

Deformação espacial em um modelo espaço-temporal bayesiano para respostas matriz-variadas

XIX Escola de Modelos de Regressão

Rodrigo de Souza Bulhões

Trabalho em conjunto com Marina Paez e Dani Gamerman

IME/UFBA & IM/UFRJ

João Pessoa/PB, 29 de outubro de 2025.



INSTITUTO DE
**MATEMÁTICA
E ESTATÍSTICA**
UFBA



INSTITUTO DE MATEMÁTICA
Universidade Federal do Rio de Janeiro

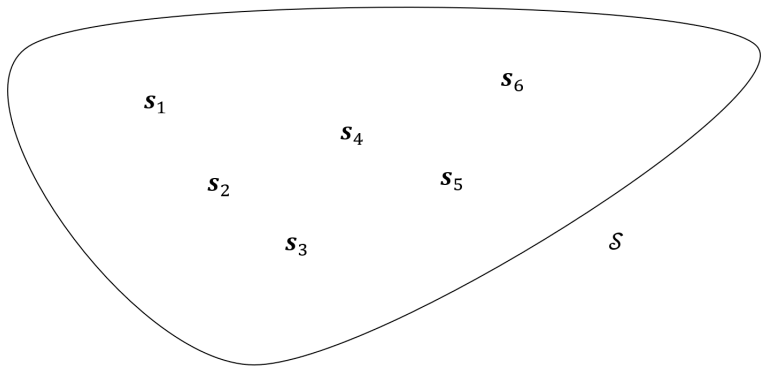
Tópicos

- 1 Introdução
- 2 Deformação espacial em um modelo espaço-temporal bayesiano para respostas matriz-variadas
- 3 Modelagem bayesiana de matrizes de respostas incompletas
- 4 Aplicação
- 5 Considerações finais e trabalhos futuros

Introdução

Introdução

Considere uma região geográfica \mathcal{S} que possui N estações de monitoramento.



Seja $\mathbf{s}_n = (\text{lon}_n, \text{lat}_n) \in \mathcal{S} \subset \mathbb{R}^2$ o par de coordenadas da n -ésima estação, $n \in \{1, \dots, N\}$.

Isotropia e anisotropia

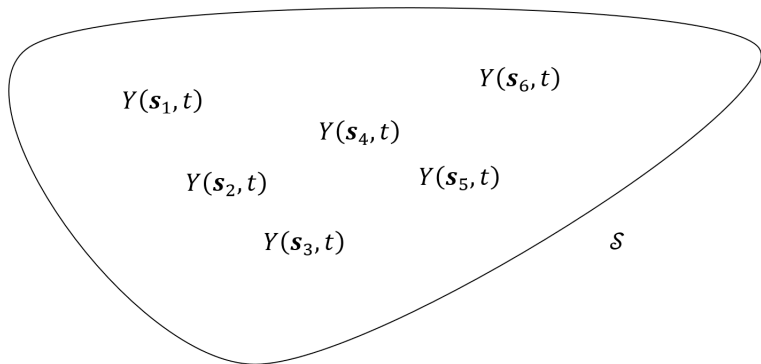
- Muitos trabalhos assumem **isotropia** ao descrever a estrutura de dependência espacial.
- Isso significa que a relação espacial entre quaisquer dois pontos depende apenas da distância que os separa.
- É razoável supor que as respostas obtidas em locais próximos sejam mais correlacionadas, mas não depende só disto.
- Na **anisotropia**, o ângulo (direção) também exerce influência!

Exemplo

- Pense na dispersão de um poluente no ar liberado por uma chaminé e o efeito do vento.
- O poluente se espalhará muito mais na direção do vento (e.g., Sudoeste) do que contra o vento ou lateralmente.
- A estrutura de correlação é mais elíptica do que circular.

Introdução

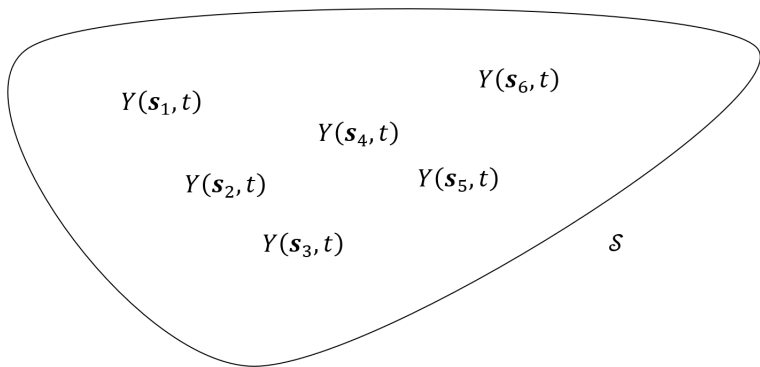
Morales *et al.* (2013) desenvolveram um modelo para observações **unidimensionais** (e.g., PM_{10}) que variam ao longo do tempo e do espaço, com $Y(\underline{s}_n, t) \in \mathcal{Y} \subset \mathbb{R}$.



No referido trabalho, eles modelaram a anisotropia espacial.

Introdução

As observações podem ser **multidimensionais** (e.g., PM₁₀ e PM_{2.5}), digamos $Y(\mathbf{s}_n, t) = (Y_1(\mathbf{s}_n, t), \dots, Y_q(\mathbf{s}_n, t)) \in \mathcal{Y} \subset \mathbb{R}^q$.



