

# Aula1\_3\_TendenciaSazonalidade

July 14, 2020

## 1 Tendência e Sazonalidade

por Cíbele Russo

ICMC/USP - São Carlos SP

Baseado em

- Moretting, P.A.; Toloi, C.M.C. “Análise de Séries Temporais”. Blucher, 2004.
- Ehlers, R.S. (2009) Análise de Séries Temporais, <http://www.icmc.usp.br/~ehlers/stemp/stemp.pdf>. Acessado em 28/06/2020.

Considere as observações  $\{Z_t, t = 1, \dots, N\}$  de uma série temporal.

Podemos escrever um modelo de decomposição de  $Z_t$  como uma soma de três componentes não-observáveis:

$$Z_t = T_t + S_t + a_t$$

em que

- $T_t$  representa uma componente de **tendência**
- $S_t$  representa uma componente de **sazonalidade**
- $a_t$  é uma **componente aleatória** (resíduos ou ruído), com média 0 e variância  $\sigma_a^2$ .

Se  $\{a_t\}$  for um ruído branco, então  $Cov(a_t, a_s) = 0$  com  $s \neq t$ , mas podemos eventualmente relaxar essa suposição supondo  $\{a_t\}$  um processo estacionário. Segue-se que, em geral,  $\{Z_t\}$  é uma série não-estacionária.

### 1.1 Tendência

Suponha inicialmente que a componente sazonal  $S_t$  não esteja presente. Um modelo inicial é

$$Z_t = T_t + a_t$$

em que  $a_t$  é um ruído branco, com variância  $\sigma_a^2$ .

Queremos estimar  $T_t$ . Para isso, considere algumas possibilidades:

- i. ajustar uma função no tempo, como um polinômio, uma exponencial ou outra função suave de  $t$ ;
- ii. suavizar (ou filtrar) os valores da série ao redor de um ponto, para estimar a tendência naquele ponto
- iii. suavizar os valores da série através de sucessivos ajustes de retas de mínimos quadrados ponderados ('lowess')

Estimada a tendência,  $\hat{T}_t$ , podemos obter a série ajustada para tendência ou livre de tendência

$$Y_t = Z_t - \hat{T}_t$$

Para eliminar a tendência de uma série, também é possível considerar as diferenças. Normalmente para séries econômicas, a primeira diferença

$$\Delta Z_t = Z_t - Z_{t-1}$$

já é estacionária.

### 1.1.1 Tendência polinomial

Suponha que

$$T_t = \beta_0 + \beta_1 t + \dots + \beta_m t^m$$

onde o grau do polinômio  $m$  é bem menor do que o número de observações  $N$ .

