

4

Modelos em Espaço de Estado e Filtro de Kalman

4.1

Modelo em Espaço de Estado

As aplicações do filtro e suavizador de Kalman requerem que os modelos a serem estimados estejam representados na forma em Espaço de Estado (EE). Neste trabalho nos restringiremos apenas aos modelos lineares, para os quais seguem uma definição geral e flexível [cf. Pizzinga (2004)]:

Definição: Seja $Y_t = (Y_{t,1}, Y_{t,2}, \dots, Y_{t,p})$ para $t = 1, 2, \dots$, um processo estocástico p-variado observável definido em algum espaço de probabilidade (Ω, Λ, P) . Diz-se que Y_t tem uma representação em Espaço de Estado Linear, se existem duas equações, uma das medidas e outra do estado, dadas respectivamente por:

$$Y_{t \ (px1)} = Z_{t \ (pxm)} \alpha_{t \ (mx1)} + d_{t \ (px1)} + \varepsilon_{t \ (px1)} \quad , \quad \varepsilon_t \sim WN(0_{(px1)}, H_{t \ (pxp)}) \quad (30)$$

$$\alpha_{t+1 \ (mx1)} = T_{t \ (mxm)} \alpha_{t \ (mx1)} + c_{t \ (mx1)} + R_{t \ (mxr)} \eta_{t \ (rx1)} \quad , \quad \eta_t \sim WN(0_{(rx1)}, Q_{t \ (rxr)})$$

$$t = 1, 2, \dots$$

para as quais:

- (i) α_t é um processo estocástico m-variado, em geral não observável, chamado de vetor de estado, ou puramente, estado.
- (ii) α_1 é um vetor aleatório m-variado de segunda ordem, chamado estado inicial, com vetor de médias $a_{\cdot 1}$ e matriz de covariâncias não negativa definida P_1 .
- (iii) ε_t é um processo estocástico p-variado, em geral não observável e de segunda ordem, chamado de erro da equação das medidas, com vetor de médias nulo e matriz de covariâncias não negativa definida H_t .

(iv) η_t é um processo estocástico r-variado, em geral não observável, chamado de erro da equação do estado, com vetor de médias nulo e matriz de covariâncias não negativa definida Q_t .

(v) ε_t, η_s e α_1 são vetores aleatórios não correlacionados para quaisquer instantes t e s.

(vi) $Z_t, d_t, H_t, T_t, c_t, R_t, e Q_t$ são ditas matrizes do sistema e têm natureza determinística, podendo depender de termos passados do processo Y_t .

A representação acima será referida como forma em Espaço de Estado “wide sense” dada a generalização com que foi definida. Representações com propriedades especiais surgem com a imposição de restrições. Seguem abaixo alguns desses modelos, cujas nomenclaturas são largamente citadas na literatura.

- Modelo em Espaço de Estado Linear Gaussiano: Sejam as equações em (30) válidas. Dizemos que Y_t segue um modelo em Espaço de Estado Linear Gaussiano [cf. Durbin e Koopman, 2001] se valem as seguintes condições:

$$(i) \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim NID \left(\begin{bmatrix} 0_{px1} \\ 0_{rx1} \end{bmatrix}, \begin{bmatrix} H_{t \text{ } p \times p} & 0_{p \times r} \\ 0_{r \times p} & Q_{t \text{ } r \times r} \end{bmatrix} \right), \quad (31)$$

$$(ii) \alpha_1 \sim N(a_1, P_1) \text{ e}$$

$$(iii) \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \text{ e } \alpha_1 \text{ são independentes,}$$

- Modelo em Espaço de Estado Linear Condicionalmente Gaussiano: Denomina-se modelo em Espaço de Estado Linear Condicionalmente Gaussiano [cf Harvey (1989)] aos modelos em Espaço de Estado que apresentam termo(s) dependente(s) de observações passadas de Y_t em alguma(s) matriz(es) do sistema.

