NOTAS DE AULA – III

Componente de tendência em modelos estruturais

⇒ Modelos Estruturais

• Os ME fazem parte de uma classe particular de modelos pertencentes à abordagem de componentes não observáveis (CNO).

• Nos ME a ST é decomposta em <u>componentes ortogonais</u>, que possuem interpretação direta, tais como: tendência, sazonalidade e ciclo.

• As componentes são processos estocásticos não observáveis, sendo cada uma delas definida por uma equação de evolução.

• Por exemplo:

$$
\begin{array}{lll}
-\text{eq. das observ.:} & y_t = \mu_t + \varepsilon_t & \varepsilon_t \sim N(0, \sigma^2_\varepsilon) \\
-\text{eq. do estado:} & \mu_t = \mu_{t-1} + \eta_t & \eta_t \sim N(0, \sigma^2_n)
\end{array}
$$

$$\text{onde } E(\varepsilon_t \eta_s) = 0, \quad \forall\ t, s.$$

• Em termos de diagrama:

$$
\begin{aligned}
\text{eq. das observações: } y_t &= \text{Tend} + \text{Saz} + \text{Ciclo} + \text{Irreg} \\
&= \mu_t + \gamma_t + \psi_t + \varepsilon_t
\end{aligned}
$$

$$
\text{eqs. do estado:} \quad
\begin{array}{l}
\mu_t \Rightarrow \text{eq1} \\
\gamma_t \Rightarrow \text{eq2} \\
\psi_t \Rightarrow \text{eq3}
\end{array}
$$

• É importante observar que as componentes nos ME são processos estocásticos, i.e., evoluem dinamicamente no tempo.

• Esta característica torna as componentes “locais” ao invés de ”globais”, dando a flexibilidade necessária para o modelo seguir a evolução da ST ao longo do tempo.

• Os modelos onde as componentes são determinísticas, ou “globais”, que seriam estimados por MQO, são casos particulares dos ME.

eq.das obs: $y_t = \text{Tend} + \text{Saz} + \text{Irreg}$

$$= \mu_t + \gamma_t + \varepsilon_t$$

eqs. do estado: $\mu_t = a + bt$

$$\gamma_t = \sum_{i=1}^{s-1} [\gamma_j \cos(\omega_j t) + \gamma_j^* \text{sen}(\omega_j t)], \omega_j = 2\pi j/S$$

$$\varepsilon_t \sim NID(0, \sigma^2)$$

• Portanto, os ME podem ser vistos como modelos de regressão onde os regressores são funções do tempo (t, cos(wt), sen(wt), etc) e os coeficientes dos regressores (a, b, γ ‘s, γ *’s) evoluem estocásticamente no tempo.

- Uma vez especificada as equações do estado, o vetor de componentes $\alpha_t = \{\mu_t, \gamma_t, \psi_t\}$ é estimado seqüencialmente no tempo através do filtro de Kalman (FK).

- É importante observar que a estimação é condicionante os conjunto de informações:

$$E(\alpha_t | Y_{t-1}) = \hat{\alpha}_{t|t-1} \rightarrow \text{previsão}$$

$$E(\alpha_t | Y_t) = \hat{\alpha}_t \rightarrow \text{filtragem}$$

$$E(\alpha_t | Y_T) = \hat{\alpha}_{t|T} \rightarrow \text{alisamento (smoothing)}$$

- A exposição dos ME seguirá a seguinte lógica:
  - apresentação dos modelos para cada componente;
  - forma do modelo em espaço de estados;
  - filtro de Kalman;
  - verossimilhança concentrada;
  - algoritmo de suavização

⇒ Modelos de tendência

- A definição de tendência nos ME segue o mesmo espírito da decomposição de B&N, sendo definida como a componente que domina a previsão s-passos à frente.

- Observando que, geralmente nos modelos estruturais de tendência estocástica o “drift” é zero, segue que, para uma série dessazonalizada a componente de tendência será aquela que :

$$\lim_{s\to\infty}(\hat{y}_{t+s|t} - \hat{\mu}_t) = 0$$

- Ou seja, é a componente que, quando extrapolada, domina a função de previsão.

- Na seqüência iremos apresentar dois ME p/ tendência:

  – o modelo de nível local (já apresentado);
  – o modelo linear local

I. Modelo de nível local (*local level model, RW plus noise, MNL*)

- É um modelo de tendência estocástica, adequado para ST onde o nível não é constante, tal com, desemprego no Brasil.

