Frances Blan

# Modelos Estruturais p/ Séries Temporais

2011.1

Prof: Cristiano Fernandes

## 1ª parte (P1)

Aula 02/03

## Modelo Espaço-Estado

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t$$ (equação de medida)

$$\alpha_{t+1} = T_t \alpha_t + \cancel{c_t} + R_t \eta_t$$ (evolução do estado) (equações de estado)

$y_t$ se relaciona com $\alpha_t$

$\alpha_t$ se relaciona com $\alpha_{t-1}$

Ex. 1) $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$

1ª equação de medida

normalmente, tem que ter $y_t = y_t$

$$y_t = \underbrace{(1 \quad 0)}_{Z_t} \underbrace{\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}}_{\alpha_t} + \varepsilon_t$$

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

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

$$\underbrace{\begin{pmatrix} y_{t+1} \\ y_t \end{pmatrix}}_{\alpha_{t+1}} = \underbrace{\begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix}}_{T_t} \underbrace{\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}}_{\alpha_t} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_{t+1}$$

Outra representação, de forma que $\alpha_t^* = B\alpha_t$ onde $\det(B) \neq 0$

⇒ fção verossimilhança será a m/me

$$y_t = (1 \quad 0) \begin{pmatrix} y_t \\ \phi_2 y_{t-1} \end{pmatrix} + \varepsilon_t$$

$$\underbrace{\begin{pmatrix} y_{t+1} \\ \phi_2 y_t \end{pmatrix}}_{\alpha_{t+1}^*} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} y_t \\ \phi_2 y_{t-1} \end{pmatrix}}_{\alpha_t^*} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_t$$

$B \Rightarrow \det B \neq 0$

$$\Rightarrow \alpha_t^* = \begin{pmatrix} 1 & 0 \\ 0 & \phi_2 \end{pmatrix} \alpha_t$$

Ex. 2) $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$

$$\begin{cases} y_t = Z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \end{cases} \rightarrow \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}$$

$$y_t = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} y_{t+1} \\ y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ 1 & 0 & \dots & 0 \\ & \ddots & & \vdots \\ 0 & & & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \varepsilon_{t+1}$$

---

$$\hookrightarrow y_{t+1} = \phi_1 y_t + \phi_2 y_{t-1} + \dots + \phi_p y_{t-p+1} + \varepsilon_{t+1}$$

---

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

$y_t = \theta \varepsilon_{t-1} + \varepsilon_t \Rightarrow y_{t+1} = \theta \varepsilon_t + \varepsilon_{t+1}$

$$\begin{cases} y_t = Z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \end{cases} \rightarrow \begin{pmatrix} y_t \\ \varepsilon_t \end{pmatrix}$$

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \varepsilon_t \end{pmatrix} + \varepsilon_t$$

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

$\varepsilon_{t+1}$

<u>Escrever o ARMA.</u>

Ex. 4) $y_t = \phi y_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t$ (ARMA)

$$\begin{cases} y_t = Z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \end{cases} \rightarrow \begin{matrix} y_t \\ \varepsilon_t \\ \varepsilon_{t-1} \end{matrix}$$

$y_{t+1} = \phi y_t + \theta \varepsilon_t +$

$$y_t = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} y_{t+1} \\ \varepsilon_{t+1} \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} \phi & \theta & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \varepsilon_{t+1}$$

---

Ex. 5) Regressão múltipla

---

ver solução caderno de outra pessoa.

Aula
16/03.

## Modelo

$$\begin{cases} y_t = Z_t \alpha_t + d_t + \varepsilon_t , & \varepsilon_t \sim N(0, H_t) \\ \alpha_{t+1} = T_t \alpha_t + c_t + R_t \eta_t , & \eta_t \sim N(0, Q_t) \end{cases}$$

$$\begin{cases} \alpha_1 \sim N(a_1, P_1) & \Rightarrow \text{cond. inicial} \\ E[\varepsilon_t \eta_s'] = 0 \quad \forall\, t, s & \Rightarrow \text{choques descorrelaci dos (pode ser relaxad} \\ E[\eta_t' \alpha_1) = E[\varepsilon_t' \alpha_1] = 0 \quad \forall\, t & \Rightarrow \text{jamais pode relaxa} \\ E[\varepsilon_t \varepsilon_s') = 0 \quad \forall\, s, t & \Rightarrow \text{ruídos brancos} \\ E[\eta_t \eta_s'] = 0 \quad \forall\, s, t & \end{cases}$$

Propriedades dos MEE lineares a Tempo Discreto.

(1) Estacionariedade de 2ª ordem

(2) Observabilidade

(1). Estacionariedade de 2ª ordem

Processo estocástico $X_t$ é dito estacionário sss as seguintes condições são satisfeitas:

(i) $E[X_t] = \mu \quad \forall t, \ \mu \leq \infty$

(ii) $E[(x_t - E[x_t])(x_{t+h} E[x_{t+h}])') = \Gamma(h) \quad \forall t, \ \forall h = 0, 1, 2 \ldots$

Equações do MEE:

$$\begin{cases} y_t = Z \alpha_t + \varepsilon_t & (\text{fazendo } d_t \text{ e } c_t = 0 \text{ spg}) \\ \alpha_{t+1} = T \alpha_t + R \eta_t & (\text{fazendo } Z_t = Z, T_t = T, R_t = R) \end{cases}$$

Cond (i): $E[y_t] = \underset{\substack{\downarrow \\ \text{depende de } E[\alpha_t]}}{Z E[\alpha_t]} + E[\underset{0}{\cancel{\varepsilon_t}}]$

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

avaliar $x = \Gamma_y(h)$

- $y_t - E[y_t] = Z\alpha_t + \varepsilon_t - Z E[\alpha_t]$

$= Z(\alpha_t - E[\alpha_t]) + \varepsilon_t$

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

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

$= [Z(\alpha_t - E[\alpha_t]) + \varepsilon_t][(\alpha_{t+h} - E[\alpha_{t+h}])' Z' + \varepsilon'_{t+h}]$

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

$+ Z(\alpha_t - E[\alpha_t])\varepsilon'_{t+h} + \varepsilon_t(\alpha_{t+h} - E[\alpha_{t+h}])' Z' +$

$+ \varepsilon_t \varepsilon'_{t+h}$

Passando valor esperado:

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

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

$= Z \Gamma_\alpha(h) Z'$

- $E[Z(\alpha_t - E[\alpha_t])\varepsilon'_{t+h}) =$

$= Z E[(\alpha_t - E[\alpha_t])\varepsilon'_{t+h}] =$

$= Z \underbrace{E[\alpha_t \varepsilon'_{t+h}]}_{0} - Z E[\alpha_t] E[\underset{0}{\cancel{\varepsilon_{t+h}}}] = 0$

porquê? ⊛

$$\cdot\ E[\varepsilon_t(\alpha_{t+h} - E[\alpha_{t+h}])'Z'] =$$

$$= E[\underset{0}{\varepsilon_t \alpha_{t+h}}]Z' - E[\underset{0}{\varepsilon_t}]E[\alpha_{t+h}]Z' = 0$$

por quê? ⊛

$$\cdot\ E[\varepsilon_t \varepsilon'_{t+h}) = \begin{cases} 0 & \text{se } h \neq 0 \\ H & \text{se } h = 0 \end{cases}$$

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

$$\alpha_2 = T\alpha_1 + R\eta_1$$

$$\alpha_3 = T\alpha_2 + R\eta_2$$
$$= T(T\alpha_1 + R\eta_1) + R\eta_2$$
$$= T^2\alpha_1 + TR\eta_1 + R\eta_2$$

$$\alpha_4 = T\alpha_3 + R\eta_4$$
$$= T(T^2\alpha_1 + TR\eta_1 + R\eta_2) + R\eta_3$$
$$= T^3\alpha_1 + T^2R\eta_1 + TR\eta_2 + R\eta_3$$

$$\Rightarrow \boxed{\alpha_{t+1} = T^t\alpha_1 + \sum_{i=0}^{t-1} T^i R \eta_{t-i}}$$

$$E[\alpha_t \varepsilon_{t+h}] =$$

$$E\left[\left(T^{t-1}\alpha_1 + \sum_{i=0}^{t-2} T^i R \eta_{t-i}\right)\varepsilon'_{t+h}\right]$$

$\Rightarrow$ Termos com $E[\alpha_1 \varepsilon'_{t+h}] = 0$.

com $E[\eta_{t-i}\varepsilon'_{t+h}] = 0$.

Logo:

$$E[(y_t - E[y_t])(y_{t+h} - E[y_{t+h}])'] = \begin{cases} Z\Gamma_\alpha(0,t)Z' + H & \text{se } h = 0 \\ Z\Gamma_\alpha(h,t)Z' & \text{se } h \neq 0 \end{cases}$$

EX1:

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

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

$|\phi| < 1.$

(*) Iterando $\mu_{t+1}$.

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

$\mu_3 = \phi\mu_2 + \eta_2$

$= \phi^2\mu_1 + \phi\eta_1 + \eta_2$

$\mu_4 = \phi\mu_3 + \eta_3$

$= \phi^3\mu_1 + \phi^2\eta_1 + \phi\eta_2 +$

$\Rightarrow E(\mu_{t+1}) = \underbrace{\phi^t E(\mu_1)}_{\text{tende a zero se } |\phi|<1} + 0$

$\mu_t$: segue AR(1)

porque (*) $\Rightarrow$ $E[\mu_t] = 0$

$\text{var}[\mu_t] = \gamma_\mu(0) = \dfrac{\gamma_\eta^2}{(1-\phi^2)}$

dúvida de $\mu_t$ e $\mu_{t+1}$

$\gamma_\mu(h) = \phi^h \dfrac{\sigma_\eta^2}{(1-\phi^2)}$

funçaõ só de h

$\rho_\mu(h) = \dfrac{\gamma(h)}{\gamma(0)} = \phi^h$

Estudar modelos AR (Hamilton).

$\Rightarrow$ AR(1) é estacionário se $|\phi| < 1$.

(senaõ $\gamma_\mu(h)$ explode). mas teremos $\lim_{h\to\infty} (\gamma_\mu(h)) < \infty$.

- $E[y_t] = E[\mu_t] + E[\varepsilon_t] = 0.$

- $\text{var}[y_t] = \text{var}[\mu_t] + \text{var}[\varepsilon_t]$

obs: $\mu_t$ e $\varepsilon_t$ independentes

$\text{var}[\mu_t + \varepsilon_t] = \text{var}(\mu_t) + \text{var}(\varepsilon_t)$

variância da v.a. no mesmo inst.

$\gamma_y(0) = \gamma_\mu(0) + \sigma_\varepsilon^2$

$= \dfrac{\sigma_\eta^2}{(1-\phi^2)} + \sigma_\varepsilon^2$

covariância da variável com lag

- $\gamma_y(h) = E[(y_t - \underset{0}{E[y_t]})(y_{t+h} - \underset{0}{E[y_{t+h}]})']$

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

$= E[(\mu_t + \varepsilon_t)(\mu_{t+h} + \varepsilon_{t+h})']$

$= E[\mu_t\mu_{t+h}] + \underset{0}{E[\mu_t\varepsilon_{t+h}']} + \underset{0}{E[\varepsilon_t\mu_{t+h}']} + \underset{0}{E[\varepsilon_t\varepsilon_{t+h}']}$

$= \gamma_\mu(h)$

2) $y_t = \mu_t + \varepsilon_t$

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

obs: mesmo anterior, mas $\phi = 1$.

cond. inicial

$\mu_1 \sim N(a_1, P_1)$

e todas as outras inicialmente expostas

- $E[\mu_t] = ?$

Iterando $\mu_{t+1}$:

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

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

$$\mu_4 = \mu_3 + \eta_3 = \mu_1 + \eta_1 + \eta_2 + \eta_3 \cdots$$

$$\mu_{t+1} \boxed{\mu_1 + \sum_{i=1}^{t} \eta_i} \Rightarrow E[\mu_{t+1}] = E[\mu_1] = a_1 \Rightarrow$$ da cond. inicia

$$\therefore E[\mu_{t+1}] = a_1$$

- $\text{var}[\mu_{t+1}] = \gamma_\mu(0)$

$$= E[(\mu_{t+1} - a_1)(\mu_{t+1} - a_1)')$$

$$= E\left[\left(\mu_1 + \sum_{i=1}^{t} \eta_i - a_1\right)\left(\mu_1 + \sum_{i=1}^{t} \eta_i - a_i\right)'\right]$$

$$= E[\mu_1 \mu_1'] + \underset{0}{E\left[\mu_1 \sum_{i=1}^{t} \eta_i'\right]} + \underset{0}{E\left[\sum_{i=1}^{t} \eta_i \mu_1'\right]} + E\left[\left(\sum_{i=1}^{t} \eta_i\right)\left(\sum_{j=1}^{t} \eta_j\right)'\right]$$

p/ $i = j \Rightarrow \sigma_\eta^2$

p/ $i \neq j \Rightarrow \varnothing$

Dúvida

$\text{var}(\mu_t) = P_1 + (t-1)\sigma_\eta^2$ ?

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

$\Rightarrow$ var $(\mu)$ depende de $t$

Processo não será estacionário.

- $E[y_t] = E[\mu_t] + E[\varepsilon_t] = a_1$

$$\cdot\ \text{var}[y_t] = \gamma_y(0)$$

$$= \text{var}(\mu_t) + \text{var}(\varepsilon_t)$$

$$= \gamma_\mu(0) + \sigma_\varepsilon^2$$

$$= P_1 + \underbrace{(t-1)} \sigma_\eta^2 + \sigma_\varepsilon^2$$

No caderno está t.
Não deveria ser t-1?

⇒ depende de t: y não é estacionário.

Conclusão: Estacionariedade dependerá sempre do comportamento de $\boxed{\alpha_t}$.

---

— " —

Voltando à demonstração:

$$\cdot\ E[y_t] = Z\, E[\alpha_t]$$

$$\cdot\ \Gamma_y(h,t) = \begin{cases} Z\, \Gamma_\alpha(h,t)\, Z', & h \neq 0 \\ Z\, \Gamma_\alpha(0,t)\, Z' + H, & h = 0. \end{cases}$$

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

1ª condição: sobre média

$$\cdot\ E[\alpha_{t+1}] = T^t E[\alpha_1] = T^t a_1 = f(t)\ ?$$

↑
queremos avaliar se $E[\alpha_{t+1}]$ é função do tempo.

Supondo que todos os autovalores de T são distintos.

Lembrando: $T f_i = \lambda_i f_i \quad i = 1, 2, \ldots, m$

onde $\lambda_i$ autovalor
$f_i$ autovetor.

$$\Rightarrow \quad TF_i - \lambda_i F_i = 0$$

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

Como $F_i \neq 0$.

$\Rightarrow \det(T - \lambda_i I) = 0$ : equação característica dará $\textcircled{m}$ autovalores $\neq$s.

supondo ($\alpha$ m-variad

Vale a decomposição espectral:

$$T = F \Lambda (F^t) = F^{-1} \text{ (transp = inversa pois matriz ortogonal)}$$

Prova-se que:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \Big( ae + bg$$

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

$$\Rightarrow E[\alpha_{t+1}] = T^t a_1 = F \Lambda^t F^{-1} a_1$$

$$E\left[\begin{pmatrix} \alpha_{1,t+1} \\ \alpha_{2,t+2} \\ \vdots \\ \alpha_{m,t+1} \end{pmatrix}\right] = \underbrace{\begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & & \cdots & f_{2m} \\ & & & \vdots \\ f_{m1} & & \cdots & f_{mm} \end{pmatrix}}_{m \times m} \underbrace{\begin{pmatrix} \lambda_1^t & & \\ & \ddots & \\ 0 & & \lambda_m^t \end{pmatrix}}_{m \times m} \underbrace{\begin{pmatrix} f^{11} & f^{12} & f^{1m} \\ & & \\ f^{m1} & & f^{mm} \end{pmatrix}}_{m \times m} \Big($$

↓ Transposta de F

$$= \underbrace{\begin{pmatrix} f_{11}\lambda_1^t & f_{12}\lambda_2^t & \cdots & f_{1m}\lambda_m^t \\ f_{21}\lambda_1^t & f_{22}\lambda_2^t & \cdots & f_{2m}\lambda_m^t \\ \vdots & & & \vdots \\ f_{m1}\lambda_1^t & & \cdots & f_{mm}\lambda_m^t \end{pmatrix}}_{m \times m} \underbrace{\begin{pmatrix} f^{11} & f^{12} & \cdots & f^{1m} \\ f^{21} & f^{22} & \cdots & f^{2m} \\ \vdots & & & \\ f^{m1} & & \cdots & f^{mm} \end{pmatrix}}_{m \times m} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

$\textcircled{\text{fazer}}$ ↓

$$= \Big( \Sigma$$

Para cada $\alpha_j$ temos o seguinte p/ $\alpha_1$

$$E[\alpha_{1,t+1}] = \delta_1 \lambda_1^t + \delta_2 \lambda_2^t + \dots + \delta_m \lambda_m^t$$

onde $\delta_i = f_{1i} f^{1i} a_{1i}$, $i = 1, \dots, m$

$\downarrow$

depende explicitamente de t, mas...

se $|\lambda_i| < 1 \quad \forall i = 1, \dots, m$

fazendo: $\lim_{t \to \infty} E[\alpha_{i,t+1}] = 0.$

logo: $|\lambda_i| < 1$ $\forall i = 1, \dots m$ : condição necessária e suficiente p/ estacionariedade de 1ª ordem.

obs: tb será p/ de 2ª ordem

---

2ª condição: sobre variância

- $\Gamma_\alpha(h,t) = ?$

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

Explicitando $\alpha_{t+1}$ em função de $\alpha_1$.

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

$$E[\alpha_{t+1}] = T^t E[\alpha_t] = T^t a_1$$

$$\rightarrow \alpha_{t+1} - E[\alpha_{t+1}] = T^t(\alpha_1 - a_1) + \sum_{j=0}^{t-1} T^j R \eta_{t-j}$$

$$\rightarrow \alpha_{t+h+1} - E[\alpha_{t+h+1}] = T^{t+h}(\alpha_1 - a_1) + \sum_{j=0}^{t+h-1} T^j R \eta_{t+h-j}$$

Então

$$\Gamma_\alpha(h,t) = E\left[\left[T^t(\alpha_1 - a_1) + \sum_{j=0}^{t-1} T^j R \eta_{t-j}\right]\left[T^{t+h}(\alpha_1 - a_1) + \sum_{j=0}^{t+h-1} T^j R \eta_{t+h-j}\right]'\right.$$

$$= E\left[\left[T^t(\alpha_1 - a_1) + \sum_{j=0}^{t-1} T^j R \eta_{t-j}\right]\left[(\alpha_1 - a_1)' T^{t+h'} + \sum_{j=0}^{t+h-1} \eta'_{t+h-j} R' T^{j'}\right.\right.$$

Temos os seguintes termos:

- $E\left[T^t(\alpha_1 - a_1)(\alpha_1 - a_1)' T^{t+h'}\right] =$

$$= T^t E\left[(\alpha_1 - a_1)(\alpha_1 - a_1)'\right] T^{t+h'} =$$

$$= T^t P_1 T^{t+h'}$$

- $E\left[T^t(\alpha_1 - a_1) \sum_{j=0}^{t+h-1} \eta'_{t+h-j} R' T^{j'}\right]$

dará 0

- $E\left[\sum_{j=0}^{t-1} T^j R \eta_{t-j} (\alpha_1 - a_1)' T^{t+h'}\right]$

dará 0

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

Para $h=0$

$$\Gamma_\alpha(0,t) = T^t P_1 T^{t+h'} + \sum_{j=0}^{t-1} T^j (RQR') T'^j$$

$\vdots$

Procurar demo

Podemos demonstrar também que condição se resumirá a:

$|\lambda_i| < 1 \quad \forall i = 1, \dots, m$

e $t \to \infty$ $\Rightarrow \Gamma_\alpha(0,t) \to \Sigma(0)$ independente d

Obs:

Fazer condições de estacionariedade p/ AR(1):

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

· $\alpha_t = \alpha_1 \phi^t + \Sigma \phi^j \eta_{t-j}$

· $\alpha_{t+h} = \alpha_1 \phi^{t+h} + \Sigma \phi^j \eta_{t+h-j}$

Calcular $\gamma(h,t) = E\left[(\alpha_t - E(\alpha_t))(\alpha_{t+h} - E(\alpha_{t+h}))'\right)$

Supondo cond. de estacionariedade de 2ª ordem satisfeitas.

$|\lambda_i(t)| < 1$ e $t \to \infty$.

Obter expressão geral p/ a matriz de covariância de $\alpha_t$:

Já vimos que: $E[\alpha_{t+1}] = 0$.

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

obs: operações em AL: ver caderno.

$$\begin{cases} vec(A) \\ A \otimes B \\ vec(ABC) = (C' \otimes A)\, vec B \end{cases}$$

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

$$var(\alpha_{t+1}) = T \underbrace{var(\alpha_t)}_{\Sigma_\alpha(0)} T' + R \underbrace{var(\eta_t)}_{Q\text{ : ct. por hipótese.}} R'$$

$\downarrow$

$= var(\alpha_t)$

$= \Sigma_\alpha(0)$

$$\Sigma_\alpha(0) = T\,\Sigma_\alpha(0)\,T' + RQR'$$

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

$$= (T \otimes T)\, vec(\Sigma_\alpha(0)) + (R \otimes R')\, vec\, Q$$

$$(I - (T \otimes T))\, vec(\Sigma_\alpha(0)) = (R \otimes R')\, vec\, Q$$

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

Dúvida: como e quando usa?

Exercícios:

Considere os seguintes modelos em EE.

(1) $$y_t = \underset{\downarrow\, Z}{[1 \quad 0]} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

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

$$\underset{\downarrow}{\tilde{\eta}_t} = \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \Rightarrow E[\tilde{\eta}_t \tilde{\eta}_t'] = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix} = \underset{\downarrow}{Q}$$

da notação anterior

da notação anterior

**obs:** Equações
$$\begin{cases} y_t = \mu_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \beta_t + \eta_t \\ \beta_{t+1} = \beta_t + \xi_t \end{cases}$$
Modelo linear local
tendência linear estocástica

---

se $\sigma_\eta^2 = \sigma_\xi^2 = 0 \Rightarrow y_t = a + \underset{\downarrow}{\boxed{bt}} + \varepsilon_t$

porquê? a vai somando a cada

---

Calculando autovalores de T

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

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

$$\lambda_1 = \lambda_2 = 1.$$

$\Rightarrow$ Processo não estacionário

(2) $$y_t = \overbrace{[1 \quad 0]}^{Z_t} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

$$\underbrace{\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix}}_{\alpha_{t+1}} = \underbrace{\begin{bmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix}}_{\tilde{\eta}_t}$$

Analogamente:

$$\underset{\sim}{\eta}_t = \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \Rightarrow E\left[\underset{\sim}{\eta}_t \underset{\sim}{\eta}_t'\right] = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix} = Q$$

Equações:

$$\begin{cases} y_t = \mu_t + \varepsilon_t \\ \mu_{t+1} = \phi_1 \mu_t + \beta_t + \eta_t \\ \beta_{t+1} = \phi_2 \beta_t + \xi_t \end{cases}$$

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} \quad T = \begin{pmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{pmatrix} \quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \eta_t = \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix}$$

$$Q = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{pmatrix}$$

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

(a) Sob que condições $|\lambda_i(T)| < 1 \quad i = 1, 2$

Auto valores de T:

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

$$\left| \begin{pmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix} \phi_1 - \lambda & 1 \\ 0 & \phi_2 - \lambda \end{pmatrix} \right| = (\phi_1 - \lambda)(\phi_2 - \lambda) = 0$$

$$\phi_1\phi_2 - \phi_1\lambda - \phi_2\lambda + \lambda^2 = 0$$

$$\lambda^2 - (\phi_1 + \phi_2)\lambda + (\phi_1\phi_2) = 0$$

$$\lambda = \frac{(\phi_1 + \phi_2) \pm \sqrt{(\phi_1 + \phi_2)^2 - 4(\phi_1\phi_2)}}{2}$$

$$\lambda = (\phi_1 + \phi_2) \pm \sqrt{(\phi_1 - \phi_2)^2}$$

$$= \phi_1 + \phi_2 \pm (\phi_1 - \phi_2)$$

Queremos $|\lambda| < 1$

$$|\cancel{\phi_1} + \phi_2 - \cancel{\phi_1} + \phi_2| < 1 \quad = \quad 2|\phi_2| < 1 \Rightarrow \boxed{|\phi_2| < 1/2}$$

$$|\phi_1 + \phi_2 + \phi_1 - \phi_2| < 1 \quad \Rightarrow \quad \boxed{|\phi_1| < 1/2}$$

b) Avaliar $\lim_{t \to \infty} \Gamma_\alpha(0,t)$

$\lim_{t \to \infty} \Gamma_\alpha(h,t)$

$\lim_{t \to \infty} \Gamma_y(0,t)$

$\lim_{t \to \infty} \Gamma_y(h,t)$

$$\Gamma_\alpha(0,t) = \underbrace{T^t P_1 T^{t'}}_{\text{mostrar}} + \Sigma T^j (R.Q.R)'$$

- $\Gamma_\alpha(h,t) = ?$

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

- $E[\alpha_{t+1}] = E\begin{bmatrix} \mu_{t+1} \\ \beta_{t+1} \end{bmatrix}$

$$= E\begin{bmatrix} \phi_1 \mu_t + \beta_t + \eta_t \\ \phi_2 \beta_t + \xi_t \end{bmatrix}$$

$\beta_t$ : AR(1) $\Rightarrow$ $E[\beta_{t+1}] = 0$ $\Rightarrow$ $E[\alpha_t] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$E[\mu_{t+1}] = 0$

$$\Gamma_\alpha(0,t) = T^t P_1 T^{t'} + \Sigma T^j (RQR') T^{j'}$$

$$T^2 = \begin{pmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{pmatrix} \begin{pmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{pmatrix} = \begin{pmatrix} \phi_1^2 & \phi_1 + \phi_2 \\ 0 & \phi_2^2 \end{pmatrix}$$

$$T^3 = \begin{pmatrix} \phi_1^2 & \phi_1 + \phi_2 \\ 0 & \phi_2^2 \end{pmatrix} \begin{pmatrix} \phi_1 & 1 \\ 0 & \phi_2 \end{pmatrix} =$$

$$= \begin{pmatrix} \phi_1^3 & \phi_1^2 + \phi_1 \phi_2 + \phi_2^2 \\ 0 & \phi_2^3 \end{pmatrix}$$

$$T^4 = \begin{pmatrix} \phi_1^3 & \phi_1^2 + \phi_1 \phi_2 + \phi_2^2 \\ 0 & \phi_2^3 \end{pmatrix} \left( \begin{matrix} \phi_1 & 1 \\ 0 & \phi \end{matrix} \right.$$

aula 23/03

## Observabilidade

MEE linear, discreto, invariante no tempo.

$$\begin{cases} y_t = \overset{p\times m}{Z}\alpha_t \\ \alpha_{t+1} = \overset{m\times m}{T}\alpha_t + R\eta_t \end{cases}$$

sistema observável se $\alpha_0$ pode ser determinado a partir de conjunto de medidas de $y$.

