

Filtro de Kalman

Pedro Valls*

CEQEF e EESP-FGV

18 de agosto de 2022

Sumário

1	Introdução	2
1.1	Modelo de Expectativa Adaptativa com Coeficiente de Ajuste Constante.	3
1.2	Modelo de Expectativa Adaptativa com Coeficiente de Ajuste Variando no Tempo.	4
1.3	Extração da Variância Condicional de uma série	5
2	Representação em Espaço de Estado Linear	5
2.1	Exemplos	6
2.1.1	AR(1) + Ruído	6
2.1.2	Modelo de Volatilidade Estocástica com Efeitos Calendários . . .	7
2.1.3	Modelo $ARMA(p, q)$	7
2.1.4	Casos Particulares	8
2.2	Exercícios	11
3	Filtragem, Suavizamento e Previsão - Filtro de Kalman	12
3.1	Filtragem	12
3.1.1	Programa no EVIEWS para Estimar Modelos por Filtro de Kalman	14
3.2	Previsão	23
3.3	Suavizamento (a ser modificado)	24
3.3.1	Alguns exemplos	24
3.3.2	Comando no Eviews para gerar componentes filtrados e suavizados	29
3.3.3	Estimando o Modelo de Nível Local e de Tendência Local no Ox-Metrics usando o STAMP	34
3.3.4	Exercícios	41

*Sao Paulo School of Economics - FGV e CEQEF-FGV, Rua Itapeva 474 s.1006, 01332-000, , São Paulo, SP, Brasil. E-mail:pedro.valls@fgv.br. © 2022 - Pedro Valls

4	Sazonalidade	42
4.1	Introdução	42
4.2	Modelos <i>ARMA</i> Sazonais	42
4.3	Modelos <i>ARIMA</i> Sazonais	45
4.4	Sazonalidade Determinística	49
4.5	Componente Sazonal no Modelo Estrutural	50
4.5.1	Usando o STAMP para Estimar o Modelo Nível Local + Sazonalidade Estocástica Bens de Capital	52
5	Componente Cíclico	54
5.1	Componente Determinístico	54
5.2	Componente Cíclico Estocástico	55
5.2.1	Usando o STAMP para Estimar o Modelo Nível Local + Sazonalidade Estocástica+Ciclo Estocástico para Bens de Capital	56
6	Modelo Estrutural Básico	58
7	Modelo Estrutural com Variáveis Exogenas	58
8	Valores Aberrantes e Mudanças Estruturais	59
9	Apêndice I - Distribuição Condicional	62

1 Introdução

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

- (i) modelos, em geral, relacionam variáveis observadas à variáveis não observadas. Por exemplo no modelo linear de regressão relaciona-se, uma variável observada a variáveis não observadas, os parâmetros e a função de ligação é linear nos regressores.

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

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

$$\begin{aligned} Y_t &= X_t\beta_t + \varepsilon_t \\ \beta_t &= \beta_{t-1} + \eta_t \end{aligned}$$

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 maximizar esta função. Maximizando esta função obtém-se 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. 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.

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:

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

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

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

mas por (1) temos:

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

que substituindo em (2) temos:

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

$$\Updownarrow$$

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

$$\Updownarrow$$

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

que é o **modelo de expectativas adaptativas com coeficiente de ajuste θ 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 médio quadrático de previsão.

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

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

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:

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

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

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

Agora suponha que o pesos decaem exponencialmente, isto é:

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

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:

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

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

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

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

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

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 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 $E[s_t | I_{t-1}] = 0$, $E[\pi_t s_{t-j}] = 0$ 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 logaritmo do quadrado dos retornos, a um não observado o logaritmo 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

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:

$$\begin{matrix} y_t & = & Z_t & \alpha_t & + & c_t & + & G_t & \varepsilon_t \\ (Nx1) & & (Nxm) & (mx1) & & (Nx1) & & (NxN) & (Nx1) \end{matrix} \text{ para } t = 1, \dots, T \quad (17)$$

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:

$$\begin{matrix} \alpha_t & = & T_{t-1} & \alpha_{t-1} & + & d_{t-1} & + & H_t & \eta_t \\ (mx1) & & (mxm) & (mx1) & & (mx1) & & (mxg) & (gx1) \end{matrix} \text{ para } t = 1, \dots, T \quad (18)$$

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) $E[\varepsilon_t \eta_s'] = 0 \forall t, s$
- (iii) $E[\varepsilon_t \alpha_0'] = E[\eta_t \alpha_0'] = 0 \forall t$.

2.1 Exemplos

2.1.1 AR(1) + Ruído

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

$$\begin{aligned} y_t &= \mu_t + \sigma_\xi \xi_t \text{ com } \xi_t \sim NI(0, 1) \\ \mu_t &= \phi \mu_{t-1} + \sigma_\eta \eta_t \text{ com } \eta_t \sim NI(0, 1) \end{aligned}$$

que corresponde ao modelo descrito na seção 1.3 se:

$$\begin{aligned} y_t &= \ln r_t^2 \\ \mu_t &= \ln \sigma_t^2 \\ \xi_t &= \ln \varepsilon_t^2 \\ \sigma_\xi &= 1 \end{aligned}$$

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

Observe que se $\sigma_\xi = 0$, teríamos um AR(1) sem ruído, que poderia ser estimado por Mínimos Quadrados Condicionais.

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,

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

Temos então o seguinte modelo:

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

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

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:

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

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

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

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

com alguns dos ϕ 's ou θ 's iguais a zero.

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

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

e a equação de transição será dada por:

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

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

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) $AR(2)$

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

$$\alpha_t = \begin{bmatrix} y_t \\ l_t^{(1)} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ 0 \end{bmatrix} \varepsilon_t \quad (25)$$

observe que (25) nós dá:

$$y_t = \phi_1 y_{t-1} + l_{t-1}^{(1)} + \sigma_\varepsilon \varepsilon_t \quad (26)$$

$$l_t^{(1)} = \phi_2 y_{t-1} \quad (27)$$

e substituindo (27) em (26) temos

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

.

Caso o vetor de estado for definido da seguinte forma, $\alpha'_t = [y_t, y_{t-1}]'$ a equação (25) pode ser re-escrita da seguinte forma:

$$\alpha_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ 0 \end{bmatrix} \varepsilon_t \quad (28)$$

(3) *MA(1)*

$$y_t = (1, 0) \alpha_t \quad (29)$$

$$\alpha_t = \begin{bmatrix} y_t \\ l_t^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \sigma_\varepsilon \\ \sigma_\varepsilon \theta_1 \end{bmatrix} \varepsilon_t \quad (30)$$

observe que (30) implica em:

$$y_t = l_{t-1}^{(1)} + \sigma_\varepsilon \varepsilon_t \quad (31)$$

$$l_t^{(1)} = \theta_1 \sigma_\varepsilon \varepsilon_t \quad (32)$$

e substituindo (32) em (31) 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 Local¹

$$y_t = \mu_t + \sigma_\varepsilon \varepsilon_t \quad (33)$$

$$\mu_t = \mu_{t-1} + \sigma_\eta \eta_t \quad (34)$$

\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 (34) pode ser re-escrito da seguinte forma:

$$\Delta \mu_t = \sigma_\eta \eta_t \quad (35)$$

e aplicando o operador diferença em (33) temos

$$\Delta y_t = \Delta \mu_t + \sigma_\varepsilon \Delta \varepsilon_t \quad (36)$$

e substituindo (35) em (36) obtém-se:

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

$$\Delta y_t = \sigma_\eta \eta_t + \sigma_\varepsilon \Delta \varepsilon_t \quad (37)$$

e o lado direito de (37) pode ser escrito como $\sigma_\xi \xi_t + \theta \sigma_\xi \xi_{t-1}^2$, um $MA(1)$ cujos parâmetros são dados pela solução do seguinte sistema:

$$(1 + \theta^2) \sigma_\xi^2 = \sigma_\eta^2 + 2\sigma_\varepsilon^2 \quad (38)$$

$$\theta \sigma_\xi^2 = -\sigma_\varepsilon^2 \quad (39)$$

dividindo (38) por (39) obtém-se:

$$\begin{aligned} \frac{(1 + \theta^2)}{\theta} &= -\frac{\sigma_\eta^2 + 2\sigma_\varepsilon^2}{\sigma_\varepsilon^2} \\ \frac{(1 + \theta^2)}{\theta} &= -(q + 2) \end{aligned} \quad (40)$$

onde $q = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$ e a solução de (40) é dada pela solução da seguinte equação do segundo grau:

$$\theta^2 + (q + 2)\theta + 1 = 0 \quad (41)$$

(5) Modelo de Tendência Local

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

\Downarrow

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

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

$$\Delta \beta_t = \sigma_\xi \xi_t \implies \beta_t = \frac{\sigma_\xi \xi_t}{(1 - L)} \quad (42)$$

e substituindo (42) na equação do nível temos:

$$\Delta \mu_t = \beta_{t-1} + \sigma_\xi \xi_t \implies \Delta \mu_t = \frac{\sigma_\xi \xi_{t-1}}{(1 - L)} + \sigma_\eta \eta_t \quad (43)$$

e substituindo (43) na equação das observações temos:

$$\begin{aligned} \Delta y_t &= \Delta \mu_t + \sigma_\varepsilon \Delta \varepsilon_t \implies \Delta y_t = \frac{\sigma_\xi \xi_{t-1}}{(1 - L)} + \sigma_\eta \eta_t + \sigma_\varepsilon \Delta \varepsilon_t \\ \Delta \Delta y_t &= \sigma_\xi \xi_{t-1} + \sigma_\eta \Delta \eta_t + \sigma_\varepsilon \Delta \Delta \varepsilon_t \end{aligned} \quad (44)$$

²Estamos assumindo que $\xi_t \sim NI(0, 1)$

onde o lado direito de (44) pode ser escrito como $\sigma_\zeta \zeta_t + \theta_1 \sigma_\zeta \zeta_{t-1} + \theta_2 \sigma_\zeta \zeta_{t-2}$,³ um $MA(2)$ cujos parâmetros estão relacionados as variâncias do modelo de tendência local pelas seguintes equações:

$$(1 + \theta_1^2 + \theta_2^2)\sigma_\zeta^2 = \sigma_\xi^2 + 2\sigma_\eta^2 + 6\sigma_\varepsilon^2 \quad (45)$$

$$(\theta_1 + \theta_1\theta_2)\sigma_\zeta^2 = -\sigma_\eta^2 - 4\sigma_\varepsilon^2 \quad (46)$$

$$\theta_2\sigma_\zeta^2 = \sigma_\varepsilon^2 \quad (47)$$

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 somente 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 em um software (EViews, OxMetrics ou 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 $y_1 = \varepsilon_1 = 0$.

4. Escreva um programa no (EViews, OxMetrics ou 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?

³Estamos assumindo que $\zeta_t \sim NI(0, 1)$.

3 Filtragem, Suavizamento e Previsão - Filtro de Kalman

Existem várias referências para os algoritmos de filtragem, suavizamento e previsão para o Filtro de Kalman. Entre elas podemos citar:

1. Harvey [1993] é a segunda edição do livro inicialmente lançado em 1981. Semelhante a apresentação deste capítulo;
2. Durbin and Koopman [2012] mais avançado que o anterior tem abordagem clássica. Completo na parte de algoritmos de suavizamento;
3. Hamilton [1994] também mais avançado que Harvey [1993] mas semelhante ao descrito neste capítulo.
4. Tsay and Chen [2018] para uma abordagem clássica e bayesiana.

3.1 Filtragem

Seja a_t o estimador ótimo de α_t usando a informação até o instante t , isto é, 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 preditor de α_{t+1} usando a informação até t é denotado por $a_{t+1|t}$ e será dado por:

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

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

$$a_{t+1|t} = T_t a_t + d_t \quad (48)$$

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

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

então a variância deste erro é dada por:

$$\begin{aligned} P_{t+1|t} &= E_t[(a_{t+1|t} - \alpha_{t+1})(a_{t+1|t} - \alpha_{t+1})'] = \\ &= T_t E_t[(a_t - \alpha_t)(a_t - \alpha_t)'] T_t' + \\ &\quad + H_{t+1} E_t[\eta_{t+1} \eta_{t+1}'] H_{t+1}' - 2 T_t E_t[(a_t - \alpha_t) \eta_{t+1}'] H_{t+1}' \end{aligned} \quad (49)$$

mas

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

então (49) reduz-se a:

$$P_{t+1|t} = T_t P_t T'_t + H_{t+1} H'_{t+1} \quad (50)$$

ERRO DE PREVISÃO E SUA VARIÂNCIA

Dado (48) podemos obter o previsor das observação usando a informação até o instante $t - 1$ que será dado por:

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

$$\Downarrow$$

$$\tilde{y}_{t+1|t} = Z_{t+1}a_{t+1|t} + c_{t+1} \quad (51)$$

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

$$\begin{aligned} \nu_{t+1} &= y_{t+1} - \tilde{y}_{t+1|t} = \\ &= y_{t+1} - Z_{t+1}a_{t+1|t} - c_{t+1} \\ &= Z_{t+1}(\alpha_{t+1} - a_{t+1|t}) + G_{t+1}\varepsilon_{t+1} \end{aligned} \quad (52)$$

e a variância deste erro é dada por:

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

$$\Downarrow$$

$$F_{t+1} = Z_{t+1}P_{t+1|t}Z'_{t+1} + G_{t+1}G'_{t+1} \quad (53)$$

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:

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

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

para facilitar vamos assumir distribuição normal para as observações, temos então:

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

e a distribuição de $\alpha_{t+1}|y_{t+1}, I_t$ será dada por⁴:

$$\alpha_{t+1}|y_{t+1}, I_t \sim N(a_{t+1|t} + P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}\nu_{t+1}, P_{t+1|t} - P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}Z_{t+1}P_{t+1|t}) \quad (54)$$

deste modo temos que:

$$a_{t+1} = E_t[\alpha_{t+1}] = a_{t+1|t} + P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}\nu_{t+1} \quad (55)$$

e

$$P_{t+1} = Var_t[\alpha_{t+1}] = P_{t+1|t} - P_{t+1|t}Z'_{t+1}(F_{t+1})^{-1}Z_{t+1}P_{t+1|t} \quad (56)$$

