Modelos SARIMA (Auto-Regressivo Integrado de Médias Móveis com Sazonalidade): parte 2

Prof. Caio Azevedo

Previsão

- Considere $Y_1, ..., Y_t$ uma processo $SARIMA(p, d, q)(P, D, Q)_s$, ou seja:

$$\begin{aligned} \Phi(B^s)\phi(B)(1-B^s)^D(1-B)^d Y_t &= \delta + \Theta(B^s)\theta(B)\epsilon_t, \\ \epsilon_t &\sim RB(0, \sigma^2) \end{aligned}$$

- Uma forma de calcular previsões, é utilizarmos resultados desenvolvidos para os modelos ARMA da seguinte forma (próximo slide, e nos seguintes, apresentaremos alguns exemplos):

Previsão

1. Defina: $X_t = (1-B)^d(1-B^s)^D Y_t$, então $\{X_t\}$ é um processo ARMA(p+Ps,q+Qs) com restrições (aqui), ou seja:

$$\Phi(B^s)\phi(B)X_t = \delta + \Theta(B^s)\theta(B)\epsilon_t, \quad (1)$$

de modo que já sabemos como calcular as previsões $X_n(k)$ (aqui). (Note que a amostra disponível para tal tarefa é $x_{d+Ds+1}, ..., x_n$).

2. A partir de (1) encontramos uma expressão para $Y_n(k)$ que depende de $\{Y_n, Y_{n-1}, ...\}$ e de $X_n(t), t \leq k$

Exemplos (previsão)

- ARIMA(1,1,0) causal: $(1-\phi B)(1-B)Y_t = \epsilon_t$. Previsão:
  1. Seja $X_t = (1-B)Y_t$, então $X_t \sim AR(1)$ e, portanto $X_n(k) = \phi^k X_n$.
  2. Como $X_t = Y_t - Y_{t-1}$, então, aplicando o operador $\mathcal{E}\left[.|\mathcal{F}_n\right]$ em X_{n+k}, obtemos $X_n(k) = Y_n(k) - Y_n(k-1)$. Logo, $Y_n(k) = X_n(k) + Y_n(k-1)$. Portanto,

$$
\begin{aligned}
Y_n(1) &= \phi X_n + Y_n \\
Y_n(2) &= \phi(1+\phi)X_n + Y_n \\
&\vdots \text{ por indução} \\
Y_n(k) &= \frac{\phi(1-\phi^k)}{1-\phi}\left(Y_n - Y_{n-1}\right) + Y_n
\end{aligned}
$$

- Assim, se $k \to \infty$, então $Y_n(k) \to \left(\frac{1}{1-\phi}\right) Y_n + \left(-\frac{\phi}{1-\phi}\right) Y_{n-1}$

Exemplos (previsão)

- Modelo Airline: SARIMA$(0,1,1)(0,1,1)_{12}$:

$$(1-B)(1-B^{12})Y_t = (1+\theta B)(1+\Theta B^s)\epsilon_t.$$

- Previsão:

1. Seja $X_t = (1-B)(1-B^{12})Y_t$, então

$$\begin{aligned} X_t &= (1+\theta B)(1+\Theta B^s)\epsilon_t \\ X_t &= \epsilon_t + \theta\epsilon_{t-1} + \Theta\epsilon_{t-12} + \theta\Theta\epsilon_{t-13}. \end{aligned} \tag{2}$$

Vamos encontrar previsões para $\{X_t\}$. Uma vez que:

$$\begin{aligned} X_n(k) &= \mathcal{E}(\epsilon_{n+k}|\mathcal{F}_n) + \theta\mathcal{E}(\epsilon_{n+k-1}|\mathcal{F}_n) + \Theta\mathcal{E}(\epsilon_{n+k-12}|\mathcal{F}_n) \\ &+ \theta\Theta\mathcal{E}(\epsilon_{n+k-13}|\mathcal{F}_n). \end{aligned}$$

Cont.

1. (Cont.) Precisamos estimar $\mathcal{E}(\epsilon_t|\mathcal{F}_n)$ quando $t \leq n$. Seja ϵ_t^n o estimador de $\mathcal{E}(\epsilon_t|\mathcal{F}_n)$, então:

$$\begin{aligned}
X_n(1) &= \theta\epsilon_n^n + \Theta\epsilon_{n-11}^n + \theta\Theta\epsilon_{n-12}^n \\
X_n(2) &= \Theta\epsilon_{n-10}^n + \theta\Theta\epsilon_{n-11}^n \\
X_n(3) &= \Theta\epsilon_{n-9}^n + \theta\Theta\epsilon_{n-10}^n \\
&\vdots \\
X_n(12) &= \Theta\epsilon_n^n + \theta\Theta\epsilon_{n-1}^n \\
X_n(13) &= \theta\Theta\epsilon_n^n \\
X_n(k) &= 0, k \geq 14
\end{aligned}$$

Cont.

