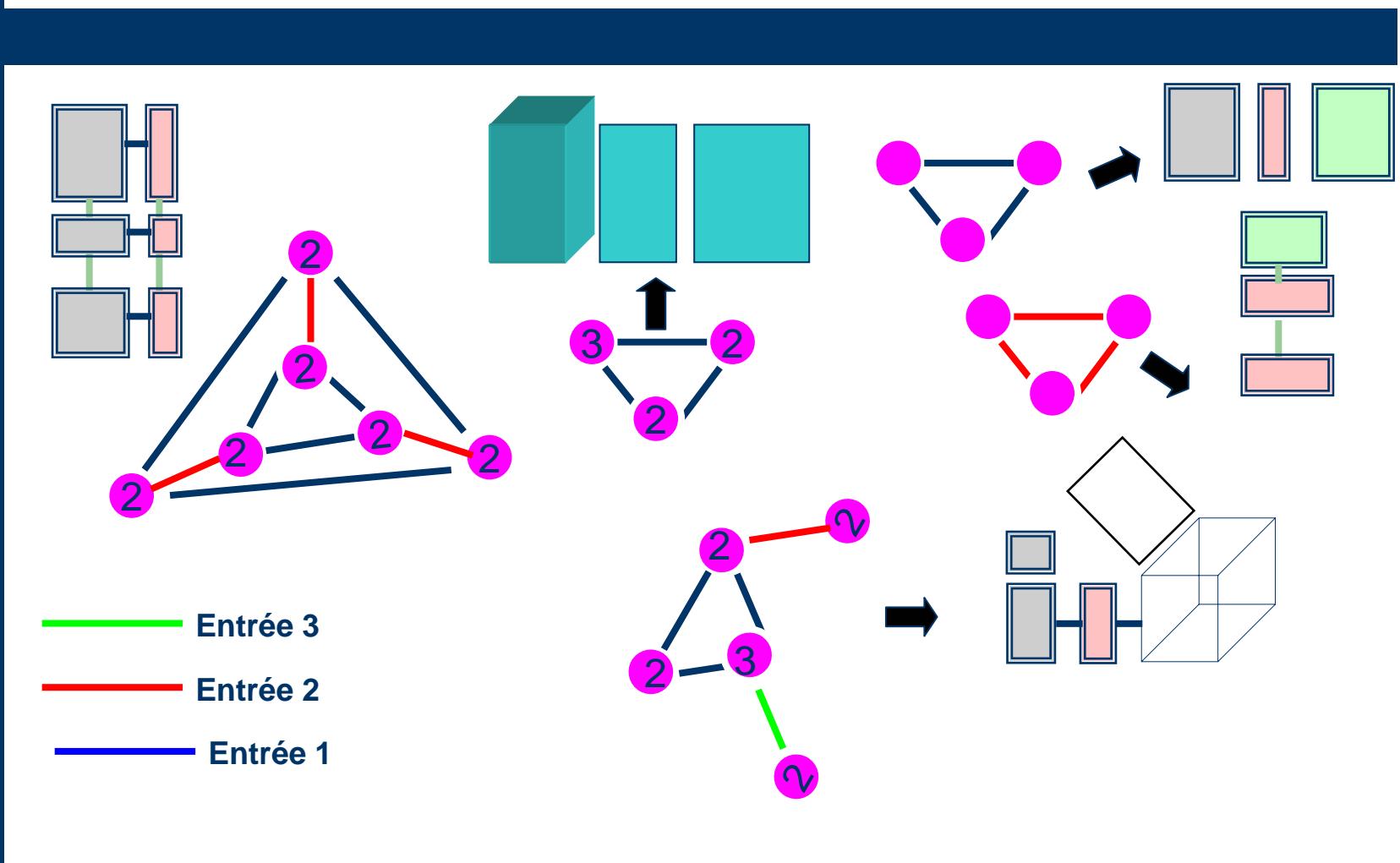
# Données structurées en blocs

- Graphe associé à un ensemble de tableaux
  - $\mathbf{S} = \{\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_L\}$  (pas nécessairement de même ordre)
  - $\mathbf{A} = \{(\mathbf{T}_l, \mathbf{T}_j) \mid \mathbf{T}_l \mathcal{R} \mathbf{T}_j\} \subseteq \mathbf{S} \times \mathbf{S}$ 
$$\forall (l \neq j) \quad \mathbf{T}_l \mathcal{R} \mathbf{T}_j \Leftrightarrow$$
les deux tableaux  $\mathbf{T}_j$  et  $\mathbf{T}_l$  ont une entrée commune
  - $(\mathbf{S}, \mathbf{A})$  est un graphe
- Un tableau structurée en blocs est un ensemble de tableaux dont le graphe associé est connexe

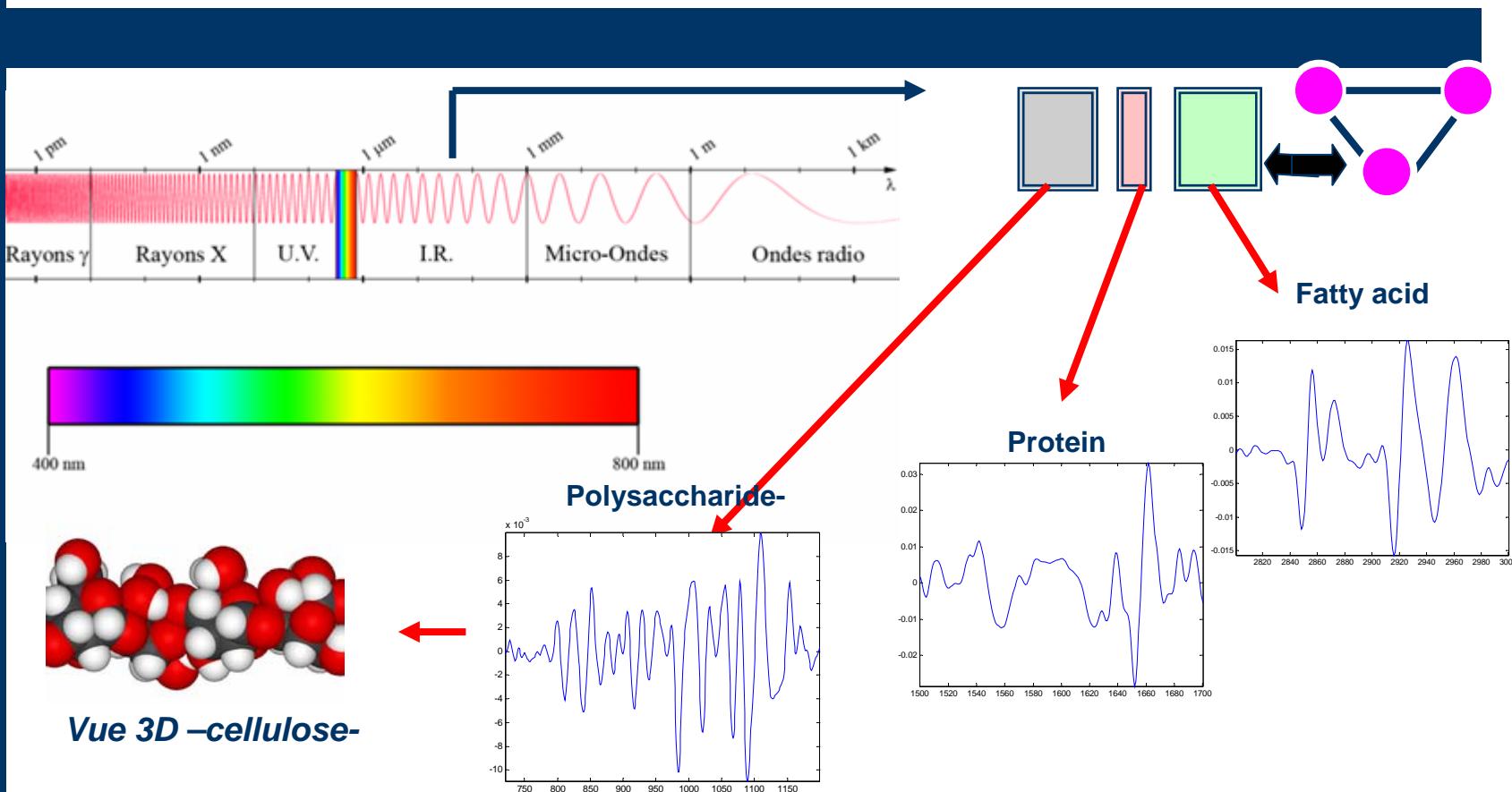
# Exemples



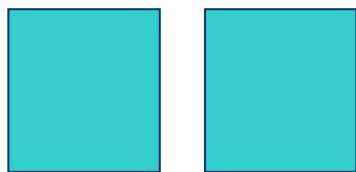
# structurées en blocs et multi graphe



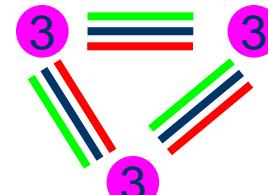
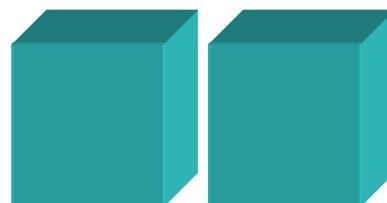
# Exemple de données partagées



# Remarque



3



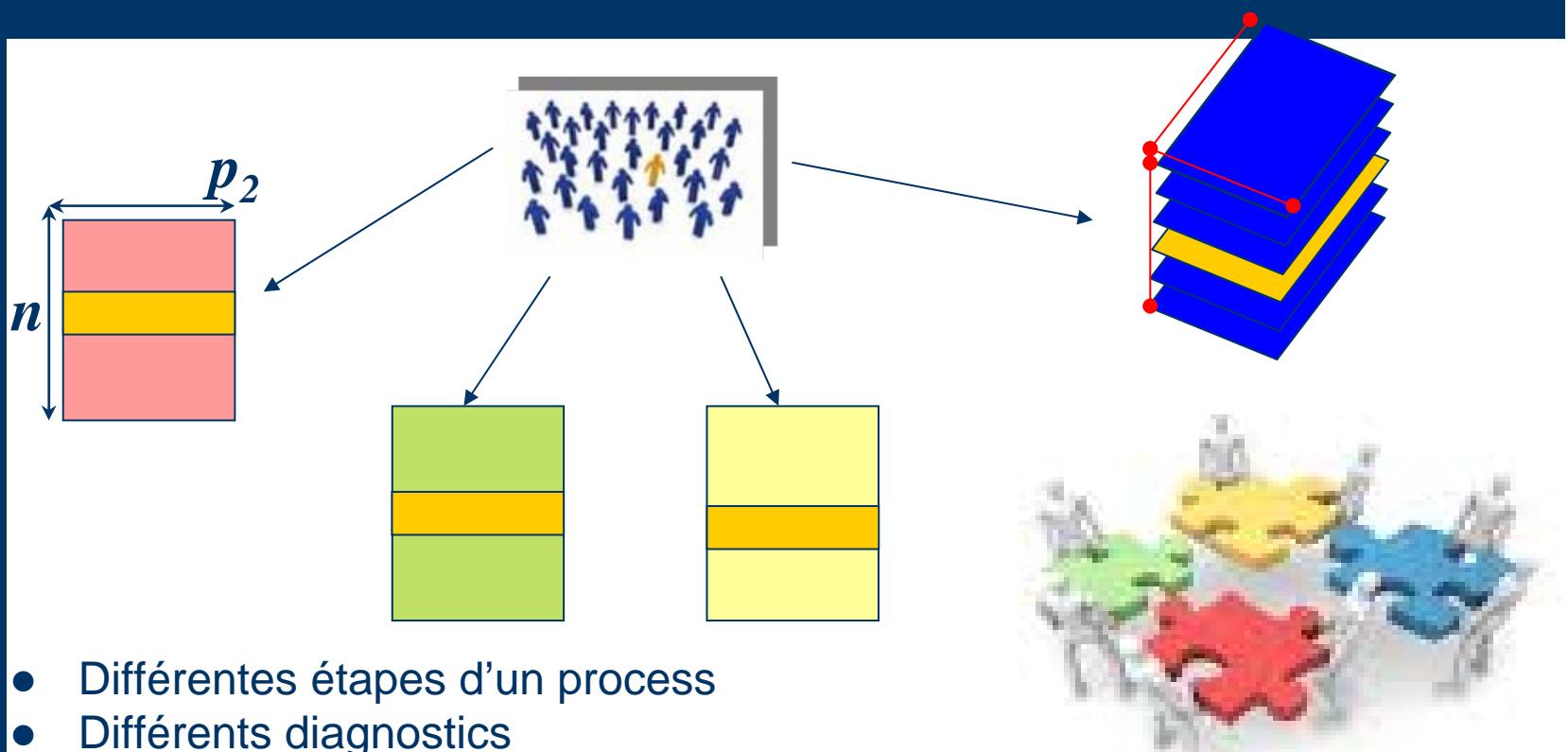
4

# Approche systémique ou intégrative

- Considère l'échantillon comme un système complexe constitué de différentes composantes.
- Les composantes interviennent et interagissent à différents niveaux des organisations structurales et fonctionnelles des échantillons étudiées.
- Prise en compte globale (intégration) de l'ensemble des données disponibles.



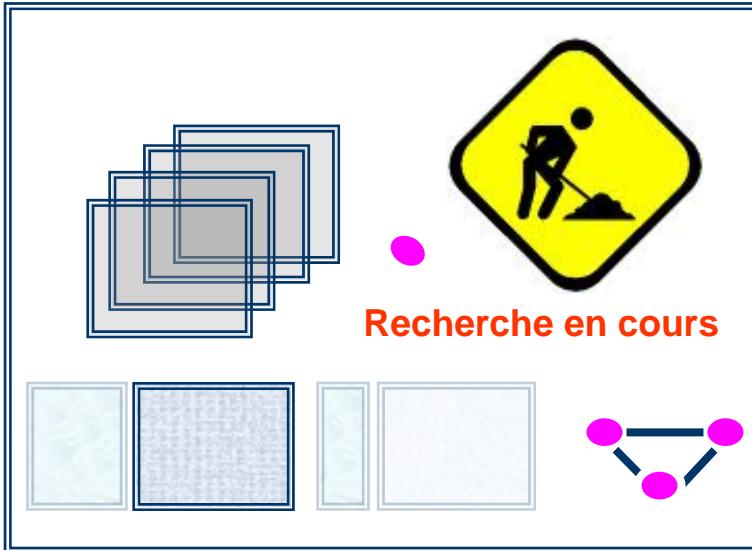
# Approche systémique ou intégrative



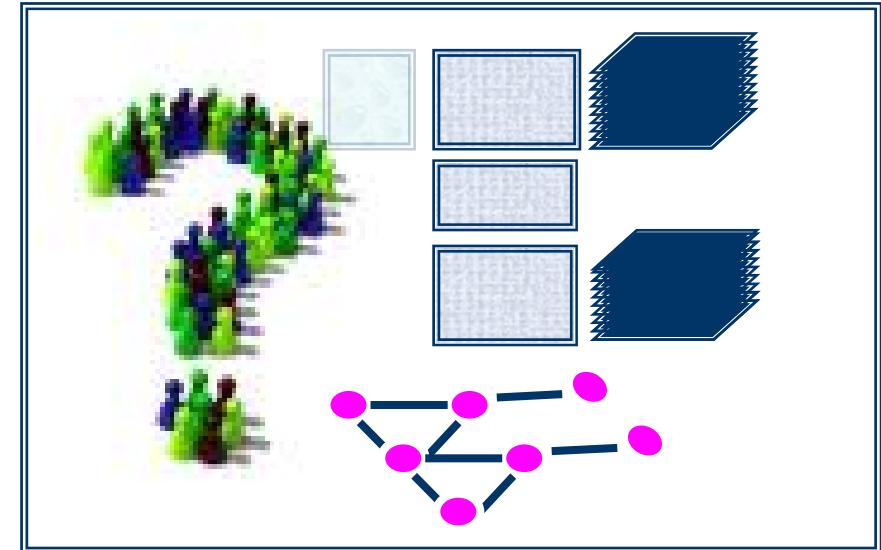
# Méthodologie d'analyse des données structurées en blocs

- Se partage avec la méthodologie d'Analyse des Données Multidimensionnelles (ADM) les même objectifs de traitement ( synthèse, visualisation, discrimination, régression, classification, etc.. )
- Se différencie de la méthodologie ADM par la structure de données à analyser.
- Etend les techniques ADM usuelles en présence de données structurées en blocs.
- Combine des outils et des concepts mathématiques, statistiques et informatiques dans le but :
  - d' apprécier de manière efficace les informations structurelles telles qu'elles figurent dans différents blocs de données ?
  - de relier différents blocs de données?
  - d' exploiter, in fine, au mieux des données acquises tout en préservant l'intégrité de chacun des différents blocs?

# Bilan et état de l'art

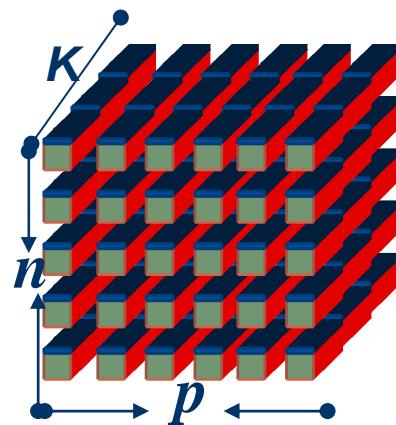


problématiques traitées  
1930-à ce jour  
deux communautés séparées



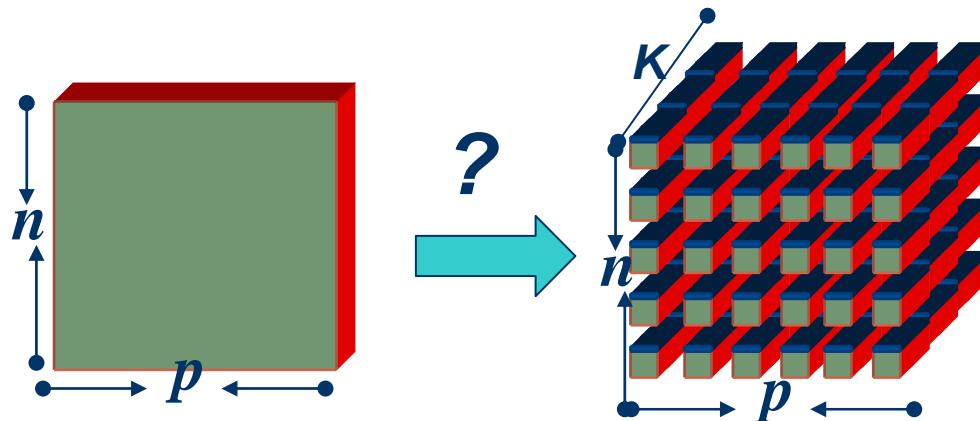
Problématique émergente  
de la généricité des données à la généricité  
des modèles

# PARTIE 3

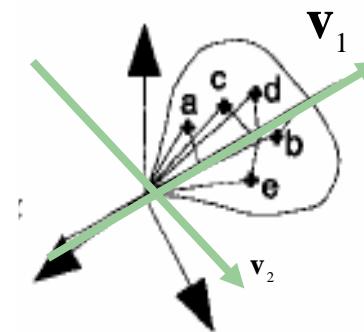
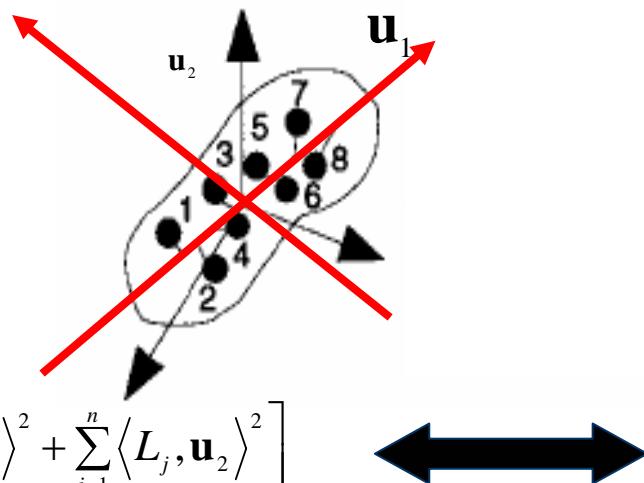
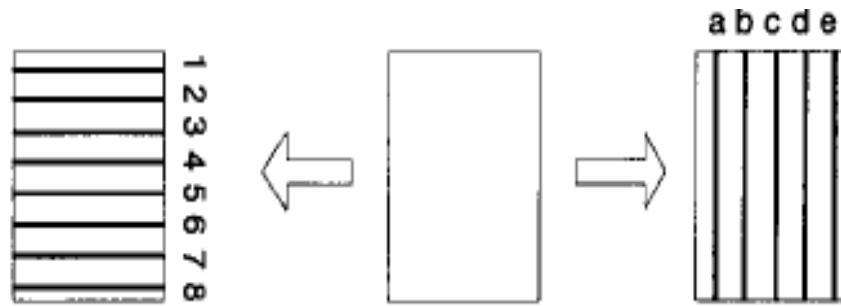


# Problématique

- Etendre de techniques usuelles spécifiques à l'analyse des tableaux à données à deux entrées aux tableaux à trois entrées et plus.
- Exemple : l'Analyse en Composantes Principales.



