

6

Modelos em Espaço de Estado Não Lineares e Filtros Não Lineares Aproximados

Neste capítulo, o modelo em EE definido na seção 2.1 é devidamente generalizado, de forma que suas equações – a das medidas e a do estado – possam ter um aspecto funcional não linear. Tal generalização ganha sentido nesta Dissertação, a partir do momento que é reconhecida sua utilidade no processo de imposição de restrições não lineares às componentes de um modelo em EE.

Também são discutidos alguns procedimentos de estimação *aproximada* do vetor de estado, dentre os quais o Filtro de Kalman Estendido [Wishner, Tabaczynski e Athans (1969), Anderson e Moore (1979), Harvey (1989), Fuhrer (1992) e, mais recentemente, Tanizaki (1996)] e o Filtro de Simulação de Monte Carlo [Tanizaki (1996) e Tanizaki e Mariano (1996)] desempenham papel central na apresentação. Alguns resultados teóricos, no que diz respeito a vantagens e desvantagens destes estimadores aproximados, são enunciados, eventualmente demonstrados e discutidos.

6.1

Definição Geral de um Modelo Não Linear

Perseguindo o mesmo encaminhamento utilizado na seção 2.1, apresenta-se uma formulação geral, ainda sob um aspecto axiomático, de um modelo em EE *intrinsecamente* não linear, que é de notação bastante próxima daquela proposta em Durbin e Koopman (2001a), página 184.

Definição 6.1. *Seja $Y_t = (Y_{t1}, \dots, Y_{tp})$, $t = 1, 2, \dots$, 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 não linear se existem duas equações, uma das medidas e outra do estado, dadas respectivamente por*

$$\begin{aligned} Y_t &= Z_t(\alpha_t) + \varepsilon_t \\ \alpha_{t+1} &= T_t(\alpha_t) + R_t\eta_t \\ t &= 1, 2, \dots \end{aligned} \quad (6.1)$$

valendo exatamente todos os pressupostos anteriores enunciados na Definição 2.1, com exceção de que Z_t e T_t , antes matrizes representando transformações lineares, agora podem ser quaisquer transformações mensuráveis (com respeito a Borel).

Vale dizer que todas as observações apresentadas na seção 2.1 após a definição de um modelo em EE linear se aplicam diretamente aqui, com a necessidade de óbvias modificações. No entanto, seguem-se algumas *novas* observações relevantes, quais sejam:

- 1) Pressupõe-se que todas os vetores aleatórios envolvidos no modelo e possíveis transformações (mensuráveis!) dos mesmos são incondicionalmente integráveis (isto deve ser postulado, uma vez não sendo diretamente dedutível a partir de transformações não lineares dos erros ε_t e η_t);
- 2) Em Anderson e Moore (1979) e Harvey (1989), uma especificação um pouco mais geral de um modelo não linear é abordada, na qual $R_t = R_t(\alpha_t)$, isto é, a matriz R_t pode depender do estado do instante t . A função que caracteriza esta dependência, assim como Z_t e T_t , deve ser mensurável;
- 3) Quando ε_t e η_t são considerados Gaussianos, então, os parâmetros desconhecidos do modelo não linear poderão ser estimados por uma maneira heurística que Fuhrer (1992) utilizou e denominou de *máxima verossimilhança aproximada*. Este procedimento, também corroborado e igualmente denominado por Tanizaki (1996) e Tanizaki e Mariano (1996), é discutido no próximo capítulo desta Dissertação; e
- 4) De forma exatamente igual ao que fora feito no capítulo 2, pode-se adotar as mesmas definições de *componente*, *componente natural* e *componente de interesse* ao se trabalhar com o modelo em EE não linear proposto acima.

Antes de terminar esta seção, discute-se uma questão importante, a qual concerne à maneira de como definir um modelo em EE não linear. Em Tanizaki (1996) e Tanizaki e Mariano (1996), é proposta uma formulação diferente daquela dada na expressão em (6.1), a qual se caracteriza pelas equações

$$\begin{aligned} Y_t &= Z_t(\alpha_t, \varepsilon_t) \\ \alpha_{t+1} &= T_t(\alpha_t, \eta_t) \\ t &= 1, 2, \dots, \end{aligned} \quad (6.2)$$

nas quais Z_t e T_t ainda são funções mensuráveis arbitrárias e os erros ε_t e η_t ainda satisfazem os mesmos pressupostos anteriores. No entanto, este não é um motivo de preocupação, pois a formulação exibida anteriormente em (6.1) se configura como um caso particular desta última em (6.2), e vice-versa. Este fato é consumado na seguinte

Proposição 6.2. *As duas formulações apresentadas são “equivalentes”.*

Prova: Construtiva.

(6.1) \Rightarrow (6.2): Dadas as equações dadas em (6.1), basta que se definam as funções relativas a (6.2), já devidamente avaliadas no vetor de estado e nos respectivos erros, como sendo $Z_t^*(\alpha_t, \varepsilon_t) = Z_t(\alpha_t) + \varepsilon_t$ e $T_t^*(\alpha_t, \eta_t) = T_t(\alpha_t) + R_t \eta_t$.