1. (Cont.) Os valores $\epsilon_n^n, \epsilon_{n-1}^n, ..., \epsilon_{n-12}^n$ são calculados recursivamente. Assim, de (2), temos que:

$$\epsilon_t = X_t - \theta\epsilon_{t-1} - \Theta\epsilon_{t-12} - \theta\Theta\epsilon_{t-13},$$

então

$$\epsilon_t^n = X_t - \theta\epsilon_{t-1}^n - \Theta\epsilon_{t-12}^n - \theta\Theta\epsilon_{t-13}^n.$$

Note que amostra disponível em termos do processo $\{X_t\}$, é $X_{14}, X_{15}, ..., X_n$. Portanto, considerando que $\epsilon_j^n = 0$, $j \leq 13$, obtemos (próximo slide):

Cont.

1. (Cont.)

$$
\begin{aligned}
\epsilon_{14}^{n} &= X_{14} \\
\epsilon_{15}^{n} &= X_{15} - \epsilon_{14}^{n} \\
&\vdots \\
\epsilon_{27}^{n} &= X_{27} - \theta\epsilon_{26}^{n} - \Theta\epsilon_{15}^{n} - \theta\Theta\epsilon_{14}^{n} \\
&\vdots \\
\epsilon_{n}^{n} &= X_{27} - \theta\epsilon_{n-1}^{n} - \Theta\epsilon_{n-12}^{n} - \theta\Theta\epsilon_{n-13}^{n}.
\end{aligned}
$$

Cont.

2. Como $X_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$, então:

$$Y_n(k) = X_n(k) + Y_n(k-1) + Y_n(k-12) + Y_n(k-13)$$

3. Por exemplo,

$$\begin{aligned} Y_n(1) &= X_n(1) + Y_n + Y_{n-11} - Y_{n-12} \\ Y_n(k) &= Y_n(k-1) + Y_n(k-12) - Y_n(k-13), k \geq 14 \end{aligned}$$

Previsão

- Modelo ARIMA(0,1,1) e EWMA ("Exponentially Weighted Moving Average"). Seja o modelo:

$$Y_t = Y_{t-1} - \lambda \epsilon_{t-1} + \epsilon_t,$$

em que $|\lambda| < 1$, $Y_0 = 0$, $\epsilon_t \sim RB(0, \sigma^2)$. Considere

$$Z_t = \epsilon_t - \lambda \epsilon_{t-1},$$

então $Y_t = Y_{t-1} + Z_t$. Logo,

$$Z_t = Y_t - Y_{t-1}. \tag{3}$$

Previsão

- Como $\{Z_t\}$ corresponde a um processo MA(1), portanto é invertível e pode ser (re)escrito como

$$Z_t = -\sum_{j=1}^{\infty} \lambda^j Z_{t-j} + \epsilon_t. \tag{4}$$

- Substituindo (3) em (4), temos que:

$$\begin{aligned} Y_t &= (1-\lambda)\sum_{j=1}^{\infty} \lambda^{j-1} Y_{t-j} + \epsilon_t \\ &= \widetilde{Y}_t + \epsilon_t, \\ \widetilde{Y}_t &= (1-\lambda)\sum_{j=1}^{\infty} \lambda^{j-1} Y_{t-j}. \end{aligned} \tag{5}$$

Previsão

- Assim, (5) é uma EWMA ("Exponentially Weighted Moving Average") e pode ser (re)escrito como :

$$\widetilde{Y}_{t+1} = (1-\lambda)\,Y_t + \lambda\widetilde{Y}_t$$

- Assim, $\widetilde{Y}_{t+1}$ é atualizado de forma simples, através de uma média ponderada entre a última observação (Y_t) e o valor (predito) prévio ($\widetilde{Y}_t$).
- O parâmetro (λ) é uma constante de suavização que pode ser determinada a priori (assumindo valores no intervalo [0,1], para algumas aplicações) ou estimada através da minimização de $S(\lambda) = \sum_t \left(Y_t - \widetilde{Y}_t\right)^2$.

