Regularized Estimation of Spatial Patterns

Wen-Ting Wang

June 18, 2017

Joint work with Hsin-Cheng Huang @ Academia Sinica

Outline

1. Background
2. Proposed Method: Spatial MCA
3. Numerical Example
4. Summary

Climate Change

Climate Change

increases the odds of **extreme weather events** occuring,

Flood

Drought

Climate Change

affects **human health** and **quality of life**

Drought in the East Africa

2011

10,000

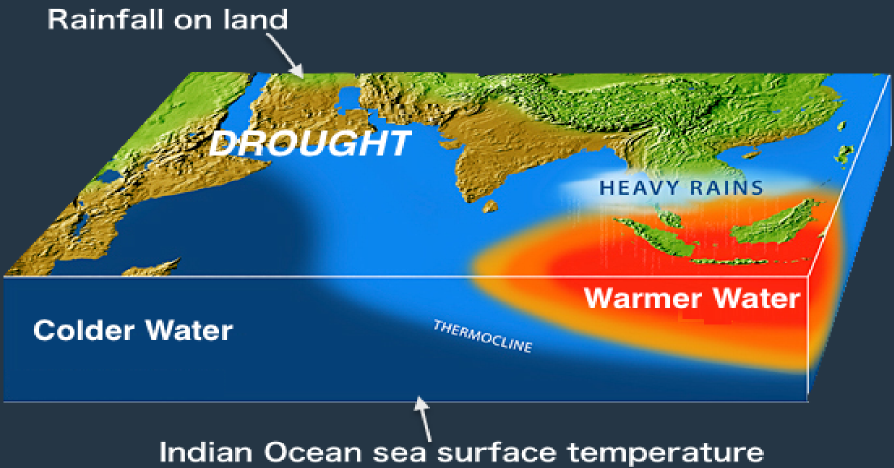
People have been killed
by the worst drought in 60 years.

Climate Change

are associated with atmospheric dynamics

Atmospheric dynamics

can be studied through spatial patterns


Rainfall on land
DROUGHT
HEAVY RAINS
Warmer Water
Colder Water
THERMOCLINE
Indian Ocean sea surface temperature

How rainfalls in East Africa
are affected by
sea surface temperature in the Indian Ocean?

- Analyze this problem via their coupled spatial patterns
- Ref: Omondi et al.,2013

Outline

1. **Background**
2. Proposed Method: Spatial MCA
3. Numerical Example
4. Summary

Background

- Bivariate spatial processes:

$$\{(\eta_{1i}(\boldsymbol{s}_1), \eta_{2i}(\boldsymbol{s}_2)) : \boldsymbol{s}_1 \in D_1, \boldsymbol{s}_2 \in D_2\}; \quad i = 1, \dots, n$$

  - $D_1, D_2 \subset \mathbb{R}^d$
  - $\eta_{11}(\boldsymbol{s}_1), \dots, \eta_{1n}(\boldsymbol{s}_1)$: uncorrelated and mean zero
  - $\eta_{21}(\boldsymbol{s}_2), \dots, \eta_{2n}(\boldsymbol{s}_2)$: uncorrelated and mean zero
  - common spatial covariance function:
    - $C_{11}(\boldsymbol{s}_1, \boldsymbol{s}_1^*) = \text{cov}(\eta_{1i}(\boldsymbol{s}_1), \eta_{1i}(\boldsymbol{s}_1^*))$
    - $C_{22}(\boldsymbol{s}_2, \boldsymbol{s}_2^*) = \text{cov}(\eta_{2i}(\boldsymbol{s}_2), \eta_{2i}(\boldsymbol{s}_2^*))$
    - $C_{12}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \text{cov}(\eta_{1i}(\boldsymbol{s}_1), \eta_{2i}(\boldsymbol{s}_2))$

Background

- Data at locations $\boldsymbol{s}_{11}, \ldots, \boldsymbol{s}_{1p_1} \in D_1$ and $\boldsymbol{s}_{21}, \ldots, \boldsymbol{s}_{2p_2} \in D_2$
    - $$Y_{1i}(\boldsymbol{s}_{1j}) = \eta_{1i}(\boldsymbol{s}_{1j}) + \epsilon_{1ij};\, j = 1, \ldots, p_1$$
        - $\epsilon_{1ij} \sim (0, \sigma_1^2)$
        - ϵ_{1ij}: uncorrelated with $\eta_1(\cdot)$
    - $$Y_{2i}(\boldsymbol{s}_{2j}) = \eta_{2i}(\boldsymbol{s}_{2j}) + \epsilon_{2ij};\, j = 1, \ldots, p_2$$
        - $\epsilon_{2ij} \sim (0, \sigma_1^2)$
        - ϵ_{2ij}: uncorrelated with $\eta_2(\cdot)$
- $i = 1, \ldots, n$

Aims

- Find dominant coupled patterns between $\eta_{1i}(\cdot)$ and $\eta_{2i}(\cdot)$
  - to study how variations of $\eta_{1i}(\cdot)$ affect $\eta_{2i}(\cdot)$

Rank-K Cross-Covariance Model

- $D_1, D_2 \subset \mathbb{R}^d$: continuous domain
- Azaïez and Belgacem (2015),

$$C_{12}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \text{cov}(\eta_{1i}(\boldsymbol{s}_1), \eta_{2i}(\boldsymbol{s}_2)) = \sum_{k=1}^{\infty} d_k u_k(\boldsymbol{s}_1) v_k(\boldsymbol{s}_2)$$

  - nonnegative singular values: $d_1 \geq d_2 \geq \cdots$
  - $\{u_k(\cdot)\}$ and $\{v_k(\cdot)\}$: sets of orthonormal basis functions
  - similar to the Karhunen-Loéve expansion

Rank-K Cross-Covariance Model

- $D_1, D_2 \subset \mathbb{R}^d$: continuous domain
- Azaïez and Belgacem (2015),

$$C_{12}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \text{cov}(\eta_{1i}(\boldsymbol{s}_1), \eta_{2i}(\boldsymbol{s}_2)) = \sum_{k=1}^{\infty} d_k u_k(\boldsymbol{s}_1) v_k(\boldsymbol{s}_2)$$

  - nonnegative singular values: $d_1 \geq d_2 \geq \cdots$
  - $\{u_k(\cdot)\}$ and $\{v_k(\cdot)\}$: sets of orthonormal basis functions
  - similar to the Karhunen-Loéve expansion

- Assume $d_{K+1} = 0$.
- $(u_1(\cdot), v_1(\cdot)), \ldots, (u_K(\cdot), v_K(\cdot))$: K dominant coupled patterns

Goal

- Find $(u_1(\cdot), v_1(\cdot)), \ldots, (u_K(\cdot), v_K(\cdot))$ as K coupled patterns

Goal

- Find $(u_1(\cdot), v_1(\cdot)), \ldots, (u_K(\cdot), v_K(\cdot))$ as K coupled patterns
- Common approach: maximum covariance analysis (MCA)

