

### 3

## Modelos em espaço de estado e o filtro de Kalman

### 3.1. O modelo em espaço de estado linear

Um *modelo em espaço de estado linear*, sem efeitos de regressão e/ou intervenções, é definido pelas *matrizes do sistema*  $\{Z_t, T_t, R_t, H_t, Q_t\}$ . Suas equações são descrito abaixo:

$$\begin{aligned} Y_t &= Z_t \gamma_t + \varepsilon_t, \quad \varepsilon_t \sim WN(0, H_t) \\ \gamma_{t+1} &= T_t \gamma_t + R_t \eta_t, \quad \eta_t \sim WN(0, Q_t) \\ \gamma_1 &\sim (a_1, P_1) \end{aligned} \quad (1)$$

A primeira equação do sistema é chamada *equação das medidas* ou *das observações*. A segunda equação, nomeada por *equação de transição* ou *do estado*, representa a dinâmica do processo  $\gamma_t$ , chamado de *vetor de estado*, o qual é m-variado e, em geral, não observável, e cujo componente  $\gamma_1$ , chamado de *vetor de estado inicial*, possui esperança matemática e matriz de covariância incondicionais dados por  $a_1$  e  $P_1$ , respectivamente. Quando  $\gamma_1$  tem distribuição Gaussiana e é independente de processo  $(\varepsilon_t', \eta_t')$  e, além disso,

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim NID \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} H_t & 0 \\ 0 & Q_t \end{bmatrix} \right) \quad (2)$$

então, o modelo (1) é chamado modelo em *espaço de estado linear Gaussiano* (cf. Harvey, 1989; e Durbin e Koopman, 2001). Sobre as matrizes do sistema, pode ser dito que estas podem depender de termos passados de  $Y_t$  e/ou podem ser invariantes no tempo.

Para diversos modelos em séries temporais, existe pelo menos uma representação da forma espaço de estados. Dentre os principais, enumeram-se: ARIMA, VAR, modelos de regressão com coeficientes variante no tempo, modelos estruturais univariados e multivariados, modelos de volatilidade

estocástica, modelos de duração, modelos de *spline* etc. Aprofundamentos sobre como estes e outros modelos de séries temporais devem ser reconhecidos em formas do tipo descrito em (1) podem ser encontrados em Harvey (1989), Hamilton (1994), West e Harrison (1997), Durbin e Koopman (2001), Brockwell e Davis (2002), e Shumway e Stoffer (2006).

### 3.2. O filtro de Kalman

Para o que se segue, considere a versão Gaussiana<sup>6</sup> de (1). Um conjunto de equações recursivas para a obtenção de estimadores do vetor de estado, para cada instante de tempo e com boas propriedades, é o *filtro de Kalman*. Antes de que se apresentem as recursões do filtro de Kalman, considere-se a seguinte notação:

- $n$  é numero de observações extraídas do processo  $Y_t$ ;
- $\mathfrak{I}_j$  é a  $\sigma$ -álgebra produzida pelas medidas até o instante  $j$ , isto é,  $\mathfrak{I}_j \equiv \sigma(Y_1, Y_2, \dots, Y_j)$ ;
- $\hat{\gamma}_{t|j} \equiv E(\gamma_t | \mathfrak{I}_j)$ ; e
- $P_{t|j} \equiv E[(\gamma_t - \hat{\gamma}_{t|j})(\gamma_t - \hat{\gamma}_{t|j})' | \mathfrak{I}_j]$ .

As recursões do filtro de Kalman para os momentos em que  $j = t - 1$  (*previsão*),  $j = t$  (*atualização*) e  $j = n$  (*suavização*), são dadas, respectivamente, por:

- *Equações de Previsão*

$$\begin{aligned}\hat{\gamma}_{t+1|t} &= T_t \hat{\gamma}_{t|t} \\ P_{t+1|t} &= T_t P_{t|t} T_t' + R_t Q_t R_t'\end{aligned}\quad (3)$$

---

<sup>6</sup> Menciona-se que, embora esta hipótese possa ser um pouco restritiva, tudo que é aqui apresentado mantém grande generalidade, pois, fora de normalidade, expressões do filtro de Kalman representam estimadores lineares ótimos e correspondentes matrizes de erros médios quadráticos. Livros que perseguem esta leitura mais geral da teoria são Brockwell e Davis (2002) e Shumway e Stoffer (2006).