Previsão

- O método EWMA é considerado um método de previsão **ad-hoc**. Aqui foi demonstrado que a respectiva fórmula surge a partir de um modelo ARIMA.
- Contudo, ao utilizarmos a abordagem ARIMA temos maior flexibilidade para a modelagem (estimação/previsão).
- De qualquer forma, um método bastante útil para predizer a volatilidade em séries financeiras é baseado no EWMA.

Previsão

- Se o processo é não estacionário então a incerteza das previsões aumenta à medida que aumenta o horizonte de previsão.
- Podem ser calculados intervalos de previsão aproximados supondo normalidade do ruído branco (veja aqui e aqui).
- Comparações em termos de previsão (para valores observados e futuros) podem ser feitos através das metodologias vistas aqui.

Modelagem

- Devemos, primeiramente, entender o problema, levantando aspectos teóricos, da literatura e informações dos responsáveis pelo problema.
- Devemos observar os gráficos de ST, FAC e FACP, tentando visualizar:
  - Se a ST parecer ser estacionária ou não estacionária (na média e/ou na variância).
  - Os padrões da FAC e da FACP.
  - Existência de sazonalidade/ciclos (esta(s), pode(m) ser, eventualmente, modelada(s) usando modelos ARIMA) e sua(s) ordem(ns).
  - Estrutura e ordem da dependência.
- Nos próximos slides, detalharemos as etapas acima.

Modelagem

- No estágio de identificação de modelos, primeiro observamos se a série apresenta tendência e/ou sazonalidade (tentando determinar o valor do período sazonal, s). Ferramentas para fazer isso incluem o gráfico da série e os gráficos da FAC e FACP.
- Na análise da parte sazonal, nas: FAC e FACP, se observarmos que o decaimento das autocorrelações nas defasagens múltiplos de s é muito devagar, então pode ser conveniente diferenciar a ST, sazonalmente.
- Na parte regular, se observarmos que as autocorrelações decaem (muito) vagarosamente, então pode ser conveniente diferenciar.

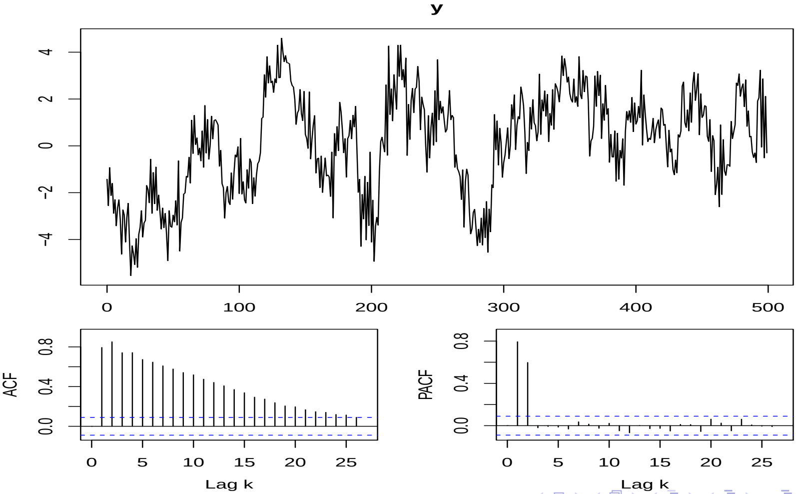
Modelagem

- Outra evidência ("no curso") de que é necessário diferenciar a ST é quando ela apresentar tendência.
- Na prática são escolhidos valores de (d, D) iguais a 0, 1 ou 2. A série obtida após diferenciar a ST original, em relação as partes sazonal e regular, tem que apresentar comportamento estacionário.
- Devemos ter cuidado com a sobre-diferenciação.
- Com base na série original e nas FAC e FACP, conjecturamos um (ou mais) conjunto(s) de valores para p, P, q, Q. Claro que, uma vez que descritivamente é difícil determinar tais valores, alguns modelos podem ser ajustados e comparados.