Paez *et al.* (2008) trabalharam com dados multivariados que variam no tempo e no espaço, assumindo isotropia.

Introdução

Principais ideias do trabalho

- Partindo do trabalho de Paez *et al.* (2008), desejamos uma extensão para lidar com dados anisotrópicos.
- Com as devidas ressalvas, o trabalho de Morales *et al.* (2013) faz a extensão que queremos para $q = 1$.
- Neste trabalho, queremos combinar as ideias de Paez *et al.* (2008) e Morales *et al.* (2013) para lidar com dados anisotrópicos multidimensionais ($q \geq 1$).

Introdução

Objetivo do trabalho

Para analisar observações matriz-variadas feitas no tempo e no espaço, estamos propondo um modelo espaço-temporal com as seguintes características:

- 1 **Estrutura de dependência temporal:** Captura via modelo dinâmico bayesiano matriz-normal;
- 2 **Estrutura de dependência espacial:** Incorporação da anisotropia via deformação espacial.

Deformação espacial em um modelo espaço-temporal bayesiano para respostas matriz-variadas

Deformação espacial

Deformação espacial

- Temos interesse no método de deformação espacial desenvolvido por Sampson e Guttorp (1992).
- **Ideia central:** mapear as coordenadas geográficas originais para um novo espaço latente onde a isotropia seja válida.
- O *processo de deformação* é um processo aleatório bidimensional $d(\cdot) = \{d(\underline{s}) : \underline{s} \in \mathcal{S}\}$ que mapeia as coordenadas da região \mathcal{S} para um espaço \mathcal{D} , em que:
 - $\underline{s} = (\text{lon}, \text{lat}) \in \mathcal{S} \subset \mathbb{R}^2$ é um ponto de coordenadas geográficas situado na região de interesse; e
 - $d(\underline{s}) = (d_1(\underline{s}), d_2(\underline{s})) \in \mathcal{D} \subset \mathbb{R}^2$ é um ponto bivariado em um espaço latente onde a isotropia é válida.
- Trabalhamos com uma versão bayesiana deste método que supõe que $d(\cdot)$ é um processo gaussiano, proposta por Schmidt e O'Hagan (2003).

Deformação espacial

Parâmetro \mathbf{D} e sua distribuição *a priori*

- Seja $\mathbf{d}_n = d(\mathbf{s}_n) = (d_1(\mathbf{s}_n), d_2(\mathbf{s}_n))$ para todo $n \in \{1, \dots, N\}$.
- Se $\mathbf{d}_1, \dots, \mathbf{d}_N$ são N pontos de interesse no espaço latente \mathcal{D} , escreva a seguinte matriz $2 \times N$ formada por eles:

$$\mathbf{D} = \begin{bmatrix} \mathbf{d}_1 & \cdots & \mathbf{d}_N \end{bmatrix}$$

- Como um processo gaussiano *a priori* foi especificado para $d(\cdot)$, pela definição de processos aleatórios multidimensionais temos que \mathbf{D} tem distribuição *a priori* matriz-normal com:
 - Matriz $2 \times N$ de médias $\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_N \end{bmatrix}$;
 - Matriz 2×2 de covariâncias à esquerda $\boldsymbol{\sigma}_d^2$; e
 - Matriz $N \times N$ de correlação à direita \mathbf{R}_d .
- Isto é denotado por $\mathbf{D} \sim \mathbf{N}_{2 \times N}(\mathbf{S}, \boldsymbol{\sigma}_d^2, \mathbf{R}_d)$.

Deformação espacial

Especificação da matriz de covariâncias σ_d^2

- A matriz diagonal σ_d^2 (2×2) calibra a estrutura de covariância *a priori* do sistema de coordenadas no espaço \mathcal{D} .
- Seus elementos diagonais controlam o nível de distorção no mapeamento do espaço \mathcal{S} ao espaço \mathcal{D} .
- Como usualmente σ_d^2 não é bem estimado, decidimos fixá-lo.
- Optamos pela especificação empírica $\sigma_d^2 = \tau \cdot \text{diag}\{\hat{\sigma}_{\text{lon}}^2, \hat{\sigma}_{\text{lat}}^2\}$ para evitar más especificações, com $0 < \tau \leq 1$.

Deformação espacial

Especificação da matriz de correlação \mathbf{R}_d

- Seja $\mathbf{R}_d = \begin{bmatrix} R_{n,n'} \end{bmatrix}_{N \times N}$ uma matriz de correlação com entradas dadas por uma função de correlação gaussiana:

$$R_{n,n'} = \rho_d(\mathbf{s}_n - \mathbf{s}_{n'}) = \begin{cases} \exp\{-\psi \|\mathbf{s}_n - \mathbf{s}_{n'}\|^2\}, & \text{se } n \neq n' \\ 1, & \text{se } n = n' \end{cases}$$

- \mathbf{R}_d dá a estrutura de correlação *a priori* dos locais medidos em \mathcal{D} e controla o grau de suavidade do processo gaussiano.
- O termo $\psi > 0$ é conhecido, a ser controlado.
- Pontos mais distantes de \mathcal{S} têm menos correlação *a priori* em \mathcal{D} e suas distâncias podem ser mais distorcidas em \mathcal{D} .