- Fazendo $\eta_t = 0$ $\forall t$ em I:

$$\alpha_{t+1} = T\alpha_t \Rightarrow \alpha_{t+1} = T^t\alpha_1 \text{ (recursivamente)}$$

$$\boxed{\alpha_t = T^t\alpha_0}$$

$$y_t = Z\alpha_t \therefore \boxed{y_t = ZT^t\alpha_0}$$

$$\begin{aligned} \underset{p\times 1}{y_0} &= \overset{p\times m}{Z}\overset{m\times 1}{\alpha_0} \\ \underset{p\times 1}{y_1} &= ZT\alpha_0 \\ y_2 &= ZT^2\alpha_0 \\ &\vdots \\ y_{m-1} &= ZT^{m-1}\alpha_0 \end{aligned} \qquad \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{m-1} \end{pmatrix} = \underbrace{\begin{pmatrix} \overset{p\times m}{Z} \\ ZT \\ ZT^2 \\ \vdots \\ ZT^{m-1} \end{pmatrix}}_{pm\times m} \underbrace{\begin{pmatrix} \alpha_{0,1} \\ \alpha_{0,2} \\ \vdots \\ \alpha_{0,m} \end{pmatrix}}_{\tilde{\alpha}_0 \ m\times 1}$$

(pois cada $Z$: $p\times m$ e matriz tem $m$ linhas de matrizes $ZT^i$)

$$\underset{pm\times 1}{\tilde{y}} = \underset{pm\times m}{\tilde{M}}\ \underset{m\times 1}{\tilde{\alpha}_0}$$

- Para determinarmos os elementos de $\alpha_0$ de forma única, a cond. necessária e suficiente será:

posto $(M) = m$ $\Rightarrow$ é exatamente o tamanho de $\tilde{\alpha}_0$ e faz com que haja solução única p/ $\tilde{\alpha}_0$ no sistema acima.

Como posto $M$ = posto $M'$

posto $M'$ = posto $(Z', T'Z', \ldots, (T^{m-1})'Z') = m$

Se $p=1$, i.e. $y_t$ é escalar $\Rightarrow$ $M : m \times m$

cond de observabilidade pode ser investigada checando $\det M \neq 0$

Exs:

$Z : p \times m$ ; $ZT : p \times m$
$T : m \times m$

1) $y_t = \mu_t + \varepsilon_t$

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

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

$p=1$, $m=2$ ; $M \Rightarrow pm \times m : 2 \times 2$

$Z \Rightarrow 1 \times 2$ ; $ZT \Rightarrow 1 \times 2$

$T \Rightarrow 2 \times 2$ $\quad T'Z' \Rightarrow 2 \times 1$

Em forma EE:

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

$$\underline{\underline{m=2}} \quad \begin{pmatrix} \mu_{t+1} \\ \beta_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix}$$

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} \Rightarrow Z' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Rightarrow T' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$T'Z' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Logo: $M' = (Z', T'Z')$

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

$\det M' = 1 \neq 0 \Rightarrow$ posto $(M) = 2 = m$ cheio ↳ observável.

2). $y_t = \mu_t + \varepsilon_t$

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

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

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

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

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} \rightarrow \underset{1 \times 2}{p \times m} \quad \Rightarrow \quad Z' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \underset{2 \times 2}{m \times m} \quad \Rightarrow \quad T' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$T'Z' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad M' = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \det M' = 0$$

sistema não observável

3) $y_t = \mu_t + \beta_t + \varepsilon_t$

$\mu_{t+1} = \phi_1 \mu_t + \eta_t$

$\beta_{t+1} = \phi_2 \beta_t + \zeta_t$

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

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

$$Z = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

$$Z' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$T = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} \Rightarrow T' = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix}$$

$$\Rightarrow T'Z' = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$\Rightarrow M' = \begin{pmatrix} 1 & \phi_1 \\ 1 & \phi_2 \end{pmatrix} \Rightarrow \det M' = \phi_2 - \phi_1 \neq 0$$

Para sistema observável: $\phi_1 \neq \phi_2$.
p/ termos posto $(M') = 2$.

4) Seja $y_t = \mu_t + \beta_t + \varepsilon_t$ → p/ o caso (3) com $\phi_1 = \phi_2 = 1$

$y_t = \mu_t + \beta_t + \varepsilon_t$
$\mu_{t+1} = \mu_t + \eta_t$
$\beta_{t+1} = \beta_t + \xi_t$
sistema não observável

Seja $\theta_t = \mu_t + \beta_t$

$\psi_t = \mu_t - \beta_t$

← outras formas de ver

- $y_t = \theta_t + \varepsilon_t$

- $$\begin{aligned} \theta_{t+1} &= \mu_{t+1} + \beta_{t+1} \\ &= (\mu_t + \eta_t) + (\beta_t + \xi_t) \\ &= (\mu_t + \beta_t) + \underbrace{(\eta_t + \xi_t)}_{\delta_{1,t}} \\ &= \theta_t + \delta_{1,t} \end{aligned}$$

- $$\begin{aligned} \psi_{t+1} &= \mu_{t+1} - \beta_{t+1} \\ &= (\mu_t - \beta_t) + \underbrace{(\eta_t - \xi_t)}_{\delta_{2t}} \\ &= \psi_t - \delta_{2,t} \end{aligned}$$

→ $y_t = [1 \quad 0] \begin{bmatrix} \theta_t \\ \psi_t \end{bmatrix}$ → sobra na especificação

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

↓ Nesta parametrização fica + fácil de ver que sistema é não observável.

<u>Obs</u>: → observabilidade: será obrigatória p/ estimação dos parâmetros

→ estacionariedade: não " , vai depender da definição do modelo.

será importante definir nas cond. iniciais quais componentes são ou não estacionárias.

cont. aula 23/03

## Estimação de MEE lineares e Gaussianos

$$y_t = Z_t \alpha_t + \varepsilon_t$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad (t = 1, \dots, n)$$

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

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

Obs:
- equações do FK não exigem que sist seja invariante no tempo
- $n$: tamanho da série (ou $T$)
- FK não exige que choques tenham dist. Gaussiana, mas que sejam definidas sobre os 2 primeiros momentos.

c/ Gaussianidade: erro médio quadrático será ótimo

s/ ": ótimo local (melhor estimador linear

→ O que queremos?

(1) Estimar $\alpha_t$ dado certo conjunto de informação (observável)

↓

Equivale a obter $p(\alpha_t \mid \underset{\sim}{Y}_j)$ : densidade condicional de $\alpha$ dado $Y$.

$\therefore E[\alpha_t \mid Y_j] = \dots$

$\text{var}[\alpha_t \mid Y_j] = \dots$

Conjuntos possíveis de $Y_j$:

$Y_{t-1} = (y_1\ y_2 \dots y_{t-1}) \Rightarrow j = t-1$ : previsão ⎫ usados no FK

$Y_t = (y_1\ y_2 \dots y_t) \Rightarrow j = t$ : atualização ⎭

$Y_n = (y_1\ y_2 \dots y_{t-1}\ y_t \dots y_n) \Rightarrow j = n$ : suavização

(2) Estimar hiperparâmetros
(por MV nas matrizes do sistema)

⇒ Estimar { vetor de estado usando previsão e atualização

Exemplo:

considere o modelo linear local

$$\begin{cases} y_t = \mu_t + \varepsilon_t & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_{t+1} = \mu_t + \eta_t & \eta_t \sim N(0, \sigma_\eta^2) \end{cases}$$

$\mu_1 \sim (a_1, p_1)$

$E[\varepsilon_t \mu_1] = E[\eta_t \mu_1] = 0 \ \forall t$

$E[\varepsilon_t, \eta_s] = 0 \ \forall t, s.$

Obter $p(\alpha_t | Y_t)$ sem fazer uso do FK, apenas recorrendo a propriedades de vetores normais multivariados

→ Para $y$

$y_1 = \mu_1 + \varepsilon_1$

$y_2 = \mu_2 + \varepsilon_2 = \mu_1 + \eta_1 + \varepsilon_1$

$\vdots$

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

Seja $y = (y_1 \ y_2 \ \dots \ y_n)'$

. $E[y] = \begin{pmatrix} a_1 \\ \vdots \\ a_1 \end{pmatrix} = \mathbb{1} a_1$

. $\text{var}[y] = E[(y - E[y])(y - E[y])'] = \Omega$

matriz de var/covar

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

variâncias:

- $\text{var}(y_1) = E[(y_1 - a_1)^2] =$

$= E[(\mu_1 + \varepsilon_1 - a_1)^2] = \text{var}(\mu_1) + E(\varepsilon_1^2)$

$= p_1 \quad \sigma_{\varepsilon_1}^2$

- $\text{var}(y_2) = E[(\mu_1 + \eta_1 + \varepsilon_1 - a_1)^2] = \text{var}(\mu_1) + E(\varepsilon_1^2) + E[\eta_1^2]$

- $var(y_n) = p_1 \quad (n-1)\sigma_\eta^2 + \sigma_\varepsilon^2$

covariâncias

- $cov(y_1, y_2) = E[(y_1 - a_1)(y_2 - a_1)]$

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

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

$$= E[(\mu_1 - a_1)^2) = p_1$$

Logo:

$$\Sigma_{yy} = \begin{bmatrix} p_1 + \sigma_\varepsilon^2 & p_1 & \cdots & p_1 \\ p_1 & p_1 + \sigma_\eta^2 + \sigma_\varepsilon^2 & & \vdots \\ \vdots & p_1 + \sigma_\eta^2 & \ddots & \vdots \\ \vdots & p_1 + 2\sigma_\eta^2 & & \\ p_1 & \cdots & & p_1 + (n-1)\sigma_\eta^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

$$= \mathbb{1}\mathbb{1}' p_1 + \Sigma_{ij} \quad \text{onde} \quad \underset{\text{(diagonal)}}{\Sigma_{ij}} = \begin{cases} (i-1)\sigma_\eta^2 & i < j \\ \sigma_\varepsilon^2 + (i-1)\sigma_\eta^2 & i = j \\ (j-1)\sigma_\eta^2 & i > j \end{cases}$$

$cov(y_2, y_3) = E[(y_1 - a_1)(y_3 - a_1)]$

$$= E[(\mu_1 + \eta_1 + \varepsilon_2 - a_1)(\mu_1 + \eta_1 + \eta_2 + \varepsilon_3 - a_1)]$$

$$= E[(\mu_1 - a_1)^2] + E[\eta_1^2)$$

$$= p_1 + \sigma_\eta^2$$

Obs. $n$ cresce com o número de observações. Do ponto de vista operacional é um problema.

→ Para $\mu$

$\mu_2 = \mu_1 + \eta_1$

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

$\vdots$

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

seja $\mu = (\mu_1 \; \mu_2 \cdots \mu_n)'$

- $E[y] = \begin{pmatrix} a_1 \\ \vdots \\ a_1 \end{pmatrix} = \mathbb{1} a_1$

. $\mathrm{var}(\mu) = E[(\mu - E[\mu])(\mu - E[\mu])') = \Sigma_\mu$

$$\Sigma_{\mu\mu} = \begin{bmatrix} \mathrm{var}(\mu_1) & \mathrm{cov}(\mu_1, \mu_2) & \cdots & \mathrm{cov}(\mu_1, \mu_n) \\ \mathrm{cov}(\mu_2 \mu_1) & \mathrm{var}\,\mu_2 & & \vdots \\ \vdots & & & \\ \mathrm{cov}(\mu_n \mu_1) & & & \mathrm{var}(\mu_n) \end{bmatrix}$$

→ variâncias:

- $\mathrm{var}(\mu_1) = E[(\mu_1 - a_1)^2) = p_1$
- $\mathrm{var}\ (\mu_2 = E[(\mu_2 - a_1)^2) =$

$$= E[(\mu_1 + \eta_1 - a_1)^2]$$

$$= E[(\mu_1 - a_1)^2] + E[\eta_1^2) = p_1 + \sigma_\eta^2$$

obs: $\mu_t = \mu_1 + \sum_{i=1}^{t-1} \eta_i \longrightarrow (\mu_1 - a_1)^2 + 2\underbrace{(\mu_1 - a_1)\sum_{i=1}^{t-1}\eta_i}^{=0}$

$+ (\underbrace{\sum_{i=1}^{t-1}\eta_i}_{\mathrm{cov}(\eta_i, \eta_j) = 0 \ p/\, i \neq j})^2$

$\therefore \mathrm{var}(\mu_t) = E[(\mu_1 + \sum_{i=1}^{t-1}\eta_i - a_1)^2] =$

$$= E[(\mu_1 - a_1)^2] + \sum_{i=1}^{t-1} E[\eta_i^2] =$$

$$= p_1 + \sum_{i=1}^{t-1} \sigma_\eta^2$$

ou seja:

$$\mathrm{var}(\mu_t) = p_1 + (i-1)\sigma_\eta^2 \qquad i = 1, 2, \dots n$$

→ covariâncias ⇒ dependerá da posição relativa entre i e j

- $\mathrm{cov}(\mu_1, \mu_2) = E[(\mu_1 - a_1)(\mu_2 - a_1))$

$$= E[(\mu_1 - a_1)(\mu_1 + \eta_1 - a_1)) =$$

$$= E[(\mu_1 - a_1)^2] + E[\underbrace{(\mu_1 - a_1)(+\eta_1)}_{0}) = p_1$$

- $\mathrm{cov}(\mu_1, \mu_3) = E[(\mu_1 - a_1)(\mu_3 - a_1))$

$$= E[(\mu_1 - a_1)(\mu_1 + \eta_1 + \eta_2 - a_1)] =$$

$$= E[(\mu_1 - a_1)^2) + E[(\mu_1 - a_1)\eta_1) + E[(\mu_1 - a_1)\eta_2)$$

$$= p_1$$

- $\mathrm{cov}(\mu\ \ \mu) = p_1$

- $\mathrm{cov}(\mu_2, \mu_3) = E[(\mu_2 - a_1)(\mu_3 - a_1)]$

$$= E[(\mu_1 + \eta_1 - a_1)(\mu_1 + \eta_1 + \eta_2 - a_1)]$$

$$= E[((\mu_1 - a_1)^2 + \eta_1)((\mu_1 - a_1)^2 + \eta_1 + \eta_2)] =$$

$$= E[(\mu_1 - a_1)^2] + \underbrace{E[\eta_1(\mu_1 - a_1)^2]}_{0} + \underbrace{E[\eta_2(\mu_1 - a_1)^2]}_{0} +$$

$$+ \underbrace{E[\eta_1(\mu_1 - a_1)^2]}_{0} + \underbrace{E[\eta_1 \eta_1]}_{\sigma_\eta^2} + \underbrace{E[\eta_1 \eta_2]}_{0}$$

$$= p_1 + \sigma_\eta^2$$

- $\mathrm{cov}(\mu_2, \mu_t) = p_1 + \sigma_\eta^2$ para $t > 2$

- Generalizando: $j > i$

$$\mu_i = \mu_1 + \sum_{r=1}^{i-1} \eta_r$$

$$\mu_j = \mu_1 + \sum_{s=1}^{j-1} \eta_s$$

$$\mathrm{cov}(\mu_i, \mu_j) = E\left[\left(\mu_1 + \sum_{r=1}^{i-1} \eta_r - a_1\right)\left(\mu_1 + \sum_{s=1}^{j-1} \eta_s - a_1\right)\right] =$$

$$= E[(\mu_1 - a_1)^2] + E\left[\sum_{r=1}^{i-1} \eta_r \cdot \sum_{s=1}^{j-1} \eta_s\right] =$$

$$= p_1 + E[\dots]$$

$$E\left[\sum_{r=1}^{i-1} \eta_r + \sum_{s=1}^{j-1} \eta_s\right] =$$

$$= E[(\eta_1 + \eta_2 + \dots + \eta_{i-1})(\eta_1 + \eta_2 + \dots \eta_{j-1})]$$

( se $i < j \Rightarrow$ temos $(i-1)$ termos $E[\eta^2]$

$$= (i-1)\sigma_\eta^2$$

se $i > j$

$$= (j-1)\sigma_\eta^2$$

Logo: $\mathrm{cov}(\mu_i, \mu_j) = \begin{cases} p_1 + (i-1)\sigma_\eta^2 & \text{se } i < j \\ p_1 + (j-1)\sigma_\eta^2 & \text{se } i > j \end{cases}$

( ~~e $p_1$ se $i = j$~~ )

Logo:

$$\Sigma_{\mu\mu} = \Sigma_{ij} \quad \text{onde } \Sigma_{ij} = \begin{cases} p_1 + (i-1)\sigma^2_\eta & \text{se } i<j \\ p_1 + (i-1)\sigma^2_\eta & \text{se } i=j \\ p_1 + (j-1)\sigma^2_\eta & \text{se } i>j \end{cases}$$

→ Para $y$ e $\mu$ conjuntamente:

$$\underbrace{cov(y, \mu)}_{\Sigma_{y\mu}} = E\left[ (y - \mathbb{1}a_1)(\mu - \mathbb{1}a_1)' \right]$$

cada termo individualmente:

$$cov(y_i, \mu_j) = E[(y_i - a_1)(\mu_j - a_1)]$$

$$y_i = \mu_1 + \sum_{r=1}^{i-1} \eta_r + \varepsilon_i$$

$$\mu_j = \mu_1 + \sum_{s=1}^{j-1} \eta_s$$

$$E\left[(y_i - a_1)(\mu_j - a_1)\right] = E\left[\left((\mu_1 - a_1) + \sum_{r=1}^{i-1}\eta_r + \varepsilon_i\right)\left((\mu_1 - a_1) + \sum_{s=1}^{j-1}\eta_s\right)\right)$$

$$= E\left[(\mu_1 - a_1)^2\right] + \underbrace{E\left[\sum_{r=1}^{i-1}\eta_r \cdot \sum_{s=1}^{j-1}\eta_s\right]}_{\text{já vimos que}}$$

$$= p_1 + \begin{cases} (i-1)\sigma^2_\eta & \text{se } i \boxed{\leq} j \\ (j-1)\sigma^2_\eta & \text{se } j > 1 \end{cases}$$

$$\text{Logo: } \Sigma_{y\alpha} = \Sigma_{ij} \quad \text{onde } \Sigma_{ij} = \begin{cases} p_1 + (i-1)\sigma^2_\eta & \text{se } i \leq j \\ p_1 + (j-1)\sigma^2_\eta & \text{se } i > j \end{cases}$$

**Assim**

Podemos caracterizar probabilisticamente os vetores aleatórios $y$ e $\mu$ através do seguinte resultado:

$$\begin{pmatrix} \underset{\sim}{\mu} \\ \underset{\sim}{y} \end{pmatrix} \sim N \left[ \begin{pmatrix} \mathbf{1} a_1 \\ \mathbf{1} a_1 \end{pmatrix}, \begin{pmatrix} \Sigma_{\mu\mu} & \Sigma_{\mu y} \\ \Sigma_{y\mu} & \Sigma_{yy} \end{pmatrix} \right]$$

onde $\mathbf{1} a_1 = (a_1 \ldots a_1)'$ $n \times 1$

$$\Sigma_{\mu\mu} \to \Sigma_{\mu\mu}^{ij} = \begin{cases} (i-1)\sigma_\eta^2 + p_1 & \text{se } i \leq j \\ (j-1)\sigma_\eta^2 + p_1 & \text{se } j > i \end{cases}$$

$$\Sigma_{yy} \to \Sigma_{yy}^{ij} \begin{cases} (i-1)\sigma_\eta^2 + p_1 & i < j \\ (i-1)\sigma_\eta^2 + p_1 + \sigma_\varepsilon^2 & i = j \\ (j-1) + p_1 & i > j \end{cases}$$

$$\Sigma_{\mu y} \to \Sigma_{\mu y}^{ij} = \begin{cases} (i-1)\sigma_\eta^2 + p_1 & \text{se } i \leq j \\ (j-1)\sigma_\eta^2 + p_1 & \text{se } j > i \end{cases} = \Sigma_{\mu\mu}$$

Queremos estimar $\underset{\sim}{\alpha}$ dada a informação da série observada $y$. Assim, devemos obter a densidade condicional:

$$p(\underset{\sim}{\alpha} \mid \underset{\sim}{y})$$

A partir dela, calculamos $E[\underset{\sim}{\alpha} \mid y]$, $\text{var}[\underset{\sim}{\alpha} \mid y]$ etc. e assim caracterizar probabilisticamente $\underset{\sim}{\alpha}$ condicional a $y$.

Por exemplo:

É natural que adotemos

$$\hat{\alpha} = E[\underset{\sim}{\alpha} \mid y]$$

Para avançarmos, faremos uso do seguinte resultado p/ vetores de variáveis aleatórias com dist Normal multivariada

<u>Teorema</u>: $\begin{pmatrix} \underset{\sim}{X} \\ \underset{\sim}{Y} \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}\right]$, $\Sigma_{xy} = \Sigma_{yx}'$

então: $f(X \mid Y = y) \sim N\left(E(X \mid Y = y), \text{var}[X \mid Y = y]\right)$

$$E[X \mid Y = y] = \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y)$$

$$\text{Var}[X \mid Y = y] = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}$$

Fazendo: $\underset{\sim}{X} = \underset{\sim}{\mu}$ e $Y = y$, segue que:

$$f(\underset{\sim}{\mu} | Y = y) \sim N\left(E[\underset{\sim}{\mu} | Y = y],\ var[\underset{\sim}{\mu} | Y = y]\right) \text{ onde}$$

$$E[\underset{\sim}{\mu} | Y = y] = E[\underset{\sim}{\mu}] + \Sigma_{\mu y} \Sigma_{yy}^{-1} (y - a_1 \mathbb{1})$$

$$= \mathbb{1} a_1 + \Sigma_{\mu y} \Sigma_{yy}^{-1} (y - a_1 \mathbb{1})$$

$$= \mathbb{1} a_1 + \Sigma_0$$

Observações

Na prática, esta forma de se obter estimativa p/ o vetor de estados $\underset{\sim}{\mu}$ não é recomendável pois

i. As matrizes $\Sigma_{yy}$ e $\Sigma_{\mu\mu}$ possuem dimensão $n \times n$ onde $n$ é o tamanho da amostra.