Diversos modelos podem ser estruturados na forma em Espaço de Estado, entre eles, os modelos ARMA, ARIMA, SARIMA, modelos de regressão com coeficientes constantes ou variantes no tempo, modelos estruturais uni ou

multivariados, Vetores Auto-Regressivos (VAR) e modelos de volatilidade estocástica, dentre outros. Conforme listado em Marques (2009), aprofundamentos sobre a passagem destes e outros modelos de séries temporais para esta forma podem ser encontrados em Harvey (1989), Hamilton (1994), Durbin e Koopman (2001), Brockwell e Davis (2002), e Shumway e Stoffer (2006).

4.2

O Filtro de Kalman

O filtro de Kalman é um algoritmo que fornece estimadores não tendenciosos e consistentes do vetor de estado, para cada instante de tempo.

O filtro gera um estimador linear ótimo sob o ponto de vista do erro quadrático médio. Quando a hipótese de normalidade dos erros do estado e das medidas é assumida, o estimador obtido é ótimo não apenas entre os lineares⁸, coincidindo com a média do vetor α_t condicionada à σ -álgebra \mathfrak{F}_j ⁹.

Sob uma perspectiva geométrica, o filtro gera a projeção ortogonal das coordenadas univariadas do vetor de estado α_t no subespaço gerado pelas coordenadas univariadas de todas as medidas até o instante j [cf. Pizzinga (2004)], ou informalmente, do conjunto de informações disponíveis até o momento j .

De acordo com a σ -álgebra condicionante, podemos obter os seguintes tipos de estimação do estado:

- se $j < t$, o estimador realiza uma previsão do estado (a_t)
- se $j = t$, o estimador realiza uma atualização ou filtragem do estado

$$(a_{t/t})$$

- se $j > t$, o estimador realiza a suavização do estado ($\tilde{\alpha}_t$).

⁸ Estimadores formados pela combinação linear das variáveis $1, Y_1, Y_2, \dots, Y_j$,

⁹ Denomina-se \mathfrak{F}_j à σ -álgebra gerada pelas variáveis Y_1, Y_2, \dots, Y_j convencionalmente denotada por $\sigma(Y_1, Y_2, \dots, Y_j)$

É intuitivo que, quanto maior o uso de informações para a estimação, maior a precisão do estimador obtido pelo filtro de Kalman. Desta forma, é fácil mostrar que a variância do estado suavizado (\tilde{V}_t) é inferior (ou igual) à do estado atualizado ($P_{t/t}$), que por sua vez é inferior (ou igual) à do estado previsto (P_t) [vide Pizzinga (2004) e Marques (2009)].

O filtro de Kalman é construído a partir de um conjunto de equações recursivas, que serão apresentadas após a exposição da seguinte notação básica:

- n : número total de observações presente em cada uma das séries do vetor Y ;
- $\mathbf{a}_{t/j} \equiv \Pi_{S_j}(\alpha_t)$: projeção ortogonal de α_t no subespaço gerado pelo conjunto de todas as combinações lineares das coordenadas de $S_j = (1, Y_1, Y_2, \dots, Y_j)$, que sob normalidade, equivale a média condicional de α_t dada \mathfrak{T}_j ;
- $\mathbf{a}_t \equiv \mathbf{a}_{t/t-1}$: estado previsto;
- $\mathbf{a}_{t/t}$: estado atualizado;
- $\mathbf{a}_{t/n}$: estado suavizado;
- $P_{t/j} = MSE(\mathbf{a}_{t/j}) = E[(\alpha_t - \mathbf{a}_{t/j})(\alpha_t - \mathbf{a}_{t/j})' / \mathfrak{T}_j]$: matriz de erros quadráticos médios de α_t condicionados a \mathfrak{T}_j , que sob normalidade equivale a matriz de variâncias e covariâncias de α_t condicionada a \mathfrak{T}_j ;
- $P_t = P_{t/t-1}$
- $V_t = P_{t/n}$;
- $v_t \equiv Y_t - E(Y_t / \mathfrak{T}_{t-1})$: inovações
- $F_t \equiv Var(v_t / \mathfrak{T}_{t-1})$: matriz de variâncias e covariâncias condicionais das inovações

Seguem as recursões do filtro de Kalman para realização da previsão, atualização e suavização (pela abordagem de intervalo fixo) do estado, quando este possui apenas variáveis estacionárias de segunda ordem:

