

# REVUE DE STATISTIQUE APPLIQUÉE

M. VIVIEN

R. SABATIER

## **Une extension multi-tableaux de la régression PLS**

*Revue de statistique appliquée*, tome 49, n° 1 (2001), p. 31-54

[http://www.numdam.org/item?id=RSA\\_2001\\_\\_49\\_1\\_31\\_0](http://www.numdam.org/item?id=RSA_2001__49_1_31_0)

© Société française de statistique, 2001, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## UNE EXTENSION MULTI-TABLEAUX DE LA RÉGRESSION PLS

M. Vivien et R. Sabatier

Laboratoire de Physique Moléculaire et Structurale, UMR 9921, Faculté de Pharmacie,  
15 av. Ch. Flahault, 34060 Montpellier Cedex 2, France

### RÉSUMÉ

Nous présentons ici une extension de la méthode PLS, l'ACIMO-PLS, permettant de modéliser « globalement » un ensemble de  $K$  tableaux de données mesurés sur les mêmes  $n$  individus. L'optimisation du critère proposé revient à la résolution de PLS traditionnel, mais engendre des objets et aides à l'interprétation supplémentaires. Un exemple de données réelles sera traité.

**Mots-clés :** *régression PLS, analyse sensorielle, validation croisée, analyse multi-tableaux.*

### ABSTRACT

In this article, we are going to present an extension for the PLS method, OMCIA-PLS, which allows a « global » modelling of  $K$  data sets measured on the same  $n$  individuals. The optimization of the criterion that we propose amounts to resolving usual PLS, but gives extra objects and help to interpretation. An example based on real data will be processed.

**Keywords :** *PLS regression, sensor analysis, cross-validation, multiway analysis.*

### 1. Introduction

Tout au long de cet article, nous allons nous placer dans la situation dans laquelle on dispose de  $K$  groupes de variables, notés  $\{Y_k\}_{k=1,\dots,K}$ , que l'on désire expliquer « globalement » à l'aide d'un seul groupe de prédicteurs, noté  $X$ . Ces  $K + 1$  tableaux étant mesurés sur les mêmes  $n$  individus.

Peu de méthodes relatives à ce problème ont été développées. Wold *et coll.* (1987) ont proposé une méthode appelée « multiway PLS », MPLS, qui revient, en fait à faire la régression PLS « usuelle » d'un cube de variables ( $n \times p \times K$ ) réponses en le dépliant en une matrice de taille  $n \times pK$ , sur un cube de variables ( $n \times q \times K'$ ) déplié en matrice de taille  $n \times qK'$ .

Plus récemment, Bro (1996) a développé la PLS multi-linéaire, N-PLS, qui est basé sur la maximisation d'un critère de covariance généralisant celui de PLS « usuelle » au cas de cubes de données.

Citons aussi la méthode Wold *et coll.* (1996), la PLS hiérarchique, H-PLS, qui est destinée à expliquer  $K$  sous-groupes d'un ensemble de variables réponses par  $K'$  sous-groupes d'un ensemble de prédicteurs. Elle calcule des composantes de type PLS pour chaque sous-groupes de variables, et fait une régression PLS des composantes réponses sur les composantes explicatives, afin d'obtenir des composantes que nous dirons « globales ». Précisons que le calcul des composantes de chaque sous-groupe dépend de la composante « globale », ce qui montre la complexité de l'algorithme et nous amène à penser que cette technique ne maximise aucun critère.

Qannari, Courcoux & Maunit (1997) ont proposé une méthode inspirée de l'Analyse Procruste Généralisée (Gower, 1975).

La méthode que nous proposons, l'Analyse de Co-Inertie Multiple Orthogonale - Partial Least Squares, ACIMO-PLS, maximise un critère qui peut être vu comme une généralisation à plusieurs tableaux de la régression PLS « usuelle » de deux tableaux, telle que l'a définie Tenenhaus (1995). Ce critère est à mi-chemin de l'Analyse Canonique Généralisée (Carroll, 1968) et de STATIS (Lavit, 1988).

Il se trouve que la solution est un problème aux valeurs propres, ce qui présente l'avantage d'être facile à résoudre compte tenu de la stabilité des algorithmes existants.

De plus, les solutions de l'ACIMO-PLS peuvent s'obtenir par le rajout de quelques éléments et de nouveaux graphiques au programme de PLS à deux tableaux. Elle est donc facile à obtenir.

D'un point de vue pratique, cette méthode présente un intérêt certain en Analyse sensorielle, où les techniques traditionnelles ne permettent pas d'obtenir  $K$  modèles comparables. Ce qui est possible avec l'ACIMO-PLS : en effet, les  $K$  modèles que fournit l'ACIMO-PLS sont construits en fonction des mêmes composantes des variables explicatives, c'est-à-dire des mêmes combinaisons linéaires des variables explicatives (*cf.* parties 4 et 5).

Après avoir précisé nos principales notations, nous présentons l'aspect mathématique de l'ACIMO-PLS et explicitons ses points communs et ses avantages par rapport à une régression PLS à deux tableaux.

Enfin, nous l'appliquons à un problème d'Analyse Sensorielle afin de montrer « son intérêt ».

## 2. Notations utilisées

- $A'$  est la matrice transposée de la matrice  $A$ .
- $(Y_k, Q_k, D)$ ,  $k = 1, \dots, K$   $K$  triplets statistiques. Un tableau  $Y_k$ ,  $n \times p_k$ , représente un ensemble de  $p_k$  variables à expliquer (variables réponses) mesurées sur les mêmes  $n$  individus.  $Q_k$  est la métrique,  $p_k \times p_k$ , des individus dans  $\mathbb{R}^{p_k}$  et  $D$  la métrique, en général diagonale, des poids des individus dans l'espace des variables  $\mathbb{R}^n$ . En général, et sauf avis contraire,  $Y_k$  est centré en colonne pour la métrique  $D$ .

–  $W_k = Y_k Q_k Y_k'$  est la matrice,  $n \times n$ , des produits scalaires des individus du tableau  $Y_k$  au sens de la métrique  $Q_k$ .

–  $(Y, Q_Y, D)$  triplet statistique où  $Y$  est la matrice de taille  $n \times \sum_k p_k$  obtenue par juxtaposition des  $K$  tableaux  $Y_k : Y = [Y_1 | \dots | Y_K]$ .  $Q_Y$  est la métrique des individus dans  $\mathbb{R}^{\sum_k p_k}$ , donnée par la matrice diagonale par blocs dans laquelle le  $k^{i\text{ème}}$  bloc est la métrique  $Q_k$ .

–  $(X, Q, D)$  est le triplet statistique où  $X$ ,  $n \times q$ , représente  $q$  variables explicatives mesurées sur les mêmes  $n$  individus que les tableaux  $Y_k$ .  $Q$  est la métrique dans  $\mathbb{R}^q$ . En général,  $X$  est centré en colonne pour la métrique  $D$ .

$$- W = X Q X'.$$

–  $V_{kX} = V'_{Xk} = Y'_k D X$  est la matrice des covariances entre les variables de  $Y_k$  et celles de  $X$ .

–  $s = 1, \dots, S$  indice des étapes.  $S$  est le nombre d'axes retenus, qui est aussi le rang des modèles successifs emboîtés.

–  $b_{k,s}$  (*resp.*  $u_{k,s}$ ) et  $a_s$  (*resp.*  $t_s$ ) les axes (*resp.* composantes) de co-inertie d'ordre  $s$  associés respectivement aux tableaux  $Y_k$  et  $X$ , qui vont être définis ci-après.

### 3. Nouvelle approche pour PLS : ACIMO-PLS <sup>1</sup>

L'Analyse de Co-Inertie Multiple Orthogonale - Partial Least Squares (ACIMO-PLS), que nous allons développer, est une méthode linéaire itérative basée sur la maximisation sous contraintes d'un critère généralisant celui de PLS.

On va chercher successivement  $s$  ( $s = 1, \dots, S$ ) composantes (vecteurs de  $\mathbb{R}^n$  combinaisons linéaires des variables)  $u_{k,s}$  pour les tableaux  $Y_k$  et  $t_s$  pour le tableau  $X$  telles que :

- les composantes  $\{t_s\}_s$  forment un système  $D$ -orthogonal de  $\mathbb{R}^n$  et soient le plus explicatives des variables de  $X$ ,
- les composantes  $\{u_{k,s}\}_s$  soient aussi représentatives que possible de leur propre groupe de variables ( $Y_k$ ) tout en étant proches des composantes  $t_s$ .

Les composantes  $t_s$  seront ensuite utilisées pour modéliser les tableaux  $Y_k$  et permettre d'obtenir une suite de modèles emboîtés  $\hat{Y}_k^{(s)}$ . Ces modèles sont linéaires en les composantes de  $X$ , lesquelles tiennent compte implicitement de l'information apportée par les  $K$  groupes de variables réponses.

