Filtro de Kalman

Pedro Valls[1]
CEQEF e EESP-FGV

[1]CEQEF e EESP-FGV, Rua Dr Plinio Barreto 365 s.1319, 01313-020, , São Paulo, SP, Brasil. E-mail:pedro.valls@fgv.br.

Sumário

Lista de Tabelas

Lista de Figuras

1 Introdução

A idéia de escrever modelos na representação em espaço de estado está relacionada ao seguintes fatos:

(i) modelos, em geral, relacionam variáveis observadas à variávies não observadas. Por exemplo o modelo de regressão relaciona uma variável observada a variáveis não observadas, parâmetros.

$$\underbrace{Y_t}_{\text{var.obs}} = \underbrace{X_t}_{\text{var. obs}} \underbrace{\beta_t}_{\text{parâmetro}} + \varepsilon_t$$

(ii) modelos podem ter parâmetros que variam no tempo. Por exemplo no modelo de regressão as variáveis não observadas, parâmetros, não precisam ser invariantes no tempo.

$$\begin{aligned}
\underbrace{Y_t}_{\text{var.obs}} &= \underbrace{X_t}_{\text{var. obs}} \underbrace{\beta_t}_{\text{parâmetro}} + \varepsilon_t \\
\beta_t &= \beta_{t-1} + \eta_t
\end{aligned}$$

A representação em espaço de estado nos permite obter um mecanismo recursivo para obter o erro de previsão um passo à frente e a variância deste erro de previsão. Desta forma, é possível obter a função de verossimilhança do modelo e através de métodos numéricos é possível maximizar esta função e obter os estimadores de máxima verossimilhança (ou quase-máxima verossimilhança) para os parâmetros. Como subproduto os erros de previsão podem ser usados para se fazer diagnóstico da adequabilidade do modelo.

A seguir vamos apresentar três motivações para a representação em espaço de estado. A primeira está relacionada, na literatura econômica, ao modelo de expectativas adaptativas que na literatura estatística está relacionada ao modelo de suavizamento exponencial (veja Brown (1959)). O segundo exemplo está relacionada ao modelo de expectativas adaptativas onde o coeficiente de ajuste varia no tempo. O último está relacionado a extração da variância condicional de uma série que é conhecido em finanças como o modelo de volatilidade estocástica(veja Ghysels et al. (1996)).

$$\underbrace{Y_t}_{\text{var observada}} = \underbrace{X_t\beta}_{\text{var. obs*parametro}} + \varepsilon_t$$

1.1 Modelo de Expectativa Adaptativa com Coeficiente de Ajuste Constante.

Suponha que y_t representa o log da renda e que este processo pode ser modelado por um $ARIMA(0,1,1)$ $(IMA(1,1))$, isto é:

$$y_t = y_{t-1} + \varepsilon_t - (1-\theta)\varepsilon_{t-1} \tag{1}$$

Seja $\mathbf{I}_{t-1} = \{y_{t-j}, \varepsilon_{t-j},\ j \geq 1\}$ o conjunto de informação disponível no instante $t-1$ e defina $\tilde{y}_{t|t-1} = E[y_t \mid \mathbf{I}_{t-1}]$, isto é o previsor ótimo da renda usando o conjunto de informação $\mathbf{I}_{t-1}$. Então aplicando o operador $E[\bullet \mid \mathbf{I}_{t-1}]$ em (1) obtém-se:

$$\tilde{y}_{t|t-1} = y_{t-1} - (1-\theta)\varepsilon_{t-1} \tag{2}$$

mas por (1) temos:

$$\varepsilon_t = \frac{\Delta y_t}{1-(1-\theta)L} \tag{3}$$

que substituindo em (2) temos:

$$\tilde{y}_{t|t-1} = y_{t-1} - (1-\theta)\frac{\Delta y_{t-1}}{1-(1-\theta)L} \tag{4}$$

$$\Updownarrow$$

$$\tilde{y}_{t|t-1} - (1-\theta)\tilde{y}_{t-1|t-2} = y_{t-1} - (1-\theta)y_{t-2} - (1-\theta)\Delta y_{t-1}$$

$$\Updownarrow$$

$$\tilde{y}_{t|t-1} - \tilde{y}_{t-1|t-2} = \theta(y_{t-1} - \tilde{y}_{t-1|t-2}) \tag{5}$$

que é o modelo de expectativas adaptativas com coeficiente de ajuste θ que é fixo e está relacionado ao coeficiente da parte MA do processo.

Observe que a otimalidade neste caso faz com que as expectativas, em média, estejam corretas, isto é são ótima no sentido de minimizar o erro quadrático médio de previsão.

Observe que (5) pode ser reescrito da seguinte forma:

$$\tilde{y}_{t|t-1} = (1-\theta)\tilde{y}_{t-1|t-2} + \theta y_{t-1} \tag{6}$$

e esta formulação está relacionada ao modelo de suavizamento exponencial que é um dos modelos ingênuos para estimar o nível de uma série como iremos mostrar a seguir.

Suponha que o estatístico deseja estimar o nível de uma série dando peso maior as observações mais recentes. Para fazer isto ele define μ_T , o nível da série no instante T por:

$$\mu_T = \sum_{j=0}^{T-1} \omega_j y_{T-j} \text{ com } \sum_{j=0}^{T-1} \omega_j = 1 \tag{7}$$

então o previsor das observações futuras, usando a informação até o instante T será dado por:

$$\tilde{y}_{T+l|T} = \mu_T \text{ para } l \geq 1 \tag{8}$$

Agora suponha que o pesos decaem exponencialmente, isto é:

$$\omega_j = \theta(1-\theta)^j \text{ com } 0 < \theta \leq 1 \tag{9}$$

onde θ é chamado de coeficiente de suavizamento. Para T grande, a condição $\sum_{j=0}^{T-1} \omega_j = 1$ será satisfeita uma vez que:

$$\lim_{T \to \infty} = \sum_{j=0}^{T-1} \theta(1-\theta)^j = \theta \sum_{j=0}^{\infty} (1-\theta)^j = \theta \frac{1}{1-(1-\theta)} = 1$$

Agora substituindo os pesos (9) em (7) temos:

$$\mu_T = \theta \sum_{j=0}^{T-1} (1-\theta)^j y_{T-j} \tag{10}$$

e fazendo $T = t$ em (10) temos:

$$\mu_t = \theta \sum_{j=0}^{t-1} (1-\theta)^j y_{t-j} = (1-\theta)\mu_{t-1} + \theta y_t \text{ para } t = 1, \cdots, T \tag{11}$$

assumindo que $\mu_0 = 0$.

Mas em (8) fazendo $T = t$ e $l = 1$ temos que $\mu_t = \tilde{y}_{t+1|t}$ e fazendo $T = t-1$ e $l = 1$ temos $\mu_{t-1} = \tilde{y}_{t||t-1}$ e então (11) pode ser reescrito da seguinte forma:

$$\tilde{y}_{t+1|t} = (1-\theta)\tilde{y}_{t|t-1} + \theta y_t$$

que é o **modelo de suavizamento simple exponencial de Brown**.

1.2 Modelo de Expectativa Adaptativa com Coeficiente de Ajuste Variando no Tempo.

Suponha que y_t , o logaritmo da renda, é composto de dois componentes, π_t o componente permanente que é não observado e s_t o componente inesperado com média zero.

Suponha que os agentes tem uma estimativa a priori para π_0 e querem atualizar esta estimativa usando a informação que se torna disponível para a renda. É necessário descrever de que forma o componente permanente evolui. Assumiremos que esta evolução se dá de forma suave, por exemplo, este componente segue uma passeio aleatório. Por outro lado, os agentes percebem que uma fração k_t de s_t é incorporada ao componente permanente da renda e a fração $(1-k_t)s_t$ é incorporada ao componente transitório. Tem-se então:

$$\begin{aligned} y_t &= \pi_t + (1-k_t)s_t \\ \pi_t &= \pi_{t-1} + k_t s_t \end{aligned}$$

com $E[s_t \mid \mathbf{I}_{t-1}] = 0$, $E[\pi_t s_{t-j}] = 0$ para $\forall j \geq 0$, isto é, s_t é um processo de inovação não correlacionado com o componente permanente. Agora a forma reduzida do modelo acima é dada por:

$$\begin{aligned} \Delta y_t &= k_t s_t + \Delta[(1-k_t)s_t] \\ \Delta y_t &= k_t s_t + (1-k_t)s_t + (1-k_{t-1})s_{t-1} \\ \Delta y_t &= s_t + (1-k_{t-1})s_{t-1} \end{aligned} \tag{12}$$

onde (12) é um IMA(1,1) com coeficiente variando no tempo.

1.3 Extração da Variância Condicional de uma série

Seja p_t o preço de um ativo. Define-se o retorno composto deste ativo por:

$$r_t = \Delta \ln(p_t)$$

e vamos assumir que a distribuição destes retornos, condicional a informação passada, é normal com média zero e variância condicional dada por σ_t^2 . Além disto vamos assumir que o logaritmo desta variância pode ser descrito por um processo $AR(1)$ com média não nula δ . A racionalidade para este modelo vem do fato que este processo reverte para a média. Tem-se então o seguinte modelo para descrever os retornos:

$$r_t = \varepsilon_t \sigma_t \quad \text{com} \quad \varepsilon_t \sim NI(0,1) \tag{13}$$

$$\ln \sigma_t^2 = \delta + \phi \ln \sigma_{t-1}^2 + \sigma_\eta \eta_t \quad \text{com} \quad \eta_t \sim NI(0,1) \tag{14}$$

Agora tomando a transformação $\ln(\cdot)^2$ em (13), o sistema fica:

$$\ln r_t^2 = \ln \sigma_t^2 + \ln \varepsilon_t^2 \tag{15}$$

$$\ln \sigma_t^2 = \delta + \phi \ln \sigma_{t-1}^2 + \sigma_\eta \eta_t \tag{16}$$

onde (15) relaciona as observações, o logarítmo do quadrado dos retornos, a um não observado o logarítmo da variância condicional e (16) dá de que forma esta variância evolui ao longo do tempo. Este modelo é conhecido na literatura de finanças por modelo de volatilidade estocástica.

2 Representação em Espaço de Estado Linear[1]

Em todos os exemplos acima o processo era representado por duas equações. A primeira, chamada de **equação de observação** ou **de medida**, relacionava as observações a um não observado e a segunda, chamada de **equação de transição,** que dava a evolução deste não observado. Estas duas equações são chamadas de **representação em espaço de estado** do modelo.

A formulação mais geral é a seguinte: seja y_t um vetor $N \times 1$ de observações e α_t um vetor $m \times 1$ de variáveis não observadas chamado de **vetor de estado**. A equação de observação é dada por:

[1] Esta formulação é baseada em Durbin and Koopman (2012), que difere da formulação original de Harvey (1993), que será usada em alguns dos exemplos.

$$\underset{(Nx1)}{y_t} = \underset{(Nxm)}{Z_t} \underset{(mx1)}{\alpha_t} + \underset{(Nx1)}{c_t} + \underset{(NxN)}{G_t} \underset{(Nx1)}{\varepsilon_t} \quad \text{para } t = 1, \cdots, T \tag{17}$$

onde c_t contém variáveis determinísticas observadas que podem afetar as observações e ε_t tem média zero, é serialmente não correlacionado e tem matriz de variância-covariância $\mathbf{I}$.

Como α_t em geral é não observado vamos assumir que este pode ser gerado por um processo Markoviano de primeira ordem, isto é:

$$\underset{(mx1)}{\alpha_t} = \underset{(mxm)}{T_{t-1}} \underset{(mx1)}{\alpha_{t-1}} + \underset{(mx1)}{d_{t-1}} + \underset{(mxg)}{H_t} \underset{(gx1)}{\eta_t} \quad \text{para } t = 1, ..., T \tag{18}$$

onde d_{t-1} são variáveis observadas que podem afetar o vetor de estado e η_t tem média zero, é serialmente não correlacionado e tem matriz de variância-covariância dada por $\mathbf{I}$.

As equações (17-18) são a representação em espaço de estado linear para o processo y_t .

Algumas hipóteses adicionais são necessária para completar a especificação desta representação, a saber:

(i) $E[\alpha_0] = a_0$, $Var[\alpha_0] = P_0$

(ii) $E[\varepsilon_t \eta\prime_s] = 0 \ \forall t, s$

(iii) $E[\varepsilon_t \alpha\prime_0] = E[\eta_t \alpha\prime_0] = 0 \ \forall t$.

2.1 Exemplos

2.1.1 AR(1) + Ruído

O modelo AR(1) + Ruído é equivalente ao Modelo de Volatilidade estocástica com $\gamma = 0$.

$y_t = \mu_t + \sigma_\xi \xi_t$ com $\xi_t \sim NI(0,1)$

$\mu_t = \phi \mu_{t-1} + \sigma_\eta \eta_t$ com $\eta_t \sim NI(0,1)$

que corresponde ao modelo descrito na seção 1.3 se:

$y_t = \ln r_t^2$

$\mu_t = \ln \sigma_t^2$

$\xi_t = \ln \varepsilon_t^2$

$\sigma_\xi = 1$

e Z_t, T_t, H_t e G_t são invariantes no tempo.

2.1.2 Modelo de Volatilidade Estocástica com Efeitos Calendários

É provável que a volatilidade seja afetada por dia-da-semana porque existem três noites entre sexta-feira e segunda-feira e devem acarretar um aumento da volatilidade assim como no mercado brasileiro as quintas-feiras estavam associadas ao dia do boato implicando também numa volatilidade diferente também para este dia. Devido a estas possíveis regularidades o modelo apresentado na seção 1.3 não é de todo adequado dado que não contempla estes efeitos calendários.

Uma forma de incorporar tais efeitos é permitir que a volatilidade seja modelada por:

$$h_t^* = h_t + \beta' x_t$$

onde x_t são variáveis dummies semanais, isto é,

$$\begin{aligned}
x_t &= (x_{1t}, x_{2t}, ..., x_{5t}) \\
x_{1t} &= \begin{cases} 1 \text{ se } t = \text{segunda} \\ 0 \text{ caso contrário} \end{cases} \\
x_{2t} &= \begin{cases} 1 \text{ se } t = \text{terça} \\ 0 \text{ caso contrário} \end{cases} \\
& \,\,\vdots \\
x_{5t} &= \begin{cases} 1 \text{ se } t = \text{sexta} \\ 0 \text{ caso contrário} \end{cases}
\end{aligned}$$

Temos então o seguinte modelo:

$$y_t = \xi_t \exp(\frac{h_t^*}{2}) \tag{19}$$

$$h_t = \phi h_{t-1} + \sigma_\eta \eta_t \tag{20}$$

e fazendo a transformação $\log(\cdot)^2$ em (19) tem-se:

$$\log(y_t^2) = h_t + \beta' x_t + \ln(\xi_t^2)$$

$$h_t = \phi h_{t-1} + \sigma_\eta \eta_t$$

que é a representação em espaço de estado linear para este modelo e observe que na equação de observação aparece $c_t = \beta' x_t$ que tenta captar o efeito calendário na volatilidade, dado que esta é modelada agora por h_t^* .

2.1.3 Modelo $ARMA(p, q)$

Considere o modelo $ARMA(p, q)$, isto é:

$$y_t = \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \sigma_\varepsilon \varepsilon_t + \theta_1 \sigma_\varepsilon \varepsilon_{t-1} + ... + \theta_q \sigma_\varepsilon \varepsilon_{t-q} \qquad (21)$$

com $\varepsilon_t \sim NI(0, 1)$

Seja $m = max(p, q + 1)$ então (21) pode ser reescrito da seguinte forma:

$$y_t = \phi_1 y_{t-1} + ... + \phi_m y_{t-m} + \sigma_\varepsilon \varepsilon_t + \theta_1 \sigma_\varepsilon \varepsilon_{t-1} + ... + \theta_{m-1} \sigma_\varepsilon \varepsilon_{t-m+1} \qquad (22)$$

com alguns dos $\phi' s$ ou $\theta' s$ iguais a zero.

Defina então o vetor $\alpha_t' = [y_t, l_t^{(1)}, ..., l_t^{(m-1)}]'$ e $Z_t = (1; \mathbf{0}_{m-1})$ temos então que a equação de observação é dada por:

$$y_t = Z_t \ \alpha_t \Leftrightarrow y_t \equiv y_t \qquad (23)$$

e a equação de transição será dada por:

$$\alpha_t = \begin{bmatrix} \phi_1 & 1 & 0 & . & . & . & 0 & 0 \\ \phi_2 & 0 & 1 & . & . & . & 0 & 0 \\ \phi_3 & 0 & 0 & . & . & . & 0 & 0 \\ . & & . & . & . & & & . \\ . & & & . & . & . & & . \\ . & & & & . & . & . & . \\ . & & & & & . & . & . & 1 \\ \phi_m & 0 & . & . & . & 0 & 0 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ \sigma_\varepsilon \theta_1 \\ \sigma_\varepsilon \theta_2 \\ . \\ . \\ . \\ . \\ \sigma_\varepsilon \theta_{m-1} \end{bmatrix} \varepsilon_t \qquad (24)$$

Observe que a função das variáveis $l_t^{(i)}$ para $i = 1, ..., m-1$ é recuperar as defasagens de y_t e ε_t.

Shumway and Stoffer definem a representação em espaço de estado de um modelo $ARMA(p,q)$ com os dois erros da equação de medida e de transição perfeitamente correlacionados. A equação é definida para α_{t+1} em vez de α_t. É assumido que $p \geq q$, e o vetor de estado $\alpha'_{t+1} = [y_{t+1}, y_t, \cdots, y_{t-(p-1)}]'$ e a equação de transição é dada por:

$$\alpha_{t+1} = \begin{bmatrix} \phi_1 & 1 & 0 & . & . & . & 0 & 0 \\ \phi_2 & 0 & 1 & . & . & . & 0 & 0 \\ \phi_3 & 0 & 0 & . & . & . & 0 & 0 \\ . & . & . & . & & & & . \\ . & & . & . & . & & & . \\ . & & & . & . & . & & . \\ . & & & & . & . & . & 1 \\ \phi_m & 0 & . & . & . & 0 & 0 & 0 \end{bmatrix} \alpha_t + \begin{bmatrix} \sigma_\varepsilon(\theta_1 + \phi_1) \\ . \\ . \\ . \\ \sigma_\varepsilon(\theta_q + \phi_q) \\ \sigma_\varepsilon \phi_{q+1} \\ . \\ . \\ . \\ \sigma_\varepsilon \phi_p \end{bmatrix} \varepsilon_t \tag{25}$$

a equação de medida é dada por:

$$y_t = Z_t \ \alpha_t + \sigma_\varepsilon \varepsilon_t \tag{26}$$

A respresentação dada por (23) e (24) será usada nos casos particulares abaixo e a representação dad por (26) e (25) serão usadas na seção (3.1.1) para estimar modelos por Filtro de Kalman.

2.1.4 Casos Particulares

(1) $AR(1)$

$$y_t = (1) \ \alpha_t$$

$$\alpha_t = \phi_1 \alpha_{t-1} + \sigma_\varepsilon \varepsilon_t$$

(2) $AR(2)$

$$y_t = (1, 0) \ \alpha_t$$

$$\alpha_t = \begin{bmatrix} y_t \\ l_t^{(1)} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ 0 \end{bmatrix} \varepsilon_t \tag{27}$$

observe que (27) nós dá:

$$y_t = \phi_1 y_{t-1} + l_{t-1}^{(1)} + \sigma_\varepsilon \varepsilon_t \tag{28}$$

$$l_t^{(1)} = \phi_2 y_{t-1} \tag{29}$$

e substituindo (29) em (28) temos

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \sigma_\varepsilon \varepsilon_t$$

(3) $MA(1)$

$$y_t = (1, 0)\ \alpha_t \tag{30}$$

$$\alpha_t = \begin{bmatrix} y_t \\ l_t^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ \sigma_\varepsilon \theta_1 \end{bmatrix} \varepsilon_t \tag{31}$$

observe que (31) implica em:

$$y_t = l_{t-1}^{(1)} + \sigma_\varepsilon \varepsilon_t \tag{32}$$

$$l_t^{(1)} = \theta_1 \sigma_\varepsilon \varepsilon_t \tag{33}$$

e substituindo (33) em (32) temos:

$$y_t = \theta_1 \sigma_\varepsilon \varepsilon_{t-1} + \sigma_\varepsilon \varepsilon_t = \sigma_\varepsilon \varepsilon_t + \theta_1 \sigma_\varepsilon \varepsilon_{t-1}$$

(4) Modelo de Nível Local[2]

$$y_t = \mu_t + \sigma_\varepsilon \varepsilon_t \tag{34}$$

$$\mu_t = \mu_{t-1} + \sigma_\eta \eta_t \tag{35}$$

[2]Estamos assumindo que $\varepsilon_t \sim NI(0,1)$ assim como $\eta_t \sim NI(0,1)$ e ambos são independentes entre si.

$$\Downarrow$$

$$\begin{aligned} y_t &= (1)\alpha_t + \sigma_\varepsilon \varepsilon_t \\ \alpha_t &= [\mu_t] = [1]\,\alpha_{t-1} + \sigma_\eta \eta_t \end{aligned}$$

onde μ_t é o nível da série.

