

Formal modelo em Espaço de Estado (terminologia)

↓ problema real descrito  
em equações matemáticas,  
acabando

podem ser colocados em modelos	}	ARMA
E-E		ARIMA
	}	VARMA
		REGRESSÃO (SLM, etc.)

Espaço → ideia veio da Física

conjunto mínimos de info足以 descrever sistema de forma completa.

obj: Na mecânica → dois eixos足以 descrever o sistema: posição e velocidade

( $x, y, z$ ) ( $v_x, v_y, v_z$ )  
Na Termodin. → temperatura, pressão, velocidade ( $T, P, V$ )

Em particular, origem está no processamento de sinais

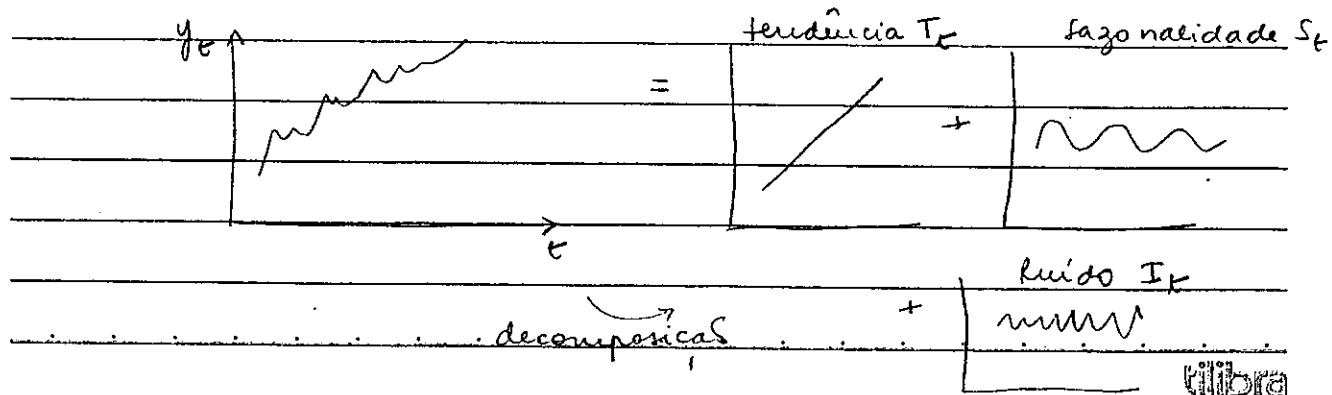
$$y_t = \theta_t + \epsilon_t$$

recebe "mensagem" erro (medo)  
nas observadas

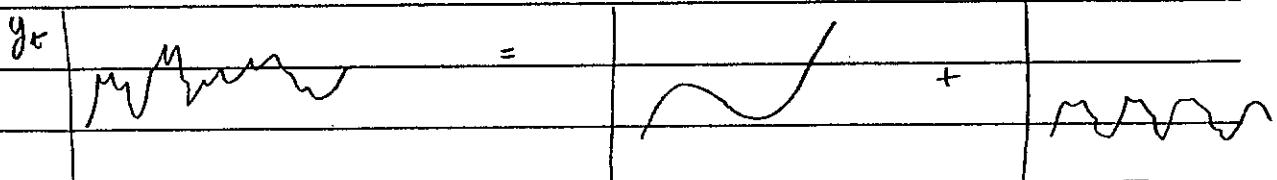
Como, observando  $y_t$ , podemos filtrar a msg  $\theta_t$ ?

O que interessa, a "mensagem", é nas observadas.

Ex: Observando uma série de tempo



ou



(comportamento  
um pouco + matemático)

$T_t$  +  $S_t$

\ /  
componentes são  
estocásticos.

$$y_t = T_t + S_t + I_t$$

$$T_t = a + bt$$

$$S_t = \sum_{j=1}^n \gamma_j D_{jt}, \quad D_{jt} = \begin{cases} 1 & j=t \\ 0 & j \neq t \end{cases} \quad (\text{dummis ou zeros e consos})$$

$$I_t = E_t \sim N(0, \sigma^2) \quad \text{ruído branco}$$

$$y_t = (a + bt) + \left( \sum_{j=1}^n \gamma_j D_{jt} \right) + E_t \quad : \text{é um modelo de regressão}\dots \text{particulares de } y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt}$$

↳ como estimamos este modelo?

Precisamos: conjunto de observações  $\{y_t\}_{t=1}^T$

+ Método (de estimacões)

↳ MLE (primeiro método que pensamos)

Vetor de parâmetros desconhecidos:  $\Psi = (a, b, \gamma_1, \gamma_2, \dots, \gamma_n, \sigma^2)$   
(a serem estimados)

tribo

funções objetivo

$$S(\psi^*) = \sum_{t=1}^T (y_t - \hat{y}_t)^2 = \sum (y_t - [(\alpha + \beta t) + \sum_{j=1}^J \delta_j D_{jt}])^2$$

onde  $\psi^* = \{\psi\} - \sigma^2$

↳ pois se o  $\sigma^2$  não conseguisse estimar  $\sigma^2$ , já que não aparece na função objetivo

$\psi^*$  tal que  $\min S(\psi^*)$

$$\frac{dS}{d\psi^*} = 0 \quad \therefore \quad \left( \frac{\partial S}{\partial \psi_1}, \frac{\partial S}{\partial \psi_2}, \dots, \frac{\partial S}{\partial \psi_J} \right)^T = 0^T$$

$$\hat{\psi}^* = (X'X)^{-1} X' y \quad y = (y_1, \dots, y_T)^T$$

Mas esta questão é questionável.

Componentes de séries não são normalmente determinísticas.  
(por ex, séries macroeconômicas).

Ex: Série de Desemprego

Componente razoável determinística não é ruim  
mas componente tendência "é ruim"

No Eviews só estimar: quick

Equation Estimation

desemp-sp c @trend @mas(1) @ras(2) ... @ras(11)

c → tend. → dumais razoáveis em relação ao  
a      b      r      mês 12.

Obs: deixamos todas as dumais de razoabilidade

mesmo que elas não sejam significativas  
estatisticamente.

11

obs: sazonalidade

$$E[y^{\text{dez}}] = a + bt^{\text{dez}}$$

todas as dummies de mês 1 a 11 saem zero

$$\text{e } E[\epsilon_0] = 0$$

$$E[y^{\text{Jan}}] = a + bt^{\text{jan}} + \gamma_1$$

$$\Delta^{\text{Jan-dez}} = E[y^{\text{Jan}}] - E[y^{\text{dez}}]$$

$$= b \underbrace{(t^{\text{Jan}} - t^{\text{dez}})}_{1} + \gamma_1 = b + \gamma_1$$

Para um mês qualquer:

$$\Delta^{\text{mês}_j, \text{dez}} = b_j + \gamma_j$$

$H_0: \gamma_j = 0$  → meses que não são estatisticamente significantes

$H_1: \gamma_j \neq 0$  (seja como trabalhar com  
uma lista de meses de referência:  
dez + todos não significativamente ≠)

→ Para se ter a ideia de existência de ciclo,  
podermos olhar o erro da regressão em relações  
à tendência e à sazonalidade e verificar se há  
alguma repetição em períodos  $\geq 1$  ano.

⇒ Modelo não é satisfatório.

Autocorrelações dos resíduos: há alguma não-estacionariedade  
(é bem forte) no resíduo.

Tendência não é determinística como percebemos (é estocástica)  
Podemos até forçar a ver ciclos macroeconômicos.

(obs: E-Views não faz ajuste pp tendéncia estocástica).

tilibra

/ /

Ainda trabalhando com tendência e sazonalidade:

$$y_t = \underbrace{(a_t + b_t t)}_{T_t} + \underbrace{\sum \delta_j D_j t}_{S_t} + \underbrace{e_t}_{I_t}$$

pp cada  $t \neq$  parâmetro é diferente

Propondo modelos onde constante  $\neq$  sas funções do tempo.

Modelos estruturais: modelos de regressão onde modelos explicativas e coeficientes sas funções do tempo.

Supondo que estejamos trabalhando c/ séries anuais:  
(não tem sazonalidade)

$$\begin{cases} y_t = a_t + b_t t + e_t & \rightarrow \text{mas pp cada } t = 1 \text{ a } T \text{ só} \\ a_t = " " & \text{teríamos 1 observação que} \\ b_t = " " & \text{não permitiria estimar} \end{cases}$$

precisa colocar alguma estrutura pp  $a_t$  e  $b_t$ .

$$\cdot y_t = \mu + e_t \quad \Rightarrow \hat{\mu} = \frac{1}{T} \sum_{i=1}^T y_i$$

~~$\mu$~~  → (desinteressante)

$$\cdot y_t = \mu_t + e_t \quad \downarrow \text{precisa propor processo de evolução de } \mu$$

$$\therefore \mu_t = \phi \mu_{t-1} + \eta_t \Rightarrow \text{processo estocástico que define evolução de } \mu$$

$$1) \text{ se } |\phi| < 1 \Rightarrow \text{AR}(1)$$

equilíbrio

como  $\mu_t$  estacionário  $\left\{ \begin{array}{l} y_t \text{ estacionário} \\ \text{e } E_t \text{ "} \end{array} \right.$

2) se  $\phi = 1 \Rightarrow \mu_t \sim \text{processo aleatório}$

(não estacionário)

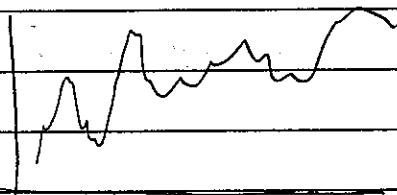
Neste caso  $\Rightarrow y_t$  é não estacionário

Normalmente, usamos  $\phi=1$

Modelo de  $\left\{ \begin{array}{l} y_t = \mu_t + \varepsilon_t \\ \text{nível local} \end{array} \right.$

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

(Local level  
model)



obs 1: A fórmula de previsão deste modelo é um EWMA (método de amorteamento exponencial) para o modelo que

Caracteriza tendência  
estocástica.

$$\hat{y}_{t+1|t} = E[y_{t+1} | Y_t] = \lambda y_t + (1 - \lambda) \hat{y}_{t|t}, \quad 0 < \lambda < 1$$

→ É um esquema, um método de previsão (com amortecimento) para séries de sazonalidade exponencial

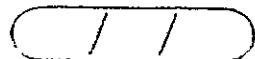
obs 1: a real função das parâmetros do filtro de Kalman

$$\underline{\text{obs 2:}} \lim_{t \rightarrow \infty} \lambda_t \rightarrow \bar{\lambda}$$

obs 3: Se  $\sigma_{\eta_t}^2 \rightarrow 0$ :  $\mu_t = \mu_{t-1}$

Modelo tende à tendência determinística

como caso particular



Outros modelos: (tb de sazonalidade, apenas tendência)

Tendência linear local

$$\left\{ \begin{array}{l} y_t = \mu_t + \varepsilon_t \quad \rightarrow \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad \rightarrow \quad \eta_t \sim N(0, \eta^2) : \mu_t \text{ já é a tendência} \\ \beta_t = \beta_{t-1} + \beta_t \quad \rightarrow \quad \beta_t \sim N(0, \beta^2) : \text{ passo aleatório} \\ \text{algo que compõe a tendência} \end{array} \right.$$

obs: notações  $y_t = \mu_t + \varepsilon_t \Rightarrow$  NO Koopman  
 $\mu_{t+1} = \mu_t + \beta_t + \eta_t \Rightarrow$  aparece assim.  
 $\beta_{t+1} = \beta_t + \beta_t \Rightarrow$  fórmulas do FK  
 nas mudanças

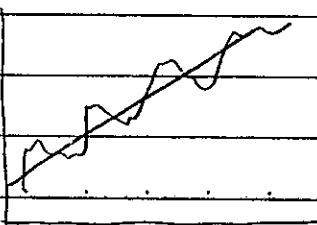
Dois caminhos p/ verificar o modelo:  
 (testar)

- | Analítico
- | Empírico (por simulações)

A partir deste modelo, podemos gerar vários casos particulares:

$$\begin{aligned} 1) \quad \sigma_\eta^2 = \sigma_\beta^2 = 0 \quad \Rightarrow \quad \mu_t &= \mu_{t-1} + \beta_0 \\ &\mu_t = \mu_0 + \beta_0 t \quad (\text{lineando}) \\ \therefore y_t &= \mu_0 + \beta_0 t + \varepsilon_t \quad \Rightarrow \text{tendência linear} \end{aligned}$$

obs: tendência determinística  $\Rightarrow$  chamada global  
 " estocástica  $\Rightarrow$  " local



Hibra

11

modelos com razoabilidade e ciclo

$$y_t = \mu_t + \delta_t + \psi_t + \epsilon_t \Rightarrow \text{equações da medida}$$

$$\left\{ \begin{array}{l} \mu_t = \dots \\ \delta_t = \dots \\ \psi_t = \dots \end{array} \right\} \quad \begin{array}{l} \text{equações} \\ \text{dos estados} \end{array}$$

obs: como estimar os parâmetros?

MFO pressupõe parâmetros fixos no tempo.

Neste caso, os parâmetros fixos são chamados de hiperparâmetros

Pf estimar hiperparâmetros  $\Rightarrow$  MFO

Pf estimar parâmetros variáveis  $\Rightarrow$  Filtros de kalman

Mas pf usar FK precisamos de hiperparâmetros e pf  
estimar hiperparâmetros, precisamos de FK.

Definição de uma representação em espaço estado linear:

$\rightarrow$  mais geral. Na literatura = o normalmente

$$y_t = Z_t x_t + d_t + \epsilon_t \quad (\text{espera-se } m \ll p)$$

$p_{x_t} = p_{x_0} \otimes \dots \otimes p_{x_t}$

$\downarrow$   
choque

$x_t$ : vetor de estados não observados.

$\rightarrow$  mais flexibilidade na representação

$$x_{t+1} = T_t x_t + C_t + R_t \eta_t \quad \text{tendo + choques}$$

$$\epsilon_t \sim WN(0, H_t)$$

$$\eta_t \sim WN(0, Q_t)$$

/ /

$$\alpha_i \sim N(a_i, P_i)$$

(especificar  $a_i$  e  $P_i$ : condições iniciais)

Podemos considerar:

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H_t & 0 \\ 0 & Q_t \end{pmatrix} \right) \quad \text{descorrelatados.}$$

(da definição  
do modelo)

Dizemos também que:

$$E(\varepsilon_t \alpha_i^T) = 0$$

$$E(\eta_t \alpha_i^T) = 0$$

Dado modelo, precisaremos do algoritmo de estimacões

(Filtro de Kalman).

Muitos modelos econômicos / ST podem ser colocados na forma EE (não existe uma forma única)

Se  $\alpha_t^* = B\alpha_t$ , onde  $\det(B) \neq 0 \Rightarrow$  função de verossimilhança será a mesma.

Ex.1:  $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$  (AR-2 sem drift, podemos ignorar  $c_t \in dt$ )

duas representações:

$$1) y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t = \alpha_{1t} \\ \phi_2 y_{t-1} = \alpha_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{1t} \\ \alpha_{2t} \end{pmatrix} \Rightarrow \text{equações das medidas}$$

$$\begin{pmatrix} y_{t+1} \\ \phi_2 y_t \end{pmatrix} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \phi_2 y_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_{t+1} \Rightarrow \text{equações dos estados}$$

$\alpha_{t+1}$        $T$        $\alpha_t$        $R_t$        $\eta_t$       **equilíbrio**

$$2) \quad y_t = (1 \quad 0) \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} \xrightarrow{\alpha^*} \Rightarrow \text{equações das medidas}$$

$$\begin{pmatrix} y_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_{t+1} \Rightarrow \text{equações dos estados}$$

$$\text{Logo : } \alpha_t^* = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1/\phi_2 \end{pmatrix}}_B \underbrace{\begin{pmatrix} y_t \\ \phi_2 y_{t-1} \end{pmatrix}}_{\text{estado da 1ª representação}}$$

- Quando colocamos uma equação em uma representação espaço de estados: entramos com vetores e parâmetros em um pacote computacional e pedimos pra estimar o estado sequencialmente

$\Rightarrow$  estimará parâmetros fixos através das verossimilhanças  
(neste caso,  $\phi_1$ ,  $\phi_2$  e  $\sigma_e^2$ )

- Qualquer modelo  $A_t$  pode ser estimado na forma de espaço de estado.

Ex. 2:  $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \rightarrow$  Modelos AR-P  
 Raízes do polinômio  
 características de um processo estacionário  
 círculo unitário  
 $\Rightarrow$  processo estacionário  
 de 2º ordem

$$\phi(l) = 1 - \phi_1 l + \phi_2 l^2 - \dots - \phi_p l^p \quad (\text{raízes para o círculo})$$

$$\phi(z) = z^p - \phi_1 z^{p-1} + \phi_2 z^{p-2} - \dots - \phi_p. \quad (\text{raízes dentro dos círculos})$$

/ /

$$y_t = (1 \ 0 \ 0 \dots 0) \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{pmatrix} + \alpha_t$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t$$

$$\begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \\ \vdots \\ y_{t+1-p} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 1 & 0 & 0 \end{pmatrix}}_J \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \epsilon_{t+1}$$

Ex 3)  $y_t \sim MA(1)$ 

$$y_t = \theta \epsilon_{t-1} + \epsilon_t$$

$$\Rightarrow y_t = (1 \ 0) \begin{pmatrix} y_t \\ \theta \epsilon_t \end{pmatrix} + \alpha_t$$

$$\begin{pmatrix} y_{t+1} \\ \theta \epsilon_{t+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \theta \epsilon_t \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \epsilon_{t+1} =$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t$$

Se conseguimos escrever AR  $\Leftrightarrow$  MA, conseguimos escrever um ARMA  
(Exercício)

$$y_t = \phi y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \quad (y_{t+1} = \phi y_t + \theta \epsilon_{t-1} + \epsilon_t)$$

a)  $y_t = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \phi y_{t-1} \\ \theta \epsilon_{t-1} \end{bmatrix}$

$\xrightarrow{\alpha_t}$

cuidado com índice de  $\epsilon_t$ !  
escrever as equações de espaço de estado.

1º elemento

termo que

representa

equações

$$\begin{bmatrix} y_{t+1} \\ \phi y_t \\ \theta \epsilon_t \end{bmatrix} = \begin{bmatrix} \phi & 0 & 1 \\ \phi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \phi y_{t-1} \\ \theta \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \epsilon_{t+1}$$

#### Ex. 4) Regressão múltipla

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_p x_{pt} + \epsilon_t$$

$$y_t = (1 \ x_{1t} \ x_{2t} \ \dots \ x_{pt}) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \epsilon_t$$

$$\left\{ \begin{array}{l} y_t = x_t^\top \beta + \epsilon_t \\ \downarrow \\ \text{seja o vetor de estado} \\ \beta_t = \beta_{t-1} \end{array} \right.$$

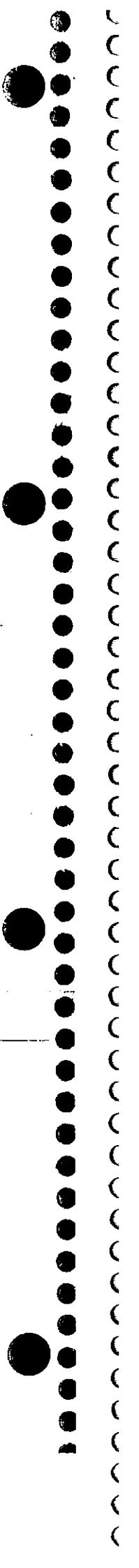
$$T_t = I, R_t = 0$$

Filtros de Kalman estimarão  $\beta$

7  
//

$$y_t = \beta_0 + \beta_1 x_{1t} +$$

cont.



tilbra

Se for tempo contínuo:

ao invés de s.p. de diferenças estocásticas,  
teríamos s.t. de equações diferenciais

em tempo discreto  
e dist. Gaussiana 16/03/11

## Modelo Espaço de Estado Linear (MEE linear)

processo estocástico  
que depende de

$$y_t = z_t \alpha_t + d_t + \varepsilon_t \quad \text{pxp} \quad \text{obs: riva, d}$$

$$x_{t+1} = T_t x_t + C_t + R_t \eta_t \quad \eta_t \sim N(0, Q_t) \quad \text{nxn}$$

equações de

evolutivas

(de estado)

↓

estado pode ter

interpretação direta

ou não (nos modelos atuais)

não tem, mas em modelos estruturais tem)

$$\left\{ \begin{array}{l} E(\varepsilon \varepsilon') = 0 \\ E(\eta \eta') = 0 \end{array} \right. \quad \alpha_0 \sim N(a_0, P_0)$$

$$\left\{ \begin{array}{l} E(\varepsilon \eta') = 0 \\ E[\varepsilon_t \alpha_0] = 0 \end{array} \right. \quad \forall t, s$$

$$\left\{ \begin{array}{l} E[\eta_t \alpha_0] = E[\varepsilon_t \alpha_0] = 0 \\ \text{(ruídos brancos)} \end{array} \right. \quad \forall t$$

cond. inicial

de node

se pxt ou

algo contra

muitas vezes

mas aparece

na formula

cat, mas s.

é mais

geral

faç parte da

definição do modelo EE.

já não pode ser relaxada

(a desconectar  
entre choques  
pode ser relaxada)

ideia: colocar no formato EE e estimar parâmetros

olhas s.p. como processos estocásticos e estender propriedades.

Propriedades dos modelos EE lineares a Tempo Discreto

(1) Estacionariedade de 2ª ordem

(2) Observabilidade

Definição: As matrizes  $z_t, d_t, H_t, T_t, C_t, R_t, Q_t$  são  
chamadas matrizes do sistema

Vamos assumir que algumas matrizes são invariantes no tempo

Isto não invalidará resultados do Filtro de Kalman.

Por hipótese, nas deduções subsequentes, o sistema será tido como homogêneo invariante no tempo.

De reja, vamos abandonar o subscrito  $t$  das matrizes do sistema (pode ser relaxada para  $c_t$  e  $d_t$ )

ex:  $C_t = \beta X_t$  se  $X_t$  for série estacionária de 2<sup>a</sup> ordem  
(não afeta propriedade)

$$Ex: y_t = \mu_t + \epsilon_t \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

systema invariante

$$\mu_{t+1} = \phi \mu_t + \eta_t \quad \eta_t \sim N(0, \sigma_\eta^2) \quad \text{no tempo}$$

### (i) Estacionariedade de 2<sup>a</sup> ordem

processo-matrizial

definição: um processo estocástico  $z(t)$  é dito estacionário de 2<sup>a</sup> ordem se satisfizer as seguintes condições:  
e somente se

→ médios do vetor de médias é constante

$$(i) E[z_t] = \mu \quad \forall t, \mu < \infty$$

$$(ii) E[(z_t - E[z_t])(z_{t+h} - E[z_{t+h}])'] = \Gamma(h) \quad \forall h = 0, 1, 2, \dots$$

→ não depende

matriz de autocovariâncias de  $h$

de lag  $h$

(se  $h=0 \Rightarrow$  matriz de var covar)

Utilizando as equações do MEE segue que:

$$cso: E[y_t] = z E[\alpha_t] + E[d_t]$$

↓  
de  $E[\alpha_t]$

spg, faremos  $d_t=0$

mas podemos ter  $d_t$  cte ou

até  $d_t = p x_t$  (com variável

explicativa exógena) desde

que  $x_t$  seja série estacionária  
de 2º ordem

$$cso: E[(y_t - E[y_t])(y_{t+h} - E[y_{t+h}])'] = \Gamma_y(h) ?$$

$$\left\{ \begin{array}{l} y_t = z \alpha_t + \varepsilon_t \\ E[y_t] = z E[\alpha_t] \end{array} \right.$$

valores atuais se é verdade.

$$\left\{ \begin{array}{l} y_t = z \alpha_t + \varepsilon_t \\ E[y_t] = z E[\alpha_t] \end{array} \right.$$

$$\Rightarrow y_t - E[y_t] = z (\alpha_t - E[\alpha_t]) + \varepsilon_t$$

$$y_{t+h} - E[y_{t+h}] = z (\alpha_{t+h} - E[\alpha_{t+h}]) + \varepsilon_{t+h}$$

E assim, segue que:

$$(y_t - E[y_t])(y_{t+h} - E[y_{t+h}])' =$$

$$= [z(\alpha_t - E[\alpha_t]) + \varepsilon_t] [(\alpha_{t+h} - E[\alpha_{t+h}])' z' + \varepsilon_{t+h}'] =$$

$$= z(\alpha_t - E[\alpha_t])(\alpha_{t+h} - E[\alpha_{t+h}])' z' + z(\alpha_t - E[\alpha_t]) \varepsilon_{t+h}' +$$

valor esperado cruzado será  $\Gamma_z$  valor esperado cruzado

$$+ \varepsilon_t (\alpha_{t+h} - E[\alpha_{t+h}])' z' + \varepsilon_t \varepsilon_{t+h}'$$

Passando valor esperado:

$$\Gamma_y(h) = z \Gamma_z(h) z' + z (E[\alpha_t \varepsilon_{t+h}'] - E[\alpha_t] E[\varepsilon_{t+h}']) + o +$$

$$\underbrace{E[\varepsilon_t \varepsilon_{t+h}']}_{=0 \text{ se } h=0}$$

$$e=H \text{ se } h \neq 0$$

equilíbrio

para  
 caídas  
 de valen-  
 tina  
 operati-  
 vidad

Observación: Da eq. de obs  $\alpha_{t+1} = T_t \alpha_t + c_t + R_t \eta_t$

recursivamente:

$$\alpha_{t+1} = T^t \alpha_1 + \sum_{j=0}^{t-1} T^j R_j \eta_{t-j}$$

$$\Rightarrow E[\alpha_t \alpha_{t+h}] = 0 \text{ por}$$

$$E[\alpha_{t+h} \eta_t] = 0$$

$$\text{e } E[\alpha_{t+h} \alpha_t] = 0.$$

Logo:  $\Gamma(h) = \sum_{h,t} T_h(\eta_t) \geq' n \neq 0$ .

$$= \sum_{0,t} T_\alpha(\eta_t) \geq' + H \neq h = 0 \quad , \quad \Gamma(h, t)$$

para y estacionario  
depende de  $\alpha$ "

( $\Gamma_{\alpha}(h)$   
mas de  
de do  
temp)

explicativamente

$$(1) \quad y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\text{proceso estacionario} \quad \mu_{t+1} = \phi \mu_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \quad |\phi| < 1$$

Proceso AR(1)

$\therefore y_t$ : é um AR(1) + um choque

que é AR(1) é a componente  $\mu_t$

$$E[\mu_t] = 0$$

$$\text{var}[\mu_t] = \delta(0) = \frac{\sigma_\eta^2}{(1-\phi^2)}$$

$$r(h) = \phi^h \sigma_\eta^2$$

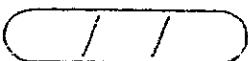
$$(1-\phi^2)$$

equilíbrio

$$r(h) = \frac{\delta(h)}{r(0)} = \phi^h ?$$

infos para

AR(1)



$$\Rightarrow E[y_t] = E[\mu_t] = 0$$

$$\text{var}[y_t] = \gamma_y(0) = \sigma_\eta^2 + \sigma_\epsilon^2$$

$(1-\phi^2)$

$$\gamma_y(h) = \gamma_\mu(h)$$

Na realidade, basta aplicarmos a fórmula anterior obtida de forma geométrica se caso matricial

processo  
não estacionário (II)

$$\left\{ \begin{array}{l} y_t = \mu_t + \epsilon_t \quad \Rightarrow \text{chamado de tendência local} \\ \mu_{t+1} = \mu_t + \eta_t \quad \rightarrow \text{tendência com drift} \\ \qquad \qquad \qquad (\text{passo aleatório}) \end{array} \right.$$

$$E[\mu_t] = ?$$

Obs: Ituando  $\mu_{t+1}$ ,

$$\mu_2 = \mu_1 + \eta_1$$

$$\mu_3 = \mu_2 + \eta_2 = \mu_1 + \eta_1 + \eta_2$$

$$\Rightarrow \mu_{t+1} = \mu_1 + \sum_{s=1}^t \eta_s$$

$$\text{Logo, } E[\mu_{t+1}] = a_1 \quad (\text{da cond. inicial})$$

$$\text{var}[\mu_{t+1}] = P_1 + t\sigma_\eta^2$$

$\text{var}(P_1) \quad \text{soma de variâncias}$

de choques independentes

$\Rightarrow \text{var}(\mu)$  - depende de  $t$

: processo  $y$  será mais estacionário

$$y_t(0,t) = (\rho_1 + t\sigma_y^2) + \sigma_E^2$$

divida  
Nascerão  
 $x_t(t, \cdot)$

Logo:

Estacionariedade dependerá sempre de como o processo  $\alpha$  comporta

— “ —

Da demonstração, obtemos:

queremos olhar este processo.

$$E[y_t] = z E[\alpha_t]$$

$$\Gamma_y(h,t) = \begin{cases} z \Gamma_z(h,t) z' & h \neq 0 \\ z \Gamma_z(0,t) z' + H & h = 0 \end{cases}$$

$$\alpha_{t+1} = T^t \alpha_t + \sum_{j=0}^{t-1} T^j R \eta_j$$

$$E[\alpha_{t+1}] = T^t E[\alpha_t] = T^t a_i = f(t)?$$

queremos avaliar se é ou não função do tempo.

$\Rightarrow$  Separar que os autovalores de  $T$  são todos distintos

autovalor

$$\text{Lembando: } TF_i = \lambda_i F_i \quad i = 1, 2, \dots, m$$

$\downarrow$   
autovetor

$$TF_i - \lambda_i F_i = 0 \therefore (T - \lambda_i I) F_i = 0$$

Como  $F_i \neq 0$

$\Rightarrow \det(T - \lambda_i I) = 0$ : equações características

dará  $m$ . autovalores  $\neq s$

último

(queremos achar as condições p/ a qual  $\alpha_t$  seja estacionário de 2º orden)

/ /

Vale a decomposição espectral:

$\Rightarrow F^t$  (transp. = inversa pois matriz ortogonal)

$$T = F \Lambda F^t$$

$$\text{máx } m \times m \quad \text{diagonal } (\lambda_1, \lambda_2, \dots, \lambda_m) = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix}$$

(autovalores)

obs: se algum autovalor igual, há outras decomposições

Prova-se que:

$$T^t = F \Lambda^t F^{-1}$$

Assim segue que:

$$E[\alpha_{t+1}] = T^t a_i = F \Lambda^t F^{-1} a_i =$$

$$\left( \begin{array}{c} E[\alpha_{1,t+1}] \\ E[\alpha_{2,t+1}] \\ \vdots \\ E[\alpha_{m,t+1}] \end{array} \right) = \left( \begin{array}{cccc} f_{11} & f_{12} & \dots & f_{1m} \\ f_{21} & f_{22} & \dots & f_{2m} \\ \vdots & \vdots & & \vdots \\ f_{m1} & f_{m2} & \dots & f_{mm} \end{array} \right) \left( \begin{array}{cccc} \lambda_1^t & & & \\ & \lambda_2^t & & \\ & & \ddots & \\ & & & \lambda_m^t \end{array} \right) \left( \begin{array}{c} f^{11} & f^{12} & \dots & f^{1m} \\ f^{21} & & & \vdots \\ \vdots & & & f^{m1} & \dots & f^{mm} \end{array} \right)^{-1} a_i$$

$$\Rightarrow E[\alpha_{i,t+1}] = \delta_1 \lambda_1^t + \delta_2 \lambda_2^t + \dots + \delta_m \lambda_m^t, \quad \delta_i = f_{i1} f^{11} a_{i1}, \quad i=1,2,\dots$$

vale para todos os  $\alpha_i$

$\Rightarrow E[\alpha_{i,t+1}]$  depende explicitamente de  $t$

Mas...

Se  $|\lambda_i| < 1$ ,  $\forall i = 1, \dots, m$ , para  $t \rightarrow \infty$ :  $E[\alpha_i, t] \rightarrow 0$

cond. necessária e suficiente p/ estacionariedade (será tb p/ 2ª ordem).

dade de 1ª ordem

equilíbrio

obs: Essas simulações  $\Rightarrow$  "warming up": x afastar da origem  
 descartar primeiras simulações  
 pf garantir estacionariedade  
 "análogo a  $t \rightarrow \infty$ .

Queremos agora

$\Gamma_\alpha(h, t) = ?$  para avaliar a estacionariedade  
 de 2º ordem

$$\Gamma_\alpha(h, t) = E[(\alpha_{t+h} - E[\alpha_{t+h}]) (\alpha_{t+h+h} - E[\alpha_{t+h+h}])']$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t = T^t \alpha_1 + \sum_{j=0}^{t-1} T^j R \eta_{t-j}$$

$$E[\alpha_{t+1}] = T E[\alpha_t] = T^t \alpha_1$$

$$\alpha_{t+1} - E[\alpha_{t+1}] = T^t (\alpha_1 - \alpha_1) + \sum_{j=0}^{t-1} T^j R \eta_{t-j}$$

$$\alpha_{t+h+1} - E[\alpha_{t+h+1}] = T^{t+h} (\alpha_1 - \alpha_1) + \sum_{j=0}^{t+h-1} T^j R \eta_{t+j}$$

$$\Rightarrow \Gamma_\alpha(h, t) = E[(\alpha_{t+1} - E[\alpha_{t+1}]) (\alpha_{t+h+1} - E[\alpha_{t+h+1}])']$$

$$= E \left[ \left[ T^t (\alpha_1 - \alpha_1) + \sum_{j=0}^{t-1} T^j R \eta_{t-j} \right] \left[ (\alpha_1 - \alpha_1)' T^{t+h} + \sum_{j=0}^{t+h-1} \eta_{t+j}' R' T^j \right] \right]$$

obs:  $E[\cdot]$  das

termos cruzados  $= 0$

$$= T^t E[(\alpha_1 - \alpha_1)(\alpha_1 - \alpha_1)'] T^{t+h} + E \left[ \sum_{j=0}^{t-1} T^j R \eta_{t-j} \cdot \sum_{j=0}^{t+h-1} \eta_{t+j}' R' T^j \right]$$

$P_1$

/ /

$$\text{Para } h=0 : \Gamma_x(0, t) = T^t P_i T^{t'} + \sum_{j=0}^{t-1} T^j (R Q R') T^j$$

Poderemos mostrar que fazendo

$$|\lambda_i| < 1$$

e  $t \rightarrow \infty$

poderemos tentar con  
derar a decomp. espectral  
e fazer  $T = F \Lambda F^{-1}$   
para demonstrar

$$\Rightarrow \Gamma_x(0, t) \rightarrow \Sigma(0) \text{ matriz que independe de } t$$

obs: No Harvey-Koopman, não há preocupações com isso pois  
tratam de processos não estacionários.

Oby 2: Esta demo pode ser mais intuitiva se pensarmos em  
um processo autoregressivo simples

$$\alpha_t = \phi \alpha_{t-1} + \eta_t$$

$\downarrow T$

$$\alpha_t = \alpha_0 \phi^t + \sum \phi^j \eta_{t-j}$$

$$\Rightarrow E[\alpha_t] = \alpha_0 \phi^t \rightarrow 0$$

$$\alpha_{t+h} = \alpha_0 \phi^{t+h} + \sum \phi^j \eta_{t+h-j}$$

$$\delta(h) = E[(\alpha_t - E(\alpha_t))(\alpha_{t+h} - E(\alpha_{t+h}))] = \dots \rightarrow \text{converge}$$

se  $|\phi| < 1$ .