Notações

- Defina $Y_{n,i,t}$ como o valor observado da variável resposta i no local \underline{s}_n e no tempo t , em que $t \in \{1, \dots, T\}$, $n \in \{1, \dots, N\}$ e $i \in \{1, \dots, q\}$.
- Para cada $t \in \mathcal{T}$, defina a matriz de respostas ($N \times q$) como:

$$\begin{aligned}\mathbf{Y}_t &= \begin{bmatrix} Y_{1,1,t} & \cdots & Y_{1,i,t} & \cdots & Y_{1,q,t} \\ Y_{2,1,t} & \cdots & Y_{2,i,t} & \cdots & Y_{2,q,t} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Y_{N,1,t} & \cdots & Y_{N,i,t} & \cdots & Y_{N,q,t} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{\tilde{Y}}_{1,t} & \cdots & \mathbf{\tilde{Y}}_{i,t} & \cdots & \mathbf{\tilde{Y}}_{q,t} \end{bmatrix},\end{aligned}$$

onde $\mathbf{\tilde{Y}}_{i,t} = \begin{bmatrix} Y_{1,i,t} & \cdots & Y_{N,i,t} \end{bmatrix}^\top$ é a i -ésima coluna de \mathbf{Y}_t .

- Veremos um modelo **para cada coluna** de $\mathbf{Y}_{i,t}$ (i.e., $\mathbf{\tilde{Y}}_{i,t}$).
- Em seguida, apresentaremos um modelo para a matriz \mathbf{Y}_t .

Modelo proposto

Para modelar a i -ésima variável resposta, assumimos que elas têm a mesma estrutura de dependência espacial dada por uma matriz \mathbf{B} e suas próprias estruturas de dependência temporal:

$$\tilde{\mathbf{Y}}_{i,t} \mid \tilde{\boldsymbol{\beta}}_{i,t}, V, \phi, \mathbf{D} \sim \mathbf{N}_N(\mathbf{X}_t \tilde{\boldsymbol{\beta}}_{i,t}, V \cdot \mathbf{B}), \quad t \in \{1, \dots, T\}$$

$$\mathbf{B} = \left[B_{n,n'} \right]_{N \times N}$$

$$B_{n,n'} = \begin{cases} \exp\{-\phi \|d(\mathbf{s}_n) - d(\mathbf{s}_{n'})\|\}, & \text{se } n \neq n' \\ 1, & \text{se } n = n' \end{cases}$$

$$\tilde{\boldsymbol{\beta}}_{i,t} \mid \tilde{\boldsymbol{\beta}}_{i,t-1}, V \sim \mathbf{N}_p(\mathbf{G}_t \tilde{\boldsymbol{\beta}}_{i,t-1}, V \cdot \mathbf{W}), \quad t \in \{1, \dots, T\}$$

$$\tilde{\boldsymbol{\beta}}_{i,0} \mid V \sim \mathbf{N}_p(\mathbf{M}_{i,0}, V \cdot \mathbf{C}_0)$$

$$V \sim \text{IG}(a_V, b_V)$$

$$\phi \sim \text{G}(a_\phi, b_\phi)$$

$$\mathbf{D} \sim \mathbf{N}_{2 \times N}(\mathbf{S}, \boldsymbol{\sigma}_d^2, \mathbf{R}_d)$$

Modelo proposto

Modelo espaço-temporal matriz-variado

- Considere uma matriz Σ ($q \times q$) para acomodar a relação entre pares de variáveis respostas, assumindo-se que ela não varia no tempo.
- Similar a $\mathbf{Y}_t = \begin{bmatrix} \mathbf{Y}_{1,t} & \cdots & \mathbf{Y}_{i,t} & \cdots & \mathbf{Y}_{q,t} \end{bmatrix}$, defina as matrizes $p \times q$ a seguir:
 - $\mathbf{M}_0 = \begin{bmatrix} \mathbf{M}_{1,0} & \cdots & \mathbf{M}_{i,0} & \cdots & \mathbf{M}_{q,0} \end{bmatrix}$
 - $\boldsymbol{\beta}_0 = \begin{bmatrix} \boldsymbol{\beta}_{1,0} & \cdots & \boldsymbol{\beta}_{i,0} & \cdots & \boldsymbol{\beta}_{q,0} \end{bmatrix}$
 - $\boldsymbol{\beta}_t = \begin{bmatrix} \boldsymbol{\beta}_{1,t} & \cdots & \boldsymbol{\beta}_{i,t} & \cdots & \boldsymbol{\beta}_{q,t} \end{bmatrix}$, com $t \in \{1, \dots, T\}$.
- Com esses elementos, podemos escrever um modelo para as matrizes de resposta \mathbf{Y}_t .