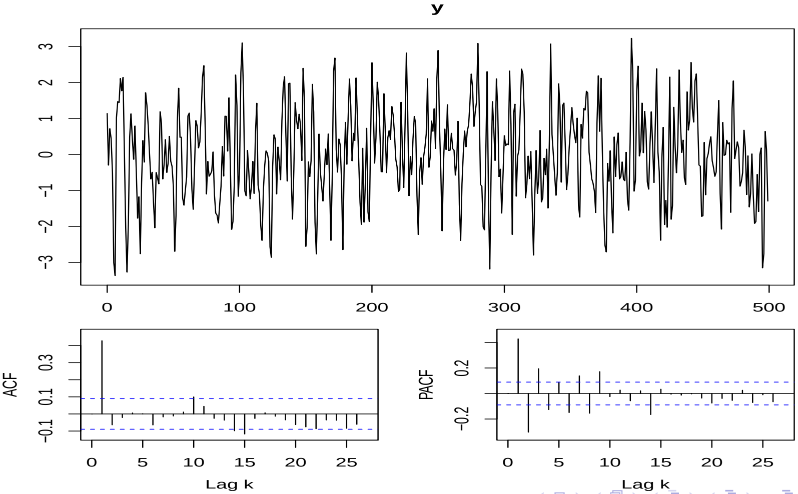
Modelagem

- A estimação dos modelos competidores é feita, usualmente, por máxima verosimilhança.
- Posteriormente, cada modelo deve ser submetido à uma análise residual apropriada. No curso, pelo menos a ACF dos resíduos deve mostrar que todas elas são não significativas.
- Com relação à comparação de modelos, podemos utilizar os critério de informação (aqui) e os de qualidade da previsão (slide 14 deste material).
- A seguir, apresentamos gráficos de algumas ST simuladas (sem a pretensão de ser exaustivo).

