Regularized Multivariate Two-way Functional Principal Component Analysis

Mobina Pourmoshir, Dr. Mehdi Maadooliat
Marquette University
Department of Mathematical and Statistical Sciences

I. Background

- **Functional data are ubiquitous.**
  Modern sensors yield curves, images, and surfaces observed over time/space.
  Functional PCA (FPCA) summarizes such data with a few principal functions for interpretation and modeling (Ramsay & Silverman, 2005).
- **Extensions exist but are isolated.**
  - *Smoothed FPCA*: roughness penalties produce smoother, less noisy components (Huang, Shen, & Buja, 2008; Silverman, 1996).
  - *Sparse FPCA*: sparsity zeros out unimportant regions, improving interpretability (Nie & Cao, 2020; Shen & Huang, 2008).
  - *Multivariate FPCA (MFPCA)*: captures shared variation across multiple functional variables (Happ & Greven, 2018).
  - *Two-way functional data* (e.g., time × space): require structure in *both* domains.
- **Limitations.**
  Classical FPCA is noise-sensitive and can yield rough, dense patterns; many methods address *either* smoothness *or* sparsity, or are limited to univariate data.
- **Motivation.**
  Develop a regularized FPCA framework that (i) handles **multivariate** and **two-way** functional structures,

  (ii) imposes **smoothness** (noise reduction) and **sparsity** (feature selection) *simultaneously* on scores and loadings,
  and (iii) yields low-rank, interpretable, and stable components.
  Recent work (e.g., ReMFPCA) points in this direction but leaves room for a unified treatment and broader applicability (Haghbin, Zhao, & Maadooliat, 2025).

II. Methodology

1) *Multivariate FPCA Formulation:* Concatenate p functional variables into $\mathbf{X} \in \mathbb{R}^{n\times M}$, where $M = \sum_{i=1}^{p} m_i$.
We estimate a rank-one structure with penalties:

$$\min_{u,v} \| \mathbf{X} - uv^\top \|_F^2 + \alpha\, v^\top \boldsymbol{\Omega} v + p_\gamma(v), \tag{1}$$

where $\boldsymbol{\Omega} = \mathrm{diag}(\Omega_1, \ldots, \Omega_p)$ encodes **roughness** and $p_\gamma(\cdot)$ induces **sparsity** (soft, hard, or SCAD) (Huang et al., 2008; Nie & Cao, 2020; Shen & Huang, 2008).
2) *Sequential Power Algorithm:* Let $S(\alpha) = (I + \alpha\boldsymbol{\Omega})^{-1}$. Iterate:

1. **Initialize:** v via rank-one SVD of $\mathbf{X}$.
2. **Repeat:**
   - $u \leftarrow \mathbf{X}v$
   - $v \leftarrow S(\alpha)\, h_\gamma(\mathbf{X}^\top u)$
   - $v \leftarrow v / \| v \|$
3. **Deflate:** $\mathbf{X} \leftarrow \mathbf{X} - \sigma uv^\top$ to extract additional components.

Tuning:
Choose γ by K-fold cross-validation.
Choose α by generalized cross-validation (GCV):

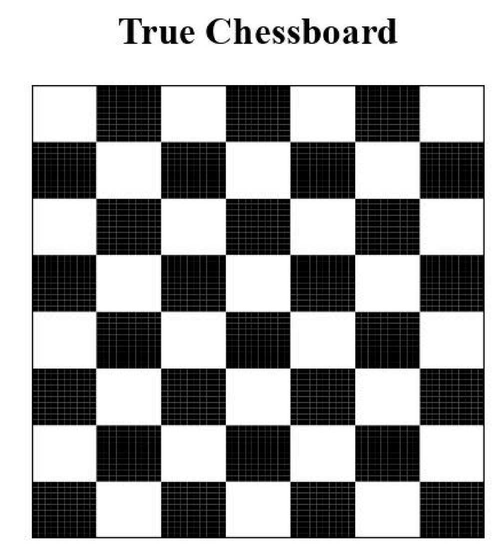
$$\mathrm{GCV}(\alpha) = \frac{\| (I - S(\alpha))(\mathbf{X}^\top u) \|^2 / M}{\left(1 - \frac{1}{M}\,\mathrm{tr}\, S(\alpha)\right)^2}. \tag{2}$$