Bivariate Data Vector

- $$\begin{pmatrix} \boldsymbol{Y}_{1i} \\ \boldsymbol{Y}_{2i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\eta}_{1i} \\ \boldsymbol{\eta}_{2i} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_{1i} \\ \boldsymbol{\epsilon}_{2i} \end{pmatrix}; i = 1, \ldots, n$$
  - $$\begin{pmatrix} \boldsymbol{Y}_{1i} \\ \boldsymbol{Y}_{2i} \end{pmatrix} = \begin{pmatrix} (Y_{1i}(\boldsymbol{s}_{11}), \ldots, Y_{1i}(\boldsymbol{s}_{1p_1}))' \\ (Y_{2i}(\boldsymbol{s}_{21}), \ldots, Y_{2i}(\boldsymbol{s}_{2p_2}))' \end{pmatrix}$$
  - $$\begin{pmatrix} \boldsymbol{\eta}_{1i} \\ \boldsymbol{\eta}_{2i} \end{pmatrix} = \begin{pmatrix} (\eta_{1i}(\boldsymbol{s}_{11}), \ldots, \eta_{1i}(\boldsymbol{s}_{1p_1}))' \\ (\eta_{2i}(\boldsymbol{s}_{21}), \ldots, \eta_{2i}(\boldsymbol{s}_{2p_2}))' \end{pmatrix}$$
  - $$\begin{pmatrix} \boldsymbol{\epsilon}_{1i} \\ \boldsymbol{\epsilon}_{2i} \end{pmatrix} = \begin{pmatrix} (\epsilon_{1i1}, \ldots, \epsilon_{1ip_1})' \\ (\epsilon_{2i1}, \ldots, \epsilon_{2ip_2})' \end{pmatrix}$$
- Assume $p_1 \geq p_2$

Maximum Covariance Analysis (MCA)

- Bivariate data vector

$$\begin{pmatrix} \boldsymbol{Y}_{1i} \\ \boldsymbol{Y}_{2i} \end{pmatrix} \stackrel{i.i.d.}{\sim} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right)$$

  - $\boldsymbol{\Sigma}_{12} = \text{cov}(\boldsymbol{Y}_{1i}, \boldsymbol{Y}_{2i}) = \text{cov}(\boldsymbol{\eta}_{1i}, \boldsymbol{\eta}_{2i})$

- Idea: find $\boldsymbol{u} \in \mathcal{R}^{p_1}$ and $\boldsymbol{v} \in \mathcal{R}^{p_2}$, with $\|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2 = 1$, which maximize

$$d = \text{cov}(\boldsymbol{u}'\boldsymbol{Y}_{1i}, \boldsymbol{v}'\boldsymbol{Y}_{2i}) = \boldsymbol{u}'\boldsymbol{\Sigma}_{12}\boldsymbol{v}$$

Maximum Covariance Analysis (MCA)

- Bivariate data vector

$$\begin{pmatrix} \boldsymbol{Y}_{1i} \\ \boldsymbol{Y}_{2i} \end{pmatrix} \overset{i.i.d.}{\sim} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right)$$

  - $\boldsymbol{\Sigma}_{12} = \text{cov}(\boldsymbol{Y}_{1i}, \boldsymbol{Y}_{2i}) = \text{cov}(\boldsymbol{\eta}_{1i}, \boldsymbol{\eta}_{2i})$

- Singular value decomposition (SVD): $\boldsymbol{\Sigma}_{12} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}'$
  - Singular values: $\boldsymbol{D}_{K \times K} = \text{diag}(d_1, \ldots, d_K); d_1 \geq \cdots \geq d_K > 0$
  - Left singular vectors: $\boldsymbol{U}_{p_1 \times K} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K\}$
  - Right singular vectors: $\boldsymbol{V}_{p_2 \times K} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K\}$

Maximum Covariance Analysis (MCA)

- Bivariate data vector

$$\begin{pmatrix} \boldsymbol{Y}_{1i} \\ \boldsymbol{Y}_{2i} \end{pmatrix} \overset{i.i.d.}{\sim} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right)$$

  - $\boldsymbol{\Sigma}_{12} = \text{cov}(\boldsymbol{Y}_{1i}, \boldsymbol{Y}_{2i}) = \text{cov}(\boldsymbol{\eta}_{1i}, \boldsymbol{\eta}_{2i})$

- Singular value decomposition (SVD): $\boldsymbol{\Sigma}_{12} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}'$
  - Singular values: $\boldsymbol{D}_{K\times K} = \text{diag}(d_1, \ldots, d_K); d_1 \geq \cdots \geq d_K > 0$
  - Left singular vectors: $\boldsymbol{U}_{p_1\times K} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K\}$
  - Right singular vectors: $\boldsymbol{V}_{p_2\times K} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K\}$
- Coupled pattern: $(\boldsymbol{u}_1, \boldsymbol{v}_1), \ldots, (\boldsymbol{u}_K, \boldsymbol{v}_K)$

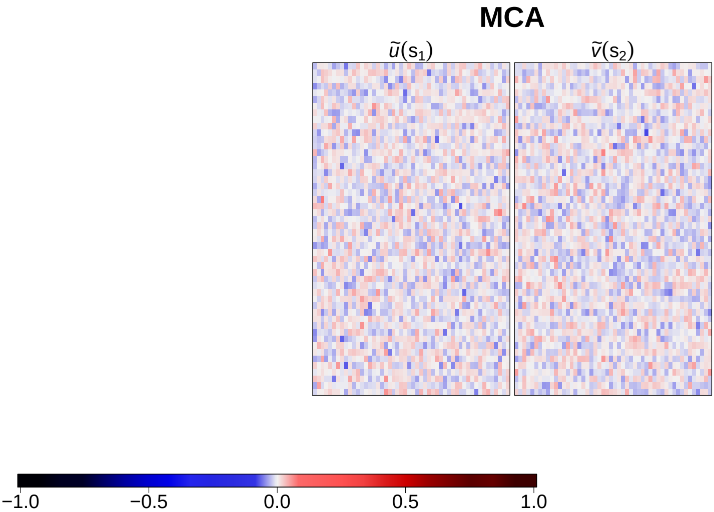
Sample Maximum Covariance Analysis

- Bivariate data matrix: $(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$
  - $\boldsymbol{Y}_1 = (\boldsymbol{Y}_{11}, \ldots, \boldsymbol{Y}_{1n})'$; $\boldsymbol{Y}_2 = (\boldsymbol{Y}_{21}, \ldots, \boldsymbol{Y}_{2n})'$
- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{Y}_1'\boldsymbol{Y}_2/n$
- Singular value decomposition (SVD): $\boldsymbol{S}_{12} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}'$
  - $\tilde{\boldsymbol{D}}_{p_2 \times p_2} = \text{diag}(\tilde{d}_1 \ldots \tilde{d}_{p_2})$; $\tilde{d}_1 \geq \cdots \geq \tilde{d}_{p_2} \geq 0$
  - $\tilde{\boldsymbol{U}}_{p_1 \times p_2} = \{\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_{p_2}\}$
  - $\tilde{\boldsymbol{V}}_{p_2 \times p_2} = \{\tilde{\boldsymbol{v}}_1, \ldots, \tilde{\boldsymbol{v}}_{p_2}\}$
- $(\tilde{\boldsymbol{u}}_1, \tilde{\boldsymbol{v}}_1), \ldots, (\tilde{\boldsymbol{u}}_K, \tilde{\boldsymbol{v}}_K)$: estimates of $(\boldsymbol{u}_1, \boldsymbol{v}_1), \ldots, (\boldsymbol{u}_K, \boldsymbol{v}_K)$