# Analyse en Composante Principales



# Point de départ

Le problème d'Eckart-Young

$$\underset{\text{rang}(\hat{\mathbf{T}})=r}{\text{Min}} \|\mathbf{T} - \hat{\mathbf{T}}\|^2$$

Eckart C. et Young G. (1936). Psychometrika, I, 211-218  
« Approximations d'une matrice par une autre de rang inférieur »,

# Formulation de l'ACP en terme du rang

$$\underset{\text{rang}(\hat{\mathbf{T}})=2}{\text{Min}} \|\mathbf{T} - \hat{\mathbf{T}}\|^2 = \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$

$$\tilde{\mathbf{T}} = \lambda^{(1)} \mathbf{a}_1 \mathbf{b}_1 + \lambda^{(2)} \mathbf{a}_2 \mathbf{b}_2$$

$$\tilde{\mathbf{T}} = \lambda^{(1)} \mathbf{a}_1 \mathbf{b}_1 + \lambda^{(2)} \mathbf{a}_2 \mathbf{b}_2 = \tilde{\mathbf{a}}_1 \tilde{\mathbf{b}}_1 + \tilde{\mathbf{a}}_2 \tilde{\mathbf{b}}_2 = \tilde{\mathbf{A}} \tilde{\mathbf{B}} = \tilde{\mathbf{A}} \Theta (\mathbf{B} \Theta^{-1})$$

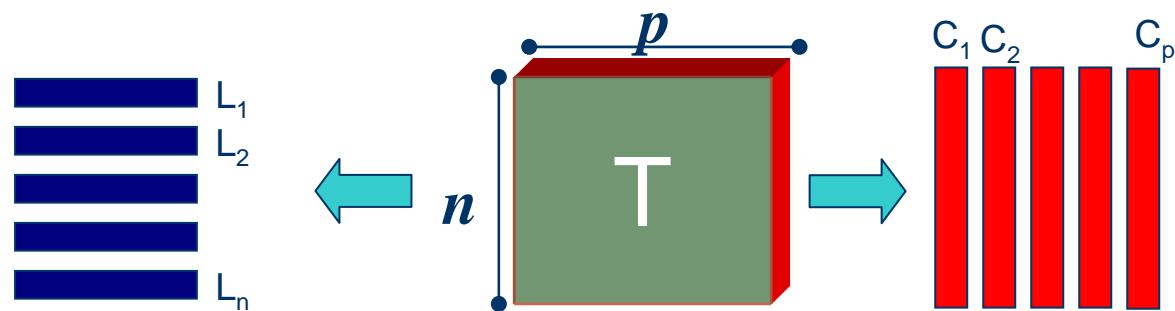
$$\underset{\lambda_i, \|\mathbf{u}_i\| = \|\mathbf{v}_i\| = 1}{\text{Min}} \|\mathbf{T} - \lambda_1 \mathbf{v}_1 \mathbf{u}_1 - \lambda_2 \mathbf{v}_2 \mathbf{u}_2\|^2 \Leftrightarrow$$

$$\underset{\|\mathbf{v}\|=1}{\text{Max}} \left[ \sum_{j=1}^p \langle C_j, \mathbf{v}_1 \rangle^2 + \sum_{i=1}^n \langle C_j, \mathbf{v}_2 \rangle^2 \right] \Leftrightarrow \underset{\|\mathbf{v}\|=1}{\text{Max}} \left[ \sum_{j=1}^p \langle L_i, \mathbf{u}_2 \rangle^2 + \sum_{i=1}^n \langle L_i, \mathbf{u}_2 \rangle^2 \right]$$

# Démarche

- **Etape 1 : Généraliser la notion du rang d'un tableaux à deux entrées à plusieurs entrées**
- **Etape 2 : Formuler le problème d'Eckat-Young**
- **Etape 3: Résoudre du problème et exhiber des propriétés de la solution (Recherches en cours).**

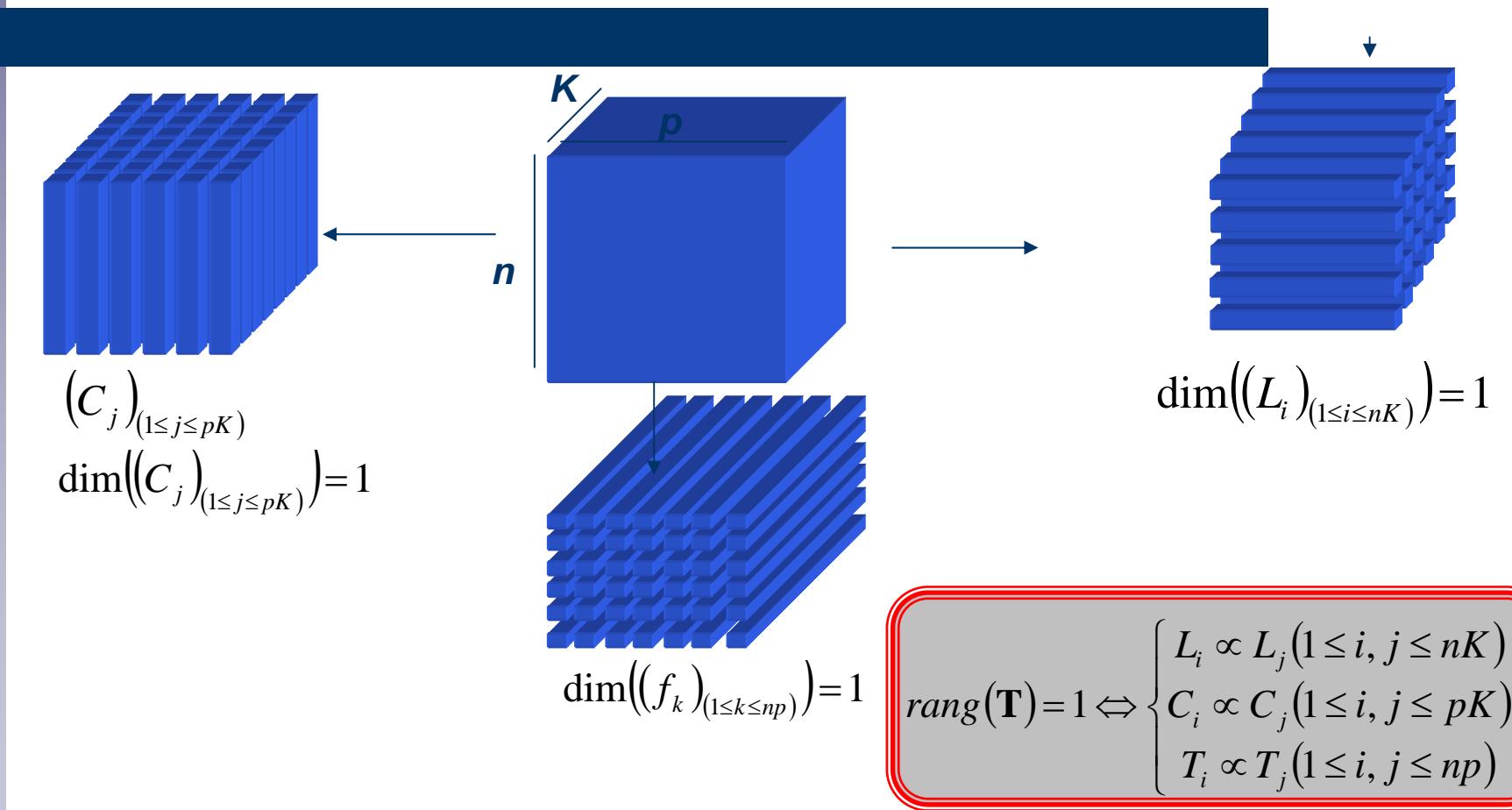
# Tableaux deux entrées de rang 1



$$\begin{aligned} \text{rang}(\mathbf{T}) = 1 &\Leftrightarrow \dim(L_i \ (1 \leq i \leq n)) = \dim(C_j \ (1 \leq j \leq p)) = 1 \\ \text{rang}(\mathbf{T}) = 1 &\Leftrightarrow L_i \propto L_j \ (1 \leq i, j \leq n) \Leftrightarrow C_i \propto C_j \ (1 \leq i, j \leq p) \end{aligned}$$

$$\text{rang}(\mathbf{T}) = 1 \Leftrightarrow x_{ij} = \lambda u_i v_j \Leftrightarrow \mathbf{X} = \lambda \mathbf{u} \mathbf{v}^T = \lambda \mathbf{u} \circ \mathbf{v} \quad (\|\mathbf{u}\| = \|\mathbf{v}\|)$$

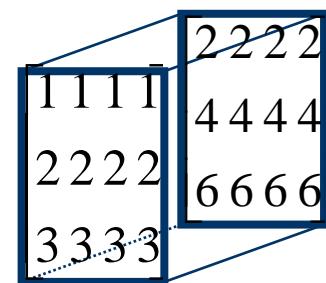
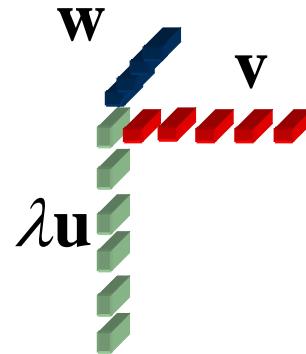
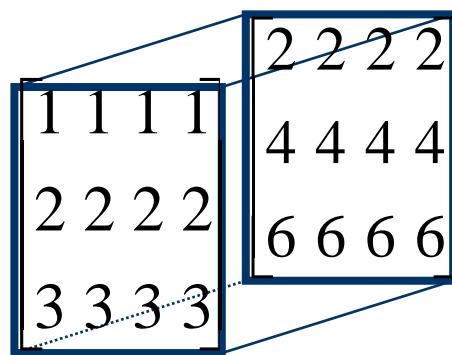
# Tableaux à trois entrées de rang 1



# Exemple et notations

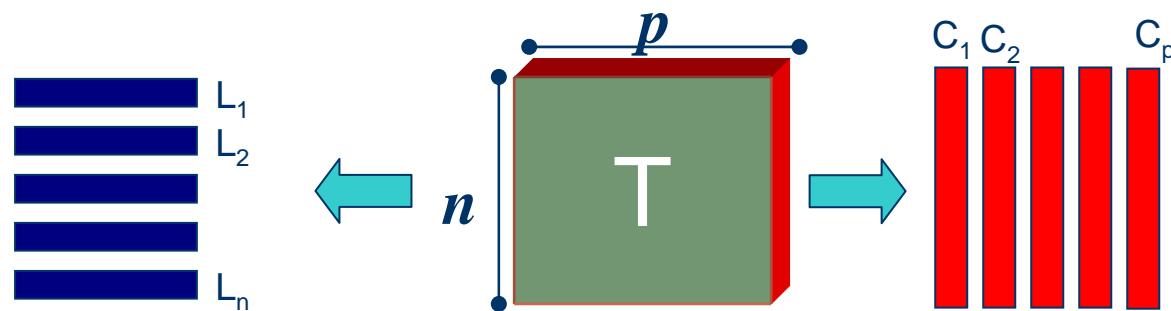
$$\begin{bmatrix} 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 6 & 6 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$



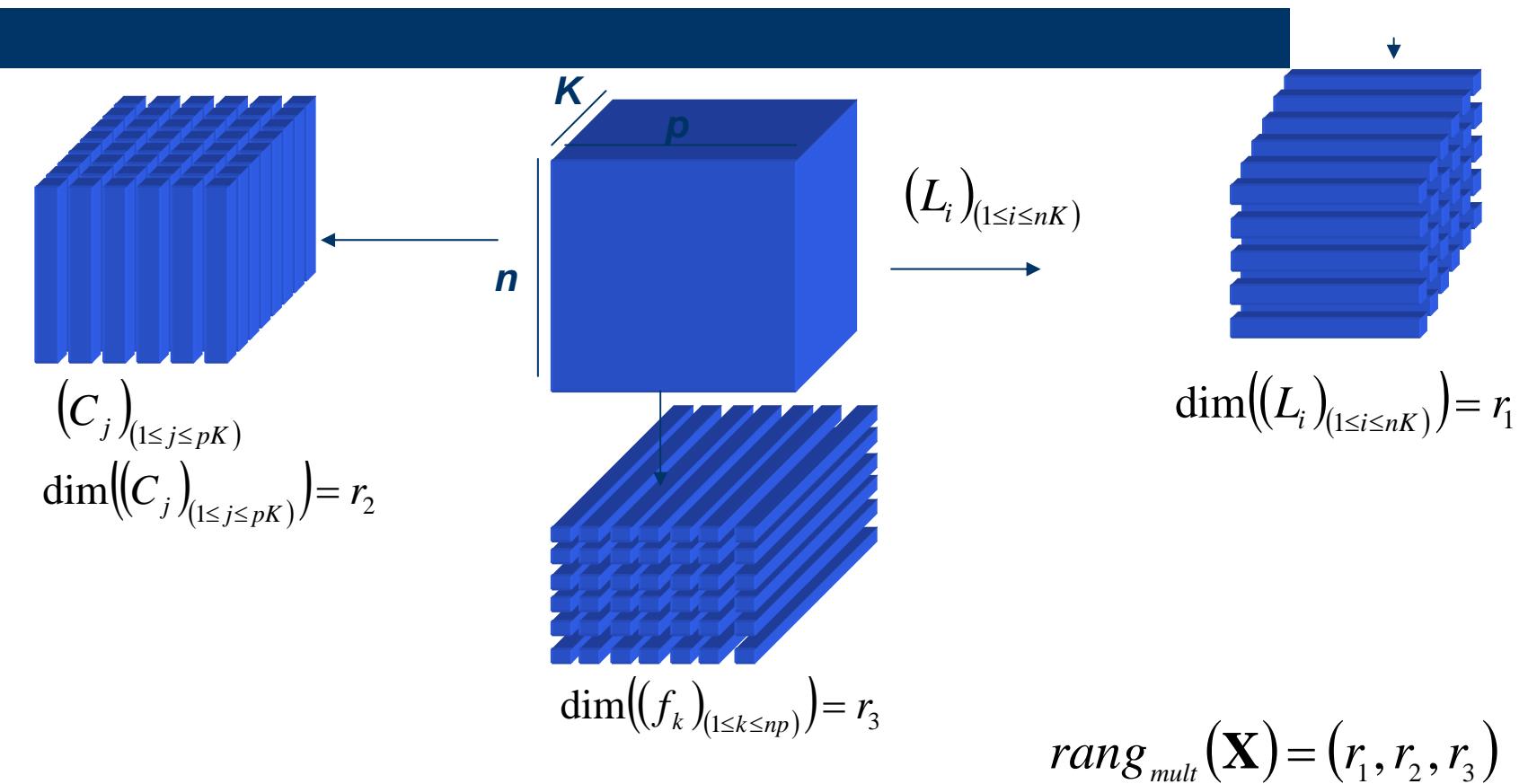
$$\lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w} = \lambda [u_i v_j w_k] \quad (\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| = 1)$$

# Tableaux deux entrées de rang $r$



$$\text{rang}(\mathbf{T}) = r \Leftrightarrow \dim((L_i) (1 \leq i \leq n)) = \dim(C_j (1 \leq j \leq p)) = r$$

# Proposition 1



# Tableaux à deux entrées de rang r

- Décomposition d'un tableau : toute écriture du tableau comme somme finie ( $L$ ) de tableaux tous de rang 1.

$$T = \lambda^{(1)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_1 \circ v_1 + \lambda^{(2)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_2 \circ v_2 + \dots + \lambda^{(L)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_L \circ v_L$$

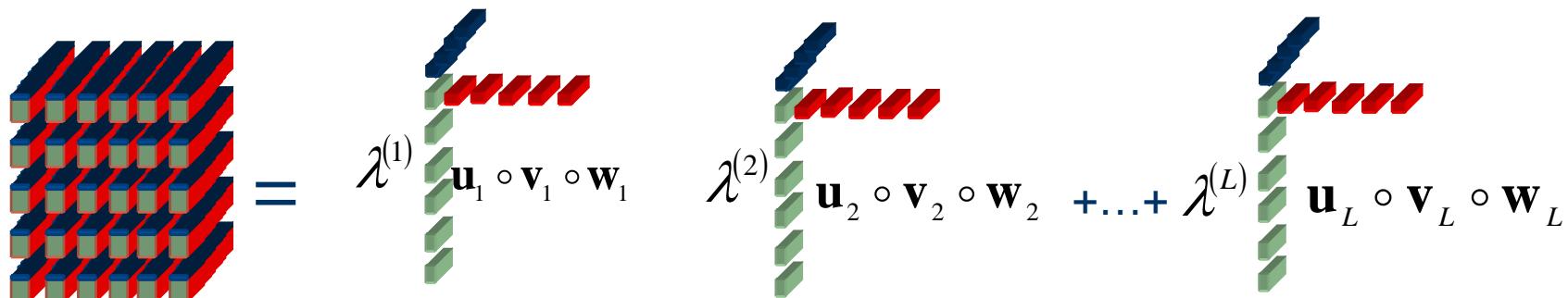
$L$  = longueur de la décomposition.

- Une décomposition est dite minimale lorsque sa longueur est minimale.
- Un tableau est de rang égale à  $r$  si la longueur minimale de toute décomposition de celui-ci est  $r$

$$T = \lambda^{(1)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_1 \circ v_1 + \lambda^{(2)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_2 \circ v_2 + \dots + \lambda^{(r)} \begin{array}{c} \text{---} \\ | \\ \text{---} \end{array} u_r \circ v_r$$

# Proposition 2

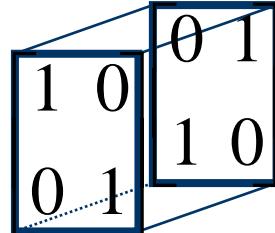
- Décomposition d'un tableau : toute écriture du tableau comme somme finie ( $L$ ) de tableaux tous de rang 1.



**L = longueur de la décomposition.**

- Une décomposition est dite minimale lorsque sa longueur est minimale.
- Un tableau est de rang égale à  $r$  si la longueur minimale de toute décomposition de celui-ci est  $r$
- Notation  $\text{rang}(X)$

## Exemple :



$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Décomposition de longueur L=4.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 0 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

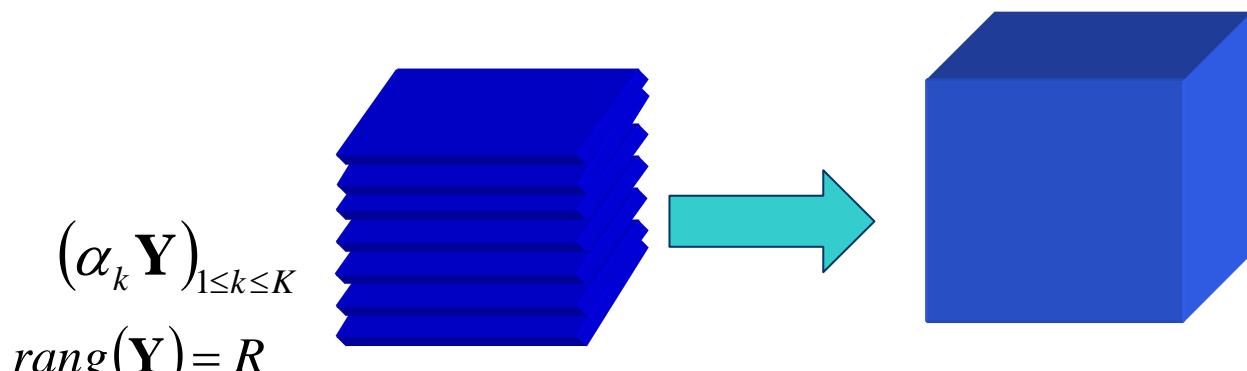
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Décomposition minimale de longueur L=2.