O esforço computacional de estimação seria proporcional a $n$.

$$E[\underset{\sim}{\alpha} | Y = y] = \begin{bmatrix} E(\alpha_1 | \underset{\sim}{Y_n}) \\ \vdots \\ E[\alpha_n | \underset{\sim}{Y_n}] \end{bmatrix}$$

$\Rightarrow$ a cada instante $t = 1, \dots, n$ as estimativas são calculadas utilizando toda a informação da amostra pois

$$Y_n = (y_1, y_2 \dots y_n)$$

ii. Se a série de observações vai crescendo no tempo, ou seja, a partir de $n$ novas obs. vão sendo adicionadas, as matrizes $\Sigma_{yy}$ e $\Sigma_{\mu\mu}$ crescerão

copiar

23/03.

## Passo da Previsaõ no FK

(a) $y_t = Z_t \alpha_t + \varepsilon_t$ , $\varepsilon_t \sim N(0, H_t)$

(b) $\alpha_{t+1} = T_t \alpha_t + R_t \eta_t$ , $\eta_t \sim N(0, Q_t)$

$$E[\varepsilon_t' \eta_s] = 0 \quad \forall\, s, t$$

$$\alpha_1 \sim N(a_1, P_1)$$

$$E[\alpha_1 \varepsilon_t] = E[\alpha_1 \eta_t] = 0 \quad \forall\, t$$

Sejam: valor esperado do vetor de estado no período t+1 dada observaçaõ até t

$a_{t+1} = E[\alpha_{t+1} \mid \underset{\sim}{Y}_t]$ onde $Y_t = (y_1\ y_2 \cdots y_{t-1}\ y_t)$

$P_{t+1} = var[\alpha_{t+1} \mid Y_t]$

De (b), segue que:

- $E[\alpha_{t+1} \mid \underset{\sim}{Y}_t] = E[T_t \alpha_t + R_t \eta_t \mid \underset{\sim}{Y}_t]$

$$= T_t E[\alpha_t \mid \underset{\sim}{Y}_t] = T_t a_{t|t}$$

↳ Notaçaõ do Koopman equivale a $T_t E[\alpha_t | Y_t$

$\Rightarrow \boxed{a_{t+1} = T_t a_t}$

valor esperado

obs: Se no modelo $\alpha_{t+1} = T_t \alpha_t + c_t + R_t \eta_t \Rightarrow a_{t+1} = T_t a_t + \boxed{c_t}$

- $\underbrace{var[\alpha_{t+1} \mid \underset{\sim}{Y}_t]}_{P_{t+1}} = E\left[ (\alpha_{t+1} - \underbrace{E[\alpha_{t+1} \mid \underset{\sim}{Y}_t]}_{a_{t+1}})(\alpha_{t+1} - \underbrace{E[\alpha_{t+1} \mid \underset{\sim}{Y}_t]}_{a_{t+1}})' \,\middle|\, \underset{\sim}{Y}_t \right.$

imp

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

$a_{t+1} = T_t a_{t|t}$ $\Rightarrow \alpha_{t+1} - a_{t+1} = T_t(\alpha_{t+1} - a_{t|t}) + R_t \eta_t$

$$\therefore P_{t+1} = E\left[ \left( T_t(\alpha_t - a_{t|t}) + R_t \eta_t \right) \left( T_t(\alpha_t - a_{t|t}) + R_t \eta_t \right)' \,\middle|\, \underset{\sim}{Y}_t \right.$$

$$= E\left[ \left( T_t(\alpha_t - a_{t|t}) + R_t \eta_t \right) \left( (\alpha_t - a_{t|t})' T_t' + \eta_t' R_t' \right) \middle| \underset{\sim}{Y}_t \right.$$

$$\therefore \boxed{\begin{aligned} P_{t+1} = {} & E_t\left[T_t(\alpha_t - a_{t|t})(\alpha_t - a_{t|t})' T_t'\right] + E_t\left[T_t(\alpha_t - a_t)\eta_t' R_t'\right] + \\ & + E_t\left[R_t \eta_t (\alpha_t - a_{t|t})' T_t'\right] + E_t\left[R_t \eta_t \eta_t' R_t'\right] \end{aligned}}$$

onde $E_t[\cdot] = E\left[(\ )\mid \underset{\sim}{Y}_t\right]$

Escrevendo a equação (b) por subst. recursiva, temos:

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

$$\alpha_2 = T_1 \alpha_1 + R_1 \eta_1$$

$$\alpha_3 = T_2 \alpha_2 + R_2 \eta_2 = T_2(T_1\alpha_1 + R_1\eta_1) + R_2\eta_2$$

$$\alpha_4 = T_3 \alpha_3 + R_3 \eta_3 = T_3\left(T_2(T_1\alpha_1 + R_1\eta_1)\right) + R_3\eta_3$$

$$= \underbrace{T_3 T_2}_{} T_1 \alpha_1 + T_3 T_2 R_1 \eta_1 + R_3 \eta_3$$

entender

$$\prod_{i=1}^{3} T_1 \alpha_1$$

$$\Rightarrow \quad \alpha_{t+1} = \sum_{j=1}^{t-1}\left[\prod_{i=0}^{j-1} T_{t-i}\right] R_{t-j}\,\eta_{t-j} + R_t \eta_t \qquad (b_1)$$

Seja:

$$\cdot\ A_t = E_t\left[T_t(\alpha_t - a_t)(\alpha_t - a_t)' T_t'\right] =$$

$$= T_t \underbrace{E_t\left[(\alpha_t - a_t)(\alpha_t - a_t)'\right]}_{} T_t' =$$

$$= T_t P_{t|t} T'$$

$$\cdot\ B_t = E_t\left[T_t(\alpha_t - a_t)\eta_t' R_t'\right] =$$

$$= T_t E_t\left[(\alpha_t - a_t)\eta_t'\right] R_t' =$$

$$= T_t E_t\left[\alpha_t \eta_t'\right] R_t' - T_t a_t \underset{0}{E[\cancel{\eta_t'}]} R_t'$$

$$= T_t E_t\left[\alpha_t \eta_t'\right] R_t'$$

mas: $E_t[\alpha_t \eta_t'] \Leftrightarrow E_{t+1}[\alpha_{t+1}\eta_{t+1}']$

Substituindo $\alpha_{t+1}$ encontrado em (51), temos:

$$E_{t+1}[\alpha_{t+1}\eta_{t+1}'] = E_{t+1}\left[\left(\sum_{j=1}^{t-1}\left[\prod_{i=0}^{j-1} T_{t-i}\right]R_{t-j}\eta_{t-j} + R_t\eta_t\right)\eta_{t+1}'\right] =$$

$$= \left(\sum_{j=1}^{t-1}\left[\prod_{i=0}^{j-1}T_{t-i}\right]R_{t-j}\,E_{t+1}[\eta_{t-j}\eta_{t+1}']\right) + R_t\,E_{t+1}[\eta_t\eta_{t+1}']$$

mas:

- $E_{t+1}[\eta_{t-j}\eta_{t+1}'] = E[\eta_{t-j}\eta_{t+1}']$ $\quad j=1,2,\ldots,t-1$

  $=0$ pois, por hipótese, $E[\eta_t\eta_s']=0 \ \forall t\neq s$.

- $E_{t+1}[\eta_t\eta_{t+1}'] = E[\eta_t\eta_{t+1}'] = 0$ de forma análoga.

Logo: $B_t \equiv 0 \quad \forall t$.

- $C_t = E_t[R_t\eta_t(\alpha_t - a_{t|t})T_t']$

$$= \left(R_t \underbrace{E_t[\eta_t\alpha_t']}_{0} - R_t \underbrace{E[\eta_t]}_{0}a_{t|t}'\right)T_t' = 0.$$

- $D_t = E_t[R_t\eta_t\eta_t'R_t'] = R_t E_t[\eta_t\eta_t']R_t'$

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

$$= R_t Q_t R_t'$$

Portanto, segue que:

$$\boxed{P_{t+1} = \mathrm{var}[\alpha_{t+1}|\underset{\sim}{y}_t] = T_t P_{t|t} T_t' + R_t Q_t R_t'}$$

Aula 30/03

(I) $y_t = Z_t \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, H_t)$

(II) $\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim N(0, Q_t)$

Da aula anterior:

→ Passo da previsão:

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

$$P_{t+1} = var(\alpha_{t+1} | Y_t)$$
$$= T_t P_{t|t} T_t' + R_t Q_t R_t'$$

Queremos agora:

$$\begin{cases} a_{t|t} = E[\alpha_t | \underset{\sim}{Y_t}] \\ P_{t|t} = var[\alpha_t | \underset{\sim}{Y_t}] \Rightarrow \text{matriz de var|covar} \end{cases}$$

Ainda da aula passada:

(1) $$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz} \end{pmatrix}\right]$$

Supondo $\mu_z = 0$ e $\Sigma_{zy} = 0$ $(\Sigma_{yz} = 0)$, temos:

$$E[X | Y = y, Z = z] = \mu_{x|y} + \Sigma_{xz} \Sigma_{zz}^{-1} z$$

$$var[X | Y = y, Z = z] = \Sigma_{x|y} - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx}$$

(2) Se X e Y são dois vetores aleatórios e $g: \mathbb{R}^p \to \mathbb{R}^p$ (bijeti

então $f_{X|Y} = f_{X|g(Y)} \Rightarrow E[X | Y = y] = E(X | g(Y) = g(y))$

Calculemos $a_{t|t}$ e $P_{t|t}$

$a_{t|t}$

- $a_{t|t} = E[\alpha_t | \underset{\sim}{Y_t}]$ → decompondo

$= E[\alpha_t | \underset{\sim}{Y_{t-1}}, y_t] =$

$= E[\alpha_t | \underset{\sim}{Y_{t-1}}, v_t]$ onde $v_t = y_t - E[y_t | \underset{\sim}{Y_{t-1}}]$

erro de previsão 1 passo a frente (inovação)

Da q. (I) do modelo:

ostrar que pode os usar $_t$ ao invé de $y_t$ usando s. do ultado 2).

$$E[y_t \mid \underset{\sim}{Y}_{t-1}] = E[Z_t\alpha_t + \varepsilon_t \mid \underset{\sim}{Y}_{t-1}]$$
$$= Z_t\, E[\alpha_t \mid \underset{\sim}{Y}_{t-1}]$$
$$= Z_t a_t$$

Logo: $v_t = y_t - Z_t a_t$ ⇒ conteúdo informacional de $y_t$ não é modificado se usarmos $v_t$ podemos fazer isso pois $v_t$ é trns bijetiva

$$\Rightarrow \boxed{\begin{aligned} a_{t|t} &= E[\alpha_t \mid \underset{\sim}{Y}_t] = \\ &= E[\alpha_t \mid \underset{\sim}{Y}_{t-1}, v_t] \end{aligned}}$$

Usando resultado (1) e fazendo $x = \alpha_t$, $Y = \underset{\sim}{Y}_{t-1}$, $Z = v_t$

- Primeiro precisamos provar que $\mu_Z = \mu_{v_t} = 0$ e $\Sigma_{ZY} = \Sigma_{v_t \underset{\sim}{Y}_{t-1}} = 0$ (p/ usar resultado)

Mostrar que podemos usar resultado (1)

$* \quad \mu_Z = E[Z] = 0$

$\therefore E[v_t] = 0$

mas $E[v_t] = E[E[v_t \mid \underset{\sim}{Y}_{t-1}]]$

e $E[v_t \mid \underset{\sim}{Y}_{t-1}] = E[y_t - Z_t a_t \mid \underset{\sim}{Y}_{t-1}]$

$$= E[y_t \mid \underset{\sim}{Y}_{t-1}] - Z_t a_t$$
$$\overset{\text{do modelo}}{=} Z_t a_t - Z_t a_t = 0.$$

$\therefore E[v_t] = E[0] = 0$. ok!

$* \quad \Sigma_{YZ} = 0$

$\therefore \Sigma_{Y_{t-1} v_t} = cov(Y_{t-1}, v_t) = 0$

mas $cov(Y_{t-1}, v_t) = E[(Y_{t-1} - E[Y_{t-1}])(v_t - E[v_t])']$

$$= E[Y_{t-1} \cdot v_t'] - E[Y_{t-1}] \cdot \underset{0}{\overset{0}{E[v_t']}}$$
$$= E[Y_{t-1} \cdot v_t']$$

$$\underset{\sim}{Y}_{t-1} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_1 \end{bmatrix}$$

⇓

matriz cujos elementos são

$\{E[y_{t-j}\, v_t']\}_{j=1,2}$

$$E[y_{t-j}\nu_t'] = E[E[y_{t-j}\nu_t' \mid \underset{\sim}{Y}_{t-1}]]$$

é passado, pode sair de $E[\cdot]$

$$E[y_{t-j}\nu_t' \mid \underset{\sim}{Y}_{t-1}] = y_{t-j}\,E[\underset{0}{\nu_t'} / \underset{\sim}{Y}_{t-1}]$$

$$= 0.$$

$$\Rightarrow E[y_{t-1}\nu_t'] = 0 \qquad \text{ok!}$$

Portanto, podemos usar o resultado e escrever:

$$E[\alpha_t \mid \underset{\sim}{y}_t] = \underbrace{E[\alpha_t \mid \underset{\sim}{Y}_{t-1}]}_{a_t} + \text{cov}[\alpha_t, \nu_t]\,(\text{var}[\nu_t])^{-1}\nu_t$$

$a_t$ (já sabemos quem é do passo da previsão em t-1)

Precisamos calcular estas componentes da equação.

hegar
a expressão
no lado
de $a_{t+1}$

- Calculando var$[\nu_t]$

cular
mponente

$$\text{var}[\nu_t] = E[(\nu_t - \underset{0}{E[\nu_t]})(\nu_t - \underset{0}{E[\nu_t]})']$$

$$= E[\nu_t\nu_t']$$

$$= E[\underbrace{E[\nu_t\nu_t' \mid \underset{\sim}{Y}_{t-1}]}]$$

Sabemos que $\nu_t = y_t - Z_t a_t$

$$= Z_t\alpha_t + \varepsilon_t - Z_t a_t$$

$$= Z_t(\alpha_t - a_t) + \varepsilon_t$$

$$\Rightarrow E[\nu_t\nu_t' \mid \underset{\sim}{Y}_{t-1}] =$$

$$= E[(Z_t(\alpha_t - a_t) + \varepsilon_t)(Z_t(\alpha_t - a_t) + \varepsilon_t)' \mid \underset{\sim}{Y}_{t-1}]$$

$$= E[Z_t(\alpha_t - a_t)(\alpha_t - a_t)'Z_t' \mid \underset{\sim}{Y}_{t-1}] +$$

$$+ \underbrace{E[Z_t(\alpha_t - a_t)\varepsilon_t' \mid \underset{\sim}{Y}_{t-1}]}_{=0} + E[\varepsilon_t(\alpha_t - a_t)'Z_t' \mid \underset{\sim}{Y}_{t-1}]$$

$$+ E[\varepsilon_t\varepsilon_t' \mid \underset{\sim}{Y}_{t-1}]$$

$$Z_t E[\alpha_t \varepsilon_t' \mid \underset{\sim}{Y}_{t-1}] - Z_t a_t E[\underset{0}{\varepsilon_t'} \mid \underset{\sim}{Y}_{t-1}]$$

funções de $\alpha_1$ e $\eta_{t-j}$ $\to 0$

$$\therefore\ E[\nu_t \nu_t' \mid \underset{\sim}{Y}_{t-1}] = Z_t \underbrace{E[(\alpha_t - a_t)(\alpha_t - a_t)' \mid \underset{\sim}{Y}_{t-1}]}_{P_t} Z_t' + \underbrace{E[\varepsilon_t \varepsilon_t' \mid \underset{\sim}{Y}_{t-1}]}_{H_t} .$$

$$= Z_t P_t Z_t' + H_t$$

Logo:

$$\boxed{var(\nu_t') = E[\nu_t \nu_t'] = Z_t P_t Z_t' + H_t \hat{=} F_t}$$

- calculando $cov(\alpha_t, \nu_t)$

da definição:

$$cov(\alpha_t, \nu_t) = E[(\alpha_t - E(\alpha_t))(\nu_t - \underset{0}{E(\nu_t)})]$$

$$= E[\alpha_t \nu_t'] - E(\alpha_t)\underset{0}{E(\nu_t)}$$

$$= E[\alpha_t \nu_t']$$

$$= E[E(\alpha_t \nu_t' \mid \underset{\sim}{Y}_{t-1})]$$

$$\downarrow$$

$$\alpha_t \nu_t' = \alpha_t(y_t' - a_t' Z_t')$$

$$= \alpha_t[(\alpha_t' - a_t')Z_t' + \varepsilon_t']$$

$$\Rightarrow E[\alpha_t \nu_t' \mid \underset{\sim}{Y}_{t-1}] = E[(\alpha_t (\alpha_t - a_t)' Z_t' + \varepsilon_t') \mid \underset{\sim}{Y}_{t-1}] =$$

$$= E[\alpha_t (\alpha_t - a_t)' Z_t' \mid \underset{\sim}{Y}_{t-1}] + E[\underset{0}{\varepsilon_t' \mid \underset{\sim}{Y}_{t-1}}]$$

truque: subtrai $a_t$ ↗

$$= E[(\alpha_t - a_t)(\alpha_t - a_t)' Z_t' \mid \underset{\sim}{Y}_{t-1}]$$

$$= \underbrace{E[(\alpha_t - a_t)(\alpha_t - a_t)' \mid \underset{\sim}{Y}_{t-1}]}_{P_t} \cdot Z_t'$$

$$= P_t Z_t'$$

$$\therefore\ \boxed{cov(\alpha_t, \nu_t) = E(P_t Z_t') = P_t Z_t' = M_t}$$

Usando os resultados p/ $var(Y_t')$ e $cov(\alpha_t, \nu_t)$, podemos escrever:

$$E[\alpha_t \mid \underset{\sim}{Y}_t] = E(\alpha_t \mid \underset{\sim}{Y}_{t-1}) + \operatorname{cov}(\alpha_t, v_t)(\operatorname{var}(v_t))^{-1} v_t$$

Resultado p/ $a_{t|t}$

$$\boxed{a_{t|t} = E(\alpha_t \mid \underset{\sim}{Y}_t) = a_t + M_t F_t^{-1} v_t}$$

onde $\begin{cases} M_t = P_t Z_t' \\ F_t = Z_t P_t Z_t' + H_t \\ v_t = y_t - Z_t a_t \end{cases}$

- $P_{t|t} = \operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_t]$

$t|t)$

$$= \operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_{t-1}, y_t] =$$

$$= \operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_{t-1}, v_t]$$

↳ podemos usar $v_t$ ao invés de $y_t$

e podemos usar resultado (1) para var

s que
demos
sar $v_t$ e
sultado
p/ var

Novamente fazendo $x = \alpha_t$, $y = Y_{t+1}$ e $z = v_t$

$$\operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_t] = \operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_{t-1}, v_t]$$

$$= \underbrace{\operatorname{var}[\alpha_t \mid \underset{\sim}{Y}_{t-1}]}_{P_t} - \underbrace{\operatorname{cov}[\alpha_t, v_t]}_{M_t} \underbrace{(\operatorname{var}(v_t))^{-1}}_{F_t^{-1}} \underbrace{\operatorname{cov}(\alpha_t, v_t)'}_{M_t'}$$

$$= P_t - M_t F_t^{-1} M_t'$$

$$= P_t - P_t Z_t' F_t^{-1} (P_t Z_t')'$$

Resultado p/ $P_{t|t}$

$$\therefore \boxed{P_{t|t} = \operatorname{var}(\alpha_t \mid \underset{\sim}{Y}_t) = P_t - P_t Z_t' F_t^{-1} Z_t P_t'}$$

→

# RESUMINDO

Equações do FK

Previsão: $a_{t+1} = T_t a_{t|t}$

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

Atualização: $a_{t|t} = a_t + M_t F_t^{-1} v_t$    onde $v_t = y_t - Z_t a_t$

$$\underset{\substack{\downarrow \\ \text{atualizada}}}{P_{t|t}} = \underset{\substack{\downarrow \\ \text{prevista}}}{P_t} - \underset{\substack{\downarrow \\ \text{correção}}}{P_t Z_t' F_t^{-1} Z_t P_t}$$

$F_t = Z_t P_t Z_t' + H_t$

$M_t = P_t Z_t'$

→ Na prática:

Temos uma previsão sequencial

$\alpha_1 \sim N(a_1, P_1)$ : distribuição a priori inicial conhecida

É o passo de previsão inicial

Começamos na atualização usando $\underset{\substack{\downarrow \\ v_1}}{y_1}$ (1ª observação)

$$\begin{cases} a_{t|t} = a_1 + M_1 F_1^{-1} \textcircled{$v_1$} \text{ usa } y_1 \\ P_{t|t} = P_1 + \dots \end{cases}$$

Com $a_{t|t}$ e $P_{t|t}$, jogamos nas fórmulas da previsão

— // —

Podemos reescrever as fórmulas em função apenas de $a_{t+1}/a_t$ e $P_{t+1}/P_t$ se substituirmos $a_{t|t}$ e $P_{t|t}$ das equações de atualização nas de previsão.