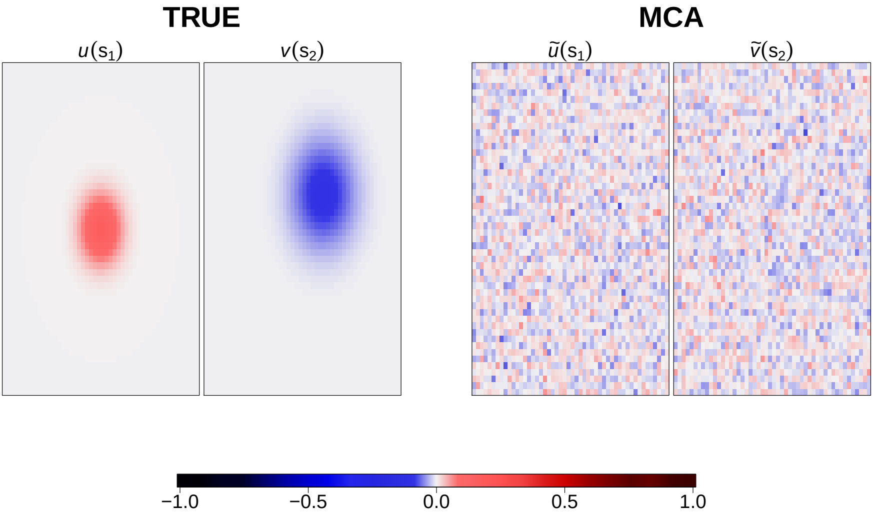
Sample Maximum Covariance Analysis

- Bivariate data matrix: $(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$
  - $\boldsymbol{Y}_1 = (\boldsymbol{Y}_{11}, \ldots, \boldsymbol{Y}_{1n})'$; $\boldsymbol{Y}_2 = (\boldsymbol{Y}_{21}, \ldots, \boldsymbol{Y}_{2n})'$
- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{Y}_1'\boldsymbol{Y}_2/n$
- Singular value decomposition (SVD): $\boldsymbol{S}_{12} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}'$
  - $\tilde{\boldsymbol{D}}_{p_2 \times p_2} = \text{diag}(\tilde{d}_1 \ldots \tilde{d}_{p_2})$; $\tilde{d}_1 \geq \cdots \geq \tilde{d}_{p_2} \geq 0$
  - $\tilde{\boldsymbol{U}}_{p_1 \times p_2} = \{\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_{p_2}\}$
  - $\tilde{\boldsymbol{V}}_{p_2 \times p_2} = \{\tilde{\boldsymbol{v}}_1, \ldots, \tilde{\boldsymbol{v}}_{p_2}\}$
- $(\tilde{\boldsymbol{u}}_1, \tilde{\boldsymbol{v}}_1), \ldots, (\tilde{\boldsymbol{u}}_K, \tilde{\boldsymbol{v}}_K)$: estimates of $(\boldsymbol{u}_1, \boldsymbol{v}_1), \ldots, (\boldsymbol{u}_K, \boldsymbol{v}_K)$
- Problem:
  - high estimation variability: n small, p_1 or p_2 large
    - noisy patterns $\rightarrow$ low interpretation
  - without a spatial structure of $(\boldsymbol{u}, \boldsymbol{v})$

Example:



Example:



Motivation

To enhance the interpretability, we implement MCA by incorporating

Motivation

To enhance the interpretability, we implement MCA by incorporating

1. spatial structure of $(u_k(\cdot), v_k(\cdot))$

Motivation

To enhance the interpretability, we implement MCA by incorporating

1. spatial structure of $(u_k(\cdot), v_k(\cdot))$
2. spatial localized patterns of $(u_k(\cdot), v_k(\cdot))$

Motivation

To enhance the interpretability, we implement MCA by incorporating

1. spatial structure of $(u_k(\cdot), v_k(\cdot))$
2. spatial localized patterns of $(u_k(\cdot), v_k(\cdot))$

retain that orthogonal constraints for $(\boldsymbol{u}, \boldsymbol{v})$.

Outline

Quick Recap

- Data: $\boldsymbol{Y}_{\ell i} = (Y_{\ell i}(\boldsymbol{s}_{\ell 1}), \ldots, Y_{\ell i}(\boldsymbol{s}_{\ell p_\ell}))'$, $i = 1, \ldots, n$
  - $$Y_{\ell i}(\boldsymbol{s}_{\ell j}) = \eta_{\ell i}(\boldsymbol{s}_{\ell j}) + \epsilon_{\ell ij};\ j = 1, \ldots, p_\ell$$
    - $\epsilon_{\ell ij} \sim (0, \sigma_\ell^2)$
    - $\epsilon_{\ell ij}$: uncorrelated with $\eta_\ell(\cdot)$
  - $\ell = 1, 2$

Quick Recap

- Data: $\boldsymbol{Y}_{\ell i} = (Y_{\ell i}(\boldsymbol{s}_{\ell 1}), \ldots, Y_{\ell i}(\boldsymbol{s}_{\ell p_\ell}))'$, $i = 1, \ldots, n$
  - $$Y_{\ell i}(\boldsymbol{s}_{\ell j}) = \eta_{\ell i}(\boldsymbol{s}_{\ell j}) + \epsilon_{\ell ij};\ j = 1, \ldots, p_\ell$$
    - $\epsilon_{\ell ij} \sim (0, \sigma_\ell^2)$
    - $\epsilon_{\ell ij}$: uncorrelated with $\eta_\ell(\cdot)$
  - $\ell = 1, 2$
- Spatial cross-covariance function:

$$C_{12}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \text{cov}(\eta_{1i}(\boldsymbol{s}_1), \eta_{2i}(\boldsymbol{s}_2)) = \sum_{k=1}^{K} d_k u_k(\boldsymbol{s}_1) v_k(\boldsymbol{s}_2)$$

  - $d_1 \geq \cdots \geq d_K \geq 0$
  - $u_1(\cdot), \ldots, u_K(\cdot)$: K unknown orthonormal functions
  - $v_1(\cdot), \ldots, v_K(\cdot)$: K unknown orthonormal functions

MCA (alternative version)

- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{Y}_1'\boldsymbol{Y}_2/n$
- MCA: perform SVD of $\boldsymbol{S}_{12}$
- Alternative method:

$$(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) = \underset{\boldsymbol{U},\boldsymbol{V}}{\arg\max}\, \mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}),$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

  - $\boldsymbol{U}_{p_1 \times K} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K)$ with $u_{jk} = u_k(\boldsymbol{s}_{1j})$
  - $\boldsymbol{V}_{p_2 \times K} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K)$ with $v_{jk} = v_k(\boldsymbol{s}_{2j})$