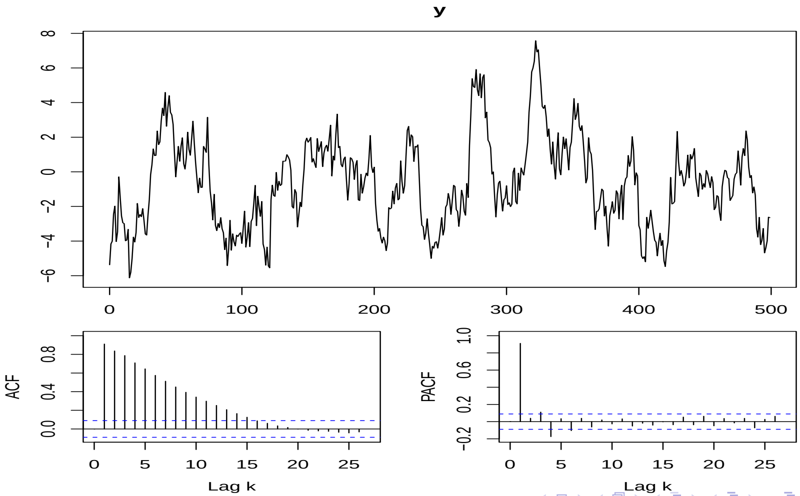
$SARIMA(2,0,0)(0,0,0)_{s=0} = AR(2)$



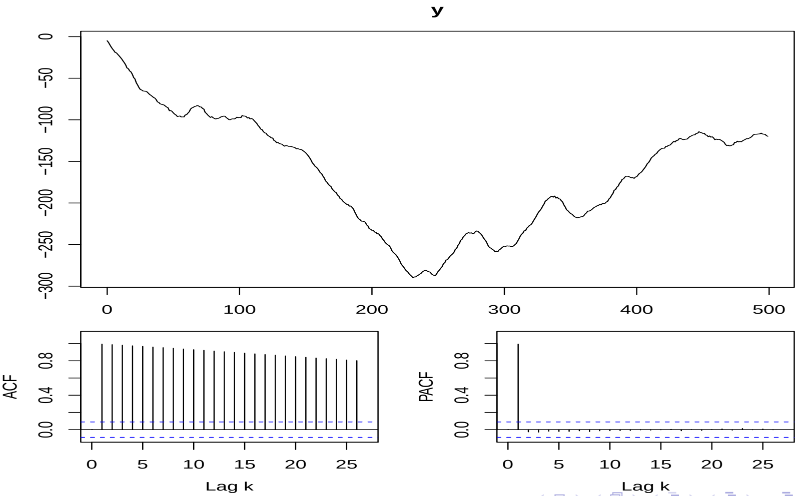
$SARIMA(0,0,2)(0,0,0)_{s=0} = MA(2)$



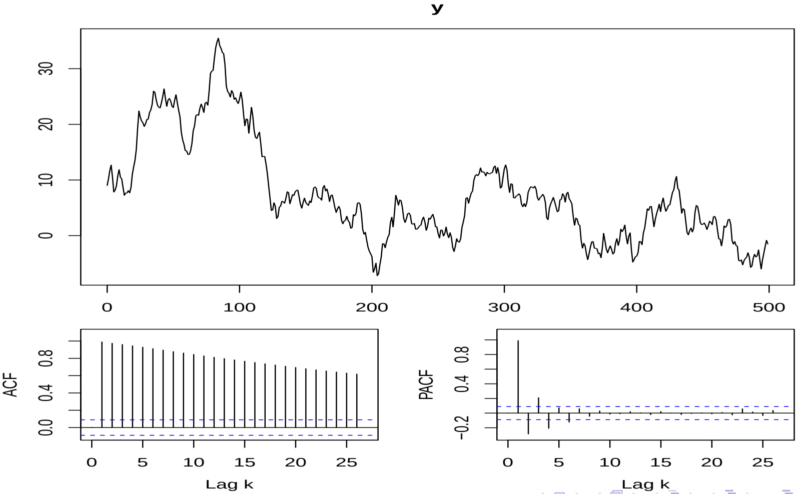
$SARIMA(2,0,2)(0,0,0)_{s=0} = ARMA(2,2)$



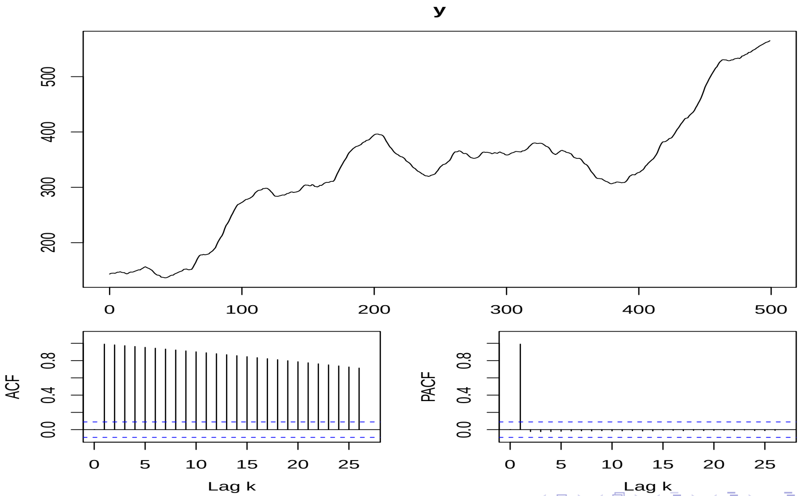
$SARIMA(2,1,0)(0,0,0)_{s=0} = ARIMA(2,1,0)$



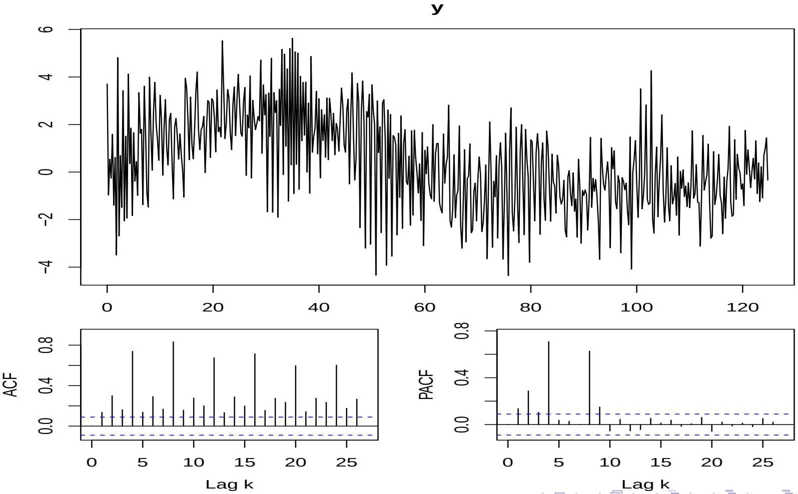
$SARIMA(0,1,2)(0,0,0)_{s=0} = ARIMA(0,1,2)$



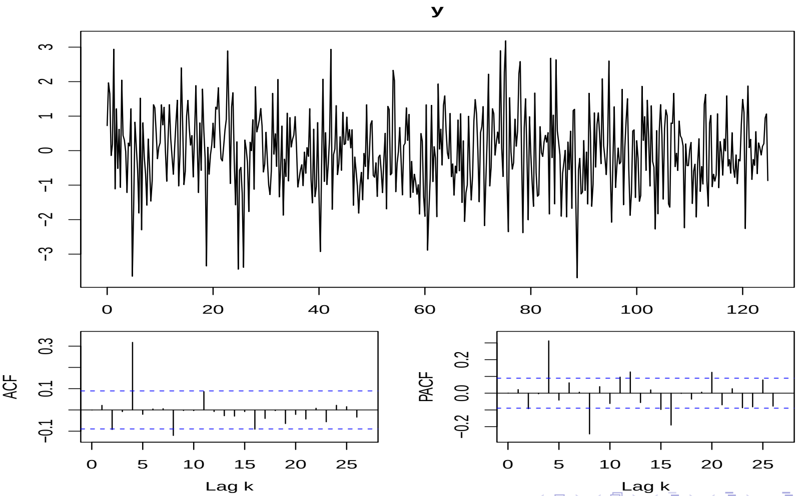
$SARIMA(2,1,2)(0,0,0)_{s=0} = ARIMA(2,1,2)$



