Tensor Generalized Canonical Correlation Analysis

Séminaire de statistique appliquée du CNAM

F. Girka, A. Gloaguen, L. Le Brusquet,
V. Zujovic & A. Tenenhaus

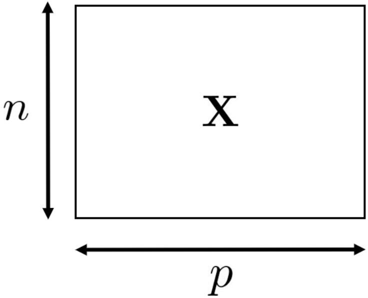
Laboratoire des Signaux et Systèmes

27 octobre 2023

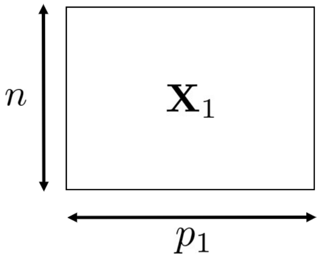
Table of Contents

Table of Contents

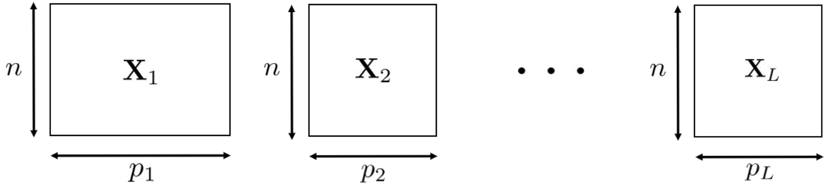
Most common type of data



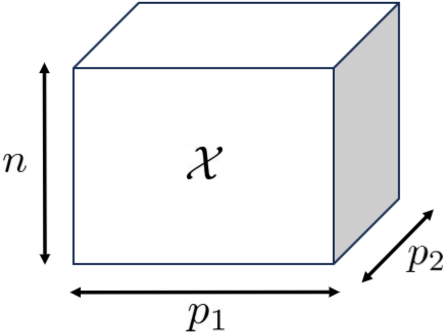
Multiblock data: 2 blocks



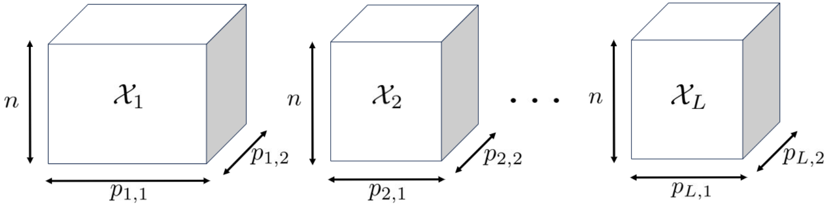
Multiblock data: L blocks



Tensor data



Multiblock tensor data: L blocks



The example of multiple sclerosis

Data from the Paris Brain Institute

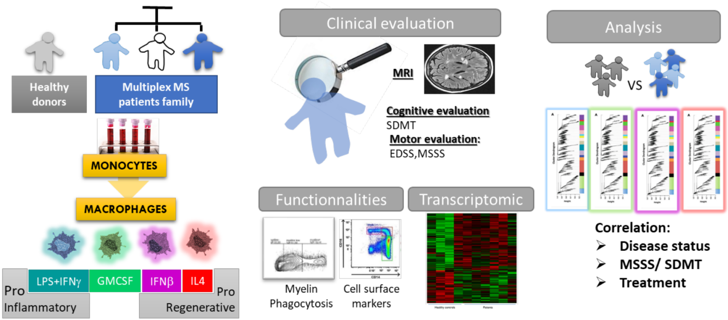


Figure: Graphical abstract from Fransson et al. (2021)

Our goal

Objective

Summarize the joint information between tensor blocks of data (collected on n individuals).

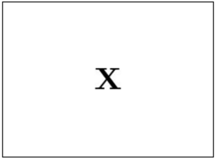
Table of Contents

Principal Component Analysis

PCA

Summarize the variance present in a block of data.

$$\underset{\mathbf{a}}{\text{maximize}} \ \text{Var}(\mathbf{X}\mathbf{a}) \quad \text{s.t.} \quad \|\mathbf{a}\| = 1. \tag{1}$$

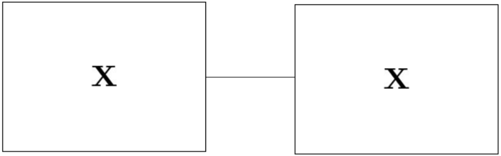


Principal Component Analysis

PCA

Summarize the variance present in a block of data.

