Partial Least Squares Algorithms for PCA, Component-Based Regression and Predictive Path Modeling: the joining ring and their Non-Metric extension.

Giorgio Russolillo

Laboratoire CEDRIC-CNAM

giorgio.russolillo@cnam.fr

le cnam

What is PLS?

→It is an acronym standing for Partial Least Squares

→It refers to a family of ALGORITHMS implementing a suite of METHODS of multivariate analysis for analyzing one, two or several blocks of variables

→All these algorithms are iterative algorithms

→These algorithms consist of various extensions of the Nonlinear estimation by Iterative PArtial Least Squares (NIPALS) algorithm

le cnam

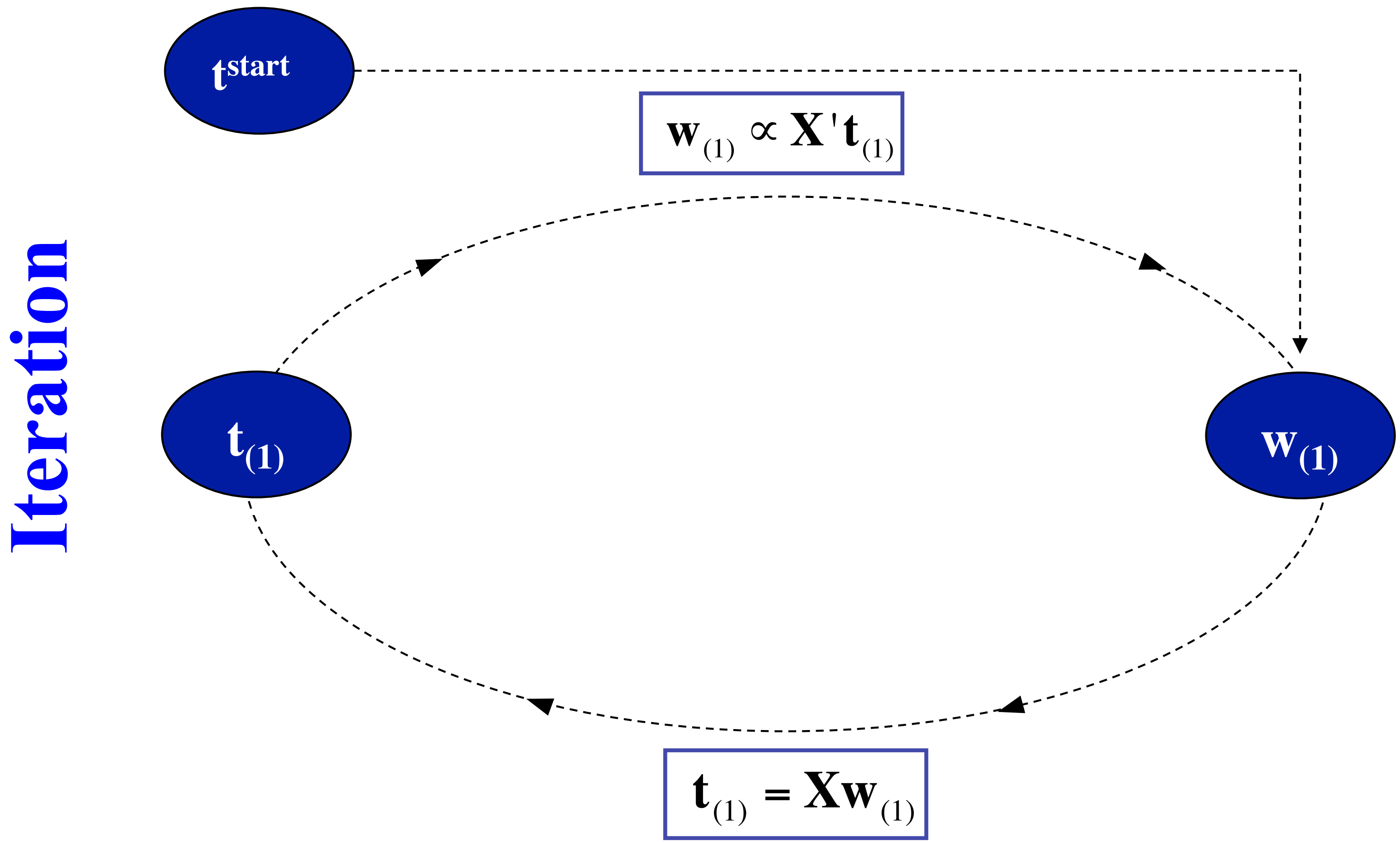
Nonlinear Iterative Partial Least Squares

NIPALS (Wold H., 1966), algorithm is an iterative algorithm for implementing a PCA to a block **X** of variables.

The peculiarity of this algorithm is that it calculates principal components by means of an iterative sequence of simple ordinary least squares (OLS) regressions.

This feature enables us to overcome computational problems due to missing data or landscape data matrices, i.e. matrices having more columns than rows.

NIPALS Algorithm



Deflation: Find $\mathbf{E}_{1(1)}$ as the residual matrix of the regression of $\mathbf{X}$ on $\mathbf{t}_{(1)}$

The PLS family

By using NIPALS-based algorithms, i.e. PLS algorithms, applied on one or several blocks of variables is possible to perform a large number of well known MDA methods, such as:

→ PCA (Hotelling, 1933)

→ CCA (Hotelling, 1936)

→ Tucker Interbactery Analysis (Tucker, 1958)

→ Redundancy Analysis (Van de Wollenberg, 1977)

→ Horst and Carroll GCCA (Horst,1965; Carroll, 1968)

→ Multi-tables analysis (Kettenring, 1971)

→ Analyse Factorielle Multiple (Escofier, 1994)

….. among others

le cnam

The two PLS cultures

PLS Regression models (PLS-R)

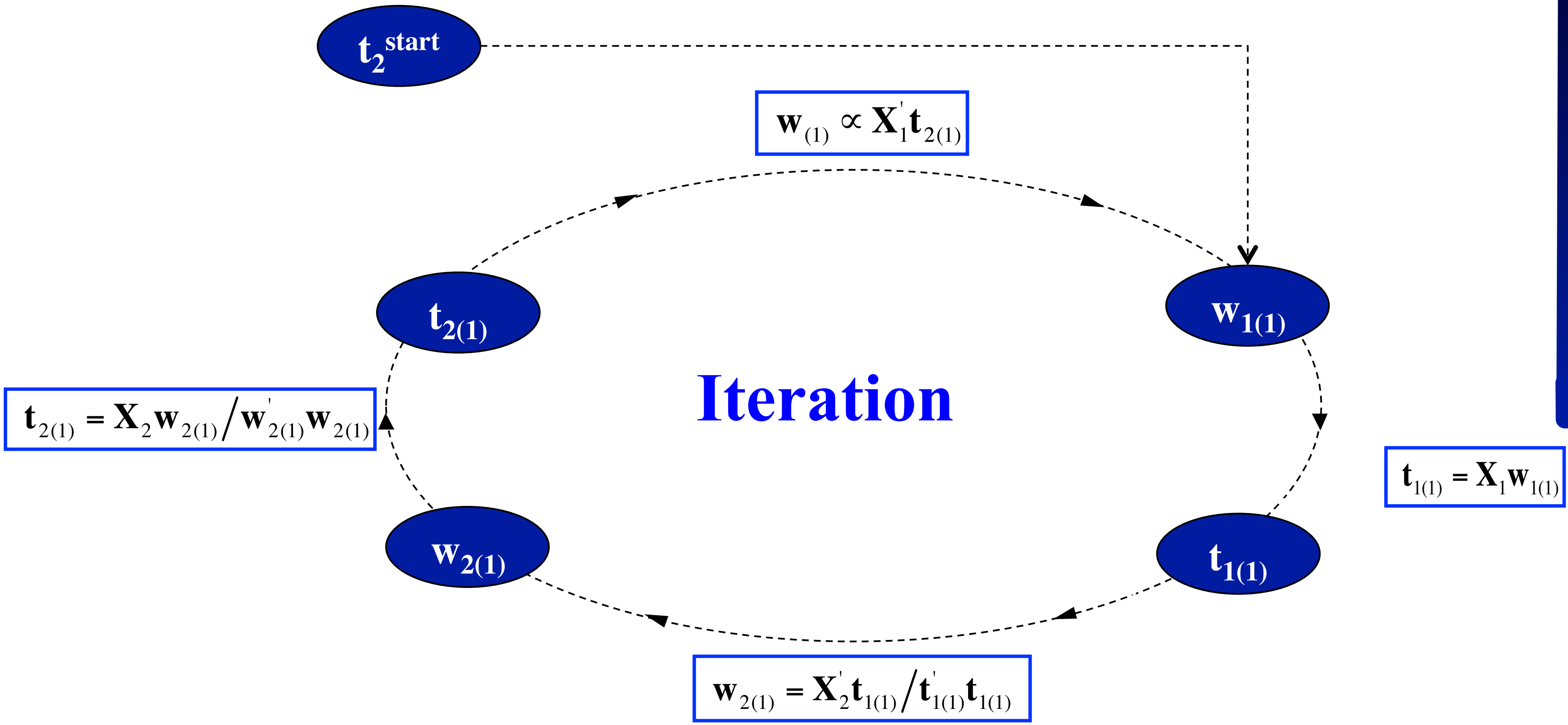
- 2 blocks of variables ($\mathbf{X}_1$, $\mathbf{X}_2$) (at least originally, then L-PLSR, U-PLSR, etc.)
- Component-based Regularized regression tool
- Chemometrics and related areas, Biometrics, Anthropology
- Some references: (non exhaustive list)
  - Wold S., Martens H. and Wold H., 1983
  - Martens H. and Næs T. , 1991
  - Tenenhaus M., 1998

PLS Path Modeling (PLS-PM)

- Several blocks of variables ($\mathbf{X}_1$,…, $\mathbf{X}_Q$) linked by interdependent regressions
- Predictive approach to network of dependence relationships (SEM), hierarchical models (MTA)
- Econometrics, Social Sciences, Marketing, Strategic Management
- Some references: (non exhaustive list)
  - Wold H., 1982, 1985, 1966, 1975, 1977
  - Fornell C., Bookstein F.L., 1982
  - Lohmöller J.B., 1987, 1989
  - Chin W.W., 1998
  - Tenenhaus M. et al., 2005
  - Esposito Vinzi V. et al., 2010

le cnam

PLS-R Algorithm



Deflation:

Find $\mathbf{E}_1$ and $\mathbf{F}_1$ as the residual matrix of the regressions of $\mathbf{X}_1$ and $\mathbf{X}_2$ on $\mathbf{t}_{1(1)}$

PLS Regression criterion

Research of H (chosen by **cross-validation**) **orthogonal** components $\mathbf{t}_{1(h)} = \mathbf{X}_1\mathbf{w}_{1(h)}$ and $\mathbf{t}_{2(h)} = \mathbf{X}_2\mathbf{w}_{2(h)}$ as **correlated** between them as possible **and explanatory** of their own groups.

$$\boxed{\begin{gathered}\mathbf{Cov}^2(\mathbf{t}_{1(h)}, \mathbf{t}_{2(h)}) = \\ \mathbf{Cor}^2(\mathbf{t}_{1(h)}, \mathbf{t}_{2(h)}) * \mathbf{Var}(\mathbf{t}_{1(h)}) * \mathbf{Var}(\mathbf{t}_{2(h)})\end{gathered}}$$

PLS2 regression leads to a **compromise** between **a canonical correlation analysis** between **Y** and **X** and **two principal component analyses** of **X** (orthogonal) and **Y** (oblique = non orthogonal components).