Le choix du rang  $S$  des modèles (nombre de composantes  $t_s$  de  $X$  entrant dans le modèle) se fait par validation croisée comme pour les méthodes PLS (Wold, 1966; Tenenhaus *et coll.*, 1995).

<sup>1</sup> Le nom ACIMO (*cf.* (Vivien, 1999)) provient d'une méthode plus générale, Analyse de Co-Inertie Multiple Orthogonale de  $K + 1$  tableaux dont l'ACIMO-PLS est une variante.

La Figure 1 résume ces propos.

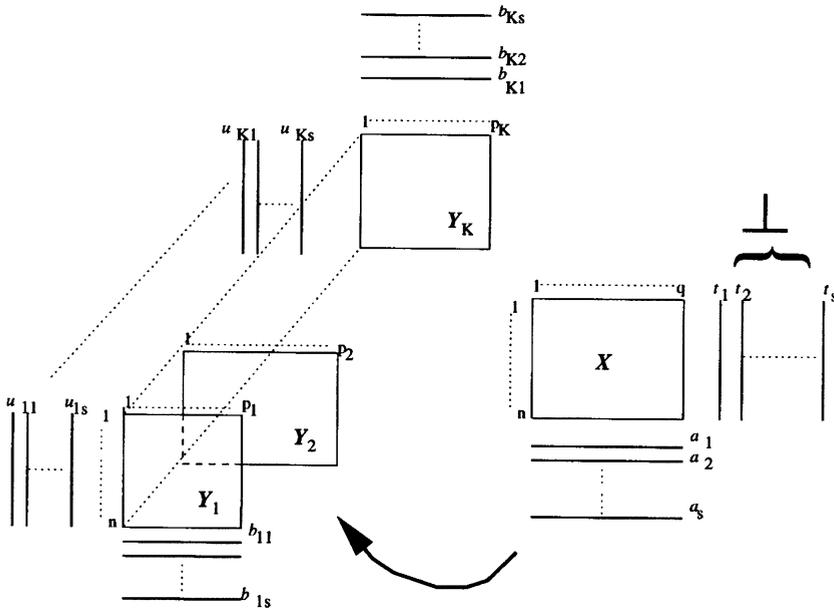


FIGURE 1  
L'ACIMO-PLS

### 3.1. Solutions d'ordre 1

À la première étape de la méthode, on va chercher  $K$  composantes  $u_{k,1} = Y_k Q_k b_{k,1}$  et la composante  $t_1 = X Q a_1$  de  $X$ .

On appelle ACIMO-PLS d'ordre 1, la maximisation de la fonction objectif suivante, somme des carrés des covariances entre  $u_k$  et  $t$ ,

$$f(b_1, \dots, b_K, a) = \sum_{k=1}^K cov^2(Y_k Q_k b_k, X Q a)$$

sous les contraintes usuelles de norme :  $\|b_k\|_{Q_k}^2 = 1, \forall k$  et  $\|a\|_Q^2 = 1$ .

Les solutions sont déterminées grâce à la méthode des multiplicateurs de Lagrange.

Le Lagrangien utilisé est

$$L(\{b_k\}, a, \lambda, \{\mu_k\}) = \sum_{k=1}^K (a' Q V_{Xk} Q_k b_k)^2 - \lambda (a' Q a - 1) - \sum_{k=1}^K \mu_k (b_k' Q_k b_k - 1),$$

où  $\lambda$  et  $\{\mu_k\}$  sont, respectivement, les multiplicateurs associés aux contraintes de norme pour les axes de co-inerties  $a$  et  $b_k$ .

On en tire les équations normales :

$$\frac{1}{2} \frac{\partial L}{\partial a} = \sum_{k=1}^K (a' Q V_{Xk} Q_k b_k) Q V_{Xk} Q_k b_k - \lambda Q a = 0 \quad (1)$$

$$\frac{1}{2} \frac{\partial L}{\partial b_k} = (a' Q V_{Xk} Q_k b_k) Q_k V_{kX} Q a - \mu_k Q_k b_k = 0 \quad k = 1, \dots, K \quad (2)$$

$$\frac{\partial L}{\partial \lambda} = -a' Q a + 1 = 0 \quad (3)$$

$$\frac{\partial L}{\partial \mu_k} = -b_k' Q_k b_k + 1 = 0 \quad k = 1, \dots, K. \quad (4)$$

En multipliant à gauche l'équation (1) par  $a'$  et les équations (2) à gauche par  $b_k'$ , on trouve :

$$\begin{aligned} \mu_k &= \text{cov}^2(X Q a, Y_k Q_k b_k) \quad k = 1, \dots, K \\ \lambda &= \sum_{k=1}^K \mu_k. \end{aligned} \quad (5)$$

Enfin, en utilisant ces résultats dans les équations (2) et (1), on obtient les relations :

$$b_k = \frac{1}{\sqrt{\mu_k}} V_{kX} Q a \quad k = 1, \dots, K \quad (6)$$

$$\sum_{k=1}^K V_{Xk} Q_k V_{kX} Q a = \lambda a, \quad (7)$$

$b_k$  étant de norme unité, on a  $\sqrt{\mu_k} = \|V_{kX} Q a\|_{Q_k}$ .

L'axe de co-inertie  $a_1$  recherché est donc le premier vecteur propre de la matrice  $(Q$ -symétrique, positive, donc diagonalisable et à éléments propres réels)

$\sum_{k=1}^K V_{Xk} Q_k V_{kX} Q$  associé à la première valeur propre  $\lambda_1$ . Les axes  $\{b_{k,1}\}_k$  sont ensuite déterminés par les formules de transition (6).

Les solutions obtenues sont bien le maximum de la fonction objectif, car les axes  $a$  et  $b_k$  sont caractérisés par  $\lambda$  (5) qui est égal au critère, et qui est choisi maximum.

On en déduit le même type de relations pour les composantes  $u_{k,1} = Y_k Q_k b_{k,1}$ ,  $k = 1, \dots, K$ , et  $t_1 = X Q a_1$  associées respectivement aux tableaux  $Y_k$ ,  $k = 1, \dots, K$ , et  $X$  :

$$WD \left( \sum_{k=1}^K W_k D \right) t_1 = \lambda t_1 \quad (8)$$

$$W_k D t_1 = \sqrt{\mu_{k,1}} u_{k,1}, \quad (9)$$

avec  $\mu_{k,1} = \|V_{kX} Q a_1\|_{Q_k}^2$ , valeur de  $\mu_k$  associée à  $a_1$ .

Donc, l'équation (8) montre que  $t_1$  est le premier vecteur propre, non normé à l'unité, de  $WD \sum_{k=1}^K W_k D$ , associé à sa plus grande valeur propre  $\lambda_1$ .

L'équation (9) montre que l'image de  $t_1$  (donc composante associée à  $X$ ) par l'opérateur associé à  $Y_k$  est proportionnel à la composante  $u_{k,1}$ .

### 3.2. Solutions d'ordre $s$ ( $s > 1$ )

On définit les solutions d'ordre  $s$ ,  $s > 1$ , de l'ACIMO-PLS, par les solutions d'ordre 1 de l'ACIMO-PLS des triplets  $(Y_1^{(s-1)}, Q_1, D), \dots, (Y_K^{(s-1)}, Q_K, D), (X^{(s-1)}, Q, D)$ , où  $X^{(0)} = X$ ,  $Y_k^{(0)} = Y_k$  et  $X^{(s-1)}$ , (*resp.*  $Y_k^{(s-1)}$ ) est le résidu de la projection  $D$ -orthogonale de  $X$  (*resp.* de  $Y_k$ ) sur l'espace  $T^{(s-1)}$  engendré par les vecteurs  $t_1, \dots, t_{s-1}$ .

C'est-à-dire, que l'axe  $a_s$  est le premier (*i.e.* associé à la plus grande valeur propre  $\lambda_s$ ) vecteur propre du problème suivant :

$$\sum_{k=1}^K V_{Xk}^{(s-1)} Q_k V_{kX}^{(s-1)} Q a_s = \lambda_s a_s,$$

où  $V_{Xk}^{(s-1)} = V_{kX}^{(s-1)'}$  désigne  $X^{(s-1)'} D Y_k^{(s-1)}$ .

Les axes  $b_{k,s}$  sont alors obtenus par  $b_{k,s} = \frac{1}{\sqrt{\mu_{k,s}}} Y_k^{(s-1)'} D X Q a_s$ , avec  $\mu_{k,s} = \|V_{kX}^{(s-1)} Q a_s\|_{Q_k}^2$ .

Ainsi, par définition, on obtient à l'étape  $s$  un système de composantes  $\{t_l = X^{(l-1)} Q a_l\}_{l=1, \dots, s}$   $D$ -orthogonal dans  $\mathbb{R}^n$ .

