Generalizing Principal Components Analysis

Genevera I. Allen

Department of Pediatrics-Neurology, Baylor College of Medicine, & Department of Statistics, Rice University & Jan and Dan Duncan Neurological Research Institute, Texas Children’s Hospital

October, 2011
1 Introduction
   • Motivation

2 Generalized Least Squares Matrix Decomposition
   • Problem and Solution
   • Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   • Problem & Solution
   • Algorithms & Examples

4 Results
   • Real fMRI Example
   • Case Study: NMR Data
   • Future Work
Principal Components Analysis (PCA):

- Dimension reduction.
- Exploratory data analysis.
- PCA Problem:

\[
\text{maximize } \mathbf{v}_k^T \mathbf{X}^T \mathbf{X} \mathbf{v}_k \quad \text{subject to } \mathbf{v}_k^T \mathbf{v}_k = 1 \text{ and } \mathbf{v}_k^T \mathbf{v}_{k'} = 0 \quad \forall \ k' < k.
\]

PC: \( \mathbf{z}_k = \mathbf{X} \mathbf{v}_k \).

Given by the singular value decomposition (SVD) of the data matrix: \( \mathbf{X} = \mathbf{U} \mathbf{D} \mathbf{V}^T \), then \( \mathbf{Z} = \mathbf{X} \mathbf{V} \).
When does PCA (SVD) fail?

1. High-dimensional data.
   - Fix: Sparsity - Sparse PCA (Johnstone and Lu, 2004).

2. Structured Factors.
   - Fix: Sparsity - Sparse PCA (Jolliffe et. al, 2003).

Transposable Data: Dependencies among the rows and/or column of a data matrix.
When does PCA (SVD) fail?

1. High-dimensional data.
   - Fix: Sparsity - Sparse PCA (Johnstone and Lu, 2004).

2. Structured Factors.
   - Fix: Sparsity - Sparse PCA (Jolliffe et. al, 2003).

3. Strong dependencies among variables? Or Two-way dependencies in transposable data?
   - *Transposable Data*: Dependencies among the rows and/or column of a data matrix.
Motivating Example: Functional MRIs

Functional MRIs:
- Rows: Voxels ($64 \times 64 \times 30$).
- Columns: Time points ($\approx 1,000 - 10,000$).
Motivating Example: Functional MRIs

Multivariate analysis techniques used for finding Regions of Interest and Activation Patterns, but . . .

(Viviani et. al, 2005)
PCA and Correlated Noise

True Signal  Signal + Independent Noise  rank 1 SVD  FPCA Component 1

True Signal  Signal + Independent Noise  rank 1 SVD  SPCA Component 1
PCA and Correlated Noise
SVD Model

SVD Model: $$X_{n \times p} = U_{n \times K} D_{K \times K} V^T_{p \times K} + E_{n \times p}.$$  
- $$U D V^T$$ is the low rank signal. $$U$$ and $$V$$ are the left and right factors (singular vectors). Diagonal of $$D$$ are the singular values.
- $$E$$ is the noise. $$E_{ij} \sim iid (0, \sigma^2)$$, or $$\text{Cov} (\text{vec}(E)) = \sigma^2 I(p) \otimes I(n)$$.

SVD Loss Function: $$\| X - U D V^T \|_F^2.$$  
- $$\| \cdot \|_F$$ is the Frobenius norm (sums of squared errors).
- Error terms weighted equally.
- Cross-product errors between elements $$ij$$ and $$i'j'$$ are ignored.
Our Generative Model

\[
X_{n \times p} = \sum_{k=1}^{K} d_k u_k v_k^T + E_{n \times p}.
\]

