

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

ALBERSON DA SILVA MIRANDA

**RECONCILIAÇÃO ÓTIMA PROBABILÍSTICA EM SÉRIES  
TEMPORAIS HIERÁRQUICAS**

Vitória

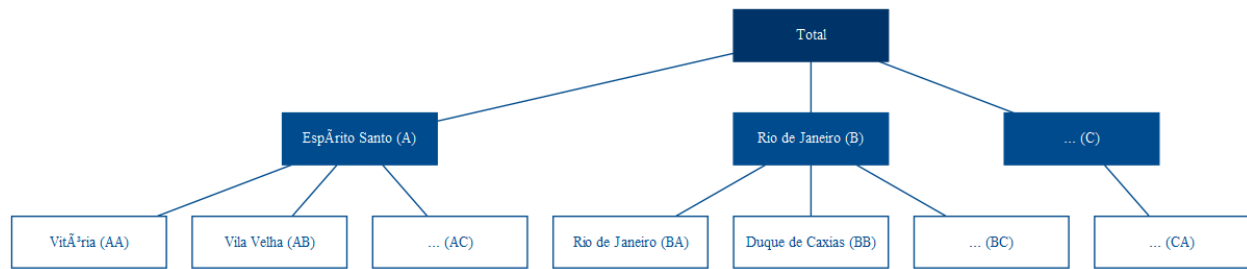
2022

## INTRODUÇÃO

Neste trabalho,

### SÉRIES HIERÁRQUICAS E SÉRIES AGRUPADAS

Séries temporais hierárquicas são aquelas que podem ser agregadas ou desagregadas naturalmente em uma estrutura aninhada (Hyndman and 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).

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

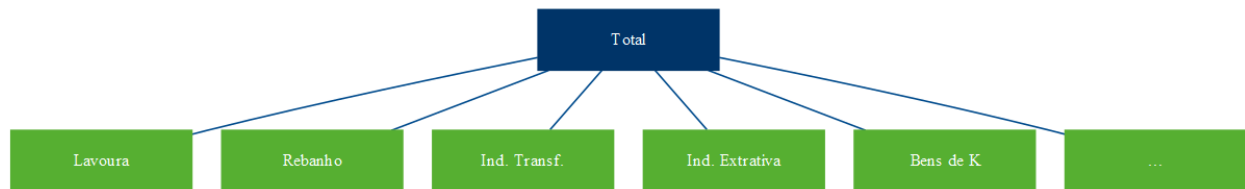
$$y_t = y_{AA,t} + y_{AB,t} + y_{AC,t} + y_{BA,t} + y_{BC,t} + y_{CA,t} \quad (2)$$

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

Alternativamente, podemos descrever a estrutura completa de forma matricial:

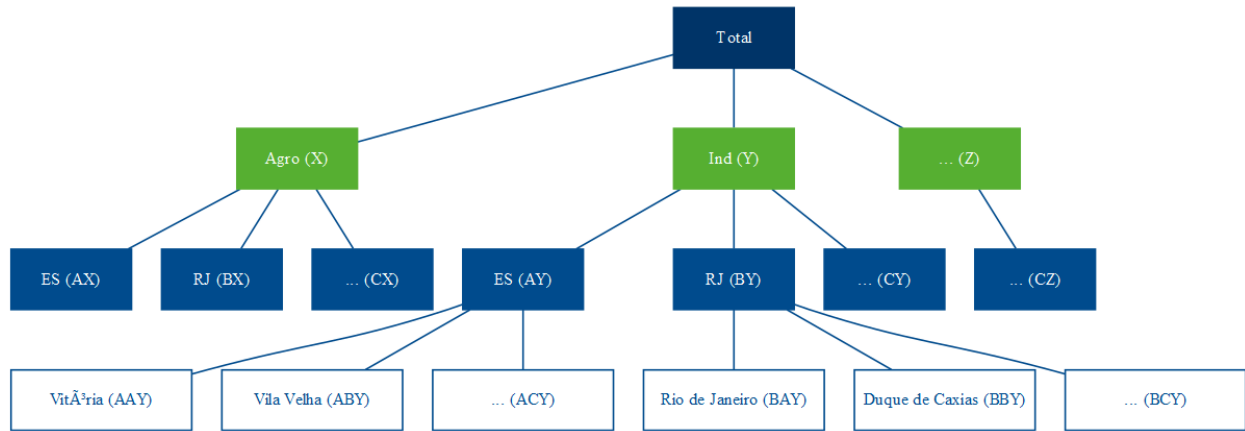
$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \\ y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \\ y_{BC,t} \\ y_{CA,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \\ y_{BC,t} \\ y_{CA,t} \end{bmatrix} \quad (4)$$

