

ALBERSON DA SILVA MIRANDA

**RECONCILIAÇÃO ÓTIMA PROBABILÍSTICA EM SÉRIES  
TEMPORAIS HIERÁRQUICAS**

Projeto de dissertação apresentado ao Programa  
de Pós-Graduação em Economia da Universidade  
Federal do Espírito Santo.

Orientador: Prof. Dr. Guilherme A. A. Pereira

Vitória

2022

MIRANDA, Alberson da Silva. **Reconciliação Ótima Probabilística em Séries Temporais Hierárquicas**. 2022. 6 folhas. Projeto de dissertação (Programa de Pós-Graduação em Economia) – Universidade Federal do Espírito Santo, Vitória, 2022.

## RESUMO

No máximo 500 palavras em espaço simples e sem parágrafos. Deve apresentar de forma concisa os objetivos, metodologia e os resultados alcançados, utilizar o verbo na voz ativa. Espaçamento simples, sem recuo de parágrafos.

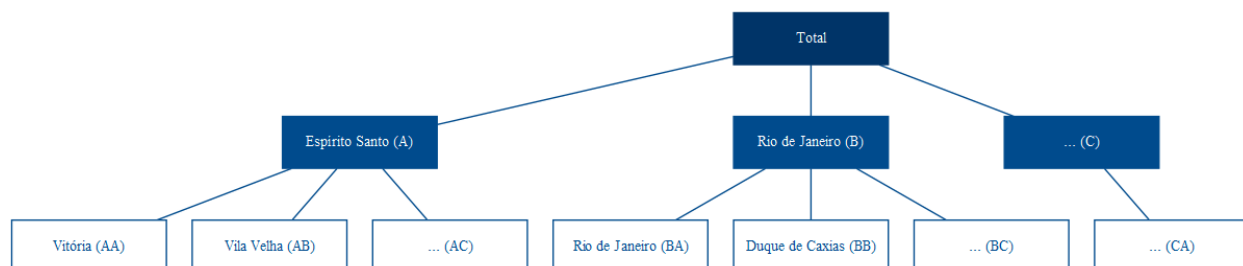
**Palavras-chave:** Palavra 1. Palavra 2. Palavra 3. Palavra 4. Palavra 5.

## 1 INTRODUÇÃO

Em minha pesquisa no mestrado, pretendo trabalhar no problema da reconciliação ótima de séries temporais hierárquicas e agrupadas no contexto probabilístico.

### 1.1 CONTEXTUALIZAÇÃO DA PESQUISA

Séries temporais hierárquicas são aquelas que podem ser agregadas ou desagregadas naturalmente em uma estrutura aninhada (HYNDMAN; ATHANASOPOULOS, 2021). Para ilustrar, tome a série do Pib brasileiro. Ela pode ser desagregada por estado que, por sua vez, pode ser desagregada por município.



**Figura 1** – Séries Hierárquicas

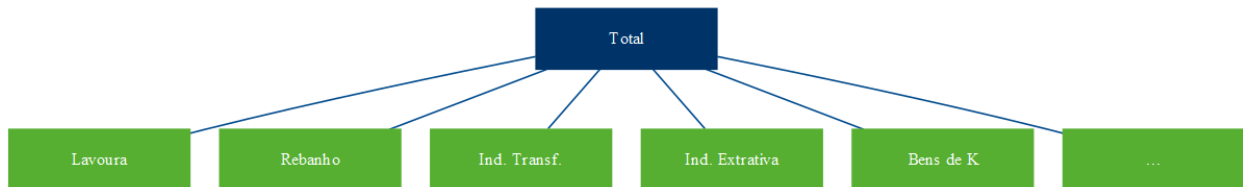
Essa estrutura pode ser representada por equações para qualquer nível de agregação. Assim, o agregado nacional pode ser representado apenas pelos agregados dos estados, através de (1), ou como o agregado dos municípios (2). Já o agregado para o estado do Espírito Santo é representado por (3).

$$y_t = y_{A,t} + y_{B,t} + y_{C,t} \quad (1)$$

$$y_t = y_{AA,t} + y_{AB,t} + y_{AC,t} + y_{BA,t} + y_{BC,t} + y_{CA,t} \quad (2)$$

$$y_{A,t} = y_{AA,t} + y_{AB,t} + y_{AC,t} \quad (3)$$

Por outro lado, o Pib pode ser também desagregado de forma cruzada de acordo com a atividade econômica — lavoura, rebanho, indústria de transformação, extrativa, bens de capital, bens intermediários, comércio de vestuário, automotivos, serviços etc. Essa estrutura não pode ser desagregada naturalmente de uma única forma, como é a hierarquia de estados e municípios. Não pode ser aninhada por um atributo como a própria geografia. A esse tipo de estrutura dá-se o nome de séries agrupadas.