- **Previsão:**

$$\begin{aligned} a_{t+1} &= T_t a_{t/t} + c_t \\ P_{t+1} &= T_t P_{t/t} T_t' + R_t Q_t R_t' \end{aligned} \quad (32)$$

$$\text{Sendo: } \begin{cases} a_1 = (I_m - T_1)^{-1} c_1 \\ \text{vec}(P_1) = (I_{m^2} - T_1 \otimes T_1')^{-1} \text{vec}(Q_1) \end{cases}$$

- **Atualização:**

$$\begin{aligned} a_{t/t} &= a_t + P_t Z_t' F_t^{-1} v_t \\ P_{t/t} &= P_t - P_t Z_t' F_t^{-1} Z_t P_t \end{aligned} \quad (33)$$

para F_t^{-1} qualquer inversa generalizada de F_t

Observe que os *inputs* das equações de previsão são os vetores atualizados passados, enquanto os *inputs* das equações de atualização são os vetores previstos correntes. Como o vetor de estado é composto unicamente por componentes estacionárias de segunda ordem, a_1 e P_1 são conhecidos *a priori*, o que permite que o filtro seja iniciado com as equações de atualização para obtenção de $a_{1/1}$ e $P_{1/1}$. Uma vez calculados, podemos obter o segundo estado previsto a_2 e a matriz P_2 , os quais servirão de *inputs* para a obtenção dos vetores atualizados $a_{2/2}$ e $P_{2/2}$ e assim sucessivamente.

Perceba que substituindo $a_{t/t}$ e $P_{t/t}$ de (33), nas equações de previsão em (32), obtemos as seguintes recursões para o filtro, conhecidas como recursões do filtro de Kalman “2 em 1”:

$$\begin{aligned} a_{t+1} &= T_t (a_t + P_t Z_t' F_t^{-1} v_t) + c_t \\ P_{t+1} &= T_t (P_t - P_t Z_t' F_t^{-1} Z_t P_t) T_t' + R_t Q_t R_t' \end{aligned} \quad (34)$$

- **Suavização:**

$$\begin{aligned}\hat{\alpha}_t &= a_{t/t} + A_t(\hat{\alpha}_{t/t} - a_{t+1}) \\ V_t &= P_{t/t} - A_t(P_{t+1} - V_{t+1})A_t'\end{aligned}\quad (35)$$

para $t = n-1, n-2, \dots, 1$

sendo: $A_t = P_{t/t}T_t'P_{t+1}^-$ e P_{t+1}^- qualquer inversa generalizada de P_{t+1} .

Outra representação para o suavizador é dada por:

$$\begin{aligned}\hat{\alpha}_t &= a_t + P_t r_{t-1} \\ r_{t-1} &= Z_t F_t^- v_t + (T_t - T_t P_t Z_t' F_t^- Z_t)' r_t \\ V_t &= P_t - P_t N_{t-1} P_t \\ N_{t-1} &= Z_t' F_t^- Z_t + (T_t - T_t P_t Z_t' F_t^- Z_t)' N_t (T_t - T_t P_t Z_t' F_t^- Z_t) \\ &\text{com} \\ r_n &= 0, \\ V_n &= 0, \\ &F_t^- \text{ qualquer inversa generalizada de } F_t\end{aligned}\quad (36)$$

As vantagens da segunda representação foram exaustivamente exploradas em De Jong (1989) e, entre elas, cita-se a eficiência computacional.

Quando as matrizes no sistema não possuem parâmetros desconhecidos (caso teórico) a estimação do vetor de estado pode ser feita diretamente através da aplicação do algoritmo de Kalman. Caso contrário, será necessário o uso de métodos de estimação específicos para obtenção dos hiperparâmetros. Neste trabalho, utilizaremos o método de *Quasi* Máxima Verossimilhança (QMV) para obtenção de estimativas dos hiperparâmetros.