Teremos FK 2 em 1

- $a_{t+1} = T_t(a_t + M_t F_t^{-1} v_t)$

$$= T_t a_t + \underbrace{T_t M_t F_t^{-1}}_{K_t} v_t$$

onde $K_t$ é o ganho de Kalman.

$$\therefore \boxed{a_{t+1} = T_t a_t + K_t \underbrace{v_t}}$$

observaçõs entram aqui

- $$P_{t+1} = T_t \Big( P_t - \underbrace{P_t Z_t'}_{M_t} F_t^{-1} \underbrace{Z_t P_t}_{M_t'} \Big) T_t' + R_t Q_t R_t'$$

$$= T_t P_t \left( I - \cancel{M_t F_t^{-1}} M_t' \right) T_t' + R_t Q_t R_t'$$

↓ dúvida

$$\boxed{P_{t+1} = T_t P_t L_t' + R_t Q_t R_t'}$$ onde $L_t = T_t - K_t Z_t$

↓

Não depende das observações.
Dependendo do caso, converge p/ um valor (steady state). $\bar{P}$

cont.

aula 06/04

- terminar
- Ricatti

Passo da suavização

$$y = (y_1' \; y_2' \cdots y_n')' = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \sim (n \times p) \times 1$$

→ cada elemento é um vetor $p \times 1$

$y$

↓

$\underset{\sim}{Y}_n$

Queremos: $\hat{\alpha}_t = E[\alpha_t | y] \quad t = 1, 2, \ldots, n$

(antes $a_t$)

$V_t = var[\alpha_t | y]$

(antes $P_t$)

antes íamos até t. Agora, até n.

cálculo do valor esperado do suavizado →

$$\hat{\alpha}_t = E[\alpha_t | y] = E[\alpha_t | \underset{\sim}{Y}_{t-1}, v_t, v_{t+1}, \ldots, v_n]$$

No vetor multivariado, do resultado, da aula passada:

$$\begin{cases} x = \alpha_t \\ y = \underset{\sim}{Y}_{t-1} \\ z = (v_t, v_{t+1}, \ldots, v_n)' \end{cases}$$

$$\therefore \hat{\alpha}_t = E[\alpha_t | y] = \underbrace{E[\alpha_t | Y_{t-1}]}_{a_t \text{ (do FK)}} + cov(\alpha_t, (v_t, v_{t+1}, \ldots, v_n)') \cdot (var(v_t \cdots v_n)')^{-1} \cdot (v_t, v_{t+1}, \ldots, v_n)'$$

- Cálculo de $var(v_t, v_{t+1}, \ldots, v_n)^{-1}$

Mostraremos que $v_i$'s são independentes

Faremos isso mostrando que $p(v_1, v_2, \ldots, v_n) = \prod_{i=1}^{n} p(v_i)$

Para $y_i$, podemos escrever:

$$p(y_1, y_2, \ldots y_n) = p(y_n, \underbrace{y_{n-1}, \ldots, y_2, y_1}_{Y_{n-1}})$$

$$= p(y_n | \underset{\sim}{Y}_{n-1}) \, p(\underset{\sim}{Y}_{n-1}) =$$

$$= p(y_n | Y_{n-1}) \cdot p(y_{n-1}, \underset{\sim}{Y}_{n-2}) =$$

$$= p(y_n | \underset{\sim}{Y}_{n-1}) \cdot p(y_{n-1} | \underset{\sim}{Y}_{n-2}) \cdot p(Y_{n-2}) \ldots =$$

$$= p(y_1) \prod_{t=2}^{n} p(y_t | \underset{\sim}{Y}_{t-1}) \quad \text{onde } p(y_t | Y_{t-1}) \sim N(Z_t a_t, F_t$$

Mas queremos $p(\nu_1, \nu_2 \dots \nu_n)$

e Sabemos que $\nu_t = y_t - Z_t a_t$

Resultado Auxiliar:

Se $X$ é vetor aleatório $n \times 1$

e $Y = g(X)$, $g$ diferenciável e com inversa.

$$f_Y(y) = f_X(g^{-1}(y)) \, |J(y)| \quad \text{onde } J(y) = \begin{pmatrix} \partial y_1/\partial x_1 & \dots & \partial y_1/\partial x_n \\ \vdots & & \vdots \\ \partial y_n/\partial x_1 & \dots & \partial y_n/\partial x_n \end{pmatrix}$$

$|J(y)|$ é o determinante

Fazendo $X = (y_1\ y_2 \dots y_n)'$

$Y = (\nu_1, \nu_2 \dots \nu_n)$

$$\Rightarrow p(\nu_1, \nu_2 \dots \nu_n) = p(y_1, y_2, \dots y_n)\Big|_{y_i = \nu_i + Z_i a_i} |J|$$

$$\text{onde } \frac{\partial y_j}{\partial \nu_i} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \Rightarrow J = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots\ 1 \end{pmatrix} \Rightarrow |J| = 1$$

$$\text{Logo: } p(\nu_1, \nu_2, \dots \nu_n) = p(y_1, y_2, \dots y_n)\Big|_{y_i = \nu_i + Z_i a_i}$$

$$= p(y_1) \prod_{t=2}^{n} p(y_t | \underset{\sim}{Y}_{t-1})\Big|_{y_t = \nu_t + Z_t a_t}$$

$\downarrow$

dist. multivariada pois $y_t \sim p \times 1$

$$\text{Sendo } p(y_t | \underset{\sim}{Y}_{t-1}) = \frac{1}{(2\pi)^{p/2} |F_t|^{1/2}} \exp\left\{ -\frac{1}{2} \underbrace{(y_t - Z_t a_t)'}_{\nu_t'} F_t^{-1} \underbrace{(y_t - Z_t a_t)}_{\nu_t} \right\}$$

$$\therefore p(y_t | \underset{\sim}{Y}_{t-1})\Big|_{y_t = \nu_t + Z_t a_t} = \frac{1}{(2\pi)^{p/2} |F_t|^{1/2}} \exp\left\{ -\frac{1}{2} \nu_t' F^{-1} \nu_t \right\} =$$

$$= p(\nu_t)$$

$t = 2, \dots n$

$$\Rightarrow p(\nu_1\ \nu_2 \dots \nu_n) = p(y_1) \cdot \prod_{t=2}^{n} p(\nu_t)$$

Calculemos $p(y_1)$:

$t = 1 \Rightarrow y_t = Z_t \alpha_t + \varepsilon_t \therefore y_1 = Z_1 \alpha_1 + \varepsilon_1$

$\cdot\ E[y_1] = Z_1 E[\alpha_1] = Z_1 a_1$

$\cdot\ var[y_1] = E[(y_1 - E[y_1])(y_1 - E[y_1])']$

$= E[(Z_1(\alpha_1 - a_1) + \varepsilon_1)(Z_1(\alpha_1 - a_1) + \varepsilon_1)']$

$= E[Z_1(\alpha_1 - a_1)(\alpha_1 - a_1)' Z_1'] + E[\varepsilon_1 \varepsilon_1'] =$

$= Z_1 P_1 Z_1' + H_1 = F_1$

como $v_1 = y_1 - Z_1 a_1 \Rightarrow E[v_1] = 0$
$var[v_1] = F_1$

$$\therefore p(y_1) = \frac{1}{(2\pi)^{p/2}|F_1|^{1/2}} \exp\left\{-\frac{1}{2}(y_1 - Z_1 a_1)' F_1^{-1}(y_1 - Z_1 a_1)\right\}$$

$$\therefore p(v_1) = \frac{1}{(2\pi)^{p/2}|F_1|^{1/2}} \exp\left\{-\frac{1}{2} v_1' F_1^{-1} v_1\right\}$$

Logo: $p(v_1, v_2, \dots, v_n) = \prod_{t=1}^{n} p(v_t) \Rightarrow$ (i) $v_t$'s são independentes entre si
(ii) $f(v_t | \underset{\sim}{Y}_{t-1}) = f(v_t)$

$\Rightarrow (v_t, v_{t+1}, \dots v_n)$ são também Normais Multivariadas e independentes

Teremos então que $E[v_i v_j'] = 0 \quad \forall\ i, j = t, t+1, \dots, n \quad i \neq j$

Esta independência entre $v_i$'s implica em:

$$var(v_t, v_{t+1}, \dots, v_n) = E\left[\begin{pmatrix} v_t \\ v_{t+1} \\ \vdots \\ v_n \end{pmatrix}(v_t' \ v_{t+1}' \ \dots \ v_n')\right]$$

$$= E\begin{bmatrix} v_t v_t' & v_t v_{t+1}' & \dots & v_t v_n' \\ \vdots & & & \\ v_n v_t' & & & v_n v_n' \end{bmatrix}$$

$$\therefore \quad var(v_t, v_{t+1} \cdots v_n) = \begin{pmatrix} F_t & 0 & \cdots & 0 \\ 0 & F_{t+1} & & \vdots \\ \vdots & & & \\ 0 & \cdots & & F_n \end{pmatrix}$$

- Cálculo de $cov(\alpha_t, (v_t, v_{t+1} \cdots v_n)')$

$$cov(\alpha_t, (v_t, v_{t+1} \cdots v_n)') = E\Big[\underset{m\times 1}{\alpha_t}(\underset{1\times p}{v_t'}, v_{t+1}' \cdots v_n')\Big] \quad \text{pois } E[v_i]=$$

$$= \Big[\underbrace{E(\alpha_t v_t')}_{m\times p} \;\; E[\alpha_t v_{t+1}'] \cdots E[\alpha_t v_n']\Big]$$

⇒ Na expressão para $\hat{\alpha}_t$, temos:

$$cov(\alpha_t, (v_t, v_{t+1} \cdots v_n)') \cdot var(v_t \cdots v_n)' \cdot (v_t \; v_{t+1} \cdots v_n)' =$$

$$= \Big[E(\alpha_t v_t') \;\; E[\alpha_t v_{t+1}') \cdots E[\alpha_t v_n')\Big] \begin{bmatrix} F_t^{-1} & & 0 \\ & \ddots & \\ 0 & & F_n^{-1} \end{bmatrix} \begin{bmatrix} v_t \\ v_{t+1} \\ \vdots \\ v_n \end{bmatrix} =$$

$$= \Big[E[\alpha_t v_t') \cdots E[\alpha_t v_n']\Big] \begin{bmatrix} F_t^{-1} v_t \\ F_{t+1}^{-1} v_{t+1} \\ \vdots \\ F_n^{-1} v_n \end{bmatrix} =$$

$$= \sum_{j=t}^{n} E[\alpha_t v_j') F_j^{-1} v_j$$

**Logo:** $$\boxed{\underset{\substack{\downarrow \\ \text{estimativa} \\ \text{suavizada}}}{\hat{\alpha}_t} = E[\alpha_t | Y_n] = \underset{\substack{\downarrow \\ \text{estimativa} \\ \text{prevista}}}{a_t} + \sum_{j=t}^{n} E[\alpha_t v_j') F_j^{-1} v_j}$$

→ Outros resultados dos quais precisaremos:

Seja $x_t \triangleq \alpha_t - a_t$ (como se fosse um erro de previsão).

$var(x_t) = var(\alpha_t) = P_t$

Escrevendo $v_t$ em função de $x_t$

$$v_t = y_t - Z_t a_t$$
$$= Z_t \alpha_t + \varepsilon_t - Z_t a_t$$
$$= Z_t (\alpha_t - a_t) + \varepsilon_t$$
$$\boxed{v_t = Z_t x_t + \varepsilon_t} \quad (a)$$

Temos ainda que: (escrever $x_{t+1}$ em função de $x_t$)

$$x_{t+1} = \alpha_{t+1} - a_{t+1}$$
$$= \underbrace{(T\alpha_t + R_t \eta_t)}_{\text{do modelo EE}} - \underbrace{(T_t a_t + K_t v_t)}_{\text{do FK}} =$$
$$= T_t \underbrace{(\alpha_t - a_t)}_{x_t} + R_t \eta_t - K_t (Z_t x_t + \varepsilon_t) =$$
$$= T_t x_t + R_t \eta_t - K_t Z_t x_t + K_t \varepsilon_t$$
$$= \underbrace{(T_t - K_t Z_t)}_{L_t} x_t + R_t \eta_t - K_t \varepsilon_t =$$
$$\therefore \boxed{x_{t+1} = L_t x_t + R_t \eta_t - K_t \varepsilon_t} \quad (b)$$

As equações (a) e (b) podem ser vistas como uma representação em inovações do MEE:

$$\begin{cases} v_t = Z_t x_t + \varepsilon_t \\ x_{t+1} = L_t x_t + \underbrace{(R_t \eta_t - K_t \varepsilon_t)}_{\text{ñ são descorrelatadas}} \end{cases} \quad (=) \quad \begin{matrix} y_t = Z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} = T_t \alpha_t + R_t \eta_t \end{matrix}$$

- Em $\hat{\alpha}_t$, onde tem-se $E[\alpha_t v_j']$, podemos escrever:

$$E[\alpha_t v_j'] = E[\alpha_t (Z_j x_j + \varepsilon_j)'] =$$
$$= E[\alpha_t x_j'] Z_j' \qquad j = t \ldots n$$

calculando $E[\alpha_t x_j']$:

$$E[\alpha_t x_t'] = E[E(\alpha_t x_t' \mid \underset{\sim}{y}_{t-1})])$$

$$= E[E[\alpha_t (\alpha_t - a_t)' \mid Y_{t-1}]]) \;\}\; \text{truque}$$

$$= E[E[(\alpha_t - a_t)(\alpha_t - a_t)' \mid Y_{t-1}]])$$

$$= E(P_t] = P_t$$

calculando sucessivamente:

$$\rightarrow E(\alpha_t x_{t+1}') = E[E[\alpha_t x_{t+1}' \mid \underset{\sim}{y}_{t-1}))=$$

$$= E[E(\alpha_t (L_t x_t + R_t \eta_t - K_t \varepsilon_t)' \mid Y_{t-1}]) =$$

$$= E[E[\alpha_t (x_t' L_t' + \eta_t' R_t' - \cancel{\varepsilon_t' K_t'}) \mid \underset{\sim}{Y}_{t-1}]] =$$

$$= E[E((\alpha_t x_t' L_t' \mid \underset{\sim}{Y}_{t-1})] =$$

$$= E[E(\alpha_t (\alpha_t - a_t)' L_t' \mid \underset{\sim}{Y}_{t-1}])) =$$

$$= E[E[(\alpha_t - a_t)(\alpha_t - a_t)' L_t' \mid \underset{\sim}{Y}_{t-1})] =$$

$$= P_t L_t'$$

$$\rightarrow E(\alpha_t x_{t+2}') = E[E[\alpha_t x_{t+2}' \mid \underset{\sim}{y}_{t-1}]]) =$$

$$= E[E(\alpha_t (L_{t+1} x_{t+1} + R_{t+1} \eta_{t+1} - K_{t+1} \varepsilon_{t+1})' \mid \underset{\sim}{Y}_{t-1}]]) =$$

$$= E[E[\alpha_t (x_{t+1}' L_{t+1}' + \eta_{t+1}' R_{t+1}' - \varepsilon_{t+1}' K_{t+1}') \mid \underset{\sim}{Y}_{t-1}])) =$$

$$= \underbrace{E[E[\alpha_t x_{t+1}' \mid \underset{\sim}{Y}_{t-1}]]}_{} L_{t+1}'$$

$$= P_t L_t' L_{t+1}'$$

$$\vdots$$

$$\rightarrow E[\alpha_t x_n'] = P_t L_t' L_{t+1}' \cdots L_{n-1}' \quad \text{(por indução)}$$

Podemos reescrever:

$$\hat{\alpha}_t = a_t + \sum_{j=t}^{n} \underbrace{E[\alpha_t v_j']}_{} F_j^{-1} v_j$$

$$E[\alpha_t x_j'] \, z_j' =$$

$$= P_t L_t' L_{t+1}' \cdots L_{j-1}' z_j'$$

$$\therefore \quad \hat{\alpha}_t = a_t + P_t Z_t' F_t^{-1} v_t + P_t L_t' Z_{t+1}' F_{t+1}' v_{t+1} + \ldots + P_t L_t' L_{t+1}' \cdots L_n' Z_n' F_n'$$

Quando $t = n \Rightarrow$ passo de atualizações coincide com o de smoothing

$$\hat{\alpha}_n = E[\alpha_n | y_1 \ldots y_n]$$

$$a_{n|n} = E[\alpha_n | \underset{\sim}{Y}_n]$$

Logo: $\hat{\alpha}_n = a_n + P_n Z_n^{-1} F_n^{-1} v_n$ $\quad (a_{t|t} = a_t + M_t F_t^{-1} v_t$ onde $M_t = P_t Z_t)$

Voltando um passo atrás p/ $t = n-1$

$$\hat{\alpha}_{n-1} = a_{n-1} + \sum_{j=n-1}^{n} E[\alpha_t v_j'] F_j^{-1} v_j \quad (2 \text{ termos})$$

$(n-1$ em $\hat{\alpha}_t)$

$$= a_{n-1} + P_{n-1} Z_{n-1}' F_{n-1}^{-1} v_{n-1} + P_{n-1} L_{n-1}' Z_n' F_n^{-1} v_n$$

**Portanto**

$$\hat{\alpha}_t = a_t + P_t Z_t' F_t^{-1} v_t + P_t L_t' Z_{t+1}' F_{t+1}' v_{t+1} + \ldots + P_t L_t' L_{t+1}' \cdots L_n' Z_n' F_n^{-1} v_n$$

$$\Rightarrow \boxed{\hat{\alpha}_t = a_t + P_t r_{t-1}}$$

onde $r_{t-1} = Z_t' F_t^{-1} v_t + L_t' Z_{t+1}' F_{t+1}^{-1} v_{t+1} + \ldots + L_t' L_{t+1}' \cdots L_n' Z_n F_n^{-1} v_n$

$t = n-2, n-3, \ldots \underline{\underline{n}}$

P/ $t = n \Rightarrow r_{n-1} = Z_n' F_n^{-1} v_n$

P/ $t = n-1 \Rightarrow r_{n-2} = Z_{n-1}' F_{n-1}^{-1} v_{n-1} + L_{n-1}' Z_n' F_n^{-1} v_n$

onde $r_{t-1}$: soma ponderada das inovações $v_j$ ocorridas após o tempo $t$.

$$r_t = Z_{t+1}' F_{t+1}^{-1} v_{t+1} + L_{t+1}' Z_t' F_t' v_t + \ldots + L_{t+1}' \cdots L_{n-1}' Z_n' F_n^{-1} v_n$$

com $r_n \equiv 0$ pois inovações até $t = n$.

Escrevendo de forma recursiva, temos:

$$\boxed{r_{t-1} = Z_t' F_t^{-1} v_t + L_t' r_t \ , \ r_n \equiv 0.}$$

A estimativa suavizada do vetor de estado será dada por:

$$\underbrace{\hat{\alpha}_t}_{E[\alpha_t | \underset{\sim}{y}_n]} = \underbrace{a_t}_{E[\alpha_t | \underset{\sim}{y}_{t-1}]} + P_t \underbrace{r_{t-1}}_{\text{soma ponderada de inovações entre } t \text{ e } n}$$

onde: $r_{t-1} = Z_t' F_t^{-1} v_t + L_t' r_t$ com $r_n \equiv 0$

↳ suavizador de intervalo fixo.

Cálculo da variância suavizada →

$$V_t = \text{var}(\alpha_t | \underset{\sim}{y}_n)$$

$$= \text{var}[\alpha_t | \underset{\sim}{y}_{t-1}, v_t, v_{t+1}, \ldots, v_n]$$

No resultado p/ vetor multivariado, temos: → $X = \alpha_t$, $Y = \underset{\sim}{y}_{t-1}$, $Z = (v_t \ldots v_n)'$

$$V_t = \text{var}(\alpha_t | \underset{\sim}{y}_{t-1}) - \text{cov}(\alpha_t, (v_t, v_{t+1} \ldots v_n)') \underbrace{[\text{var}(v_t \ldots v_n)']^{-1}}_{\text{diagonal}} \text{cov}(\alpha_t, (v_t, v_{t+1} \ldots v_n)')'$$

$$\underbrace{V_t}_{\text{variância suavizada}} = \underbrace{P_t}_{\text{variância prevista}} - \sum_{j=t}^{n} \text{cov}(\alpha_t v_j) F_j^{-1} \text{cov}(\alpha_t, v_j)'$$

$\Rightarrow$ $V_t$ é mais preciso do que $P_t$

- Já vimos que:

$$\text{cov}(\alpha_t, v_j) = E[\alpha_t v_j'] = E[\alpha_t x_j'] Z_j' \qquad j = t, \ldots, n$$

$$\Rightarrow V_t = P_t - \sum_{j=t}^{n} E[\alpha_t x_j'] Z_j' F_j^{-1} Z_j (E[\alpha_t x_j'])'$$

$$= P_t - [E[\alpha_t x_t'] Z_t' F_t^{-1} Z_t (E(\alpha_t x_t'))' + E[\alpha_t x_{t+1}'] . Z_{t+1}' F_{t+1}^{-1} Z_{t+1} . [E(\alpha_t x_{t+1}'$$

$$+ \ldots + E[\alpha_t x_n'] Z_n' F_n^{-1} v_n (E[\alpha_t x_n'])']$$

- Já vimos também que

$$E[\alpha_t x_t'] = P_t$$

$$E[\alpha_t x_{t+1}'] = P_t L_t'$$

$$\vdots$$

Logo:

$$V_t = P_t - P_t Z_t' F_t^{-1} Z_t P_t' - P_t L_t' Z_{t+1}' F_{t+1}^{-1} Z_{t+1} L_t P_t' - \ldots$$

$$- P_t L_t' \ldots L_{n-1}' Z_n' F_n^{-1} v_n L_{n-1} \ldots L_t P_t$$

$$= P_t - P_t ( Z_t' F_t^{-1} Z_t - L_t' Z_{t+1}' F_{t+1}^{-1} Z_{t+1} L_t - \ldots ) P_t$$

$\therefore \quad V_t = P_t - P_t N_{t-1} P_t$

onde $N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' Z_{t+1}' F_{t+1}^{-1} Z_{t+1} L_t + \dots + L_t' \dots L_{n-1}' Z_n' F_n^{-1} Z_n L_{n-1} \dots L_t$

$N_t = Z_{t+1}' F_{t+1}^{-1} Z_{t+1} + L_{t+1}' Z_{t+2}' F_{t+2}^{-1} Z_{t+2} L_{t+1} + \cdots$

$$\Rightarrow \quad \boxed{N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \qquad t = n, \dots, 1}$$

Em $t = n-1$ : começa recursão $\Rightarrow$ com $N_n = 0$

## Portanto

$$\left\{ \begin{array}{l} V_t = P_t - P_t N_{t-1} P_t \\ \downarrow \\ \text{var}[\alpha_t \mid \underset{\sim}{Y_n}] \\ \text{onde } N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \qquad N_n = 0. \end{array} \right.$$

$N_{t-1}$ ↓ em $t = n$ : dá = a estimativa da previsão.

## Observação

Para o $r_t$, temos: $\left\{ \begin{array}{l} E(r_t) = 0 \\ \text{var}(r_t) = N_t \end{array} \right.$

## Resumo

$\hat{\alpha}_t = a_t + P_t r_{t-1}$

$r_{t-1} = Z_t' F_t^{-1} v_t + L_t' r_t$

$r_n = 0$

$V_t = P_t - P_t N_{t-1} P_t$

$N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t$

$N_n = 0.$

Aula 06/04

## Equações de Ricatti (steady state FK)

Equações p/

$$\text{Var}[\alpha_{t+1} \mid \underset{\sim}{Y_t}] = P_{t+1}$$

$\Rightarrow P_{t+1} = T_t P_t L_t' + R_t Q_t R_t'$ $\Rightarrow$ Não depende de observações

Considerando sistemas invariantes no tempo (matrizes do sistema independem de t)

$$P_{t+1} = T P_t L_t' + RQR'$$

onde $L_t = T - K_t Z$

e $K_t = T P_t Z' F_t^{-1}$

e $F_t = Z P_t Z' + H$

$$\Rightarrow P_{t+1} = T P_t \left[T - (T P_t Z' F_t^{-1}) Z\right]' + RQR'$$

$$\boxed{P_{t+1} = T P_t T' - T P_t Z' F_t^{-1} Z P_t T' + RQR'}$$

Equações de Ricatti

Supondo que $\lim_{t \to \infty} P_{t+1} = P_t = \bar{P}$

$\Rightarrow \bar{P} = T \bar{P} T' - T \bar{P} Z' \bar{F}^{-1} Z \bar{P} T' + RQR$ e $\bar{F} = Z \bar{P} Z' + H$

Precisamos resolver p/ $\bar{P}$

Exemplo:

1)

$y_t = \mu_t + \varepsilon_t$  $\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$

$\mu_{t+1} = \mu_t + \eta_t$  $\eta_t \sim N(0, \sigma^2_\eta)$

Nível local

$T = Z = 1$  $R = 1$

$$\bar{P} = \bar{P} - \bar{P} \underset{(\bar{P} + \sigma^2_\varepsilon)^{-1}}{\underset{\downarrow}{\bar{F}^{-1}}} \bar{P} + \sigma^2_\eta \quad \therefore \quad \bar{P} = \bar{P}\left(1 - \frac{\bar{P}}{\bar{P} + \sigma^2_\varepsilon}\right) + \sigma^2_\eta$$

∴ $x^2 - xh - h = 0$ onde $\bar{x} = \frac{\bar{P}}{\sigma_\varepsilon^2}$

$h = \sigma_\eta^2 / \sigma_\varepsilon^2 = q$

Solução: $x = \frac{h + \sqrt{h^2+4h}}{2}$ desde que $h > 0$.

2) $y_t = y_{t-1}\varphi_t + \varepsilon_t$

$\varphi_t = \phi\varphi_{t-1} + \eta_t$

⇒ sistema não é invariante no tempo.

Jamais haverá solução de Ricatti

3) Tendência local

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

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

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

Na forma EE:

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

Fazer

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

$Z = \begin{pmatrix} 1 & 1 \end{pmatrix}$ $\quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\quad Q = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_k^2 \end{pmatrix}$

$T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ $\quad H = \sigma_\varepsilon^2 I$

$$\bar{P} = T\bar{P}T' - T\bar{P}Z'\bar{F}^{-1}Z\bar{P}T' + RQR'$$

$$\bar{F} = \underset{1\times 2}{Z}\,\underset{2\times 2}{\bar{P}}\,\underset{2\times 1}{Z'} + H = \begin{pmatrix} 1 & 1 \end{pmatrix} \bar{P} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sigma_\varepsilon^2$$

aula 06/04 - cont

Observações faltantes

- Em $t = \tau, \tau+1, \dots, \tau^*-1 \Rightarrow$ observações de $y_t$ não existem
- Como aplicar FK | Smoothing p/ observações faltantes?

Como preencher estas observações (dependerá do MEE usado)?