$$\underset{\mathbf{a}}{\text{maximize}} \ \mathrm{Cov}(\mathbf{Xa}, \mathbf{Xa}) \quad \text{s.t.} \quad \|\mathbf{a}\| = 1. \tag{2}$$

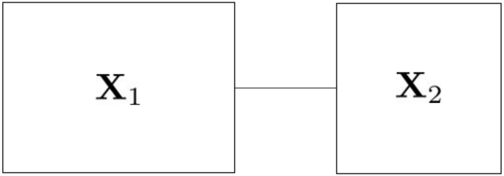


Partial Least Squares

PLS

Summarize the covariance between two blocks of data.

$$\underset{\mathbf{a}_1,\mathbf{a}_2}{\text{maximize}} \ \mathrm{Cov}(\mathbf{X}_1\mathbf{a}_1, \mathbf{X}_2\mathbf{a}_2) \quad \text{s.t.} \quad \begin{cases} \|\mathbf{a}_1\| = 1, \\ \|\mathbf{a}_2\| = 1. \end{cases} \tag{3}$$

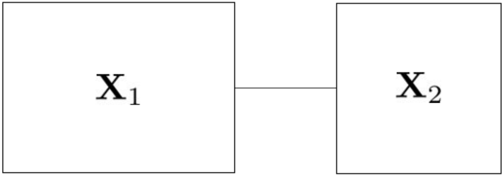


Canonical Correlation Analysis

CCA

Summarize the correlation between two blocks of data.

$$\underset{\mathbf{a}_1,\mathbf{a}_2}{\text{maximize}}\ \text{Cov}(\mathbf{X}_1\mathbf{a}_1, \mathbf{X}_2\mathbf{a}_2) \quad \text{s.t.} \quad \begin{cases} \text{Var}(\mathbf{X}_1\mathbf{a}_1) = 1, \\ \text{Var}(\mathbf{X}_2\mathbf{a}_2) = 1. \end{cases} \tag{4}$$

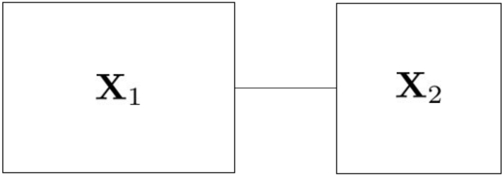


Regularized Canonical Correlation Analysis

RCCA

Summarize the joint information between two blocks of data.

$$\underset{\mathbf{a}_1,\mathbf{a}_2}{\text{maximize}}\ \text{Cov}(\mathbf{X}_1\mathbf{a}_1, \mathbf{X}_2\mathbf{a}_2) \text{ s.t. } \begin{cases} \tau_1\|\mathbf{a}_1\|^2 + (1-\tau_1)\text{Var}(\mathbf{X}_1\mathbf{a}_1) = 1, \\ \tau_2\|\mathbf{a}_2\|^2 + (1-\tau_2)\text{Var}(\mathbf{X}_2\mathbf{a}_2) = 1. \end{cases} \quad (5)$$



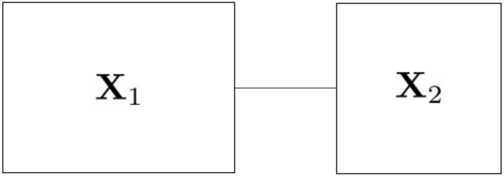
Regularized Canonical Correlation Analysis

RCCA

Summarize the joint information between two blocks of data.

$$\underset{\mathbf{a}_1,\mathbf{a}_2}{\text{maximize}} \ \mathrm{Cov}(\mathbf{X}_1\mathbf{a}_1, \mathbf{X}_2\mathbf{a}_2) \quad \text{s.t.} \quad \begin{cases} \mathbf{a}_1^\top \mathbf{M}_1 \mathbf{a}_1 = 1, \\ \mathbf{a}_2^\top \mathbf{M}_2 \mathbf{a}_2 = 1. \end{cases} \tag{6}$$

$\mathbf{M}_l$ is a symmetric positive definite matrix.

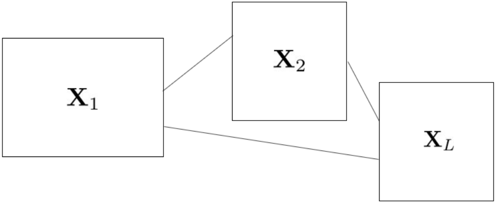


Regularized Generalized Canonical Correlation Analysis

RGCCA

Summarize the joint information between blocks of data.