- *Equações de Atualização*

$$\begin{aligned}\hat{\gamma}_{t|t} &= \hat{\gamma}_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} v_t \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}\end{aligned}\quad (4)$$

- *Equações de Suavização*

$$\begin{aligned}\hat{\gamma}_{t|n} &= \hat{\gamma}_{t|t-1} + P_{t|t-1} r_{t-1} \\ r_{t-1} &= Z_t' F_t^{-1} v_t + (T_t - T_t P_{t|t-1} Z_t' F_t^{-1} Z_t)' r_t \\ P_{t|n} &= P_{t|t-1} - P_{t|t-1} N_{t-1} P_{t|t-1} \\ N_{t-1} &= Z_t' F_t^{-1} Z_t + (T_t - T_t P_{t|t-1} Z_t' F_t^{-1} Z_t)' N_t (T_t - T_t P_{t|t-1} Z_t' F_t^{-1} Z_t) \\ r_n &= 0 \quad e \quad N_n = 0\end{aligned}\quad (5)$$

Nota-se que, para o cálculo das recursões acima, é necessário que se calculem as *inovações* e sua variância condicional. Essas são obtidas por:  
 $v_t \equiv Y_t - E(Y_t | \mathfrak{I}_{t-1}) = Y_t - Z_t \hat{\gamma}_{t|t-1}$  e  $F_t \equiv \text{Var}(v_t | \mathfrak{I}_{t-1}) = Z_t P_{t|t-1} Z_t' + H_t$ .

Referências sobre as deduções das recursões de Kalman, sob a hipótese de normalidade, podem ser encontradas em Harvey (1989, 1993), e Durbin e Koopman (2001).

### 3.3. Imposição de restrições no filtro de Kalman

Conforme descrito em Pizzinga et al. (2009), o *filtro de Kalman restrito reduzido* (cf. Pizzinga, 2009) é uma das possibilidades de se impor a restrição de carteira – referente à análise de estilo semi-forte – aos modelos de fatores dinâmicos, na forma de espaço de estado.

Em termos gerais, e relegando os detalhes da mencionada particularização ao contexto da análise dinâmica de estilo ao capítulo 4 desta Dissertação, tem-se que esta versão do filtro de Kalman sob restrições consiste em reescrever algumas coordenadas do vetor de estado em função de outras. O resultado é substituído na equação das medidas de forma apropriada.

*Especificando:* Primeiramente, supõe-se que o vetor de estado, no instante  $t$ , satisfaz restrições lineares, da forma  $A_t \gamma_t = q_t$ , na qual  $A_t$  é uma matriz fixa e conhecida de dimensão  $k \times m$  e  $q_t$  é um vetor possivelmente aleatório, certamente

observável e  $\mathfrak{I}_t$ -*mensurável*<sup>7</sup>. A descrição do algoritmo, para a transferência de tais restrições à estimação do estado via filtro de Kalman, é dada abaixo, sob notações análogas a de Pizzinga et al. (2009):

1. Particione o vetor de estado e reescreva a restrição linear, sendo que a matriz  $A_{1,t}$   $k \times k$  é invertível:

$$A_{1,t}\gamma_{1,t} + A_{2,t}\gamma_{2,t} = q_t \quad (6)$$

2. Resolva a equação (6) para  $\gamma_{1,t}$  e obtenha:

$$\gamma_{1,t} = A_{1,t}^{-1}q_t - A_{1,t}^{-1}A_{2,t}\gamma_{2,t} \quad (7)$$

3. Substitua, na equação das medidas do modelo (1), a expressão de  $\gamma_{1,t}$  deduzida na equação (7):

$$Y_t^* = Y_t - Z_{1,t}A_{1,t}^{-1}q_t = (Z_{2,t} - Z_{1,t}A_{1,t}^{-1}A_{2,t})\gamma_{2,t} + \varepsilon_t = Z_{2,t}^*\gamma_{2,t} + \varepsilon_t \quad (8)$$

4. Proponha uma equação de transição para  $\gamma_{2,t}$  de tal forma que se preserve um modelo em espaço de estado linear, isto é:

$$\gamma_{2,t} = T_{2,t}\gamma_{2,t} + R_{2,t}\eta_{2,t} \quad (9)$$

5. Para o modelo definido anteriormente em (8) e (9), basta aplicar as equações de atualização e de suavização do filtro de Kalman para obter estimativas de  $\gamma_{2,t}$ . Finalmente, utilize as estimativas encontradas de  $\gamma_{2,t}$  para calcular estimativas de  $\gamma_{1,t}$ :

$$\hat{\gamma}_{1,t|j} = A_{1,t}^{-1}q_t - A_{1,t}^{-1}A_{2,t}\hat{\gamma}_{2,t|j} \quad \forall t \leq j \quad (10)$$

