



MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

Alberson Miranda, PPGeco/UFES

junho de 2023



SUMÁRIO

1 CONTEXTUALIZAÇÃO DA PESQUISA

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

▶ REFERÊNCIAS



ANÁLISE DE SÉRIES TEMPORAIS EM ECONOMIA BANCÁRIA

1 CONTEXTUALIZAÇÃO DA PESQUISA

- Problemas populares na literatura: risco de crédito, detecção de anomalias (fraude), preços de ativos, alocação ótima de numerários
- Pouco sobre projeção de saldos e rendas (*guidance, budgeting*)
- Previsões equivocadas = percepção de incompetência, redução de reputação e receitas
- *Budgeting* em empresas de muitas filiais: alta dimensionalidade e dúvidas



ANÁLISE DE SÉRIES TEMPORAIS EM ECONOMIA BANCÁRIA

1 CONTEXTUALIZAÇÃO DA PESQUISA

- Banestes: 96 agências, 70 municípios do ES¹, 13 microrregiões, 4 mesorregiões
- Qual a melhor forma de realizar projeções?
 - Individualmente por agência?
 - Apenas o total e distribuir proporcionalmente?
 - As informações por mesorregião ou microrregião são importantes?
 - Como usar toda a informação?

¹Postos de atendimento não contam como agência.



SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

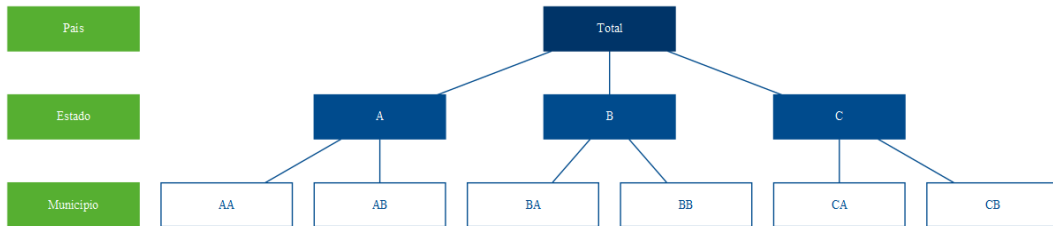
1 CONTEXTUALIZAÇÃO DA PESQUISA

Séries temporais hierárquicas são aquelas que podem ser agregadas ou desagregadas naturalmente em uma estrutura aninhada



SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

1 CONTEXTUALIZAÇÃO DA PESQUISA





SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

1 CONTEXTUALIZAÇÃO DA PESQUISA

- Cada nível traz informação diferente
- É razoável supor que exista covariância entre as previsões dos diferentes níveis
- Toda informação da estrutura pode ser útil



SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

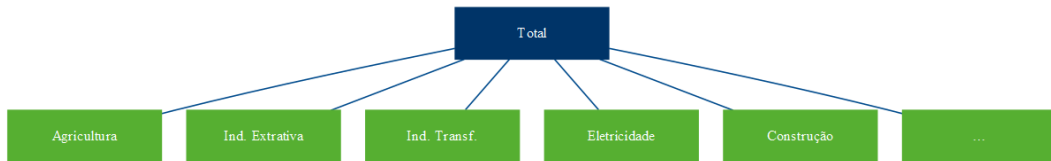
1 CONTEXTUALIZAÇÃO DA PESQUISA

Séries temporais agrupadas são aquelas que não impõem uma única estrutura hierárquica



SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

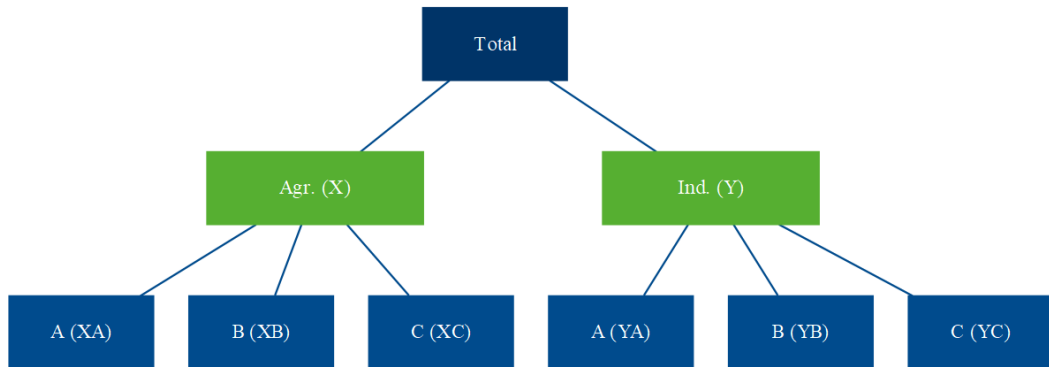
1 CONTEXTUALIZAÇÃO DA PESQUISA





SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

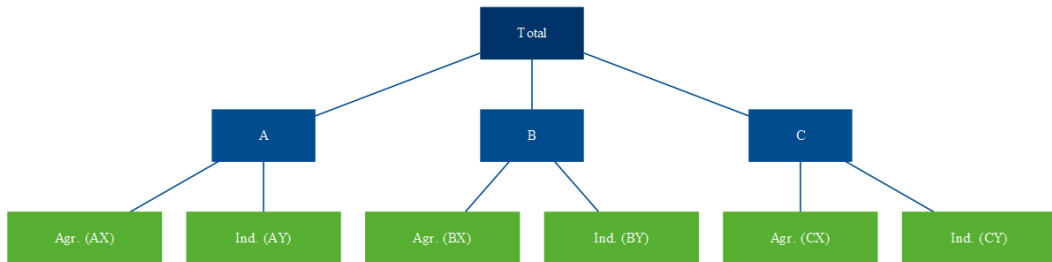
1 CONTEXTUALIZAÇÃO DA PESQUISA





SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

1 CONTEXTUALIZAÇÃO DA PESQUISA





COERÊNCIA E RECONCILIAÇÃO

1 CONTEXTUALIZAÇÃO DA PESQUISA

- Coerência: cada nó da hierarquia deve totalizar os nós filhos
- Não há razão para que as previsões individuais (e.g., Arima, ETS) sejam coerentes

RECONCILIAÇÃO ÓTIMA

Reescrever as previsões de forma que elas sejam coerentes, utilizando toda a informação disponível na estrutura hierárquica, ao mesmo tempo em que minimiza a variância do erro de previsão



SUMÁRIO

2 OBJETIVOS

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► ABORDAGENS DE NÍVEL ÚNICO

► MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

► REFERÊNCIAS



OBJETIVO GERAL

2 OBJETIVOS

Estudar o problema da reconciliação ótima de previsões pontuais a partir de métodos de *machine learning*



OBJETIVOS ESPECÍFICOS

2 OBJETIVOS

1. Aplicar o método *elastic net* para reconciliação ótima de previsões pontuais de séries temporais hierárquicas e agrupadas e atestar sua validade em dados de alta dimensionalidade;
2. Comparar a performance, em termos de funções de custo e tempo de processamento, do método *elastic net* nas configurações $\alpha = 0$ (*lasso*), $\alpha = 1$ (*ridge*) e uma combinação de ambos obtida via reamostragem;
3. Comparar a performance, em termos de funções de custo e tempo de processamento, do método *elastic net* com os métodos analíticos *bottom-up*, *top-down* e *MinT Shrink*;



OBJETIVOS ESPECÍFICOS

2 OBJETIVOS

4. Comparar a performance, em termos de funções de custo e tempo de processamento, das estratégias de reamostragem validação cruzada *k-fold* e *holdout* para o método *elastic net*;
5. Verificar se a aplicação de regressão regularizada resulta em algum padrão reconhecível na estrutura (e.g., se a regularização tende a favorecer os coeficientes de séries temporais que compartilham o mesmo nó pai ou, em outras palavras, se tende a zerar os coeficientes de séries temporais que não compartilham o mesmo nó pai na hierarquia);



SUMÁRIO

3 ABORDAGENS DE NÍVEL ÚNICO

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

▶ REFERÊNCIAS



BOTTOM-UP

3 ABORDAGENS DE NÍVEL ÚNICO

- informações apenas dos níveis mais desagregados
- soma-se para obter os níveis agregados



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- apenas informações do nível mais agregado



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- apenas informações do nível mais agregado
- há dezenas de métodos para distribuir proporcionalmente as previsões agregadas



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- apenas informações do nível mais agregado
- há dezenas de métodos para distribuir proporcionalmente as previsões agregadas
 - proporções históricas

$$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t}$$



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- apenas informações do nível mais agregado
- há dezenas de métodos para distribuir proporcionalmente as previsões agregadas
 - proporções históricas

$$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t}$$

- médias históricas

$$p_j = \frac{\sum_{t=1}^T \frac{y_{j,t}}{T}}{\sum_{t=1}^T \frac{y_t}{T}}$$



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- apenas informações do nível mais agregado
- há dezenas de métodos para distribuir proporcionalmente as previsões agregadas
 - proporções históricas

$$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t}$$

- médias históricas

$$p_j = \frac{\sum_{t=1}^T \frac{y_{j,t}}{T}}{\sum_{t=1}^T \frac{y_t}{T}}$$

- proporções das previsões (ATHANASOPOULOS; AHMED; HYNDMAN, 2009)

$$p_j = \prod_{\ell=0}^{K-1} \frac{\hat{y}_{j,h}^{(\ell)}}{\hat{S}_{j,h}^{(\ell+1)}}$$



