

Processos autoregressivos (parte 4)

Prof. Caio Azevedo

Introdução

- Discutiremos a respeito de diagnóstico do modelo e previsão (para os instantes observados e futuros).
- Métodos de diagnósticos são de fundamental importância no processo inferencial. Essencialmente, caso as suposições do modelo não sejam satisfeitas:
 - 1 Os resultados desenvolvidos (EP, IC, TH etc) não são válidos, parcial ou totalmente e outro modelo deve ser utilizado (modelo não robusto).
 - 2 Os resultados desenvolvidos (EP, IC, TH etc) continuam válidos, ainda que com alguma perda e pode-se continuar, ainda que com ressalvas, com o modelo em questão (**modelo robusto**)

Introdução

- Uma revisão sobre resíduos para MRNLH (modelos de regressão normais lineares homocedásticos) e para MLG (modelos lineares generalizados) podem ser vistas [aqui](#) e [aqui](#), respectivamente.
- Por outro lado, um modelo de regressão (probabilístico) precisa ser capaz de reproduzir de forma adequada os dados observados (mecanismo gerador), bem como (em particular para ST) prever observações futuras (dado que o regime da série não mudará e/ou fatores não contemplados, não venham afetar seu comportamento).
- Primeiramente, falaremos sobre análise de diagnóstico (resíduos).

Análise de residual em ST

- Lembrando que o modelo AR(p) (Gaussiano) é dado por:

$$Y_t = \mu + \phi_1 (Y_{t-1} - \mu) + \cdots + \phi_p (Y_{t-p} - \mu) + \epsilon_t, \epsilon_t \sim NID(0, \sigma^2).$$

- Essencialmente para os modelos AR(p) (e para as outras classes de modelos de ST que veremos), temos de verificar se os erros (através de algum resíduos apropriado) são um ruído branco Gaussiano.
- Ou seja, temos que verificar se os resíduos são não correlacionados, homocedásticos e normalmente distribuídos.

Análise de residual em ST

- Basicamente, há dois tipos de resíduos (valor predito do erro) utilizados para os modelos que estudaremos: os resíduos ordinários e os padronizados.

- Resíduo ordinário: $\hat{\epsilon}_t^* = Y_t - \hat{Y}_t, t = 1, 2, \dots, n$

- Resíduo padronizado (RP): $\hat{\epsilon}_t = \frac{\hat{\epsilon}_t^*}{\sqrt{\hat{P}_t}} = Y_t - \hat{Y}_t, t = 1, 2, \dots, n$ em

que :

- $\hat{Y}_t = \hat{\mu} + \hat{\phi}_1 (Y_{t-1} - \hat{\mu}) + \dots + \hat{\phi}_p (Y_{t-p} - \hat{\mu}),$

- \hat{P}_t é o erro estimado associado à $\hat{\epsilon}_t^*$ e

- $(\hat{\mu}, \hat{\phi}_1, \dots, \hat{\phi}_p)'$ é o vetor de estimadores obtidos através de algum método (MM, MQC, MV, [aqui](#)).

Análise residual em ST

- Falaremos sobre \widehat{P}_t mais à frente (que está associado ao, assim chamado, erro de previsão).
- Com relação à \widehat{Y}_t temos que:
 - $\widehat{Y}_1 = \widehat{\mu}$.
 - $\widehat{Y}_2 = \widehat{\mu} + \widehat{\phi}_1 (Y_1 - \widehat{\mu})$
 - $\widehat{Y}_3 = \widehat{\mu} + \widehat{\phi}_1 (Y_1 - \widehat{\mu}) + \widehat{\phi}_2 (Y_2 - \widehat{\mu})$
 - \vdots
 - $\widehat{Y}_n = \widehat{\mu} + \widehat{\phi}_1 (Y_1 - \widehat{\mu}) + \widehat{\phi}_2 (Y_2 - \widehat{\mu}) + \dots + \widehat{\phi}_p (Y_{n-1} - \widehat{\mu})$
- Se as hipóteses forem satisfeitas, para uma dada ST, então, para n suficientemente grande:

$$\widehat{\epsilon}_t \approx NID(0, 1)$$

Análise residual e ST

- Para verificar a validade do resultado acima, podemos utilizar:
 - Gráfico do RP \times índice ([aqui](#)).
 - Gráfico do RP \times valores ajustados ([aqui](#)).
 - Histograma, Box-plot ([aqui](#)).
 - Gráfico de quantil-quantil com envelope ([aqui](#)). Neste caso, não é imprescindível construir um envelope simulado, pois a distribuição de referência do resíduo é conhecida.
 - Teste de ruído branco com gráficos ([aqui](#)).
 - Testes para estacionariedade (não vistos): Dickey-Fuller, Dickey-Fuller aumentado, Kwiatkowski-Phillips-Schmidt-Shin (KPSS), Phillips-Perron entre outros ([aqui](#)).