Modelo proposto

Para $t \in \mathcal{T} = \{1, \dots, T\}$, propomos um modelo espaço-temporal matriz-variado dado por:

$$\begin{aligned}\mathbf{Y}_t \mid \boldsymbol{\beta}_t, V, \phi, \mathbf{D}, \boldsymbol{\Sigma} &\sim \mathbf{N}_{N \times q}(\mathbf{X}_t \boldsymbol{\beta}_t, V \cdot \mathbf{B}, \boldsymbol{\Sigma}) \\ B_{n,n'} &= \begin{cases} \exp\{-\phi \|d(\mathbf{s}_n) - d(\mathbf{s}_{n'})\|\}, & n \neq n' \\ 1, & n = n' \end{cases} \\ \boldsymbol{\beta}_t \mid \boldsymbol{\beta}_{t-1}, V, \boldsymbol{\Sigma} &\sim \mathbf{N}_{p \times q}(\mathbf{G}_t \boldsymbol{\beta}_{t-1}, V \cdot \mathbf{W}, \boldsymbol{\Sigma}) \\ \boldsymbol{\beta}_0 \mid V, \boldsymbol{\Sigma} &\sim \mathbf{N}_{p \times q}(\mathbf{M}_0, V \cdot \mathbf{C}_0, \boldsymbol{\Sigma}) \\ V &\sim \text{IG}(a_V, b_V) \\ \boldsymbol{\Sigma} &\sim \text{IW}_q(a_{\boldsymbol{\Sigma}}, \mathbf{b}_{\boldsymbol{\Sigma}}) \\ \phi &\sim \text{G}(a_{\phi}, b_{\phi}) \\ \mathbf{D} &\sim \mathbf{N}_{2 \times N}(\mathbf{S}, \boldsymbol{\sigma}_d^2, \mathbf{R}_d)\end{aligned}$$

Modelo proposto

Conjunto de parâmetros e sua distribuição *a priori*

- O conjunto de parâmetros é $\theta = \{\phi, \mathbf{D}, V, \Sigma, \beta_0, \beta\}$, em que $\beta = \{\beta_1, \dots, \beta_T\}$.
- Assumindo-se que ϕ , \mathbf{D} , V e Σ são independentes e usando a propriedade markoviana dos modelos dinâmicos, tem-se a distribuição *a priori* a seguir:

$$f(\theta) = f(\phi)f(\mathbf{D})f(V)f(\Sigma)f(\beta_0 | V, \Sigma) \prod_{t=1}^T f(\beta_t | \beta_{t-1}, V, \Sigma)$$

Modelo proposto

Função de verossimilhança e distribuição *a posteriori*

- Defina $\mathbf{Y} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_T\}$ o conjunto de respostas matriz-variadas, assumindo-se que $\mathbf{Y}_1, \dots, \mathbf{Y}_T$ são observações condicionalmente independentes dado $\boldsymbol{\theta}$.
- Então, a função de verossimilhança de $\boldsymbol{\theta}$ é:

$$l(\boldsymbol{\theta}; \mathbf{y}) = f(\mathbf{y} \mid \boldsymbol{\theta}) = \prod_{t=1}^T f(\mathbf{y}_t \mid \beta_t, V, \phi, \mathbf{D}, \boldsymbol{\Sigma})$$

- Pelo teorema de Bayes, a densidade *a posteriori* de $\boldsymbol{\theta}$ é:

$$\begin{aligned} f(\boldsymbol{\theta} \mid \mathbf{y}) &\propto f(\phi) f(\mathbf{D}) f(V) f(\boldsymbol{\Sigma}) f(\beta_0 \mid V, \boldsymbol{\Sigma}) \\ &\times \left[\prod_{t=1}^T f(\beta_t \mid \beta_{t-1}, V, \boldsymbol{\Sigma}) f(\mathbf{y}_t \mid \beta_t, V, \phi, \mathbf{D}, \boldsymbol{\Sigma}) \right] \end{aligned}$$

Modelo proposto

Estimação de parâmetros

- Uma vez que $f(\boldsymbol{\theta} \mid \mathbf{y})$ não tem forma fechada, recorreremos ao método MCMC para obter amostras dos parâmetros do modelo.
- Para realizar sua implementação, será necessário amostrar da distribuição condicional completa de cada uma das componentes.

Distribuição condicional completa de $\beta_0, \beta_1, \dots, \beta_T$

Usamos o algoritmo FFBS para amostrar da distribuição a *posteriori* dos parâmetros $\beta_0, \beta_1, \dots, \beta_T$.

Distribuição condicional completa de ϕ

- Como a densidade condicional completa de ϕ não tem forma conhecida, usaremos o algoritmo Metropolis-Hastings.
- Fixamos $a_\phi = 1$ e $b_\phi = 0.3/\zeta$ (Fonseca & Steel, 2011), em que ζ é a mediana das distâncias entre as localizações.

Distribuição condicional completa de V

- Elicitando-se $V \sim \text{IG}(a_V, b_V)$, temos a seguinte conjugação:

$$V \mid \mathbf{Y} = \mathbf{y}, \beta_0, \beta, \phi, \mathbf{D}, \Sigma \sim \text{IG}(a'_V, b'_V),$$

em que a'_V e b'_V têm forma conhecida.

- Fizemos escolhas não informativas para a_V e b_V .

Distribuição condicional completa de Σ

- Elicitando-se $\Sigma \sim \text{IW}_q(a_\Sigma, \mathbf{b}_\Sigma)$, temos a seguinte conjugação:

$$\Sigma \mid \mathbf{Y} = \mathbf{y}, \beta_0, \beta, V, \phi, \mathbf{D} \sim \text{IW}_q(a'_\Sigma, \mathbf{b}'_\Sigma),$$

em que a'_Σ e \mathbf{b}'_Σ têm forma conhecida.