Resumindo as recursões do Filtro de Kalman

(1) Equações de previsão:

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

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

(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 necessárias condições iniciais para a_0 e P_0 . Para os modelos estacionários estas condições são dadas por:

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

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

Os estimadores de máxima verossimilhança dos parâmetros são facilmente obtidos usando-se a decomposição do erro de previsão, isto é, como $\nu_t|I_{t-1} \sim N(0, F_t)$ 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 EVIEWS para Estimar Modelos por Filtro de Kalman

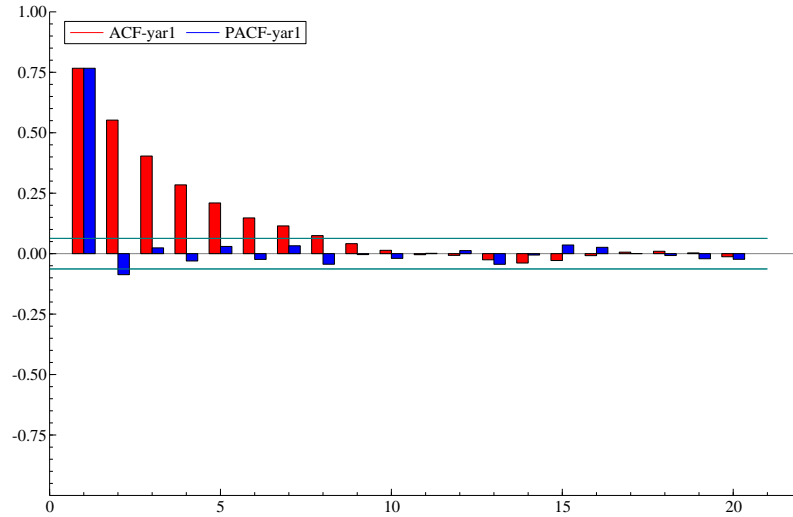
Os programas abaixo simulam e estimam modelos: $AR(1)$, $AR(4)$, $MA(1)$, $MA(4)$, $ARMA(1,1)$, nível local, tendência local e modelo de volatilidade estocástica.

Github-Eviews

⁴Veja Apêndice I para a obtenção destes resultados para o caso geral.

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

Figura 1: Correlação da Série *yar1*



Abaixo são apresentadas as estimativas dos parâmetros da série *yar1*. Observe que a estimativa da variância do erro é dada por $\exp[C(1)]$ que é próxima de um e a estimativa do parâmetro autorregressivo, $C(2)$, é próxima do valor verdadeiro, 0.75.

Tabela 1: Estimativa do Modelo $AR(1)$ com $\phi = 0.75$

Sspace: SSAR1				
Method: Maximum likelihood (BFGS / Marquardt steps)				
Sample: 1 1000				
Included observations: 1000				
Convergence achieved after 18 iterations				
Coefficient covariance computed using outer product of gradients				
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.003851	0.042960	-0.089649	0.9286
C(2)	0.769570	0.020814	36.97387	0.0000
	Final State	Root MSE	z-Statistic	Prob.
SV1	-1.318996	0.998076	-1.321539	0.1863
Log likelihood	-1417.461	Akaike info criterion	2.838923	
Parameters	2	Schwarz criterion	2.848738	
Diffuse priors	0	Hannan-Quinn criter.	2.842653	

Modelo $AR(4)$ O correlograma na figura(2) tem uma FAC que decae exponencialmente e uma FACP com as quatro primeiras autocorrelações parciais diferentes de zero caracterizando um modelo $AR(4)$.

Na tabela(2) são apresentadas as estimativas dos parâmetros do modelo *yar4*. Observe que todas as estimativas são estatisticamente iguais as parâmetros verdadeiros.

Figura 2: Correlação da Série *yar4* com $\phi_1 = 0.32$, $\phi_2 = -0.23$, $\phi_3 = 0.43$, $\phi_4 = -0.11$

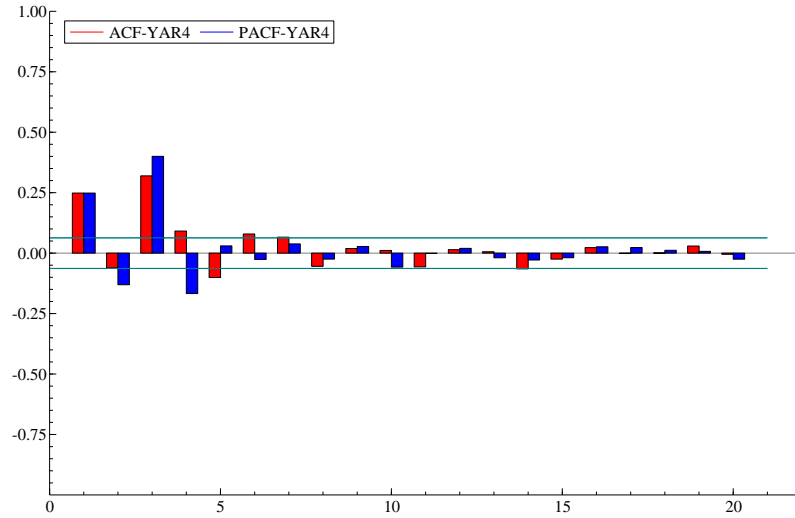


Tabela 2: Estimativa do Modelo $AR(4)$ com $\phi_1 = 0.32$, $\phi_2 = -0.23$, $\phi_3 = 0.43$, $\phi_4 = -0.11$

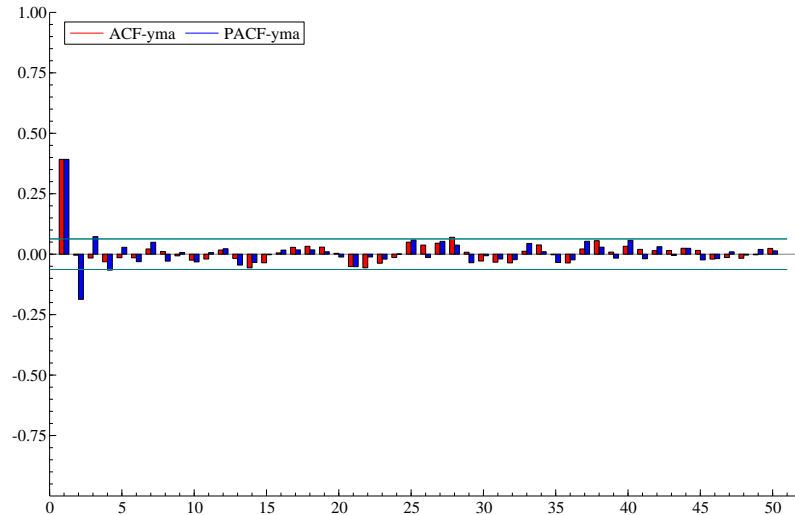
Sspace: SSAR4
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 19:55
Sample: 5 1000
Included observations: 996
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

		Coefficient	Std. Error	z-Statistic	Prob.
	C(1)	-0.011882	0.042968	-0.276519	0.7821
	C(2)	0.406965	0.031232	13.03049	0.0000
	C(3)	-0.285796	0.033460	-8.541387	0.0000
	C(4)	0.458051	0.030773	14.88462	0.0000
	C(5)	-0.170347	0.031979	-5.326929	0.0000
		Final State	Root MSE	z-Statistic	Prob.
	SV1	-0.687373	0.994077	-0.691469	0.4893
	SV2	-1.533775	0.000000	NA	0.0000
	SV3	-0.152687	0.000000	NA	0.0000
	SV4	-0.078300	0.000000	NA	0.0000
Log likelihood		-1407.717	Akaike info criterion		2.836781
Parameters		5	Schwarz criterion		2.861398
Diffuse priors		0	Hannan-Quinn criter.		2.846139

Modelo $MA(1)$ Valls Pereira [1987a] ou Valls Pereira [1987b]

Na figura(3) é apresentado o correlograma da série yma . Observe que somente a primeira autocorrelação é diferente de zero e a FACP decae exponencialmente caracterizando um processo $MA(1)$.

Figura 3: Correlação da Série $yma1$ com $\theta = 0.4$



Na tabela (3) abaixo são apresentadas as estimativas dos parâmetros para o modelo $MA(1)$ e estas são estatisticamente iguais as parâmetros verdadeiros.

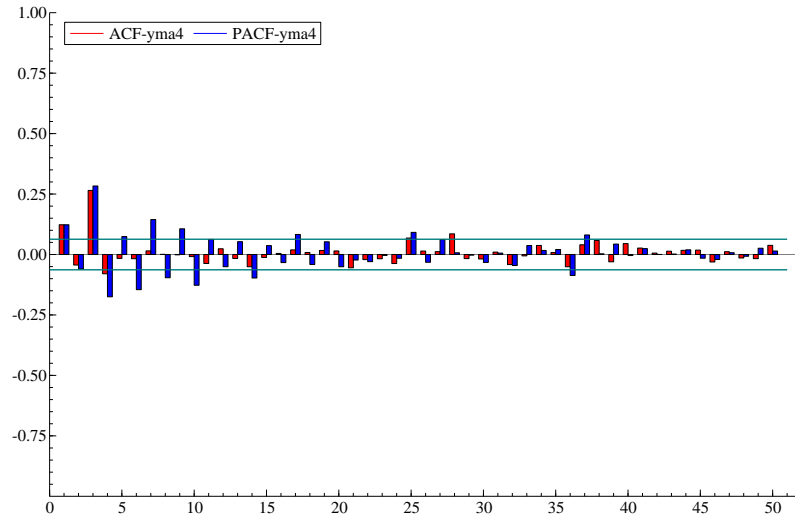
Tabela 3: Estimativa do Modelo $MA(1)$ com $\theta_1 = 0.4$

Sspace: SSMA1
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 20:25
Sample: 2 1000
Included observations: 999
Convergence achieved after 10 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.495383	0.028076	17.64463	0.0000
C(2)	-0.012869	0.042881	-0.300106	0.7641
	Final State	Root MSE	z-Statistic	Prob.
SV1	0.000000	0.993586	0.000000	1.0000
SV2	-1.555619	1.83E - 08	-85238834	0.0000
Log likelihood	-1411.232	Akaike info criterion		2.829294
Parameters	2	Schwarz criterion		2.839117
Diffuse priors	0	Hannan-Quinn criter.		2.833028

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

Figura 4: Correlação da Série y_{ma1} com $\theta_1 = 0.32$, $\theta_2 = -0.23$, $\theta_3 = 0.43$, $\theta_4 = -0.11$



Na tabela (4) abaixo são apresentadas as estimativas dos parâmetros para o modelo $MA(14)$ e estas são estatisticamente iguais as parâmetros verdadeiros.

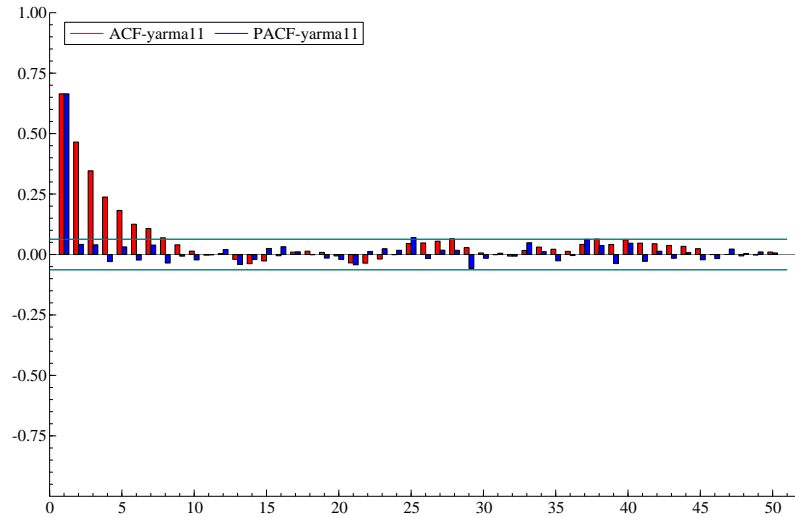
Tabela 4: Estimativa do Modelo $MA(1)$ com $\theta_1 = 0.32$, $\theta_2 = -0.23$, $\theta_3 = 0.43$, $\theta_4 = -0.11$

Sspace: SSMA4
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 20:25
Sample: 5 1000
Included observations: 996
Convergence achieved after 63 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.330684	0.031021	10.66005	0.0000
C(2)	-0.169641	0.033482	-5.066651	0.0000
C(3)	0.361829	0.031507	11.48417	0.0000
C(4)	-0.068766	0.031613	-2.175222	0.0296
C(5)	0.067028	0.042574	1.574394	0.1154
	Final State	Root MSE	z-Statistic	Prob.
SV1	0.000000	1.034082	0.000000	1.0000
SV2	-1.753486	2.11E - 08	-83208385	0.0000
SV3	-0.173280	2.11E - 08	-8222656.	0.0000
SV4	-0.141566	2.11E - 08	-6717742.	0.0000
SV5	0.024763	2.11E - 08	1175085.	0.0000
Log likelihood	-1447.572	Akaike info criterion		2.916811
Parameters	5	Schwarz criterion		2.941428
Diffuse priors	0	Hannan-Quinn criter.		2.926169

Modelo $ARMA(1, 1)$ Na figura (5) abaixo é apresentado o correlograma da série y_{arma11} . Observe as autocorrelações decaem exponencialmente e a FACP decae exponencialmente caracterizando um processo $ARMA(1, 1)$.

Figura 5: Correlação da Série *yarma11* com $\phi_1 = 0.75$, $\theta_1 = -0.2$



Na tabela (5) abaixo são apresentadas as estimativas dos parâmetros para o modelo $ARMA(1, 1)$ e a duas são estatisticamente são parecidas com os parâmetros verdadeiros, $C(3)$ corresponde ao parâmetro autorregressivo, que é proóximo do paraâmetro verdadeiro, e $C(1)$ mas para o parâmetros de média-móvel, mas que não é próximo do valor verdadeiro.