Le rang est 2.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

# Deux propositions bien différentes



$$rang(\mathbf{T}) = R$$

$$rang_{mult}(\mathbf{T}) = (R, R, 1)$$

# Formulation du problème d'Eckart-Young pour tableaux à trois entrées

$$\underset{rang_{Mult}(\tilde{\mathbf{T}}) = (r_1, r_2, r_3)}{Min} \left\| \mathbf{T} - \tilde{\mathbf{T}} \right\|^2$$

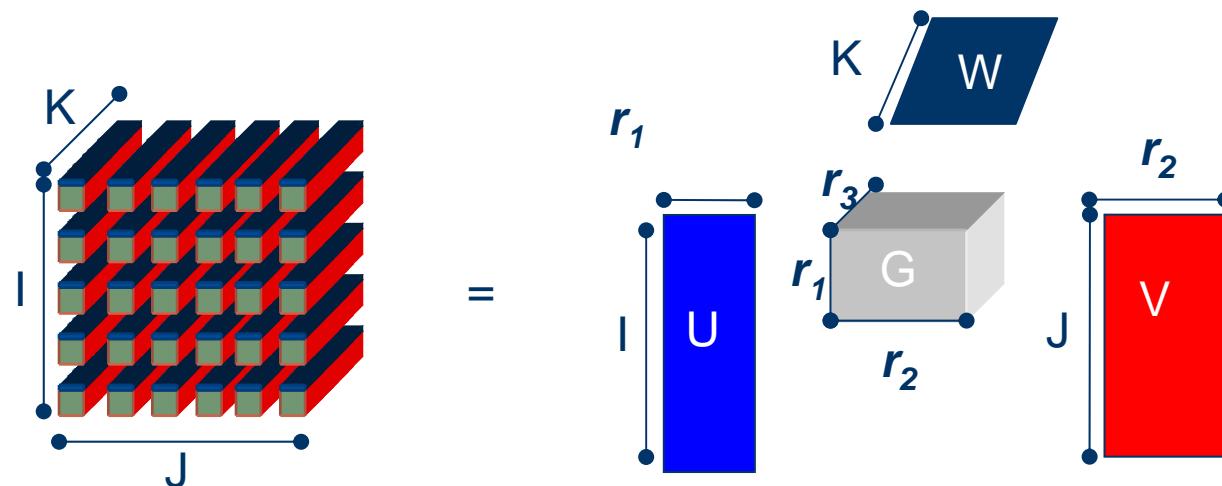
$$\underset{rang(\tilde{\mathbf{T}}) = r_0}{Min} \left\| \mathbf{T} - \tilde{\mathbf{T}} \right\|^2$$

# Formulation du problème d'Eckart-Young pour tableaux à trois entrées

$$\underset{rang_{mult}(\tilde{\mathbf{T}}) = (r_1, r_2, r_3)}{Min} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$

- Three-mode factor analysis (3MFA/Tucker3) Tucker, 1966
- Three-mode PCA (3MPCA) Kroonenberg and De Leeuw, 1980
- $N$ -mode PCA Kapteyn et al., 1986
- Higher-order SVD (HOSVD) De Lathauwer et al., 2000

$$rang_{mult}(\tilde{\mathbf{T}}) = (r_1, r_2, r_3)$$

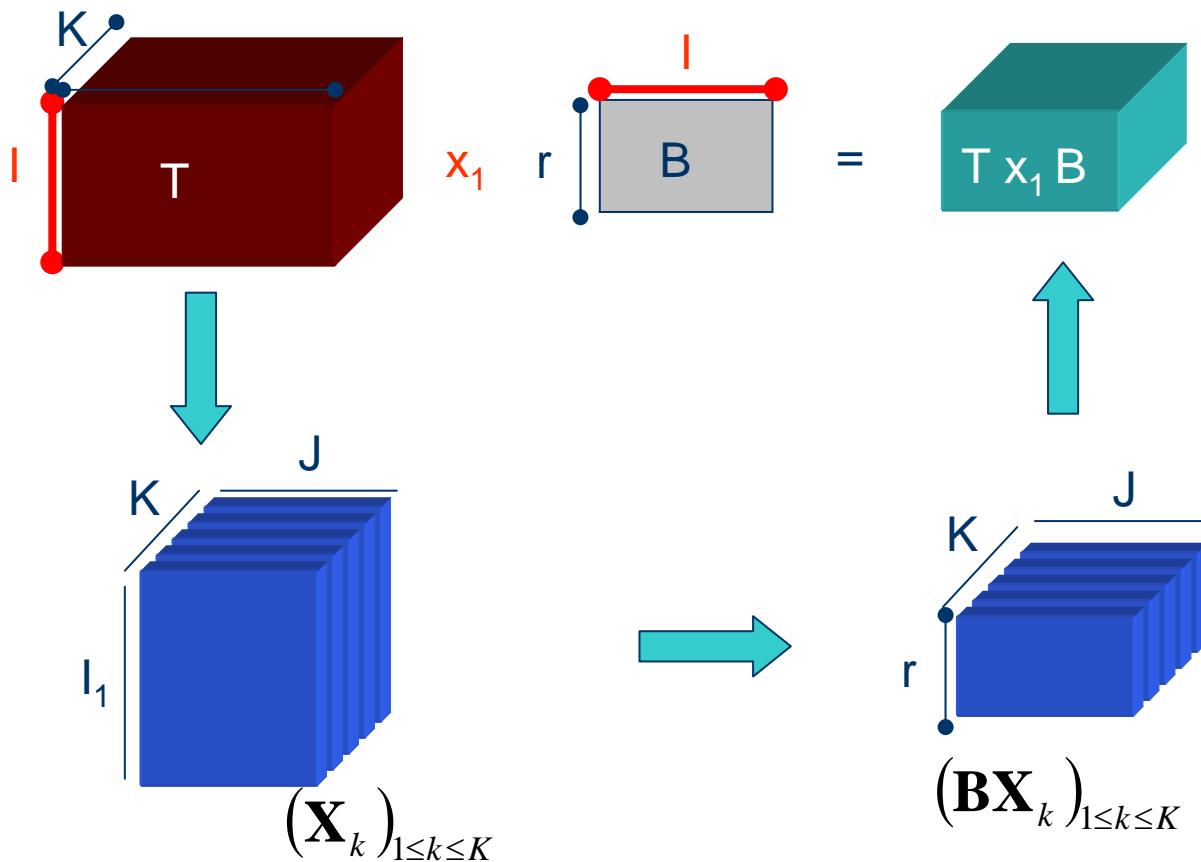


$$\tilde{\mathbf{T}} = \mathbf{G} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}$$

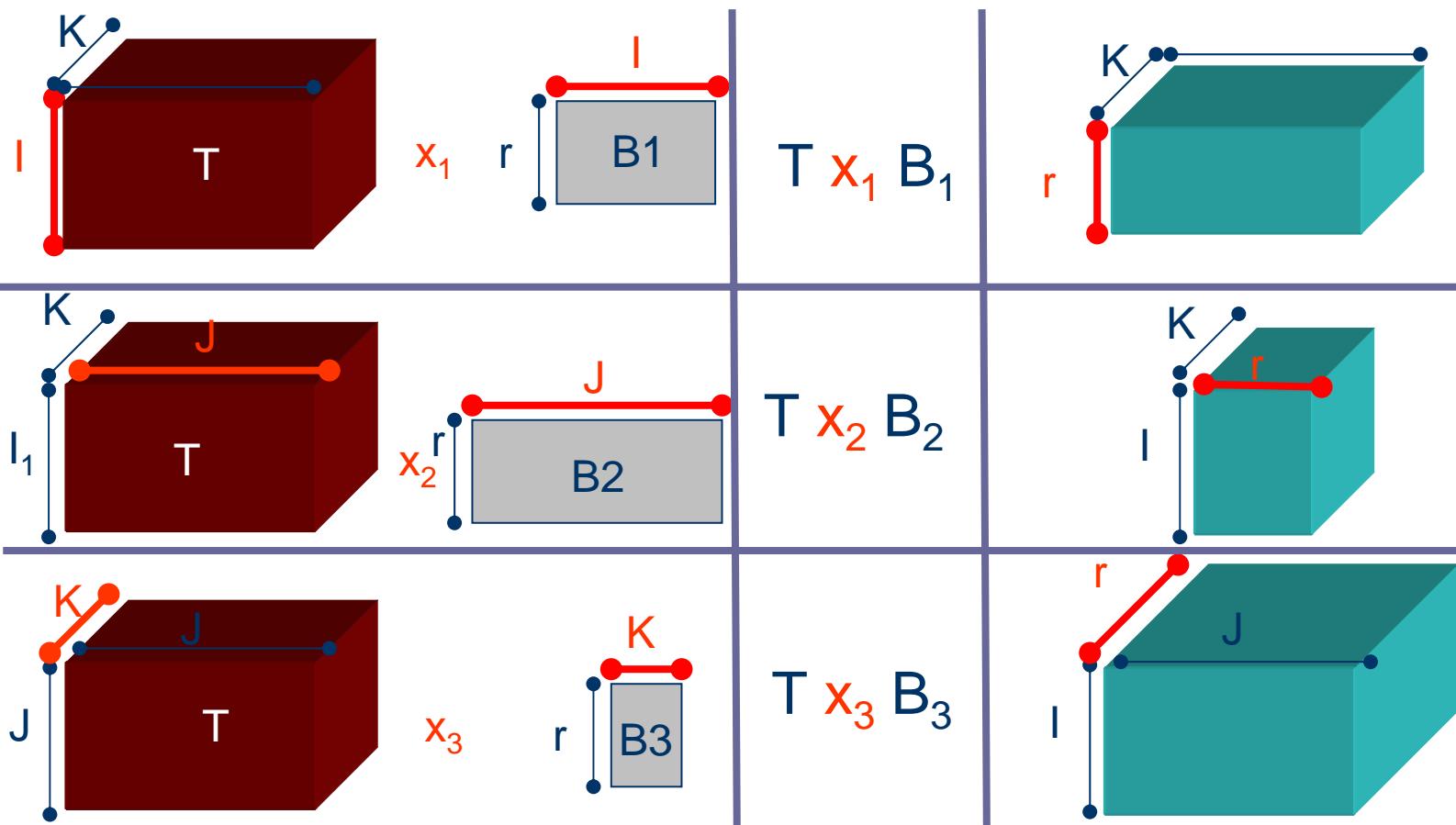
$$\mathbf{U}'\mathbf{U} = \mathbf{I}_{r_1}, \mathbf{V}'\mathbf{V} = \mathbf{I}_{r_2}, \mathbf{W}'\mathbf{W} = \mathbf{I}_{r_3}$$

Higher-order SVD (HOSVD) De Lathauwer et al., 2000

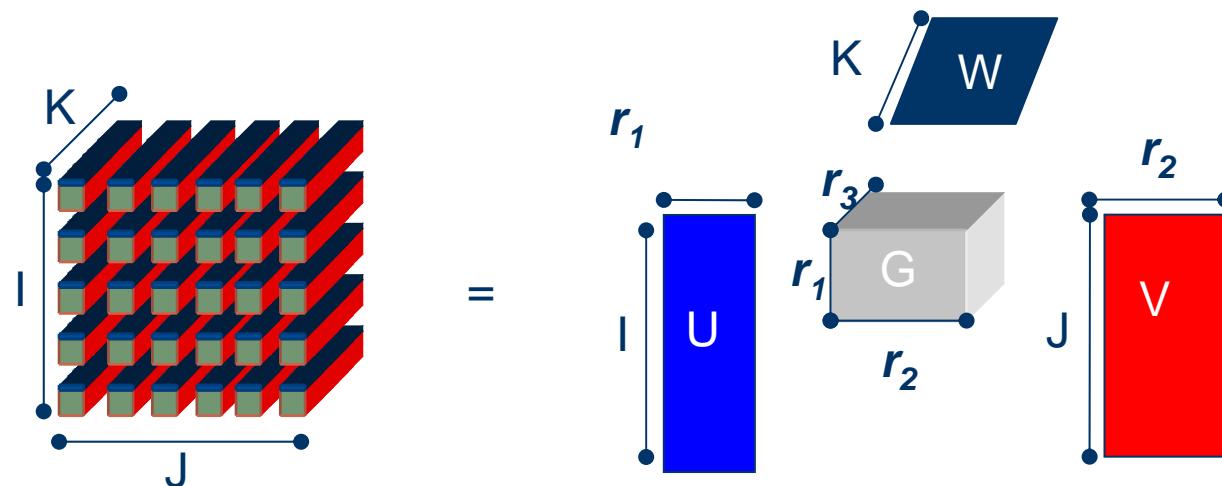
# Illustration



# Produit n mode



$$rang_{mult}(\tilde{\mathbf{T}}) = (r_1, r_2, r_3)$$

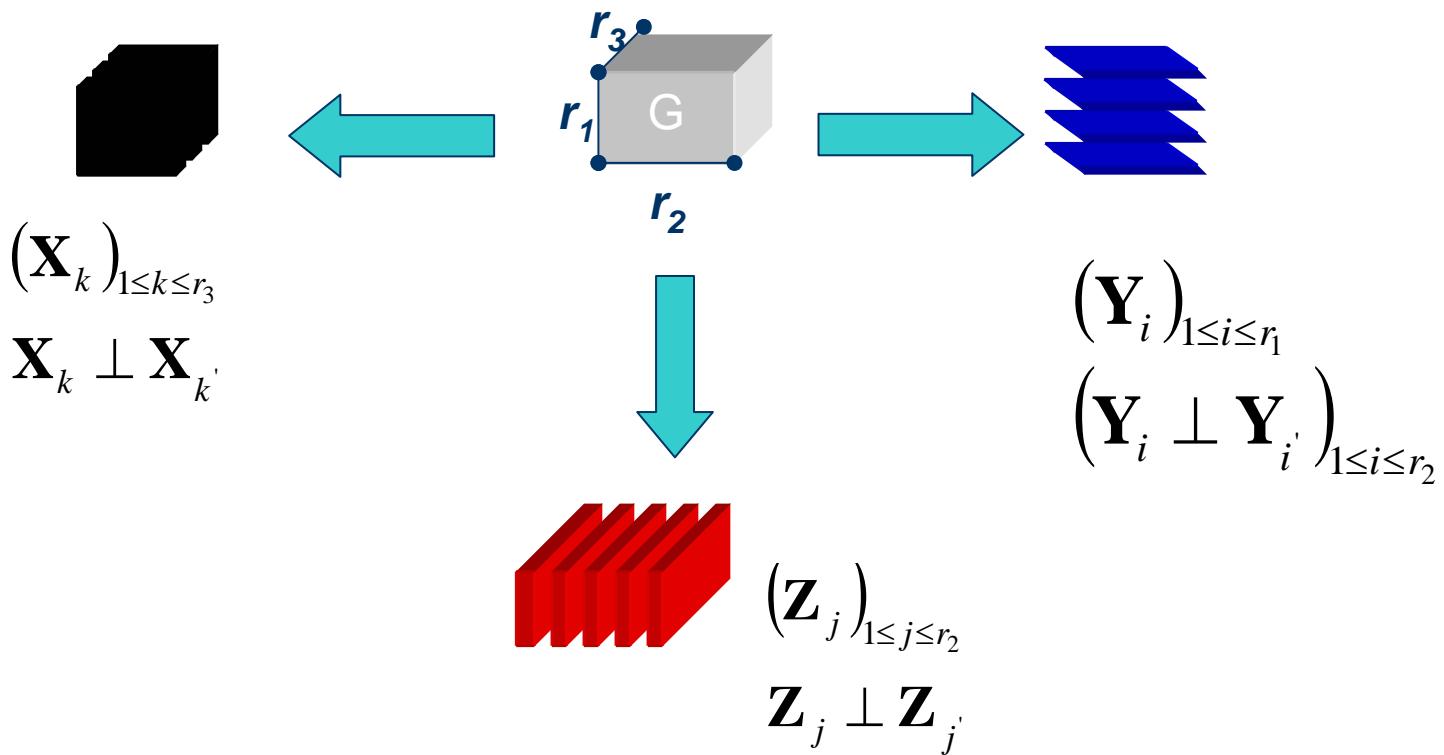


$$\tilde{\mathbf{T}} = \mathbf{G} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}$$

$$\mathbf{U}'\mathbf{U} = \mathbf{I}_{r_1}, \mathbf{V}'\mathbf{V} = \mathbf{I}_{r_2}, \mathbf{W}'\mathbf{W} = \mathbf{I}_{r_3}$$

Higher-order SVD (HOSVD) De Lathauwer et al., 2000

# Structure de G



# Formulation du problème d'Eckart-Young pour tableaux à trois entrées

$$\underset{\text{rang}(\tilde{\mathbf{T}})=r_0}{\text{Min}} \left\| \mathbf{T} - \tilde{\mathbf{T}} \right\|^2$$

$$\text{rang}(\tilde{\mathbf{T}}) = r_0 \Rightarrow \tilde{\mathbf{T}} = \sum_{l=1}^{r_0} \lambda^{(l)} \mathbf{a}_l \circ \mathbf{b}_l \circ \mathbf{c}_l$$

- Polyadic form of a tensor (Hitchcock, 1927)
- PARAFAC (parallel factors) (Harshman, 1970)
- CANDECOMP or CAND (canonical decomposition) (Carroll and Chang, 1970)
- Topographic components model (Möcks, 1988)
- CP (CANDECOMP/PARAFAC) (Kiers, 2000)

# Formulation du problème d'Eckart-Young pour tableaux à trois entrées

$$\underset{\text{rang}(\tilde{\mathbf{X}})=r_0}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$

$$\tilde{\mathbf{T}} = \sum_{l=1}^{r_0} \lambda^{(1)} \mathbf{a}_l \circ \mathbf{b}_l \circ \mathbf{c}_l$$

Un problème mal posé car le minimum n'est pas atteint.

$$\mathbf{X} = \mathbf{u}_1 \circ \mathbf{v}_1 \circ \mathbf{w}_2 + \mathbf{u}_1 \circ \mathbf{v}_2 \circ \mathbf{w}_1 + \mathbf{u}_2 \circ \mathbf{v}_1 \circ \mathbf{w}$$

$$\mathbf{X}_n = \left( \mathbf{u}_1 + \frac{1}{n} \mathbf{u}_2 \right) \circ \left( \mathbf{v}_1 + \frac{1}{n} \mathbf{v}_2 \right) \circ \left( \mathbf{w}_1 + \frac{1}{n} \mathbf{w}_2 \right) - n \mathbf{u}_1 \circ \mathbf{v}_1 \circ \mathbf{w}_1$$

$$\mathbf{X} - \mathbf{X}_n = \frac{1}{n} \left[ \mathbf{u}_2 \circ \mathbf{v}_2 \circ \mathbf{w}_1 + \mathbf{u}_2 \circ \mathbf{v}_1 \circ \mathbf{w}_2 + \mathbf{u}_1 \circ \mathbf{v}_2 \circ \mathbf{w}_2 + \frac{1}{n} \mathbf{u}_2 \circ \mathbf{v}_2 \circ \mathbf{w}_2 \right]$$

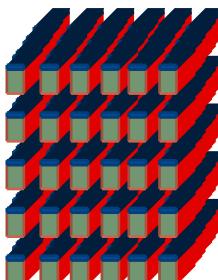
# Alternative.

$$\underset{\lambda, \|\mathbf{w}\|=\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Min}} \|\mathbf{T} - \lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}\|^2$$

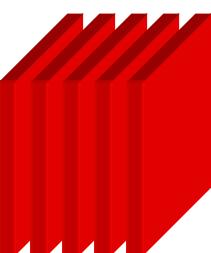


$$(\mathbf{X}_k)_{1 \leq k \leq K}$$

$$\underset{\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Max}} \sum_{k=1}^K \langle \mathbf{X}_k \mathbf{u}, \mathbf{v} \rangle^2$$

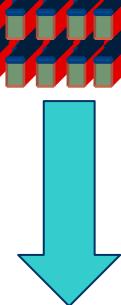


$$(\mathbf{Y}_i)_{1 \leq i \leq n} \underset{\|\mathbf{v}\|=\|\mathbf{w}\|=1}{\text{Max}} \sum_{i=1}^n \langle \mathbf{Y}_i \mathbf{v}, \mathbf{w} \rangle^2$$