(6.2) \Rightarrow (6.1): Dadas as equações em (6.2), escreva-se

$$Y_t = Z_t(\alpha_t, \varepsilon_t) + \varepsilon_t^*, \quad \varepsilon_t^* \sim (0,0), \quad e$$

$$\begin{pmatrix} \alpha_{t+1} \\ \varepsilon_{t+1} \\ \eta_{t+1} \end{pmatrix} = \begin{pmatrix} T_t(\alpha_t, \eta_t) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ I_p & 0 \\ 0 & I_r \end{pmatrix} \begin{pmatrix} \tilde{\varepsilon}_t \\ \tilde{\eta}_t \end{pmatrix}, \quad \begin{pmatrix} \tilde{\varepsilon}_t \\ \tilde{\eta}_t \end{pmatrix} \sim \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H_{t+1} & 0 \\ 0 & Q_{t+1} \end{pmatrix} \right).$$

$$\text{Defina-se, então, } H_t^* \equiv 0, \quad \alpha_t^* \equiv \begin{pmatrix} \alpha_t \\ \varepsilon_t \\ \eta_t \end{pmatrix}, \quad Z_t^*(\alpha_t^*) \equiv Z_t(\alpha_t, \varepsilon_t),$$

$$T_t^*(\alpha_t^*) \equiv \begin{pmatrix} T_t(\alpha_t, \eta_t) \\ 0 \\ 0 \end{pmatrix}, \quad R_t^* \equiv \begin{pmatrix} 0 & 0 \\ I_p & 0 \\ 0 & I_r \end{pmatrix}, \quad \eta_t^* \equiv \begin{pmatrix} \tilde{\varepsilon}_t \\ \tilde{\eta}_t \end{pmatrix} \text{ e } Q_t^* \equiv \begin{pmatrix} H_{t+1} & 0 \\ 0 & Q_{t+1} \end{pmatrix}.$$

Com base nas quantidades descritas acima, torna-se possível escrever o modelo dado pelas equações de (6.2) na forma

$$\begin{aligned} Y_t^* &= Z_t^*(\alpha_t^*) + \varepsilon_t^* \\ \alpha_{t+1}^* &= T_t^*(\alpha_t^*) + R_t^* \eta_t^* \end{aligned}, \quad \begin{pmatrix} \varepsilon_t^* \\ \eta_t^* \end{pmatrix} \sim \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H_t^* & 0 \\ 0 & Q_t^* \end{pmatrix} \right),$$

a qual é um caso particular da formulação dada em (6.1). \square

No transcurso deste capítulo, todas as metodologias (referentes aos filtros aproximados) são desenvolvidas para a formulação referente a (6.1), associada à Definição 6.1. Se, porventura, o modelo em EE não linear a ser proposto esteja genuinamente na formulação referente a (6.2), então há-de se fazer a conversão descrita na prova da Proposição 6.2, no intuito de que as técnicas, aqui concebidas, possam ser usadas. Para apresentação destas últimas com base na formulação de (6.2), sugerem-se Tanizaki (1996) e Tanizaki e Mariano (1996).

Como exemplo de um modelo que possa se enquadrar nestes tipos expressão, cita-se uma proposta de generalização do modelo de fatores para classes de ativos [Sharpe (1988), Sharpe (1992) e Varga e Valli (1998)] usado para determinação do estilo de gestão em fundos de investimento, de forma que as participações em cada tipo de mercado possam variar (estocasticamente!) ao

longo do tempo. Para manter a perspectiva original do modelo *estático*, há-de se fazer com que as participações somem um e sejam positivas, para todo os instantes de tempo. No oitavo capítulo, serão apresentadas considerações adicionais sobre este tipo de modelo de fatores, juntamente com uma proposta de formulação das equações em (6.1), de maneira que tais restrições sejam satisfeitas.

Outro exemplo, considerado, talvez, o mais canônico de todos, seria o modelo estrutural com tendência e sazonalidade multiplicativas. Este foi discutido em Durbin e Koopman (2001a) e, com relativa profundidade, em Shephard (1994).

Para finalizar, cita-se um terceiro exemplo, contextualizado, como o primeiro exemplo, no tema da Dissertação: imagine-se que haja interesse em um modelo para uma série que apresente nível estocástico, sendo este restrito a assumir valores entre -1 e 1 . Neste caso, a componente de interesse seria, justamente, o nível e este deveria ser dado por uma transformação não linear do estado – nesta situação, univariado – que satisfizesse o requerimento em tela. Uma

possível candidata poderia ser $Z_t(\alpha_t) = \frac{\alpha_t}{1 + |\alpha_t|}$. O estado, por sua vez, poderia

evoluir como um passeio aleatório sem *drift*, ou seja, $\alpha_{t+1} = \alpha_t + \eta_t$. Note-se que a função Z_t é diferenciável quase certamente (com efeito: observe-se que α_t é variável aleatória contínua e que $\{0\}$ é boreliano com medida de Lebesgue nula; logo, α_t assume o “valor crítico” 0 com probabilidade nula!). Harvey (1989) propõe um modelo não linear semelhante, mas que postula a restrição ao intervalo $(0,1)$.

Outros exemplos de modelos em EE não lineares, inclusive com não linearidade na equação do estado, podem ser encontrados em Anderson e Moore (1979), Harvey (1989) e Tanizaki (1996).