Tabela 5: Estimativa do Modelo $ARMA(1, 1)$ com $\phi = 0.75$ e $\theta_1 = -0.2$

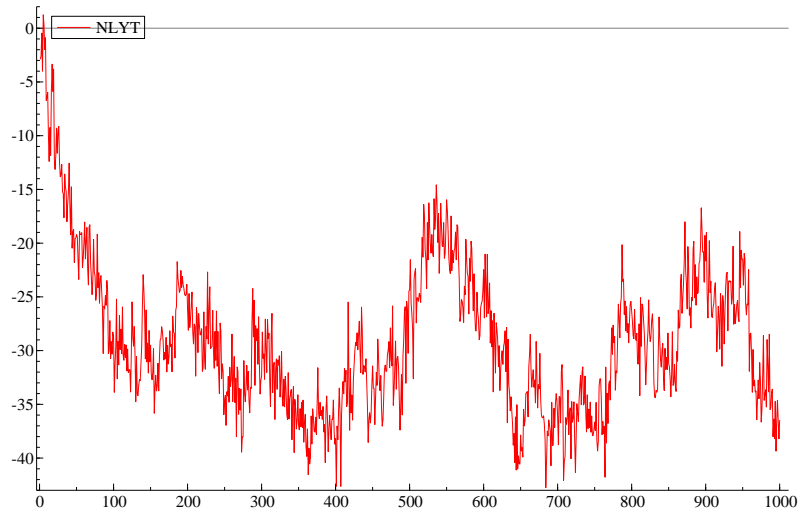
Sspace: SSARMA11
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 20:25
Sample: 2 1000
Included observations: 999
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.069824	0.048906	-1.427715	0.1534
C(2)	-0.011179	0.042870	-0.260757	0.7943
C(3)	0.706445	0.035974	19.63789	0.0000
	Final State	Root MSE	z-Statistic	Prob.
SV1	-1.199412	0.994426	-1.206135	0.2278
SV2	-1.697814	0.000000	NA	0.0000
Log likelihood	-1412.233	Akaike info criterion		2.833300
Parameters	3	Schwarz criterion		2.848035
Diffuse priors	0	Hannan-Quinn criter.		2.838901

Modelo de Nível Local A seguir é apresentado, na figura(6) o um modelo de nível local. Observe que esta série é não estacionária uma vez que não retorna frequentemente para a sua média que é zero.

Abaixo são apresentadas as estimativas dos parâmetros, a variância do nível e das

Figura 6: Série de Nível Local



observações e observe ambas são próximas dos valores verdadeiros, um para o erro do nível e quatro para o erro das observações.

Tabela 6: Estimativa do Modelo de Nível Local

Sspace: NIVELLOCAL
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 20:25
Sample: 1 1000
Included observations: 1000
Convergence achieved after 19 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.374134	0.060446	22.73308	0.0000
C(2)	0.062488	0.134928	0.463121	0.6433
	Final State	Root MSE	z-Statistic	Prob.
SV1	-36.79935	1.628233	-22.60079	0.0000
Log likelihood	-2369.558	Akaike info criterion		4.743115
Parameters	2	Schwarz criterion		4.752931
Diffuse priors	1	Hannan-Quinn criter.		4.746846

Observe que também foi estimada a forma reduzida deste modelo cujos resultados estão apresentados na tabela(7)

Tabela 7: Estimativa da Forma Reduzida do Modelo de Nível Local

Dependent Variable: D(NLYT)
Method: ARMA Maximum Likelihood (BFGS)
Date: 08/15/22 Time: 20:25
Sample: 2 1000
Included observations: 999
Convergence achieved after 3 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.599073	0.025169	-23.80240	0.0000
SIGMASQ	6.609392	0.298781	22.12120	0.0000
R-squared	0.256020	Mean dependent var		-0.033603
Adjusted R-squared	0.255274	S.D. dependent var		2.982069
S.E. of regression	2.573451	Akaike info criterion		4.730818
Sum squared resid	6602.783	Schwarz criterion		4.740641
Log likelihood	-2361.043	Hannan-Quinn criter.		4.734551
Durbin-Watson stat	1.975520			
Inverted MA Roots	.60			

onde a relação entre os parâmetros da forma reduzida e da forma estrutural são obtidos através de (39) e (41), isto é,

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

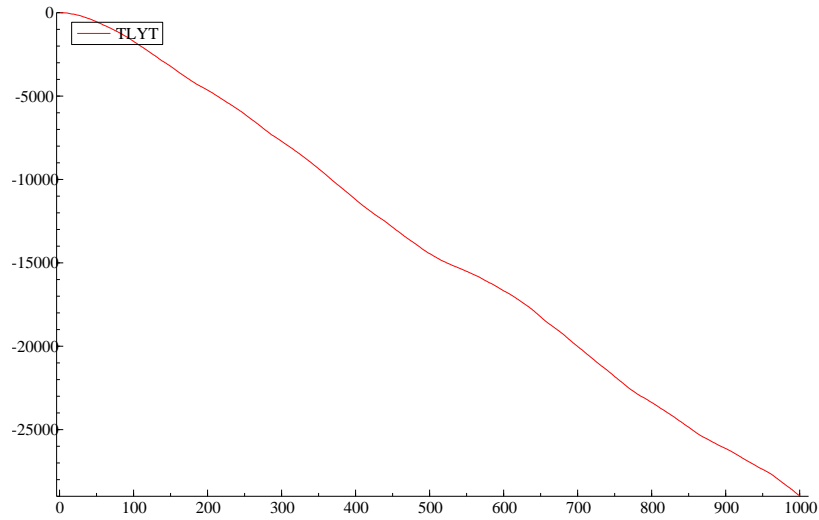
$$\theta^2 + \frac{9}{4}\theta + 1 = 0$$

que implica que $\theta = -0.61$ e portanto $\sigma_{\xi}^2 = 6.5574$ cujo desvio padrão é dado por 2.5607, que estão próximo dos valores estimados que são de $\theta = -0.59907$ e $\sigma_{\xi} = 2.5708$.

Mas se forem usados os valores estimados, isto é, $\sigma_{\eta}^2 = \exp(0.062488) = 1.0645$ e $\sigma_{\varepsilon}^2 = \exp(1.374134) = 3.9516$, teríamos as seguinte soluções $\theta = -0.5985$ e com $\sigma_{\xi} = 2.5852$, e estes valores também estão próximos dos valores estimados.

Modelo de Tendência Local Na figura(7) abaixo é apresentado o gráfico do modelo de tendência local. Observe que esta série é não estacionária é provavelmente é integrada de ordem 2.

Figura 7: Série de Tendência Local



Na tabela(8) abaixo são apresentadas as estimativas dos parâmetros do modelo de tendência local. Observe que estatisticamente os valores não são diferentes dos valores verdadeiros, uma vez que $\sigma_\varepsilon^2 = \exp(2.261629) = 9.5987$ tendo um intervalo de confiança igual a $[\exp(2.261629 - 2 * .100442) : \exp(2.261629 + 2 * .100442)] = [7.8518 : 11.734]$; $\sigma_\eta^2 = \exp(1.506388) = 4.5104$ tendo um intervalo de confiança igual a $[\exp(1.506388 - 2 * .334430) : \exp(1.506388 + 2 * .334430)] = [2.3106 : 8.8044]$ e $\sigma_\xi^2 = \exp(-0.034034) = 0.96654$ tendo um intervalo de confiança igual a $[\exp(-0.034034 - 2 * .170112) : \exp(-0.034034 + 2 * .170112)] = [0.68780 : 1.3582]$.

Tabela 8: Estimativa do Modelo de Tendência Local

Sspace: TENDENCIALLOCAL
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/15/22 Time: 20:25
Sample: 1 1000
Included observations: 1000
Convergence achieved after 54 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.015580	0.112690	17.88605	0.0000
C(2)	1.774245	0.249953	7.098327	0.0000
C(3)	-0.112017	0.178661	-0.626981	0.5307
	Final State	Root MSE	z-Statistic	Prob.
SV1	-29015.97	4.154208	-6984.717	0.0000
SV2	-35.88536	2.042824	-17.56654	0.0000
Log likelihood	-3036.976	Akaike info criterion		6.079952
Parameters	3	Schwarz criterion		6.094676
Diffuse priors	2	Hannan-Quinn criter.		6.085548

A seguir é apresentada a estimativa da forma reduzida do modelo de tendência local onde a relação entre os parâmetros da forma reduzida e da forma estrutural são obtidos através de (45-47), isto é,

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

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

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

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

cujas soluções são: 0.350 31, 2. 854 6, 0.897 54+0.440 93*i* e 0.897 54−0.440 93*i* e escolhemos a primeira que não é estatisticamente diferente do valor estimado que é 0.356447 cujo intervalo de confiança é dado por $[0.356447 - 2 * 0.029656 : 0.356447 + 2 * 0.029656] = [0.297 14 : 0.415 76]$.

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 estatisticamente diferente do valor estimado que é −1.168419 cujo intervalo de confiança é dado por $[-1.168419 - 2 * 0.029684 : -1.168419 + 2 * 0.029684] = [-1. 109 1 : -1. 227 8]$.

Para obter o valor de σ_ζ^2 devemos resolver

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

que não é estatisticamente diferente do valor estimado que é 5.227593.

3.2 Previsão

A previsão do vetor de estado é dada por:

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

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:

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

com um erro quadrático médio dado por:

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

3.3 Suavizamento (a ser modificado)

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:

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

$$P_{t|T} = P_t + P_t^*(P_{t+1|T} - P_{t+1|t})P_t^{*'}$$

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

para $t = T - 1, \dots, 1$, 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 $P_0 = \frac{\sigma^2}{1-\phi_1^2}$

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

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

$$\nu_1 = y_1$$

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

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

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

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

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

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

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

$$f_2 = \sigma^2$$

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

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

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:

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

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

(2) $MA(1)$ esta derivação obtida em Valls Pereira [1987a] ou Valls Pereira [1987b]

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:

$$\nu_1 = y_1$$

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

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

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

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

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

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

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

então para $t \geq 2$ temos

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

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

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 substitui θ_1 por $\frac{1}{\theta_1}$. Deste modo a maximização da verossimilhança dá a solução invertível do processo.

(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:

$$b_t = \left(\sum_{j=1}^t x_j x_j' \right)^{-1} \sum_{j=1}^t x_j y_j = (X_t' X_t)^{-1} (X_t' \mathbf{y}_t) \quad (57)$$

para $t > k$.

Agora observe que

$$\begin{aligned} (X_t' X_t)^{-1} &= (X_{t-1}' X_{t-1} + x_t x_t')^{-1} \\ &= (X_{t-1}' X_{t-1})^{-1} - \frac{(X_{t-1}' X_{t-1})^{-1} x_t x_t' (X_{t-1}' X_{t-1})^{-1}}{1 + x_t' (X_{t-1}' X_{t-1})^{-1} x_t} \end{aligned} \quad (58)$$

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 (58) em (57) e usando que $(X_t' \mathbf{y}_t) = (X_{t-1}' \mathbf{y}_{t-1} + x_t' y_t)$ tem-se:

$$b_t = b_{t-1} + (X_{t-1}' X_{t-1})^{-1} x_t \frac{(y_t - x_t' b_{t-1})}{f_t} \quad (59)$$

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

Para iniciar o processo faça $t = k$ em (57) obtendo:

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

Observe que (58) corresponde, a menos de σ^2 , a atualização da matriz de covariância do estimador dos parâmetros e (59) 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:

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

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

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

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:

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

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:

$$m_t = m_{t|t-1} + p_{t|t-1} \frac{(y_t - m_{t|t-1})}{1 + p_{t|t-1}} \quad (60)$$

$$p_t = p_{t|t-1} - \frac{p_{t|t-1}^2}{1 + p_{t|t-1}} \quad (61)$$

As condições iniciais são

$$m_1 = y_1$$

$$p_1 = 1$$

Observe que (60) pode ser reescrito da seguinte forma:

$$m_t = (1 - \lambda_t)m_{t-1} + \lambda_t y_t \quad (62)$$

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 (61) 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 previsora 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 previsora das observações futuras será dado por:

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

com erro quadrático de previsão dado por:

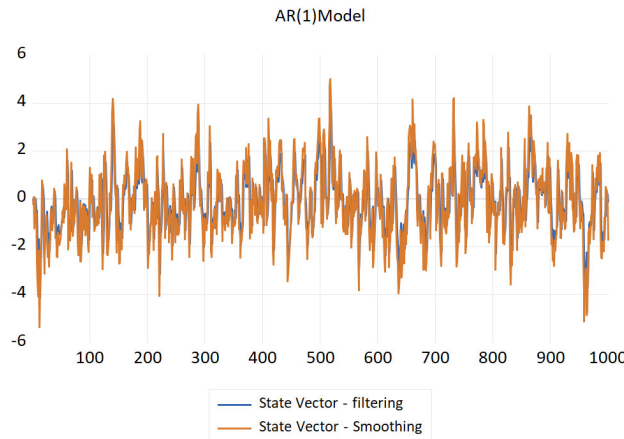
$$EQP(\tilde{y}_{T+l|T}) = p_T + lq + \sigma_\varepsilon^2$$

3.3.2 Comando no Eviews para gerar componentes filtrados e suavizados

No programa em Github-Eviews após a estimação do modelo $AR(1)$ se forem usado os comando das linhas 21 a 27 do programa é possível extrair os componentes filtrados e suavizados para o vetor de estado.

O gráfico 8 abaixo apresenta os componentes filtrados e suavizados para o modelo $AR(1)$ gerado acima. Observe que como não temos erro na equação de observação o componente suavizado é igual a própria série.