Supondo que as condições de estacionariedade de 2ª ordem  
sao satisfeitas ( $|\lambda_i(t)| < 1$  e  $t \rightarrow \infty$ ), pode-se obter uma  
expressão geral para a matriz de covariância de  $\alpha_t$ :

$$\alpha_{t+1} = T\alpha_t + R\eta_t$$

Já vimos que:  $E[\alpha_{t+1}] = 0$ .

$$\text{var}[\alpha_{t+1}] = ?$$

Algumas operações em Álgebra Matricial

i. Seja  $A \sim mxn$

$$A = \begin{pmatrix} (a_{11}) & (a_{12}) & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & & & a_{mn} \end{pmatrix}$$

$$\Rightarrow \text{vec}(A) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

propriedade  
de empilhamento

ii. Seja  $A \sim mxn$

$$B \sim p \times q$$

cada elemento da matriz A  
multiplica a matriz B

$$A \otimes B \stackrel{\text{Kronecker}}{\equiv} \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}_{mp \times nq}$$

/ /

$$\text{iii) } \text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$$

Agora, usando que  $\alpha_{t+1} = T\alpha_t + R\eta_t$  e condições de estacionariedade de 2ª ordem:

$$\text{var}(\alpha_{t+1}) = \text{var}(T\alpha_t) + \text{var}(R\eta_t) = \text{var}(\alpha_t) = \Sigma_\alpha(0)$$

$$\text{var}(\alpha_{t+1}) = T\text{var}(\alpha_t)T' + R\text{var}(\eta_t)R'$$

$$\text{var}(\alpha_{t+1}) = T\text{var}(\alpha_t)T' + RQR'$$

↓ ctc. por hipótese do sistema

$$= \text{var}(\alpha_t) = \Sigma_\alpha(0)$$

$$\text{vec} \Sigma_\alpha(0) = \text{vec}(T\Sigma_\alpha(0)T') + \text{vec}(RQR')$$

$$= (T \otimes T) \text{vec}(\Sigma_\alpha(0)) + (R \otimes R) \text{vec}(Q)$$

$$\text{vec} \Sigma_\alpha(0) - (T \otimes T) \text{vec}(\Sigma_\alpha(0)) = (R \otimes R) \text{vec}(Q)$$

$$(I - T \otimes T) \text{vec}(\Sigma_\alpha(0)) = (R \otimes R) \text{vec}(Q)$$

$$\therefore \text{vec}(\Sigma_\alpha(0)) = (I - T \otimes T)^{-1} \text{vec}(Q)$$

Exercício:

Considere os seguintes modelos em EE

$$\begin{aligned} y_t &= z \alpha_t + \varepsilon_t \\ &\quad x_{t+1} = T \alpha_t + R \eta_t \\ (1) \quad y_t &= \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \varepsilon_t \end{aligned}$$

$$\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} \quad \text{tilibra}$$

obs:  $\text{det}(T) = 1$

$$\eta_t = \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \rightarrow E[\eta_t \eta_t^T] = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix} = \Phi$$

Equações:  $\left\{ \begin{array}{l} y_t = \mu_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \beta_t + \eta_t \\ \beta_{t+1} = \beta_t + \xi_t \end{array} \right.$

modelo linear local

tendência linear estocástica.

Obs: Se  $\sigma_\eta^2 = \sigma_\xi^2 = 0 \Rightarrow y_t = a + bt + \varepsilon_t$  : tendência determinística  
é caso particular

Calculando autovalores de  $T$ :

$$|T - \lambda I| = 0$$

$$\left| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = \left| \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} \right| = (1-\lambda)^2 = 0$$

$$\lambda_1 = \lambda_2 = 1$$

$\Rightarrow$  processo não estacionário

ds: se existir tiver-se 3ª componente  $\Rightarrow$  eq. característica  
seria polinomial de  
3º grau.

Vetor de não é

estacionário de 2º orden

Não adianta procurarmos

$\Gamma_y(0)$  e  $\Gamma_\alpha(0)$ , pois  
dependem de  $t$ .

tilibra

/ /

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

$$\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}$$

a) Estabelecer sob que condições  $|\lambda_i(t)| < 1 \quad i=1,2,\dots$

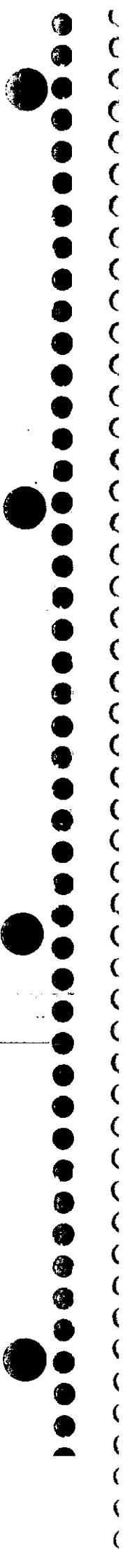
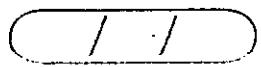
(em  $\phi_1 \in \Phi$ )

(processo nra' estoc. de 2º ordem)

b) Avaliar  $\lim_{t \rightarrow \infty} \Gamma_x(0, t)$ ,  $\Gamma_x(h, t)$  e consequentemente

$\lim_{t \rightarrow \infty} \Gamma_y(0, t)$ ,  $\Gamma_y(h, t)$  mostrando que condições

impostas nos  $\phi$ 's são suficientes para estabelecer  
estacionariedade de 2º ordem para  $x_t$  e  $y_t$ .



Killbora

só fazem sentido pf sistemas invariantes no tempo.

23/03/11

No caso da estacionariedade, vemos que depende dos autovalores de  $T$ .

### Propriedades de MEE (cont.)

Observabilidade: motivada pela seguinte pergunta:

ado que  
observamos y  
nos últimos nun,  
como posso ver  
minha trajetória  
de x no mesmo  
intervalo.

"Em que condições é possível estabelecer, em um intervalo de tempo finito, a trajetória temporal do estado  $x_t$  de um sistema dinâmico, dada a história do vetor de medidas  $y_t$  no mesmo intervalo de tempo?"

Seja um MEE linear, discreto, invariante no tempo.

$\rightarrow p \times 1 \rightarrow p \times m$   $\rightarrow$  Naq se é necessário considerar sistema p-varável.

$$\begin{cases} y_t = 2x_t - mx_0 & (I) \\ x_{t+1} = Tx_t + R\eta_t \end{cases}$$

$\downarrow mx_0 \quad \downarrow mx_m \quad \rightarrow mx \times 1$

Este sistema será dito observável se  $x_0$  pode ser determinado a partir de um conjunto de medidas de  $y$ .

Vale pf A condição necessária e suficiente será dita a seguir:  
qualquer tempo, mas fazemos  $\eta_t = 0 \forall t$  em I, segue que:

$$x_{t+1} = Tx_t \Rightarrow x_{t+1} = T^t x_0$$

$x_t = T^t x_0$  - pf linkar com eq. das obs.

$$\text{Mas } y_t = 2x_t = 2T^t x_0$$

$(x \sim mx \times 1)$  - temos que obter m medidas do vetor  $y$ .

$$p \times 1 \rightarrow y_0 = 2x_0 \xleftarrow{mx \times 1}$$

$$p \times 1 \rightarrow y_1 = 2T x_0$$

$$y_2 = 2T^2 x_0$$

⋮

$$y_{m-1} = 2T^{m-1} x_0$$

$$\Rightarrow \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{m-1} \end{pmatrix} = \begin{pmatrix} 2 \\ 2T \\ 2T^2 \\ \vdots \\ 2T^{m-1} \end{pmatrix} \begin{pmatrix} x_0, 1 \\ x_0, 2 \\ x_0, 3 \\ \vdots \\ x_0, m \end{pmatrix}$$

$x_0 \rightarrow mx \times 1$  libra

essa solução tem que ser única pf todas as medidas

$$\begin{matrix} y = M \alpha_0 \\ \downarrow \quad \downarrow \quad \downarrow \\ p_m x_1 \quad p_m x_m \end{matrix}$$

- Assim, segue que para determinarmos os elementos de  $\alpha_0$  de forma única que a condição necessária e suficiente seja:

$\text{posto}(M) = m$  (é exatamente o tamanho de  $\alpha_0$  e faz com que haja solução única para  $\alpha_0$  no sistema acima)

$\text{posto } M = \# \text{máximo de linhas / colunas linearmente independentes}$

- Como  $\text{posto } M = \text{posto } M'$ , então

$$\text{posto } M' = \text{posto } (z', T'z', \dots, (T^{m-1})'z') = m$$

- Se  $p=1$ , i.e.,  $y_t$  é escalar  $M \sim m \times m$ , e, assim, a condição de observabilidade pode ser investigada checando se  $\det(M) \neq 0$ . → pois  $p=1 \Rightarrow M$  será uma matriz quadrada

Eex:

(Geralmente, nesta parte do curso)

$$y_t = \mu_t + \varepsilon_t$$

trataremos modelos univariados.

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t$$

$p=1 \Rightarrow$  uma única série temporal

$$\text{modelos} \quad \beta_{t+1} = \beta_t + \xi_t$$

para verificar, neste caso

de tendência linear

checar se  $\det \neq 0$  ou se

estocástica

há uma linha que

$$\Rightarrow y_t = a + b t + \varepsilon_t$$

seja cl das outras

(tendência determinística)

$$\Rightarrow y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} z \\ \alpha_t \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix}$$

tilibria

⇒ Questão: Será que conseguimos recuperar  $\alpha_t$ 's dados  $y_t$ 's?

11

$$z = \begin{bmatrix} 1 & 0 \end{bmatrix} \Rightarrow z' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\therefore T'z' = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$M' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ tem posto } 2 = m \Rightarrow \text{é observável.}$$

2)  $y_t = \mu_t + \epsilon_t$

$$\mu_{t+1} = \mu_t + \eta_t$$

$$\beta_{t+1} = \beta_t + \xi_t$$

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

$$\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix}$$

$$M' = (z', T'z') \quad T'z' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

 $\ell_1, \ell_2$ 

$$\Rightarrow M' = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \Rightarrow \text{posto}(M') = 1, \text{ pois } \ell_1 = \ell_2 : \text{sistema} \neq 2 \text{ não observável}$$

3) Chequem para

não podemos estimar

modelos que não é

observável. Modelos

não faz sentido.

$$\left\{ \begin{array}{l} y_t = \mu_t + \beta_t + \epsilon_t \\ \mu_{t+1} = \phi_1 \mu_t + \eta_t \\ \beta_{t+1} = \phi_2 \beta_t + \xi_t \end{array} \right.$$

parâmetros

trivacos

11

Estabelecer condições para  $\phi_1$  e  $\phi_2$  de forma que o sistema seja observável.

- Se  $\phi_1 = \phi_2 = 1 \rightarrow$  não é observável  $\Rightarrow M' = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow \text{rank}(M') = 1$

$$\theta_t = \mu_t + \beta_t$$

$$\psi_t = \mu_t - \beta_t$$

$$\therefore y_t = \mu_t + \beta_t + \varepsilon_t$$

$$y_t = \theta_t + \varepsilon_t$$

$$\theta_{t+1} = \theta_t + \delta_{1t}$$



$$\theta_{t+1} = \mu_{t+1} + \beta_{t+1} = (\mu_t + \beta_t) + (\eta_t + \varepsilon_t) \\ = \theta_t + \delta_{1t}$$

$$\text{Também: } \psi_{t+1} = \mu_{t+1} - \beta_{t+1} \\ = (\mu_t - \beta_t) + (\eta_t - \varepsilon_t) \\ = \psi_t - \delta_{2t}$$

$$\Rightarrow \begin{cases} y_t = [1 \ 0] [\theta_t] \\ \psi_t \end{cases}$$

ML2

→ sobra na especificação

y não depende de  $\psi$ . Nem  $\theta$  e  $\psi$  se relacionam

parametrizações

$$\begin{bmatrix} \theta_{t+1} \\ \psi_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_t \\ \psi_t \end{bmatrix} + \begin{bmatrix} \delta_{1t} \\ \delta_{2t} \end{bmatrix}$$

Nesta parametrização fica + fácil de verificar.

Nos modelos que usaremos aqui  $\Rightarrow$  termos que se observáveis sejam adequados.

/ /

obs: No exercício 1, se tivermos trabalhando com  $p=2$   
 $(y \sim 2x_1)$

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ p_{1t} \\ p_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim N(0, \Sigma)$$

Nesse

caso não se

comunicam aqui,

mas podem se comunicar nos  $E$ .

— II —

A questão da observabilidade é obrigatória na estimacão dos parâmetros do modelo.

A questão da estacionariedade nas. É da definição do modelo.

Na cond. inicial será importante definir quais componentes

Nas estacionariedades e quais não.

Nas estacionariedades

(obs: cond. inicial  $\alpha_i \sim N(\alpha_i, P_i)$ )

traz problemas

$\begin{pmatrix} \alpha_{1,0} \\ \alpha_{1,1} \\ \alpha_{1,2} \end{pmatrix}$  componentes

nos modelos de

de  $\alpha$ ,

reflexos, etc.

obs: Existe uma correspondência entre modelos estruturais e modelos ARIMA de B&J.

— III —

Estimacões de MEE lineares e Gaussianas

ep. do Filtro de Kalman nas exigem que o sistema

$$y_t = z_t \alpha_t + \varepsilon_t \quad \text{seja invariante no tempo.}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$$

De forma geral,  $z_t$

$(t = 1, \dots, n)$  tamanho  
da série (out)

onde  $\varepsilon_t \sim N(0, H_t)$

$$\eta_t \sim N(0, Q_t)$$

obs: Filtro de Kalman II

Mas considerando Gaussianidade, esse é o único que é ótimo.

nas exige dist. Gaussiana  
dos choques, apenas dependendo  
sobre os 1º momentos →

Se Gaussianidade  $\Rightarrow$  FK dará ótimo local  
(melhor estimativa linear)

1 / 1

O que queremos?

- (1) Estimar  $x_t$  dado certo conjunto de informações ( $y_t$ )  
↳ observável.

Filtros de Kalman

$Y_{t-1} = (y_1, y_2, \dots, y_{t-1}) \rightarrow$  previsão (usa dados até  $t-1$ )

$\tilde{Y}_t = (y_1, y_2, \dots, y_t) \rightarrow$  atualização (atualiza com dado novo)

$Y_n = (y_1, y_2, \dots, y_t, y_{t+1}, \dots, y_n) \rightarrow$  suavização (passo que usa todos os dados disponíveis)

$\tilde{y}_j = \begin{cases} j = t \\ j = t-1 \\ j = n \end{cases}$

Xine para fazer extracções de componentes (económica, comum temporal e sazonalidade)

$p(x_t | y_j)$ ?  $\rightarrow$  estimativas de  $x_t$  é ter esta densidade  
(densidade condicional a

↓  
 $E[x_t | y_j]$   
 $\text{var}[x_t | y_j]$

partir do que observamos em  $y$ )

No caso: veremos a estimativa usando previsão e atualizações  $\Rightarrow$  Filtro de Kalman  
veremos um tipo de suavização.

Toda dedução do FK, pressupõe que quaisquer elementos desconhecidos em alguma matriz serão considerados conhecidos. (faltas quase certezas nas matrizes de covariância)

(2) Hiperparâmetros

↳ nas matrizes do sistema por MV  
↳ dados faltantes nestas matrizes.

/ /

usando pesos

Dois âmbitos de estimativas } do vetor de estados } e atualização  
 das hiperparâmetros das matrizes } do sistema

Considere o modelo linear local

$$\begin{cases} y_t = \mu_t + \epsilon_t & \epsilon_t \sim N(0, \sigma_e^2) \\ \mu_{t+1} = \mu_t + \eta_t & \eta_t \sim N(0, \sigma_\eta^2) \end{cases} \quad \mu_t \sim N(a_t, P_t)$$

$$E[\epsilon_t | \mu_t] = E[\eta_t | \mu_t] = 0 \quad \forall t$$

$$E[\epsilon_t, \eta_s] = 0 \quad \forall t, s$$

Em princípio, podemos obter  $p(x_t | y_t)$  sem fazer uso do FK, apenas recorrendo a propriedades de vetores normais multivariados.

Para  $y$

$$y_1 = \mu_1 + \epsilon_1$$

$$y_2 = \mu_2 + \epsilon_2 = \mu_1 + \eta_1 + \epsilon_2$$

:

$$y_n = \mu_1 + (\eta_1 + \eta_2 + \dots + \eta_{n-1}) + \epsilon_n$$

$$y = (y_1, y_2, \dots, y_n)^T \Rightarrow E[y] = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{1}a_i = E[y]$$

$$\text{var}(y) = E[(y - E[y])(y - E[y])^T] = R$$

$$= \begin{bmatrix} \text{var}(y_1) & \text{cov}(y_1, y_2) & \dots & \text{cov}(y_1, y_n) \\ \text{cov}(y_2, y_1) & \text{var}(y_2) & & \\ \vdots & & \ddots & \\ & & & \text{var}(y_n) \end{bmatrix}_{n \times n}$$

1 /

## • Variâncias

$$\text{var}(y_1) = p_1 + \sigma_e^2$$

$$\text{var}(y_2) = p_1 + \sigma_\eta^2 + \sigma_e^2$$

:

$$\text{var}(y_n) = p_1 + (n-1) \sigma_\eta^2 + \sigma_e^2$$

## • covariâncias

$$\text{cov}(y_1, y_2) = E[(y_1 - a_1)(y_2 - a_2)]$$

$$= E[(\mu_1 + \varepsilon_1 - a_1)(\mu_1 + \eta_1 + \varepsilon_2 - a_2)] = p_1$$

Poderemos escrever:

$$R = 11' p_1 + \Sigma_{ij} \quad \text{onde } \Sigma_{ij} = \begin{cases} (i-1) \sigma_\eta^2 & , i < j \\ \sigma_e^2 + (i-1) \sigma_\eta^2 & , i = j \\ (j-1) \sigma_\eta^2 & , i > j \\ \end{cases} \quad i = 1, \dots, n$$

↓  
cresce com  $n$   
(com o  $n$ : de observações)

Do ponto de vista operacional,

estariamos usando

matrizes enormes

se tivemos muitas

observações.

(Nas será este caminho  
a ser usado).

$$\mu_{t+1} = \mu_t + \eta_t$$

$$\mu_2 = \mu_1 + \eta_1$$

$$\mu_3 = \mu_2 + \eta_2 = \mu_1 + \eta_1 + \eta_2$$

:

$$\mu_n = \mu_1 + (\eta_1 + \eta_2 + \dots + \eta_{n-1})$$

$$\Rightarrow \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \Rightarrow E[\mu] = 1 a_1$$

/ /

$\text{var}(\mu)$  ... terá a ver com  
termos  $p_i$  e  $\sigma_{\eta}^2$  na  
matriz  $R$  (var de  $y$ )

Desenvolvendo, podemos escrever:

$$\begin{pmatrix} y \\ \mu \end{pmatrix} \sim N \left( \begin{pmatrix} 1a_1 \\ 1a_1 \end{pmatrix}, \begin{pmatrix} R & \Sigma_{y\mu} \\ \Sigma_{\mu y} & \Sigma_{\mu\mu} \end{pmatrix} \right) \rightarrow E[(y - E[y])(\mu - E[\mu])']$$

$$E[(\mu - E[\mu])(\mu - E[\mu])']$$

Resultado: (Teorema)

seja  $(x, y) \sim N \left[ \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \right]$

Será i.  
exterior  
variancia  
usado

Então:

$$p(x | y = y) \sim N \left( \underbrace{\mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y)}_{\mu_{x|y}}, \underbrace{\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}}_{\Sigma_{x|y}} \right)$$

Faça  $x = \mu$ ,  $y = y$  e assim:

$$p(\mu | y = y) \sim N(E[\mu | y = y], \text{var}[\mu | y = y]) \quad \text{Fazer!}$$

etc.  $\Rightarrow$  obter expressão explícita.

→ mesmo nas usando FK, p/ este modelo apresentado,  
enviando de forma multivariada e estimando ( $\alpha$  na fronteira  
 $\sigma_{\epsilon}^2 = \sigma_{\eta}^2$  por MV, obtemos o resultado acima p/  
vetores multivariados normais.  $\Rightarrow$  Dist. condicional  
de  $\mu$  dado  $y$ . Vida

$$\begin{matrix} Y = m \times 1 \\ Z = n \times 1 \end{matrix} \quad Y'Z' = 1 \times m \times 1 \times n$$

Este processo de estimativas do desconhecido é válido, mas não é factível do ponto de vista operacional. Nas é sequencial. Apresenta matrizes muito grandes.

Neste caso, estamos fazendo estimativa suavizada:

$$E[\mu|Y] = \begin{pmatrix} E[\mu_1|Y] \\ E[\mu_2|Y] \\ \vdots \\ E[\mu_n|Y] \end{pmatrix}$$

usamos todos  
as informações  
→ usaremos o resultado  
anterior bastante agora

Lema do Resultado anterior:

Seja  $\begin{array}{|c|c|} \hline X & \sim N \left( \begin{pmatrix} \mu_X \\ \mu_Y \\ \vdots \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz} \end{pmatrix} \right) \\ \hline Y & \\ Z & \end{array}$

$X, Y, Z:$  vetores de dimensões arbitrárias. Então:

particularizações  
será usada no FK.

$$(x|y=y, z=z) \sim N [E(x|y=y, z=z), \text{var}(x|y=y, z=z)]$$

onde:

$$E(x|y=y, z=z) = \mu_{x|y} + \Sigma_{xz} \Sigma_{zz}^{-1} z$$

$$\text{var}(x|y=y, z=z) = \Sigma_{xx} - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx}$$

Prova: aplique o teorema anterior substituindo  $Y$  por  $(Y'Z')$

$(1 \times m \times 1 \times n)$

$(1 \times m)$   
 $(1 \times n)$

Fazer!

1 / 1

$$\begin{cases} y_t = z_t \alpha_t + \varepsilon_t & (a) \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t & (b) \end{cases}$$

Notações:  $(\alpha_t | Y_{t-1}) \Rightarrow E[\alpha_t | Y_{t-1}] = \alpha_t$   
 $\text{var}[\alpha_t | Y_{t-1}] = P_t$

ou  $\alpha_{t+1} = E[\alpha_{t+1} | Y_t] ; \alpha_{t|t} = E[\alpha_t | Y_t]$

$P_{t+1} = \text{var}[\alpha_{t+1} | Y_t] ; P_{t|t} = \text{var}[\alpha_t | Y_t]$

1º passo: distribuições da previsão de  $\alpha_{t+1} \Rightarrow$  usar eq. (a)

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$$

$$\therefore E[\alpha_{t+1} | Y_t] = T_t E[\alpha_t | Y_t]$$

$$\boxed{\alpha_{t+1} = T_t \alpha_{t|t}} \rightarrow \text{No koopman } T^T E[\alpha_t | Y_t]$$

(questas de notações)

$$\text{var}[\alpha_{t+1} | Y_t] = P_{t+1} = \text{var}[\underbrace{T_t \alpha_t + R_t \eta_t}_{A_t} | Y_t]$$

$$= E[(A_t - E[A_t | Y_t])(A_t - E[A_t | Y_t])']$$

Mas:  $A_t = T_t \alpha_t + R_t \eta_t$

$$\therefore E[A_t | Y_t] = T_t \alpha_{t|t}$$

$$\therefore A_t - E[A_t | Y_t] = \underbrace{T_t \alpha_t + R_t \eta_t - T_t \alpha_{t|t}}_{B_t} = T_t (\alpha_t - \alpha_{t|t} + R_t \eta_t)$$

$$\therefore [A_t - E[A_t | Y_t]]' = \eta_t' R_t' + (\alpha_t - \alpha_{t|t})' T_t'$$

Logo:

$$\begin{aligned}
 E[B_t B_t^T] &= P_{t+1} = \\
 &\quad \underset{0}{\text{(termos anulares)}} \quad P_{t+1} \\
 &= E[T_r (\alpha_t - a_{t+1}) \eta_r^T R_t^T] + E[T_r (\alpha_t - a_{t+1}) (\alpha_t - a_{t+1})^T T_r^T] + \\
 &\quad + E[R_t \eta_t^T (\alpha_t - a_{t+1})] + E[R_t \eta_t^T \eta_{t+1}^T R_t^T] \\
 &\quad \underset{0}{\text{acertar as}} \quad \underset{\text{expressões.}}{\text{expressões.}}
 \end{aligned}$$

$$\therefore P_{t+1} = T_t P_{t|t} T_t' + R_t Q_t R_t'$$

### Previsas

### No modelo de Nível Local

$$\begin{aligned} \text{1) } y_t &= \mu_t + \epsilon_t \xrightarrow{\epsilon \text{ puro } \phi: \text{ Passeio aleatório}} N(0, \sigma_e^2) \\ &\quad \downarrow \\ \text{depois} \quad \mu_{t+1} &= \mu_t + \eta_t \xrightarrow{\eta \text{ puro } \phi: \text{ Tendência mas com stochasticidade}} N(0, \sigma_\eta^2) \end{aligned}$$

d sinus  
nas estacio-  
nárias

Importante: sempre importante ver como é o comportamento  
na variância

### 2) Forma Reduzida

$$\Phi_p(L) D^d y_t = c + \Theta_q(L) \epsilon_t \Rightarrow \text{ARIMA}(p, d, q) + c$$

→ questões de identificação pode ocorrer em modelos estruturais

$$\text{Por exemplo } \rightarrow y_t = \alpha + \beta x_t + \epsilon_t \quad \text{se } \epsilon_t \sim N(0, \sigma_e^2)$$

$$E(y_t | x_t) = (\alpha + c) + \beta x_t$$

Na regressão

$$\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 x_t \Rightarrow \hat{r}_0 = \alpha + c \quad \begin{cases} 2 \text{ equações,} \\ \hat{\gamma}_1 = \beta \quad 3 \text{ intervalos (pa-} \\ \text{metros)} \end{cases} \Rightarrow \text{N identificável}$$

Todo modelo ARIMA, por construção, é sempre identificável



escrevemos modelos estruturais na forma ARIMA  $\gamma$  e ver  
se é identificável ou não.

11

pg 10

Forma reduzida:

$$\left\{ \begin{array}{l} y_t = \mu_t + \epsilon_t \\ \mu_{t+1} = \mu_t + \eta_t \end{array} \right.$$

$$\hookrightarrow \text{vale } \mu_t = \mu_{t-1} + \eta_t$$

Tanto faz  
pois dist  $\eta_t$   
é estacionária

Tira 1ª diferença:

$$\Delta y_t = \Delta \mu_t + \Delta \epsilon_t$$

$$\Delta \mu_t = \eta_t$$

$$\Rightarrow \Delta y_t = \eta_t + \Delta \epsilon_t$$

c.l. de processos estacionários  $\Rightarrow$  podemos prosseguir

forma ARIMA

leve a esta transformação

$$\rightarrow E[\Delta y_t] = 0$$

$$\rightarrow \text{var}(\Delta y_t) = \text{var}(\eta_t) + \text{var}(\Delta \epsilon_t) = \sigma_\eta^2 + 2\sigma_\epsilon^2$$

$$\hookrightarrow \text{var}(\epsilon_t - \epsilon_{t-1}) = \text{var}(\epsilon_t) + \text{var}(\epsilon_{t-1}) - 2\text{cov}(\epsilon_t, \epsilon_{t-1})$$

E como saber qual o termo correspondente?

$$\text{MA}(q) \Rightarrow p(k) \left\{ \begin{array}{l} \neq 0 \quad k=1, 2, \dots, q \\ = 0 \quad k=q+1, \dots \end{array} \right.$$

$$\text{AR}(p) \Rightarrow p(k) \left\{ \begin{array}{l} \text{exponencial ou} \\ \text{nóide amortecida} \end{array} \right.$$

$$\rightarrow \text{FAC: } p(k) = \frac{\text{cov}(\Delta y_t, \Delta y_{t-k})}{\text{var}(\Delta y_t)}$$

/ /

 $k=1:$ 

$$\text{cov}(\Delta y_t, \Delta y_{t+1}) = E[\Delta y_t \Delta y_{t+1}] \rightarrow \text{média} = 0 \text{ e } t$$

$$= E[(\eta_t + \epsilon_t - \epsilon_{t-1})(\eta_{t+1} + \epsilon_{t+1} - \epsilon_{t-2})] = \\ = -\sigma_\epsilon^2$$

$$\therefore \rho(1) = \frac{-\sigma_\epsilon^2}{\sigma_\eta^2 + 2\sigma_\epsilon^2} = \frac{-1}{q+2}$$

onde  $q = \sigma_\eta^2/\sigma_\epsilon^2$ 

$$\text{Para } k \geq 2: \text{ cov}(\Delta y_t, \Delta y_{t+k}) = 0$$

↓

$$E[(\eta_t + \epsilon_t - \epsilon_{t-1})(\eta_{t+j} + \epsilon_{t+j} - \epsilon_{t-j-1})]$$

↓

pt  $j \geq 2$ , mas há nenhum par no termo anterior

$\Rightarrow E$  é um processo MA(1) —  $F_{k+1} \neq 0$  pt  $k=1$

$= 0$  pt  $k \geq 2$ .

Resultado: Se  $y_t$  é MNL  $\Rightarrow \Delta y_t \sim \text{MA}(1)$

Suponha que vamos estimar a 1ª diferença do processo como MA(1)

Seja  $z_t \sim \text{MA}(1)$

$$\therefore z_t = w_t + \theta w_{t-1}, \quad w_t \sim N(0, \sigma_w^2)$$

$$\begin{cases} E[z_t] = 0 \\ \text{var}[z_t] = (1 + \theta^2) \sigma_w^2 \\ \rho(1) = E[z_t z_{t-1}] = \theta \sigma_w^2 \end{cases}$$

$$\Rightarrow \rho(1) = \frac{\theta}{(1 + \theta^2)}$$

2010

1.1.

Queremos responder:

$$\text{Pf } \delta y_t \Rightarrow 0 < q < \infty \quad (q \text{ é o parâmetro a ser estimado na forma reduzida})$$
$$\text{Pf } z_t^2 \Rightarrow -1 < \theta < 1$$

$$\text{Pf } \delta y_t \Rightarrow p(1) = \frac{-1}{q+2}$$

$$\text{Pf } z_t^2 \Rightarrow p(1) = \frac{\theta}{1+\theta^2}$$

queremos ver se fazendo igualdade e

variando  $q$  de  $0$  a  $\infty$ ,  $\theta$  variará de  $-1$  a  $1$

$$p(1) = \frac{\theta}{1+\theta^2} = \frac{-1}{q+2}$$

$$\therefore \theta = -\frac{(2+q) \pm \sqrt{q^2 + 4q}}{2} \quad \text{onde } q \in (0, \infty)$$

$$q \rightarrow 0 : \theta = -\frac{2 \pm 0}{2} = -1$$

$$q \rightarrow \infty : \text{Truque algébrico: } x \in \mathbb{R} \text{ por } \sqrt{x}$$

$$\theta = 0$$

Espaço paramétrico. Pf  $\theta$  será restrito:  $-1 < \theta < 0$

↓

e' o que encontramos

quando envermos  $\delta y_t$  como MNL

Se realmente auditarmos que modelo melhora o usado  
por MNL  $\Rightarrow$  sua restrição maior é natural

tilibra

/ /

se supusermos  $E[\eta_t \eta_{t-1}] = \sigma_{\eta\eta}$  (nas sas + descorrelatadas)

$$\Delta y_t = \eta_t + \Delta \varepsilon_t$$

$\Rightarrow E[\Delta y_t] = 0$  — Nas são afetadas pela nova hipótese

$$\text{var}(\Delta y_t) = \sigma_\eta^2 + 2\sigma_\varepsilon^2 + 2\sigma_{\eta\varepsilon}$$



$$\begin{aligned} E[(\Delta y_t - E[\Delta y_t])^2] &= E[\Delta y_t^2] = \\ &= E[(\eta_t + \varepsilon_t - \varepsilon_{t-1})^2] \\ &= E[(\eta_t^2 + 2\eta_t(\varepsilon_t - \varepsilon_{t-1}) + (\varepsilon_t - \varepsilon_{t-1})^2)] \\ &= \sigma_\eta^2 + 2\sigma_{\eta\varepsilon} + 2\sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} r(1) &= E[\Delta y_t \Delta y_{t-1}] \\ &= E[(\eta_t + \varepsilon_t + \varepsilon_{t-1})(\eta_{t-1} + \varepsilon_{t-1} - \varepsilon_{t-2})] \\ &= -\sigma_\varepsilon^2 \quad \text{— Nas são afetadas também.} \end{aligned}$$

$$\text{Logo } r(1) = \frac{-\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_{\eta\varepsilon} + 2\sigma_\varepsilon^2} = \theta$$



na  $\Delta y_t$  temos

3 parâmetros:

$$\sigma_\eta^2, \sigma_\varepsilon^2, \sigma_{\eta\varepsilon}$$

na  $\Delta y_t$  temos 2 parâmetros:

$$\theta \text{ e } \sigma_\varepsilon^2$$

Este modelo de covariância nas é identificável.

↓ diz respeito a  $\sqrt{\text{grau de liberdade}}$  parâmetros

(n é a mesma coisa da observabilidade → que tem a ver com as componentes)

Nas faixas muito uso.

### Forma UCARIMA

$$\Delta y_t = \eta_t + \Delta \varepsilon_t$$

$$\therefore y_t = \frac{\eta_t}{\Delta} + \varepsilon_t$$

chama de tendência

$$= T_t + \varepsilon_t$$

$\therefore T_t = \eta_t$   $\therefore \Delta T_t = \eta_t$   $\because T_t = T_{t-1} + \eta_t \Rightarrow$  componente  
de passo abertos

Notas 13

### Modelo de Tendência linear local

$$y_t = \mu_t + \varepsilon_t \sim \sigma^2_{\varepsilon}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \sim \sigma^2_{\eta}$$

$$\beta_t = \beta_{t-1} + \gamma_t \sim \sigma^2_{\gamma}$$

Definir 2 razõesinal ruído:  $q_{\eta} = \frac{\sigma^2_{\eta}}{\sigma^2_{\varepsilon}}$ ,  $q_{\gamma} = \frac{\sigma^2_{\gamma}}{\sigma^2_{\varepsilon}}$

$\mu_t \Rightarrow$  tendência estocástica da série

$\mu_t$ : é uma tendência linear estocástica  
(+ complicada do que um RW)

é modificada por 2 choques: um na própria tendência  
e outro na inclinação.

$$\text{Se } \sigma^2_{\eta} = \sigma^2_{\gamma} = 0 \Rightarrow y_t = a + bt + \varepsilon_t$$

tilibra

(Importante: Olhar  $\mu_t$   
casos particulares)

/ /

⇒ Ver casos particulares na apostila.

ex: 1) Se fizermos estimativas e definirmos a  $\sigma_e^2 = 10^{-5}$  (muito baixa)

⇒ pode desconsiderar estocasticidade da inclinação

2) Se  $\sigma_\beta^2 \neq 0$  e  $\sigma_\eta^2 = 0$ :

$$y_t = \mu_t + \epsilon_t$$

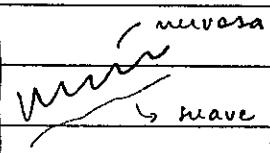
$$\mu_t = \mu_{t-1} + \beta_{t-1}$$

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

1ª diferença de  $\mu_t$ :  $\Delta \mu_t = \beta_{t-1}$

2ª " " :  $\Delta^2 \mu_t = \Delta \beta_{t-1} = \xi_{t-1}$

O processo terá uma tendência suave.



⇒ 2ª derivada da componente de  
tendência é um ruído branco.

$$\frac{d^2 \mu_t}{dt^2} = \xi_t \Rightarrow \mu_t = \int \int \xi_{t-1} dt$$

→ operações de soma no ruído branco.  
Passa a suavizá-lo.

Elimina componente de alta frequênci

Este modelo de tendência suave se comunica com filtro HP

(Modelo para separar tendência

de ciclos de sínus macroeconómico)

FORMA de Espaço de Estado:

↓

pt modelos estruturais q rea' invariante no tempo

1 /

$$y_t = z\alpha_t + \beta_t = (1 \ 0) \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \epsilon_t$$

$$\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_T \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_R \begin{pmatrix} \eta_t \\ \tau_t \end{pmatrix}$$

$$\alpha_t = mx1 \quad mxm \quad (mxn) \quad nx1$$

(Nesse caso  $m = n$ )

Mas pode ser que tenhamos  
menos choques do que compõem  
os.