4.3

Estimador de *Quasi* Máxima Verossimilhança e Aplicação do Filtro de Kalman

Em exercícios teóricos é comum supor o conhecimento *a priori* da função de distribuição condicional de Y_t dada a σ -álgebra \mathfrak{F}_{t-1} . Mais especificamente, costuma-se atribuir a distribuição Normal para o processo

Y_t / \mathfrak{F}_{t-1} dada a vantagem do estimador resultante ser ótimo¹⁰ não somente entre os estimadores lineares e poder ser interpretado como a média condicional de α_t dada a σ -álgebra \mathfrak{F}_{t-1} .

Em um ambiente teórico onde a distribuição dos dados é conhecida, o estimador que maximiza a função de verossimilhança possui propriedades desejáveis como consistência e eficiência. Entretanto, quando passamos “da teoria para a prática”, lidamos com o desconhecimento da função de distribuição dos dados. Ao supor normalidade para dados que na verdade não possuem tal distribuição, gera-se uma função de verossimilhança diferente daquela que seria obtida através da distribuição verdadeira das séries. Conseqüentemente, o estimador que maximiza a função de verossimilhança otimiza uma função que não reflete a realidade. Este estimador é dito então de *Quasi* Máxima Verossimilhança (QMV) e, embora não seja eficiente, possui boas propriedades assintóticas, consistência e invariância em relação a funções bijetivas.

Posto isso, imporemos a suposição de Normalidade à distribuição condicional de Y_t :

$$Y_t / \mathfrak{F}_{t-1} \sim N[E(Y_t / \mathfrak{F}_{t-1}), V(Y_t / \mathfrak{F}_{t-1})],$$

de forma que a função de verossimilhança definida por:

$$L(\psi) = p(Y_1, Y_2, \dots, Y_p) = \prod_{t=1}^n p(Y_t / \mathfrak{F}_{t-1})$$

(onde: ψ é o vetor dos parâmetros e $\mathfrak{F}_0 = \{\phi, \Omega\}$ é a σ -álgebra minimal),

pode ser escrita, de acordo com a teoria da distribuição Normal Multivariada, como:

$$L(\psi) = \frac{1}{(2\pi)^p * \det[V(Y_t / \mathfrak{F}_{t-1})]^{1/2}} * \exp\left\{-\left(\frac{1}{2}\right) * (Y_t - E(Y_t / \mathfrak{F}_{t-1}))' * (V(Y_t / \mathfrak{F}_{t-1}))^{-1} * (Y_t - E(Y_t / \mathfrak{F}_{t-1}))\right\} \quad (37)$$

$$L(\psi) = \frac{1}{(2\pi)^p * \det[F_t]^{1/2}} * \exp\left\{-\left(\frac{1}{2}\right) * v_t' * (F_t)^{-1} * v_t\right\} \quad (38)$$

$$L(\psi) = \frac{1}{(2\pi)^p * \det[Z_t' * P_t * Z_t' + H_t]^{1/2}} * \exp\left\{-\left(\frac{1}{2}\right) * (Z_t' * a_t - d_t)' * (Z_t' * P_t * Z_t' + H_t)^{-1} * (Z_t' * a_t - d_t)\right\} \quad (39)$$

¹⁰ Possui variância mínima.

Perceba que a equação (39) descreve a verossimilhança como uma função de a_t , P_t e das matrizes de transição.

Para efeito de simplificação dos cálculos da otimização, optamos¹¹ por maximizar a função $\ell(\psi) = \ln[L(\psi)]$ ao invés de $L(\psi)$.

$$\ell(\psi) = -\frac{np}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n [\ln(\det(F_t) + v_t' F_t^{-1} v_t)] \quad (40)$$

$$\ell(\psi) = -\frac{np}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n [\ln(\det(Z_t P_t Z_t' + H_t) + (Z_t a_t - d_t)' (Z_t P_t Z_t' + H_t)^{-1} (Z_t a_t - d_t))] \quad (41)$$

Observe que os parâmetros desconhecidos estão presentes nas matrizes do sistema, as quais aparecem como *inputs* das recursões do filtro de Kalman, o qual por sua vez gera a_t e P_t , *inputs* da função de verossimilhança. Desta forma, a estimação dos parâmetros terá de ser feita concomitantemente com a estimação da média e variância condicionais do estado (supondo normalidade) pelo filtro de Kalman, como ilustrado a seguir:

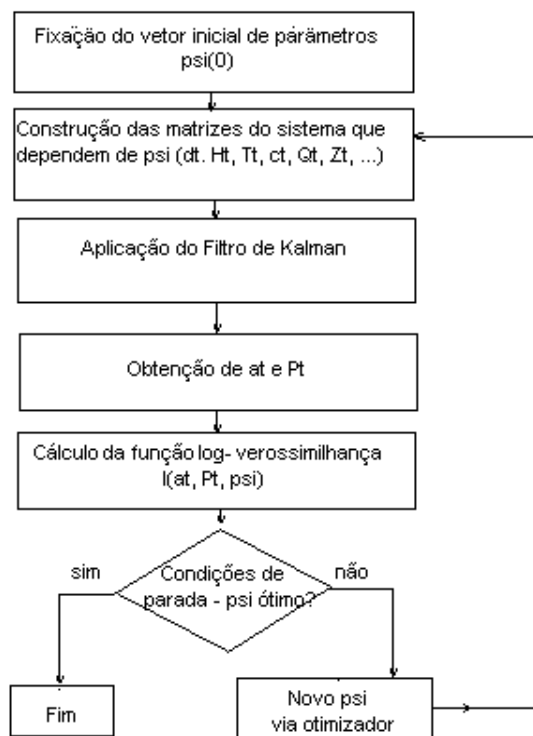


Figura 2 – Fluxo da estimação dos hiperparâmetros em um modelo de EE linear gaussiano

¹¹ Tal opção faz-se possível em decorrência da monotonicidade crescente e estrita da função $f(x) = \ln(x), x \in \mathcal{R}_+^*$

Após a obtenção do vetor ótimo de parâmetros, calcula-se pela última vez as matrizes do sistema e aplicam-se o filtro e o suavizador de Kalman para obtenção dos estados finais previsto e suavizado.

4.4

Inicialização do filtro de Kalman

Na presença de variáveis não estacionárias de segunda ordem no estado, o vetor a_1 e a matriz P_1 (que sob normalidade denotam respectivamente a média e variância incondicionais do vetor inicial α_1) não são bem definidos, violando a condição (ii) presente na definição da forma em Espaço de Estado Linear e invalidando, portanto, todas as recursões subseqüentes. Para comprovar isto, basta observar que para a_1 e P_1 não definidos, as recursões em (32) e (33) (ou (34)) apresentam problema de inicialização, isto é, não é possível obter os estados previstos seguintes.

Há diferentes propostas para lidar com este inconveniente. Duas delas estão apresentadas a seguir:

- Inicialização Difusa¹² Aproximada – “Big Kappa” [cf. Harvey e Phillips (1979)]
- Inicialização Difusa Exata [cf. Koopman (1997) e Durbin e Koopman (2001)]

Antes da apresentação formal das abordagens acima, segue notação especial [cf. Durbin e Koopman (2001)] para representação do vetor de estado inicial, na qual se basearão as explanações posteriores.

Seja α_1 o vetor de estado inicial. Representa-se α_1 da seguinte forma:

$$\alpha_{1 \times 1} = a_{0 \times 1} + A_{m \times q} \delta_{q \times 1} + R_{0 \times (m-q)} \eta_{0(m-q) \times 1} \quad (42)$$

$$\eta_{0(m-q) \times 1} \sim (0_{(m-q) \times 1}, Q_{0(m-q), (m-q)})$$

¹² Adotaremos aqui o termo vetor aleatório difuso para tratar de variáveis como δ , com distribuição $N(0, \kappa I)$ quando $\kappa \rightarrow \infty$. De forma informal, são variáveis não estacionárias que apresentam variâncias arbitrariamente grandes. Desta forma, conforme Durbin e Koopman (2001, p. 101), denomina-se inicialização difusa do filtro de Kalman a qualquer procedimento que vise a aplicação do filtro de Kalman quando houver no vetor inicial pelo menos uma coordenada difusa.

