

Decomposição de Séries Temporais

Profa. R. Ballini

Bibliografia Básica:

- Pesaran, M. H. (2015). *Time Series and Panel Data Econometrics*. Cap. 12 – 12.5.
- Morettin, P. A. (2004). *Análise de Séries Temporais*. Cap. 3.

Introdução a Série Temporal

Definição

Uma série temporal é uma sequencia de observações ao longo do tempo.

Principais objetivos em se estudar séries temporais:

- Descrição. Descrever propriedades da série, ou seja, a tendência de longo prazo, componente sazonal, componente cíclico, observações discrepantes (*outliers*), alterações estruturais (mudanças no padrão da tendência), etc.
- Predição: prever valores futuros com base em valores passados. Aqui assume-se que o futuro envolve incerteza, ou seja as previsões não são perfeitas.
- Explicação. Usar a variação em uma série para explicar a variação em outra série.

Introdução a Série Temporal

Abordagens para tratar séries temporais:

1. Técnicas Descritivas: gráficos, identificação de padrões, etc.
2. Métodos não paramétricos: decomposição, alisamento ou suavização
3. Modelos Probabilísticos: Seleção, comparação e adequação de modelos, estimação, predição. Ferramenta básica é a função de autocorrelação e critérios de informação.
4. Outras Abordagens: modelos de espaço de estados, modelos não lineares, séries multivariadas, processos de longa dependência, modelos para volatilidade, etc.

Decomposição Clássica

Considere uma série temporal $\{Z_t, t = 1, \dots, N\}$ que pode ser representada como a soma dos componentes:

- Tendência (T_t): componente de longo prazo associado ao movimento da variável no tempo;
- Ciclo (C_t): componente de médio prazo associado a períodos de expansão ou recessão econômica;
- Sazonalidade (S_t): componente de curto prazo associado a variações provocadas por épocas do ano (feriados, estação do ano, etc);
- Resíduo (ϵ_t): componente que não se pode explicar, componente aleatório, também denominada de ruído.

Exemplos de séries temporais

Faça os gráficos das seguintes séries:

1. Dados mensais de total de passageiros em linhas aéreas internacionais nos EUA entre 1949 a 1960 (Box, Jenkins, Reinsel & Ljung, 2016)
2. Número anual de lincas capturados em armadilhas entre 1821 a 1934 no Canadá (Brockwell & Davis, 1991)
3. Medições anuais de vazões do Rio Nilo em Ashwan entre 1871 e 1970 (Cobb, G. W., 1978)
4. Consumo de Gás no Reino Unido entre o primeiro trimestre de 1960 e o quarto trimestre de 1986 (Durbin, J. & Koopman, S. J., 2001)
5. PIB anual da Bélgica (arquivo *PIB_Anuual_Belgica.xlsx*), de 1953 a 2010

Decomposição Clássica

Tipos de modelos:

1. Aditivo: variâncias estabilizadas no tempo (homocedástico)

$$Z_t = T_t + S_t + C_t + \epsilon_t \quad (1)$$

2. Multiplicativo: variâncias crescentes no tempo (heterocedástica)

$$Z_t = T_t \times S_t \times C_t + \epsilon_t \quad (2)$$

Tendência

Vamos iniciar supondo que a série Z_t seja formada apenas pelos componentes de tendência e aleatório:

$$Z_t = T_t + \epsilon_t \quad (3)$$

Estimando a tendência \hat{T}_t , obtemos uma série ajustada para tendência ou livre de tendência, ou seja:

$$\hat{\epsilon}_t = Z_t - \hat{T}_t \quad (4)$$

em que $\hat{\epsilon}_t$ é a estimativa do componente aleatório.

Na literatura há vários métodos para estimar a tendência, sendo que os frequentemente usados são:

1. Ajuste de uma função do tempo, como um polinômio;
2. Suavizar (ou filtrar) os valores da série, por meio de médias móveis;

Tendência Polinomial

Vamos supor que a tendência T_t possa ser representada por um polinômio de ordem m , ou seja:

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_m t^m \quad (5)$$

em que $\beta_0, \beta_1, \dots, \beta_m$ são parâmetros a serem estimadas e $t = 1, 2, \dots, N$ representa o tempo. Para estimar os parâmetros $\beta_j, j = 0, 1, \dots, m$ usamos o método dos mínimos quadrados (MQO).

Geralmente, uma função linear ($m = 1$) ou quadrática ($m = 2$) será apropriada para representar uma tendência monotonicamente crescente ou decrescente.

A tendência na equação (5) é uma função determinística (função matemática) do tempo e algumas vezes é chamada de *tendência global* (i.e. vale para toda a série).