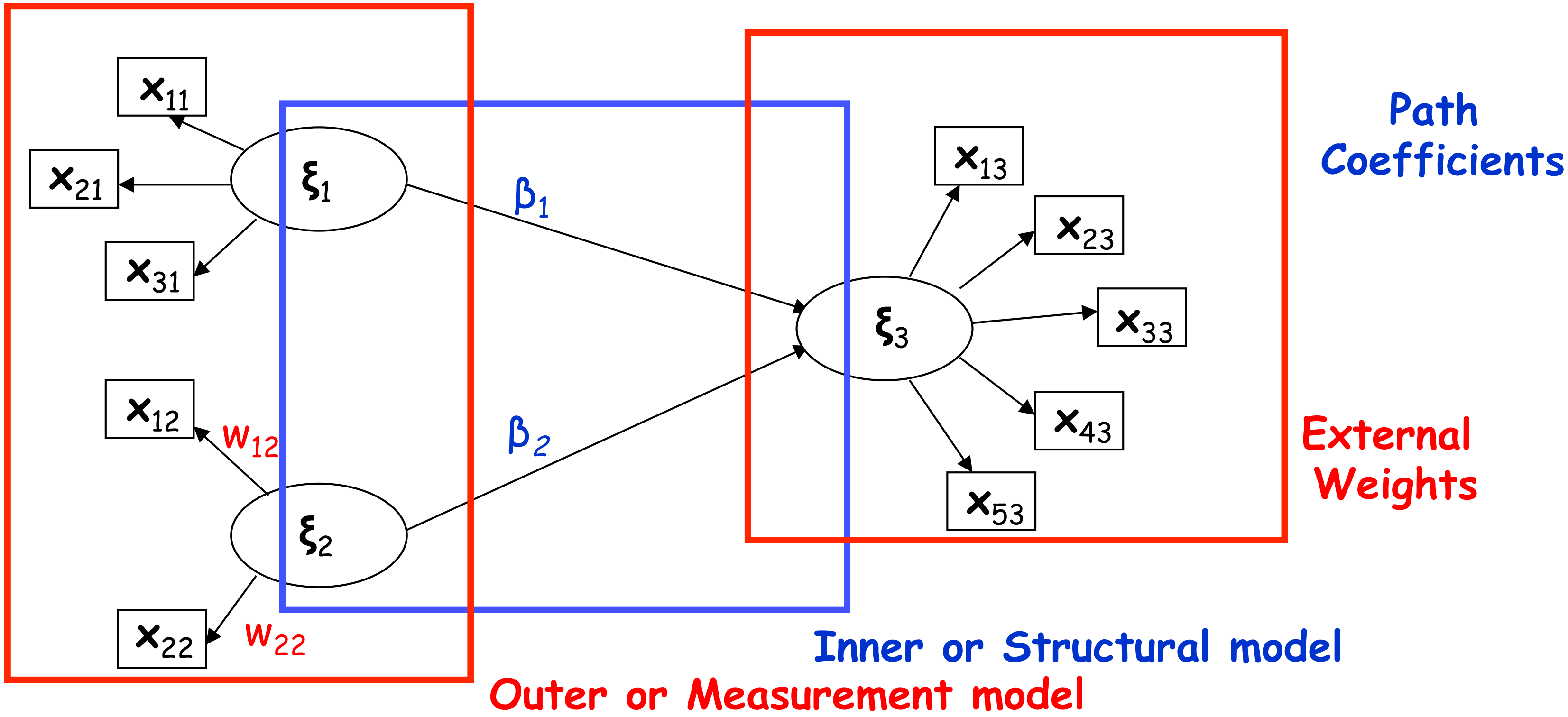
le cnam

PLS-R is useful for..

1. Prediction (Regression equation)
2. Visualization (PLS = Projection on Latent Structures)
3. Regularization (Full component PLS1 coefficients = OLS coefficients)

le cnam

Structural Equation Models: notations

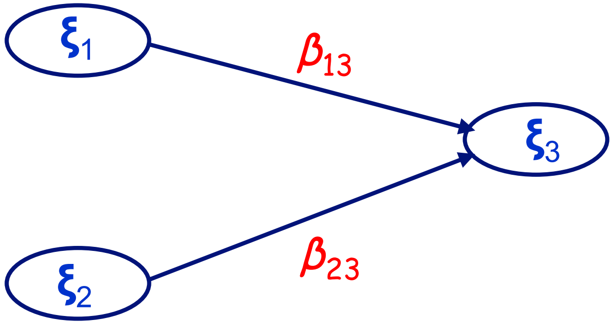


- P **manifest variables** (MVs)observed on N units → $\mathbf{x}_{pq}$ generic MV
- Q **latent variables** (LVs) → ξ_q generic LV
- Q **blocks** composed by each LV and the corresponding MVs

  → in each *q-th* block p_q manifest variables $\mathbf{x}_{pq}$, with $\sum_{q=1}^{Q} p_q = P$

Structural Equation Models: inner model

The structural model describes the causations among the latent variables.



For each endogenous LV in the model it can be written as:

$$\xi_j = \sum_{m=1}^{M} \beta_{mj} \xi_m + \zeta_j$$

where:
- B_{mj} is the path-coefficient linking the *m-th* LV to the *j-th* endogenous LV
- M is the number of the explanatory LVs impacting on ξ_j

le cnam

Structural Equation Models: outer model

Mode A

$$\mathbf{x}_{pq} = \lambda_{pq}\xi_q + \varepsilon_{pq}$$

- **OLS simple regressions** not affected by multicollinearity
- **LV is a principal component of its MV's** (minimizes outer residual variances) under the constraint of being the best neighbor of its adjacent LV's (minimizes inner residual variances)

$$\xi_q = \sum_{p=1}^{P_q} \omega_{pq}\mathbf{x}_{pq} + \delta_q$$

- **OLS multiple regression** affected by multicollinearity
- **LV is the best predictor of its MV's** under the constraint of minimizing the trace of the residual variances in the structural model
- Aim at **minimizing residuals** in structural relationships (explanation of unobserved variance)

Mode B

PLS-PM Algorithm (Mode A)

Outer Estimation

$$\mathbf{t}_q \propto \mathbf{X}_q \mathbf{w}_q$$

Updating* of the inner weights $e_{q'q}$

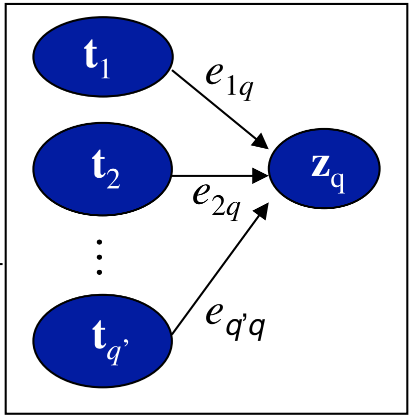
*Choice of inner weights $e_{qq'}$:
- Centroid: correlation signs
- Factorial: correlation coefficients
- Path weighting scheme: OLS regression coefficients ***or*** correlations

Inizialization

$$\mathbf{w}_q$$

Iteration

Inner Estimation

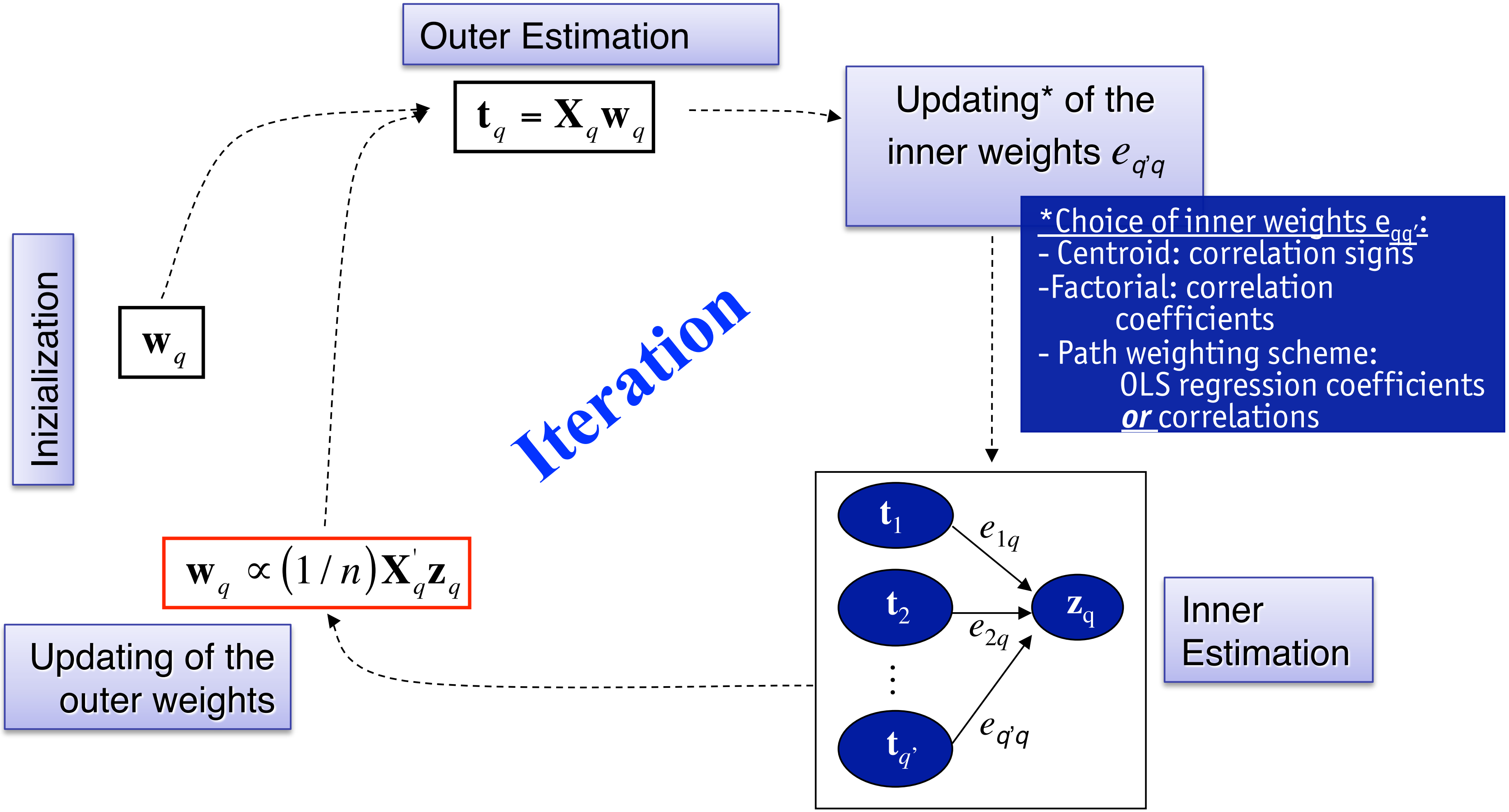


$$\mathbf{w}_q = (1/n)\mathbf{X}_q'\mathbf{z}_q$$

Updating of the outer weights

After convergence: OLS simple/multiple regressions on LV scores for path coefficients