  – eq. das observ.: $y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \sim N(0,\sigma^2_\varepsilon)$ (i)

  - eq. do estado: $\mu_{t+1} = \mu_t + \eta_t \quad \eta_t \sim N(0,\sigma^2_n)$ (ii)

  ,onde:

  $E(\varepsilon_t \eta_s)=0, \ \forall\ t,s.$

  $E(\varepsilon_t \mu_1)=E(\eta_t \mu_1)=0, \ \forall\ t.$

  $\mu_1 \sim N(a_1,p_1).$

- A razão sinal ruído é definida como:

$$q = \sigma^2_\eta / \sigma^2_\varepsilon, \ \text{ ou}$$
$$\sigma^2_\eta = q\sigma^2_\varepsilon, \ 0 < q < \infty.$$

- Portanto, a "razão sinal-ruído" (rsn) representa um parâmetro que governa a suavidade da componente de nível local em relação ao ruído do modelo:

  – se q é grande (> 1), a variação do nível supera a variação do ruído da equação das observações;
  – se q é pequeno (< 1), o nível do processo fica "engolfado" pelo ruído.

- Observe os seguintes casos particulares da rsn:

$$\rightarrow \sigma_{\eta}^{2} = 0 \therefore q = 0. \quad \text{Portanto:}$$

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim N(0,\sigma_{\varepsilon}^{2}) \qquad (i)$$

$$\mu_t = \mu_{t-1} = \mu \qquad (ii)$$

- nesta situação o nível é constante, sendo estimado por MQO, produzindo a seguinte estimativa:

$$\hat{\mu}_t = (1/t)\sum_{i=1}^{t} y_i$$

- podemos, também mostrar que:

$$\hat{\mu}_{t+s|t} = \hat{\mu}_t$$

$$Var(y_{t+s|t}) = \sigma^2$$

$$\rightarrow \sigma_{\varepsilon}^{2} = 0 \therefore q = \infty. \quad Por\tan to:$$

$$y_t = \mu_t \qquad (i)$$

$$\mu_t = \mu_{t-1} + \eta_t \qquad \eta_t \sim \mathrm{N}(0,\sigma_{\eta}^{2}) \qquad (ii)$$

Ou seja y é um passeio aleatório:

$$y_t = y_{t-1} + \eta_t \qquad \eta_t \sim \mathrm{N}(0,\sigma_{\eta}^{2})$$

podemos, também mostrar que

$$\hat{y}_{t+s|t} = y_t$$
$$Var(\hat{y}_{t+s|t}) = s\sigma^2$$

• Caso geral:

$E(y_t) = E(\mu_t) = E(\mu_0) = a_0.$

$Var(y_t) = t\sigma_\eta^2 + \sigma_\varepsilon^2 + p_0.$

Portanto, o processo não é estacionário.

Previsão s-passos à frente:

$$y_{t+s} = \mu_{t+s} + \varepsilon_{t+s}$$
$$\mu_{t+s} = \mu_t + \sum_{i=1}^{s} \eta_{t+i}$$

donde segue que:

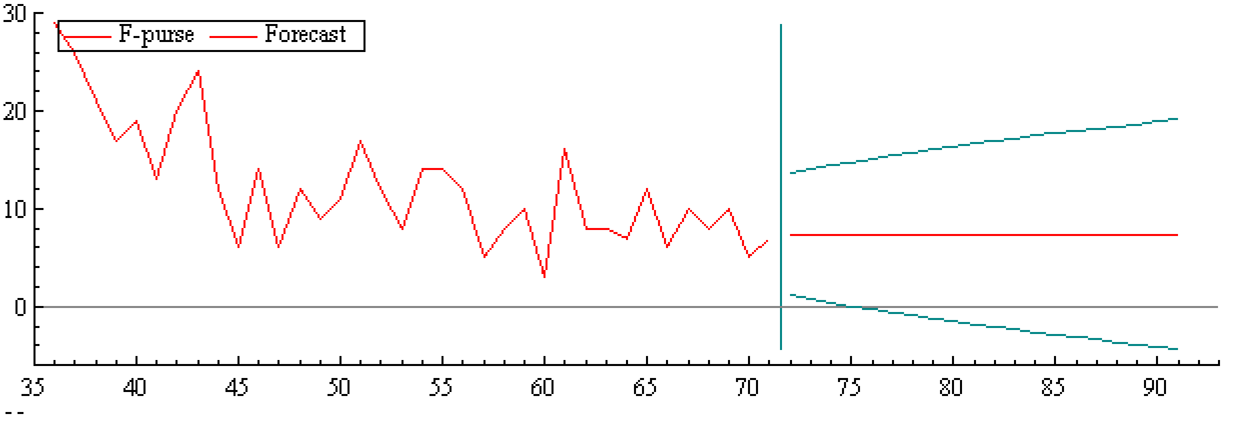
$\hat{y}_{t+s|t} = E(y_{t+s}|Y_t) = E(\mu_{t+s}|Y_t) = E(\mu_t|Y_t) = \hat{\mu}_{t|t} \rightarrow$ estimado pelo FK

$Var(y_{t+s|t}) = p_{t|t} + s\sigma_\eta^2 + \sigma_\varepsilon^2,$ $\quad p_{t|t} \rightarrow$ estimado pelo FK