---

<sup>7</sup> O uso deste termo, neste contexto de modelos lineares Gaussianos, é equivalente a dizer

Uma vantagem metodológica de se utilizar o filtro de Kalman restrito reduzido é a de que, durante o processo de modelagem, é possível a comparação entre vários modelos concorrentes mediante os critérios de informação dos tipos AIC, BIC e outros. E, além disso, existe também uma óbvia vantagem operacional, a qual está relacionada à redução do tamanho do vetor estado. Isto pode, em muitos casos, gerar uma eficiência computacional durante a estimação do modelo. Para maiores discussões sobre esta maneira de se imporem restrições lineares em modelos em espaço de estado lineares, assim como comparações com outras abordagens concorrentes de estimação restrita, vide Pizzinga (2009).

### 3.4. Inicialização do filtro de Kalman

Para a utilização das recursões de Kalman, a esperança e a variância incondicionais  $a_1$  e  $P_1$  do vetor de estado inicial  $\gamma_1$  devem ser conhecidas. No caso em que existe pelo menos uma coordenada não-estacionária do vetor de estado  $\gamma_1$ , faz-se necessário o uso de procedimentos que auxiliem na inicialização do filtro de Kalman, uma vez que tais condições iniciais são de difícil definição. As abordagens mais conhecidas na literatura de séries temporais para inicialização do filtro de Kalman são: abordagem *big kappa* (cf. Harvey e Phillips, 1979), o *filtro de Kalman inicial exato* (cf. Koopman, 1997; Durbin e Koopman, 2001, cap. 5; e Koopman e Durbin, 2003) e o *filtro de Kalman aumentado* (cf. de Jong, 1991; Harvey, 1989; Durbin e Koopman, 2001, cap. 5; de Jong e Chu-Chun-Lin, 2003). Os dois primeiros métodos serão aqui descritos e, na parte aplicada da Dissertação (capítulo 5), considerados e implementados. Para tanto, adota-se o seguinte *modelo geral* para o vetor de estado inicial, sob a notação de Durbin e Koopman (2001):

$$\gamma_1 = a + A\delta + R_0\eta_0, \quad \eta_0 \sim N(0, Q_0) \quad (11)$$

---

que  $q_t$  é função afim de  $(Y_1', \dots, Y_t')$ .

no qual  $a$  é um vetor  $m \times 1$  de constantes,  $\delta$  é vetor aleatório  $q \times 1$  não-observável com  $E(\delta) = 0$ ,  $Var(\delta) = \kappa I_q$  e  $E(\delta \eta_0') = Cov(\delta \eta_0') = 0$ . Os termos  $A$  e  $R_0$  são matrizes que selecionam as coordenadas de  $\gamma_1$  associadas às componentes não-estacionárias e estacionárias, respectivamente, de  $\gamma_t$ . Também, admite-se que  $Q_0 > 0$  e as coordenadas de  $a$ , associadas às coordenadas não-nulas  $A\delta$ , são nulas. Com o modelo geral em (11), é possível atribuir total ignorância sobre as condições iniciais referentes às coordenadas não-estacionárias de  $\gamma_t$ . O procedimento é fazer com a variância de  $\delta$  se torne arbitrariamente grande.

Na *abordagem big kappa*, que foi sugerida originalmente em Harvey e Phillips (1979), e de certa forma defendida em Harvey (1989), págs: 121 e 122, a total ignorância é traduzida ao se especificar um valor numericamente grande para  $\kappa$ . No entanto, esta é uma metodologia heurística, sob a qual a determinação de uma constante de  $\kappa$  razoavelmente grande é variante conforme a exigência de cada analista, e cuja estabilidade computacional pode ficar seriamente comprometida.