PLS-PM Algorithm (New Mode A)



After convergence: OLS simple/multiple regressions on LV scores for path coefficients

Three PLS Criteria

- **NIPALS:** $\arg\max_{\|\mathbf{w}\|=1}\left\{\operatorname{var}(\mathbf{Xw})\right\}$
- **PLS-R:** $\arg\max_{\|\mathbf{w}_1\|=\|\mathbf{w}_2\|=1}\left\{\operatorname{cov}^2(\mathbf{X}_1\mathbf{w}_1,\mathbf{X}_2\mathbf{w}_2)\right\}$
- **Full Mode A PLS-PM** ?

- **Full New Mode A PLS-PM (Wold algorithm):**

$$\arg\max_{\|\mathbf{w}_q\|=1}\left\{\sum_{q\neq q'} c_{qq'}\, g\left(\operatorname{cov}\left(\mathbf{X}_q\mathbf{w}_q,\mathbf{X}_{q'}\mathbf{w}_{q'}\right)\right)\right\}$$

(*Tenenhaus & Tenenhaus*, 2009)

PLS algorithms: the joining ring (1)

Each iteration of a PLS algorithm can be resumed in three steps:

1. Compute weights
2. Normalize weights
3. Calculate the score vector(s)

NIPALS iteration	PLS-R iteration	new Mode A PLS-PM iteration
$\mathbf{w} = \mathbf{X}'(\mathbf{X}\mathbf{w})$	$\mathbf{w}_q = \mathbf{X}'_q(\mathbf{X}_{q'}\mathbf{w}_{q'})$	$\mathbf{w}_q = \mathbf{X}'_q(\sum_{q'} e_{qq'}\mathbf{X}_{q'}\mathbf{w}_{q'})$
$\mathrm{norm}(\mathbf{w})$	$\mathrm{norm}(\mathbf{w}_q)$	$\mathrm{norm}(\mathbf{w}_q)$
$\mathbf{t} = \mathbf{X}\mathbf{w}$	$\mathbf{t}_q = \mathbf{X}_q\mathbf{w}_q$	$\mathbf{t}_q = \mathbf{X}_q\mathbf{w}_q$

$q \in \{1,2\};\ q \neq q'$

$q \in \{1,2,\ldots,Q\};\ q \neq q'$

$e_{qq'}$ is the generic element of a squared matrix of order Q that is null if $x_{q'}$ is connected to x_q; otherwise, it represents the corresponding inner weight.

PLS algorithms: the joining ring (2)

In all of the previously showed PLS methods, when working on standardized variables, **weights are calculated as Pearson's product-moment correlation coefficients between each variable and a latent construct**.

In homage to what Hayashi called *External Criterion* in its quantification methods, we call this latent construct *Latent Criterion* (LC, noted as γ)

$$\boxed{w_{pq} \propto \mathrm{cor}\left(\mathbf{x}_{pq}, \gamma_q\right)}$$

For each PLS method different LCs are considered:

- In NIPALS, the LC to bear in mind is the first PC.
- In PLS-R, we have to keep into account two LCs: the vector scores in predictor space for the response variables and the vector score in response space for the independent variables.
- In new Mode A PLS-PM, a LC is considered for each block of manifest variables, i.e. the corresponding inner estimate.

A limit of PLS methods

This leads to two basic hypotheses underlying PLS models:

- **Each variable is measured on a interval (or ratio) scale**.
- **Relations** between variables and latent constructs **are linear and**, consequently, **monotone**.

As a consequence, **standard PLS methods cannot handle data** which are measured on a scale which **does not have metric properties**, **nor non-linear relations**.

Solution: the Non-Metric approach

A new class of PLS algorithms making the PLS iteration able to work as optimal scaling algorithms, calculating iteratively both scaling and model parameters: we name them

Non-Metric Partial Least Squares (NM-PLS) algorithms

because they are able to provide optimally scaled data ($\hat{\mathbf{X}}$) with a new metric structure, which does not depend on the metric properties of the raw data ($\mathbf{X}^*$). In other words, NM-PLS methods yield a metric to non-metric data, and a new metric to metric data, linearizing the relations between variables and latent constructs, as required by the hypothesis of standard PLS models.

These methods could be named non-linear PLS methods as well, since they discard the intrinsic linearity hypothesis of standard PLS methods.

Optimal Scaling

According to Young [1981], "*Optimal scaling (OS) is a data analysis technique which assigns numerical values to observation categories in a way which maximizes the relation between the observations and the data analysis model while respecting the measurement character of the data*"

Hence, scalings provided by OS methods are optimal, if they satisfy two conditions:

- they optimize the same criterion of the analysis in which the Optimal Scaling is involved
- they respect the constraints defining which properties of the original measurement scale we want to preserve.

Non-Metric PLS criteria

NM-PLS algorithms optimize NIPALS, PLS-R and New Mode A PLS-PM criteria under two sets of parameters: the model parameters and the scaling parameters, constrained to the restrictions due to the scaling level chosen for each raw variable $\mathbf{x}^*$

- **NM-NIPALS:** $$\underset{\substack{\|\mathbf{w}\|=1,\\ \mathrm{var}(\hat{\mathbf{x}}_p=1)}}{\arg\max}\left\{\mathrm{var}\left(\hat{\mathbf{X}}\mathbf{w}\right)\right\}$$
- **NM-PLSR:** $$\underset{\substack{\|\mathbf{w}_1\|=\|\mathbf{w}_2\|=1,\\ \mathrm{var}(\hat{\mathbf{x}}_{p1})=\mathrm{var}(\hat{\mathbf{x}}_{p2})=1}}{\arg\max}\left\{\mathrm{cov}^2\left(\hat{\mathbf{X}}_1\mathbf{w}_1,\hat{\mathbf{X}}_2\mathbf{w}_2\right)\right\}$$
- **Full New Mode A NM-PLSPM (Wold algorithm):** $$\underset{\substack{\|\mathbf{w}_q\|=1,\\ \mathrm{var}(\hat{\mathbf{x}}_{pq})=1\\ \forall p,q}}{\arg\max}\left\{\sum_{q\neq q'} c_{qq'}\, g\left(\mathrm{cov}\left(\hat{\mathbf{X}}_q\mathbf{w}_q,\hat{\mathbf{X}}_{q'}\mathbf{w}_{q'}\right)\right)\right\}$$

le cnam

Non-Metric PLS scaling levels

NM-PLS methods satisfy model criteria under three possible levels of scaling analysis:

Nominal

Ordinal

Polynomial

To each level of scaling analysis, it corresponds an *ad hoc* scaling function.

Nominal scaling and its properties

- In a nominal analysis, a variable is quantified as orthogonal projection of γ in the space spanned by the columns of the indicator matrix $\tilde{\mathbf{X}}_n$ (Richardson & Kuder, 1933):

$$\boxed{Q\left(\tilde{\mathbf{X}}^{n},\gamma\right)=\tilde{\mathbf{X}}^{n}\left(\tilde{\mathbf{X}}^{n'}\tilde{\mathbf{X}}^{n}\right)^{-1}\tilde{\mathbf{X}}^{n'}\gamma}$$

- Quantification function $Q\left(\tilde{\mathbf{X}}^{n},\gamma\right)$ optimizes model criterion while respecting the grouping constraint that, for each pair of observations i and i':

$$\left(x_i^* \sim x_{i'}^*\right) \Rightarrow \left(\hat{x}_i = \hat{x}_{i'}\right)$$

→ Relation between γ and $\mathbf{x}^*$ in terms correlation ratio can be expressed as Pearson's correlation coefficient between the scaled variable and γ:

$$\mathrm{cor}(\hat{\mathbf{x}},\gamma)=\eta_{\gamma|\mathbf{x}^*}$$

Ordinal scaling and its properties

- In a ordinal analysis, a variable is quantified as orthogonal projection of γ in the space spanned by the columns of the indicator matrix $\tilde{\mathbf{X}}_{\mathrm{o}}$, built by Kruskal's secondary least squares monotonic transformation

$$\boxed{Q\left(\tilde{\mathbf{X}}^{\mathrm{o}},\gamma\right)=\tilde{\mathbf{X}}^{\mathrm{o}}\left(\tilde{\mathbf{X}}^{\mathrm{o}\prime}\tilde{\mathbf{X}}^{\mathrm{o}}\right)^{-1}\tilde{\mathbf{X}}^{\mathrm{o}\prime}\gamma}$$

- Quantification function $Q\left(\tilde{\mathbf{X}}^{\mathrm{o}},\gamma\right)$ optimizes model criterion under the order constraint that, for each pair of observations *i* and *i'*:

$$\left(x_i^* \sim x_{i'}^*\right) \Rightarrow \left(\hat{x}_i = \hat{x}_{i'}\right) \quad \text{and} \quad \left(x_i^* \prec x_{i'}^*\right) \Rightarrow \left(\hat{x}_i \leq \hat{x}_{i'}\right)$$

→ Relation between γ and $\hat{\mathbf{x}}$ in terms of linear correlation can be interpreted as a measure of the approaching monotonicity of the relation between $\mathbf{x}^*$ and the LC, as it is strictly related to Kruskal's STRESS index

Polynomial scaling and its properties

- In functional scaling we suppose that we know the degree D of a polynomial relation between a raw numerical variable and the LC.