6.2

O Filtro de Kalman Estendido

O primeiro de um conjunto de possíveis tratamentos *aproximados* a serem discutidos neste capítulo para modelos que atendem as especificações postuladas

na Definição 6.1 é o chamado *Filtro de Kalman Estendido*, ou *Filtro Não Linear de Primeira Ordem*, ou ainda *Filtro de Primeira Ordem*. Este se baseia em uma linearização inicial das funções Z_t e T_t apresentadas nas equações em (6.1) através de uma expansão em Taylor primeira ordem (Pressuposto adicional ao modelo dado na Definição 6.1: diferenciabilidade de primeira ordem para as funções Z_t e T_t). Como referências clássicas do assunto, citam-se Wishner, Tabaczynski e Athans (1969), Anderson e Moore (1979) e Harvey (1989). Como referências contemporâneas, citam-se o tratamento estatístico dado por Tanizaki (1996) e outro mais adequado a problemas da engenharia dado por Chui e Chen (1999).

6.2.1

As Recursões do Filtro de Kalman Estendido

O Filtro de Kalman Estendido se constitui de um procedimento heurístico, visando a um paralelo com as recursões de previsão e de atualização do Filtro de Kalman, apresentadas no capítulo 3 em (3.2) e (3.3), respectivamente, para um modelo linearizado, mas que, como será visto, preserva a não linearidade original em todos os instantes de tempo.

A técnica pode ser sintetizada pela seguinte seqüência de passos:

1º Passo) Para um dado instante t , aproximar as funções Z_t e T_t em expansão de Taylor de primeira ordem, em torno, respectivamente, de a_t e $a_{t/t}$, sendo que estes últimos possuem o mesmo significado de antes – isto é, vetor de estado *previsto* e *atualizado* no instante t , respectivamente. São obtidas, então,

$$\begin{aligned} Z_t(\alpha_t) &\cong Z_t(a_t) + \dot{Z}_t(\alpha_t - a_t) \\ T_t(\alpha_t) &\cong T_t(a_{t/t}) + \dot{T}_t(\alpha_t - a_{t/t}) \end{aligned} \quad , \quad (6.3)$$

nas quais \dot{Z}_t é matriz Jacobiana da função Z_t avaliada em a_t e \dot{T}_t é a matriz Jacobiana da função T_t avaliada em $a_{t/t}$.

2º Passo) Estabelecer a seguinte aproximação linear para o modelo não linear dada por

$$\begin{aligned} Y_t &= \dot{Z}_t \alpha_t + d_t + \tilde{\varepsilon}_t \\ \alpha_{t+1} &= \dot{T}_t \alpha_t + c_t + \tilde{\eta}_t \end{aligned} \quad (6.4)$$

na qual $d_t = Z_t(a_t) - \dot{Z}_t a_t$, $c_t = T_t(a_{t/t}) - \dot{T}_t a_{t/t}$ e os novos erros $\tilde{\varepsilon}_t$ e $\tilde{\eta}_t$ são dados pela soma dos erros originais ε_t e $R_t \eta_t$ com resíduos da expansão em Taylor, os quais dependem de α_t e do passado das medidas até o instante $t-1$.

3º Passo) Para a aproximação linear dada no passo anterior, aplicar as recursões tradicionais de Kalman dadas em (3.2) e (3.3), com as seguintes mudanças nas equações de atualização e de previsão do estado, dadas por

$$\begin{aligned} a_{t/t} &= a_t + P_t \dot{Z}_t' F_t^{-1} (Y_t - Z_t(a_t)) \\ a_{t+1} &= T_t(a_{t/t}) \end{aligned} \quad (6.5)$$

Observe-se que: (i) a não linearidade da especificação original é recuperada na definição da *pseudo-inovação* $\tilde{v}_t = Y_t - Z_t(a_t)$ e na equação de previsão do estado; (ii) as quantidades d_t e c_t nunca são calculadas, como pertinentemente salientado em Harvey (1989), na primeira linha da página 162; e (iii) as matrizes P_t , $P_{t/t}$ e F_t não podem ser calculadas independentemente de a_t e $a_{t/t}$, como destacado em Anderson e Moore (1979).

As equações dadas em (6.5), juntamente com as referentes às matrizes de erros médios quadráticos P_t , $P_{t/t}$ e F_t dadas em (3.2) e (3.3), constituem o Filtro de Kalman Estendido. Para inicialização deste processo recursivo, são necessárias quantidades iniciais a_1 e P_1 .

Em Anderson e Moore (1979), há uma proposta de aprimoramento do Filtro de Kalman Estendido, a qual se baseia na expansão em Taylor da função Z_t em torno de $a_{t/t}$. Este é denominado de *Filtro de Kalman Estendido Iterado*.

6.2.2

Propriedades do Filtro de Kalman Estendido

Embora sendo uma tecnologia heurística, é possível que se estabeleçam, de forma rigorosa, algumas propriedades, boas e ruins, do Filtro de Kalman Estendido. A maioria delas são deduzidas com base nos pressupostos mais fortes

de que os vetores de erros ε_t e η_t são Gaussianos e independentes, os quais são adotados a partir deste momento. Sob normalidade, as quantidades computadas pelo Filtro de Kalman Estendido podem ser interpretadas como sendo *aproximações* de vetores de médias e matrizes de covariâncias condicionais. Sob um ambiente mais geral, poder-se-ia pensar que $a_{t/t}$ pode ser visto como aproximações de imagens da única projeção ortogonal das componentes univariadas do estado no subespaço gerado pelas componentes univariadas passadas das medidas, e que $P_{t/t}$ pode ser vista como a respectiva matriz aproximada de erros médios condicionais (vide discussões análogas nas seções 3.1 e 4.4 desta Dissertação).