Regularized MCA

- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{X}'\boldsymbol{Y}/n$
- $\boldsymbol{U}_{p_1 \times K} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K)$ with $u_{jk} = u_k(\boldsymbol{s}_{1j})$
- $\boldsymbol{V}_{p_2 \times K} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K)$ with $v_{jk} = v_k(\boldsymbol{s}_{2j})$

- Objective function:

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V})$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

Regularized MCA

- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{X}'\boldsymbol{Y}/n$
- $\boldsymbol{U}_{p_1 \times K} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K)$ with $u_{jk} = u_k(\boldsymbol{s}_{1j})$
- $\boldsymbol{V}_{p_2 \times K} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K)$ with $v_{jk} = v_k(\boldsymbol{s}_{2j})$
- Objective function:

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K} \left\{ \tau_{1u} J(u_k) + \tau_{2u} \sum_{j=1}^{p_1} |u_k(\boldsymbol{s}_{1j})| + \tau_{1v} J(v_k) + \tau_{2v} \sum_{j=1}^{p_2} |v_k(\boldsymbol{s}_{2j})| \right\}$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

Regularized MCA

- Sample cross-covariance matrix: $\boldsymbol{S}_{12} = \boldsymbol{X}'\boldsymbol{Y}/n$
- $\boldsymbol{U}_{p_1 \times K} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K)$ with $u_{jk} = u_k(\boldsymbol{s}_{1j})$
- $\boldsymbol{V}_{p_2 \times K} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K)$ with $v_{jk} = v_k(\boldsymbol{s}_{2j})$
- Objective function:

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K} \left\{ \tau_{1u} J(u_k) + \tau_{2u} \sum_{j=1}^{p_1} |u_k(\boldsymbol{s}_{1j})| + \tau_{1v} J(v_k) + \tau_{2v} \sum_{j=1}^{p_2} |v_k(\boldsymbol{s}_{2j})| \right\}$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

- $J(u_k(\cdot)) = \displaystyle\sum_{z_1+\cdots+z_d=2} \int_{R^d} \left(\frac{\partial^2 u_k(\boldsymbol{s})}{\partial x_1^{z_1} \ldots \partial x_d^{z_d}} \right)^2 d\boldsymbol{s}$
  - $\boldsymbol{s} = (x_1, \ldots, x_d)'$
- τ_{1u}, τ_{1v}: smoothness parameter
- τ_{2u}, τ_{2v}: sparseness parameter

Spatial MCA (SpatMCA)

- $J(u_k(\cdot)) = \boldsymbol{u}_k' \boldsymbol{\Omega}_1 \boldsymbol{u}_k, J(v_k(\cdot)) = \boldsymbol{v}_k' \boldsymbol{\Omega}_2 \boldsymbol{v}_k$
  - $\boldsymbol{\Omega}_1 \succ \boldsymbol{0}$: determined only by $\boldsymbol{s}_{11}, \ldots, \boldsymbol{s}_{1p_1}$
  - $\boldsymbol{\Omega}_2 \succ \boldsymbol{0}$: determined only by $\boldsymbol{s}_{21}, \ldots, \boldsymbol{s}_{2p_2}$
  - Green and Silverman (1994)

Spatial MCA (SpatMCA)

- $J(u_k(\cdot)) = \boldsymbol{u}_k'\boldsymbol{\Omega}_1\boldsymbol{u}_k,\ J(v_k(\cdot)) = \boldsymbol{v}_k'\boldsymbol{\Omega}_2\boldsymbol{v}_k$
  - $\boldsymbol{\Omega}_1 \succ \mathbf{0}$: determined only by $\boldsymbol{s}_{11}, \dots, \boldsymbol{s}_{1p_1}$
  - $\boldsymbol{\Omega}_2 \succ \mathbf{0}$: determined only by $\boldsymbol{s}_{21}, \dots, \boldsymbol{s}_{2p_2}$
  - Green and Silverman (1994)
- SpatMCA: $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}})$ maximizes:

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K}\left\{\tau_{1u}\boldsymbol{u}_k'\boldsymbol{\Omega}_1\boldsymbol{u}_k + \tau_{2u}\sum_{j=1}^{p_1}|u_{jk}| + \tau_{1v}\boldsymbol{v}_k'\boldsymbol{\Omega}_2\boldsymbol{v}_k + \tau_{2v}\sum_{j=1}^{p_2}|v_{jk}|\right\}$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

Spatial MCA (SpatMCA)

- $J(u_k(\cdot)) = \boldsymbol{u}_k'\boldsymbol{\Omega}_1\boldsymbol{u}_k$, $J(v_k(\cdot)) = \boldsymbol{v}_k'\boldsymbol{\Omega}_2\boldsymbol{v}_k$
  - $\boldsymbol{\Omega}_1 \succ \mathbf{0}$: determined only by $\boldsymbol{s}_{11}, \ldots, \boldsymbol{s}_{1p_1}$
  - $\boldsymbol{\Omega}_2 \succ \mathbf{0}$: determined only by $\boldsymbol{s}_{21}, \ldots, \boldsymbol{s}_{2p_2}$
  - Green and Silverman (1994)

- SpatMCA: $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}})$ maximizes:

$$\text{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K}\left\{\tau_{1u}\boldsymbol{u}_k'\boldsymbol{\Omega}_1\boldsymbol{u}_k + \tau_{2u}\sum_{j=1}^{p_1}|u_{jk}| + \tau_{1v}\boldsymbol{v}_k'\boldsymbol{\Omega}_2\boldsymbol{v}_k + \tau_{2v}\sum_{j=1}^{p_2}|v_{jk}|\right\}$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

- As $\tau_{1u} = \tau_{1v} = \tau_{2u} = \tau_{2v} = 0$, $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}) = (\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}})$ (sample MCA)

SpatMCA: $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \ldots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$

- $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \ldots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$ maximize

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K} \left\{ \tau_{1u} J(u_k) + \tau_{2u} \sum_{j=1}^{p_1} |u_k(\boldsymbol{s}_{1j})| + \tau_{1v} J(v_k) + \tau_{2v} \sum_{j=1}^{p_2} |v_k(\boldsymbol{s}_{2j})| \right\},$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

SpatMCA: $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \dots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$

- $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \dots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$ maximize

$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K}\left\{\tau_{1u}J(u_k)+\tau_{2u}\sum_{j=1}^{p_1}|u_k(\boldsymbol{s}_{1j})| + \tau_{1v}J(v_k)+\tau_{2v}\sum_{j=1}^{p_2}|v_k(\boldsymbol{s}_{2j})|\right\},$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

- $$\hat{u}_k(\boldsymbol{s}_1) = \sum_{i=1}^{p_1} a_{1i}g(\|\boldsymbol{s}_1 - \boldsymbol{s}_{1i}\|) + b_{10} + \sum_{j=1}^{d} b_{1j}x_{1j}$$
$$\hat{v}_k(\boldsymbol{s}_2) = \sum_{i=1}^{p_2} a_{2i}g(\|\boldsymbol{s}_2 - \boldsymbol{s}_{2i}\|) + b_{20} + \sum_{j=1}^{d} b_{2j}x_{2j}$$