Pode-se utilizar o método de mínimos quadrados para estimar  $\beta_0, \dots, \beta_m$  e obter  $\hat{\beta}_0, \dots, \hat{\beta}_m$ .

```
[1]: # Exemplo tendência polinomial - Carregando as bibliotecas
```

```
import pandas as pd
import numpy as np
import matplotlib.dates as mdates
import matplotlib.pyplot as plt
```

```
[2]: covidSP = pd.read_csv('covidSP.csv', index_col=0, parse_dates=True)
```

```
# Aqui não trabalharemos com as datas mas sim com um índice inteiro das
→ observações
x = np.array(range(len(covidSP.index)))

y = np.array(covidSP['deaths'])

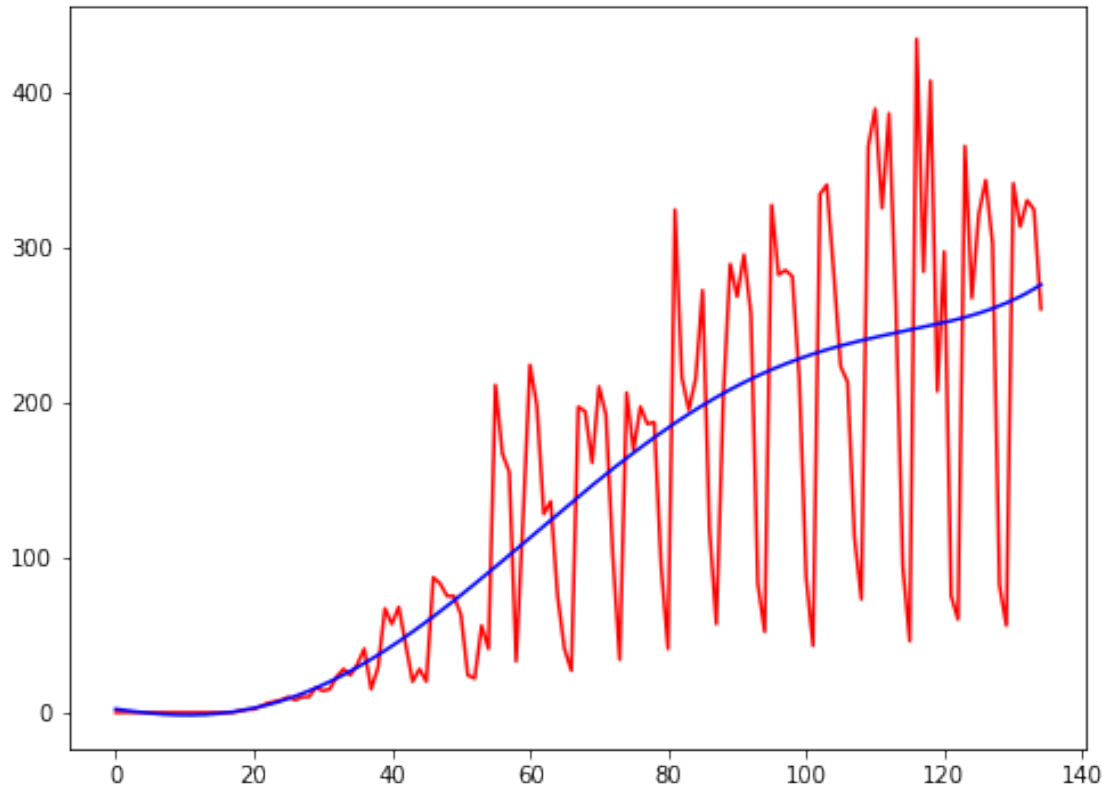
polynomial = np.polyfit(x, y, 5)
p5 = np.poly1d(polynomial)
```

```
xx = np.linspace(x.min(), x.max(), len(y))

plt.rcParams['figure.figsize'] = [8,6]

plt.plot(x,y,label="Real data",color="red") # Plota os dados
plt.plot(xx,p5(xx),label="Polinomio",color="blue") # Polinômio
```

[2]: [<matplotlib.lines.Line2D at 0x7f5ed654da10>]



## 1.2 Sazonalidade

Comportamento da série temporal tende a se repetir a cada s periodos de tempo.

Por exemplo, para dados de COVID-19, os picos em geral ocorrem às terças-feiras e aos fins de semana os registros diminuem, dada a sistemática das notificações.

### Tipos de Sazonalidade

- **Aditiva.** A série apresenta flutuações sazonais mais ou menos constantes não importando o nível global da série.
- **Multiplicativa.** O tamanho das flutuações sazonais varia dependendo do nível global da série.

Considere as observações  $\{Z_t, t = 1, \dots, N\}$  de uma série temporal.

Supondo que não existe tendência mas somente sazonalidade :

$$Z_t = S_t + a_t$$

em que

- $S_t$  representa uma componente de **sazonalidade**
- $a_t$  é um ruído branco com média 0 e variância  $\sigma_a^2$ .