III. Two-way Regularized MFPCA

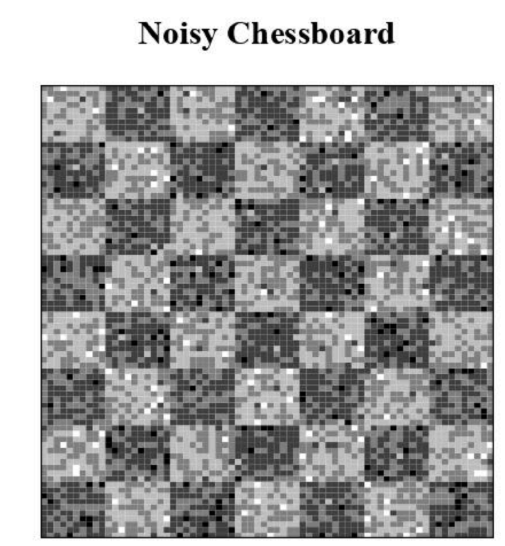
- **Two-way functional data:**
  Two-way functional data consist of a data matrix whose row and column domains are both structured. Classical FPCA focuses on one domain and penalizes only one set of components, often ignoring structure in the second direction.
- **Framework & Penalty:**

$$\min_{u,v} \| \boldsymbol{X} - uv^\top \|_F^2 + \sum_{j}^{J} \mathcal{P}_j^{[\theta]}(u,v) \tag{3}$$

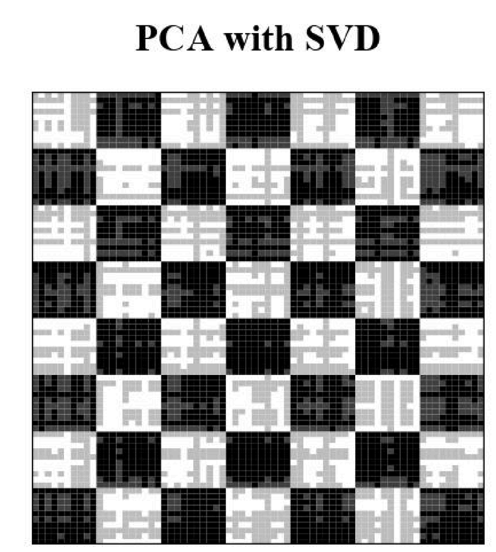
- where J is the number of penalty components, and θ is the vector of all tuning parameters. The composite penalty $\sum_{j=1}^{J} \mathcal{P}_j^{(\theta)}(u,v)$ lets us mix regularizers, e.g., smoothness with $\theta = (\boldsymbol{\alpha}_u, \boldsymbol{\alpha}_v)$ and sparsity with $\theta = (\gamma_u, \gamma_v)$ (controlling sparsity), and can include other structures as needed.
- **Sequential Power Algorithm:**
  1. Initialize u, v using rank-one SVD of $\mathbf{X}$.
  2. Update u with smoothing and sparsity transformations:
     $u \leftarrow S_u^{[\alpha_u]} h_u^{[\gamma_u]}(\mathbf{X}v)$
  3. Update v similarly:
     $v \leftarrow S_v^{[\boldsymbol{\alpha}_v]} h_v^{[\gamma_v]}(\mathbf{X}^\top u)$
  4. Normalize v and deflate $\mathbf{X}$ to extract further components.
- **Tuning:**
  - $\boldsymbol{\alpha}_u, \boldsymbol{\alpha}_v \longrightarrow$ control **smoothness** of scores and loadings.
  - $\gamma_u, \gamma_v \longrightarrow$ control **sparsity** of scores and loadings.
  - **Strategy (Conditional Tuning):**
    1. Initialize with no penalties.
    2. Use **cross-validation (CV)** to tune sparsity parameters.
    3. Use **generalized cross-validation (GCV)** to tune smoothness parameters.
    4. Alternate steps 2–3 until convergence for an optimal balance of smoothness and sparsity.