- Optimal parameters for the polynomial transformation (Young, 1981) are found by projecting $\boldsymbol{\gamma}$ in the conic space spanned by the columns of matrix $\tilde{\mathbf{X}}^{\mathrm{p}}$, built with a row for each observation and with $(D + 1)$ columns, each column being an integer power of the vector $\mathbf{x}^*$

$$\boxed{Q(\tilde{\mathbf{X}}^{\mathrm{p}},\gamma) = \tilde{\mathbf{X}}^{\mathrm{p}}\left(\tilde{\mathbf{X}}^{\mathrm{p}\prime}\tilde{\mathbf{X}}^{\mathrm{p}}\right)^{-1}\tilde{\mathbf{X}}^{\mathrm{p}\prime}\gamma}$$

- If we suppose that the variable and the LC are linked by a linear relation, we just have to put D = 1. If this is the case for all of the variables, NM-PLS methods provide the same results of the standard PLS methods.

le cnam

How to implement all of these scaling functions?

Problem:
As previously shown, scaling are obtained as a function of LCs.. But LCs, on their turn, are functions of scaling values!

Solution:
Non-Metric PLS algorithms

How do NM-PLS algorithms work?

In NM-PLS algorithms, model and scaling parameters are alternately optimized in a modified PLS loop where a quantification step is added.

- In standard PLS steps the model parameters are optimized for given scaling parameters.
- In the quantification step, instead, the scaling parameters are optimized for given model parameters: raw variables are properly transformed through scaling functions Q, then they are normalized to unitary variance

NM-NIPALS iteration	NM-PLSR iteration	NM-PLSPM iteration
$\widehat{\mathbf{x}}_p \propto \mathcal{Q}(\mathbf{x}^*_p, \mathbf{t})$ $\mathbf{w} = \widehat{\mathbf{X}}'(\widehat{\mathbf{X}}\mathbf{w})$ $\text{norm}(\mathbf{w})$ $\mathbf{t} = \widehat{\mathbf{X}}\mathbf{w}$	$\widehat{\mathbf{x}}_{pq} \propto \mathcal{Q}(\mathbf{x}^*_{pq}, \mathbf{t}_{q'})$ $\mathbf{w}_q = \widehat{\mathbf{X}}'_q(\widehat{\mathbf{X}}_{q'}\mathbf{w}_{q'})$ $\text{norm}(\mathbf{w}_q)$ $\mathbf{t}_q = \widehat{\mathbf{X}}_q\mathbf{w}_q$	$\widehat{\mathbf{x}}_{pq} \propto \mathcal{Q}(\mathbf{x}^*_{pq}, \mathbf{z}_q)$ $\mathbf{w}_q = \widehat{\mathbf{X}}'_q(\sum_{q'} e_{qq'}\widehat{\mathbf{X}}_{q'}\mathbf{w}_{q'})$ $\text{norm}(\mathbf{w}_q)$ $\mathbf{t}_q = \widehat{\mathbf{X}}_q\mathbf{w}_q$

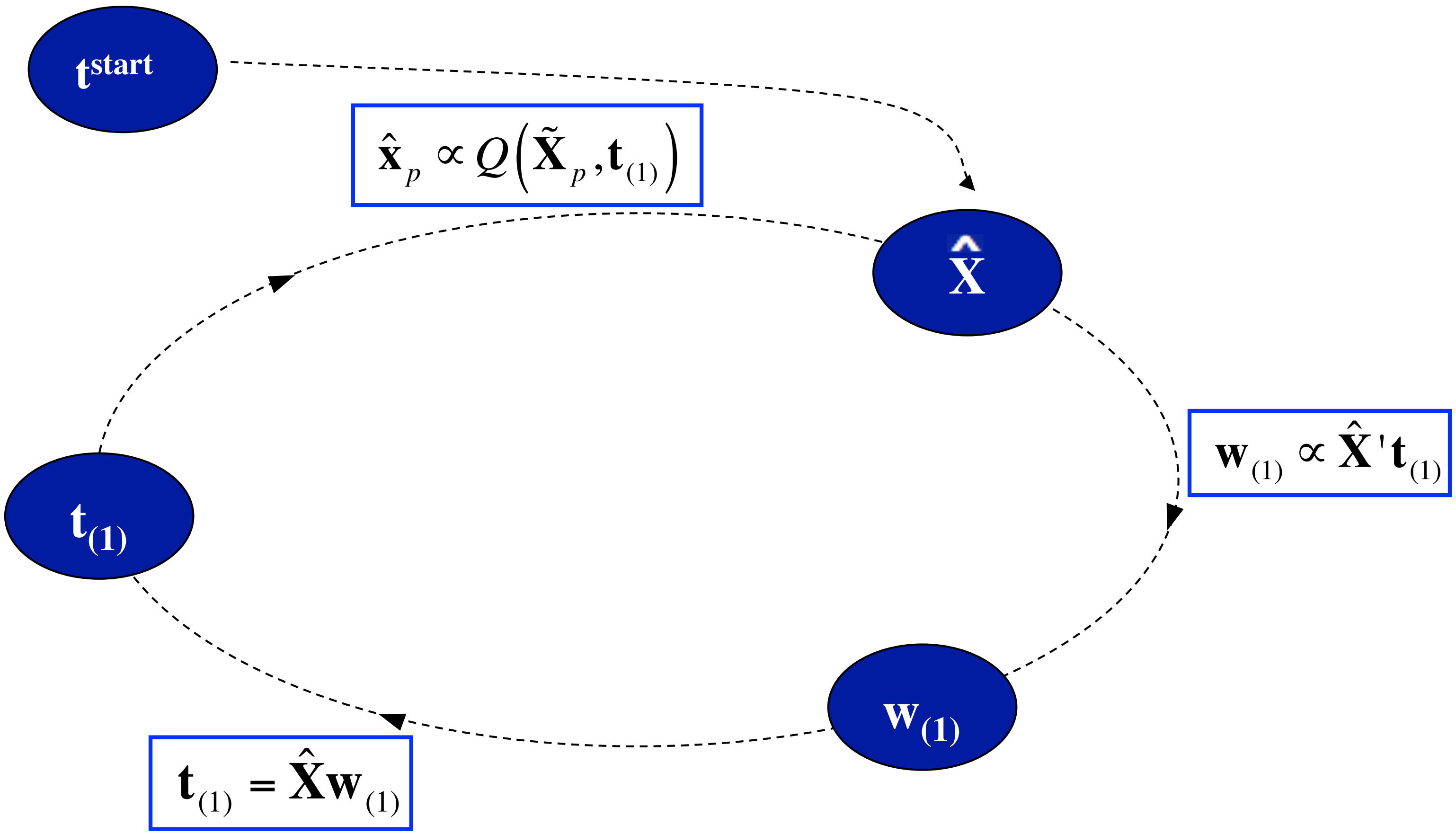
$q \in \{1,2\};\ q \neq q'$

$q \in \{1,2,\ldots,Q\};\ q \neq q'$

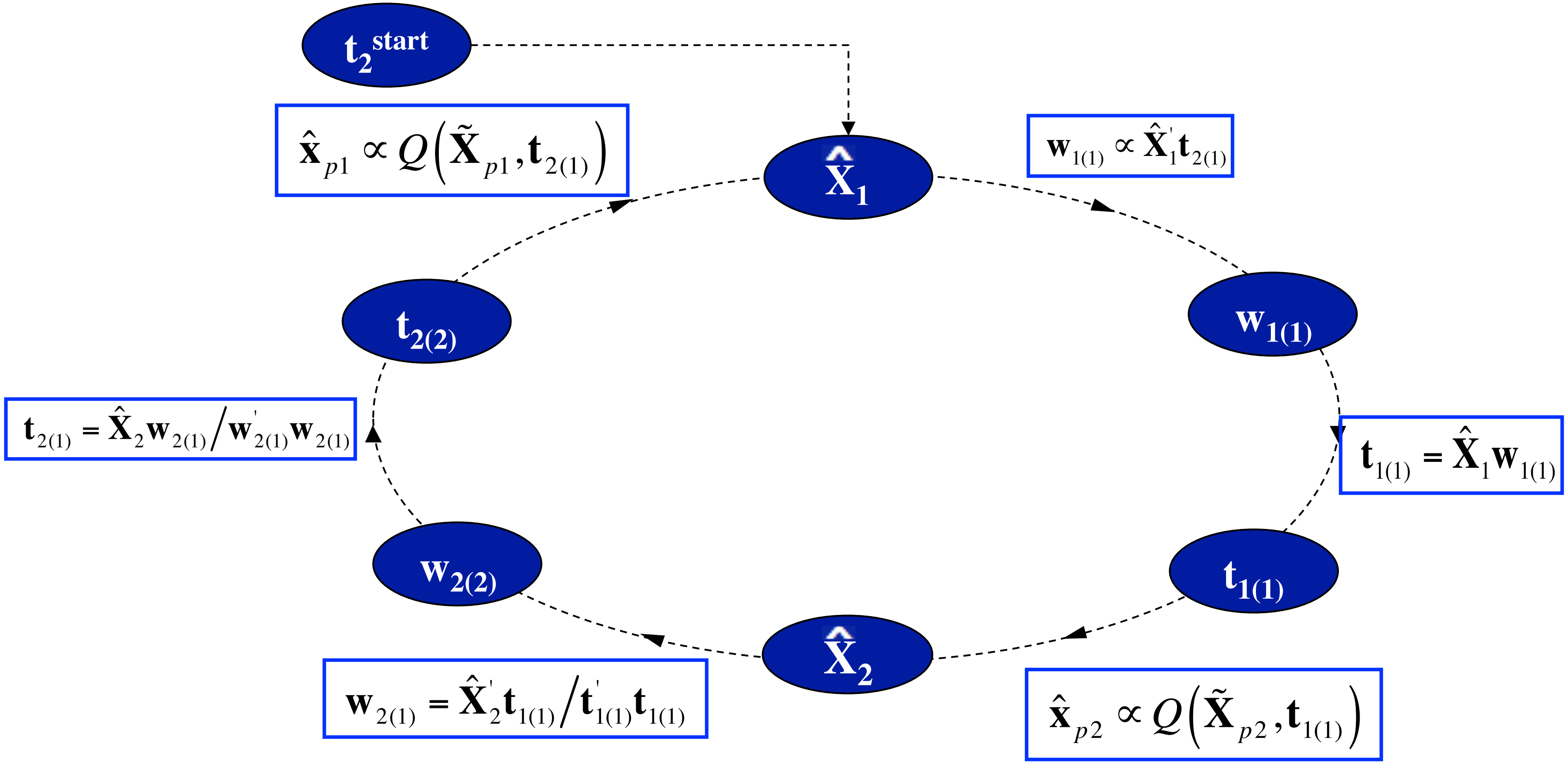
$e_{qq'}$ is the generic element of a squared matrix of order Q that is null if $x_{q'}$ is connected to x_q; otherwise, it represents the corresponding inner weight.