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.

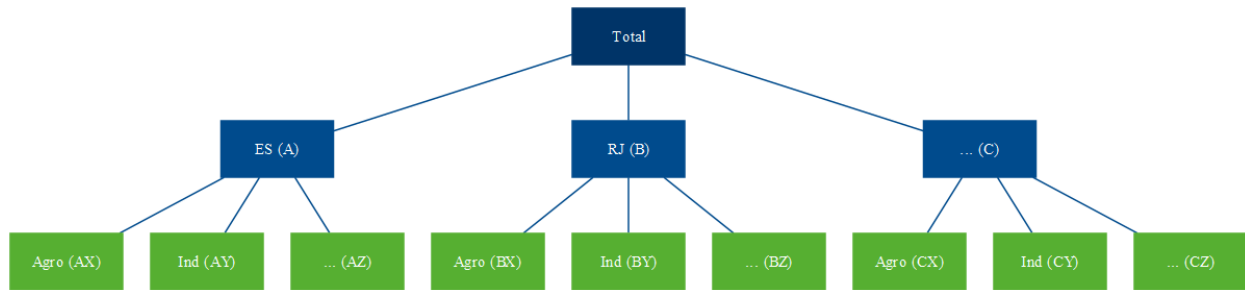


**Figura 2 – Séries Agrupadas**

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 (figura 3) quanto abaixo (figura 4) da hierárquica.



**Figura 3 – S ries Hier rquicas Agrupadas (a)**



**Figura 4 – S ries Hier rquicas Agrupadas (b)**

Na not    matricial, a estrutura da figura 4   representada como abaixo. Formalmente, o primeiro membro da igualdade   composto pelo vetor  $y_t$   $n$ -dimensional com todas as observa  es no tempo  $t$  para todos os n veis da hierarquia. O segundo membro   composto pela matriz de soma  $S$  de dimens    $n \times m$  que define as equa  es para todo n vel de agrega   , e pela matriz  $b_t$  composta pelas s ries no n vel mais desagregado.

$$y_t = Sb_t$$

$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \\ y_{X,t} \\ y_{Y,t} \\ y_{Z,t} \\ y_{AX,t} \\ y_{AY,t} \\ y_{AZ,t} \\ y_{BX,t} \\ y_{BY,t} \\ y_{BZ,t} \\ y_{CX,t} \\ y_{CY,t} \\ y_{CZ,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AX,t} \\ y_{AY,t} \\ y_{AZ,t} \\ y_{BX,t} \\ y_{BY,t} \\ y_{BZ,t} \\ y_{CX,t} \\ y_{CY,t} \\ y_{CZ,t} \end{bmatrix} \quad (5)$$

## ABORDAGENS TOP-DOWN E BOTTOM-UP

Talvez as formas mais intuitivas de se pensar em previsões para esses tipos de estrutura sejam as abordagens top-down e bottom-up. Tome a estrutura descrita na figura 1, por exemplo. Podemos realizar a previsão para o horizonte de tempo  $h$  do agregado do Pib brasileiro, representado no topo da hierarquia por *Total* (6), e então distribuir os valores previstos proporcionalmente entre os estados e municípios.