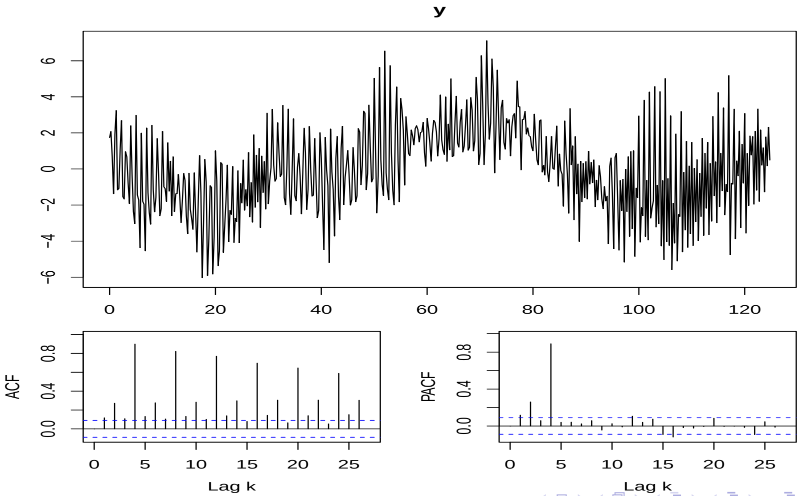
$SARIMA(0,0,0)(2,0,0)_{s=4} = SARMA(2,0)_{s=4}$



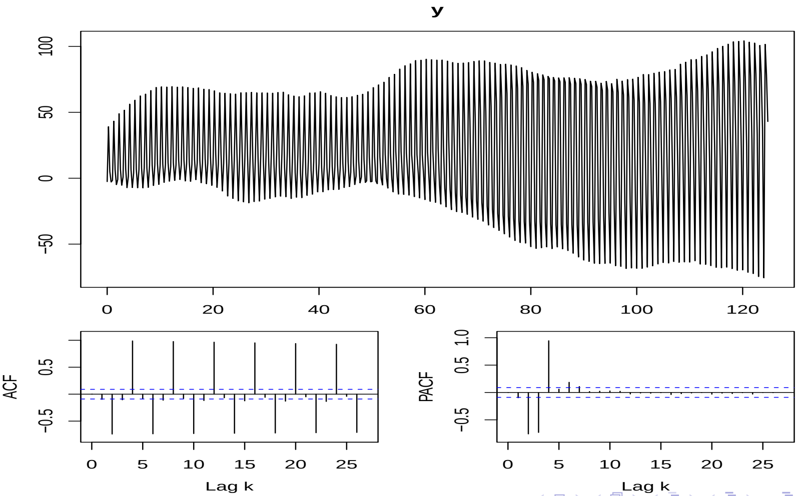
$SARIMA(0,0,0)(0,0,2)_{s=4} = SARMA(0,2)_{s=4}$



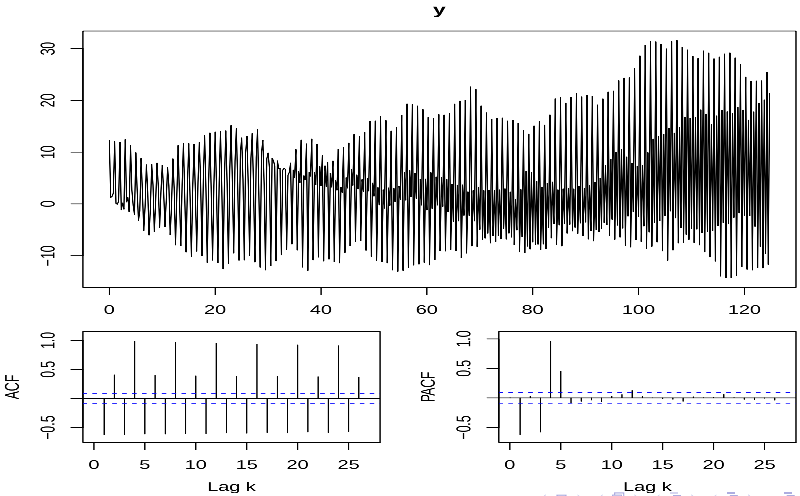
$SARIMA(0,0,0)(2,0,2)_{s=4} = SARMA(2,2)_{s=4}$