SpatMCA: $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \ldots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$

- $(\hat{u}_1(\cdot), \hat{v}_1(\cdot)), \ldots, (\hat{u}_K(\cdot), \hat{v}_K(\cdot))$ maximize

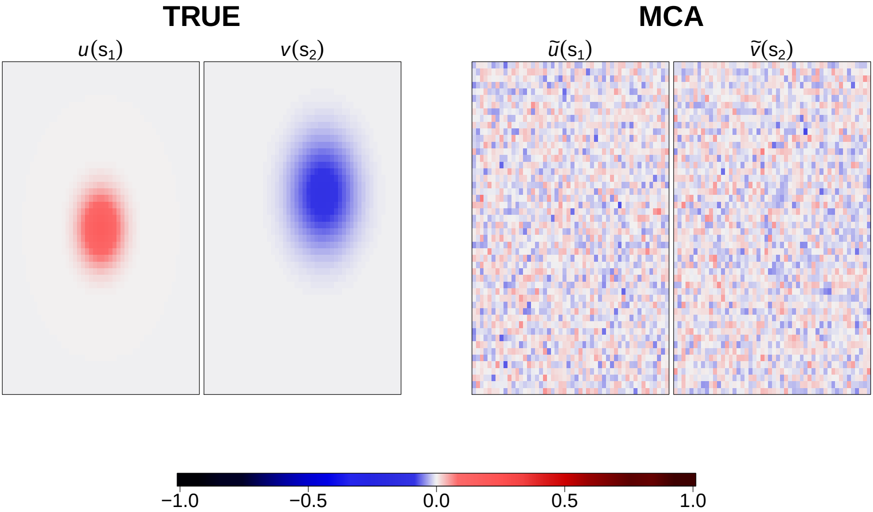
$$\mathrm{tr}(\boldsymbol{U}'\boldsymbol{S}_{12}\boldsymbol{V}) - \sum_{k=1}^{K}\left\{\tau_{1u}J(u_k)+\tau_{2u}\sum_{j=1}^{p_1}|u_k(\boldsymbol{s}_{1j})| + \tau_{1v}J(v_k)+\tau_{2v}\sum_{j=1}^{p_2}|v_k(\boldsymbol{s}_{2j})|\right\},$$

subject to $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_K$

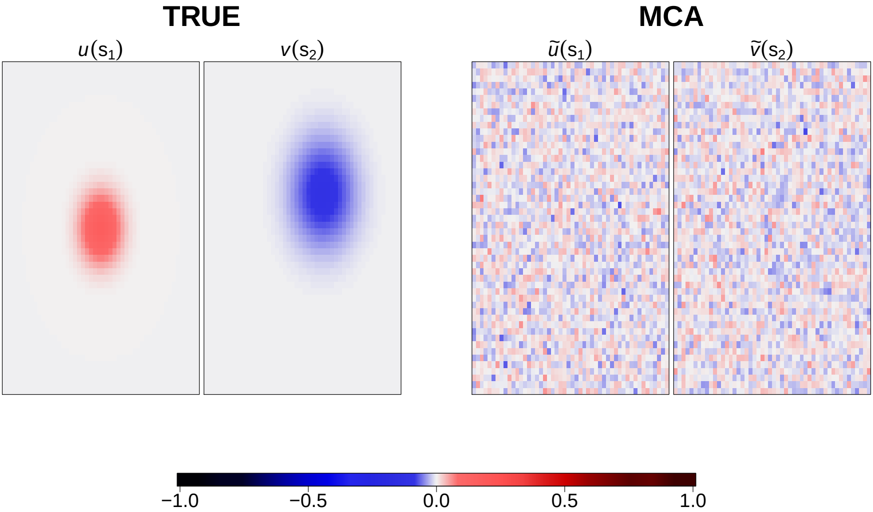
- $$\hat{u}_k(\boldsymbol{s}_1) = \sum_{i=1}^{p_1} a_{1i}g(\|\boldsymbol{s}_1 - \boldsymbol{s}_{1i}\|) + b_{10} + \sum_{j=1}^{d} b_{1j}x_{1j}$$
  $$\hat{v}_k(\boldsymbol{s}_2) = \sum_{i=1}^{p_2} a_{2i}g(\|\boldsymbol{s}_2 - \boldsymbol{s}_{2i}\|) + b_{20} + \sum_{j=1}^{d} b_{2j}x_{2j}$$
  - $\boldsymbol{s}_1 = (x_{11}, \ldots, x_{1d})'$; $\boldsymbol{s}_2 = (x_{21}, \ldots, x_{2d})'$
  - $g(r) = \begin{cases} \frac{1}{16\pi} r^2 \log r; & \text{if } d = 2, \\ \frac{\Gamma(d/2-2)}{16\pi^{d/2}} r^{4-d}; & \text{if } d = 1, 3, \end{cases}$
  - $\boldsymbol{a}_1 = (a_{11}, \ldots, a_{1p_1})'$ and $\boldsymbol{b}_1 = (b_{10}, b_{11}, \ldots, b_{1d})'$ based on $\hat{\boldsymbol{u}}_k$
  - $\boldsymbol{a}_2 = (a_{21}, \ldots, a_{2p_2})'$ and $\boldsymbol{b}_2 = (b_{20}, b_{21}, \ldots, b_{2d})'$ based on $\hat{\boldsymbol{v}}_k$

Why roughness and Lasso penalties?