Previsão

- Voltemos ao modelo AR(p) estacionário e causal.
- O problema de previsão consiste em determinar um valor para Y_{n+k} com $k > 0$, dada a informação: $\mathcal{F}_n = \{Y_n, Y_{n-1}, \dots, Y_1\}$ Denotemos a respectiva previsão por $Y_n(k)$ onde k é o horizonte de previsão. A v.a. que representa o valor a ser observado é Y_{n+k}
- Previsões são calculadas de forma que alguma função de perda seja minimizada.

Previsão

- Neste curso adotaremos a minimização do Erro Quadrático médio de previsão.
- Assim, temos que minimizar $\mathcal{E}(\epsilon_n^2(k))$ onde $\epsilon_n(k) = Y_{n+k} - Y_n(k)$ é o erro de previsão para o horizonte k .
- Pode ser demonstrado que o preditor ótimo é a esperança condicional:
$$Y_n(k) = \mathcal{E}(Y_{n+k} | \mathcal{F}_n).$$

Previsão para o $AR(1)$

- Seja o processo $AR(1)$ estacionário e causal (com $\mu = 0$):

$$Y_t = \phi Y_{t-1} + \epsilon_t, \epsilon_t \sim RB(0, \sigma^2)$$

- Para $k=1$, como $Y_{n+1} = \phi Y_n + \epsilon_{n+1}$, então:

$$Y_n(1) = \mathcal{E}(Y_{n+1} | \mathcal{F}_n) = \phi \mathcal{E}(Y_n | \mathcal{F}_n) + \mathcal{E}(\epsilon_{n+1} | \mathcal{F}_n) = \phi Y_n$$

pois o processo é causal $\epsilon_{n+1} \perp \mathcal{F}_t$

Previsão para o $AR(1)$

- Para $k = 2$, como $Y_{n+2} = \phi Y_{n+1} + \epsilon_{n+2}$, então

$$Y_n(2) = \mathcal{E}(Y_{n+2} | \mathcal{F}_n) = \phi \mathcal{E}(Y_{n+1} | \mathcal{F}_n) + \mathcal{E}(\epsilon_{n+2} | \mathcal{F}_n) = \phi Y_n(1) = \phi^2 Y_n$$

- Por indução, para k geral, temos que:

$$Y_n(k) = \phi^k Y_n$$

Previsão para o $AR(1)$

- Variância dos erros de previsão. Seja

$$\sigma_n^2(k) = \mathcal{V}(\epsilon_n(k))$$

então

$$\begin{aligned}\mathcal{V}(\epsilon_n(1)) &= \mathcal{V}[(Y_{n+1} - Y_n(1))] = \mathcal{V}[\phi Y_n + \epsilon_{n+1} - \phi Y_n] \\ &= \mathcal{V}[\epsilon_{n+1}] = \sigma^2\end{aligned}$$

Previsão para o $AR(1)$

- Analogamente:

$$\sigma_n^2(2) = \mathcal{V}(\phi\epsilon_{n+1} + \epsilon_{n+2}) = \sigma^2(1 + \phi^2)$$

- Por indução, temos que:

$$\sigma_n^2(k) = \sigma^2(1 + \phi^2 + \dots + \phi^{2(k-1)}) = \sigma^2 \frac{1 - \phi^{2k}}{1 - \phi^2}$$

- A longo prazo ($k \rightarrow \infty$), temos que:

$$Y_n(k) \rightarrow 0, \sigma_n^2(k) \rightarrow \sigma^2 \frac{1}{1 - \phi^2}$$

Previsão para o $AR(1)$

- Para os processos $AR(1)$ com média diferente de zero ou com drift, primeiro defina o processo: $X_t = Y_t - \mu$. Depois, obtenha as previsões $X_n(k)$ e, posteriormente, calcule:

$$Y_n(k) = X_n(k) + \mu$$

(substituindo μ por $\hat{\mu}$ se aquele for desconhecido)

- Outra forma é aplicar diretamente a esperança condicional no modelo definido com drift (ou $\mu \neq 0$) (exercício).