A forma deduzida do modelo de nível local é um $ARIMA(0,1,1)$ uma vez que (35) pode ser re-escrito da seguinte forma:

$$\Delta\mu_t = \sigma_\eta \eta_t \tag{36}$$

e aplicando o operador diferença em (34) temos

$$\Delta y_t = \Delta\mu_t + \sigma_\varepsilon \Delta\varepsilon_t \tag{37}$$

e substituindo (36) em (37) obtém-se:

$$\Delta y_t = \sigma_\eta \eta_t + \sigma_\varepsilon \Delta\varepsilon_t \tag{38}$$

e o lado direito de (38) pode ser escrito como $\sigma_\xi \xi_t + \theta\sigma_\xi \xi_{t-1}$[3], um $MA(1)$ cujos parâmetros são dados pela solução do seguinte sistema:

$$(1+\theta^2)\sigma_\xi^2 = \sigma_\eta^2 + 2\sigma_\varepsilon^2 \tag{39}$$

$$\theta\sigma_\xi^2 = -\sigma_\varepsilon^2 \tag{40}$$

dividindo (39) pode (40) obtém-se:

$$\begin{aligned} \frac{(1+\theta^2)}{\theta} &= -\frac{\sigma_\eta^2 + 2\sigma_\varepsilon^2}{\sigma_\varepsilon^2} \\ \frac{(1+\theta^2)}{\theta} &= -(q+2) \end{aligned} \tag{41}$$

onde $q = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$ e a solução de (41) é dada pela solução da seguinte equação de segundo grau:

[3]Estamos assumindo que $\xi_t \sim NI(0,1)$

$$\theta^2 + (q+2)\theta + 1 = 0 \tag{42}$$

(5) Modelo de Tendência Local

$$\begin{array}{rcllll} y_t & = & \mu_t & & & + \ \sigma_\varepsilon \varepsilon_t \\ \mu_t & = & \mu_{t-1} & + & \beta_{t-1} & + \ \sigma_\eta \eta_t \\ \beta_t & = & \beta_{t-1} & & & + \ \sigma_\xi \xi_t \end{array}$$

$$\Downarrow$$

$$\begin{array}{l} y_t = (1,0)\alpha_t + \sigma_\varepsilon \varepsilon_t \\ \alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\eta & 0 \\ 0 & \sigma_\xi \end{bmatrix} \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix} \end{array}$$

onde μ_t é o nível da série e β_t a taxa de crescimento

A forma reduzida deste modelo é um $ARIMA(0,2,2)$ uma vez que

$$\Delta\beta_t = \sigma_\xi \xi_t \Longrightarrow \beta_t = \frac{\sigma_\xi \xi_t}{(1-L)} \tag{43}$$

e substituindo (43) na equação do nível temos:

$$\Delta\mu_t = \beta_{t-1} + \sigma_\xi \xi_t \Longrightarrow \Delta\mu_t = \frac{\sigma_\xi \xi_{t-1}}{(1-L)} + \sigma_\eta \eta_t \tag{44}$$

e substituindo (44) na equação das observações temos:

$$\begin{array}{rcl} \Delta y_t & = & \Delta\mu_t + \sigma_\varepsilon \Delta\varepsilon_t \Longrightarrow \Delta y_t = \dfrac{\sigma_\xi \xi_{t-1}}{(1-L)} + \sigma_\eta \eta_t + \sigma_\varepsilon \Delta\varepsilon_t \\ \Delta\Delta y_t & = & \sigma_\xi \xi_{t-1} + \sigma_\eta \Delta\eta_t + \sigma_\varepsilon \Delta\Delta\varepsilon_t \end{array} \tag{45}$$

onde o lado direito de (45) pode ser escrito como $\sigma_\varsigma \varsigma_t + \theta_1 \sigma_\varsigma \varsigma_{t-1} + \theta_2 \sigma_\varsigma \varsigma_{t-2}$, [4] $MA(2)$ cujos parâmetros estão relacionados as variâncias do modelo de tendência local pelas seguintes equações:

$$(1 + \theta_1^2 + \theta_2^2)\sigma_\varsigma^2 = \sigma_\xi^2 + 2\sigma_\eta^2 + 6\sigma_\varepsilon^2 \tag{46}$$

[4] Estamos assumindo que $\varsigma_t \sim NI(0,1)$.

$$(\theta_1 + \theta_1\theta_2)\sigma_\varsigma^2 = -\sigma_\eta^2 - 4\sigma_\varepsilon^2 \tag{47}$$

$$\theta_2\sigma_\varsigma^2 = \sigma_\varepsilon^2 \tag{48}$$

cujas soluções dão os parâmetros do $MA(2)$.

2.2 Exercícios

1. Considere o seguinte modelo de regressão:

$$y_t = x_t'\beta + \varepsilon_t$$

e assumindo que β representa o vetor de estado, escreva este modelo em representação em espaço de estado.

2. Uma extensão da classe de modelo ARMA é a classe ARMAX que permite que variáveis exógenas sejam usadas no modelo ARMA. Podemos ter o ARMAX(1,0,1) onde a parte autorregressiva tem ordem um, a parte média móvel tem ordem zero e temos sómente uma variável exógena. Este modelo é dado por:

$$y_t = \phi y_{t-1} + x_t\beta + \varepsilon_t$$

Assumindo que β é um dos componentes do vetor de estado, escreva este modelo em representação em espaço de estado.

3. Escreva um programa no R para gerar o seguinte ARMA(1,1) mas escrito em representação em espaço de estado

$$y_t = 0.5 * y_{t-1} + \varepsilon_t + 0.3 * \varepsilon_{t-1}$$

assumindo que $\varepsilon_t \sim NI(0,1)$ e com condições iniciais $y_1 = \varepsilon_1$.

4. Escreva um programa no R para gerar um modelo de tendência local mas escrito em representação em espaço de estado.

5. Considere o seguinte modelo não linear, chamado de limiar autorregressivo, que pode caracterizar assimetrias do ciclo economico, dado por

$$y_t = \begin{cases} \phi_1 y_{t-1} + \varepsilon_{1t} & \text{se } y_{t-1} < 0 \\ \phi_2 y_{t-1} + \varepsilon_{2t} & \text{se } y_{t-1} \geq 0 \end{cases}$$

com $\varepsilon_{1t} \sim NI(0, \sigma_1^2)$ e $\varepsilon_{2t} \sim NI(0, \sigma_2^2)$. Como você poderia escrever este modelo em representação em espaço de estado?

3 Filtragem, Suavizamento e Previsão - Filtro de Kalman.[5]

3.1 Filtragem

Seja a_t o estimador ótimo de α_t usando a informação até o instante t, isto é, $\mathbf{I}_t$ e P_t uma matriz $m \times m$ de variância-covariância associada ao erro, isto é:

$$P_t = E_t[(\alpha_t - a_t)(\alpha_t - a_t)']$$

EQUAÇÕES DE PREVISÃO

No instante t, a_t e P_t são conhecidos, então o melhor previsor de α_{t+1} usando a informação até t é denotado por $a_{t+1|t}$ e será dado por:

$$\begin{aligned} a_{t+1|t} &= E_t[T_t\alpha_t + d_t + H_{t+1}\eta_{t+1}\ ] \\ &= T_t E_t[\alpha_t\ ] + d_t + H_{t+1}E_t[\eta_{t+1}\ ] \end{aligned}$$

$$\Updownarrow E_t[\eta_{t+1}\ ] = 0$$

$$a_{t+1|t} = T_t a_t + d_t \tag{49}$$

Agora o erro de previsão para o vetor de estado é dado por:

$$\begin{aligned} a_{t+1|t} - \alpha_{t+1} &= T_t a_t + d_t - T_t\alpha_t - d_t - H_{t+1}\eta_{t+1} \\ &= T_t(a_t - \alpha_t) - H_{t+1}\eta_{t+1} \end{aligned}$$

[5]Estamos usando a notação de Harvey (1993).

então a variância deste erro é dada por:

$$
\begin{aligned}
P_{t+1|t} &= E_t[(a_{t+1|t} - \alpha_{t+1})(a_{t+1|t} - \alpha_{t+1})'] = \\
&= T_t E_t[(a_t - \alpha_t)(a_t - \alpha_t)']T_t' + \\
&\quad + H_{t+1} E_t[\eta_{t+1}\eta_{t+1}']H_{t+1}' - 2T_t E_t[(a_t - \alpha_t)\eta_{t+1}']H_{t+1}' \qquad (50)
\end{aligned}
$$

mas

$$E_t[(a_t - \alpha_t)\eta_{t+1}'] = 0$$

então (50) reduz-se a:

$$P_{t+1|t} = T_t P_t T_t' + H_{t+1} H_{t+1}' \qquad (51)$$

ERRO DE PREVISÃO E SUA VARIÂNCIA

Dado (49) podemos obter o previsor das observação usando a informação até o instante $t-1$ que será dado por:

$$
\begin{aligned}
\widetilde{y}_{t+1|t} &= E_t[y_{t+1}] = E_t[Z_{t+1}\alpha_{t+1} + c_{t+1} + G_{t+1}\varepsilon_{t+1}] \\
&= Z_{t+1}E_t[\alpha_{t+1}] + c_{t+1} + G_{t+1}E_t[\varepsilon_{t+1}]
\end{aligned}
$$

$$\Updownarrow$$

$$\widetilde{y}_{t+1|t} = Z_{t+1} a_{t+1|t} + c_{t+1} \qquad (52)$$

logo, o erro de previsão um passo à frente será dado por:

$$
\begin{aligned}
\nu_{t+1} &= y_{t+1} - \widetilde{y}_{t+1|t} = \\
&= y_{t+1} - Z_{t+1} a_{t+1|t} - c_{t+1} \\
&= Z_{t+1}(\alpha_{t+1} - a_{t+1|t}) + G_{t+1}\varepsilon_{t+1} \qquad (53)
\end{aligned}
$$

e a variância deste erro é dada por:

$$
\begin{aligned}
F_{t+1} &= E_t[\nu_{t+1}\nu_{t+1}'] = E_t[Z_{t+1}(\alpha_{t+1} - a_{t+1|t})(\alpha_{t+1} - a_{t+1|t})' Z_{t+1}' + G_{t+1}\varepsilon_{t+1}\varepsilon_{t+1}' G_{t+1}' \\
&\quad + 2Z_{t+1}(\alpha_{t+1} - a_{t+1|t})\varepsilon_{t+1}']
\end{aligned}
$$

$$\Updownarrow$$

$$F_{t+1} = Z_{t+1}P_{t+1|t}Z'_{t+1} + G_{t+1}G'_{t+1} \tag{54}$$

EQUAÇÕES DE ATUALIZAÇÃO

Agora uma nova observação se torna disponível, isto é, y_{t+1} é observada e desejamos obter a atualização do vetor de estado levando em consideração esta nova observação. Observe então que podemos definir as seguintes equações:

$$\alpha_{t+1} = a_{t+1|t} + (\alpha_{t+1} - a_{t+1|t})$$

$$y_{t+1} = Z_{t+1}a_{t+1|t} + c_{t+1} + Z_{t+1}(\alpha_{t+1} - a_{t+1|t}) + G_{t+1}\varepsilon_{t+1}$$

como estamos assumindo distribuição normal para os choques temos que:

$$\begin{pmatrix} \alpha_{t+1} & \mid I_t \\ y_{t+1} & \end{pmatrix} \sim N\left[\begin{pmatrix} a_{t+1|t} \\ Z_{t+1}a_{t+1|t} + c_{t+1} \end{pmatrix}, \begin{pmatrix} P_{t+1|t} & P_{t+1|t}Z'_{t+1} \\ Z_{t+1}P_{t+1|t} & Z_{t+1}P_{t+1|t}Z'_{t+1} + G_{t+1}G'_{t+1} \end{pmatrix}\right]$$

e a distribuição de $\alpha_{t+1}|y_{t+1}, I_t$ será dada por[6]:

$$\alpha_{t+1}|y_{t+1}, I_t \sim N(a_{t+1|t} + P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}\nu_{t+1}\ , P_{t+1|t} - P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}Z_{t+1}P_{t+1|t}) \tag{55}$$

deste modo temos que:

$$a_{t+1} = E_t[\alpha_{t+1}] = a_{t+1|t} + P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}\nu_{t+1} \tag{56}$$

e

$$P_{t+1} = Var_t[\alpha_{t+1}] = P_{t+1|t} - P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}Z_{t+1}P_{t+1|t} \tag{57}$$

Resumindo as recursões do Filtro de Kalman

(1) Equações de previsão:

$a_{t+1|t} = T_t a_t + d_t$

$P_{t+1|t} = T_t P_t T'_t + H_{t+1}H'_{t+1}$

[6] Veja Apêndice I para a obtenção destes resultados para o caso geral.

(2) Erro de Previsão
$\nu_{t+1} = y_{t+1} - Z_{t+1}a_{t+1|t} - c_{t+1}$
(3) Variância do Erro de Previsão
$F_{t+1} = Z_{t+1}P_{t+1|t}Z'_{t+1} + G_{t+1}G'_{t+1}$
(4) Equações de atualização
$a_{t+1} = a_{t+1|t} + P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}\nu_{t+1}$
$P_{t+1} = P_{t+1|t} - P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}Z_{t+1}P_{t+1|t}$

Observe que são necessária condições iniciais para a_0 e P_0 que para modelos estacionários são dadas por:

(i) $a_0 =$ média não condicional do processo $= (I - T)^{-1}d_0$

(ii) $P_0 =$ variância não condicional do processo $\Leftrightarrow$ a solução da equação vetorial $P_0 = TP_0T' + HH'$

Os estimadores de máxima verossimilhança dos parâmetros são facilmente obtidos usando-se a decomposição do erro de previsão, isto é, como $\nu_t|\mathbf{I}_{t-1} \sim N(0, F_t)$ temos:

$$l(\Psi) = -\frac{NT}{2}ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T} ln|F_t| - \frac{1}{2}\sum_{t=1}^{T} \nu'_t F_t^{-1}\nu_t$$

3.1.1 Programa no R para Estimar Modelos por Filtro de Kalman

Os programas abaixo simulam e estimam modelos: $AR(1)$, $AR(4)$, $MA(1)$, $ARMA(1,1)$, nível local, tendência local e modelo de volatilidade estocástica

Modelo $AR(1)$

O programa para estimar um modelo $AR(1)$ no R está no link abaixo:

AR(1) usando pacote astsa

O correlograma para a série `yar1` é apresentado abaixo e a FAC decae exponencialmente e a FACP sómente a primeira autocorrelaçao parcial é diferente de zero implicando num modelo $AR(1)$.

Abaixo são apresentadas as estimativas dos parâmetros da série $yar1$. Observe que a estimativa da variância do erro é dada por σ_ε que é próxima de um e a estimativa do parâmetro autorregressivo, $AR(1)$, é próxima do valor verdadeiro, 0.75.

Figura 1: FAC e FACP para AR(1) com $\phi = 0.75$

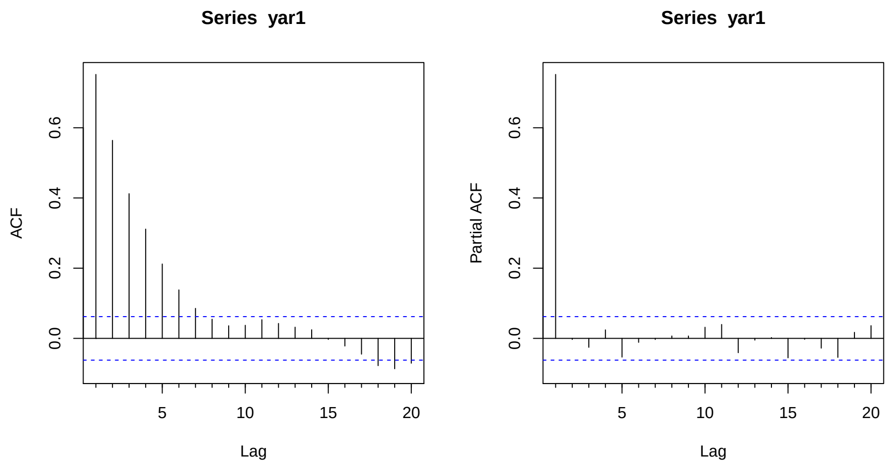


Tabela 1: Modelling AR(1) with $\phi = 0.75$ and $\sigma_\varepsilon = 1$

EQ(1) Modelling $AR(1)$ by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$AR(1)$	0.7520	0.0208	36.1464	0.000
σ_ε	0.9939	0.2222	44.7215	0.000
no of observations: 1000				
no of parameters: 2				
AIC = -0.9846				
BIC = -0.9610				
log-likelihood = 494.3142				

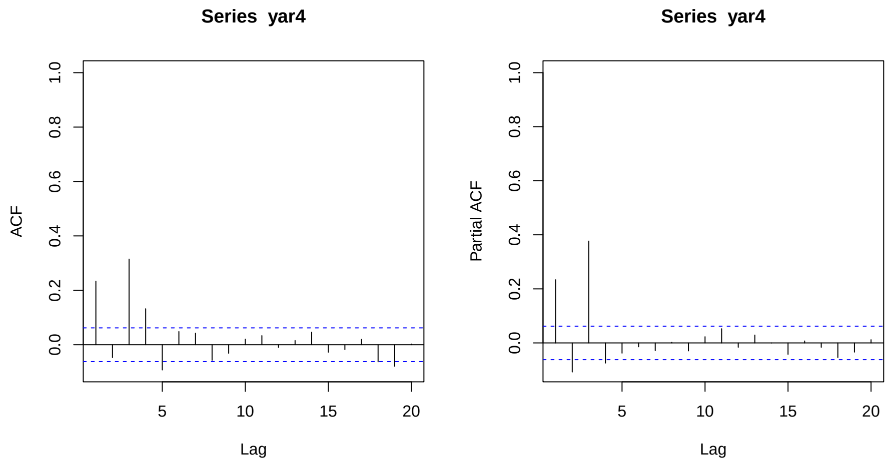
Modelo $AR(4)$

O programa para estimar um modelo $AR(4)$ no R está no link abaixo:

AR(4) usando pacote astsa

O correlograma abaixo tem uma FAC que decae exponencialmente e uma FACP com as quatro primeiras autocorrelações parciais diferentes de zero caracterizando um modelo $AR(4)$.

Figura 2: FAC e FACP para AR(4) com $\phi_1 = 0.32$, $\phi_2 = -0.23$, $\phi_3 = 0.43$ e $\phi_4 = -0.11$



Abaixo são apresentadas as estimativas dos parâmetros do modelo $yar4$. Observe que todas as estimativas são estatisticamente iguais as parâmetros verdadeiros.

Tabela 2: Modelling AR(4) with $\phi_1 = 0.32$, $\phi_2 = -0.23$, $\phi_3 = 0.43$, $\phi_4 = -0.11$ and $\sigma_\varepsilon = 1$

EQ(2) Modelling $AR(4)$ by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$AR(1)$	0.3251	0.0316	10.2827	0.0000
$AR(2)$	-0.2204	0.0308	-7.1564	0.0000
$AR(3)$	0.3980	0.0308	12.9223	0.0000
$AR(4)$	-0.0705	0.0313	-2.2522	0.0012
σ_ε	0.9924	0.0223	44.5640	0.0000
no of observations: 996				
no of parameters: 5				
AIC = -0.9840				
BIC = -0.9129				
log-likelihood = 496.0558				

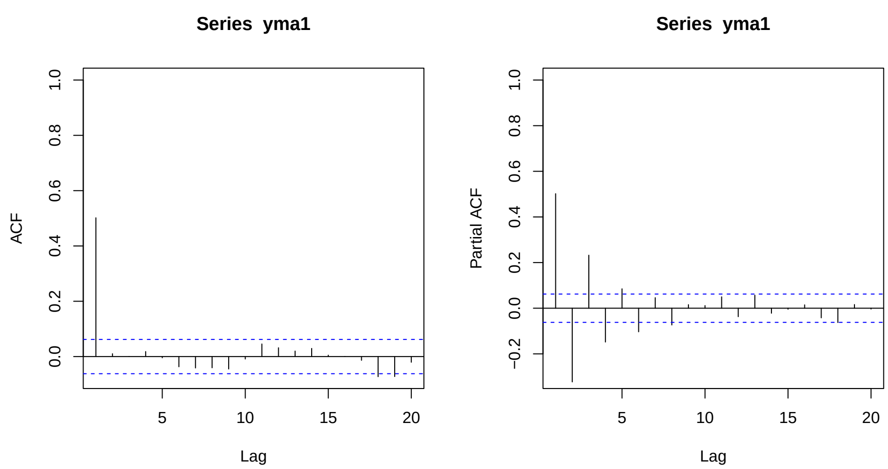
Modelo $MA(1)$

O programa para estimar um $MA(1)$ no R está no link abaixo:

MA(1) usando pacote astsa

Abaixo é apresentado o correlograma da série yma. Observe que sómente a primeira autocorrelação é diferente de zero e a FACP decaem exponencialmente caracterizando um processo $MA(1)$.

Figura 3: FAC e FACP para MA(1) com $\theta_1 = 0.8$



Abaixo são apresentadas as estimativas dos parâmetros para o modelo $MA(1)$ e a estimativa de θ_1 é estatísticamente igual ao parâmetro verdadeiro, mas a estimativa de σ_ε difere do valor verdadeiro.[7]

[7] Caso o OxMetrics fosse usado a estimativa de σ^2_ε seria 0.9861 próxima ao valor verdadeiro, veja o batch MA1 estiamdo no Ox via Batch ou via Ox Code MA1 estimado no Ox através de OxCode.