Funções de previsões:

↓

Nas tem nada a ver com  $F^K$

Vêm sempre a partir da equações dos modelos

Dois tipos de fazer:

$$I) \quad y_t = \mu_t + \epsilon_t, \quad t=1, \dots, n$$

$$II) \quad y_t = z_t \alpha_t + \epsilon_t$$

$$y_{t+k} = \mu_{t+k} + \epsilon_{t+k}$$

$$\hat{y}_{t+k|t} = z E[\alpha_{t+k} | y_t]$$

$$\hat{y}_{t+k|t} = E[y_{t+k} | y_t]$$

$$= E[\mu_{t+k} | y_t] + E[\epsilon_{t+k} | y_t]$$

$$\downarrow \\ E[\epsilon_t] = 0 \quad \forall t$$

Temos que  $\alpha_t = T\alpha_{t-1} + \eta_t$

$$\alpha_{t+k} = T^k \alpha_t + (\dots + \eta_t)$$

$$\Rightarrow E[\alpha_{t+k} | y_t] = T^k E[\alpha_t | y_t]$$

Veja de  $F^K$

$$= T^k \begin{pmatrix} \hat{\mu}_{t|t} \\ \hat{\beta}_{t|t} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

tilibra

$$= \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{\mu}_{t|t} \\ \hat{\beta}_{t|t} \end{pmatrix} = \begin{pmatrix} \hat{\mu}_{t|t+k} \\ \hat{\beta}_{t|t+k} \end{pmatrix}$$

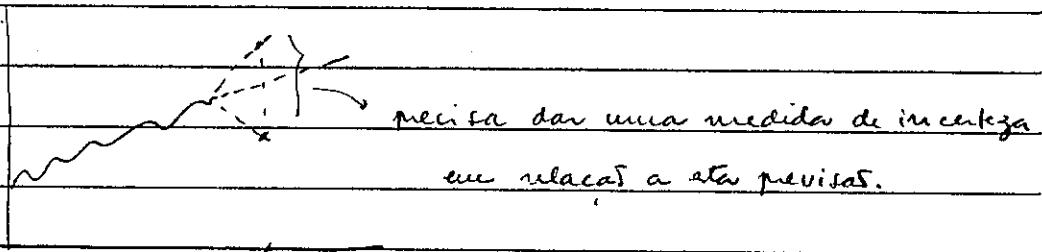
$$\therefore \hat{y}_{t+k} = E[\hat{x}_{t+k} | \hat{y}_t] = (1 \ 0) \begin{pmatrix} \hat{\mu}_{t+k} \\ \hat{\beta}_{t+k} \end{pmatrix}$$

$$= \hat{\mu}_{t+k} + k \hat{\beta}_{t+k}$$

↓

equações de credibilidade linear

Exercício:  $\hat{\mu}_{t+k}$  e  $\hat{\beta}_{t+k}$  dadas por amortecimento exponencial simples.



Obs: Nessa equação, trabalhamos com  $\hat{\Phi}$  (hiperparâmetros fixos conhecidos, estimados)  $\Rightarrow$  plug-in

A incerteza a respeito de  $\hat{\Phi}$  não é transmitida p/ as componentes calculadas com  $F_k$ .

Intuitivamente nos modelos ARIMA:

$$y_t = \phi y_{t-1} + \epsilon_t$$

$$\hat{y}_{t+k|t} = \hat{\phi}^k y_t \rightarrow \text{Já usamos } y_t = \hat{\phi} y_{t-1} + \epsilon_t \text{ naturalmente}$$

MSE da previsão:

$$\text{MSE}(\hat{y}_{t+k|t}) = E[(y_{t+k} - \hat{y}_{t+k|t})^2]$$

Para o modelo:

$$y_{t+2} = \mu_{t+2} + \varepsilon_{t+2}$$

$$\hat{y}_{t+1|t} = \hat{\mu}_{t+1|t} + s\hat{\beta}_{t+1|t}$$

$$\Rightarrow \text{MSE}(\hat{y}_{t+1|t}) = E[(y_{t+2} - \hat{y}_{t+1|t})^2] =$$

$$= E[((\mu_{t+2} - \hat{\mu}_{t+1|t}) + \varepsilon_{t+2} + s\hat{\beta}_{t+1|t})^2 | Y_t]$$

matricialmente, temos uma expressão mais simples:

(

ver pag. 20 (notas III)

Fazer cálculos.

Ver ex. pag. 21: duas previsões feitas a partir de 3 modelos atocintado  
"determinística"

### Forma reduzida

objetivo: obter modelo estacionário e ver se pode ser  
escrito como modelo ARIMA

E ver qual " " pode representar o modelo  
estrutural.

Tb pode resolver a questão de modelos não identificáveis.

Toma 1ª dif:  $\Delta y_t = \Delta \mu_t + \Delta \varepsilon_t$

e  $\Delta \mu_t = \beta_{t-1} + \eta_t$  — Nas é estacionárias

pois  $\beta_t$  é um RW

Toma 2ª dif:

tilibra

/ /

$$\Delta^2 y_{t-2} = \Delta \beta_{t-1} + \Delta \eta_t + \Delta^2 \varepsilon_t$$

$$\therefore \Delta^2 y_t = \gamma_{t-1} + \Delta \eta_t + \Delta^2 \varepsilon_t \Rightarrow \text{aqui chegamos a um processo}$$

só termos RB do lado direito

Agora: precisa identificar o modelo ARIMA ao qual se relaciona.

Ver F.A.C's.

Tiraremos  $E(\cdot)$ ,  $\text{var}(\cdot)$  e  $\gamma(k)$

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

$\hookrightarrow (1-L)^2 \varepsilon_t = (1-2L+L^2) \varepsilon_t$

$$\rightarrow E(w_t) = 0$$

$$\begin{aligned} \rightarrow \text{var}(w_t) &= \gamma(0) = \text{var}[\gamma_{t-1} + \eta_t - \eta_{t-1} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}] \\ &= \sigma_\eta^2 + 2\sigma_\eta^2 + 6\sigma_\varepsilon^2 \quad (\text{ver cálculos}) \end{aligned}$$

$$\rightarrow \text{cov}(w_t, w_{t-k}) = \begin{cases} -\sigma_\eta^2 - 4\sigma_\varepsilon^2, & k=1 \\ \sigma_\varepsilon^2, & k=2 \\ 0, & k \geq 3 \end{cases} \Rightarrow \text{MA}(2)$$

-TLL  $\Rightarrow$  MA(2)  $\Rightarrow$  3 hiperparâmetros  $\therefore \Psi^{\text{TLL}} = (\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\varepsilon^2)$

$$z_t = \theta_1 w_{t-1} + \theta_2 w_{t-2} + w_t \quad w_t \sim N(0, \sigma_w^2) : \Psi^2 = (\theta_1, \theta_2, \sigma_w^2)$$

3 hiperparâmetros  $\Rightarrow$  precisamos de 3 equações.

variancia,  $\gamma(1) \in \gamma(2)$

$$E[w_t^2] = E[z_t^2]$$

$$E[w_t w_{t-1}] = E[z_t z_{t-1}]$$

$$E[w_t w_{t-2}] = E[z_t z_{t-2}]$$

↑  
forma reduzida  
do TLL

↑  
MA(2)

filadéia

Restrições sobre  $\theta_1$  e  $\theta_2$  para MA(2) inviável

→ obter  $w_t$  em função de  $z_t$

$$z_t = (1 + \theta_1 L + \theta_2 L^2) w_t$$

$$\Rightarrow \Theta(L) = 1 + \theta_1 L + \theta_2 L^2 \equiv 0$$

$$(\lambda_1, \lambda_2) \mid |\lambda_i| > 1 \quad \forall i=1,2$$

↓

Essas restrições podem ser mapeadas em  $\theta_1$  e  $\theta_2$

$$\text{entrevendo } (L - \lambda_1)(L - \lambda_2)$$

e veremos que  $\theta_1$  e  $\theta_2$  têm que estar numa certa região.

Ver pg 25.

(Na realidade, entrevendo

$$f(v) = v^2 + \theta_1 v + \theta_2 v \quad \text{e } |V_i| < 1$$

$\theta_1$  e  $\theta_2$  não vão poder estar em toda a região se forem compatíveis com MLL.

Ver área hachurada na pg. 25.

Se acharmos  $\theta_1$  e  $\theta_2$  para dessa região → não vai ser

índice aproximado

por MLL.

Ver pg 27 ⇒ outros modelos de tendência

— 11 —

Sazonalidade → Notas IV de aula.

(eventos periódicos

período + ou - regular.

precisa vir dentro de 1 ano.

equilíbrio

Diferença sazonalidade e ciclo: no período



dentro de 1 ano  
por questões climáticas,  
convenções sociais/ adm

periodicidade não é tão

bem definida. Ex: ciclos  
económicos.

(Tanto questões do período quanto  
quanto de amplitude)

PF questões naturais e mais bem  
definido.

$\Rightarrow$  o que é sazonal é esperado.

temos que tirar da série as variações esperadas por conta da  
sazonalidade pf avaliar tendência.

Não há maneira única de decompor série em tend. e sazonalidade.

Se vir série desazonalizada  $\Rightarrow$  precisa saber qual foi o método  
de desazonalização.

Pode saber se série foi pouca desazonalizada  $\Rightarrow$  Ver pela FAC se

ainda há ciclo em k=12

$$y_t^{(a)} = y_t - \delta_t$$

$y_t^{(a)}$  não é a tendência  $\Rightarrow$  há o erro ainda.

$$y_t = \mu_t + \delta_t + \varepsilon_t \Rightarrow y_t - \delta_t = \mu_t + \varepsilon_t$$

$$\underbrace{y_t^{(a)}}_{\downarrow} \quad \left\{ \begin{array}{l} \\ \end{array} \right.$$

é mais suave.

(picos da sazonalidade são retirados da série)

$\hookrightarrow$  pg 3 das Notas IV

Dois tipos de sazonalidade: aditiva e multiplicativa

$$y_t = \mu_t + \delta_t + \epsilon_t$$

↓      ↓  
sazonalidade  
tendência

→ Veremos principais formas de estimar a sazonalidade

→ Primeiros tratarmos no contexto determinístico.

Sazonalidade determinist. sugiria dar esto cástica qd alguma  $\sigma^2$  por a zero.

Três modelos:

Obs: No caso de ST, ARIMA(1,2)

$$y_t = \phi y_{t-1} + \phi_{12} y_{t-12} + \epsilon_t$$

{

sazonalidade tratada internamente

Não dá p/ filtrar a sazonalidade  
(não usarmos)

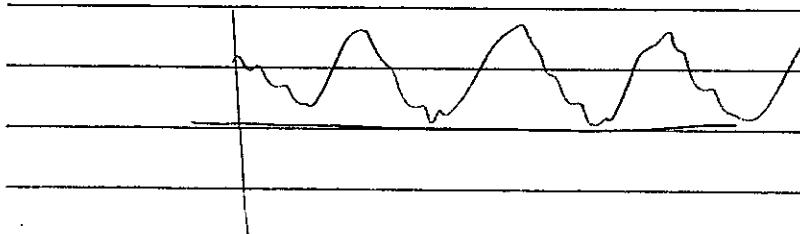
i) dummy

ST trimestral

1 dummy p/ cada trimestre

p intercepto

4 variáveis explicativas



/ /

	$t$	$D_{1t}$	$D_{2t}$	$D_{3t}$	$D_{4t}$
2010	Q1	1	0	0	0
	Q2	0	1	0	0
	Q3	0	0	1	0
	Q4	0	0	0	1
2011	Q1		etc.		
	Q2				
	Q3				
	Q4				

Modelo poderia ser trivialmente estimado por MLE

$$y_t = \beta + \sum_{j=1}^4 \gamma_j D_{jt} + \epsilon_t$$

problema  $\Rightarrow$  não podemos colocar tantas dummys quanto os períodos da série, não conseguiremos resolver o problema por MLE por conta de multicolinearidade perfeita

$$(x'x) \Rightarrow \hat{\beta} = (x'x)^{-1}x'y$$

$\checkmark$   
Nas reais inversível

Ver parametrizações possíveis

(i), (ii), (iii)

$\rightarrow$  Usaremos parametrização (iii)

tilibra

pg 19.



De forma geral, teremos:

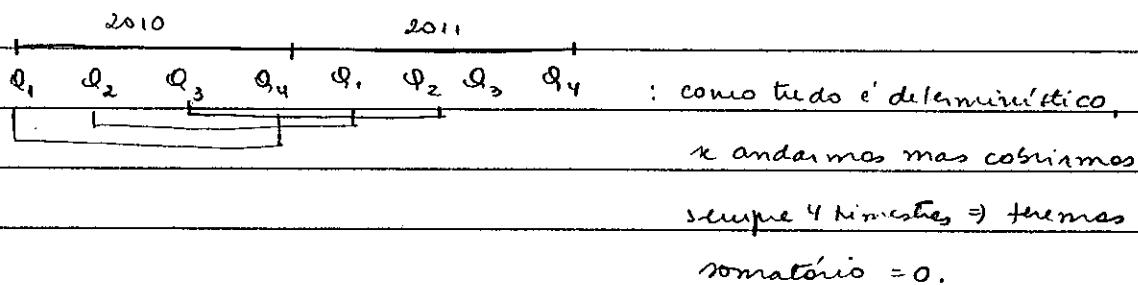
$$\begin{cases} y_t = \mu_t + \gamma_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \gamma_t \\ \gamma_t = ? \end{cases}$$

↳ como explicitaremos?

Nas  
seus  
notas

### Ponto de Partida

$$\sum_{j=1}^s \gamma_j = 0 \quad (\text{da parametrização iii})$$



s=4

$$\Leftrightarrow \sum_{j=0}^{s-1} \gamma_{t-j} = 0$$

$$\text{pois } \gamma_t + \gamma_{t-1} + \gamma_{t-2} + \gamma_{t-3} = 0$$

Qualquer que seja  $t$ : cobrimos o período.

ds:  $\gamma | \Theta \Rightarrow$  coeficientes que x dígitos razoáveis (nos ex.)

(nas fac. a ver com Y do ME).

Faremos:  $\sum_{j=0}^{s-1} \gamma_{t-j} = w_t$ ,  $w_t \sim N(0, \sigma_w^2)$   
(flutuações stocásticas)

$$(obs: E[\sum \gamma_{t-j}] = 0)$$

$$\delta_t = - \sum_{j=1}^{s-1} \delta_{t-j} + w_t$$

— Dummies ficam implícitas

Nos termos mais  $D_{1t}, D_{2t}, \dots$

Fazemos apenas  $\delta_t$  correspondente

ao período em que estamos trabalhando.

$$\Rightarrow \begin{cases} y_t = \mu_t + \delta_t + \epsilon_t \\ \mu_{t+1} = \mu_t + \gamma_t \\ \delta_t = - \sum_{j=1}^{s-1} \delta_{t-j} + w_t \end{cases}$$

<u>t</u>	determinado trimestre	<u>t+1</u>	(dados trimestrais)
Q1 2010 (p. ex.)		⇒	Q2 2010

$$\begin{array}{lcl} \delta_{1,t} & = & \delta_{2,t+1} \\ \delta_{2,t} & = & \delta_{3,t+1} \\ \delta_{3,t} & = & \delta_{4,t+1} \\ \delta_{4,t} & = & \delta_{1,t+1} \end{array} \quad \left. \begin{array}{l} \text{rearranjo dos fatores} \\ \text{sazonais no trimestre} \end{array} \right\}$$

• Como  $\delta_{1,t+1} = \delta_{4,t}$  e  $\delta_{4,t} + \delta_{3,t} + \delta_{2,t} + \delta_{1,t} = 0$   
 $\Rightarrow \delta_{4,t} = -(\delta_{3,t} + \delta_{2,t} + \delta_{1,t})$

$$\Rightarrow \delta_{1,t+1} = -\delta_{4,t} = -\delta_{1,t} + \delta_{2,t} - \delta_{3,t} + w_t$$

Coloca drogue  $w_t$

•  $\delta_{2,t+1} = \delta_{1,t}$  m. estatístico.

$$\delta_{3,t+1} = \delta_{2,t}$$

Tratando apenas a sazonalidade em  $y_t$ :

$$y_t = (1 \ 0 \ 0) \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} \delta_{1,t+1} \\ \delta_{2,t+1} \\ \delta_{3,t+1} \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix} + \begin{pmatrix} w_t \\ 0 \\ 0 \end{pmatrix}$$

(8)

No modelo original:

$$Y_t = \sum_{j=1}^{s-1} Y_{t-j} + w_t \Rightarrow \delta_t = -\delta_{t-1} - \delta_{t-2} - \delta_{t-3} + w_t$$

$$\delta_{t+1} = -\delta_t - \delta_{t-1} - \delta_{t-2} + w_t$$

$$\Rightarrow y_t = \delta_t + \varepsilon_t$$

$$\delta_{t+1} = -\delta_t - \delta_{t-1} - \delta_{t-2} + w_t$$

$$\Rightarrow \begin{pmatrix} \delta_{t+1} \\ \delta_t \\ \delta_{t-1} \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \delta_t \\ \delta_{t-1} \\ \delta_{t-2} \end{pmatrix} + \begin{pmatrix} w_t \\ 0 \\ 0 \end{pmatrix}$$

(9)

### Última parametrização

Esta é a forma + simples de incluir sazonalidade no ME.

obs: Se considerarmos tendência  $\Rightarrow$  matriz T é diagonal em blocos.

Não haverá influência das comp. de tendência na sazonalidade e vice-versa.

→ Sazonalidade por funções trigonométricas

Fator sazonal em t reia dado por

$$Y_t = \sum_{j=1}^{s_2} (r_j \cos(\lambda_j t) + r_j^* \sin(\lambda_j t))$$

Exemplo:

$$\left\{ \begin{array}{l} Y_t = Y_t + \epsilon_t \\ Y_t = \sum_{j=1}^s [r_j \cos(\lambda_j t) + r_j^* \sin(\lambda_j t)] \end{array} \right. \quad s=4: \text{ST trimestral}$$

$$\lambda_j = \frac{2\pi j}{s} = \frac{2\pi j}{4} = \frac{\pi}{2} j, \quad j=1, 2$$

Abrindo expressas para  $\gamma_t$ :

$$Y_t = [r_1 \cos(\lambda_1 t) + r_1^* \sin(\lambda_1 t)] + [r_2 \cos(\lambda_2 t) + r_2^* \cancel{\sin(\lambda_2 t)}]$$

Para  $j=2 \Rightarrow \pi j = \pi \therefore \sin(\lambda_2 t) = \sin(\pi t) = 0 \quad \forall t$

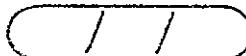
$$\Rightarrow Y_t = \underbrace{r_1 \cos(\lambda_1 t)}_{x_{1t}} + \underbrace{r_1^* \sin(\lambda_1 t)}_{x_{2t}} + \underbrace{r_2 \cos(\lambda_2 t)}_{x_{3t}} + \epsilon_t$$

Precisaremos estimar 3 parâmetros (série trimestral, associados à sazonalidade)

$r_1, r_1^* \in r_2$

como fizemos no caso determinístico).

(Nas é coincidência).



Atenções!  $\hat{\gamma}_1$ ,  $\hat{\gamma}_1^*$  e  $\hat{\gamma}_2$  são os fatores racionais.

Eles são dados por:

$$\hat{\gamma}_t = \underbrace{\hat{\gamma}_1 \cos \lambda_1 t}_{\pi/2 t} + \underbrace{\hat{\gamma}_1^* \sin \lambda_1 t}_{\pi/2 t} + \underbrace{\hat{\gamma}_2 \cos \lambda_2 t}_{\pi t}$$

Para  $Q_1 \rightarrow t=1$

$Q_2 \rightarrow t=2$

$Q_3 \rightarrow t=3$

$Q_4 \rightarrow t=4$

$Q_5 \rightarrow t=5$  daí a volta.

Se fizermos  $t=1 \Rightarrow \hat{\gamma}_1$  será o fator do 1º tri

(e numericamente será = ao calculado

no modelo determinístico de param. iii)

Obs: No somatório

Se  $S=12 \Rightarrow$  somatório terá 6 cos e 6 sen

Para  $j=1$ : 1º harmônico (fundamental)

Para  $j=2$ : 2º harmônico

:

Na física  $\Rightarrow$  série de temperatura é a + comportada em relações à razionalidade.

Neste caso, não será necessário usar todos os harmônicos. Os 1º ditarão o comportamento.

$\hat{\gamma}_1^{+0}$  + harmônicos usamos  $\Rightarrow$  mais esquinhos serão os comportamentos racionais

(j harmônicos + altas n̄ forem muito importantes

$\Rightarrow$  mas serão estatisticamente

significantes).

mas não ser abandonados.

tilibra

Questão: como escrevemos este modelo na forma EE?

pg 26

$$y_t = \mu_t + \delta_t + \epsilon_t$$

$$\delta_t = \sum_{j=1}^{12} \delta_{jt} \quad \text{onde } \delta_{jt} = \delta_j \cos(\lambda_j t) + \delta_j^* \sin(\lambda_j t)$$



Podemos escrever  $\delta_{jt+1}$  e obter como funções de  $\delta_{jt}$   
(ver apostila).

— II —

Modelo Aréline

✓  
Sazonalidade multiplicativa

(amplitude aumentando no tempo)

→ Pacote STAMP

Formulate a model

Select components

Level:  $\mu_t$       Se ambos esto estatísticos  $\Rightarrow$  MLE

Slope:  $\rho_t$   
 $\Rightarrow$  Slope  $\Rightarrow$  MNL

Seasonal: (n aparece opção de dummy / trigon.)

Estimation model: Maximum likelihood

No resultado  $\Rightarrow$  Summary statistics: do resíduos

Estatística

Normality  $\Rightarrow$  teste Jarque-Bera.

H: rejeita hipótese

de autocorrelação dos resíduos

de homocedasticidade.

$R^2 = 0.2$

Variâncias de扰动:

$\sigma_e^2 \Rightarrow$  pequena em relação aos outros pois os 3 componentes explicam praticamente tudo

$\sigma_{\eta}^2$

$\sigma_k^2$

$\sigma_w^2$

— " —

obs: M extracções de componentes

$\Rightarrow$  faz sempre smoothing.

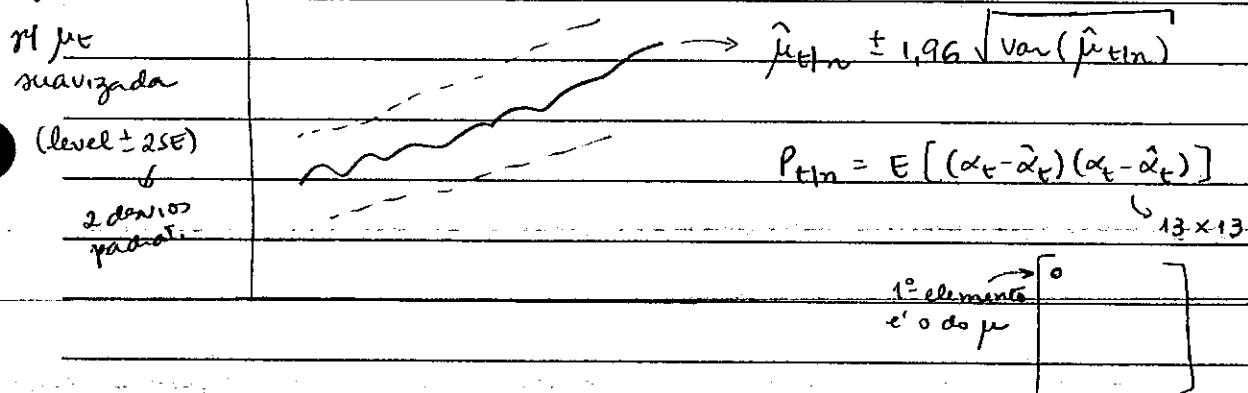
Ruído não está legal: crescendo e muita autocorrelação

(modelo mal especificado)

— " —

Olhando praticamente

gráfico



Temos gráfico  $\neq \beta_t$

Temos "  $\neq$  irregular  $\Rightarrow$  praticamente zero pois suavizada "rouba" praticamente toda info.

" " " " sazonalidade

— " —

Reestimar o modelo  $\Rightarrow \log(\text{airline})$

• Olhando summary statistics  $\Rightarrow H$  cai

• Olhando distorções (variações)

level praticamente fixo.

seasonal " determinístico

Pede Anti-log analysis

$$\text{Fez } \log y_t = \mu_t + \gamma_t + \epsilon_t$$

Depois que estimamos,

$$\log y_t = \hat{\mu}_t + \hat{\gamma}_t$$

$$\Rightarrow y_t = e^{\hat{\mu}_t} e^{\hat{\gamma}_t}$$

$$\mu_t \quad \gamma_t'$$

Resultados:

$$\cdot \text{resíduos padronizado: } \hat{v}_t^* = \frac{y_t - \hat{y}_{t+1}}{(V)} \quad (\text{outra original da série})$$

( $F_t$ )

tilibra

25/05/H

Aula passada:

- Sazonalidade
  - (estocástica)
  - determinística)
    - caso particular da estocástica
- dummies
- trigonométricas

Notas  
pg 3.

MEB: modelo Estacional Básico

$$y_t = \mu_t + \gamma_t + \epsilon_t$$

duas componentes padrão (tendência + sazonalidade)

$\mu_t$ : tend. linear estoc.

$\gamma_t$ : sazonal estocástica

Forma reduzida de um MEB:

↳ é identificável.

Muito conhecido.

mas qual o modelo SARIMA equivalente?

↳ tem que ser SARIMA (por conta da sazon.)

SARIMA  $(p, d, q) \times (P, D, Q)$ ?

operador de diferenciações sazonal

$\Delta_s \cong I - L^{\circledcirc}$  — fct. da sazonalidade e' sempre confluente  
(trimestral, mensal etc)

$$(\neq \Delta = I - L)$$

$\Delta_s \cdot y_t = y_t - y_{t-s}$  → se sazonalidade fosse determinística, esta operação tiraria totalmente a tábua sazonalidade.

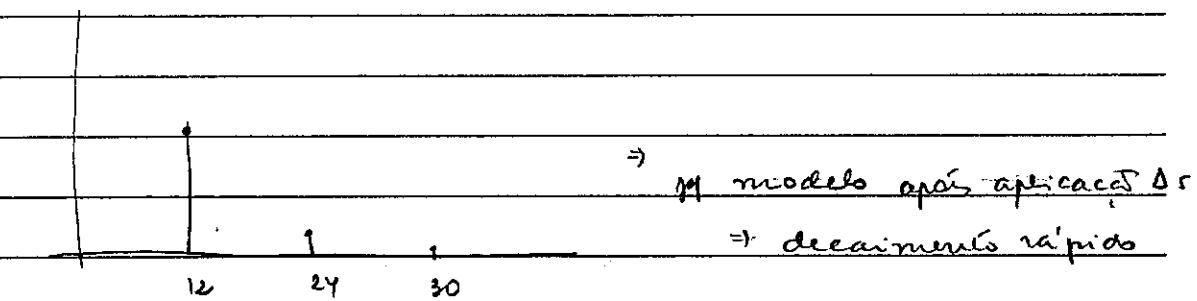
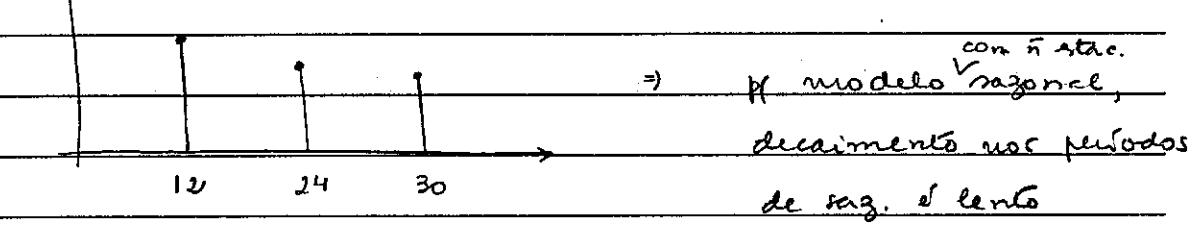
Só se deve operar a saz. estocástica numa série já diferenciada em relação à tendência.

$\Delta$  = retira a estacionariedade na tend.

$\Delta_s$  = " " " na sazonalidade

Para a séie  $y_t$

$n(t)$



É import. sazonalidade estacionária no contexto Bt

No contexto ME. não é necessário (nem pf tend.) mas

vários precisam verificar pf outra forma reduzida.

Precisamos definir tb:

$$S_s(L) \hat{=} 1 + L + L^2 + \dots + L^{s-1} \rightarrow \text{operador: vamos precisar dele}$$

$$\Delta S_s(L) = (1 - L) * ( " ) = \Delta_s = (1 - L^s)$$

Vai tirando diferenças.

Quando para?  $\Rightarrow$  quando obtivermos do lado direito apenas funções das erros.

PG<sup>4</sup>

② Operadores  $S(L)$  e  $\Delta$  comutam.

Nas expressões:

$$\underbrace{S(L) \Delta \Delta y_t}_{} = \Delta_s \Delta y_t$$

$$\Delta S(L)$$

$$\Delta_s \Delta y_t$$

O termo diferenciado da sazonalidade será:

$$\rightarrow \text{Como } S(L) \gamma_t = w_t \Rightarrow \text{Divida}$$

$$\Rightarrow \Delta^2 S(L) \gamma_t = \Delta^2 w_t$$

Recapitulando até aqui:

$\rightarrow$  Partida: MEB

$\rightarrow$  Chegada: forma reduzida

$$y_t = \mu_t + \delta_t + \varepsilon_t$$

$$\Delta \Delta s y_t = \Delta_s \gamma_t + \Delta \Delta_s \varepsilon_t + \Delta^2 w_t + S(L) \gamma_t$$

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

(

apenas uma equação

$$\text{e' um } \rightarrow \beta_t = \beta_{t-1} + \delta_t$$

(equivalente a modelo ARIMA)

$$\text{máx } S(L) \gamma_t = w_t$$

e' componente  
(não aparece  
em  $y_t$ )

$$\gamma_t = ? \quad (\text{sairia correspondente})$$

$$E[\gamma_t] = 0$$

$$\gamma(0) = E[\gamma_t^2] = -$$

$$\gamma(j) = E[\gamma_t \gamma_{t-j}] = -$$

ilibra

Precita desenvolver tudo:

$$z_t = \eta_t - \eta_{t-s} + (1-L)(\varepsilon_t - \varepsilon_{t-s}) + (1-L)^2 w_t + (1+L+\dots+L^{s-1}) \{_{t-1}$$

: ver resultados na pg 4

— //

Obs: Quando modelamos por B&T, raramente modelamos por um modelo MA(s+1).

A classe + comum de modelos usados é SARIMA multiplicativo

No caso que vimos  $\Rightarrow$  modelos de saz. aditiva

~~Observando as autocorrelações:~~

termos que correspondem a relações de CP e a relações sazonais (e as intermédias)

Se, por ex., tivéssemos  $s=4$ :

$$\Delta \Delta_s y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} + \theta_5 \varepsilon_{t-5}$$

MA(s+1)  $\rightarrow$  saz. aditiva

Mas raramente usamos este modelo dito de saz. aditiva.

Mais comum usar SARIMA multiplicativo:

$$\Delta \Delta_s y_t = \left[ \frac{\Theta_q(L^s)}{\Phi_p(L^s)} \right] \eta_t$$

$\hookrightarrow$  caracteriza dep. sazonal

$$\begin{aligned} \Theta_q(L) &= 1 + \Theta_1 L^s + \Theta_2 L^{2s} + \dots + \Theta_q L^{qs} \\ \Phi_p(L) &= 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps} \end{aligned} \quad \left. \begin{array}{l} \text{ajust. dep. sazonal,} \\ \text{mas } \overbrace{\text{à relações de CP}}^{\text{dependência}} \end{array} \right\}$$

11

$\Rightarrow \eta_t$  ainda não será RB.

Preciso ajustar tb.

$$\varphi(L) \eta_t = \theta(L) \varepsilon_t \Rightarrow \eta_t = \frac{\theta_q(L)}{\varphi_p(L)} \varepsilon_t \quad (\text{agora sim } \eta_t \text{ é RB})$$

$$\Rightarrow \Delta \Delta_s y_t = \frac{\Theta_q(L^s)}{\Phi_p(L^s)} \eta_t$$

$$\Rightarrow \boxed{\Phi_p(L^s) \varphi_p(L) \Delta \Delta_s y_t = \Theta_q(L^s) \theta_q(L) \varepsilon_t} \Rightarrow \text{SARIMA}$$

$$\text{Exemplo: } d = 0, 1, 2 \Rightarrow p = q = 0$$

$$D = 0, 1, 2 \quad d = D = 1$$

$$q = Q = 1$$

fazendo  $s = 4$

nos temos  
parte AR

$$\begin{aligned} \Delta \Delta_s y_t &= (1 + \Theta_1 L^4) (1 + \theta_1 L) \varepsilon_t \\ &= (1 + \Theta_1 L^4) (\varepsilon_t + \theta_1 \varepsilon_{t-1}) \\ &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-4} + \Theta_1 \theta_1 \varepsilon_{t-5} \end{aligned}$$

✓

apenas  $\theta_1$  e  $\Theta_1$  como parâmetros

(aparece termo multiplicativo  $\Theta_1 \theta_1$ , das o nome

SARIMA multiplicativo).

• FACP

obs: Olha FACP. Pode ver diretamente MA ou AR.

Se não  $\Rightarrow$  mai ACPMA

usa critérios de identificação (AKAIKE, BIC etc.)

Aqui será a mesma coisa, mas teremos que definir nos

to' p, q, mas tb p, q.

tilibra

Olkaremos AIC, BIC  $\Rightarrow$  encontra o melhor ajuste

$\Rightarrow$  Modelos multiplicativos são + parcimoniosos do que aditivos.  
Neste deve-se olhar aditivos e calcular AIC e BIC M  
escolher melhor modelo.

No exemplo anterior

SARIMA:  $(0,1,1) \times (0,1,1)$   $\Rightarrow$  modelo Airline

(é um ponto de partida

M modelagem de ST

no anexo de B&J)

Airline  $\Rightarrow$  modelo multiplicativo da parte sazonal e

parte n̄ sazonal:  $\Phi_p \times \varphi_p$  e  $\Theta_q \cdot \theta_q$

Voltando às notas de aula: (pg 5)

Para modelo Airline (SARIMA  $(0,1,1) \times (0,1,1)$ )

$$E[z_t] = 0$$

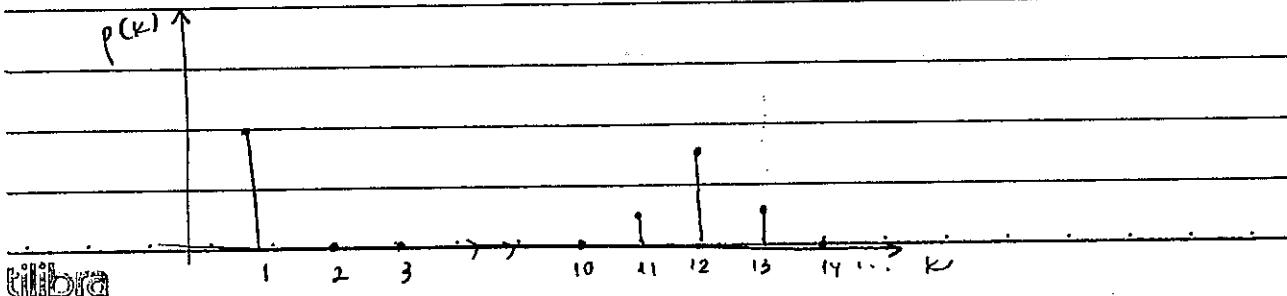
$$\gamma(0) = E[z_t^2]$$

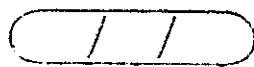
$$\gamma(k) = E[z_t z_{t-k}] \Rightarrow \gamma(k) = 0 \quad \forall k=2 \text{ a } S-2$$

$$\gamma(S-1) = \theta \Theta \sigma^2$$

$$\gamma(S) = \Theta (1+\theta^2) \sigma^2$$

$$\gamma(S+1) = :$$





Que tipo de restrições básicas devemos impor ao MEB para se comunicar com AIRLINE?

forma reduzida do MEB ( $\sigma_\eta^2 = \sigma_\epsilon^2 = 0$ )



inclusão razoabilidade  
cas

\*\* pg 5:

$$\text{MEB com } \sigma_\eta^2 = 0 \text{ e } \sigma_\epsilon^2 = 0 \Rightarrow \begin{cases} w_t = 0 \quad \forall t \\ \xi_t = 0 \quad \forall t \end{cases}$$

(tomar eq. reduzida geral e fazer  $\sigma_\eta^2 = 0$  e  $\sigma_\epsilon^2 = 0$ ):

$$\Delta \Delta_s y_t = \Delta_s \eta_t + \Delta_s \Delta \epsilon_t \\ = \Delta_s (\eta_t + \Delta \epsilon_t)$$

$$\text{MA}(1) \Rightarrow z_t = \eta_t + \Delta \epsilon_t = \eta_t + \epsilon_t - \epsilon_{t-1}$$

$\sim \text{MA}(1)$   $\Rightarrow$  prova usando F.A.C

$$= \theta \epsilon_{t-1} + \epsilon_t$$

$$= \Delta_s (1 + \theta L) \epsilon_t$$

$$\Delta \Delta_s y_t = (1 - L^s) (1 + \theta L) \epsilon_t$$

$$\text{P/ o Airline, temos: } \Delta \Delta_s y_t = (1 + \theta_1 L) (1 + \theta_2 L^s) \epsilon_t$$

A correspondência se dá quando  $\theta_2 = -1$



Modelo não é inversível na parte razoável.