De plus, on a les propriétés suivantes analogues à celles de PLS ordinaire :

- $(a_s | a_{s'})_D = \delta_{ss'}$ ,  $s, s' = 1, \dots, S$ .

- $(u_{k,s}|t_l)_D = 0, \quad l = 1, \dots, s-1, \text{ avec } u_{k,s} = Y_k^{(s-1)} Q_k b_{k,s},$   
 $k = 1, \dots, K.$
- $(b_{k,l}|b_{k,s})_{Q_k} = (u_{k,l}|t_s)_D \text{cov}^2(t_s, u_{k,s}), \quad l = 1, \dots, s-1, \quad k = 1, \dots, K.$

Les démonstrations sont fournies dans (Vivien, 1999).

### 3.3. Lien avec PLS « usuelle »

#### 3.3.1. Rappels succincts sur la régression PLS

Rappelons en quoi consiste la méthode de régression PLS d'un triplet  $(Y, Q_Y, D)$  sur  $(X, Q_X, D)$ , les deux triplets étant mesurés sur les mêmes  $n$  individus. L'introduction des métriques dans les espaces des individus est due à Cazes (1997). C'est une simple généralisation de la régression PLS telle que l'ont défini Wold (1966) et Tenenhaus *et coll.* (1995) au cas où les métriques  $Q_X$  et  $Q_Y$  ne sont pas égales à l'identité.

La régression PLS est une méthode itérative basée sur la maximisation de la fonction objectif  $g(w, c) = \text{cov}(Y Q_Y c, X Q_X w)$  sous les contraintes de norme sur les axes  $\|c\|_{Q_Y}^2 = \|w\|_{Q_X}^2 = 1$ .

À chaque étape  $s$ , on maximise cette fonction  $g$  en remplaçant respectivement les tableaux «  $Y$  » et «  $X$  » par les résidus  $(Y^{(s-1)})$  et  $(X^{(s-1)})$  des projections  $D$ -orthogonales de  $Y$  et  $X$  sur l'espace engendré par les composantes PLS de  $X$  obtenues aux étapes précédentes.

L'axe  $w_s$ , d'ordre  $s$ , associé à  $X$  est le vecteur propre associé à la plus grande valeur propre  $\nu^2$  de  $X^{(s-1)'} D Y^{(s-1)} Q_Y Y^{(s-1)'} D X^{(s-1)} Q_X$ .  $\nu$  est égale au maximum du critère  $g$ .

L'axe d'ordre  $s$ ,  $c_s$ , associé à  $Y$  est alors obtenu par la formule de transition  $c_s = \frac{1}{\nu} Y^{(s-1)'} D X^{(s-1)} Q_X w_s$ .

Le système des  $S$  composantes  $\{X Q_X w_s\}_{s=1, \dots, S}$  est alors  $D$ -orthogonal, ce qui permet d'écrire des modèles linéaires d'ordre  $S$  pour chacune des variables à expliquer de  $Y$ .

#### 3.3.2. Relation ACIMO-PLS - Régression PLS « usuelle »

Commençons par préciser quelques notations supplémentaires. On désignera par :

- $b_s$ , vecteur de dimension  $\sum_k p_k \times 1$ , le  $s^{\text{ème}}$  axe de  $Y$  issu de la régression PLS du triplet  $(Y, Q_Y, D)$  sur  $(X, Q, D)$ ,
- $u_s = Y^{(s-1)} Q_Y b_s$ , vecteur de  $\mathbb{R}^n$ , la  $s^{\text{ème}}$  composante PLS de  $Y$ ,
- $b_s[k]$  le  $k^{\text{ième}}$  bloc de longueur  $p_k$  du vecteur  $b_s$ .

Nous avons les correspondances suivantes :

**Propriété.** — *Les axes de co-inerties  $a_s$  (resp. les composantes  $t_s$ ) de l'ACIMO-PLS associés à  $X$  sont identiques aux axes d'ordres  $s$  (resp. aux composantes d'ordre  $s$ ) de  $X$  issus de la régression PLS de  $(Y, Q_Y, D)$  sur  $(X, Q, D)$ .*

*De plus, on a une relation entre les axes de co-inertie des tableaux  $Y_k$  et l'axe PLS,  $b_s$ , associé à  $Y$  :  $b_{k,s} = b_s[k] / \|b_s[k]\|_{Q_k}$ .*

*Enfin, la composante PLS,  $u_s$ , associée à  $Y$  est une somme pondérée des composantes de co-inertie associées aux tableaux  $Y_k$  :  $u_s = \sum_k \|b_s[k]\|_{Q_k} u_{k,s}$ .*

En effet, les axes associés à  $X$  issus des deux méthodes vérifient le même problème aux valeurs propres. Ce qui implique l'égalité respective des axes et des composantes d'ordre 1. Ensuite, les procédures de calculs des objets d'ordre supérieur étant identiques, il est facile de montrer que les axes (resp. les composantes) PLS et les axes (resp. les composantes) de co-inertie associés à  $X$  correspondent à tout ordre.

La propriété relative aux axes associés à  $Y_k$  est en partie due à la formule de transition entre  $b_s$  et  $a_s$  :

$$b_s = \frac{1}{\nu_s} Y^{(s-1)'} D X^{(s-1)} Q a_s \quad \text{avec} \quad \nu_s = \sqrt{\lambda_s}.$$

*Attention : dans ce qui suit et jusqu'au paragraphe 3.4, afin d'alléger les notations, nous noterons  $Y_k$  et  $X$ , au lieu de  $Y_k^{(s-1)}$  et  $X^{(s-1)}$ , les tableaux à l'étape  $s$ .*

À partir de cette relation, on peut exprimer les  $K$  blocs de longueur  $p_k$  du vecteur  $b_s$  par  $b_s[k] = \frac{1}{\nu_s} Y_k' D X Q a_s$ , et finalement  $b_s[k] / \|b_s[k]\|_{Q_k} = Y_k' D X Q a_s / \|Y_k' D X Q a_s\|_{Q_k}$ .

Or, nous avons montré (cf. paragraphe 3.1) que  $\|Y_k' D X Q a_s\|_{Q_k} = \sqrt{\mu_{k,s}}$ . D'où,

$$b_{k,s} = b_s[k] / \|b_s[k]\|_{Q_k}.$$

Il est facile de montrer, de plus, que  $\|b_s[k]\|_{Q_k} = \sqrt{\mu_{k,s}} / \sqrt{\lambda_s}$ .

Passons maintenant à la propriété sur les composantes. Par définition, nous avons  $u_s = Y Q_Y b_s$ , ce qui peut s'écrire aussi de la forme  $u_s = \sum_k Y_k Q_k b_s[k]$ .

D'après l'égalité obtenue précédemment, on déduit le résultat.

Cette expression de  $u_s$  et l'égalité  $\|b_s[k]\|_{Q_k} = \sqrt{\mu_{k,s}} / \sqrt{\lambda_s}$  montrent que plus la covariance (qui est positive) entre  $u_{k,s}$  et  $t_s$  est faible, moins la composante  $u_{k,s}$  entre en jeu dans le calcul de  $u_s$ . Autrement dit, moins la composante  $t_s$  est explicative (proche) de  $u_{k,s}$ , moins  $u_s$  en tient compte. On peut dire que  $u_s$  néglige de l'information au bénéfice de l'information la plus importante.

### 3.4. Les modèles $\{\widehat{Y}_k^S\}_{k=1,\dots,K}$

La modélisation de rang  $S$  des tableaux  $Y_k$  par  $X$  est immédiate en appliquant les résultats de la régression PLS, car  $\widehat{Y}^S = [\widehat{Y}_1^S | \dots | \widehat{Y}_K^S]$ .

Nous rappelons ci-après l'expression de ces modèles linéaires ( $\{\widehat{Y}_k^S\}_{k=1,\dots,K}$ ) et une propriété, généralisant celle de PLS « usuelle » (Tenenhaus, 1995), permettant de les obtenir.

**Propriété** — *Comme en PLS de deux tableaux, les composantes  $t_s = X^{(s-1)}Qa_s$  peuvent s'écrire sous la forme  $t_s = X\omega_s^*$  pour tout  $s = 1, \dots, S$  et  $\omega_s^*$  est donné par :*

$$\begin{aligned}\omega_1^* &= Qa_1 \\ \omega_s^* &= Qa_s - \sum_{l=1}^{s-1} \frac{\omega_l^* \omega_l^{*'}}{\|t_l\|_D^2} X' D X Q a_s \quad s \geq 2.\end{aligned}$$

**Propriété.** — *Pour chacun des tableaux  $Y_k$ , on a le modèle linéaire suivant :*

$$\begin{aligned}Y_k &= \sum_{s=1}^S P_{t_s} Y_k + Y_k^{(S)} \doteq \widehat{Y}_k^S + Y_k^{(S)} \\ \widehat{Y}_k^S &= X \widehat{\beta}_k^S\end{aligned}$$

avec

$$\widehat{\beta}_k^S = \sum_{s=1}^S \frac{\omega_s^* \omega_s^{*'}}{\|t_s\|_D^2} X' D Y_k.$$

Les démonstrations sont disponibles dans (Vivien, 1999).