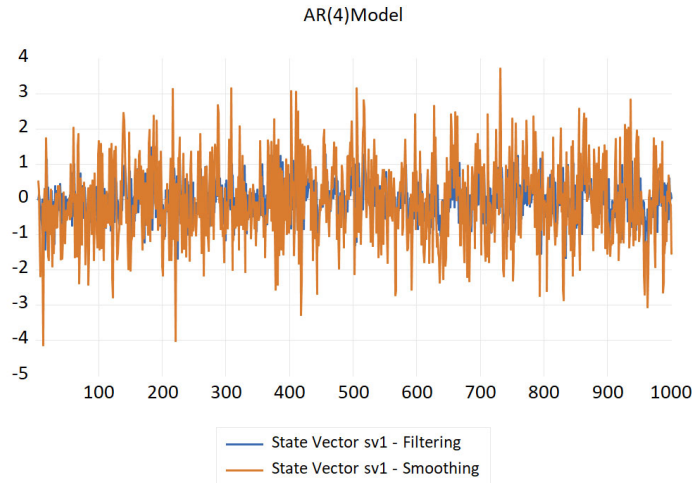
Figura 8: Componente Filtrado e Suavizado para o $AR(1)$



O mesmo para o modelo $AR(4)$ linhas 36 até 42 do programa é possível extrair os componentes filtrados e suavizados para o vetor de estado.

O gráfico 9 abaixo apresenta os componentes filtrados e suavizados do primeiro elemento do vetor de estado para o modelo $AR(4)$ gerado acima. Observe que como não temos erro na equação de observação o componente suavizado é igual a própria série.

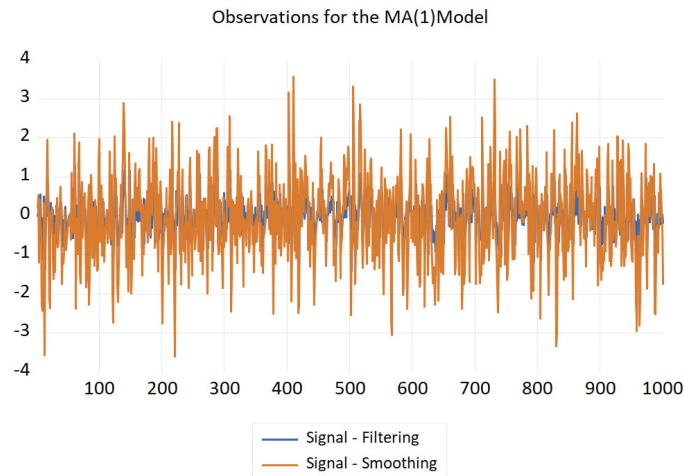
Figura 9: Componente Filtrado e Suavizado para o $AR(4)$



Para o modelo $MA(1)$, as linhas 66 até 73 do programa é possível extrair os componentes filtrados e suavizados para o vetor de estado.

O gráfico 10 abaixo apresenta os componentes filtrados e suavizados para o sinal da representação em espaço de estado do modelo $MA(1)$ gerado acima. Observe que por construção o componente suavizado é igual a própria série.

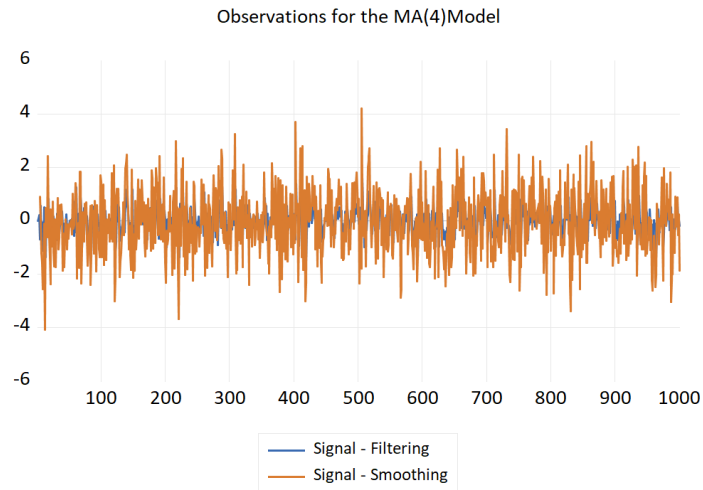
Figura 10: Componente Filtrado e Suavizado para o $MA(1)$



Para o modelo $MA(4)$ linhas 92 até 99 do programa é possível extrair os componentes filtrados e suavizados para o vetor de estado.

O gráfico 11 abaixo apresenta os componentes filtrados e suavizados para o sinal da representação em espaço de estado do modelo $MA(4)$ gerado acima. Observe que por construção o componente suavizado é igual a própria série.

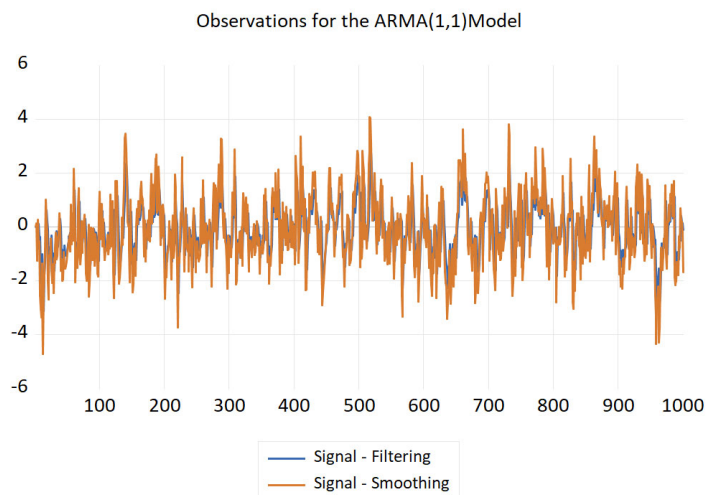
Figura 11: Componente Filtrado e Suavizado para o $MA(4)$



Para o modelo $ARMA(1,1)$ linhas 114 até 121 do programa é possível extrair os componentes filtrados e suavizados para o vetor de estado.

O gráfico 12 abaixo apresenta os componentes filtrados e suavizados para o sinal da representação em espaço de estado do modelo $ARMA(1,1)$ gerado acima. Observe que por construção o componente suavizado é igual a própria série.

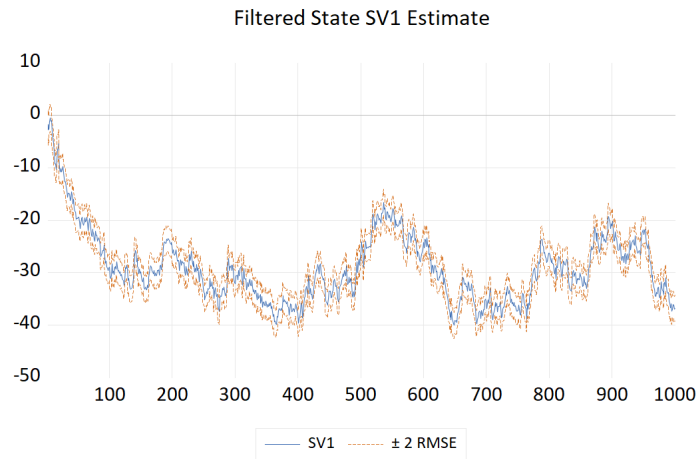
Figura 12: Componente Filtrado e Suavizado para o $ARMA(1,1)$



Para o modelo de nível local linhas 114 até 121 do programa é possível extrair os componentes filtrados para o nível μ_t e para o sinal y_t .

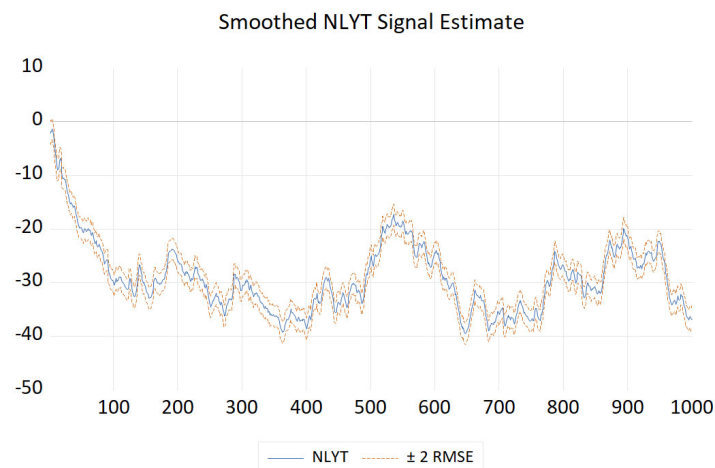
A seguir é, gráfico 13, apresentado o componente filtrado do vetor de estado do modelo de nível local. Observe que este componente representa o nível da série original.

Figura 13: Estimativa Filtrada para o Vetor de Estado do Modelo de Nível Local



A seguir é, gráfico 14, apresentado o componente filtrado das observações no modelo de nível local.

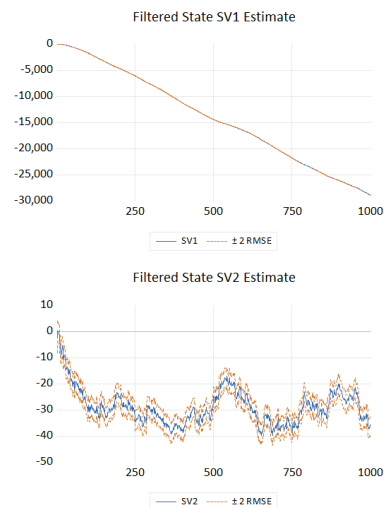
Figura 14: Estimativa Filtrada para as Observações no Modelo de Nível Local



Para o modelo de tendência local as linhas 175 até 180 do programa é possível extrair os componentes filtrados para o nível μ_t e para o sinal y_t . :

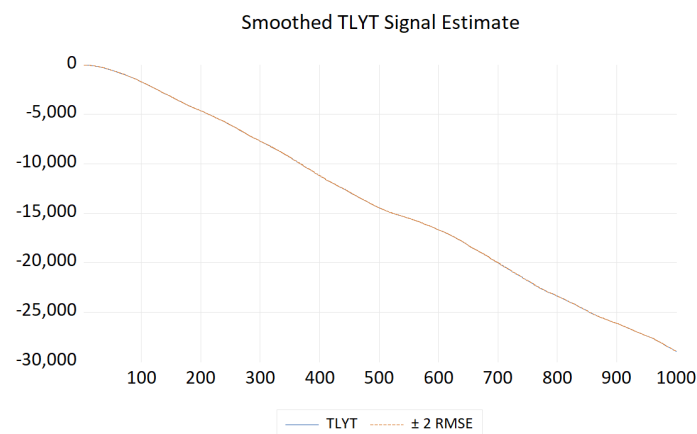
A seguir, gráficos 15, são apresentados os componentes filtrados do vetor de estado do modelo de tendência local. Observe que o primeiro componente representa o nível da série original e o segundo a taxa de crescimento.

Figura 15: Componentes Filtrados do Modelo de Tendência Local



A seguir, gráfico 16, é apresentado o componente filtrado das observações no modelo de tendência local.

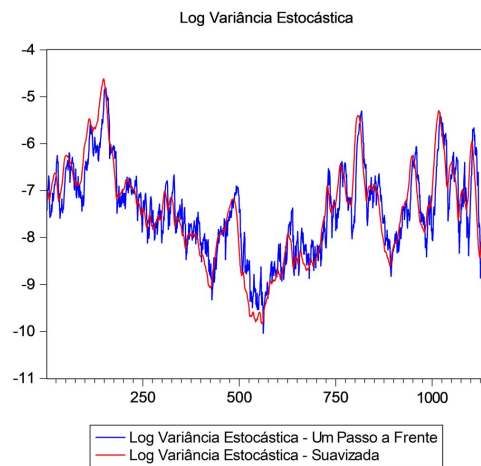
Figura 16: Componente Filtrado das Observações no Modelo de Tendência Local



Para o modelo de volatilidade estocástica as linhas 200 até 226 do programa é possível extrair os componentes filtrados e suavizados e as bandas pra o Modelo de Volatilidade Estocástica.

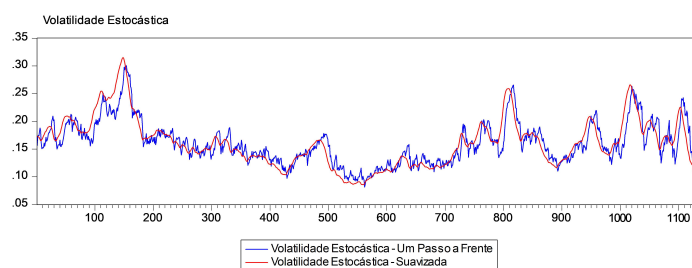
O gráfico 17, a seguir, apresenta o componente filtrado (um passo à frente) e o suavizado para o Logaritmo da Variância Estocástica

Figura 17: Log Variância para o Modelo de Volatilidade Estocástica



O gráfico a seguir apresenta os componentes filtrados e suavizados do Desvio Padrão Condicional claculado pelo modelo de Volatilidade Estocástica.

Figura 18: Volatilidade Estocástica

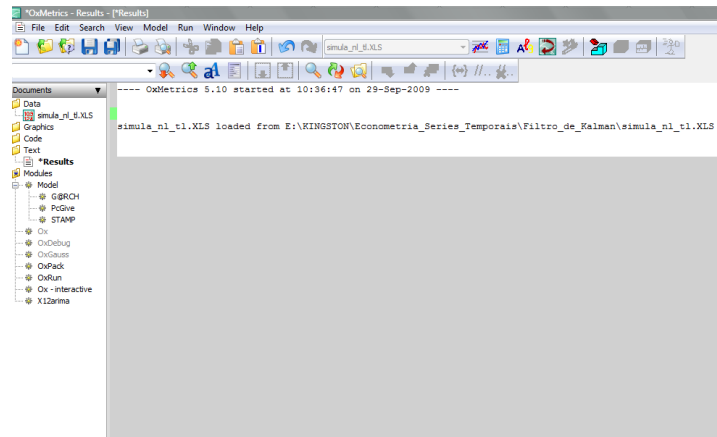


3.3.3 Estimando o Modelo de Nível Local e de Tendência Local no OxMetrics usando o STAMP

A tela de entrada do OxMetrics é apresentada a seguir onde já foi importado o arquivo `simula.nl.tl.xls` que contém os dados simulados para o modelo de nível local e de tendência

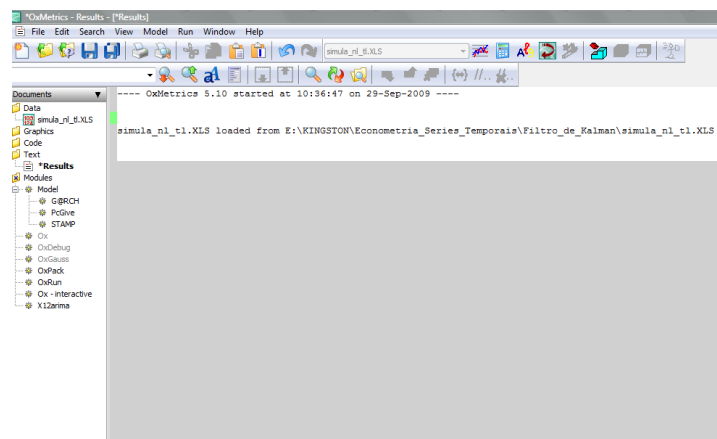
local.