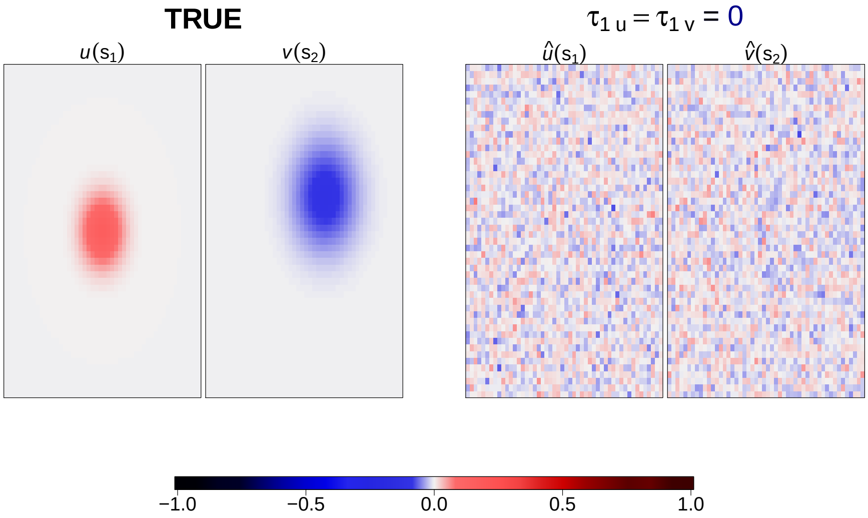
2D Example



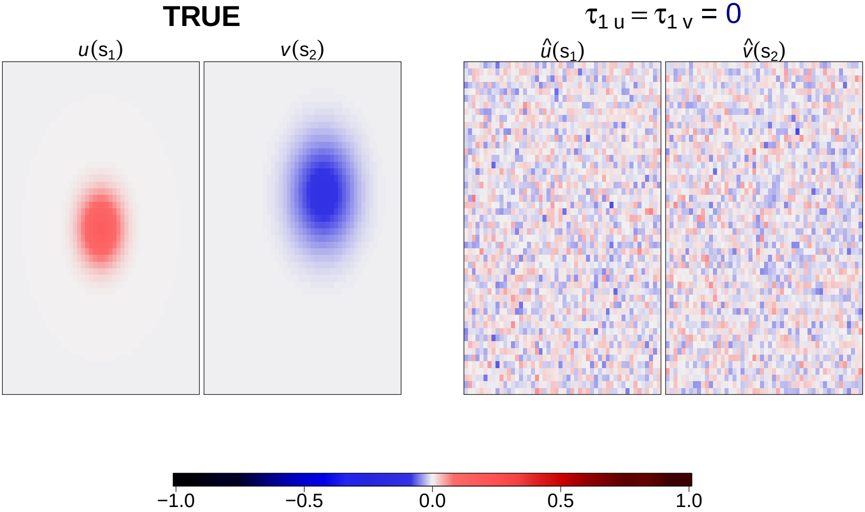
2D Example



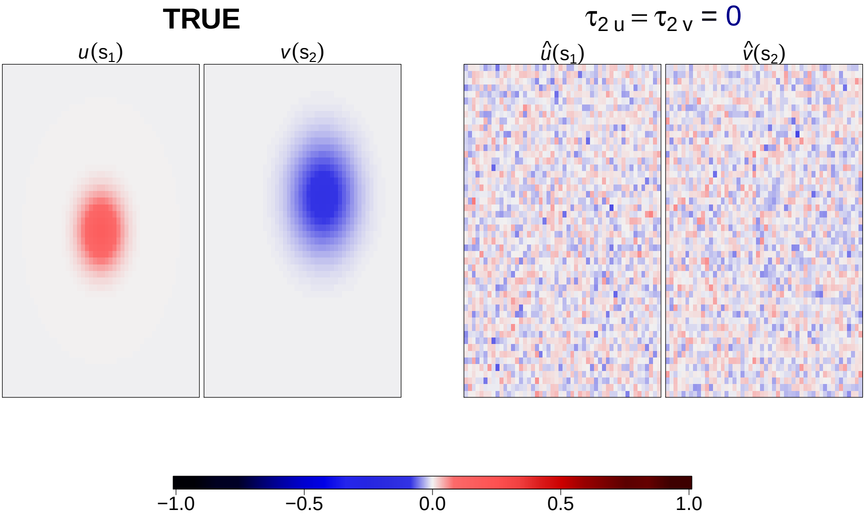
Case 1: $\tau_{2u} = \tau_{2v} = 0$ (only smoothness)



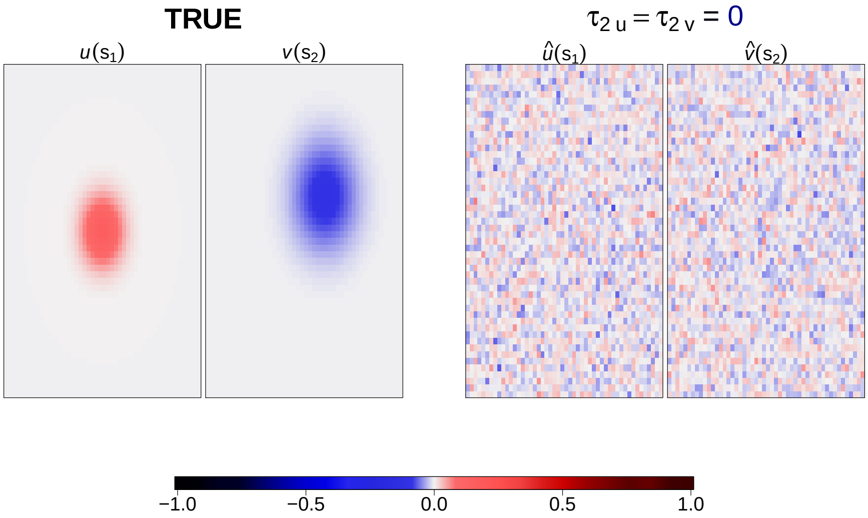
Case 1: $\tau_{2u} = \tau_{2v} = 0$ (only smoothness)



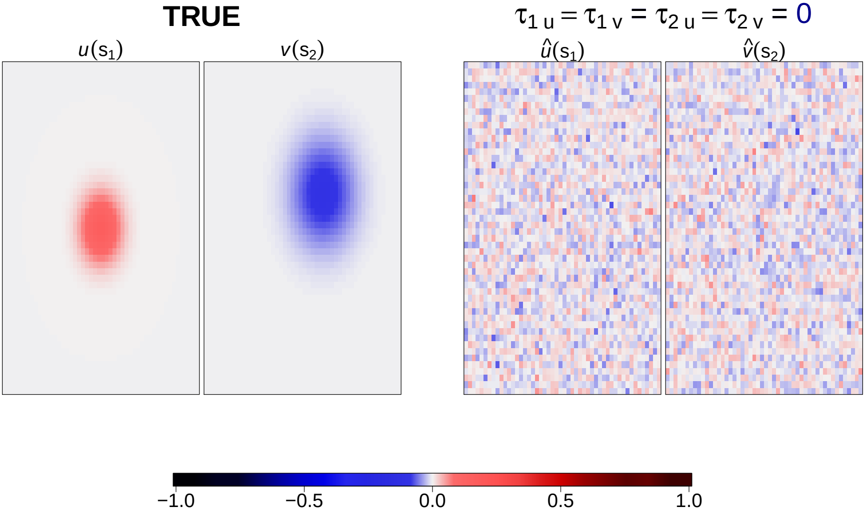
Case 2: $\tau_{1u} = \tau_{1v} = 0$ (only sparseness)



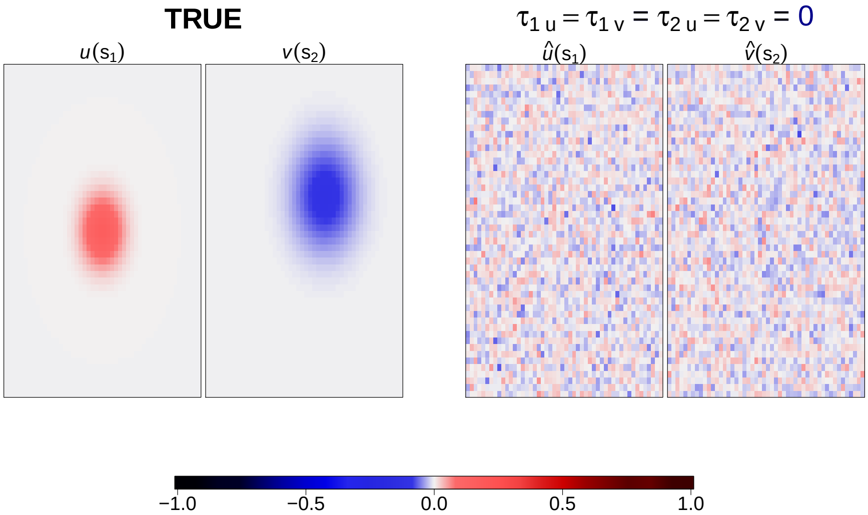
Case 2: $\tau_{1u} = \tau_{1v} = 0$ (only sparseness)