### 3.5. Commentaires et liens avec d'autres méthodes

– L'ACIMO-PLS met en évidence une certaine optimalité de la régression PLS de  $Y$  sur  $X$ .

En effet, l'ACIMO-PLS montre que les modèles obtenus pour les  $\sum_k p_k$  variables réponses par la régression PLS de  $(Y, Q_Y, D)$  sur  $(X, Q, D)$  sont ajustés par tableau. Les composantes associées aux prédicteurs sont les plus explicatives et les plus corrélées avec les composantes les plus explicatives  $(u_{k,s})$  des variables réponses.

– En plus, l'ACIMO-PLS fournit des composantes par tableaux. On peut alors représenter sur un même graphique les points de coordonnées  $(u_s, t_s)$ ,  $\{(u_{k,s}, t_s)\}_k$ , ce qui donne une vision globale et/ou partielle (par tableau) de la qualité des

régressions partielles de  $Y$  et  $Y_k$  sur  $t_s$  à l'étape  $s$ . Si le nuage de points est proche d'une droite, la relation entre les variables de  $Y$  et celles de  $X$  est bien linéaire et justifie l'emploi de cette méthode. L'ACIMO-PLS permet de faire une étude plus approfondie quant à la validité des modèles pour chacun des groupes de variables réponses.

– Comme les composantes PLS associées aux variables réponses, les composantes  $\{u_{k,s}\}_s$  ne forment pas un système  $D$ -orthogonal dans  $\mathbb{R}^n$ . Dans notre problème, qui est en fait une étude sur la modélisation, c'est l'orthogonalité des composantes  $\{t_s\}_s$  qui est importante ainsi que la relation entre les composantes  $u_k$  et  $u$  : c'est cela qui permet l'obtention des modèles linéaires. Cependant, on peut choisir des orthogonalités supplémentaires, par exemple sur les composantes  $\{u_{k,s}\}_s$ . Nous avons appelé cette méthode, dont le but est différent, Analyse de Co-Inertie Multiple Orthogonale (Vivien, 1999).

– Le rang  $S$  des modèles ne peut dépasser le rang de  $X$ .

– Notons que le critère utilisé,  $\sum_k cov^2(t, u)$ , peut aussi s'écrire  $\sum_k r^2(t, u) var(t)var(u)$ . Ainsi, en le maximisant, on maximise à la fois la corrélation entre les composantes et leur variance respective. De plus, celui-ci est une généralisation « naturelle » du critère de la régression PLS « usuelle ». En effet, maximiser  $cov(YQ_Yc, XQ_Xw)$  ou  $cov^2(YQ_Yc, XQ_Xw)$  sous les contraintes de norme unité pour  $w$  et  $c$ , donne un résultat identique.

– Si  $k = 1$ , l'ACIMO-PLS correspond à la régression PLS de  $Y_1$  sur  $X$ .

– Lorsqu'on utilise les métriques usuelles dans les espaces des individus, c'est-à-dire les métriques identités, les modèles obtenus sont ceux de Wold *et coll.* (1987) avec la multiway PLS.

– Les solutions d'ordre 1 de l'ACIMO-PLS correspondent aux solutions d'ordre 1 de l'Analyse de la Concordance définie par Lafosse et Hanafi (1997).

– Les  $K$  composantes de co-inertie  $u_{k,1}$  de l'ACIMO-PLS peuvent se retrouver, à la norme près, avec l'approche PLS définie par Wold (1985) et reprise par Tenenhaus (1999). En suivant le réseau de causalité décrit Figure 2 faisant intervenir  $K + 1$  variables latentes (1 exogène et  $K$  endogènes) et  $K + 1$  blocs réfléchifs, et en utilisant le mode A et le schéma centroïde (ou factoriel), les estimations externes (notées «  $Y_j$  » par Tenenhaus (1999)) des  $K$  variables latentes endogènes correspondent, à la norme près, aux composantes de co-inertie  $u_{k,1}$  de l'ACIMO-PLS.

Si, lors de la maximisation du critère, on avait imposé aux composantes  $t$  et  $u_k$  d'être de norme unité au lieu de l'imposer aux axes de co-inertie, on aurait bien évidemment retrouvé les résultats de l'approche PLS. Mais, dans ce cas, le critère à maximiser serait une somme de corrélations au carré, ce qui ne serait plus du tout le même problème.

– Avec un choix particulier de  $X$ ,  $Q_k$  ou  $Q$ , les solutions d'ordre 1 de l'ACIMO-PLS correspondent aux solutions d'ordre 1 de l'Analyse de CO-inertie Multiple, ACOM, (Chessel et Hanafi, 1996), de Analyse Canonique Généralisée de Carroll,

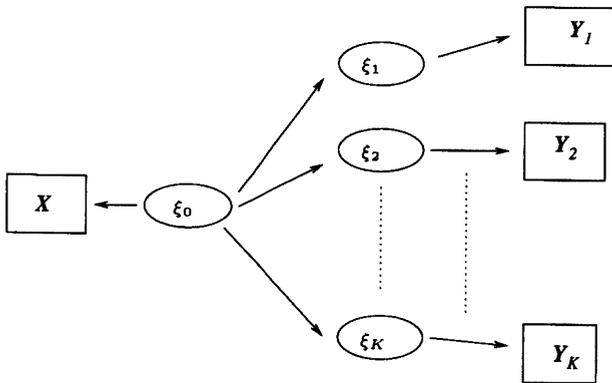


FIGURE 2

*Réseau de causalité pour l'approche PLS*

ACG, (Carroll, 1968) ou encore de l'Analyse en Composantes Principales sur Variables Instrumentales, ACPVI, (Rao, 1964). Ces résultats étant plus théoriques que pratiques, et finalement assez techniques, nous n'en donnons pas ici les démonstrations. On peut les trouver dans (Vivien, 1999).

– D'un point de vue pratique, il suffit d'utiliser une régression PLS « usuelle » en y ajoutant quelques calculs.

– Enfin, toujours d'un point de vue pratique, on peut comparer la variance d'une composante de co-inertie  $u_{k,s}$  (resp.  $t_s$ ) avec la variance de la composante d'inertie (issue de l'ACP) de même ordre issue de l'Analyse en Composantes Principales (ACP) du triplet  $(Y_k, Q_k, D)$  (resp. du triplet  $(X, Q, D)$ ). Plus elles sont proches, meilleure est l'explication du tableau  $Y_k$  (resp.  $X$ ) par l'axe de co-inertie correspondant.

Les graphiques appelés «  $w^*c$  » par Tenenhaus (1998) sont bien sûr faisables ici, ainsi que toutes les aides à l'interprétations que l'on utilise en PLS. Mais nous ne les ferons pas dans l'exemple traité par la suite, pour une économie de place. De plus, il semble naturel de comparer la variance des composantes  $u_{k,s}$  à celle des composantes de même ordre issues de la régression PLS « usuelle » de  $Y_k$  sur  $X$ , ainsi qu'à la variance des composantes de même ordre de l'ACPVI de  $(Y_k, Q_k, D)$  par rapport à  $X$ . Ces dernières analyses visant à expliquer indépendamment chaque tableau  $Y_k$  par  $X$ , les composantes associées ont toutes les raisons d'être plus explicatives de chaque tableau que les composantes de l'ACIMO-PLS, si bien sûr,  $X$  explique un tant soit peu le tableau  $Y_k$ . Les composantes de l'ACIMO-PLS expliquent chacun des tableaux en tenant compte de l'information apportée par les autres variables réponses, d'où l'intérêt en Analyse Sensorielle (ce que nous expliciterons aux paragraphes 4 et 5), et dans ce cadre, cette méthode permet la comparaison des différents modèles.

#### 4. L'Analyse sensorielle et l'ACIMO-PLS

En Analyse Sensorielle, on peut être confronté à deux types de situations : soit modéliser les appréciations hédoniques des consommateurs quant aux produits étudiés à l'aide des réponses de  $K$  juges (tableaux des propriétés sensorielles) (Robin, 1996), soit expliquer les propriétés sensorielles des produits en fonction des caractéristiques physico-chimiques de ces mêmes produits, afin d'étudier leur influence sur les réponses données par les juges (Qannari, Courcoux & Maunit, 1997).