Figura 19: Tela inicial do OxMetrics



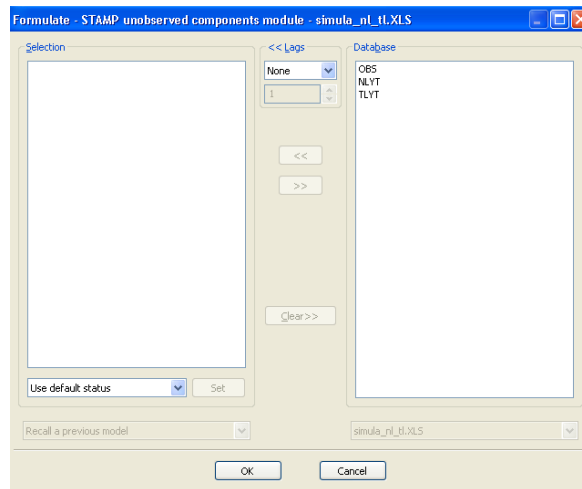
Para ativar os módulos do OxMetrics, clique em **Model** na barra a esquerda que aparecerá a seguinte tela:

Figura 20: Modulos do OxMetrics



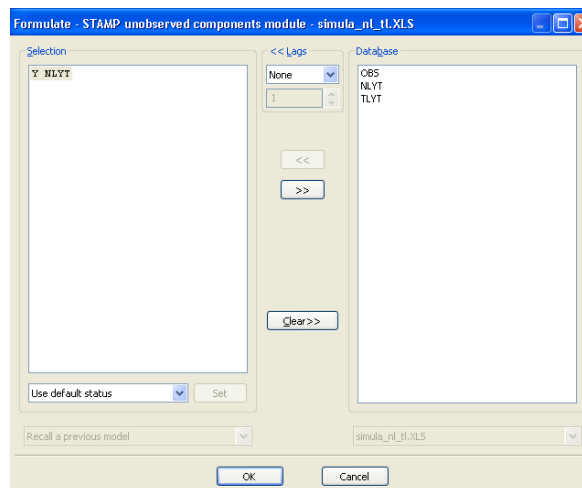
A seguir, ative o modulo STAMP, clicando em **Formulate** e aparecer á a seguinte caixa de diálogo

Figura 21: Seleção das Variáveis



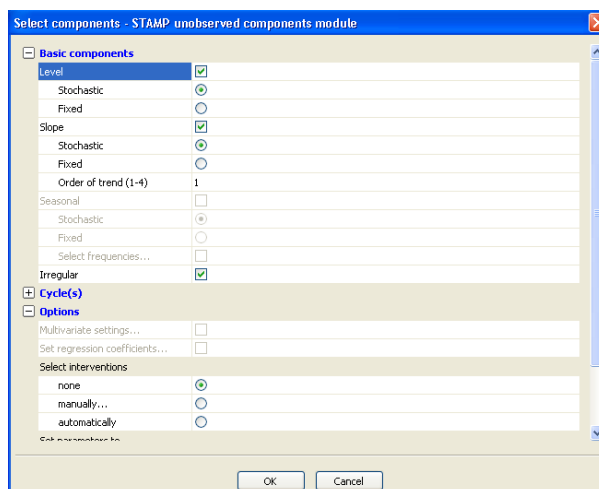
e clicando na variável a ser modelada na caixa **Database** o botão será ativado e aparecerá a seguinte caixa de diálogo

Figura 22: Escolha da variável a ser modelada



Após clicar em **OK** aparecerá a caixa de diálogo que especificação o modelo estrutural de séries temporais, isto é,

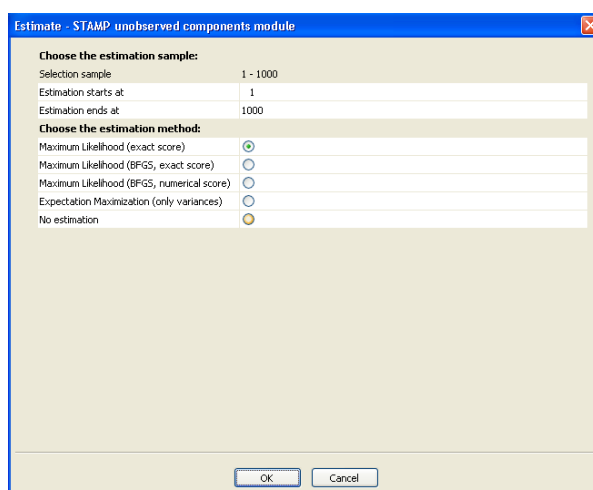
Figura 23: Formulação do Modelo Estrutural



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 (BFGS⁵ exact score) e Maximum Likelihood (BFGS numerical score)) e Estimação só das variâncias (Expectation Maximization (only variances)).

Figura 24: Métodos de Estimação



Obtém-se os seguintes resultados

⁵BFGS é um método de estimação não linear do tipo Quase-Newton.

```

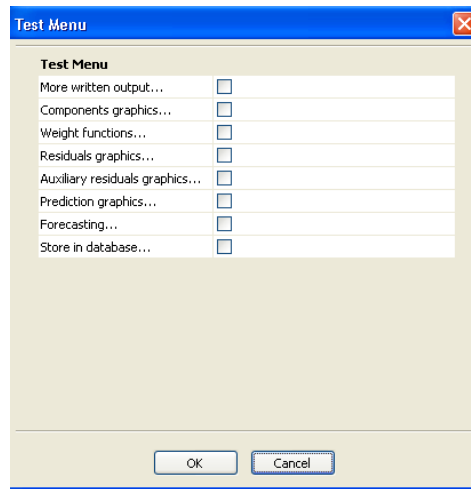
Summary statistics
  NLYT
  T 1000.Log-Likelihood is -946.8826.
Prediction error variance is 6.5965
Summary statistics
  NLYT
  T      1000.0
  p      2.0000
  std.error 2.5710
  Normality 0.3457
  H(332)   1.0293
  DW       1.9762
  r(1)     0.0083301
  q        32.000
  r(q)     -0.027046
  Q(q,q-p) 24.892
  Rd^2     0.25821
  Variances of disturbances:
            Value   (q-ratio)
Level      1.06100 ( 0.2679)
Slope      0.000000 ( 0.0000)
Irregular  3.96060 ( 1.000)
State vector analysis at period 1000
            Value   Prob
Level -35.85080 [0.0000]
Slope -0.03465 [0.28862]
Goodness-of-fit based on Residuals NLYT
                                     Value
Prediction error variance (p.e.v)    6.5965
Prediction error mean deviation (m.d) 5.2801
Ratio p.e.v. / m.d in squares        0.99362
Coefficient of determination R^2      0.86358
... based on differences Rd^2        0.25821
Information criterion Akaike (AIC)    1.8925
... Bayesian Schwartz (BIC)          1.9073

```

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 -0.03465 que não é estatisticamente significativo.

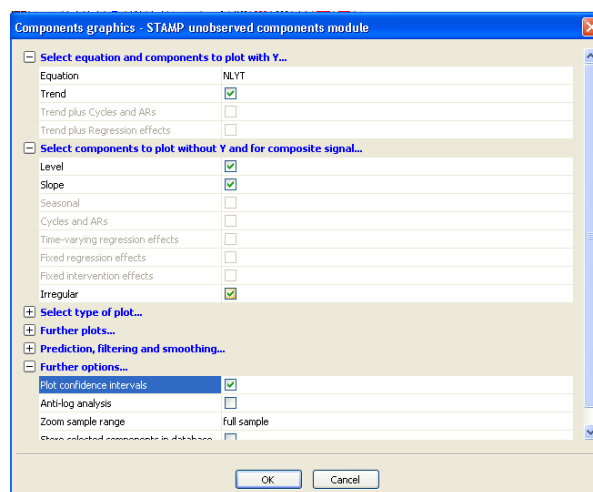
Para se obter as estimativas dos componentes do vetor de estado e do irregular temos clicar em **Model** e depois em **Test** obtendo a seguinte caixa de diálogo

Figura 25: Test Menu



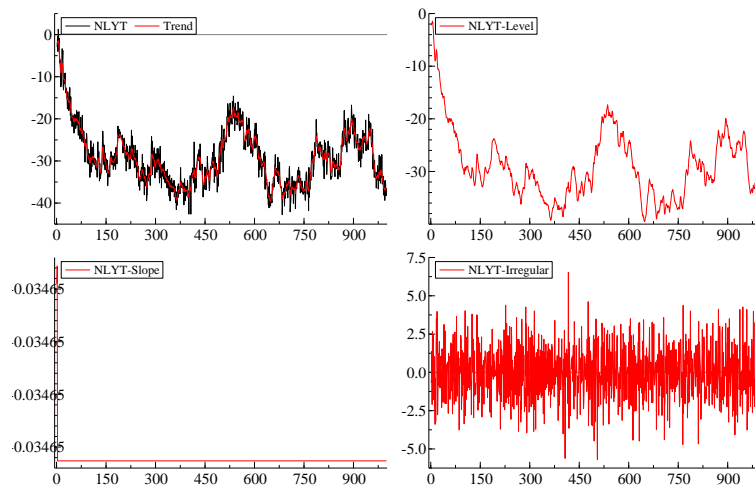
Clicando em **Components graphics**, obtém-se a seguinte caixa de diálogo

Figura 26: Menu do Gráfico dos Componentes



Temos então os gráficos da 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 27: Componentes Suavizados Tendencia Local



Estimando o modelo de nível local, obtém-se as seguintes estimativas

Log-Likelihood is -943.5239

Prediction error variance is 6.6028

Summary statistics

NLYT

T 1000.0

p 1.0000

std.error 2.5708

Normality 0.43727

H(333) 1.0203

DW 1.9755

r(1) 0.0111113

q 31.000

r(q) -0.034358

Q(q,q-p) 22.971

R² 0.86331

Variances of disturbances:

Value (q-ratio)

Level 1.06241 (0.2683)

Irregular 3.95951 (1.000)

State vector analysis at period 1000

Value Prob

Level -36.7998 [0.0000]

Goodness-of-fit based on Residuals NLYT

Value

Prediction error variance (p.e.v) 6.6028

Prediction error mean deviation (m.d) 5.2819

Ratio p.e.v. / m.d in squares 0.99482

Coefficient of determination R² 0.86331

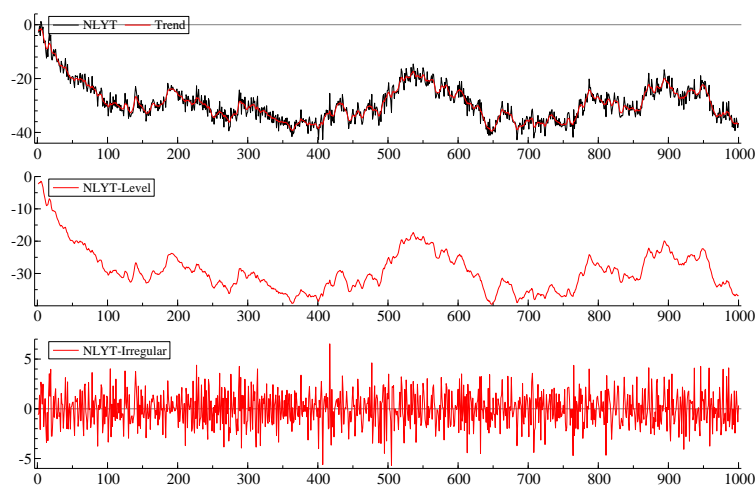
... based on differences Rd² 0.25676

Information criterion Akaike (AIC) 1.8915

... Bayesian Schwartz (BIC) 1.9013

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 28: Componentes Suavizados Nível Local



3.3.4 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 função de previsão e o erro médio quadrático de previsão.
5. Escreva um programa no EVIEWS para obter o previsor pontual assim como o erro médio quadrático de previsão para um modelo $AR(1)$.
6. Escreva um programa no EVIEWS 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.wf1`, estime um modelo de nível local e também um modelo de tendência local. Qual dos dois você escolheria e porque?
8. No arquivo `bolsa.xls` 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 o período sazonal. Podemos pensar que a função de previsão tem a propriedade de somar zero a cada s períodos. Este componente pode ser determinístico 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 dos 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:

$$y_t = \phi_4 y_{t-4} + \varepsilon_t \quad (63)$$

Para que este modelo seja estacionário precisamos 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:

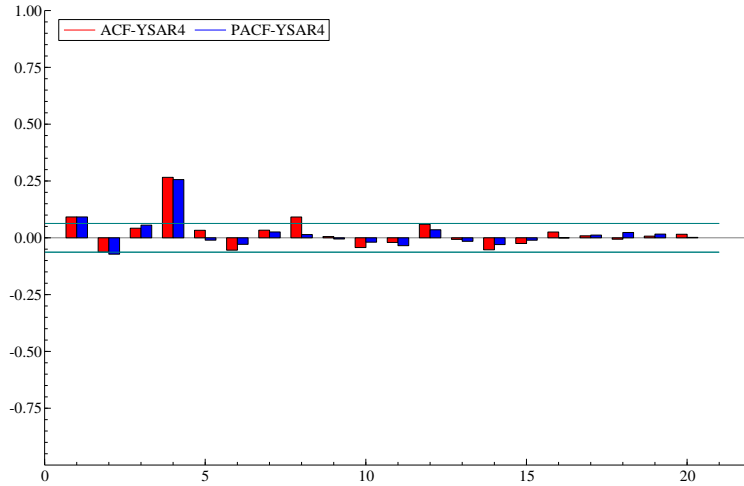
$$\rho(\tau) = \begin{cases} \phi_4^\tau & \text{se } \tau = 0, 4, 8, \dots \\ 0 & \text{caso contrário} \end{cases} \quad (64)$$

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

No programa em Github-Eviews nas linhas 230 até 245 simula um *AR*(4) sazonal, faz a FAC e FACP para até 12 defasagens e estima em representação em espaço de estado o *AR*(4).