- Random: \(d_k\) & Fixed: \(U = [u_1, \ldots, u_K]\) and \(V = [v_1, \ldots, v_K]\).
- Signal factors: \(U^T \Sigma^{-1} U = I\) & \(V^T \Delta^{-1} V = I\).
- Noise: Two-way (separable) dependencies:

\[
\text{Cov}(\text{vec}(E)) = \Delta \otimes \Sigma,
\]

with \(\Delta \in \mathbb{R}^{p \times p}\) the column covariance and \(\Sigma \in \mathbb{R}^{n \times n}\) the row covariance.

\[
\Delta \otimes \Sigma = \begin{pmatrix}
\Delta_{11} \Sigma & \Delta_{12} \Sigma & \ldots & \Delta_{1p} \Sigma \\
\Delta_{21} \Sigma & \Delta_{22} \Sigma & \ldots & \Delta_{2p} \Sigma \\
\vdots & \vdots & \ddots & \vdots \\
\Delta_{p1} \Sigma & \ldots & \Delta_{pp} \Sigma
\end{pmatrix}
\]
Our Contributions

We develop . . .

- A matrix decomposition that directly accounts for known two-way dependencies.

- A regularized matrix factorization that directly accounts for known two-way dependencies.

- Algorithms to compute these decompositions for massive data sets.
Application of our methods:

1. Finding **Regions of Interest**.
2. Exploring activation patterns.
1 Introduction
   • Motivation

2 Generalized Least Squares Matrix Decomposition
   • Problem and Solution
   • Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   • Problem & Solution
   • Algorithms & Examples

4 Results
   • Real fMRI Example
   • Case Study: NMR Data
   • Future Work
Background: \( Q, R \)-norm

SVD Problem

\[
\begin{align*}
\text{minimize} & \quad \| X - U D V^T \|_F^2 \\
\text{subject to} & \quad U^T U = I_K, \quad V^T V = I_K \quad \& \quad \text{diag}(D) \geq 0.
\end{align*}
\]

How can we modify the Frobenius norm?

G. I. Allen (BCM & Rice)  Generalizing PCA  October, 2011  12 / 47
Background: $Q$, $R$-norm

Definition: $Q$, $R$-norm

- $Q \in \mathbb{R}^{n \times n}: Q \succeq 0$ (positive semi-definite).
- $R \in \mathbb{R}^{p \times p}: R \succeq 0$.

$Q$, $R$-norm: $\|X\|_{Q,R} = \sqrt{\text{tr}(QXRX^T)}$.

- Call $Q$, $R$-norm a transposable quadratic norm.
- Norm induced by an inner product space.
- Generalization of the Frobenius norm: $\|X\|_{I,I} = \|X\|_F$. 
Background: \( Q, R \)-norm

Where does the \( Q, R \)-norm arise?

- Let \( X \) be matrix-variate normal: \( X \sim N_{n,p}(UDVT, Q^{-1}, R^{-1}) \), or \( \text{vec}(X) \sim N(\text{vec}(UDVT), R^{-1} \otimes Q^{-1}) \).

- Log-likelihood:

\[
\ell(X \mid Q^{-1}, R^{-1}) \propto \text{tr} \left( Q(X - UDVT)R(X - UDVT)^T \right) \\
= \| X - UDVT \|^2_{Q,R}.
\]

- Frobenius norm: \( \| X \|_F \propto \ell(X \mid I(n), I(p)) \).

- \( Q, R \)-norm: \( \| X \|_{Q,R} \propto \ell(X \mid Q^{-1}, R^{-1}) \).
GMD Problem

Generalized Least Squares Matrix Decomposition (GMD) Problem

\[
\begin{align*}
\text{minimize} & \quad \| X - UDV^T \|^2_{Q,R} \\
\text{subject to} & \quad U^T Q U = I_K, \quad V^T R V = I_K \quad \& \quad \text{diag}(D) \geq 0.
\end{align*}
\]

- \( U \) and \( V \) the left and right GMD factors.
- Diagonal elements of \( D \), the GMD values.
- \( Q \) and \( R \) the left and right quadratic operators.
Let \( \text{rank}(Q) = l \) and \( \text{rank}(R) = m \).