Tabela 3: Modelling MA(1) with $\theta_1 = 0.8$ and $\sigma_\varepsilon = 1$

EQ(3) Modelling $MA(1)$ by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$MA(1)$	0.7807	0.0610	12.7917	0.0000
σ_ε	1.6517	0.0610	27.0622	0.0000
no of observations: 999				
no of parameters: 2				
AIC = -3.0102				
BIC = -2.9865				
log-likelihood = 1505.57508				

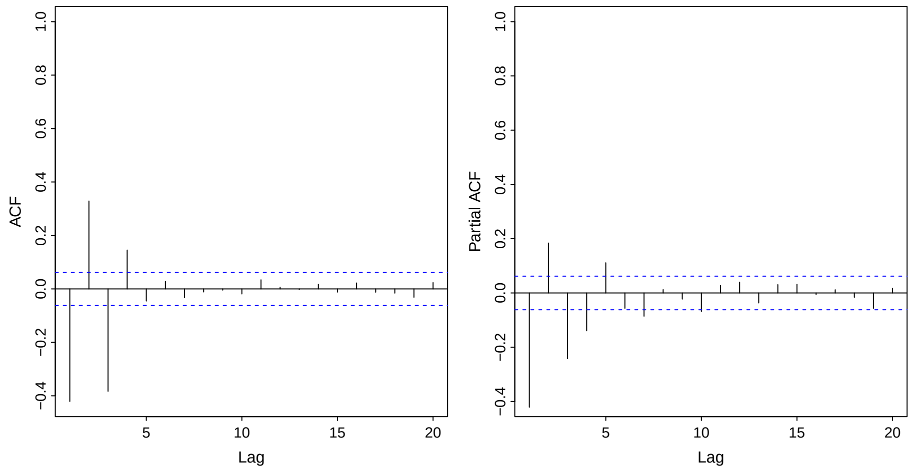
Modelo $MA(4)$

O programa para estimar um $MA(4)$ no R está no link abaixo:

MA(4) usando pacote astsa

Abaixo é apresentado o correlograma da série $yma4$. Observe que sómente as quatro primeiras autocorrelações são diferentes de zero e a FACP decae exponencialmente caracterizando um processo $MA(4)$.

Figura 4: FAC e FACP para MA(4) com $\theta_1 = -0.32$, $\theta_2 = 0.23$, $\theta_3 = -0.43$ e $\theta_4 = -0.1$



Abaixo são apresentadas as estimativas dos parâmetros para o modelo $MA(4)$ e a primeira não é estatísticamente igual ao parâmetro verdadeiro enquanto que as três últimas são.

Tabela 4: Modelling MA(4) with $\theta_1 = -0.32$, $\theta_2 = 0.23$, $\theta_3 = -0.43$, $\theta_4 = 0.11$ and $\sigma_\varepsilon = 1$

EQ(4) Modelling $MA(4)$ by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$MA(1)$	-0.3098	0.03084	-10.0429	0.0000
$MA(2)$	0.2397	0.02958	8.1111	0.0000
$MA(3)$	-0.4629	0.0309	-14.9970	0.0000
$MA(4)$	0.1474	0.0314	4.6902	0.0012
σ_ε	0.9918	0.0223	44.5657	0.0000
no of observations: 996				
no of parameters: 5				
AIC = -0.9832				
BIC = -0.9121				
log-likelihood = 495.6349				

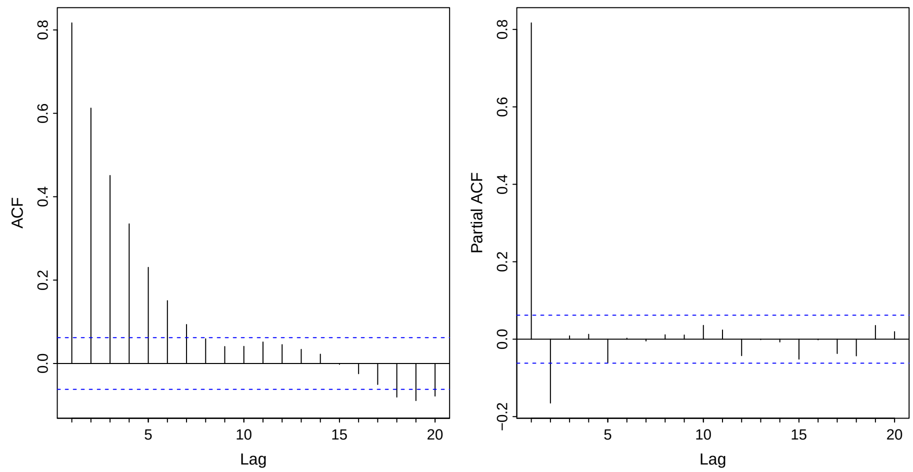
Modelo $ARMA(1,1)$

O programa para estimar um $ARMA(1,1)$ no R está no link abaixo:

ARMA(1,1) usando pacote astsa

Abaixo é apresentado o correlograma da série $yarma11$. Observe as autocorrelações decaem exponencialmente e a FACP decae exponencialmente caracterizando um processo $ARMA(1,1)$.

Figura 5: FAC e FACP para ARMA(1,1) com $\phi_1 = 0.75$ e $\theta_1 = 0.2$



Abaixo são apresentadas as estimativas dos parâmetros para o modelo $ARMA(1,1)$ e da variância dos erros. As três são estatísticamente iguais as parâmetros verdadeiros.

Tabela 5: Modelling ARMA(1,1) with $4\phi_1 = 0.75$, $\theta_1 = 0.2$ and $\sigma_\varepsilon = 1$

EQ(5) Modelling $ARMA(1,1)$ by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$AR(1)$	0.7545	0.0247	30.5031	0.0000
$MA(1)$	0.1942	0.0358	0.0000	
σ_ε	0.9921	0.0222	44.6557	0.0000
no of observations: 998				
no of parameters: 3				
AIC = -0.9835				
BIC = -0.9480				
log-likelihood = 493.769				

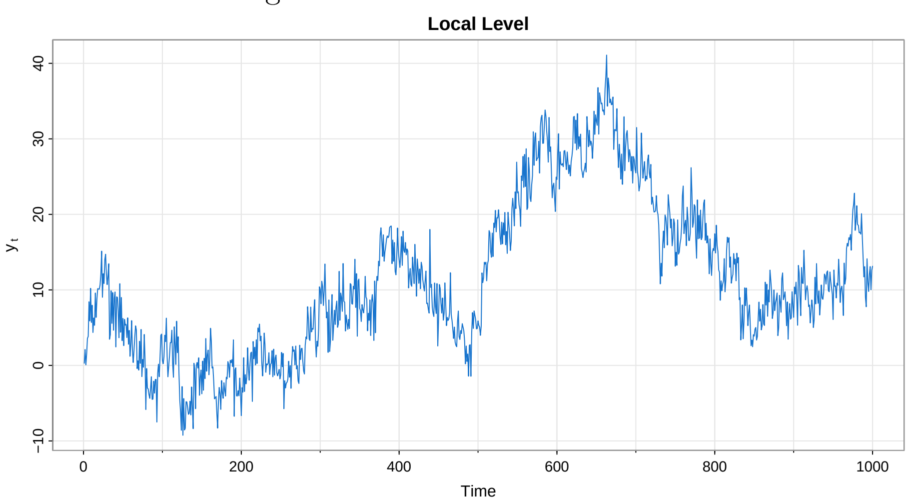
Modelo de Nível Local

O programa para estimar um Modelo de Nível Local no R está no link abaixo:

Local Level usando pacote astsa

A seguir é apresentado de o gráfico do modelo de nível local gerado pelo programa acoma. Observe que esta série é não estacionária uma vez que não retorna frequentemente para a sua média que é zero.

Figura 6: Modelo de Nível Local



Abaixo são apresentadas as estimativas dos parâmetros, a variância do nível e das observações e observe ambas são próximas dos valores verdadeiros, um para o erro do nível e quatro para o erro das observações.

Tabela 6: Modelling Local Level with $\sigma_\eta = 1$ and $\sigma_\varepsilon = 2$

EQ(6) Modelling Local Level by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
σ_η	1.0941	0.0707	15.4789	0.0000
σ_ε	1.9702	0.0620	0.0000	
no of observations: 1000				
no of parameters: 2				
AIC = -2.9006				
BIC = -2.8769				
log-likelihood = 1452.278				

Observe que também foi estimada a forma reduzida deste modelo cujos resultados estão apresentados abaixo

Tabela 7: Modelling Reduced Form of Local Level as an $ARIMA(0,1,1)$

EQ(6) Modelling Local Level by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$MA(1)$	-0.5782	0.0258	-22.4528	0.0000
no of observations: 998				
no of parameters: 2				
$\sigma_\xi^2 = 6.7201$				
AIC = 4.7475				
BIC =4.7573				
log-likelihood = -2369.381				

onde a relação entre os parâmetros da forma reduzida e da forma estrutural são obtidos através de (40) e (42), isto é,

$$\theta\sigma_\xi^2 = -4$$

$$\theta^2 + \frac{9}{4}\theta + 1 = 0$$

que implica que $\theta = -0.61$ e portanto $\sigma^2_{\xi} = 6.\,557\,4$ cujo desvio padrão é dado por $2.\,560\,7$, que estão próximo dos valores estimados que são de $\theta = -0.58$ e $\sigma_{\xi} = 2.5925$.

Mas se forem usados os valores estimados, isto é, $\sigma^2_{\eta} = 1.1970$ e $\sigma^2_{\varepsilon} = 3.8815$, teriamos as seguinte soluções $\theta = -0.578$ e com $\sigma_{\xi} = 2.5914$, e estes valores também estão próximos dos valores estimados.

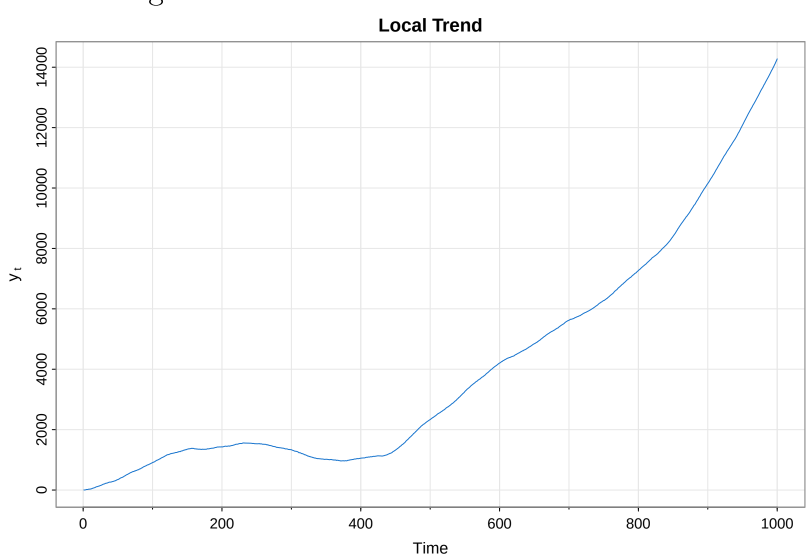
Modelo de Tendência Local

O programa para estimar um Modelo de Nível Local no R está no link abaixo:

Local Trend usando pacote astsa

Abaixo é apresentado o gráfico do modelo de tendência local gerado pelo programa acima. Observe que esta série é não estacionária é provavelmente é integrada de ordem 2.

Figura 7: Modelo de TendênciaNível Local



Abaixo são apresentadas as estimativas dos parâmetros do modelo de tendência local. Observe que estatísticamente os valores não são diferentes dos valores verdadeiros, uma vez que $\sigma^2_\varepsilon = 8.5685$ tendo um intervalo de confiança igual a $[6.8602 : 10.4665]$; $\sigma^2_\eta = 4.0136$ tendo um intervalo de confiança igual a $[1.6078 : 7.5010]$ e $\sigma^2_\xi = 1.0908$ tendo um intervalo de confiança igual a $[0.7786 : 1.4554]$.

Tabela 8: Modelling Local Trend with $\sigma_\xi = 1$, $\sigma_\eta = 2$ and $\sigma_\varepsilon = 3$

EQ(7) Modelling Local Trend by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
σ_ξ	1.0444	0.08103	128891	0.0000
σ_η	2.0034	0.3677	5.4489	0.0000
σ_ε	2.9272	0.1540	19.0076	0.0000
no of observations: 1000				
no of parameters: 3				
AIC = -4.2266				
BIC = -4.1912				
log-likelihood = 2116.3080				

A seguir é apresentada a estimativa da forma reduzida do modelo de tendência local

Tabela 9: Modelling Reduced Form of Local Trend as an $ARIMA(0, 2, 2)$

EQ(8) Modelling Local Trend by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
$MA(1)$	-1.1312	0.0324	-34.9501	0.0000
$MA(2)$	0.3389	0.0323	10.4892	0.0000
no of observations: 998				
no of parameters: 3				
$\sigma^2_\xi = 25.2816$				
AIC = 6.0754				
BIC = 6.0902				
log-likelihood = -3028.657				

onde a relação entre os parâmetros da forma reduzida e da forma estrutural são obtidos através de (46-48), isto é,

$$(1 + \theta_1^2 + \theta_2^2)\sigma_\varsigma^2 = 63$$

$$(\theta_1 + \theta_1\theta_2)\sigma_\varsigma^2 = -40$$

$$\theta_2\sigma_\varsigma^2 = 9$$

$$63 \times \theta_2 - 9 - 9 \times \left[\frac{-40 \times \theta_2}{9 \times (1 + \theta_2)}\right]^2 - 9 \times \theta_2^2 = 0$$

cujas soluções são: $0.350\,31$, $2.\,854\,6$, $0.897\,54 + 0.440\,93i$ e $0.897\,54 - 0.440\,93i$ e escolhemos a primeira que não é estatísticamente diferente do valor estimado que é 0.3389 cujo intervalo de confiança é dado por $[0.3389 - 2 * 0.0323 : 0.3389 + 2 * 0.0323] = [0.2743 : 0.4035]$.

Para obter o valor de θ_1 devemos resolver

$$\theta_1 = \frac{-40 \times \theta_2}{9 \times (1 + \theta_2)} = \frac{-40 \times (0.350\,31)}{9 \times (1 + (0.350\,31))} = -1.\,153$$

que não é estatísticamente diferente do valor estimado que é -1.1312 cujo intervalo de confiança é dado por $[-1.196 : -1.0664]$.

Para obter o valor de σ_ς^2 devemos resolver

$$\sigma_\varsigma = \sqrt{\frac{9}{\theta_2}} = \sqrt{\frac{9}{0.350\,31}} = 5.\,068\,7$$

que não é estatísticamente diferente do valor estimado que é 5.028081.

Modelo de Volatilidade Estocástica

A formulação do Modelo de Volatilidade Estocástica no programa `astsa` é diferente do apresentado acima iremos utilizar esta formulação, onde r_t são o log-retorno de algum ativo e $x_t = log(\sigma_t^2)$ a log volatilidade deste ativo, a equação de transição é dada por:

$$x_{t+1} = \gamma r_t + \phi x_t + \sigma\omega_t \tag{58}$$

E a equação de observação onde $y_t = log(r_t^2)$,[8] é dada por:

$$y_t = \alpha + x_t + \eta_t \tag{59}$$

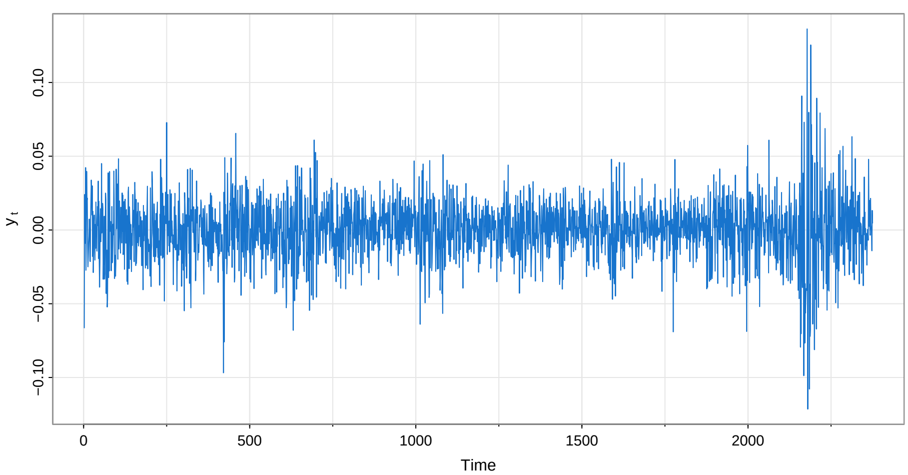
onde η_t é uma mistura de duas normais, uma centada no zero que tenta aproximar a distribuição da $log(\chi_1^2)$.É permitido correlação entre ω_t e η_t e é denotada por ρ que é necessária de existir feedback, isto é $\gamma \neq 0$ em (58).

O programa para estimar um Modelo de Volatilidadde Estocástica para o log do Indice IBOVESPA especificado por (59) e (58) no R está no link abaixo:

SV QMLE usando pacote astsa

O gráfico do log retorno do IBOVESPA é apresentado abaixo.

Figura 8: Log Retorno do IBOVESPA Centrado



Abaixo são apresentadas as estimativas dos parâmetros, média, coeficiente autorregressivo e variância para o modelo do logaritmo da variância.

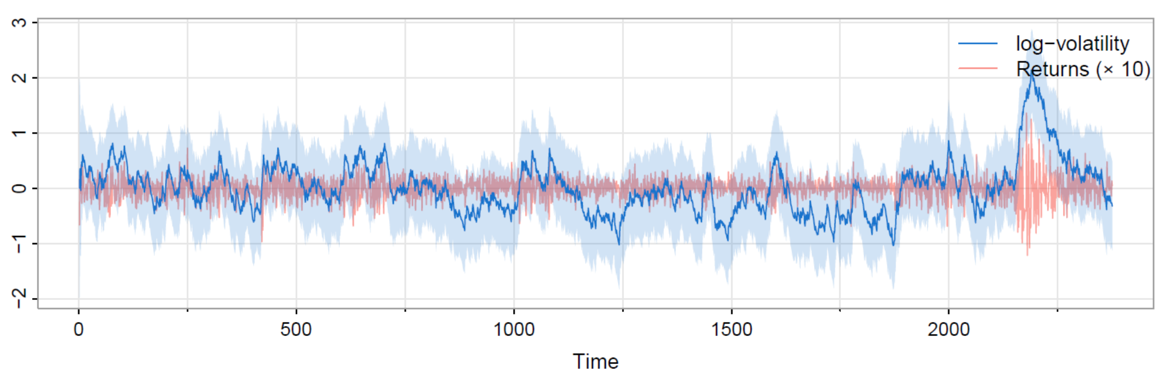
[8]Como r_t pode ser zero usamos $r_t - \overline{r}_t$ onde $\overline{r}_t$ representa a média de r_t.

Tabela 10: Modelling SV Model using QMLE

EQ(10) Modelling SV Model by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
ϕ	0.9824	0.0087	112,7971	0.0000
σ_ω	0.1276	0.0298	4.2824	0.0000
α	-8.12	0.1916	-42.4196	0.0000
σ_0	0.9768	0.0451	21.6526	0.0000
σ_1	2.5153	0.0864	29.1175	0.0000
no of observations: 2374				
no of parameters: 6				
AIC = -1.2079				
BIC = -1.1734				
log-likelihood = 1439.7720				

A seguir é apresentado o gráfico da log-volatilidade junto com o log retorno do IBOVESPA centrado

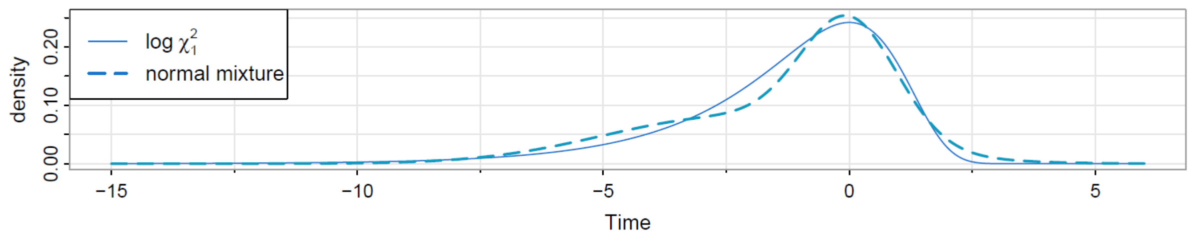
Figura 9: Log-Volatilidadde e log Retorno do IBOVESPA Centrado



observe que nos momentos de crse há um aumento de volatilidade.

Como a distribuição do erro observacional foi aproximada pela mistura de duas normais o gráfico abaixo apresenta d sitribuição verdadeira ($log(\xi_1^2)$ e da mistura das duas normais.

Figura 10: Log(ξ_1^2) e mistura das normais



3.2 Previsão

A previsão do vetor de estado é dada por:

$$a_{T+l|T} = T_{T+l-1}a_{T+l-1|T} + d_{T+l-1} \text{ para } l = 1, 2, ...$$

com $a_{T|T} = a_T$.