Se a sazonalidade for multiplicativa o modelo é definido como

$$Z_t = S_t a_t$$

### Métodos para estimar a sazonalidade

Para obter a sazonalidade ajustada  $\hat{S}_t$ , podemos utilizar diferentes métodos:

- **Método da regressão** (método determinístico) em que as covariáveis são variáveis periódicas, por exemplo seno, cosseno ou variáveis ‘dummy’, que são variáveis indicadoras
- **Médias móveis** (método estocástico)

E então obtemos a série livre de sazonalidade (ou sazonalmente ajustada)

$$Y_t = Z_t - \hat{S}_t, \text{ no caso aditivo, ou}$$

$$Y_t = Z_t / \hat{S}_t \text{ no caso multiplicativo}$$

Aqui, nós usaremos a função `seasonal_decompose` do pacote `statsmodels`.

```
[3]: import pandas as pd

from pandas import read_csv
from matplotlib import pyplot

pkgdir = '/home/cibele/CibelePython/AprendizadoDinamico/Data'

# Air Passengers: Passageiros aéreos
passageiros = pd.read_csv(f'{pkgdir}/airline_passengers.csv', index_col=0,
                          parse_dates=True)
```

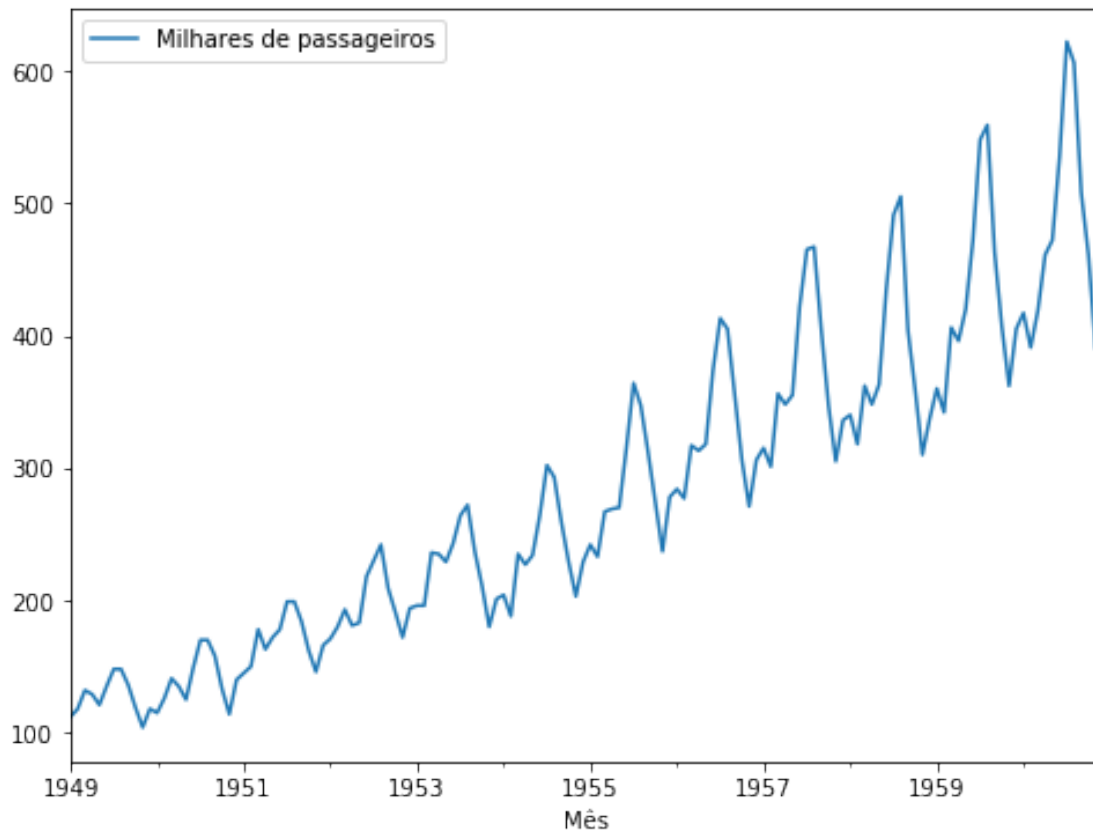
```
[4]: # Para essa análise, os dados faltantes devem ser excluídos!

passageiros = passageiros.dropna()

passageiros.plot()
```

```
# A sazonalidade é aditiva ou multiplicativa?
```

```
[4]: <matplotlib.axes._subplots.AxesSubplot at 0x7f5ed5485f90>
```