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}} \sum_{l,k=1}^{L} \text{Cov}(\mathbf{X}_l\mathbf{a}_l, \mathbf{X}_k\mathbf{a}_k) \quad \text{s.t.} \quad \mathbf{a}_l^\top \mathbf{M}_l \mathbf{a}_l = 1. \tag{7}$$

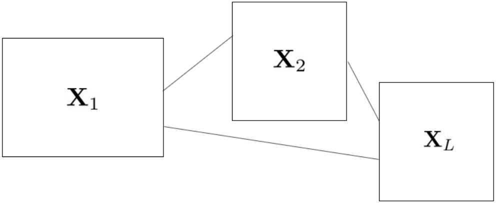


Regularized Generalized Canonical Correlation Analysis

RGCCA

Summarize the joint information between blocks of data.

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}} \sum_{l,k=1}^{L} c_{kl}\text{Cov}(\mathbf{X}_l\mathbf{a}_l, \mathbf{X}_k\mathbf{a}_k) \quad \text{s.t.} \quad \mathbf{a}_l^\top \mathbf{M}_l\mathbf{a}_l = 1. \tag{8}$$



Regularized Generalized Canonical Correlation Analysis

RGCCA

Summarize the joint information between blocks of data.

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}} \quad \sum_{l,k=1}^{L} c_{kl}\,g\big(\mathrm{Cov}(\mathbf{X}_l\mathbf{a}_l, \mathbf{X}_k\mathbf{a}_k)\big) \quad \text{s.t.} \quad \mathbf{a}_l^\top \mathbf{M}_l \mathbf{a}_l = 1. \qquad (9)$$

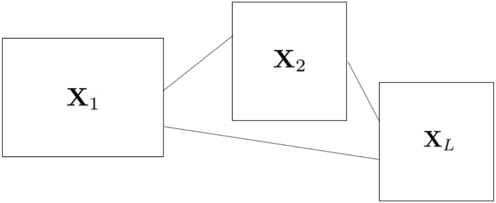


Table of Contents

Tensor Generalized Canonical Correlation Analysis

TGCCA

Summarize the joint information between tensor blocks of data.

- ☐ What is the equivalent of $\mathbf{X}_l \mathbf{a}_l$?
- ☐ What kind of constraints on $\mathbf{a}_l$?

We could just unfold the tensors and apply RGCCA!

Matrix rank

Let $\mathbf{X} \in \mathbb{R}^{p_1 \times p_2}$ be a rank-R matrix,
There exists R triplets $\{(\lambda^{(r)}, \mathbf{a}_1^{(r)}, \mathbf{a}_2^{(r)}) \in \mathbb{R}^{1 \times p_1 \times p_2}\}_{r=1}^{R}$ with $\|\mathbf{a}_1^{(r)}\| = \|\mathbf{a}_2^{(r)}\| = 1$ such that:

$$\mathbf{X} = \sum_{r=1}^{R} \lambda^{(r)} \mathbf{a}_1^{(r)} \mathbf{a}_2^{(r)\top}. \tag{10}$$

Or equivalently,

$$\mathrm{Vec}(\mathbf{X}) = \sum_{r=1}^{R} \lambda^{(r)} \mathbf{a}_2^{(r)} \otimes \mathbf{a}_1^{(r)}. \tag{11}$$

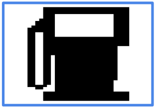
Tensor rank

Let $\mathcal{X} \in \mathbb{R}^{p_1 \times \cdots \times p_d}$ be a rank-R tensor,
There exists R tuples $\{(\lambda^{(r)}, \mathbf{a}_1^{(r)}, \ldots, \mathbf{a}_d^{(r)}) \in \mathbb{R}^{1 \times p_1 \times \cdots \times p_d}\}_{r=1}^{R}$ with $\|\mathbf{a}_1^{(r)}\| = \cdots = \|\mathbf{a}_d^{(r)}\| = 1$ such that:

$$\mathrm{Vec}(\mathcal{X}) = \sum_{r=1}^{R} \lambda^{(r)} \mathbf{a}_d^{(r)} \otimes \cdots \otimes \mathbf{a}_1^{(r)}. \tag{12}$$

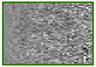
Why should we go further?

Structure of the matrix

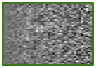


Noisy observations of the matrix

X_1

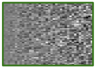


X_2

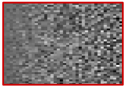


.
.
.

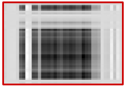
X_n



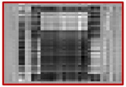
Structure extracted by different methods



“Classical” method: too high degree of freedom, noise is modeled



Rank-1 CP decomposition: (too) simplified extracted structure



Rank-3 CP decomposition: richer model with still a reasonable degree of freedom (compromise)

Tensor Generalized Canonical Correlation Analysis

TGCCA

Summarize the joint information between tensor blocks of data.