E da matriz de covariância por:

$$P_{T+l|T} = T_{T+l-1}P_{T+l-1|T}T'_{T+l-1} + H_{T+l}H'_{T+l}$$

O previsor das observações é dado por:

$$\widetilde{y}_{T+l|T} = Z_{T+l}a_{T+l|T} + c_{T+l}$$

com um erro quadrático médio dado por:

$$EMQ(\widetilde{y}_{T+l|T}) = Z_{T+l}P_{T+l-1|T}Z'_{T+l}$$

3.3 Suavizamento[9]

Usando toda a informação amostral para obter o melhor estimador da média e da matriz de covariância do vetor de estado. As recursões são dadas por:

$$a_{t|T} = a_t + P_t^*(a_{t+1|T} - T_{t+1}a_t - c_{t+1})$$

[9]Em Durbin and Koopman (2012) são apresentados outros algorítmos de suavizamento que são mais eficiente do que este.

$$P_{t|T} = P_t + P_t^*(P_{t+1|T} - P_{t+1|t})P_t^{*\prime}$$

$$P_t^* = P_t T_{t+1}' P_{t+1|t}^{-1}$$

para $t = T-1, ...1$, com as seguintes condições iniciais:

$$a_{T|T} = a_T$$

$$P_{T|T} = P_T$$

3.3.1 Alguns exemplos

(1) $AR(1)$

A representação em espaço de estado de um AR(1) é dada por:

$$y_t = (1)\ \alpha_t$$

$$\alpha_t = \phi_1 \alpha_{t-1} + \sigma \varepsilon_t$$

e para iniciar $a_0 = 0$ e $P_0 = \frac{\sigma^2}{1-\phi_1^2}$

$$a_{1|0} = 0$$

$$P_{1|0} = \phi_1^2 \frac{\sigma^2}{1-\phi_1^2} + \sigma^2 = \frac{\sigma^2}{1-\phi_1^2}$$

$$\nu_1 = y_1$$

$$f_1 = \frac{\sigma^2}{1-\phi_1^2}$$

$$a_1 = 0 + \frac{\sigma^2}{1-\phi_1^2}\left(\frac{\sigma^2}{1-\phi_1^2}\right)^{-1} y_1 \Rightarrow a_1 = y_1$$

$$P_1 = \frac{\sigma^2}{1-\phi_1^2} - \frac{\sigma^2}{1-\phi_1^2}\left(\frac{\sigma^2}{1-\phi_1^2}\right)^{-1}\frac{\sigma^2}{1-\phi_1^2} = 0$$

$$a_{2|1} = \phi_1 y_1$$

$$a_{2|1} = \phi_1 y_1$$

$$P_{2|1} = 0 + \sigma^2 \Rightarrow P_{2|1} = \sigma^2$$

$$\nu_2 = y_2 - \phi_1 y_1$$

$$f_2 = \sigma^2$$

$$a_2 = \phi_1 y_1 + \sigma^2 \left(\sigma^2\right)^{-1} (y_2 - \phi_1 y_1) \Rightarrow a_2 = y_2$$

$$P_2 = \sigma^2 - \sigma^2 \left(\sigma^2\right)^{-1} \sigma^2 = 0$$

então é fácil ver que para $t \geq 2$,

$$a_{t|t-1} = \phi_1 y_{t-1}$$

$$P_{t|t-1} = \sigma^2$$

$$\nu_t = y_t - \phi_1 y_{t-1}$$

$$f_t = \sigma^2$$

$$a_t = y_t$$

$$P_t = 0$$

e a função de verossimilhança é dada por:

$$\begin{aligned} l(\Psi) &= -\frac{T}{2} ln(2\pi) - \frac{T}{2} ln\sigma^2 + \frac{1}{2} ln(1-\phi_1^2) - \frac{1}{2\sigma^2} y_1^2 (1-\phi_1^2) \\ &\quad -\frac{1}{2\sigma^2} \sum_{t=2}^{T} (y_t - \phi_1 y_{t-1})^2 \end{aligned}$$

que é a verossimilhança exata para um $AR(1)$.

(2) $MA(1)$

A representação em espaço de estado de um $MA(1)$ é dada por:

$$y_t = (1:0)\ \alpha_t$$

$$\alpha_t = \begin{bmatrix} y_t \\ l_t^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma \\ \sigma\theta_1 \end{bmatrix} \varepsilon_t$$

e para iniciar observe que $\alpha_0 = \begin{bmatrix} y_0 \\ \sigma\theta_1\varepsilon_0 \end{bmatrix}$ então $a_0 = E(\alpha_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ e

$$P_0 = E \begin{bmatrix} y_0^2 & y_0(\sigma\theta_1\varepsilon_0) \\ (\sigma\theta_1\varepsilon_0)y_0 & \sigma^2\theta_1^2\varepsilon_0^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} (1+\theta_1^2) & \theta_1 \\ \theta_1 & \theta_1^2 \end{bmatrix}$$

então

$$a_{1|0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} P_{1|0} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma^2(1+\theta_1^2) & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} \sigma \\ \sigma\theta_1 \end{bmatrix} \begin{bmatrix} \sigma & \sigma\theta_1 \end{bmatrix} \\ &= \begin{bmatrix} \sigma^2\theta_1^2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma^2 & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma^2(1+\theta_1^2) & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} \end{aligned}$$

, transpose:

$$\nu_1 = y_1$$

$$f_1 = \sigma^2(1+\theta_1^2)$$

$$a_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \sigma^2(1+\theta_1^2) \\ \sigma^2\theta_1 \end{bmatrix} \frac{y_1}{\sigma^2(1+\theta_1^2)} \Rightarrow a_1 = \begin{bmatrix} y_1 \\ \frac{\theta_1\ y_1}{1+\theta_1^2} \end{bmatrix}$$

$$
\begin{aligned}
P_1 &= \begin{bmatrix} \sigma^2(1+\theta_1^2) & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} - \begin{bmatrix} \sigma^2(1+\theta_1^2) \\ \sigma^2\theta_1 \end{bmatrix} \frac{1}{\sigma^2(1+\theta_1^2)} \begin{bmatrix} \sigma^2\left(\theta_1^2+1\right) & \sigma^2\theta_1 \end{bmatrix} \\
&= \begin{bmatrix} \sigma^2(1+\theta_1^2) & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} - \begin{bmatrix} \sigma^2\left(1+\theta_1^2\right) & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\frac{\theta_1^2}{\theta_1^2+1} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2\frac{\theta_1^4}{\theta_1^2+1} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
a_{2|1} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \frac{\theta_1\ y_1}{1+\theta_1^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\theta_1\ y_1}{1+\theta_1^2} \\ 0 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
P_{2|1} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2\frac{\theta_1^4}{\theta_1^2+1} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} \sigma \\ \sigma\theta_1 \end{bmatrix} \begin{bmatrix} \sigma & \sigma\theta_1 \end{bmatrix} \\
&= \begin{bmatrix} \sigma^2\frac{\theta_1^4}{\theta_1^2+1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma^2 & \sigma^2\theta_1 \\ \sigma^2\theta_1 & \sigma^2\theta_1^2 \end{bmatrix} \\
&= \sigma^2 \begin{bmatrix} \frac{1+\theta_1^2+\theta_1^4}{1+\theta_1^2} & \theta_1 \\ \theta_1 & \theta_1^2 \end{bmatrix}
\end{aligned}
$$

$$
\nu_2 = y_2 - \frac{\theta_1\ y_1}{1+\theta_1^2} = y_2 - \frac{\theta_1\ \nu_1}{\frac{f_1}{\sigma^2}}
$$

$$
f_2 = \sigma^2\left(\frac{1+\theta_1^2+\theta_1^4}{1+\theta_1^2}\right) = \sigma^2(1+\frac{\theta_1^4}{1+\theta_1^2})
$$

então para $t \geq 2$ temos

$$
\nu_t = y_t - \frac{\theta_1\ \nu_{t-1}}{\frac{f_{t-1}}{\sigma^2}}
$$

$$
f_t = \sigma^2\left(1+\frac{\theta_1^{2t}}{1+\theta_1^2+...+\theta_1^{2(t-1)}}\right)
$$

e a função de verossimilhança é uma função não linear em θ_1. Observe que ν_t e f_t resultam na mesma expressão quando se subtitui θ_1 por $\frac{1}{\theta_1}$. Deste modo a maximização da verossimilhança dá a solução invertível do processo, como indicado em Valls Pereira (1987).

(3) Mínimos Quadrados Recursivos

Considere o seguinte modelo de regressão

$$y_t = x_t'\beta_t + u_t$$

$$\beta_t = \beta_{t-1}$$

onde o vetor de parâmetros tem dimensão $k \times 1$.

O estimador de β_t usando a informação até o instante t, denotado por b_t, é dado por:

$$b_t = \left(\sum_{j=1}^{t} x_j x_j'\right)^{-1} \sum_{j=1}^{t} x_j y_j = (X_t' \ X_t)^{-1}(X_t'\mathbf{y}_t) \tag{60}$$

para $t > k$.

Agora observe que

$$\begin{aligned}(X_t' \ X_t)^{-1} &= (X_{t-1}' \ X_{t-1} + x_t x_t')^{-1} \\ &= \left(X_{t-1}' \ X_{t-1}\right)^{-1} - \frac{\left(X_{t-1}' \ X_{t-1}\right)^{-1} x_t x_t' \left(X_{t-1}' \ X_{t-1}\right)^{-1}}{1 + x_t' \ \left(X_{t-1}' \ X_{t-1}\right)^{-1} \ x_t}\end{aligned} \tag{61}$$

onde foi usado o seguinte resultado $(A + bb')^{-1} = \left(A^{-1} - \frac{A^{-1}bb'A^{-1}}{1+b'A^{-1}b}\right)$. Agora substituindo (61) em (60) e usando que $(X_t'\mathbf{y}_t) = (X_{t-1}'\mathbf{y}_{t-1} + x_t' y_t)$ tem-se:

$$b_t = b_{t-1} + \left(X_{t-1}' \ X_{t-1}\right)^{-1} x_t \frac{(y_t - x_t' b_{t-1})}{f_t} \tag{62}$$

onde $f_t = 1 + x_t' \ \left(X_{t-1}' \ X_{t-1}\right)^{-1} \ x_t$.

Para iniciar o processo faça $t = k$ em (60) obtendo:

$$b_k = (X_k' \ X_k)^{-1}(X_k'\mathbf{y}_k) = X_k^{-1}\mathbf{y}_k$$

Observe que (61) corresponde, a menos de σ^2, a atualização da matriz de covariância do estimador dos parâmetros e (62) é a atualização da média do vetor de estado que corresponde ao estimador dos parâmetros.

O erro de previsão um passo a frente é dado por:

$$\nu_t = y_t - x_t' b_{t-1}$$

cuja variância é f_t. Logo podemos definir o erro de previsão padronizado por:

$$\widetilde{\nu}_t = \frac{\nu_t}{\sqrt{f_t}}$$

Observe também a soma dos quadrados dos resíduos também pode ser obtida recursivamente usando-se:

$$SQRes_t = SQRes_{t-1} + \nu_t^2$$

(4) Modelo de Nível Local

O Modelo de Nível Local é definido por:

$$\begin{array}{ccccc} y_t & = & \mu_t & + & \varepsilon_t \\ \mu_t & = & \mu_{t-1} & + & \eta_t \end{array}$$

com ε_t e η_t ruídos brancos independentes com variâncias σ_ε^2 e $\sigma_\eta^2 = q\sigma_\varepsilon^2$

Observe que este modelo já está em representação em espaço de estado.

As equações de previsão são dadas por:

$$m_{t|t-1} = m_{t-1}$$

$$p_{t|t-1} = p_{t-1} + q$$

O erro de previsão é dado por:

$$\nu_t = y_t - m_{t|t-1}$$

e a variância do erro de previsão é dada por:

$$f_t = 1 + p_{t|t-1}$$

As equações de atualização são dadas por:

$$m_t = m_{t|t-1} + p_{t|t-1}\frac{(y_t - m_{t|t-1})}{1 + p_{t|t-1}} \tag{63}$$

$$p_t = p_{t|t-1} - \frac{p^2_{t|t-1}}{1 + p_{t|t-1}} \tag{64}$$

As condições iniciais são

$$m_1 = y_1$$

$$p_1 = 1$$

Observe que (63) pode ser reescrito da seguinte forma:

$$m_t = (1 - \lambda_t)m_{t-1} + \lambda_t y_t \tag{65}$$

onde $\lambda_t = \frac{p_{t-1}+q}{1+p_{t-1}+q}$, isto é um modelo de suavizamento exponencial com coeficiente não constante. Observe que (64) pode ser reescrito da seguinte forma:

$$p_t = p_{t-1} + q - \frac{(p_{t-1} + q)^2}{1 + p_{t-1} + q}$$

Observe também que a previsor do vetor de estado em $T + l$ usando a informação até o instante T é dada por:

$$m_{T+l|T} = m_T$$

com erro quadrático associado dado por:

$$p_{T+l|T} = p_T + lq$$

Logo o previsor das observações futuras será dado por:

$$\widetilde{y}_{T+l|T} = m_T$$

com erro quadrático de previsão dado por:

$$EQP(\widetilde{y}_{T+l|T}) = p_T + lq + \sigma^2_\varepsilon$$

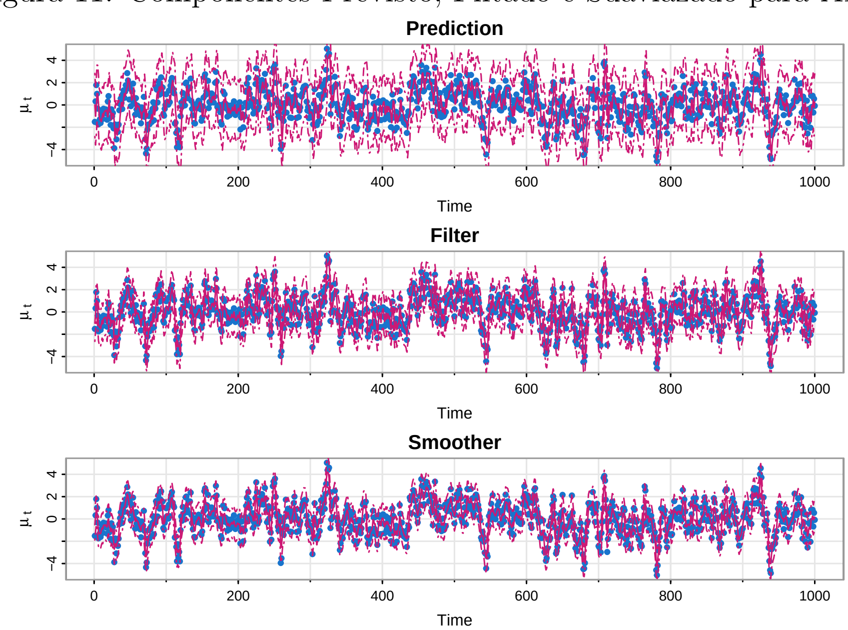
3.3.2 Comando no R para gerar componentes previsto, filtrados e suavizados

No programa acima após a estimação do modelo $AR(1)$ se forem usado os comando abaixo é possível extrair os componentes filtrados e suavizados para o vetor de estado.

```
# filter and smooth (Ksmooth does both)
mu0 = 0
sigma0 = 1
phi = 1
sQ = 1
sR = 1
ks = Ksmooth(yar1, A=1, mu0, sigma0, phi, sQ, sR)
```

O gráfico abaixo apresenta os componentes previsto ($yar1_{t|t-1}$), filtrados ($yar1_{t|t}$) e suavizados ($yar1_{t|T}$) para o modelo $AR(1)$ gerado acima. Observe que como temos erro na equação de observação o componente suavizado não é igual a própria série.

Figura 11: Componentes Previsto, Filtado e Suaviazado para $AR(1)$



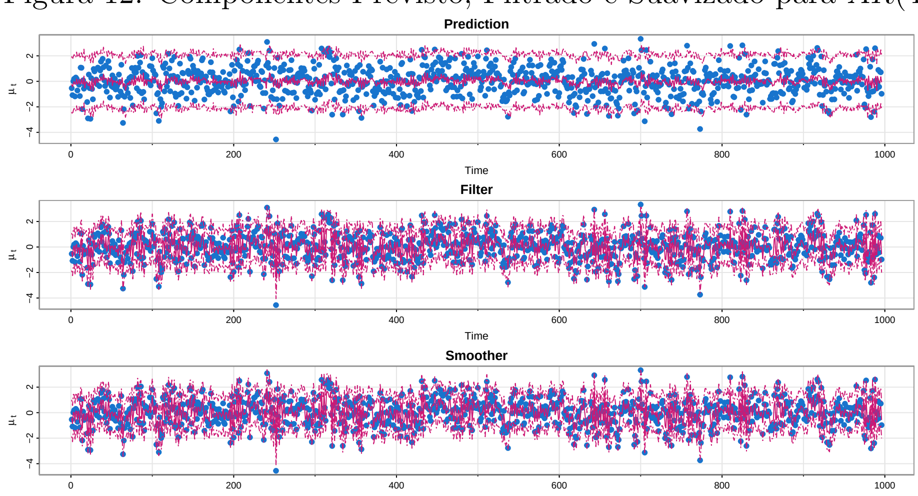
O mesmo para o modelo $AR(4)$

No programa acima após a estimação do modelo $AR(4)$ se forem usado os comando abaixo é possível extrair os componentes previstos, filtrados e suavizados para o vetor de estado.

```
# filter and smooth (Ksmooth does both)
num = length(y)
A = matrix(c(1,0,0,0),1,4)
mu0=matrix(c(0,0,0,0),4,1)
Sigma0 = diag(100,4)
phi = matrix(c(ests$par[1],ests$par[2],ests$par[3],ests$par[4],
1,0,0,0, 0,1,0,0, 0,0,1,0),4)
phi=t(phi)
sR = ests$par[5]
sQ = matrix(c(1,0,0,0),4)*sR
ks = Ksmooth(yar4, A, mu0, Sigma0, phi, sQ, sR)
```

O gráfico abaixo apresenta os componentes filtrados e suavizados do primeiro elemento do vetor de estado para o modelo $AR(4)$ gerado acima. Observe que como temos erro na equação de observação o componente suavizado não é igual a própria série.

Figura 12: Componentes Previsto, Filtrado e Suavizado para $AR(4)$



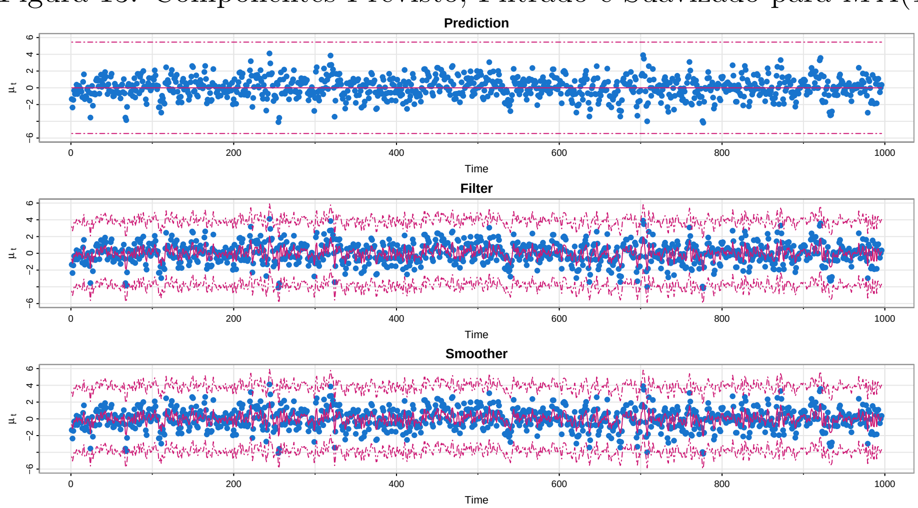
Para o modelo $MA(1)$:

No programa acima após a estimação do modelo $AR(4)$ se forem usado os comando abaixo é possível extrair os componentes previstos, filtrados e suavizados para o vetor de estado.

```
# filter and smooth (Ksmooth does both)
num = length(y)
A = matrix(c(1,0),1,2)
theta1 =ests$par[1]
mu0=matrix(c(0,0),2,1)
Sigma0 = diag(100,2)
phi = matrix(c(0,1,0,0),2)
S = 1
sR = ests$par[2]
sQ = matrix(c(1,ests$par[2]),2)*sR
ks = Ksmooth(yma1, A, mu0, Sigma0, phi, sQ, sR)
```

O gráfico abaixo apresenta os componentes previsto, filtrados e suavizados para o sinal da representação em espaço de estado do modelo $MA(1)$ gerado acima.

Figura 13: Componentes Previsto, Filtrado e Suavizado para $MA(1)$



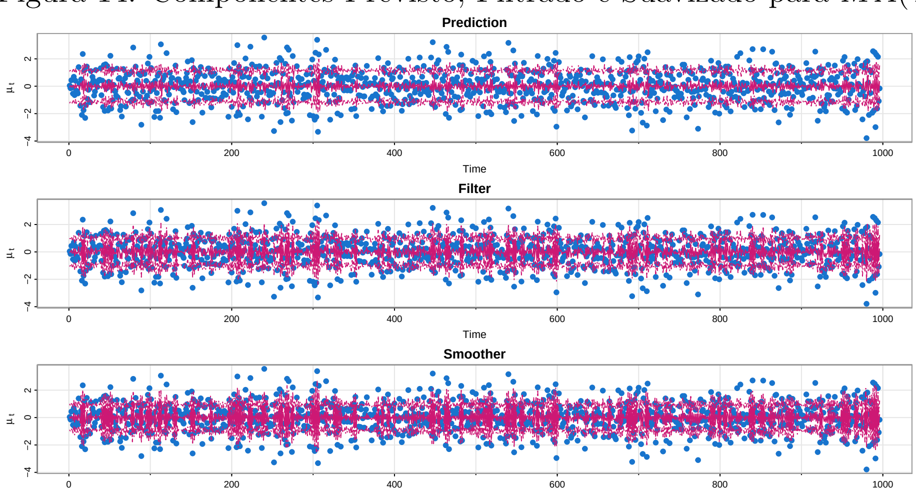
Para o modelo $MA(4)$:

No programa acima após a estimação do modelo $AR(4)$ se forem usado os comando abaixo é possível extrair os componentes previstos, filtrados e suavizados para o vetor de estado.

```
# filter and smooth (Ksmooth does both)
num = length(y)
A = matrix(c(1,0,0,0),1,4)
theta1 =ests$par[1]; theta2 = ests$par[2]; theta3 = ests$par[3];
theta4 = ests$par[4]
mu0=matrix(c(0,0,0,0),4,1)
Sigma0 = diag(100,4)
phi = matrix(c(0,0,0,0,1,0,0,0,0,1,0,0,0,0,1,0),4)
S = 1
sR = ests$par[5]
sQ = matrix(c(theta1,theta2,theta3,theta4),4)*sR
ks = Ksmooth(yma4, A, mu0, Sigma0, phi, sQ, sR)
```

O gráfico abaixo apresenta os componentes previsto, filtrados e suavizados para o sinal da representação em espaço de estado do modelo $MA(4)$ gerado acima.