le cnam

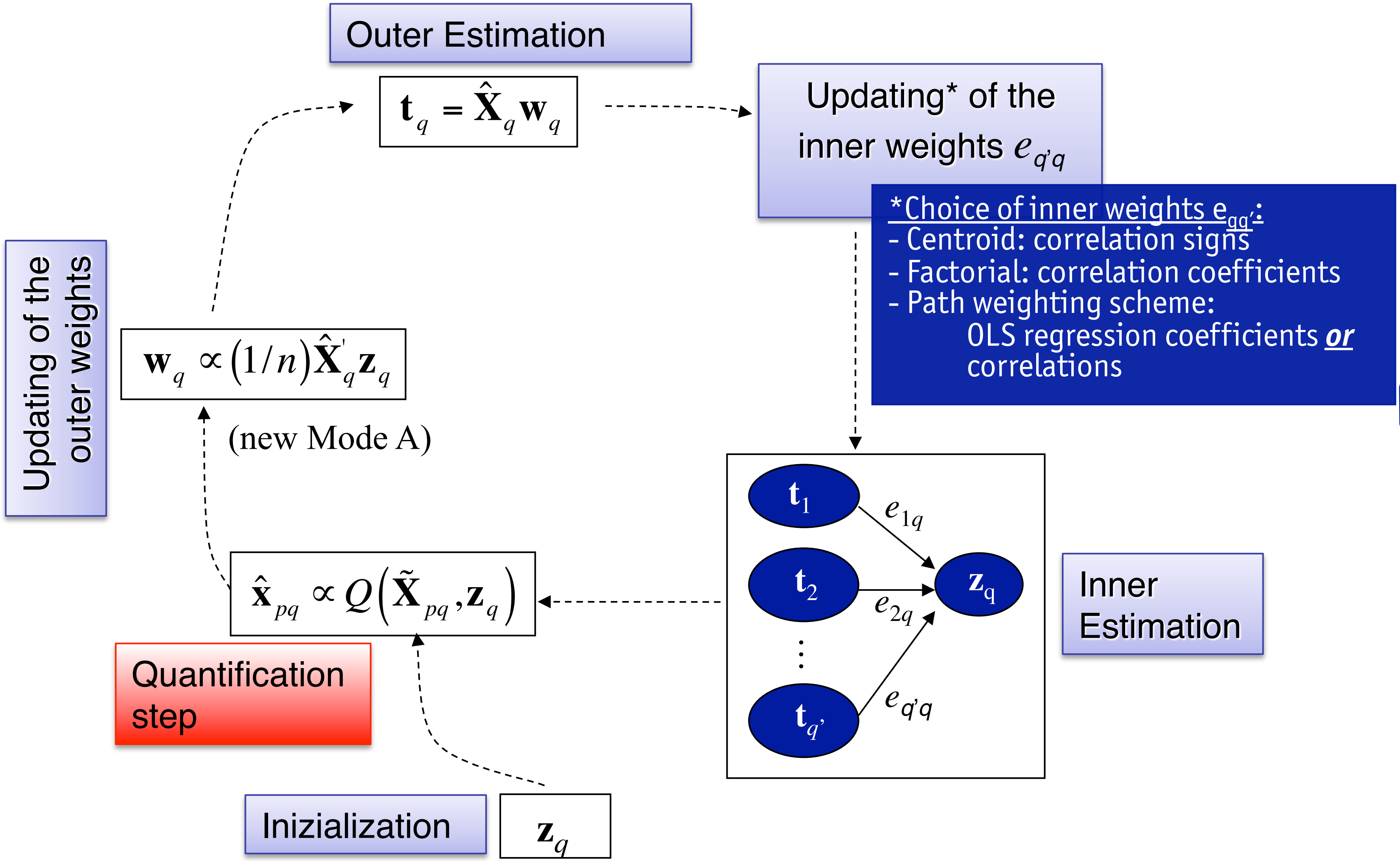
NM-NIPALS loop



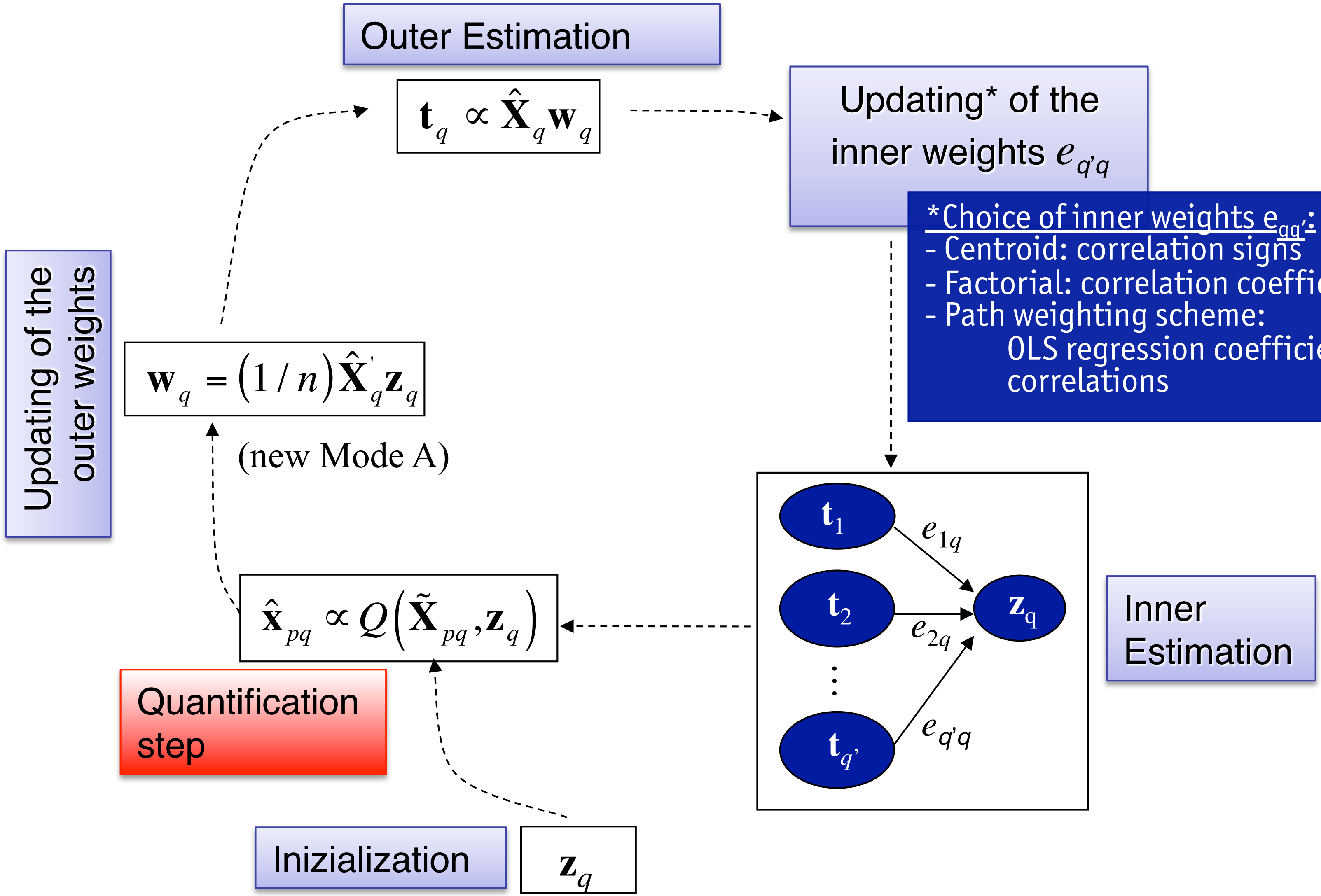
NM-PLSR loop


t_2 start
$\hat{\mathbf{x}}_{p1} \propto Q\left(\tilde{\mathbf{X}}_{p1}, \mathbf{t}_{2(1)}\right)$
$\hat{X}_1$
$\mathbf{w}_{1(1)} \propto \hat{\mathbf{X}}_1^{'}\mathbf{t}_{2(1)}$
$w_{1(1)}$
$\mathbf{t}_{1(1)} = \hat{\mathbf{X}}_1\mathbf{w}_{1(1)}$
$t_{1(1)}$
$\hat{\mathbf{x}}_{p2} \propto Q\left(\tilde{\mathbf{X}}_{p2}, \mathbf{t}_{1(1)}\right)$
$\hat{X}_2$
$\mathbf{w}_{2(1)} = \hat{\mathbf{X}}_2^{'}\mathbf{t}_{1(1)} / \mathbf{t}_{1(1)}^{'}\mathbf{t}_{1(1)}$
$w_{2(2)}$
$\mathbf{t}_{2(1)} = \hat{\mathbf{X}}_2\mathbf{w}_{2(1)} / \mathbf{w}_{2(1)}^{'}\mathbf{w}_{2(1)}$
$t_{2(2)}$

NM-PLSPM loop (New Mode A)



NM-PLSPM loop (Mode A)



*Choice of inner weights $e_{qq'}$:
- Centroid: correlation signs
- Factorial: correlation coefficients
- Path weighting scheme:
  OLS regression coefficients ***or*** correlations

le cnam

Non-Metric PLS features

Optimality: NM-PLS algorithms optimizes NIPALS, PLS-R and New Mode A PLS-PM criteria under two sets of parameters: the model parameters and the scaling parameters, constrained to the restrictions due to the scaling level chosen for each raw variable $\mathbf{x}^*$

Flexibility: NM-PLS satisfies model criterion under several levels of scaling, depending on the properties of the raw variable we want to retain.