## modos p/ ciclo (pg +)

$$y_t = \mu_t + \psi_t + \epsilon_t$$

\ /                           : decomposições  
? ?                           mas é única.

componentes competem p/ explicar  $y_t$

Há modelos em que choque de  $\mu_t$  e  $\psi_t$

são correlados ou não correlados



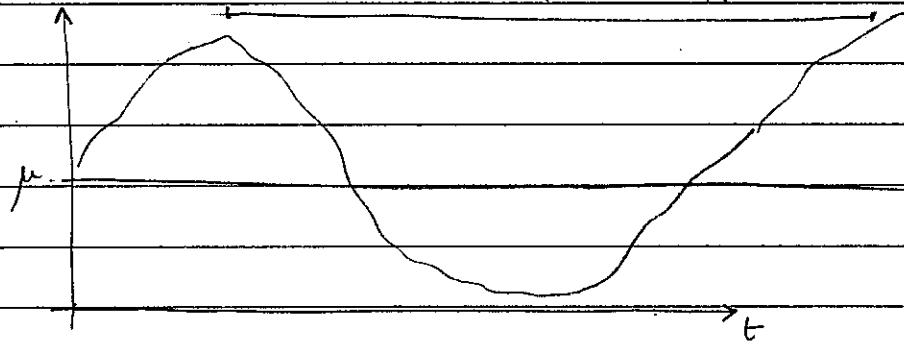
se for, geralmente não é identificável

(verificam se for perfeitamente

(D) vida) → correlado +?

Versão determinística.

$T_c$ : período (p/ o ciclo x repetit).



$$y_t = \mu + \psi_t + \epsilon_t$$



Na razionalidade, temos + harmonônicas.

No ciclo, p/ cada ciclo seu particular → 1 frequência

Seja:  $\psi_t = \sum_{j=1}^{s/2} [\delta_j \cos(\lambda_j t) + \delta_j^* \sin(\lambda_j t)]$

fácil de ser estimado pois

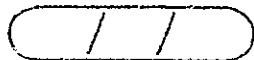
$s$  é conhecido.

$$\lambda_j = \frac{2\pi j}{s}$$

Transforma:

$\cos(\lambda_j t)$  e  $\sin(\lambda_j t)$   
em variáveis explicativas  
⇒ estimar por M&O

tilbra



Na questão do ciclo,  $T_c$  é geralmente desconhecido

(não sabemos direito qual o  $T$ , periódico, do ciclo)

mas pode-se ter alguma ideia, pois quando vamos atrás de um já temos uma ideia pré-concebida do que buscamos.

Se formos estimar por MFO:

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

nas reais funções conhecidas no tempo.

$$y_t = \mu + \underbrace{\alpha \cos \lambda_c t}_{x_{1t}} + \underbrace{\beta \sin \lambda_c t}_{x_{2t}} + \varepsilon_t \quad t = 1, \dots, T$$

$$\lambda_c = \frac{2\pi}{T}$$

x conhecemos  $T$ , teríamos  
regressão linear trivial y 2  
variáveis explicativas.

$$\text{Por MFO: } S(\mu, \alpha, \beta, T_c) = \sum_{t=1}^T [y_t - (\mu + \cos(2\pi/T)t + \sin(2\pi/T)t)]^2$$

$$\frac{dS}{dT} = 0 \rightarrow \bar{N} \text{ é linear.}$$

Nas poderemos usar MFO.

Teremos MFO Nas lineares

(montamos como MFO e resolvemos como MFO NL)

Depois de obter o modelo, devemos obter  $R^2$ .

Ex: Achamos  $T = 13,4$  anos e  $R^2 = 25\%$   $\Rightarrow$  componente

cíclica só explica 25% da variação (o resto poderia ser completamente aleatório).

tilibra

Nos podemos acuidar no T encontrado.

É uma informação problemática.  
pode ser espúria.

→ Tem sempre que tentar linkar com a área de estudo,  
de trabalho.

O achado estatístico do ciclo deve ter repaldo teórico  
bem embasado.

— II —

Para sazonalidade, fizemos:

$$\left\{ \begin{array}{l} y_t = \mu + \gamma_t + \epsilon_t \\ \gamma_t = \sum_{j=1}^{s_{12}} \gamma_{jt} \end{array} \right.$$

$$\begin{pmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{pmatrix} = \begin{pmatrix} \cos \lambda & \sin \lambda \\ -\sin \lambda & \cos \lambda \end{pmatrix} \begin{pmatrix} \gamma_{j,t-1} \\ \gamma_{j,t+1}^* \end{pmatrix} + \begin{pmatrix} w_t \\ w_t^* \end{pmatrix}$$

síntese  $\sigma_w^2$   
obs: neste modelo,  
variação  $\sigma_w^2$  de todos  
os harmônicos.

Para o ciclo: só temos uma frequência ( $j=1$ )

Nos verá mais  $\gamma_t$  e não  $\psi_t$  pois estamos tratando de ciclo.

$$y_t = \mu + \psi_t + \epsilon_t$$

$$\begin{pmatrix} \psi_t \\ \psi_t^* \end{pmatrix} = \rho \begin{pmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{pmatrix} + \begin{pmatrix} k_t \\ k_t^* \end{pmatrix}$$

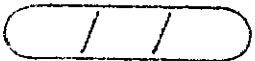
$\sigma_k^2$

$$|\rho| < 1$$

contribui  $\rho$  que  
ciclo fique vivo.

↳ contribui  $\rho$  amortecimento  
ao choque

tilibra



Se tivermos  $\sigma_k^2 = 0$  (determinístico)

Caso particular: modelo de ciclo determinístico.

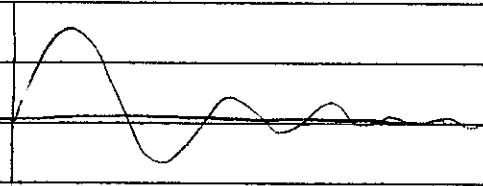
$$y_t = \mu + \psi_t + \epsilon_t$$

$$\psi_t = \alpha \cos \lambda_c t + \beta \sin \lambda_c t$$

Obt:  $p$  - amortecimento

$k_p$  - choque: mantém ciclo vivo

Fora da amostra, mas tem choque. Previsão terá cara:



Como o ciclo é estocástico, faz sentido o comportamento se variar com o tempo.

A estimativa de  $p$  dirá isso. Se  $p$  for próximo de 1, ciclo será + persistente.

Na FEE:

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_t \\ \psi_t^* \end{pmatrix} + \epsilon_t$$

$$\begin{pmatrix} \psi_t \\ \psi_t^* \end{pmatrix} = p \begin{pmatrix} \cos \lambda_c t & \sin \lambda_c t \\ -\sin \lambda_c t & \cos \lambda_c t \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{pmatrix} + \begin{pmatrix} k_t \\ k_t^* \end{pmatrix}$$

$T$

$T = T(p) \quad \lambda_c = 2\pi/T \Rightarrow$  matriz  $T$  tem parâmetros desconhecidos agora.  
 $(p, T)$ .

Observações:

i) Se  $\gamma_0 = 0$  ou  $\pi'$

$$\begin{pmatrix} \psi_t \\ \psi_t^+ \end{pmatrix} = \rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \psi_{t-1}^+ \end{pmatrix} + \begin{pmatrix} k_t \\ k_{t-1}^+ \end{pmatrix},$$

$$\Rightarrow \psi_t = \rho \psi_{t-1} + k_t \Rightarrow \text{processo AR(1)}$$

$|\rho| < 1.$

por que é importante termos um AR(1)

Ex: se modelo  $y_t = \mu + \gamma_t + \epsilon_t$

se tomarmos  $v_t^* = y_t - \hat{y}_{t-1}$  inovações

$V_t$

e F.A.C ainda apresente resíduos



$\Rightarrow$  usar  $\psi_t$  como AR(1) pode ajudar a melhorar o modelo.

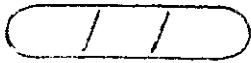
Obs: ~~\*\*~~ pág 10

$\psi_t$ : não é exatamente determinístico.

$$y_t = \mu + (\alpha \cos \omega t + \beta \sin \omega t) + \epsilon_t$$

$$(\alpha, \beta) \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right)$$

$\alpha, \beta$  podem ser fixas e desconhecidas, podem ter multivariadas estocásticas.  
há 7 formas de modelar.



## SOFTWARE

STAMP

Rainbeaz (faz previsão de chuvas em Fortaleza).

Adotando modelo:  $y_t = \mu_t + \psi_t + \epsilon_t$

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

( $\eta$  inclinação)

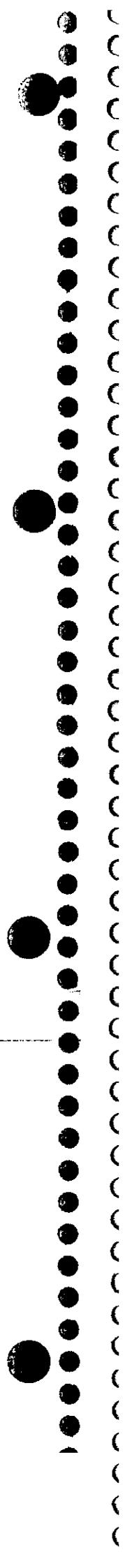
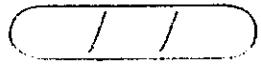
Formulate a model

Select components:

Level (stochastic)

Intrapolar

Cycles (pode ter ciclos de periodicidade  $\neq$ )



**tilibra**

## Variáveis Explanatórias e Intervenções em ME

$$X_t = (x_{1t} \ x_{2t} \ \dots \ x_{kt}) \sim K \times 1$$

$$\gamma = (\gamma_1 \ \gamma_2 \ \dots \ \gamma_K) \sim K \times 1$$

Ex:  $\begin{cases} y_t = \mu_t + \gamma' x_t + \epsilon_t & \epsilon_t \sim N(0, \sigma^2) \\ \mu_{t+1} = \mu_t + \beta_t + \eta_t \\ \beta_{t+1} = \beta_t + \xi_t \end{cases}$

duas possibilidades:

1) Escreve o modelo na forma

depois de  $\gamma_t$ :  $y_t = z_t^* x_t + c_t + \epsilon_t$  onde  $c_t = \gamma' x_t$  por m.v. parâmetros fixos associados a  $x_t$

$$v_t = y_t - \hat{y}_{t|t-1} = y_t - (z_t^* \alpha_t + c_t)$$

previstas | passo a frente

$$v_t = v_t(\gamma, r) \Rightarrow l(\gamma, r) \propto -\frac{1}{2} \sum \ln F_t + v_t^2 | F_t$$

contém  $r$

maximiza em relações a  $r$ .

II) Incorpora  $r$  no vetor de estado

depois do vetor de estado

$$y_t = \mu_t + \gamma' x_t + \epsilon_t$$

Novo caso,

estaremos implicitamente considerando  $r$  invariante no tempo:  $r_{t+1} = r_t$

$$y_t = \begin{pmatrix} 1 & 0 & x_t \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \\ r_t \end{pmatrix} + \epsilon_t$$

$K \times 1$ ; incluímos  $K$  componentes a mais.

Em  $t=T \Rightarrow \gamma_T$ : é isto que ilibra rumos pefas.

$$\begin{pmatrix} \mu_{t+1} \\ \beta_{t+1} \\ \delta_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \\ \delta_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix}$$

$\tilde{\gamma}_{t+1} = \tilde{\gamma}_t \Rightarrow$  não há variações em  $\tilde{\gamma}_t$  (poderia haver!)

Ex:

precisa dizer

$$\tilde{\gamma}_{t+1} = \phi \tilde{\gamma}_t + w_t$$

$$\begin{pmatrix} \delta_{1,t+1} \\ \delta_{2,t+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} \begin{pmatrix} \delta_{1,t} \\ \delta_{2,t} \end{pmatrix} + \begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} \leftarrow \begin{array}{l} \text{como varia no} \\ \text{tempo (ex: de avaliações} \\ \text{do coeficiente, estrutura)} \end{array} \Rightarrow \text{VAR(1), } w_t \sim N(0, Q_w) \downarrow \begin{array}{l} \text{cheia} \\ (\text{covariâncias} \\ \text{no dia-a-dia}) \end{array}$$

Pode-se mostrar que as estimativas de  $\tilde{\gamma}$  obtidas via FKF<sup>atualizada</sup> são equivalentes às de MS Recursivos



$$\tilde{\gamma}_t = \tilde{\gamma}_{t-1} + ( )$$

Diz: se quisermos extrair esse princípio as componentes <sup>no</sup> inerentes a uma série pp buscar explicações econômicas / físicas, mas devemos colocar variáveis explicativas

Se o objetivo  
não for inter-  
pretar, mas

concorrência com as componentes,  
que nos quer + as mesmas

é mais previsões,  
a inclusão de  
variáveis expli-  
cativas pode  
melhorar  
o modelo.

(ex: análise de venda de cerveja,  
incluindo temperatura como var. explicativa)

↳ comp. sazonal n'teria numa explicação  
que não

→ Pode-se 1º fazer o procedimento pf var. explicativa

Se por exemplo resíduos ainda n'rem RB, tem correlações

⇒ pode incluir variável explicativa

Ainda assim, o significado da sazonalidade nas

será mais o mesmo que o verificado ante da variável  
explicativa.

utiliza

Poderemos querer separar por exemplo a influência de 2 fatores na razoabilidade:

$$(i) y_t = \mu_t + \gamma_t + \epsilon_t$$

$\gamma_t$ : raz. total  $\Rightarrow$  temp + feriados etc.

$$(ii) y_t = \mu_t + \gamma_t + \delta x_t + \epsilon_t$$

$\gamma_t$ : temperatura

$\delta x_t$  (expulsando temp.)  $\Rightarrow$  feriado

Obs 2: Essa variável explicativa pode ser uma dummy ou ainda, intervenções

- Por exemplo, raz modelagem e análise de resíduos ainda apresenta problema, como por exemplo, por conta de outliers

$\Rightarrow$  Modela intervenções a posteriori

O que é atípico não pode ser decidido a priori. É preciso modelar e verificar.

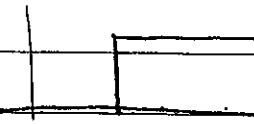
Pl o modelo dinâmico, só se pode identificar o que é atípico a posteriori. Muito algo que a princípio não de rodar o modelo parece estranho, pode ser absorvido pelo modelo.

Por isso, esta análise do que é atípico só pode ser feita a posteriori, olhando o comportamento dos resíduos.

Exs. de intervenções:

i) Pulso  $x_t = \begin{cases} 1 & t = t^* \\ 0 & t \neq t^* \end{cases}$

ii) Escadas  
(step function).

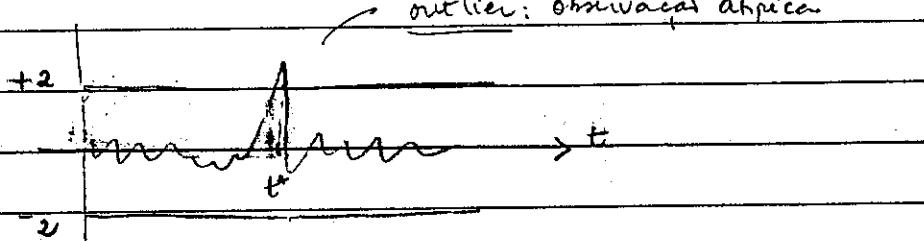


$$x_t = \begin{cases} 0 & t < t^* \\ 1 & t \geq t^* \end{cases}$$

última

Qual tipo de funções usar para intervenções dependerá da análise de resíduos na inovação.

Ex1:  $\hat{v}_t^* = \frac{v_t}{f_t^{1/2}} \sim N(0,1)$



Nesse caso, coloca variável explicativa como função pulso.

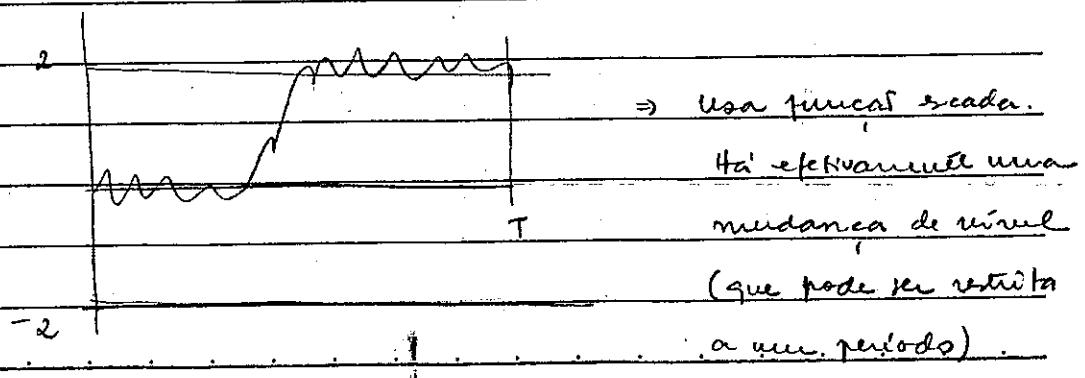
Como esse  $t^*$  ocorreu algo atípico, muito fora do padrão, modela como dummy PI melhorar o ajuste do modelo in-sample.

Este ponto atípico pode estar comprometendo o modelo no que se refere a heterocedasticidade, normalidade, outros fatores.

Poderia estar condenando o modelo por conta de uma observação.

Com intervenções  $\Rightarrow$  formas de acomodar obs. ao modelo.

Ex2:



Uilibria

Há outras funções elaboradas.

(importante  $\Rightarrow$  distinção entre efeito transitório e permanente.

6  
↓  
até agora  
observações nesse  
período t.  
não uma  
mudança  
de estrutura.

Não sempre é possível / fácil verificar efeito transitório permanente na análise dos resíduos das inovações.

Uma forma é olhar se resíduos no nível das variáveis de estado.



Só nos ME podemos fazer esta distinção, tratando resíduos no nível das variáveis de estado

(No modelo, um efeito transitório em  $\eta_t$  é permanente em  $y_t$ , devido a estrutura AR de  $\mu_t$ )

$$\mu_t = \mu_0 + \varepsilon_{\mu_t}$$

---

### Diagnósticos

Modelo:  $y_t = z_t \alpha_t + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, \sigma_e^2)$

$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$ ,  $\eta_t \sim N(0, \sigma_\eta^2)$

Hipóteses: i) linearidade das componentes aditivas

x combinação de forma linear aditiva  
p/ permanecer  $y_t$

ii)  $\varepsilon_t \rightarrow$  i) Normal

ii) homocedástico

iii) descorrelacionado  $\rightarrow E[\varepsilon_t, \varepsilon_{t-s}] = 0, \forall s \neq 0$

Método de ajuste :  $R^2$ , MAPE, RMSE etc

(adequação)

tilibra

↳ se o modelo é adequado.

### Residuals

Seja  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$  o modelo estimado.  
Agora queremos ver se os resíduos  
sao iid.

Verificando se os resíduos sao iid.

$$\text{Resíduo} = y_t - \hat{y}_t = \varepsilon_t \quad (\text{estimado})$$

No teste de autocorrelação o algoritmo  
calcula a soma das autocorrelações  
de ordens 1, 2, ..., k para os resíduos.

$$E_t = y_t - \hat{y}_t x_t \Rightarrow E_t \text{ não observado.}$$

$$\hat{E}_t = y_t - \hat{y}_t \hat{x}_{t+1} = \hat{\nu}_t \text{ não observado, estimado de } E_t$$

Sabemos que  $E(\nu_t) = 0$

$$E(\nu_t^2) = f_t$$

$$\Rightarrow \text{movimento indeterminado. } \tilde{\nu}_t = \frac{\hat{\nu}_t}{\hat{f}_t^{1/2}} \sim N(0, 1)$$

Então em vez de usar

os resíduos  $\varepsilon_t$  usamos  $\tilde{\nu}_t$ .

Assim

se os resíduos  $\tilde{\nu}_t$  forem iid, então os resíduos  $\varepsilon_t$  tambem.

Uma das diagnósticos é verificar o quanto  $\text{var}(\tilde{Y})$  está  $\approx 1$

$\text{var}(\tilde{Y}) \approx 1 \Rightarrow \text{Bom!}$

⇒ Normalidade:

Quais procedimentos podemos usar para verificar?

Histograma

Q-Q plot ①

testes estatísticos:  $\left\{ \begin{array}{l} \text{Jarque-Bera (simetria e} \\ \text{assimetria)} \end{array} \right.$

A Darling

$$JB = \frac{n}{6} (\hat{s}^2 - 1) + \frac{n}{24} (\hat{k}^2 - 3) \sim \chi^2(2)$$

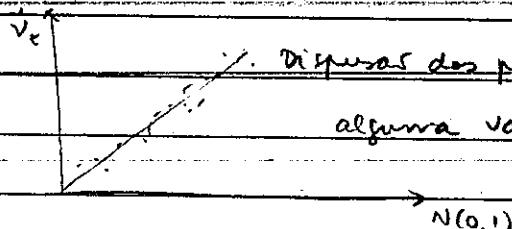
(estudar!)

$AD = \int (F_n - F_0)^2 dF$  (relações entre densidade empírica

$F_0(1-F_0)$  e teórica)

⇒ tem mais  
potência/poder  
sob  $H_0$ .

② Q-Q plot



• Dispersão dos pontos em volta de 45° com  
alguma variabilidade estatística

⇒ Homocedasticidade: variação do erro cte.

Pode ser feita de 2 formas

→ No Stata, busca um tipo específico de homocedasticidade

Obs: pf série trimestral  $\Rightarrow$  usa-se  $m=3$   $\rightarrow$  pega tb sazonalidade.  
" " mensal  $\Rightarrow$  usa-se  $m=12$

(  
Na F.A.C., olha tb pontual

$r_j$  = autocorrelações das inovações

- - -

Pf saber se há tb efeito ARCH, outra possibilidade é  
trabalhar com estatística Ljung-Box para  $v_t^2$

(b) teste Q para  $v_t^2 \Rightarrow$  investigar efeito ARCH em  $v_t$

(  
F.A.C. pf  $v_t^2 \Rightarrow v_t^2 = \alpha + \beta v_{t-1}^2 + \epsilon_t$   
 $\text{cov}(v_t^2, v_{t-1}^2) \neq 0.$

- - -

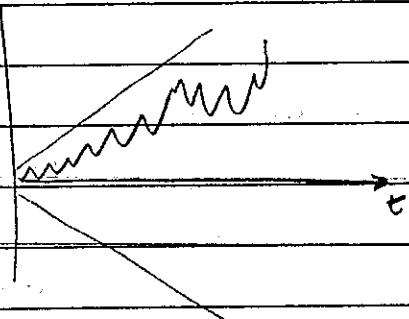
Outros tipos de teste

- Soma acumulada das inovações<sup>2</sup>

b) se modelo bem especificado,

fica dentro de uma banda

se cair para, pode ser sinal de  
heteroscedasticidade



## Resíduos Auxiliares

e Inovações:  $v_t = \hat{e}_{t|t-1} = y_t - \hat{y}_{t|t-1}$

Resíduos auxiliares:  $\hat{e}_t = E[\hat{e}_t | Y_n]$

(cap. 4, sec. 4.4)

parece com inovações,

mas inovação

baseada  
até info de  
 $t-1$ .

R. Aux. é  
concorrente  
inovações  
suavizada.

Mostre que:  $\hat{e}_t = H_t \mu_t$ ,  $t = n, \dots, 1$

onde  $\mu_t = F_t^{-1} v_t - K_t' r_t$  → do algoritmo de suavização  
(calculado de forma recursiva)

$H_t = E[\hat{e}_t \hat{e}_t']$  No caso univariado =  $\sigma_e^2$

$$\Rightarrow \text{var}(\hat{e}_t | Y_n) = H_t - H_t D_t H_t'$$

$$D_t = F_t^{-1} + K_t' + N_t K_t$$

→  $+ t$  do algoritmo de suavização.

Temos também que:

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$$

$\hat{\eta}_t = E[\eta_t | Y_n] \Rightarrow$  estimativa suavizada de  $\eta_t$  (do resíduo no nível da variável de estado).

$$= Q_t R_t' r_t$$

$$Q_t = E[\eta_t \eta_t']$$

geralmente é constante

$$\text{var}[\hat{\eta}_t | Y_n] = Q_t - Q_t R_t' N_t R_t Q_t$$

Além disso:  $\hat{e}_t$  e  $\hat{\eta}_t$  não são descorrelacionados, nem descorrelacionados entre si.

⇒ Resíduos auxiliares podem ser facilmente calculados a partir do algoritmo de suavização.



08/06/11

Nos é tal formalizado  
como teste estatístico.  
nos está bem documentada em  
livros antigos.

## Resíduos auxiliares como ferramenta de diagnóstico

constroi  $\hat{\epsilon}_t = E[\epsilon_t | y]$  } estruturas suavizadas  
2 tipos de resíduos, além dos resíduos de inovações

da inovação,  $\hat{\eta}_t = E[\epsilon_t | y]$  } dos erros:  
específicos de mE } (1) sar intra-correlacionadas (pois tas  
parte do algoritmo de suavizações } suavizad  
calcular os resíduos de inovação  
toda instante t } (2) sar inter-correlacionados  
de info em comum p/ calcular os

obs: Nos outros mE suavizad  
componente de pt ou da sazonalidade. também usamos a notação

$$y = (y_1, y_2, \dots, y_T) \quad t=1, 2, \dots, T$$

$$Y_T = (y_T, y_{T-1}, \dots, y_1) \equiv y$$

Nos trazemos  
mas mE diagnósticos (mas per-  
mentas e nos usados em testes  
estatísticos).

Estratégia: Usar  $\hat{\eta}_t$  e  $\hat{\epsilon}_t$  para distinguir entre choques transientes  
e choques permanentes em  $y_t$ , de forma a mel-

hor caracterizar as análises de intervenções.

(choque detectado padras anormais em resíduos, observações atípicas).

Super modelo nível local  $y_t = \mu_t + \epsilon_t$

(exceções simples)

$$\mu_t = \mu_{t-1} + \eta_t \Rightarrow \mu_t = \mu_1 + \sum_{i=1}^t \eta_i$$

"choque"

m ajuste

a posteriori

depois

de obs

var

componente

tempo segun-

do o modelo.

pode haver ocorrido

uma situaçao

que o modelo

não seja capaz

de detectar

mais tipos de choques  $\rightarrow$  em  $\epsilon_t$ : transiente

afetará  $y_t$  apenas em  $t$

$\rightarrow$  outlier: tratado via função peso

$$D_t = \begin{cases} 1 & t=t^* \\ 0 & t \neq t^* \end{cases} \Rightarrow y_t = \mu_t + \delta D_t + \epsilon_t$$

em  $\eta_t$ : Dada recursividade de  $\mu_t$ , em  
virtude de sua natureza markoviana.

Nos é uma permanente  
eterno. Pode ser compensando com  $\eta_t$  paus  
quentes. Dependendo do  
nível do choque, pode  
ser carregado.

e raiz unitária, um choque em  $\eta_j$

$1 \leq j \leq t$  será incorporado a  $\mu_t$  de forma

permanente e, assim, afeta  $y_t$  também

de forma permanente.

$\Rightarrow$  Roda profunda, calculando  $\hat{\epsilon}_t$  e  $\hat{\eta}_t$ . Detectar períodos onde  $\hat{\epsilon}_t \geq K\sigma_{\hat{\epsilon}_t}$  e  $\hat{\eta}_t \geq K\sigma_{\hat{\eta}_t}$   
se for esse  $\hat{\epsilon}_t \rightarrow$  coloca função peso e estima parâmetro  $\delta$ . "Ajuda" externa ao modelo.  
Isto só ajuda mE melhor estatísticas do modelo insomte.

(\*) Se detecta observações anómalas em  $\eta_t$ : trata outliers esse  $D_t$  ou escada em  $y_t$ .  
 ⇒ Esta é a forma padrão que o stamp tratará intervenções. Itália formas + sofisticadas de tratar.

Ex. Suponha que em  $t = t^*$  haja um choque em  $\eta_{t^*}$ .  $\rightarrow$  informalmente k+2 características anômalas  
 Portanto  $\eta_{t^*}$  é "grande" =  $K\alpha_1$ ,  $K=2, 2.5, 3$  etc.

Assim, para  $t \geq t^*$ ,  $y_t$  contará este choque e, assim,

$$\begin{cases} y_t = \mu_t + \eta_t^* + \eta_{t-1} + \dots + \eta_1, & t = t^* \\ y_t = \mu_t + \eta_t + \eta_{t-1}^* + \eta_{t-1} + \dots + \eta_1, & t > t^* \end{cases}$$

Ou seja, se  $t \geq t^* \Rightarrow y_t$  mudará de nível de forma permanente.

Assim, ao detectarmos um "outlier" em  $\eta_t$ , podemos tratá-lo como tal em  $\eta_t$ :

(a) via funções pulso

(b) como uma função escada  $\begin{cases} 1 & t \geq t^* \text{ na equação de } y_t \\ 0 & t < t^* \end{cases}$

$$(a) y_t = \mu_t + \epsilon_t$$

$$\mu_t = \mu_{t-1} + \delta D_t + \eta_t$$

$$D_t = \begin{cases} 1 & t = t^* \\ 0 & t < t^* \end{cases}$$

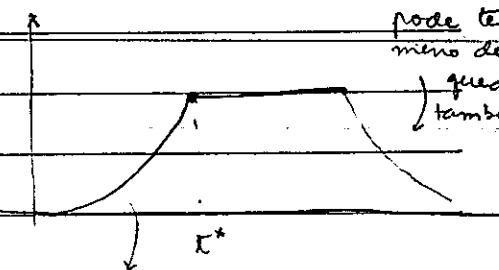
$$(b) y_t = \mu_t + \delta E_t + \epsilon_t$$

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

$$E_t = \begin{cases} 1 & t \geq t^* \\ 0 & t < t^* \end{cases}$$

f

Outra alternativa:



Obs: Dist. dos

resíduos auxiliares  
também é assintoticamente normal.

Cuidado! Teste de Normalidade

Torque-Beta padrões  
não iid.

No caso dos resíduos auxiliares,  
mas fas iid ⇒ perita usar teste  
de Jarque-Bera corrigido.

pode haver uma cota infecciosa

Ex: lei de cinto de segurança introduzida,  
Análise do n.º de acidentados  
no trânsito (letalidade)

Uilibra

$$\text{Obj 2: } y_t = \phi y_{t-1} + \delta D_t + \epsilon_t$$

↳ piso em  $t=t^*$

Mais formas de estimar

Dividir

MV cond  $\Rightarrow \mu_0$

$$\Rightarrow \mu_{t^*} = 0$$

$$\text{MV incond} \Rightarrow y_0 \sim N\left(0, \frac{\sigma^2}{1-\phi^2}\right) \Rightarrow \mu_0 \approx 0 \text{ (Nasceria zero)}$$

Exemplo: série anual

1871 - 1970

Vazão do Nilo em  $10^8 \text{ m}^3$

importante conhecer a série pf entender questões dos outliers.

| 1899  $\Rightarrow$  construções da barragem de Assuan  $\Rightarrow$  séria permanente (\*1)

| 1913  $\Rightarrow$  obs. atípica  $\Rightarrow$  transitória (\*2)

Estratégia: (\*1) em  $\hat{\eta}_t$

(\*2) em  $\hat{\epsilon}_t$

$\rightarrow$  Olha a série. Data Plot

$\rightarrow$  Model

basic components: level  $\rightarrow$  stochastic

models

nível local

Estimation method: MV

$\rightarrow$  Tat Mense

Auxiliary residual graphics

More written outputs  $\Rightarrow$  write <sup>longer</sup> absolute values

..... exceeding the value of [3]

Irregular:  $\hat{\epsilon}_t$

level:  $\hat{\eta}_t$

tilibra

Vieus > Results

Values larger  $\Rightarrow$  p value alto?

(prob.)  $\Rightarrow$  rejeitar hipótese nula

Select interventions

Irregular  $\Rightarrow$  1913

Level  $\Rightarrow$  1899

Test  $\xrightarrow{\text{menos}}$  component graphic

Nível praticamente cte. (linha reta) a menos da intervenção mudando o patamar.

Provavelmente  $\sigma^2 \approx 0$  ( $\mu_t = \mu_{t-1} + \eta_t$ )

No Manual: pp estimar models

e' colocado um medium cycle (10 anos)

Obs: Ter a possibilidade de intervenções transitórias e permanentes é inerente a m.

Esse modelo ARIMA, p. ex., tem uma fonte apenas de choque.

Modelos Estruturais multivariados

- Generalizações de modelos estruturais univariados.

Mais complexos em relações a estimativas ( $f_k$  e  $m_v$ ), identificacões paramétrica, interpretações e diagnósticos.

$$\text{Univariado} \Rightarrow y_{it} = \mu_{it} + \gamma_{it} + \varepsilon_{it}$$

↳ diagnósticos

$$\frac{v_{it}}{F_{it}^{1/2}} = \frac{y_{it} - \tilde{y}_{it+1t-1}}{F_{it}^{1/2}}$$

(testes JB, B-Ljung, Ptc.)

$$\text{Multivariado} \Rightarrow y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{kt} \end{pmatrix} = \mu_t + \gamma_t + \varepsilon_t$$

$$\Rightarrow v_t = \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \sim N(0, F_t)$$

tests se satisfeçõe

normais bivariadas.

(se satisfeçõe  $N$  bi  $\Rightarrow$  cada marginal

satisfaz  $N$  univariado, mas isto não

garante  $N$  univariado nas estimativas

verificam se vale normal bivariada)

pode resumir:

Testes de diagnósticos precisam ser bi, tri, ..., multivariados.

= Teste bivariado para: normalidade

Auto-correlações  $\rightarrow$  inovações inter e intra  
descorrelatados  
etc.  
(testes conjuntos)

• Motivação:

flexibilidade → estimar componentes comuns entre ST's.

→ melhorar inferência e previsões / simulações.

$$z_t = f(y_{1t}, y_{2t})$$

↓

✓ fatores de risco

resultado é

Variável comum

(ex: dados financeiros)

ex:  $z_t$ : lucro Pene

$y_{1t}$ : preço petróleo

$y_{2t}$ : " gasolina.

Se tratar  $y_{1t}$  e  $y_{2t}$  de forma separada, perde possibilidade de ter influências / ligações diretas entre os 2.

É indicado por razões econômicas / financeiras o tratamento de fatores de risco de forma multivariada.

↓

de forma mais geral, pode ser complexo e se tornar custoso (mais acadêmico) mas na prática indicam-se alguns agrupamentos.

1) Modelo SUTSE : seemingly unrelated TS equations

$$\begin{aligned} & \xrightarrow{\text{pxL}} & & \xrightarrow{\text{PXP}} \\ y_t &= \mu_t + \varepsilon_t & , \quad \varepsilon_t \sim N(0, \Sigma_\varepsilon) \\ \mu_{t+1} &= \mu_t + \eta_t & , \quad \eta_t \sim N(0, \Sigma_\eta) \end{aligned}$$

↓ p x p

Se  $p=5 \Rightarrow 15$  parâmetros de  $\Sigma_\varepsilon$  e

" " de  $\Sigma_\eta$

obs: N° de parâmetros é  $n + \binom{n}{2}$

Variações, covariâncias.

equilíbrio

Então, se não tivermos nenhuma restrição, estaremos falando de 30 parâmetros.