• Como veremos, oportunamente, a previsão 1-passo à frente deste modelo é um EWMA, i.e., :

$$\hat{y}_{t+1|t} = E(y_{t+1}|Y_t) = (1-\lambda_t)y_t + \lambda_t \hat{y}_{t|t-1}.$$

- Observe que a função de previsão deste modelo é uma constante:



- Resumo dos casos especiais do modelo de nível local

nível	σ^2_ε	σ^2_η
- constante	*	0
- nível local	*	*
- passeio aleatório p/ y	0	*

(*) indica qualquer valor positivo

⇒ <u>Forma reduzida</u>: é obtida pela diferenciação apropriada de um ME, até torná-lo um processo estacionário de 2ª ordem.

- A forma reduzida, terá um equivalente ARIMA(p,d,q) com restrições adicionais nos parâmetros (parte deste material já foi visto). Serve também como um procediemnto para verificar se um dado

- Forma reduzida do modelo de nível local:

$$\Delta y_t^{(0)} = \Delta\mu_t + \Delta\varepsilon_t = \eta_t + \Delta\varepsilon_t$$

donde segue que:

$$E(\Delta y_t^{(0)}) = 0$$

$$Var(\Delta y_t^{(0)}) = \sigma_n^2 + 2\sigma_\varepsilon^2 \qquad (a)$$

$$Cov(\Delta y_t^{(0)}, \Delta y_{t+1}^{(0)}) = -\sigma_\varepsilon^2 \quad i = 1 \qquad (b)$$

$$= 0 \quad i \geq 2$$

- Portanto, podemos reescrever Δy como:

$$\Delta y_t = a_0 + e_t + \theta e_{t-1}, \quad e_t \sim N(0, \sigma_e^2)$$

donde segue que:

$$Var(\Delta y_t) = (1 + \theta^2)\sigma_e^2 \qquad (a1)$$

$$Cov(\Delta y_t, \Delta y_{t-1}) = \theta\sigma_e^2 \qquad (b1)$$

⇒ Forma reduzida: é obtida pela diferenciação apropriada de um ME, até torná-lo um processo estacionário de 2ª ordem.

- A forma reduzida, terá um equivalente ARIMA(p,d,q) com restrições adicionais nos parâmetros .

- Serve também como um procedimento para verificar se um dado ME é identificável, pois como todo modelo ARIMA é identificável, ao igualarmos os momentos do ARIMA equivalente como os da forma reduzida de um ME temos que ser capazes de resolver este sistema de eqs. para os parâmetros do ME de forma única a partir dos parâmetros do ARIMA equivalente.

- Se essa solução não for única o ME será não identificável, e assim sendo teremos que impor restrições adicionais aos hiperparâmetros do modelo, por exemplo, assumindo que correlações entre choques são nulas, o que interceptos são nulos.

- Forma reduzida do modelo de nível local:

$$\Delta y_t^{(0)} = \Delta\mu_t + \Delta\varepsilon_t = \eta_t + \Delta\varepsilon_t$$

donde segue que:

$$E(\Delta y_t^{(0)}) = 0$$
$$Var(\Delta y_t^{(0)}) = \sigma_n^2 + 2\sigma_\varepsilon^2 \qquad (a)$$
$$Cov(\Delta y_t^{(0)}, \Delta y_{t+1}^{(0)}) = -\sigma_\varepsilon^2 \quad i = 1 \qquad (b)$$
$$= 0 \quad i \geq 2$$

- Das propriedades (a) e (b) segue que $z_t = \Delta y_t$ ~ MA(1), ou seja, y_t ~ ARIMA(0,1,1).

$$z_t = e_t + \theta e_{t-1}, \quad e_t \sim N(0, \sigma_e^2), |\theta| < 1$$

donde segue que:

$$E(z_t) = 0$$
$$Var(z_t) = (1+\theta^2)\sigma_e^2 \quad (a1)$$
$$Cov(z_t, z_{t-1}) = \theta\sigma_e^2 \quad (b1)$$

- Igualando as eq (a) com eq (a1) e eq (b) com eq(b1), temos um sistema de 2 equações a duas incógnitas, e assim podemos obter, de forma única, estimativas de σ_n^2 e σ_ε^2 como função das estimativas de σ_e^2 e θ .

- Iremos agora observar que tipo de restrição a forma reduzida de um ME implica no seu equivalente ARMA.

- Igualando a FAC de ordem 1 do ME na forma reduzida com o ARIMA equivalente, obtemos:

$$\frac{\theta}{(1+\theta^2)} = \frac{-\sigma_\varepsilon^2}{\sigma_n^2 + 2\sigma_\varepsilon^2} = \frac{-1}{q+2}, \quad donde\ \ segue\ \ que:$$

$$\theta = [-(2+q) \pm \sqrt{q^2 + 4q}\,]/2.$$

$$Por\ \ fim\ usando\ \ que\ \ q \in (0,\infty),\ segue\ que\ \theta \in (-1,0).$$

$\Rightarrow$ Forma UCARIMA: representa um passo intermediário entre a forma estrutural e a reduzida.