6.2.2.1

Vantagens

Uma vantagem interessante deste filtro não linear, já citada no capítulo 1, é a relativa facilidade de implementação que este apresenta. Pode-se dizer que o esforço computacional demandado para tal tarefa é praticamente o mesmo que aquele relativo à implementação do Filtro de Kalman tradicional para modelos em EE lineares. O único ônus adicional a este último é a necessidade de serem implementados métodos de derivação numérica; no entanto, este não se faria presente se as equações fossem implementadas utilizando-se diretamente derivadas analíticas.¹

Outra considerável vantagem, agora relativa à qualidade da estimação do estado, tanto em termos de previsão quanto de atualização, é dada pelo seguinte

Teorema 6.3. *Seja o modelo em EE não linear dado da Definição 6.1 com vetores de erro Gaussianos e independentes, e com vetor de estado inicial Gaussiano com vetor de médias a_1 e matriz de covariâncias P_1 e independente dos vetores de erros para todo instante de tempo t . Então:*

¹ Ao se trabalhar com derivadas analíticas, embora proporcionando velocidade computacional, a implementação resultante do Filtro de Kalman Estendido se torna automaticamente restrita a uma única especificação de modelo em EE não linear, o que pode não ser interessante para fins de generalizações.

- (a) Se $\alpha_t/Y_1, \dots, Y_{t-1} \sim N(a_t, P_t)$, tem-se que $\alpha_t/Y_1, \dots, Y_t \approx N(a_{t/t}, P_{t/t})$, sempre que $P_t \approx 0$, uniformemente em α_t e em Y_t .
- (b) Se $\alpha_t/Y_1, \dots, Y_t \sim N(a_{t/t}, P_{t/t})$, tem-se que $\alpha_{t+1}/Y_1, \dots, Y_t \approx N(a_{t+1}, P_{t+1})$, sempre que $P_{t/t} \approx 0$.

Prova: Vide Anderson e Moore (1979). □

Então, sob normalidade, sempre que se estima o vetor de estado com base no Filtro de Kalman com *precisão considerável* através do previsto (atualizado), observa-se que o atualizado (previsto) imediatamente seguinte é próximo de uma média condicional. Em outras palavras, sob estes pressupostos, o Filtro de Kalman Estendido fornece estimadores do estado aproximadamente ótimos. No entanto, Anderson e Moore (1979) argumentam que o Teorema 6.3 é de utilidade limitada, uma vez que as condições suficientes enunciadas (o vetor de estado inicial deve ter sempre distribuição Gaussiana e as aproximações $P_t \approx 0$ e $P_{t/t} \approx 0$ devem ser satisfeitas para todo instante de tempo t) são um tanto improváveis na prática. Mas, ainda assim, este é um resultado teórico relevante que pode ser um argumento a favor da adoção do Filtro de Kalman Estendido, pois nada impede que estas condições suficientes ocorram pelo menos localmente (isto é, para alguns instantes de tempo).

6.2.2.2

Desvantagens

Quando se pensa em Filtro de Kalman Estendido, ou em qualquer outro procedimento baseado em expansões de Taylor de ordens superiores, assunto a ser discutido mais adiante, o que se tem logo em mente é a seguinte aproximação, em geral *inapropriada*, envolvendo a esperança condicional:

$$E(g(X)/\Lambda) \cong g(E(X/\Lambda)) \quad , \quad (6.6)$$

na qual X é um vetor aleatório integrável definido em algum espaço de probabilidade apropriado $(\Omega, \mathfrak{F}, P)$, Λ é uma sigma álgebra contida em \mathfrak{F} e g é

uma função mensurável Borel qualquer, mas tal que $g(X)$ seja integrável. Se g é linear, então é sabido que a equação dada em (6.6) é válida na forma de uma igualdade P-quase certamente. No entanto, em geral, esta relação não é válida. Por exemplo, se X é univariado e g é uma função convexa não afim com valores reais, então a Desigualdade de Jensen garante que $E(g(X)/\Lambda) \geq g(E(X/\Lambda))$ P-quase certamente.

Logo, quando se admite que

$$\begin{aligned} \mathbf{a}_t &\equiv E(\alpha_t/Y_1, \dots, Y_{t-1}) = E(T_{t-1}(\alpha_{t-1}) + R_{t-1}\eta_{t-1}/Y_1, \dots, Y_{t-1}) \\ &\cong T_{t-1}(E(\alpha_{t-1}/Y_1, \dots, Y_{t-1})) = T_{t-1}(\mathbf{a}_{t-1/t-1}) \end{aligned} \quad (6.7)$$

e que

$$\begin{aligned} \hat{Y}_{t/t-1} &\equiv E(Y_t/Y_1, \dots, Y_{t-1}) = E(Z_t(\alpha_t) + \varepsilon_t/Y_1, \dots, Y_{t-1}) \\ &\cong Z_t(E(\alpha_t/Y_1, \dots, Y_{t-1})) = Z_t(\mathbf{a}_t) \end{aligned} \quad (6.8)$$

sendo estas aproximações inadequadas, Tanizaki (1996) postula que, em geral, o modelo aproximado em (6.4) – cujas expressões no instante t dependem das recursões do Filtro de Kalman Estendido calculadas nos instantes anteriores – apresenta os seguintes problemas para todo instante de tempo t :