Para  $p=2$

$y_{1t}$  só depende de  $\mu_{1t}$

$y_{2t}$  " " de  $\mu_{2t}$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

$\Rightarrow$  Nas x comunicam explicitamente.

$$\begin{pmatrix} \mu_{1,t+1} \\ \mu_{2,t+1} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

A comunicação vem das covariâncias. Se fizer teste de hipótese e covariâncias estatisticamente nulas

$$\epsilon_t \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\epsilon_1}^2 & \sigma_{\epsilon_1 \epsilon_2} \\ \sigma_{\epsilon_1 \epsilon_2} & \sigma_{\epsilon_2}^2 \end{pmatrix} \right)$$

$\Rightarrow$  devem tratar as x separadamente

$$\eta_t \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\eta_1}^2 & \sigma_{\eta_1 \eta_2} \\ \sigma_{\eta_1 \eta_2} & \sigma_{\eta_2}^2 \end{pmatrix} \right)$$

$\Rightarrow$  Nesta forma geral, o SUTSE é de difícil estimacão.

-- --

Queremos um modelo SUTSE + simples, com menos parâmetros.

Definições: Modelos homogêneos

= Processo estocástico de 2ª ordem, p-variado, é dito homogêneo se todas as combinações lineares de suas componentes possuem as mesmas propriedades estocásticas de 2ª ordem.

Em outras palavras,  $y_t = (y_{1t}, y_{2t}, \dots, y_{pt})$  é homogêneo se

$\forall \alpha \in \mathbb{R}^p$ , o processo univariado estacionário  $z_t = (\alpha' y_t)$

possui FAC independente de  $\alpha$ .

• pode ser

que y nas sejam estacionárias (não necessário dizer isso)

como a FAC é a média  $\mu$

• qualquer combinação, mas não depende de  $\alpha$ .

$$\text{Ex: } z_t = 1y_{1t} + 2y_{2t} \quad | \quad \text{FAC não depende de } \alpha \\ 0,5y_{1t} + 0,7y_{2t}$$

titúlio

nas é difícil de mostrar. Faz  $\eta = p=2$ .

Proposição 1: Se  $y_t$  segue um SUTSE de nível local, então  $y_t$  será homogêneo SSS

$$\Sigma_{\eta_t} = q \Sigma_{\varepsilon_t}, \quad q \in \mathbb{R}^+$$

$\& p=5 \quad \left\{ \begin{array}{l} \text{SUTSE Geral: } \#4=3 \\ \text{SUTSE homog: } \#4= \\ (15+q) \end{array} \right.$

$\mu_t, \delta_t, \gamma_t \text{ etc.}$

Proposição 2: Se  $y_t$  segue um SUTSE com  $j$  tipos de componentes ortogonais, então  $y_t$  será homogêneo SSS

$$\Sigma_{\eta_t} = \begin{pmatrix} \Sigma_{\eta_1} & & & \\ & \ddots & & \\ & & \Sigma_{\eta_j} & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} q_1 \Sigma_{E_{11}} & & & \\ & q_2 \Sigma_{E_{22}} & & \\ & & \ddots & \\ & & & q_j \Sigma_{E_{jj}} \end{pmatrix}$$

como processo  
ortogonal  
só há comunicação (?)

1 - Tendência

2 - Racionalidade

$$\text{onde } \Sigma_{\varepsilon_t} = h \Sigma_{\varepsilon}$$

DÚVIDA  
cada  $\Sigma_{\eta_i}$  é que é  
um bloco por componentes  
de  $\eta$  e vai de 1 a  $p$   
não é só para  $\eta$ , mas para  
tudo que é para  $\eta$

### Propriedades Empíricas de PH's. (processos homogêneos)

- i) Pode haver teorias "a priori" que justifiquem a aderção de estrutura de proporcionalidade entre as covariâncias
- ii) As regras de Kalman podem ser implementadas para cada uma das  $p$ -equações separadamente (ver Harvey)  
(livro azul, + grano.)
- iii) Há menos risco de não identificacão paramétrica pela baixa dimensão de  $\Psi$ .  
→ processo homogêneo é sempre identificável.
- iv) Estimacão por MV é consideravelmente simplificada.

Tudo pode ser estudado se nenhuma complicação  
H MLL, q razoabilidade, d' ciclo etc.  
for será cobrado MNL.

22/06/11

## MNL - multivariado:

i) SUTSE - Irrestrito - modelos, t qual no que diz respeito a parâmetros.  
muitos parâmetros p/ estima.

problemas de convergência  
e instabilidades.

uma simplificação é  
usar modelos homogêneos.  
 $\Sigma_{\epsilon} = q \Sigma_{\epsilon}$

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma_{\epsilon})$$

$\downarrow p \times 1 \quad \downarrow p \times 1 \quad \downarrow p \times p$

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim N(0, \Sigma_{\eta})$$

$$\downarrow p \times p \quad \downarrow p \times p \quad \downarrow p \times p$$

$$\Sigma_{\eta} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{(k \times p)}$$

ii) Parametrizações do SUTSE → na aula passada, avaliamos H p=2.

Poderemos sempre redefinir  $\alpha_t^+ = \beta \alpha_t$

→ definição do vetor de estado  $\alpha_t$

é totalmente arbitrária. Podemos

sempre ter uma outra

parametrização que leve a

uma outra interpretação, mas

que mantenha coerência e

verossimilhança.

*Ver desenho*  
diferentes  
niveis  
nos dias  
e aulas  
e aulas  
passadas

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} I_K & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} \mu_t^+ \\ \bar{\mu}_t \end{pmatrix}$$

$K \times 1$   
 $(p-k) \times 1$

$$\begin{pmatrix} \mu_t^+ \\ \bar{\mu}_t \end{pmatrix} = \begin{pmatrix} \mu_{t-1}^+ \\ \bar{\mu}_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t^+ \\ \bar{\eta}_t \end{pmatrix}$$

$(p-k) \times 1$

$$\hookrightarrow N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{\eta}^+ & 0 \\ 0 & \bar{\Sigma} \end{pmatrix} \right)$$

Parametrizações  
não p/ difundir à  
questões de componentes  
comuns.

$$\text{onde } \Sigma_{\eta}^+ = \Sigma_{11} \quad e \quad \bar{\Sigma} = \Sigma_{22} \Sigma_{11}^{-1}$$

iii) Modelos com níveis comuns → agora determinar cond. suficiente  
p/ existência de comp. comuns (nesse  
caso do MNL, estamos buscando os níveis

se posto  $(\Sigma_{\eta}) = k < p \Rightarrow \mu_t^+$  contém k níveis comuns

tem informação sendo  $\Sigma_{\eta}^+$  p.d. e  $\bar{\Sigma} = 0$ , então pode-se escrever (II) como:  
casos importantes

p/ dia

quuntas

componentes LI

tem independentes.

Se nesse posto cheio ⇒ não há

$$\bar{\mu}_t = \bar{\mu}_{t-1} \Rightarrow \text{ct. determinística} = \bar{\mu}$$

tilibra

$$\begin{pmatrix} K \times 1 \\ \vdots \\ y_{1t} = \mu_t^+ + \epsilon_{1t} \\ \vdots \\ y_{2t} = \Pi \mu_t^+ + \bar{\mu} + \epsilon_{2t} \\ \vdots \\ \mu_t^+ = \mu_{t-1}^+ + \eta_t^+, \quad \eta_t^+ \sim N(0, \Sigma_\eta^+) \end{pmatrix}$$

Fórmula operativa do MLE onde  
1º k componente sal direcio.

madas por  $\mu_t^+$  e outras ( $p - k$ )  
sas cl de  $\mu_t^+$

$\mu_t^+ = \mu_{t-1}^+ + \eta_t^+, \quad \eta_t^+ \sim N(0, \Sigma_\eta^+) \Rightarrow$  apenas k componentes evoluem  
segundo processo linear local

Aíris, na prática, saber se um determinado ME possui

"componentes comuns" pode ser inferido observando-se o posto  
núm. da matriz covariância dos choques dessa componente (no caso  
razoável:  
dade:  $p=2 \Leftrightarrow \rho_{\eta_1, \eta_2} = 1$ ) → na prática, partindo de um SUTSE genérico,  
ciclo.

Queremos ver se há comp. comuns (\*\*)

(Outra pt posto 2)

Nyblom & Harvey (2001) desenvolvem um teste  $\tau_p$  de componentes comuns  
dentro do arcabouço de ME. → Teste + formulais  $\tau_p$  verificam se  
posto  $\Sigma_\eta$  é  $\leq p$ .

-11-

### Modelos Rotacionados

que possuem k fatores comuns

Muitas vezes, o modelo tipo (IIa) não é de fácil interpretação.

Entretanto, como as componentes NO nas sas únicas, podemos  
procurar um outro conjunto de componentes comuns que a-  
presentem melhor interpretação.

Considere as equações do modelo (IIa)

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} I_K \\ \Pi \end{pmatrix} \mu_t^+ + \epsilon_t$$

$$\text{e } \mu_t^+ = \mu_{t-1}^+ + \eta_t^+ \quad (*)$$

para obter uma matriz de var  
cova diagonal.

Seja  $\mu_t^* = (\Sigma_{\eta}^+)^{-1/2} \mu_t^+$  a "nova componente"

$$\Rightarrow \mu_t^* = (\Sigma_{\eta}^+)^{-1/2} \mu_t^+$$

de (\*):

$$(\Sigma_{\eta}^+)^{-1/2} \mu_t^+ = (\Sigma_{\eta}^+)^{-1/2} \mu_{t-1}^+ + (\Sigma_{\eta}^+)^{-1/2} \eta_t^+$$

$$\Rightarrow \mu_t^* = \mu_{t-1}^* + \eta_t^* \quad \text{onde } E[\eta_t^* \eta_t^{*'}] = I_K$$

$$\mu_t^* = \mu_{t-1}^* + \eta_t^* \quad (\text{Prova!})$$

Por outro lado:

$$y_t = \begin{pmatrix} I_K \\ \pi \end{pmatrix} \mu_t^* + \begin{pmatrix} 0 \\ \bar{\mu} \end{pmatrix} + \varepsilon_t$$

$$= \begin{pmatrix} I_K \\ \pi \end{pmatrix} \Sigma_{\eta}^{1/2} \mu_t^* + \begin{pmatrix} 0 \\ \bar{\mu} \end{pmatrix} + \varepsilon_t \quad (\text{a})$$

Seja  $\Pi^+ = (I, \pi')$

$$\textcircled{1} = \Pi^+ \Sigma_{\eta}^{-1/2}$$

$$\mu_0 = (0, \bar{\mu})'$$

Então (a) pode ser resrito como:

$$\left\{ \begin{array}{l} y_t = \textcircled{1} \mu_t^* + \mu_0 + \varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma_{\varepsilon}) \end{array} \right.$$

$$\mu_t^* = \mu_{t-1}^* + \eta_t^* \quad \eta_t^* \sim N(0, I_K)$$

É o ponto de partida se começarmos a fazer rotacões (a partir  
da obtenção das componentes ortogonais) tilibra

Nessa nova parametrização (notas), os fatores comuns nas saídas correlacionadas entre si e possuem variância unitária.

Se  $(\Sigma_{\eta}^+)^{1/2}$  é triangular superior, então  $\theta$  será do tipo ~~ts~~ também  
 Ex: 
$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} - & - & - \\ - & - & - \\ - & - & - \end{pmatrix} \begin{pmatrix} \mu_{1,t} \\ \mu_{2,t} \\ \mu_{3,t} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix}$$
 (~~comum!~~)  
 $p=3 \quad K=2 \quad \theta_{ij} = 0 \text{ para } j > i$

obs: CORREÇÕES nas NOTAS:

$$PF 15: \mu_t^+ = \mu_{t-1}^+ + \eta_t^+$$

$$\circlearrowleft (\Sigma_{\eta}^+)^{-1/2} \mu_t^+ \dots$$

até aqui

$$\Rightarrow \mu_t^+ = \mu_{t-1}^+ + \eta_t^+$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} \theta_{11}, 0 \\ \theta_{21}, \theta_{22} \\ \theta_{31}, \theta_{32} \end{pmatrix} \begin{pmatrix} \mu_{1t}^+ \\ \mu_{2t}^+ \\ \mu_{3t}^+ \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix}$$

$$\Rightarrow y_{1t} = \mu_{1t}^+ + \epsilon_{1t} \rightarrow \text{poderia ter wr}$$

(influencia da para)

$$y_{2t} = \theta_{21} \mu_{1t}^+ + \theta_{22} \mu_{2t}^+ + \epsilon_{2t}$$

$$y_{3t} = \theta_{31} \mu_{1t}^+ + \theta_{32} \mu_{2t}^+ + \epsilon_{3t} \text{ poderia ter duração}$$

(\*\*) Na prática, partindo de um WTSF comum, se  $p=2$ ,

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \Rightarrow \hat{p} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \text{ se } \hat{p} \approx \pm 1 \Rightarrow \text{há uma componente comum.}$$

(obs: se  $p=2$  podemos ter até uma comp. comum).

se genérico  $\Rightarrow$  pode ter  $K$  comp. comuns).

Se tivermos modelando simultaneamente 3 séries:

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} \text{ petróleo quando há comp. comum}$$

. . . . .  $\Rightarrow$  há co-integrações.  
 tólibra . . . . . (movimentos no preço das jazidas.  
 . . . . . Transferidos do petróleo pf derivados)

Obs: p-valor

espectro RB  $\Rightarrow$  cte.

MARPE:  $p_f$  para da amostra

$\Rightarrow$  deixar 6 obs. para da amostra

— — —

Exemplo:

Dentro do arrebatamento ME

- Tornando 2 séries, como avaliar se cointegram

- Seja usav. SUTSE, tem a tendência e verifica correlação entre módos.

Usando SUTSE: olhar matriz de variancia de  $\mu$ .

— — —

Ex. 4: com modelos

Gerar série c/ 300 obs por MC (série artificial)

Como sabe valor verdadeiro, usa como valor inicial algo perto do valor verdadeiro.

$$\text{ex: } \sigma_e^2 = 10^{-3}$$

como  $\Rightarrow 0,01$

$$\sigma_\eta^2 = 0,01$$

$y(0) \quad 0,1$

$$\sigma_z^2 = 0,0001$$

0,1

$$\sigma_w^2 = 0,2$$

0,1

outras

coisa,

gera

série com

tend. e razona-

lidade determinística

otimizações das op. estatísticas.

Mas valores EIR

$\Rightarrow$  usa log verossimil:  $\ell = l + \exp((\eta))$

(a cada iteração)

$\underbrace{\eta}_{\text{EIR}}$  libra

$\sigma_\eta$

/ /

Outra coisa: escalarar a verossimilhança  $\Rightarrow l/n$   
quando passar p/  
a retina (as  
acabar iteracões) (diridir pa n° de obs)

Ál siver parte p/ a série real.

→ Outro ponto: concentra verossimilhança

## MLN SUTSE homogêneos

$$y_t = \mu_t + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, \Sigma_{\varepsilon})$$

$$\mu_{t+1} = \mu_t + \eta_t \quad , \quad \eta_t \sim N(0, \Sigma_{\eta})$$

precisam ser estimadas.

$$\Sigma_{\eta} = q \Sigma_{\varepsilon} \Leftrightarrow \begin{pmatrix} \sigma_{\eta_1}^2 & \sigma_{\eta_1, \eta_2} \\ \sigma_{\eta_2, \eta_1} & \sigma_{\eta_2}^2 \end{pmatrix} = q \begin{pmatrix} \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_1, \varepsilon_2} \\ \sigma_{\varepsilon_2, \varepsilon_1} & \sigma_{\varepsilon_2}^2 \end{pmatrix}$$

Os outros por construções iterações automaticamente obtidos.

$$\begin{cases} p=2 \\ \text{p. exemplo} \\ y_{it} = \mu_{it} + \varepsilon_{it} \end{cases} \Rightarrow \Delta y_{it} = \eta_{i,t-1} + \Delta \varepsilon_{it} \quad , \quad i=1,2$$

$$\mu_{it+1} = \mu_{it} + \eta_{it}$$

$\sim MA(1)$  (I)

$$\text{Assim: } z_t = w_t + \theta w_{t-1} \quad , \quad w_t \sim N(0, \sigma_w^2)$$

$$E[z_t] = 0$$

$$\text{não faz } E[z_t^2] = (1+\theta^2) \sigma_w^2 \Rightarrow E[z_t z_{t-1}] = \frac{E[z_t z_{t-1}]}{\text{var}[z_t]} = 0$$

sentido.

$$\text{Deveria} \\ \text{outra MA bisarada. } E[z_t z_{t-1}] = \theta \sigma_w^2 \quad (\text{II})$$

$$\text{de (I): } E[\Delta y_{it}] = 0$$

$$E[\Delta y_{it}^2] = \sigma_{\eta_i}^2 + 2\sigma_{\varepsilon_i}^2 \Rightarrow \rho_{\Delta y_{it}}(1) = -\frac{\sigma_{\varepsilon_i}^2}{(\sigma_{\eta_i}^2 + 2\sigma_{\varepsilon_i}^2)}$$

FAC das  
primeiras  
reduzida  
de cada  
uma  
das duas  
colunas  
do  
SUTSE

$$\text{Mas por homogeneidade: } \sigma_{\eta_i}^2 = (q) \sigma_{\varepsilon_i}^2 \quad , \quad i=1,2$$

$$\Rightarrow \rho_{\Delta y_{it}}(1) = -\frac{\sigma_{\varepsilon_i}^2}{q\sigma_{\varepsilon_i}^2 + 2\sigma_{\varepsilon_i}^2} = -\frac{1}{q+2} \Rightarrow \text{independe de } i \quad (i=1,2)$$

(é a mesma p1 y1t, y2t. se tornarmos t6

C. l. de n'ries.  $\Rightarrow$  não mudaria).

$$V_t = \alpha_1 \Delta y_{1t} + \alpha_2 \Delta y_{2t} \Rightarrow \text{FAC: } p(1) \text{ também é ilíbra}$$

MA(1) MA(1) independe de  $\alpha_1$  e  $\alpha_2$ .

cada um dos MA  
faz a mesma dinâmica.

Portanto, na forma reduzida do SUTSE homogêneo, os séries  
possuem a mesma dinâmica

⇒ trata-se de um modelo bem restritivo.

$$\text{Estimacões} \Rightarrow \Psi = (\sigma_{e1}^2, \sigma_{e2}^2, \sigma_{e,t+1}, g), \quad \tilde{y}_{t-1} = (\tilde{y}_{1,t-1}, \tilde{y}_{2,t-1})$$

$$l(\Psi) = \log L(\Psi) = \sum_{t=1}^n \log p(y_{1t}, y_{2t} | \tilde{y}_{t-1}) \quad \begin{matrix} \text{funções densidade} \\ \text{conjunta bivariada.} \end{matrix}$$

$$= \sum_{t=1}^n \log \left[ \frac{1}{(2\pi)^{\frac{n}{2}} |F_t|^{1/2}} \exp \left\{ -\frac{1}{2} \tilde{v}_t' F_t^{-1} \tilde{v}_t \right\} \right]$$

Usando inicializa-  
ção difusa exata.

$$= -n \log 2\pi - \frac{1}{2} \sum_{t=1}^d w_t - \frac{1}{2} \sum_{t=d+1}^n \log |F_t| - \frac{1}{2} \sum_{t=d+1}^n \tilde{v}_t' F_t^{-1} \tilde{v}_t$$

De usarmos bif.  
kappa, vrs

mas fast afetados pela inici. exata.

na forma  $\sum w_t$ .

$$w_t = \begin{cases} \log |F_{0,t}| & , F_{0,t} \text{ é pos. definida.} \\ \log |F_{*,t}| + v_t^{(0)} F_{*,t}^{-1} v_t^{(0)} & , F_{0,t} = 0 \end{cases}$$

~~Divide~~ para a ser por  $t=1$   
on  $t=d$ ? (quantas componentes fas mas stoc.)

Obs: Modelo SUTSE pode ser generalizado se o caso q+ componentes

$$y_t = \mu_t + \gamma_t$$

$$y_{1t} = \mu_{1t} + \gamma_{1t}$$

covariâncias entre

Variâncias de tendências

utiliza

Aula passada: modelos linear SUTSE  $\Rightarrow y_t = \mu_t + \varepsilon_t$   $y (p \times 1)$  multivariado

Cada  $y_{it}$  depende de sua componente de seu erro, bem como as componentes.

A dependência vem da estrutura das matrizes varcovas ( $\Sigma_E, \Sigma_\eta$ )

Modelos SUTSE (continuações)

Precisamos calcular algumas variâncias e de parâmetros a ser estimado. Ex: Homogeneidade ( $\Sigma_\eta = q \Sigma_E$ )

É possível generalizar a estrutura SUTSE de forma a incorporar outros componentes, e.g., linear estocástica, sazonalidade e ciclo.

Vamos super ortogonalidade entre componentes de interações diferentes ( $\mu_t \perp \delta_t$ ,  $\mu_t \perp \gamma_t$  etc.)

considere  $p=2$  (p modelo q nível e ciclo, onde nível segue modelo linear estocástico).

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \gamma_{1t} \\ \gamma_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

3 parâmetros

$\downarrow \varepsilon_t \sim N(0, \Sigma_E)$

$$\begin{pmatrix} \mu_{1,t+1} \\ \mu_{2,t+1} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \beta_{1t} \\ \beta_{2t} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

matrizes tan cheias

$\downarrow \eta_t \sim N(0, \Sigma_\eta)$

3 par.

$$\begin{pmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \end{pmatrix} = \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \end{pmatrix} + \begin{pmatrix} \zeta_{1t} \\ \zeta_{2t} \end{pmatrix}, \quad \zeta_t \sim N(0, \Sigma_\zeta)$$

3 par.

$$\begin{pmatrix} \psi_{1,t+1} \\ \psi_{2,t+1} \\ \psi_{1,t+1}^* \\ \psi_{2,t+1}^* \end{pmatrix} = P \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \otimes I_2 \begin{pmatrix} \psi_{1,t} \\ \psi_{2,t} \\ \psi_{1,t}^* \\ \psi_{2,t}^* \end{pmatrix} + K_t \begin{pmatrix} K_{1t} \\ K_{2t} \\ K_{1t}^* \\ K_{2t}^* \end{pmatrix}$$

$\lambda = 2\pi / T_c$  (a mesma mas 2 séries)

$$K_t \sim N(0, \Sigma_K)$$

3 par.

$\downarrow$   
 $2 \times 2$

$$\begin{bmatrix} P[\cdot] & 0 \\ 0 & P[\cdot] \end{bmatrix}_{4 \times 4}$$

$$(\text{pois } \sigma_k^2 = \sigma_{k^*}^2)$$

tilibra

Obs: cada matriz  $\Sigma$  apresenta 3 parâmetros q sejam estimados.

Colocando o modelo de tendência linear estocástica + ciclo

SUTSE na forma EE:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}}_{Z} \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ \beta_{1t} \\ \beta_{2t} \\ \Psi_{1t} \\ \Psi_{2t} \\ \Psi_{1t}^* \\ \Psi_{2t}^* \end{pmatrix} + \begin{pmatrix} E_{1t} \\ E_{2t} \end{pmatrix}$$

importante a ordem  
de entrada do vetor  
de estado.

(Se muda, encontrará  
uma nova forma EE)

$$\begin{array}{c} T \quad R \\ \hline \mu_{1,t+1} & | & 1 & 0 & 1 & 0 & | & \mu_{1t} & | & 1 \\ \mu_{2,t+1} & | & 0 & 1 & 0 & 1 & | & \mu_{2t} & + & 1 \\ \beta_{1,t+1} & = & 0 & 0 & 1 & 0 & | & \beta_{1t} & & 1 & 0 \\ \beta_{2,t+1} & | & 0 & 0 & 0 & 1 & | & \beta_{2t} & & 1 & \\ \Psi_{1,t+1} & | & a & 0 & b & 0 & | & \Psi_{1t} & | & 0 & + \\ \Psi_{2,t+1} & | & 0 & a & 0 & b & | & \Psi_{2t} & & \sim & 1 \\ \Psi_{1,t+1}^* & | & \sim & c & 0 & a & 0 & | & \Psi_{1t}^* & & 1 \\ \Psi_{2,t+1}^* & | & & 0 & c & 0 & a & | & \Psi_{2t}^* & & 1 \\ \hline & & 8x1 & & 8x8 & & & & & & 8x8 & \\ & & & & & & & & & & & I_8 \end{array}$$

$$a = p \cos \lambda_c$$

$$b = p \sin \lambda_c \quad c = -b$$

$$c = -p \sin \lambda_c$$

Considere as matrizes do sistema do modelo univariado  
equivalente:

**tilibra** (É interessante ver uma outra forma de escrever a forma EE,  
baseando nas matrizes do modelo univariado original.

equivalentes p/ o modelo tomado /  
com as mesmas componentes. ( $y_t = \mu_t + \epsilon_t + \varepsilon_t$ )

$$Z_v = (1 \ 0 \ 1 \ 0) \quad T_v = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & a \end{pmatrix} \quad R_v = I_4$$

Então, segue que:

$$- Z_v \otimes I_2 = (1 \ 0 \ 1 \ 0) \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

dimensões p  
(do vetor  $y_t$ )

$= Z$  (matriz  $Z$  multivariada futur)

$$- T_v \otimes I_2 = T \quad (\text{provar!})$$

$$- R = R_v \otimes I_p \underset{\substack{\sim \\ I_4}}{\equiv} I_p$$

E assim podemos escrever, de uma forma geral,

$\forall$  SUTSE como: (qualquer que seja a combinação de componentes do SUTSE)

$$\left\{ \begin{array}{l} y_t = (Z_v \otimes I_p) \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(\lambda, \Sigma_\varepsilon) \\ \alpha_{t+1} = (T_v \otimes I_p) \alpha_t + (R_v \otimes I_p) \eta_t, \quad \eta_t \sim N(0, \Sigma_\eta) \end{array} \right. \Rightarrow \text{Notações} \\ + \text{conota}$$

Observação: Produto de Kronecker

Divida:  
matrix é  
dividida entre  
componentes?

Se  $A \sim m \times n$

$B \sim p \times q$

$$A \otimes B \hat{=} \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix} \sim mp \times rq$$

tilibra

## Fatores Comuns e cointegradas - (apostila)

### 1) Modelos de nível local

Inicialmente considere o SUTSE bivariado.

$$(I) \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \quad \varepsilon_t \sim N(0, \Sigma_\varepsilon) \quad \text{dependência entre} \\ \text{do fato que matrizes} \\ \text{são cheias (nenhuma)} \\ \text{comp. de uma série} \\ \text{"transferida" para} \\ \text{outra)}$$

$$\begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t-1} \\ \mu_{2,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}, \quad \eta_t \sim N(0, \Sigma_\eta) \quad \therefore \sigma_{\eta_1, \eta_2} = \rho_\eta \sigma_{\eta_1} \sigma_{\eta_2}$$

O modelo (I) pode ser resrito como:

$$(II) \begin{aligned} y_{1t} &= \mu_{1t} + \varepsilon_{1t} && \rightarrow \text{sistema (II) é reparametrizado} \\ y_{2t} &= \bar{\mu} y_{1t} + \bar{\mu}_t + \varepsilon_{2t} && \text{do (I) mas preserva invariante} \\ \mu_{1t} &= \mu_{1,t-1} + \eta_{1t} && \text{algumas coisas importantes} \oplus \\ \bar{\mu}_t &= \bar{\mu}_{t-1} + \bar{\eta}_t \end{aligned}$$

$$\text{onde } \bar{\mu} = \rho \sigma_{\eta_2}, \quad \text{var} \begin{pmatrix} \eta_{1t} \\ \bar{\eta}_t \end{pmatrix} = \begin{pmatrix} \sigma_{\eta_1}^2 & 0 \\ 0 & (1-\rho^2) \sigma_{\eta_2}^2 \end{pmatrix} \quad \text{Aqui, a matriz é diagonal} \\ \text{(mas cheia)}$$

Os sistemas (I) e (II) são equivalentes, pois:  $\Rightarrow$  mostram (i) e (ii)

e estas

(i) a estrutura de dependências <sup>dos y's</sup> permanece inalterada <sup>de prova!</sup>

/ /

bivariada.

(ii) a previsões dos  $y_t$ 's e a função de verossimilhança permanecem inalteradas.

(e calcularmos forma reduzida, FAC etc.  $\Rightarrow$  tudo será igual).

Essencialmente, o sistema (II) é uma nova parametrização do sistema (I), obtida por uma transformação do tipo:

$$\alpha_t = B \alpha_t^*, \quad \det(B) \neq 0.$$

(transformações nas singular do vetor de estado)

Prova:

$$\alpha_t = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ \pi & 1 \end{pmatrix}}_B \begin{pmatrix} \mu_{1t} \\ \bar{\mu}_t \end{pmatrix}$$

(I)

B

(II)

Através de uma nova parametrização fica + explícita a identificação de componentes comuns.

$$(\det B = 1 \neq 0)$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \pi & 1 \end{pmatrix} \begin{pmatrix} \mu_{1t} \\ \bar{\mu}_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$\text{obs: } \alpha_t \text{ (I)} \quad \mu_{2t} = \mu_{2,t-1} + \eta_{2t}$$

Nas notas, assim está definido  $\bar{\eta}_t$ .

$$\text{mas } \bar{\eta}_t = \eta_{2t} - \pi \eta_{1t} \quad \therefore \eta_{2t} = \bar{\eta}_t + \pi \eta_{1t}$$

mostramos que  $\bar{\eta}_t$  é o que  
varia ( $\frac{\eta_{1t}}{\eta_{2t}}$ ) + diafones

$$\text{logo: } \mu_{2t} = \mu_{2,t-1} + \bar{\eta}_t + \pi \eta_{1t}$$

do modelo (I)

$$\therefore \mu_{2t} = \mu_{2,t-1} + \bar{\eta}_t + \pi (\mu_{1t} - \mu_{1,t-1})$$

$$\therefore \mu_{2t} - \pi \mu_{1t} = (\mu_{2,t-1} - \pi \mu_{1,t-1}) + \bar{\eta}_t$$

$$\therefore \bar{\mu}_t = \bar{\mu}_{t-1} + \bar{\eta}_t$$

tilibra

$$\bar{\mu}_t = \mu_{2,t} - \pi \mu_{1,t} \quad ; \quad \mu_{2,t} = \bar{\mu}_t + \pi \mu_{1,t}$$

↓

usando as linhas componentes

dos 2 sistemas (I) e (II)

$$\therefore \mu_{2,t} = (\pi \quad 1) \begin{pmatrix} \mu_{1,t} \\ \bar{\mu}_t \end{pmatrix}$$

$$\text{Logo: } \begin{pmatrix} \mu_{1,t} \\ \bar{\mu}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \pi & 1 \end{pmatrix} \begin{pmatrix} \mu_{1,t} \\ \bar{\mu}_t \end{pmatrix}$$

daí:

$$\text{do sistema (I): } \mu_t = \mu_{t-1} + \eta_t$$

$$\text{Da transformada: } \bar{\mu}_t = B \bar{\mu}_{t-1} \quad ; \quad \bar{\mu}_t = B^* \bar{\mu}_t$$

$$\Rightarrow \bar{B}^* \bar{\mu}_t = B^* \bar{\mu}_{t-1} + B^* \eta_t$$

$$\bar{\mu}_t = \bar{\mu}_{t-1} + \eta_t^*$$

$$\text{neste caso: } B^* = \begin{pmatrix} 1 & -\pi \\ 0 & 1 \end{pmatrix}$$

$$\text{Com o } B^* \Rightarrow \text{obtemos } \bar{\eta}_t = \eta_{2,t} - \pi \eta_{1,t}$$

$\Rightarrow$  A parametrização (II) nos permite investigar

copia

correlações entre os erros das 2 tendências.

(pg 10 das notas)

Se  $\rho = \pm 1 \Rightarrow$  Existiria uma componente comum necessariamente.

$$\bullet \text{ Se } \rho_{\eta} = \pm 1 \Rightarrow \text{var}(\bar{\eta}_t) = (1 - \rho^2) \sigma_{\eta_2}^2$$

$$\therefore \text{var}(\bar{\eta}_t) = 0$$

(cond. suficiente p/ que haja comp. de tendência comum)

$\hookrightarrow \bar{\mu}_t$  será constante

$$\Rightarrow \bar{\mu}_t = \bar{\mu}_{t-1} = \bar{\mu} + t$$

E, assim, teremos apenas 1 componente comum, pois:

$$y_{1t} = \mu_t^+ + \varepsilon_{1t}$$

$$y_{2t} = \pi \mu_t^+ + \bar{\mu} + \varepsilon_{2t}$$

$$\mu_t^+ = \mu_{t-1}^+ + \eta_{t-1}^+, \quad \eta_t^+ \sim N(0, \sigma_{\eta_1}^2)$$

existe componente comum. em  $y_{1t}, y_{2t}$

$$\bullet \text{ Se } \pi = 1 \text{ (dado que } \rho_{\eta} = 1)$$

$\hookrightarrow$  se pode ser +, mas  $\pi = \rho \frac{\sigma_{\eta_2}}{\sigma_{\eta_1}}$

Teremos

$$\text{moldo chomada} \quad y_{1t} = \mu_t^+ + \varepsilon_{1t}$$

$$\text{de balanced growth} \quad y_{2t} = \mu_t^+ + \bar{\mu} + \varepsilon_{2t}$$

$$= \mu_t^+ + \varepsilon_{2t} \quad \rightarrow \mu_t^+ - \bar{\mu} = \bar{\mu} : \text{diferença entre as tendências é constante} + t$$

$\Rightarrow$  "balanced growth"

$$\bullet \text{ Se } \bar{\mu} = 0 : \mu_t^+ = \mu_t$$

$\Rightarrow$  tendências idênticas

- - -

Nível comum ( $\Leftrightarrow$ ) co-integração

(Veremos def. geral p/ co-integrações e vamos apresentar p/ este caso <sup>especial</sup> particular),

tilibra

## Definições de co-integradas:

Um processo estocástico p-variado  $y_t = (y_{1t}, y_{2t}, \dots, y_{pt})'$

é dito co-integrado de ordens  $d$  e  $b$ ,  $b \leq d$ , i.e.

↳ números inteiros.

$y_t \sim CI(d, b)$  se

cada um dos elementos é integrado da mesma ordem.

(i)  $y_{jt} \sim I(d)$ ,  $j = 1, 2, \dots, p$  (ordem: n.º de vezes que precisa diferenciar para tornar estacionária)

(ii)  $\exists \alpha \in \mathbb{R}^p - \{0\}$  tal que  $\alpha'y_t \sim I(d-b)$  ( $\Delta^d y_t$  estac.)  
(Nem todos os elementos de  $\alpha = 0$ )  $(d=0) \sim I(0)$

Exemplo: Se  $y_{jt} \sim I(1)$  e são CI, então  $\exists \alpha'y \sim I(0)$

Combinação linear prévia  
de ordens de  
integridade  
que cada série individualmente.

$$d=1 \\ b=1$$

↳ existe comb. linear  
destas séries que será  
estacionária.

## Intuições:

Séries não estacionárias não tem achatos

Mas há muitas dicas que podem variar de forma não estacionária, mas podem estar em sincronia, mas divergir muito (andam ± em fase) por conta de forças (econômicas, etc.) comuns que atuam sobre elas.

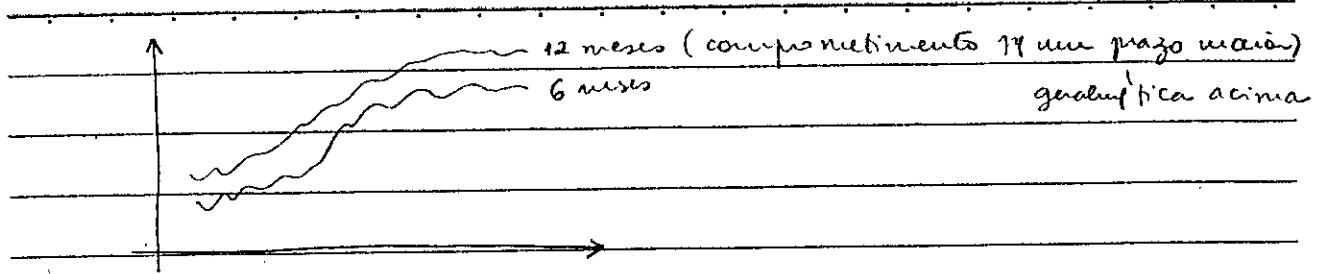
Serão  $p$  séries multivariadas, não estacionárias.

(No caso de estação não se aplica este conceito. Nesse caso são correlacionadas ou não).