- Para o modelo de nível local temos que:

$$\Delta y_t^{(0)} = \Delta\mu_t + \Delta\varepsilon_t$$
$$= \eta_t + \Delta\varepsilon_t$$

donde segue que:

$$y_t^{(0)} = \eta_t/\Delta + \varepsilon_t$$
$$= T_t + I_t,$$

que são as componentes de tendencia e a irregular.

- Portanto segue que a tendência é um passeio aleatório, ou ARIMA(0,1,0), como esperado.

$$T_t = \eta_t / \Delta, \text{ ou } \Delta T_t = \eta_t$$
$$T_t = T_{t-1} + \eta_t$$

- A utilidade das formas UCARIMA esta na possibilidade de representar um modelo de CNO como a soma de processos ARIMA(p,d,q).

- Esta abordagem foi utilizada por Nerlove e Carvalho na década de 70, e mais recentemente por Maravall *et al.*

II. Modelo de tendência linear local (*local linear trend, MLL*)

- O MLL possui uma tendência linear estocástica, sendo adequado p/ capturar tendências em ST de agregados macroeconômicas (PIB, prod. Industrial), séries associadas ao crescimento da população, preços e salários (nominais).

$$
\begin{array}{llll}
-eq.\ das\ observ.: & y_t = \mu_t + \varepsilon_t & \varepsilon_t \sim \mathrm{N}(0,\sigma^2_\varepsilon) & (i) \\
-eq.\ do\ \ estado: & \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t & \eta_t \sim N(0,\sigma^2_n) & (ii) \\
 & \beta_t = \beta_{t-1} + \zeta_t & \zeta_t \sim \mathrm{N}(0,\sigma^2_\xi) & (iii)
\end{array}
$$

$onde:$

$$E(\varepsilon_t\eta_s) = E(\varepsilon_t\zeta_s) = E(\eta_t\zeta_s) = 0, \ \ \forall\ t,s.$$

$$E(\varepsilon_t\alpha_0) = E(\eta_t\alpha_0) = E(\zeta_t\alpha_0)0, \ \forall\ t.$$

$$Se\ \alpha_t = (\mu_t,\ \beta_t)',\ então\ \ \alpha_o \sim N(a_0, P_0).$$

- De forma análoga ao modelo de nível local, podemos definir a razão sinal ruído para cada uma das componentes do MLL:

$$\mathrm{q}_\eta = \sigma^2_\eta / \sigma^2_\varepsilon \quad \mathrm{e} \quad \mathrm{q}_\xi = \sigma^2_\xi / \sigma^2_\varepsilon$$

- Interpretação das componentes:

  - μ_t= é a tendência estocástica da série;

  - β_t= é a inclinação estocástica da série;se a série estiver em logs, então 100xβ_t % é a taxa de crescimento instantâneo de y.

$\Rightarrow$ Casos particulares do MLL:

i. Se $\sigma^2_\eta = \sigma^2_\xi = 0$, então a tendência torna-se determinística, i.e., $y_t = \alpha + \beta t + \varepsilon_t$, onde $\alpha = \mu_0$ e $\beta = \beta_0$.

ii. Se $\sigma^2_\xi = 0$, então a inclinação da ST é fixa no tempo. Teremos, portanto o seguinte processo:

$$\mu_t = \mu_{t-1} + \beta + \eta_t \rightarrow \quad \text{RW} + \text{drift}$$

iii. Se $\sigma^2_\eta = 0$, então o modelo é conhecido como tendência suave.

$$\mu_t = \mu_{t-1} + \beta_{t-1} \rightarrow tendência \; suave$$

$$\beta_t = \beta_{t-1} + \zeta_t , \; \zeta_t \sim N(0, \sigma^2_\zeta)$$

Explicação: a segunda diferença da tendência é um ruído branco. Portanto a tendência é uma “integral” dupla de um rb, o que a torna suave, pois cada operação de integral equivale a um filtro de passa baixo.

iv. relação com o filtro HP.

• É possível obter o filtro HP como um caso particular do MLL.

• O entendimento detalhado desta relação exige que parametrizemos a tendência e o ciclo do procedimento de HP através de processos não estacionários, que se tornam estacionários após um número adequado de diferenciações.

• A melhor estimativa da tendência e do ciclo nesta situação, no sentido de erro quadrático médio, coincide com o resultado de HP, e foi demonstrado por Bell (84).