C'est dans cette dernière situation que nous allons nous placer. Les méthodes traditionnellement utilisées pour résoudre ce problème, sont basées, tout d'abord, sur la recherche d'un tableau « moyen », résumant les propriétés sensorielles à l'aide de la méthode STATIS (Lavit, 1988; Schlich, 1993) ou de méthodes Procrustes Généralisées (Gower, 1975; Qannari *et coll.*, 1997) puis sur l'utilisation de méthodes d'analyses de deux tableaux.

La méthode la plus immédiate, qui consiste à modéliser le tableau conjoint  $Y = [Y_1 | \dots | Y_K]$  ( $K$  tableaux des propriétés sensorielles) par le tableau  $X$  (tableau des propriétés physico-chimiques des produits sujets à l'étude) grâce à la régression PLS, semble ignorée. Or, celle-ci possède des propriétés d'optimalités qui la rendent très proche de l'ACIMO-PLS, comme nous avons pu le voir dans le paragraphe précédent.

L'ACIMO-PLS permet d'étudier le tableau  $X$  et d'explicitier son influence sur les études sensorielles de manière générale.

On peut donner un modèle  $\widehat{Y}_k^S$  pour chacun des tableaux tenant compte implicitement de tous les tableaux. Ces modèles présentent un intérêt : ils tiennent compte des réponses de tous les juges, et ont tendance à « améliorer » un tableau ayant des réponses éloignées des autres juges. De plus, les modèles pour les  $K$  tableaux sont comparables, car ils sont expliqués par les mêmes composantes  $t$  et sont de même rang. Alors qu'en faisant  $K$  régressions PLS séparées (par exemple), on n'a *a priori* pas de raison d'obtenir à chaque fois les mêmes composantes pour  $X$ , ni des modèles de même dimension.

Ici, si les réponses des juges sont cohérentes, les modèles devraient être « presque » proportionnels.

Un des buts final est de prédire les réponses des juges. Pour cela, on peut, comme dans PLS, utiliser un critère du type validation croisée pour valider de façon « interne » un modèle.

#### 5. Application à l'Analyse Sensorielle

##### 5.1. Présentation des données utilisées

Les données (voir l'Annexe ) que nous allons étudier (brièvement) sont celles utilisées par Gonzalez (1982) dans sa thèse.

Elles ont été recueillies à l'ENSBANA (Dijon) auprès de 7 juges qui seront ici les tableaux  $Y_k (K = 7)$ .

Il s'agit d'expliquer les propriétés sensorielles de cafés boissons à l'aide de propriétés physico-chimiques des cafés, ces dernières étant associées au tableau  $X$ .

**Dix-huit** échantillons de café ont été préparés suivant trois paramètres, *température*, *proportion mouture-eau*, *granulométrie*, ayant respectivement trois, trois et deux modalités.

**Quatre** de ces échantillons, ayant des propriétés physico-chimiques relativement proches, ont été utilisés comme témoins lors de la comparaison par paires (décrite ci-après) avec **treize** des échantillons restants. Le dernier échantillon n'a pas été utilisé.

Chaque juge a comparé trois fois chacun des treize échantillons avec un témoin et a répondu à **six** questions : *lequel des deux est le plus caractéristique en parfum, le plus intense en parfum, le plus amer, le plus acide, le plus caractéristique en arôme et le plus intense en arôme*, tout en précisant si la différence lui a semblé *très facile, facile, délicate, difficile, ou quasiment impossible* à percevoir.

Une échelle de notation<sup>2</sup> numérotée de 1 à 5, a été mise au point afin de résumer l'opinion des 7 juges sur les treize échantillons vis-à-vis des questions :

- 1 : le témoin est nettement supérieur à l'échantillon,
- 2 : le témoin est supérieur à l'échantillon,
- 3 : le juge n'a pas perçu de différences,
- 4 : l'échantillon est supérieur au témoin,
- 5 : l'échantillon est nettement supérieur au témoin.

Ce sont ces notes que l'on retrouve dans les 7 tableaux de données. Les 7 tableaux sont donc tous constitués de valeurs homogènes et du même ordre de grandeur. Les notations sont explicitées dans le Tableau 1.

#### *Préparation des données :*

Tous les tableaux sont  $D$ -centrés avec la métrique uniforme  $D = Id_{13}/13$ .

Dans les espaces des individus (les cafés) de chaque tableau, on utilise les métriques identité :  $Q_k = Id_6 \quad \forall k$  ( $p_k = 6$  propriétés sensorielles) et  $Q = Id_{10}$  ( $q = 10$  propriétés physico-chimiques).

De plus, les tableaux ont été réduits, car, comme le montre le Tableau 2, les inerties des tableaux associés sont disparates : le premier a une très faible inertie par rapport aux autres tableaux. Après réduction, les 7 tableaux  $Y_k$  ont ainsi la même inertie à savoir 6. Le tableau  $X$  (propriétés physico-chimiques) a une inertie de 10.

<sup>2</sup> Pour une description détaillée de l'échelle de notation, se référer à la thèse de Gonzalez (1982).

TABLEAU 1  
Les données

Variables des tableaux  $X$  et  $Y_k$  :

<i>Les propriétés physico-chimiques</i>	
Extrait à sec	EXS
Taux d'extraction	TEE
PH	PHH
Acidité	CID
Densité optique à 430n.m	DO4
Densité optique à 510n.m	DO5
Conductance	CDT
Caféine	CAF
Viscosité	VIS
Capacité de rétention d'eau dans la mouture	CPR
<i>Les propriétés sensorielles</i>	
Parfum qualité	PAC
Parfum intensité	PAI
Amertume	AME
Acidité	ACI
Arôme qualité	ARC
Arôme intensité	ARI

Propriétés d'extraction de chacun des individus (cafés) :

$n^{\circ}$ individu	Température ( $^{\circ}C$ )	Prop. mouture/eau (%)	Granulométrie
1	105	25	2.25
2	70	6	0.5
3	105	10	2.25
4	105	10	0.5
5	95	10	0.5
6	70	10	2.25
7	70	10	0.5
8	95	6	0.5
9	95	6	2.25
10	105	6	2.25
11	105	6	0.5
12	70	25	0.5
13	70	25	2.25

TABLEAU 2  
Inertie des 7 tableaux juges

$k$	1	2	3	4	5	6	7
Inertie	3.148	11.550	8.982	9.361	10.213	10.343	9.207

### 5.2. Application de l'ACIMO-PLS

Nous allons modéliser à l'aide de l'ACIMO-PLS les tableaux des 7 juges en fonction des propriétés physico-chimiques des cafés.

L'application a été réalisée après programmation sous SPLUS ® (Venables & Ripley, 1996).

#### 5.2.1. Résultats de l'analyse proprement dite

Le Tableau 3 nous permet d'étudier la qualité des composantes  $u_{k,1}$  et  $u_{k,2}$  qui, rappelons-le, ne sont pas orthogonales, pour chaque tableau  $Y_k$ .

TABLEAU 3

Résultats de l'ACIMO-PLS d'ordre 1 et 2 :  $\lambda_s$  donne le maximum du critère à l'étape  $s$ .  $cov^2$  donne les covariances carrée partielles  $cov^2(u_{k,s}, t_s)$ . Variance ACP donne la variance des composantes principales de chaque  $Y_k$ . Variance  $u_k$  donne la variance des composantes  $u_k$  de l'ACIMO-PLS. Variance PLS donne la variance des composantes PLS des  $Y_k$  issues de la régression PLS de  $Y_k$  sur  $X$ . Variance ACPVI représente la variance des composantes de l'ACPVI de  $(Y_k, Id_6, D)$  par rapport à  $X$ .

$k$	1	2	3	4	5	6	7	
	<i>Étape 1</i>						$\lambda_1 = 60.801$	
$cov^2$	10.830	8.251	13.770	5.799	9.234	1.631	11.286	
Variance ACP	3.143	3.023	3.348	2.230	2.562	3.260	2.721	
Variance $u_{k,1}$	3.120	2.575	3.143	2.091	2.325	2.010	2.585	
Variance PLS	3.118	2.582	3.167	2.115	2.318	2.150	2.579	
Variance ACPVI	2.925	2.903	3.134	2.046	2.354	1.947	2.502	
	<i>Étape 2</i>						$\lambda_2 = 5.433$	
$cov^2$	1.015	0.887	1.217	0.155	0.893	0.070	1.196	
Variance ACP	1.186	1.308	1.021	1.601	1.475	1.576	1.321	
Variance $u_{k,2}$	1.083	1.210	1.333	1.140	1.052	1.404	1.113	
Variance PLS	0.988	1.647	1.232	1.385	1.149	2.010	1.231	
Variance ACPVI	0.986	1.254	0.858	1.516	1.242	0.892	1.252	

Nous pouvons voir que, pour tous les juges, les variances des différentes composantes sont relativement proches, sauf pour les juges 2, 4 et 6. Il est clair que  $X$  n'explique pas bien les réponses du juge 6 : la variance de la composante ACPVI, qui a un sens ici, est de 1.947 sur le premier axe, alors que la variance du premier axe d'inertie est de 3.260. La variance sur l'axe 1 de la composante de l'ACIMO-PLS se situe, en général, entre la variance de la composante ACPVI et de la composante PLS, les 3 étant assez proches. Il ne semble pas y avoir une relation très nette entre les réponses du juge 6 et les propriétés physico-chimiques des cafés. Nous pouvons remarquer que pour les juges 1 et 7, avec des composantes ACIMO-PLS très proches des composantes de  $X$ , les variances des différentes composantes sont très proches. Le juge 3 est celui dont les composantes sont les plus covariées avec  $t$  (composante de  $X$ ). Mais la variance de l'axe d'inertie est un peu plus élevée que les trois autres variances (qui sont très proches).

