CO-INERTIA ANALYSIS

Analisi Statistica dei Dati Multidimensionali[1]

[1]Corso di Laurea in Scienze Statistiche e Attuariali

Facoltà di Scienze Economiche e Aziendali
Università degli Studi del Sannio

Indice

Co-Inertia Analysis: History

- *Co-Inertia Analysis* is known, even if not by that name, and practiced in several fields.
- It is very famous in **ecology** by the papers of Chessel and Mercier (1993) and Chessel and Doledec (1994).
- In the **atmospheric sciences**, where it is well known as *Singular Value Decomposition Analysis* (SVD), has been popularized by Bretherton *et al.* (1992) and Wallace *et al.* (1992) in order to detect temporally synchronous spatial patterns even if its first use in climatology was apparently by Prohaska (1976).
- In this context, *Predictor Analysis* (Thacker, 1999) can be regarded as a metric-based SVD of the cross-correlation matrix.

Co-Inertia Analysis: History

- It is very popular also in the **social sciences** where it belongs to a class of methods of matching matrices.
- Van de Geer (1984) referred to it as the *MAXDIFF criterion*.
- Previously, Tucker (1958) introduced this method, with the name *Inter-Battery Factor Analysis*, in order to find common factors in two batteries of tests presented to the same group of statistical units.
- Finally, it is well known also in the study of **behavioral teratology** with the name *Partial Least Squares–SVD* where Sampson *et al.* (1989) introduced it in a study of the relationship between fetal alcohol exposure and neurobehavioral deficits.

Basic definitions and Notations

- Let $(\mathbf{X}, \mathbf{Q}_\mathbf{X}, \mathbf{D})$ be a statistical study associated with the matrix $\mathbf{X}$ of order $(n \times p)$, collecting a set of p quantitative/qualitative variables observed on n statistical units.
- $\mathbf{Q}_\mathbf{X}$ is a $(p \times p)$ non negative definite metric of statistical units in $\Re^p$ so that the distance between two statistical units $\mathbf{x}_j$ and $\mathbf{x}_k$ is $(\mathbf{x}_j - \mathbf{x}_k)^T \mathbf{Q}_\mathbf{X} (\mathbf{x}_j - \mathbf{x}_k)$ with $\mathbf{X}^T$ the transpose of matrix $\mathbf{X}$.
- Let $\mathbf{D}$ be the (diagonal) weights metric into vectorial space of variables $\Re^n$.

Basic definitions and Notations

- In the same way, let $(\mathbf{Y}, \mathbf{Q_Y}, \mathbf{D})$ be the statistical study associated with the matrix $\mathbf{Y}$ of order $(n \times q)$, collecting a set of q (quantitative/qualitative) variables observed on the same n statistical units.
- $\mathbf{Q_Y}$ is the $(q \times q)$ non negative definite metric of the statistical units in $\Re^q$.
- We assume that all the variables are columnwise $\mathbf{D}$-centered.

Basic definitions and Notations

- These two statistical triplets are characterized by the same statistical units on which are observed two sets of different variables, so that, statistical units belong to different spaces.
- In this sense, two statistical triplets $(\mathbf{X}, \mathbf{Q_X}, \mathbf{D})$ and $(\mathbf{Y}, \mathbf{Q_Y}, \mathbf{D})$ are said to be matched when they describe the same statistical units
- They are defined to be *fully matched* if they describe the same n statistical units by means of the same p variables (Lafosse, 1985, 1989).