Let \( Q = \tilde{Q}\tilde{Q}^T \) such that \( \tilde{Q} \in \mathbb{R}^{n \times l} \) has full column rank.

Let \( R = \tilde{R}\tilde{R}^T \) such that \( \tilde{R} \in \mathbb{R}^{p \times m} \) has full column rank.

Define \( \tilde{X} = \tilde{Q}^TX\tilde{R} \).

SVD of \( \tilde{X} \): 
\[
\tilde{X} = \tilde{U}\tilde{D}\tilde{V}^T.
\]

**Theorem: GMD Solution**

GMD solution, \( \hat{X} = U^* D^* (V^*)^T \), given by:

\[
U^* = \tilde{Q}^{-1}\tilde{U}, \quad V^* = \tilde{R}^{-1}\tilde{V}, \quad \& \quad D^* = \tilde{D}.
\]
Sphere or whiten the data so that the SVD is appropriate.

1. \( \tilde{Q} \) behaves like the square root of the inverse ROW covariance matrix.
2. \( \tilde{R} \) behaves like the square root of the inverse COLUMN covariance matrix.
3. Elements of \( \tilde{X} = \tilde{Q}^T X \tilde{R} \) approximating independent.

Reconstructs the GMD factors in the original coordinates.

GMD values are the same as the singular values of the sphered data.

- \( \tilde{Q} \) and \( \tilde{R} \) can be found via the full eigenvalue decomposition.
- Inherits many of the nice properties of the SVD.
1 Introduction
   - Motivation

2 Generalized Least Squares Matrix Decomposition
   - Problem and Solution
   - Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   - Problem & Solution
   - Algorithms & Examples

4 Results
   - Real fMRI Example
   - Case Study: NMR Data
   - Future Work
Computing the GMD Solution

Problem:
- Computing solution requires factoring and inverting $Q$ and $R$.
- Taking the SVD of $\tilde{X}$.
- fMRI data: voxels, $n = O(10^6)$ and time points, $p = O(10^4)$.

Solution: Algorithm based on the Power Method.
GMD Algorithm

GMD Algorithm (Power Method)

1. Let $\hat{X}^{(1)} = X$ and initialize $u_1$ and $v_1$.
2. For $k = 1 \ldots K$:
   1. Repeat until convergence:
      1. Set $u_k = \frac{\hat{X}^{(k)} R v_k}{\|\hat{X}^{(k)} R v_k\|_Q}$.
      2. Set $v_k = \frac{(\hat{X}^{(k)})^T Q u_k}{\|((\hat{X}^{(k)})^T Q u_k\|_R}$.
   2. Set $d_k = u_k^T Q \hat{X}^{(k)} R v_k$.
   3. Set $\hat{X}^{(k+1)} = \hat{X}^{(k)} - u_k d_k v_k^T$.
3. Return $U^* = [u_1, \ldots, u_K]$, $V^* = [v_1, \ldots, v_K]$ and $D^* = \text{diag}(d_1, \ldots, d_K)$.

- Converges to the **GLOBAL** solution!
- Do not need to diagonalize $Q$ and $R$ or compute SVD!
GMD power method updates:

\[
\begin{align*}
\text{minimize}_v & \quad \| X - (d' u') v^T \|_Q^2, \\
\text{minimize}_u & \quad \| X^T -(d' v') u^T \|_R^2.
\end{align*}
\]

Alternating generalized least squares problems.
Generalized PCA

Classical PCA Problem

\[ \begin{align*}
& \text{maximize} \quad v_k^T X^T X v_k \\
& \text{subject to} \quad v_k^T v_k = 1 \quad \& \quad v_k^T v_{k'} = 0 \quad \forall \quad k' < k.
\end{align*} \]

PC: \( z_k = X v_k \).
Generalized PCA

Classical PCA Problem

maximize \( v_k^T X^T X v_k \)
subject to \( v_k^T v_k = 1 \) & \( v_k^T v_{k'} = 0 \) \( \forall k' < k \).