Figura 14: Componentes Previsto, Filtrado e Suavizado para $MA(4)$

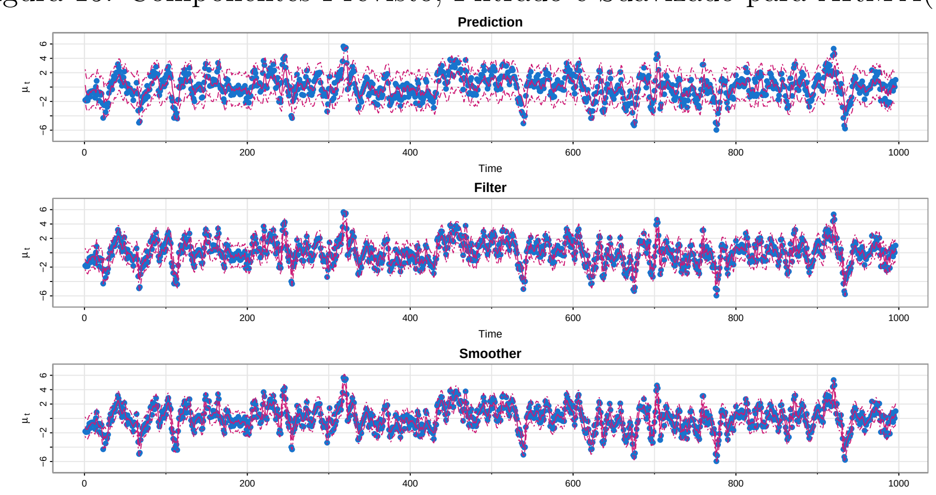


Para o modelo $ARMA(1,1)$:

```
# filter and smooth (Ksmooth does both)
num = length(y)
A = matrix(c(1),1,1)
phi = est$par[1]
theta1 = est$par[2]
mu0=matrix(c(0),1,1)
sigma0 = diag(100,1)
Phi = matrix(c(phi),1)
S = 1
sR = est$par[3]
sQ = matrix(c(phi+theta1),1)*sR
ks = Ksmooth(yarma11, A=1, mu0, sigma0, Phi, sQ, sR)
```

O gráfico abaixo apresenta os componentes previsto, filtrados e suavizados para o sinal da representação em espaço de estado do modelo $ARMA(1,1)$ gerado acima.

Figura 15: Componentes Previsto, Filtrado e Suavizado para $ARMA(1,1)$



Para o modelo de nível local:

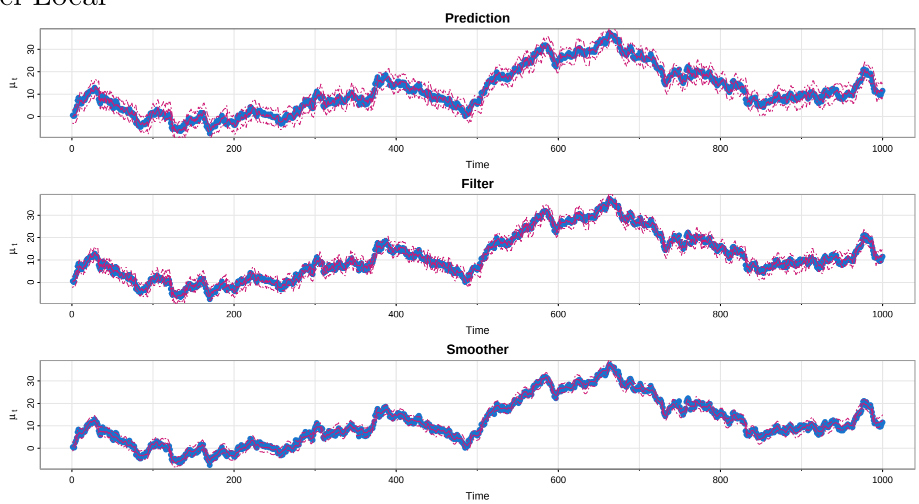
No programa acima após a estimação do modelo de Nível Local se forem usado os comando abaixo é possível extrair os componentes previstos, filtrados e suavizados para o vetor de estado.

```
# filter and smooth (Ksmooth does both)
mu0 = 0; sigma0 = 1; phi = 1; sQ = ests$par[1]; sR = ests$par[2]
ks = Ksmooth(y, A=1, mu0, sigma0, phi, sQ, sR)
```

A seguir são apresentados os componentes previsto, filtrado e suavizado do vetor de estado do modelo de nível local. Observe que este componente representa o nível da série original, quano se usa a inforamção até $t-1$, t e T, respectivamente.

A seguir é apresentado o componente filtrado das observações no modelo de nível local.

Figura 16: Componentes Previsto, Filtrado e Suavizado para o Modelo de Nível Local



Para o modelo de tendência local:

No programa acima após a estimação do modelo de Tendência Local se forem usado os comando abaixo é possível extrair os componentes previstos, filtrados e suavizados para o vetor de estado.

```
A = cbind(1,0)
mu0=rbind(0,0)
Sigma0=diag(100,2)
Phi = diag(0,2); Phi[1,1]=1; Phi[1,2]=1; Phi[2,2]=1
cQ1=ests$par[1]; cQ2=ests$par[2]; cR=ests$par[3]
```

```
cQ=diag(0,2); cQ[1,1]=cQ1; cQ[2,2]=cQ2;
ks = Ksmooth(y, A, mu0, Sigma0, Phi, cQ, cR)
```

A seguir são apresentados os componentes previstos, filtrados e suavizados do vetor de estado do modelo de tendência local. Observe que o primeiro componente representa o nível da série original e o segundo a taxa de crescimento.

Figura 17: Componentes Previsto, Filtrado e Suavizado para o Nível no Modelo de Tendência Local

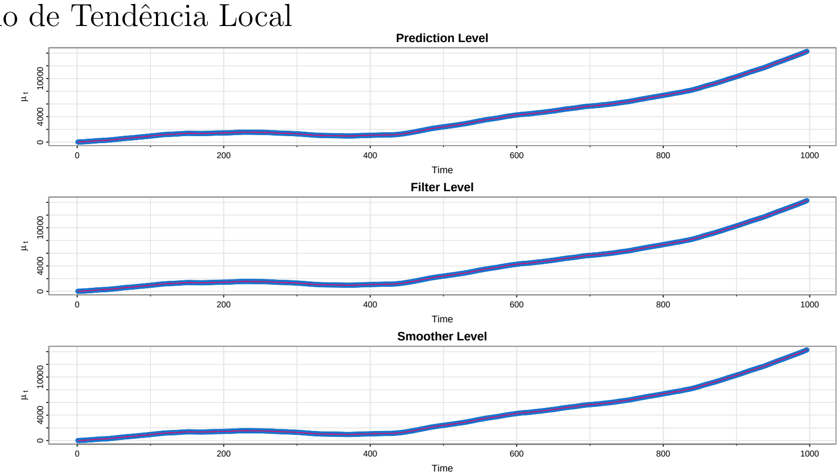
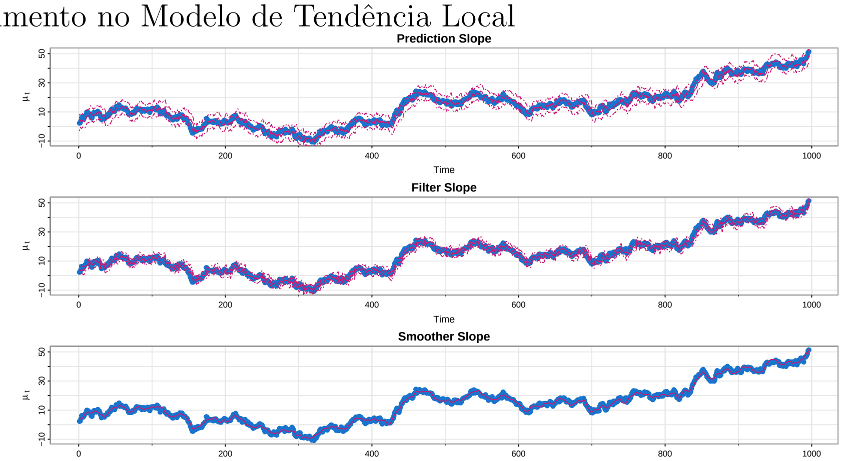


Figura 18: Componentes Previsto, Filtrado e Suavizado para a Taxa de Crescimento no Modelo de Tendência Local

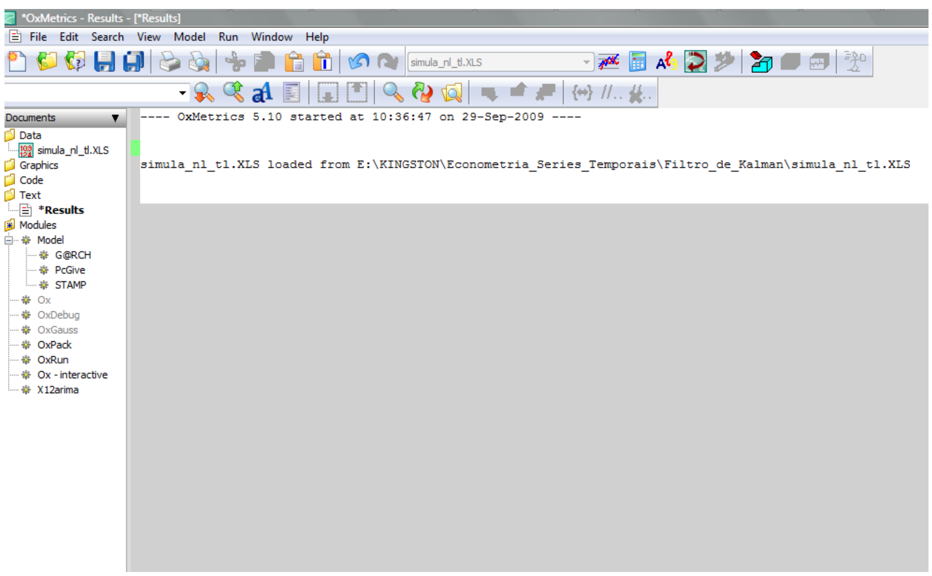


Para o modelo de volatilidade estocástica a obtenção dos componentes previsto, filtrado e suavizado será feita usando o STAMP na seção 3.3.4

3.3.3 Estimando o Modelo de Nível Local e de Tendência Local no OxMetrics usando o STAMP

A tela de entrada do OxMetrics é apresentada a seguir onde já foi importado os arquivos simula_nl.xlsx e simula_tl.xlsxque contém os dados simulados para o modelo de nível local e de tendência local.

Figura 19: Tela de Entrada do OxMetrics



Para ativar os módulos do OxMetrics, clique em `Model` na barra a esquerda que aparecerá a seguinte tela:

A seguir, ative o modulo STAMP, clicando em `Formulate` e aparecerá a seguinte caixa de diálogo

Figura 20: Tela dos Modelos no OxMetrics

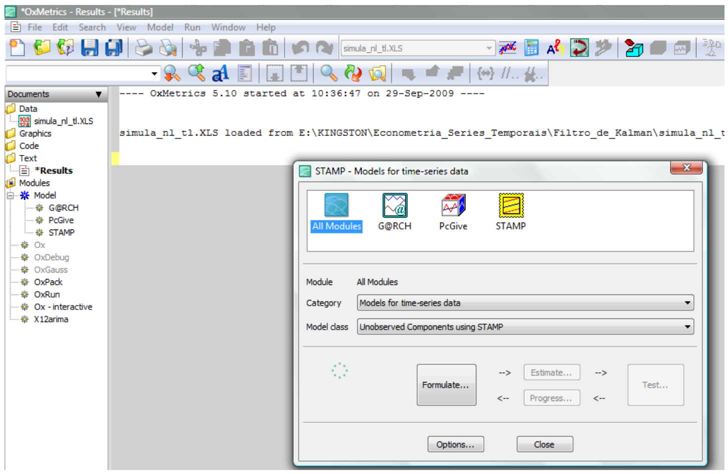
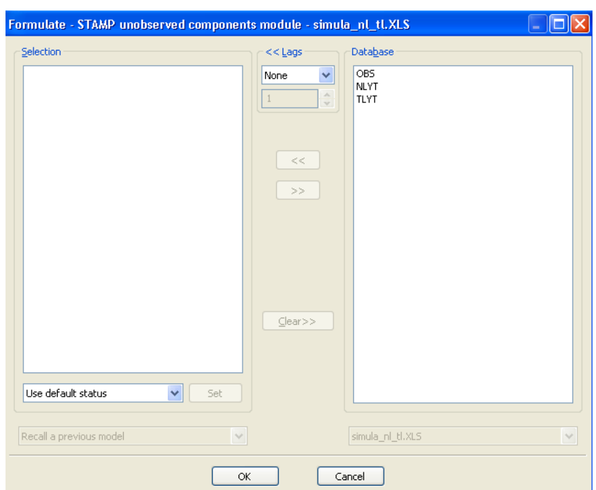


Figura 21: Tela do STAMP no OxMetrics

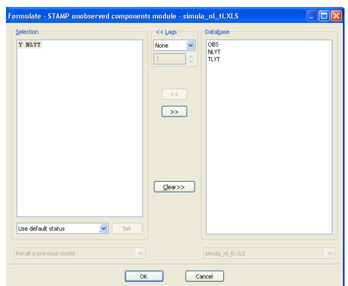


e clicando na variável a ser modelada na caixa `Database` o botão será ativado e aparecerá a seguinte caixa de diálogo

Após clicar em `OK` aparecerá a caixa de diálogo que especificação o modelo

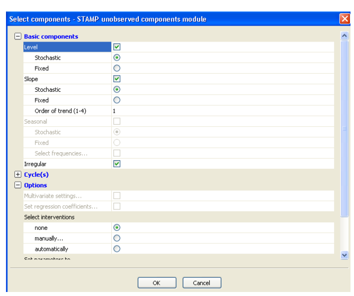
Figura 22: Botão de Seleção de Variávies no STAMP

Figura 23: Variável NL foi selecionada no STAMP



estrutural de séries temporais, isto é

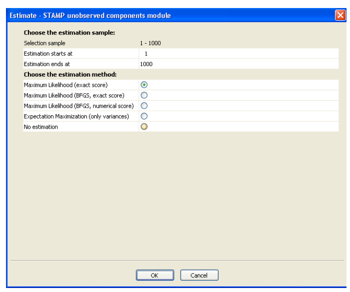
Figura 24: Modelo Estrutural para a Variável NL



Neste caso o *default* é um modelo de tendência local (`Level-stochastics; Slope-stochastics)` com erro observacional (`Irregular).`

Após clicar em `OK` aparecerá a caixa de diálogo que especifica a amostra a ser usada e o método de estimação, Filtro de Kalman (Máxima Verossimilhança Exata (`Maximum Likelihood (exact score)`, `Maximum Likelihood (BFGS`[10] `exact score)` e `Maximum Likelihood (BFGS numerical score))` e Estimação só das variâncias `(Expectation Maximization (only variances)).`

Figura 25: Metodo de Estimação do Modelo Estrutural para a Variável NL



Obtém-se os seguintes resultados

```
Summary statistics
 NLYT
Log-Likelihood is -2371.7714
Akaike Information Criterion (AIC): 4.7515
Bayesian Information Criterion (BIC): 4.7712
Prediction error variance is 6.7147
Summary statistics
 NLYT
 T          1000.0
 Normality 2.6757
 H(332)     092762
 DW         2.0245
 r(1)       -0.013094
 q         32.000
```

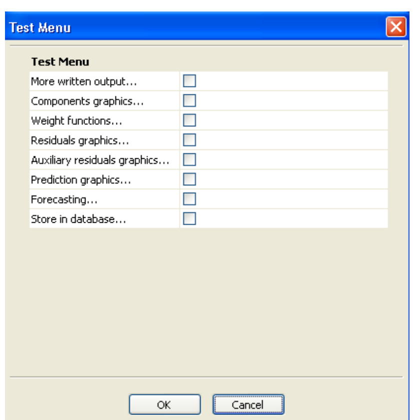
[10]BFGS é um método de estimação não linear do tipo Quase-Newton.

```
 p          2.0000
 r(q)       -0.027589
 Q(q,q-p) 23.447
 Rd^2       0.25937
 Variances of disturbances:
            Value    (q-ratio)
Level     1.20517  ( 0.3106)
Slope     0.000000 ( 0.0000)
Irregular 3.87967  ( 1.000)
State vector analysis at period 1000
        Value   Prob
Level 12.37842 [0.0000]
Slope 0.0108 [0.7515]
```

Observe que a variância da taxa de crescimento (`Slope`) é igual a zero e também que o valor do vetor de estado no último instante de tempo é igual a `0.01108` que não é estatísticamente significativo.

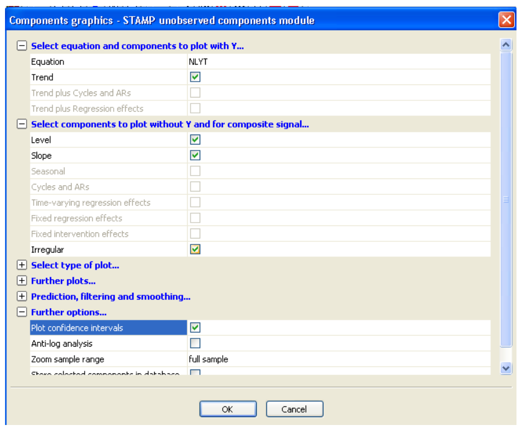
Para se obter as estimativas dos componentes do vetor de estado e do irregular temos clicar em `Model` e depois em `Test` obtendo a seguinte caixa de diálogo

Figura 26: Estimativa dos componentes do vetor de estado e do irregular para Modelo Estrutural da Variável NL



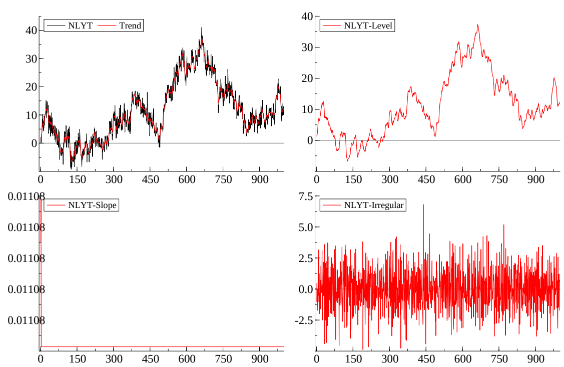
Clicando em `Components graphics`, obtem-se a seguinte caixa de diálogo

Figura 27: Gráfico dos componentes do vetor de estado e do irregular para Modelo Estrutural da Variável NL



Temos então os gráficos do série original e do componente de nível, o componente de nível com os intervalos de confiança, o componente de taxa de crescimento e o irregular.

Figura 28: Gráfico dos componentes do vetor de estado e do irregular para Modelo Estrutural da Variável NL usando um Modelo de Tendência Local

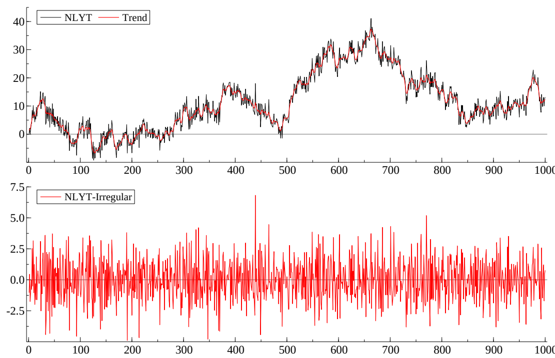


Como a taxa de crescimento é não significativa podemo simplificar o modelo para o modelo de nível local. Estimando o modelo de nível local, obtém-se as seguintes estimativas

```
Log-Likelihood is -2369.3806
Akaike Information Criterion (AIC): 4.7448
Bayesian Information Criterion (BIC): 4.7595
Prediction error variance is 6.7142
Summary statistics
 NLYT
 T         1000.0
 Normality 2.0186
 H(333)     092416
 DW         2.0225
 r(1)      -0.011657
 q         31.000
 p         1.0000
 r(q)       0.042742
 Q(q,q-p)  23.194
 R^2        0.93591
 Variances of disturbances:
        Value (q-ratio)
Level     1.19586 ( 0.3077)
Irregular 3.88588 ( 1.000)
State vector analysis at period 1000
       Value   Prob
Level 12.63074 [0.0000]
```

Observe que há uma redução do Schwartz indicando que este modelo é melhor do que o anterior. Os gráficos dos componentes para este modelo são apresentados abaixo

Figura 29: Gráfico dos componentes do vetor de estado e do irregular para Modelo Estrutural da Variável NL usando um Modelo de Nível Local



3.3.4 Estimando o Modelo de Volatilidade Estocástica para IBOVESPA no OxMetrics usando o STAMP

3.3.5 Exercícios

1. Para o modelo $AR(1)$ obtenha a função de previsão e o erro médio quadrático de previsão.

2. Para o modelo $MA(1)$ obtenha a função de previsão e o erro médio quadrático de previsão.

3. Para o modelo de nível local obtenha a função de previsão e o erro médio quadrático de previsão.

4. Para o modelo de tendência local obtenha a funçao de previsão e o erro médio qudarático de previsão.

5. Escreva um programa no R para obter o previsor pontual assim como o erro médio quadrático de previsão para um modelo $AR(1)$.

6. Escreva um programa no R para obter o previsor pontual assim como o erro médio quadrático de previsão para um modelo $MA(1)$.

7. Como os dados do PIB Brasileiro de 1861 até 2008 que estão no arquivo pib1861_2008.xlsx, estime um modelo de nível local e também um modelo de tendência local. Qual dos dois você escolheria e porque?