Basic definitions and Notations

- The study of a statistical triplet $(\mathbf{X}, \mathbf{Q_X}, \mathbf{D})$ (Escoufier,1987) is equivalent, from a geometrical point of view, to search the inertia axes of a cloud of n points of $\Re^p$ (principal axes) or, in similar way, looking for the inertia axes of a cloud of p points in $\Re^n$ (principal components).
- The solutions involve to diagonalize the inertia operators $\mathbf{X}^T\mathbf{D}\mathbf{X}\mathbf{Q_X}$ and $\mathbf{X}\mathbf{Q_X}\mathbf{X}^T\mathbf{D}$, respectively.
- The total inertia $trace(\mathbf{X}^T\mathbf{D}\mathbf{X}\mathbf{Q_X})$ is a global measure of the variability of the data and it can be decomposed on a set of p orthogonal $\mathbf{Q_X}$-normed vectors $\mathbf{w}_k$:

$$\sum\nolimits_{k=1}^{p} \mathbf{w}_k^T\mathbf{Q_X}\mathbf{X}^T\mathbf{D}\mathbf{X}\mathbf{Q_X}\mathbf{w}_k$$

- The total inertia of matrix $\mathbf{Y}$ can be also decomposed on a set of q orthogonal $\mathbf{Q_Y}$-normed vectors $\mathbf{c}_j \quad (j = 1, ..., q)$

Co-Inertia Analysis

- The mathematical model of the co-inertia analysis may be examined by using the statistical triplets $(\mathbf{X}, \mathbf{Q_X}, \mathbf{D})$ and $(\mathbf{Y}, \mathbf{Q_Y}, \mathbf{D})$.
- These triplets are characterized by the same statistical units on which are observed two different sets of variables, so that, statistical unit vectors belong to different spaces $\Re^p$ and $\Re^q$, respectively.
- In order to study the common geometry of the two clouds Chessel and Mercier (1993) suggested the Co-Inertia Analysis (hereafter COA) which is a symmetric coupling method.

Co-Inertia Analysis

- Aim of COA is to find linear combinations of the data $\mathbf{t}_i = \mathbf{X}\mathbf{Q}_\mathbf{X}\mathbf{w}_i$ and $\mathbf{u}_i = \mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{c}_i$ with the maximum covariance

$$\mathrm{cov}^2(\mathbf{t}_i, \mathbf{u}_i)_\mathbf{D} = < \mathbf{t}_i, \mathbf{u}_i >^2_\mathbf{D} = (\mathbf{w}_i^T \mathbf{Q}_\mathbf{X} \mathbf{X}^T \mathbf{D} \mathbf{Y} \mathbf{Q}_\mathbf{Y} \mathbf{c}_i)^2$$

  satisfying the constraints

$$\mathbf{w}_i^T \mathbf{Q}_\mathbf{X} \mathbf{w}_i = \mathbf{c}_i^T \mathbf{Q}_\mathbf{Y} \mathbf{c}_i = 1$$

  and

$$\mathbf{w}_i^T \mathbf{Q}_\mathbf{X} \mathbf{w}_j = \mathbf{c}_i^T \mathbf{Q}_\mathbf{Y} \mathbf{c}_j = 0 \text{ for } i \neq j$$

  and $i = 1, ..., s$ with $s = \min(p,\ q)$.
- The solution to this problem is the same (Cherry, 1997) as the solution to the simultaneous orthogonal rotation problem (Cliff, 1966).

- The criterion $\text{cov}^2(\mathbf{t}_i, \mathbf{u}_i)_{\mathbf{D}}$ of the COA problem measures the proximity between the variables of both sets.
- This methods searches then for the components (under constraints on the weighting vector norms) associated with each groups of variables that are the most close, in order to point out and to explain the proximities.
- The criterion $\text{cov}^2(\mathbf{t}_i, \mathbf{u}_i)_{\mathbf{D}}$ is equal to

$$\text{cov}^2(\mathbf{t}_i, \mathbf{u}_i)_{\mathbf{D}} = \text{cor}^2(\mathbf{t}_i, \mathbf{u}_i)_{\mathbf{D}} \times \text{var}(\mathbf{t}_i)_{\mathbf{D}} \times \text{var}(\mathbf{u}_i)_{\mathbf{D}}$$

  so maximizing this criterion we also maximize the correlation between the components $(\mathbf{t}_i, \mathbf{u}_i)$ and their respective variances, simultaneously.