Uma alternativa teórica e operacionalmente mais interessante à abordagem big kappa é o *filtro de Kalman inicial exato*, devido a Koopman (1997), e revisitado em outros aspectos por Koopman e Durbin (2003) – também se recomenda o livro de Durbin e Koopman (2001), cap. 5. Este tipo de inicialização difusa *exata* consiste basicamente em reconhecer que as equações do filtro de Kalman “2 em 1”<sup>8</sup> e do suavizador de Kalman (equações (5)) dependem de  $\kappa$  na forma de funções suaves, escrever estas funções como fórmulas de Maclaurin, e tomar  $\kappa \rightarrow \infty$  nestas expressões.

---

<sup>8</sup> O filtro de Kalman “2 em 1” é obtido mediante a substituição das equações em (4) nas equações em (3), e muitas vezes é considerado na literatura com o próprio filtro de Kalman (cf. Durbin e Koopman, 2001, cap. 4). Suas expressões são dadas por:

$$\begin{aligned}\hat{\gamma}_{t+1|t} &= T_t \hat{\gamma}_{t|t-1} + T_t P_{t|t-1} Z_t' F_t^{-1} v_t & t = 1, 2, \dots, n \\ P_{t+1|t} &= T_t P_{t|t-1} (T_t - T_t P_{t|t-1} Z_t' F_t^{-1})' + R_t Q_t R_t'\end{aligned}$$

Os referentes coeficientes matriciais da parte polinomial destas fórmulas de Maclaurin, no caso do filtro de Kalman “2 em 1”, constituem o filtro de Kalman inicial exato<sup>9</sup>; suas recursões são dadas na seqüência:

$$\begin{aligned}
 F_{\infty,t} &= Z_t P_{\infty,t} Z_t' & F_{*,t} &= Z_t P_{*,t} Z_t' + H_t \\
 M_{*,t} &= P_{*,t} Z_t' & M_{\infty,t} &= P_{\infty,t} Z_t' \\
 F_t^{(1)} &= F_{\infty,t}^{-1} & F_t^{(2)} &= -F_{\infty,t}^{-1} F_{*,t} F_{\infty,t}^{-1} \\
 K_t^{(0)} &= T_t M_{\infty,t} F_t^{(1)} & K_t^{(1)} &= T_t M_{\infty,t} F_t^{(2)} + T_t M_{*,t} F_t^{(1)} \\
 L_t^{(0)} &= T_t - K_t^{(0)} Z_t' & L_t^{(1)} &= -K_t^{(1)} Z_t' \\
 v_t^{(0)} &= Y_t - Z_t \hat{\gamma}_t^{(0)} & \hat{\gamma}_{t+1}^{(0)} &= T_t \hat{\gamma}_t^{(0)} + K_t^{(0)} v_t^{(0)} \\
 P_{\infty,t+1} &= T_t P_{\infty,t} L_t^{(0)'} & P_{*,t+1} &= T_t P_{\infty,t} L_t^{(1)'} + T_t P_{*,t} L_t^{(0)'} + R_t Q_t R_t'
 \end{aligned} \tag{12}$$

$$\hat{\gamma}_1^{(0)} = a, \quad P_{\infty,1} = AA' \quad P_{*,1} = R_0 Q_0 R_0'$$

É importante mencionar que as recursões em (12) coincidirão com as equações usuais do filtro de Kalman “2 em 1” a partir do momento em que  $P_{\infty,t}$  se anula. O último momento em que a referida matriz é não-nula é denotado por  $d$ , cuja existência é garantida teoricamente (cf. Koopman, 1997). Logo, a partir de  $d+1$ , o filtro de Kalman inicial exato em (12) deve ser substituído pelo filtro de Kalman “2 em 1”, ou pelas equações de atualização e de previsão, dadas respectivamente em (3) e (4), com as condições iniciais  $P_{d+1} = P_{*,d+1}$  e  $a_{d+1} = \hat{\gamma}_{d+1}^{(0)}$ .

De forma análoga ao procedimento adotado para obter o filtro de Kalman inicial exato, os coeficientes matriciais da parte polinomial das fórmulas de Maclaurin, para as equações de suavização em (3), são dados, no formato de recursões<sup>10</sup>, por:

---

<sup>9</sup> Nos instantes em que  $F_{\infty,t} = 0$ , estas recursões possuem formatos diferentes para alguns termos. Como tais casos nunca aparecerão nos modelos em espaço de estado de interesse desta Dissertação, as versões modificadas não serão aqui apresentadas. Para aprofundamentos, o leitor interessado é convidado a estudar uma referência bem completa no assunto, qual seja, Koopman e Durbin (2003).