O correlograma é apresentado abaixo e podemos observar uma FAC que tem o padrão de (64) e a FACP sómente a quarta autocorrelação parcial é diferente de zero implicando num modelo *ARMA*(1, 0)₄.

Figura 29: FAC e FACP para $ARMA(1,0)_4$



Abaixo é apresentada a estimação deste modelo por Filtro de Kalman. Observe que $\exp[C(1)]$ é a estimativa da variância do choque que é próxima do valor verdadeiro que é um e $C(2)$ representa a estimativa do coeficiente do autorregressivo que é próximo do valor verdadeiro que é 0,3.

Sspace: SSSAR4
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/17/22 Time: 17:46
Sample: 5 1000
Included observations: 996
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.002385	0.042973	-0.055491	0.9557
C(2)	0.268039	0.030528	8.780221	0.0000
	Final State	Root MSE	z-Statistic	Prob.
SV1	-0.067445	0.998808	-0.067525	0.9462
SV2	-1.341113	0.000000	NA	0.0000
SV3	-0.502444	0.000000	NA	0.0000
SV4	0.204124	0.000000	NA	0.0000
Log likelihood	-1412.224	Akaike info criterion		2.839808
Parameters	2	Schwarz criterion		2.849655
Diffuse priors	0	Hannan-Quinn criter.		2.843551

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

$$y_t = \varepsilon_t + \theta_4 \varepsilon_{t-4} \quad (65)$$

neste caso a função de autocorrelação é dada por:

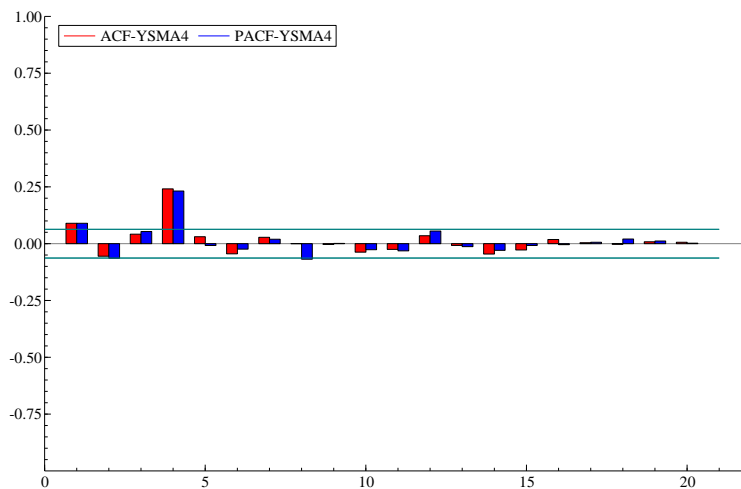
$$\rho(\tau) = \begin{cases} 1 & \text{se } \tau = 0 \\ \frac{\theta_4}{1-\theta_4^2} & \text{se } \tau = 4 \\ 0 & \text{caso contrário} \end{cases} \quad (66)$$

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

No programa em Github-Eviews nas linhas 249 até 265 simula um $MA(4)$ sazonal, faz a FAC e FACP para até 12 defasagens e estima em representação em espaço de estado o $MA(4)$

O correlograma é apresentado abaixo e podemos observar uma FAC que tem o padrão de (66) 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 30: FAC e FACP para $ARMA(0, 1)_4$



Abaixo é apresentada a estimação deste modelo por Filtro de Kalman. Observe que $\exp[C(2)] = 0.99734$ é a estimativa da variância do choque que é próxima do valor verdadeiro que é um e $C(1)$ representa a estimativa do coeficiente da média-móvel que é próximo do valor verdadeiro que é 0,3.

Sspace: SSSMA4
Method: Maximum likelihood (BFGS / Marquardt steps)
Date: 08/17/22 Time: 17:46
Sample: 5 1000
Included observations: 996
Convergence achieved after 10 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.265008	0.030607	8.658376	0.0000
C(2)	-0.002661	0.042929	-0.061989	0.9506
	Final State	Root MSE	z-Statistic	Prob.
SV1	0.000000	0.998670	0.000000	1.0000
SV2	-1.560702	1.49E - 08	-1.05E + 08	0.0000
SV3	-0.352358	1.49E - 08	-23646378	0.0000
SV4	0.028102	1.49E - 08	1885870.	0.0000
SV5	-0.055741	1.49E - 08	-3740697.	0.0000
Log likelihood	-1412.083	Akaike info criterion		2.839524
Parameters	2	Schwarz criterion		2.849371
Diffuse priors	0	Hannan-Quinn criter.		2.843268

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

$$\Phi(L^s) y_t = \Theta(L^s) \xi_t \quad (67)$$

onde

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

e

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

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:

$$\xi_t = \frac{\Theta(L)}{\Phi(L)} \varepsilon_t \quad (68)$$

então o modelo geral é:

$$\Phi(L^s) \Phi(L) y_t = \Theta(L^s) \Theta(L) \varepsilon_t \quad (69)$$

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 (69). 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:

$$y_t = \gamma_t + v_t \quad (70)$$

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

Nesta representação os componentes, sazonais e não sazonais, são aditivos enquanto que em (69) eles são multiplicativos.

4.3 Modelos *ARIMA* Sazonais

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

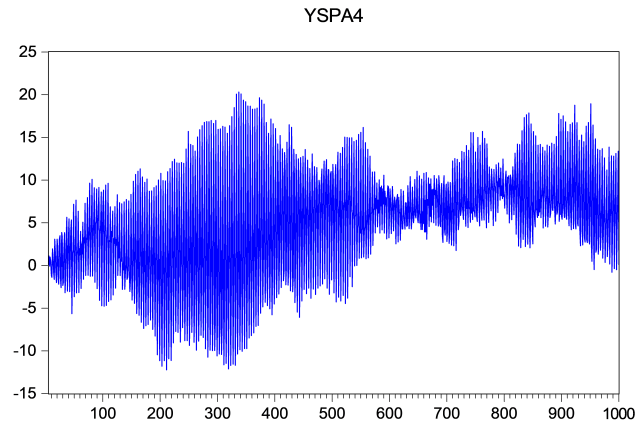
$$\Delta_s = (1 - L^s) \quad (71)$$

Uma série y_t que tenha uma raiz unitária sazonal, se tornará estacionária após a aplicação do operador definido em (71). Um caso particular é apresentado a seguir

$$\begin{aligned} \Delta_s y_t &= \varepsilon_t \\ y_t &= y_{t-s} + \varepsilon_t \end{aligned} \quad (72)$$

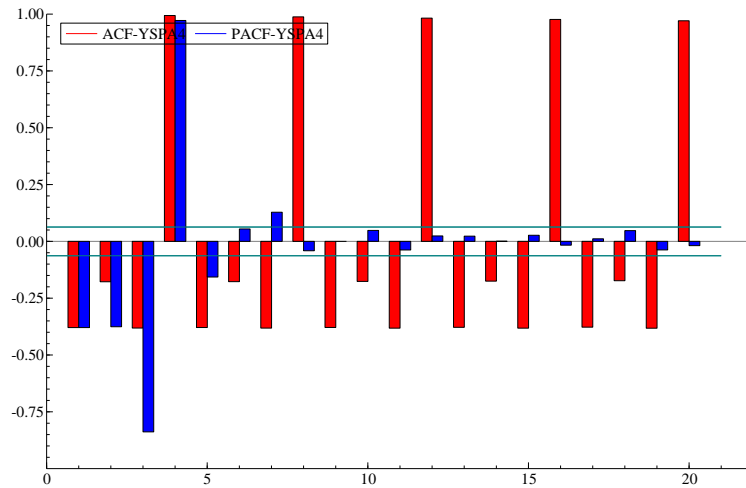
No programa em Github-Eviews nas linhas 267 até 275 simula um "passeio aleatório sazonal", assumindo que $s = 4$;

Figura 31: Diferença Sazonal com $s = 4$ - Passeio Aleatório sazonal



A FAC e FACP para esta série é apresentada a seguir

Figura 32: FAC e $FACP$ para "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 $FACP$ somente a quarta é significativa.

Então (67) 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 (69) 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:

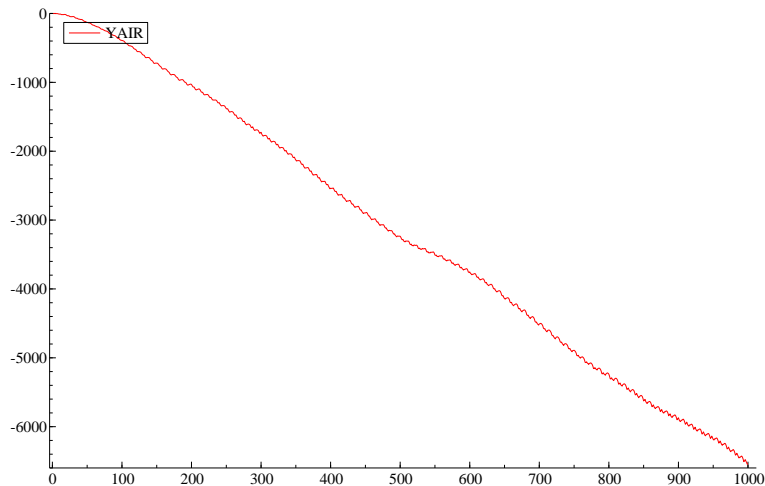
$$\Phi(L^s) \Phi(L)(1 - L)^d(1 - L^s)^D y_t = \Theta(L^s)\Theta(L)\varepsilon_t$$

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

No programa em Github-Eviews nas linhas 279 até 300 simula o modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$

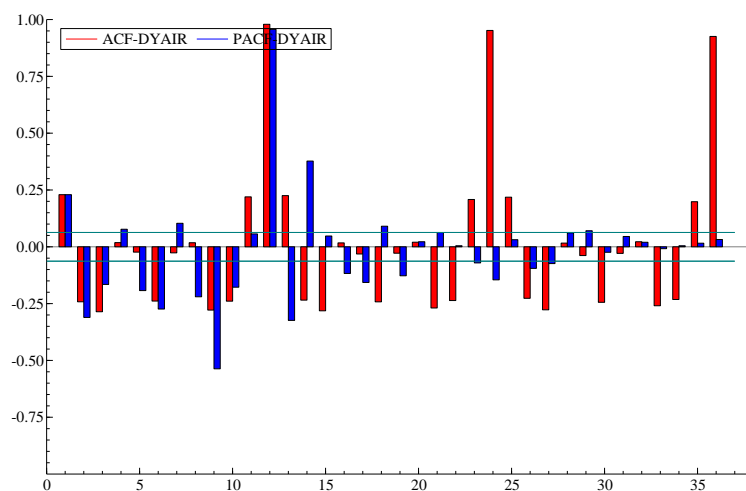
O gráfico a seguir apresenta a série gerada segundo o modelo *Airline Passanger*. Observe que esta série é claramente não estacionária, mas não é possível saber qual a ordem de integração.

Figura 33: $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$



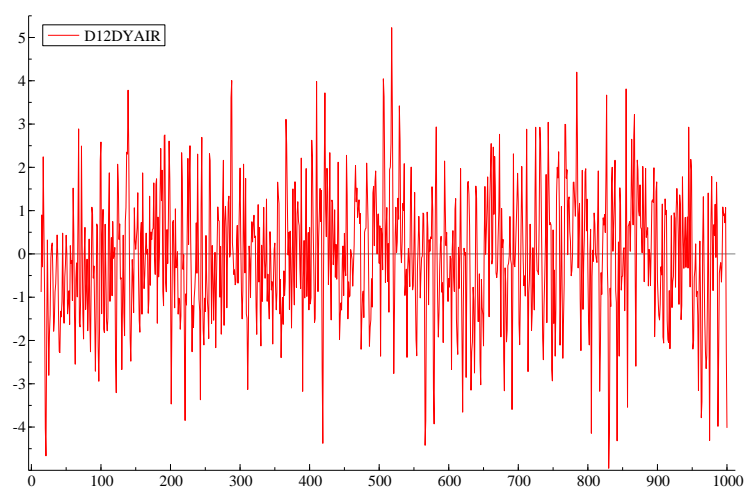
Após diferenciar a série, fizemos a FAC e $FACP$ e observamos que a série parece ter uma raiz unitária em frequência sazonal, uma vez que a FAC na defasagem 12 é muito perto de um e o decaimento em múltiplo de 12 é bem lento. Desta forma vamos diferenciar a série usando a diferença sazonal.