- Fizemos escolhas não informativas para a_Σ e \mathbf{b}_Σ .

Caso particular (para fins comparativos)

- Suponha que $\Sigma = \text{diag}\{\Sigma_{1,1}, \dots, \Sigma_{q,q}\}$.
- Elicitando-se $\Sigma_{i,i} \sim \text{IG}(a_{\Sigma_{i,i}}, b_{\Sigma_{i,i}})$ com a suposição de que $\Sigma_{1,1}, \dots, \Sigma_{q,q}$ são independentes *a priori*, mostra-se que as distribuições condicionais completas de $\Sigma_{1,1}, \dots, \Sigma_{q,q}$ são independentes com:

$$\Sigma_{i,i} \mid \mathbf{Y} = \mathbf{y}, \beta_0, \beta, V, \phi, \mathbf{D} \sim \text{IG}(a'_{\Sigma_{i,i}}, b'_{\Sigma_{i,i}}).$$

Distribuição condicional completa de \mathbf{D}

- A densidade condicional completa de \mathbf{D} não tem forma conhecida.
- Propomos estimar cada entrada de \mathbf{D} usando um slice sampler univariado.
- Seguindo uma abordagem mais tradicional, pode-se utilizar o algoritmo de Metropolis-Hastings para fazer isso.
- Assim como é feito em outros trabalhos que usam deformação espacial sob o paradigma bayesiano, nós fixamos as localizações de dois locais em \mathcal{D} para evitar o problema da falta de identificabilidade.

Caso particular (para fins comparativos)

- Se quisermos ajustar um modelo isotrópico, basta fixarmos $\mathbf{D} = \mathbf{S}$ em todas as iterações.
- Trata-se da transformação identidade, com $d(\underline{s}) = \underline{s}$.

Inferência bayesiana via MCMC

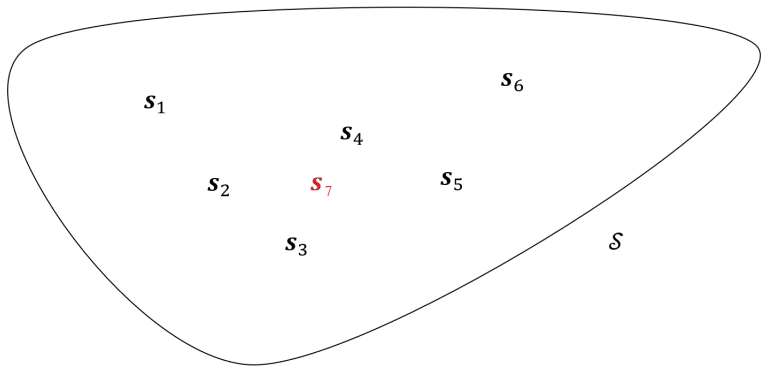
Algoritmo MCMC híbrido

- 1 Defina valores iniciais para todos os parâmetros (*i.e.*,
 $\theta^{(0)} = \{V^{(0)}, \Sigma^{(0)}, \beta_0^{(0)}, \beta_1^{(0)}, \dots, \beta_T^{(0)}, \phi^{(0)}, \mathbf{D}^{(0)}\} \in \Theta$)
- 2 Faça $j \leftarrow 1$.
- 3 Repita
 - 1 Gere $V^{(j)}$ e $\Sigma^{(j)}$ a partir de suas distribuições condicionais completas conhecidas.
 - 2 Gere $\beta_0^{(j)}, \beta_1^{(j)}, \dots, \beta_T^{(j)}$ usando o algoritmo FFBS.
 - 3 Gere $\phi^{(j)}$ usando o algoritmo Metropolis-Hastings.
 - 4 Gere $\mathbf{D}^{(j)}$ usando slice sampler (ou o algoritmo Metropolis-Hastings, alternativamente).
 - 5 Faça $j \leftarrow j + 1$.

até que a convergência seja alcançada.

Interpolação

Considere uma região geográfica \mathcal{S} que possui N estações de monitoramento.



O objetivo é interpolar a matriz de respostas nesse novo local não medido.

Interpolação

Interpolação

- Desejamos interpolar as respostas nos N^* locais $\mathbf{z}_{N+1}, \dots, \mathbf{z}_{N+N^*}$ em que não houve medição.
- Temos interesse em interpolar o conjunto de matrizes de respostas $\mathbf{Y}_{\text{int}} = \{\mathbf{Y}_1^*, \dots, \mathbf{Y}_T^*\}$, em que \mathbf{Y}_t^* é $N^* \times q$.
- Conseguimos aproximar $f(\mathbf{y}_{\text{int}} \mid \mathbf{y})$ usando integração de Monte Carlo:
 - ① Amostre de $f(\boldsymbol{\theta} \mid \mathbf{y})$ usando o algoritmo híbrido.
 - ② Amostre da distribuição conhecida $\mathbf{D}^* \mid \mathbf{D}$.
 - ③ Amostre da distribuição conhecida $\mathbf{Y}_t^* \mid \mathbf{Y}_t, \beta_t, V, \phi, \mathbf{D}, \mathbf{D}^*, \boldsymbol{\Sigma}$.