- Maximize the criterion of RGCCA,
- With a low-rank tensor structure applied on $\mathbf{a}_l$.

We also add orthogonality constraints since the set of tensors with rank at most R is not closed!

Tensor Generalized Canonical Correlation Analysis

TGCCA

Summarize the joint information between tensor blocks of data.

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}} \sum_{l,k=1}^{L} c_{lk} \mathrm{g}\left(\mathrm{Cov}(\mathbf{X}_l\mathbf{a}_l, \mathbf{X}_k\mathbf{a}_k)\right) \tag{13}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{a}_l^\top \mathbf{M}_l \mathbf{a}_l = 1, \\ \mathbf{a}_l = \sum_{r=1}^{R_l} \lambda_l^{(r)} \mathbf{a}_{l,d_l}^{(r)} \otimes \cdots \otimes \mathbf{a}_{l,1}^{(r)}, \\ \mathbf{a}_{l,1}^{(r)\top} \mathbf{K}_l \mathbf{a}_{l,1}^{(s)} = 0 \text{ for } r \neq s. \end{cases}$$

$\mathbf{X}_l$ is the mode-1 matricization of $\boldsymbol{\mathcal{X}}_l$.

$\mathbf{K}_l$ is a symmetric positive definite matrix.

Models highly related to TGCCA

MGCCA (Gloaguen et al., 2020)

Summarize the joint information between tensor blocks of data.

$$\underset{\mathbf{a}_1,\dots,\mathbf{a}_L}{\text{maximize}} \sum_{l,k=1}^{L} c_{lk}\,\mathrm{g}\left(\mathrm{Cov}(\mathbf{X}_l\mathbf{a}_l, \mathbf{X}_k\mathbf{a}_k)\right) \tag{14}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{a}_l^\top \mathbf{M}_l \mathbf{a}_l = 1, \\ \mathbf{a}_l = \lambda_l \mathbf{a}_{l,d_l} \otimes \cdots \otimes \mathbf{a}_{l,1}. \end{cases}$$

Pros: L blocks, RGCCA criterion,
Cons: Rank-1 tensors, matrices $\mathbf{M}_l$ separable.

Models highly related to TGCCA

TCCA (Min et al., 2019)

Summarize the correlation between two tensor blocks of data.

$$\underset{\mathbf{a}_1, \mathbf{a}_2}{\text{maximize}} \ \text{Cov}(\mathbf{X}_1 \mathbf{a}_1, \mathbf{X}_2 \mathbf{a}_2) \tag{15}$$

$$\text{s.t.} \quad \begin{cases} \text{Var}(\mathbf{X}_1 \mathbf{a}_1) = 1, \\ \text{Var}(\mathbf{X}_2 \mathbf{a}_2) = 1, \\ \mathbf{a}_1 = \sum_{r=1}^{R_1} \lambda_1^{(r)} \mathbf{a}_{1,d_1}^{(r)} \otimes \cdots \otimes \mathbf{a}_{1,1}^{(r)}, \\ \mathbf{a}_2 = \sum_{r=1}^{R_2} \lambda_2^{(r)} \mathbf{a}_{2,d_2}^{(r)} \otimes \cdots \otimes \mathbf{a}_{2,1}^{(r)}. \end{cases}$$

Pros: Rank-R tensors, no orthogonality.
Cons: Two blocks, CCA criterion, no orthogonality.

Separability

Separable matrix

A matrix $\mathbf{M}_l \in \mathbb{R}^{p_l \times p_l}$ is said to be separable if there exists submatrices $\{\mathbf{M}_{l,m} \in \mathbb{R}^{p_{l,m} \times p_{l,m}}\}_{m=1}^{d_l}$ such that:

$$\mathbf{M}_l = \mathbf{M}_{l,d_l} \otimes \cdots \otimes \mathbf{M}_{l,1},$$

$$\prod_{m=1}^{d_l} p_{l,m} = p_l.$$

Simplification when $\mathbf{M}_l$ is separable

If $\mathbf{M}$ is separable and $\mathbf{a} = \sum_{r=1}^{R} \lambda^{(r)} \mathbf{a}_d^{(r)} \otimes \cdots \otimes \mathbf{a}_1^{(r)}$, we have:

$$\begin{aligned}
\mathbf{a}^\top \mathbf{M} \mathbf{a} &= \sum_{r,s=1}^{R} \lambda^{(r)} \lambda^{(s)} (\mathbf{a}_d^{(r)} \otimes \cdots \otimes \mathbf{a}_1^{(r)})^\top (\mathbf{M}_d \otimes \cdots \otimes \mathbf{M}_1)(\mathbf{a}_d^{(s)} \otimes \cdots \otimes \mathbf{a}_1^{(s)}), \\
&= \sum_{r,s=1}^{R} \lambda^{(r)} \lambda^{(s)} ({\mathbf{a}_d^{(r)}}^\top \mathbf{M}_d \mathbf{a}_d^{(s)}) \ldots ({\mathbf{a}_1^{(r)}}^\top \mathbf{M}_1 \mathbf{a}_1^{(s)}).
\end{aligned}$$

We choose $\mathbf{K} = \mathbf{M}_1$ to remove the crossed terms.

TGCCA for separable regularization matrices

TGCCA for separable matrices

By making a change of variable and choosing $\mathbf{K}_l = \mathbf{M}_{l,1}$, we get:

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}} \sum_{l,k=1}^{L} c_{lk} \mathrm{g}\left(\mathrm{Cov}(\mathbf{X}_l^* \mathbf{a}_l, \mathbf{X}_k^* \mathbf{a}_k)\right) \tag{16}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{a}_l = \sum_{r=1}^{R_l} \lambda_l^{(r)} \mathbf{a}_{l,d_l}^{(r)} \otimes \cdots \otimes \mathbf{a}_{l,1}^{(r)}, \\ {\mathbf{a}_{l,1}^{(r)}}^\top \mathbf{a}_{l,1}^{(s)} = 0 \text{ for } r \neq s, \\ \|\mathbf{a}_{l,m}^{(r)}\| = 1 \text{ for } m \in \{1, \ldots, d_l\}, \\ \|\boldsymbol{\lambda}_l\| = 1. \end{cases}$$

$\mathbf{X}_l^* = \mathbf{M}_l^{-\frac{1}{2}} \mathbf{X}_l$.

Table of Contents

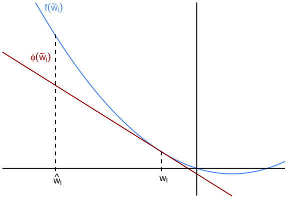
Optimization procedure

Block Coordinate Ascent (BCA, de Leeuw, 1994)

$$\underset{\mathbf{a}_1,\ldots,\mathbf{a}_L}{\text{maximize}}\ f(\mathbf{a}_1,\ldots,\mathbf{a}_L) \quad \text{s.t.} \quad \mathbf{a}_l \in \Omega_l,\ l = 1,\ldots,L. \qquad (17)$$

- f is a multi-convex continuously differentiable function,
- $f(\mathbf{a}_1,\ldots,\mathbf{a}_{l-1},\tilde{\mathbf{a}}_l,\mathbf{a}_{l+1},\ldots,\mathbf{a}_L) \geq f(\mathbf{a}) + \nabla_l f(\mathbf{a})^\top(\tilde{\mathbf{a}}_l - \mathbf{a}_l)$.

$$\hat{\mathbf{a}}_l = \underset{\tilde{\mathbf{a}}_l\in\Omega_l}{\text{argmax}}\ \nabla_l f(\mathbf{a})^\top \tilde{\mathbf{a}}_l = \underset{\tilde{\mathbf{a}}_l\in\Omega_l}{\text{argmax}}\ \phi(\tilde{\mathbf{a}}_l). \qquad (18)$$



Maximize a scalar product under constraints

Core TGCCA optimization problem

$$\underset{\mathbf{a}_l}{\text{maximize}}\ \mathbf{q}^\top \mathbf{a}_l \tag{19}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{a}_l = \sum_{r=1}^{R_l} \lambda_l^{(r)} \mathbf{a}_{l,d_l}^{(r)} \otimes \cdots \otimes \mathbf{a}_{l,1}^{(r)}, \\ \mathbf{a}_{l,1}^{(r)\top} \mathbf{a}_{l,1}^{(s)} = 0 \text{ for } r \neq s, \\ \|\mathbf{a}_{l,m}^{(r)}\| = 1 \text{ for } m \in \{1, \ldots, d_l\}, \\ \|\boldsymbol{\lambda}_l\| = 1. \end{cases}$$

Solve this by alternating between $\boldsymbol{\lambda}_l$, and the $\{\mathbf{a}_{l,m}^{(r)}\}_{r=1}^{R_l}$ for $m \in \{1, \ldots, d_l\}$.