Modelo

$$y_t = Z_t \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, H_t)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim N(0, Q_t)$$

Nos períodos em que $\exists$ observações, as equações FK usadas são

$$FK \begin{cases} a_{t+1} = T_t a_t + K_t v_t & t \leq \tau - 1 \\ P_{t+1} = T_t P_t L_t' + R_t Q_t R_t' & t \geq \tau^* \\ L_t = T_t - K_t Z_t & \\ K_t = T_t P_t Z_t' F_t^{-1} & \end{cases}$$

Nos períodos de dados faltantes $(t = \tau, \tau+1, \dots, \tau^*-1)$, não usamos expressões de FK pois $\nexists y_t \therefore \nexists v_t$, $\nexists F_t$ $\Rightarrow$ $v_t = 0$, $K_t = 0$ $\Big\}$ $L_t = T_t$

- Valor Esperado Iremos então "extrapolar o estado", realizando a substituição:

$$a_t \longrightarrow a_{\tau+j} \qquad j = 1, 2, \dots$$

(p/ $t \in (\tau, \tau^*-1)$

onde $a_{\tau+j} \hat{=} E[\alpha_{\tau+j} | \underset{\sim}{Y}_{\tau-1}]$ ↘ última observação

Supondo sistema invariante no tempo

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

Iterando:

$$\alpha_{t+2} = T\alpha_{t+1} + R\eta_{t+1} = T^2\alpha_t + TR\eta_t + R\eta_{t+1}$$

$$\vdots$$

P/ j passos a frente

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

Fazendo $t = \tau$

$$\alpha_{\tau+j} = T^j\alpha_\tau + \sum_{i=1}^{j} T^{j-i} R\eta_{\tau+i-1}$$

Assim:

$$E[\alpha_{\tau+j} | Y_{\tau-1}] = T^j E[\alpha_\tau | \underset{\sim}{Y}_{\tau-1}]$$

↳ não tem obs. em $\tau$ mas tem estimativa direta

$$a_{\tau+j} = E[\alpha_{\tau+j} | Y_\tau] = E[\alpha_{\tau+j} | Y_{\tau-1}]$$

$$\therefore \boxed{a_{\tau+j} = \tau^j a_\tau} \quad j = 1, 2, \ldots, \tau^* - \tau - 1$$

O FK nesse período será substituído por $a_{\tau+j}$, o que equivale

$$\boxed{a_{t+1} = T a_t} \quad (t = \tau, \tau+1, \ldots, \tau^* - 1) \quad \underline{\underline{\text{ou}}} \quad a_{t+1} = T_t a_t$$

pois

$$a_{\tau+1} = T a_\tau$$

$$a_{\tau+2} = T a_{\tau+1} = T^2 a_{\tau+2} \ldots$$

<u>Na prática</u>

Fazer $v_t = 0$ na equação do FK

- Variância

Tb projetaremos a variância j passos a frente

↳ Atenção! É uma matriz

Queremos a expressão para

$$P_{\tau+j} = \text{var}(\alpha_{\tau+j} | Y_{\tau-1}) = E\left[(\alpha_{\tau+j} - a_{\tau+j})(\alpha_{\tau+j} - a_{\tau+j})' \,|\, \underset{\sim}{Y}_{\tau-1}\right]$$

mas:

$$\alpha_{\tau+j} = T^j \alpha_\tau + \sum_{i=1}^{j} T^{j-i} R \eta_{\tau+i-1}$$

$$a_{\tau+j} = T^j a_\tau$$

$$\Rightarrow (\alpha_{\tau+j} - a_{\tau+j}) = T^j(\alpha_t - a_t) + \sum_{i=1}^{j} T^{j-i} R \eta_{t+i-1}$$

$$\therefore P_{\tau+j} = E\Big[\Big(T^j(\alpha_\tau - a_\tau) + \sum_{i=1}^{j} T^{j-i} R \eta_{t+i-1}\Big)\Big((\alpha_\tau - a_\tau)' T^{j'} + \sum_{i=1}^{j} \eta'_{t+i-1} R' T^{j-i'}\Big)$$

$$= T^j E\big[(\alpha_\tau - a_\tau)(\alpha_\tau - a_\tau)' | \underset{\sim}{Y}_{\tau-1}\big] T^{j'} + \sum_{i=1}^{j} T^{j-i} R E\big[\eta_{\tau+i-1}\eta'_{\tau+i-1} | \underset{\sim}{Y}_{\tau-1}\big] R$$

(T

(termos cruzados são nulos)

$$= T^j P_\tau T^{j'} + \sum_{i=1}^{j} T^{j-i} R Q R' (T^{j-i})' \qquad j = 1, 2, \ldots, \tau^* - \tau - 1$$

o que é equivalente a:

$$\boxed{P_{t+1} = T P_t T' + R Q R'}$$

pois $P_{\tau+1} = T P_\tau T' + R Q R'$

$$P_{\tau+2} = T P_{\tau+1} T' + R Q R' = T\,[T P_\tau T' + R Q R']T' + R Q R'$$

$$= T^2 P_\tau (T')^2 + T R Q R' T' + R Q R'$$

$$\vdots$$

Na prática

Fazer $K_t = 0$ (pois $F_t^{-1} = 0$) no FK

$\therefore L_t = T_t$

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

$$\vdots$$

RESUMINDO: usamos FK p/ os períodos de dados faltantes, fazendo $v_t = 0$ e $F_t^{-1} = 0 \Rightarrow K_t = 0$

Ⓐ $$\therefore \begin{cases} a_{t+1} = T_t a_t \\ P_{t+1} = T_t P_t T_t' + R_t Q_t R_t' \end{cases}$$

//

## Para Suavização

As equações são:

$$\begin{cases} \hat{\alpha}_t = a_t + P_t r_{t-1} \\ r_{t-1} = Z_t F_t^{-1} v_t + L_t' r_t \end{cases} \qquad \begin{cases} V_t = P_t - P_t N_{t-1} P_t \\ N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \end{cases}$$

$r_n = 0$ , $t = n, n-1, \ldots, 1$ $\qquad$ $N_n = 0$ , $L_t = T_t - K_t Z_t$

Como $v_t = 0$ e $F_t^{-1} = 0$

Ⓑ $$\Rightarrow \begin{cases} \hat{\alpha}_t = a_t + P_t r_{t-1} \\ r_{t-1} = L_t' r_t = T_t' r_t \end{cases}$$

Ⓒ $$\begin{cases} V_t = P_t - P_t N_{t-1} P_t \\ N_{t-1} = T_t' N_t T_t \end{cases}$$

$t = \tau^* - 1, \ldots, \tau$

(Nos períodos de dados faltantes $a_t$ e $P_t$ são estimados com Ⓐ)

## Preenchimento das obs. faltantes

Podemos substituir as obs. faltantes através da seguinte estratégia:

$y_t$ subst. por $E[y_t | \underset{\sim}{Y}_n]$ em $t = \tau, \tau+1, \ldots, \tau^* - 1$

Como $y_t = Z_t \alpha_t + \varepsilon_t$ , $\forall t$

$$E[y_t | \underset{\sim}{Y}_n] = Z_t E[\alpha_t | \underset{\sim}{Y}_n] + E[\varepsilon_t | \underset{\sim}{Y}_n]$$

$$= Z_t \hat{\alpha}_t + \hat{\varepsilon}_t \Rightarrow$$ atenção! Não é zero

Não confundir com $E[\varepsilon_t | \underset{\sim}{Y}_{t-1}]$

Atenção!

$\hat{\varepsilon}_\tau = E[\varepsilon_t | \underset{\sim}{Y}_n]$ : na pg 73, por 4.4 $\Rightarrow$ $\hat{\varepsilon}_t = H_t \mu_t$

onde $H_t = \text{var}(\varepsilon_t)$

$\mu_t = F_t^{-1} v_t - K_t' r_t$

Mas p/ obs. faltantes: $v_t = 0$, $F_t^{-1} = 0$, $K_t = 0$

$$\therefore E[y_t | \underset{\sim}{Y}_n] = \hat{y}_t = Z_t \hat{\alpha}_t$$

$\downarrow$ dado por Ⓑ $\qquad t = \tau, \tau+1. \ \tau^* - 1$

## Exemplo

1) $\begin{cases} y_t = \mu_t + \varepsilon_t & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_{t+1} = \mu_t + \eta_t & \eta_t \sim N(0, \sigma_\eta^2) \end{cases}$

$T = Z = R = 1$

$H = \sigma_\varepsilon^2$

$\varphi = \sigma_\eta^2$

- Usando equações Ⓐ : $\begin{cases} a_{t+1} = a_t \\ P_{t+1} = P_t + \sigma_\eta^2 \end{cases}$ $\quad t = \tau, \tau+1, \dots, \tau^*-1$

FK

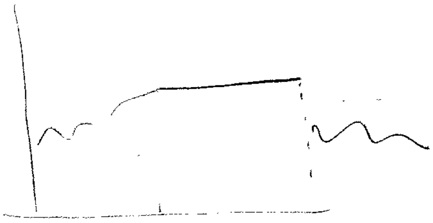
estimada por $a_{\tau-1}$

Iterando: $\Rightarrow a_{\tau+1} = a_\tau$

$a_{\tau+2} = a_{\tau+1} = a_\tau \Rightarrow a_{\tau+j} = a_\tau \quad j = 1, 2, \dots$

$\Rightarrow P_{\tau+1} = P_\tau + \sigma_\eta^2$

$P_{\tau+2} = P_{\tau+1} + \sigma_\eta^2 = P_\tau + 2\sigma_\eta^2 \Rightarrow P_{\tau+j} = P_\tau + j\sigma_\eta^2$



- Por Smoothing : $L = 1$

$v_t = 0$

$F_t^{-1} = 0$

$$\left.\begin{array}{l} \hat{\alpha}_t = a_t + P_t r_{t-1} \\ r_{t-1} = r_t \Rightarrow r_\tau = r_{\tau-1} \end{array}\right\} \quad \begin{aligned} \hat{\alpha}_{\tau+j} &= a_\tau + P_{\tau+j} r_{\tau+j-1} \\ &= a_\tau + (\underbrace{P_\tau + j\sigma_\eta^2}_{\text{do FK}}) r_\tau \\ &= \underbrace{(a_\tau + P_\tau r_\tau)}_{a} + \underbrace{(\sigma_\eta^2 r_\tau)}_{b} j \end{aligned}$$

Preenchimento $\hat{\alpha}_{\tau+j} = a + bj \;,\; j = 1, 2, \dots, \tau^* - 1 - \tau$

$\Rightarrow \quad \hat{y}_t = \hat{\alpha}_t \quad \underline{\underline{\text{ou}}} \quad \hat{y}_\tau = \hat{\alpha}_\tau = a_\tau + P_\tau \wedge_\tau$

$(= E[\alpha_t | \underset{\sim}{y}_n])$

$\hat{y}_{\tau+1} = \hat{\alpha}_{\tau+1} = a + b$

$\hat{y}_{\tau+2} = \hat{\alpha}_{\tau+2} = a + \textcircled{2b}$

mas entendi $\underline{\underline{\text{ok}}}$

2) $y_t = \mu_t + \varepsilon_t$

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

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

Material Adicional - 06/04

## Inicializações

$\alpha_1 \sim N(a_1, P_1)$

Na prática, não sabemos quem são $a_1$ e $P_1$

$\alpha_1 \sim (\alpha_1 | Y_0)$ : a priori

No FK:

$\nu_t = y_t - a_t$

? → $(P_t + \sigma_\varepsilon^2 = F_1 \quad (Z_t P_t Z_t' + \sigma_\varepsilon^2 = F_t)$ ?

$$\begin{cases} a_{t+1} = a_t + k_t \nu_t \\ P_{t+1} = P_t(1-k_t) + \sigma_\eta^2 \\ k_t = P_t / F_1 \end{cases}$$

Para $t=1 \Rightarrow \nu_1 = y_1 - a_1$

$F_1 = P_1 - \sigma_\varepsilon^2$

Para $t=2 \Rightarrow a_2 = a_1 + k_1 \nu_1$

$$= a_1 + \frac{P_1}{F_1}\nu_1 = a_1 + \frac{P_1}{P_1+\sigma_\varepsilon^2}(y_1 - a_1)$$

e $P_2 = P_1(1-k_1) + \sigma_\eta^2$

$$= P_1\left(1 - \frac{P_1}{P_1+\sigma_\varepsilon^2}\right) + \sigma_\eta^2 = \frac{P_1}{P_1+\sigma_\varepsilon^2} \cdot \sigma_\varepsilon^2 + \sigma_\eta^2$$

Fazendo $P_1 \to \infty$ obtemos:

$$\begin{cases} a_2 = a_1 + (y_1 - a_1) = y_1 \quad \forall a_1 \\ P_2 = \sigma_\varepsilon^2 + \sigma_\eta^2 \end{cases}$$

ou seja, para $t = 2, \ldots, n$, as equações do FK não apresentam problema pois as distribuições são próprias (bem definidas)

Observação importante: Inicializar o FK em $t=1$ com $a_1$, $P_1 = k \to \infty$ equivale a adotar a priori

$(\alpha_2 | y_1) \sim N(a_2 = y_1, P_2 = \sigma_\varepsilon^2 + \sigma_\eta^2)$

→ Efeito das condições iniciais no alisamento

estimativa suavizada •
$$\hat{\alpha}_t = a_t + P_t r_{t-1}$$
$$= a_t + P_t \left( v_t F_t^{-1} + L_t r_t \right) \quad \text{onde } L_t = 1 - k_t = \frac{\sigma_\varepsilon^2}{F_1} = \frac{\sigma_\varepsilon^2}{P_t + \sigma_\varepsilon^2}$$

$$\therefore \hat{\alpha}_t = a_t + P_t \left[ \frac{(y_t - a_t)}{P_t + \sigma_\varepsilon^2} + \frac{\sigma_\varepsilon^2}{P_t + \sigma_\varepsilon^2} r_t \right]$$

Para $\hat{\alpha}_1$:

$$\hat{\alpha}_1 = a_1 + \frac{P_1 v_1}{P_1 + \sigma_\varepsilon^2} + \frac{P_1 \sigma_\varepsilon^2 r_1}{P_1 + \sigma_\varepsilon^2}$$

$$\text{se } P_1 \to \infty : \boxed{\hat{\alpha}_1 = a_1 + v_1 + \sigma_\varepsilon^2 r_1}$$

variância suavizada •
$$V_t = P_t - P_t^2 N_{t-1}$$
$$= P_t - P_t^2 \left( F_t^{-1} + L_t^2 N_t \right)$$
$$= P_t - P_t^2 \left( F_t^{-1} + \left( \frac{\sigma_\varepsilon^2}{P_t + \sigma_\varepsilon^2} \right) N_t \right)$$

Para $V_1$:
$$V_1 = P_1 - P_1^2 \left( \frac{1}{P_1 + \sigma_\varepsilon^2} + \frac{\sigma_\varepsilon^2}{P_1 + \sigma_\varepsilon^2} N_1 \right)$$

$$= \left( \frac{P_1}{P_1 + \sigma_\varepsilon^2} \right) \sigma_\varepsilon^2 - \sigma_\varepsilon^4 N_1 \left( \frac{P_1}{P_1 + \sigma_\varepsilon^2} \right)^2$$

$$\text{Se } P_1 \to \infty : \boxed{V_1 = \sigma_\varepsilon^2 - \sigma_\varepsilon^4 N_1}$$

⇒ Desta forma, a média do distúrbio suavizada é:

? → $$\hat{\varepsilon}_t = \sigma_\varepsilon^2 \mu_t \quad \therefore \quad \hat{\varepsilon}_1 = \hat{\sigma}_\varepsilon^2 \mu_1$$

$$\mu_t = F_t^{-1} v_t - K_t r_t \qquad \text{onde } K_t = \frac{P_t}{F_t} = \frac{P_t}{\sigma_\varepsilon^2 + P_t}$$
$$\mu_1 = F_1^{-1} v_1 - K_1 r_1$$
$$= \left( \frac{1}{P_1 + \sigma_\varepsilon^2} \right) v_1 - \left( \frac{P_1}{P_1 + \sigma_\varepsilon^2} \right) r_1$$

$$\lim \mu_1 = 0 - r_1 = -r_1 \quad \therefore \quad \lim_{P_1 \to \infty} \hat{\varepsilon}_t = - r_1 \hat{\sigma}_\varepsilon^2$$

Material Adicional - 06/04

## Estimação dos Hiperparâmetros

→ MV: $p(y) = p(y_1, \dots, y_n) = \prod_{t=1}^{n} p(y_t | Y_{t-1})$

onde $p(y_1 | Y_0) = p(y_1)$

$p(y_t | Y_{t-1}) \sim N(\overset{\nearrow z_t a_t}{a_t}, F_t)$

e $\nu_t = y_t - a_t$

$$p(y_t | Y_{t-1}) = \frac{1}{(2\pi F_t)^{1/2}} \exp\left\{ -\frac{1}{2} \frac{(y_t - a_t)^2}{F_t} \right\}$$

$y$ escalar $(p=1)$

$$\therefore p(y) = \prod_{t=1}^{n} p(y_t | Y_{t-1}) = \frac{1}{(2\pi)^{n/2}} \left( \prod_{t=1}^{n} \frac{1}{F_t^{1/2}} \right) \prod_{t=1}^{n} \exp\left\{ -\frac{1}{2} \frac{\nu_t^2}{F_t} \right\}$$

$$= \frac{1}{(2\pi)^{n/2}} \cdot \prod_{t=1}^{n} \frac{1}{F_t^{1/2}} \cdot \exp\left\{ -\frac{1}{2} \sum_{t=1}^{n} \frac{\nu_t^2}{F_t} \right\}$$

Logo:

$$\log L(\psi | y) = \log p(y) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} \log F_t - \frac{1}{2} \sum_{t=1}^{n} \frac{\nu_t^2}{F_t}$$

$$\therefore \log L = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} \left( \log F_t + \frac{\nu_t^2}{F_t} \right)$$

onde $\nu_t$ e $F_t$ são calculados usando o FK.

De forma alternativa, o log da verossimilhança pode ser obtido usando a representação em $y$ do modelo de nível local

$y \sim N(a_1 \mathbb{1}, \Omega)$ → Normal Multivariado (com n observações)

$$p(y) = \frac{1}{(2\pi)^{n/2} |\Omega|^{1/2}} \exp\left\{ -\frac{1}{2} \left[ (y - a_1 \mathbb{1})' \Omega^{-1} (y - a_1 \mathbb{1}) \right] \right\}$$

$$L(\psi) = \log p(y) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Omega| - \frac{1}{2}\left[(y-a_1 1)'\Omega^{-1}(y-a_1 1)\right]$$

Podemos escrever:

dúvida

$\Omega = C F C'$ onde $|C| = 1$ e $CC' = I$

$(C' = C^{-1})$

$\downarrow$ variâncias de cada t na diag.

C: cte do modelo de nível local.

$$F = \begin{pmatrix} F_1 & & 0 \\ & F_2 \ddots & \\ \boxed{0} & & F_n \end{pmatrix}$$

cov = 0

$$\Omega^{-1} = (CFC')^{-1} = (C')^{-1}F^{-1}C^{-1} = C'F^{-1}C$$

$\Rightarrow$ i) $\log|\Omega| = \log|CFC'| = \log\left[|C||F||C|\right]$

$$= \log|F| = \log\prod_{t=1}^{n} F_t = \sum_{t=1}^{n}\log F_t$$

ii) $(y-a_1 1)'\Omega^{-1}(y-a_1 1)$ onde $v = C(y-a_1 1)$

$$= (y-a_1 1)'C'F^{-1}C(y-a_1 1) =$$

$$= \left[C(y-a_1 1)\right]'F^{-1}\left[C(y-a_1 1)\right] = v'F^{-1}v =$$

$$= (v_1 \dots v_n)\begin{pmatrix} F_1^{-1} & & 0 \\ & \ddots & \\ 0 & & F_n^{-1}\end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n\end{pmatrix} =$$

$$= \sum_{t=1}^{n}\frac{v_t^2}{F_t}$$

(e encontramos
a mesma expressão
para log-verossimilhança)

→ Inicializações Difusa e MV

$$\log L_D \triangleq \lim_{P_1 \to \infty}\left(\log L + \frac{1}{2}\log P_1\right)$$

da cond. inicial

Dúvida: $L + P_1^{-1/2}$ ?

$$\triangleq -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(\log F_t + v_t^2/F_t\right) + \frac{1}{2}\log P_1 =$$

não depende de $\sigma_\varepsilon^2$ ou $\sigma_\eta^2$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}\left(\log F_1 + \frac{v_1^2}{F_1}\right) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + \frac{v_t^2}{F}\right) + \frac{1}{2}\log P_1 =$$

$$= -\frac{1}{2}\left(\log\frac{F_1}{P_1} + \frac{v_1^2}{F_1}\right) - \frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + \frac{v_t^2}{F_t}\right)$$

Mas temos que:

i) $\frac{F_1}{P_1} = \frac{P_1+\sigma_\varepsilon^2}{P_1} = 1 + \frac{\sigma_\varepsilon^2}{P_1}$

$\lim_{P_1\to\infty} \frac{F_1}{P_1} = 1$ . $\lim_{P_1\to\infty} \log\left(\frac{F_1}{P_1}\right) = 0$

ii) $\frac{v_1^2}{F_1} = \frac{(y_1-a_1)^2}{P_1+\sigma_\varepsilon^2}$

$\lim_{P_1\to\infty} \frac{(y_1-a_1)^2}{P_1+\sigma_\varepsilon^2} = 0$

Logo:

$$\log L_D = -\frac{1}{2}\lim_{P_1\to\infty}\underbrace{\left(\log\frac{F_1}{P_1} + \frac{v_1^2}{F_1}\right)}_{0} - \frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

$$\log L_D = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

obs: Na prática, reparametrizamos $\sigma_\varepsilon^2 = e^{\psi_\varepsilon}$, $\sigma_\eta^2 = e^{\psi_\eta}$, $\psi_\varepsilon, \psi_\eta \in \mathbb{R}$

→ Concentração de MV

concentração = reparametrização do modelo em EE para reduzir a dimensionalidade da busca numérica

$$y_t = \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$
$$\alpha_{t+1} = \alpha_t + \eta_t, \quad \eta_t \sim N(0, q\sigma_\varepsilon^2)$$

$(\sigma_\varepsilon^2, \sigma_\eta^2) \to (\sigma_\varepsilon^2, q)$, $q = \sigma_\eta^2/\sigma_\varepsilon^2$ · razão sinal ruído

Defina $P_t^* = \frac{P_t}{\sigma_\varepsilon^2}$, $F_t^* = \frac{F_t}{\sigma_\varepsilon^2}$

Para esta nova parametrização, o FK difuso será:

$\rightarrow v_t = y_t - a_t$

$\rightarrow a_{t+1} = a_t + k_t v_t$

Logo: $\cdot\; k_t^* = \dfrac{P_t^*}{F_t^*} = \dfrac{P_t/\sigma_\varepsilon^2}{F_t/\sigma_\varepsilon^2} = \dfrac{P_t}{F_t} \therefore \boxed{k_t^* = k_t}$

$\cdot\; F_t = P_t + \sigma_\varepsilon^2$

$\therefore \dfrac{F_t}{\sigma_\varepsilon^2} = P_t + 1 \therefore \boxed{F_t^* = P_t^* + 1}$ (depende apenas de q)

$\rightarrow P_{t+1} = P_t(1 - k_t) + \sigma_\eta^2$

$\therefore \dfrac{P_{t+1}}{\sigma_\varepsilon^2} = \dfrac{P_t}{\sigma_\varepsilon^2}(1 - k_t) + \dfrac{\sigma_\eta^2}{\sigma_\varepsilon^2}$

$\therefore P_{t+1}^* = P_t^*(1 - k_t) + q$

Inicialização: $\begin{cases} a_2 = y \\ P_2^* = 1 + q \end{cases}$ ? Dúvida

Com a nova parametrização, temos p/ o $\log L_D$:

$$\log L_D(\psi) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right) \quad \Big\downarrow F_t = F_t^*\sigma_\varepsilon^2$$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\Big(\underbrace{\log F_t^*\sigma_\varepsilon^2}_{\log F_t^* + \log\sigma_\varepsilon^2} + \frac{v_t^2}{F_t^*\sigma_\varepsilon^2}\Big)$$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}(n-1)\log\sigma_\varepsilon^2 - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t^* + \frac{v_t^2}{F_t^*\sigma_\varepsilon^2}\right)$$

$\rightarrow$ Observe que:

(i) $F_t^*$ e $v_t$ não envolvem $\sigma_\varepsilon^2$

(ii) Maximizando para $\sigma_\varepsilon^2$:

$$\frac{\partial \log L_D}{\partial \sigma_\varepsilon^2} = 0 \quad \therefore \quad -\frac{1}{2}(n-1)\frac{1}{\sigma_\varepsilon^2} + \frac{1}{(\sigma_\varepsilon^2)^2}\cdot\frac{1}{2}\sum_{t=2}^{n}\frac{\nu_t^2}{F_t^*} = 0$$

$$\therefore \quad \frac{1}{\sigma_\varepsilon^2}\cdot\frac{1}{2}\sum_{t=2}^{n}\frac{\nu_t^2}{F_t^*} = \frac{1}{2}(n-1)$$

$$\therefore \quad \boxed{\hat{\sigma}_\varepsilon^2 = \left(\frac{1}{n-1}\right)\sum_{t=2}^{n}\frac{\hat{\nu}_t^2}{F_t^*}}$$

Finalmente, substituindo $\sigma_\varepsilon^2$ por $\hat{\sigma}_\varepsilon^2$ em $\log L_d(\Psi)$ resulta na log verossimilhança concentrada e difusa, dada por:

$$\log L_{dc}(\Psi) = -\frac{n}{2}\log 2\pi - \frac{1}{2}(n-1)\log \hat{\sigma}_\varepsilon(q) - \frac{1}{2}\sum_{t=2}^{n}\log F_t^*(q)$$

Assim, na prática, a maximização é realizada apenas com respeito a $q$, ou seja, a busca dimensional é realizada com uma dimensão em $\mathbb{R}^+$ pois $q \in (0,\infty)$

→ Steady State (FK estacionário)

Como $P_{t+1}$ não depende de $y$, eventualmente: $\lim_{t\to\infty} P_{t+1} = \bar{P}$

$$P_{t+1} = P_t(1-K_t) + \sigma_\eta^2 \quad ; \quad K_t = \frac{P_t}{F_t} = \frac{P_t}{P_t + \sigma_\varepsilon^2}$$

$$\lim_{t\to\infty} \quad \bar{P} = \bar{P}(1-\bar{K}) + \sigma_\eta^2 \Rightarrow \text{Equação de Ricatti onde } \bar{K} = \frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}$$

$$\bar{P} = \bar{P}\left(1 - \frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}\right) + \sigma_\eta^2$$

$$\bar{P} = \frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}(\cancel{\bar{P}} + \sigma_\varepsilon^2 - \cancel{\bar{P}}) + \sigma_\eta^2$$

$$\therefore \quad \bar{P} = \frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}\sigma_\varepsilon^2 + \sigma_\eta^2 \quad \therefore \quad \bar{P}^2 + \cancel{\bar{P}\sigma_\varepsilon^2} = \cancel{\bar{P}\sigma_\varepsilon^2} + \sigma_\eta^2\bar{P} + \sigma_\eta^2\sigma_\varepsilon^2$$

$$\bar{P}^2 - \sigma_\eta^2\bar{P} - \sigma_\eta^2\sigma_\varepsilon^2 = 0$$

$$\therefore \quad \frac{\bar{P}^2}{(\sigma_\varepsilon^2)^2} - \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}\frac{P}{\sigma_\varepsilon^2} - \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} = 0$$

sejam $x = \frac{\bar{P}}{}$ ; $h = \frac{\sigma_\eta^2}{} \cdot q$

$$\Rightarrow \quad x^2 - qx - q = 0$$

Soluções: $x = \dfrac{q \pm \sqrt{q^2+4q}}{2}$ $\Rightarrow$ só faz sentido solução $>0$ pois $x$ é $>0$.

$$x = \frac{q + \sqrt{q^2+4q}}{2}$$

$$\bar{P} = \sigma_\varepsilon^2 x = \sigma_\varepsilon^2 \left( \frac{q + \sqrt{q^2+4q}}{2} \right)$$

aula 16/04.