Tendência Linear e Previsão

Supondo $T_t = \beta_0 + \beta_1 t$ (ou seja, $m = 1$), temos:

$$Z_t = \beta_0 + \beta_1 t + \epsilon_t \quad (6)$$

que é um modelo de regressão linear simples, com Z_t a variável dependente e t a variável independente. Aplicando o método de MQO, obtêm-se a série livre de tendência dada por:

$$\hat{\epsilon}_t = Z_t - \hat{T}_t = Z_t - (\hat{\beta}_0 + \hat{\beta}_1 t)$$

A partir do modelo de tendência estimado, podemos prever valores da série, para o instante $t = N + 1, N + 2, \dots$

Exemplo — PIB per capita, Brasil

Considerando a série anual do PIB - paridade do poder de compra (PPC) - per capita - US\$ de 2017, Brasil, período 1980 a 2021, disponível no IPEADATA (arquivo *PIB-Brasil.xlsx*), faça:

1. Gráfico da série;
2. Estime o componente de tendência, supondo que seja linear;
3. Obtenha a série livre de tendência;
4. Faça um gráfico com a série original, a série de tendência e a série livre de tendência;
5. Obtenha o valor previsto para o ano de 2022-2024.

Exercício - PIB da Bélgica

Considerando a série anual do PIB da Bélgica (arquivo *PIB_Anual_Belgica.xlsx*), de 1953 a 2010, faça:

1. Gráfico da série;
2. Estime o componente de tendência;
3. Obtenha a série livre de tendência;
4. Faça um gráfico com a série original, a série de tendência estimada e série livre de tendência.
5. Obtenha o valor previsto para os anos de 2011 a 2020.

Empiricamente, defini-se o componente sazonal como fenômenos que ocorrem regularmente de ano para ano.

Exemplos de séries com componentes sazonais: vendas do comércio na época do Natal e períodos de chuva e de seca no ano que afetam a safra agrícola, etc.

Possíveis modelos sazonais:

1. Variáveis dummies (binárias). O coeficiente de cada variável *dummy* representa o fator sazonal do respectivo mês, trimestre, etc;
2. Médias Móveis para se obter os fatores sazonais.

Sazonalidade por meio de Dummies

Vamos iniciar supondo que a série Z_t seja formada apenas pelos componentes de sazonalidade e o aleatório:

$$Z_t = S_t + \epsilon_t \quad (7)$$

Por simplicidade, vamos assumir que temos uma série temporal com periodicidade trimestral. Neste caso, o componente sazonal será representado pela seguinte regressão:

$$S_t = \beta_0 + \beta_1 D_2 + \beta_3 D_3 + \beta_4 D_4$$

Incluimos *dummies* para cada trimestre exceto o primeiro, isto para não termos problemas de linearidade perfeita. Isto também significa que todos os efeitos medidos por D_i serão relativos ao primeiro trimestre.

Removendo Sazonalidade por meio de Dummies

Ajustada a regressão por MQO podemos fazer um teste F para verificarmos se a presença de uma sazonalidade trimestral é significativamente diferente de zero.

$$Z_t = \hat{\beta}_0 + \hat{\beta}_1 D_2 + \hat{\beta}_3 D_3 + \hat{\beta}_4 D_4 + \hat{\epsilon}_t \quad (8)$$

A rejeição da hipótese nula significa que a série Y_t apresenta sazonalidade.

O modelo (8) pode ser representado por:

$$Z_t = \hat{S}_t + \hat{\epsilon}_t \quad (9)$$

A série original Z_t menos a parte que capta os efeitos da sazonalidade S_t é igual a parte “filtrada” da série, ou seja, a série dessazonalizada (Z^{SA}):

$$\hat{\epsilon}_t = Z_t^{SA} = Z_t - \hat{S}_t$$

Exemplo - Taxa de Desemprego

Para a série mensal de taxa de desemprego - RMSP (arquivo TaxaDesemprego.xlsx), a partir de janeiro de 1985:

- a) Remova a sazonalidade usando a técnica de variáveis dummies.
- b) Obtenha os fatores sazonais. Esses fatores são conjuntamente, significativamente diferentes de zero?
- c) Faça os gráficos da série original, da série sazonal e da série dessazonalizada.