PC: \( z_k = X v_k \).

Generalized PCA Problem

maximize \( v_k^T R X^T Q X R v_k \)
subject to \( v_k^T R v_k = 1 \) & \( v_k^T R v_{k'} = 0 \) \( \forall k' < k \).

GPC: \( z_k = X R v_k \).

Inner product spaces are those induced by the \( Q, R \)-norm.
Interpretations of Quadratic Operators

1. Matrix-variate Normal: $Q$ and $R$ behave like inverse row and column covariances.

2. Covariance Decomposition:
   - $X = S + E = \text{signal} + \text{noise}.$
   - $S = \sum d_k u_k v_k^T,$ $u_k, v_k$ are fixed.
   - $\text{vec}(E) \sim (0, R^{-1} \otimes Q^{-1})$

   $$\text{Cov}(\text{vec}(X)) = \sum \text{Var}(d_k)(v_k v_k^T) \otimes (u_k u_k^T) + R^{-1} \otimes Q^{-1}.$$ Where $V^T R V = I$ and $U^T Q U = I.$

3. Smoothing Matrices: Factors $U$ and $V$ as smooth as the smallest eigenvectors of $Q$ and $R.$

4. Weighting Matrices.
How do we choose $Q$ and $R$?

Estimated from data:
- Precision matrix of a known model for the noise.
  - Example: fMRI data assumed to follow an autoregressive process.
- Precision matrix of a Gaussian Markov random field.

Determined via known structure:
- Graph Laplacians.
  - $L = D - A$. (Laplacian = Degree matrix - Adjacency matrix).
  - $Q$ and $R$ enforce smoothness with respect to a graph structure.
- Smoothing matrix.
  - Example: kernel smoothers over a distance function.

Note: For numerical reasons, $Q$ and $R$ are taken to have operator norm $\leq 1$. 
Simulated Example

Simulation Model:

- Inspired by fMRI data.
- 2 Spatial factors, $\mathbf{U}$: Three regions of interest on a $16 \times 16$ grid with $\mathbf{U}^T \mathbf{U} = \mathbf{I}$.
- 2 Temporal factors, $\mathbf{V}$: Sinusoidal activation patterns with 200 time points with $\mathbf{V}^T \mathbf{V} = \mathbf{I}$.
- Noise, $\mathbf{E}$: Autoregressive spatio-temporal Gaussian Markov random field.
- Strength of Factor 1 vs. Factor 2: $\mathbf{D} = \text{diag}(1.0, 0.8)$.
- Model: $\mathbf{Y} = \phi \mathbf{U} \mathbf{D} \mathbf{V}^T + \mathbf{E}$.
- Scalar, $\phi$ such that the signal to noise ratio (SNR) = 1.
Simulated Example

True Spatial 1

SVD Spatial 1

GMD Spatial 1

True Spatial 2

SVD Spatial 2

GMD Spatial 2

Signal + Spatio-Temporal Noise

SVD Spatial 3

GMD Spatial 3
Simulated Example

True Temporal Factors

SVD Temporal Factors 1-3

GMD Temporal Factors 1-3
1 Introduction
   • Motivation

2 Generalized Least Squares Matrix Decomposition
   • Problem and Solution
   • Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   • Problem & Solution
   • Algorithms & Examples

4 Results
   • Real fMRI Example
   • Case Study: NMR Data
   • Future Work
Single Factor SVD Problem

\[
\text{maximize} \quad \mathbf{u}^T \mathbf{X} \mathbf{v} \\
\text{subject to} \quad \mathbf{u}^T \mathbf{u} = 1, \quad \& \quad \mathbf{v}^T \mathbf{v} = 1.
\]
Background

Existing two-way penalized matrix factorizations:


$$\max_{v,u} u^T X v \text{ subject to } u^T u \leq 1, \ v^T v \leq 1, \ P_1(u) \leq c_1, \ & P_2(v) \leq c_2.$$  