TGCCA Algorithm

Algorithm 1: TGCCA algorithm

Result: $\mathbf{a}^t = (\mathbf{a}_1^t, \dots, \mathbf{a}_L^t)$
Initialization: $\mathbf{a}_l^0 \in \Omega_l$, $l = 1, \dots, L$, ε;
t = 0;
repeat
    for $l = 1$ **to** L **do**
        Compute $\boldsymbol{\lambda}_l$ and update $\mathbf{a}_l^{t+1}$;
        for $m = 1$ **to** d_l **do**
            Compute $\{\mathbf{a}_{l,m}^{(r)}\}_{r=1}^{R_l}$ and update $\mathbf{a}_l^{t+1}$;
        end
    end
    $t = t + 1$;
until $f(\mathbf{a}_1^{t+1}, \dots, \mathbf{a}_L^{t+1}) - f(\mathbf{a}_1^t, \dots, \mathbf{a}_L^t) < \varepsilon$;

TGCCA Algorithm - Convergence property

Proposition

Let Φ be the mapping associated with the outer for loop of the TGCCA Algorithm and let $\{\mathbf{a}^t\}_{t=0}^{\infty}$ be any sequence generated by the recurrence relation $\mathbf{a}^{t+1} = \Phi(\mathbf{a}^t)$ with $\mathbf{a}^0 \in \Omega$. Then, the limit of any convergent subsequence of $\{\mathbf{a}^t\}$ is a stationary point of the TGCCA optimization problem.

Update of $\boldsymbol{\lambda}_l$

$$\mathbf{q}^\top \mathbf{a}_l = \mathbf{q}^\top \mathbf{A}_l \boldsymbol{\lambda}_l \text{ where the r}^{\text{th}} \text{ column of } \mathbf{A}_l \text{ is } \mathbf{a}_{l,d_l}^{(r)} \otimes \cdots \otimes \mathbf{a}_{l,1}^{(r)}.$$

$$\hat{\boldsymbol{\lambda}}_l = \underset{\boldsymbol{\lambda}_l, \|\boldsymbol{\lambda}_l\|=1}{\operatorname{argmax}} \mathbf{q}^\top \mathbf{A}_l \boldsymbol{\lambda}_l = \frac{\mathbf{A}_l^\top \mathbf{q}}{\|\mathbf{A}_l^\top \mathbf{q}\|}. \tag{20}$$

Update of the $\{\mathbf{a}_{l,m}^{(r)}\}_{r=1}^{R_l}$

Goal: express $\mathbf{q}^\top \mathbf{a}_l$ as a function of the $\{\mathbf{a}_{l,m}^{(r)}\}_{r=1}^{R_l}$.

Let $\mathcal{Q} \in \mathbb{R}^{p_1 \times \cdots \times p_{d_l}}$ be the folded version of $\mathbf{q}$,

$$
\begin{aligned}
\mathbf{q}^\top \mathbf{a}_l &= \sum_{r=1}^{R} \lambda_l^{(r)} \mathbf{q}^\top \mathbf{a}_l^{(r)} \\
&= \sum_{r=1}^{R} \lambda^{(r)} \mathcal{Q} \times_1 \mathbf{a}_{l,1}^{(r)} \cdots \times_m \mathbf{a}_{l,m}^{(r)} \cdots \times_d \mathbf{a}_{l,d_l}^{(r)} \\
&= \sum_{r=1}^{R} \lambda^{(r)} \mathrm{Vec}\left(\mathcal{Q} \times_1 \mathbf{a}_{l,1}^{(r)} \cdots \times_m \mathbf{I}_{p_{l,m}} \cdots \times_{d_l} \mathbf{a}_{l,d_l}^{(r)}\right)^\top \mathbf{a}_{l,m}^{(r)} \\
&:= \sum_{r=1}^{R_l} \mathbf{q}_m^{(r)\top} \mathbf{a}_{l,m}^{(r)}.
\end{aligned}
$$

Update of the $\{\mathbf{a}_{l,m}^{(r)}\}_{r=1}^{R_l}$

$$\mathbf{q}^\top \mathbf{a}_l = \sum_{r=1}^{R_l} \mathbf{q}_m^{(r)\top} \mathbf{a}_{l,m}^{(r)}.$$

For $m > 1$, $\hat{\mathbf{a}}_{l,m}^{(r)} = \frac{\mathbf{q}_m^{(r)}}{\|\mathbf{q}_m^{(r)}\|}$.

For $m = 1$, the vectors $\hat{\mathbf{a}}_{l,m}^{(r)}$ are the columns of $\mathbf{U}\mathbf{V}^\top$ where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ is the Singular Value Decomposition of the matrix which columns are the vectors $\mathbf{q}_m^{(r)}$ (Orthogonal Procrustes problem, Everson, 1997).