Ex. de séries macroeconómicas co-integradas

- títulos de dívida pública e vencimentos +

(ambas variam de acordo com estruturas de juros).



Elas são co-integradas pois faz E(.) de uma série de taxas comuns.

No processo formador desta estrutura a termo há um fator comum e relevante que afeta as séries.

Ex. Relações entre consumo e renda agregada.

Tb por razões macroeconómicas espera-se que sejam co-integradas

Ex. Petróleo em várias maturidades

Petróleo e seus derivados.

O fato de serem co-integradas é a existência de um fator comum (ex: componente comum, como tendências por exemplo).

Voltando ao modelo:

$$\begin{cases} y_{1t} = \mu_t^+ + \epsilon_{1t} \\ y_{2t} = \pi \mu_t^+ + \bar{\mu} + \epsilon_{2t} \\ \mu_t^+ = \mu_{t-1}^+ + \eta_{t-1}^+ \end{cases}$$

$$\Rightarrow \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_t^+ \\ \pi \mu_t^+ + \bar{\mu} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \quad \rightarrow \text{cada } y_{jt} \text{ é um RW (I(1))}$$

Se  $y_{1t}$  e  $y_{2t}$  sas CI  $\Rightarrow \alpha_1 y_{1t} \sim I(0)$  estacionário

$y_{1t} \sim I(1)$  (trivial)

$$(\alpha_1 \alpha_2) \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \alpha_1 y_{1t} + \alpha_2 y_{2t} = \alpha_1 \mu_t^+ + \alpha_2 \pi \mu_t^+ + \alpha_2 \bar{\mu} + (\alpha_1 \epsilon_{1t} + \alpha_2 \epsilon_{2t}) \sim I(1) ?$$

sas arbitrárias

Se conseguirmos escolher  $\alpha_1$  e  $\alpha_2$  tal que srie seja  $I(0)$

$\Rightarrow$  provaremos que sas co-integradas.

Escolhendo  $\alpha_1 = -\pi$  } conveniente de forma a anular  
 $\alpha_2 = 1$  processo  $\pi$  estacionário  $\mu_t^+$

$$\Rightarrow (-\pi + 1) \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = -\pi \mu_t^+ + \pi \mu_t^+ + \bar{\mu} + (-\pi \epsilon_{1t} + \epsilon_{2t})$$
$$= \bar{\mu} + \epsilon_t \sim I(0)$$

$\downarrow$   
sries sas co-integradas

pois cl tem ordem menor

do que cada uma separadamente

$\Rightarrow$  conclusao:  $\rho = 1$  e condicões suficiente p/ existéncia  
de componente comum  
(sries picam co-integradas)

( $\pi = 1$  mas é necessário. Ele torna a diferença entre  
tendências cl.)

Generalizações do que vimos até o momento  
bivariado. Agora  $y_t$  é p-variado

1 / 1

1) Modelos de Nível Local Geral  $\Rightarrow y_t \sim p \times 1$ : quebramos em  
particais  $\begin{cases} y_{1t} \rightarrow k \times 1 \\ y_{2t} \rightarrow n \times 1 \end{cases}$

$$(I) \quad y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma_\varepsilon)$$

$$\text{SUTSE} \quad \mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim N(0, \Sigma_\eta)$$

p-variado  
de nível  
local.

considere a seguinte partição para  $\Sigma_\eta = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$\begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t-1} \\ \mu_{2,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \Rightarrow \text{matriz}$$

onde  $n = p - k$

de var covar  $\Sigma_\eta$

também sofrerá partição  
como consequência.

$$\text{Seja } L \equiv \begin{pmatrix} I_k & 0 \\ -\Pi & I_n \end{pmatrix} \Rightarrow \text{Transformações nas níveis}$$

(propriedades fundamentais  
permanecem invariantes)

$$\hookrightarrow n \times k, \quad \Pi = \Sigma_{22} \Sigma_{11}^{-1}$$

Usamos novamente  
esta em particular

$$\therefore L \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} = L \begin{pmatrix} \mu_{1,t-1} \\ \mu_{2,t-1} \end{pmatrix} + L \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

Para resolver e  
avaliar a questão  
de componentes  
comuns,  
correlação

$$\begin{pmatrix} I & 0 \\ -\Pi & I \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{2t} \end{pmatrix} = \begin{pmatrix} I & 0 \\ -\Pi & I \end{pmatrix} \begin{pmatrix} \mu_{1,t-1} \\ \mu_{2,t-1} \end{pmatrix} + \begin{pmatrix} I & 0 \\ -\Pi & I \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

$$I \mu_{1t} = I \mu_{1,t-1} + I \eta_{1t} \Rightarrow \mu_{1t} = \mu_{1,t-1} + \eta_{1t}$$

$$-\Pi \mu_{1t} + I \mu_{2t} = -\Pi \mu_{1,t-1} + I \mu_{2,t-1} - \Pi \eta_{1t} + \eta_{2t}$$

$$\therefore \underbrace{\mu_{1t} - \Pi \mu_{1t}}_{\bar{\mu}_t} = \underbrace{\mu_{2,t-1} - \Pi \mu_{2,t-1}}_{\bar{\mu}_{2,t-1}} + \underbrace{(\eta_{1t} - \Pi \eta_{1t})}_{\eta_{1t}}.$$

equilíbrio

$$\bar{\mu}_t = \bar{\mu}_{t-1} + \bar{\eta}_t$$

E assim segue que:

$$L(\mu_t) = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix} \Rightarrow L\mu_t = \mu_t^*$$

Mostrar-se que  $\text{cov} \begin{pmatrix} \eta_{1t} \\ \eta_t \end{pmatrix} = \begin{pmatrix} \bar{\Sigma}_{\eta_1} & 0 \\ 0 & \bar{\Sigma}_{\eta} \end{pmatrix}$

↓  
provar!

$$\text{onde } \bar{\Sigma}_{\eta} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$

Por outro lado,

$$y_t = \mu_t + \varepsilon_t$$

$$y_t = \underbrace{L^{-1}L}_{\mu_t^*} \mu_t + \varepsilon_t$$

$$\therefore y_t = L^{-1} \mu_t^* + \varepsilon_t$$

$$\therefore \left\{ \begin{array}{l} y_t = \begin{pmatrix} I_K & \Pi \\ 0 & I_n \end{pmatrix} \begin{pmatrix} \mu_{1t} \\ \bar{\mu}_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \Rightarrow \text{E o (II) } \text{aparametrizado} \end{array} \right.$$

$$\left( \begin{matrix} \mu_{1t} \\ \bar{\mu}_t \end{matrix} \right) = \left( \begin{matrix} \mu_{1,t-1} \\ \bar{\mu}_{t-1} \end{matrix} \right) + \left( \begin{matrix} \eta_{1t} \\ \bar{\eta}_t \end{matrix} \right), \quad y_t = \begin{pmatrix} \eta_{1t} \\ \bar{\eta}_t \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Sigma}_{\eta_1} & 0 \\ 0 & \bar{\Sigma}_{\eta} \end{pmatrix} \right)$$

$$\bar{\Sigma}_{\eta} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$

30.103/11

$$(I) \quad y_t = \alpha_t + \epsilon_t \quad \epsilon_t \sim N(0, H_t)$$

$$(II) \quad \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \eta_t \sim N(0, Q_t)$$

$m \times n \approx x_1$

Passo da Previsão:  $\hat{\alpha}_{t+1} = E[\alpha_{t+1} | Y_t] = T_t \hat{\alpha}_t | t$  Harvey  
 (da aula) Notações do Koopman

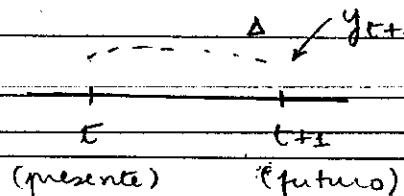
$$\begin{aligned} P_{t+1} &= \text{Var}(\alpha_{t+1} | Y_t) = \\ &= T_t P_t | t T_t' + R_t Q_t R_t' \end{aligned}$$

Queremos  $\hat{\alpha}_{t+1}$  e  $P_{t+1}$

↓      ↓

$E[\alpha_t | Y_t]$      $\text{Var}(\alpha_t | Y_t)$  (matriz de var covar)

A situação na prática é:



O passo  $E[\alpha_t | Y_t]$  é a revisão em  $t+1$  das previsões que havia sido feita

- |  $(\alpha_t | Y_{t-1})$ : estamos em  $t-1$  e queremos prever  $y_t$
- |  $(\alpha_t | Y_t)$ : chega info  $y_t$  e queremos revisar a previsão passada.

Queremos epis que linkem as 2 funções dist. de prob. dos 2 períodos.

obs: só olhamos média e variância pois estamos trabalhando com ambiente Gaussiano. Se não tivéssemos Gaussianidade, teríamos que recuperar a fdp a cada instante.

11

base p/ nossas deduções

Resultados (da aula passada):

$$(1) \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \sim N \begin{pmatrix} \mu_X \\ \mu_Y \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} & \Sigma_{XZ} \\ \Sigma_{YX} & \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZX} & \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}$$

$X, Y, Z$ : vetores

de dimensões

arbitraria

Supor: (i)  $\mu_Z = 0$

(ii)  $\Sigma_{2y} = 0 \Rightarrow \Sigma_{yz} = 0$

Dado isso, prova-se que:

$$E[x | y=y, z=z] = \mu_x|_y + \Sigma_{xz} \Sigma_{zz}^{-1} z \quad (a)$$

$$\text{var}[x | y=y, z=z] = \Sigma_{xx} - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx} \quad (b)$$

(2) Se  $X$  e  $Y$  são dois vetores aleatórios e

$q_{x1} \quad p_{x1}$

$g: \mathbb{R}^p \rightarrow \mathbb{R}^p$  (transf. bijetiva - existe inversa:  $\exists g^{-1}$ )

cond. necessária e  
suficiente

então  $F_{x|y} = F_{x|g(y)}$

$$\Rightarrow E[x | y=y] = E[x | g(y) = g(y)]$$

Calculando

descompondo em  $y_{t-1}$  e  $y_t$

$$\Rightarrow E[x_t | y_t] = E[x_t | y_{t-1}, y_t] =$$

$$= E[x_t | y_{t-1}, v_t]$$

onde  $v_t \stackrel{\text{def}}{=} y_t - E[y_t | y_{t-1}]$

última

inovação (modo revisar 1 passo a frente)

$$\bullet \text{var}(v_t) = E[(v_t - E[v_t])(v_t - E[v_t])'] = E[v_t v_t']$$

$$\text{mas } E[v_t v_t'] = E[E[v_t v_t' | Y_{t-1}]]$$

$$\text{Sabemos que } v_t = y_t - z_t \alpha_t$$

$$= z_t \alpha_t + \epsilon_t - z_t \alpha_t$$

$$= z_t (\alpha_t - \hat{\alpha}_t) + \epsilon_t$$

$$\therefore v_t' = (\alpha_t - \hat{\alpha}_t)' z_t' + \epsilon_t'$$

$$\Rightarrow E[v_t v_t' | Y_{t-1}] = E[(z_t (\alpha_t - \hat{\alpha}_t) + \epsilon_t) ((\alpha_t - \hat{\alpha}_t)' z_t' + \epsilon_t') | Y_{t-1}]$$

$$= E[z_t (\alpha_t - \hat{\alpha}_t) (\alpha_t - \hat{\alpha}_t)' z_t' | Y_{t-1}] + E[\epsilon_t \epsilon_t' | Y_{t-1}]$$

✓

(termos cruzados se anulam)

$$\text{Ex: } E[z_t (\alpha_t - \hat{\alpha}_t) \epsilon_t' | Y_{t-1}] =$$

$$= E[z_t \alpha_t \epsilon_t' | Y_{t-1}] - E[z_t \hat{\alpha}_t \epsilon_t' | Y_{t-1}]$$

$$= z_t E[\alpha_t \epsilon_t' | Y_{t-1}] - z_t \hat{\alpha}_t E[\epsilon_t' | Y_{t-1}]$$

↓                      ↓                      ↓  
se iterarmos      dado passado      ε conhecido  
ε seja função      de α, e η<sub>t-g</sub> ⇒ p

obj. A cond. inicial  $\alpha_0 \sim N(\alpha_0, P_0)$  descrevem os dados que

faz com que termos cruzados se anulam

$$\therefore E[v_t v_t' | Y_{t-1}] = z_t P_t z_t' + H_t$$

(valor condicional)

↓  
matriz de var/covar estimada no passo anterior

$$\text{Logo: } E[v_t v_t'] = E[E[v_t v_t' | Y_{t-1}]] \Rightarrow E[v_t v_t'] = z_t P_t z_t' + H_t \stackrel{?}{=} F_t$$

. matriz var.-covar. da inovação  
ilíbia

$$\circ \text{cov}[\alpha_t, v_t] = M_t$$

Da definição:

$$\begin{aligned} M_t &= E[(\alpha_t - E[\alpha_t])(v_t - E[v_t])^T] \\ &= E[\alpha_t v_t^T] - E[\alpha_t] \cdot E[v_t]^T = \\ &= E[\alpha_t v_t^T] \xrightarrow{\text{lei das exp. iteradas}} \\ &\quad = E[E[\alpha_t v_t^T | y_{t-1}]] \end{aligned}$$

$$\text{Calculemos } \alpha_t v_t^T = \alpha_t [(\alpha_t - a_t)' z_t^T + \epsilon_t^T]$$

calculado no passo anterior

$$\text{Logo: } E[\alpha_t v_t^T | y_{t-1}] = E[\alpha_t (\alpha_t - a_t)' z_t^T | y_{t-1}] + E[\alpha_t \epsilon_t^T | y_{t-1}]$$

iterado:  $f(\alpha_t)$

$$= E[(\alpha_t - a_t)(\alpha_t - a_t)' z_t^T | y_{t-1}]$$

$$\text{obs: } (E[\alpha_t \alpha_t^T | y_{t-1}] - E[\alpha_t | y_{t-1}] a_t^T) z_t^T \xrightarrow{\text{é a m/maco}}$$

$$(E[\alpha_t \alpha_t^T | y_{t-1}] - a_t a_t^T) z_t^T$$

$$\therefore E[\alpha_t v_t^T | y_{t-1}] = E[(\alpha_t - a_t)(\alpha_t - a_t)' | y_{t-1}] z_t^T =$$

$$= P_t z_t^T$$

$$\text{Portanto: } M_t = E[E[\alpha_t v_t^T | y_{t-1}]] = P_t z_t^T$$

Usando os resultados obtidos p/  $\text{var}(v_t)$  e  $\text{cov}(\alpha_t, v_t)$ , podemos escrever:

(1)

$$E[\alpha_t | Y_t] = \hat{\alpha}_t + M_t F_t^{-1} v_t \quad \text{onde } M_t = P_t Z_t'$$

$$F_t = Z_t P_t Z_t' + H_t$$

$$v_t = y_t - Z_t \hat{\alpha}_t$$

Agora queremos:

$$\text{var}[\alpha_t | Y_t] = ?$$

$$\text{Tínhamos feito } x = \alpha_t, y = y_{t+1}, z = v_t$$

⇒ Usando a fórmula (b):

$$\text{var}[\alpha_t | Y_t] = \text{var}[\alpha_t | Y_{t-1}, v_t]$$

$$= \text{var}[\alpha_t | Y_{t-1}] - \text{cov}(\alpha_t, v_t) \text{var}(v_t) \text{cov}(\alpha_t, v_t)$$

$$\therefore P_{t|t} = P_t - M_t F_t^{-1} M_t'$$

$$= P_t - M_t F_t^{-1} (P_t Z_t')' =$$

$$= P_t - P_t Z_t' F_t^{-1} Z_t P_t$$

$$\Rightarrow P_{t|t} = P_t - P_t Z_t' F_t^{-1} Z_t P_t$$

e' simétrica

Resumindo:

Equações do FK

$$\text{Previsões: } \hat{\alpha}_{t+1} = T_t \hat{\alpha}_t$$

$$P_{t+1} = T_t P_{t|t} T_t' + R_t Q_t R_t'$$

$$\text{onde: } v_t = y_t - Z_t \hat{\alpha}_t$$

$$\text{Atualizações: } (\hat{\alpha}_{t|t}) = (\hat{\alpha}_t) + M_t F_t^{-1} \xrightarrow{(y_t - Z_t \hat{\alpha}_t)} v_t$$

$$F_t = Z_t P_t Z_t' + H_t$$

$$P_{t|t} = P_t - P_t Z_t' F_t^{-1} Z_t P_t$$

↑      ↓  
atualizada    prevista    correcções

equilíbrio

na prática temos uma previsão sequencial

$x_t \sim w(a_t, p_t)$ : dist. a priori inicial  $\Rightarrow$  é o passo de previsão inicial  
(conhecido)

Começamos na atualização usando  $y_t$  ( $1^{\text{a}} \text{ obs}$ ).

Com  $a_{t|t}$  e  $P_{t|t}$ , jogamos nas fórmulas da previsão.

Se rescrevermos as fórmulas em funções apenas de  $a_t$  e  $a_{t+1}$  ( $p_t$  e  $P_{t+1}$ ), rest.  $a_{t|t}$  e  $P_{t|t}$  das eq. de atualização nas eqs. de previsão, temos:

$$F_t \text{ é em } 1 \Rightarrow a_{t+1} = T_t(a_t + M_t F_t^{-1} v_t)$$

(passo da

atualização  
desaparece)  $\therefore a_{t+1} = T_t a_t + \underbrace{T_t M_t F_t^{-1} v_t}_{K_t}$  onde  $K_t$ : ganho de Kalman

$$\therefore a_{t+1} = T_t a_t + K_t v_t$$

observações entram aqui

fazendo pf:  $P_{t+1} = T_t P_t L_t^{-1} + R_t Q_t R_t^{-1}$  onde  $L_t = T_t - K_t z_t$

$\downarrow$  não depende das observações

Dependendo do caso, converge para

um valor. Se for assim, podemos

trabalhar com  $\bar{P}$  (steady state value)  
e paramos de computar.

(11)

Exemplo:

$$\begin{cases} y_t = \mu_t + \epsilon_t & (m=p=1) \\ \mu_{t+1} = \mu_t + \eta_t \end{cases}$$

des:  $E[y_t | Y_{t-1}] =$   
 $= E[\mu_t + \epsilon_t | Y_{t-1}] =$   
 $= \mu_t$

$$\begin{cases} z = T = R = 1 \text{ (escalares)} \\ H = \sigma_e^2 \\ Q = \sigma_\eta^2 \end{cases}$$

(\*)

•  $K_t = T_t M_t F_t^{-1} = P_t F_t^{-1} = \frac{P_t}{F_t}$

$\downarrow P_t z_t$        $F_t - \text{escalar}$

$\therefore K_t = \frac{P_t}{P_t + \sigma_e^2} = 0 < K_t < 1$

•  $v_t = y_t - z_t a_t = y_t - a_t$

$$\begin{aligned} a_{t+1} &= a_t + K_t v_t \\ &= a_t + K_t(y_t - a_t) \\ &= a_t + K_t y_t - K_t a_t \end{aligned}$$

$\therefore a_{t+1} = (1 - K_t)(a_t) + K_t y_t$

$\rightarrow E[y_t | Y_{t-1}]$  (\*)

$\hat{y}_{t+1 t} = K_t y_t + (1 - K_t) \hat{y}_{t t-1}$	$\rightarrow \text{EWMA}$
---	---------------------------

FK está fazendo Ajustamento Exponencial

No caso clássico:

$$\hat{y}_{t+1|t} = \lambda y_t + (1 - \lambda) \hat{y}_{t|t-1} \quad 0 < \lambda < 1$$

fazendo  $w_j = \lambda(1-\lambda)^j \Rightarrow \sum_{j=0}^{T-1} w_j^j y_{t-j}$ : média ponderada <sup>Wibra</sup>  
 nesse caso

Quando adotamos modelos de nível local, a função de previsões é um EWMA

Na hora que estimamos  $\sigma^2_\epsilon$  temos um  $\sigma^2_\epsilon$  ótimo.

Os pesos do EWMA estarão otimizados.

Estes métodos garantem a optimidade da escolha dos parâmetros.

Obs: Também  $y_t = \mu_t + \gamma_t + \epsilon_t$

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t$$

$$\beta_{t+1} = \beta_t + \xi_t$$

$$\gamma_t = \sum_{j=1}^{12} \gamma_{t+j} + w_t$$

Modelo

(

de previsões

sazonalidade

usando fórmula convencional

pf Holt Winters

Como estamos trabalhando com modelos lineares,

todos os métodos acabarão trabalhando com médias das observações passadas.

Os pesos estarão sendo objetivamente calculados de forma ótima (parâmetros serão estimados por MV).

Sob Normalidade:

usamos média central como estimador de  $\mu$   
→ (= moda = mediana)

Usualmente fazemos isso, pois a média tem  
uma propriedade ótima ⇒ minimiza o erro quadrático  
mínimo

(1/1)

$\hat{\alpha}_{t+1|t}^*$   $\Rightarrow$  estimativa

$$\text{Queremos } \min_{\hat{\alpha}_{t+1|t}^*} E \left[ (\alpha_{t+1} - \hat{\alpha}_{t+1|t}^*) (\alpha_{t+1} - \hat{\alpha}_{t+1|t}^*)^\top \mid Y_t \right]$$

$\downarrow \text{MSE}(\hat{\alpha}_{t+1|t}^*)$

$$\Rightarrow \hat{\alpha}_{t+1|t}^* = E[\alpha_{t+1} \mid Y_t]$$

$$= \hat{\alpha}_{t+1}$$

O estimador que minimiza o MSE é a média condicional (sob norma lida).

### Nas gaussianas

(Se usarmos as primeiras obtidas nessa situação,  
o que serão os resultados que obtivemos?)

$\hat{\alpha}_{t+1}$ : melhor estimador linear de  $\alpha_{t+1}$

$$\hat{\alpha}_{t+1|t} = \sum_{j=0}^{t-1} w_j y_{t-j}$$

queremos que este estimador linear seja ótimo no sentido de min. MSE

Se fizermos isso, obtemos o equivalente ao FK que calculamos.

Nas sarà a média, mas terá uma característica de optimalidade local.

(sob gaussianidade, temos o melhor estimador entre os lineares e nas lineares, como vimos)

Na prática: (para implementar o FK)

$$\begin{aligned} \sigma_e^2 &= \langle \cdot \rangle & a_i = \langle a_i \rangle & \left. \right\} \text{valores dados} \\ \sigma_n^2 &= \langle \cdot \rangle & p_i = \langle p_i^+ \rangle & \end{aligned}$$

$$\text{for } t=1, n \quad \Rightarrow \quad a_{t+1} = a_t + k_t v_t$$

$$v_t = y_t - \alpha_t \quad p_{t+1} = p_t (1 - K_t) + \sigma_\eta^2$$

$$K_t = \frac{P_t}{P_t + \sigma^2 \epsilon}$$

continue (loop ate'n)

$$F_t = P_t + \sigma_{\varepsilon}^2$$

Temos uma ST qualquer e queremos modelá-la

Começa q a, e p;<sup>\*</sup>

Calcular  $V_1$ ,  $K_1$

Calcular  $a_1$  e  $b_1$  ) loop  $n \times$  (teremos  $n$  valores)

Mas precisamos saber com que valores entramos p/  
parâmetros fixos

Tenemos una foto de Verasimilhanca:

$$l_t = l - 1 \left( \ln f_t + v_t^2 / f_t \right)$$

(começa q  $\ell=0$  e vai acumulando em  $\ell$ )

Y en ótimo, daremos la máxima.

Temos que ter um algoritmo que varia

os parâmetros fixos para otimizar  $\ell \Rightarrow$  gera novas

In this case  $\sigma_{\varepsilon}^2 < \sigma_{\eta}^2$ . values.

$$\text{Ex: } \Psi^{(i+1)} = \Psi^{(i)} - \lambda H^{-1}(\Psi^{(i)}) g(\Psi^{(i)}). \quad \therefore \text{ forma itera-}$$

Heriano      gradiente      +va p/ gerar  
 $\frac{\partial \Psi}{\partial x_2}$        $\frac{\partial \Psi}{\partial x_1}$       novo valor

(11)

Iterações (i)       $\ell^{(i)}$ 0                   $\ell^{(0)}$ 1                   $\ell^{(1)}$ 2                   $\vdots$ 

I

$\Rightarrow$  a cada iteração temos que checar as condições de optimização  
 (gradiente = 0 e Hessiana)

Difícil: implementar um algoritmo que converja para valores ótimos dos parâmetros.

de otimização

Três se problema tiver restrições:

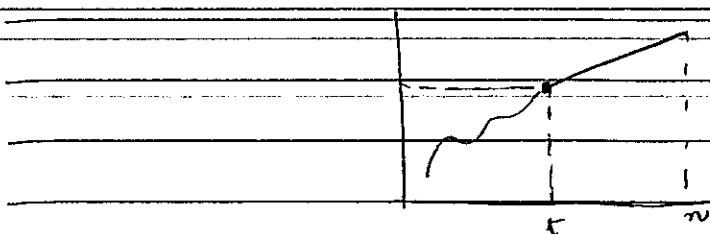
- . se  $\sigma_e^2$  sempre  $> 0 \Rightarrow$  podemos fazer  $\sigma_e^2 = e^\theta$  e otimizar por  $\theta$ .

— u —

Estimativa suavizada ou alisada do estado  
 (smoothing)

$$f(x_t | Y_n) \sim N(\hat{x}_t | V_t) \quad \rightarrow \text{notas do koopman}$$

$\downarrow$  fdp de  $x_t$  dado o conjunto de observações completo



Notas:  $y = Y_n$  (do koopman)

$$= (y_1, y_2, y_3, \dots)$$

(1) → média suavizada do estado

$$\hat{x}_t = E[x_t | y] = E[x_t | y_{t-1}, y_t, y_{t+1}, \dots, y_n] \quad \left. \begin{array}{l} \text{como fizemos} \\ \text{em previsões} \\ \text{atualizadas} \end{array} \right\}$$
$$= E[x_t | y_{t-1}, v_t, v_{t+1}, \dots, v_n]$$

Temos o resultado:

$$E[X | y = y, z = z] = E[X | y] + \sum_{x,z} \sum_{z,z}^1 z$$

Faremos:  $x = x_t$

erro de previsões um passo a frente

$$y = y_{t-1}$$

$$z = (v_t, v_{t+1}, \dots, v_n) \rightsquigarrow \text{dimensão: } (n-t+1) p \times 1$$

$$\therefore E[x_t | y] = E[x_t | y_{t-1}] + \text{cov}(x_t, (v_t, v_{t+1}, \dots, v_n)') \cdot \text{var}(v_t, \dots, v_n)$$

$$\begin{pmatrix} v_t \\ v_{t+1} \\ \vdots \\ v_n \end{pmatrix}$$

(Ainda no passo de atualização)

(06/04/11)

Equações de Riccati (steady state KF)Observar a equação para var( $\alpha_{t+1}, \gamma_t$ ) =  $P_{t+1}$ 

$$P_{t+1} = T_t P_t L_t' + R_t Q_t R_t' \rightarrow F_K \text{ é em } \perp \text{ (embutindo passo de previsão)}$$

sa matriz  
as é adaptada  
à novas obs.  
côs (apenas  
meus dias).  
por conta  
da norma  
lidade.

Considere sistemas invariantes no tempo (matrizes do sistema independem de t)

$$P_{t+1} = T P_t L_t' + R Q R'$$

$$L_t = T - K_t Z = T - (T P_t Z' F_t^{-1}) Z \quad (K \text{ e } K_t \text{ não depende de } P_t)$$

$$F_t = Z P_t Z' + H$$

desta forma:

$$P_{t+1} = T P_t [T - (T P_t Z' F_t) Z] + R Q R'$$

$$\therefore P_{t+1} = T P_t T' - T P_t Z' F_t^{-1} P_t T' + R Q R' \quad \begin{array}{l} \text{chamada} \\ \text{equações de} \\ \text{Riccati} \end{array}$$

Suponha que lim<sub>t→∞</sub>  $P_{t+1} = P_t = \bar{P}$  → computacionalmente, nos passos são recalculados (mais eficiente)

$$\bar{P} = T \bar{P} T' - T P_t Z' F_t^{-1} \bar{P} T' + R Q R \quad \text{e} \quad \bar{F} = Z \bar{P} Z' + H$$

resolver para  $\bar{P}$ equação matricial  $\bar{P} = \bar{P}$ 

$$\text{Exemplo: } y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

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

obs: Existem algumas cond. p/ qd. de Riccati ter soluções. No nosso caso só nos preocupamos com a checagem.

$$T = Z = 1$$

$$\therefore \bar{P} = \bar{P} - \bar{P} \bar{F}^{-1} \bar{P} + \sigma_\eta^2$$

$$= \bar{P} - \frac{\bar{P}^2}{\bar{F}} + \sigma_\eta^2 \Rightarrow \bar{P} = \bar{P} \left( 1 - \frac{\bar{P}}{\bar{P} + \sigma_\epsilon^2} \right) + \sigma_\eta^2$$

$$\Rightarrow x^2 - xh - h = 0 \quad \text{onde } x = \bar{P}/\sigma_\epsilon^2$$

fazer  
sinal ruí  
signal to  
noise ratio

$$h = \sigma_\eta^2 / \sigma_\epsilon^2 = q$$

$$\text{soluções: } x = \frac{h \pm \sqrt{h^2 + 4h}}{2}, \text{ desde que } h > 0 \text{ ou }$$

$\rightarrow z_t$  depende do tempo

$$\text{obs: } y_t = y_{t-1} \varphi_t + \varepsilon_t \quad \text{Jamais haverá}$$

$$\varphi_t = \varphi \varphi_{t-1} + \eta_t \quad \text{soluções de Riccati}$$

(ver lista)

Obs 2: computacionalmente

A cada iteração, verificaremos se diferença entre  $P_t$  e  $P_{t+1}$

é menor que um certo valor (definir desigualdade pp matriz).

Se for assim pp os n primeiros casos  $\Rightarrow$  abandona cálculo de  $P_{t+1}$

(Ao invés de calcular se há solução da eq. de Riccati)

Passo de suavizações (ver folha)

O que vimos até agora:

$$\hat{\alpha}_j = E[\alpha_t | \tilde{y}_j]$$



at  $\rightarrow j = t-1$

at/b  $\rightarrow j = t$

Agora, vamos expandir o conjunto de observações  $\tilde{y}$  o qual tomamos a esperança.

Faremos  $j = n$

(tria passado,  $t-1$ , presente,  $t$ , e futuro, ate' n)

Existem vários tipos de smoothing.

Veamos um desenvolvido por Koopmans.

(obs.: Na literatura de Box, os livros que dão origem à Teoria de Espaço de Estado: Anderson & Moore; Jagwinsky.

obs 1:  $p$  reúpciona que  $y_{t+1}$  é independente

temos que olhar  $p$  diff. conjunta.

Como é serie de tempo

$$p(y_1, \dots, y_n) = \prod_{t=1}^n p(y_t | \tilde{y}_{t-1}), \quad p(y_1 | y_0) = p(y_1)$$

$$\text{mas } y_t = z_t \alpha_t + \epsilon_t$$

$$p(y_t | \tilde{y}_{t-1}) = \int_{-\infty}^{\infty} p(y_t, \alpha_t | \tilde{y}_{t-1}) d\alpha_t =$$

$$= \int_{-\infty}^{\alpha} p(y_t | \alpha_t) \cdot p(\alpha_t | \tilde{y}_{t-1}) d\alpha_t$$

Normal

$\hat{y}_{t|t-1}$

$$E[y_t | \hat{y}_{t-1}] = z_t a_t$$

$$F_t = \text{var}(y_t - \hat{y}_{t-1}) = z_t P_t z_t + H_t$$

Se pudemos provar que prob. conjunta das inovações = produto das marginais

$\Rightarrow$  inovações são independentes

— II —

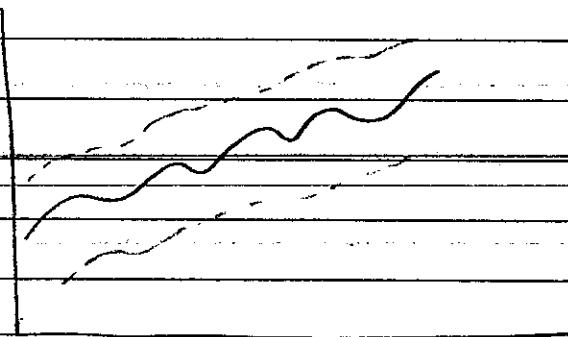
Intervalos de confiança à tendência

$$\text{Ex: } y_t = \mu_t + \gamma_t + \epsilon_t$$

$$\begin{aligned} x_t &= \begin{pmatrix} \mu_t \\ \gamma_t \end{pmatrix} \text{ tendência} & \equiv & \alpha_{1,t} \\ &&& \alpha_{2,t} \\ &\text{sazonalidade} && \end{aligned}$$

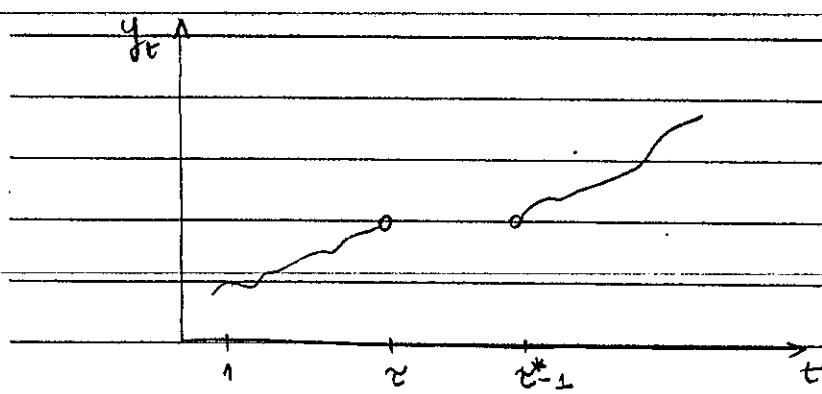
$$\hat{\alpha}_{1,t} \pm 1.96 \sqrt{V_{1,t}}$$

da variação suavizada



11

Observações faltantes em MEE



Observações faltantes em  $t = \tau, \tau+1, \dots, \tau^*-1$

$(\tau^* - \tau)$  observações faltantes

$F_K$

$$a_{t+1} = T_t a_t + K_t v_t$$

$$P_{t+1} = T_t P_t L_t' + R_t Q_t R_t'$$

$$L_t = T_t - K_t Z_t$$

Nas inst. em que tiverem obs faltantes faremos  $v_t = 0$  (mas manter os s. pf suas inovações correspondentes).

E como n̄ vai ter ganho (pois n̄ há inovações)  $\Rightarrow K_t = 0$ .

$$\text{Em } t = \tau, \tau+1, \dots, \tau^*-1 \Rightarrow v_t = 0 \quad \left. \begin{array}{l} \\ \end{array} \right\} L_t = T_t \\ K_t = 0$$

$$\therefore a_{t+1} = T_t a_t$$

$$P_{t+1} = T_t P_t T_t' + R_t Q_t R_t'$$

último t com observação é  $t = T - 1$

Supondo sistema invariante no tempo (pr contas + simples)

$$\alpha_{t+1} = T\alpha_t$$

$$P_{t+1} = TP_tT' + RQR'$$

$$t+1 = \gamma \Rightarrow \alpha_\gamma = T\alpha_{\gamma-1}$$

$$\alpha_{\gamma+1} = T\alpha_\gamma = T^2\alpha_{\gamma-1}$$

:

$$\alpha_{\gamma+k} = T^k\alpha_{\gamma-1}$$

Da equações do estado:

$$\alpha_{t+1} = T\alpha_t + R\eta_t$$

Iterando até  $t+j$ :

$$\alpha_{t+j} = T^j\alpha_t + (\dots)$$

$$\text{Previsões: } E[\alpha_{t+j} | y_t] = T^j E[\alpha_t | y_t] = T^j \alpha_{t|t}$$

Nos períodos de previsão, o FKF projeta j passos à frente.