8. No arquivo bolsa.xlsx use os dados de fechamento do Ibovespa e estime um modelo de nível e de tendência local. Qual dos dois você escolheria e porque?

4 Sazonalidade

4.1 Introdução

Sazonalidade é um padrão que se repete a cada s períodos, onde s é a frequênica sazonal. Podemos pensar que a função de previsão tem a propriedade de somar zero a cada s períodos. Este componente pode ser determinísticos ou estocástico, estacionário ou não estacionário.

A forma mais simples de definir sazonalidade é através da sazonalidade determinística que pode ser representada por dummies sazonais ou por funções trigonométricas em frequências sazonais.

Outra forma de definir sazonalidade é generalizar os modelos $ARIMA$ de Box-Jenkins de forma a captar estes padrões que se repetem a cada s períodos. E uma outra forma é através do modelos estruturais que transformar a sazonalidade determinística em estocástica permitindo que estas variem suavemente. A seguir apresentamos as três abordagens.

4.2 Modelos $ARMA$ Sazonais

Na metodologia de Box-Jenkins sazonalidade pode ser estacionária ou não estacionária. No caso estacionário temos modelos do tipo AR, MA ou $ARMA$ onde o padrão das autocorrelações dependem da frequência dos dados. Por exemplo podemos considerar o modelo $AR(4)$ definido da seguinte forma:

$$y_t = \phi_4 \; y_{t-4} + \varepsilon_t \tag{66}$$

Para que este modelo seja estacionário necessitamos da condição:

$$|\phi_4| < 1$$

e também que a origem dos tempos é em $t = -\infty$ ou então que a distribuição não condicional de y_1 é dada por:

$$y_1 \sim N\left(0, \frac{\sigma^2}{1-\phi_4^2}\right)$$

Observe que a função de autocorrelação neste caso é dada por:

$$\rho(\tau) = \begin{cases} \phi_4^\tau & \text{se } \tau = 0, 4, 8, ... \\ 0 & \text{caso contrário} \end{cases} \tag{67}$$

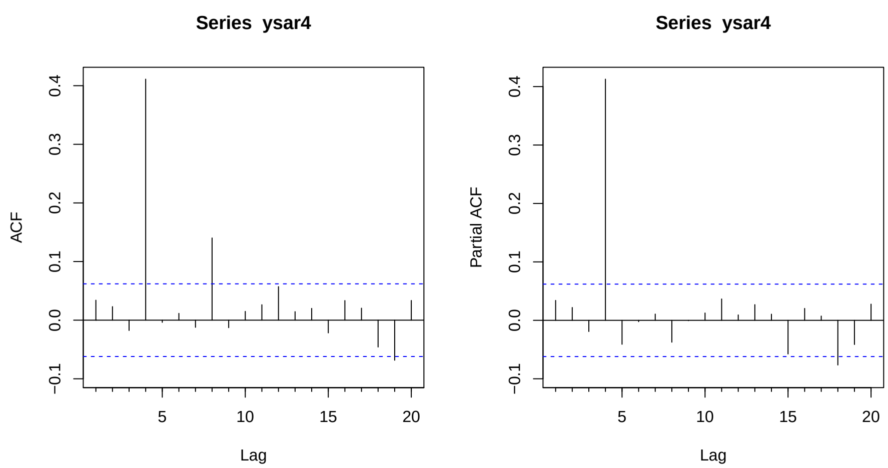
Como o modelo é estacionário, temos que os padrões sazonais decaem ao longo do tempo e portanto os efeitos sazonais não são determinísticos.

O programa abaixo gera um $AR(4)$ sazonal, faz a FAC e FACP para até 12 defasagens e estima em representação em espaço de estado o $AR(4)$, com $\phi_4 = 0.4$.

AR(4) Sazonal usando pacote astsa

O correlograma é apresentado abaixo e podemos observar uma FAC que tem o padrão de (67) e a FACP sómente a quarta autocorrelação parcial é diferente de zero implicando num modelo $ARMA(1,0)_4$.

Figura 30: FAC e FACP para AR(4) Sazonal



Abaixo é apresentada a estimação deste modelo por Filtro de Kalman.

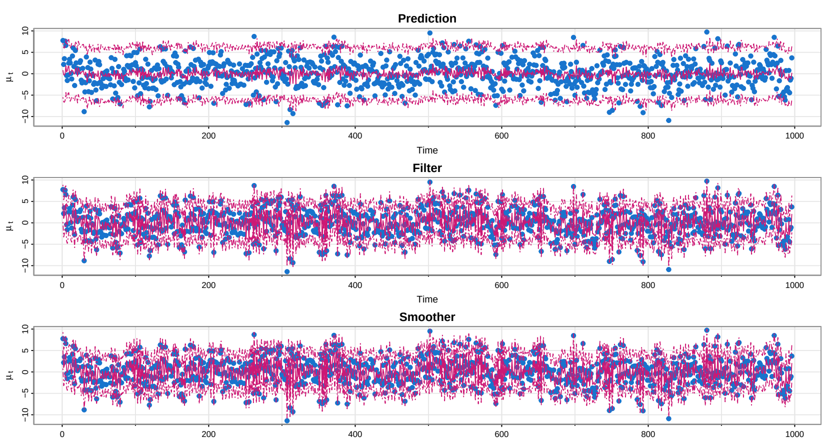
Observe que $\hat{\sigma}_\omega^2$ é a estimativa da variância do choque que não é próxima do valor verdadeiro que é um e a estimativa do coeficiente do autorregressivo que é próximo do valor verdadeiro que é $0,4$.

Tabela 11: Modelling Seasonal AR(4) Model using Kalman Filter

EQ(11) Modelling Seasonal AR(4) Model by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
ϕ_4	0.4114	0.0289	14.2288	0.0000
σ^2_ω	2.9778	0.06668	44.6558	0.0000
no of observations: 1000				
no of parameters: 2				
AIC = -3.1722				
BIC = -3.1013				
log-likelihood = 1592.101				

A seguir são apresentados os componentes previsto, filtrado e suavizado para o $AR(4)$ sazonal.

Figura 31: Componentes Previsto, Filtrado e Suavizado para o $AR(4)$ Sazonal



Para um modelo MA sazonal, isto é, um $MA(4)$ a especificação do modelo é dada por:

$$y_t = \varepsilon_t + \theta_4 \ \varepsilon_{t-4} \tag{68}$$

neste caso a função de autocorrelação é dada por:

$$\rho(\tau) = \begin{cases} 1 \text{ se } \tau = 0 \\ \frac{\theta_4}{1-\theta_4^2} \text{ se } \tau = 4 \\ 0 \text{ caso contrário} \end{cases} \tag{69}$$

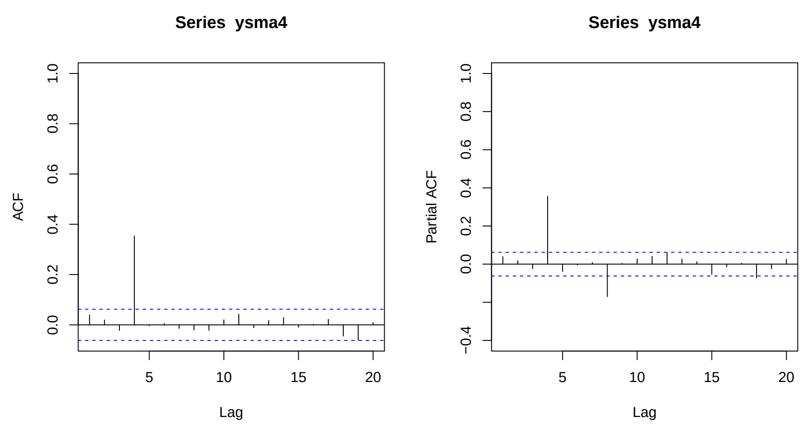
da mesma forma que no modelo anterior o padrão sazonal é zero após a quarta defasagem e portanto os efeitos sazonais não são determinísticos.

O programa abaixo gera um $MA(4)$ sazonal, faz a FAC e FACP para até 12 defasagens e estima em representação em espaço de estado o $MA(4)$, com $\theta_4 = 0.4$.

MA(4) Sazonal usando pacote astsa

O correlograma é apresentado abaixo e podemos observar uma FAC que tem o padrão de (69) e a FACP decae exponencialmente sendo que a quarta e oitava autocorrelações parciais são diferente de zero implicando num modelo $ARMA(0,1)_4$.

Figura 32: FAC e FACP para o $MA(4)$ Sazonal



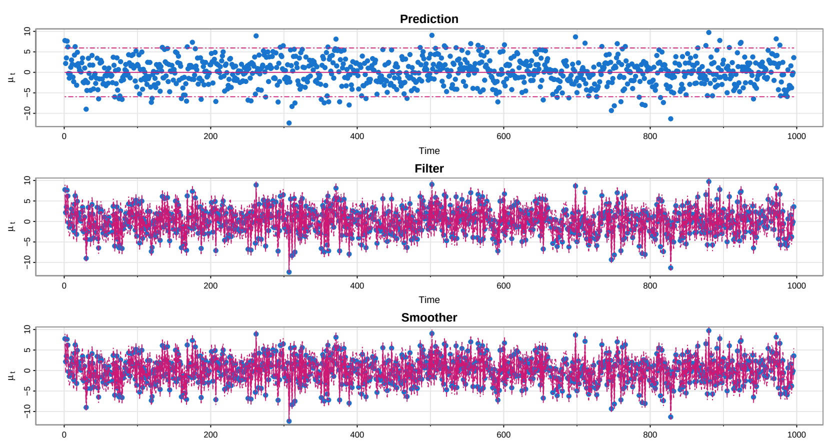
Abaixo é apresentada a estimação deste modelo por Filtro de Kalman. Observe que $\hat{\sigma}^2_\omega$ é a estimativa da variância do choque que é próxima do valor verdadeiro que é um e $\hat{\theta}_4$ representa a estimativa do coeficiente da média-móvel que é próximo do valor verdadeiro que é $0,4$.

Tabela 12: Modelling Seasonal MA(4) Model using Kalman Filter

EQ(12) Modelling Seasonal AR(4) Model by Kalman Filter				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
θ_4	0.43448	0.1531	2.8401	0.0022
σ^2_ω	1.2933	0.0907	14.2492	0.0000
no of observations: 1000				
no of parameters: 2				
AIC = -3.1812				
BIC = -3.1575				
log-likelihood = 1592.582				

A seguir são apresentados os componentes previsto, filtrado e suavizado para o $MA(4)$ sazonal.

Figura 33: Componentes Previsto, Filtrado e Suavizado para o $MA(4)$ Sazonal



O modelo estacionário geral, isto é, um $ARMA(P,Q)_s$ é dado por:

$$\Phi(L^s)\ y_t = \Theta(L^s)\ \xi_t \tag{70}$$

onde

$$\Phi(L^s) = (1 - \Phi_1\ L^s - ... - \Phi_P\ L^{Ps})$$

e

$$\Theta(L^s) = (1 + \Theta_1\ L^s + ... + \Theta_Q\ L^{Qs})$$

e s é o período da sazonalidade que em geral é 4 ou 12, se os dados são trimestrais ou mensais respectivamente.

Se ξ_t for um ruído branco o modelo é puramente sazonal, mas se ξ_t for uma $ARMA(p,q)$, isto é:

$$\xi_t = \frac{\Theta(L)}{\Phi(L)}\varepsilon_t \tag{71}$$

então o modelo geral é:

$$\Phi(L^s)\ \Phi(L)y_t = \Theta(L^s)\Theta(L)\varepsilon_t \tag{72}$$

que é denotado por $ARMA(p,q) \times (P,Q)_s$.

Uma forma alternativa de especificar o modelo $ARMA(p,q) \times (P,Q)_s$ é através de componentes não observados, embora a forma reduzida não seja equivalente a (72). Neste caso a série é decomposta como a soma de dois componentes sendo um não sazonal modelado por um $ARMA(p,q)$ e outro sazonal modelado por um $ARMA(P,Q)_s$, temos então:

$$y_t = \gamma_t + \upsilon_t \tag{73}$$

onde $\nu_t \sim ARMA(p,q)$ e $\gamma_t \sim ARMA(P,Q)_s$.

Nesta representação os componentes, sazonais e não sazonais, são aditivos enquanto que em (72) eles são multiplicativos.

4.3 Modelos $ARIMA$ Sazonais

Os modelos apresentados na seção anterior eram todos estacionário, mas em geral séries economicas ou financeiras são não estacionárias. No caso de modelos sazonais a fonte de não estacionaridade pode ser na frequência sazonal, isto é, a raíz unitária é no polinomio L^s, e podemos definir a diferença sazonal pelo operador

$$\Delta_s = (1 - L^s) \tag{74}$$

Uma série y_t que tenha uma raíz unitária sazonal, se tornará estacionária após a aplicação do operador definido em (74). Um caso particular é apresentado a seguir

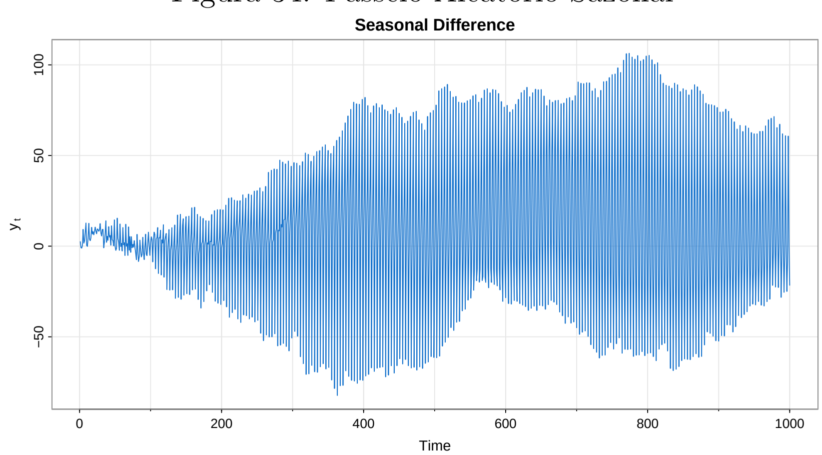
$$\begin{aligned} \Delta_s y_t &= \varepsilon_t \\ y_t &= y_{t-s} + \varepsilon_t \end{aligned} \tag{75}$$

Para gerar este "passeio aleatório sazonal", temos os seguintes comandos no R assumindo que $s = 4$

Passeio Aleatório Sazonal

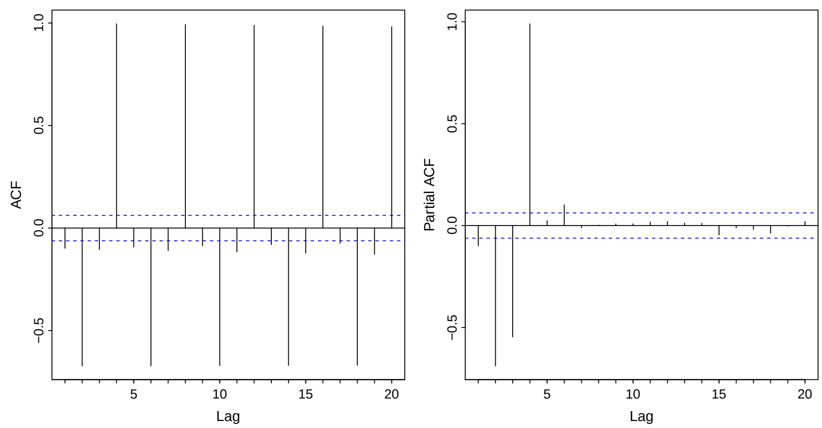
O gráfico desta série é apresentado a seguir.

Figura 34: Passeio Aleatório Sazonal



A FAC e FACP para esta série é apresentada a seguir

Figura 35: FAC e FACP do Passeio Aleatório Sazonal



E observe que a FAC tem autocorrelações significativas em múltiplos de 4 começando com $0,987$ com decaimento lento e para a $FACP$ sómente a quarta é significativa.

Então (70) pode ser generalizado para um modelo $ARIMA(P, D, Q)_s$ que tem a seguinte expressão

$$\Phi(L^s)\ (1 - L^s)^D y_t = \Theta(L^s)\ \xi_t$$

onde os polinômios AR e MA sazonais foram definidos acima é a $D - ésima$ diferença sazonal é dada por

$$(1 - L^s)^D = \underbrace{(1 - L^s) \times (1 - L^s) \cdots \times (1 - L^s)}_{D \text{ vezes}}$$

e a generalização de (72) permitindo uma ordem de integração não sazonal de d será o modelo $ARIMA(p, d, q) \times (P, D, Q)_s$ cujo processo gerador é dado por:

$$\Phi(L^s)\ \Phi(L)(1 - L)^d(1 - L^s)^D y_t = \Theta(L^s)\Theta(L)\varepsilon_t$$

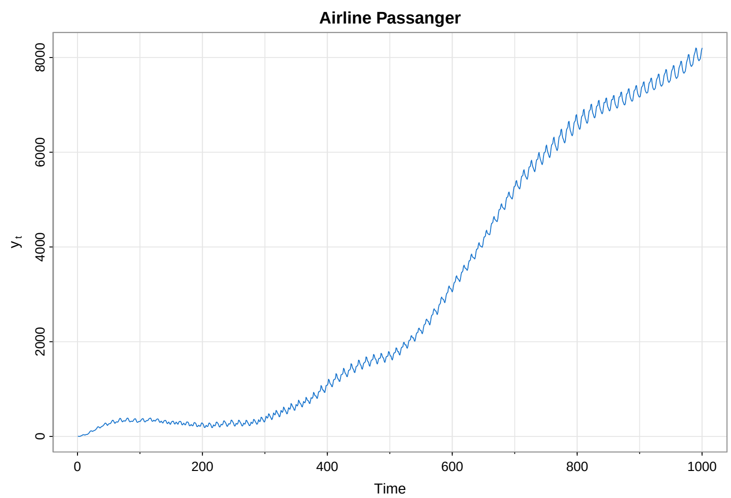
Um caso particular é o modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_s$ que é conhecido na literatura de pelo nome de *"AIRLINE PASSANGER MODEL"*, veja Box and Jenkins (1970).

O programa abaixo simula um modelos $ARIMA(0,1,1) \times (0,1,1)_{12}$ com $\theta_1 = 0.5$ e $\Theta_1 = 0.8$

Modelo Airline Passanger

O gráfico a seguir apresenta a série gerada segundo o modelo *Airline Passanger.* Observe que esta série é claramente não estacionária, mas não é possível saber qual a ordem de integração.

Figura 36: Airline Passanger



Após diferenciar a série, fizemos a FAC e $FACP$ e observamos que a série parece ter uma raíz unitária em frequencia sazonal, uma vez que a FAC na defasagem 12 é muito perto de um e o decaimento em multiplo de 12 é bem lento. Desta forma vamos diferenciar a série usando a diferença sazonal.

O gráfico contra o tempo da série $(1-L)(1-L^{12})y_t$ é apresentado abaixo

e parece que está série é estacionária.

A seguir é apresentado a FAC e $FACP$ da série $(1-L)(1-L^{12})y_t$

Observe que agora a série apresenta um comportamento estacionário. Pelas primeiras autocorrelações temos uma indicação de que a parte não sazonal segue um $MA(1)$. Ao se observar a FAC e $FACP$ para os múltiplos de 12 temos uma indicação de que a série tem um comportamento média-móvel sazonal de primeira ordem. Deste modo o modelo que parece adequado a estas observações é um $ARIMA(0,1,1) \times (0,1,1)_{12}$.