Previsão para o $AR(1)$

- Supondo que o processo é Gaussiano ($\epsilon_t \sim NID(0, \sigma^2)$), intervalos de previsão de $\gamma\%$ de confiança para Y_{n+k} , são da forma:

$$IP(Y_{n+k}) = Y_n(k) \pm z_{\frac{1+\gamma}{2}} \sqrt{\sigma^2(k)}$$

em que $P\left(Z \leq z_{\frac{1+\gamma}{2}}\right)$, $Z \sim N(0, 1)$.

- Se os parâmetros $(\mu, \phi', \sigma^2)'$ forem conhecidos, o IP acima é exato. Se pelo menos um não for, mesmo sob a validade das hipóteses (inclusive a normalidade) eles são assintóticos.
- À medida que as suposições deixarem de ser válidas, os resultados apresentados também o deixam.

Previsão para o $AR(2)$

- Seja o processo $AR(2)$ estacionário e causal (com $\mu = 0$):

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t, \epsilon_t \sim RB(0, \sigma^2)$$

- Para $k=1$, como $Y_{n+1} = \phi_1 Y_n + \phi_2 Y_{n-1} + \epsilon_{n+1}$, então

$$\begin{aligned} Y_n(1) &= \mathcal{E}(Y_{n+1} | \mathcal{F}_n) = \phi_1 \mathcal{E}(Y_n | \mathcal{F}_n) + \phi_2 \mathcal{E}(Y_{n-1} | \mathcal{F}_n) + \mathcal{E}(\epsilon_{n+1} | \mathcal{F}_n) \\ &= \phi_1 Y_n + \phi_2 Y_{n-1} \end{aligned}$$

pois o processo é causal ($\epsilon_{n+1} \perp \mathcal{F}_t$).

Previsão para o $AR(2)$

- Para $k=2$, como $Y_{n+2} = \phi_1 Y_{n+1} + \phi_2 Y_n + \epsilon_{n+2}$, então

$$\begin{aligned} Y_n(2) &= \mathcal{E}(Y_{n+2}|\mathcal{F}_n) = \phi_1 \mathcal{E}(Y_{n+1}|\mathcal{F}_n) + \phi_2 \mathcal{E}(Y_n|\mathcal{F}_n) + \mathcal{E}(\epsilon_{n+2}|\mathcal{F}_n) \\ &= \phi_1 Y_n(1) + \phi_2 Y_n \end{aligned}$$

- Para $k=3$, como $Y_{n+3} = \phi_1 Y_{n+2} + \phi_2 Y_{n+1} + \epsilon_{n+3}$, então

$$\begin{aligned} Y_n(3) &= \mathcal{E}(Y_{n+3}|\mathcal{F}_n) = \phi_1 \mathcal{E}(Y_{n+2}|\mathcal{F}_n) + \phi_2 \mathcal{E}(Y_{n+1}|\mathcal{F}_n) + \mathcal{E}(\epsilon_{n+3}|\mathcal{F}_n) \\ &= \phi_1 Y_n(2) + \phi_2 Y_n(1) \end{aligned}$$

Previsão para o $AR(2)$

- Para o caso geral temos que resolver a equação em diferenças de segunda ordem

$$Y_n(k) = \phi_1 Y_n(k-1) + \phi_2 Y_n(k-2) \quad (1)$$

$$(1 - B - B^2) Y_n(k) = 0$$

- Dadas estimativas dos parâmetros, a Equação (1) pode ser resolvida de forma recursiva, $k=1,2,\dots$

Previsão para o $AR(2)$

- Variâncias dos erros de previsão. Exercício: calcule $\sigma_n^2(k) = \mathcal{V}(\epsilon_n(k))$, $k = 1, 2$.
- Para os processos $AR(2)$ com média diferente de zero ou drift, valem as observações apresentadas para os modelos $AR(1)$. Idem para os intervalos de previsão.

Previsão para o $AR(p)$

- Seja o processo $AR(p)$ estacionário e causal (com $\mu = 0$):

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t, \epsilon_t \sim RB(0, \sigma^2)$$

- Pode-se provar que:

$$Y_t(k) = \phi_1 Y_n(k-1) + \dots + \phi_p Y_n(k-p)$$

em que $Y_n(s) = Y_{n-s}, s \leq 0$

Previsão para o $AR(p)$

- Para os processos $AR(p)$ com média diferente de zero ou drift, valem as observações apresentadas para os modelos $AR(1)$. Idem para os intervalos de previsão.
- Expressões para $\sigma_n^2(k)$ serão discutidas na parte de Processos ARMA.

Pacote astsa

- astsa: acrônimo para “Applied Statistical Time Series Analysis”
- apresenta diversas funções para: análise exploratória, estimação, testes de hipótese, critérios de informação, diagnóstico e previsão para diversas classes de modelos de ST (e outras).
- Encapsula algumas funções existentes no R como a função “sarima”.
- Links de interesse se encontram na página do curso na seção [Links/Material adicional](#).