Pega-se o último vetor de estados e projeta a média  
ao longo do período de dados faltantes

Para a variância valerá a mesma coisa.

$$P_{t+j} = TP_{t+j}T' + RQR'$$

$$= T(TP_tT' + RQR')T' + RQR =$$

$$= T^2 P_t (T')^2 + TRQT' + RQR$$

$$\text{Ex: } y_t = \alpha_t + \epsilon_t$$

$$\alpha_{t+1} = \alpha_t + \eta_t$$

missing

 $\tau-1 \quad \tau \quad \tau+1 \dots \tau^*-1$ 

$$T = z = 1$$

$$\alpha_t \equiv \alpha_\tau + \sum_{j=\tau}^{t-1} \eta_j, \quad t = \tau, \tau+1, \dots, \tau^*-1 \quad \textcircled{*}$$

$$E[\alpha_t | Y_{t-1}] = E[\alpha_t | Y_{\tau-1}]$$

$$= a_\tau \quad \begin{matrix} \uparrow \\ \text{temos obs ate' } \tau-1 \end{matrix}$$

→ última previsão ante de faltar dados

Para os períodos de obs faltantes:

$$\text{pp } a_\tau \text{ ainda} \quad a_{t+1} = a_t, \quad \tau, \dots, \tau^*-1$$

$$\text{podemos} \quad a_{\tau+1} = a_\tau$$

$$\text{pois } a_\tau = E[a_t | Y_{\tau-1}] \quad a_{\tau+2} = a_{\tau+1} = a_\tau$$

;

$$\textcircled{*} \quad \text{var}(\alpha_t | Y_{t-1}) = \text{var}(\alpha_t | Y_{\tau-1}) + \text{var}\left(\sum_{j=\tau}^{t-1} \eta_j\right)$$

$$= P_\tau + (t-\tau) \sigma_\eta^2$$

variância tb é projetada

Nos algoritmos de smoothing:

$$\text{Equações originais: } \hat{\alpha}_t = \alpha_t + P_t n_{t-1} \quad e - V_t = P_t - P_t N_{t-1} P_t$$

$$n_{t-1} = F_t^{-1} Y_t + L_t \quad N_{t-1} = F_t^{-1} Z_t + L_t^T N_t L_t$$

$$n_n = 0$$

$$N_n = 0$$

$$\text{Para } t = \tau_1, \tau_1+1, \dots, \tau^*-1 \Rightarrow K_t = 0 \Rightarrow L_t = T_t \\ V_t = 0$$

$$\therefore \begin{cases} N_{t-1} = L_t^T N_t = T_t^T N_t \\ N_{t-1} = T_t^T N_t L_t \end{cases}$$

$$\text{Example: } y_t = \alpha_t + \varepsilon_t$$

$$\alpha_{t+1} = \alpha_t + \eta_t$$

$$\cdot \hat{\alpha}_t = \alpha_t + p_t n_{t-1} \quad (\text{no pierde d. obs. faltantes})$$

$$\cdot \hat{\alpha}_{\tau+j} = \alpha_\tau + p_{\tau+j} + n_{\tau+j-1}, \quad j = 1, 2, \dots, \tau^*-1$$

$$= \alpha_\tau + (p_\tau + j\sigma_\eta^2) n_\tau$$

↓

do resultado anterior

medo  $j$  o gap de dados

faltantes

$$= (\underbrace{\alpha_\tau + (p_\tau n_\tau)}_a) + j(\underbrace{\sigma_\eta^2 n_\tau}_b)$$

$$\therefore \hat{\alpha}_{\tau+j} = a + b j$$

$$y_t = \alpha_t + \varepsilon_t$$

$$y_{\tau+j} = \alpha_{\tau+j} + \varepsilon_{\tau+j}$$

$$\text{smoothing} \quad E[y_{\tau+j} | y_m] = E[\alpha_{\tau+j}] \\ (\text{interpolac.}) \quad = \alpha_{\tau+j}$$

tildeira

é característica  
do modelo.  
foi final da amostra via ch. tb  
(anterior)

a e b com dados passados et u

## Inicializações do Fk

pois tecnicamente não é significante

mas no Fk precisamos olhar atentamente.

Pf processos estacionários, cond. inicial perde importância com  $t \rightarrow \infty$ .

Pf processos n'estacionários  $\Rightarrow$  isso não é verdade.

Vemos

$\checkmark$  mas maneiras de fazer inicializações

$$\left\{ \begin{array}{l} y_t = z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \end{array} \right.$$

$$\alpha_1 = T_0 \alpha_0 + R_0 \eta_0$$

$$t=1 : \alpha_2 = T_1 \alpha_1 + R_1 \eta_1$$

↓

$N(\alpha_i, P_i) \rightarrow$  até agora, tratamos como conhecidos

↓ ↓

vetor px, matriz var, covar  $p \times p$

O que determina como tratar cond. inicial é a estacionariedade ou não de cada componente do vetor de estado.

Pf comp. estacionária  $\Rightarrow$  usa como cond. a dist. de LP do processo.

if comp. nos estacionários  $\Rightarrow$  dist. a priori difusa

$$y_t = \mu_t + \epsilon_t$$

$$\mu_{t+1} = \mu_t + \eta_t$$

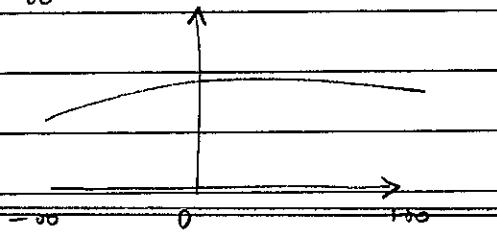
$$\alpha_i = \mu_i \sim N(a_i, P_i)$$

(

já que não sabemos nada sobre  $\alpha_i$ , mas devemos priorizar nenhum valor sobre outro.

Seria como tomar  $\alpha_i = p_i$  dist. uniforme

Equivale a fazer  $P_i = K \rightarrow \infty$



obs: mas do ponto de vista estatístico é desapadrinhada  
pois esta é uma dist. imprópria, ou seja:

$$\lim_{K \rightarrow \infty} \int_{-\infty}^{\infty} f(\alpha_i) d\alpha_i \neq 0.$$

Prior difusa é imprópria pois tem var infinita  
e  $\lim_{K \rightarrow \infty} \int f(\alpha_i) d\alpha_i$  não converge.

Na prática, nas reais problemas pois só começaremos a trabalhar os dados quando o processo tiver prior própria (deixar de pendurar a prior difusa)

obs: A apresentação da forma que vemos é particular  
do livro de Koopman.

(prova!)

11

- Fazer cálculos / manipulações pp modelos + geral

$$y_t = \mu_t + \varepsilon_t$$

$$\alpha_t \sim N(0, K\sigma_e^2)$$

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t$$

$$\beta_{t+1} = \beta_t + \varsigma_t$$

tabalha com  
matriz var/covar  
 $K=oo$  diagonal

Neste processo, pode-se mostrar que a partir de  $t=3$   
teremos priors bem definidas (próprias) e assim o  
cômputo das equações do FK não apresentará problema.

Cf d componentes estacionárias) a partir de  $t+1$  podemos  
computar FK de forma  
adequada

pg(8): Algumas contas:

Tendência linear local

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_e^2) \Rightarrow \begin{cases} y_t = \mu_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \eta_t \\ \beta_{t+1} = \beta_t + \varsigma_t \end{cases}$$

$$\begin{pmatrix} \mu_{t+1} \\ \beta_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_t \\ \varsigma_t \end{pmatrix}$$

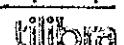
$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix}, Q_t = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\varsigma^2 \end{pmatrix} = \sigma_e^2 \begin{pmatrix} q_\eta & 0 \\ 0 & q_\varsigma \end{pmatrix}$$

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$R_t = I_2$$

$$H_t = \sigma_e^2$$

$$\text{onde } \frac{\sigma_\eta^2}{\sigma_e^2} = q_\eta, \quad \frac{\sigma_\varsigma^2}{\sigma_e^2} = q_\varsigma$$

Razões nível ruído. 

Mais fácil se estimar  $\hat{q}_y, \hat{q}_z$  por  $mV$   
 (reparametrizando o sistema pode deixar  
 um parâmetro a menos ser estimado)

$$\hat{\sigma}_e^2 = \frac{1}{T-2} \sum_{t=3}^T v_t^2 / f_t$$

$$v_t = f(\hat{q}_y, \hat{q}_z)$$

$$f_t = g(\hat{q}_y, \hat{q}_z)$$

A cond. inicial é dada por:

$$\alpha_1 = a + A\delta + R_0 \gamma_0$$

$$P_1 = K P_{00} + P_x$$

$$\text{onde } P_{00} = AA'$$

$R_0 = 0$  pois não há componentes estacionárias.

$$\Rightarrow \alpha_1 = \begin{pmatrix} \mu_1 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}$$

$$\text{onde } \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, K \right]$$

$$\therefore P_{00} = AA' = I_2 I_2' = I_2 = P_{00,1}$$

$$P_x = 0$$

$$= P_{x,1}$$

ligado a  $\mu$

ligado a  $\beta$

to suas iteradas

$$\Rightarrow P_1 = K P_{00} = K I_2$$

Primeira estimativa:

$$P_1 \circ F_K = 1 \text{ usamos } \alpha_{t+1} \circ P_{t+1}$$

Enquanto estivermos na fase de inicialização exata, usaremos notações diferentes:  $\hat{\alpha}_2^{(0)}$

Quando atingirmos  $P_{00,t} = 0 \Rightarrow t = d$ : aí começaremos a considerar  $F_K$  padronas.

$$\hat{\alpha}_2^{(0)} = \hat{T} \hat{\alpha}_1^{(0)} + k_1^{(0)} v_1^{(0)} - k_1^{(0)} v_1^{(0)}$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}^{(0)} \\ \alpha_{2,1}^{(0)} \end{pmatrix}$$

fazer contas como na folha

— — —

### Estimacões de hiperparâmetros

Além das componentes que são estimadas pelo FK

Há alguns hiperparâmetros do modelo que precisam ser estimados.

são constantes fixas desconhecidas nas matrizes  $\{T_t, T_e, H_t, Q_t\}$

Normalmente R nas tem  
(porque é matriz de seleção)

Em modelos estruturais, normalmente em q e H (podem ser cheias ou diagonais)

↓  
mais comum

$\alpha_F \rightarrow FK$

$\Psi = \{ \text{hiperparâmetros} \} \Rightarrow \{T_t, T_e, H_t, Q_t\}$

↓  
usaremos MV

Quando dist iid  $\Rightarrow$  prova-se que estimadores MV

1. consistentes

2. assintoticamente N vizinhas

3. eficientes

4. dist. normal assintótica

iid  $\Rightarrow$  1. consistência:  $\text{plim } \hat{\theta} = \theta$

2. assintoticamente nas vizinhadas:  $\lim_{n \rightarrow \infty} E[\hat{\theta}_n] \rightarrow \theta$

3. eficientes:  $\text{Var}(\hat{\theta})$  é mínima

4. dist. normal assintótica  $\Rightarrow \hat{\theta} \sim N(\theta, \text{Var}(\hat{\theta}))$

↓  
dados dependentes

Quando os dados não são iid, a prova é mais complicada, mas tem as mesmas propriedades ( desde que, não estejam na fronteira do espaço paramétrico) parâmetros

$\Rightarrow$  Por isso usamos Verossimilhança

obs: pf pequenas amostras, propriedades assintóticas não são muitas boas  $\Rightarrow$  usa bootstrap

Verossimilhança:  $\rightarrow$  parte de uma densidade conjunta, mas

o argumento da função

$$L(\psi) = p(y_1, y_2, \dots, y_n; \psi)$$

é diferente (os parâmetros). Neste caso,

$$= \prod_{i=1}^n p(y_i | y_{i-1}; \psi)$$

seus parâmetros não.



precisa ser = 1

densidade

preditiva

(pf dados dependentes)

duas formas de obter  $p(y_t | y_{t-1})$ :

Nos modela em ET

$$1) p(y_t | \underline{y}_{t-1}) = \int p(y_t, \alpha_t | \underline{y}_{t-1}) d\alpha_t =$$

$$= \int p(y_t | \alpha_t) p(\alpha_t | \underline{y}_{t-1}) d\alpha_t \Rightarrow \text{resultado geral independente}$$

$$\text{Mas, } y_t = z_t \alpha_t + \epsilon_t, \epsilon_t \sim N(0, H_t)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \eta_t \sim N(0, Q)$$

de estamos em ambiente Gaussiano

$$\Rightarrow p(y_t | \alpha_t) \sim N(z_t \alpha_t, (H_t))$$

$$p(\alpha_t | \underline{y}_{t-1}) \sim N(\alpha_t, P_t)$$

Logo:  $N(z_t \alpha_t, H_t) \cdot N(\alpha_t, P_t) \Rightarrow$  convoluções da normal.

(

Trabalhando esta expressão, achamos

$$E[\cdot] = V[\cdot]$$

Mas podemos olhar diretamente as eqs:

$$E[y_t | \underline{y}_{t-1}] = z_t \alpha_t$$

$$\text{var}[y_t | \underline{y}_{t-1}] = E_t P_t E_t' + R_t \cancel{R_t'} = F_t \\ H_t$$

$$\Rightarrow p(y_t | \underline{y}_{t-1}) \sim N(z_t \alpha_t, F_t)$$

$$\therefore p(y_t | \underline{y}_{t-1}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|F_t|^{1/2}} \exp \left\{ -\frac{1}{2} \underbrace{(y_t - z_t \alpha_t)' F_t^{-1} (y_t - z_t \alpha_t)}_{V_t} \right\}$$

p-variaada

$$= (2\pi)^{-p/2} |F_t|^{-1/2} e^{-\frac{1}{2} V_t' F_t^{-1} V_t}$$

$$\therefore L(\psi) = p(y_1, y_2, \dots, y_n, \psi) = \prod_{t=1}^n p(y_t | y_{t-1}, \psi) =$$

$$= \prod_{t=1}^n (2\pi)^{-p_{12}} |F_t|^{-1/2} e^{-\frac{1}{2} v_t^T F_t^{-1} v_t}$$

$$\log L(\psi) = l(\psi) = -\frac{pn}{2} \log(2\pi) - \frac{1}{2} \sum_t (\log |F_t| + v_t^T F_t^{-1} v_t)$$

( entra  $F_t$  (do FK))



estimar  $\psi$  depende do FK  
e FK depende de  $\psi$

cond. 1ª ordem (necessária)

$$\frac{\partial l}{\partial \psi} = 0 \quad \therefore \left( \frac{\partial l}{\partial \psi_1}, \frac{\partial l}{\partial \psi_2}, \frac{\partial l}{\partial \psi_3}, \dots, \frac{\partial l}{\partial \psi_l} \right)' = 0$$

vetor nulo

respondo  $\psi$  com dimensão  $l$

cond. 2ª ordem (suficiente)

$$H = \begin{pmatrix} \frac{\partial^2 l}{\partial \psi \partial \psi'} \end{pmatrix}_{\text{exc.}} \quad \text{é negativa definida} \Rightarrow z^T H z < 0$$

(negativa)

$$H z \neq 0$$

→ todos autovalores São < 0

matriz Hessiana

$$x_i(H) < 0 \quad i=1, 2, \dots, l$$

Nos vamos conseguir encher  $\psi$  como soluções analítica de  $y_t$ 's.  
(teremos que usar métodos não-lineares de otimização)

Se  $l$  for cônica  $\Rightarrow$  resolvido problema encontraremos

maximum global

Mas  $l$  não é cônica

Quando calcularmos

stática

Na funções  $\ell(\psi)$ :  $F_t = F_t(\psi)$

$$v_t = v_t(\psi)$$

Inicializar: duas formas  $\xrightarrow{(i)} \text{difusa por big kappa}$   
 $\xrightarrow{(ii)} \text{Exata}$

(i) . Difusa por big kappa

Computar verosimilhança:

for  $t = 1, \dots, n$   $\leftarrow$  pega cond. inicial  $a_0, P_0$

$$a_{t+1} =$$

$$P_{t+1} =$$

$$v_t =$$

No modelo linear

$$\text{local, p ex: } a_i \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$P_t = K I_{2 \times K^2}$$

$$F_t =$$

$$l = l - \frac{1}{2} \log |F_t| - \frac{1}{2} v_t' F_t^{-1} v_t$$

continua

$$y_t = z_t a_t$$

$$z_t = P_t z_t' + H_t$$

enquanto

estas infos não forem

bem definidas, não podemos

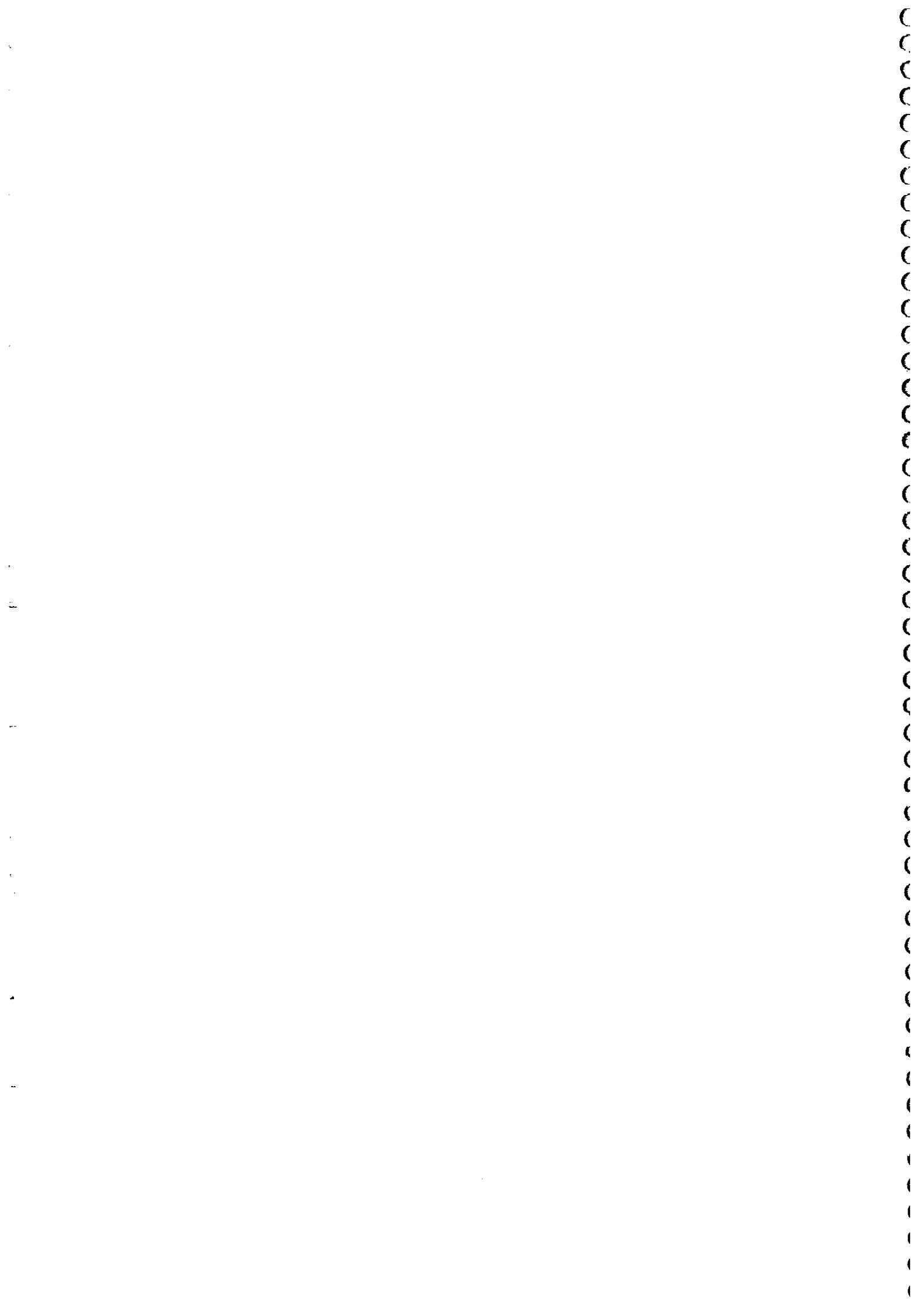
computar verosimilhança.

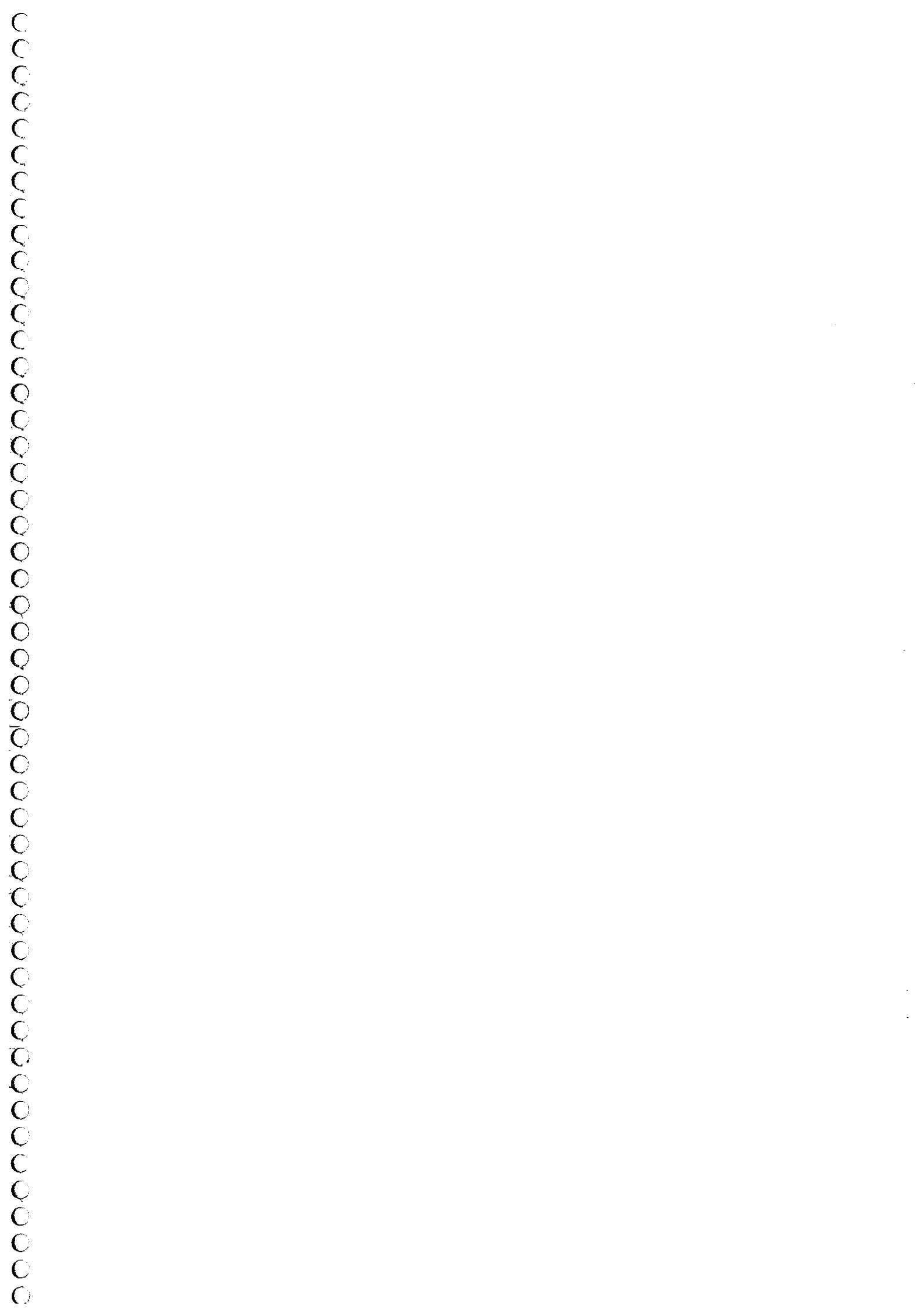
foi vamos calculá-la a partir

de  $t = d$ .

leusões de AK são calculadas desde o início, mas

$\ell(\psi)$  apenas quando dist. difusa bem definida:  $(\sum_{t=d}^n)$





contribuições  
dada pelas  
termos  $\log |F_t|$

### (ii). Difusa Exata

Funções Verossimilhança:  $\ell(4) \propto -q \log k$

Verossimilhança  
padras  
nas e  
adequada  
para trabalhar.  
Termos que  
reservem os  
comos

$$\ell_d = \ell(4) + q \log k$$

$q$ : n.º componentes nas estacionárias  
do modelo.

$$= -np \log(2\pi) - \frac{1}{2} \sum_{t=1}^d w_t - \frac{1}{2} \sum_{t=d+1}^n [v_t F_t^{-1} v_t + \log |F_t|]$$

↓ de  $t=d+1$  a  $n$ : forma padras

$d=t=1$  a  $d$ : 1<sup>as</sup> observações tratadas de forma diferente

$$w_t = \log |F_{00,t}| \propto F_{00,t} \text{ é p.d.}$$

$$\log |F_{x,t}| + v_t^{(0)} F_{t,x}^{-1} v_t^{(0)} \propto F_{00,t=0}$$

Otimização: (método numérico, tipo quase Newton)

6

Usando info de  
cond. de 1ª e 2ª

orden (gradiente,  
Hessiano)

ver software

Stamp

dois casos  
particulares:

BFGS,  $\Rightarrow$  Forma geral de Algoritmos:

BHHH

$$\psi^{(i+1)} = \psi^{(i)} - \gamma H^{-1}(\psi^{(i)}) g(\psi^{(i)}) \quad i=1, 2, \dots, I_{\max}$$

$(\ell x_1) \quad (\ell x_1) \quad \uparrow$   
scalar    esc    esc

começa com solução inicial:  $i=1$   $\sigma_\eta^2 = 0,002$

$$\sigma_E^2 = 0,1$$

$\gamma$ : passo. p/ controlar a busca pela solução

tilde

( / / )

se  $H(\psi)$  for sempre neg. definida  $\Rightarrow$  garante que resultado  
mais econômico (não  
fim de paros)

$\Rightarrow$  A cada instante de  $t \rightarrow$  teremos um  $\psi^{(i+1)}$

Temos que avaliar  $x$  chegamos no ótimo

Há critérios de parada (ver notas scanneadas sobre otimizações)

$$1. \text{ Verossimilhança: } \frac{l(\psi^{(i)}) - l(\psi^{(i-1)})}{l(\psi^{(i)})} < \epsilon$$

$\xrightarrow{\text{mais importante}}$  2. Gradiente:  $\frac{dl}{d\psi_i} = 0$

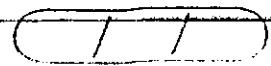
Imaginando que temos:

$$\begin{pmatrix} \frac{dl}{d\psi_1} \\ \frac{dl}{d\psi_2} \\ \vdots \\ \frac{dl}{d\psi_n} \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \end{pmatrix}$$

$$\text{se } \sum_{i=1}^n \left| \frac{dl}{d\psi_i} \right| \approx 0$$

Se a média for muito pequena  $\Rightarrow$   
podemos considerar que grad = 0.

3. Parâmetros:



tilibra

(20/04/11)

### Estimacões de Hiperparâmetros

$$\hookrightarrow \text{Para MLE} \quad \hat{\psi} \sim \mathcal{N}(\psi, \hat{\Omega}_{\psi})$$

Estimados pelo princípio da máxima Verossimilhança  
(apresentam propriedades desejáveis)

- 1) Nas viciadas  $\lim_{n \rightarrow \infty} E[\hat{\psi}_n] \rightarrow \psi$
- 2) Consistência plausível  $\hat{\psi}_n = \psi$
- 3) Eficiência  $\text{var}(\hat{\psi}_n) \rightarrow \text{LCR} (\text{Cramer Rao})$
- 4) Dist. Assintótica normal  $\hat{\psi} \stackrel{d}{\sim} N(\psi, \hat{\Omega}_{\psi})$   
(importância p/ IC's e TH's)  
p/ os estimadores

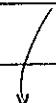
Problema p/ considerar amostras  
muito grandes (o que reia  $n \rightarrow \infty$ )

Método Bootstrap: p/ avaliar as distribuições  
p/ diferentes tamanhos de amostra.

Bootstrap: técnicas p/ obter dist. de estimadores p/ tamanhos  
de amostra aleatórios.

$$L(\psi) = \prod_{t=1}^n p(y_t | y_{t-1}, \psi)$$

$$\text{Vemos que } p(y_t | y_{t-1}) = (2\pi)^{-p/2} |F_t|^{-1/2} V_t' F_t^{-1} V_t$$



$$V_t = y_t - z_t a_t$$

ponta do F\_t p/ implementar verossimilhança

Mais adiante trabalhar com  $\ell(\psi) = \log L(\psi)$

$$\frac{d\ell}{d\psi} = \frac{1}{L(\psi)} \cdot \frac{\partial L}{\partial \psi} = 0 \Rightarrow \text{Max de } L(\psi) = \text{Max de } \ell(\psi)$$

$$\ell(\psi) = \log L(\psi) = \sum_{t=1}^n \log p(y_t | y_{t-1})$$

$y_t$  → contribuições da  $t$ -ésima componente em  $\ell(\psi)$

$$\therefore \ell(\psi) = -p_n \log(2\pi) - \frac{1}{2} \sum_{t=1}^n [\log |F_t| + v_t' F_t^{-1} v_t]$$

(não interessa)

pois é cte.  
(não depende dos hiperparâmetros)

Kernel das funções de Verossimilhança

Verossimilhança por  
chamada de decomposições do erro de previsão

inovações  
(verossimilhança muita tendo  $v_t$  como parâmetro)

$$\text{onde: } v_t = y_t - z_t a_t$$

$$z_t = z_t P_t z_t' + H_t$$

Problema: como funciona a questão da inicialização  
do  $F_k$  e a verossimilhança?

Como será tratada?

Inicializações do  $F_k$  na verossimilhança:

duas formas → prior difusa via "big kappa"  
inicializações exata

tilibra

$$p_i \propto \alpha_i \sim N(\bar{\alpha}_i, \bar{P}_i)$$

### i) Prior difusa

Ex:  $y_t = \mu_t + \epsilon_t$ ,  $\epsilon_t \sim N(0, \sigma_e^2)$

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

$$\text{Hyperparâmetros} \Rightarrow \Psi = (\sigma_e^2, \sigma_\eta^2) \text{ ou } (q, \sigma_e^2)$$

Vantagem:

$$\hat{\sigma}_e^2 = \frac{1}{n} \sum V_t^T F_t^{-1} V_t \rightarrow \text{so depen de } q$$

Procura da verossimilhança pode ser feita só em relação a  $q$  e encontrar  $\sigma_e^2$  em funções de  $q$ .

Reparametrizações: Rescreve  $P_t$  e  $F_t$  em funções de  $\sigma_e^2$  (dividindo por) acha  $P_t^*$  e  $F_t^*$

funções de  $q$

$$l(\sigma_e^2, q) \rightarrow dL = 0 \rightarrow \text{terá termo } \frac{d\sigma_e^2}{d\sigma_e^2} V_t^T F_t^{-1} V_t$$

apenas como função de  $q$ .

No exemplo:  $p = m = 1$

$$Z = T = R = 1$$

$$l(\Psi) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \left[ \log F_t + \frac{V_t^2}{F_t} \right] \quad \begin{matrix} \text{corresponde a} \\ \sum V_t^T F_t^{-1} V_t \text{ unidimensional} \end{matrix}$$

Inicializações difusa:  $\alpha_t \sim N(0, k)$ ,  $k \sim 10^3$

Em  $t=1$ : distribuição  $\bar{m}$  é bem definida

$\Rightarrow$  considera tudo começando em  $t=2$

$$P|t=1 \Rightarrow \text{calcula } \bar{F}_t = P_t + \sigma_e^2 \rightarrow P_{t+1}$$

$$\alpha_{t+1}$$

$$l(\psi) = - \frac{1}{2} (n-1) \log(2\pi) - \frac{1}{2} \sum_{t=2}^n (\log F_t + v_t^2/F_t)$$

Outra questão é o número de componentes nas estacionárias no vetor de estado.

Caso mais comum é ter comp. nas estacionárias

$x_t$ : tem q componentes nas estacionárias

$x_t \sim m \times 1$  onde  $q \leq m$

$$\Rightarrow l(\psi) = - \frac{1}{2} p(n-q) \log(2\pi) - \frac{1}{2} \sum_{t=q+1}^n (\log |F_t| + v_t' F_t^{-1} v_t)$$

Na implementação, só para a considerar  $t \geq t=q+1$

$a=0 \quad R=k$

$$t=0 \quad \rightarrow v_0 = y_0 - a\psi$$

$$\text{for } t=1, n \quad P_t = \dots$$

$$l = l - \frac{1}{2} (\log F_t + v_t^2/F_t)$$

If  $t=1$  a q

(computa informações, mas não a  
verossimilhança)

2) Inicializações Exatas

normalmente coincide com nº de  
variáveis n estacionárias

Faz de:  $t=1$   $d$   $d+1$   $\dots$   $n$

$F_k$  "equitativo" Rada.  $F_k$  padas.

Podem acontecer 2 coisas:  $F_{0,t}$  é p.d. → mais comum

$$F_{0,t} = 0$$

particular  
det. nula

Na verossimilhança:

$$\ell(\psi) = -\frac{pn}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^d \left( \dots \right) - \frac{1}{2} \sum_{t=d+1}^n \left( \dots \right)$$

usa eq.

cl alterações

p. o FK (que

vimos na

aula passada)

usa equações

do FK padrão

$$\ell_d(\psi) \hat{=} \lim_{k \rightarrow \infty} \left( \ell(\psi) + \frac{q}{2} \log k \right)$$

a contribuições em  $\ell(\psi)$  de  $\alpha_i$  é

dada por  $-\frac{q}{2} \log k$

(koopman in motion,

apenas informa)

Temos i anulado e fica apenas  $\ell(\psi)$

$$\therefore \ell_d(\psi) = -\frac{(np)}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^d w_t + \frac{1}{2} \sum_{t=d+1}^n (\log |F_t| + v_t^\top F^{-1} v_t)$$

$w_t = \begin{cases} \log |F_{0,t}|, & F_{0,t} \text{ é p.d.} \rightarrow \text{caso mais comum} \\ \end{cases}$

$\log |F_{0,t}| + v_t^{(0)} F_{0,t} v_t^{(0)}, & F_{0,t} = 0.$

obs: interessante fazer no prox exercício inic. bfg kappa,

initialização exata, n.veros., e ver q faz muitas diferenças

LIBRA

Hamilton: tem uma recaída

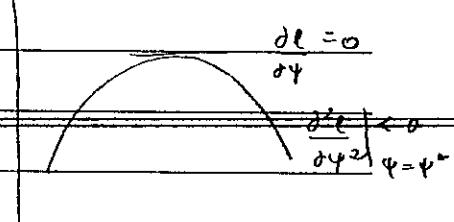
## Otros tipos de Números

## de optimizac̄as numérica

(mas nem N modelos estruturais.)

$$\hat{\varphi} = \arg \max_{\varphi \in \Sigma} \ell(\varphi)$$

No caso univariado:



Mas função é não é necessariamente côncava.

Podeem fer màximes locals.

$$\text{2º ordem : } H(4) = -\delta^2 \ell \quad | \quad > 0$$

$\downarrow$   
Mariano

$$\delta \Psi \delta \Psi' \quad | \quad \Psi = \Psi'$$

$$\text{onde } H(\vec{\psi}) = -J^2 \vec{\ell} = -\begin{pmatrix} \frac{\partial^2 \ell}{\partial \psi_1^2} & \frac{\partial^2 \ell}{\partial \psi_1 \partial \psi_2} & \dots & \frac{\partial^2 \ell}{\partial \psi_1 \partial \psi_v} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \ell}{\partial \psi_v \partial \psi_1} & \dots & \frac{\partial^2 \ell}{\partial \psi_v^2} \end{pmatrix}$$

$$\Rightarrow H(\hat{\psi}) = H|_{\psi=\hat{\psi}}$$

L'Ufficio

Como  $H = -\frac{\partial^2 \ell}{\partial \psi \partial \bar{\psi}}$ , se precisaremos ter  $H$  partindo

*Entrada:  $\mathbf{z}' \mathbf{H} \mathbf{g} > 0$ , i.e.,  $\mathbf{H}$  é positiva definida (p.d.)*

*$\mathbf{z}$  é vetor  $\mathbf{v} \times \mathbf{1} \neq 0$ .*

$$\text{Se } \mathbf{v} = \mathbf{z} \Rightarrow \Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \Rightarrow h_{11} = \frac{\partial^2 \ell}{\partial \Psi_1^2} > 0$$

$$h_{22} = \frac{\partial^2 \ell}{\partial \Psi_2^2} > 0$$

↓

*valores da diagonal de  $\mathbf{H}$*

Entretanto, dificilmente temos soluções analítica.