- 1) $E(\tilde{\varepsilon}_t/Y_1, \dots, Y_{t-1}) \neq 0$ e $E(\tilde{\eta}_t/Y_1, \dots, Y_t) \neq 0$
 - 2) $\text{Cov}(\tilde{\varepsilon}_t, \alpha_t/Y_1, \dots, Y_{t-1}) \neq 0$ e $\text{Cov}(\tilde{\eta}_t, \alpha_t/Y_1, \dots, Y_t) \neq 0$
 - 3) $\text{Cov}(\tilde{\varepsilon}_t, \tilde{\eta}_{t-1}/Y_1, \dots, Y_{t-1}) \neq 0$
 - 4) $\tilde{\varepsilon}_t$ e $\tilde{\eta}_{t-1}$ são condicionalmente não normais, dado Y_1, \dots, Y_{t-1}
- (6.9)

A argumentação deste fato é intuitiva, pois se baseia na composição dos erros $\tilde{\varepsilon}_t$ e $\tilde{\eta}_t$, os quais são funções de α_t e do passado das medidas, e no uso de das aproximações inadequadas dadas em (6.7) e (6.8).

Uma vez que os erros das equações das medidas e do estado violam vários pressupostos, como pode ser visto em (6.9), tem-se, diretamente, ausência de garantia de que o Filtro de Kalman Estendido proporcione estimadores não viciados, tampouco de variância mínima, uma vez que sua tentativa de aproximar o vetor de médias e matriz de covariâncias condicionais do vetor de estado pode trivialmente “fracassar”. Em outras palavras, *espera-se, em geral, que o Filtro de*

Kalman Estendido seja um estimador viciado do vetor de estado. Cita-se, também, que as pseudo-inovações v_t , mesmo padronizadas, também podem apresentar média condicional não nula e correlação serial incondicional não nula para alguns *lags*, além de não normalidade, fato que pode prejudicar a prática de diagnósticos (vide próximo capítulo) e que corrobora, ainda mais, a possível presença de viés no processo de estimação.

Outro problema apresentado pelo Filtro de Kalman Estendido é falta de teoria desenvolvida para a prática de inicializações difusas - aproximadas ou exatas – como discutido na seção 3.6 no tratamento de modelos em EE lineares estimados pelo Filtro de Kalman.

6.3

O Filtro de Segunda Ordem

Uma maneira de se ajustar o modelo em EE não linear dado pela Definição 6.1 ainda de forma aproximada, mas com relativamente pequena demanda computacional, seria através de um filtro não linear baseado em expansões de Taylor de ordem maior, diga-se de segunda ordem, para as funções Z_t e T_t (Pressuposto adicional ao modelo dado na Definição 6.1: diferenciabilidade de segunda ordem para as funções Z_t e T_t). As expressões resultantes constituem o chamado *Filtro Não Linear de Segunda Ordem*, ou simplesmente *Filtro de Segunda Ordem*. Citam-se como referências sobre o assunto Wishner, Tabaczynski e Athans (1969), Gelb (1974) e Tanizaki (1996).

6.3.1

Recursões do Filtro de Segunda Ordem

As expressões do Filtro de Segunda Ordem são bem mais complicadas do que aquelas referentes ao Filtro de Kalman Estendido, pois envolvem expansões de Taylor de segunda ordem para funções vetoriais; algo complexo sob perspectiva algébrica. O pesquisador interessado nestas expressões está convidado a consultar as páginas 52, 53 e 54 de Tanizaki (1996) para a apresentação da versão linearizada, necessária ao método, do modelo não linear dado pelas

expressões em (6.1) e das recursões do Filtro de Segunda Ordem². Outra referência indicada é Wishner, Tabaczynski e Athans (1969), páginas 490 e 491.

6.3.2

Propriedades do Filtro de Segunda Ordem

Como vantagens principais do Filtro de Segunda Ordem, citam-se

- Esforço computacional ainda não muito grande, embora de implementação, em termos de linhas de programação, mais trabalhosa do que a do Filtro de Kalman Estendido; e
- Melhor aproximação de médias condicionais (ou imagens de projeções ortogonais) e matrizes de covariâncias condicionais (ou matrizes de erros médios quadráticos condicionais), uma vez que mais termos da expansão em Taylor são considerados.

Como desvantagem principal, cita-se que, como este é um filtro que também explora a aproximação, muitas vezes inadequada, referida em (6.6), então, em geral, espera-se que *todos* os problemas enunciados em (6.9) aconteçam para a versão linearizada do modelo, implicando, possivelmente, em viés para os estimadores do estado, da mesma forma como ocorre com o Filtro de Kalman Estendido. Outra desvantagem considerável é também a questão da inicialização, a qual carece de uma teoria mais contundente na literatura estatística.

6.4

O Filtro de Simulação de Monte Carlo

Tanizaki (1996) e Tanizaki e Mariano (1996) propõem uma tecnologia menos heurística para estimação recursiva do vetor de estado em um modelo em EE não linear, baseada no “cálculo” de vetores de médias e matrizes de covariâncias, através de integração estocástica via simulações de Monte Carlo. A metodologia resultante ficou denominada de *Filtro de Simulação de Monte Carlo*,

² Como Tanizaki (1996) adota as expressões dadas em (6.2), seriam necessárias algumas modificações para que as recursões se tornem adequadas às expressões de (6.1). Este tema não será discutido na presente Dissertação.

ou *Filtro de Primeira Ordem com Simulação de Monte Carlo*. Em linhas gerais, este filtro se objetiva em melhorar o Filtro de Kalman Estendido discutido anteriormente, no sentido de calcular (leia-se “aproximar”) melhor algumas esperanças condicionais, rejeitando o uso da relação dada em (6.6).