Interpretability: The weight of a scaled variable can be interpreted as a measure of the statistical relation between $\mathbf{x}^*$ and the LV

Sensorial data: the Tea dataset

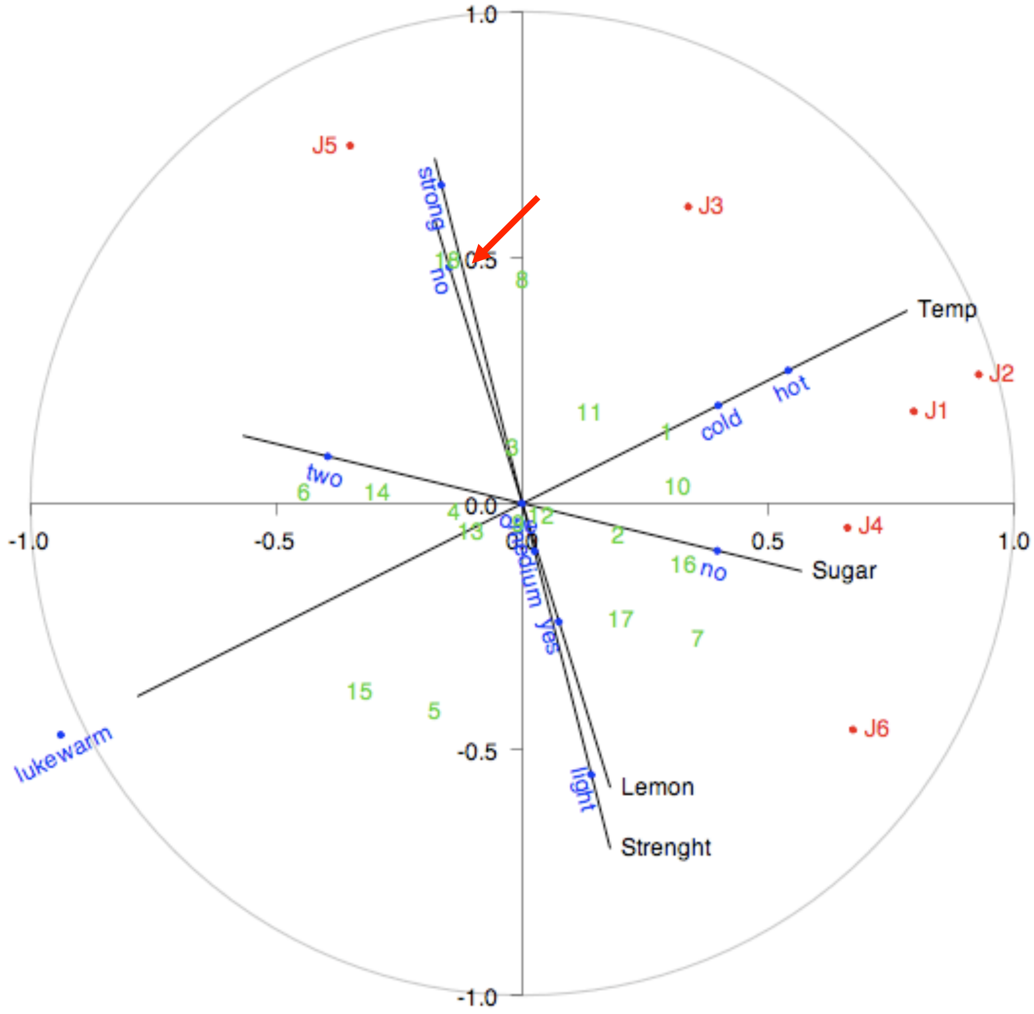
Obs	J1	J2	J3	J4	J5	J6	Temp	Sugar	Strength	Lemon
1	4	2	4	3	13	5	Hot	No	Str	Y
2	2	8	1	9	10	8	Hot	One	Med	Y
3	6	10	13	18	5	6	Hot	Two	Light	N
4	13	13	10	5	2	12	LW	No	Med	N
5	14	16	17	12	16	9	LW	One	Light	Y
6	15	18	12	15	8	16	LW	Two	Str	Y
7	7	3	14	2	18	2	Cold	No	Light	Y
8	11	6	5	7	3	17	Cold	One	Str	N
9	10	11	6	13	12	7	Cold	Two	Med	Y
10	3	1	11	4	6	4	Hot	No	Light	N
11	1	7	2	10	7	14	Hot	One	Str	Y
12	5	12	3	17	9	13	Hot	Two	Med	Y
13	17	14	16	6	11	18	LW	No	Str	Y
14	18	15	9	11	1	11	LW	One	Med	N
15	16	17	18	16	15	10	LW	Two	Light	Y
16	8	4	8	1	14	1	Cold	No	Med	Y
17	9	5	15	8	17	3	Cold	One	Light	Y
18	12	9	7	14	4	15	Cold	Two	Str	N

The regression coefficients matrix

Var	J1	J2	J3	[illegible]	J5	J6
Temp(H)	-0.35	[illegible]	-0.35	[illegible]	-0.15	-0.08
Temp(LW)	0.49	0.51	[illegible]	[illegible]	-0.01	0.26
Temp(C)	-0.15	-0.18	[illegible]	[illegible]	0.15	-0.18
Sugar(0)	-0.21	-0.27	[illegible]	-0.22	0.23	-0.26
Sugar(1)	-0.03	-0.02	[illegible]	0.01	-0.08	0.04
Sugar(2)	0.25	0.29	[illegible]	0.21	-0.15	0.23
Strenght(S)	0.01	0.05	[illegible]	0.10	-0.24	0.16
Strenght(M)	-0.07	[illegible]	[illegible]	0.01	-0.14	0.05
Strenght(L)	0.07	[illegible]	[illegible]	[illegible]	0.38	-0.21
Lemon(Y)	-0.00	[illegible]	[illegible]	[illegible]	0.29	-0.19
Lemon(N)	0.00	[illegible]	-0.18	[illegible]	-0.29	0.19

Var	**J1**	**J2**	**J3**	**J4**	**J5**	**J6**
Temp	0.70	0.83	0.50	0.5	0.01	0.35
Sugar	0.43	0.49	0.11	0.38	-0.30	0.45
Strenght	0.01	-0.02	-0.36	-0.15	-0.57	0.44
Lemon	0.03	0.01	-0.29	-0.15	-0.48	0.39

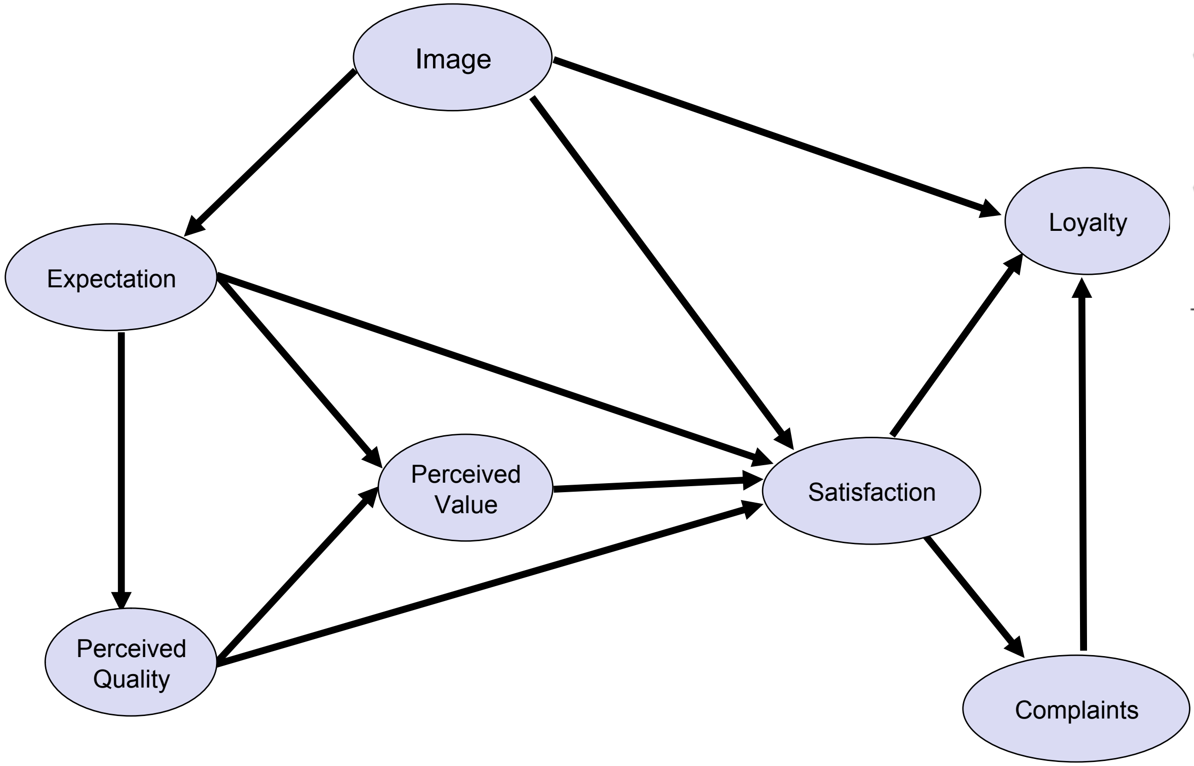
Mapping the observations on the Correlation Circle: the Biplot



Mobile Data – ECSI model

All the items are from 1 to 10.
1 expresses a very negative point of view on the product while 10 a very positive opinion.

Latent variables	Manifest variables
Image (ξ_1)	(a) It can be trusted in what it says and does (b) It is stable and firmly established (c) It has a social contribution for the society (d) It is concerned with customers (e) It is innovative and forward looking
Customer expectations of the overall quality (ξ_2)	(a) Expectations for the overall quality of "your mobile phone provider" at the moment you became customer of this provider (b) Expectations for "your mobile phone provider" to provide products and services to meet your personal need (c) How often did you expect that things could go wrong at "your mobile phone provider"
Perceived quality (ξ_3)	(a) Overall perceived quality (b) Technical quality of the network (c) Customer service and personal advice offered (d) Quality of the services you use (e) Range of services and products offered (f) Reliability and accuracy of the products and services provided (g) Clarity and transparency of information provided
Perceived value (ξ_4)	(a) Given the quality of the products and services offered by "your mobile phone provider" how would you rate the fees and prices that you pay for them? (b) Given the fees and prices that you pay for "your mobile phone provider" how would you rate the quality of the products and services offered by "your mobile phone provider"?
Customer satisfaction (ξ_5)	(a) Overall satisfaction (b) Fulfillment of expectations (c) How well do you think "your mobile phone provider" compares with your ideal mobile phone provider?
Customer complaints (ξ_6)	(a) You complained about "your mobile phone provider" last year. How well, or poorly, was your most recent complaint handled or (b) You did not complain about "your mobile phone provider" last year. Imagine you have to complain to "your mobile phone provider" because of a bad quality of service or product. To what extent do you think that "your mobile phone provider" will care about your complaint?
Customer loyalty (ξ_7)	(a) If you would need to choose a new mobile phone provider how likely is it that you would choose "your provider" again? (b) Let us now suppose that other mobile phone providers decide to lower their fees and prices, but "your mobile phone provider" stays at the same level as today. At which level of difference (in %) would you choose another mobile phone provider? (c) If a friend or colleague asks you for advice, how likely is it that you would recommend "your mobile phone provider"?