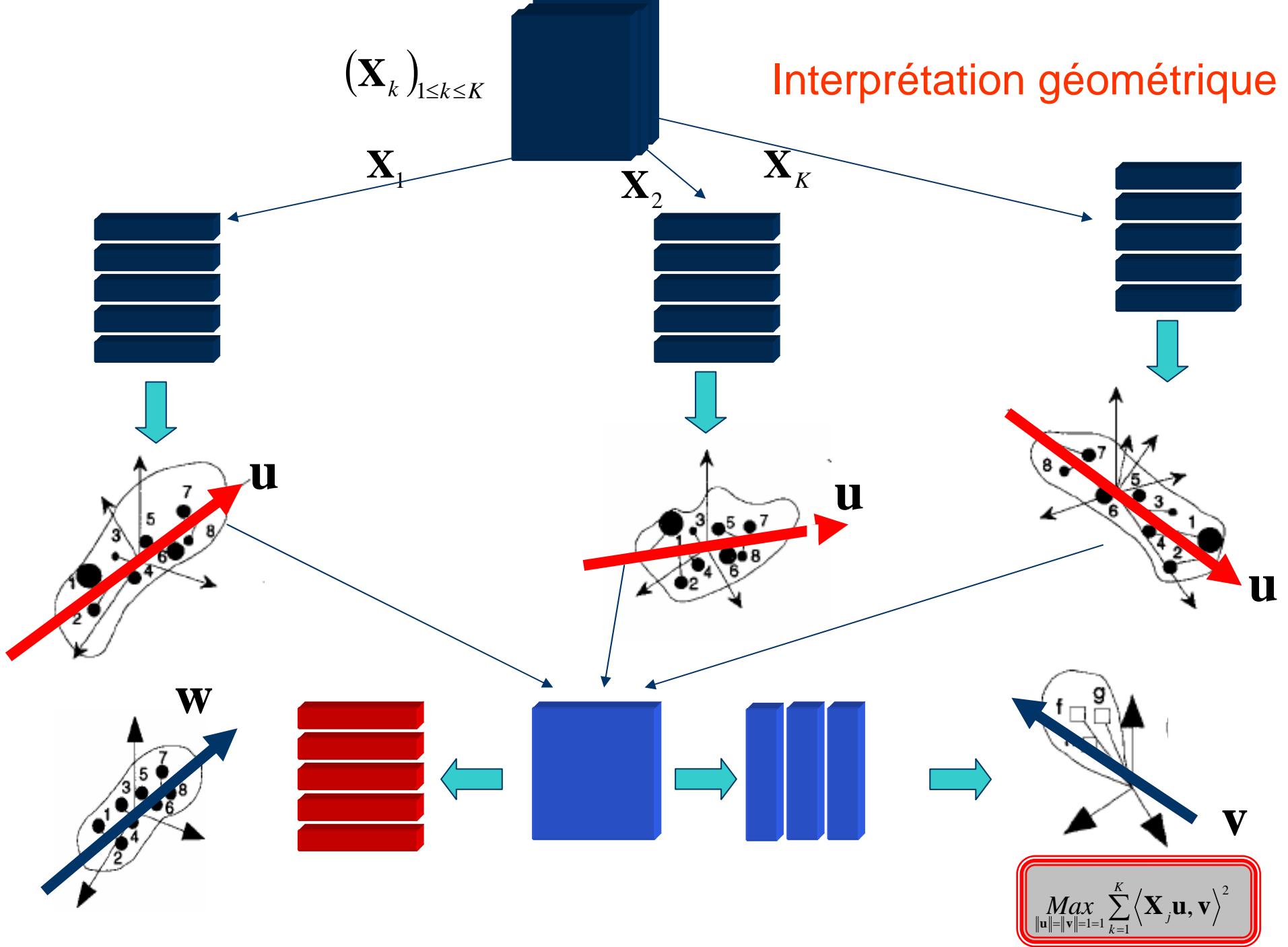
$$(\mathbf{Z}_j)_{1 \leq j \leq p}$$

$$\underset{\|\mathbf{w}\|=\|\mathbf{u}\|=1}{\text{Max}} \sum_{j=1}^p \langle \mathbf{Z}_j \mathbf{w}, \mathbf{u} \rangle^2$$



$$(\mathbf{X}_k)_{1 \leq k \leq K}$$

Interprétation géométrique



# Résolution et remarques.

$$\left[ \sum_{i=1}^n \mathbf{X}_j \mathbf{u}^{(s)} \mathbf{u}^{(s)'} \mathbf{X}_j' \right] \mathbf{v}^{(s+1)} = \lambda_{\max}^{(s)} \mathbf{v}^{(s+1)}$$

$$\left[ \sum_{i=1}^n \mathbf{X}_j \mathbf{v}^{(s+1)} \mathbf{v}^{(s+1)'} \mathbf{X}_j' \right] \mathbf{u}^{(s+1)} = \lambda_{\max}^{(s+1)} \mathbf{u}^{(s+1)}$$

- $\lambda_{\max}^{(s+1)} > \lambda_{\max}^{(s)}$

- $\lambda_{\max}^{(s+1)} \xrightarrow{s \rightarrow \infty} \lambda_{\max}^{(*)}$



Ne garantit pas le calcul de la solution globale

%	79%	4%
Valeur	0.48	0.30



- $\mathbf{u}^{(s)} \xrightarrow{s \rightarrow \infty} \mathbf{u}^{(*)}(?)$

Résultats sur de 100 simulations sur un jeu (10,10,8) de norme = 1

# Dans la pratique ...

$$\underset{rang(\tilde{\mathbf{X}})=r_0}{Min} \left\| \mathbf{T} - \tilde{\mathbf{T}} \right\|_F^2$$

$$\tilde{\mathbf{T}} = \sum_{l=1}^{r_0} \lambda^{(1)} \mathbf{a}_l \circ \mathbf{b}_l \circ \mathbf{c}_l$$

$$rg(\mathbf{T}_n) = r_0$$

$$\delta_n = \left\| \mathbf{T} - \mathbf{T}_n \right\|^2$$

$$\delta_n \rightarrow 0$$

$$\underset{rang_{Mult}(\tilde{\mathbf{X}})=(r_1, r_2, r_3)}{Min} \left\| \mathbf{T} - \tilde{\mathbf{T}} \right\|_F^2$$

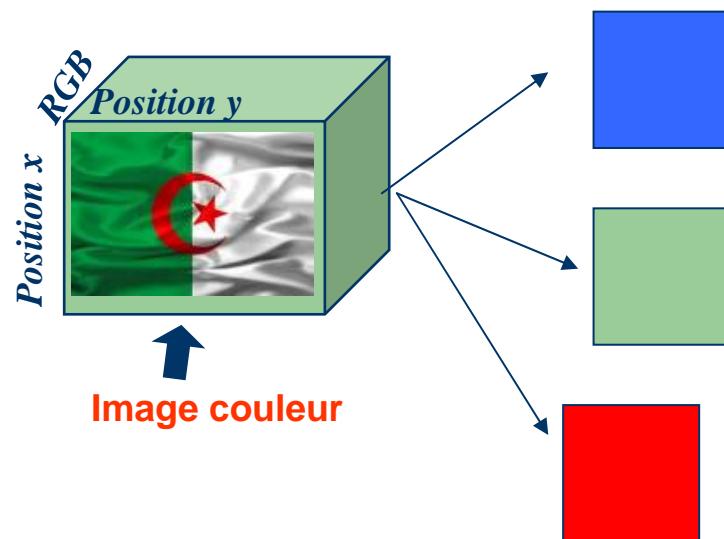
$$\tilde{\mathbf{T}} = \mathbf{G} \times_1 \mathbf{U} \times_1 \mathbf{V} \times_3 \mathbf{W}$$

$$rang_{Mult}(\mathbf{T}_n) = (r_1, r_2, r_3)$$

$$\delta_n = \left\| \mathbf{T} - \mathbf{T}_n \right\|^2$$

$$\delta_n \rightarrow 0$$

# Exemple





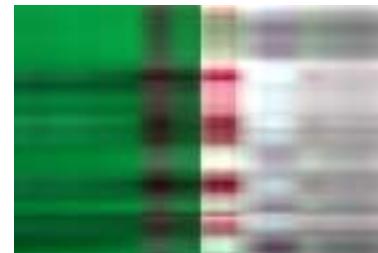
R=1



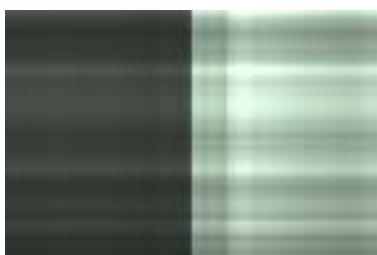
R=2



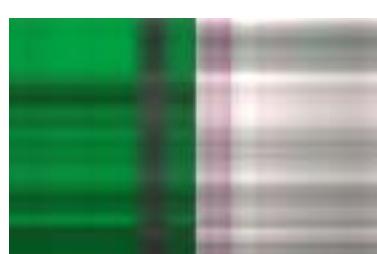
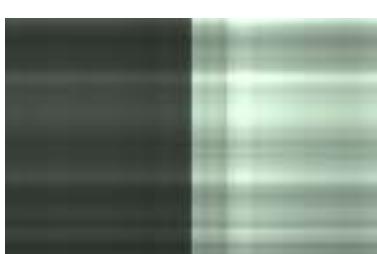
R=3



$$\underset{\text{rang}(\tilde{\mathbf{T}})=R}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



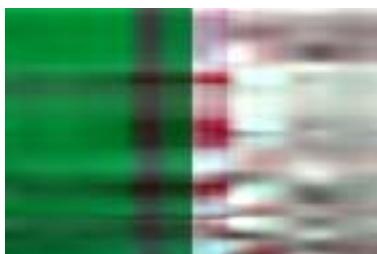
$$\underset{\lambda, \|\mathbf{w}\|=\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Min}} \|\mathbf{T} - \lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}\|^2$$



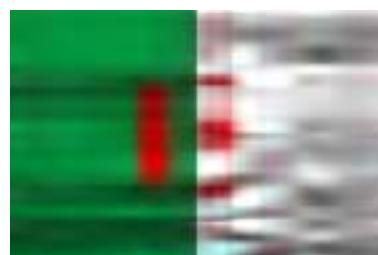
$$\underset{\text{rang}_{\text{mult}}(\tilde{\mathbf{T}})=(R,R,3)}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



R=4



R=5



R=6



$$\underset{\text{rang}(\tilde{\mathbf{T}})=R}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$

$$\underset{\lambda, \|\mathbf{w}\|=\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Min}} \|\mathbf{T} - \lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}\|^2$$

$$\underset{\text{rang}_{\text{mult}}(\tilde{\mathbf{T}})=(R,R,3)}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



R=7



R=8



R=9



$$\underset{\text{rang}(\tilde{\mathbf{T}})=R}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



$$\underset{\lambda, \|\mathbf{w}\|=\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Min}} \|\mathbf{T} - \lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}\|^2$$



$$\underset{\text{rang}_{\text{mult}}(\tilde{\mathbf{T}})=(R,R,3)}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



R=10



$$\underset{\text{rang}(\tilde{\mathbf{T}})=r_0}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$



$$\underset{\lambda, \|\mathbf{w}\|=\|\mathbf{u}\|=\|\mathbf{v}\|=1}{\text{Min}} \|\mathbf{T} - \lambda \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}\|^2$$

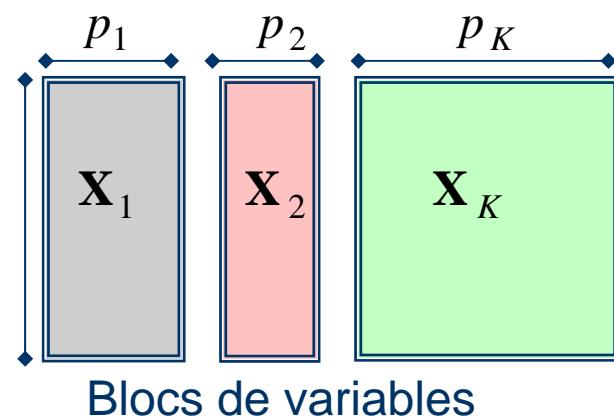


$$\underset{\text{rang}_{\text{mult}}(\tilde{\mathbf{T}})=(r_1, r_2, r_3)}{\text{Min}} \|\mathbf{T} - \tilde{\mathbf{T}}\|^2$$

# Conclusions.

- Pour en savoir davantage
  - A.K. Smilde, R. Bro, P. Geladi. Multiway Analysis. - 2004 John Wiley and Sons.
  - Pieter Kroonenberg . Applied Multiway data analysis. 2008. Wiley series in probability and statistics.
- Outils logiciels
  - Tensor toolbox (Sandia National Laboratories, USA)
  - CuBatch interface software (NwayQual European Project porté par TOTAL)
- Colloques dédiés :
  - TRICAP, Tensor Decompositions.
- Questions méthodologiques émergentes.
  - Extensions dans le contexte de la régression et de la discrimination

# PARTIE 3



# Un exemple

	rouge	bleu	doré	intensité
V1	7	4	2	6
V2	2	0	6	6
V3	6	3	4	7
V4	2	0	6	4
V5	5	1	5	6
V6	3	0	5	5
V7	2	0	4	3
V8	4	0	4	5

	Profond	frais	brillant	
V1	9	7	9	
V2	8	6	7	
V3	10	6	7	
V4	7	7	8	
V5	8	7	8	
V6	8	8	10	
V7	6	5	10	
V8	8	9	10	

	rouge	bleu	doré	intensité
V1	7	4	2	6
V2	2	0	6	6
V3	6	3	4	7
V4	2	0	6	4
V5	5	1	5	6
V6	3	0	5	5
V7	2	0	4	3
V8	4	0	4	5

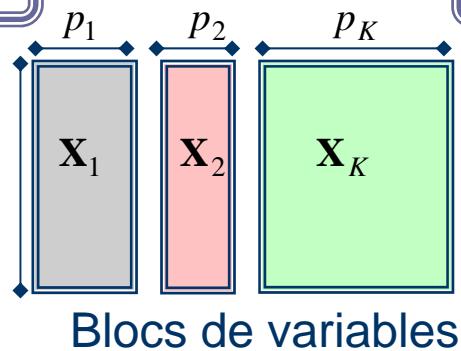
	Profond	frais	brillant	
V1	9	7	9	
V2	8	6	7	
V3	10	6	7	
V4	7	7	8	
V5	8	7	8	
V6	8	8	10	
V7	6	5	10	
V8	8	9	10	

# Concepts méthodologiques

Articulation de la méthodologie sur des généralisations  
de concepts anciens

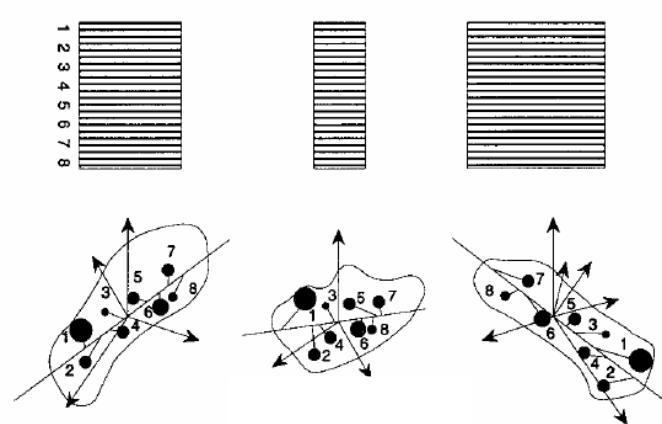
Corrélation canonique

Rotation orthogonale

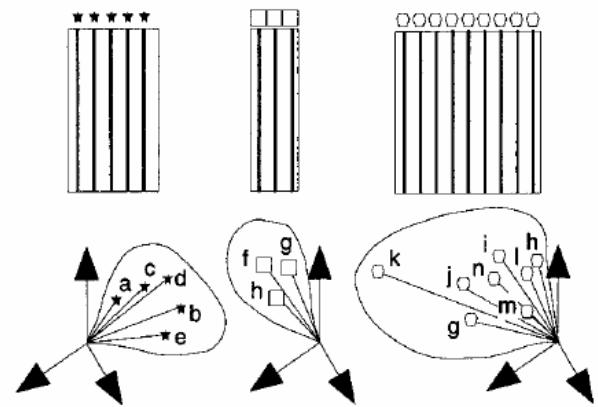


Algorithme NIPALS

# Géométrie du problème



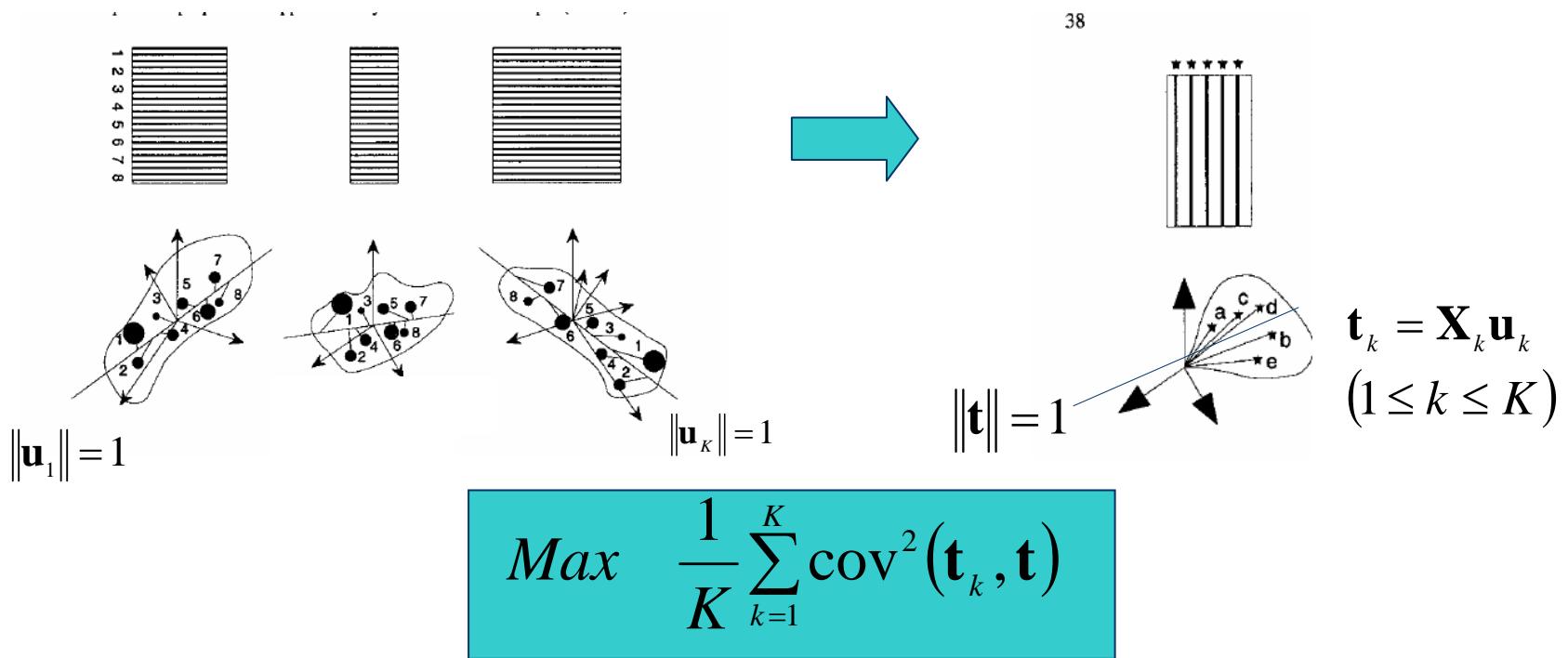
Espaces des individus



Espace des variables

Plusieurs blocs de données

# Démarche



Multiple COinertia Analysis (MCOA); Hierarchical PLS Path Model;  
Consensus Principal Component Analysis (CPCA)

1. SUMCOR (Horst, 1961)  $\text{Max} \sum_{j,k} \text{Cor}(F_j, F_k)$  or  $\text{Max} \sum_j \text{Cor}(F_j, \sum_k F_k)$
2. MAXVAR (Horst, 1961) or GCCA (Carroll, 1968):  $\text{Max } \lambda_{\text{first}}[\text{Cor}(F_j, F_k)]$  or  $\text{Max} \sum_j \text{Cor}^2(F_j, F)$
3. SsqCor (Kettenring, 1971):  $\text{Max} \sum_{j,k} \text{Cor}^2(F_j, F_k)$
4. GenVar (Kettenring, 1971):  $\text{Min} \det[\text{Cor}(F_j, F_k)]$
5. MINVAR (Kettenring, 1971):  $\text{Min} \lambda_{\text{last}}[\text{Cor}(F_j, F_k)]$
6. Lafosse (1989):  $\text{Max} \sum_j \text{Cor}^2(F_j, \sum_k F_k)$
7. Mathes (1993), Hanafi (2004):  $\text{Max} \sum_{j,k} |\text{Cor}(F_j, F_k)|$
8. ACOM (Hanafi & Chessel, 1996):  $\underset{\text{All } w_j = 1, F \text{ standardized}}{\text{Maximize}} \sum_j \text{Cov}^2(X_j w_j, F)$
9. MAXBET (Van de Geer, 1984 & Ten Berge, 1988):  $\underset{\text{All } w_j = 1}{\text{Max}} [\sum_j \text{Var}(X_j w_j) + \sum_{j \neq k} \text{Cov}(X_j w_j, X_k w_k)]$
10. MAXDIFF (Van de Geer, 1984 & Ten Berge, 1988):  $\underset{\text{All } w_j = 1}{\text{Max}} [\sum_{j \neq k} \text{Cov}(X_j w_j, X_k w_k)]$
11. Generalized PCA (Casin, 2001):  $\text{Max} \sum_j R^2(F, X_j) \sum_h \text{Cor}^2(x_{jh}, X_j (X_j' X_j)^{-1} X_j' F)$
12. Split-PCA (Lohmöller, 1989), SUM-PCA (Smilde et al., 2003):  $\underset{F, P_j}{\text{Min}} \sum_j \|X_j - FP_j'\|^2$
13. Multiple Factor Analysis (Escofier & Pagès, 1994):  $\underset{F, P_j}{\text{Min}} \sum_j \frac{1}{\lambda_{\text{first}}[\text{Cor}(x_{jh}, x_{jh})]} \|X_j - FP_j'\|^2$
14. Oblique Maximum Variance method (Horst, 1965):  $\underset{F, P_j}{\text{Min}} \sum_j \|X_j (\frac{1}{n} X_j' X_j)^{-1/2} - FP_j'\|^2$

# Hierarchical Principal Component Analysis (HPCA)

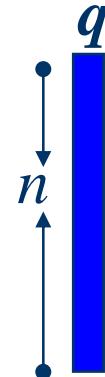
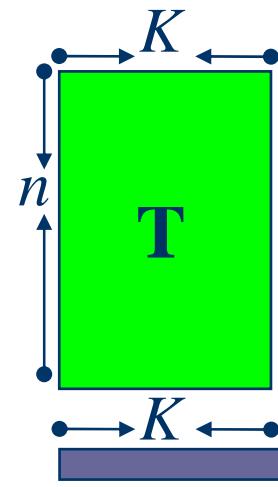
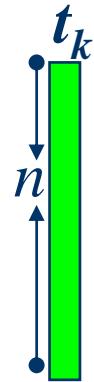
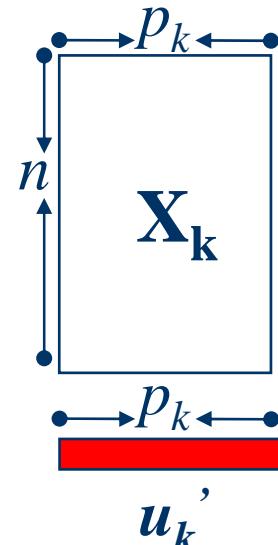
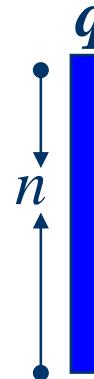
**HPCA une extension intuitive à plusieurs blocs de NIPALS pour ACP.**

- introduite par Wold et al(1996).
  - *Wold S, Tjessem K. J. Chemometrics 1996;10:463-482*
- Comparée à quelques méthodes.
  - *Westerhuis JA, Kourti T, Macgregor JF. J. Chemometrics 1998;12:301-321*
  - *Smilde AK, Westerhuis JA, De Jong S. J. Chemometrics 2003;17:323-337*

# l'algorithme HPCA

- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$

$$\mathbf{X}_k = \mathbf{X}_k - \mathbf{q} \mathbf{q}' \mathbf{X}_k$$



# Problèmes

- Quelles propriétés de convergence.
- Quel critère d'optimisation.