Table of Contents

First dataset

Simulated data

Creation of a dataset to make sure the TGCCA algorithm solves the TGCCA optimization problem.

- Generated dataset with 5 correlated blocks using a latent factor model (Bach and Jordan, 2005; Min et al., 2019);
- One canonical vector $\mathbf{a}_l$ per block with a low-rank tensor structure;
- Independent noise is added to each block.

Compared methods on the first dataset

Table: Summary of the methods included in the comparison.

Name	Number of blocks	CP rank	Separable version	Non-separable version
2DCCA Chen et al. (2021)	2	1	$\times$	✓
MGCCA Gloaguen et al. (2020)	L	1	✓	$\times$
TCCA Min et al. (2019)	2	R	✓	✓
TGCCA	L	R	✓	✓
RGCCA Tenenhaus et al. (2017)	L	NA	$\times$	✓
SVD	1	NA	$\times$	✓

- Prefix "sp" for separable versions, rank added as a suffix;
- MGCCA is equivalent to spTGCCA1.

Compared methods on the first dataset

Table: Summary of the methods included in the comparison.

Name	Number of blocks	CP rank	Separable version	Non-separable version
2DCCA Chen et al. (2021)	2	1	×	✓
MGCCA Gloaguen et al. (2020)	L	1	✓	×
TCCA Min et al. (2019)	2	R	✓	✓
TGCCA	L	R	✓	✓
RGCCA Tenenhaus et al. (2017)	L	NA	×	✓
SVD	1	NA	×	✓

- Prefix "sp" for separable versions, rank added as a suffix;
- MGCCA is equivalent to spTGCCA1.

Results on the first dataset

Table: Cosine between the true and the estimated canonical vectors for rank-1 models.

Model	Gas	Cross (small)	Computation time (s)
2DCCA1	0.30 (0.01, 0.89)	0.43 (0.16, 0.85)	3.09 (2.76, 4.50)
TCCA1	0.89 (0.22, 0.90)	0.85 (0.32, 0.86)	7.72 (7.38, 9.17)
TGCCA1	0.89 (0.87, 0.90)	0.85 (0.83, 0.86)	8.60 (8.36, 9.12)
spTCCA1	0.89 (0.22, 0.90)	0.85 (0.32, 0.86)	7.43 (7.24, 8.04)
MGCCA	0.89 (0.87, 0.90)	0.86 (0.83, 0.86)	5.08 (4.85, 5.32)

- All rank-1 models behave similarly;
- Unfair to 2DCCA1 because run only once using a so-called "effective" strategy.

Results on the first dataset

Table: Cosine between the true and the estimated canonical vectors for rank-3 models. MGCCA is used as a reference.

Model	Gas	Cross (small)	Computation time (s)
MGCCA	0.89 (0.87, 0.90)	0.86 (0.83, 0.86)	5.08 (4.85, 5.32)
2DCCA3	0.04 (0.01, 0.21)	0.13 (0.05, 0.31)	1.50 (1.41, 3.09)
TCCA3	0.89 (0.87, 0.90)	0.85 (0.83, 0.86)	7.93 (7.73, 8.92)
spTCCA3	0.89 (0.87, 0.90)	0.85 (0.83, 0.86)	7.32 (7.22, 7.51)
TGCCA3	**0.91 (0.78, 0.94)**	**0.92 (0.79, 0.96)**	11.32 (10.12, 15.94)
spTGCCA3	**0.92 (0.82, 0.94)**	**0.93 (0.83, 0.96)**	5.66 (5.52, 6.39)

- 2DCCA3 is not able to retrieve the canonical vectors because look for 3 canonical components of rank 1;
- TCCA3 and spTCCA3 produce almost collinear vectors.

Results on the first dataset

Table: Cosine between the true and the estimated canonical vectors for spTGCCA3, RGCCA and SVD.

Model	Gas	Cross (small)	Computation time (s)
spTGCCA3	**0.92 (0.82, 0.94)**	**0.93 (0.83, 0.96)**	5.66 (5.52, 6.39)
RGCCA	0.17 (0.05, 0.26)	0.11 (0.06, 0.20)	13.12 (12.67, 14.07)
SVD	0.00 (0.00, 0.01)	0.01 (0.00, 0.03)	5.78 (5.44, 6.07)

RGCCA and SVD model the noise of the blocks.

Second dataset

Data from Paris Brain Institute

The goal is to characterize the differences between multiple sclerosis (MS) patients and controls using TGCCA.

- ☐ Dataset with 4 blocks;
- ☐ Each block is a third-order tensor.