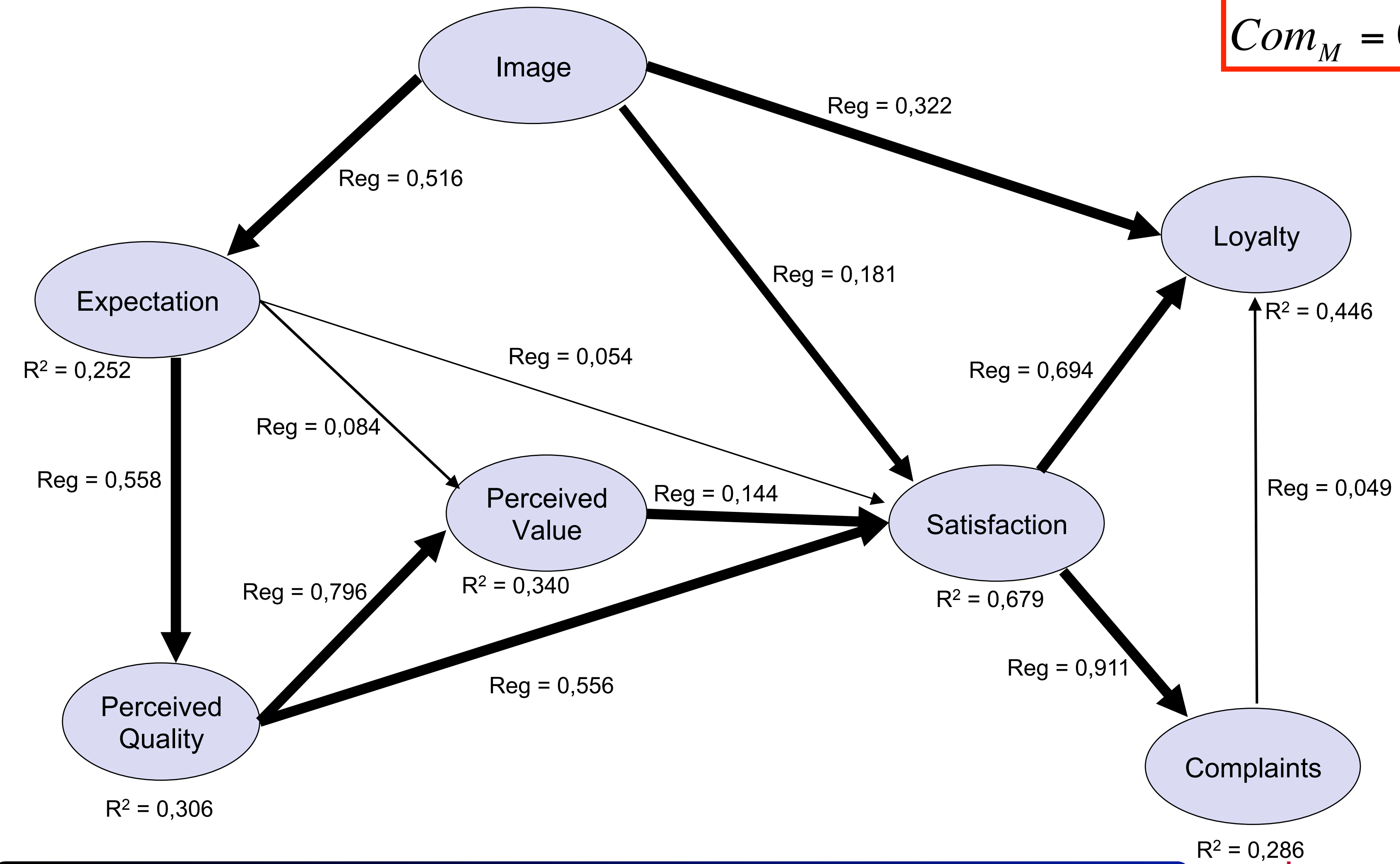
-> Standardized MVs

-> All reflective blocks

-> Centroid scheme has been used

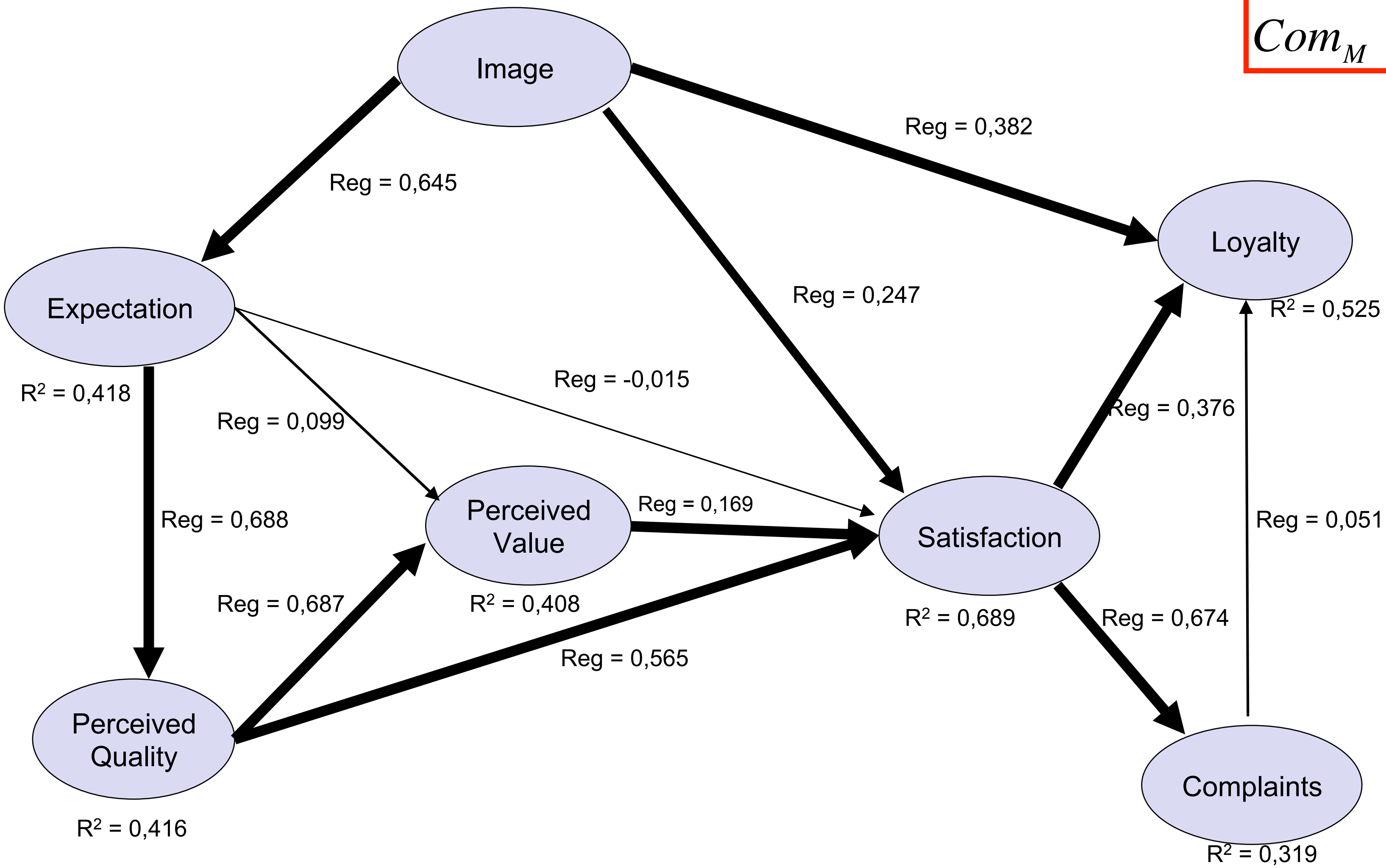
Model with standard linearity hypothesis

$$GoF = 0.471$$
$$\overline{R}^2 = 0.387$$
$$Com_M = 0.574$$



NM-PLSPM (ordinal scaling)

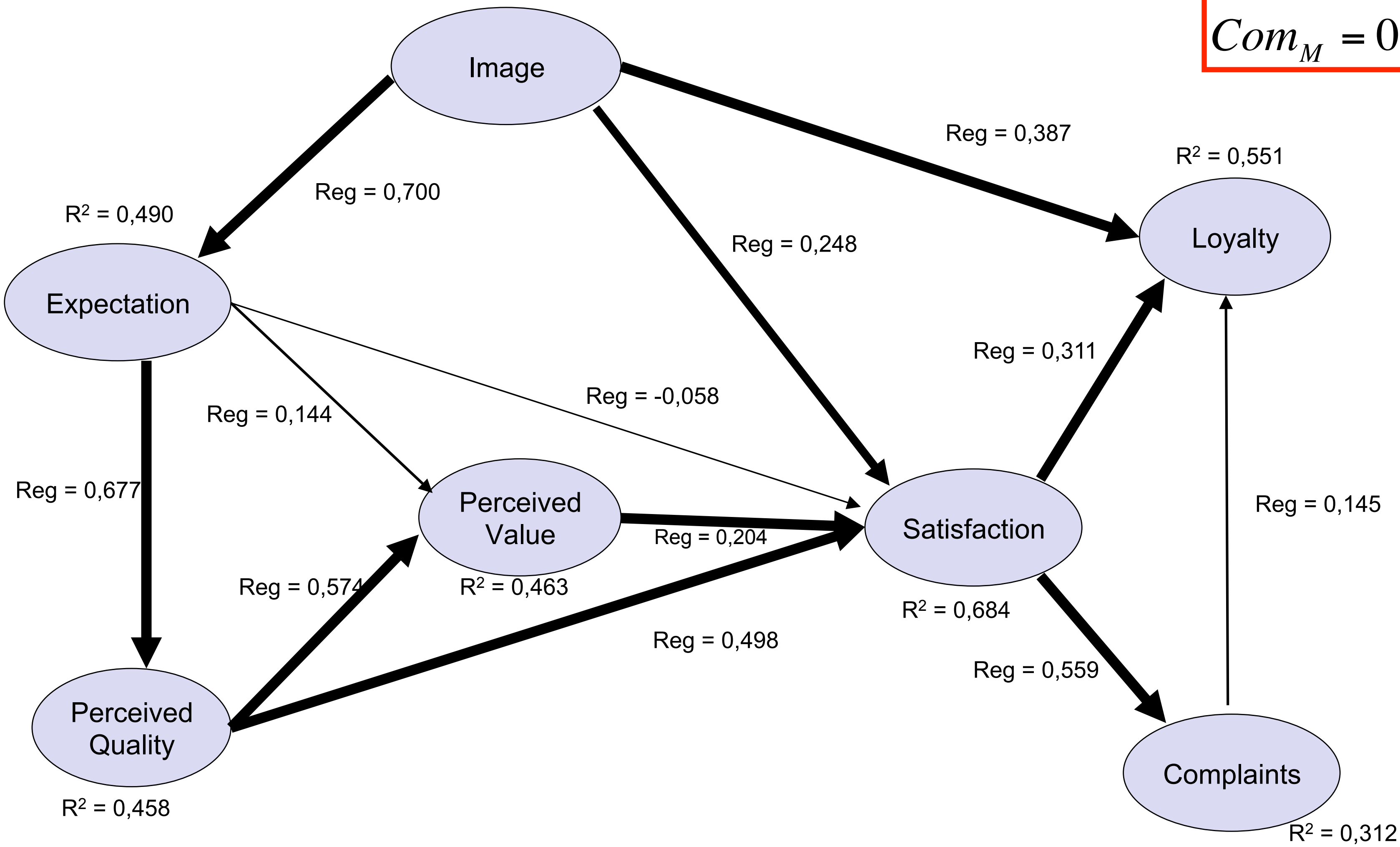
$$GoF = 0.526$$
$$\overline{R}^2 = 0.464$$
$$Com_M = 0.597$$