Métricas de avaliação

Métricas de avaliação

- DIC – Deviance Information Criterion.
- PMSE – Predictive Mean Squared Error

$$\text{PMSE} = \frac{1}{N^* q T} \sum_{n=1}^{N^*} \sum_{i=1}^q \sum_{t=1}^T (\hat{Y}_{N+n,i,t} - Y_{N+n,i,t})^2,$$

em que $\hat{Y}_{N+n,i,t}$ é a média preditiva *a posteriori*.

- IS – Interval Score, para quantificar as larguras dos intervalos preditivos (faixa do percentil 2.5 ao percentil 97.5) e penalizar quando os valores verdadeiros caem fora.

Modelagem bayesiana de matrizes de respostas incompletas

Mod. p/ matrizes de resp. incompletas

Modelagem

- É comum que o número de valores faltantes varie por coluna da matriz de respostas, como aqui:

$$\mathbf{Y}_t = \begin{bmatrix} \tilde{\mathbf{Y}}_{1,t} & \tilde{\mathbf{Y}}_{2,t} \end{bmatrix} = \begin{bmatrix} Y_{1,1,t} & \text{NA} \\ Y_{2,1,t} & \text{NA} \\ \text{NA} & Y_{3,2,t} \\ Y_{4,1,t} & Y_{4,2,t} \end{bmatrix} = \begin{bmatrix} Y_{1,1,t}^{\text{obs}} & Y_{1,2,t}^{\text{mis}} \\ Y_{2,1,t}^{\text{obs}} & Y_{2,2,t}^{\text{mis}} \\ Y_{3,1,t}^{\text{mis}} & Y_{3,2,t}^{\text{obs}} \\ Y_{4,1,t}^{\text{obs}} & Y_{4,2,t}^{\text{obs}} \end{bmatrix}$$

- Note que
 $\text{vec } \mathbf{Y}_t = (Y_{1,1,t}^{\text{obs}}, Y_{2,1,t}^{\text{obs}}, Y_{3,1,t}^{\text{mis}}, Y_{4,1,t}^{\text{obs}}, Y_{1,2,t}^{\text{mis}}, Y_{2,2,t}^{\text{mis}}, Y_{3,2,t}^{\text{obs}}, Y_{4,2,t}^{\text{obs}})$.
- Com a permutação $\mathbf{P}_t \text{vec } \mathbf{Y}_t$, podemos separar as partes observadas e não observadas:

$$(Y_{1,1,t}^{\text{obs}}, Y_{2,1,t}^{\text{obs}}, Y_{4,1,t}^{\text{obs}}, Y_{3,2,t}^{\text{obs}}, Y_{3,1,t}^{\text{mis}}, Y_{1,2,t}^{\text{mis}}, Y_{2,2,t}^{\text{mis}}).$$

Mod. p/ matrizes de resp. incompletas

Para $\underline{\mathbf{Y}}_t = \text{vec } \mathbf{Y}_t$, a versão vetorizada do modelo proposto é:

$$\begin{aligned}\underline{\mathbf{Y}}_t \mid \beta_t, V, \phi, \mathbf{D}, \Sigma &\sim \mathbf{N}_{Nq}([\mathbf{I}_q \otimes \mathbf{X}_t] \text{vec } \beta_t, V \cdot [\Sigma \otimes \mathbf{B}]) \\ \mathbf{B} &= \begin{bmatrix} B_{n,n'} \end{bmatrix}_{N \times N} \\ B_{n,n'} &= \begin{cases} \exp\{-\phi \|d(\underline{\mathbf{s}}_n) - d(\underline{\mathbf{s}}_{n'})\|\}, & \text{se } n \neq n' \\ 1, & \text{se } n = n' \end{cases} \\ \text{vec } \beta_t \mid \beta_{t-1}, V, \Sigma &\sim \mathbf{N}_{pq}([\mathbf{I}_q \otimes \mathbf{G}_t] \text{vec } \beta_{t-1}, V \cdot [\Sigma \otimes \mathbf{W}]) \\ \text{vec } \beta_0 \mid V, \Sigma &\sim \mathbf{N}_{pq}(\text{vec } \mathbf{M}_0, V \cdot [\Sigma \otimes \mathbf{C}_0]) \\ V &\sim \text{IG}(a_V, b_V) \\ \Sigma &\sim \text{IW}_q(a_\Sigma, \mathbf{b}_\Sigma) \\ \phi &\sim \text{G}(a_\phi, b_\phi) \\ \mathbf{D} &\sim \mathbf{N}_{2 \times N}(\mathbf{S}, \sigma_d^2, \mathbf{R}_d)\end{aligned}$$

Usando propriedades da distribuição normal multivariada, podemos amostrar da densidade $f(\underline{\mathbf{y}}_{\text{mis}} \mid \boldsymbol{\theta}, \underline{\mathbf{y}}_{\text{obs}})$.

Procedimento de inferência

Abordagem informal

Para amostrar de $[\text{Parameters}, \text{Missing data} \mid \text{Observed data}]$, use o método de dados aumentados:

- Amostre de $[\text{Missing data} \mid \text{Parameters}, \text{Observed data}]$;
- Amostre de $[\text{Parameters} \mid \text{Missing data}, \text{Observed data}]$ (ou seja, de $[\text{Parameters} \mid \text{Complete data}]$).