6.4.1

Recursões do Filtro de Simulação de Monte Carlo

O Filtro de Simulação de Monte Carlo é inteiramente baseado numa aproximação linear das expressões dadas em (6.1) referentes ao modelo em EE não linear, a qual compreende vetores de erros que apresentam apenas os problemas 3 e 4 listados em (6.9) – e, eventualmente, apenas o problema 4, como garante a Proposição 6.6 mais adiante. Isto é relevante, pois se eliminam fontes de viés que poderiam se fazer presentes no processo recursivo.

A versão linearizada do modelo em EE não linear, que se constitui no bloco construtor do Filtro de Simulação de Monte Carlo, é dada por

$$\begin{aligned} Y_t &= Y_{t/t-1} + M_{t/t-1} P_t^{-1} (\alpha_t - a_t) + \tilde{\varepsilon}_t \quad (\text{Equação das Medidas}) \\ \alpha_{t+1} &= a_{t+1} + N_{t+1/t} P_{t/t}^{-1} (\alpha_t - a_{t/t}) + \tilde{\eta}_t \quad (\text{Equação do Estado}) \quad , \quad (6.10) \\ & \quad t = 1, 2, \dots \end{aligned}$$

no qual³:

- a_t e $a_{t/t}$ são as médias condicionais de α_t , dado todo o passado das medidas até t-1 e até t, respectivamente;
- P_t e $P_{t/t}$ são as matrizes de covariâncias condicionais de α_t , dado todo o passado das medidas até t-1 e até t, respectivamente;
- $Y_{t/t-1}$ é a média condicional de Y_t , dado todo o passado das medidas até o instante t-1;
- $M_{t/t-1}$ é matriz de covariâncias condicionais entre Y_t e α_t , dado todo o passado das medidas até t-1;

³ Não se deve confundir o cálculo de algumas destas quantidades com aqueles discutidos no decorrer do capítulo 3, para o caso de modelos lineares.

- $N_{t+1/t}$ é matriz de covariâncias condicionais entre α_{t+1} e α_t , dado todo o passado das medidas até t ; e
- $\tilde{\varepsilon}_t$ e $\tilde{\eta}_t$ são vetores de erro que dependem, de forma apropriada, de ε_t e $R_t \eta_t$, do vetor de estado α_t e de termos envolvendo o passado das medidas até $t-1$ e t , respectivamente.

Inicie-se a argumentação a favor desta aproximação proposta, estabelecendo-se o seguinte

Lema 6.4. *Sejam dois vetores aleatórios Y e X de segunda ordem definidos em um mesmo espaço de probabilidade, com primeiros e segundos momentos dados*

$$\text{por } E \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \text{ e por } \text{Var} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}, \text{ com } \Sigma_{YY} \text{ invertível.}$$

Defina-se $\tilde{X} \equiv X - \Sigma_{XY} \Sigma_{YY}^{-1} (Y - \mu_Y)$. Então, \tilde{X} e Y são vetores aleatórios não correlacionados.

$$\text{Prova: } \text{Cov}(\tilde{X}, Y) = E\tilde{X}Y' - E\tilde{X}EY'$$

$$\begin{aligned} &= E \left[(X - \Sigma_{XY} \Sigma_{YY}^{-1} (Y - \mu_Y)) Y' \right] - E(X - \Sigma_{XY} \Sigma_{YY}^{-1} (Y - \mu_Y)) \mu_Y' \\ &= EXY' - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YY} - \mu_X \mu_Y' + \Sigma_{XY} \Sigma_{YY}^{-1} E(Y - \mu_Y) \mu_Y' \\ &= EXY' - \mu_X \mu_Y' - \Sigma_{XY} + 0 = \Sigma_{XY} - \Sigma_{XY} = 0. \quad \square \end{aligned}$$

Fazendo-se as devidas associações notacionais com a aproximação linear dada em (6.10), evidencia-se que:

- 1) dado todo o passado das medidas até $t-1$, o vetor de erros $\tilde{\varepsilon}_t$ e o vetor de estado α_t satisfazem os pressupostos do Lema 6.4, além de que $\tilde{\varepsilon}_t$ tem média condicional nula; e
- 2) dado todo o passado das medidas até t , o vetor de erros $\tilde{\eta}_t$ e o vetor de estado α_t também satisfazem os pressupostos do Lema 6.4, além de que $\tilde{\eta}_t$ tem média condicional nula.