Pacote astsa

- Por enquanto utilizaremos:
 - `sarima(xdata,p,d,q,...)` (deixaremos o restante dos argumentos como default)
 - `xdata`: ST de interesse
 - `p,d,q`: respectivamente ordem do processo AR, grau de diferenciação (processos ARIMA), ordem do processo MA (média movel). Por enquanto, $d=0,q=0$
 - Essencialmente utiliza estimativas de MQC (não iterativas), como valores iniciais em um processo iterativo para obter as estimativas de MQC. De posse destas, performa outro processo iterativo a fim de obter as estimativas de MV.
 - Tem implementado os mecanismos inferenciais vistos até o momento (e outros).

Pacote astsa

- Por enquanto utilizaremos:
 - `sarima.for(xdata,n.ahead,p,d,q,...)` (deixaremos o restante dos argumentos como default)
 - Todos os argumentos são como para a função `sarima`, exceto `n.ahead` que corresponde à janela de previsão.
 - Fornece as previsões com barras de erro correspondentes à ± 1 e 2 erros de previsão.

Aplicação 1: dados sobre o hormônio luteinizante

- Disponível no pacote R (default) sob o nome “lh”.
- Uma série temporal regular relativa à concentração (unidade de medida não especificada) do hormônio luteinizante em amostras de sangue, medidas em intervalos de 10 minutos, de uma mulher , $n = 48$ (amostras).
- Veja também [aqui](#).
- Estudar o comportamento da ST, incluso fazendo previsões.

Análise exploratória

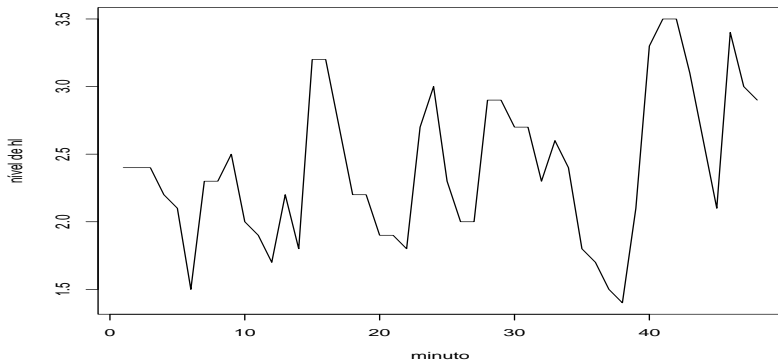


Figura: Gráfico de ST sobre os dados de hl.

Análise exploratória

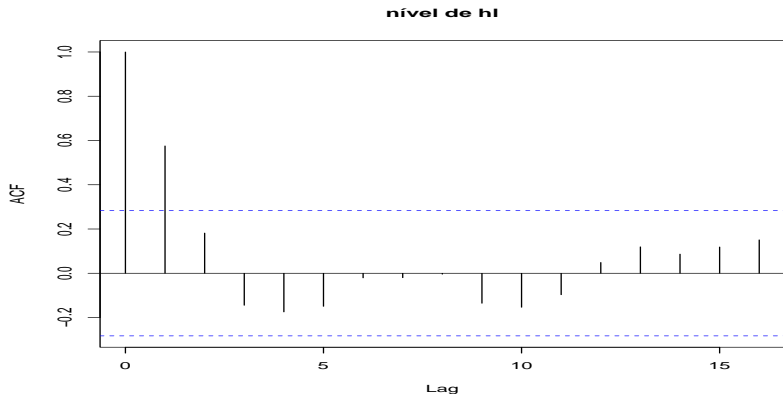


Figura: FAC para a ST sobre o hl.

Análise exploratória

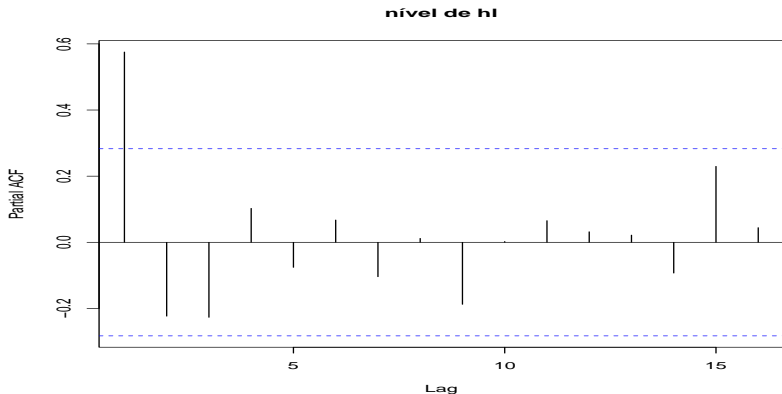


Figura: FACP para a ST sobre o hl.

