Análise de Séries Temporais

Modelo ARIMA Sazonal (SARIMA)

José Augusto Fiorucci

- Seja “s” o número de observações por ciclo sazonal (geralmente anual).
  - Exemplo:
    - série mensal: s=12
    - série bimestral: s=6
    - série trimestral: s=4

- Operador de diferenças sazonais (∇_s)

$$\nabla_s x_t = x_t - x_{t-s}$$

  - Exemplo:
    - série mensal: $\nabla_{12} x_t = x_t - x_{t-12}$
    - série bimestral: $\nabla_6 x_t = x_t - x_{t-6}$
    - série trimestral: $\nabla_4 x_t = x_t - x_{t-4}$

  - No caso de ordens maiores, o operador é aplicado recursivamente.
    - Exemplo: $\nabla_s^2 x_t = \nabla_s(\nabla_s x_t) = \nabla_s x_t - \nabla_s x_{t-s} = x_t - 2x_{t-s} + x_{t-2s}$

  - Em termos do operador de retardo, note que

$$\nabla_s x_t = (1 - B^s) x_t$$

  e portanto,

$$\nabla_s^D x_t = (1 - B^s)^D x_t$$

- **sAR(P)** - Sazonal Puro Autoregressivo de ordem P

$$x_t = \varphi_1 x_{t-s} + \varphi_2 x_{t-2s} + \cdots + \varphi_P x_{t-Ps} + \varepsilon_t$$

em que $\varphi_1, \varphi_2, \ldots, \varphi_P$ são os coeficientes autoregressivos sazonais.

Equivalentemente, podemos escrever

$$\Phi_P(B^s)x_t = \varepsilon_t$$

em que $\Phi_P(B^s) = 1 - \varphi_1 B^s - \varphi_2 B^{2s} - \cdots - \varphi_P B^{Ps}$ é o polinômio autoregressivo sazonal.

  - note que o modelo sAR(P) consiste basicamente em um modelo AR de ordem $p = P \times s$, em que os coeficientes não sazonais são assumidos como sendo zero.
  - condição de estacionariedade: a mesma do modelo AR($p = P \times s$)

- **sMA(Q)** - Sazonal Puro Médias Moveis de ordem Q

$$x_t = \varepsilon_t + \vartheta_1 \varepsilon_{t-s} + \vartheta_2 \varepsilon_{t-2s} + \cdots + \vartheta_Q \varepsilon_{t-Qs}$$

  em que $\vartheta_1, \vartheta_2, \ldots, \vartheta_Q$ são os coeficientes de médias móveis sazonais.

  Equivalentemente, podemos escrever

$$x_t = \Theta_Q(B^s)\varepsilon_t$$

  em que $\Theta_Q(B^s) = 1 + \varphi_1 B^s + \varphi_2 B^{2s} + \cdots + \varphi_Q B^{Qs}$ é o polinômio de médias moveis sazonal.

  - note que o modelo sMA(Q) consiste basicamente em um modelo MA de ordem $q = Q \times s$, em que os coeficientes não sazonais são assumidos como sendo zero.
  - condição de estacionariedade: sempre estacionário
  - condição de inversibilidade: a mesma do modelo MA($q = Q \times s$)

Modelos Sazonais Puros

- **sARMA(P,D,Q)** - Sazonal Puro ARMA(P,Q)

$$\Phi_P(B^s)\,x_t = \Theta_Q(B^s)\,\varepsilon_t$$

  - note que o modelo sARMA(P,Q) consiste basicamente em um modelo ARMA(p,q) de ordem $p = P \times s$ e $q = Q \times s$, em que os coeficientes não sazonais são assumidos como sendo zero.
  - condição de estacionariedade: a mesma do modelo AR($p = P \times s$)
  - condição de inversibilidade: a mesma do modelo MA($q = Q \times s$)

- **sARIMA(P,Q)** - Sazonal Puro ARIMA(P,D,Q)

$$\Phi_P(B^s)\,\nabla_s^D\,x_t = \Theta_Q(B^s)\,\varepsilon_t$$

  - note que $w_t = \nabla_s^D\,x_t$ é um processo estacionário sARMA(P,Q).
  - generaliza todos os modelos sazonais puros vistos.