*(Smilde AK, Westerhuis JA, De Jong S. J. Chemometrics 2003;17:323-337)*

- Quels liens avec les autres méthodes .

# Notations

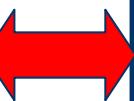
- |  |   |
|--|---|
| • $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$                       | $\mathbf{u}_k^{(1)}, \mathbf{u}_k^{(2)}, \dots, \mathbf{u}_k^{(s)}$ |
| • $\mathbf{u}_k \leftarrow \mathbf{X}_k' \mathbf{q} / \mathbf{q}' \mathbf{q}$              | $\mathbf{t}_k^{(1)}, \mathbf{t}_k^{(2)}, \dots, \mathbf{t}_k^{(s)}$ |
| • $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$                                      | $\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \dots, \mathbf{q}^{(s)}$       |
| • $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K]$ | $\mathbf{w}^{(1)}, \mathbf{w}^{(2)}, \dots, \mathbf{w}^{(s)}$       |
| • $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$                  |   |
| • $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$  |   |

$w_k^{(s)}$  composantes du vecteur  $\mathbf{w}^{(s)}$

$\mathbf{x}_k^{(j)}$  j<sup>th</sup> variable de  $\mathbf{X}_k$

# Forme compacte de l'algorithme

- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \cdots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$



## Propriété 1

$$\sum_{k=1}^K w_k^{(s)} \mathbf{X}_k \mathbf{X}_k' \mathbf{q}^{(s)} = \lambda^{(s)} \mathbf{q}^{(s+1)}$$

$$w_k^{(s)} = \mathbf{q}^{(s)'} \mathbf{X}_k \mathbf{X}_k' \mathbf{q}^{(s)}$$

$$\|\mathbf{q}^{(s)}\| = \|\mathbf{q}^{(s+1)}\| = 1$$

# Properties (à l'iteration s )

$n^4 \times \left[ \sum_{j=1}^{p_k} \text{cov}^2(\mathbf{x}_k^{(j)}, \mathbf{q}^{(s)}) \right]^2$	• Variance totale du block $k$ expliquée par la composante globale
$n^2 \times \text{cov}^2(\mathbf{t}_k^{(s)}, \mathbf{q}^{(s)})$	• Variance de la composante partielle expliquée par la composante globale
$\ \mathbf{X}_k \mathbf{X}_k'\ ^2 - \left\  \mathbf{X}_k \mathbf{X}_k' - w_k^{(s)} \mathbf{q}^{(s)} (\mathbf{q}^{(s)})' \right\ ^2$	<u>Meilleur approximation de rang 1</u>

## Propriété 2

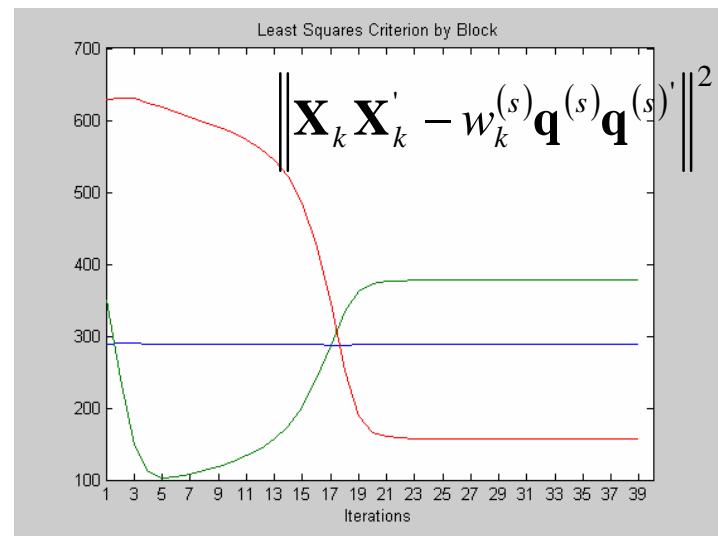
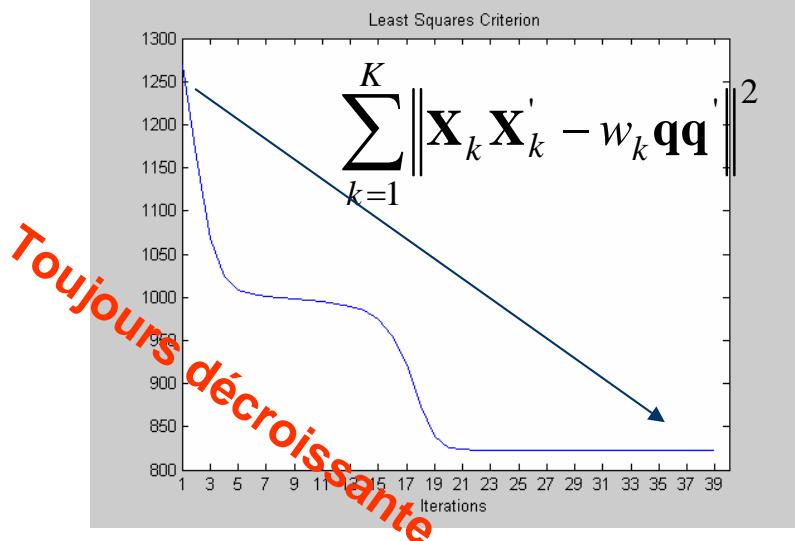
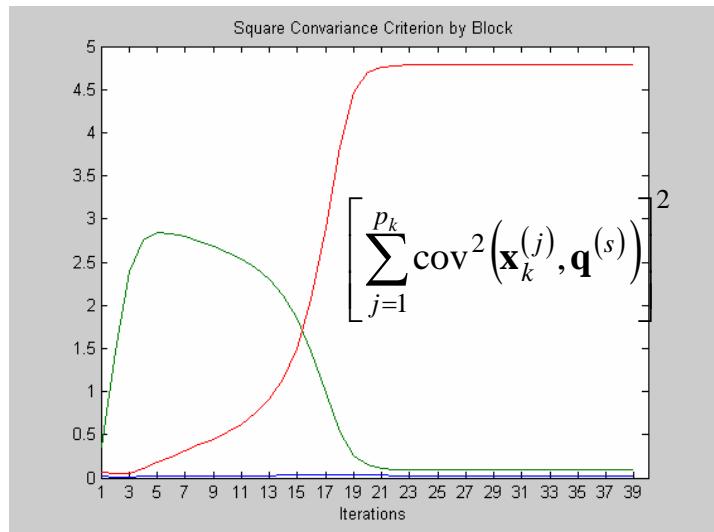
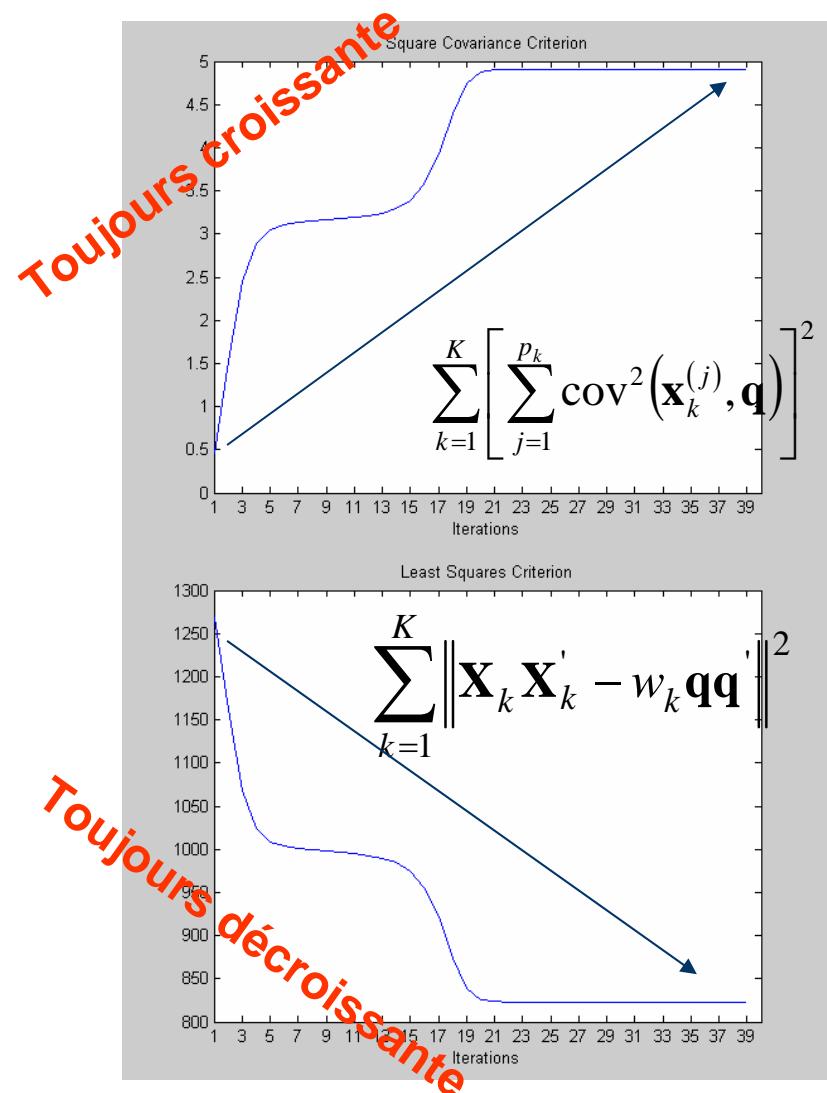
$$n^4 \times \left[ \sum_{j=1}^{p_k} \text{cov}^2(\mathbf{x}_k^{(j)}, \mathbf{q}^{(s)}) \right]^2 = n^2 \times \text{cov}^2(\mathbf{t}_k^{(s)}, \mathbf{q}^{(s)}) = \|\mathbf{X}_k \mathbf{X}_k'\|^2 - \left\| \mathbf{X}_k \mathbf{X}_k' - w_k^{(s)} \mathbf{q}^{(s)} (\mathbf{q}^{(s)})' \right\|^2$$

## Propriété 3

$$n^4 \times \sum_{k=1}^m \left[ \sum_{j=1}^{p_k} \text{cov}^2(\mathbf{x}_k^{(j)}, \mathbf{q}^{(s)}) \right]^2 = n^2 \times \sum_{k=1}^K \text{cov}^2(\mathbf{t}_k^{(s)}, \mathbf{q}^{(s)}) = \sum_{k=1}^K \|\mathbf{X}_k \mathbf{X}_k'\|^2 - \sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k^{(s)} \mathbf{q}^{(s)} (\mathbf{q}^{(s)})' \right\|^2$$

# Propriété de monotonie

X <sub>1</sub>	X <sub>2</sub>			X <sub>3</sub>			
154 -3 -91	-147	51	62	106	-116	68	13
-35 57 -150	-57	1	-109	-75	-167	-67	-28
-14 -15 59	40	56	-192	-211	-37	123	-38
29 -47 135	-18	58	-151	-115	-105	-115	221
-114 -21 202	-187	-58	70	-1	26	38	-32
-82 84 -63	-172	220	-67	-131	-39	-103	155
88 -42 -65	14	40	15	296	134	-38	22
164 123 126	-70	56	202	-33	-1	38	4
-156 36 25	-145	-11	-107	-18	135	-89	51
-121 -20 19	89	111	180	-101	-19	73	158



### Property 4

$$\sum_{k=1}^K \left[ \sum_{j=1}^{p_k} \text{cov}^2(\mathbf{x}_k^{(j)}, \mathbf{q}^{(s)}) \right]^2 \leq \sum_{k=1}^K \left[ \sum_{j=1}^{p_k} \text{cov}^2(\mathbf{x}_k^{(j)}, \mathbf{q}^{(s+1)}) \right]^2$$

### Property 5

$$\sum_{k=1}^K \text{cov}^2(\mathbf{X}_k \mathbf{u}_k^{(s)}, \mathbf{q}^{(s)}) \leq \sum_{k=1}^K \text{cov}^2(\mathbf{X}_k \mathbf{u}_k^{(s+1)}, \mathbf{q}^{(s+1)})$$

### Property 6

$$\sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k^{(s)} \mathbf{q}^{(s)} \mathbf{q}^{(s)'} \right\|^2 \geq \sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k^{(s+1)} \mathbf{q}^{(s+1)} \mathbf{q}^{(s+1)'} \right\|^2$$

HPCA

Hanafi M., A. Kohler, E.M. Qannari (in press).

Shedding new light on Hierarchical Principal Component Analysis. *Journal of Chemometrics.*

### Résultats similaires pour Consensus Principal Component Analysis

Hanafi, M., A. Kohler, E. M. Qannari (in press).

Connections between Multiple Co-inertia Analysis and Consensus Principal Component Analysis.

*Chemometrics and Intelligent Laboratory Systems.*

CPCA

### Property 7

$$\sum_{k=1}^m \left[ \sum_{j=1}^{p_k} \text{cov}^2 \left( \mathbf{x}_k^{(j)}, \mathbf{q}^{(s)} \right) \right]^2 \xrightarrow[s \rightarrow \infty]{} a^{(*)}$$

### Property 8

$$\sum_{k=1}^K \text{cov}^2 \left( \mathbf{X}_k \mathbf{u}_k^{(s)}, \mathbf{q}^{(s)} \right) \xrightarrow[s \rightarrow \infty]{} a^{(*)}$$

### Property 9

$$\sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k^{(s)} \mathbf{q}^{(s)} \mathbf{q}^{(s)'} \right\|^2 \xrightarrow[s \rightarrow 0]{} \sum_{k=1}^m \left\| \mathbf{X}_k \mathbf{X}_k' \right\|^2 - a^{(*)}$$

$$\sum_{k=1}^m \left[ \sum_{j=1}^{p_k} \text{cov}^2 \left( \mathbf{x}_k^{(j)}, \mathbf{q}^{(s)} \right) \right]^2 \xrightarrow[s \rightarrow \infty]{} a^{(*)}$$

$$\sum_{k=1}^K \text{cov}^2 \left( \mathbf{X}_k \mathbf{u}_k^{(s)}, \mathbf{q}^{(s)} \right) \xrightarrow[s \rightarrow \infty]{} a^{(*)}$$

$$\sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k^\top - w_k^{(s)} \mathbf{q}^{(s)} \mathbf{q}^{(s)\top} \right\|^2 \xrightarrow[s \rightarrow 0]{} \sum_{k=1}^m \left\| \mathbf{X}_k \mathbf{X}_k^\top \right\|^2 - a^{(*)}$$

$\Rightarrow$

$$\mathbf{q}^{(s)} \xrightarrow[s \rightarrow \infty]{} \mathbf{q}^{(*)}$$

**Perspective**  
« More difficult »

$$\mathbf{q}^{(s)} \xrightarrow[s \rightarrow \infty]{?} \mathbf{q}^{(*)}$$

## Critères équivalents pour HPCA

Propriété 10

$$\underset{\|\mathbf{q}\|=1}{\operatorname{Max}} \sum_{k=1}^m \left[ \sum_{j=1}^{p_k} \operatorname{cov}^2 \left( \mathbf{x}_k^{(j)}, \mathbf{q} \right) \right]^2$$

Propriété 11

$$\underset{\|\mathbf{q}\|=1}{\operatorname{Max}} \sum_{k=1}^K \left( \mathbf{q}^\top \mathbf{X}_k \mathbf{X}_k^\top \mathbf{q} \right)^2$$

Propriété 12

$$\underset{\|\mathbf{q}\|=1}{\operatorname{Min}} \sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k^\top - w_k \mathbf{q} \mathbf{q}^\top \right\|^2$$

ACCPS

# ACCPS

- Qannari E. M., Wakeling I., Courcoux Ph., MacFie M.F.(2000) *Food Quality and Preference.* 11, 151-154.
- Hanafi M., Qannari E. M,(2008). *Revue de la Société Française de Statistique.*(In french)
- Hanafi M., Mazerolles G., Dufour E., Qannari E. M. (2006). *Journal of Chemometrics.* Vol 20, 5, 172-183
- Qannari E. M., Courcoux Ph., Vigneau E. *Food Quality and Preference .*(2000). 12, 365-368.
- Pram Nielsen J., Bertrand D., Micklander E., Courcoux Ph., Munck L. .(2000) *Journal of Near Infrared Spectroscopy.* 9, 275 –285.
- Courcoux Ph., Devaux F., Bouchet B. .(2000) *Chemometrics and Intelligent Laboratory Systems.* 63, 57-68
- Mazerolles G., Devaux F, Dufour E., Qannari E.M., Courcoux Ph. .(2002) *Chemometrics and Intelligent Laboratory Systems.* 63, 57-68.
- Mazerolles G., Hanafi M., Dufour E., Qannari E. M., Bertrand D. (2006). *Chemometrics and Intelligent Laboratory Systems.* 81, 41-49.