Co-Inertia Analysis

- The first step ($i = 1$) of this approach consists then in computing the weight vectors $\mathbf{w}_1$ and $\mathbf{c}_1$ of the components $\mathbf{t}_1 = \mathbf{X}\mathbf{Q}_\mathbf{X}\mathbf{w}_1$ and $\mathbf{u}_1 = \mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{c}_1$, respectively.
- The weight vectors $\mathbf{w}_1$ and $\mathbf{c}_1$ are the solutions of order 1 ($i = 1$) of the following objective function:

$$\begin{cases} \max\limits_{\mathbf{w}_i\mathbf{c}_i}(\mathbf{w}_i^T\mathbf{Q}_\mathbf{X}\mathbf{X}^T\mathbf{D}\mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{c}_s)^2 \\ \|\mathbf{w}_i\|^2_{\mathbf{Q}_\mathbf{X}} = 1 \\ \|\mathbf{c}_i\|^2_{\mathbf{Q}_\mathbf{Y}} = 1 \end{cases}$$

where $(\mathbf{w}_i^T\mathbf{Q}_\mathbf{X}\mathbf{X}^T\mathbf{D}\mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{c}_i)^2 = (\mathbf{t}_i^T\mathbf{D}\mathbf{u}_i)^2 = \text{cov}^2(\mathbf{t}_i, \mathbf{u}_i)_\mathbf{D}$ is the squared $\mathbf{D}$-covariance between $\mathbf{t}_i$ and $\mathbf{u}_i$, and $\|\mathbf{w}_s\|^2_{\mathbf{Q}_\mathbf{X}} = \mathbf{w}_i^T\mathbf{Q}_\mathbf{X}\mathbf{w}_i$ is the squared $\mathbf{Q}_\mathbf{X}$-norm of $\mathbf{w}_i$.

- Due to the symmetry between the properties concerning $\mathbf{X}$ and $\mathbf{Y}$ for this solution we refer all the formulas to the weight vector $\mathbf{w}_s$ avoiding a double explanation.
- Solutions are obtained by Lagrange multipliers method.
- The Lagrangian function L is defined as

$$L = (\mathbf{w}_1^T \mathbf{Q}_\mathbf{X} \mathbf{X}^T \mathbf{D} \mathbf{Y} \mathbf{Q}_\mathbf{Y} \mathbf{c}_1)^2 - \lambda(\mathbf{w}_1^T \mathbf{Q}_\mathbf{X} \mathbf{w}_1 - 1) - \mu(\mathbf{c}_1^T \mathbf{Q}_\mathbf{Y} \mathbf{c}_1 - 1)$$

- By differentiating L with respect to $\mathbf{w}_1$ and $\mathbf{c}_1$ and setting the result equal to zero, we obtain the following normal equations

$$
\begin{aligned}
\frac{1}{2}\frac{\partial L}{\partial \mathbf{w}_1} &= (\mathbf{w}_1^T \mathbf{Q_X X}^T \mathbf{DYQ_Y c}_1)\mathbf{Q_X X}^T \mathbf{DYQ_Y c}_1 - \lambda \mathbf{Q_X w}_1 = 0 \\
\frac{1}{2}\frac{\partial L}{\partial \mathbf{c}_1} &= (\mathbf{w}_1^T \mathbf{Q_X X}^T \mathbf{DYQ_Y c}_1)\mathbf{Q_Y Y}^T \mathbf{DXQ_X w}_1 - \mu \mathbf{Q_Y c}_1 = 0 \\
\frac{\partial L}{\partial \lambda} &= -\mathbf{w}_1^T \mathbf{Q_X w}_1 + 1 = 0 \\
\frac{\partial L}{\partial \mu} &= -\mathbf{c}_1^T \mathbf{Q_Y c}_1 + 1 = 0
\end{aligned}
$$