Huang et. al (2009): Two-way Functional PCA.

$$\max_{v,u} u^T X v - \frac{\lambda}{2} P_1(u) P_2(v).$$

Lee et. al (2010): Sparse SVD.

$$\max_{v,u} u^T X v - \frac{1}{2} u^T u v^T v - \frac{\lambda_u}{2} P_1(u) - \frac{\lambda_v}{2} P_2(v).$$

Problems:

1. Scaling of the factors.
2. Which penalties?
Single Factor GPMF Problem

\[
\begin{align*}
\text{maximize} \quad & u^T Q X R v - \lambda_v P_1(v) - \lambda_u P_2(u) \\
\text{subject to} \quad & u^T Q u \leq 1 \quad \& \quad v^T R v \leq 1.
\end{align*}
\]

- Similar to the Lagrangian form of Witten et. al (2009) problem.
- Inner product spaces induced by the $Q, R$-norm.
- Multi-factor problem solved via power method.
- Simple alternating coordinate ascent algorithm.
GPMF Solution

\[ \hat{v} = \arg\min_v \left\{ \frac{1}{2} \left\| X^T Q u' - v \right\|_R^2 + \lambda_v P_1(v) \right\} \]

\[ \hat{u} = \arg\min_u \left\{ \frac{1}{2} \left\| X R v - u \right\|_Q^2 + \lambda_u P_2(u) \right\} . \]

Theorem

- Assume \( P_1() \) and \( P_2() \) convex and homogeneous of order one: 
  \( P(cx) = cP(x) \quad \forall \ c > 0. \)
- Coordinate-wise update to single factor GPMF problem:
  \[
  v^* = \begin{cases} 
  \hat{v} / \left\| \hat{v} \right\|_R & \text{if } \left\| \hat{v} \right\|_R > 0 \\
  0 & \text{otherwise,} \end{cases} \quad \& \quad u^* = \begin{cases} 
  \hat{u} / \left\| \hat{u} \right\|_Q & \text{if } \left\| \hat{u} \right\|_Q > 0 \\
  0 & \text{otherwise.} \end{cases} \]

Solution as simple as solving penalized regression problems!
1 Introduction
   - Motivation

2 Generalized Least Squares Matrix Decomposition
   - Problem and Solution
   - Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   - Problem & Solution
   - Algorithms & Examples

4 Results
   - Real fMRI Example
   - Case Study: NMR Data
   - Future Work
Sparsity: LASSO Penalty

$\mathbf{R}$-norm Lasso Problem:

$$\frac{1}{2} \| \mathbf{X}^T \mathbf{Q} \mathbf{u} - \mathbf{v} \|_\mathbf{R}^2 + \lambda \| \mathbf{v} \|_1$$

- Can be solved via coordinate descent:

$$\hat{v}_j = \frac{1}{R_{jj}} S \left( R_{rj} \mathbf{X}^T \mathbf{Q} \mathbf{u} - R_{j,\neq j} \hat{v}_{\neq j}, \lambda \right).$$

- Or ISTA (and FISTA):

$$\hat{v}^{(t+1)} = S \left( \hat{v}^{(t)} + \frac{1}{L} \left( \mathbf{R} \mathbf{X}^T \mathbf{Q} \mathbf{u} - \mathbf{R} \hat{v}^{(t)} \right), \frac{\lambda}{L} \right).$$

- Do not need to diagonalize $\mathbf{R}$!

$S(\cdot, \lambda)$ is the soft-thresholding operator: $S(\cdot, \lambda) = \text{sign}(\cdot)(|\cdot| - \lambda)_+$. 

$S(\cdot, \lambda)$ is the soft-thresholding operator: $S(\cdot, \lambda) = \text{sign}(\cdot)(|\cdot| - \lambda)_+$. 