Case 3: $\tau_{1u} = \tau_{1v} = \tau_{2u} = \tau_{2v}$



Case 3: $\tau_{1u} = \tau_{1v} = \tau_{2u} = \tau_{2v}$



Estimation of D

Given the SpatMCA estimate $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}})$,

$$\hat{\boldsymbol{D}} = \underset{d_1,\ldots,d_K \geq 0}{\arg\min} \|\boldsymbol{S}_{12} - \hat{\boldsymbol{U}}\boldsymbol{D}\hat{\boldsymbol{V}}'\|_F^2 = \mathrm{diag}(\hat{d}_1, \ldots, \hat{d}_K)$$

- $\hat{d}_k = \min(\hat{u}_k' \boldsymbol{S}_{12} \hat{v}_k, 0)$

Selection of $(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v})$

- The proposed CV criterion is

$$\begin{aligned} &\mathrm{CV}(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v}) \\ &= \frac{1}{M} \sum_{m=1}^{M} \| \boldsymbol{S}_{12}^{(m)} - \hat{\boldsymbol{U}}_{\tau_{1u},\tau_{2u}}^{(-m)} \hat{\boldsymbol{D}}_{\tau_{1u},\tau_{2u},\tau_{1v},\tau_{2v}}^{(-m)} (\hat{\boldsymbol{V}}_{\tau_{1v},\tau_{2v}}^{(-m)})' \|_F^2 \end{aligned}$$

Selection of $(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v})$

- The proposed CV criterion is

$$\begin{aligned} &\text{CV}(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v}) \\ &= \frac{1}{M}\sum_{m=1}^{M} \|\boldsymbol{S}_{12}^{(m)} - \hat{\boldsymbol{U}}_{\tau_{1u},\tau_{2u}}^{(-m)} \hat{\boldsymbol{D}}_{\tau_{1u},\tau_{2u},\tau_{1v},\tau_{2v}}^{(-m)} (\hat{\boldsymbol{V}}_{\tau_{1v},\tau_{2v}}^{(-m)})' \|_F^2 \end{aligned}$$

  - Partition $\{(\boldsymbol{Y}_{11}, \boldsymbol{Y}_{21}), \ldots, (\boldsymbol{Y}_{1n}, \boldsymbol{Y}_{2n})\}$ into M parts with equal size n_M
  - $\boldsymbol{S}_{12}^{(m)} = (\boldsymbol{Y}_1^{(m)})' \boldsymbol{Y}_2^{(m)} / n_M$ based on the m-th part data $(\boldsymbol{Y}_1^{(m)}, \boldsymbol{Y}_2^{(m)})$
  - $\hat{\boldsymbol{U}}_{\tau_1,\tau_2}^{(-m)}, \hat{\boldsymbol{V}}_{\tau_1,\tau_2}^{(-m)}, \hat{\boldsymbol{D}}_{\tau_{1u},\tau_{2u},\tau_{1v},\tau_{2v}}^{(-m)}$: based on $(\boldsymbol{Y}_1^{(-m)}, \boldsymbol{Y}_2^{(-m)})$
    - $(\boldsymbol{Y}_1^{(-m)}, \boldsymbol{Y}_2^{(-m)})$: remaining data, i.e., $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ excluding $(\boldsymbol{Y}_1^{(m)}, \boldsymbol{Y}_2^{(m)})$

Selection of $(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v})$

- High computation cost to select $\{\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v}\}$ simultaneously

Selection of $(\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v})$

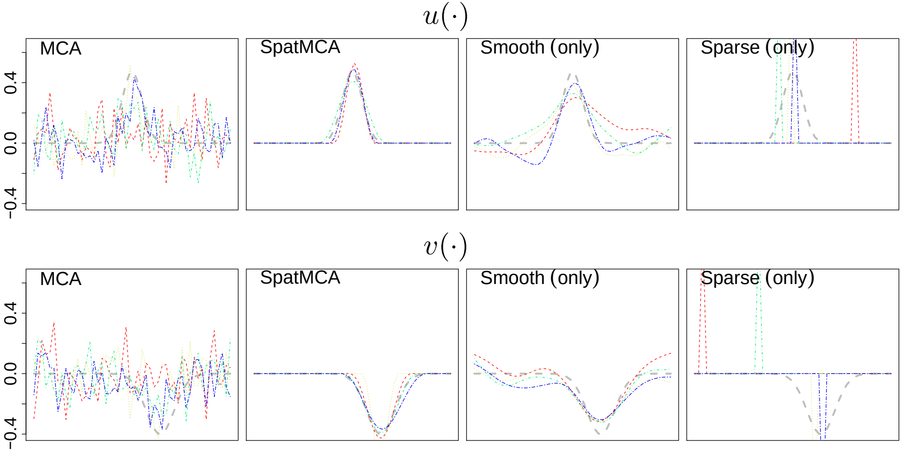
- High computation cost to select $\{\tau_{1u}, \tau_{2u}, \tau_{1v}, \tau_{2v}\}$ simultaneously
- Two-step procedure:
  1. Select smoothness parameters τ_{1u} and τ_{1v}:

$$(\hat{\tau}_{1u}, \hat{\tau}_{1v}) = \underset{\{\tau_{1u}, \tau_{1v}\} \subset [0,\infty)^2}{\arg\min} \mathrm{CV}(\tau_{1u}, 0, \tau_{1v}, 0),$$

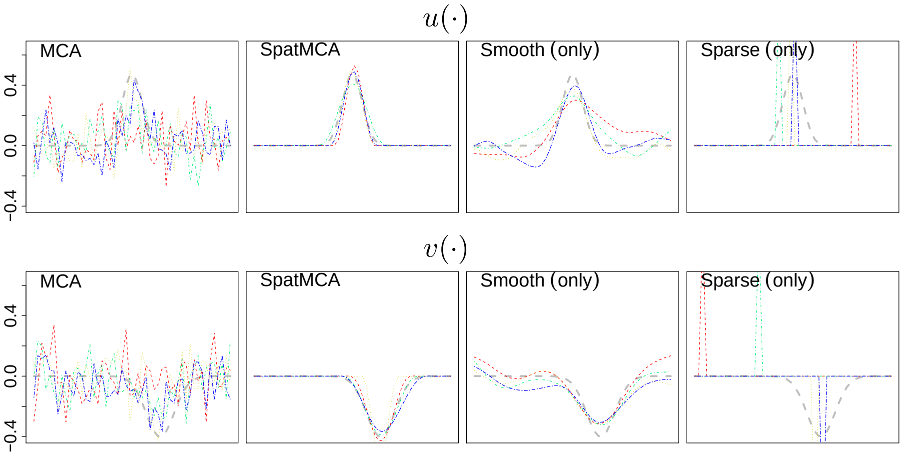
  2. Select sparseness parameters τ_{2u} and τ_{2v}:

$$(\hat{\tau}_{2u}, \hat{\tau}_{2v}) = \underset{\{\tau_{2u}, \tau_{2v}\} \subset [0,\infty)^2}{\arg\min} \mathrm{CV}(\hat{\tau}_{1u}, \tau_{2u}, \hat{\tau}_{1v}, \tau_{2v}).$$