- The pairs of axes $\mathbf{w}_i$ and $\mathbf{c}_j$ are obtained by the eigenvectors $\tilde{\mathbf{w}}_i$ (or $\tilde{\mathbf{c}}_j$) associated to the decomposition of the operators

$$\mathbf{Q}_X^{1/2}\mathbf{X}^T\mathbf{D}\mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{Y}^T\mathbf{D}\mathbf{X}\mathbf{Q}_X^{1/2} \text{ with } (p < q)$$

or

$$\mathbf{Q}_Y^{1/2}\mathbf{Y}^T\mathbf{D}\mathbf{X}\mathbf{Q}_\mathbf{X}\mathbf{X}^T\mathbf{D}\mathbf{Y}\mathbf{Q}_Y^{1/2} \text{ with } (q < p)$$

respectively, linked to the same maximum eigenvalue

$$\lambda = (\mathbf{w}_i^T\mathbf{Q}_\mathbf{X}\mathbf{X}^T\mathbf{D}\mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{c}_i)^2$$

where $\sqrt{\lambda}$ is the covariance between $\mathbf{t}_i$ and $\mathbf{u}_i$.
- After diagonalization s principal axes are preserved.

- All the remaining COA weight vectors $\mathbf{w}_s$ and $\mathbf{c}_s$ ($s = 1, ..., \min(p, q)$) are obtained in a single stage by

$$\mathbf{W}_s = \mathbf{Q}_\mathbf{X}^{-1/2}\tilde{\mathbf{W}}_s$$

and

$$\mathbf{C}_s = \mathbf{Y}^T\mathbf{D}\mathbf{X}\mathbf{Q}_\mathbf{X}^{1/2}\tilde{\mathbf{W}}_s\mathbf{\Lambda}_s^{-1/2}$$

respectively, such that $\mathbf{W}_s^T\mathbf{Q}_\mathbf{X}\mathbf{W}_s = \mathbf{I}_s$ and $\mathbf{C}_s^T\mathbf{Q}_\mathbf{Y}\mathbf{C}_s = \mathbf{I}_s$.

- Moreover, $\mathbf{X}$ and $\mathbf{Y}$ COA row scores are given by

$$\mathbf{T}_s = \mathbf{X}\mathbf{Q}_\mathbf{X}\mathbf{W}_s = \mathbf{X}\mathbf{Q}_\mathbf{X}^{1/2}\tilde{\mathbf{W}}_s$$

and

$$\mathbf{U}_s = \mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{C}_s = \mathbf{Y}\mathbf{Q}_\mathbf{Y}\mathbf{Y}^T\mathbf{D}\mathbf{T}_s\mathbf{\Lambda}_s^{-1/2}$$

respectively.

- Computationally $\text{COA}(\mathbf{X}, \mathbf{Y})_{\mathbf{Q_X},\mathbf{Q_Y}}$ amounts to the GSVD of the matrix $\mathbf{X}^T\mathbf{DY}$with the row metric $\mathbf{Q_X}$ and the column metric $\mathbf{Q_Y}$ and it is denoted $\text{GSVD}(\mathbf{X}^T\mathbf{DY})_{\mathbf{Q_X},\mathbf{Q_Y}}$.
- This method is also defined by the analysis of the statistical triplet $(\mathbf{Y}^T\mathbf{DX}, \mathbf{Q_X}, \mathbf{Q_Y})$.
- The solution to this problem is the same (Cherry, 1997) as the solution to the simultaneous orthogonal rotation problem (Cliff, 1966).
- The number of components cannot be larger than the minimum of the ranks of $\mathbf{X}$ and $\mathbf{Y}$. This implies that this method leads to only one component when $\mathbf{X}$ or $\mathbf{Y}$ is limited to one variable.