$$\hat{y}_{T+h|T} = E[y_{T+h}|\Omega_T] \quad (6)$$

Essa é a abordagem top-down. Nela, a previsão para os níveis mais desagregados da hierarquia são determinadas por uma proporção  $p_i$  do nível agregado. Por exemplo, as previsões para Vitória são

dadas pela equação 7.

$$\tilde{y}_{AA,T+h|T} = p_1 \hat{y}_{T+h|T} \quad (7)$$

Para isso, temos de definir uma matriz com todos esses pesos, que, seguindo a formulação de [Hyndman and Athanasopoulos \(2021\)](#), vamos chamar de  $G$ :

$$G = \begin{bmatrix} p_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (8)$$

$G$  é uma matriz  $m \times n$  que multiplica a matriz  $\hat{y}_{T+h|T}$  que, por sua vez, é composta pelas previsões base — as previsões individuais para todos os níveis de agregação. A equação para a abordagem *top-down* será, então:

$$\tilde{y}_{T+h|T} = SG \hat{y}_{T+h|T} \quad (9)$$

Na notação matricial para a estrutura da figura 1, temos:

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{C,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \\ \tilde{y}_{BC,t} \\ \tilde{y}_{CA,t} \end{bmatrix} = \mathbf{S} \begin{bmatrix} p_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{y}_{T+h|T} \\ \hat{y}_{A,T+h|T} \\ \hat{y}_{B,T+h|T} \\ \hat{y}_{C,T+h|T} \\ \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \\ \hat{y}_{BC,T+h|T} \\ \hat{y}_{CA,T+h|T} \end{bmatrix} \quad (10)$$

O que nos dá uma proporção do total para cada elemento no nível mais desagregado.

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{C,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \\ \tilde{y}_{BC,t} \\ \tilde{y}_{CA,t} \end{bmatrix} = \mathbf{S} \begin{bmatrix} p_1 \hat{y}_{T+h|T} \\ p_2 \hat{y}_{T+h|T} \\ p_3 \hat{y}_{T+h|T} \\ p_4 \hat{y}_{T+h|T} \\ p_5 \hat{y}_{T+h|T} \\ p_6 \hat{y}_{T+h|T} \\ p_7 \hat{y}_{T+h|T} \end{bmatrix} \quad (11)$$

Substituindo a matriz  $\mathbf{S}$ , temos as equações que definem cada previsão da estrutura em função de proporções da previsão do agregado.

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{C,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \\ \tilde{y}_{BC,t} \\ \tilde{y}_{CA,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 \hat{y}_{T+h|T} \\ p_2 \hat{y}_{T+h|T} \\ p_3 \hat{y}_{T+h|T} \\ p_4 \hat{y}_{T+h|T} \\ p_5 \hat{y}_{T+h|T} \\ p_6 \hat{y}_{T+h|T} \\ p_7 \hat{y}_{T+h|T} \end{bmatrix} \quad (12)$$

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  $\mathbf{G}$ .

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

O que resulta nas equações desejadas. Portanto,  $\mathbf{G}$  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.



$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{C,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \\ \tilde{y}_{BC,t} \\ \tilde{y}_{CA,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \\ \hat{y}_{BC,T+h|T} \\ \hat{y}_{CA,T+h|T} \end{bmatrix} \quad (14)$$

## COERÊNCIA E RECONCILIAÇÃO

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  $\tilde{y}_t$  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  $SG$ .

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* e, como vimos, são definidos na matriz  $G$ . 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 — por isso, apenas a primeira coluna em (8) é

diferente de zero. Já na abordagem *bottom-up*, utiliza-se apenas as informações dos níveis mais desagregados, o que resulta na submatriz identidade  $m \times m$  em (13), enquanto as colunas que representam os níveis mais agregados são nulas.

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*.

## 1 REVISÃO DA LITERATURA

## REFERÊNCIAS

Hyndman, R. and Athanasopoulos, G. (2021). *Forecasting: principles and practice*. Melbourne, Australia, 3 edition. Acessado em 14/09/20.