Figura 37: FAC e FACP DAirline

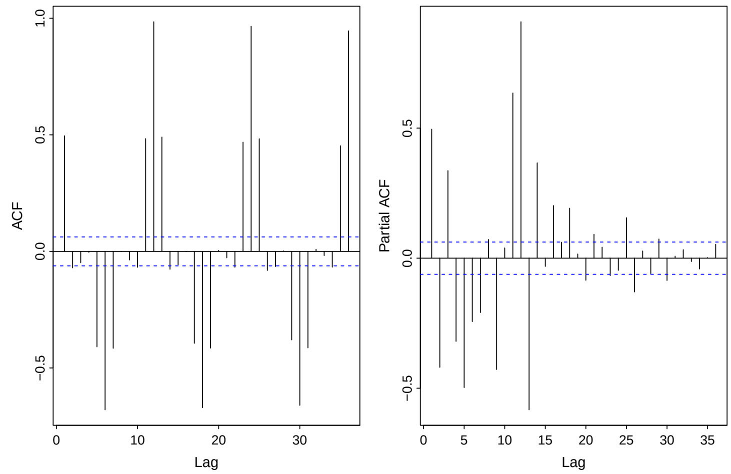
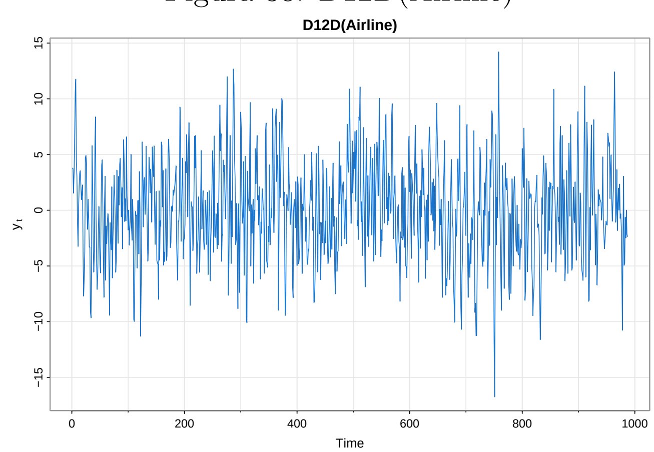


Figura 38: D12D(Airline)



A seguie é apresentada a estimativas dos parâmetros deste modelo

Figura 39: FAC e FACP de D12D(Airline)

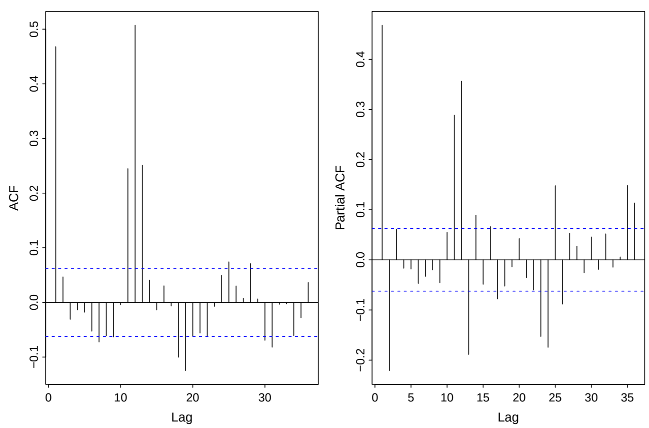


Tabela 13: Modelling Airline Passanger

EQ(13) Modelling Airline Passanger				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
θ_1	0.5181	0.0254	20.4215	0.0000
Θ_1	0.8048	0.0214	37.5207	0.0000
no of observations: 987				
no of parameters: 3				
$\sigma^2 = 8.767722$				
$AIC = 5.0280$				
$BIC = 5.0429$				
log-likelihood = -2467.987				

4.4 Sazonalidade Determinística

Historicamente sazonalidade era definida através de variáveis binárias, definidas da seguinte forma:

$$z_{jt} = \begin{cases} 1 \text{ se } t = j,\ j+s,\ j+2s, ... \\ 0 \text{ se caso contrário} \end{cases}$$

e a extração da sazonalidade era feita através de uma regressão nesta dummies, isto é, o seguinte modelo de regressão era ajustado ao dados:

$$y_t = \sum_{j=1}^{s} \gamma_j \ z_{jt} + \varepsilon_t \tag{76}$$

ou então:

$$y_t = \mu + \sum_{j=1}^{s-1} \gamma_j \ z_{jt} + \varepsilon_t \tag{77}$$

A diferença entre as duas especificações é que na primeira temos os efeitos de cada período sazonal em separado e da segunda através da constante temos o efeito global e através das outras $s-1$ dummies os acréscimos ou decréscimos aos $s-1$ períodos sazonais.

O problema com esta especificação é que os padrões sazonais não somam zero a cada s períodos. Uma especificação alternativa que satisfaz esta propriedade é apresentada a seguir. Defina as dummies centradas da seguinte forma:

$$z_{jt} = \begin{cases} 1 \text{ se } t = j, \ j+s, \ j+2s, ... \\ -1 \text{ se } t = s, \ 2s, \ 3s, ... \\ 0 \text{ se caso contrário} \end{cases} \tag{78}$$

e a especificação usada é a (77).

Uma forma equivalente para a sazonalidade determinística é a utilização de funções trigonométricas em frequências sazonais. O modelo de regressão é definido por:

$$y_t = \sum_{j=1}^{[s/2]} \left(\gamma_j \cos \lambda_j t + \gamma_j^* \sin \lambda_j t \right) + \varepsilon_t \tag{79}$$

com $\lambda_j = \frac{2\pi j}{s}$ para $j = 1, ..., \left[\frac{s}{2}\right]$.

4.5 Componente Sazonal no Modelo Estrutural

Na metodologia de Box & Jenkins uma forma de extrair a sazonalidade é através do operador diferença sazonal, isto é, $(1 - L^s)\gamma_t$. Uma forma de permitir que este componente varie suavemente é adicionar um choque, um ruído branco, a esta especificação.Temos então a seguinte especificação para o componente sazonal:

$$(1 - L^s)\gamma_t = \omega_t \Leftrightarrow \gamma_t = \gamma_{t-s} + \omega_t \tag{80}$$

e a equação de observação na representação em espaço de estado é dada por:

$$y_t = \gamma_t + \varepsilon_t \tag{81}$$

Esta especificação tem problemas uma vez que o operador $(1-L^s)$ mistura nível com componente sazonal uma vez que $(1 - L^s) = (1 - L)(1 + L + L^2 + ... + L^{s-1})$ e o primeiro operador extrai a tendência e o segundo tem a propriedade de somar zero a cada s períodos que é uma característica do componente sazonal. O segundo problema com o operador $(1 - L^s)$ é não somar zero a cada s períodos o que faz com que esta especificação não tenha as características desejáveis para um componente sazonal.

Uma forma de restringir os fatores a somarem zero a cada s períodos é especificar o componente sazonal da seguinte forma:

$$\gamma_t + \gamma_{t-1} + ... + \gamma_{t-s+1} = 0 \tag{82}$$

e permitindo que (82) varie suavemente temos que esta especificação pode ser re-escrita da seguinte forma:

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \Leftrightarrow \gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \tag{83}$$

Observe que $\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \Leftrightarrow S(L)\gamma_t = \omega_t$ e é este operador, $S(L)$, que garante que os fatores somem zero a cada s períodos.

Observe que a função de previsão deste modelo é dada por:

$$\gamma_{T+l \mid T} = -\sum_{j=1}^{s-1} \gamma_{T+l-1 \mid T} \tag{84}$$

que tem a propriedade de somar zero a cada s períodos. Observe também que esta função de previsão irá se repetir a cada s períodos e será constante, isto é, o melhor previsor dos fatores sazonais são os últimos fatores.

Por exemplo se o Processo Gerador dos Dados (P.G..D.) fosse um modelo de nível local + dummies estocásticas teriamos a seguinte representação em espaço de estado:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{85}$$

$$\mu_t = \mu_{t-1} + \eta_t \tag{86}$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \tag{87}$$

Observe que a não estacionaridade de y_t é decorrência da não estacionaridade de μ_t vinda do operador $(1-L)$ e da não estacionaridade de γ_t vinda do operador $S(L)$ e combinando os dois operadores teremos que $(1-L)S(L) = (1-L^s)$.

A representação (85-87) em forma matricial é dada por:

$$y_t = (1:1:0_{s-2})\alpha_t + \varepsilon_t \tag{88}$$

$$\alpha_t = \begin{bmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ . \\ . \\ . \\ \gamma_{t-s+2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & -1 & -1 & . & . & . & -1 \\ 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \gamma_{t-1} \\ \gamma_{t-2} \\ . \\ . \\ . \\ \gamma_{t-s+1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \omega_t \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} \tag{89}$$

É fácil mostrar que a forma reduzida é dada por:

$$(1-L)S(L)y_t = S(L)\eta_t + (1-L)\omega_t + (1-L)S(L)\varepsilon_t \tag{90}$$

onde o lado direito de (90) pode ser escrito como um $MA(s)$ e, portanto, a forma reduzida para y_t é um $ARIMA(0,1,1)_s$ mas com o MA sazonal tendo coeficientes não nulos para defasagens menores do que s.

Caso o P.G.D. fosse um modelo de tendência local + dummies estocásticas a forma reduzida seria $\Delta\Delta_s y_t \sim MA(s+1)$ que é semelhante ao modelo $ARIMA(0,1,1) \times (0,1,1)_s$ que é o "*AIRLINE PASSANGER MODEL* "de

Box and Jenkins (1970), que é considerado um dos modelos mais parcimoniosos para dados sazonais.

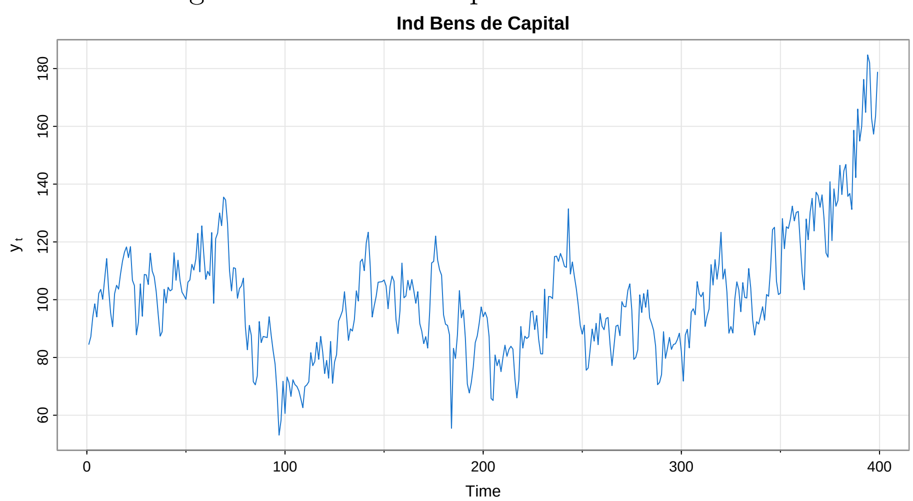
4.5.1 Usando o R para Estimar o Modelo Nível Local + Sazonalidade Estocástica Bens de Capital

O programa abaixo estima o modelo de nível local, sazonalidade estocástica e Irregular para a série de Indústria Bens de Capital pa(beara o Brasl.

BCAP = Level+Seasonal+Irregular

A série de logaritmo da Produção Industrial Bens de Capital para o período de janeiro de 1975 até março de 2008 é apresentada abaixo

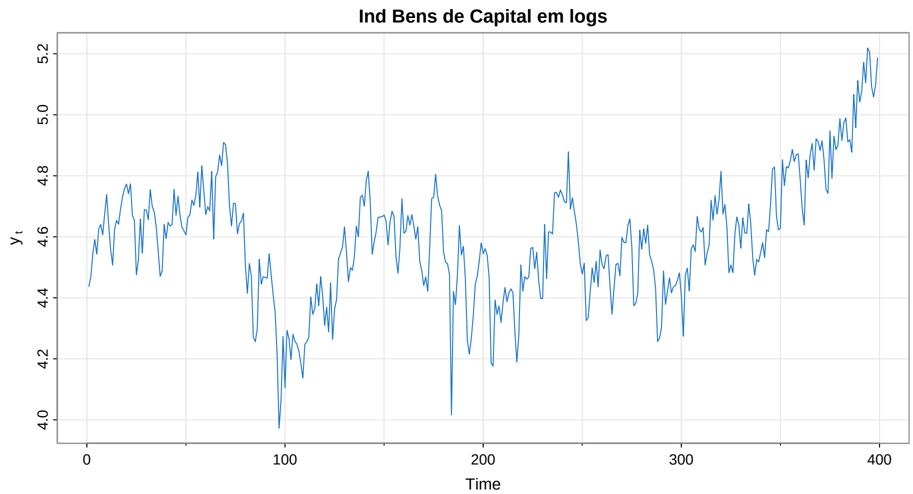
Figura 40: Bens de Capital - Indústria Brasil



Observe que esta série tem nível que varia, portanto um modelo de nível local deve ser adequado para modelar a tendência da série e parece existe uma sazonalidade que está também mudando ao longo do tempo.

Para diminuir a variabilidade, será utilizada a transformção logarítmica. Temos então o seguinte gráfico.

Figura 41: Bens de Capital - Indústria Brasil em logs

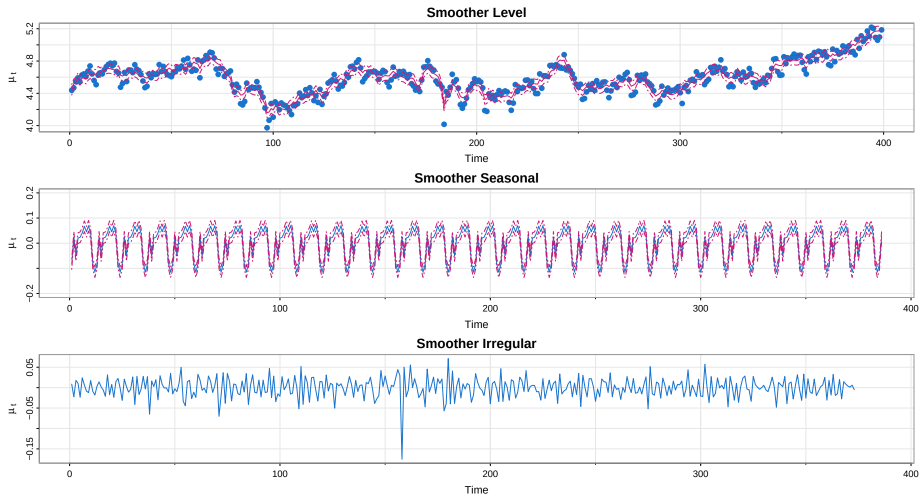


Estimando este modelo temos os seguintes resultados

Tabela 14: Local Level+Seasonal+ Irregular for LBCAP

EQ(14) Local Level+Seasonal+ Irregular for LBCAP				
The estimation sample is: 1 - 399				
	Coefficient	Std. Error	t-Statitistics	p-value
σ_η	0.0490	0.0044	11.1010	0.0000
σ_ω	0.0000	0.0016	0.0000	0.5000
σ_ε	0.0382	0.0033	11.5110	0.0000
no of observations: 399				
no of parameters: 3				
AIC = 3.0843				
BIC =3.1593				
log-likelihood = -612.3204				

Figura 42: Componentes Estimados



5 Componente Cíclico

5.1 Componente Determinístico

O componente cíclico é aquela parte do modelo que representa flutuações de longo prazo. Como estas flutuações podem ser periódicas, podemos, em primeira aproximação, assumir que este componente pode ser representado por uma função periódica, por exemplo, um conseno. Temos, então:

$$y_t = \cos(\lambda t) \tag{91}$$

onde λ representa a freqüência e o período é dado por $\frac{\lambda}{2\pi}$.

Uma representação alternativa a (91) é aquela em que esta função periódica tem uma amplitude e uma fase. Então (91) pode ser reescrita da seguinte forma:

$$y_t = \rho \cos(\lambda t - \theta) \tag{92}$$

onde ρ representa a amplitude e θ a fase.

Observe que (92) pode ser reescrito da seguinte forma[11]:

$$y_t = \alpha \cos(\lambda t) + \beta \sin(\lambda t) \tag{93}$$

[11]Usando a seguinte igualdade trigonométrica: $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$.

com $\alpha = \rho\cos(\theta)$ e $\beta = \rho\sin(\theta)$.

Podemos então especificar o componente cíclico da seguinte forma:

$$y_t = \psi_t + \varepsilon_t \quad \text{para} \quad t = 1, ..., T \tag{94}$$

e

$$\psi_t = \alpha\cos(\lambda_c\ t) + \beta\sin(\lambda_c\ t) \tag{95}$$

onde $(\alpha^2 + \beta^2)^{1/2}$ é a amplitude e $\tan^{-1}(\frac{\beta}{\alpha})$ é a fase.

Se λ_c for conhecido os parâmetros de (94-95) podem ser estimados por M.Q.O., como no modelo linear de regressão já que os regressores são não estocásticos. Observe que (94-95) é uma representação de um componente cíclico determinístico.

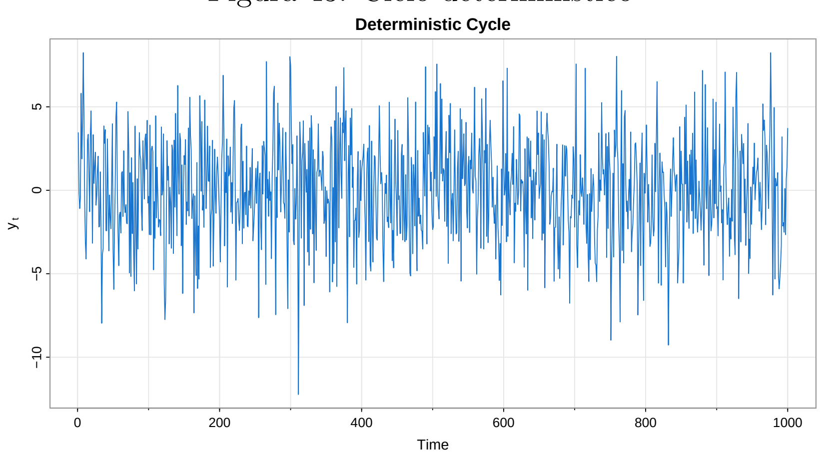
5.1.1 Programa para Gerar um Componente Cíclico Determinístico

Abaixo é apresentado um programa no R para gerar este modelo, com $\alpha = 0.6618$, $\beta = 0.6815$ e $\lambda = 0.8$.

Ciclo Determinístico

O gráfico da ciclo determinístico é apresentado a seguir

Figura 43: Ciclo determinístico



A estimação dos parâmetros é feita por mínimos quadrados não linear. Temos os eguintes resultados

Tabela 15: Estiamção do Cíclo determinístico

EQ(15) Deterministic Cycle				
The estimation sample is: 1 - 1000				
	Coefficient	Std. Error	t-Statitistics	p-value
α	0.8691	0.1827	4.758	0.0000
β	0.5598	0.2353	2.379	0.0175
λ	0.7994	0.0004	1791.089	0.0000
no of observations: 1000				
no of parameters: 3				
$\sigma_\varepsilon = 2.976$				
$AIC = 2186.959$				
$BIC = 2222.405$				

5.2 Componente Cíclico Estocástico

Podemos permitir que este componente cíclico determinístico evolua ao longo do tempo, isto é, α e β variem da seguinte forma:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} \cos(\lambda_c) & \sin(\lambda_c) \\ -\sin(\lambda_c) & \cos(\lambda_c) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} \text{ para } t = 1, ..., T \tag{96}$$

Observe que a forma reduzida de (96) é dada por:

$$\psi_t = 2\cos(\lambda_c)\psi_{t-1} - \psi_{t-2} \tag{97}$$

e quando $\alpha = \psi_0$ e $\beta = \psi_0^*$ temos que (97) é equivalente a (95).

Uma forma de permitir que este componente seja estocástico é introduzir dois choques em (96) não correlacionados entre si mas com a restrição de que a variância é igual. Observe que pela forma reduzida dada em (97) o componente cíclico seria não estacionário uma vez que uma das raízes do polinômio de segunda ordem é unitária. Para tornar este componente estacionário vamos introduzir um fator de decaimento, ρ, de tal sorte que $0 \leq \rho < 1$, temos então:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos(\lambda_c) & \sin(\lambda_c) \\ -\sin(\lambda_c) & \cos(\lambda_c) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{98}$$

onde κ_t e κ_t^* são dois ruídos brancos independentes com mesma variância σ_κ^2. Mas é mais conveniente reparametrizar de tal sorte que a variância do ciclo σ_ψ^2 fique constante e $\sigma_\kappa^2 = (1-\rho^2)\sigma_\psi^2$, que se reduzz a um ciclo determinístico quando $\sigma_\kappa^2 \to 0$ quando $\rho \to 1$.

A forma reduzida de (98) é dada por:

$$\psi_t - 2\rho\cos(\lambda_c)\psi_{t-1} + \rho^2 \ \psi_{t-2} \ = \kappa_t - \cos(\lambda_c)\kappa_{t-1} + \sin(\lambda_c)\kappa_{t-1}^* \qquad (99)$$

onde o lado direito de (99) representa um processo $MA(1)$, logo temos que a forma reduzida é um $ARMA(2,1)$ onde as raízes da parte autorregressiva são complexas conjugadas com módulo igual a ρ.

Agora a forma reduzida para (94) e (98) é dada por:

$$y_t - 2\rho\cos(\lambda_c)y_{t-1} + \rho^2 \ y_{t-2} \ = (1-2\rho\cos(\lambda_c)L + \rho^2\cos(\lambda_c)L^2)\varepsilon_t + (1-\cos(\lambda_c)L)\kappa_t + \sin(\lambda_c)\kappa_{t-1}^* \qquad (100)$$

onde o lado direito de (100) representa um processo $MA(2)$, logo temos que

a forma reduzida é um $ARMA(2,2)$ estacionário.

Observe que (94) e (98) já estão em representação em espaço de estado e os parâmetros podem ser estimados pelo Filtro de Kalman.

A função de previsão para o componente cíclico é dada por:

$$\psi_{T+l|T} = \rho^l(\cos(\lambda_c l)\psi_T + \ \sin(\lambda_c l)\psi_T^*) \text{ para } l = 1, 2, ... \qquad (101)$$

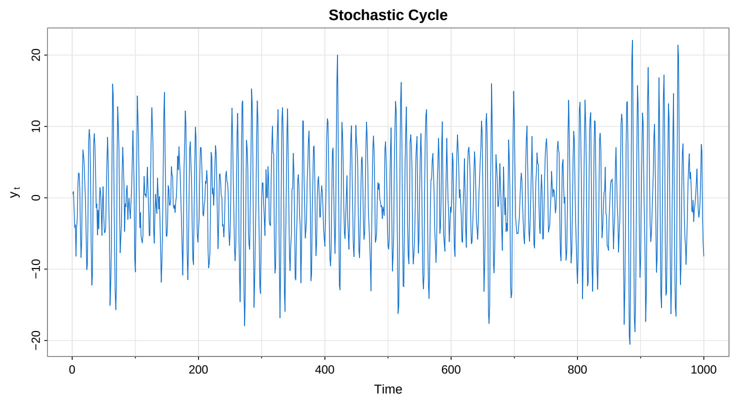
5.2.1 Programa para Gerar um Componente Cíclico Estocástico

Abaixo é apresentado um programa no R para gerar este modelo

Ciclo Estocástico

O gráfico a seguir apresenta o ciclo estocástico

Figura 44: Ciclo Estocástico

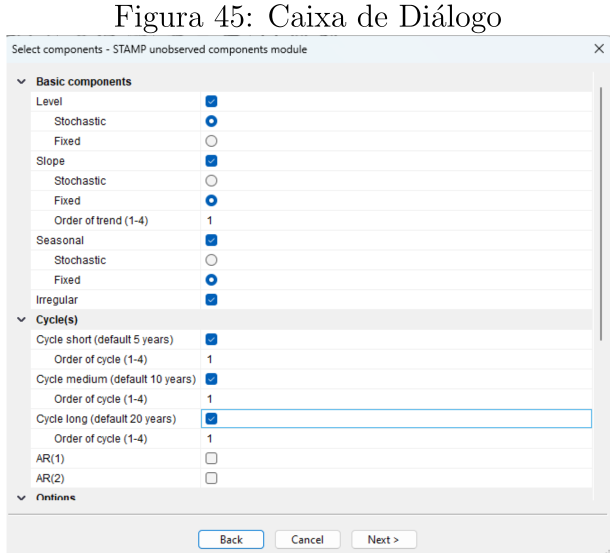


5.2.2 Usando o STAMP para Estimar o Modelo Nível Local + Sazonalidade Estocástica+Ciclo Estocástico para Bens de Capital

No STAMP é permitido três componentes cíclicos, a saber de curto, médio e longo prazos que estão associados aos períodos destes ciclos.