Análise exploratória

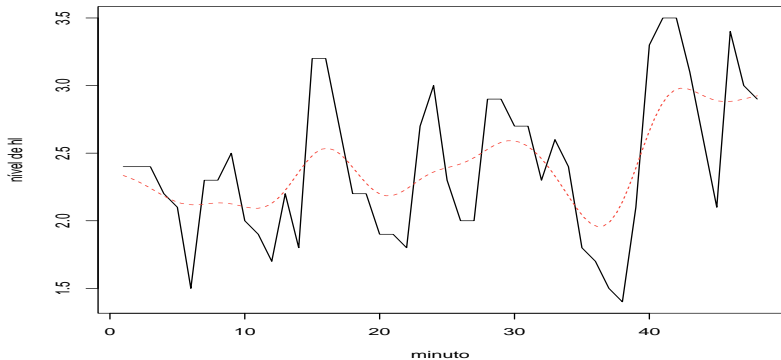
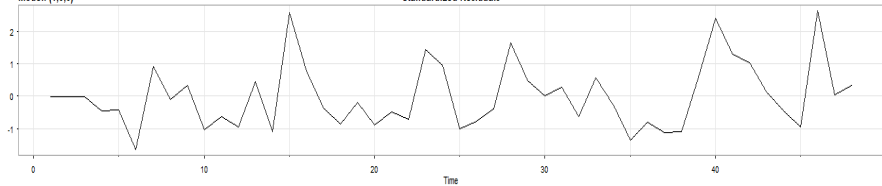


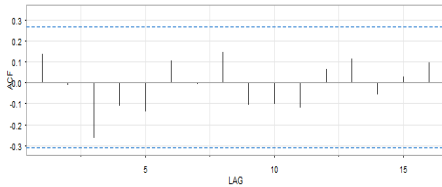
Figura: Suavização por kernel ($b=6$) para a ST sobre o hl.

Model: (1,0,0)

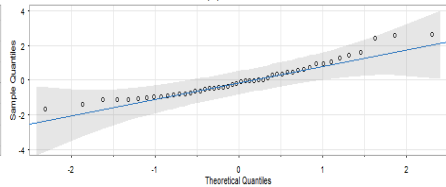
Standardized Residuals



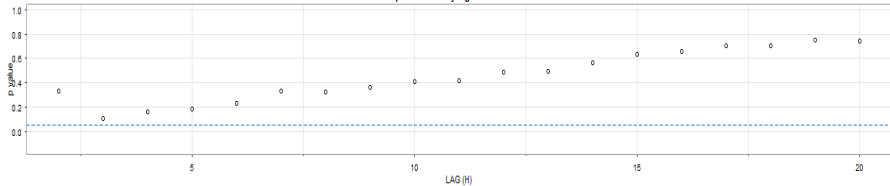
ACF of Residuals



Normal Q-Q Plot of Std Residuals



p values for Ljung-Box statistic



Comentários (RP)

- Auto-correlações, essencialmente, não significativas.
- Aparentemente não há tendência nem indicação de heterocedasticidade (talvez, de forma leve).
- Os resíduos parecem se aproximar de um ruído branco.
- Contudo, os resíduos apresentam assimetria positiva.
- O modelo não parece estar bem ajustado.
- Sugestão: apenas modificar a distribuição condicional ($Y_t|y_{t-1}$) (distribuição positiva).