• Assumindo que:

$$
\begin{aligned}
y_t &= \mu_t + c_t \\
\delta_\mu(L)\mu_t &= \xi_t, \quad \xi_t \sim NID(0,\sigma^2_\xi) \\
\delta_c(L)c_t &= \upsilon_t, \quad \upsilon_t \sim NID(0,\sigma^2_\upsilon) \\
\delta_y(L)y_t &= \omega_t, \quad \omega_t \sim NID(0,\sigma^2_\omega)
\end{aligned}
$$

• Bell, mostra que a solução da estimação de μ_t é dada por

$$E(\mu_t|Y_t) = \frac{\delta_c(L)\delta_c(L^{-1})\gamma_\xi(L^{-1})}{\gamma_\omega(L^{-1})} y_t$$

$$onde\ \gamma_\omega(L) = \delta_c(L)\delta_c(L^{-1})\gamma_\xi(L^{-1}) + \delta_\mu(L)\delta_\mu(L^{-1})\gamma_\upsilon(L^{-1})$$

- Os γ's são as funções geradoras de autocovariância de cada ruído branco.

- Utilizando que no filtro de HP:

$$\mu_t \sim I(2) \therefore \delta_\mu(L) = (1-L)^2$$
$$c_t \sim I(0) \therefore \delta_c(L) = 1, \textit{segue que}:$$

$$E(\mu_t \mid Y_t) = \frac{\gamma_\xi(L^{-1}) / \gamma_\upsilon(L^{-1})}{\gamma_\xi(L^{-1}) / \gamma_\upsilon(L^{-1}) + (1-L)^2(1-L^{-1})^2} y_t$$

- Finalmente utilizando que as funções geradoras de autocovariância dos ruídos brancos, $\gamma_\xi(k) = \gamma_\xi(0) = \sigma^2_\xi$ e $\gamma_\upsilon(k) = \gamma_\upsilon(0) = \sigma^2_\upsilon$, segue que:

$$\mathrm{E}(\mu_t | \mathrm{Y}_t) = \frac{\lambda^{-1}}{\lambda^{-1} + (1-\mathrm{L})^2(1-\mathrm{L}^{-1})^2} \mathrm{y}_t, \text{ onde}$$

$$\lambda^{-1} = \sigma_\xi^2 / \sigma_\upsilon^2$$

- Observe que esta é exatamente a solução de HP.

- Resta-nos agora estabelecer a relação entre o filtro de HP e os modelos estruturais.

- O ponto de partida é o modelo com tendência suave, onde σ^2_η=0:

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim N(0,\sigma^2_\varepsilon)$$
$$\mu_t = \mu_{t-1} + \beta_{t-1}$$
$$\beta_t = \beta_{t-1} + \zeta_t \qquad \zeta_t \sim N(0,\sigma^2_\xi).$$

- A componente ε_t concorda com a versão de Bell para a componente cíclica do filtro HP (c_t), i.e., são variáveis aleatórias normais independente e identicamente distribuídas com média zero, variância cte e função geradora de autocovariância δ_c (L)=1.

- Por outro lado, a componente de tendência μ_t ~ I(2), pois:

Usando que:

$$\mu_t = \mu_{t-1} + \beta_{t-1}$$
$$\beta_t = \beta_{t-1} + \zeta_t \qquad \zeta_t \sim N(0,\sigma^2_\xi).$$

Segue que

$$(1-L)^2\mu_t = \Delta(\Delta\mu_t) = \Delta\beta_{t-1} = \zeta_{t-1}.$$

$$E(\mu_t|Y_t) = \frac{q_\zeta}{q_\zeta + (1-L)^2(1-L^{-1})^2} y_t, \text{ onde}$$

$q_\zeta = \sigma_\zeta^2 / \sigma_\upsilon^2$ é a razão sinal ruído para a inclinação.

- Portanto, na prática para implementar o filtro HP, rodamos um modelo estrutural com as seguintes características:

  - $\sigma^2_\eta = 0$;
  - $\sigma^2_\zeta / \sigma^2_\varepsilon = 1/1600$.

$\Rightarrow$ <u>Forma de espaço de estados</u>

- Um modelo estrutural é colocado na forma de espaço de estados (EE) qdo se utiliza a notação matricial. Para o MLL a forma de EE é a seguinte:

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix}$$

$$Ou: \qquad y_t = Z' \alpha_t + \varepsilon_t$$
$$\alpha_t = \mathrm{T}\, \alpha_{t-1} + R\eta_t$$

$\Rightarrow$ <u>Função de previsão</u>:

$$\hat{y}_{t+s|t} = E(y_{t+s} \,|\, Y_t) = Z'\, E(\alpha_{t+s}|\mathrm{Y_t}) = Z'\, \hat{\alpha}_{t+s|t}$$

Mas:

$$\hat{\alpha}_{t+s|t} = \mathrm{T^s}\, E(\alpha_t \,|\, Y_t) = \mathrm{T^s} \begin{bmatrix} \hat{\mu}_{t|t} \\ \hat{\beta}_{t|t} \end{bmatrix} = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\mu}_{t|t} \\ \hat{\beta}_{t|t} \end{bmatrix} = \begin{bmatrix} \hat{\mu}_{t|t} + s\, \hat{\beta}_{t|t} \\ \hat{\beta}_{t|t} \end{bmatrix}$$