### ACCPS et HPCA sont identiques

$$(w_1^{(1)}, w_2^{(1)}, \dots, w_K^{(1)}, \mathbf{q}_1) \Rightarrow \underset{w_2, w_2, \dots, w_K, \|\mathbf{q}\|=1}{\text{Minimize}} \sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k \mathbf{q} \mathbf{q}' \right\|^2$$

$$\mathbf{X}_k^{(r)} = \mathbf{X}_k - \sum_{j < r} \mathbf{q}_j \mathbf{q}_j' \mathbf{X}_k$$

$$(w_1^{(r)}, w_2^{(r)}, \dots, w_K^{(r)}, \mathbf{q}_r) \Rightarrow \underset{w_1, w_2, \dots, w_K, \|\mathbf{q}\|=1}{\text{Minimize}} \sum_{k=1}^K \left\| \mathbf{X}_k^{(r)} \mathbf{X}_k^{(r)'} - w_k \mathbf{q} \mathbf{q}' \right\|^2$$

## HPCA-ACCPs et PARAFAC

$$\mathbf{X}_k \mathbf{X}_k' = \mathbf{w} \tilde{\mathbf{q}}' + \mathbf{E}$$

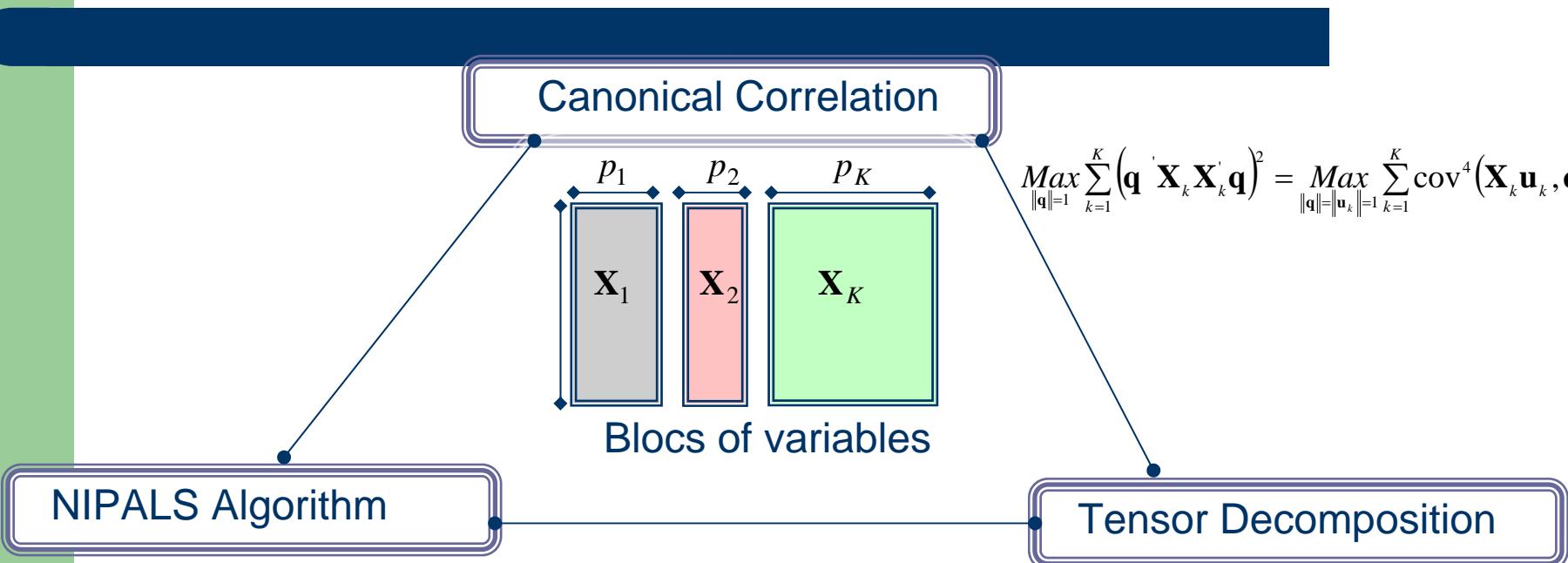
$$\left( w_1^{(1)}, w_2^{(1)}, \dots, w_k^{(1)}, \mathbf{q}^{(1)}, \tilde{\mathbf{q}}^{(1)} \right) \leftarrow \underset{\lambda_1, \lambda_2, \dots, \lambda_K, \|\tilde{\mathbf{q}}\|=\|\mathbf{q}\|=1}{\text{Minimize}} \sum_{k=1}^K \left\| \mathbf{X}_k \mathbf{X}_k' - w_k \mathbf{q} \tilde{\mathbf{q}}' \right\|^2$$

$$\left( \mathbf{q}^{(1)} = \tilde{\mathbf{q}}^{(1)} \right)$$

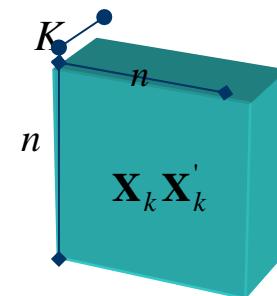
Ten Berge J. M. F., Kiers, H. A. L., Krijnen, W. P. (1993).  
*Journal of classification.* (10), 115-124.

**HPCA-CCSWA = meilleur approximation de rang 1**

## Gap entre trois approches

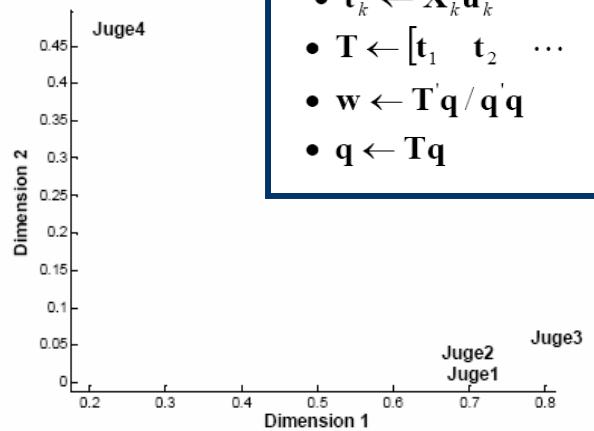


- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$

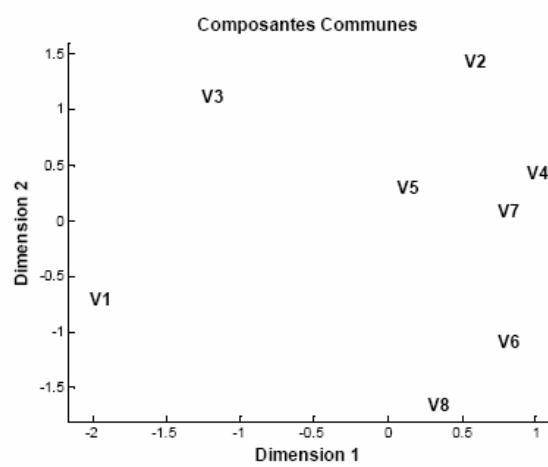


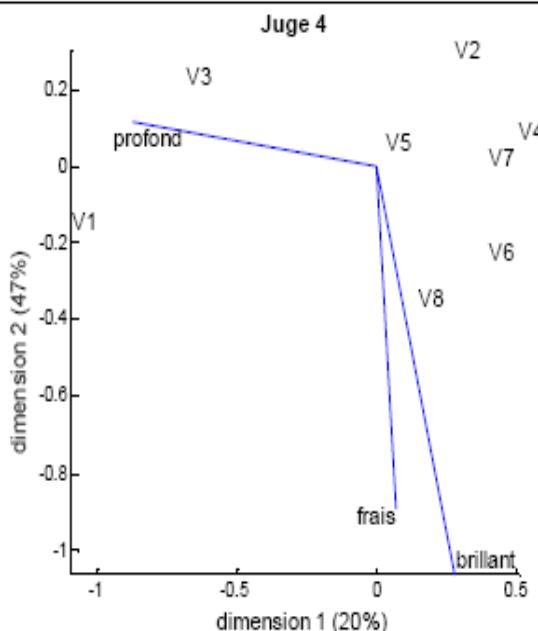
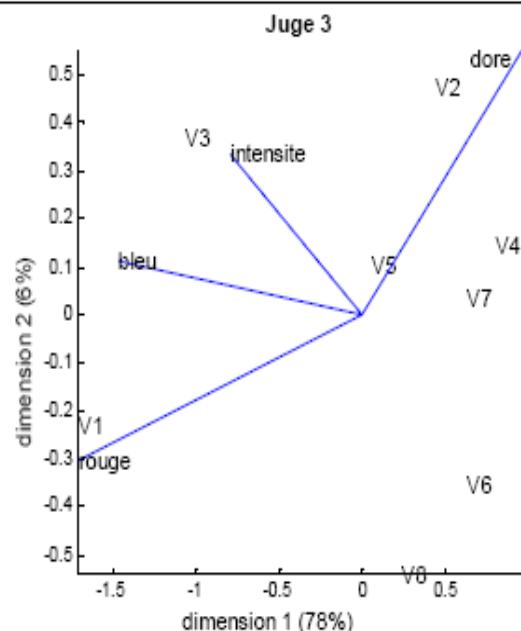
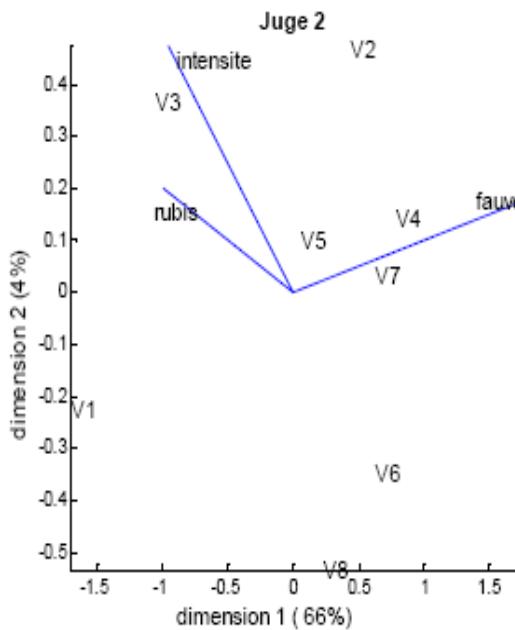
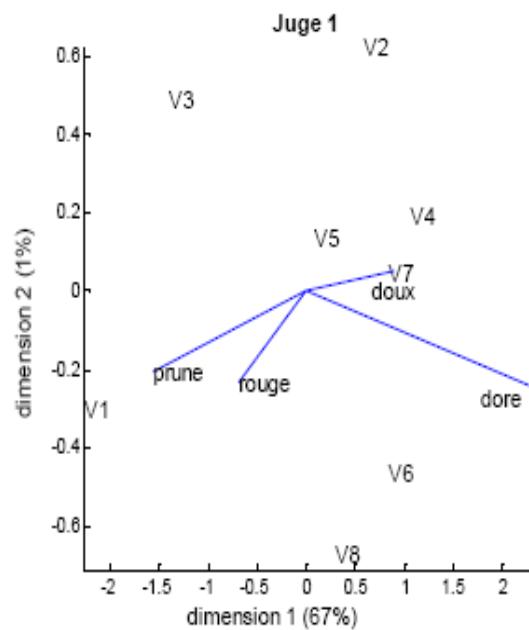
- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$

dimension		1	2	3	4	5	6	7
Juge 1	Poids spécifiques	0.67	0.01	0.19	0.08	0.02	0.01	0.02
	% Inertie Cummulée	0.67	0.68	0.87	0.95	0.97	0.98	1.00
Juge 2	Poids spécifiques	0.66	0.04	0.22	0.03	0.01	0.03	0.01
	% Inertie Cummulée	0.66	0.70	0.92	0.95	0.96	0.99	1.00
Juge 3	Poids spécifiques	0.78	0.06	0.11	0.01	0.02	0.01	0.01
	% Inertie Cummulée	0.78	0.84	0.95	0.96	0.98	0.99	1.00
Juge 4	Poids spécifiques	0.21	0.47	0.28	0.03	0.01	0.00	0.00
	% Inertie Cummulée	0.21	0.68	0.96	0.99	1.00	1.00	1.00



- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \cdots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$





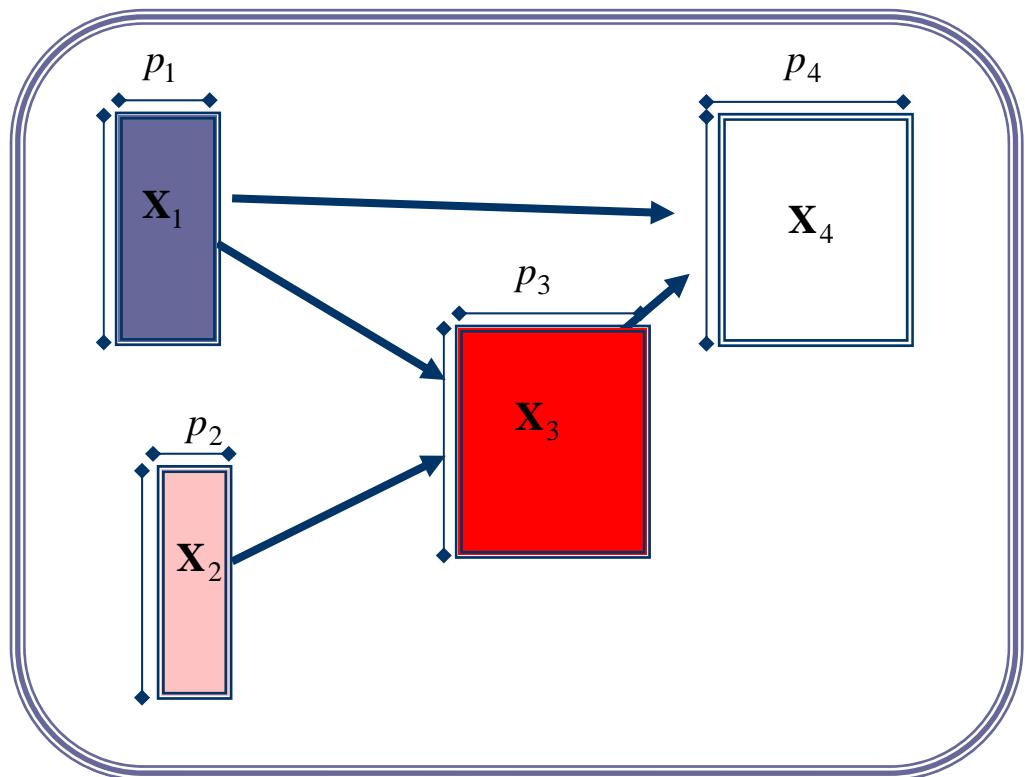
- $\mathbf{q} \leftarrow \mathbf{q} / \sqrt{\mathbf{q}' \mathbf{q}}$
- $\mathbf{u}_k \leftarrow \mathbf{X}_k' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{t}_k \leftarrow \mathbf{X}_k \mathbf{u}_k$
- $\mathbf{T} \leftarrow [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K]$
- $\mathbf{w} \leftarrow \mathbf{T}' \mathbf{q} / \mathbf{q}' \mathbf{q}$
- $\mathbf{q} \leftarrow \mathbf{T} \mathbf{q}$

Figure 3. Représentation simultanée vins-descrivteurs par juge

# Structural modelling with Partial Least Squares

- **Jan-Bernd Lohmöller (1989).** Latent variable path modelling with partial least squares. Physica-Verlag, Heidelberg
- **Herman Wold (1985).** Partial Least Squares. Encyclopedia of statistical sciences , vol 6 Kotz, S & Johnson, N.L(Eds), John Wiley & Sons, New York, pp 581-591.

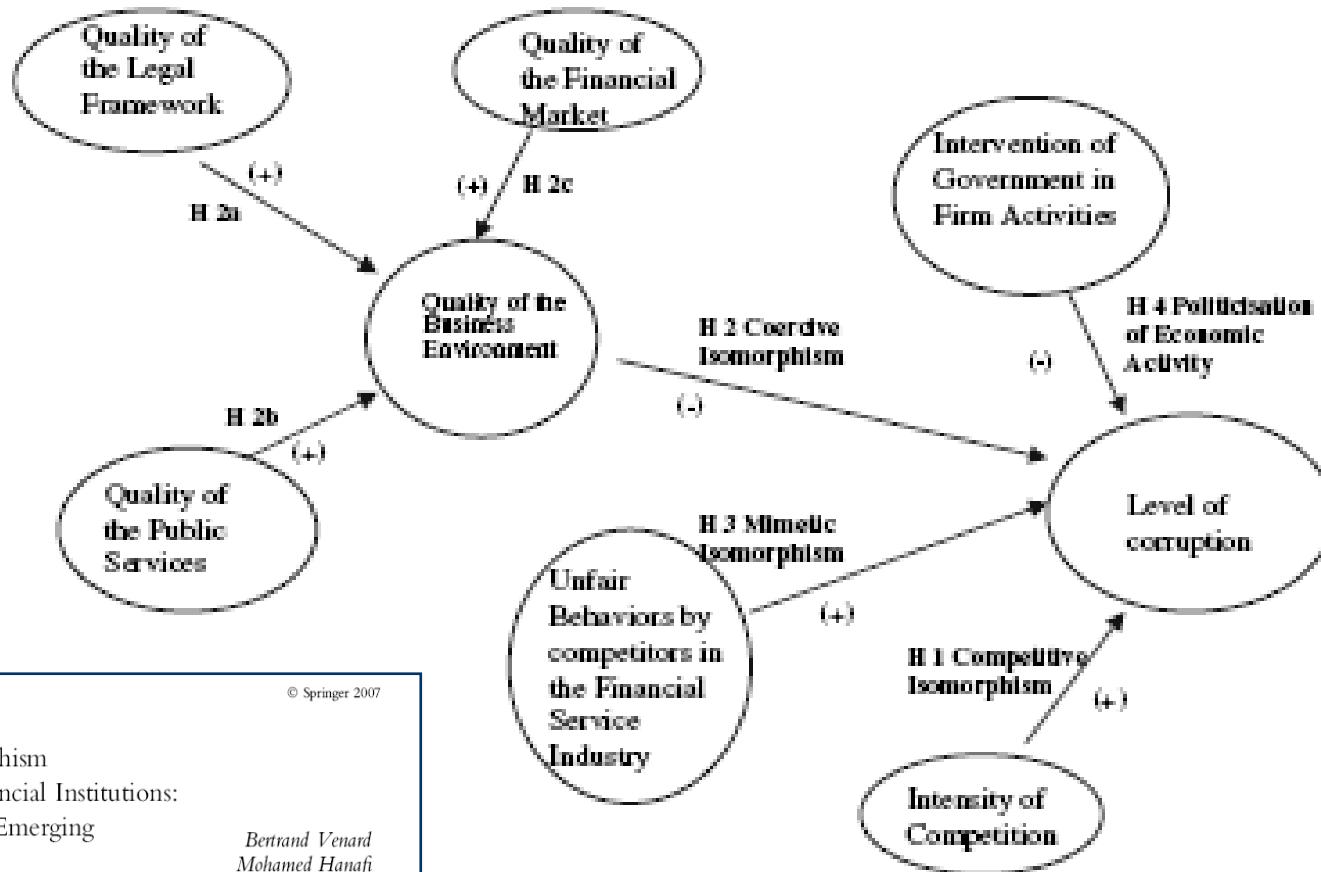
# Démarche en trois étapes



- C matrice binaire (K,K)
- éléments diagonaux 0

$$C = [c_{kl}] = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

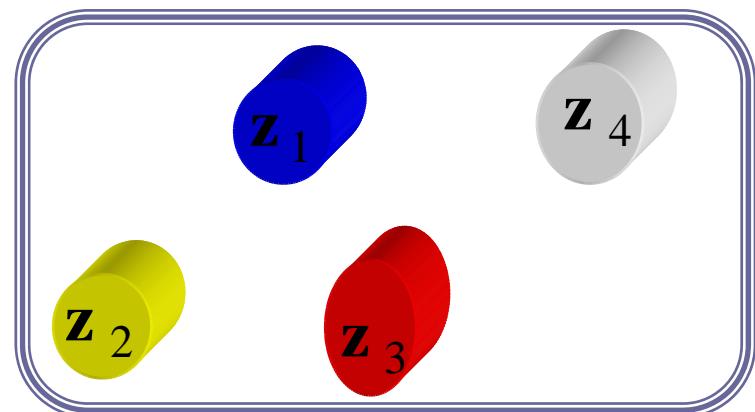
# Example



# Etape 1.

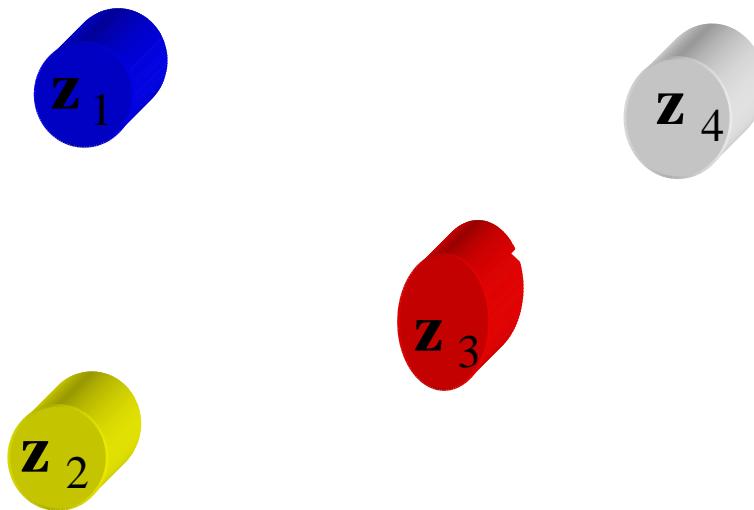
## Détermination des variables latentes

- Variable latente : combinaison linéaire standardisée par bloc.
- Calcul par un procédé itératif (intégration des spécifications des liens entre blocs par les procédures)



## Etape 2.

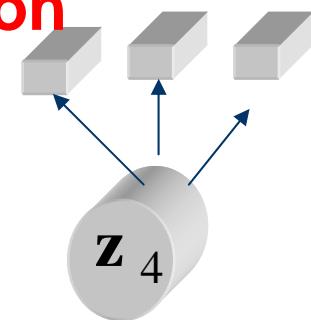
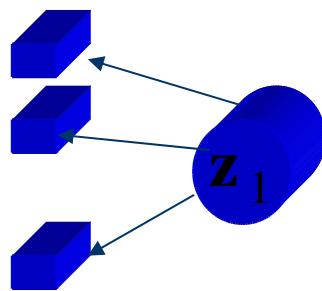
Estimation des coefficients de régression  
(Inner Model)



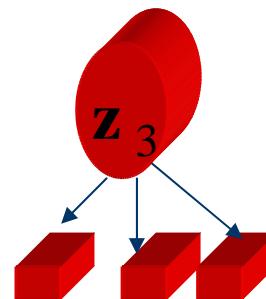
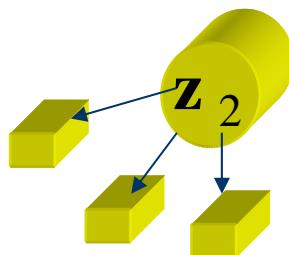
Inner Model(Structural model, Path model)

## Etape 3.

Estimation des coefficients de régression  
(Outer Model)



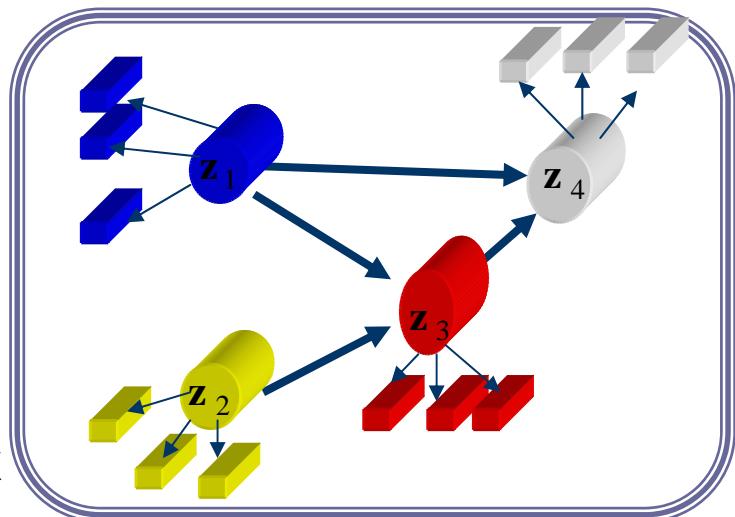
Outer Model ( Factor model, measurement model)



# Problématiques

Constat : rôle central des variables latentes dans l'approche PLS

- Problèmes :
  - Pas de critères qui caractérisent les variables latentes
  - Pas de propriétés associées aux procédures.