Modelo ARIMA Sazonal - SARIMA

- A mistura dos modelos *Sazonal Puros ARIMA* com os modelos *ARIMA não sazonais* da origem aos modelos conhecidos como *ARIMA Sazonal*, ou simplesmente, SARIMA.

- O modelo SARIMA(p,d,q)×(P,D,Q) é definido como

$$\Phi_P(B^s)\,\Phi_p(B)\,\nabla_s^D\,\nabla^d\,x_t = \Theta_Q(B^s)\,\Theta_q(B)\,\varepsilon_t$$

  - Generaliza todos os modelos da família ARIMA
  - Permite modelar séries estacionárias e não estacionárias, bem como séries sazonais e não sazonais

- Exemplo 1: Se $\{x_t\}$ segue um processo SARIMA(1,0,1)$\times$(1,0,1), qual a função da média condicional e dos resíduos do modelo?
  - Abrindo a equação do modelo

$$
\begin{aligned}
&\Phi_1(B^s)\,\Phi_1(B)\,x_t = \Theta_1(B^s)\,\Theta_1(B)\,\varepsilon_t \\
\Rightarrow\ & (1-\varphi B^s)(1-\phi B)\,x_t = (1+\vartheta B^s)(1+\theta B)\varepsilon_t \\
\Rightarrow\ & (1-\phi B-\varphi B^s+\phi\varphi B^{s+1})\,x_t = (1+\theta B+\vartheta B^s+\theta\vartheta B^{s+1})\varepsilon_t \\
\Rightarrow\ & x_t-\phi\,x_{t-1}-\varphi\,x_{t-s}+\phi\varphi\,x_{t-s-1} = \varepsilon_t+\theta\varepsilon_{t-1}+\vartheta\varepsilon_{t-s}+\theta\vartheta\varepsilon_{t-s-1} \\
\Rightarrow\ & x_t = \phi\,x_{t-1}+\varphi\,x_{t-s}-\phi\varphi\,x_{t-s-1}+\varepsilon_t+\theta\varepsilon_{t-1}+\vartheta\varepsilon_{t-s}+\theta\vartheta\varepsilon_{t-s-1}
\end{aligned}
$$

  - Portanto, a média condicional é dada por

$$\mu_{t|t-1} = \phi\,x_{t-1}+\varphi\,x_{t-s}-\phi\varphi\,x_{t-s-1}+\theta\varepsilon_{t-1}+\vartheta\varepsilon_{t-s}+\theta\vartheta\varepsilon_{t-s-1}$$

  - Se $\widehat{\beta}$ é a estimativa do vetor de parâmetros, então os resíduos são dados por

$$\varepsilon_t(\widehat{\beta}) = x_t-\mu_{t|t-1}(\widehat{\beta}), \qquad t = s+2, ..., n$$

  - Obs: Para inicializar o modelo, geralmente é tomado $\varepsilon_1 = \cdots = \varepsilon_{s+1} = 0$, ficando assim perdidos os erros do primeiro ciclo sazonal. Desta forma $\mu_{t|t-1}$ pode ser calculado no intervalo $s+1 < t \leq n+1$.