Com isso, é acabado de se demonstrar o

Corolário 6.5. *Para a aproximação em debate dada em (6.10), os problemas 1 e 2 listados em (6.9) estão automaticamente resolvidos.*

Ainda há uma outra situação em que esta aproximação se torna mais interessante, no sentido de que o problema 3 dado em (6.10) também é resolvido. Esta idéia é sumarizada na seguinte

Proposição 6.6. *Considere-se o modelo em EE não linear estabelecido na Definição 6.1. Se a função Z_t for afim (isto é, linear e transladada por um vetor) e se os vetores de erros ε_t e η_s forem independentes⁴ para quaisquer t e s , então $\tilde{\varepsilon}_t$ e $\tilde{\eta}_{t-1}$ associados à aproximação linear dada em (6.10) são condicionalmente não correlacionados, dado todo o passado das medidas até o instante $t-1$.*

Prova: [Adaptada de Tanizaki (1996)] Rearrumando-se as expressões de (6.10), tem-se que

$$\begin{aligned}\tilde{\varepsilon}_t &= Y_t - Y_{t/t-1} - M_{t/t-1} P_t^{-1} (\alpha_t - a_t) \\ &= Y_t - E(Z_t(\alpha_t)/Y_1, \dots, Y_{t-1}) - M_{t/t-1} P_t^{-1} (\alpha_t - a_t)\end{aligned}$$

e que

$$\begin{aligned}\tilde{\eta}_{t-1} &= \alpha_t - a_t - N_{t/t-1} P_{t-1/t-1}^{-1} (\alpha_{t-1} - a_{t-1/t-1}) \\ &= \alpha_t - E(T_{t-1}(\alpha_{t-1})/Y_1, \dots, Y_{t-1}) - N_{t/t-1} P_{t-1/t-1}^{-1} (\alpha_{t-1} - a_{t-1/t-1})\end{aligned}$$

Se $Z_t(\alpha_t) \equiv Z_t \alpha_t + d_t$, na qual Z_t é uma matriz $p \times m$ e d_t é um vetor $p \times 1$, então, da penúltima expressão, deduz-se que

⁴ Note-se que o pressuposto original era de apenas matriz de correlações nula!

$$\begin{aligned}
\tilde{\varepsilon}_t &= Y_t - Z_t a_t - Z_t P_t P_t^{-1} (\alpha_t - a_t) \\
&= Z_t (\alpha_t - a_t) + \varepsilon_t - Z_t (\alpha_t - a_t) \\
&= \varepsilon_t
\end{aligned} \tag{6.11}$$

Como ε_t é independente de α_{t-1} e de η_{t-1} dado as medidas até t-1, tem-se que ε_t e $\tilde{\eta}_{t-1}$ também serão independentes dado as medidas até t-1.

Conclusão: Linearidade na equação das medidas e independência dos erros induz ausência de correlação condicional entre os erros dados na aproximação linear (6.10). \square

Logo, quando se tem um modelo em EE “não linear no estado” com choques independentes, o modelo aproximado proposto em (6.10) é mais interessante ainda, pois este preserva maior número de pressupostos, isto é, o problema (ou violação) 3 da lista dada em (6.9) é eliminado.

Se, porventura, os dois problemas restantes forem “ignorados” – ou o último restante, caso se tenha não linearidade apenas na equação do estado – pode ser provado, com base em propriedades sobejamente conhecidas da distribuição normal multivariada e na própria estrutura da aproximação linear dada em (6.10), que

$$\begin{aligned}
a_{t/t} &= a_t + M'_{t/t-1} F_t^{-1} (Y_t - Y_{t/t-1}) \\
P_{t/t} &= P_t - M'_{t/t-1} F_t^{-1} M_{t/t-1}
\end{aligned}, \tag{6.12}$$

nas quais $a_{t/t}$ e $P_{t/t}$ são, como antes, o vetor de médias e a matriz de variâncias condicionais do vetor de estado α_t , dado todas as medidas até t, e F_t é a matriz de covariâncias condicional de Y_t , dado todas as medidas até t-1.

A idéia central do Filtro de Simulação de Monte Carlo é justamente fazer com que a_t , $M_{t/t-1}$, F_t , $Y_{t/t-1}$ e P_t sejam calculados (leia-se estimados) – com o mínimo possível de viés, algo que, por construção, não se consegue com tanta desenvoltura pelos filtros baseados em expansão de Taylor, posto que estes utilizam a indesejável relação dada em (6.6). Uma vez que estas quantidades estejam livres, ou quase livres de vícios, as expressões em (6.12) também estarão livres, ou quase livres.

As recursões do Filtro de Simulação de Monte Carlo são dadas por:

$$\begin{aligned}\hat{Y}_{v/t-1} &= \frac{1}{K} \sum_{j=1}^K (Z_t(\alpha_{i,t}) + \varepsilon_{i,t}) \\ \hat{F}_t &= \frac{1}{K} \sum_{j=1}^K \left((Z_t(\alpha_{i,t}) + \varepsilon_{i,t}) - \hat{Y}_{v/t-1} \right) \left((Z_t(\alpha_{i,t}) + \varepsilon_{i,t}) - \hat{Y}_{v/t-1} \right)' \\ \hat{M}_{v/t-1} &= \frac{1}{K} \sum_{j=1}^K \left((Z_t(\alpha_{i,t}) + \varepsilon_{i,t}) - \hat{Y}_{v/t-1} \right) (\alpha_{i,t} - \hat{a}_t)'\end{aligned}\quad (6.13)$$

sendo que os vetores $(\alpha_{i,t}, \varepsilon_{i,t})'$ são "sorteios", independentes

entre si, da distribuição $N\left(\begin{pmatrix} \hat{a}_t \\ 0 \end{pmatrix}, \begin{pmatrix} \hat{P}_t & 0 \\ 0 & H_t \end{pmatrix}\right)$, $\forall i = 1, \dots, K$, e