## Inicialização do FK

P/ estimar MEE via FK, será necessário especificar cond inicial

$$\alpha_1 \sim N(\underset{p\times 1}{a_1}, \underset{p\times p}{P_1})$$

P/ processos estacionários ⇒ cond. inicial perde importância em $t \to \infty$

" não-estacionários ⇒ não é verdade.

A forma de especificar a condição inicial irá depender da naturez estacionária da componente do vetor de estado.

Exemplo:

$$\begin{cases} y_t = \mu_t + \psi_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \eta_t \\ \psi_{t+1} = \phi\psi_t + \xi_t \end{cases} \Rightarrow \quad y_t = \begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} \alpha_{1,t} \\ \alpha_{2,t} \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} \alpha_{1,t+1} \\ \alpha_{2,t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \phi \end{pmatrix}\begin{pmatrix} \alpha_{1,t} \\ \alpha_{2,t} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix}$$

⇒ $\alpha_{1,t} = \mu_t$ é não-estacionário

$\alpha_{2,t} = \psi_t$ é estacionário

→ Quando a componente for estacionária, então é natural que a distribuição inicial seja incondicional deste elemento no vetor de estado.

(É a dist. de LP do processo. No LP ⇒ processo converge p/ esta condição,

No exemplo:

Dist. incondicional AR(1)

→ $E[\psi_{t+1}] = \phi E[\psi_t)$

como $E[\psi_{t+1}] = E[\psi_t] = \mu$ ⇒ $\mu = \phi\mu \therefore \mu = 0$

(queremos média ct).

→ $\text{var}[\psi_{t+1}] = \phi^2 \text{var}(\psi_t) + \sigma_\eta^2$

Como $\text{var}[\psi_{t+1}] = \text{var}(\psi_t) = \sigma_\psi^2 \Rightarrow \sigma_\psi^2 = \phi^2\sigma_\psi^2 + \sigma_\eta^2$

(queremos var. ct.)

$$\therefore \quad \sigma_\psi^2 = \frac{\sigma_\eta^2}{1-\phi^2}$$

obs: P/ autocovar, temos: $E[\psi_t \psi_{t+h}] = \phi^h \cdot \frac{\sigma_\eta^2}{(1-\phi^2)}$

Assim, a dist. incondicional de $\psi_{t+1} \sim N(0, \sigma_\eta^2/1-\phi^2)$

No vetor $\underset{\sim}{\alpha}_1 = (\alpha_{1,1} \;\; \alpha_{2,1}) \Rightarrow a_{2,1} = 0$

$$P_{2,1} = \frac{\sigma_\eta^2}{(1-\phi^2)}$$

atenção! Cond. inicial deve ser função de parâmetros desconhecidos ⇒ estimados por MV.

De forma geral, se todas componentes de $\alpha_t$ forem estacionárias: (ou subvetor de $\alpha_t$), a dist. incondicional será dada por:

$$\alpha_{t+1} = T\alpha_t + R\eta_t \qquad \eta_t \sim N(0, Q)$$

– P/ ser estacionário, sistema tem que ser invariante no tempo.

Supondo $|\lambda_i(T)| < 1 \quad i = 1, \dots, p$

$$\Rightarrow E[\alpha_{t+1}] = TE[\alpha_t] \therefore E[\alpha_t] = 0.$$

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

$$\therefore vec[var(\alpha_{t+1})] = vec[T var(\alpha_t) T'] + vec[RQR']$$

$$= (T \otimes T)\, vec[var(\alpha_t)] + R \otimes R'\, vec Q$$

$$(I - T\otimes T)\, vec[var(\alpha_t)] = [R \otimes R'\, vec Q]$$

$$vec[var(\alpha_t)] = [I - T\otimes T]^{-1} [R\otimes R'\, vec(Q)]$$

Dúvida
Como calcula na prática

$$\Rightarrow \underset{p\times 1}{a_1} \sim N(\underset{p\times 1}{0}, \underset{p\times p}{var(\alpha_t)})$$

(T e Q podem envolver parâmetros desconhecidos, a serem estimados por MV).

→ Para as componentes não estacionárias, pode-se adotar uma distribuição difusa (ou prior difusa). Na ausência de informação inicial, é razoável supor que todos os valores de $\alpha_1$ são equiprováveis

1) Big Kappa.

Densidade seria como uma uniforme nos reais.

$$\alpha_{1,1} \sim N(a_{1,1}, p_{1,1}) \text{ onde } a_{1,1} \text{ é arbitrário (geralmente } = 0)$$

$$p_{1,1} = k, \; k \to \infty$$

(Big Kappa)

Como $\int_{-\infty}^{\infty} f(\alpha_{1,1})\, d\alpha_{1,1} \to \infty$ (quando $k \to \infty$) ⇒ diz-se que é impróp

Isto terá efeito apenas localizado no FK.

Ele será bem definido para $t = d+1, d+2, \dots$, onde $d$ é a dimensão do sub-vetor não estacionário de $\alpha_t$.

Exemplo: Nível local

$$y_t = \mu_t + \varepsilon_t \qquad \mu_1 = \alpha_1 \sim N(a_1, P_1)$$

$$\mu_{t+1} = \mu_t + \eta_t \qquad a_1 = 0$$

$$P_1 = k,\ k \to \infty.$$

$Z = 1$
$T = 1$
$R = 1$
$K = TMF^{-1}$
$F_t = ZPZ' + H$
$= P_t + \sigma_\varepsilon^2$
$M = P_t Z_t' = P_t$
$\Rightarrow K = \frac{1 P_t}{P_t + \sigma_\varepsilon^2}$

$T.P.L' =$
$1.P_t(1-K) + \sigma_\eta^2$
$\searrow \frac{\sigma_\varepsilon^2}{P_t +}$

Equações de FK p/ este modelo:

$$a_{t+1} = a_t + k_t v_t \qquad P_{t+1} = (1-k_t)P_t + \sigma_\eta^2$$

$$v_t = y_t - a_t$$

$$k_t = \frac{P_t}{F_t} = \frac{P_t}{P_t + \sigma_\varepsilon^2}$$

- Para $t = 1$:

$\Rightarrow$ Fazer p/ tendência local e ver que a partir de $t = 3$ teremos priori bem definida

$$a_2 = a_1 + K_1 v_1, \quad v_1 = y_1 - a_1$$

$$K_1 = \frac{P_1}{P_1 + \sigma_\varepsilon^2}$$

$$\therefore a_2 = a_1 + \frac{P_1}{P_1 + \sigma_\varepsilon^2}(y_1 - a_1)$$

$$\text{e} \quad P_2 = (1-K_1)P_1 + \sigma_\eta^2 = \underbrace{\left(\frac{\sigma_\varepsilon^2}{P_1 + \sigma_\varepsilon^2}\right)}_{(1-K)} P_1 + \sigma_\eta^2 \qquad \left\{ \begin{array}{l} P_1 = k \\ \boxed{a_1 = 0} \end{array} \right.$$

Quando $k \to \infty$:

$$\begin{cases} a_2 = a_1 + (y_1 - a_1) = y_1 \\ P_2 = \sigma_\varepsilon^2 + \sigma_\eta^2 \end{cases} \qquad \text{pois } \frac{P_1}{P_1 + \sigma_\varepsilon^2} \to 1 \text{ quando } P_1 = k \to \infty.$$

Em $t = 1$:

$$\begin{cases} a_2 = y_1 \\ P_2 = \sigma_\varepsilon^2 + \sigma_\eta^2 \end{cases} \Rightarrow \text{densidade } p(\alpha_{t+1} | Y_t) \text{ bem definida, própria.}$$

P/ $t = 2, 3, \dots$ não há problema p/ calcular FK

É como esquecer 1º passo e começar FK p/ $t = 2$ (*)

Big Kappa ⇒ pode ñ ser estável computacionalmente
erros de arredondamento.

2) Inicialização Exata

outra forma:
Realizar expansão de produtos de matrizes de FK em termos de $K^{-1}$ e reter apenas os primeiros 2 ou 3 termos.
fazer $K \to \infty$ e obter termo dominante.

Expressão Geral p/ $\alpha_1$:

$$\alpha_1 = \underset{\substack{m\times 1\\ \text{conhecido}}}{a} + \underset{m\times q}{A}\,\underset{q\times 1}{\delta} + \underset{m\times(m-q)}{R_0}\,\overset{(m-q)\times 1}{\eta_0} \qquad \eta_0 \sim N(0, Q_0)$$

$A, R_0$: matrizes de seleção

$R_0\eta_0$: cond. inicial de componente estacionária

$\delta$: desconhecido.

(Dos m componentes $\begin{cases} \to q \text{ ñ estac.} \\ \to m-q \text{ estac.} \end{cases}$)

obs: se todos os elementos de $\alpha_t$ são estacionários $E(\alpha_1)$ e $var(\alpha_1)$ podem ser obtidos através dos parâmetros do modelo.

Exemplo: (modelo ñ estacionário)

$$\begin{cases} y_t = \mu_t + \rho_t + \varepsilon_t \\ \mu_{t+1} = \mu_t + \nu_t + \xi_t \\ \nu_{t+1} = \nu_t + \zeta_t \\ \rho_{t+1} = \phi\rho_t + \tau_t \end{cases}$$

> tendência linear local (para $\mu_{t+1}$ e $\nu_{t+1}$)

$\rho_{t+1} = \phi\rho_t + \tau_t$ (resíduo AR(1) pois nem sempre componentes são suficientes p/ explicar variações da série).

componente estacionária de ordem 1

$$y_t = (1 \;\; 1 \;\; 0)\begin{pmatrix} \rho_t \\ \mu_t \\ \nu_t \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} \rho_{t+1} \\ \mu_{t+1} \\ \nu_{t+1} \end{pmatrix} = \begin{pmatrix} \phi & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \rho_t \\ \mu_t \\ \nu_t \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \tau_t \\ \xi_t \\ \zeta_t \end{pmatrix}$$

13/05-0 8h.

Ⓐ

Adotar difusa $\alpha_1 \sim N(0, k)$ $k \to \infty$ em $t=1$

equivale a

$y_1$ p/ x0 tal que. $y_1 = \alpha_1 + \varepsilon_1 \Rightarrow \alpha_1 = y_1 - \varepsilon_1$

$E[\alpha_1) = y_1$ (p/ x0)

$\mathrm{var}(y_1) = \sigma_\varepsilon^2$

$\Rightarrow \alpha_1 \sim N(y_1, \sigma_\varepsilon^2)$

Como $\alpha_{t+1} = \alpha_t + \eta_t \Rightarrow \alpha_2 = \alpha_1 + \eta_1$

$E(\alpha_2) = E(\alpha_1) = y_1$

$\mathrm{var}[\alpha_2) = \mathrm{var}(\alpha_1) + \mathrm{var}[\eta_1)$

$= \sigma_\varepsilon^2 + \sigma_\eta^2$

Logo: se a partir de $t=2$, usarmos $\alpha_2 \sim N(y_1, \sigma_\varepsilon^2 + \sigma_\eta^2)$

teremos o mesmo resultado do FK difuso $\alpha_1 \sim N(0, k)$

$\Rightarrow$ Numa situação mais geral, onde $\alpha_t$ possui várias componentes não estacionárias, a cond. inicial difusa será dada por:

$$\alpha_1 \sim N\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, k\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}\right] \text{ ou } \alpha_1 \sim N(0, kI_p) \quad k \to \infty$$

$\Rightarrow$ fazer exercício p/ tendência linear local e mostrar que p/ $t = 3, 4 \ldots$ as eqs. de FK serão bem definidas.

$\mu_t$ e $\nu_t$ não-estacionários

$\rho_t$ estacionário $(|\phi|<1) \to E(\rho_t)=0$, $\text{var}(\rho_t) = \dfrac{\sigma_\tau^2}{(1-\phi^2)}$ $\Big\}$ $\rho_1 \sim N(0, \sigma_\tau^2/(1-\phi^2))$

$$\Rightarrow \quad \alpha_1 = a + A\delta + R_0\eta_0 \;, \quad \eta_0 \sim N(0, Q_0) \quad (I)$$

$$\begin{pmatrix}\rho_1\\ \mu_1\\ \nu_1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} + \begin{pmatrix}0&0\\1&0\\0&1\end{pmatrix}\underset{\sim}{\delta} + \begin{pmatrix}1\\0\\0\end{pmatrix}\eta_0 \quad ; \quad q_0 = \frac{\sigma_\tau^2}{(1-\phi^2)}$$

matriz de selecção: gera comp. estacionária

caracteriza componentes não estacionárias

caracteriza componentes estacionárias

$\underset{\sim}{\delta}$ é v.a. com $\underset{\sim}{\delta} = \begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix}, kI_2\right]$ $k \to \infty$.

- Para $\alpha_1$: $\alpha_1 \sim N(a_1, P_1)$ da eq. (I).

onde $a_1 = E[\alpha_1] = a$ (=0 p/ modelos não estacionários)

$$P_1 = E[(\alpha_1 - a)(\alpha_1 - a)'] = A\,\underbrace{\text{var}(\delta)}_{kI_2}\,A' + R_0Q_0R_0'$$

$$\therefore \quad P_1 = kAA' + R_0Q_0R'$$

$$= kP_\infty + P_* \quad \text{onde faz-se } k \to \infty \text{ p/ } t \text{ apropriado}$$

associada à parte ñ estacionária

associada à parte estacionária

Se não houver componentes estacionárias: $P_* = 0$

$$\Rightarrow \quad \boxed{\begin{aligned} P_\infty &= AA' \\ P_* &= R_0Q_0R_0' \end{aligned}}$$

Pode-se mostrar que (por expansão de Taylor)

$\to$ $P_t = kP_{\infty,t} + P_{*,t} + O(k-1)$ $\quad t = 2, 3, \ldots, n$

$\to$ $P_{\infty,t} \equiv 0$ para algum $t = d$ e assim para $t = d+1, d+2, \ldots, n$ utiliza-se FK padrão.

Mostra-se que para algum $t = d$ $(d << n)$ : $P_{\infty,t} = 0$

Assim p/ $t = 1, 2, \dots, d$ : usa FK inicialização exata

$t = d+1, \dots, n$ : usa FK padrão.

FK

Equações p/ FK exato inicial

para $t = 1, 2, 3, \dots \underset{\uparrow}{d}$

1º momento em que $P_{\infty,t}$ zera

Equações do FK padrão

para $t = d+1, d+2, \dots, n$ usando $P_{d+1} = P_{*,d+1}$ ↓

ou seja, com dados de $P_{*,d}$, $P_{\infty,d}$

e $a_{d+1} = \hat{a}^{(0)}_{d+1}$

suavização

~~Algoritmo de suavização será afetado pelas cond. iniciais exatas~~

(equações de recursão dadas)

⇓

p/ $t = d, d-1, \dots, 1$

Para $t = d+1, \dots, n$ ⇒ usam-se equações padrão de suavização

Exemplo : Tendência linear local.

$$y_t = (1 \quad 0)\begin{pmatrix}\mu_t \\ \beta_t\end{pmatrix} + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

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

(contas)

$P_{\infty,3} = 0$ calculado em $t=2$ como previsão $t+1=3$.

Dúvida ⇒ $d = 3$ (onde zera $P_{\infty}$)
ou $\boxed{d = 2}$ (nº de componentes não estacionárias)

aula 20/04

## Estimação de Parâmetros

- MEE possui um conjunto de parâmetros fixos e desconhecidos (elementos reais das matrizes $Z_t, T_t, H_t$ e $Q_t$)

  $\Rightarrow$ Hiperparâmetros (vetor $\psi$)

- Estimação de $\psi$ usando princípio da verossimilhança

$$L(\psi) = p(y_1 y_2 \ldots y_n ; \psi) =$$

$$= p(y_1, \psi) \prod_{t=2}^{n} p(y_t | \underset{\sim}{Y}_{t-1} ; \psi) \quad \leftarrow \text{ver em demos anteriores.}$$

$$= \prod_{t=1}^{n} p(y_t | \underset{\sim}{Y}_{t-1} ; \psi) \quad \text{com } p(y_1 | Y_0) = p(y_1)$$

- Da eq. das observações:

$$y_t = Z_t \alpha_t + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, H_t)$$

$$\Rightarrow E[y_t | \underset{\sim}{Y}_{t-1}] = Z_t a_t$$

$$\text{var}[y_t | \underset{\sim}{Y}_{t-1}] = F_t = Z_t P_t Z_t' + H_t$$

$$v_t = y_t - Z_t a_t$$

Como $y_t | \underset{\sim}{Y}_{t-1}$ é normal:

$$p(y_t | \underset{\sim}{Y}_{t-1}) = \frac{1}{(2\pi)^{p/2} |F_t|^{1/2}} \exp\left\{ -\frac{1}{2} \underbrace{(y_t - Z_t a_t)'}_{v_t'} F_t^{-1} \underbrace{(y_t - Z_t a_t)}_{v_t} \right\}$$

$\downarrow$

caso geral é $y_t \sim p \times 1$ : multivariado.

contas

$$L(\psi) = \prod_{t=1}^{n} \frac{1}{(2\pi)^{p/2} |F_t|^{1/2}} \exp\left\{ -\frac{1}{2} v_t' F_t^{-1} v_t \right\}$$

$$= \left( \frac{1}{(2\pi)^{np/2}} \right) \left( \prod_{t=1}^{n} |F_t|^{-1/2} \right) \left( \exp\left\{ -\frac{1}{2} \sum_{t=1}^{n} v_t' F_t^{-1} v_t \right\} \right)$$

$$\ell(\psi) = \log(\psi) \therefore \ell(\psi) = -\frac{np}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \log |F_t| - \frac{1}{2} \sum_{t=1}^{n} v_t' F_t^{-1} v_t$$

$$\therefore \boxed{\ell(\psi) = -\frac{np}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left(\log|F_t| + v_t' F_t^{-1} v_t\right)}$$

↓

decomposição da verossimilhança em erro de previsão

Atenção!

Devemos estar atentos p/ como a inicialização será tratada no FK pois $v_t$ e $F_t$ (que aparecem em $\ell(\psi)$) são avaliadas sequencialmente pelo FK.

$$v_t = y_t - Z_t a_t$$
$$F_t = Z_t P_t Z_t' + H_t$$

Vejamos como cada tratamento de inicialização afeta $\ell(\psi)$:

a) Inicialização Prior Difusa. (big kappa)

- Enquanto "efeito" da prior difusa estiver atuando no FK
  $\Rightarrow$ $\ell(\psi)$ não incorpora os valores de $a_t$ e $P_t$ (via $F_t$ e $v_t$)

Apenas quando $t = d+1$ tal que $\alpha_t$ tem dist. própria é que começamos a computar $\ell(\psi)$. $\Rightarrow$ ou seja, quando $a_t, P_t$ passam a ter valores próprios

Exemplo:

Nível local

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim N(0,\sigma^2_\varepsilon) \qquad t = 1,2,\dots,n$$
$$\mu_{t+1} = \mu_t + \eta_t \qquad \eta_t \sim N(0,\sigma^2_\eta)$$

Escolhendo $\alpha_1 \sim N(0,k)$ $k \to \infty$ p/ $t = 1$ $\Rightarrow$ $a_t$ e $P_t$ ainda estarão sob efeito de $k$

Para $t = 2,3,\dots,n$ não teremos mais problemas.

$$\Rightarrow \overset{bk}{\ell}(\psi) = -\frac{(n-1)}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

- Repare: p/ modelos com $q$ componentes não estacionárias no vetor de estado

$$\Rightarrow \overset{bk}{\ell}(\psi) = -\underline{(n-q)}\,p.\log(2\pi) - \frac{1}{2}\sum_{t=q+1}^{n}\left[\log|F_t| + v_t' F_t^{-1} v_t\right]$$

obs: $a_t$ e $P_t$ saõ atualizados via FK de $t=1$ a $q$, mas naõ entram em $l(\psi)$ (via $v_t$ e $F_t$). Apenas para $t=q+1,\dots,n$ é que passam a entrar no cálculo de $l(\psi)$

b) Inicializaçaõ Exata.