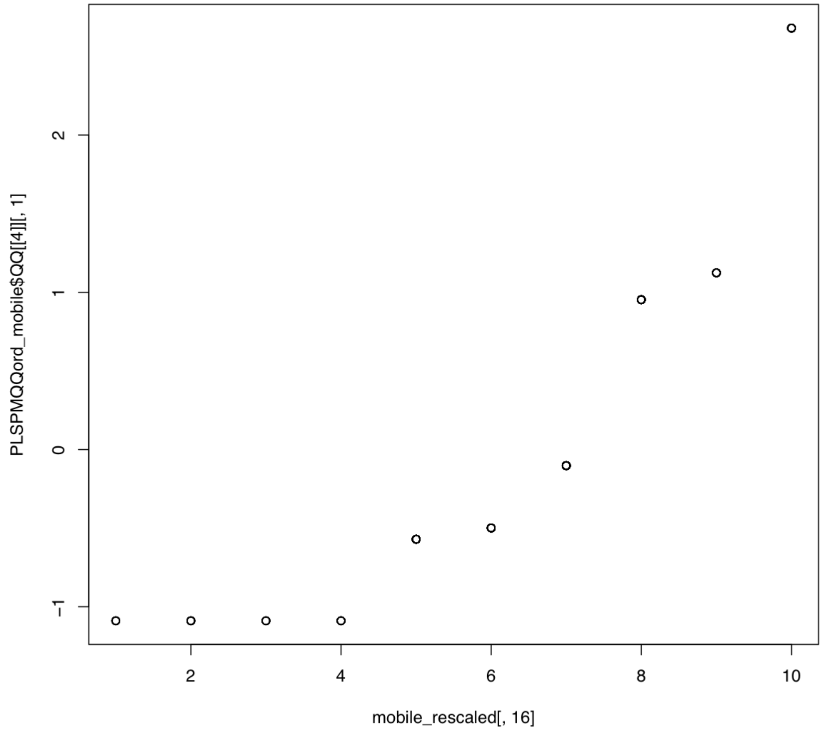
le cnam

NM-PLSPM (nominal scaling)

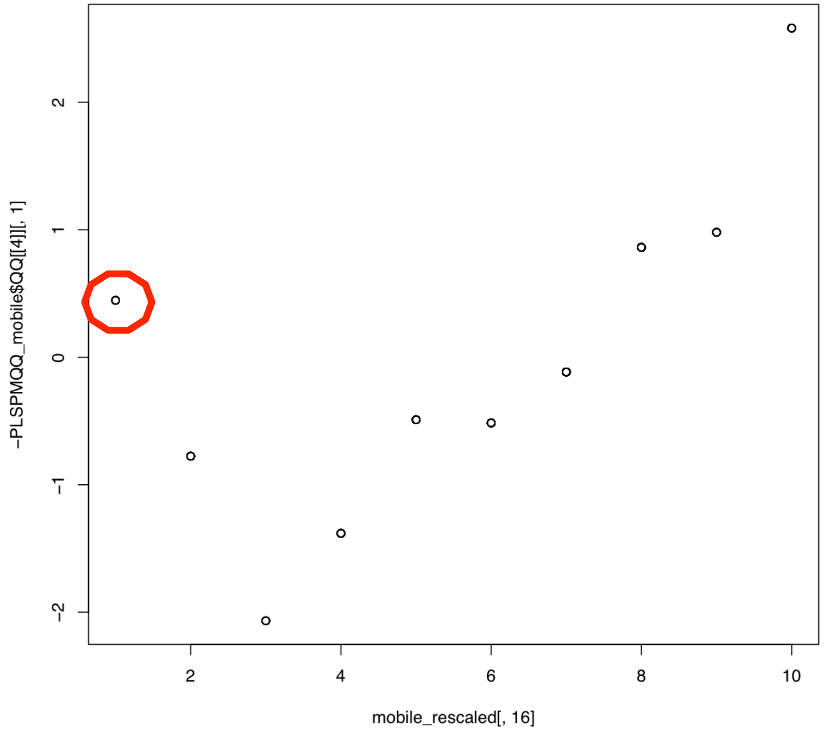
$$GoF = 0.546$$
$$\bar{R}^2 = 0.493$$
$$Com_M = 0.605$$



Ordinal and Nominal Quantification: the case of "Ratio Quality-Prize" variable



Quantification with ordinal constraints



Quantification without ordinal constraints

le cnam

Flashback

Power methods involve a suite of iterative algorithms, well known from the first half of the 20th century, used for extracting the greatest eigenvalue (in absolute value) of a matrix.

Two classes of algorithms of this type have been greatly used in Statistics to solve different problems: PLS algorithms and ALSOS algorithms.

(H. Wold)

A suite of algorithms (NIPALS, PLS-R, PLS-PM, etc.) aiming to:

- easily handle in presence of missing data
- Perform multidimensional analyses of landscape matrices
- perform a regression analysis of data affected by multicollinearity problems
- estimate SEM parameters without making strong distributional assumptions (PLS-PM)

(J. de Leeuw)

A suite of algorithms (Princals, Ascal, Morals, Corals, Overals, etc.) for multivariate analysis on one, two or several blocks of variables measured at different scale levels.

le cnam

Conclusion

Quoting F. W. Young (Psychometrika, vol. 46, n. 4, 1981):

"Certain strong correspondences exist between an ALSOS procedure and the NILES approach developed by Wold and Lyttkens.

The main difference between these metric algorithms and the nonmetric ALSOS algorithms is the optimal scaling features of the ALSOS algorithm.

The scaling feature permits the analysis of qualitative data, whereas the previous procedures can only analyze quantitative data."

This ultimate statement is no more true!

It is possible to exploit the iterative nature of PLS algorithms to transform PLS methods in optimal scaling methods

Thank you for your attention!

Main References (I)

1. **Amato, V**., *Statistica*, Cacucci Editore. vol. 1
2. **Bollen K. A.**, *Structural equations with latent variables*, Wiley, New York (1989).
3. **Benzécri J.P.**, *Analyse des données*, Dunod, Paris (1973)
4. **Breiman L., Friedman J**., Estimating Optimal Transformations for Multiple Regression and Correlation. *Journal of the American Statistical Association*, Vol. 80, No. 391, 580-598 (1985).
5. **De Leeuw, J., Young, F. W., and Takane, Y**., Additive Structure in Qualitative Data: An Alternating Least Squares Method With Optimal Scaling Features. *Psychometrika*, 41, 471-503 (1976).
6. **Gifi, A**., *Nonlinear Multivariate Analysis*, Chichester, UK: Wiley (1990).
7. **Hayashi C.**, On the quantification of qualitative data from the mathematicostatiscal point of view". *Ann. Inst. Statist. Math*. 2: 35 (1950).
8. **Hayashi C.**, On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematicostatistical point of view. *Ann. Inst. Statist. Math.* 3: 69 (1952).
9. **Kruskal, J**., Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika* 29 (1), 1-27 (1964)
10. **Michailidis G., de Leeuw J**., The Gifi System of Descriptve Multivariate Analysis, Statistical Science, 13:307-336 (1998).
11. **Richardson, M. & Kuder, G. F**., Making a rating scale that measures, *Personnel Journal* 12, 36-40 (1933)
12. **Russolillo G.**, A proposal for handling categorical predictors in PLS regression framework. in *First joint meeting of the Socit Francophone de Classification and the Classification and Data Analysis Group of the Italian Statistical Society. Book of short papers*, Edizioni Scientifiche Italiane, 401 (2008).

Main References (II)

13. **Russolillo G.**, PLS Methods for Non-Metric Data, PhD thesis, Università degli Studi di Napoli "Federico II" (2009).

14. **Tenenhaus M., Esposito Vinzi V., Chatelin Y.-M. and Lauro C.**, PLS Path Modeling, *Computational Statistics and Data Analysis*, 48, 159-205 (2005).

15. **Tenenhaus M.,** *La Régression PLS*, Technip, Paris (1998)

16. **Tenenhaus M**. (2009), A criterion based PLS approach to structural equation modelling, in *Programme and Abstract of the 6th International Conference on Partial Least Squares Methods*, p. 3

17. **Wold H**., Nonlinear Estimation by Iterative Least Squares Procedures, in F.David (Editor), *Research Papers in Statistics*, Wiley, New York, pp. 411- 444 (1966).

18**. Wold S., Albano C., Dunn III W.J., Edlund U., Eliasson B., Johansson E., Norden B., Sjöström M.**, The indirect observation of molecular chemical systems, in: *Systems Under Indirect Observation* (vols. I and II), K.G. Jöreskog & H. Wold (Eds.), North-Holland, Amsterdam, Chapter 8, 177–207 (1982) .

19. **Wold. H.**, Path Models with latent variables: the NIPALS approach, in*: Quantitative sociology: international perspectives on mathematical and statistical modeling*, H.M. Blalock et al. (Eds), Academic, 307-357 (1975).

20. **Wold. H.**, Soft modeling: the basic design and some extensions, in: *Systems under indirect observation*, K.G. Joreskog & H. Wold (Eds), North-Holland, Amsterdam, 1–54 (1982).

21**. Wold, S., Martens, H. & Wold, H**. The multivariate calibration-problem in chemistry solved by the PLS method. Lecture Notes in Mathematics 973, 286–293 (1983).

22. **Young, F**. [1975], Methods for describing ordinal data with cardinal models*, Journal of Mathematical Psychology* 12, 416-436

23. **Young F. W., Jan de Leeuw J. and Takane Y**., Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features. *Psychometrika* ,Vol. 41, No. 4, 505-529 (1976).

24. **Young F. W.,** Quantitative analysis of qualitative data. *Psychometrika*, vol. 46, n. 4, (1981).

le cnam