<sup>10</sup> Mesmo conteúdo da nota anterior.

$$\begin{aligned}
r_{t-1}^{(0)} &= L_t^{(0)'} r_t^{(0)} \\
r_{t-1}^{(1)} &= Z_t' F_t^{(1)} v_t^{(0)} + L_t^{(0)'} r_t^{(1)} + L_t^{(1)'} r_t^{(0)} \\
N_{t-1}^{(0)} &= L_t^{(0)'} N_t^{(0)} L_t^{(0)} \\
N_{t-1}^{(1)} &= Z_t' F_t^{(1)} Z_t + L_t^{(0)'} N_t^{(1)} L_t^{(0)} + L_t^{(1)'} N_t^{(0)} L_t^{(0)} + L_t^{(0)'} N_t^{(0)} L_t^{(1)} \\
N_{t-1}^{(2)} &= Z_t' F_t^{(2)} Z_t + L_t^{(0)'} N_t^{(2)} L_t^{(0)} + L_t^{(0)'} N_t^{(1)} L_t^{(1)} + L_t^{(1)'} N_t^{(0)} L_t^{(1)} + L_t^{(1)'} N_t^{(1)} L_t^{(0)} \\
\hat{\gamma}_{t|n} &= \hat{\gamma}_t^{(0)} + P_{*,t} r_{t-1}^{(0)} + P_{\infty,t} r_{t-1}^{(1)} \\
\hat{P}_{t|n} &= P_{*,t} - (P_{*,t} N_{t-1}^{(0)} P_{*,t} + P_{*,t} N_{t-1}^{(1)} P_{\infty,t} + P_{\infty,t} N_{t-1}^{(1)} P_{*,t} + P_{\infty,t} N_{t-1}^{(2)} P_{\infty,t}) \\
r_d^{(0)} &= r_d, \quad N_d^{(0)} = N_d, \quad r_d^{(1)} = 0, \quad N_d^{(1)} = N_d^{(2)} = 0.
\end{aligned} \tag{13}$$

Conforme a garantia teórica apresentada por Koopman (1997), as recursões em (13) devem ser implementadas somente para os instantes  $t = d, \dots, 1$ . Para os outros instantes de tempo, as recursões do suavizador exato coincidirão plenamente com equações revisitadas em (4), com as inicializações em  $d+1$  já discutidas.

### 3.5.

#### Estimação de parâmetros por (quasi) máxima verossimilhança

Como pode ser facilmente percebido nas seções anteriores, o cálculo das recursões de Kalman demanda o conhecimento das matrizes do sistema. Estas, por sua vez, podem depender de parâmetros que na prática são desconhecidos. Assim, as recursões de Kalman devem ser combinadas com a estimação dos parâmetros via *(quasi) máxima verossimilhança*. Denota-se por  $\psi$  o vetor com todos os parâmetros a serem estimados. A função de verossimilhança do modelo apresentado na eq. (1), para uma série temporal observada do processo  $Y_t$  para  $t = 1, 2, \dots, n$ , e avaliada em  $\psi$ , é dada por:

$$L(\psi) \equiv p(y_1, \dots, y_n) = p(y_1) \prod_{t=2}^n p(y_t | \mathfrak{I}_{t-1}). \tag{14}$$

O estimador de *(quasi) máxima verossimilhança* é definido por  $\hat{\psi} = \arg \max_{\psi \in \Psi} L(\psi)$ . Conforme Harvey (1989), para o modelo em espaço de estado Gaussiano, pode-se escrever o logaritmo da equação descrita na eq. (14) apenas em termos de  $v_t = v_t(\psi)$  e  $F_t = F_t(\psi)$ , a qual é chamada, neste contexto,



de *verossimilhança decomposta pelo erro de predição*. Sua expressão, em log, sob uma inicialização não-difusa do filtro de Kalman, é dada por:

$$\log L(\psi) \propto \sum_{t=1}^n [\log |F_t| + v_t' F_t^{-1} v_t], \quad (15)$$

já que  $Y_t \sim N(Z_t \gamma_{t|t-1}, Z_t P_{t|t-1} Z_t' + H_t)$ . Por outro lado, sob a abordagem do filtro de Kalman inicial exato, as parcelas da função de log-verossimilhança em (15), correspondentes aos instantes  $t = 1, \dots, d$ , sofrem algumas modificações. Conforme demonstrado em Durbin e Koopman (2001) e utilizando o insumo adicional  $F_{\infty,t} = F_{\infty,t}(\psi)$  resultante do filtro de Kalman inicial exato, a função de *verossimilhança difusa*<sup>11</sup>, já em log, é dada por:

$$\log L(\psi) = -\frac{np}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^d \log |F_{\infty,t}| - \frac{1}{2} \sum_{t=d+1}^n [\log |F_t| + v_t' F_t^{-1} v_t] \quad (16)$$