Figura 34: FAC e FACP $(1 - L)y_t$



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

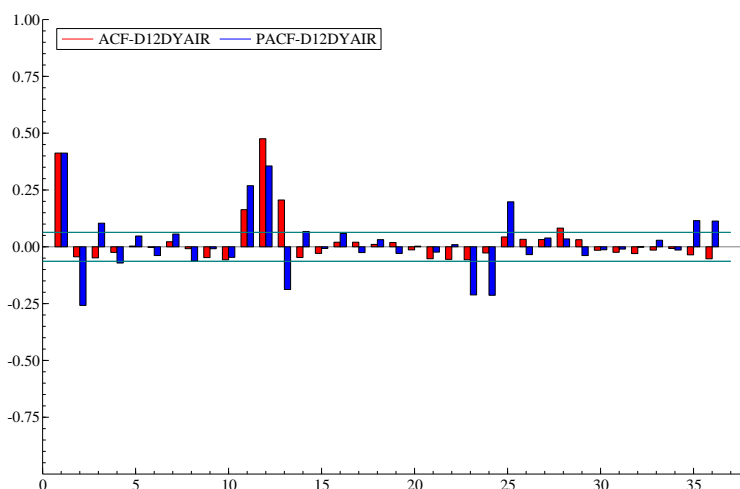
Figura 35: $(1 - L)(1 - L^{12})y_t$



e parece que está série é estacionária.

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

Figura 36: $(1 - L)(1 - L^{12})y_t$



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 $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$.

A seguir é apresentada a estimativas dos parâmetros deste modelo

Figura 37: Estimativa do $(1 - L)(1 - L^{12})y_t$

Dependent Variable: DD12YAIR
Method: ARMA Maximum Likelihood (BFGS)
Date: 08/17/22 Time: 17:46
Sample: 14 1000
Included observations: 987
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.096374	0.090570	-1.064083	0.2876
MA(1)	0.590231	0.026274	22.46414	0.0000
SMA(12)	0.808570	0.020006	40.41691	0.0000
SIGMASQ	0.992471	0.042857	23.15780	0.0000
R-squared	0.546971	Mean dependent var	-0.082957	
Adjusted R-squared	0.545589	S.D. dependent var	1.480867	
S.E. of regression	0.998253	Akaike info criterion	2.851752	
Sum squared resid	979.5685	Schwarz criterion	2.871588	
Log likelihood	-1403.339	Hannan-Quinn criter.	2.859296	
F-statistic	395.6136	Durbin-Watson stat	2.001633	
Prob(F-statistic)	0.000000			
Inverted MA Roots	.95+.25i	.95-.25i	.69-.69i	.69+.69i
	.25+.95i	.25-.95i	-.25+.95i	-.25-.95i
	-.59	-.69+.69i	-.69+.69i	-.95-.25i
	-.95+.25i			

4.4 Sazonalidade Determinística

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

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

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

$$y_t = \sum_{j=1}^s \gamma_j z_{jt} + \varepsilon_t \quad (73)$$

ou então:

$$y_t = \mu + \sum_{j=1}^{s-1} \gamma_j z_{jt} + \varepsilon_t \quad (74)$$

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:

$$z_{jt} = \begin{cases} 1 & \text{se } t = j, j + s, j + 2s, \dots \\ -1 & \text{se } t = s, 2s, 3s, \dots \\ 0 & \text{se caso contrário} \end{cases} \quad (75)$$

e a especificação usada é a (74).

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

$$y_t = \sum_{j=1}^{\lfloor s/2 \rfloor} (\gamma_j \cos \lambda_j t + \gamma_j^* \sin \lambda_j t) + \varepsilon_t \quad (76)$$

com $\lambda_j = \frac{2\pi j}{s}$ para $j = 1, \dots, \lfloor \frac{s}{2} \rfloor$.

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:

$$(1 - L^s)\gamma_t = \omega_t \Leftrightarrow \gamma_t = \gamma_{t-s} + \omega_t \quad (77)$$

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

$$y_t = \gamma_t + \varepsilon_t \quad (78)$$

Esta especificação tem problemas uma vez que o operador $(1 - L^s)$ mistura nível com componente sazonal uma vez que $(1 - L^s) = (1 - L)(1 + L + L^2 + \dots + L^{s-1})$ 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:

$$\gamma_t + \gamma_{t-1} + \dots + \gamma_{t-s+1} = 0 \quad (79)$$

e permitindo que (79) varie suavemente temos que esta especificação pode ser re-escrita da seguinte forma:

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \Leftrightarrow \gamma_t = - \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \quad (80)$$

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:

$$\gamma_{T+l} | T = - \sum_{j=1}^{s-1} \gamma_{T+l-1} | T \quad (81)$$

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 predictor 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 teríamos a seguinte representação em espaço de estado:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \quad (82)$$

$$\mu_t = \mu_{t-1} + \eta_t \quad (83)$$

$$\gamma_t = - \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \quad (84)$$

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 (82-84) em forma matricial é dada por:

$$y_t = (1 : 1 : 0_{s-2})\alpha_t + \varepsilon_t \quad (85)$$

$$\alpha_t = \begin{bmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ . \\ . \\ . \\ \gamma_{t-s+2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & -1 & -1 & . & . & . & -1 \\ 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \gamma_{t-1} \\ \gamma_{t-2} \\ . \\ . \\ . \\ \gamma_{t-s+1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \omega_t \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} \quad (86)$$

É fácil mostrar que a forma reduzida é dada por:

$$(1 - L)S(L)y_t = S(L)\eta_t + (1 - L)\omega_t + (1 - L)S(L)\varepsilon_t \quad (87)$$

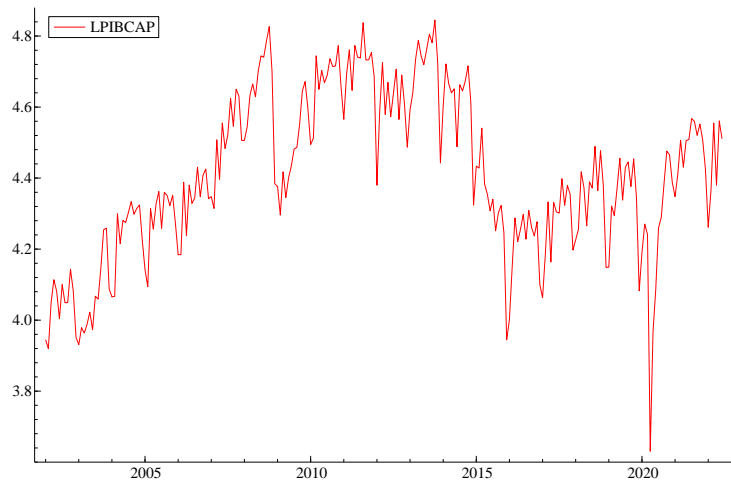
onde o lado direito de (87) pode ser escrito como um $MA(s)$ e, portanto, a forma reduzida para y_t é um $ARIMA(0, 1, 1)_s$ mas com o MA sazonal tendo coeficientes não nulos para defasagens menores do que s .

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" que é considerado um dos modelos mais parcimoniosos para dados sazonais.

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

A série de logaritmo da Produção Industrial Bens de Capital para o período de janeiro de 2002 até junho de 2022 é apresentada abaixo

Figura 38: Log do Índice de Produção Industrial - Bens de Capital



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.

Observe que na caixa de diálogo que especifica os componentes do modelo estrutural teremos a componente sazonal ativado.

Estimando este modelo temos os seguintes resultados **estes valores devem ser modificados**.

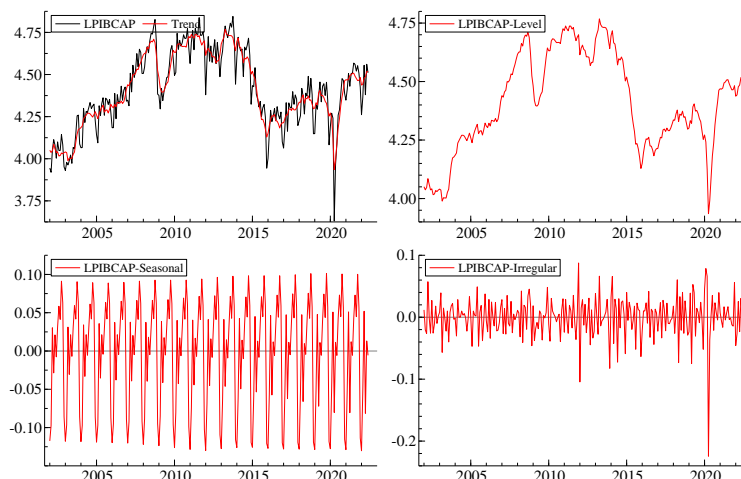
```

The model is: Y = Level + Seasonal + Irregular
Steady state..... found without full convergence
Log-Likelihood is 1019.03 (-2 LogL = -2038.07).
Prediction error variance is 0.00426359
Summary statistics
      LBCAP
T      399.00
p      2.0000
std.error 0.065296
Normality 100.79
H(129)   0.75540
DW       2.0352
r(1)     -0.019178
q       24.000
r(q)     -0.14861
Q(q,q-p) 61.365
Rs^2     0.15244
Variances of disturbances:
      Value      (q-ratio)
Level    0.00176279 ( 1.098)
Seasonal 0.000000   ( 0.0000)
Irregular 0.00160514 ( 1.000)
State vector analysis at period 2008(3)
      Value      Prob
Level    5.17035 [0.00000]
Seasonal chi2 test 352.64977 [0.00000]
Seasonal effects:
Period   Value      Prob
  1      -0.12454 [0.00000]
  2      -0.08890 [0.00000]
  3       0.02260 [0.02171]
  4      -0.05041 [0.00000]
  5       0.02234 [0.02424]
  6       0.02714 [0.00630]
  7       0.04536 [0.00001]
  8       0.07170 [0.00000]
  9       0.04904 [0.00000]
 10       0.07173 [0.00000]
 11       0.03687 [0.00022]
 12      -0.08292 [0.00000]
Goodness-of-fit based on Residuals LBCAP
Value
Prediction error variance (p.e.v)      0.0042636
Prediction error mean deviation (m.d)   0.0032427
Ratio p.e.v. / m.d in squares          1.1006

```

Coefficient of determination R^2	0.89814
... based on differences R_d^2	0.51494
... based on diff around seas mean R_s^2	0.15244
Information criterion Akaike (AIC)	-5.3925
... Bayesian Schwartz (BIC)	-5.2625

Figura 39: Componentes do Log do Índice de Produção Industrial - Bens de Capital



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 cosseno. Temos, então:

$$y_t = \cos(\lambda t) \quad (88)$$

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

Uma representação alternativa a (88) é aquela em que esta função periódica tem uma amplitude e uma fase. Então (88) pode ser reescrita da seguinte forma:

$$y_t = \rho \cos(\lambda t - \theta) \quad (89)$$

onde ρ representa a amplitude e θ a fase.

Observe que (89) pode ser reescrito da seguinte forma⁶:

$$y_t = \alpha \cos(\lambda t) + \beta \sin(\lambda t) \quad (90)$$

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

⁶Usando a seguinte igualdade trigonométrica: $\cos(a - b) = \cos(a) \cos(b) + \sin(a) \sin(b)$.

Podemos então especificar o componente cíclico da seguinte forma:

$$y_t = \psi_t + \varepsilon_t \quad \text{para } t = 1, \dots, T \quad (91)$$

e

$$\psi_t = \alpha \cos(\lambda_c t) + \beta \sin(\lambda_c t) \quad (92)$$

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 (91-92) podem ser estimados por M.Q.O., como no modelo linear de regressão já que os regressores são não estocásticos. Observe que (91-92) é uma representação de um componente cíclico determinístico.

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:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} \cos(\lambda_c) & \sin(\lambda_c) \\ -\sin(\lambda_c) & \cos(\lambda_c) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} \quad \text{para } t = 1, \dots, T \quad (93)$$

Observe que a forma reduzida de (93) é dada por:

$$\psi_t = 2 \cos(\lambda_c) \psi_{t-1} - \psi_{t-2} \quad (94)$$

e quando $\alpha = \psi_0$ e $\beta = \psi_0^*$ temos que (94) é equivalente a (92).

Uma forma de permitir que este componente seja estocástico é introduzir dois choques em (93) não correlacionados entre si mas com a restrição de que a variância é igual. Observe que pela forma reduzida dada em (94) 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:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos(\lambda_c) & \sin(\lambda_c) \\ -\sin(\lambda_c) & \cos(\lambda_c) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \quad (95)$$

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 reduz a um ciclo determinístico quando $\sigma_\kappa^2 \rightarrow 0$ quando $\rho \rightarrow 1$.

A forma reduzida de (95) é dada por:

$$\psi_t - 2\rho \cos(\lambda_c) \psi_{t-1} + \rho^2 \psi_{t-2} = \kappa_t - \cos(\lambda_c) \kappa_{t-1} + \sin(\lambda_c) \kappa_{t-1}^* \quad (96)$$

onde o lado direito de (96) 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 (91) e (95) é dada por:

$$y_t - 2\rho \cos(\lambda_c) y_{t-1} + \rho^2 y_{t-2} = (1 - 2\rho \cos(\lambda_c) L + \rho^2 \cos(\lambda_c) L^2) \varepsilon_t + (1 - \cos(\lambda_c) L) \kappa_t + \sin(\lambda_c) \kappa_{t-1}^* \quad (97)$$

onde o lado direito de (97) representa um processo $MA(2)$, logo temos que a forma

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

Observe que (91) e (95) 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:

$$\psi_{T+l|T} = \rho^l (\cos(\lambda_c l) \psi_T + \sin(\lambda_c l) \psi_T^*) \text{ para } l = 1, 2, \dots \quad (98)$$

5.2.1 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.

Temos as seguintes estimativas **Estew valores te que ser modificados**

The model is: Y = Trend + Seasonal + Irregular + Cycle 1 + Cycle 2 + Cycle 3

Steady state..... found without full convergence

Log-Likelihood is 1013.86 (-2 LogL = -2027.73).