Séries temporais observada e predita

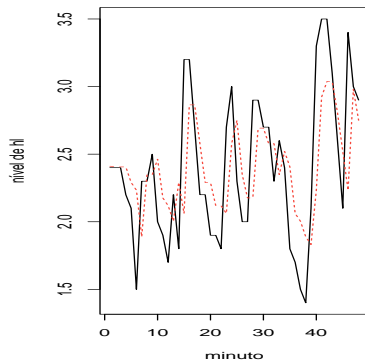
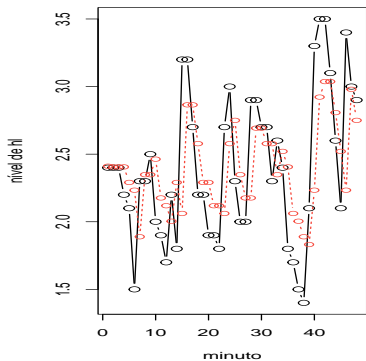
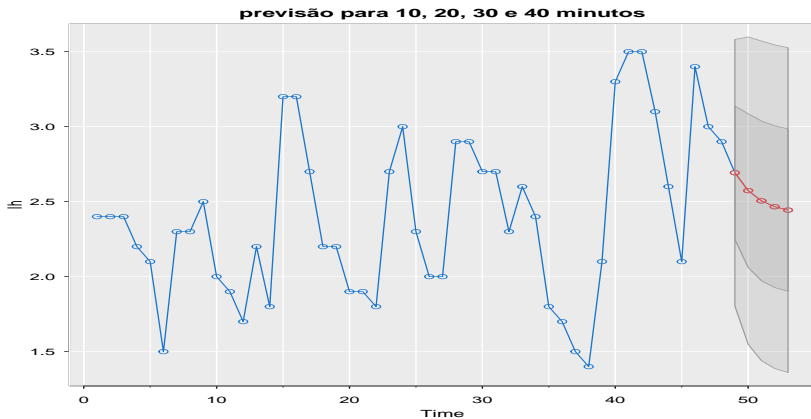


Figura: Preto - observado, vermelho - predito.

Previsão para uma janela de 50 minutos



Estimativas

Parâmetro	Est.	EP	IC(95%)	Est. t	p-valor
ϕ	0,57	0,12	[0,35 ; 0,80]	4,94	< 0,0001
μ	2,41	0,15	[2,13 ; 2,70]	16,46	< 0,0001
σ^2	0,20	-	-	-	-

A correlação entre as observações (ϕ) mostrou-se significativa e positiva.

ACF's observada e predita

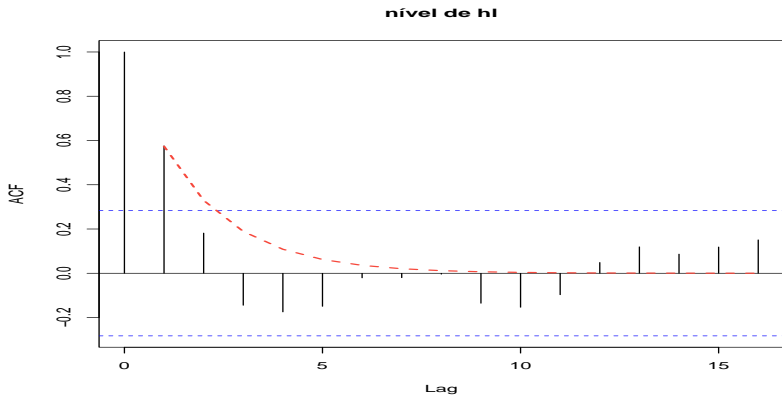


Figura: Preto - observado, vermelho - predito.

Comparação com o modelo AR2

■ Estimativas (AR(2))

Parâmetro	Est.	EP	IC(95%)	Est. t	p-valor
ϕ_1	0,70	0,14	[0,42;0,97]	5,00	< 0,0001
ϕ_2	-0,21	0,14	[-0,49;0,06]	-1,52	0,13439
μ	2,40	0,12	[2,17;2,64]	20,03	< 0,0001
σ^2	0,19	-	-	-	-

■ Critérios de informação

Modelo	AIC	AICc	BIC
AR1	1,349	1,355	1,466
AR2	1,344	1,355	1,500

Comentários

- A predição dos valores observados pode melhorar em se utilizando um modelo mais apropriado.
- É possível melhorar o ajuste usando um modelo da classe ARMA, utilizando uma distribuição positiva.

Aplicação 2: dados sobre índice de emprego no Canadá

- Índice de emprego trimestral (IET) no Canadá (1960 a 1993).
- Os dados foram **ajustados sazonalmente** (retirou-se a sazonalidade).
- Os dados não apresentam tendência (aparentemente) mas continuam a apresentar autocorrelação.

Análise exploratória



Figura: Gráfico de ST sobre os dados de IET.