- Exemplo 2: Repita o exemplo 1 para o modelo SARIMA(1,1,1)×(1,1,1).
  - Modelo
  $$\Phi_1(B^s)\,\Phi_1(B)\nabla_s\nabla\, x_t = \Theta_1(B^s)\,\Theta_1(B)\,\varepsilon_t$$
  - Série estacionária
  $$w_t = \nabla_s\nabla\, x_t = \nabla_s\,(x_t - x_{t-1}) = x_t - x_{t-s} - x_{t-1} + x_{t-s-1}, \quad t > s+1$$
  - Sabemos que $\{w_t\}$ é SARIMA(1,0,1)×(1,0,1). Do exemplo 1, temos que a média condianal do $w_{t|t-1}$ é dada por
  $$\mu^{(w)}_{t|t-1} = \phi\, w_{t-1} + \varphi\, w_{t-s} - \phi\varphi\, w_{t-s-1} + \theta\varepsilon_{t-1} + \vartheta\varepsilon_{t-s} + \theta\vartheta\varepsilon_{t-s-1}, \quad t = 2s+3, ..., n$$
  e portanto os resíduos do modelo
  $$\varepsilon_t(\widehat{\beta}) = w_t - \mu^{(w)}_{t|t-1}(\widehat{\beta}), \qquad t = 2s+3, ..., n$$
  - Das equações anteriores, segue que a média condicional do x_t é dada por
  $$\begin{aligned}\mu_{t|t-1} &= E_{t-1}[x_t] \\ &= E_{t-1}[w_t + x_{t-s} + x_{t-1} - x_{t-s-1}] \\ &= E_{t-1}[w_t] + x_{t-s} + x_{t-1} - x_{t-s-1} \\ &= \mu^{(w)}_{t|t-1} + x_{t-s} + x_{t-1} - x_{t-s-1}\end{aligned}$$
  assim $\mu_{t|t-1}$ pode ser calculado recursivamente para $t = 2s+3, ..., n$.
  - Note que pelo menos dois ciclos sazonais são perdidos na inicialização do modelo, o que pode inviabilizar aplicações em séries com poucas observações.

- Def.: Um processo $\{x_t\}$ tem *raiz unitária sazonal* se o número 1 é raiz do polinômio autoregressivo sazonal.
  - Assim como os processos de raiz unitária vistos na Aula 7, processos com raiz unitária sazonal são não estacionários.
  - Exemplo: Considere o processo sazonal definido como $x_t = x_{t-s} + \varepsilon_t$
    - Note que o processo pode ser escrito como $(1 - B^s)x_t = \varepsilon_t$, e portanto, $B = 1$ é raiz do polinômio autogressivo sazonal, dado por, $\Phi(B^s) = 1 - B^s$.
    - Logo, esse processo é não estacionário.

- Obs: qualquer processo $SARIMA\,(p, d, q) \times (P, D, Q)$ em que $d \geq 1$ ou $D \geq 1$, é não estacionário.
  - Note que o processo do exemplo anterior consiste em um $SARIMA\,(0, 0, 0) \times (0, 1, 0)$

- Exercício: para o processo

$$x_t = 0.5x_{t-1} + x_{t-s} - 0.5x_{t-s-1} + \varepsilon_t - 0.3\varepsilon_{t-1}$$

identifique a ordem do SARIMA correspondente e verifique se é estacionário.

Testes para raiz unitária sazonal

A função *nsdiffs()* do pacote *forecast* verifica quantas diferenças sazonais são necessárias para tornar a série estacionária, baseado em um dos testes de raiz unitária sazonal, listados abaixo:

1. Wang, X, Smith, KA, Hyndman, RJ (2006) "Characteristic-based clustering for time series data", Data Mining and Knowledge Discovery, 13(3), 335-364.

2. Osborn DR, Chui APL, Smith J, and Birchenhall CR (1988) "Seasonality and the order of integration for consumption", Oxford Bulletin of Economics and Statistics 50(4):361-377.

    - função: *forecast::ocsb.test()*

3. Canova F and Hansen BE (1995) "Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability", Journal of Business and Economic Statistics 13(3):237-252.

    - função: *uroot::ch.test()*

4. Hylleberg S, Engle R, Granger C and Yoo B (1990) "Seasonal integration and cointegration.", Journal of Econometrics 44(1), pp. 215-238.

    - função: *uroot::hegy.test()*

Obs: de modo geral, se for necessário utilizar diferenças sazonais, então $D = 1$ será suficiente para a grande maioria das séries.