MIDDLE-OUT

3 ABORDAGENS DE NÍVEL ÚNICO

- informações apenas de um nível intermediário
- soma-se para obter os níveis agregados e distribui-se proporcionalmente para os níveis mais desagregados



SUMÁRIO

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

▶ REFERÊNCIAS



NOTAÇÃO MATRICIAL

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

$$\tilde{y}_t = SG\hat{y}_t$$

Em que

- \tilde{y}_t é o vetor de previsões reconciliadas
- S é a matriz de soma
- G é a matriz de reconciliação
- \hat{y}_t é o vetor de previsões base



MATRIZ DE SOMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

S mapeia a estrutura hierárquica a partir da soma dos elementos mais desagregados

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{n \times 1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{n \times m} \begin{bmatrix} \hat{y}_{AA,t} \\ \hat{y}_{AB,t} \\ \hat{y}_{AC,t} \\ \hat{y}_{BA,t} \\ \hat{y}_{BB,t} \end{bmatrix}_{m \times 1}$$

exemplo 1: matriz de soma



MATRIZ DE RECONCILIAÇÃO

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

G mapeia o nível mais desagregado a partir das previsões de todos os níveis da hierarquia, garantindo a coerência

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{n \times 1} = S_{n \times m} \begin{bmatrix} p_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{m \times n} \begin{bmatrix} \hat{y}_{T+h|T} \\ \hat{y}_{A,T+h|T} \\ \hat{y}_{B,T+h|T} \\ \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \end{bmatrix}_{n \times 1}$$

exemplo 2: matriz de reconciliação top-down



MATRIZ DE RECONCILIAÇÃO

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{n \times 1} = \mathbf{S}_{n \times m} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{m \times n} \begin{bmatrix} \hat{y}_{T+h|T} \\ \hat{y}_{A,T+h|T} \\ \hat{y}_{B,T+h|T} \\ \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \end{bmatrix}_{n \times 1}$$

exemplo 3: matriz de reconciliação bottom-up



MATRIZ DE RECONCiliaÇÃO

4 MÉTODOS ANALÍTICOS DE RECONCiliaÇÃO ÓTIMA

E se quisermos utilizar toda a informação e não apenas parte do espaço das previsões base?

O PROBLEMA DE PESQUISA DA RECONCiliaÇÃO ÓTIMA

Estimar uma matriz de reconciliação G que utilize toda a informação disponível e com o menor erro de previsão



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

PROBLEMA DE REGRESSÃO

- Minimização do traço da matriz de covariância dos erros das previsões reconciliadas (MinT)

$$\min_{\tilde{e}} y_{T+h} - \tilde{y}_{T+h}$$

- Mínimos Quadrados Generalizados (MQG)

$$G = (S' W_h^\dagger S)^{-1} S' W_h^\dagger$$



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- 1ª abordagem: $W_h = k_h I$
 - Mínimos Quadrados Ordinários (MQO)
 - Não há qualquer razão ou justificativa para supor homoscedasticidade
 - Erros descorrelacionados e equivariantes não fazem sentido em séries hierárquicas



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- 2ª abordagem: $W_h = k_h \text{diag}(\hat{W}_1)$; $\hat{W}_1 = \frac{1}{T} \sum_{t=1}^T \hat{e}_{T+1} \hat{e}_{T+1}'$
 - Mínimos Quadrados Ponderados (MQP)
 - $\hat{e}_{T+1} = \mathbf{y}_{T+1} - \hat{\mathbf{y}}_{T+1}$
 - Não há qualquer razão ou justificativa para supor homoscedasticidade
 - Erros descorrelacionados e equivariantes não fazem sentido em séries hierárquicas



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- 3ª abordagem: $W_h = k_h(\lambda_D)$
 - Mínimos Quadrados Ponderados (MQP)
 - $\hat{e}_{T+1} = \mathbf{y}_{T+1} - \hat{\mathbf{y}}_{T+1}$
 - Não há qualquer razão ou justificativa para supor homoscedasticidade
 - Erros descorrelacionados e equivariantes não fazem sentido em séries hierárquicas



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

PROBLEMA DE REGRESSÃO



SUMÁRIO

5 REFERÊNCIAS

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► ABORDAGENS DE NÍVEL ÚNICO

► MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

► REFERÊNCIAS



REFERÊNCIAS

5 REFERÊNCIAS



ATHANASOPOULOS, G.; AHMED, R. A.; HYNDMAN, R. J. Hierarchical forecasts for Australian domestic tourism. en. **International Journal of Forecasting**, v. 25, n. 1, p. 146–166, jan. 2009. ISSN 0169-2070. DOI: 10.1016/j.ijforecast.2008.07.004. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169207008000691>>. Acesso em: 11 jan. 2023.



MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

Alberson Miranda, PPGeco/UFES

junho de 2023