Pode-se mostrar que nesse caso $l(\psi)$ será:

usa FK difuso

usa FK padrão

$$l_d(\psi) = \log L_d(\psi) =$$

$$= -\left(\frac{np}{2}\right)\log(2\pi) - \frac{1}{2}\sum_{t=1}^{d} w_t - \frac{1}{2}\sum_{t=d+1}^{n}\left[\log|F_t| + v_t' F_t^{-1} v_t\right]$$

$$\text{onde } w_t = \begin{cases} \log|F_{\infty,t}| & \text{se } F_{\infty,t} \text{ é pos. definida} \\ \log|F_{*,t}| + v_t^{(0)'} F_{*,t}^{-1} v_t^{(0)} & \text{se } F_{\infty,t} = 0. \end{cases}$$

Prova:

Ponto de partida é definiçaõ da verossimilhança difusa

$$l_d(\psi) = \log L_d(\psi) =$$

$$= \lim_{k\to\infty}\left[\log L(\psi) + \frac{q}{2}\log k\right]$$

a contribuiçaõ em $l(\psi)$ de $\alpha_1$ é dada por $-\frac{q}{2}\log k$

$$\text{onde } L(\psi) = -\frac{np}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left(\log|F_t| + v_t' F_t^{-1} v_t\right)$$

. Desmembrar $L(\psi)$ de $t=1\dots d$ e $t=d+1,\dots,n$

. Entre $t=1\dots d$, usamos o FK difuso, considerando as 2 possibilidades $\begin{cases} F_{\infty,t} \text{ pos. definida} \\ F_{\infty,t} = 0. \end{cases}$

a) Quando $F_{\infty,t}$ é pos. def, mostra-se que p/ $t=1,2,\dots,d$

$$\Rightarrow \lim_{k\to\infty}\left[-\log|F_t| + p\log k\right) = -\log|F_{\infty,t}|$$

$$\lim_{k\to\infty} v_t' F_t^{-1} v_t = 0.$$

b) Quando

?

## Outros comentários sobre estimação de parâmetros

- Estimadores pelo princípio da verossimilhança.

  Apresentam propriedades desejáveis:

  a) Não viciados

  $$\lim_{n \to \infty} E[\hat{\psi}_n) = \psi$$

  b) consistentes

  $$\text{plim } \hat{\psi}_n = \psi$$

  c) Eficientes

  $$\text{var}(\hat{\psi}_n) \to \text{LCR}$$

  d) Dist. Assintótica Normal

  $$\hat{\psi} \overset{a}{\sim} N(\psi_1, I^{-1}(\hat{\psi}))$$

  } importante p/ termos info de IC e TH.

- Para modelo de nível local

→ Duas formas de obter $\ell(\psi)$:

(a) Pelo produto das dist de $y_t$ escritas em função de $a_t$ e $F_t$

$$p(y_t | Y_{t-1}) \sim N(a_t, F_t)$$

pois $E[y_t | \underset{\sim}{Y}_{t-1}] = Z a_t = a_t$

$\text{var}[y_t | \underset{\sim}{Y}_{t-1}) = Z P_t Z' + H = F_t$

$= P_t + \sigma_\varepsilon^2 = F_t$

e $\nu_t = y_t - a_t$

$$\therefore p(y_t | Y_{t-1}) = \frac{1}{(2\pi F_t)^{-1/2}} \exp\left\{ \frac{-1}{2} \frac{(y_t - a_t)^2}{F_t} \right\}$$

$$\therefore p(y) = \prod_{t=1}^{n} p(y_t | \underset{\sim}{Y}_{t-1}) = (2\pi F_t)^{-n/2} \exp\left\{ -\frac{1}{2} \sum \frac{(\overbrace{y_t - a_t}^{\nu_t})^2}{F_t} \right\} \cdot \prod_{t=1}^{n} F_t^{-1/2}$$

$$\Rightarrow \log p(y) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum \frac{\nu_t^2}{F_t} - \frac{1}{2} \sum F_t$$

$$\therefore \ell(\psi) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} (\log F_t + \nu_t^2 / F_t)$$

calculadas usando FK

(b) Pela representação em y do modelo de nível local

$y \sim N(a_1 \mathbb{1}, \Omega)$ : Normal multivariado

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$p(y) = \frac{1}{(2\pi)^{-n/2}} \cdot \frac{1}{|\Omega|^{1/2}} \exp\left\{ -\frac{1}{2} (y - a_1\mathbb{1})' \Omega^{-1} (y - a_1\mathbb{1}) \right\}$$

$$\ell(\psi) = \log p(y) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\Omega| - \frac{1}{2}\left[ (y - a_1\mathbb{1})' \Omega^{-1} (y_1 - a_1\mathbb{1}) \right]$$

dúvida: quem é C

$\circledast$ $\Omega = \boxed{C} F C'$ e $|C| = 1$ , $CC' = I$

$\Omega^{-1} = (CFC')^{-1}$ mas $C' = C^{-1}$

$= (C')^{-1} F^{-1} C^{-1}$

$= C' F^{-1} C$

$\therefore \log|\Omega| = \log|CFC'| = \log(|C||F||C|)$

$= \log|F| = \Sigma \log F_t$

$$F = \begin{pmatrix} F_1 & & \\ & \ddots & \\ & & F_n \end{pmatrix} \Rightarrow |F| = \prod_{t=1}^{n} F_t$$

$\log|F| = \Sigma \log F_t$

$\circledast$ $(y - a_1\mathbb{1})' \Omega^{-1} (y - a_1\mathbb{1})$ onde $v = C(y - a_1\mathbb{1})$

$$\underbrace{(y - a_1\mathbb{1})' C'}_{v'} F^{-1} C (y - a_1\mathbb{1})$$

$v' F^{-1} v = v^2 / F_t$

$$v' F^{-1} v = (v_1 \; v_2 \cdots v_n) \begin{pmatrix} F_1^{-1} & & 0 \\ & \ddots & \\ 0 & & F_n^{-1} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_{t=1}^{n} v_t^2 / F_t$$

→ Inicialização Difusa e MV

$$\log L_D \triangleq \lim_{P_1 \to \infty} \left( \log L + \frac{1}{2} \log P_1 \right)$$ (obs: $P_1 = \kappa$ : Big kappa)

$$= -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} \left( \log F_t + v_t^2 / F_t \right) + \frac{1}{2} \log P_1 =$$

↓ não depende de $\sigma_\varepsilon^2$ ou $\sigma_\eta^2$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}\left(\log F_1 + v_1^2/F_1\right) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right) + \frac{1}{2}\log P_1$$

$$\Rightarrow \log L_D = -\frac{n}{2}\log 2\pi - \frac{1}{2}\left(\log F_1 - \log P_1 + v_1^2/F_1\right) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}\left(\log\frac{F_1}{P_1} + \frac{v_1^2}{F_1}\right) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}\left(\log\frac{P_1+\sigma_\varepsilon^2}{P_1} + \frac{v_1^2}{P_1+\sigma_\varepsilon^2}\right) - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

Tomando $\lim_{P_1 \to \infty}$:

$$\lim_{P_1 \to \infty} \log\frac{P_1+\sigma_\varepsilon^2}{P_1} = 0$$

$$\lim_{P_1 \to \infty} \frac{v_1^2}{P_1+\sigma_\varepsilon^2} = 0$$

$$\Rightarrow \left|\log L_D = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)\right|$$

Na prática, reparametrizaremos: $\sigma_\varepsilon^2 = e^{\psi_\varepsilon}$ onde $\psi_\varepsilon$ e $\psi_\eta \in \mathbb{R}$
$\sigma_\eta^2 = e^{\psi_\eta}$

→ Concentração de MV

⇒ Reparametrização do modelo em EE para reduzir dimensionalidade da busca

Parâmetros $\sigma_\varepsilon^2, \sigma_\eta^2 \to (\sigma_\varepsilon^2, q)$ onde $q = \sigma_\eta^2/\sigma_\varepsilon^2$ : razão sinal ruído.

Definimos: $P_t^* = \frac{P_t}{\sigma_\varepsilon^2}$ e $F_t^* = \frac{F_t}{\sigma_\varepsilon^2}$

⇒ Expressões de FK na nova parametrização

$v_t = y_t - a_t$
$a_{t+1} = a_t + K_t v_t$
$P_{t+1} = P_t(1-K_t) + \sigma_\eta^2$

- $K_t = \frac{P_t}{F_t} = \frac{P_t/\sigma_\varepsilon^2}{F_t/\sigma_\varepsilon^2} = \frac{P_t^*}{F_t^*} = K_t^*$
- $F_t = P_t + \sigma_\varepsilon^2 \therefore \frac{F_t}{\sigma_\varepsilon^2} = \frac{P_t}{\sigma_\varepsilon^2} + 1$

depende só de q?

$\therefore F_t^* = P_t^* + 1.$

$$\frac{P_{t+1}}{\sigma_\varepsilon^2} = \frac{P_t}{\sigma_\varepsilon^2}(1-K_t) + \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \quad \therefore \quad P_{t+1}^* = P_t^*(1-K_t)+q$$

Inicializaçoes:

Por Big kappa. quando $\left.\begin{matrix} a_1 = 0 \\ P_1^* = k \end{matrix}\right\} \begin{matrix} a_2 = y_1 \\ P_2^* = 1+q \end{matrix}$

obs ⊛ $P_2^* = P_1^*\left(1 - \frac{P_1^*}{F_1^*}\right) + q$

$$= P_1^*\left(1 - \frac{P_1^*}{P_1^*+1}\right) + q$$

$$= P_1^*\left(\frac{1}{P_1^*+1}\right) + q$$

Quando $P_1^* = k \to \infty \Rightarrow P_2^* = 1+q$

⊛ $a_2 = a_1 + (y_1 - a_1)$

Quando $a_1 = 0 \Rightarrow a_2 = y_1$

} cond. inicial.

Logo:

$$l_d(\psi) = \log L_d(\psi) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t + v_t^2/F_t\right)$$

Substituindo $F_t$ por $F_t^* \times \sigma_\varepsilon^2$:

$$l_d(\psi) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t^*\sigma_\varepsilon^2 + \frac{v_t^2}{F_t^*\sigma_\varepsilon^2}\right)$$

$$= -\frac{n}{2}\log 2\pi - \frac{1}{2}(n-1)\log\sigma_\varepsilon^2 - \frac{1}{2}\sum_{t=2}^{n}\left(\log F_t^* + \frac{v_t^2}{F_t^*\sigma_\varepsilon^2}\right)$$

Observar que:

(i) $F_t^*$ e $v_t$ nas envolvem $\sigma_\varepsilon^2$

(ii) $\frac{\partial \log L_d}{\partial \sigma_\varepsilon^2} = 0 \Rightarrow -\frac{1}{2}(n-1)\frac{1}{\sigma_\varepsilon^2} + \frac{1}{(\sigma_\varepsilon^2)^2}\cdot\frac{1}{2}\sum\frac{v_t^2}{F_t^*} = 0$

$$\therefore \quad \frac{1}{\sigma_\varepsilon^2} \cdot \frac{1}{2} \sum v_t^2 / F_t^* = \frac{1}{2}(n-1)$$

$$\therefore \quad \hat{\sigma}_\varepsilon^2 = \left( \frac{1}{(n-1)} \right) \sum_{t=2}^{n} \hat{v}_t^2 / F_t^*$$

Substituindo $\hat{\sigma}_\varepsilon^2$ em $\sigma_\varepsilon^2$ em $\log L_d(\psi)$, temos:

$$\log L_{dc}(\psi) = -\frac{n}{2} \log 2\pi - \frac{1}{2}(n-1) \log \hat{\sigma}_\varepsilon^2(q) - \frac{1}{2} \sum_{t=2}^{n} \log F_t^*(q)$$

↓

Maximização será realizada apenas com respeito a q.

Aula 20/04-27/04

## Otimização Numérica

- $\Psi \in \mathbb{R}^v$
- função objetivo: $\ell(\Psi) = \log L(\Psi)$, $\Psi \in$ interior do espaço paramétrico
- estimador MV: $\hat{\Psi} = \underset{\Psi \in \mathbb{R}^v}{\arg\max}\ \ell(\Psi)$

Condições p/ maximizar uma função:

1) CPO: 1ª ordem (necessária)

$$g(\Psi) = \frac{\partial \ell}{\partial \Psi} = \left( \frac{\partial \ell}{\partial \Psi_1}, \frac{\partial \ell}{\partial \Psi_2}, \ldots, \frac{\partial \ell}{\partial \Psi_v} \right)^T = \underset{\sim}{0}$$ (gradiente ou vetor score)

2) CSO: 2ª ordem

$$H(\Psi) = - \frac{\partial^2 \ell}{\partial \Psi \partial \Psi'} \Big|_{\Psi = \hat{\Psi}} > 0.$$

onde $$H(\hat{\Psi}) = - \begin{pmatrix} \partial^2 \ell / \partial \psi_1^2 & \partial^2 \ell / \partial \psi_1 \partial \psi_2 & \cdots & \partial^2 \ell / \partial \psi_1 \partial \psi_n \\ \partial^2 \ell / \partial \psi_1 \partial \psi_2 & \partial^2 \ell / \partial \psi_2^2 & \cdots & \partial^2 \ell / \partial \psi_2 \partial \psi_n \\ \vdots & & & \\ \partial^2 \ell / \partial \psi_1 \partial \psi_v & \cdots & & \partial^2 \ell / \partial \psi_v^2 \end{pmatrix}_{v \times v}$$ avaliada em $\Psi = \hat{\Psi}$ (sol. da CPO)

$\Rightarrow$ $z' H z > 0$ (ou seja, $H$ é matriz positiva definida)

Obs: como $H$ é a negativa do Hessiano $\Rightarrow$ positiva definid.
(e não negativa ")

Ex: P/ $v = 2$

$z = (z_1\ z_2)^T$

$$\Rightarrow z' H z = (z_1\ z_2) \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \Rightarrow h_{11} z_1^2 + (h_{12} z_1 z_2) \times 2 + h_{22} z_2^2 > 0.$$

$\Downarrow$
investigar

$\Rightarrow$ $h_{11}$ e $h_{22} > 0$ são
suficientes p/ garantir $H$ pos. definida.

Dificilmente teremos solução analítica.

Precisaremos de um método numérico p/ realizar a otimização

a) Métodos do Gradiente:

Algoritmos de otimização numérica definidos pela seguinte estrutura:

(i) cond. inicial $\Psi(0)$

(ii) forma iterativa que gera aproximações sucessivas $\hat{\Psi}^{(0)}, \hat{\Psi}^{(1)}, \dots, \hat{\Psi}^{(k)}$ para o ótimo $\hat{\Psi}$.

(iii) regra de parada que define o ótimo, ou seja, número máximo de iterações

obs: máx local (não global) pois $\ell(\Psi)$ ñ é geralmente côncava

b) Métodos Quasi-Newton

Expansão de Taylor de $\ell(\Psi)$ em torno de uma iteração arbitrária $\Psi^{(k)}$:
(de 2ª ordem)

$$\ell(\Psi) = \ell(\Psi^{(k)}) + \left(\frac{\partial \ell}{\partial \Psi^{(k)}}\right)'(\Psi - \Psi^{(k)}) + \frac{1}{2}(\Psi - \Psi^{(k)})' \frac{\partial^2 \ell}{\partial \Psi^{(k)} \partial \Psi^{(k)}} (\Psi - \Psi^{(k)})$$

$$= \ell(\Psi^{(k)}) + g(\Psi^{(k)})'(\Psi - \Psi^{(k)}) - \frac{1}{2}(\Psi - \Psi^{(k)}) H(\Psi^{(k)})(\Psi - \Psi^{(k)})$$

Por construção: queremos $\frac{\partial \ell}{\partial \Psi} = 0$ (CPO). Logo:

$$\frac{\partial \ell}{\partial \Psi} = g(\Psi) = g(\Psi^{(k)}) - H(\Psi^{(k)})(\Psi - \Psi^{(k)}) = 0$$

$$\vdots$$

Resolvendo p/ $\Psi$:

$$\Psi = \Psi^{(k)} + H^{-1}(\Psi^{(k)}) g(\Psi^{(k)})$$

$$\Rightarrow \boxed{\Psi^{(k+1)} = \Psi^{(k)} + \lambda H^{-1}(\Psi^{(k)}) g(\Psi^{(k)})}$$

$\lambda$: tamanho do passo.

obs. 1) Diversos algoritmos existentes: BHHH, BFGS etc.

2) Como $\ell(\psi)$ não é côncava, isto pode trazer problemas na busca do máximo.

→ Critérios de Parada:

Normalmente é usado 1 ou mais critérios simultaneamente

(i) critério do gradiente (é o mais usado)

$$c_1 = \sum_{j=1}^{\nu} \frac{|g_j(\psi^{(k)})|}{\nu} < 10\varepsilon$$

↳ média dos módulos de $\frac{\partial \ell}{\partial \psi_i}^{(k)}$

(ii) critério da verossimilhança

$$c_2 = \left| \frac{\ell(\psi^{(k)}) - \ell(\psi^{(k+1)})}{\ell(\psi^{(k)})} \right| < \varepsilon$$

↳ variação da função de verossim no máximo tem que ser muito pequena.

(iii) critério de Parâmetro

$$c_3 = \frac{1}{\nu} \sum_{j=1}^{\nu} \left| \frac{\psi_j^{(k+1)} - \psi_j^{(k)}}{\psi_j^{(k)}} \right| < 100\varepsilon$$

↳ média do valor do parâmetro não muda muito (equivale a (ii))

→ Transformação de parâmetros

Alguns parâmetros do MEE podem ter restrição de valores.

Algoritmos de otimização linear ⇒ assumem espaço paramétrico $\mathbb{R}^{\nu}$ (s/ restrição)

É necessário realizar transformação de parâmetros.

Exemplo: $\psi_1 = \sigma_\varepsilon^2 > 0 \Rightarrow$ define-se $\psi_1 = e^{2\theta}$ onde $\theta \in (-\infty, \infty)$

$\Rightarrow$ estimamos $\hat{\theta} \Rightarrow \hat{\psi}_1 = e^{2\hat{\theta}}$

Para MEE, as transformações + úteis são:

| parâmetros | restrição | transformação |
|---|---|---|
| $\sigma^2$ | $\sigma^2 > 0$ | $\psi = e^{2\theta}$ |
| $p$ | $0 < p < 1$ | $\psi = \lvert\theta\rvert / \sqrt{1+\theta^2}$ |
| $\phi$ | $-1 < \phi < 1$ | $\psi = \theta / \sqrt{1+\theta^2}$ |
| $\lambda$ | $0 < \lambda < \pi$ | $\psi = 2\pi / 2 + e^{\theta}$ |

→ onde aparece p na FMV, coloca $\psi$. Estima $\hat{\theta}$ e calcula $\hat{p} = \frac{\lvert\hat{\theta}\rvert}{\sqrt{1+\hat{\theta}^{2}}}$

Quando fizermos otimização numérica, teremos $\hat{\theta}$ (não o parâmetro original).

$$\theta^{(k+1)} = \theta^{(k)} + \lambda H^{-1}(\theta^{(k)})\, g(\theta^{(k)}) \quad \Rightarrow$$ Teremos o Hessiano avaliado em $\hat{\theta}$

$$\Rightarrow \hat{\theta} \sim v \times 1 , \quad \theta = T(H)$$

Incerteza: $\hat{\theta}_j \pm z \sqrt{\mathrm{var}(\hat{\theta}_j)}$

Assintoticamente:

$$\hat{\theta} \overset{a}{\sim} N\left(\theta, \frac{I^{-1}(\theta)}{n}\right)$$

$\mathrm{var}(\hat{\theta}) = I^{-1}(\theta)$ : informação de Fisher

onde $I(\theta) = -E\left(\frac{\partial^2 \ell}{\partial\theta\,\partial\theta'}\right) = E\left(\frac{\partial \ell}{\partial\theta}\frac{\partial \ell}{\partial\theta'}\right)$

onde $\ell = \sum_{t=1}^{n} \log p(y_t \mid Y_{t-1})$

$= -\frac{pn}{2}\ln 2\pi - \frac{1}{2}\sum\left[\ln\lvert F_t\rvert + v_t' F_t\right.$

tem $\theta$

Na prática ⇒ difícil termos expressão analítica para $I(\theta)$

Usaremos no lugar de $I(\theta)$, as seguintes estimativas amostrais:

$$\hat{I}_{2D} = -\frac{1}{n}\left.\frac{\partial^2 \ell}{\partial\theta\,\partial\theta'}\right|_{\theta=\hat{\theta}} = -\frac{1}{n}\sum_{t=1}^{n}\left.\frac{\partial^2 \ell_t}{\partial\theta\,\partial\theta'}\right|_{\theta=\hat{\theta}}$$

ou

$$\hat{I}_{op} = \frac{1}{n}\left.\frac{\partial \ell}{\partial\theta}\frac{\partial \ell}{\partial\theta'}\right|_{\theta=\hat{\theta}} = \frac{1}{n}\sum\left.\frac{\partial \ell}{\partial\theta}\frac{\partial \ell}{\partial\theta'}\right|_{\theta=\hat{\theta}}$$

$$\Rightarrow \mathrm{var}(\hat{\theta}) = E\left[(\hat{\theta}-\theta)(\hat{\theta}-\theta)\right] = \frac{\hat{I}^{-1}(\theta)}{n}$$

Na prática, o cálculo de $g(\theta) = \partial l / \partial \theta$ , H e $\hat{I}$ podem ser realizadas através de 2 procedimentos no caso de MEE:

i) utilizando derivadas numéricas

$$\frac{\partial l}{\partial \psi_i} = \lim_{h \to 0} \frac{l(\psi_i + h) - l(\psi_i)}{h}$$

⇒ fazer p/ nível local

ii) semi-analiticamente ⇒ em mente: queremos calcular $\boxed{\frac{\partial l}{\partial \psi_i}}$

$$l(\psi) = \log L(\psi)$$

Dúvida. por que precisa transformar p/ aparecer α na expressão? Fica + fácil de derivar em ψ?

Mas $L(\psi) = p(y;\psi)$ ; $y = (y_1, y_2, \ldots, y_n)$ → marginal γ

e $p(y,\alpha;\psi) = p(\alpha | y;\psi) \cdot p(y;\psi)$ — $p(A|B) = \frac{p(A \cdot B)}{p(B)}$

$$\therefore \log p(y,\alpha;\psi) = \log p(\alpha|y;\psi) + \log p(y;\psi)$$

$$\therefore \boxed{\underbrace{\log p(y;\psi)}_{l(\psi)} = \log p(y,\alpha;\psi) - \log(\alpha|y;\psi)} \quad (I)$$

outra forma de escrever $l(\psi)$ em funções de prob. conjunta e condicional.