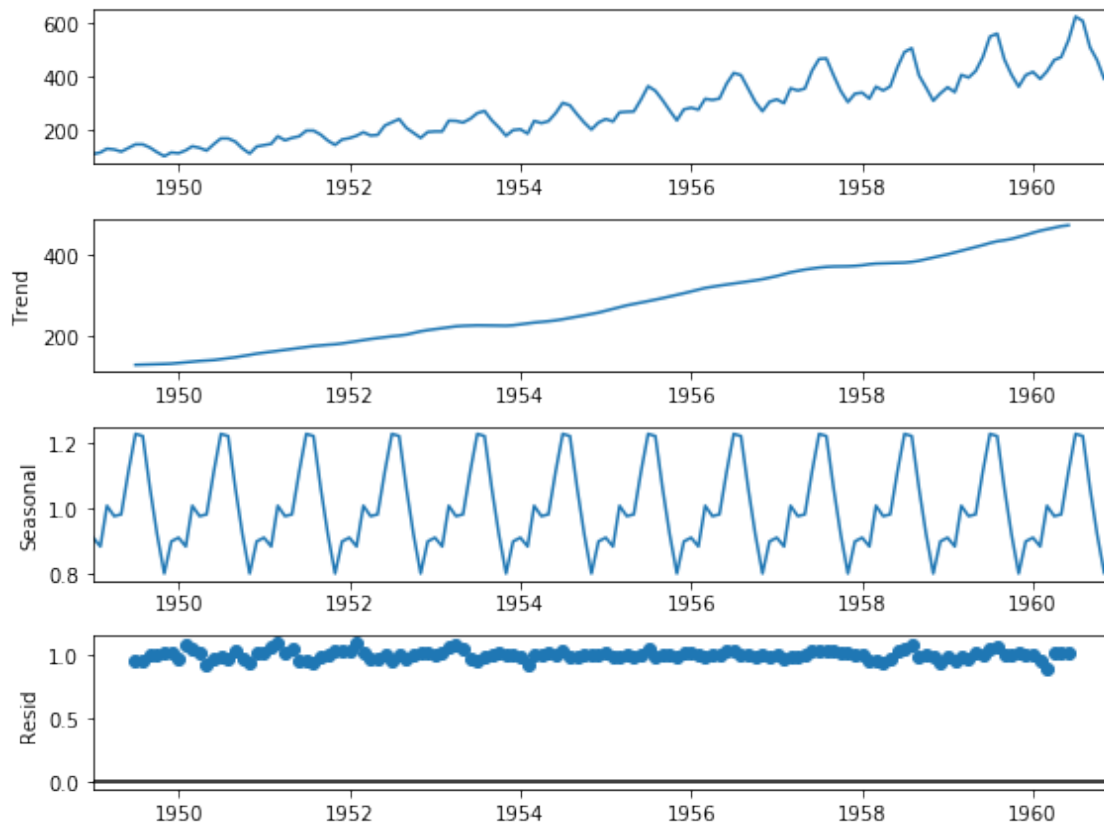


```
[5]: import statsmodels.api as sm

# Decomposição em tendência e sazonalidade

result = sm.tsa.seasonal_decompose(passageiros, model='multiplicative')

result.plot();
```



```
[6]: # Olhe para as componentes do objeto result

result.seasonal
```

```
[6]: Mês
1949-01-01    0.910230
1949-02-01    0.883625
1949-03-01    1.007366
1949-04-01    0.975906
1949-05-01    0.981378
...
1960-08-01    1.219911
1960-09-01    1.060492
1960-10-01    0.921757
1960-11-01    0.801178
1960-12-01    0.898824
Name: seasonal, Length: 144, dtype: float64
```

```
[7]: result.trend.head(15)
```