- It is easy to show that if we set $\mathbf{Q_X} = \mathbf{I}_p$, $\mathbf{Q_Y} = \mathbf{I}_q$ and $\mathbf{D} = \mathbf{I}_n$ then Tucker's approach, COA$(\mathbf{X}, \mathbf{Y})_{\mathbf{I}_p,\mathbf{I}_q}$ and Undeflated PLS (Burnham et al., 1996) lead to the same results, and the first solutions of COA$(\mathbf{X}, \mathbf{Y})_{\mathbf{I}_p,\mathbf{I}_q}$ and PLS Regression (Höskuldsson, 1988) are equal.
- We highlight that

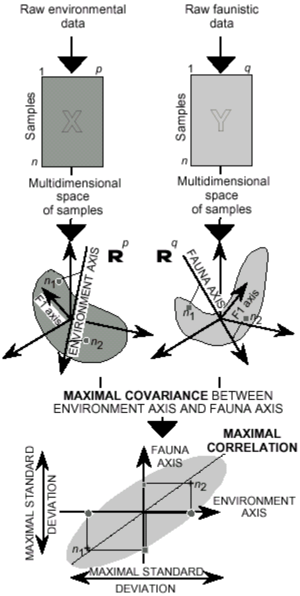
$$\text{cov}^2(\mathbf{t}_k, \mathbf{u}_j) = \text{cor}^2(\mathbf{t}_k, \mathbf{u}_j) \times var(\mathbf{t}_k) \times var(\mathbf{u}_j)$$

  Note that the square of the entity cor(.) is maximized via Canonical Correlation Analysis (hereafter CCA) while a co-inertia axis maximizes $\text{cov}^2(.)$.
- COA makes a compromise between a CCA of the two sets $\mathbf{X}$ and $\mathbf{Y}$, a PCA of $\mathbf{X}$ and a PCA of $\mathbf{Y}$

- Usually, correlation is a suitable measure of the linear association and, in this sense, CCA is more appropriate.
- If the units of measurements are the same within each set of variables and differences in covariation result to be relevant then CCA may obscure important information leading to very high correlated and uninteresting pairs of canonical variables.
- For this reason, COA seems to be more appropriate for analyzing covariance matrices rather than correlation matrices.
- Moreover, COA does not require invertibility of the variance and covariance matrix and so it can be also used when the number of the statistical units is less than the number of variables, unlike the canonical correlation analysis.

- COA and CCA should not be however considered as competing techniques due to their different goals (Cherry, 1996, 1997).
- For deeper COA features and its links with other multivariate coupling methods, see mainly (Dolédec and Chessel, 1994) and (Dray et al., 2003), respectively.
- Note that, in general, the normalized **X** scores of axis i are not correlated with all the normalized **Y** scores of axis i' with $i \neq i'$: $\mathbf{t}_i^T \mathbf{D} \mathbf{u}_{i'} = 0$. This scalar product will be clearly a covariance without normalizing data.
- Moreover, each **X** and **Y** component scores are not **D**-independent like in PCA, that is $\mathbf{t}_i^T \mathbf{D} \mathbf{t}_{i'} \neq 0$ with $i \neq i'$.

- From a geometrical point of view, Co-Inertia Analysis consists in seeking for pairs of normalized axes belonging to $\Re^p$ and $\Re^q$, respectively.
- It is equivalent in maximizing the covariance into a single space $\Re^p$ (or $\Re^q$) after a rotation.
- In this sense, Procustian Analysis and Co-Inertia Analysis are equivalent.


Raw environmental data
Raw faunistic data
1
p
1
q
Samples
X
Samples
Y
n
n
Multidimensional space of samples
Multidimensional space of samples
R^p
R^q
ENVIRONMENT AXIS
F1 axis
n1
n2
FAUNA AXIS
F1 axis
n1
n2
MAXIMAL COVARIANCE BETWEEN ENVIRONMENT AXIS AND FAUNA AXIS
MAXIMAL CORRELATION
FAUNA AXIS
ENVIRONMENT AXIS
MAXIMAL STANDARD DEVIATION
MAXIMAL STANDARD DEVIATION
n1
n2

- Furthermore, we compare the projected variability resulting from separate PCA of $\mathbf{X}$ and $\mathbf{Y}$ with that coming from COA by computing the scores of the initial PCA axes projected onto the co-inertia axes.
- Let $\mathbf{T}_r$ and $\mathbf{U}_\alpha$ be matrices collecting the p and q scores on r and α preserved axes coming from separated PCA of $\mathbf{X}$ and $\mathbf{Y}$, respectively.
- Let $\mathbf{\Lambda_T}$ and $\mathbf{\Lambda_U}$ be the associated diagonal matrices of singular values.
- $\hat{\mathbf{T}}_r = \mathbf{\Lambda_T}^{-\frac{1}{2}} \mathbf{T}_r^T \mathbf{W}_s$ and $\hat{\mathbf{U}}_s = \mathbf{\Lambda_U}^{-\frac{1}{2}} \mathbf{U}_\alpha^T \mathbf{C}_s$ will be then the sought scores of the projected PCA axes, respectively.