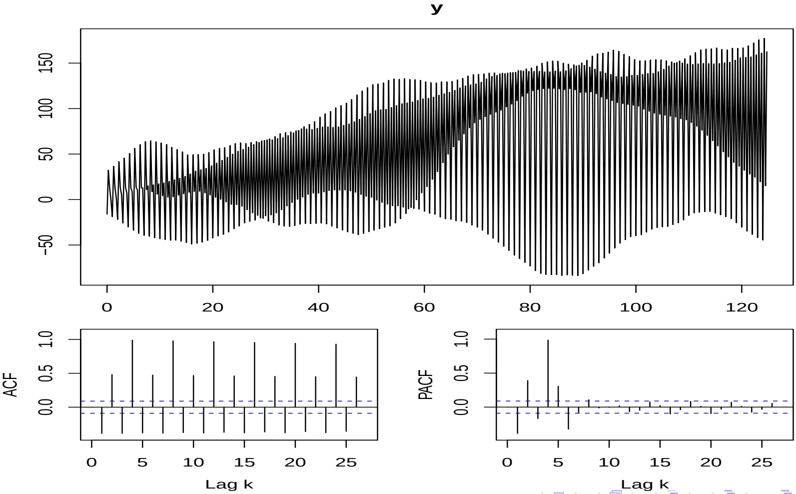
$SARIMA(0,0,0)(2,1,0)_{s=4} = SARIMA(2,1,0)_{s=4}$



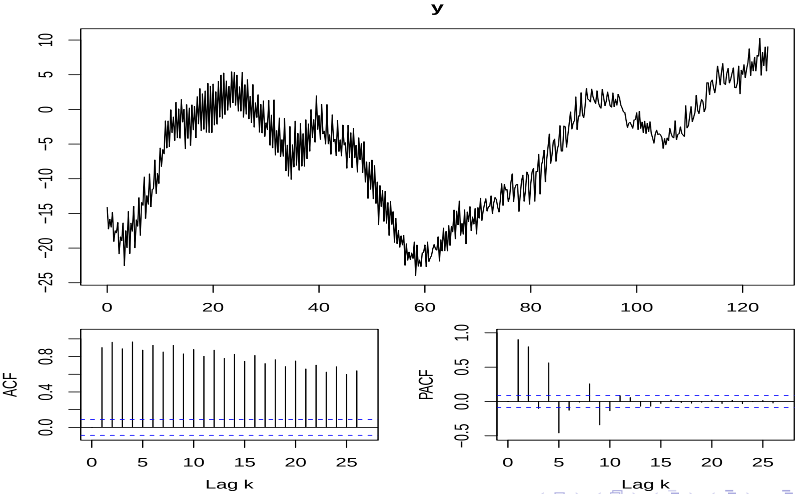
$SARIMA(0,0,0)(0,1,2)_{s=4} = SARIMA(0,1,2)_{s=4}$



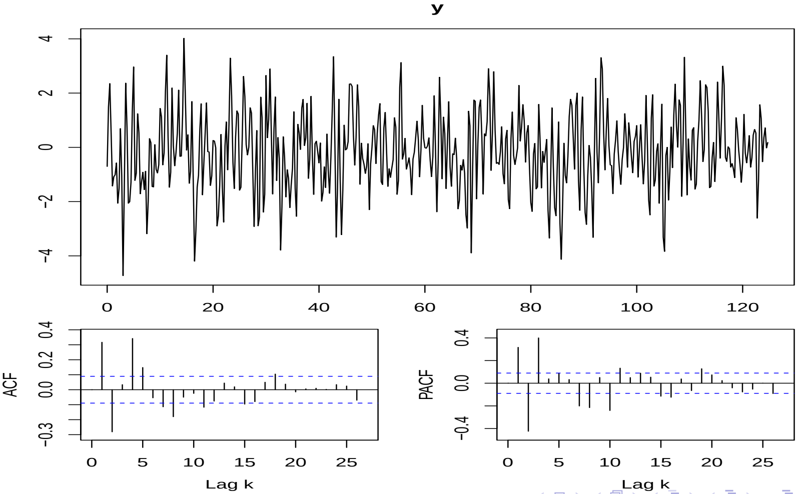
$SARIMA(0,0,0)(2,1,2)_{s=4} = SARIMA(2,1,2)_{s=4}$