La Figure 3 illustre les covariances carrées partielles : les juges 4 et 6 se distinguent des autres juges par des composantes faiblement covariées avec les composantes de  $X$ . Nous pouvons alors penser que les modèles de rang 1 et 2 pour ces 2 juges ne sont pas significatifs.

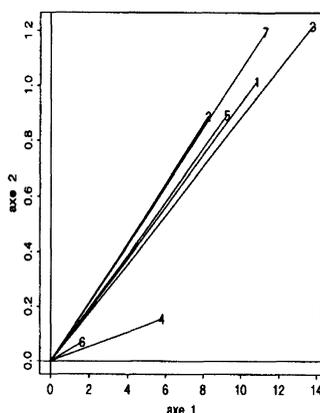


FIGURE 3

*Covariances carrées partielles entre  $u_k$  et  $t$  pour les 2 premières composantes*

### 5.2.2. Choix de la dimension commune des $K$ modèles

Si l'on utilisait les valeurs successives de  $\lambda$  comme critère d'ajustement des modèles  $\hat{Y}_k = X\hat{\beta}_k$ , nous serions amenés à choisir les modèles de rang 1 comme suffisamment explicatifs.

Une autre solution pour retenir le rang des modèles, peut consister à étudier l'évolution des pourcentages de variance expliquée par les approximations  $\hat{Y}_k^s$ , pour chacun des juges, en fonction du rang  $s$  du modèle (Tableau 4). Nous pouvons constater que la

variance expliquée<sup>3</sup> pour chacun des juges, augmente peu à partir de l'étape 1. La première composante  $t_1$  explique à elle seule plus de la moitié de l'inertie totale de  $X$  (57.93%). La deuxième composante, qui n'apporte que 18 %, environ, d'information supplémentaire, est utile pour les représentations graphiques, mais n'a pas lieu d'être retenue dans l'expression des modèles terminaux.

De plus, le pourcentage d'inertie expliqué par les propriétés physico-chimiques pour le juge 6, est très faible par rapport aux autres : les réponses du juge 6 sont peu liées aux propriétés physico-chimiques des cafés. Le juge 3, lui, répond le plus en relation avec les propriétés physico-chimiques. On retrouve là le fait que la première composante du juge 3 est la plus covariée avec  $t_1$ .

TABLEAU 4

*Pourcentage de variance de  $Y_k$  expliquée par  $\hat{Y}_k^s$   
et de  $X$  expliquée par  $t_s$  pour  $s = 1, \dots, 6$ .*

$k$	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
1	31.64	41.24	49.76	52.95	59.13	67.57
2	24.10	32.51	44.21	48.50	53.21	60.70
3	40.23	51.75	57.84	61.02	62.12	69.53
4	16.94	18.41	29.30	32.03	35.17	46.87
5	26.98	35.43	42.30	45.61	52.37	58.75
6	4.77	5.43	13.21	27.24	32.28	38.37
7	32.97	44.29	48.37	54.58	55.86	65.06
$X$	57.93	75.99	86.12	93.59	97.56	98.79

### 5.2.3. Recherche du rang du modèle par validation croisée

Un PRESS global (sur les  $K$  tableaux) nous a permis de retenir des modèles de rang 1.

On a représenté (Figure 4) le PRESS global (somme des PRESS totaux par tableau) et total par variable (somme des PRESS par variable de chacun des juges).

On peut trouver quels sont les juges qui sont les moins bien prédits en dimension 1 grâce aux PRESS totaux par tableau (Tableau 5). On peut également comparer les PRESS des tableaux entre eux, car les tableaux initiaux sont constitués du même type de valeurs. Les modèles de rang 1 pour les juges 3 et 7 ont le plus grand pouvoir prédictif. Ce sont aussi ceux dont les composantes ACIMO-PLS sont les plus proches de celle de  $X$  en dimension 1.

On remarque à nouveau que le juge 6 se distingue des autres : son PRESS est le plus élevé quel que soit la dimension du modèle.

<sup>3</sup> Variance expliquée par la modélisation de  $Y_k$  de rang  $s$  :  $\text{tr}(\hat{Y}_k^s D \hat{Y}_k^s Q_k)$ .

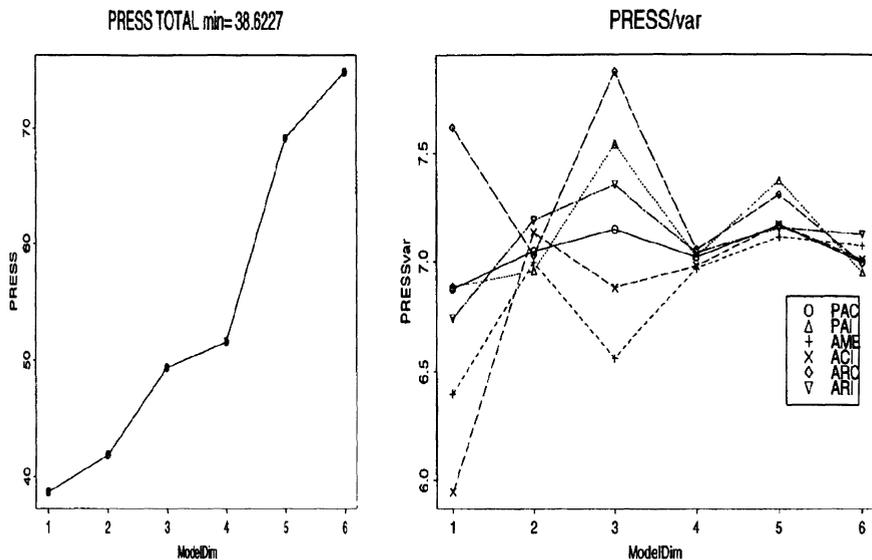


FIGURE 4

*PRESS global et décomposition  
par variable de celui-ci en fonction du rang du modèle.*

TABLEAU 5

*PRESS global et PRESS total pour chaque tableau.*

$k$	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6
1	5.88	6.11	8.55	9.25	11.34	11.18
2	5.82	6.34	7.17	7.59	11.14	12.18
3	4.38	4.18	4.44	4.43	6.04	6.06
4	5.81	6.98	6.75	7.46	9.80	11.91
5	5.26	5.51	6.38	7.06	9.13	9.57
6	6.67	7.93	10.16	9.81	12.60	15.77
7	4.81	4.82	5.92	5.97	9.05	8.18
PRESS global	38.63	41.85	49.37	51.37	69.10	74.75

## 5.2.4. Les modèles

On donne à titre d'exemple les coefficients centrés et réduits  $\hat{\beta}$  des modèles de rang 1 pour les juges 6 et 7 (Tableau 6).

TABLEAU 6  
Coefficients des modèles de rang 1 pour les juges 6 et 7.

## JUGE 6, RANG 1

	EXS	TEE	PHH	CID	DO4	DO5	CDT	CAF	VIS	CPR
PAC	0.024	0.007	-0.005	0.023	0.025	0.024	0.021	0.018	0.019	0.001
PAI	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
AME	0.061	0.017	-0.013	0.058	0.063	0.062	0.053	0.045	0.047	0.003
ACI	0.062	0.017	-0.013	0.059	0.064	0.063	0.054	0.046	0.048	0.003
ARC	-0.012	-0.003	0.002	-0.011	-0.012	-0.012	-0.010	-0.009	-0.009	-0.001
ARI	0.007	0.002	-0.001	0.007	0.007	0.007	0.006	0.005	0.005	0.000

## JUGE 7, RANG 1

	EXS	TEE	PHH	CID	DO4	DO5	CDT	CAF	VIS	CPR
PAC	0.101	0.028	-0.021	0.096	0.105	0.103	0.089	0.075	0.079	0.005
PAI	0.062	0.018	-0.013	0.059	0.065	0.064	0.055	0.047	0.049	0.003
AME	0.107	0.030	-0.022	0.102	0.111	0.109	0.094	0.080	0.083	0.005
ACI	0.123	0.035	-0.026	0.118	0.128	0.126	0.108	0.092	0.096	0.006
ARC	0.078	0.022	-0.016	0.074	0.081	0.079	0.068	0.058	0.060	0.004
ARI	0.101	0.028	-0.021	0.096	0.104	0.103	0.088	0.075	0.078	0.005

D'une manière générale ce sont les variables  $DO4$ ,  $DO5$ ,  $EXS$  et  $CID$  qui sont les plus influentes sur les réponses des juges. On note que les coefficients pour le juge 6 sont très petits par rapport à ceux des autres juges (environ 10 fois). Les signes des coefficients du juge 6 ne sont pas les mêmes que ceux des coefficients du juge 7. Parfois, ils sont opposés, ce qui signifie que les 2 juges répondent inversement l'un de l'autre pour certaines questions.