Análise exploratória

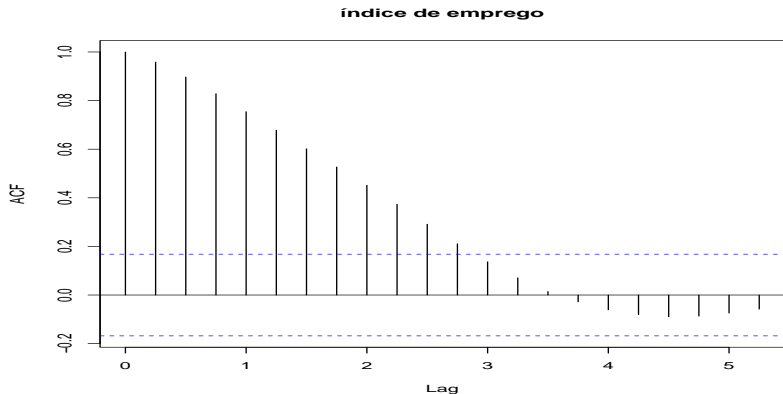


Figura: FAC para a ST sobre os dados de IET.

Análise exploratória

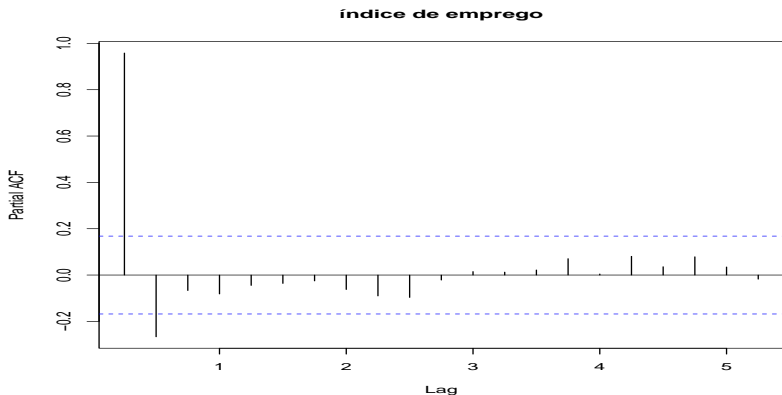


Figura: FACP para a ST sobre os dados de IET.

Análise exploratória

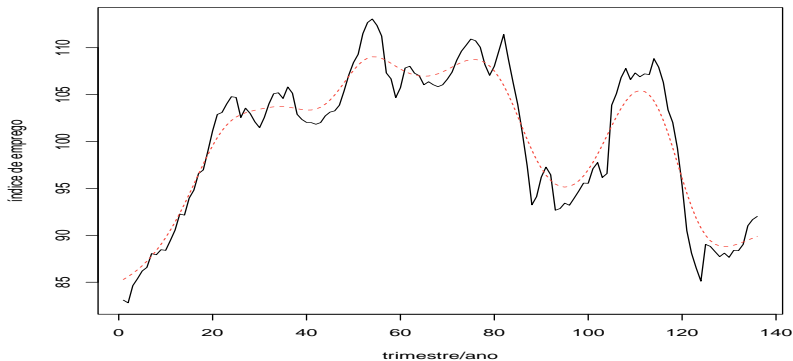
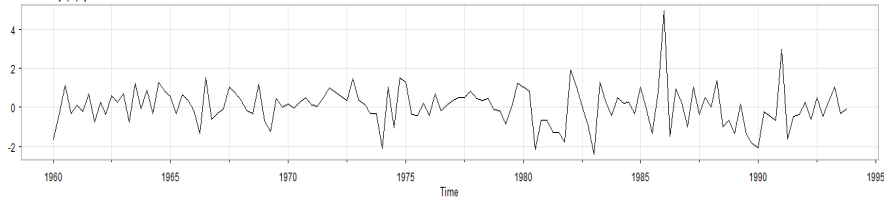


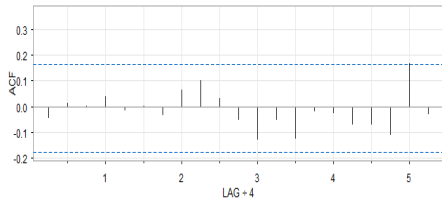
Figura: Suavização por kernel ($b=12$) para a ST sobre o IET.

Model: (2,0,0)

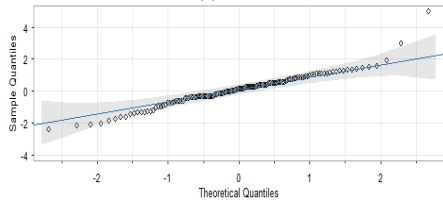
Standardized Residuals



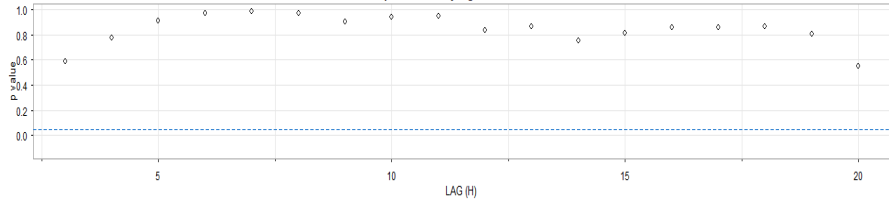
ACF of Residuals



Normal Q-Q Plot of Std Residuals



p values for Ljung-Box statistic



Comentários (RP)