Smoothness: $\Omega$-norm Penalty

$\mathbb{R}$-norm, $\Omega$-norm Penalized Problem:

$$\frac{1}{2} \| X^T Q u - v \|_R^2 + \lambda v \| \Omega v.$$

- Typical functional data penalty: $v^T \Omega v$.
- $\| v \|_\Omega = \sqrt{v^T \Omega v}$.
- $\Omega$ can be squared second or fourth differences, or taken to give spline functions.
- Can be solved via generalized gradient descent.
- Do not need to diagonalize $\mathbb{R}$!
BIC Selection Criterion:

\[
BIC(\lambda_v) = \log \left( \frac{|| X - d' u' \hat{v} ||^2_{Q,R}}{np} \right) + \frac{\log(np)}{np} \hat{df}(\lambda_v).
\]

where \( \hat{df}(\lambda_v) \) an unbiased estimate of the degrees of freedom.

- Example: \( \hat{df}(\lambda_v) = | \{ \hat{v} \} | \) for LASSO penalty.
Spatio-Temporal Simulation: Example

True Spatial 1  SPCA Spatial 1  SGPCA Spatial 1

True Spatial 2  SPCA Spatial 2  SGPCA Spatial 2

Signal + Spatio-Temporal Noise  SPCA Spatial 3  SGPCA Spatial 3
Spatio-Temporal Simulation: Example

True Temporal Factors

SPCA Temporal Factors 1-3

SGPCA Temporal Factors 1-3
1. Introduction
   - Motivation

2. Generalized Least Squares Matrix Decomposition
   - Problem and Solution
   - Algorithm and GPCA

3. Generalized Penalized Matrix Factorization
   - Problem & Solution
   - Algorithms & Examples

4. Results
   - Real fMRI Example
   - Case Study: NMR Data
   - Future Work
Axial slices of the first 8 spatial sparse GPC’s.
fMRI Results: Sparse PCA

Component 1 and 2:
1 Introduction
   • Motivation

2 Generalized Least Squares Matrix Decomposition
   • Problem and Solution
   • Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   • Problem & Solution
   • Algorithms & Examples

4 Results
   • Real fMRI Example
   • Case Study: NMR Data
   • Future Work
High-throughput technology measures all metabolites (small molecules) in a sample.

- Spectrum contains thousands of chemical resonances belonging to possibly hundreds of unknown metabolites.
- Metabolites resonate at multiple frequencies.
- Resonances of multiple metabolites may overlap.
NMR Spectroscopy Data

- Neural cell type data: 27 samples of purified neural cells measured at \( \approx 3,000 \) chemical shifts.
- Cell types: Neural Stem Cells, Neurons, Astrocytes, Oligodendrocytes, Microglia.
- Goal: Find metabolic biomarkers that differentiate the neural cell types.
PCA for NMR Data

Limitations:

(1) High-dimensional,

(2) Non-negative spectra,

(3) Strong spatial dependencies.
Sparse Non-Negative GPCA

Problem:

\[
\begin{align*}
\text{maximize} & \quad u^T X R v - \lambda \| v \|_1 \\
\text{subject to} & \quad u^T u \leq 1, \quad v^T R v \leq 1, \quad \& \quad v \geq 0.
\end{align*}
\]

Solution: Replace soft-thresholding operator,

\[S(x, \lambda) = \text{sign}(x) (|x| - \lambda)_+ ,\]

with positive operator,

\[P(x, \lambda) = (x - \lambda)_+ ,\]
Exploring Relationships Between Samples

![PCA plots](image)

- **PC1 vs PC2**
- **PC2 vs PC3**
- **PC3 vs PC4**
- **PC4 vs PC5**

![SGPC plots](image)

- **SGPC1 vs SGPC2**
- **SGPC2 vs SGPC3**
- **SGPC3 vs SGPC4**
- **SGPC4 vs SGPC5**

Legend:
- Microglia
- Astrocyte
- Oligodendrocyte
- Neuron
- Neural Stem Cell
Explaining Variance in NMR Data