$$\begin{aligned}\hat{a}_{v/t} &= \hat{a}_t + \hat{M}'_{v/t-1} \hat{F}_t^{-1} (Y_t - \hat{Y}_{v/t-1}) \\ \hat{P}_{v/t} &= \hat{P}_t - \hat{M}'_{v/t-1} \hat{F}_t^{-1} \hat{M}_{v/t-1} \\ \hat{a}_{t+1} &= \frac{1}{K} \sum_{j=1}^K (T_t(\alpha_{i,t/t}) + R_t \eta_{i,t}) \\ \hat{P}_{t+1} &= \frac{1}{K} \sum_{j=1}^K \left((T_t(\alpha_{i,t/t}) + R_t \eta_{i,t}) - \hat{a}_{t+1} \right) \left((T_t(\alpha_{i,t/t}) + R_t \eta_{i,t}) - \hat{a}_{t+1} \right)'\end{aligned}\quad (6.14)$$

sendo que os vetores $(\alpha_{i,t/t}, \eta_{i,t})'$ são "sorteios", independentes

entre si, da distribuição $N\left(\begin{pmatrix} \hat{a}_{v/t} \\ 0 \end{pmatrix}, \begin{pmatrix} \hat{a}_{v/t} & 0 \\ 0 & Q_t \end{pmatrix}\right)$, $\forall i = 1, \dots, K$.

Nas recursões acima, K é um número natural “suficientemente grande”⁵, de forma que as médias amostrais assumam valores relativamente próximos das médias teóricas. Pela *Lei Forte dos Grandes Números*, as seqüências de médias amostrais convergem quase certamente para as médias verdadeiras, à medida que K tende a infinito⁶ (é importante lembrar que, como decorrência imediata do pressuposto admitido na observação 1, dada logo após da Definição 6.1, de integralidade incondicional de todas as variáveis aleatórias e de possíveis transformadas das mesmas envolvidas no modelo, todas as médias e as

⁵ Eventualmente, a segunda expressão de (6.14) referente a $\hat{P}_{v/t}$ pode gerar uma matriz negativa definida! A explicação para tal inconveniente é que \hat{F}_t e $\hat{M}_{v/t-1}$ são provenientes de simulações, o que pode gerar evidentes más aproximações das verdadeiras matrizes. Logo, em situações práticas, K talvez tenha que ser realmente “muito grande”.

⁶ Observe-se que não são mais necessários os pressupostos adicionais de diferenciabilidade de primeira ou de segunda ordem para as funções mensuráveis Z_t e T_t requeridos pela adoção dos filtros anteriores.

covariâncias teóricas *condicionais* existem e são finitas quase certamente). Logo, K-assintoticamente, e *ignorando-se os problemas 3 e 4 restantes em (6.9)*, o Filtro de Simulação de Monte Carlo se constitui em um estimador não tendencioso e consistente do vetor de estado – tanto em termos de previsão um passo à frente, quanto em termos de atualização.

Pode-se, então, sumarizar a implementação do Filtro de Simulação de Monte Carlo como se segue:

1º Passo) Dado um modelo em EE não linear satisfazendo a Definição 6.1, adotar a aproximação linear do mesmo proposta em (6.10); e

2º Passo) Para cada instante de tempo t , aplicar as recursões em (6.13) e em (6.14) às medidas e às quantidades geradas no decorrer das mesmas.

Assim como nos outros dois filtros discutidos anteriormente (Filtro de Kalman Estendido e Filtro de Segunda Ordem), o processo recursivo necessita de quantidades iniciais a_1 e P_1 .

Também é importante ser dito que, quando H_t e/ou Q_t forem nulas para um ou mais instantes de tempo t [exemplo imediato: adoção de um modelo genuinamente enquadrado na formulação dada em (6.2) que é convertido, por intermédio da Proposição 6.2, na formulação dada em (6.1), procedimento este que induz $H_t = 0$ para todo t], o Filtro de Simulação de Monte Carlo se simplificaria, uma vez que sorteios de ε_t e/ou η_t , vetores aleatórios degenerados centrados em tais situações, não seriam mais necessários.

6.4.2

Propriedades do Filtro de Simulação de Monte Carlo

Como vantagem principal deste filtro não linear aproximado, cita-se que há menos fontes de viés em relação ao Filtro de Kalman e o Filtro de Segunda Ordem, uma vez que a aproximação linear abordada pelo mesmo apresenta bem menos problemas que aquelas baseadas em expansões de Taylor. Também foi examinado na Proposição 6.6 que essa aproximação linear é ainda mais interessante quando a não linearidade do modelo se faz presente apenas na

equação do estado. Outra possível conveniência do método é o fato de que não são demandadas diferenciabilidades de quaisquer ordens para as funções Z_t e T_t .

Como desvantagens, enumeram-se: (1) dificuldade de inicialização que, salvo maior juízo a respeito de possíveis conjugações *priori* imprópria/*posteriori* própria, deve ser não difusa; (2) há um esforço computacional considerável⁷, mesmo que ainda menor do que aquele referente à abordagem exata a ser discutida no capítulo 8 desta Dissertação; e (3) há de se supor que os problemas restantes não sanados em (6.9) possam ser desprezados para que o filtro apresente eficácia de fato.

⁷ Este problema pode ser atenuado através da adoção de variáveis antitéticas.