Precisaremos de um método numérico para realizar a otimização.

a) Métodos do gradiente: algoritmos de otimizações numérica definidas pela seguinte estrutura

i. procedimento iterativo inicializado com condições inicial

ii. iterações produzem soluções aproximadas  $\hat{\Psi}^{(1)}, \hat{\Psi}^{(2)}, \hat{\Psi}^{(k)}$

iii. a cada iteração  $k$ , a solução é feita p/ ótimo a partir de uma regra de parada

Dependendo das cond. inicial → pode parar em um máximo local.

Vale a pena mexer na cond. inicial. Se parar em pontos diferentes  $\Rightarrow$  indicações de vários máximos locais

Por exemplo, uma das formas de implementar é:

$$\Psi^{(k+1)} = \Psi^{(k)} + \lambda \underset{\mathbf{V} \times \mathbf{V}}{\mathbf{H}^{-1}} \underset{\mathbf{V} \times \mathbf{V}}{\mathbf{g}}(\Psi^{(k)}) \Rightarrow \text{p/ cada iteração coube a regra}$$

de parada.  
 $k = 1, 2, \dots, k_{\max}$

$\lambda$  é implementado de diferentes maneiras. Pode ser um  $\lambda^{(k)}$ .

ilibra

Esses tipos de algoritmos-são chamados quasi-Newton.

Tomemos  $\ell(\psi)$  e façamos expansões de Taylor de 2º ordem em torno de uma iteração arbitrária de  $\psi$  dada por  $\psi^{(k)}$ :

$$\ell(\psi) = \ell(\psi^{(k)}) + \left( \frac{de}{d\psi} \right)^T (\psi - \psi^{(k)}) + \frac{1}{2} (\psi - \psi^{(k)})^T \frac{\partial^2 \ell}{\partial \psi^{(k)} \partial \psi^{(k)}} (\psi - \psi^{(k)})$$

Por construção, queremos que  $\psi$  seja ponto de máximo de modo que satisfaça  $\frac{de}{d\psi} = 0$

Derivando a expressão:

$$g(\psi^{(k)}) - H(\psi^{(k)}) (\psi - \psi^{(k)}) = 0$$

$$H(\psi^{(k)}) (\psi - \psi^{(k)}) = g(\psi)$$

$$\boxed{\psi^{(k+1)} = \psi^{(k)} + \lambda H^{-1}(\psi^{(k)}) g(\psi^{(k)})}$$

Se formasse função quadrática, chegaria no máximo na 1ª iteração.

Mas no caso trabalhamos com iterações.

chama  $\psi$  de  $\psi^{(k+1)}$

e  $\lambda$  reai o tamanho do passo.

( / / )

Critérios de Parada:

(Normalmente é usado 1 ou mais critérios simultaneamente.  
Mais importante é o do gradiente)

(i) critério do gradiente:  $g(\psi^{(k)}) = 0$

equivale a:

$$\begin{pmatrix} \frac{\partial e}{\partial \psi_1^{(k)}} \\ \frac{\partial e}{\partial \psi_2^{(k)}} \\ \vdots \\ \frac{\partial e}{\partial \psi_n^{(k)}} \end{pmatrix} = 10^{-m}$$

dos módulos

Na prática, toma-se a média  $\bar{v}$  e verifica se é menor de cada um muito pequeno.

$$c_1 = \sum_{j=1}^n |g_j(\psi^{(k)})| < 10\varepsilon$$

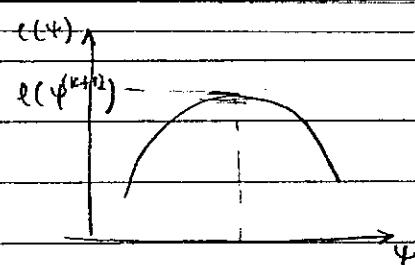
→ Nas é a única forma.

Poderia pegar o máximo do vetor  $g(\psi^{(k)})$  e ver se é  $< 10\varepsilon$

(ii) critério de verossimilhança

No max: variações da função e do ponto têm que

ser muito pequena



$$c_2 = \left| \frac{l(\psi^{(k)}) - l(\psi^{(k+1)})}{l(\psi^{(k)})} \right| < \varepsilon$$

↳ variação relativa

(iii) critério para parâmetros

$\hat{\psi}$  é a média do valor do parâmetro

ao medindo muitos

(equivalente ao que fizemos com  
a verossimilhança)

$$G_3 = \frac{1}{n} \sum_{j=1}^n \left| \frac{\psi_j^{(k+1)} - \psi_j^{(k)}}{\psi_j^{(k)}} \right| < \epsilon$$

Normalmente, algum parâmetro tem restrições de valores.

Ex:  $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \quad \sigma_e^2 > 0$

Para usar otimizações sem restrições: faz uma reparametrização  $\eta$  os estimadores que necessitam.

Ex:  $\hat{\psi}_1 = e^{2\theta} \quad \theta \in \mathbb{R}$

Variância dada por  $\sigma_e^2 = e^{2\theta}$

$$\therefore \sigma_e = e^{\theta}$$

qualquer  $g(\cdot) : (0, \infty) \rightarrow \mathbb{R}$  é permitida.

(ver nas anotações as transformações propostas)

Ex:  $\psi_t = \mu_t + \varepsilon_t \sim \sigma_e^2$        $\left. \begin{array}{l} l(\psi) \propto v_t^2 / f_t \\ f_t = p_t + e^{2\theta} \end{array} \right\}$  parâmetro  $\theta$ .

$$\mu_{t+1} = \mu_t + \gamma \varepsilon_t \sim \sigma_\gamma^2$$

Quando fizemos otimização numérica, ela estará sendo realizada não sobre os parâmetros originais, mas sim em  $\theta$ .  
Teremos:

$$\theta^{(k+1)} = \theta^{(k)} + \lambda H^{-1}(\theta^{(k)}) g(\theta^{(k)})$$

$$\hat{\theta} \sim \mathcal{N}(\mu, \Sigma), \quad \theta = T(\hat{\theta})$$

Incerteza nos estimados: IC

$$\hat{\theta}_j \pm z \sqrt{\text{var}(\hat{\theta}_j)}$$

precisaremos da estimativa de  $\text{var}(\hat{\theta}_j)$

$$\hat{\theta} \sim N\left(\theta, \frac{I'(\theta)}{n}\right) \text{ onde } I(\theta) \text{ é a informação de Fisher}$$

$$I(\theta) = E\left(\frac{\partial^2 \ell}{\partial \theta \partial \theta'}\right) = E\left(\frac{\partial \ell}{\partial \theta} \frac{\partial \ell}{\partial \theta'}\right)$$

- H

$$\left(\frac{\partial \ell}{\partial \theta}\right)'$$

$$\text{Na prática } l = l(\theta) = \sum_{t=1}^n l_t(\theta) =$$

$$= -\frac{n}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^n \left[ \ln |f_{\theta}| + v_t' f_{\theta}^{-1} v_t \right]$$

dn: n. os modelos em EE, teremos

$$E\left[\frac{\partial \ell}{\partial \theta}\right] = \underbrace{\square}_{2 \text{ partes}} + \underbrace{\square}_{\text{uma que}} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{vem}$$

⇒ cálculo desses valores gerados são computados

Tomaremos um estimador p/  $I(\theta) \rightarrow \hat{I}(\theta)$

Deas formas de construir  $\hat{I}(\theta) \rightarrow$  ver Hamilton (ver notas)

$$\begin{aligned}\hat{I}(\theta) &= \left\{ \begin{array}{l} -\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta \partial \theta'} = -\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta' \partial \theta} \end{array} \right|_{\theta=\hat{\theta}} \xrightarrow{\text{I}_{20}} \text{(I)} \\ (\text{estimativa da matriz de informações}) \quad &\quad \text{ou} \\ &= \left\{ \begin{array}{l} \frac{1}{n} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} = \frac{1}{n} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta} \left( \frac{\partial l_t}{\partial \theta} \right)' \end{array} \right|_{\theta=\hat{\theta}} \xrightarrow{\text{I}_{op}} \text{(II)} \end{aligned}$$

Na prática, essas derivadas são calculadas numericamente (nas analiticamente)

$$\frac{de}{dy_j} = \lim_{h \rightarrow 0} \frac{l(y_j+h) - l(y_j)}{h}$$

Supondo que conseguimos fazer cálculos analiticamente

$$\text{Var}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = \frac{1}{n} \hat{I}^{-1}(\theta) \quad \text{(III)}$$

por definição  
da dist. amostral  
de  $\hat{\theta}$

De (I) e (II) em (III), temos:

$$\Rightarrow \text{var}(\hat{\theta}) = I^{-1}(\theta) = \left\{ = \left[ \sum \frac{\partial^2 \ell_T}{\partial \theta \partial \theta^T} \Big|_{\theta=\hat{\theta}} \right]^{-1} \quad (\text{obs: cancelar } Y_T) \right.$$

$$\left. \left[ \sum \left( \frac{\partial \phi_i}{\partial \theta} \left( \frac{\partial \phi_i}{\partial \theta} \right)^T \right) \right]^{-1} \Big|_{\theta=\hat{\theta}} \right)$$

Estimamos a matriz de info. p/  $\theta$ , que são os parâmetros transformados

$$\hat{\theta}_j \pm z \sqrt{\text{var}(\hat{\theta}_j)}$$

veio da diagonal de  $I^{-1}(\theta)$

Mas queremos nos parâmetros originais:

$$\hat{\psi}_j \pm z \sqrt{\text{var}(\hat{\psi}_j)}$$

$\Rightarrow$  passar de  $\hat{\theta}$  p/  $\hat{\psi}$

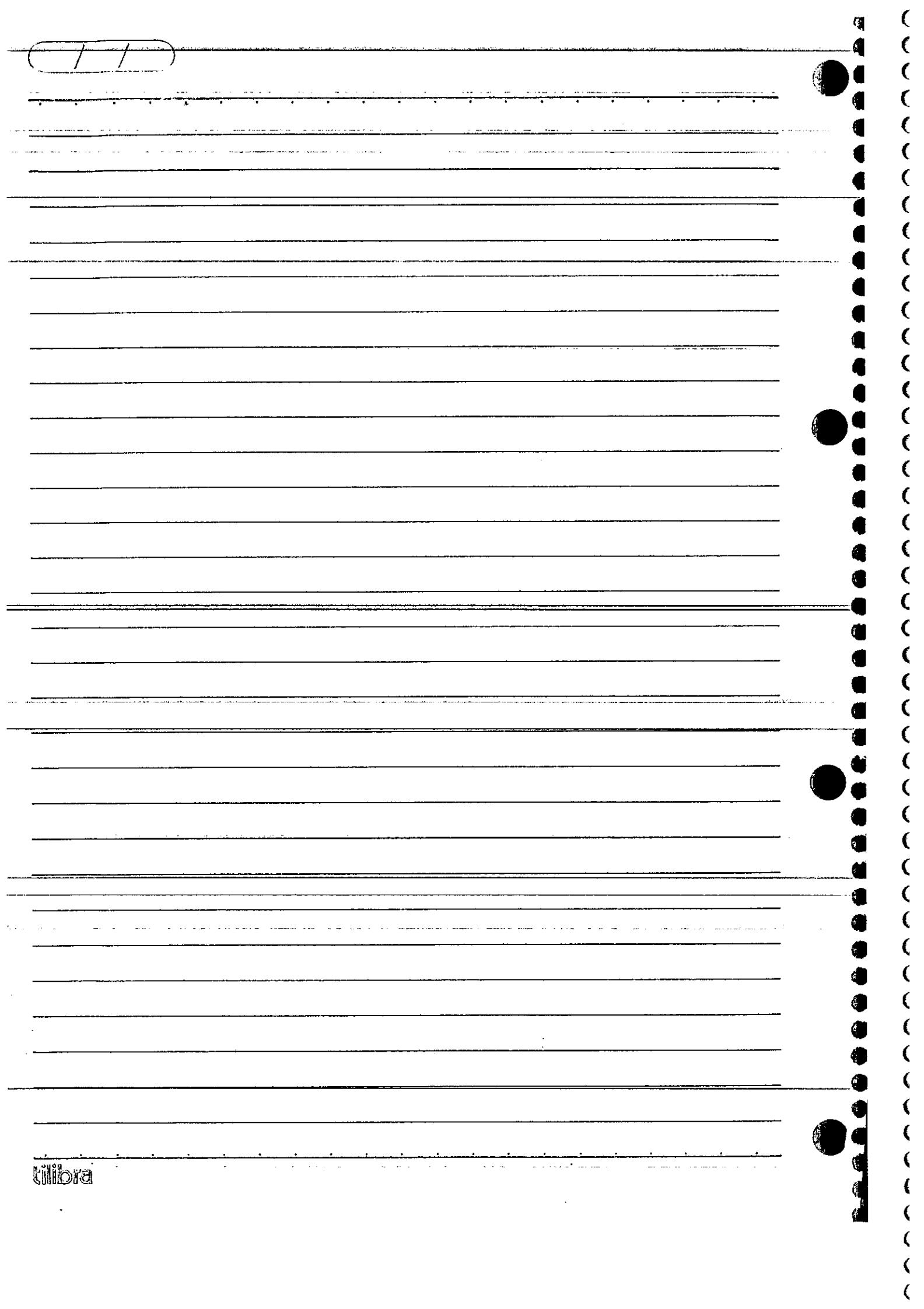
Para o vetor  $\hat{\psi}$ , temos:  $\text{var}(\hat{\psi}) = T \text{var}(\hat{\theta}) T^T$

? inversa  
(rever)

transformações de  $\text{var}(\hat{\theta})$  em  
 $\text{var}(\hat{\psi})$

obs: Para cada  $\psi_j$ , temos

$\psi_j = T_j(\theta) \Rightarrow$  transformações de  $\psi$  original p/  
 $\theta$  invertido



Ulibra

cont. otimizações

$$(*) \quad \alpha_t = T_{t-1} \alpha_{t-1} + R_{t-1} \eta_{t-1}$$

$$\therefore E[\alpha_t] = T_{t-1} \alpha_{t-1}$$

$$\text{var}(\alpha_t | \alpha_{t-1}) = E[(\alpha_t - E[\alpha_t | \alpha_{t-1}])(\alpha_t - E[\alpha_t | \alpha_{t-1}])']$$

$$\text{mas } \alpha_{t-1} - E[\alpha_t | \alpha_{t-1}] = R_{t-1} \eta_{t-1}$$

$$\therefore \text{var}(\alpha_t | \alpha_{t-1}) = E[R_{t-1} \eta_{t-1} \eta_{t-1}' R_{t-1}']$$

$$= R_{t-1} \Phi_{t-1} R_{t-1}' = \text{var}(\alpha_t | \alpha_{t-1})$$

f

Esta seria a variância de  $\alpha_t | \alpha_{t-1}$

mas não podemos escrever dessa

forma. Tem a ver com a natureza

da matriz  $R$ .

Ver exemplo (\*\*)

Temos problema para calcular  $\det(R_{t-1} \Phi_{t-1} R_{t-1}')$  que aparece na variação.

Problema:

Temos uma transformação do tipo.

$$z = a + b x$$

$$\text{para calcular } f_y(y) \Rightarrow f_y(y) = f_x \left( x = \frac{y-a}{b} \right) |J|$$

No caso, temos exatamente isso:

$$\alpha_t = T_{t-1} \alpha_{t-1} + R_{t-1} \eta_{t-1} \xrightarrow{\sim N(0, \Phi_{t-1})}$$

(11) Como  $R$  é nxn ( $\text{pois } x = nx \times 1 \text{ e } \eta = nx \times 1$ )

ou seja, temos dimensões de  $x$  & dimensões de  $\eta$ , só calcular  $p(\alpha)$  precisamos usar transformações usando jacobiano.

Nas foram apresentadas todas as contas.

Mas consideraremos sempre  $R_t$  da forma

$$R_t = [I_n \ 0']' \Rightarrow R_{t-1} R_{t-2} = I_n$$

Trataremos:  $\log |R_{t-1} R_{t-2} R_{t-1}'| \rightarrow \log R_{t-2}$ : precisa fazer o cálculo das transformações (não está explicitado)

Na verdade, teremos:  $(\alpha_t - T_{t-1} \alpha_{t-1})' (R \varphi R)^{-1} (\alpha_t - T_{t-1} \alpha_{t-1})$

↓ quando usamos o jacobiano

$$(\alpha_t - T_{t-1} \alpha_{t-1})' R_{t-2} Q^{-1} R_t' (\alpha_t - T_{t-1} \alpha_{t-1})$$

↓

$$\eta_{t-1}'$$

$$\eta_{t-1}$$

$$\text{de } \alpha_t = T_{t-1} \alpha_{t-1} + R_{t-1} \eta_{t-1}$$

$$\therefore \eta_{t-1} = R_{t-1}' (\alpha_t - T_{t-1} \alpha_{t-1})$$

— " —

(\*\*\*)  $\Rightarrow$  resultados:

É uma fórmula analítica de calcular o vetor

de score (= gradiente)

Traça utilidade em 2 situações:

a) Em otimizações, métodos de Newton

"

$$\psi^{(i+1)} = \psi^{(i)} + \lambda H^{-1}(\psi^{(i)}) g(\psi^{(i)}) \Rightarrow \text{ótimo.}$$

tilibra

b. d. (pode ser analítico, d  $\psi$  mas tbx n'tivermos info. para ser calculado)

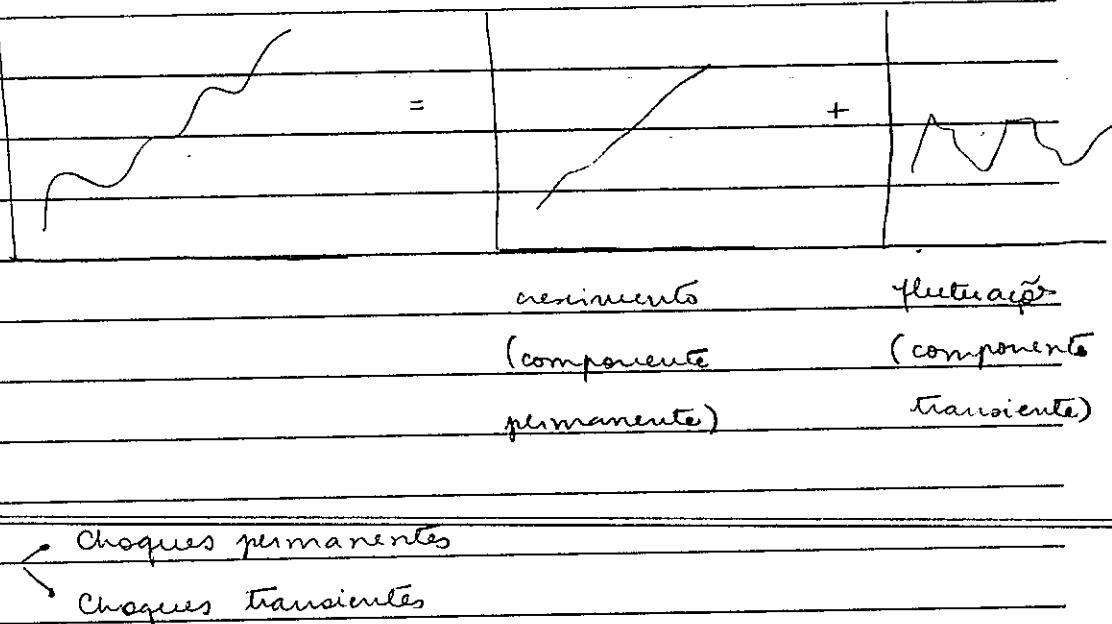
máscara de otimizações de qualquer função objetivo.

## → Macroeconomia 1:

tendência + irregular/ciclo

↳ em economia há irregularidades,

na frequência, no tempo etc.



## → Macroeconomia 2:

tendência + sazonalidade + irregular

obj:

Para esta metodologia, é preciso que se  
preocupe em estacionarizar a série:

(como fazemos nos modelos ARIMA - Box & Jenkins)

$$\hookrightarrow \hat{y}_t = \Delta^p S^d Y_t$$

Aqui, ajusta-se a série original desde que ela  
tenha as componentes propostas.

Se uma série é estacionária de 2<sup>a</sup> ordem, vale:

outro algoritmo  $\Rightarrow$  EM

por maximização



Tb é método iterativo p/ achar máximos da função, mas toma + a cara da função que está sendo maximizada

Toma a direção do máximo, mas fica lento próximo do máximo.

Normalmente: usa-se o EM p/ ter uma cond. inicial melhor.

Quando começar a ficar lento  $\Rightarrow$  joga condição inicial no algoritmo quasi-Newton.

Vantagem: ter uma cond. inicial melhor.

usar apenas 1<sup>a</sup> derivada (no EM)

No quasi-Newton entra 2<sup>a</sup> derivada tb.

## Modelos

Vamos olhar p/ ver como componentes nos observáveis.

ST = Tendência + Sazonalidade + Cíclo + Irregular



demandas

estados do

aumento de

ano, conven-

população,

cões sociais

aumento de

riqueza

representações com MA infinito:

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

$$\begin{aligned} y_t^* &= y_t - \mu = \sum \psi_j \\ &= \Phi(L) \epsilon_t \\ &= \Theta_q(L) \epsilon_t \quad \Rightarrow \text{ARMA} \end{aligned}$$

$$\Phi_p(L) y_t^* = \Theta_q(L) \epsilon_t$$

Mas x séries não estacionárias, mas há formalizações de que componentes devem ser usadas, justificativa.

Mas são ferramentas bem usadas (tendência + sazonalidade + ...)

### Fatores estilizados



de séries de retornos financeiros, por exemplo:

- excesso de autocorrelação etc.

de séries macroeconômicas

- tendência e ciclo

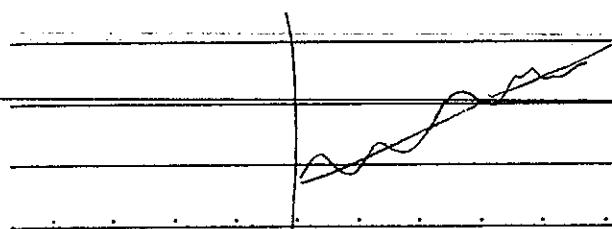
- tendência e sazonalidade

• Estas componentes mais estocásticas e locais

+ determinística

+ global

Obs: se considerarmos tendência linear



$$y_t = a_0 + b_0 t + \hat{\epsilon}_t$$

$$g_t = \hat{a}_0 + \hat{b}_0 t$$

$y_t - \hat{g}_t = \hat{\epsilon}_t \Rightarrow$  fazendo assim, amplifica o erro, o ciclo.

(11) se tendência forse estocástica  $\Rightarrow$  "erro" direcional

Nelson e Plosser (1982)  $\Rightarrow$  propõem como melhor  
tendência estocástica

com tendência e sazonalidade

$$y_t = \mu_t + \gamma_t + \epsilon_t$$

tendência sazonalidade

Série ajustada dessazonalizando:

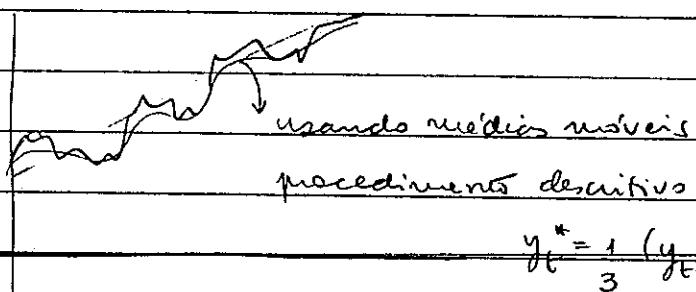
$$y_t^{(a)} = y_t - \hat{\gamma}_t \quad \left\{ \begin{array}{l} X12-ARIMA \\ SEATS-TRAMO \end{array} \right. \quad \begin{array}{l} \text{metodologia de dessazonalização} \\ (\text{IBGE}; \text{Séries dos EVA}) \end{array}$$

dependendo de como é especificado,  
pode ter séries ajustadas diferentes

### Modelos de Tendência

Componente de baixa frequência ("muda pouco ao longo do tempo")

Ex:



$$\hat{y}_t^* = \frac{1}{3} (y_{t-1} + y_t + y_{t+1})$$

eliminação de altas frequências  
Assim nos conseguimos tirar componente sazonal

11

Obs importante:

Precisa especificar a correlação entre os componentes.

Se pode ser 0 ou  $\pm \frac{1}{2}$ .

Se for  $\neq 0$ , sistema é não identificável

Dar ex. de  $F_K \Rightarrow$  achar eqs. plurais

110

Kiloid

(1)

Da equações (I) do modelo:

$$\begin{aligned} E[y_t | y_{t-1}] &= E[z_t \alpha_t + \varepsilon_t | y_{t-1}] \\ &= z_t E[\alpha_t | y_{t-1}] \\ &= z_t \alpha_t \end{aligned}$$

Logo,  $v_t = y_t - z_t \alpha_t$

intuitivas: o que queremos dizer é que o conteúdo informacional  
mas é modificado se substituirmos  $y_t$  por  $v_t$ .  
(pois  $\alpha_t$  só depende de passado)

Poderemos fazer isso pois  $g$  é bijetiva.

É como se  $g(y)$  fosse  $v_t$  (a inovação).

Entas:

$$E[\alpha_t | y_t] = \alpha_t|x = E[\alpha_t | y_{t-1}, v_t]$$

Usando (a): fazendo  $x = \alpha_t$

$$y = y_{t-1}$$

$$z = v_t$$

se usarmos (a),

precisamos provar

que a situação é de

vectores tri-variados onde  $\mu_z = 0$  e  $\Sigma_{zy} = 0$ .

Temos que provar que estas condições são satisfeitas.

$$(i) E[z] = \mu_z = 0 \Rightarrow E[v_t] = 0$$

$$E[v_t] = E[E[v_t | y_{t-1}]] \quad (\text{lei da expectativa iterada})$$

$$\begin{aligned}\therefore E[v_t | y_{t-1}] &= E[y_t - z_t a_t | y_{t-1}] = \\ &= E[y_t | y_{t-1}] - z_t a_t \\ \text{do modelo} &= z_t a_t - z_t a_t = 0.\end{aligned}$$

$$\therefore E[v_t] = E[0] = 0 \quad \text{ok! (condição (i) é satisfeita)}$$

$$(ii) \sum y_t = 0 \Rightarrow \sum_{t-1} y_t = \text{cov}(y_{t-1}, v_t) = 0$$

$$\text{cov}(y_{t-1}, v_t) = E[(y_{t-1} - E[y_{t-1}]) (v_t - E[v_t])']$$

0 (já provamos)

$$y_{t-1} = \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_1 \end{pmatrix} \quad p \times 1$$

$$= E[y_{t-1} v_t']$$

(cada elemento desta  
matriz será dada por:

$$\{ E[y_{t-j} v_t'] \}_{j=1,2,\dots,t-1}$$

$$\text{mas: } E[y_{t-j} v_t'] = E[E[(y_{t-j} v_t') | y_{t-1}]]$$

letras  
exp. iteradas  
para calcular  
valor esperado  
incondicional

como  $t-j$  é passado  
pode sair de  $E[\cdot]$

$$= E[y_{t-j} E[(v_t' | y_{t-1})]] = 0$$

= 0 (ii) ok!

Por tanto, podemos usar o resultado (a):

$$E[x_t | y_t] = E[x_t | y_{t-1}] + \text{cov}[x_t, v_t] (\text{var}[v_t])^{-1} v_t$$

$a_t$   
já sabemos

quem é

Precitamos calcular

(passo da prova.  
em  $t-1$ )

as outras componentes  
da equação.

04/05/11

Tendência e ciclo: caráter subjetivo

depende muito de onde e de que horizonte estamos olhando  
a série

Normalmente, ciclo pode ser tratado tb de forma não periódica

podemos ajustar uma tendência estocástica (e variabilidade em um ciclo).

em séries econômicas,  
por exemplo.

Tendência:

- Determinística (TS)  
(no tempo)
- Estocástica (DS)

a) Tend. determinística

processo ARMA

$$y_t = \sum_{j=0}^d \beta_j t^j + \psi(L) a_t, \quad a_t \sim NID(0, \sigma^2)$$

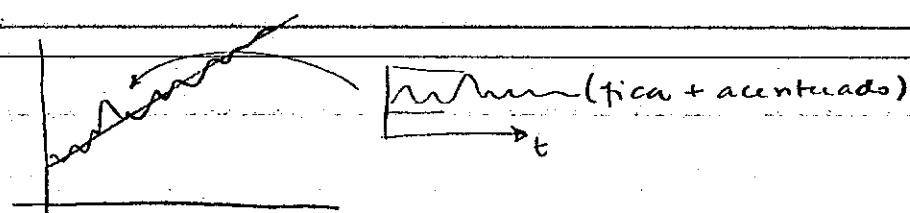
(podermos ter log  $y_t$ , ou inverso de  $y_t$ )

onde  $\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$  (n colocarmos apenas nulos bco, ainda terá correlações)

puncas  
polinomial do tempo (normalmente  $t=2$ )

9º mais suave para a tendência  $\Rightarrow$  maior é o ciclo (flutuação :  
em torno da tendência)

(mas é adequado pois nas séries econômicas, a tendência  
é essencialmente estocástica)



$\Rightarrow$  Modelo típico de ST : parte estacionária é ARMA  
" " " é tend. determ.

cov e autocovar deste modelo  $\Rightarrow$  vem da ARMA

Ex:  $d=1$  e  $a_t = AR(1)$

boa correção PRB mas ter autocorrelações.

$$y_t = \beta_0 + \beta_1 t + \alpha_t$$

$$\alpha_t = \phi \alpha_{t-1} + \epsilon_t$$

$$(1 - \phi L) \alpha_t = \epsilon_t$$

$$\therefore \alpha_t = \frac{\epsilon_t}{1 - \phi L}$$

$$y_t = \beta_0 + \beta_1 t + \frac{\epsilon_t}{1 - \phi L}$$

$$(1 - \phi L) y_t = (1 - \phi L) \beta_0 + (1 - \phi L) \beta_1 t + \epsilon_t$$

$$(1 - \phi L) y_t = (1 - \phi L) \beta_0 + \beta_1 (t - \phi(t-1)) + \epsilon_t$$

$$(1 - \phi L) y_t = (1 - \phi) \beta_0 + \beta_1 (t - \phi t + \phi) + \epsilon_t$$

$$(1 - \phi L) y_t = (1 - \phi) \beta_0 + \beta_1 \phi + (\phi - 1) \beta_1 t + \epsilon_t$$

Queremos:

$$w_t = y_t - y_{t-1}$$

$$y_t - y_{t-1} = (1 - L) y_t = (1 - L) \left( \beta_0 + \beta_1 t + \frac{\epsilon_t}{1 - \phi L} \right)$$

$$= \beta_1 (t - (t-1)) + \frac{\epsilon_t}{1 - \phi L} (1 - L)$$

$$= \beta_1 + (1 - L) \frac{\epsilon_t}{1 - \phi L}$$

$$\therefore (1 - \phi L)(1 - L) y_t = (1 - \phi) \beta_1 + (\epsilon_t - \epsilon_{t-1})$$

MA de raiz unitária  
( $\theta = 1$ )

Nas e' bom pois nas e' inversivel (problema  
na identificacao, na previsao,  
etc.)

$\Rightarrow$  Processo q tendencia deterministica; n e' recomendado  
tirar tendencia  
por 1º diferenca

(\*\*) Modelo de tendência linear estocástica

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t$$

$$\beta_{t+1} = \beta_t + \zeta_t$$

→ quando  $\sigma_\eta^2 = \sigma_\zeta^2 = 0 \Rightarrow$  passa a ser determinística.

$\left\{ \begin{array}{l} H_0: \sigma_\eta^2 = \sigma_\zeta^2 = 0 \Rightarrow \text{tão complicado de fazer pois estamos} \\ \text{na fronteira do espaço paramétrico (pois var} \\ \text{entre } 0 \text{ e } \infty) \\ H_A: \text{c.c.} \end{array} \right.$

$\left( \begin{array}{l} \sigma_\eta^2 = 0 \\ \sigma_\zeta^2 \neq 0 \end{array} \right) \text{ ou } H_0 = \text{equivalente ao processo TS}$

$H_A = " " " " " \text{ DS.}$

$\left( \begin{array}{l} \sigma_\eta^2 \neq 0 \\ \sigma_\zeta^2 = 0 \end{array} \right) \text{ ou } \left( \begin{array}{l} \sigma_\eta^2 \neq 0 \\ \sigma_\zeta^2 \neq 0 \end{array} \right)$

→ Diferente do caso das regressões em que testamos  $H_0: \beta_1 = \beta_2 = 0$

$$(y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon_t) \quad \text{pois } \beta_1, \beta_2 \in \mathbb{R}^2$$

Ajusta o modelo e estima  $\hat{\sigma}_\eta^2$  e  $\hat{\sigma}_\zeta^2$

" "  
se estimativas muito pequenas: evidência  
de tendência  
determinística

(obs: Na verdade, devemos olhar  $q_\eta$  e  $q_\zeta$ )

" "

→ Vantagem do arcabouço das ME: há uma transição suave entre  
TS e DS

é geral o suficiente para caracterizar os 2 tipos.

Na outra formatação, é necessário pré-definir

no caso ME: vai se ajustar naturalmente. Já faria na estimativa  
dos parâmetros (das variâncias)

" "

Artigo: Nelson e Plosser.

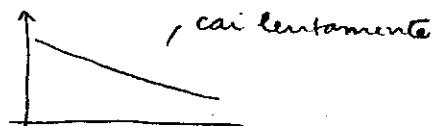
- Séries tabela 2  $\Rightarrow$  cálculo das autocorrelações de cada série  
(caem lentamente, exceto tx. desemprego)

- tabela 3  $\Rightarrow$  pega a 1<sup>a</sup> diferença

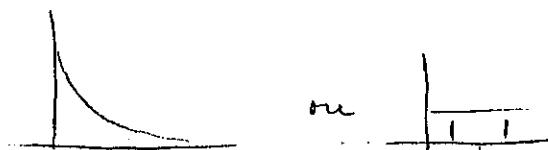
1)  $z_t = \log y_t \rightarrow$  FAC (tabela 2)

2)  $\Delta z_t = \Delta \log y_t \rightarrow$  FAC (tabela 3)

Se  $z_t = \alpha + z_{t-1} + a_t \Rightarrow$  FAC



Se  $\Delta z_t = \alpha + a_t \Rightarrow$  FAC



$\Rightarrow$  típico de processo estacionário

• Pode ser consistente com TS ou com DS.

Não é suficiente pf diferenciar ainda

3)  $z_t = \beta_0 + \beta_1 t + a_t$  (tabela 4)

$\hat{a}_t = z_t - (\hat{\beta}_0 + \hat{\beta}_1 t) \rightarrow$  FAC

$\rightarrow$  Faz regressões estimando via MPO

Estima processo DS via TS

$\rightarrow$  FAC decai lentamente pf o erro

Muita estrutura de dependência

$\gamma$  típico de séries DS  
estimada via TS

Com 1, 2, 3  $\Rightarrow$  evidência de que para séries macroeconómicas,  
grande maioria é DS (e não TS)

Além das evidências:

Faz teste da raiz unitária p/ o modelo considerando tendência estocástica e determinística.

Tabela 5: estimativa de  $\delta$  (na parcela  $\delta t$ )  $\Rightarrow$  muito pequeno  
estimativa de  $\rho$  (na parcela  $\rho \cdot z_{t-1}$ )

$$\text{Estatística: } t = \frac{\hat{\rho} - 1}{\text{SE}(\hat{\rho})}$$

$\rightarrow$  P1 maioria das séries n<sup>o</sup> pode rejeitar  $H_0: \rho = 1 \Rightarrow$  tem raiz unitária

$H_0: \delta = 0 \rightarrow$  pode ser  
testado pela  
estatística t-padrão

$\hookrightarrow$  mas pode usar estatística  
t-padrão (atarrada na  
fronteira do espaço  
paramétrico)

$t_b$  nas rejeita  $\delta = 0$  p/ algumas delas estatisticamente.  
mas em termos absolutos,  $\hat{\delta}$  é muito pequeno.  
é significante, mas não é relevante.

$\rightarrow$  Evidência de tendência estocástica.

(típico de séries macroeconómicas)