Second dataset

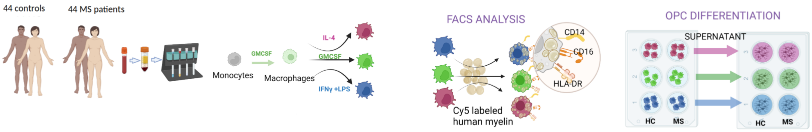


Figure: Graphical summary of the data extracted from blood samples.

Results on the second dataset

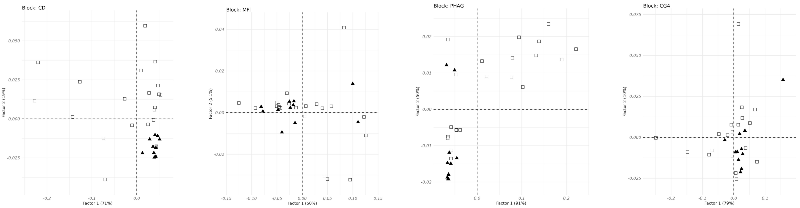


Figure: Scores of the patients obtained with TGCCA for the different blocks. Controls are represented with triangles, and MS patients with squares.

Results on the second dataset

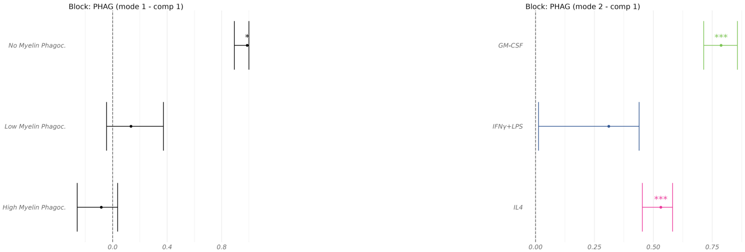


Figure: Estimated weights of the TGCCA model trained on patients and controls for the first and second modes of block PHAG. Error bars represent empirical 95% confidence intervals, and stars represent the significance levels.

Table of Contents

Conclusion

Contributions

We proposed a new GCCA optimisation problem which

- Naturally regularizes the GCCA problem;
- Drastically decreases the number of degrees of freedom;
- Achieves higher reconstruction performances at lower SNR.

Paper published in Journal of Information Fusion: Girka et al. (2024).

Conclusion

Perspectives

There are still open questions:

- How to choose the best rank?
- How to go beyond the first canonical component? Use a sequential approach or aim for a simultaneous one?
- Should we remove orthogonality constraints?

Bibliography

Bach, F. and Jordan, M. (2005). A probabilistic interpretation of canonical correlation analysis.

Chen, Y.-L., Kolar, M., and Tsay, R. S. (2021). Tensor canonical correlation analysis with convergence and statistical guarantees. *Journal of Computational and Graphical Statistics*, 0(0):1–17.

de Leeuw, J. (1994). Block-relaxation algorithms in statistics. In *Studies in Classification, Data Analysis, and Knowledge Organization*, pages 308–324. Springer Berlin Heidelberg.

Everson, R. (1997). Orthogonal, but not orthonormal, procrustes problems. In *Advances in Computational Mathematics . (Submitted). Available from http://www.ee.ic.ac.uk/research/neural/everson*.

Fransson, J., Bachelin, C., Deknuydt, F., Ichou, F., Guillot-Noël, L., Ponnaiah, M., Gloaguen, A., Maillart, E., Stankoff, B., Tenenhaus, A., Mochel, F., Fontaine, B., Louapre, C., and Zujovic, V. (2021). Dysregulated functional and metabolic response in multiple sclerosis patient macrophages correlate with a more inflammatory state, reminiscent of trained immunity. *bioRxiv*.

Bibliography (cont.)

Girka, F., Gloaguen, A., Le Brusquet, L., Zujovic, V., and Tenenhaus, A. (2024). Tensor generalized canonical correlation analysis. *Information Fusion*, 102:102045.

Gloaguen, A., Philippe, C., Frouin, V., Gennari, G., Dehaene-Lambertz, G., Brusquet, L. L., and Tenenhaus, A. (2020). Multiway generalized canonical correlation analysis. *Biostatistics*.

Min, E. J., Chi, E. C., and Zhou, H. (2019). Tensor canonical correlation analysis. *Stat*, 8(1):e253. e253 sta4.253.

Tenenhaus, M., Tenenhaus, A., and Groenen, P. J. F. (2017). Regularized generalized canonical correlation analysis: A framework for sequential multiblock component methods. *Psychometrika*, 82(3):737–777.