Vamos permitir que na série de Bens de Capital tenhamos os três ciclos. Na caixa de diálogo temos a seguinte especificação

Figura 45: Caixa de Diálogo



Temos as seguintes estimativas

```
    The model is:  Y = Trend + Seasonal + Irregular + Cycle 1
+ Cycle 2 + Cycle 3
Log-Likelihood is 466.1507.
Information criterion Akaike (AIC) -2.2213
...  Bayesian Schwartz (BIC) -1.9914
Prediction error variance is 0.00419822
Summary statistics
                   LBCAP
 T                 399.00
 Normality          97.286
 H(128)              0.7661
 DW                  2.0394
 r(1)               -0.0221
 q                  26.000
 p                   8.0000
 r(q)                0.10651
 Q(q,q-p)          67.085
```

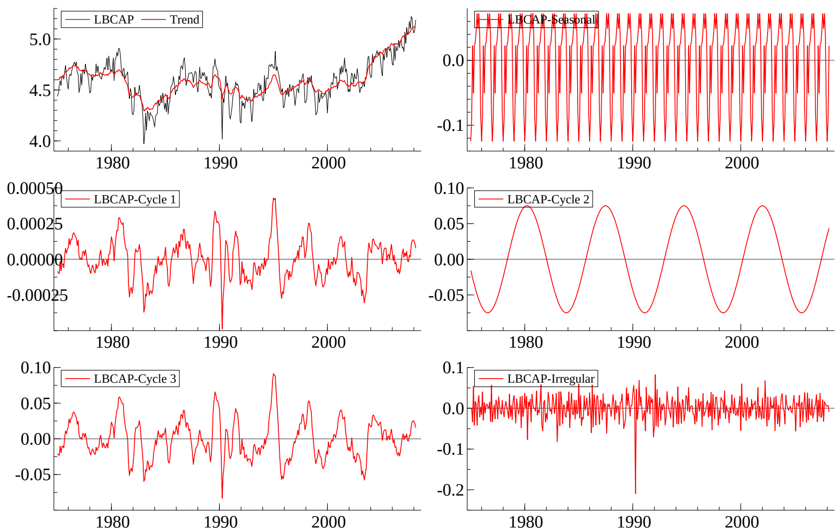
Rs^2 0.16755

```
 Variances of disturbances:
            Value        (q-ratio)
Level       0.00104559   ( 0.6447)
Slope       0.000000    ( 0.0000)
Seasonal    0.000000     ( 0.0000)
Cycle       3.83068e-06 ( 0.002362)
Cycle 2     7.42515e-09 (4.578e-06)
Cycle 3     0.000589452  ( 0.3634)
Irregular   0.00162185   ( 1.000)
Cycle other parameters:
Variance 0.00002
Period 164.9898
Period in years 13.74915
Frequency 0.03808
Damping factor 0.88029
Order 1.00000
Cycle 2 other parameters:
Variance 0.00371
Period 87.19035
```

```
Period in years 7.26586
Frequency 0.07206
Damping factor 1.00000
Order 1.00000
Cycle 3 other parameters:
Variance 0.00293
Period 52.23938
Period in years 4.35328
Frequency 0.12028
Damping factor 0.89386
Order 1.00000
State vector analysis at period 2008(3)
 Value Prob
Level 5.11385 [0.00000]
Slope 0.00129 [0.43141]
Seasonal chi2 test 353.04929 [0.00000]
Cycle 1 amplitude 0.01441 [ .NaN]
Seasonal effects:
 Period Value Prob
 1 -0.12481 [0.00000]
 2 -0.08913 [0.00000]
 3 0.02240 [0.02296]
 4 -0.05042 [0.00000]
 5 0.02241 [0.02393]
 6 0.02726 [0.00611]
 7 0.04551 [0.00001]
 8 0.07186 [0.00000]
 9 0.04918 [0.00000]
 10 0.07182 [0.00000]
 11 0.03690 [0.00022]
 12 -0.08298 [0.00000]
```

E os componentes estimados são apresentados abaixo

Figura 46: Componentes Nível + Sazonalidade + 3 Cíclos para LBCAP



6 Modelo Estrutural Básico

O modelo estrututal básico é aquele que decompõe a série em tendência estocástica, sazonalidade estocástica, cíclo estocástico e irregular.

O componente de tendência tanto pode ser tanto um modelo de nível local quanto um modelo de tendência local.

Para o componente de sazonalidade estocástica podemos ter tanto as dummies estocásticas apresentadas acima ou uma outra representação equivalente dada por:

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma^*_{j,t} \end{bmatrix} = \begin{bmatrix} \cos(\lambda_j) & \sin(\lambda_j) \\ -\sin(\lambda_j) & \cos(\lambda_j) \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma^*_{j,t-1} \end{bmatrix} + \begin{bmatrix} \omega_t \\ \omega^*_t \end{bmatrix} \tag{102}$$

para $j = 1, ..., [s/2]$ e o componente sazonal é dado por $\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}$.

O componente cíclico é dado por (98).

7 Modelo Estrutural com Variáveis Exogenas

Uma forma de introduzir variáveis "exógenas"nos modelos $ARMA$ é através dos modelos chamados de função de transferência, veja Box and Jenkins (1976). Uma especificação simples é apresentada a seguir:

$$y_t = \frac{\omega(L)}{\lambda(L)} x_t + \frac{\theta(L)}{\phi(L)\Delta^d} \varepsilon_t \Leftrightarrow \Delta^d y_t = \frac{\omega(L)}{\lambda(L)} \Delta^d x_t + \frac{\theta(L)}{\phi(L)} \varepsilon_t \tag{103}$$

Observe que neste modelo assim como nos modelos $ARIMA$ a informação do nível da série é perdida. Se $d = 2$ uma forma alternativa de especificar o modelo (103) é através do seguinte modelo estrutural:

$$\begin{aligned} y_t &= \mu_t + \delta x_t + \varepsilon_t \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + \xi_t \end{aligned} \tag{104}$$

uma vez que a forma reduzida de (104) é dada por:

$$\Delta^2 y_t = \delta \Delta^2 x_t + u_t \tag{105}$$

com

$$u_t = \Delta \eta_t + \xi_{t-1} + \Delta^2 \varepsilon_t \ \sim \ MA(2) \tag{106}$$

A diferença entre (103) e (104) é que neste último não se perde a informação do nível da série enquanto que no primeiro sim.

Veremos que um modelo alternativo ao de função de transferência é o Autorregressivo-Defasagens Distribuidas, veja Hendry (1995), que será apresentado nos modelos dinâmicos. Estes modelos tem a vantagem, por um lado, de possibilitar a incorporação de informações de curto e longo prazo, o que não acontece com os modelos de função de transferência e por outro lado, incorporam a dinâmica internamente ao modelo, isto é a parte não sistemática do modelo é um processo de inovação e nos modelos de função de transferência, parte da dinâmica é externa, isto é, os erros seguem um processo $ARMA$.

8 Valores Aberrantes e Mudanças Estruturais

Variáveis dummies podem ser usadas para controlar tanto valores aberrantes quanto mudanças estruturais. Valores aberrantes estão associados a eventos específicos que acontecem numa certa data e são em geral transitórios, por exemplo, greves, queda abrupta na produção de um produto devido a desastre. Mudanças estruturais por outro lado estão associadas a eventos permanentes, por exemplo choque do petróleo gerou uma quebra estrtutural nas economias tanto desenvolvidas quanto em desenvolvimento.

Uma forma de controlar valores aberrantes é a introdução de uma variável dummies de impulso na equação de observação onde esta variável irá assumir o valor um quando ocorrer o valor aberrantes e zero caso contrário.

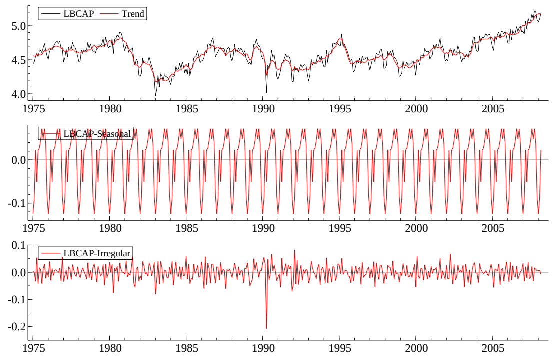
A mudança estrutural no nível da série é controlada utilizando-se uma variável do tipo degrau que assume o valor zero até a ocorrência da mudanças e após a mudança muda para de valor para um.

Uma forma de determinar se num dado período ocorreu um valor aberrante ou uma mudança estrutural, num contexto de modelos econométricos é através do procedimento proposto por Doornik (2009) denominado por Autometrics.

No contexto de modelos estruturais de séries temporais a forma de determinar valoes aberrantes ou mudanças estruturais é através dos resíduos auxiliares, veja Harvey and Koopman (1992), que são estimativas suavizadas do erro das observações e do erro do nível. No primeiro quando o resíduo for maior do que um certo número de desvios padrão este valor corresponde a valor aberrante e o segundo se o resíduo também for mais do que um certo núemro de desvios padrão este valor corresponde a uma mudança estrutural.

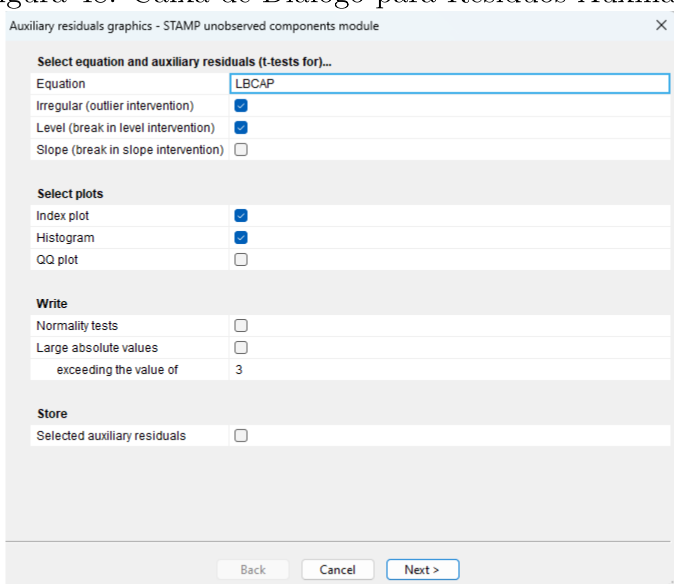
Para a série de Bens de Capital foi estimado o modelo de tendência local + sazonalidade + irregular. Os componentes suavizados são apresentados abaixo:

Figura 47: Componentes Tendência + Sazonalidade + Irregular para LBCAP



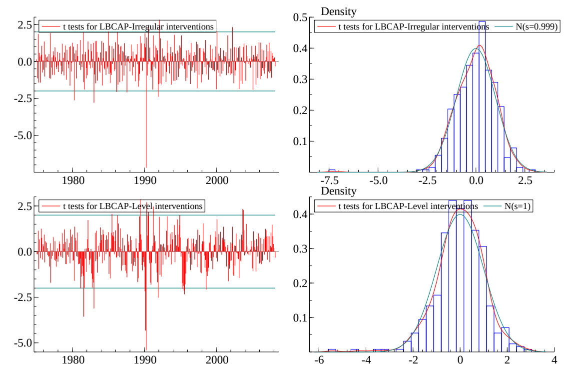
Para verificar se existem valores aberrantes ou mudanças estruturais nesta série, vamos para o menu de **Model, Test, Auxiliar Residuals graphics**. Temos então a seguinte caixa de diálogo

Figura 48: Caixa de Diálogo para Resíduos Auxiliares



Obtém-se então os seguintes gráficos

Figura 49: Caixa de Diálogo para Resí



Pelos gráficos da esquerda temos que para os resíduos do irregular, associados a valores aberrantes, um período que está fora da banda, a saber:

```
Values larger than 3 for BCAP-Irregular residual:
          Value     prob
1990(4) -7.1945 [0.00000]
```

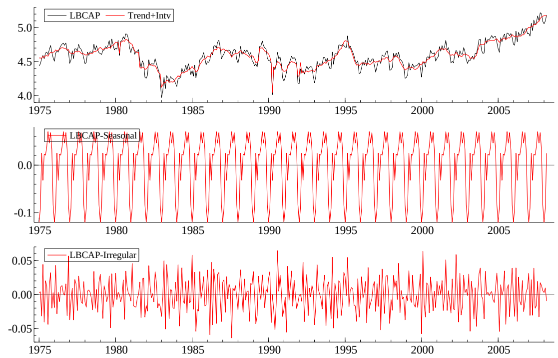
e para os resíduos do nível, associados a mudanças estruturais, temos quatro valores fora da banda.

```
Values larger than 3 for BCAP-Level residual:
          Value     prob
1981(8) -3.5664 [0.00000]
1983(1) -3.1124 [0.00099]
1990(3) -4.3240 [0.00000]
1990(4) -5.4074 [0.00000]
```

Observe que os gráficos da direita indicam que teste de normalidade seriam inadequados neste caso.

Agora re-estimando o modelo anterior mas ativando as intervenções automaticamente, temos os seguintes componentes estimados

Figura 50: Componentes com Intervenções para LBCAP



Re-estimando o modelo temos as seguintes intervenções que foram estimadas

Tabela 16: Intervenções Estimadas

Regression effects in final state at time 2008(3)					
		Coefficient	RMSE	t-value	Prob
Outlier	1980(4)	−0.16084	0.04739	−3.3940	[0.00076]
Outlier	1990(4)	−0.40852	0.04739	−8.6203	[0.00000]
Outlier	1992(2)	0.16890	0.04750	3.5556	[0.00042]
Level break	1981(8)	−0.22316	0.05273	−4.23203	[0.00003]
Level break	1983(1)	−0.19610	0.05274	−3.71848	[0.00023]
Level break	1989(6)	0.18343	0.05273	3.47842	[0.00056]
Level break	1991(12)	−0.17571	0.05291	−3.32126	[0.00098]

e observe que além do valor aberrante de 1990(4) foram identificados mais dois valores em 1980(4) e 1992(2). Quanto as mudanças estruturais foram mantidas as datas de 1981(8) e 1983(1), as datas 1990(3) e 1990(4) deixaram de ser identificadas e as datas 1989(6) e 1991(12) foram identificadas.

9 Apêndice I - Distribuição Condicional

Vamos supor que o vetor $z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}$ é um processo estocástico serialmente não correlacionado com distribuição normal. Podemos representar este vetor da seguinte forma:

$$z_t \sim NI\left(\begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{bmatrix}\right) \tag{107}$$

Observe que a distribuição conjunta depende do vetor de parâmetros $\theta = (\mu_y,\, \mu_x, \sigma_y^2, \sigma_x^2, \sigma_{yx})$ que tem dimensão cinco.

A distribuição conjunta sempre pode ser fatorada como condicional e marginal, isto é, podemos ter

$$D(z_t \,:\, \theta) = \begin{cases} D(y_t \mid x_t, \theta_1) D(x_t \mid \theta_2) \\ D(x_t \mid y_t, \psi_1) D(y_t \mid \psi_2) \end{cases} \tag{108}$$

onde as duas representações são possíveis.

A marginal de x_t na primeira representação tem como vetor de parâmetros $\theta_2 = (\mu_x, \sigma_x^2)$.

A distribuição condicional de y_t dado x_t terá média condicional dada por

$$E(y_t \mid x_t) = a + bx_t \tag{109}$$

onde

$$a = \mu_y - b\mu_x \tag{110}$$

e

$$b = \frac{cov(y_t, x_t)}{var(x_t)} = \frac{\sigma_{yx}}{\sigma_x^2} \tag{111}$$

e substituindo (110) e (111) em (109) temos:

$$E(y_t \mid x_t) = \mu_y - b\mu_x + bx_t = \mu_y + \frac{\sigma_{yx}}{\sigma_x^2}(x_t - \mu_x) \tag{112}$$

Agora a variância da distribuição condicional será dada por

$$
\begin{aligned}
Var(y_t \quad | \quad & x_t) = E(y_t - E(y_t \mid x_t) \mid x_t)^2 \\
&= E(y_t - \mu_y - \frac{\sigma_{yx}}{\sigma_x^2}(x_t - \mu_x))^2 \\
&= E(y_t - \mu_y)^2 + \left(\frac{\sigma_{yx}}{\sigma_x^2}\right)^2 E(x_t - \mu_x)^2 - 2\frac{\sigma_{yx}}{\sigma_x^2}E((y_t - \mu_y)(x_t - \mu_x)) \\
&= \sigma_y^2 + \left(\frac{\sigma_{yx}}{\sigma_x^2}\right)^2 \sigma_x^2 - 2\frac{\sigma_{yx}}{\sigma_x^2}\sigma_{yx} \\
&= \sigma_y^2 - \frac{\sigma_{yx}^2}{\sigma_x^2} \qquad (113)
\end{aligned}
$$

Observe que (113) se denotarmos as entradas da matriz de variância e covariância por ω_{ij} com $i, j = 1, 2$ esta expressão pode ser escrita da seguinte forma

$$Var(y_t \mid x_t) = \omega_{11} - \omega_{12}\omega_{22}^{-1}\omega_{21} \qquad (114)$$

que é menor do que a variânica de y_t.

Vamos considerar o caso em que o processo estocástico tem dependência temporal, neste caso a distribuição de z_t é condicional na informação passada, denotada por I_{t-1} e será dada por:

$$z_t \mid I_{t-1} \text{ ˜ } NI\left(\begin{bmatrix} \mu_{y|I_{t-1}} \\ \mu_{x|I_{t-1}} \end{bmatrix}, \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}\right) \qquad (115)$$

onde se a informação passada for só o passado mais recente, isto é, $I_{t-1} = \{y_{t-1}, x_{t-1}\}$ teremos $\mu_{y|I_{t-1}} = E(y_t|y_{t-1}, x_{t-1}) = \phi_{11}y_{t-1}+ \phi_{12}x_{t-1}$e $\mu_{x|I_{t-1}} = E(x_t|y_{t-1}, x_{t-1}) = \phi_{21}y_{t-1}+ \phi_{22}x_{t-1}$ pela normalidade.

Observe que os elementos da matriz de variância e covariância representam estas quantidades condicionais na informação passada.

Neste caso a distribuição conjunta será fatorada como condicional e marginal da seguinte forma:

$$D(z_t \mid I_{t-1}, \theta) = D(y_t \mid I_{t-1}, x_t, \theta_1)D(x_t \mid I_{t-1}, \theta_2) \qquad (116)$$

Neste caso a distribuição marginal de x_t tem como vetor de parâmetros $\theta_2 = (\mu_{x|I_{t-1}}, \omega_{22})$

Para a distribuição condicional teremos:

$$E(y_t \mid I_{t-1}, x_t, \theta_1) = a^* + b^* x_t \tag{117}$$

onde

$$a^* = \mu_{y|I_{t-1}} - b^* \mu_{x|I_{t-1}} \tag{118}$$

e

$$b^* = \omega_{12}\omega_{22}^{-1}$$

A variânica da distribuição condicional será dada por

$$Var(y_t \mid I_{t-1}, x_t, \theta_1) = \omega_{11} - \omega_{12}\omega_{22}^{-1}\omega_{21}$$

Referências

G. E. P. Box and G. M. Jenkins. *Time series analysis: forecasting and control*. Holden-Day series in time series analysis. Holden-Day, 1970. URL `https://books.google.com.br/books?id=qsnaAAAAMAAJ`.

G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

R. Brown. *Statistical Forecasting for Inventory Control*. McGraw-Hill, 1959.

J. A. Doornik. Autometrics. In J. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics: A Frestschrift in Honour of David F. Hendry*, chapter 4, pages 88–121. Oxford University Press, 2009.

J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, may 2012. doi: 10.1093/acprof:oso/9780199641178.001.0001.

E. Ghysels, A. C. Harvey, and E. Renault. Stochastic volatility. In *Handbook of Statistics*, pages 119–191. Elsevier, 1996. doi: 10.1016/s0169-7161(96)14007-4.

A. Harvey and S. J. Koopman. Diagnostic checking of unobserved-components time series models. *Journal of Business & Economic Statistics*, 10(4):377–389, 1992.

A. C. Harvey. *Time Series Models*. Harvester, 1993.

D. Hendry. *Dynamic Econometrics*. Oxford University Press, 1995. ISBN 0198283164. URL `https://www.ebook.de/de/product/3257383/david_hendry_dynamic_econometrics.html`.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications*. Springer New York. ISBN 9781441978653. doi: 10.1007/978-1-4419-7865-3.

P. L. Valls Pereira. Exact likelihood function for a regression model with MA(1) errors. *Economics Letters*, 24(2):145–149, jan 1987. doi: 10.1016/0165-1765(87)90241-2.