Onde,

- q é o número de variáveis não estacionárias no vetor estado;
- δ é um vetor cujas coordenadas são variáveis não estacionárias e ,

$$E(\delta) = 0_{q \times 1} \text{ e } V(\delta) = \kappa \mathbf{I}_{q \times q} \text{ para } \kappa \rightarrow \infty \text{ (}\kappa \text{ lê-se kappa)}$$

- η_0 é um vetor aleatório de segunda ordem, de forma que Q_0 é uma matriz positiva definida.
- Seja $a_0[i]$ a i -ésima linha do vetor a_0 . Então,

$$a_0[i] = \begin{cases} E(\alpha[i]), & \text{se } \alpha[i] \text{ é uma variável estacionária} \\ 0, & \text{se } \alpha[i] \text{ é uma variável não estacionária} \end{cases}$$

- Sejam i_1, i_2, \dots, i_q as linhas do vetor estado que possuem variáveis não estacionárias. Então a matriz A é tal que os elementos da j -ésima, coluna para $j=1, 2, \dots, q$ são:

$$A[i][j] = \begin{cases} 1 & \text{para a linha } i = i_j \\ 0 & \text{para as demais linhas} \end{cases}$$

De forma sucinta, a j -ésima coluna da matriz A corresponde à i_j -ésima coluna de uma matriz identidade de ordem m

- Sejam $i'_1, i'_2, \dots, i'_{m-q}$ as linhas do vetor estado que possuem variáveis estacionárias. Então a matriz R_0 é tal que os elementos da j -ésima coluna, para $j=1, 2, \dots, m-q$, são:

$$R_0[i][j] = \begin{cases} 1 & \text{para a linha } i = i'_j \\ 0 & \text{para as demais linhas} \end{cases}$$

Perceba, portanto, que a matriz R_0 possui as colunas da matriz identidade de ordem m que faltam na matriz A , logo $A'_{q \times m} R_{0_{m \times (q-m)}} = 0_{q \times (q-m)}$. As matrizes A e R_0 são ditas matrizes de seleção.

Dada a definição em (42) podemos obter a média e a variância incondicionais do vetor inicial:

$$\begin{aligned}
 a_1 &= E(\alpha_1) = a_0 \\
 P_1 &= V(\alpha_1) = AV(\delta)A' + R_0 Q_0 R_0' \\
 &= A(\kappa I)A' + R_0 Q_0 R_0'
 \end{aligned}
 \tag{43}$$

4.4.1

Inicialização Difusa Aproximada – “Big Kappa”

A abordagem aproximada sugerida por Harvey e Phillips (1979) consiste em fixar um valor arbitrariamente alto para κ , e em seguida substituir este valor na equação (43), obtendo P_1 bem definido. Uma vez que a_1 e P_1 estão bem definidos, as recursões em (32) e (33) (ou (34)) podem ser implementadas exatamente como no caso estacionário.

Em resumo, a introdução da constante κ na variância do bloco não estacionário (δ) do estado inicial tem como objetivo descrever a grande incerteza acerca do comportamento desta componente. A variância das coordenadas difusas é, portanto, imposta a critério do analista, sendo κ , em geral, superior a 10^3 .

A simplicidade de implementação apresentada pelo método é acompanhada de algumas desvantagens como uma potencial instabilidade computacional e estimativas possivelmente divergentes para diferentes valores de κ .

4.4.2

Inicialização Difusa Exata

Proposto por Koopman (1997), este método confere uma forma não arbitrária de lidar com a incerteza em relação à variância incondicional do estado inicial.

Este tipo de inicialização “consiste basicamente em reconhecer que as equações do filtro de Kalman ‘2 em 1’ e do suavizador de Kalman dependem de κ na forma de funções suaves, escrever estas funções como fórmulas de Maclaurin, e tomar $\kappa \rightarrow \infty$ nestas expressões.” [cf. Marques (2009)]. Tal

exercício é trabalhoso e apresenta-se detalhado em Durbin e Koopman (2001), cap. 5.