**Remarques**

– Ici, l'étude des coefficients des modèles retenus est raisonnable comme dans « PLS usuelle » et contrairement à la régression multiple aux moindres carrés ordinaire.

– Une autre manière de valider un modèle, est d'étudier l'évolution des coefficients des modèles en fonction leur rang croissant  $s$ . Lorsqu'il y a une « explosion » des coefficients du modèle de rang  $l$ , on retient un modèle de rang inférieur à  $l$ . (cf. (Durand *et coll.*, 1998)).

### 5.2.5. Information supplémentaire par rapport à la régression PLS « usuelle »

Rappelons que l'on a choisi des modèles  $\hat{Y}_k$  de rang 1, soit  $\hat{Y}_k^1$ . Ce modèle peut s'écrire  $P_{t_1} Y_k$ . C'est le premier modèle partiel de rang 1. La régression de  $u_{k,1}$  sur  $t_1$  est un indicateur de qualité de ce modèle partiel, donc ici du modèle. De même, la régression de  $u_1$  sur  $t_1$  est un indicateur de qualité de  $P_{t_1} Y = \hat{Y}^1$ . La Figure 5 représente ces régressions. D'après ce graphique, on peut dire que la qualité de la régression PLS « usuelle » est assez bonne. Or, on peut voir que la qualité du modèle pour le juge 4 n'est pas bonne. De même pour le juge 6. Les modèles pour ces deux juges ne sont donc pas significatifs. Les coefficients  $\|b_s[k]\|_{Q_k}$  sont les plus faibles pour ces juges. Avec la seule utilisation de la régression PLS de  $Y$  sur  $X$ , il aurait été impossible de constater la défaillance de ces modèles. Grâce aux composantes  $u_{k,1}$ , qui sont les plus proches de  $t_1$  parmi les mieux représentatives des tableaux  $Y_k$ , l'ACIMO-PLS apporte une information supplémentaire.

Pour conclure cet exemple, notons que l'ACIMO-PLS permet une étude plus précise de la qualité des modèles des réponses de chacun des juges. Elle permet de constater que les juges 4 et 6 ne semblent pas être d'accord avec l'ensemble des autres juges, et que les modèles retenus ne sont pas significatifs. Avec uniquement une régression PLS « usuelle », ces conclusions ne sont pas visibles, car celle-ci ne fournit que des composantes globales, combinaisons linéaires de l'ensemble de toutes les variables sensorielles.

## 6. Conclusion

L'ACIMO-PLS présente plusieurs intérêts :

- c'est une méthode simple car ses solutions successives peuvent s'obtenir grâce à une diagonalisation, ou bien grâce à quelques modifications du programme PLS pour deux tableaux.

- Les  $K$  modèles obtenus ont l'avantage d'avoir un rang commun et d'être linéaires en les mêmes combinaisons linéaires (composantes) des prédicteurs, ce qui facilite l'interprétation.

- Elle fournit des éléments supplémentaires par rapport à la régression PLS du tableau conjoint  $Y$  sur les prédicteurs  $X$ , lesquels permettent une étude de la qualité de la reconstitution des tableaux.

Cette méthode, qui semble donc très prometteuse, mérite quelques améliorations afin de passer à des modèles non linéaires et d'avoir la possibilité de l'appliquer lorsqu'il y a des données manquantes.

Le source SPLUS du programme de l'ACIMO-PLS peut être obtenu sur simple demande à [mvivien@pharma.univ-montp1.fr](mailto:mvivien@pharma.univ-montp1.fr).

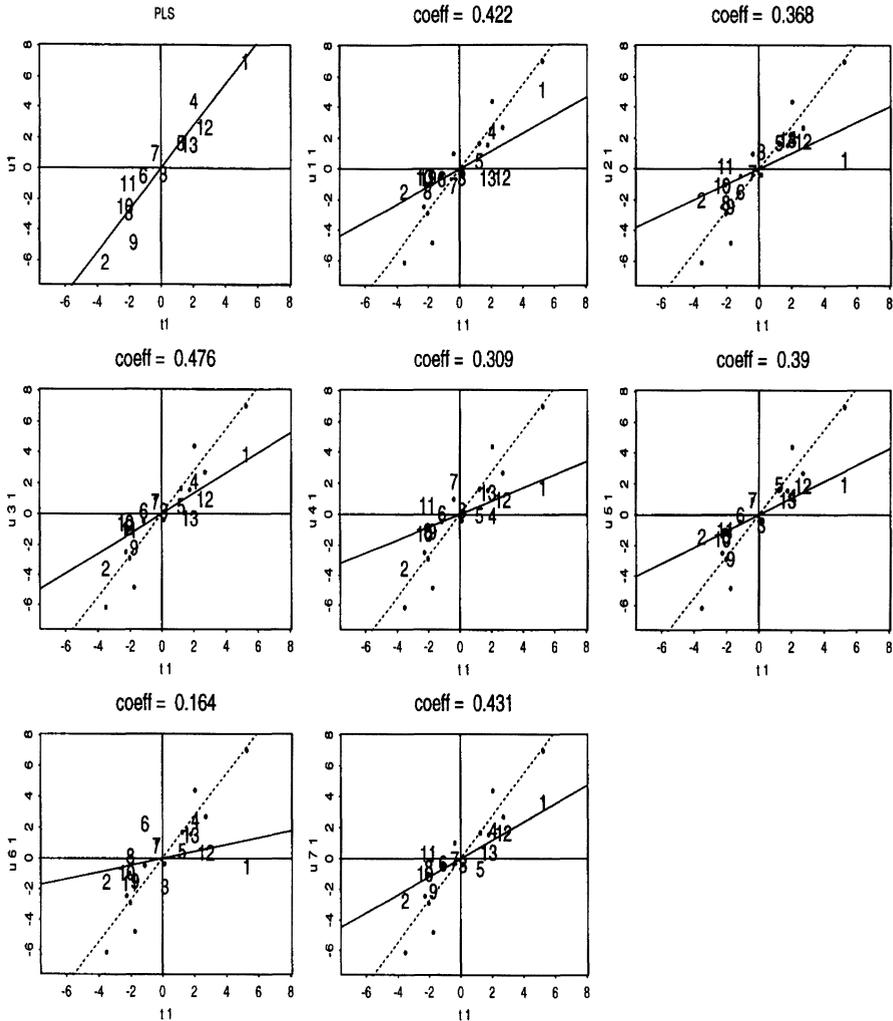


FIGURE 5

Régression de la composante PLS  $u_1$ , et des composantes par tableau  $u_{1,1}, \dots, u_{7,1}$  issues de l'ACIMO-PLS sur la composante  $t_1$ , composante associée à  $X$ . COEFF, qui est égal à  $\|b_1[k]\|_{Q_k}$ , est le coefficient de pondération de  $u_{k,1}$  dans l'expression de  $u_1$  (cf. la propriété du paragraphe 3.3.2.). La droite en pointillés est la droite de régression de  $u_1$  sur  $t_1$ . Les nombres 1 à 13 représentent les individus (les cafés).