Procedimento de inferência

Para amostrar de $f(\boldsymbol{\theta}, \mathbf{y}_{\text{mis}} \mid \mathbf{y}_{\text{obs}})$, defina valores iniciais e então execute o método iterativo:

- Amostre $\mathbf{y}_{\text{mis}}^{(j)}$ da densidade $f(\cdot \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\theta}^{(j-1)})$.
- Amostre $\boldsymbol{\theta}^{(j)} = \{V^{(j)}, \boldsymbol{\Sigma}^{(j)}, \beta_0^{(j)}, \beta_1^{(j)}, \dots, \beta_T^{(j)}, \phi^{(j)}, \mathbf{D}^{(j)}\}$ da densidade *a posteriori* $f(\cdot \mid \mathbf{y}_{\text{obs}}, \mathbf{y}_{\text{mis}}^{(j)}) = f(\cdot \mid \mathbf{y}^{(j)})$.

Aplicação

Aplicação

- Usaremos médias diárias de PM_{10} e $PM_{2.5}$ ($q = 2$) da região do Central Valley/CA, EUA.
- Consideramos o ano de 2024, totalizando $T = 366$ dias.
- Usamos dados de $N = 18$ locais para ajustar o modelo, deixando dados de $N^* = 3$ locais para avaliar a capacidade de interpolação do modelo.
- A literatura sugere anisotropia para dados de $PM_{2.5}$ dessa região.
- Não incluímos variáveis explicativas ($p = 1$).
- Há 10,71% de valores faltantes para a Resposta 1 (PM_{10}) e 1,61% de valores faltantes para a Resposta 2 ($PM_{2.5}$).
- **Fonte:** U. S. Environmental Protection Agency (EPA).

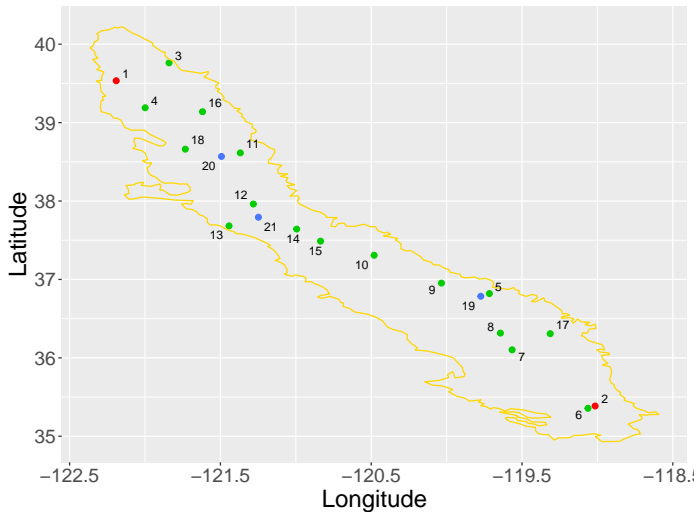


Figura: Região do Central Valley/CA, EUA. Em vermelho: pontos de ancoragem (s_1 e s_2); em verde: pontos de não ancoragem (s_3, \dots, s_{18}); e em azul: pontos usados na interpolação (s_{19} , s_{20} e s_{21}).

Comparação de modelos

Consideramos quatro modelos para fins comparativos:

- **Modelo 1** (\mathcal{M}_1): isotrópico e com respostas não correlacionadas, especificado por $d(\underline{s}) = \underline{s}$ e $\Sigma = \text{diag}\{\Sigma_{1,1}, \Sigma_{2,2}\}$.
- **Modelo 2** (\mathcal{M}_2): isotrópico e com respostas correlacionadas, especificado por $d(\underline{s}) = \underline{s}$ e $\Sigma \neq \text{diag}\{\Sigma_{1,1}, \Sigma_{2,2}\}$.
- **Modelo 3** (\mathcal{M}_3): anisotrópico e com respostas não correlacionadas, especificado por $d(\underline{s}) \neq \underline{s}$ e $\Sigma = \text{diag}\{\Sigma_{1,1}, \Sigma_{2,2}\}$.
- **Modelo 4** (\mathcal{M}_4): anisotrópico e com respostas correlacionadas, especificado por $d(\underline{s}) \neq \underline{s}$ e $\Sigma \neq \text{diag}\{\Sigma_{1,1}, \Sigma_{2,2}\}$. **Este é o modelo proposto.**

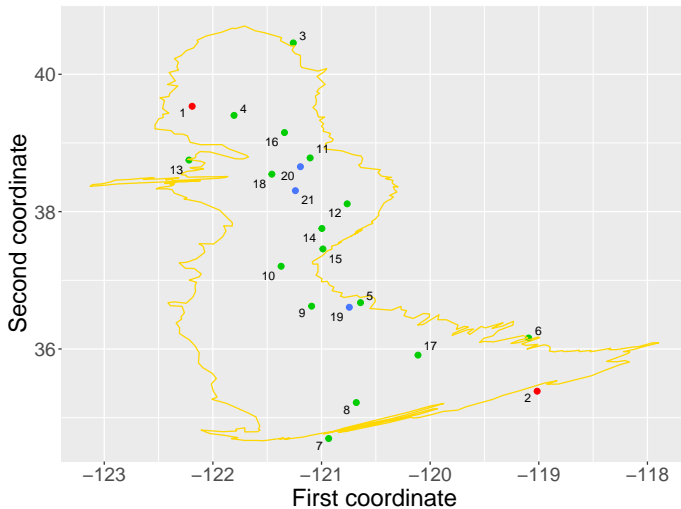


Figura: Mapa deformado sob o Modelo 3, anisotrópico e com respostas não correlacionadas.