# Lohmöller's procedure

## A. Initialisation

A.1. Choose  $K$  arbitrary vectors:  $\tilde{\mathbf{w}}_1^{(0)}, \tilde{\mathbf{w}}_2^{(0)}, \dots, \tilde{\mathbf{w}}_K^{(0)}$

A.2. Compute the vector  $\mathbf{w}_1^{(0)}, \mathbf{w}_2^{(0)}, \dots, \mathbf{w}_K^{(0)}$  as:

$$\mathbf{w}_k^{(0)} = \sqrt{n} \frac{\tilde{\mathbf{w}}_k^{(0)}}{\|\mathbf{X}_k \tilde{\mathbf{w}}_k^{(0)}\|}, \quad 1 \leq k \leq K$$

A.3. Compute the latent variables  $\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \dots, \mathbf{z}_K^{(0)}$  as:

$$\mathbf{z}_k^{(0)} = \mathbf{X}_k \mathbf{w}_k^{(0)}, \quad 1 \leq k \leq K$$

## B. Internal estimation

B.1. Compute  $r_{kl}^{(s)} = r(\mathbf{X}_k \mathbf{w}_k^{(s)}, \mathbf{X}_l \mathbf{w}_l^{(s)})$  ( $1 \leq k, l \leq K$ )

B.2. Compute  $\theta_{kl}^{(s)}$  ( $1 \leq k, l \leq K$ ) as:

$$\theta_{kl}^{(s)} = \text{sign}(r_{kl}^{(s)}) \quad (\text{if centroid weighting scheme})$$

$$\theta_{kl}^{(s)} = r_{kl}^{(s)} \quad (\text{if factorial weighting scheme})$$

B.3. Compute  $\mathbf{Z}_k^{(s)} = \sum_{l \neq k}^K c_{kl} \theta_{kl}^{(s)} \mathbf{X}_l \mathbf{w}_l^{(s)}$  ( $1 \leq k \leq K$ )

## C. External estimation

C.1. Compute  $\tilde{\mathbf{w}}_k^{(s+1)} = (\mathbf{X}'_k \mathbf{X}_k)^{-1} \mathbf{X}'_k \mathbf{Z}_k^{(s)}$  ( $1 \leq k \leq K$ )

C.2. Compute  $\mathbf{w}_k^{(s+1)} = \sqrt{n} \frac{\tilde{\mathbf{w}}_k^{(s+1)}}{\|\mathbf{X}_k \tilde{\mathbf{w}}_k^{(s+1)}\|}$  ( $1 \leq k \leq K$ )

C.3. Compute  $\mathbf{z}_k^{(s+1)} = \mathbf{X}_k \mathbf{w}_k^{(s+1)}$  ( $1 \leq k \leq K$ )

Lohmöller JB (1989). Physica-Verlag, Heidelberg

Hanafi, M. (2007). Computational Statistics, 22, 275-292.

# Wold's procedure

## A. Initialisation

- A.1. Choose K arbitrary vectors:  $\tilde{\mathbf{w}}_1^{(0)}, \tilde{\mathbf{w}}_2^{(0)}, \dots, \tilde{\mathbf{w}}_K^{(0)}$
- A.2. Compute the vector  $\mathbf{w}_1^{(0)}, \mathbf{w}_2^{(0)}, \dots, \mathbf{w}_K^{(0)}$  as:

$$\mathbf{w}_k^{(0)} = \sqrt{n} \frac{\tilde{\mathbf{w}}_k^{(0)}}{\|\mathbf{X}_k \tilde{\mathbf{w}}_k^{(0)}\|}, \quad 1 \leq k \leq K$$

- A.3. Compute the latent variables  $\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \dots, \mathbf{z}_K^{(0)}$  as:

$$\mathbf{z}_k^{(0)} = \mathbf{X}_k \mathbf{w}_k^{(0)}, \quad 1 \leq k \leq K$$

## B. Internal estimation for the block $\mathbf{X}_{k_0}$

- B.1. Compute  $r_{k_0,l}^{(s)} = \begin{cases} r(\mathbf{X}_{k_0} \mathbf{w}_{k_0}^{(s)}, \mathbf{X}_l \mathbf{w}_l^{(s+1)}) & \text{if } l < k_0 \\ r(\mathbf{X}_{k_0} \mathbf{w}_{k_0}^{(s)}, \mathbf{X}_l \mathbf{w}_l^{(s)}) & \text{if } l > k_0 \end{cases}$

- B.2. Compute  $\theta_{k_0,l}^{(s)}$

$$\theta_{k_0,l}^{(s)} = \text{sign}(r_{k_0,l}^{(s)}) \quad (\text{if centroid weighting scheme})$$

$$\theta_{k_0,l}^{(s)} = r_{k_0,l}^{(s)} \quad (\text{if factorial weighting scheme})$$

- B.3. Compute  $\mathbf{Z}_{k_0}^{(s)} = \sum_{l=1}^{k_0-1} c_{k_0,l} \theta_{k_0,l}^{(s)} \mathbf{X}_l \mathbf{w}_l^{(s+1)} + \sum_{l=k_0+1}^K c_{k_0,l} \theta_{k_0,l}^{(s)} \mathbf{X}_l \mathbf{w}_l^{(s)}$

## C. External estimation for the block $\mathbf{X}_{k_0}$ .

- C.1. Compute  $\tilde{\mathbf{w}}_{k_0}^{(s+1)} = (\mathbf{X}'_{k_0} \mathbf{X}_{k_0})^{-1} \mathbf{X}'_{k_0} \mathbf{Z}_{k_0}^{(s)}$ ,

- C.2. Compute  $\mathbf{w}_{k_0}^{(s+1)} = \sqrt{n} \frac{\tilde{\mathbf{w}}_{k_0}^{(s+1)}}{\|\mathbf{X}_{k_0} \tilde{\mathbf{w}}_{k_0}^{(s+1)}\|}$

- C.3. Compute  $\mathbf{z}_{k_0}^{(s+1)} = \mathbf{X}_{k_0} \mathbf{w}_{k_0}^{(s+1)}$

Wold, H. (1985). Encyclopaedia of statistical sciences, vol 6. Wiley, New York, 581–591

Hanafi, M. (2007). Computational Statistics, 22, 275–292.

# Formes compactes des deux procédures

$$\begin{cases} \mathbf{z}_k = \mathbf{X}_k \mathbf{w}_k \quad (1 \leq k \leq K) \\ \|\mathbf{z}_k\| = 1 \end{cases} \Leftrightarrow \begin{cases} \mathbf{z}_k = \mathbf{P}_k \mathbf{u}_k \quad (1 \leq k \leq K) \\ \|\mathbf{u}_k\| = 1 \end{cases}$$

$$\mathbf{z}_k = \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-\frac{1}{2}} \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{\frac{1}{2}} \mathbf{w}_k = \underbrace{\mathbf{P}_k}_{\left( \mathbf{X}_k^T \mathbf{X}_k \right)^{\frac{1}{2}}} \underbrace{\mathbf{u}_k}_{\left( \mathbf{X}_k^T \mathbf{X}_k \right)^{\frac{1}{2}} \mathbf{w}_k} \quad (1 \leq k \leq K)$$

$$\|\mathbf{z}_k\| = 1 \Leftrightarrow \|\mathbf{u}_k\| = 1 \quad (1 \leq k \leq K)$$

- **Pk sont des matrices orthonormées ( par exemple, les composantes principales normées )**

# Formes compactes des deux procédures

Lohmöller

$$\mathbf{u}_k^{(s+1)} = \frac{\sum_{l=1}^K \mathbf{A}_{kl}(\mathbf{u}^{(s)}) \mathbf{u}_l^{(s)}}{\left\| \sum_{l=1}^K \mathbf{A}_{kl}(\mathbf{u}^{(s)}) \mathbf{u}_l^{(s)} \right\|} \quad (1 \leq k \leq K)$$
$$\mathbf{z}_k^{(s)} = \mathbf{P}_k \mathbf{u}_k^{(s)} \quad (1 \leq k \leq K)$$

Wold

$$\mathbf{u}_k^{(s+1)} = \frac{\sum_{l < k} \mathbf{A}_{kl}(\mathbf{u}^{(s+1)}) \mathbf{u}_l^{(s+1)} + \sum_{l \geq k} \mathbf{A}_{kl}(\mathbf{u}^{(s)}) \mathbf{u}_l^{(s)}}{\left\| \sum_{l < k} \mathbf{A}_{kl}(\mathbf{u}^{(s+1)}) \mathbf{u}_l^{(s+1)} + \sum_{l \geq k} \mathbf{A}_{kl}(\mathbf{u}^{(s)}) \mathbf{u}_l^{(s)} \right\|} \quad (1 \leq k \leq K)$$
$$\mathbf{z}_k^{(s)} = \mathbf{P}_k \mathbf{u}_k^{(s)} \quad (1 \leq k \leq K)$$

# Formes compactes des deux procédures

$$\mathbf{A}(\mathbf{u}) = \begin{bmatrix} 0 & \theta_{12}(\mathbf{u}_1, \mathbf{u}_2) \mathbf{P}_1^T \mathbf{P}_2 & \cdots & \theta_{1K}(\mathbf{u}_1, \mathbf{u}_K) \mathbf{P}_1^T \mathbf{P}_K \\ \theta_{21}(\mathbf{u}_2, \mathbf{u}_1) \mathbf{P}_2^T \mathbf{P}_1 & 0 & \ddots & \theta_{2K}(\mathbf{u}_2, \mathbf{u}_K) \mathbf{P}_2^T \mathbf{P}_K \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{K1}(\mathbf{u}_K, \mathbf{u}_1) \mathbf{P}_K^T \mathbf{P}_1 & \theta_{K2}(\mathbf{u}_K, \mathbf{u}_2) \mathbf{P}_K^T \mathbf{P}_2 & \cdots & 0 \end{bmatrix}$$

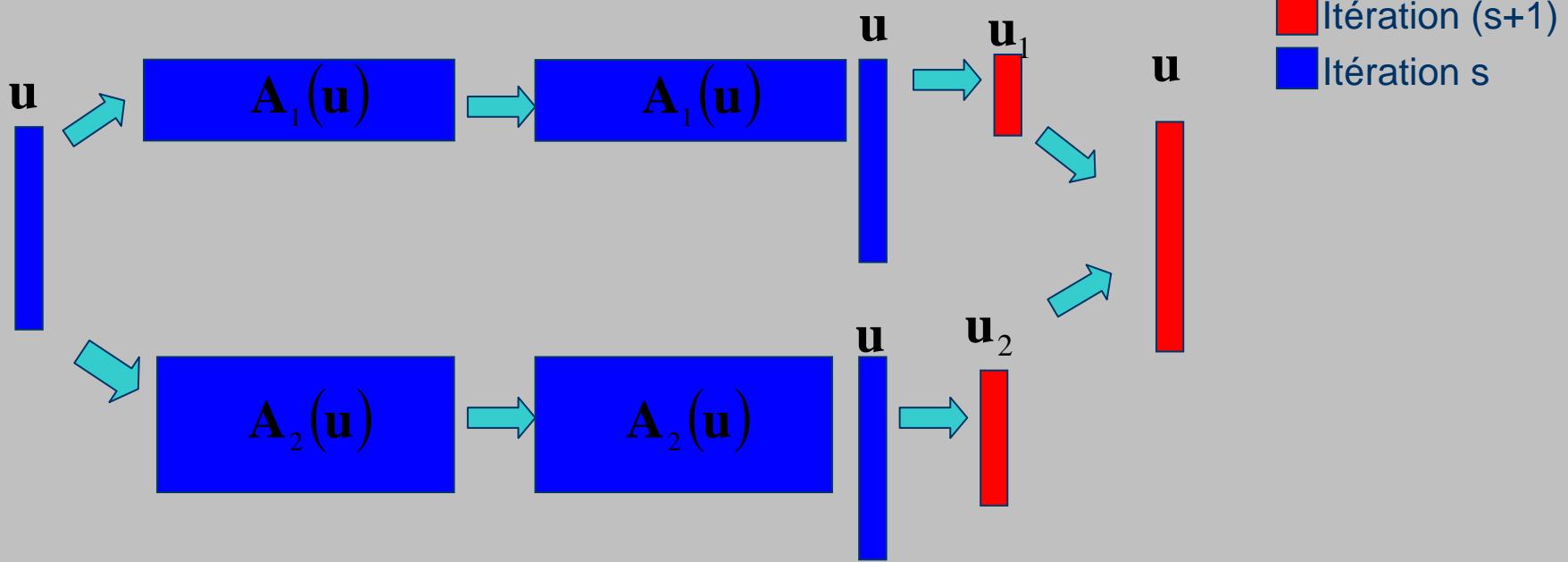
$$\mathbf{A}_{kl}(\mathbf{u}) \in \mathbf{M}_{p_k, p_l}(R)$$

$$\theta_{kl}(\mathbf{u}_k, \mathbf{u}_l) = \begin{cases} c_{kl} \text{sign}(\text{cor}(\mathbf{P}_k \mathbf{u}, \mathbf{P}_l \mathbf{u}_l)) & \text{Centroid} \\ c_{kl} \text{cor}(\mathbf{P}_k \mathbf{u}_k, \mathbf{P}_l \mathbf{u}_l) & \text{Factorial} \end{cases}$$

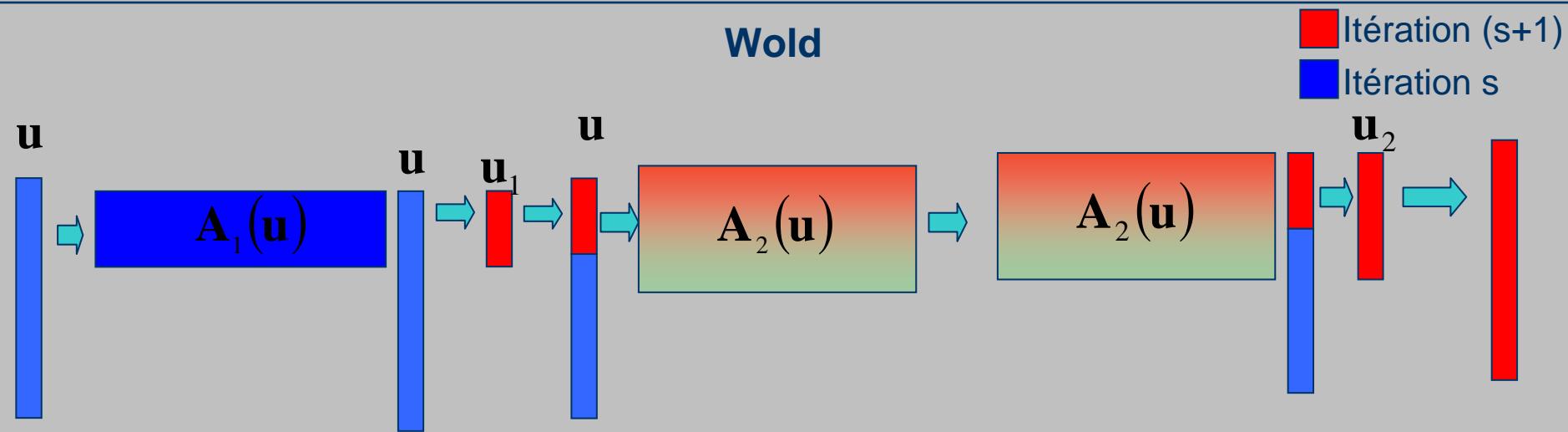
- une matrice partitionnée en (K,K) blocs
- Blocs diagonaux=[0]
- n'est pas définie positive.

$$\Rightarrow \underbrace{\mathbf{u}_{(p,1)}}_{(p,1)} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_K \end{bmatrix}, p = \sum_{k=1}^K p_k$$

## Lohmöller



## Wold



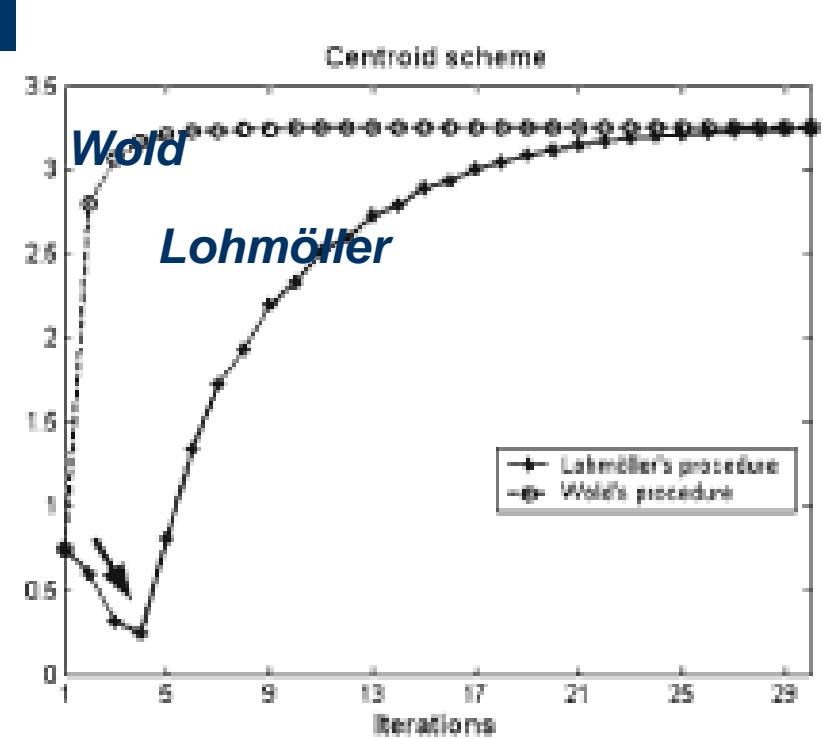
# MODE B+centroïde

- Wold → accroissement de  $f$
- N'est pas le cas de Lohmöller

ModeB + Centroid

$$\Rightarrow f(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K) = \sum_{i \neq j}^K c_{ij} |corr(\mathbf{z}_i, \mathbf{z}_j)|$$

- Résultats indépendants des données et des spécifications des liens entre blocs



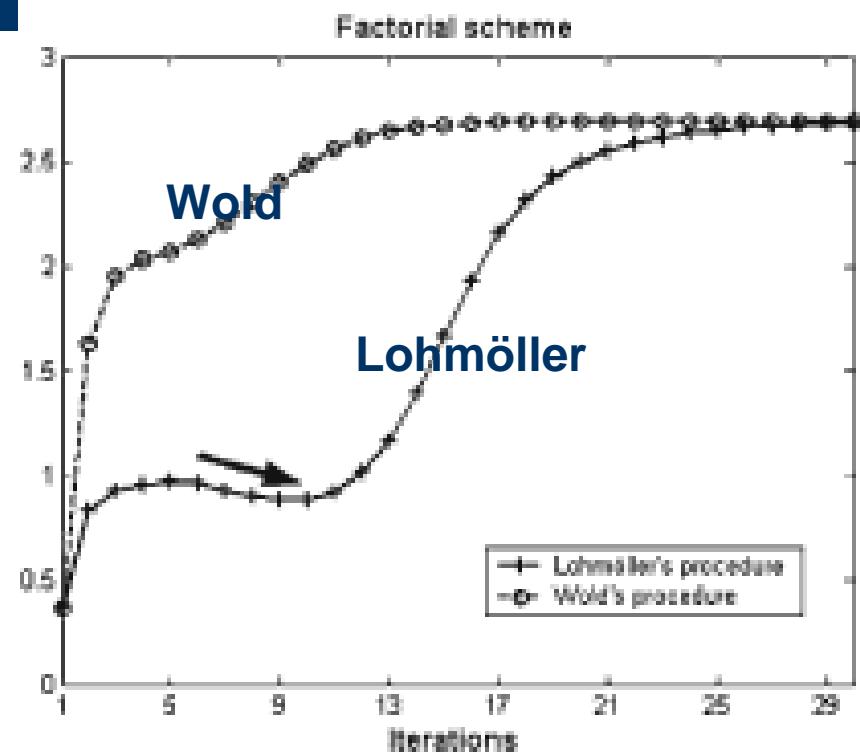
# MODE B+Factoriel

- **Wold → accroissement de  $f$**
- **N'est pas le cas de Lohmöller**

*ModeB + Factorial*

$$\Rightarrow g(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K) = \sum_{i \neq j}^K c_{ij} \text{corr}^2(\mathbf{z}_i, \mathbf{z}_j)$$

- **Résultats indépendants des données et des spécifications des liens**



# Liens avec les analyses canoniques généralisées.

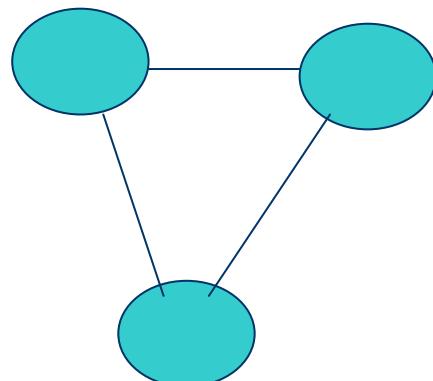
$$SUMCOR(1965) \Rightarrow \text{Max} \sum_{i \neq j, i, j=1}^m \text{corr}(\mathbf{z}_i, \mathbf{z}_j)$$

$$SSQCOR(1970) \Rightarrow \text{Max} \sum_{i \neq j, i, j=1}^m \text{corr}^2(\mathbf{z}_i, \mathbf{z}_j)$$

Analyse des correlations canoniques généralisées

$$ModeB + Centroid \Rightarrow \text{Max} \sum_{i \neq j, i, j=1}^m c_{ij} |\text{corr}(\mathbf{z}_i, \mathbf{z}_j)|$$

$$ModeB + Factorial \Rightarrow \text{Max} \sum_{i \neq j, i, j=1}^m c_{ij} \text{corr}^2(\mathbf{z}_i, \mathbf{z}_j)$$



# Convergence monotone

- $(\mathbf{z}_k^{(s)})_{s>0}$  ( $1 \leq k \leq K$ ) la suite de variables latentes générées par Wold

$$r^{(s)} = \begin{cases} \sum_{i \neq j, i,j=1}^K c_{ij} |corr(\mathbf{z}_i^{(s)}, \mathbf{z}_j^{(s)})| & ModeB + Centroid \\ \sum_{i \neq j, i,j=1}^K c_{ij} corr^2(\mathbf{z}_i^{(s)}, \mathbf{z}_j^{(s)}) & ModeB + Factorial \end{cases}$$

$$\Rightarrow r^{(s+1)} > r^{(s)} \Rightarrow (r^{(s)})_{s>0} \text{ converges}$$

**Cette propriété de monotony n'est pas valide pour Lohmöller**

# Question

Pourquoi la procédure de Lohmöller n'est pas monotone ?