## Annexe

### Les données

cafés	juge 1					juge 2					juge 3							
	PAC	PAI	AME	ACI	ARC	ARI	PAC	PAI	AME	ACI	ARC	ARI	PAC	PAI	AME	ACI	ARC	ARI
1	3	5	3	5	4	5	1	3	2	4	1	4	5	5	5	4	5	5
2	3	3	3	1	3	3	1	1	3	1	1	1	1	2	1	1	1	1
3	3	3	3	3	3	3	4	4	3	3	2	5	3	3	3	3	3	3
4	4	4	5	4	3	5	4	5	5	5	3	5	3	3	3	5	5	5
5	3	4	4	3	3	3	3	5	4	5	3	5	2	2	5	3	3	3
6	5	3	3	3	3	3	3	3	1	3	5	2	3	3	3	3	3	3
7	3	3	5	1	3	3	2	4	5	2	2	3	2	4	3	3	3	5
8	3	3	3	1	3	3	3	2	3	1	5	2	3	2	3	1	3	3
9	3	3	3	3	3	3	1	3	1	1	3	1	1	1	3	3	1	1
10	3	3	3	3	3	3	3	4	4	1	4	3	3	3	3	1	3	3
11	4	3	3	3	3	3	3	5	1	4	3	3	3	3	3	3	1	1
12	3	3	3	3	3	3	4	2	3	5	1	5	4	4	3	5	1	3
13	4	3	3	3	3	3	5	5	3	5	3	5	1	1	3	5	3	3
cafés	juge 4					juge 5					juge 6							
	PAC	PAI	AME	ACI	ARC	ARI	PAC	PAI	AME	ACI	ARC	ARI	PAC	PAI	AME	ACI	ARC	ARI
1	5	5	4	3	3	3	5	3	5	5	1	1	2	4	2	2	2	4
2	1	1	1	1	1	1	3	3	1	2	1	1	1	5	1	1	1	5
3	3	3	3	2	3	5	3	3	3	3	3	3	1	5	1	1	3	3
4	3	3	3	3	2	3	5	5	5	2	3	5	2	4	5	5	1	5
5	3	3	5	1	1	3	3	5	5	4	1	5	3	3	3	3	3	3
6	5	3	2	3	3	2	3	3	3	3	1	3	1	5	5	5	1	5
7	5	4	5	2	2	5	5	4	5	3	3	3	5	3	3	3	4	3
8	3	2	5	1	5	1	3	1	5	1	3	4	1	3	3	3	2	3
9	3	5	1	2	2	1	1	1	1	3	3	1	3	1	1	1	3	1
10	1	3	1	4	1	3	3	3	1	3	3	1	3	3	1	2	3	3
11	3	5	3	3	3	3	3	3	5	1	3	3	3	3	1	1	5	3
12	3	3	4	3	1	5	3	3	5	5	1	5	3	1	3	3	3	1
13	5	5	3	3	3	3	3	3	5	3	1	5	3	5	3	5	3	5
cafés	juge 7					caractéristiques physico-chimiques												
	PAC	PAI	AME	ACI	ARC	ARI	EXS	TEE	PHH	CID	DO4	DO5	CDT	CAF	VIS	CPR		
1	5	5	5	5	5	5	2.50	-1.18	-0.02	1.02	0.72	0.31	3.09	0.95	0.32	0.13		
2	2	5	1	1	1	1	-1.44	-4.62	0.27	-0.66	-0.42	-0.17	-2.42	-0.20	-0.11	-0.34		
3	3	5	3	1	3	3	0.18	1.42	-0.10	0.10	0.09	0.04	-0.25	0.25	-0.09	-0.70		
4	4	3	5	4	3	5	0.85	6.59	0.05	0.35	0.56	0.28	0.69	-0.05	-0.04	-0.42		
5	1	4	3	3	3	3	0.49	4.76	0.00	0.22	0.33	0.11	0.44	0.32	-0.04	-0.07		
6	3	3	3	3	3	3	-0.08	-0.15	0.05	0.00	-0.10	-0.06	-0.06	-0.20	-0.19	0.07		
7	1	3	5	3	1	5	0.18	2.56	0.10	0.01	0.01	-0.01	0.32	0.00	-0.13	0.01		
8	1	3	3	4	3	3	-1.16	-4.12	0.05	-0.57	-0.26	-0.10	0.32	0.00	-0.13	0.01		
9	3	3	3	1	1	1	-1.07	-2.76	0.00	-0.22	-0.12	-0.04	-0.64	-0.27	-0.16	-0.09		
10	3	3	3	1	3	3	-0.45	4.76	0.00	-0.50	-0.22	-0.08	-1.68	-0.50	-0.22	0.08		
11	3	3	3	3	3	5	-1.06	-2.85	-0.05	-0.63	-0.17	-0.05	-2.08	-0.60	0.09	0.00		
12	5	5	3	5	1	5	1.89	-5.82	0.15	0.78	0.28	0.12	2.30	1.05	0.05	-0.27		
13	3	3	3	5	3	3	1.45	-6.97	0.20	0.58	0.18	0.08	2.42	0.50	0.05	-0.21		

## Remerciements

Nous remercions P. Cazes pour sa lecture attentive, M. Tenenhaus pour ses remarques, qui ont contribué à l'amélioration de cet article, ainsi que l'autre rapporteur pour son efficacité à détecter les erreurs.

## Références

- BRO R. (1996), Multiway calibration. Multilinear PLS. *Journal of Chemometrics*, **10**, 47-61.
- CARROLL J.-D. (1968), A generalization of canonical correlation analysis to three or more sets of variables. *Proceedings of the 76th convention of the American Psychological Association*, number 3, (pp. 227-228).

- CAZES P. (1997), Adaptation de la régression PLS au cas de la régression après analyse en correspondances multiples. *Revue de Statistique Appliquée*, XLV(2), 89-99.
- CHEssel D. & HANAFI M. (1996), Analyses de co-inertie de  $K$  nuages de points. *Revue de Statistique Appliquée*, XLIV(2), 35-60.
- DURAND J.-F., ROMAN S. & VIVIEN M. (1998), *Guide d'utilisation de la Régression Partial Least Squares linéaire sous Splus*. Rapport technique n° 98-06, Groupe de Biostatistique et d'Analyse des Systèmes, INRA, Montpellier.
- GONZALEZ P.L. (1982), *Analyse statistique de données psycho-sensorielles*. Thèse de 3<sup>ème</sup> cycle, Université Montpellier II - 210p.
- GOWER J.-C. (1975), Generalised Procruste Analysis. *Psychometrika*, 40, 33-51.
- LAFOSSE R. & HANAFI M. (1997), Concordance d'un tableau avec  $K$  tableaux : définition de  $K + 1$  uples synthétiques. *Revue de Statistique Appliquée*, XLV(4), 111-126.
- LAVIT Ch. (1988), *Analyse conjointe de tableaux quantitatifs*. Masson, Paris - 252p.
- QANNARI E.M., COURCOUX Ph. & MAUNIT A. (1997), Analyse d'un ensemble de tableaux : application aux données sensorielles et de préférences. *Actes des 5<sup>èmes</sup> Journées Agro-industrie et Méthodes Statistiques*, Versailles.
- QANNARI E.M., COURCOUX Ph., LEJEUNE M. & MAYSTRE O. (1997), Comparaison de trois stratégies de détermination d'un compromis en évaluation sensorielle. *Revue de Statistique Appliquée*, XLV(1), 61-74.
- RAO C.R. (1964), The use and interpretation of principal component analysis in applied research. *Sankya, A*(26), 329-359.
- ROBIN F. (1996), *Méthodes statistiques permettant d'expliquer l'appréciation hédonique par les caractéristiques sensorielles*. Thèse de 3<sup>ème</sup> cycle, Université Montpellier II, 248p.
- SCHLICH P. (1993), *Contribution à la sensométrie*. Thèse de Doctorat, Université Paris Sud, 355p.
- TENENHAUS M. (1995), *A partial least squares approach to multiple regression, redundancy analysis, and canonical analysis*. Les Cahiers de la Recherche de HEC. CR 550/1995.
- TENENHAUS M. (1998), *La régression PLS : théorie et pratique*. Technip, Paris, 254p.
- TENENHAUS M. (1999), L'approche PLS. *Revue de Statistique Appliquée*, XLVII, 5-40.
- TENENHAUS M., GAUCHI J.-P., & MENARDO C. (1995), Régression PLS et applications. *Revue de Statistique Appliquée*, XLIII, 7-63.
- VENABLES W.-N. & RIPLEY B.-D. (1996), *Modern applied statistics with S-Plus*. Springer-Verlag, New York, 462p.
- VIVIEN M. (1999), *Nouvelles approches en analyse multi-tableaux*. Rapport de stage de DEA de Biostatistique, Montpellier, 96p.

- WOLD H. (1966), Estimation of principal component and related models by iterative least squares. *Multivariate Analysis*, Krishnaiah, P.R. (Ed.), Academic Press, New York, pp. 391-420.
- WOLD H. (1985), Partial Least Squares. *In Encyclopedia of Statistical Sciences*, vol. 6, Kotz, S. & Johnson, N.L. (Eds.), John Wiley & Sons, New York, pp. 581-591.
- WOLD S., GELADI P., ESBENSEN K. & ÖHMAN J. (1987), Multiway Principal Components- and PLS- Analysis. *Journal of Chemometrics*, **1**, 41-56.
- WOLD S., KETTANEH N. & TJESSEM K. (1996), Hierarchical multiblock PLS and PC models for easier interpretation and as an alternative to variable selection. *Journal of Chemometrics*, **10**, 463-482.