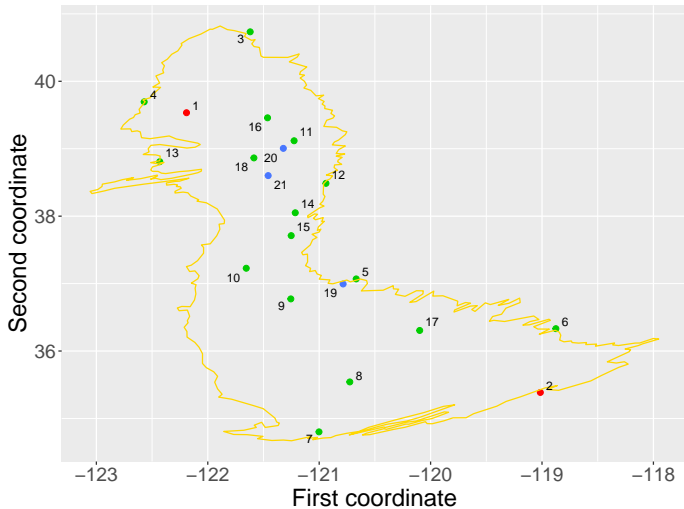


Figura: Mapa deformado sob o Modelo 4, anisotrópico e com respostas correlacionadas.

Tabela: DIC, PMSE e interval score por modelo para diferentes locais não medidos (\mathbf{s}_{19} , \mathbf{s}_{20} , e \mathbf{s}_{21}) e variáveis respostas ($i \in \{1, 2\}$).

Métrica	Local	i	\mathcal{M}_1	\mathcal{M}_2	\mathcal{M}_3	\mathcal{M}_4
DIC	-	-	81675.5	80347.0	76579.9	74882.6
PMSE	-	-	27.2789	26.0028	25.3166	22.9888
IS	\mathbf{s}_{19}	1	0.83725	0.81240	0.72242	0.72145
		2	0.31413	0.30915	0.25200	0.25269
	\mathbf{s}_{20}	1	1.09535	1.03954	0.81070	0.74511
		2	0.27022	0.26971	0.25684	0.25682
	\mathbf{s}_{21}	1	1.13926	1.12673	1.03927	1.02768
		2	0.28410	0.28160	0.25036	0.25158

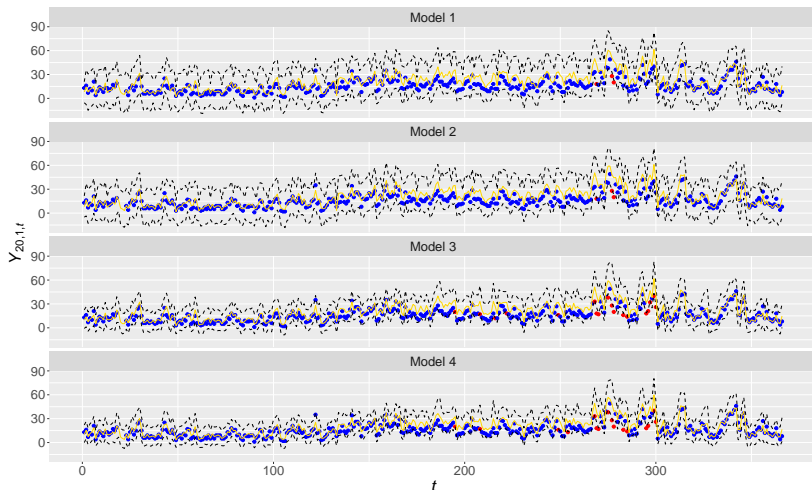


Figura: Série temporal de $Y_{n,i,t}$ para $n = 20$, $i = 1$ e $t \in \{1, \dots, 366\}$ por modelo. Os quantis 2,5 e 97,5 *a posteriori* são representados pela linha preta tracejada. Pontos representam os valores verdadeiros (em azul, se contidos no intervalo; em vermelho, c.c.).

Considerações finais e trabalhos futuros

Considerações finais

Considerações finais

- Propusemos um modelo espaço-temporal matriz-variado que incorpora anisotropia via deformação espacial.
- Ilustramos um caso em que há ganhos com o uso de deformação espacial para interpolações.
- Do ponto de vista computacional, os tempos de execução aqui obtidos podem ser considerado razoáveis.
- Ainda foi apresentada uma solução para o problema de matrizes de respostas incompletas com o método de dados aumentados.

Considerações finais

Trabalhos futuros

- Implementação de algoritmos mais eficientes (e.g., HMC e NUTS).
- O relaxamento da hipótese de normalidade é um possível trabalho futuro.
- Outra extensão consiste em permitir que cada coluna das matrizes de respostas tenha sua própria estrutura espacial (por exemplo, para estimar uma deformação própria por coluna), motivada pelos resultados de IS.
- O artigo está em fase de revisão final da escrita.

Meu site:

<https://rbulhoes.netlify.app/>

Apoio: CAPES e PRPPG/UFBA (JOVEMPESQ 2004)



PRPPG
Pró-Reitoria de Pesquisa
e Pós-Graduação

Obrigado!