- Auto-correlações, essencialmente, não significativas.
- Aparentemente não há tendência nem indicação de heterocedasticidade (talvez, de forma leve).
- Os resíduos parecem se aproximar de um ruído branco.
- Contudo, os resíduos apresentam assimetria positiva e/ou caudas pesadas.
- O modelo não parece estar bem ajustado.
- Sugestão: apenas modificar a distribuição condicional ($Y_t|y_{t-1}$) (distribuição positiva e/ou assimétrica com caudas pesadas).

Séries temporais observada e predita

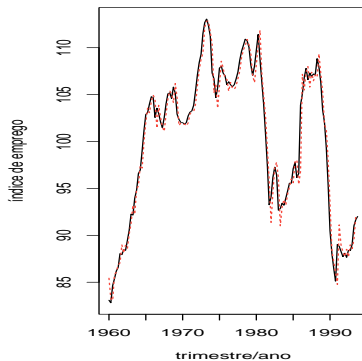
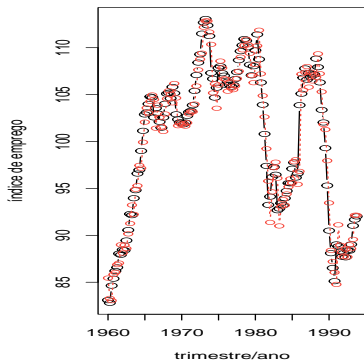
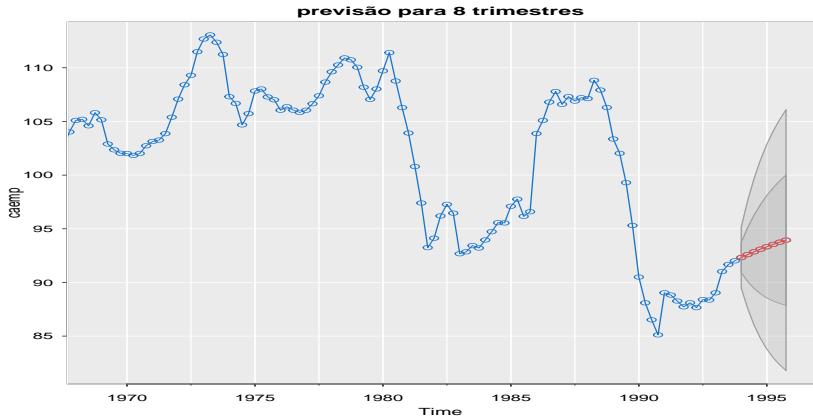


Figura: Preto - observado, vermelho - predito.

Previsão para uma janela de 8 meses (dois anos)



Estimativas

Parâmetro	Est.	EP	IC(95%)	Est. t	p-valor
ϕ_1	1,45	0,07	[1,30;1,60]	19,36	< 0,0001
ϕ_2	-0,48	0,08	[-0,63;-0,33]	-6,25	< 0,0001
μ	97,50	4,39	[88,89;106,11]	22,19	< 0,0001
σ^2	2,02	-	-	-	-

A correlação entre as observações (ϕ_1, ϕ_2) mostrou-se significativa e positiva.

Comparação com os modelos AR1 e AR3

■ Estimativas

Parâmetro	Est.	EP	IC(95%)	Est. t	p-valor
ϕ	0,99	0,01	0,96 ; 1,01	79,68	< 0,0001
μ	93,93	7,19	79,84 ; 108,03	13,06	< 0,0001
ϕ_1	1,42	0,09	1,25 ; 1,59	16,67	< 0,0001
ϕ_2	-0,38	0,15	-0,66 ; -0,09	-2,59	0,0106
ϕ_3	-0,07	0,09	-0,24 ; 0,10	-0,79	0,4309
μ	97,81	4,10	89,78 ; 105,85	23,86	< 0,0001

Comparação com os modelos AR1 e AR3

■ Critérios de informação

Modelo	AIC	AICc	BIC
AR1	3,867	3,868	3,931
AR2	3,629	3,631	3,715
AR3	3,639	3,642	3,746

Comentários

- É possível melhorar o ajuste usando um modelo da classe ARMA, utilizando uma distribuição positiva e/ou assimétrica de caudas pesadas.