Portanto:

$$\hat{y}_{t+s|t} = Z'\, \alpha_{t+s|t} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\mu}_{t|t} + s\, \hat{\beta}_{t|t} \\ \hat{\beta}_{t|t} \end{bmatrix} = \hat{\mu}_{t|t} + s\, \hat{\beta}_{t|t}$$

$$MSE(\hat{y}_{t+s|t}) = E(y_{t+s} - \hat{y}_{t+s|t})^2 = Z^{'}\ MSE(\hat{\alpha}_{t+s|t})Z + \sigma_{\varepsilon}^{2}$$

$$MSE(\hat{\alpha}_{t+s|t}) = T^{s} P_{t|t} T^{s'} + \sum_{j=0}^{s-1} T^{j} RQR^{\prime} T^{j'}$$

$$= Z^{'} T^{s} P_{t} T^{s'} Z + Z^{'} (\sum_{j=0}^{s-1} T^{j} RQR^{\prime} T^{j'}) Z + \sigma_{\varepsilon}^{2},$$

$$onde\ P_{t} = E[(\alpha - \hat{\alpha}_{t|t})(\alpha_{t} - \hat{\alpha}_{t|t})' \mid Y_{t}]$$

$$Q = E\left[\eta_{t}\eta_{t}^{\prime}\right]$$

Particularizando esta expressão para o modelo de tendência linear estocástica chega-se a:

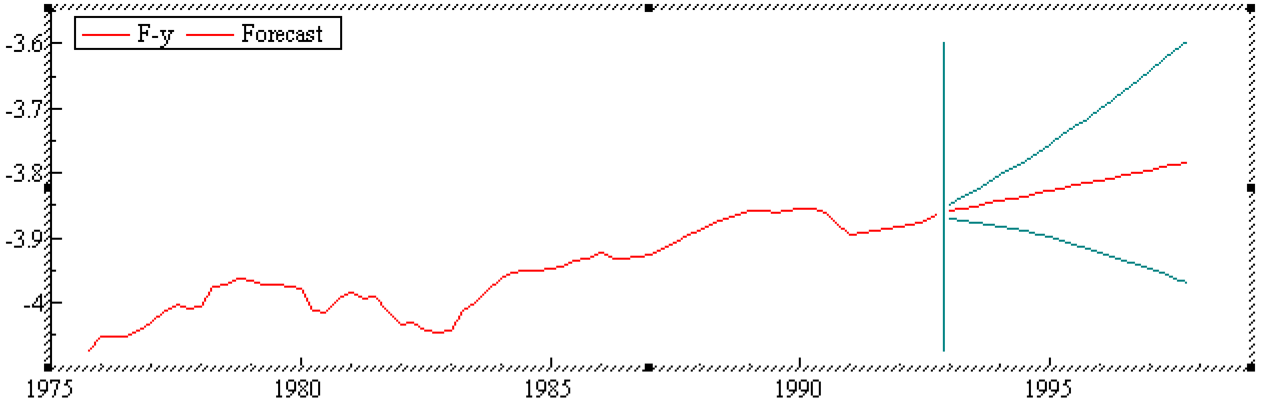
$$MSE(\hat{y}_{t+s|t}) = [(P_{t|t}{}^{(1,1)} + 2sP_{t|t}{}^{(1,2)} + s^{2} P_{t|t}{}^{(2,2)})] +$$

$$+ [sq_{\eta} + (1/6)s(s-1)(2s-1)q_{\zeta} + 1]\sigma_{\varepsilon}^{2}, \quad s = 1,2,\ldots$$

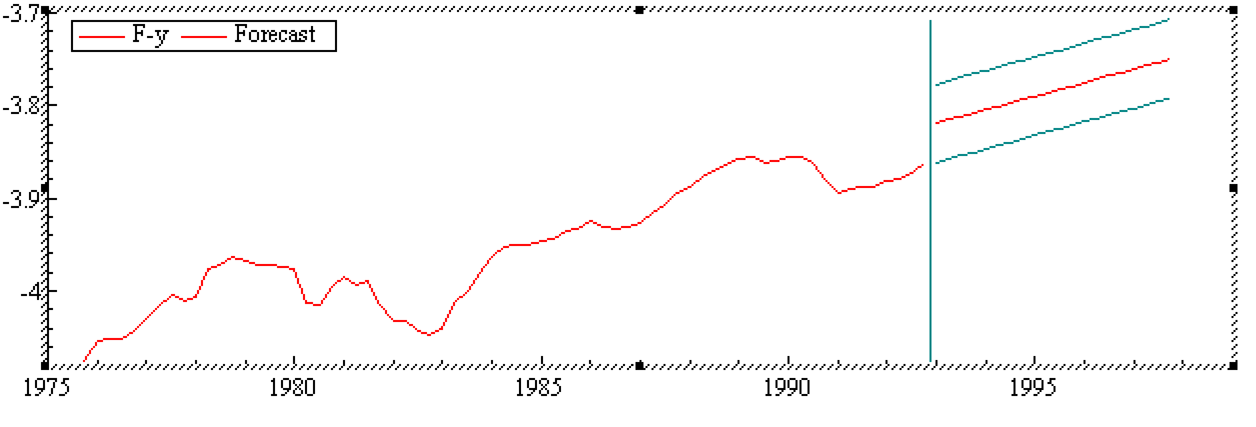
$P_{t|t}{}^{(i,j)}$ *é o ij – ésimo elemento da matriz* $P_{t|t}$.