COA Duality Diagram

Statistical Study $(\mathbf{Y}^T\mathbf{DX}, \mathbf{Q_X}, \mathbf{Q_Y})$

$$
\begin{tikzcd}
\mathfrak{R}^{p} \arrow[r, "\mathbf{Q_X}"] & \mathfrak{R}^{p*} \arrow[d, "\mathbf{Y}^T\mathbf{DX}"] \\
\mathfrak{R}^{q*} \arrow[u, "\mathbf{X}^T\mathbf{DY}"] & \mathfrak{R}^{q} \arrow[l, "\mathbf{Q_Y}"]
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathfrak{R}^{p} & \mathfrak{R}^{p} \arrow[r, "\mathbf{Q_X}"] & \mathfrak{R}^{p*} \arrow[d, "\mathbf{X}"] & \mathfrak{R}^{p*} \arrow[ddd, dashed, "\mathbf{Y}^T\mathbf{DX}"] \\
& \mathfrak{R}^{n*} \arrow[u, "\mathbf{X}^T"] & \mathfrak{R}^{n} \arrow[d, "\mathbf{D}"] & \\
& \mathfrak{R}^{n} \arrow[u, "\mathbf{D}"] & \mathfrak{R}^{n*} \arrow[d, "\mathbf{Y}^T"] & \\
\mathfrak{R}^{q*} \arrow[uuu, dashed, "\mathbf{X}^T\mathbf{DY}"] & \mathfrak{R}^{q*} \arrow[u, "\mathbf{Y}"] & \mathfrak{R}^{q} \arrow[l, "\mathbf{Q_Y}"] & \mathfrak{R}^{q}
\end{tikzcd}
$$

Orthogonal Co-Inertia Analysis

- We know that, in general, the normalized **X** scores of axis i are not correlated with all the normalized **Y** scores of axis i' with $i \neq i'$: $\mathbf{t}_i^T \mathbf{D} \mathbf{u}_{i'} = 0$. Moreover, each **X** and **Y** component scores are not **D**-independent ($\mathbf{t}_i^T \mathbf{D} \mathbf{t}_{i'} \neq 0$) like in PCA..
- In order to have **X** and **Y** components scores **D**-orthogonal in $\Re^n$, Tenenahus (1995) suggests to add the constraints $\mathbf{t}_i^T \mathbf{D} \mathbf{t}_{i'} = 0$ and $\mathbf{u}_i^T \mathbf{D} \mathbf{u}_{i'} = 0$ to the co-inertia criteria.
- In order to compute the solutions of order $s > 1$, he defines the residual matrix $\mathbf{X}_s$ (or $\mathbf{Y}_s$) as the orthogonal projection of $\mathbf{X}_s$ (or $\mathbf{Y}_s$) onto the subspaces spanned by the components $[\mathbf{t}_1, ..., \mathbf{t}_s]$ (or $[\mathbf{u}_1, ..., \mathbf{u}_s]$).