PCA

Sparse Non-Negative PCA

% Variance Explained

% Cumulative Variance Explained
Feature Selection: Sparse PCA vs. Sparse GPCA

Degree of Sparsity - Sparse Non-Negative PCA Loadings

% Features Selected

% Variance Explained

PCs

Traditional
Generalized

% Features Selected

Traditional
Generalized

G. I. Allen (BCM & Rice)
Results: Sparse Non-Negative GPCA

PC 1-3 - Baseline Height & Neural Stem Cells
Results: Sparse Non-Negative GPCA

PC 4 - 7 - Microglia, Astrocytes, Oligodendrocytes.
## Results: Sparse Non-Negative Generalized PCA

Metabolites:

<table>
<thead>
<tr>
<th>Peak Location</th>
<th>Cell Types</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.96 ppms</td>
<td>Neuron, Microglia</td>
<td>Lipid Moiety</td>
</tr>
<tr>
<td>1.19 ppms</td>
<td><strong>Microglia</strong>, Neuron</td>
<td>NAA</td>
</tr>
<tr>
<td>1.28 ppms</td>
<td><strong>Neural Stem Cell</strong>, Oligodendrocyte</td>
<td></td>
</tr>
<tr>
<td>1.48 ppms</td>
<td>Oligodendrocyte</td>
<td></td>
</tr>
<tr>
<td>2.02 ppms</td>
<td>Neuron</td>
<td></td>
</tr>
<tr>
<td>2.65 ppms</td>
<td><strong>Astrocyte</strong></td>
<td></td>
</tr>
<tr>
<td>3.01 ppms</td>
<td>Neuron</td>
<td></td>
</tr>
<tr>
<td>3.04 ppms</td>
<td>Oligodendrocyte</td>
<td></td>
</tr>
<tr>
<td>3.23 ppms</td>
<td><strong>Oligodendrocyte</strong>, Neuron, Astrocyte</td>
<td>Creatine</td>
</tr>
<tr>
<td>3.43 ppms</td>
<td>Oligodendrocyte</td>
<td>Choline</td>
</tr>
<tr>
<td>3.66 ppms</td>
<td><strong>Microglia</strong></td>
<td></td>
</tr>
</tbody>
</table>
1 Introduction
   • Motivation

2 Generalized Least Squares Matrix Decomposition
   • Problem and Solution
   • Algorithm and GPCA

3 Generalized Penalized Matrix Factorization
   • Problem & Solution
   • Algorithms & Examples

4 Results
   • Real fMRI Example
   • Case Study: NMR Data
   • Future Work
Concluding Remarks

When to use GPCA vs. PCA:
- Data in which the structure of the noise is well-understood.
- Smooth or functional data.
- Data with low signal to noise ratio.

Possible Applications:
- Neuroimaging & Image Data.
- Genomics & Proteomics.
- Times Series & Spatial Data.
Future Work

Extensions using the $Q, R$-norm:

- Multivariate Analysis Methods:
  - Canonical correlation analysis, discriminant analysis, factor analysis, multi-dimensional scaling, other matrix factorizations.

Estimating $Q$, $R$, and the GMD:

- Transposable Regularized Covariance Models (Allen and Tibshirani, 2010).
- Can we test for certain quadratic operators?
- More studies on robustness of GMD to $Q$ and $R$. 
Future Work

R Package & Matlab Toolbox

Coming Soon . . .

Matlab scripts available at www.stat.rice.edu/~gallen/software.html.
Collaborators

Thank you to . . .

Frederick Campbell, PhD Student, Rice University.

Mirjana Maletic-Savatic, Department of Pediatrics, Baylor College of Medicine & Jan and Dan Duncan Neurological Research Institute, Texas Children’s Hospital.

Logan Grosenick, Center for Mind, Brain & Computation, Stanford University.

Jonathan Taylor, Department of Statistics, Stanford University.
References