**Figura 2 – Séries Agrupadas**

Combinando as duas, temos a estrutura de séries hierárquicas agrupadas. Ao contrário da estrutura hierárquica, que só pode ser agregada de uma forma — como com os municípios abaixo dos estados —, a adição da estrutura agrupada pode ocorrer tanto acima quanto abaixo da hierárquica.

Na notação matricial, essas estruturas são representadas pelo vetor  $y_t$   $n$ -dimensional com todas as observações no tempo  $t$  para todos os níveis da hierarquia, pela matriz de soma  $S$  de dimensão  $n \times m$  que define as equações para todo nível de agregação e pela matriz  $b_t$  composta pelas séries no nível mais desagregado.

$$\mathbf{y}_t = \mathbf{S}\mathbf{b}_t \quad (4)$$

Talvez as formas mais intuitivas de se pensar em previsões para esses tipos de estrutura sejam as abordagens *top-down* e *bottom-up*. Na abordagem *top-down*, a previsão para os níveis mais desagregados da hierarquia são determinadas por uma proporção  $p_i$  do nível agregado. Para isso, temos de definir uma matriz com todos esses pesos, que, seguindo a formulação de [Hyndman e Athanasopoulos \(2021\)](#), chamo de  $\mathbf{G}$ . Já a abordagem *bottom-up* parte do raciocínio inverso e define as previsões de cada elemento da estrutura a partir das previsões dos elementos mais desagregados. Para tanto, basta modificar a matriz  $\mathbf{G}$ . Então,  $\mathbf{G}$  define a abordagem — se *top-down* ou *bottom-up* —, e  $\mathbf{S}$  define a maneira da qual as previsões são somadas para formar as equações de previsão para cada elemento da estrutura.

Seja somando as previsões do nível mais desagregado para formar os níveis superiores da hierarquia (*bottom-up*) ou distribuindo proporcionalmente as previsões do nível mais agregado (*top-down*), o vetor  $\tilde{\mathbf{y}}_t$  representa as previsões *coerentes*. Isso significa que as previsões “batem”, ou seja, são totalizadas corretamente — as previsões de cada elemento agregado corresponde ao somatório das previsões dos níveis inferiores da hierarquia. Isso é garantido pela multiplicação das matrizes  $\mathbf{S}\mathbf{G}$ .

Não fosse essa pré multiplicação, nada garantiria a coerência das previsões. Tomando a estrutura da figura 1 como exemplo, seria um acaso improvável que as previsões do agregado para o estado do Espírito Santo sejam exatamente a soma das previsões individuais de seus municípios. Isso porque cada série pode seguir um processo diferente (e.g., arima) com erros e variâncias distintas.

Os métodos de gerar previsões coerentes a partir de previsões base são chamados de métodos de *reconciliação*. Os métodos de reconciliação tradicionais apresentados, *top-down* e *bottom-up*, utilizam informação limitada. No método *top-down*, utiliza-se apenas informações do nível mais agregado, enquanto na abordagem *bottom-up* utiliza-se apenas as informações dos níveis mais desagregados.

Alternativamente, podemos pensar numa matriz  $\mathbf{G}$  qualquer que utilize toda a informação

disponível e tenha algumas propriedades que garantam que as previsões coerentes tenham a menor diferença o possível em relação às previsões base. Esse é o problema de pesquisa trabalhado na *reconciliação ótima*.

## 2 OBJETIVOS

Meu objetivo geral para a dissertação é estudar o problema da reconciliação ótima de previsões pontuais a partir de regressões quantílicas.

Como objetivos específicos, tenho:

1. Estudar métodos para estimação da matriz  $G$  para qualquer quantil desejado, e não apenas a média;
2. Identificar possíveis vantagens e limitações da abordagem quantílica na reconciliação de previsões pontuais a partir de aplicação do método estudado na previsão de saldos de crédito do Banco do Estado do Espírito Santo SA.

Não está no escopo do estudo os métodos de se obter previsões probabilísticas, ou seja, a estimação da matriz  $\hat{y}_{T+h}$  para além de previsões pontuais.

## 3 REVISÃO DA LITERATURA

Previsões pontuais de séries temporais hierárquicas não é um assunto novo. Desde a década de 70, ao menos, pesquisas foram publicadas acerca de abordagens *bottom-up* e *top-down*, suas vantagens e desvantagens, e tentativas de se definir qual é o melhor método<sup>1</sup>. Entretanto, é apenas em Hyndman, Ahmed et al. (2011) que é formalizada uma abordagem prática que utiliza toda a informação disponível, (i.e. as previsões de todos elementos de todos os níveis da hierarquia) a partir da estimação da matriz  $G$  via regressão linear por MQG.