$$IC\ de\ 95\%: \quad \hat{y}_{t+s|t} \pm 1{,}96\sqrt{MSE(\hat{y}_{t+s|t})}$$

- Ou seja, a função de previsão é uma reta, onde os coeficientes são construídos a partir de amortecimento exponencial (será visto no FK).



- Compare esta previsão com a que seria obtida através de uma tendência determinística.



- Resumo dos casos especiais do modelo de tendência linear local

tendência	σ_ε	σ_η	σ_ζ
- determinística	*	0	0
- inclinação fixa	*	*	0
- passeio aleatório c/ drift p/ y	0	*	0
- linear local	*	*	*
- tendência suave	*	0	*
- filtro HP	*	0	$.025\sigma_\varepsilon$
- segunda diferença do y	0	0	*

(*) indica qualquer valor positivo.

$\Rightarrow$ Forma reduzida do MLL:

$$y_t = \mu_t + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

Tomando a primeira diferença:
$$\Delta y_t = \Delta\mu_t + \Delta\varepsilon_t$$
$$\Delta y_t = \beta_{t-1} + \eta_t + \Delta\varepsilon_t$$

Tomando a segunda diferença:
$$\Delta^2 y_t = \Delta\beta_{t-1} + \Delta\eta_t + \Delta^2\varepsilon_t$$
$$\Delta^2 y_t = \zeta_{t-1} + \Delta\eta_t + \Delta^2\varepsilon_t \ .$$

- Seja $w_t=\Delta^2 y_t$. Segue que:

$$w_t = \zeta_{t-1} + \Delta\eta_t + \Delta^2\varepsilon_t \ .$$

$$E(w_t)=0$$
$$Var(w_t) = Var(\zeta_{t-1} + \eta_t - \eta_{t-1} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2})$$
$$= \sigma_\zeta^3 + 2\sigma_\eta^2 + 6\sigma_\varepsilon^2. \qquad \text{(a)}$$
$$Cov(w_t, w_{t-k}) = E[(\xi_{t-1} + \eta_t - \eta_{t-1} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2})$$
$$(\xi_{t-1-k} + \eta_{t-k} - \eta_{t-1-k} + \varepsilon_{t-k} - 2\varepsilon_{t-1-k} + \varepsilon_{t-2-k})]$$
$$= -\sigma_\eta^2 - 4\sigma_\varepsilon^2, \ k=1 \qquad \text{(b)}$$
$$= \sigma_\varepsilon^2, \qquad k=2 \qquad \text{(c)}$$
$$= 0, \qquad k \geq 3$$

- Ou seja $w_t=\Delta^2 y_t$ ~ MA(2) ou y_t ~ ARIMA(0,2,2). Segue portanto que o ARIMA equivalente da forma reduzida de um MLL é dado por:

$$z_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2},\ e_t \sim N(0, \sigma_e^2)$$

onde as raízes de $f(v) = v^2 + v\theta_1 + \theta_2 = 0, \lambda_i, i = 1, 2,$

devem satisfazer $|\lambda_i| < 1, i = 1, 2$ para o processo ser invertível.

$$E(z_t) = 0$$

$$Var(z_t) = (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 \qquad \text{(a1)}$$

$$Cov(z_t, z_{t-k}) = (\theta_1 + \theta_2\theta_1)\sigma_e^2 \quad k = 1 \qquad \text{(b1)}$$

$$= \theta_2\sigma_e^2 \qquad k=2 \qquad \text{(c1)}$$

- Igualando as eqs a e a1, b e b1 e c e c1, chegamos a um sistema de 3 equações a 3 incógnitas (os hiperparâmetros do MLL), e assim o sistema apresenta solução única e segue que o MLL é identificável.