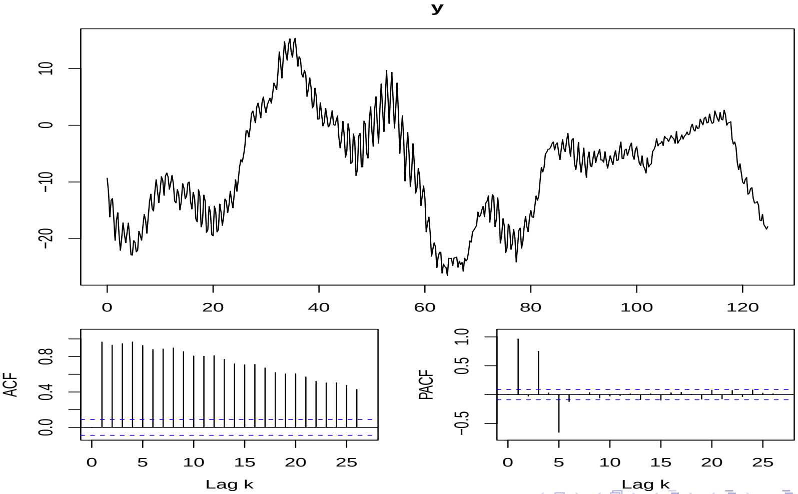
$SARIMA(2,0,0)(2,0,0)_{s=4} = SARMA(2,0)(2,0)_{s=4}$



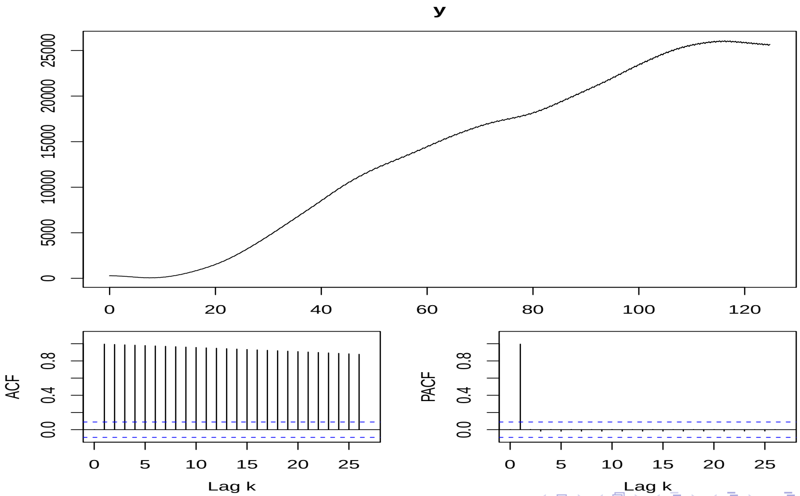
$SARIMA(0,0,2)(0,0,2)_{s=4} = SARMA(0,2)(0,2)_{s=4}$



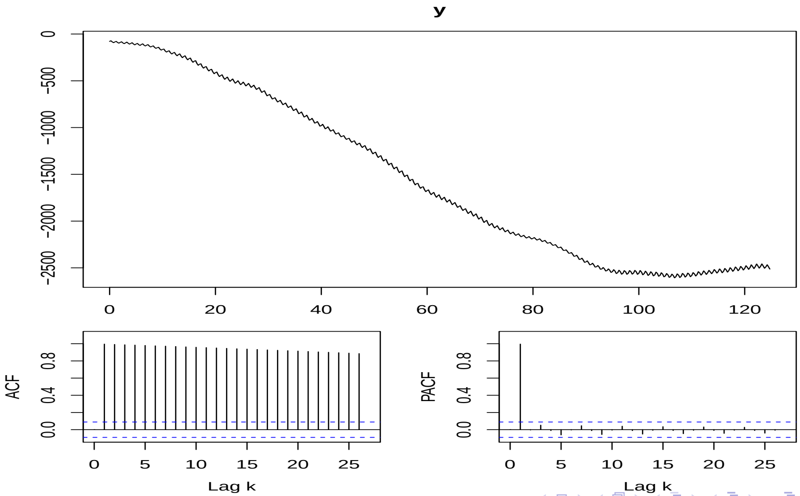
$SARIMA(2,0,2)(2,0,2)_{s=4} = SARMA(2,2)(2,2)_{s=4}$



$SARIMA(2,1,0)(2,1,0)_{s=4}$



$SARIMA(0,1,2)(0,1,2)_{s=4}$



$SARIMA(2,1,2)(2,1,2)_{s=4}$