Escreve score como fçõs de dens. conjunta e condicional (α|y)

Seja $\tilde{E} = E[\cdot]$ em relação à densidade $p(\alpha|y;\psi)$

$$\Rightarrow \tilde{E}[\log p(y;\psi)] = \tilde{E}[\log(y,\alpha;\psi)] - \tilde{E}[\log p(\alpha|y;\psi)]$$

$$\boxed{\log p(y;\psi) = \tilde{E}[\log p(y,\alpha;\psi)] - \tilde{E}[\log p(\alpha|y;\psi)]} \quad (II)$$

↳ pois $p(y;\psi)$ não depende de α

score ∴

$$\boxed{\left.\frac{\partial l}{\partial \psi}\right|_{\psi=\tilde{\psi}} = \left.\frac{\partial}{\partial \psi}\log L(\psi)\right|_{\psi=\tilde{\psi}} = \left.\frac{\partial}{\partial \psi}\log p(y;\psi)\right|_{\psi=\tilde{\psi}} = \overbrace{\tilde{E}\left[\frac{\partial}{\partial \psi}\log p(y,\alpha;\psi)\Big|_{\psi=\tilde{\psi}}\right]}^{Ⓐ} \quad \underbrace{\tilde{E}\left[\frac{\partial}{\partial \psi}\log p(\alpha|y;\psi)\Big|_{\psi=\tilde{\psi}}\right.}_{Ⓑ}} \quad III$$

→ Calculemos Ⓑ = ?

$$B = \tilde{E}\left[\frac{\partial}{\partial \psi}\log p(\alpha|y;\psi)\right] = \int_{\mathbb{R}} \frac{\partial}{\partial \psi}\log p(\alpha|y;\psi) \cdot p(\alpha|y;\tilde{\psi})\,d\alpha$$

$$= \int \frac{1}{p(\alpha|y;\psi)} \cdot \frac{\partial p(\alpha|y;\psi)}{\partial \psi} \cdot p(\alpha|y;\tilde{\psi})\,d\alpha = \int \frac{\partial p(\alpha|y;\psi)}{\partial \psi}\,d\alpha =$$

$$= \frac{\partial}{\partial \Psi} \underbrace{\int p(\alpha|y)\,d\alpha}_{=1} = \frac{\partial}{\partial \Psi} 1 = 0$$

Logo, de III:

$$\boxed{\underset{\text{score}}{\frac{\partial \log p(y;\Psi)}{\partial \Psi}}\Bigg|_{\Psi=\tilde{\Psi}} = \tilde{E}\left[\underbrace{\frac{\partial \log p(y,\alpha;\Psi)}{\partial \Psi}}_{Ⓐ}\Bigg|_{\Psi=\tilde{\Psi}}\right]} \quad IV$$

→ Precisamos calcular Ⓐ

Escrever conjunta como condicional de $y|\alpha$ e mg de $\alpha$

calculemos $p(y,\alpha;\Psi)$:

$$p(y,\alpha;\Psi) = p(y|\alpha;\Psi)\, p(\alpha;\Psi)$$

- cálculo de $p(y|\alpha;\Psi)$

$p(AB|C) = p(A|BC).\,p(B|C)$

$$p(y|\alpha;\Psi) = p(\underbrace{y_n}_{A}, \underbrace{y_{n-1} \dots y_1}_{B} \,|\, \underbrace{\alpha_n \alpha_{n-1} \dots \alpha_1}_{C})$$

$$= p(y_n | y_{n-1} \dots y_1 \boxed{\alpha_n} \dots \alpha_1)\, p(y_{n-1} \dots y_1 | \alpha_n \dots \alpha_1)$$

como processo é Markoviano, só importa $\alpha_n$

$$= p(y_n|\alpha_n).\; p(y_{n-1} y_{n-2} \dots y_1 | \alpha_n \alpha_{n-1} \dots \alpha_1)$$

⋮ iterativamente

$$= p(y_n|\alpha_n).\; p(y_{n-1}|\alpha_{n-1}) \dots\dots p(y_1|\alpha_1)$$

$$\Rightarrow \boxed{p(y|\alpha;\Psi) = \prod_{t=1}^{n} p(y_t|\alpha_t)}$$

- cálculo de $p(\alpha;\Psi)$

Também pela propriedade de processo Markoviano e uma manipulação anterior:

$$\boxed{p(\alpha;\Psi) = \prod_{t=1}^{n} p(\alpha_t|\alpha_{t-1};\Psi)}$$

Logo: $\boxed{p(y,\alpha;\Psi) = \prod_{t=1}^{n} p(y_t|\alpha_t;\Psi).\; p(\alpha_t|\alpha_{t-1};\Psi)}$ V

$$\therefore \left| \log p(y,\alpha;\psi) = \sum_{t=1}^{n} \underset{\textcircled{C}}{\log p(y_t|\alpha_t;\psi)} + \underset{\textcircled{D}}{\log p(\alpha_t|\alpha_{t-1};\psi)} \right| \quad \text{VI}$$

Consideremos um MEE linear e Gaussiano:

$$\begin{cases} y_t = Z_t\alpha_t + \varepsilon_t & \varepsilon_t \sim N(0, H_t) \quad (a) \\ \alpha_{t+1} = T_t\alpha_t + R_t\eta_t & \eta_t \sim N(0, Q_t) \quad (b) \end{cases}$$

cálculo de $\textcircled{C}$ $\log p(y_t|\alpha_t)$

- $p(y_t|\alpha_t)$

já conhecemos $\alpha_t$

$$\text{De (a)}: \; E[y_t|\alpha_t] = Z_t\alpha_t$$

$$V[y_t|\alpha_t] = var(\varepsilon_t) = H_t$$

$$\Rightarrow p(y_t|\alpha_t) \sim N(Z_t\alpha_t, H_t)$$

$$\therefore p(y_t|\alpha_t) = \frac{1}{(2\pi)^{p/2}|H_t|^{1/2}} \exp\left\{-\frac{1}{2}\overbrace{(y_t - Z_t\alpha_t)'}^{\varepsilon_t} H_t^{-1} \overbrace{(y_t - Z_t\alpha_t)}^{\varepsilon_t}\right\}$$

$$\therefore \left| \log p(y_t|\alpha_t) = -\frac{p}{2}\log(2\pi) - \frac{1}{2}\log|H_t| - \frac{1}{2}\left[\varepsilon_t' H_t^{-1}\varepsilon_t\right] \right| \quad \text{VII}$$

cálculo de $\textcircled{D}$ $\log(p(\alpha_t|\alpha_{t-1}))$

- $p(\alpha_t|\alpha_{t-1})$

$$\text{De (b)}: \; E[\alpha_t|\alpha_{t-1}] = T_{t-1}\alpha_{t-1}$$

$$V[\alpha_t|\alpha_{t-1}] = R_{t-1}Q_{t-1}R_{t-1}'$$

$$\Rightarrow p(\alpha_t|\alpha_{t-1}) \sim N(T_{t-1}\alpha_{t-1}, \; R_{t-1}Q_{t-1}R_{t-1}')$$

$\downarrow$

Atenção! esta seria a variância, mas não podemos escrever assim pois não sabemos se determinante existe

Se pudesse escrever assim:

$$p(\alpha_t|\alpha_{t-1}) = \frac{1}{(2\pi)^{m/2}|R_{t-1}Q_{t-1}R_{t-1}|^{1/2}} \exp\left\{-\frac{1}{2}(\alpha_t - T_{t-1}\alpha_{t-1})'(R_{t-1}Q_{t-1}R_{t-1})^{-1}(\alpha_t - T_{t-1}\alpha_{t-1})\right.$$

$$\therefore \left| \log p(\alpha_t|\alpha_{t-1}) = -\frac{m}{2}\log(2\pi) - \frac{1}{2}\log|R_{t-1}Q_{t-1}R_{t-1}| - \right. \quad \text{VIII}$$

$$-\frac{1}{2}\left[\underbrace{(\alpha_t - T_{t-1}\alpha_{t-1})'}_{\eta_{t-1}}(R_{t-1}Q_{t-1}R_{t-1})^{-1}(\alpha_t - T_{t-1}\alpha_{t-1})\right]$$

## Atenção

Em modelos Estruturais, a matriz $R$ será uma matriz de seleção (apenas 0's e 1's)

$$\Rightarrow \quad R_t = [\, I_r \;\; \underset{\sim}{0} \,]'$$

onde $I_r$ é matriz identidade e $\underset{\sim}{0}'$ é matriz de zeros.

Portanto: $R_t' R_t = I_r$

Exemplo:

Modelo Estrutural básico:

Tendência linear estocástica + sazonalidade estocástica via dummies

$s=4$ (obs. trimestrais)

$$\Rightarrow \quad y_t = \mu_t + \gamma_t + \varepsilon_t$$

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

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

$$\gamma_{t+1} = -(\gamma_t + \gamma_{t-1} + \gamma_{t-2}) + \omega_t$$

Escrevendo em EE:

$$y_t = [1 \;\; 0 \;\; 1 \;\; 0 \;\; 0] \begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} + \underset{m\times r}{\overset{r=3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}} \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

$$\Rightarrow \quad R_t = \begin{bmatrix} I_r \\ 0 \end{bmatrix} = [\, I_r \,\vdots\, 0' \,]'$$

$$\Rightarrow \quad \underset{r\times m}{R_t'} \, \underset{m\times r}{R_t} = \underset{(3\times 3)}{I_{r\times r}} \qquad \begin{matrix} m=5 \\ r=3. \end{matrix}$$

Olhando a expressão p/ $RQR'$ na verossimilhança de $\alpha_t|\alpha_{t-1}$:

$$RQR' = \begin{bmatrix} I_r \\ 0' \end{bmatrix} Q_{r\times r} [I_r \; 0') = Q_{r\times r}$$

No exemplo, como $m > r \Rightarrow$ densidade não existente pois $|RQR'| = 0$.

Para solucionar $\Rightarrow$ utiliza Jacobiano na transformação de densidade de $\eta_t$ p/ a de $\alpha_t|\alpha_{t-1}$.

$$\alpha_t = T_{t-1}\alpha_{t-1} + R_{t-1}\eta_{t-1}$$

$$R_{t-1}\eta_{t-1} = \alpha_t - T_{t-1}\alpha_{t-1}$$

$$\therefore \underbrace{R'_{t-1}R_{t-1}}_{I}\eta_{t-1} = R'_{t-1}(\alpha_t - T_{t-1}\alpha_{t-1})$$

$$\eta_{t-1} = R'_{t-1}(\alpha_t - T_{t-1}\alpha_{t-1})$$

Realizando o procedimento da transformação corretamente:

$$p(\alpha_t|\alpha_{t-1}) = \frac{1}{(2\pi)^{m/2}}\frac{1}{|Q_{t-1}|^{1/2}} \exp\left\{-\frac{1}{2}\eta'_{t-1}Q^{-1}_{t-1}\eta_{t-1}\right\}$$

$$\therefore \boxed{\log(p(\alpha_t|\alpha_{t-1})) = -\frac{m}{2}\log 2\pi - \frac{1}{2}\log|Q_{t-1}| - \frac{1}{2}\{\eta'_{t-1}Q^{-1}_{t-1}\eta_{t-1}\}}$$

De $\overline{\text{VII}}$ e $\overline{\text{VIII}}$ em $\overline{\text{VI}}$:

$$\boxed{\log p(y,\alpha;\psi) = cte - \frac{1}{2}\sum_{t=1}^{n}\left(\log|H_t| + \log|Q_t| + \varepsilon_t' H_t^{-1}\varepsilon_t + \eta'_{t-1}Q^{-1}_{t-1}\eta_{t-1}\right)}$$

Lembrando: Temos que calcular vetor de score.

$$\left.\frac{\partial l}{\partial \psi}\right|_{\psi=\tilde{\psi}} = \tilde{E}\left[\left.\frac{\partial}{\partial\psi}\log p(y,\alpha;\psi)\right|_{\psi=\tilde{\psi}}\right]$$

$$= E\left[\frac{\partial}{\partial\psi}\log p(y,\alpha;\psi)\,\Big|\,y\right]\Bigg|_{\psi=\tilde{\psi}}$$

$\downarrow$

valor esperado condicional

$$\frac{\partial}{\partial \psi} \log p(\alpha, y; \psi) = -\frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \left( \log|H_t| + \log|Q_t| + \varepsilon_t' H_t^{-1} \varepsilon_t + \eta_{t-1}' Q_{t-1} \eta_{t-1} \right)$$

$$\therefore \left. \frac{\partial \ell}{\partial \psi} \right|_{\psi = \tilde{\psi}} = E\left[ -\frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \left( \log|H_t| + \log|Q_t| + \varepsilon_t' H_t^{-1} \varepsilon_t + \eta_{t-1}' Q_{t-1} \eta_{t-1} \right) \middle| y \right]_{\psi = \tilde{\psi}}$$

$$= -\frac{1}{2} \frac{\partial}{\partial \psi} \Sigma \Big[ \log|H_t| + \log|Q_t| + \underbrace{E[\varepsilon_t' H_t^{-1} \varepsilon_t | y]}_{(A1)} + \underbrace{E[\eta_{t-1}' Q_{t-1} \eta_{t-1} | y]}_{(A2)} \Big]_{\psi =}$$

Atenção! É condicional.

**Obs**: três resultados para avaliar (A1) e (A2)

i) Seja $z = x'Bx$ uma forma quadrática

$$z = tr(z) = tr(x'Bx)$$
$$= tr(x'x B)$$

ii) $E[z) = E(tr(z)) = tr(E[z])$

iii) Se $x$ é vetor aleatório

$$V(x) = E[xx') - E(x)(E(x))'$$
$$\therefore E[xx'] = V(x) + E[x)(E[x])'$$

→ Cálculo de (A1) e (A2)

$$A_1 = E[\varepsilon_t' H^{-1} \varepsilon_t | y) = E[tr(\varepsilon_t \varepsilon_t' H_t^{-1} | y))] =$$
$$= tr[E(\varepsilon_t \varepsilon_t' H_t^{-1} | y)] =$$
$$= tr[E[\varepsilon_t \varepsilon_t' | y) H_t^{-1}] =$$
$$= tr[(var(\varepsilon_t | y) + E[\varepsilon_t | y] . E'[\varepsilon_t | y]) H_t^{-1}]$$

Mas $E[\varepsilon_t | y] = \hat{\varepsilon}_t$ : estimativa suavizada de $\varepsilon_t$

$Var(\varepsilon_t | y) =$ variância suavizada de $\varepsilon_t$

$$\Rightarrow A_1 = tr((\hat{\varepsilon}_t \hat{\varepsilon}_t' + var(\varepsilon_t | y)) H_t^{-1})$$

De forma análoga: $A_2 = tr((\hat{\eta}_{t-1} \hat{\eta}_{t-1}' + var(\eta_{t-1} | y)) Q_{t-1}^{-1})$

E assim:

$$\left.\frac{\partial l}{\partial \psi}\right|_{\psi=\tilde{\psi}} = -\frac{1}{2}\frac{\partial}{\partial \psi}\sum_{t=1}^{n}\Big\{\log|H_t| + \log|Q_{t-1}| + tr\big((\hat{\varepsilon}_t'\varepsilon_t + var(\varepsilon_t|y))H_t^{-1}\big)$$

$$+ tr\big((\eta_{t-1}'\hat{\eta}_{t-1} + var(\eta_t|y))Q_{t-1}'\big)\Big\} \text{ avaliado em } \psi = \tilde{\psi}$$

Observar que:

→ apenas $H_t(\psi)$ e $Q_t(\psi)$ são sujeitas a diferenciação com respeito a $\psi$

→ geralmente, $H_t$ e $Q_t$ são funções simples de $\psi$, sendo fácil calcular o vetor de score.

Dúvida

→ DK pg 145: se $\alpha_1$ possui apenas elementos difusos, o efeito da inicialização difusa exata desaparece no vetor de score.

---

Exemplo:

Modelo de nível local

$$\begin{cases} y_t = \mu_t + \varepsilon_t & \varepsilon_t \sim N(0,\sigma_\varepsilon^2) \\ \mu_{t+1} = \mu_t + \eta_t & \eta_t \sim N(0,\sigma_\eta^2) \end{cases}$$

$H_t = \sigma_\varepsilon^2 = \psi_1 \Rightarrow \log H_t = \log \psi_1 \Rightarrow \dfrac{\partial \log H_t}{\partial \psi_1} = \dfrac{1}{\psi_1} \quad \dfrac{\partial \log H_t}{\partial \psi_2} = 0$

$Q_t = \sigma_\eta^2 = \psi_2 \Rightarrow \log Q_t = \log \psi_2 \Rightarrow \dfrac{\partial \log Q_t}{\partial \psi_1} = 0 \quad \dfrac{\partial \log Q_t}{\partial \psi_2} = \dfrac{1}{\psi_2}$

Não deveria usar reparametrização aqui

**Dúvida**

$$\sigma_\varepsilon^2 = \psi_1 = e^{2\theta}$$

Outra dúvida: Na prática, usa eqs. de suavização p/ calcular $\hat{\varepsilon}$, $var(\varepsilon|y)$...

## Algoritmo EM

- Procedimento iterativo p/ maximização do log da verossimilhança (p/ cada iteração, sempre aumenta na direção do ponto de máximo)
- Apresenta convergência lenta no últimos passos (trocar por outro algorit
- Não necessita de 2ª derivada do log da verossimilhança

⇒ Ponto de partida. função densidade conjunta

$$\underbrace{L(y_1 y_2 \dots y_n, \alpha_1 \alpha_2 \dots \alpha_n ; \psi)}_{p(y, \alpha ; \psi)} = \prod_{t=1}^{n} p(y_t | \alpha_t) \, p(\alpha_t | \alpha_{t-1})$$

(calculada na otimização)

log verossimilhança conjunta

$$l^c(\psi) = \log L(y, \alpha ; \psi) =$$

$$= cte - \frac{1}{2} \sum_{t=1}^{n} \left[ \log |H_t| + \log |Q_{t-1}| + \varepsilon_t' H_t^{-1} \varepsilon_t + \eta_{t-1}' Q_{t-1}^{-1} \eta_{t-1} \right]$$

Score:

$$\frac{\partial l^c}{\partial \psi} = -\frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \left[ \log |H_t| + \log |Q_{t-1}| + \varepsilon_t' H_t^{-1} \varepsilon_t + \eta_{t-1}' Q_{t-1}^{-1} \eta_{t-1} \right]$$

Resolver $\frac{\partial l^c}{\partial \psi} = 0$ é achar soluções de MV p/ os elementos de $\psi$

**1º passo:** Realiza-se o passo E, onde o valor esperado é tomado em função de y (suavização)

valor esperado afeta aqui

$$E\left[ \frac{\partial l_c}{\partial \psi} \Big| y \right] = \boxed{\frac{\partial l}{\partial \psi}} = \; = -\frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \Big\{ \log |H_t| + \log |Q_{t-1}| + E\left[ \varepsilon_t' H_t^{-1} \varepsilon_t | y \right] +$$

e aqui

$$+ E\left[ \eta_{t-1}' Q_{t-1}^{-1} \eta_{t-1} | y \right] \Big\}$$

$$= \frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \Big\{ \log |H_t| + \log |Q_{t-1}| + tr\left[ \left( \hat{\varepsilon}_t' \varepsilon_t + var(\varepsilon_t | y) \right) H_t^{-1} \right]$$

$$+ tr\left[ \left( \hat{\eta}_{t-1} \hat{\eta}_{t-1} + var(\eta_t | y) \right) Q_{t-1}^{-1} \right] \Big\} \Big|_{\psi = \hat{\psi}}$$

Cálculos:

- $\hat{\varepsilon}_t$, $\hat{\eta}_{t-1}$, $var(\varepsilon_t|y)$ e $var(\eta_t|y)$ são calculados p/ $\psi = \tilde{\psi}$

↳ começamos com $\psi(0)$

__2º passo__: Realiza-se o passo M; igualando $E(\partial l_c / \partial y) = 0$.

$$\frac{1}{2} \frac{\partial}{\partial \psi} \sum_{t=1}^{n} \Big[ \log|H_t| + \log|Q_{t-1}| + tr\big[ (\hat{\varepsilon}_t' \hat{\varepsilon}_t + var[\varepsilon_t|y]) H_t^{-1} \big] +$$

$$+ tr\big[ (\hat{\eta}_t' \hat{\eta}_t + var[\eta_t|y]) Q_t^{-1} \big] \Big] \Big|_{\psi = \underset{\psi(0)}{\tilde{\psi}}} = 0$$

No 1º passo, afetando $\hat{\varepsilon}_t, \hat{\eta}_t$, $var[\varepsilon_t|y]$, $var(\eta_t|y)$

⇒ Não afeta $H_t$ e $Q_t$

⇒ iguala a 0 e acha expressões p/ hiperparâmetros ⇒ revisa estimativa de $\psi$ p/ $\psi^*$ (calculada c/ base em $\psi(0)$)

Faz-se $\tilde{\psi} = \psi^*$

Repete-se iterativamente passo E e passo M, até convergência

---

__Exemplo__: Modelo de nível local

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

$\psi = (\sigma_\varepsilon^2, \sigma_\eta^2)'$ ⇒ escolhe valores iniciais $\tilde{\psi}^{(0)}$

__Passo E__: Começa com $\tilde{\psi}^{(0)}$ e obtém:

$$\begin{cases} \hat{\varepsilon}_t = E[\varepsilon_t|y] \\ \hat{\eta}_{t-1} = E[\eta_{t-1}|y] \\ var[\varepsilon_t|y] \\ var[\eta_{t-1}|y] \end{cases}$$

⇒ Calculamos usando FK e suavização (DK - 20 e 21)

Avalia, com estes valores:

valores iniciais n/ afetam.

$$\tilde{E}[\log p(y, \alpha; \psi)] = -\frac{1}{2} \sum \Big\{ \log \sigma_\varepsilon^2 + \log \sigma_\eta^2 + \big[ (\hat{\varepsilon}_t^2 + var(\varepsilon_t|y)) / \sigma_\varepsilon^2 \big] +$$

$$+ \big[ \hat{\eta}_t^2 + var(\eta_t|y) / \sigma_\eta^2 \big]$$

Passo M: $\frac{\partial}{\partial \psi} \tilde{E}[\log p(y, \alpha; \psi)] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

e achamos novos valores p/ $\sigma_\varepsilon^2$ e $\sigma_\eta^2$

$$\Rightarrow \begin{bmatrix} \frac{\partial \tilde{E}}{\partial \sigma_\varepsilon^2}[\log p(y, \alpha; \psi)] \\ \frac{\partial \tilde{E}}{\partial \sigma_\eta^2}[\log p(y, \alpha; \psi)] \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

→ Em relação a $\sigma_\varepsilon^2$

são valores calculados.

$$\frac{\partial \tilde{E}}{\partial \sigma_\varepsilon^2}[\cdot] = -\frac{1}{2} \frac{\partial}{\partial \sigma_\varepsilon^2}\left[\sum_{t=1}^{n} \log \sigma_\varepsilon^2 + \frac{1}{\sigma_\varepsilon^2}\left(\hat{\varepsilon}_t^2 + \text{var}(\varepsilon_t|y)\right)\right]$$

$$= -\frac{1}{2}\left[\frac{n}{\sigma_\varepsilon^2} - \frac{1}{\sigma_\varepsilon^4} \sum\left[\hat{\varepsilon}_t^2 + \text{var}(\varepsilon_t|y)\right]\right] = 0$$

$$\therefore \frac{n}{\sigma_\varepsilon^2} - \frac{1}{\sigma_\varepsilon^4} \sum_{t=1}^{n}\left(\hat{\varepsilon}_t^2 + \text{var}(\varepsilon_t|y)\right) = 0$$

$$\therefore \hat{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_{t=1}^{n}\left[\hat{\varepsilon}_t^2 + \text{var}(\varepsilon_t|y)\right]$$

← Dúvida: calcula estas expressões apenas 1 vez. A partir daí vai substituindo recursivamente

usando expressões de DK p/ cálculo de $\hat{\varepsilon}_t^2$ e $\text{var}(\varepsilon_t|y)$:

$$\hat{\varepsilon}_t = \sigma_\varepsilon^2 u_t, \quad u_t = F_t^{-1} v_t - K_t r_t$$

$$\text{var}(\varepsilon_t|y) = \sigma_\varepsilon^2 - \sigma_\varepsilon^4 D_t, \quad D_t = F_t^{-1} + K_t^2 N_t$$

$$\Rightarrow \hat{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_{t=1}^{n}\left[\hat{\sigma}_\varepsilon^{(0)4} u_t^2 + \hat{\sigma}_\varepsilon^{(0)2} - \hat{\sigma}_\varepsilon^{4(0)} D_t\right]$$

$$= \frac{1}{n} \sum_{t=1}^{n} \hat{\sigma}_\varepsilon^{(0)4}(u_t^2 - D_t) + \hat{\sigma}_\varepsilon^{(0)2}$$

$$\boxed{\hat{\sigma}_\varepsilon^2 = \hat{\sigma}_\varepsilon^{2(0)} + \frac{\hat{\sigma}_\varepsilon^{4(0)}}{n} \sum_{t=1}^{n}(u_t^2 - D_t)}$$

passo i+1

passo i

→ Em relação a $\hat{\sigma}_\eta^2$

$$\frac{\partial \tilde{E}[\cdot]}{\partial \sigma_\eta^2} = -\frac{1}{2}\frac{\partial}{\partial \sigma_\eta^2}\left[\sum_{t=1}^{\tilde{n}} \log \sigma_\eta^2 + \frac{[\hat{\eta}_{t-1}^2 + var[\eta_t|y))}{\sigma_\eta^2}\right]$$

Por analogia:

$$\hat{\sigma}_\eta^2 = \frac{1}{n}\Sigma(\hat{\eta}_{t-1}^2 + var[\eta_t|y))$$

Usando expressões de DK para cálculo de $\hat{\eta}_t^2$ e $var[\eta_t|y)$:

$$\hat{\eta}_t = \sigma_\eta^2 r_t \Rightarrow \hat{\eta}_{t-1} = \sigma_\eta^2 r_{t-1}$$

$$var[\eta|y) = \sigma_\eta^2 - \sigma_\eta^4 N_t \Rightarrow var(\eta_{t-1}|y) = \sigma_\eta^2 - \sigma_\eta^4 N_{t-1}$$

$$\Rightarrow \hat{\sigma}_\eta^2 = \frac{1}{n-1}\sum_{t=\textcircled{2}}^{n}\left[\hat{\sigma}_t^{4(0)} r_{t-1}^2 + \hat{\sigma}_\eta^{2(0)} - \hat{\sigma}_\eta^{4(0)} N_t\right]$$

$$\therefore \boxed{\hat{\sigma}_\eta^2 = \hat{\sigma}_\eta^{2(0)} + \frac{\hat{\sigma}_\eta^{4(0)}}{n-1}\sum_{t=2}^{n}(r_{t-1}^2 - N_{t-1})}$$

⇒ Esses novos valores são inseridos no passo E e gera-se uma nova maximização, iterativamente, até convergência