- Agora queremo saber se o MA(2) equivalente terá mais restrições do que as usuais.

$$\mathrm{Corr}(w_t, w_{t-k}) = \rho_w(k)$$

$$\begin{cases} = -(4+q_\eta)/(2q_\eta + q_\zeta + 6), & k=1 \\ = 1/(2q_\eta + q_\zeta + 6), & k=2 \end{cases}$$

- Usando que z_t ~ MA(2),

$$z_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}. \text{ Segue que:}$$

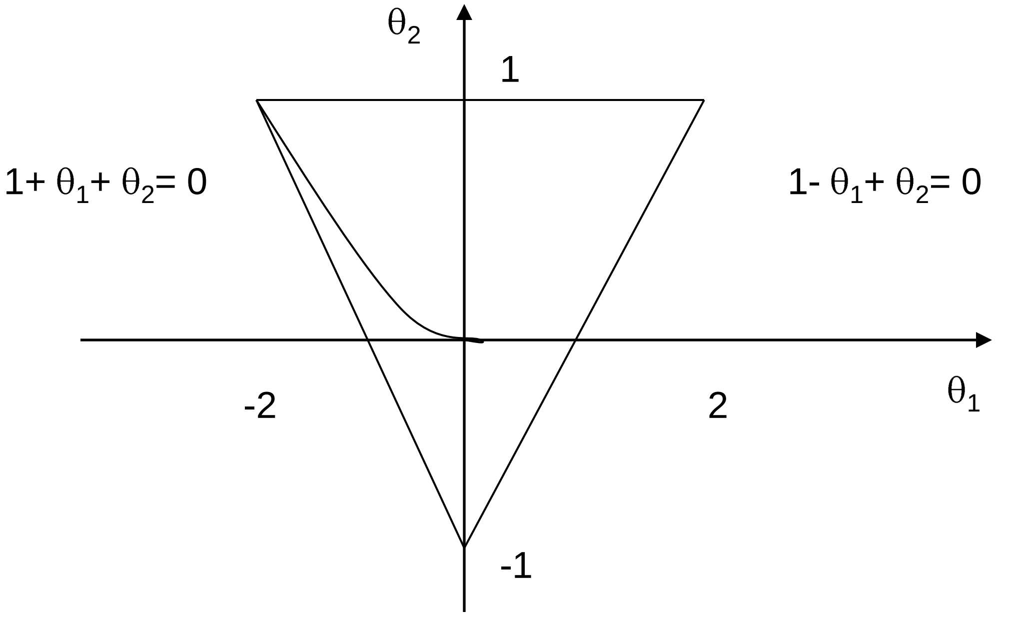
$$\rho_z(k) = (\theta_1 + \theta_1\theta_2)/(1+\theta_1^2 + \theta_2^2), \ k=1$$
$$\theta_2/(1+\theta_1^2 + \theta_2^2), \ k=2$$
$$0, \ k \geq 3.$$

- Obtendo a expressão $\rho(1)$ / $\rho(2)$ utilizando as duas fórmulas anteriores, segue que:

$$\rho(1)/\rho(2) \ _z = \rho(1)/\rho(2) \ _w$$
$$(\theta_1 + \theta_1\theta_2)/\theta_2 = -(4+q_\eta)$$

$$\text{ou, } \theta_1 = -\theta_2(4+q_\eta)/(1+\theta_2)$$

• Obtendo o gráfico θ_1 x θ_2, observamos que a região de invertibilidade do MA(2) é drásticamente reduzida.



• Observe as seguintes formas reduzidas do MLL:

i. se $\sigma^2_{\xi}= 0$ (inclinação fixa), então o modelo não será invertível.

ii. se $\sigma^2_{\xi}> 0$ e $\sigma^2_{\eta}= 0$ (tendência suave), então o modelo será sempre invertível.

$\Rightarrow$ Forma UCARIMA

$$\Delta^2 y_t = \zeta_{t-1} + \Delta\eta_t + \Delta^2\varepsilon_t \ .$$

$$y_t = (\zeta_{t-1} + \Delta\eta_t) / \Delta^2 \ + \ \varepsilon_t .$$

Definindo:

$$T_t = (\zeta_{t-1} + \Delta\eta_t) / \Delta^2 \ e$$

$$I_t = \varepsilon_t , \ segue \ que:$$

$$\Delta^2 T_t = \zeta_{t-1} + \Delta\eta_t = \omega_t \sim MA(2).$$

Portanto $T_t \sim ARIMA(0,2,2)$.

⇒ Outros modelos de tendência estocástica:

- É possível estender este procedimento para outros tipos de modelo de tendência:

  - tendência quadrática;
  - tendência amortecida;
  - curvas de crescimento

- Por exemplo, a tendência linear amortecida será dada por:

$$y_t = \mu_t + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \phi\beta_{t-1} + \zeta_t$$

onde os distúrbios aleatórios seguem os pressupostos usuais e $|\phi| < 1$.

- Pode-se demonstrar que a função de previsão deste modelo é dada por:

$$\hat{y}_{t+s|t} = \hat{\mu}_t + \hat{\beta}_t[(1-\phi^s)/(1-\phi)]$$