Orthogonal Co-Inertia Analysis

- At the end of each $s-th$ step, the residual matrices are updated.
- The orthogonal co-inertia axes $\mathbf{w}_1, ..., \mathbf{w}_S$ and $\mathbf{c}_1, ...\mathbf{c}_S$, as well as the orthogonal COA scores $\mathbf{T}_s$ and $\mathbf{U}_s$, are then sequentially determined by using the equations

$$\mathbf{Q}_X^{1/2}\mathbf{X}_s^T\mathbf{D}\mathbf{Y}_s\mathbf{Q}_\mathbf{Y}\mathbf{Y}_s^T\mathbf{D}\mathbf{X}_s\mathbf{Q}_X^{1/2} \text{ with } (p < q)$$

or

$$\mathbf{Q}_Y^{1/2}\mathbf{Y}_s^T\mathbf{D}\mathbf{X}_s\mathbf{Q}_\mathbf{X}\mathbf{X}_s^T\mathbf{D}\mathbf{Y}_s\mathbf{Q}_Y^{1/2} \text{ with } (q < p)$$

where S is the a priori fixed number of retained axes with $S \leq \min(p,\ q)$.

Orthogonal Co-Inertia Analysis: Alghorithm

Step 1

- Compute first COA solutions $\mathbf{t}_1$ and $\mathbf{u}_1$;
- Set $\mathbf{T}_1 = [\mathbf{t}_1]$ and $\mathbf{U}_1 = [\mathbf{u}_1]$;
- Let $\mathbf{P}_{\mathbf{T}_1} = \mathbf{T}_1(\mathbf{T}_1^T\mathbf{T}_1)^{-1}\mathbf{T}_1^T$ and $\mathbf{P}_{\mathbf{U}_1} = \mathbf{U}_1(\mathbf{U}_1^T\mathbf{U}_1)^{-1}\mathbf{U}_1^T$ be orthogonal projectors onto $Im(\mathbf{T}_1)$ and $Im(\mathbf{U}_1)$, respectively, and set $s = 2$;

Step 2

- Compute residual matrices

$$\mathbf{X}_s = \mathbf{X} - \mathbf{P}_{\mathbf{T}_{s-1}}\mathbf{X} = (\mathbf{I} - \mathbf{P}_{\mathbf{T}_{s-1}})\mathbf{X} = \mathbf{P}^{\perp}_{\mathbf{T}_{s-1}}\mathbf{X}$$

and

$$\mathbf{Y}_s = \mathbf{Y} - \mathbf{P}_{\mathbf{U}_{s-1}}\mathbf{Y} = (\mathbf{I} - \mathbf{P}_{\mathbf{U}_{s-1}})\mathbf{Y} = \mathbf{P}^{\perp}_{\mathbf{U}_{s-1}}\mathbf{Y}$$

Orthogonal Co-Inertia Analysis: Alghorithm

$\boxed{\textbf{Step 3}}$

- Compute the first eigenvector associated to the greatest eigenvalue of

  $$\mathbf{Q}_X^{1/2}\mathbf{X}_s^T\mathbf{D}\mathbf{Y}_s\mathbf{Q}_\mathbf{Y}\mathbf{Y}_s^T\mathbf{D}\mathbf{X}_s\mathbf{Q}_X^{1/2} \text{ with } (p < q)$$

  or

  $$\mathbf{Q}_Y^{1/2}\mathbf{Y}_s^T\mathbf{D}\mathbf{X}_s\mathbf{Q}_\mathbf{X}\mathbf{X}_s^T\mathbf{D}\mathbf{Y}_s\mathbf{Q}_Y^{1/2} \text{ with } (q < p)$$

  applied to $\mathbf{X}_s$ and $\mathbf{Y}_s$
- Set the new vectors $\mathbf{w}_s$ and $\mathbf{c}_s$

Orthogonal Co-Inertia Analysis: Alghorithm

Step 4

- Compute new COA scores $\mathbf{t}_s = \mathbf{X}_s\mathbf{Q}_\mathbf{X}\mathbf{w}_s$, $\mathbf{u}_s = \mathbf{Y}_s\mathbf{Q}_\mathbf{Y}\mathbf{c}_s$
- Update $\mathbf{T}_s = [\mathbf{t}_1, ..., \mathbf{t}_s]$, $\mathbf{U}_s = [\mathbf{u}_1, ..., \mathbf{u}_s]$;

Step 5

- Compute $\mathbf{P}_{\mathbf{T}_s}$ and $\mathbf{P}_{\mathbf{U}_s}$.
- Set $s = s + 1$ and go to step 2 until $s \leq S$.