Prediction error variance is 0.00419822

Summary statistics

	LBCAP
T	399.00
p	8.0000
std.error	0.064794
Normality	97.382
H(128)	0.76605
DW	2.0390
r(1)	-0.021930
q	26.000
r(q)	0.10651
Q(q,q-p)	67.127
Rs^2	0.16759

Variances of disturbances:

	Value	(q-ratio)
Level	0.00104496	(0.6429)
Slope	0.000000	(0.0000)
Seasonal	0.000000	(0.0000)
Cycle	8.46216e-005	(0.05206)
Cycle 2	3.81729e-009	(2.348e-006)
Cycle 3	0.000502067	(0.3089)
Irregular	0.00162548	(1.000)

Cycle other parameters:

Variance 0.00036

Period 46.01106

Period in years 3.83425

Frequency 0.13656

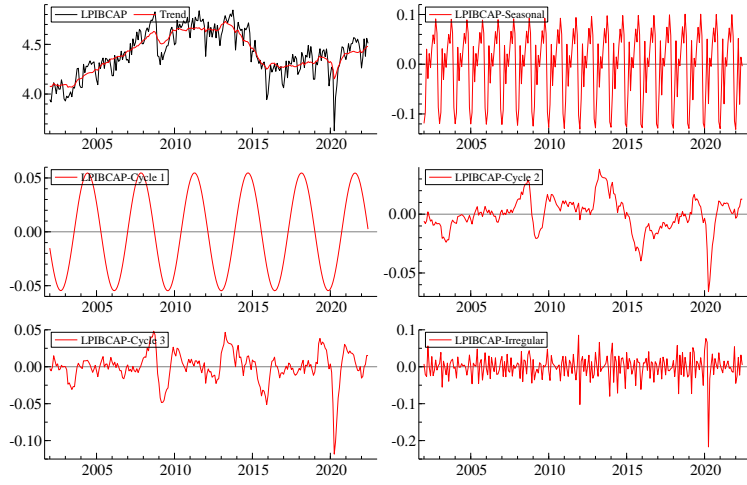
Damping factor 0.87486

Order 1.00000

Cycle 2 other parameters:
 Variance 0.00371
 Period 87.18720
 Period in years 7.26560
 Frequency 0.07207
 Damping factor 1.00000
 Order 1.00000
 Cycle 3 other parameters:
 Variance 0.00262
 Period 52.79611
 Period in years 4.39968
 Frequency 0.11901
 Damping factor 0.89907
 Order 1.00000
 State vector analysis at period 2008(3)
 Value Prob
 Level 5.11358 [0.00000]
 Slope 0.00129 [0.43141]
 Seasonal chi2 test 353.44316 [0.00000]
 Cycle 1 amplitude 0.01458 [.NaN]
 Seasonal effects:
 Period Value Prob
 1 -0.12481 [0.00000]
 2 -0.08912 [0.00000]
 3 0.02241 [0.02288]
 4 -0.05042 [0.00000]
 5 0.02241 [0.02385]
 6 0.02726 [0.00608]
 7 0.04551 [0.00001]
 8 0.07186 [0.00000]
 9 0.04918 [0.00000]
 10 0.07182 [0.00000]
 11 0.03689 [0.00022]
 12 -0.08298 [0.00000]
 Goodness-of-fit based on Residuals LBCAP
 Value
 Prediction error variance (p.e.v) 0.0041982
 Prediction error mean deviation (m.d) 0.0032013
 Ratio p.e.v. / m.d in squares 1.0949
 Coefficient of determination R^2 0.89996
 ... based on differences R_d^2 0.52361
 ... based on diff around seas mean R_s^2 0.16759
 Information criterion Akaike (AIC) -5.4029
 ... Bayesian Schwartz (BIC) -5.263

E os componentes estimados são apresentados abaixo

Figura 40: Componentes do Modelo com Ciclo



6 Modelo Estrutural Básico

O modelo estrutural básico é aquele que decompõe a série em tendência estocástica, sazonalidade estocástica, ciclo 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:

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos(\lambda_j) & \sin(\lambda_j) \\ -\sin(\lambda_j) & \cos(\lambda_j) \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_t \\ \omega_t^* \end{bmatrix} \quad (99)$$

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

O componente cíclico é dado por (95).

7 Modelo Estrutural com Variáveis Exógenas

Uma forma de introduzir variáveis "exógenas" nos modelos *ARMA* é através dos modelos chamados de função de transferência. Uma especificação simples é apresentada a seguir:

$$y_t = \frac{\omega(L)}{\lambda(L)} x_t + \frac{\theta(L)}{\phi(L) \Delta^d} \varepsilon_t \Leftrightarrow \Delta^d y_t = \frac{\omega(L)}{\lambda(L)} \Delta^d x_t + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (100)$$

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 (100) é através do seguinte modelo estrutural:

$$\begin{aligned}
y_t &= \mu_t + \delta x_t + \varepsilon_t \\
\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\
\beta_t &= \beta_{t-1} + \xi_t
\end{aligned} \tag{101}$$

uma vez que a forma reduzida de (101) é dada por:

$$\Delta^2 y_t = \delta \Delta^2 x_t + u_t \tag{102}$$

com

$$u_t = \Delta \eta_t + \xi_{t-1} + \Delta^2 \varepsilon_t \sim MA(2) \tag{103}$$

A diferença entre (100) e (101) 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 Distribuídas 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 estrutural 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.

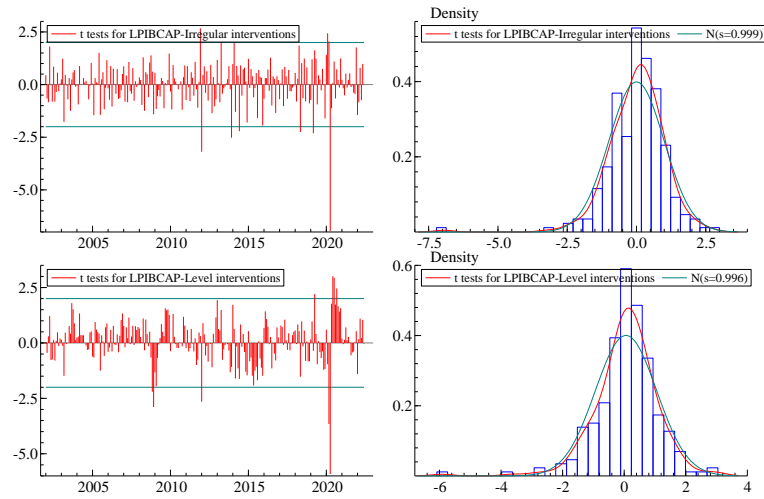
Uma forma de determinar se num dado período ocorreu um valor aberrante ou uma mudança estrutural é através dos resíduos auxiliares, veja Harvey and Koopman [1992], que são estimativas suavizadas do erro das observações e do erro do nível. No primeiro quando o resíduo for maior do que um certo número de desvios padrão este valor corresponde a valor aberrante e o segundo se o resíduo também for mais do que um certo número 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.

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

Obtém-se então os seguintes gráficos

Figura 41: Resíduos Auxiliares



Pelos gráficos da esquerda temos que para os resíduos do irregular, associados a valores aberrantes, dois períodos então fora da banda, a saber:

Tabela 9: Valores ≥ 3 para LPIBCAP - resíduos do Irregular

Date	Value	Prob
2012(1)	-3.18763	[0.00081]
2020(4)	-6.95236	[0.00000]

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

Tabela 10: Valores ≥ 3 para LPIBCAP - resíduos do Nível

Date	Value	Prob
2020(3)	-3.65335	[0.00016]
2020(4)	-5.91413	[0.00000]
2020(6)	3.01283	[0.00143]

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

Agora re-estimando o modelo anterior com as intervenções temos :

Figura 42: Estimação com intervenções

```
The database used is C:\Users\Pedro\Dropbox\Tópicos em Finanças_2022\graficos\lcap_2002_2022.in7
The selection sample is: 2002(1) - 2022(6) (N = 1, T = 246)
The dependent variable Y is: LPIBCAP
The model is: Y = Level + Seasonal + Irregular + Interventions

Profile Log-Likelihood: 587.5868
Akaike Information Criterion (AIC): -4.6227
Bayesian Information Criterion (BIC): -4.3519
Prediction error variance: 0.0047

Summary statistics:
LPICAP
T: 246
Normality: 0.0027
H(76): 1.4386
Du: 5.9969
r(1): -0.0082097
q: 24
p: 2
r(s): -0.999852
Q(q,p): 54.343
B-S: 0.35428

Variances of disturbances:
Level: 0.0016917 (0.001)
Seasonal: 6.7845e-07 (0.0003957)
Irregular: 0.00171471 (1.000)

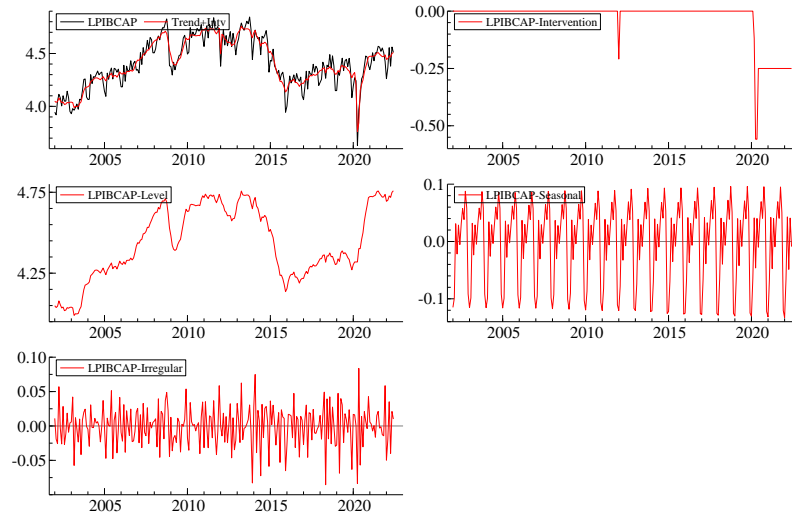
State vector analysis at period 2022(6):
Level: 4.79927 [0.00000]
Seasonal chi2 test: 127.10292 [0.00000]

Seasonal effects:
Period Value Prob
1 -0.11289 [0.00000]
2 -0.00512 [0.00193]
3 0.04289 [0.03518]
4 -0.04057 [0.01513]
5 0.00079 [0.04607]
6 -0.00894 [0.63272]
7 -0.00101 [0.60315]
8 0.00932 [0.00077]
9 0.00047 [0.04648]
10 0.00586 [0.00000]
11 0.05639 [0.01247]
12 -0.12055 [0.00000]

Regression effects in final state at time 2022(6):
Coefficient R^2 t-value Prob
Nuller 2012(1) -0.20886 0.07795 -1.08415 [0.00000]
Level break 2020(4) -0.43954 0.07097 -6.19367 [0.00000]
Level break 2020(3) -0.11912 0.07017 -1.69761 [0.09893]
Level break 2020(5) 0.30861 0.05005 4.74411 [0.00000]
```

Tem-se então os componentes estimados que são apresentados nos gráficos abaixo

Figura 43: Estimação com intervenções Final



Referências

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.

- J. D. Hamilton. *Time Series Analysis*. Princeton Univers. Press, 1994. ISBN 0691042896. URL https://www.ebook.de/de/product/3239264/james_hamilton_paterson_time_series_analysis.html.
- 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.
- R. S. Tsay and R. Chen. *Nonlinear Time Series Analysis*. Wiley & Sons, Limited, John, 2018. ISBN 9781119264057.
- P. L. Valls Pereira. Application of kalman filter. *Econometric Theory*, 3(02):306, apr 1987a. doi: 10.1017/s0266466600010392.
- P. L. Valls Pereira. Exact likelihood function for a regression model with MA(1) errors. *Economics Letters*, 24(2):145–149, jan 1987b. doi: 10.1016/0165-1765(87)90241-2.

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:

$$z_t \sim NI \left(\begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{bmatrix} \right) \quad (104)$$

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

$$D(z_t : \theta) = \begin{cases} D(y_t | x_t, \theta_1) D(x_t | \theta_2) \\ D(x_t | y_t, \psi_1) D(y_t | \psi_2) \end{cases} \quad (105)$$

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

$$E(y_t | x_t) = a + bx_t \quad (106)$$

onde

$$a = \mu_y - b\mu_x \quad (107)$$

e

$$b = \frac{\text{cov}(y_t, x_t)}{\text{var}(x_t)} = \frac{\sigma_{yx}}{\sigma_x^2} \quad (108)$$

e substituindo (107) e (108) em (106) temos:

$$E(y_t | x_t) = \mu_y - b\mu_x + bx_t = \mu_y + \frac{\sigma_{yx}}{\sigma_x^2}(x_t - \mu_x) \quad (109)$$

Agora a variância da distribuição condicional será dada por

$$\begin{aligned} Var(y_t | x_t) &= E(y_t - E(y_t | x_t) | x_t)^2 \\ &= E(y_t - \mu_y - \frac{\sigma_{yx}}{\sigma_x^2}(x_t - \mu_x))^2 \\ &= E(y_t - \mu_y)^2 + \left(\frac{\sigma_{yx}}{\sigma_x^2}\right)^2 E(x_t - \mu_x)^2 - 2\frac{\sigma_{yx}}{\sigma_x^2} E((y_t - \mu_y)(x_t - \mu_x)) \\ &= \sigma_y^2 + \left(\frac{\sigma_{yx}}{\sigma_x^2}\right)^2 \sigma_x^2 - 2\frac{\sigma_{yx}}{\sigma_x^2} \sigma_{yx} \\ &= \sigma_y^2 - \frac{\sigma_{yx}^2}{\sigma_x^2} \end{aligned} \quad (110)$$

Observe que (110) 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

$$Var(y_t | x_t) = \omega_{11} - \omega_{12}\omega_{22}^{-1}\omega_{21} \quad (111)$$

que é menor do que a variância 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:

$$z_t | I_{t-1} \sim NI \left(\begin{bmatrix} \mu_{y|I_{t-1}} \\ \mu_{x|I_{t-1}} \end{bmatrix}, \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \right) \quad (112)$$

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:

$$D(z_t | I_{t-1}, \theta) = D(y_t | I_{t-1}, x_t, \theta_1) D(x_t | I_{t-1}, \theta_2) \quad (113)$$

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:

$$E(y_t | I_{t-1}, x_t, \theta_1) = a^* + b^*x_t \quad (114)$$

onde

$$a^* = \mu_{y|I_{t-1}} - b^*\mu_{x|I_{t-1}} \quad (115)$$

e

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

A variância da distribuição condicional será dada por

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