Como observado nas equações (15) e (16), o processo de estimação via máxima verossimilhança demanda o uso de métodos numéricos devido à não-linearidade nos parâmetros. Este processo na prática pode gerar problemas de convergência numérica. Entretanto, a dificuldade em se obter convergência poderá ser mitigada ao se aplicar a versão difusa em (16). Evidências empíricas sobre esta última afirmação serão apresentadas durante o capítulo 5, destinado inteiramente às aplicações desta Dissertação.

### 3.6. Diagnósticos e seleção de modelos

Assim como outras metodologias estatísticas, a modelagem em espaço de estado abarca o uso de critérios e de medidas que auxiliam no processo de diagnósticos e na seleção de modelos.

---

<sup>11</sup> Mais uma vez, ressalta-se que, para modelos que contemplam instantes, no período difuso (ou seja, até  $d$ ), nos quais  $F_{\infty,t} = 0$ , esta verossimilhança possui algumas modificações. Para detalhes, vide Durbin e Koopman (2001), cap. 7.

No que concerne à prática de diagnósticos, as *inovações padronizadas*, definidas por  $v_t^p = F_t^{-1/2} v_t$ , são a matéria-prima essencial, pois, sob as hipóteses básicas do modelo em espaço de estado linear Gaussiano, estas constituem vetores aleatórios independentes e identicamente distribuídos, com distribuição comum  $N(0, I)$ . Violações de pelo menos uma dessas últimas duas condições indicam má especificação do modelo adotado. Como ferramentas que são comumente utilizadas nessa verificação, citam-se as diversas análises gráficas: gráfico de linha da inovação padronizada no tempo, funções de auto-correlação (FAC), Q-Q Plot etc. Além disso, o uso de estatísticas descritivas com as inovações padronizadas, tais como a de *Durbin-Watson*, a média e a variância, costuma ser, embora muito simples, útil para a averiguação de pressupostos básicos também. Por fim, também são recomendados, como procedimentos atenuantes da subjetividade, o teste *F* para heterocedasticidade, o teste *Ljung-Box* para correlação serial e os testes *Anderson-Darling* e/ou *Jarque-Bera* para normalidade.

Em caso de existirem evidências de que as inovações padronizadas são não-Gaussianas, as expressões dadas em (15) e (16) ainda costumam ser usadas como funções de *quasi log verossimilhança*, e os estimadores/estimativas, resultantes de maximização destas, são denominados estimadores/estimativas de *quasi máxima verossimilhança*<sup>12</sup>.

Quanto ao processo de seleção de modelos, Harvey (1989, cap 5) e Durbin e Koopman (2001, cap. 7) recomendam procedimentos para auxiliar a seleção de modelos e para que se verifique a qualidade do ajuste; dentre elas destacam-se:

- O valor encontrado para o logaritmo da verossimilhança, cujas possíveis versões são descritas nas equações (15) e (16);
- Critérios de Informação, tais como AIC e BIC:

$$\begin{aligned} AIC &= \frac{1}{n} [-2 \log L(\hat{\psi}) + 2(q + w)] \\ BIC &= \frac{1}{n} [-2 \log L(\hat{\psi}) + (q + w) \log(n)], \\ \text{nos quais } w &= \dim(\psi) \quad e \quad q = \dim(\delta) \end{aligned} \tag{17}$$

<sup>12</sup> No entanto, cita-se que, mesmo não sendo algo determinante para esta Dissertação em particular, alguns cuidados no que dizem respeito à inferência estatística devem ser tomados (cf. Hamilton, 1994, cap. 13).

- Medidas para avaliar o poder preditivo, tais como o *Pseudo- $R^2$*  e EQM, dados respectivamente por:

$$\begin{aligned} Pseudo-R^2 &= [corr(Y_t, \hat{Y}_{t|t-1})]^2 \\ EQM &= \frac{1}{n-d} \sum_{t=d+1}^n (Y_t - \hat{Y}_{t|t-1})^2 \end{aligned} \quad (18)$$