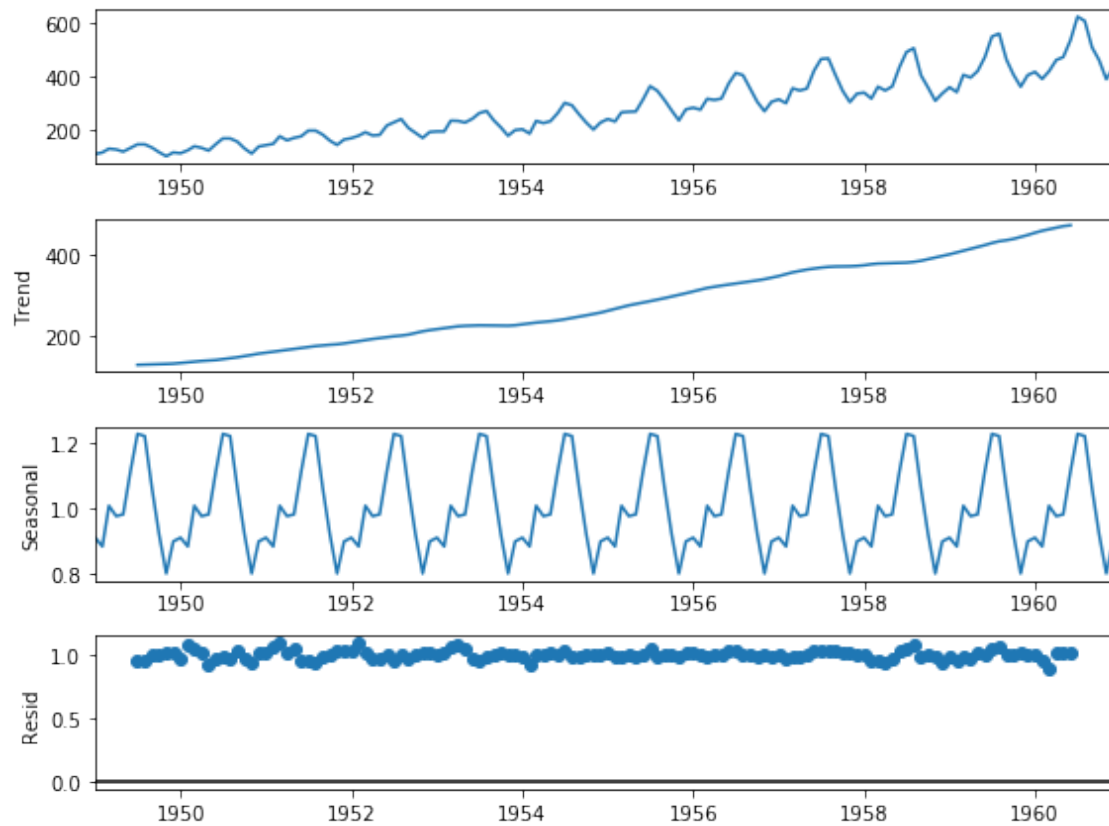
```
[7]: Mês
1949-01-01      NaN
1949-02-01      NaN
1949-03-01      NaN
1949-04-01      NaN
1949-05-01      NaN
1949-06-01      NaN
1949-07-01    126.791667
1949-08-01    127.250000
1949-09-01    127.958333
1949-10-01    128.583333
1949-11-01    129.000000
1949-12-01    129.750000
1950-01-01    131.250000
1950-02-01    133.083333
1950-03-01    134.916667
Name: trend, dtype: float64
```

### Outra forma de chamar a função

```
[8]: import pandas as pd
from statsmodels.tsa.seasonal import seasonal_decompose

result = seasonal_decompose(passageiros, model='multiplicative')

result.plot();
```



### Exercício

Considere os dados de CO2 e da Jhonson & Johnson. Como ficam as decomposições em Tendência e Sazonalidade?