Remoção da Tendência e Sazonalidade

Considerando uma série temporal com periodicidade trimestral, e incluindo *dummies* para cada trimestre exceto o primeiro, para representar o componente sazonal, ou seja:

$$S_t = \alpha_0 + \alpha_2 D_2 + \alpha_3 D_3 + \alpha_4 D_4$$

e supondo o modelo uma função linear para representar a tendência T_t dado por:

$$T_t = \beta_0 + \beta_1 t$$

estima-se por MQO os parâmetros α 's e β 's, e obtemos a estimativa da série \hat{Y}_t :

$$\hat{Y}_t = \hat{\gamma}_0 + \hat{\beta}_1 t + \hat{\alpha}_2 D_2 + \hat{\alpha}_3 D_3 + \hat{\alpha}_4 D_4$$

em que $\hat{\gamma}_0 = (\hat{\beta}_0 + \hat{\alpha}_0)$. Subtraindo a série Y_t da série \hat{Y}_t obtém-se a série do componente errático:

$$\hat{\epsilon}_t = Y_t - \hat{Y}_t$$

Exercício - PIB Agropecuária, Brasil

Para a série trimestral do PIB Agropecuária do Brasil, Fonte IPEADATA, período 2000 T1 a 2023 T3, faça:

- a) Faça o gráfico da série.
- b) Remova a tendência e sazonalidade a partir de um polinômio e variáveis dummies, respectivamente.
- c) Faça o gráfico da série original, o gráfico da série sazonal, o gráfico da série de tendência e o gráfico da série do componente aleatório.
- d) Obtenha o valor previsto para o período 4o trimestre de 2023 ao 4o trimestre de 2024.

Obtenção de Tendência por Médias Móveis

Média móvel é uma média que se movimenta, ou seja, todas as médias móveis têm como padrão um intervalo de período fixo (“janela”) onde para cada nova informação apresentada, retira-se a mais antiga e recalcula-se a média.

Esse tipo de modelo possui o objetivo de suavizar a série temporal, obtendo uma medida de tendência.

As médias móveis mais conhecidas são: Simples e Centrada.

Média Móvel Simples

A média móvel simples é dada pela média aritmética das r observações mais recentes, ou seja:

$$M_t = \frac{Y_t + Y_{t-1} + \dots + Y_{t-r+1}}{r}$$

Ano	Vendas	Média Móvel (3 períodos)
1970	5.3	
1971	7.8	
1972	7.8	6.967
1973	8.7	8.100
1974	6.7	7.733

Média Móvel Centrada

Há duas formas de calcular a média móvel centrada ao se considerar o tamanho da “janela”:

- i. Quando n é ímpar, faz-se a média de n observações consecutivas, colocando o resultado exatamente na posição central:

$$Z_t = \frac{(Y_{t-m} + Y_{t-(m-1)} + \dots + Y_{t+(m-1)} + Y_{t+m})}{n}$$

em que $m = (n - 1)/2$.

Ano	Vendas	Média Móvel (3 períodos)
1970	5.3	
1971	7.8	6.97
1972	7.8	8.10
1973	8.7	7.73
1974	6.7	

Média Móvel Centrada

- ii. Quando n é par, faz-se uma soma ponderada das $n + 1$ observações consecutivas, sendo que a primeira e a última observação têm peso $1/(2n)$, as demais observações têm peso $1/(n)$. O resultado também é colocado exatamente na posição central:

$$Z_t = \frac{Y_{t-m}}{2n} + \frac{(Y_{t-(m-1)} + \dots + Y_{t+(m-1)})}{n} + \frac{Y_{t+m}}{2n}$$

em que $m = n/2$.

Trim	Contratos	Total Móvel de 4 períodos	Total Móvel de 2 períodos	Média Móvel Centrada
I - 2003	24			
II - 2003	21			
		65		
III - 2003	11		126	15.75
		61		
IV - 2003	9			
I - 2004	20			

Índices Sazonais por Médias Móveis

Um dos métodos para se obter os índices sazonais é o método da média móvel:

1. Obter as médias móveis de ordem igual a número de períodos sazonais;
2. Obter médias móveis de 2 períodos, centradas, a partir das médias móveis calculadas no passo 1;
3. Obter os índices sazonais para cada período, dividindo os valores originais da série pelas médias móveis centradas calculadas em 2;
4. Correção dos índices sazonais: obter as medianas dos índices sazonais de cada período (por exemplo, a mediana dos índices sazonais de todos os janeiros existentes na série).

Trim	Contratos	Total Móvel de 4 períodos	Total Móvel de 2 períodos	Média Móvel Centrada	Índices Sazonais	Correção Índice
I - 2003	24					
II - 2003	21					
		65				
III - 2003	11		126	15,75	0,698413	0,626434
		61				
IV - 2003	9		121	15,125	0,595041	0,564187
		60				
I - 2004	20		116	14,5	1,37931	1,421362
		56				
II - 2004	20		109	13,625	1,46789	1,433945
		53				
III - 2004	7		101	12,625	0,554455	
		48				
IV - 2004	6		90	11,25	0,533333	
		42				
I - 2005	15		82	10,25	1,463415	
		40				
II - 2005	14		80	10	1,4	
		40				
III - 2005	5					
IV - 2005	6					

Exemplo

Considerando a série do PIB Agropecuária do Brasil, Fonte IPEADATA, período 2000 T1 a 2023 T4, obtenha os componentes de tendência e sazonal por médias móveis.