



# MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

Alberson Miranda, PPGeco/UFES

fevereiro de 2023



# SUMÁRIO

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

► CONTEXTUALIZAÇÃO DA PESQUISA

► MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA



# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

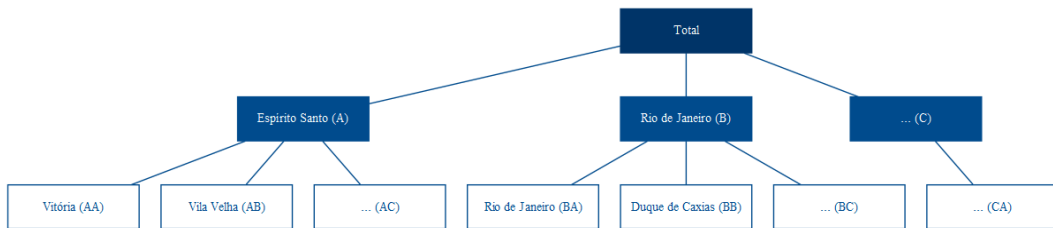
## 1 CONTEXTUALIZAÇÃO DA PESQUISA

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

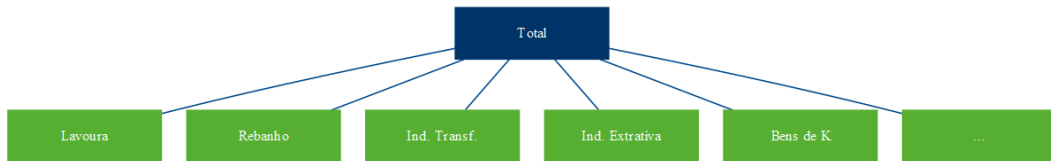
## 1 CONTEXTUALIZAÇÃO DA PESQUISA

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



# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

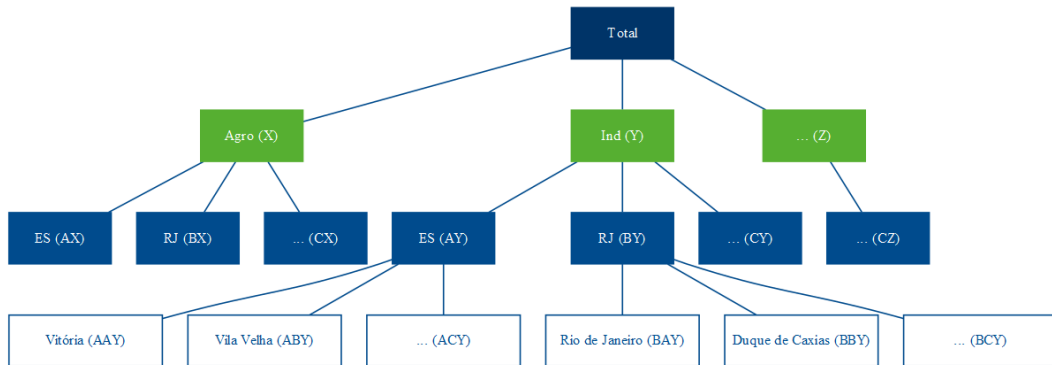
## 1 CONTEXTUALIZAÇÃO DA PESQUISA





# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

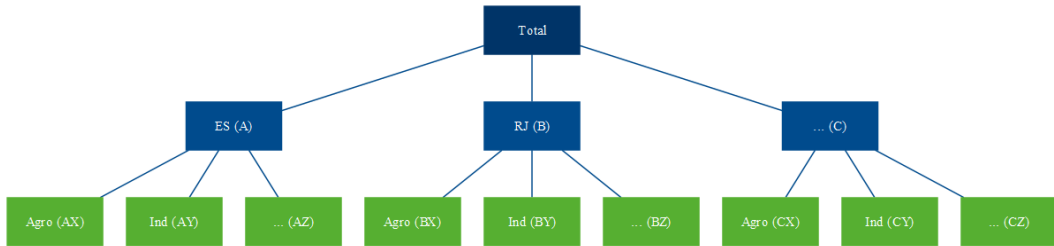






# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 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

Corrigir as previsões de forma que elas sejam coerentes, utilizando toda a informação disponível na estrutura hierárquica, ao mesmo tempo em que minimiza a variância total



# OBJETIVOS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

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

### OBJETIVOS ESPECÍFICOS

1. Estudar métodos para estimação da matriz de reconciliação aplicando algoritmos e fluxos de trabalho de machine learning, como *tuning* e *resampling*;
2. Identificar possíveis vantagens e limitações da abordagem por *machine learning* na reconciliação de previsões pontuais a partir de aplicação do método estudado na previsão de saldos de crédito do Banestes.



# SUMÁRIO

## 2 MÉTODOS ANALÍTICOS DE RECONCiliaÇÃO ÓTIMA

► CONTEXTUALIZAÇÃO DA PESQUISA

► MÉTODOS ANALÍTICOS DE RECONCiliaÇÃO ÓTIMA



# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados



# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados
- *Top-down*: apenas informações do nível mais agregado



# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados
- *Top-down*: apenas informações do nível mais agregado
  - proporções históricas

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



# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados
- *Top-down*: apenas informações do nível mais agregado
  - proporções históricas

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

- médias históricas

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





# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados
- *Top-down*: apenas informações do nível mais agregado
  - 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

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



# ABORDAGENS DE NÍVEL ÚNICO

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

- *Bottom-up*: informações apenas dos níveis mais desagregados
- *Top-down*: apenas informações do nível mais agregado
  - 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

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

- *middle-out*



# NOTAÇÃO MATRICIAL

## 2 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
- $S$  é a matriz de reconciliação
- $\hat{y}_t$  é o vetor de previsões base



# MATRIZ DE SOMA

## 2 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

## 2 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

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

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

*exemplo 3: matriz de reconciliação bottom-up*



# MATRIZ DE RECONCILIAÇÃO

## 2 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

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

### PROBLEMA DE REGRESSÃO

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

- 1ª abordagem: Mínimos Quadrados Ordinários (MQO)
  - Não há qualquer razão ou justificativa para supor

### RECONCILIAÇÃO ÓTIMA

- Mínimos Quadrados Generalizados (MQG)

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





# REFERÊNCIAS

## 2 MÉTODOS ANALÍTICOS DE RECONCILIAÇÃO ÓTIMA

HYNDMAN, R. J.; ATHANASOPOULOS, G. Forecasting: principles and practice. 3. ed. Melbourne, Austrália: OTexts, 2021.

SPILIOTIS, E. et al. Hierarchical forecast reconciliation with machine learning. Applied Soft Computing, v. 112, p. 107756, 1 nov. 2021.

WICKRAMASURIYA, S. L.; ATHANASOPOULOS, G.; HYNDMAN, R. J. Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through TraceMinimization. Journal of the American Statistical Association, v. 114, n. 526, p. 804–819, 3 abr. 2019.