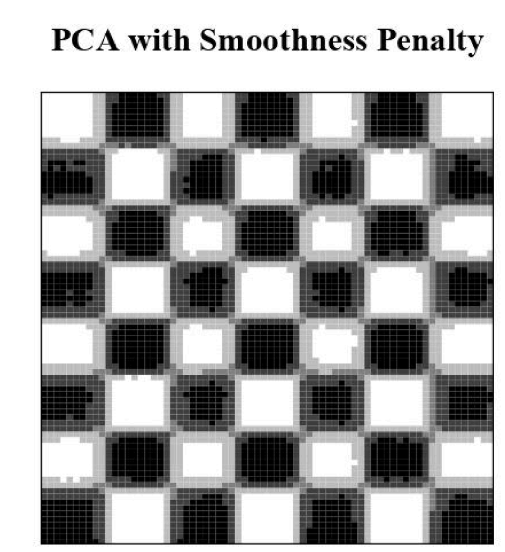
True



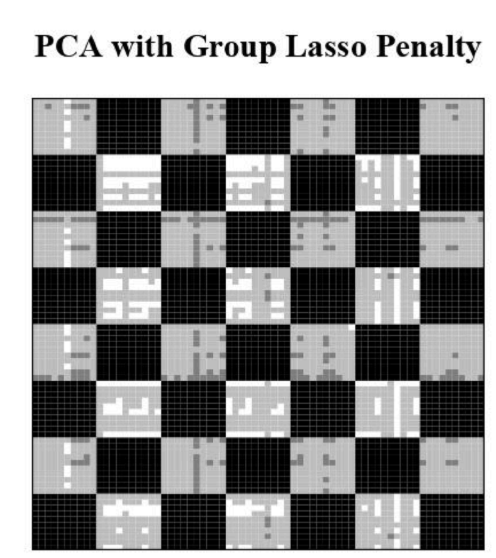
Noisy



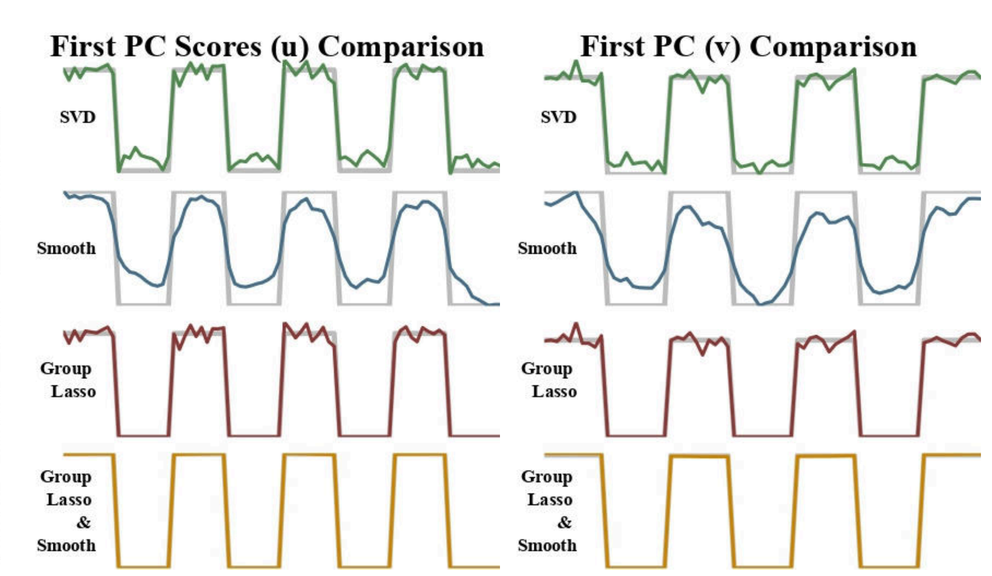
SVD reconstructed



Smoothed reconstructed



Group LASSO reconstructed



Scores (u) Loadings (v)

IV. References

Haghbin, H., Zhao, Y., & Maadooliat, M. (2025). Regularized multivariate functional principal component analysis for data observed on different domains. *Foundations of Data Science*. Retrieved from https://www.aimsciences.org/article/id/68b562c1bd10eb1421fa6ef0

Happ, C., & Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, *113*(522), 649–659. Informa UK Limited.

Huang, J. Z., Shen, H., & Buja, A. (2008). Functional principal components analysis via penalized rank one approximation. *Electronic Journal of Statistics*, *2*, 678–695. Institute of Mathematical Statistics; Bernoulli Society.

https://github.com/mobinapourmoshir/ReMPCA

CSSRFP 2025

mobina.pourmoshir@marquette.edu

Nie, Y., & Cao, J. (2020). Sparse functional principal component analysis in a new regression framework. *Computational Statistics & Data Analysis*, *152*, 107016.

Ramsay, J., & Silverman, B. W. (2005). *Functional data analysis*. Springer series in statistics. Springer.

Shen, H., & Huang, J. Z. (2008). Sparse principal component analysis via regularized low rank matrix approximation. *Journal of Multivariate Analysis*, *99*(6), 1015–1034.

Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *The Annals of Statistics*, *24*(1), 1–24. Institute of Mathematical Statistics.