A intuição que permeia as recursões do Filtro de Kalman com Inicialização difusa Exata (FKIE) é, entretanto, simples e deriva da seguinte extensão da equação (43):

$$\begin{aligned}
 P_1 = V(\alpha_1) &= AV(\delta)A' + R_0 Q_0 R_0' \\
 &= A(\kappa I)A' + R_0 Q_0 R_0' \\
 &= \kappa(AA') + (R_0 Q_0 R_0') \\
 &= \kappa P_{\infty,1} + P_{*,1}
 \end{aligned} \tag{44}$$

Observe que dada a notação do estado inicial apresentada em (44), a matriz P_1 pode ser separada em duas componentes distintas: uma “bem comportada” e outra “mal comportada” quando $\kappa \rightarrow \infty$, respectivamente representadas por $P_{*,1}$ e $\kappa P_{\infty,1}$.

Pelo FKIE, as variâncias dos estados seguintes também podem ser decompostas em duas componentes (bem e mal comportadas quando $\kappa \rightarrow \infty$), de modo que o cálculo de P_2 depende de P_1 , o cálculo de P_3 depende de P_2 , e assim sucessivamente. As recursões do FKIE permitem que a estrutura difusa do vetor inicial se propague pelos vetores de estado seguintes. Tal propagação é reduzida a cada iteração até a eventual dissipação total do componente difuso, quando então a matriz P_t se estabiliza. Denominamos o último instante para o qual o estado assume comportamento difuso (ou equivalentemente, último instante para o qual P_∞ é não nula) como instante d ¹³. A partir de $t=d+1$ a variância do estado passa a ser bem comportada ($P_t = P_{*,t}$) e, mais do que isso, as recursões do FKIE passam a coincidir com as do filtro original.

¹³ Para modelos estruturais, a teoria aponta que d coincide com o número de componentes não estacionárias no vetor de estado. Entretanto, para séries temporais curtas tal fato teórico pode não ser observado na prática.

4.5

Tratamento Univariado para Séries Multivariadas

Uma das vantagens do filtro de Kalman é a facilidade com que suas recursões podem ser aplicadas a séries multivariadas.

Entretanto, quando o vetor Y_t apresenta dimensão alta (p), as recursões do filtro de Kalman podem apresentar problemas computacionais relacionados às repetidas inversões das matrizes de variância das inovações F_t . Outro foco de ineficiência computacional da aplicação do filtro para séries multivariadas advém do processo de otimização da função de verossimilhança, a qual, sob hipótese de normalidade, também depende de F_t^{-1} (vide equação 38).

Sob a hipótese de que os erros do vetor ε_t são não correlacionados (ou, equivalentemente, que a matriz H_t é diagonal) é possível aplicar um tratamento univariado às séries multivariadas, eliminando a dificuldade mencionada acerca de inversão de matrizes.

Tal procedimento consiste em:

- (i) Seccionar as matrizes de transição Z_t e d_t :

$$\begin{bmatrix} Y_{t,1} \\ Y_{t,2} \\ \vdots \\ Y_{t,p} \end{bmatrix} = \begin{bmatrix} Z_{t,1} \\ Z_{t,2} \\ \vdots \\ Z_{t,p} \end{bmatrix} * \alpha_t + \begin{bmatrix} d_{t,1} \\ d_{t,2} \\ \vdots \\ d_{t,p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \\ \vdots \\ \varepsilon_{t,p} \end{bmatrix}, \quad (45)$$

de modo que $Z_{t,i}$ e $d_{t,i}$ representam a i -ésima linha das matrizes Z_t e d_t , respectivamente.

Observe que sob a hipótese de H_t diagonal podemos reescrever as equações do modelo multivariado clássico apresentado em (30) da seguinte forma:

$$\begin{aligned} Y_{t,j} &= Z_{t,j} \alpha_t + d_{t,j} + \varepsilon_{t,j}, & \varepsilon_{t,j} &\sim WN(0, H_{t,jj}) \text{ para } j = 1, \dots, p \\ \alpha_{t+1} &= T_t \alpha_t + c_t + R_t \eta_t, & \eta_t &\sim WN(0, Q_t) \end{aligned} \quad (46)$$

onde $H_{t,jj}$ representa o elemento da linha j e coluna j da matriz H_t e $Y_{t,j}$ é uma série univariada.