---

<sup>1</sup>Uma revisão dessa literatura pode ser encontrada em Athanasopoulos, Ahmed e Hyndman (2009).

Entretanto, para ser capaz de estimar o modelo por MQG, é necessária a matriz de variância-covariância dos erros. Hyndman, Ahmed et al. (2011) usam a matriz de erros de coerência, ou seja, a diferença entre as previsões reconciliadas e as previsões base, que tem posto incompleto e não identificada e, portanto, não pode ser estimada. Os autores contornam esse problema adotando no lugar da matriz de variância-covariância dos erros uma matriz diagonal constante, ou seja, assumem variância constante dos erros de reconciliação, e estimam a matriz  $G$  por MQO.

A estimação por esse método resulta numa reconciliação ótima que depende apenas da matriz  $G$ , ou seja, da estrutura hierárquica, e independe da variância e covariância das previsões base  $\hat{y}_{T+h}$  — o que não é uma conclusão satisfatória.

Hyndman, Lee e Wang (2016) tentam aperfeiçoar o método usando as variâncias das previsões base estimadas (dentro da amostra) como estimativa para a matriz de variância-covariância dos erros de reconciliação, de forma a as utilizar como pesos e realizar a reconciliação ótima por MQP. Assim, previsões base mais acuradas têm peso maior do que as mais ruidosas. Entretanto, não fornecem justificativa teórica para usar a diagonal da matriz de variância-covariância de  $\hat{e}_t$ .

Wickramasuriya, Athanasopoulos e Hyndman (2019) argumentam que o que de fato interessa é que as previsões reconciliadas tenham o menor erro. Então, corrigem a abordagem de reconciliação ótima para o objetivo de minimização dos erros das previsões reconciliadas  $\tilde{y}_{t+h}$ , ao invés dos erros das previsões base  $\hat{y}_{t+h}$ . Dado que isso implica na minimização da variância de  $\tilde{e}_{t+h}$ , ou seja, na minimização do somatório da diagonal, o traço, da matriz de variância-covariância de  $\tilde{e}_{t+h}$ , eles chamaram esse método de MinT (*Minimum Trace*). Paralelamente, usam desigualdade triangular para demonstrar que as previsões reconciliadas obtidas por esse método são ao menos tão boas quanto as previsões base.

Panagiotelis et al. (2021) reinterpreta a literatura de coerência e reconciliação de previsões pontuais a partir de uma abordagem geométrica, trazendo provas alternativas para conclusões anteriores ao mesmo tempo em que fornece novos teoremas. Além disso, Panagiotelis et al. (2021) então estende essa interpretação geométrica para o contexto probabilístico, fornecendo métodos paramétricos e não paramétricos (via *bootstrapping*) para reconciliação de previsões probabilísticas, ou seja, para reconciliar previsões  $\hat{y}_t$  obtidas a partir de toda a distribuição, e não apenas a média.

#### 4 METODOLOGIA

Para atingir os objetivos, será realizada pesquisa bibliográfica

#### 5 CRONOGRAMA

#### REFERÊNCIAS

- ATHANASOPOULOS, George; AHMED, Roman A.; HYNDMAN, Rob J. Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting*, v. 25, n. 1, p. 146–166, 2009. ISSN 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2008.07.004>.
- HYNDMAN, R.J.; AHMED, R.A. et al. Optimal Combination Forecasts for Hierarchical Time Series. *Computational Statistics and Data Analysis*, v. 55, p. 2579–2589, 2011.
- HYNDMAN, R.J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 3. ed. Melbourne, Australia, 2021. Acessado em 14/09/20. Disponível em: <<https://otexts.com/fpp3/>>.
- HYNDMAN, Rob J.; LEE, Alan J.; WANG, Earo. Fast computation of reconciled forecasts for hierarchical and grouped time series. *Computational Statistics and Data Analysis*, v. 97, p. 16–32, 2016. ISSN 0167-9473. DOI: <https://doi.org/10.1016/j.csda.2015.11.007>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S016794731500290X>>.
- PANAGIOTELIS, Anastasios et al. Forecast reconciliation: A geometric view with new insights on bias correction. *International Journal of Forecasting*, v. 37, n. 1, p. 343–359, 2021. ISSN 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2020.06.004>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169207020300911>>.
- WICKRAMASURIYA, Shanika L.; ATHANASOPOULOS, George; HYNDMAN, Rob J. Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. *Journal of the American Statistical Association*, Taylor & Francis, v. 114, n. 526, p. 804–819, 2019. DOI: [10.1080/01621459.2018.1448825](https://doi.org/10.1080/01621459.2018.1448825). eprint: <https://doi.org/10.1080/01621459.2018.1448825>. Disponível em: <<https://doi.org/10.1080/01621459.2018.1448825>>.