Example (1D): 5-fold CV



Example (1D): 5-fold CV



Outline

Real data analysis

- Bivariate data:
  - Sea surface temperature (SST):
    - Region: Indian Ocean (20°N and 30°S; 20°E and 120°E)
    - Number of grids: $p_1 = 3,591$
  - Rainfall:
    - Region: East African (6°N and 12°S; 20°E and 42°E)
    - Number of grids: $p_2 = 255$
  - Time period (monthly): Jan. 2011- Dec. 2015 $\rightarrow n = 60$
  - Remove monthly mean

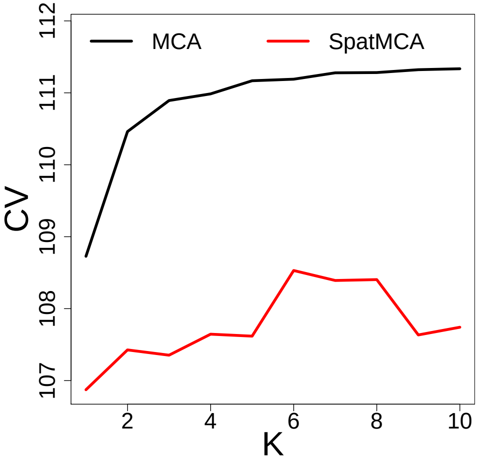
Real data analysis

- Bivariate data:
    - Sea surface temperature (SST):
        - Region: Indian Ocean (20°N and 30°S; 20°E and 120°E)
        - Number of grids: $p_1 = 3,591$
    - Rainfall:
        - Region: East African (6°N and 12°S; 20°E and 42°E)
        - Number of grids: $p_2 = 255$
    - Time period (monthly): Jan. 2011- Dec. 2015 $\rightarrow n = 60$
    - Remove monthly mean
- Goal: find coupled patterns of the SST and rainfall data
- Reference: Omondi et al. (2013)

Real Data Analysis

- Randomly decompose the data into two parts with 30 time points
  - Training data
  - Validation data
- SpatMCA: based on 5-fold CV

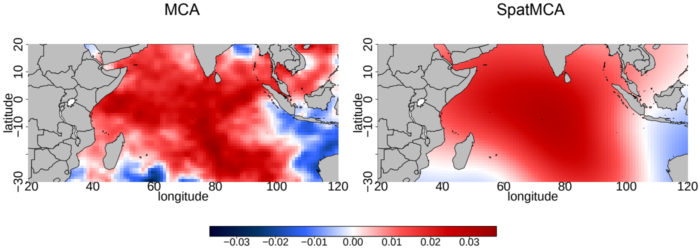
Result: CV vs. K



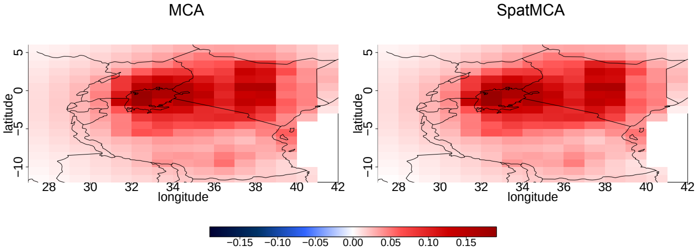
- $\hat{K} = 1$ for MCA and SpatMCA

Result: 1st Coupled Pattern

- SST

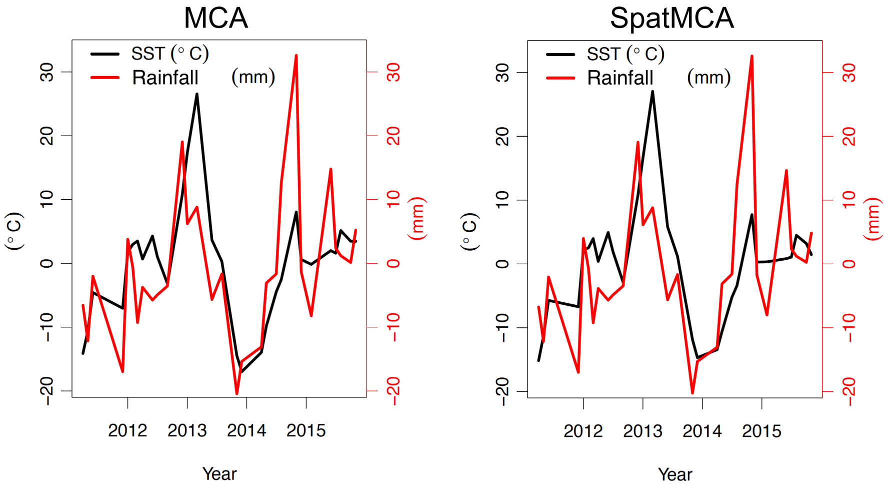


- Rainfall



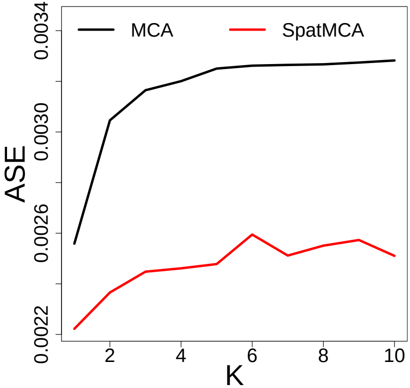
Result: 1st Maximum Covariance Variables

- 1st maximum covariance variables: $\{\hat{\boldsymbol{u}}_1' \boldsymbol{Y}_{1i}\}$; $\{\hat{\boldsymbol{v}}_1' \boldsymbol{Y}_{2i}\}$
- Pearson's correlation: 0.6 for MCA and SpatMCA



Result: Average Squared Error (ASE)

- ASE $= \frac{1}{p_1 p_2} \| \boldsymbol{S}^v_{12} - \hat{\boldsymbol{U}}_K \hat{\boldsymbol{D}}_K \hat{\boldsymbol{V}}'_K \|^2_F$
  - $\boldsymbol{S}^v_{12}$: sample cross-covariance matrix based on validation data
- Result:



Outline

Summary

SpatMCA:

- high-dimensional structure $\rightarrow$ low-dimensional structure
- with the roughness and Lasso penalties
- enhance physical interpretation, e.g., spatial localized patterns

Summary

SpatMCA:

- high-dimensional structure $\rightarrow$ low-dimensional structure
- with the roughness and Lasso penalties
- enhance physical interpretation, e.g., spatial localized patterns
- can cope with irregular spaced locations
- R package on CRAN: *SpatMCA*

Thanks for your attention!