By this algorithm, we have

$$\mathbf{t}_s^T\mathbf{D}\mathbf{t}_{s'} = \mathbf{u}_s^T\mathbf{D}\mathbf{u}_{s'} = 0$$

and

$$\mathbf{t}_s^T\mathbf{D}\mathbf{u}_{s'} \neq 0$$

with $s \neq s'$.

COUPLING SUBJECTIVE EVALUATIONS AND MEASURES OF FOODSTUFFS: AN EXAMPLE OF APPLICATION

- 28 batches of two types of fruits (peaches or nectarines) are judged by two different ways.
- They are classified in order of preference, without ex aequo, by 16 referees (J1 - J16) and they have been collected in a matrix **Y** of order (28 x 16).
- 15 quantitative variables (average of 2 judgements) describe the batches of fruits and they have been collected in a matrix **X** of order (28 x 15).
- We used uniform weights (that is 1/n) for the n statistical units.

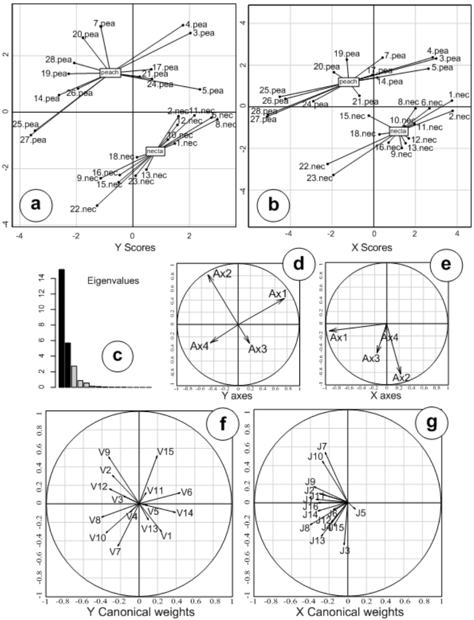
- COA is applied on the data in order to detect what kind of criteria (variables of matrix **X**) have been mainly used by the referees.
- The following table shows also how the first eigenvalue (15.134) is equal to the square of the covariance value $(\mathrm{Covar} = 3.890)$, while this last one is given by the product between the square roots of the inertia of matrices **X** $(\sigma_{\mathbf{X}} = 2.608\ )$ and **Y** $(\sigma_{\mathbf{Y}} = 1.864)$ and the correlation value $(r = 0.8002)$.

COA eigenvalues decompositions

Axis	Eigenvalue	Covar	$\sigma_{\mathbf{X}}$	$\sigma_{\mathbf{Y}}$	r
1	15.134	3.890	2.608	1.864	0.8002
2	5.704	2.388	1.551	1.776	0.8671
3	2.728	1.652	1.471	1.433	0.7832

- The following table shows the values of projected variances of the clouds onto each COA axis are close to the values of projected variances onto each axis of the standard analysis.
- Actually, by projecting the rows cloud of matrix **Y** onto the first co-inertia axis, we obtain a value of projected inertia (6.799) which is slightly smaller than the equivalent PCA inertia value (7.319) in a ratio of 0.9290.
- These losses of projected inertia are due to obtain, for the first axis, a correlation value of 0.8002 between both systems of coordinates.

	Inertia & coinertia **Y**			Inertia & coinertia **X**		
Axes	Inertia	Max	Ratio	Inertia	Max	Ratio
1	6.799	7.319	0.9290	3.476	4.392	0.7914
12	9.204	9.931	0.9268	6.630	7.620	0.8701
123	11.369	11.728	0.9694	8.685	10.053	0.8639
1234	12.313	13.028	0.9451	10.406	11.563	0.8999


a
Y Scores
b
X Scores
c
Eigenvalues
d
Y axes
Ax1
Ax2
Ax3
Ax4
e
X axes
f
Y Canonical weights
g
X Canonical weights
peach
necta