- (ii) Escrever o modelo (46) na forma de Espaço de Estado:

Embora o modelo (46) contenha as mesmas informações que o exposto em (30), o primeiro não está na forma EE, posto que há três equações de medidas para cada equação de estado.

Para colocar tal modelo na forma EE é preciso construir um novo vetor de estado $[(\alpha_t^*)_j \text{ para } j=1,2,\dots]$ tal que, para cada equação de $Y_{t,j}$, haja uma equação definindo $(\alpha_t^*)_j$. Esta construção deve ser feita cuidadosamente, de modo a preservar todas as informações fornecidas originalmente para α_t (em 30)

Observando atentamente o modelo (46), verifica-se que para um dado instante de tempo t , a equação das medidas varia com o sub-índice j , mas o estado presente no segundo termo da equação é o mesmo para todo j .

Desta forma a relação correta entre $(\alpha_t^*)_j$ e α_t é:

$$(\alpha_{t(mx1)}^*)_j = \alpha_{t(mx1)} \text{ para } j=1,2,\dots,p. \quad (47)$$

$$(\alpha_{t+1(mx1)}^*)_j = \alpha_{t+1(mx1)} \text{ para } j=1,2,\dots,p.$$

Organizando serialmente as variáveis dos vetores $Y_{t,j}$ e $(\alpha_t^*)_j$ temos:

$$\downarrow \left\{ \begin{array}{l} Y_{1,1} = 1^a \text{ coordenada de } Y_1 \\ Y_{1,2} = 2^a \text{ coordenada de } Y_1 \\ Y_{1,\dots} \\ Y_{1,p} = p^a \text{ coordenada de } Y_1 \\ Y_{2,1} = 1^a \text{ coordenada de } Y_2 \\ Y_{2,2} = 2^a \text{ coordenada de } Y_2 \\ Y_{2,\dots} \\ Y_{2,p} = p^a \text{ coordenada de } Y_2 \\ Y_{n,1} = 1^a \text{ coordenada de } Y_n \\ Y_{n,2} = 2^a \text{ coordenada de } Y_n \\ Y_{n,\dots} \\ Y_{n,p} = p^a \text{ coordenada de } Y_n \end{array} \right. \quad \downarrow \left\{ \begin{array}{l} \alpha_{1,1}^* = \alpha_1 \\ \alpha_{1,2}^* = \alpha_1 \\ \alpha_{1,\dots}^* = \alpha_1 \\ \alpha_{1,p}^* = \alpha_1 \\ \alpha_{2,1}^* = \alpha_2 \\ \alpha_{2,2}^* = \alpha_2 \\ \alpha_{2,\dots}^* = \alpha_2 \\ \alpha_{2,p}^* = \alpha_2 \\ \alpha_{n,1}^* = \alpha_n \\ \alpha_{n,2}^* = \alpha_n \\ \alpha_{n,\dots}^* = \alpha_n \\ \alpha_{n,p}^* = \alpha_n \end{array} \right. \quad (48)$$

Com base nesta organização serial, segue o modelo univariado na forma em Espaço de Estado que preserva todas as informações originais do modelo multivariado em (30):

$$\begin{array}{l}
 \text{Para } j=1, 2, \dots, p-1^{14}: \\
 \quad Y_{t,j} = Z_{t,j} \alpha_{t,j} + d_{tj1} + \varepsilon_{t,j} \\
 \quad \alpha_{t,j+1} = \alpha_{t,j} , \\
 \\
 \text{Para } j=p \\
 \quad Y_{t,p} = Z_{t,1} \alpha_{t,p} + d_{t,p} + \varepsilon_{t,p} \\
 \quad \alpha_{t+1,1} = T_t \alpha_{t,p} + c_{t,p} + R_t \eta_t
 \end{array}
 \quad \left. \vphantom{\begin{array}{l} \\ \\ \\ \\ \\ \end{array}} \right\} (49)$$

¹⁴ Observe que para $j=1,2,\dots,p-1$, assume-se que $T_t = I$, $c_t = 0$ e $R_t = 0$. Desta forma a equação do estado respeita todas as hipóteses definidas em (30).