*ModeB+Centroid*

$$\Rightarrow f(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K) = \sum_{i \neq j}^m c_{ij} |corr(\mathbf{z}_i, \mathbf{z}_j)|$$

*ModeB+Factorial*

$$\Rightarrow g(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K) = \sum_{i \neq j}^m c_{ij} corr^2(\mathbf{z}_i, \mathbf{z}_j)$$

,

•  $\left(\mathbf{z}_k^{(s)}\right)_{s>0} (1 \leq k \leq K)$  la suite de variables latentes générées par Lohmöller

$$\bullet \forall s \quad \mathbf{e}^{(s)T} [\mathbf{A}(\mathbf{u}^{(s)}) + \boldsymbol{\Gamma}^{(s)}] \mathbf{e}^{(s)} \leq \max(r^{(s+1)} - r^{(s)}, 0)$$

$\Gamma^{(s)}$

Lohmöller

$$\mathbf{u}_k^{(s+1)} = \frac{\sum_{l=1}^K \mathbf{A}_{kl} (\mathbf{u}^{(s)}) \mathbf{u}_k^{(s)}}{\left\| \sum_{l=1}^K \mathbf{A}_{kl} (\mathbf{u}^{(s)}) \mathbf{u}_k^{(s)} \right\|} \quad (1 \leq k \leq K)$$

$$\mathbf{z}_k^{(s+1)} = \mathbf{P}_k \mathbf{u}_k^{(s+1)} \quad (1 \leq k \leq K)$$

$$\Gamma^{(s)} = \begin{bmatrix} \lambda_1^{(s)} \mathbf{I}_{p_1} & 0 & \dots & 0 \\ 0 & \lambda_2^{(s)} \mathbf{I}_{p_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_K^{(s)} \mathbf{I}_{p_K} \end{bmatrix}$$

$$\lambda_k^{(s)} = \left\| \sum_{l=1}^K \mathbf{A}_{kl} (\mathbf{u}^{(s)}) \mathbf{u}_k^{(s)} \right\| \quad (1 \leq k \leq K)$$

• tous positifs.

$$\forall s \quad \mathbf{e}^{(s)T} \left[ \mathbf{A}(\mathbf{u}^{(s)})^+ \Gamma^{(s)} \right] \mathbf{e}^{(s)} \leq \max(r^{(s+1)} - r^{(s)}, 0)$$

Scenario 1.

$$\forall s \quad \mathbf{e}^{(s)T} \left[ \mathbf{A}(\mathbf{u}^{(s)})^+ \Gamma^{(s)} \right] \mathbf{e}^{(s)} > 0$$

↓

$$r^{(s+1)} > r^{(s)}$$

Scenario 2.

$$\exists s_0 \quad \mathbf{e}^{(s_0)T} \left[ \mathbf{A}(\mathbf{u}^{(s_0)})^+ \Gamma^{(s_0)} \right] \mathbf{e}^{(s_0)} \leq 0$$

↓

$$r^{(s_0+1)} \leq r^{(s_0)} \text{ or } r^{(s_0+1)} > r^{(s_0)}$$

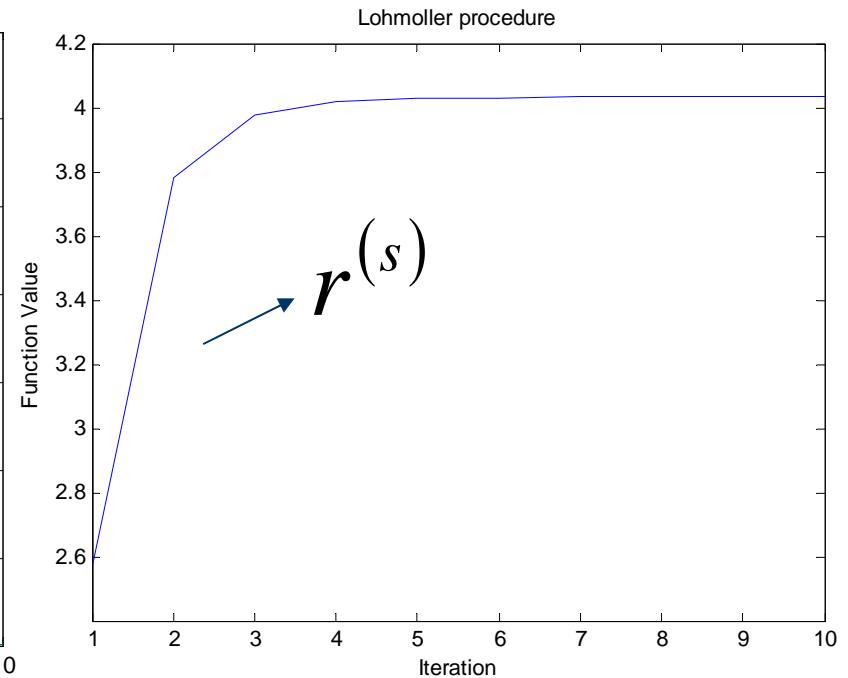
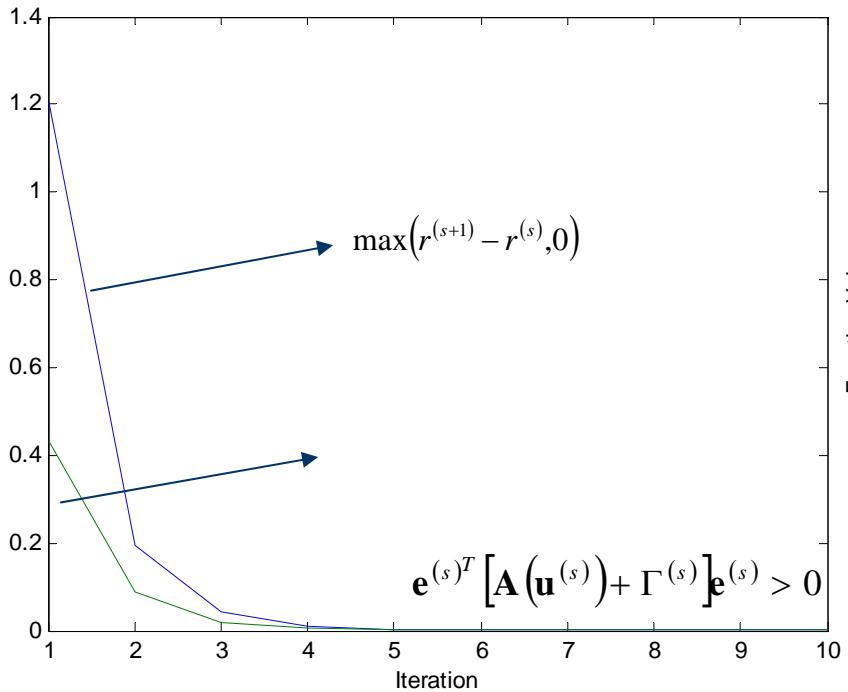
# Scénario 1.

valeurs propres de  $[A(u^{(s)}) + \Gamma^{(s)}]$

1	2	3	4	5	6	7	8	9	10
2.0631	2.6863	2.7354	2.7665	2.7651	2.7703	0.4767	2.7705	0.4769	2.7705
1.7529	2.0811	2.1777	2.1680	2.1755	2.1699	0.5126	2.1688	0.5125	2.1682
0.7987	1.2592	1.2639	1.3074	1.3029	1.3127	0.8231	1.3142	0.8219	1.3148
0.5952	0.7789	0.8301	0.8191	0.8263	0.8216	1.3116	0.8216	1.3140	0.8215
0.3204	0.4294	0.4719	0.4724	0.4761	0.4761	2.1706	0.4768	2.1688	0.4769
0.3752	0.4840	0.5095	0.5099	0.5127	0.5121	2.7693	0.5124	2.7702	0.5124

• la propriété de monotonie se réalise.

# Scénario 1.



$$\forall s \quad e^{(s)^T} [A(u^{(s)}) + \Gamma^{(s)}] e^{(s)} \leq \max(r^{(s+1)} - r^{(s)}, 0)$$

## Scénario 2 (A)

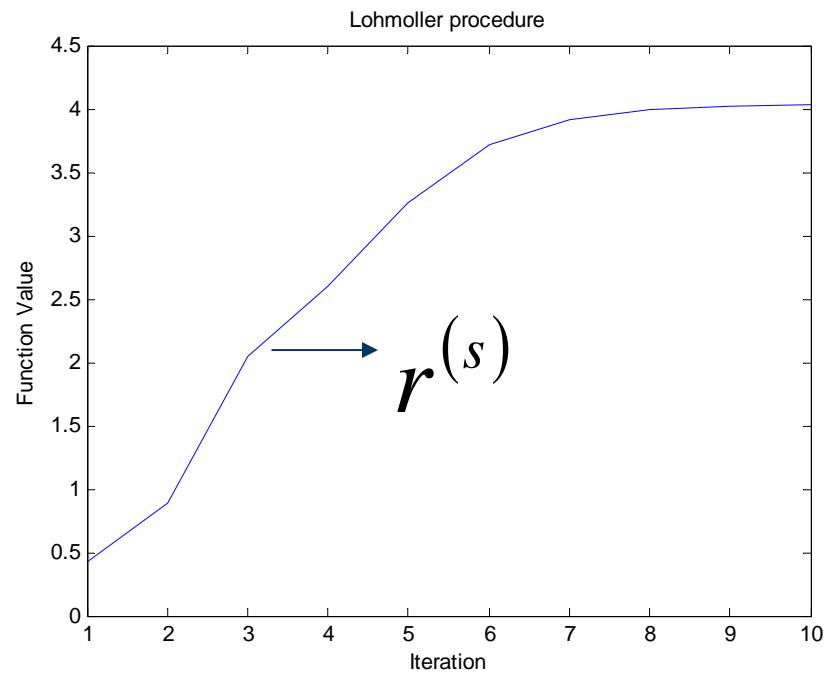
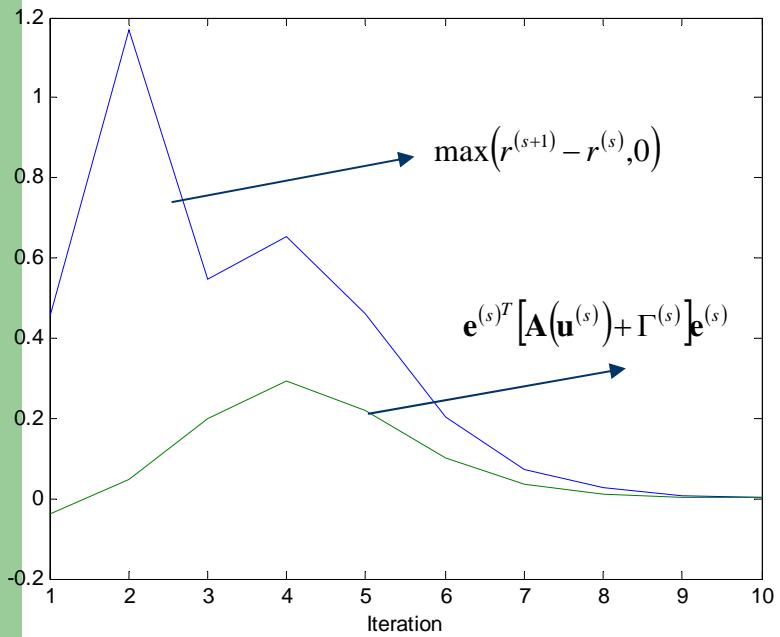
valeurs propres de  $[\mathbf{A}(\mathbf{u}^{(s)}) + \Gamma^{(s)}]$

1	2	3	4	5	6	7	8	9	10
0.7069	1.0163	-0.0999	0.1271	2.4186	2.6408	2.7314	2.7615	2.7693	2.7716
0.6167	0.8853	0.0958	0.2197	1.8496	2.0001	2.0847	2.1236	2.1445	2.1544
-0.1384	-0.1183	0.4797	0.5128	1.0852	1.2660	1.3242	1.3359	1.3313	1.3268
-0.0441	0.0197	0.5537	0.7692	0.6325	0.7202	0.7751	0.7969	0.8094	0.8145
0.2626	0.1875	1.4646	1.6715	0.2970	0.4199	0.4592	0.4736	0.4769	0.4777
0.1719	0.1420	1.7462	2.0674	0.3581	0.4488	0.4891	0.5038	0.5093	0.5111

Il existe des valeurs propres négatives mais la propriété de monotonie se réalise

# Scénario 2 (A)

$$\forall s \quad \mathbf{e}^{(s)T} [\mathbf{A}(\mathbf{u}^{(s)}) + \boldsymbol{\Gamma}^{(s)}] \mathbf{e}^{(s)} \leq \max(r^{(s+1)} - r^{(s)}, 0)$$



## Scénario 2 (B)

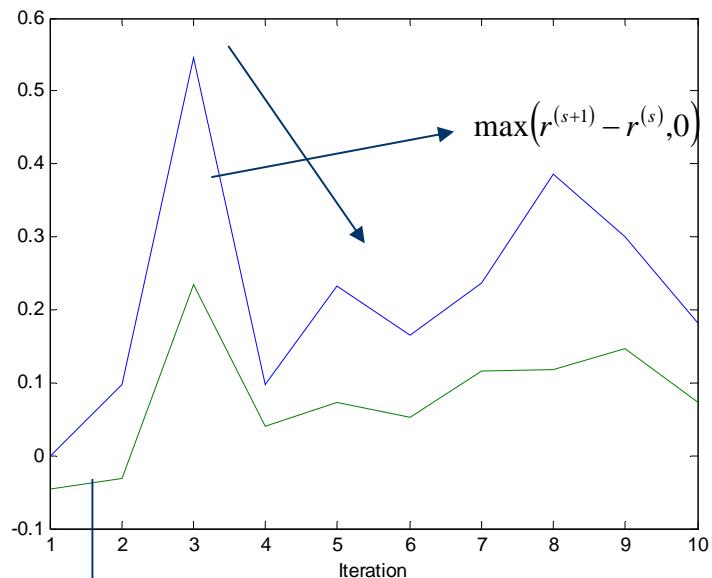
valeurs propres de  $[A(u^{(s)}) + \Gamma^{(s)}]$

1	2	3	4	5	6	7	8	9	10
1.1354	1.1085	1.1596	1.5833	1.6408	1.7897	-0.0792	1.9778	-0.0780	2.2127
0.9989	0.9692	1.0041	1.3767	1.4244	1.5729	0.0837	1.7760	0.1868	2.0300
0.2313	-0.0736	0.1426	0.1975	0.1913	0.3209	0.2480	0.5003	0.5182	-0.0458
-0.1203	0.1791	-0.1064	-0.0382	-0.0258	-0.0545	0.4630	-0.0764	0.6870	0.2356
0.0781	0.0147	-0.0201	0.0310	0.0047	0.0991	1.5979	0.2347	1.8499	0.7310
-0.0164	0.1017	0.0463	0.0160	0.0043	0.0333	1.7921	0.1060	2.0244	0.5525

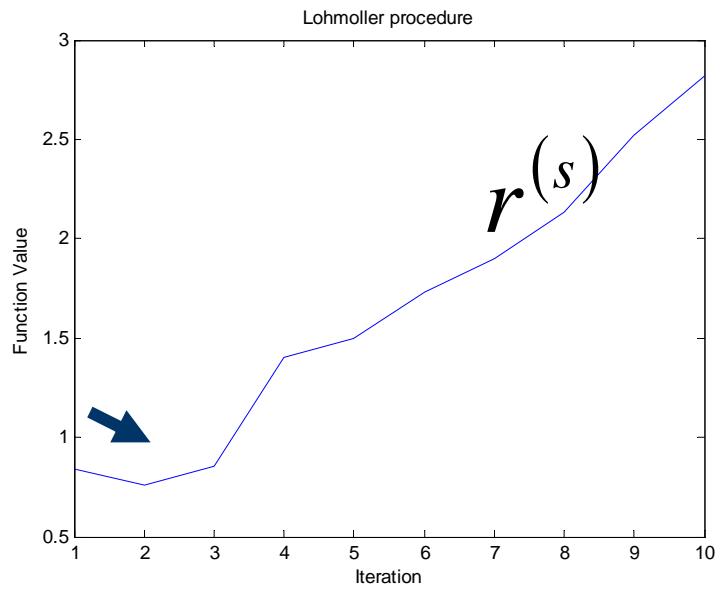
● Des valeurs propres sont négatives et la monotonie ne se réalise pas

# Scénario 2 (B)

$$\forall s \quad \mathbf{e}^{(s)T} [\mathbf{A}(\mathbf{u}^{(s)}) + \Gamma^{(s)}] \mathbf{e}^{(s)} \leq \max(r^{(s+1)} - r^{(s)}, 0)$$



$$\mathbf{e}^{(s)T} [\mathbf{A}(\mathbf{u}^{(s)}) + \Gamma^{(s)}] \mathbf{e}^{(s)} < 0$$



# Conclusions (1)

## Les contributions récentes permettent de :

- de clarifier l'étape de détermination de variables latentes
- de caractériser les variables latentes
- d'établir la convergence monotone de la procédure de Wold
- de positionner cette approche par rapport à aux généralisations des corrélations canoniques

(le mode B, PLS PM constitue une extension naturelle des généralisations de la corrélation canonique

- d'expliquer pourquoi la procédure de Lohmöller n'est pas monotone
- de proposer des procédures alternatives.

# Principale perspective algorithmique

Résultats avec le mode A ?

# Conclusions (2)

- Pour en savoir davantage

- *ESPOSITO VINZI, V.; CHIN, W.W.; HENSELER, J.; WANG, H. (EDS.) CONCEPTS, METHODS AND APPLICATIONS. SERIES: SPRINGER HANDBOOKS OF COMPUTATIONAL STATISTICS (2010)*

- Outils logiciels

- PLS Graph → <http://www.cba.uh.edu>
- SPAD → <http://www.spad.eu>
- SmartPLS → <http://www.smartpls.de>
- XLSTAT → <http://www.xlstat.com/fr/>

- Colloques dédiés :

- PLS and related methods.