

Rita Cristina Pinto de Sousa

***Estimação de estado
em modelos lineares por bocados
parcialmente observados :
comparação numérica de dois testes***



**DEPARTAMENTO DE MATEMÁTICA APLICADA
FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DO PORTO**

Dezembro de 1999

Rita Cristina Pinto de Sousa

***Estimação de estado
em modelos lineares por bocados
parcialmente observados :
comparação numérica de dois testes***



Tese submetida à Faculdade de Ciências da Universidade do Porto
para obtenção do grau de Mestre em Estatística

**DEPARTAMENTO DE MATEMÁTICA APLICADA
FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DO PORTO**

Dezembro de 1999

Dissertação realizada sob a orientação
da Professora Doutora Paula Milheiro Oliveira

Aos meus Pais e Irmã

Agradecimentos

À Professora Doutora Paula Milheiro desejo agradecer pela simpatia e disponibilidade demonstradas na orientação da tese e pelas suas sugestões e apreciações cuidadas que se revelaram extremamente úteis no desenrolar deste trabalho.

A todos os Professores do Mestrado em Estatística expresso o meu agradecimento pela disponibilidade para qualquer esclarecimento solicitado.

Aos meus colegas de Mestrado agradeço pelo companheirismo e entre ajuda demonstrados.

À minha família, em especial aos meus Pais e Irmã, quero agradecer pelo apoio que me deram e pela compreensão que demonstraram nos momentos de menor disponibilidade.

Resumo

O problema de estimação de estado em modelos lineares por bocados parcialmente observados em tempo discreto e com pequeno ruído de observação é tratado nesta tese. Assumindo hipóteses de detectabilidade é possível separar os intervalos de tempo de linearidade para a função de observação. Sobre cada um dos intervalos assim separados podemos aproximar o filtro óptimo pelo correspondente filtro de Kalman-Bucy. Na separação de intervalos de linearidade, o problema consiste em decidir qual dos filtros deve ser aplicado. Dois tipos de testes são propostos para este fim: teste da razão de verosimilhança e testes de variação quadrática. Neste trabalho apresentamos e comparamos estes testes com base em simulações numéricas.

Abstract

This thesis regards a piecewise linear discrete time filtering problem with small observation noise. Under some detectability assumption, it is possible to separate the time intervals of linearity of the observation function. Over such an interval, we can then approximate the optimal filter by the corresponding Kalman-Bucy filter. When separating the intervals of linearity, the problem is to decide which filter must be applied. Likelihood ratio and quadratic variation tests are proposed for this purpose. We present and compare both tests numerically, based on simulation results.

Índice

Índice de figuras	7
Índice de tabelas	7
Notação	8
1. Introdução	10
2. Formulação do problema	12
3. Dois estimadores de Kalman-Bucy em paralelo	17
4. Teste de detecção de passagens em zero	20
5. Testes de decisão sobre o sinal do processo não observado	25
5.1 – Teste da razão de verosimilhança	26
5.2 – Teste de variação quadrática	30
6. Exemplos de aplicação	37
6.1 – Teste de detecção de passagens em zero	39
6.2 – Testes de decisão sobre o sinal	40
7. Conclusões	50
Apêndice A	52
Apêndice B	59
Bibliografia	74

Índice de figuras

Figura 6.1 (teste da razão de verosimilhança com l_1 e l_2 fixos)	41
Figura 6.2 (teste de variação quadrática com $l_1=0.0452$ e $l_2=0.4378$)	42
Figura 6.3 (teste de variação quadrática com $l_1=0.9898$ e $l_2=1.1333$)	42
Figura 6.4 (vantagens e desvantagens do teste da razão de verosimilhança)	46
Figura 6.5 (vantagens e desvantagens do teste de variação quadrática)	47
Figura 6.6 (teste de variação quadrática, hipótese (HD1), caso 1')	48
Figura A1 (nº médio de instantes para decidir nos TV e TVQ, caso 1)	53
Figura A2 (nº médio de instantes para decidir nos TV e TVQ, caso 3)	54
Figura A3 (nº médio de instantes para decidir nos TV e TVQ, caso 10)	55

Índice de tabelas

Tabela 6.1 (resultados para o teste de detecção de passagens em zero)	39
Tabela 6.2 (TV e TVQ com diferentes valores para a função $b(x_k)$)	44
Tabela 6.3 (TV e TVQ com diferentes valores para $h(x_k)$ e $\sigma(x_k)$)	45
Tabela 6.4 (resultados para o TV sob a hipótese (HD1))	48
Tabela A1 (resultados para os testes de decisão sobre o sinal, hipótese (HD2))	56
Tabela A2 (resultados para o teste de variação quadrática, hipótese (HD1))	57
Tabela A3 (resultados para o teste da razão de verosimilhança, hipótese (HD2))	58

Notação

Δt : passo de tempo utilizado na discretização do modelo de Markov em tempo contínuo.

(HD1) : hipótese de detectabilidade constituída pela condição $h_-^2 \sigma_-^2 \neq h_+^2 \sigma_+^2$.

(HD2) : hipótese de detectabilidade constituída pela condição $h_-^2 \sigma_-^2 = h_+^2 \sigma_+^2 \wedge b_+ \neq b_-$.

T : extremo superior do intervalo de tempo contínuo para o modelo considerado.

$K = \lceil T/\varepsilon \rceil$: número máximo de instantes de simulação resultante de uma discretização com passo de tempo $\Delta t = \varepsilon$.

$\gamma_0^k = \sigma\{y_0, y_1, \dots, y_k\}$: σ -álgebra das observações até ao instante k .

A_- : acontecimento que corresponde a ter todas as variáveis de estado com sinal negativo num dado intervalo de tempo.

A_+ : acontecimento que corresponde a ter todas as variáveis de estado com sinal positivo num dado intervalo de tempo.

N : representa a lei de distribuição gaussiana.

N^- : tempo de espera no caso negativo, isto é, número de instantes necessário para decidir sobre o sinal quando ocorre A_- .

N^+ : tempo de espera no caso positivo, isto é, número de instantes necessário para decidir sobre o sinal quando ocorre A_+ .

$-l_1, l_2$: valores críticos da estatística no teste de decisão sobre o sinal.

L_n^ε : estatística do teste da razão de verosimilhança no n -ésimo instante de um intervalo em que o teste se aplica.

S_n : estatística do teste de variação quadrática no n -ésimo instante de um intervalo em que o teste se aplica.

$O(\varepsilon^p)$: ordem de grandeza de ε^p (estocástica).

$a \wedge b$: mínimo entre a e b .

TD : teste de detecção.

TV : teste da razão de verosimilhança .

TVQ : teste de variação quadrática.

α_d : nível de significância para o teste de detecção de passagens em zero.

α_s : nível de significância para o teste de decisão sobre o sinal.

$-l_1^v, l_2^v$: valores críticos para a estatística do teste da razão de verosimilhança na decisão sobre o sinal.

$-l_1^{vq}, l_2^{vq}$: valores críticos para a estatística do teste de variação quadrática na decisão sobre o sinal.

pe_d : percentagem de erro na detecção de passagens em zero.

pe_v : percentagem de erro no teste da razão de verosimilhança, na decisão sobre o sinal.

pe_{vq} : percentagem de erro no teste de variação quadrática, na decisão sobre o sinal.

pi_d : percentagem de instantes onde é aplicado o teste de detecção.

pid_v : percentagem de instantes de decisão para o teste da razão de verosimilhança.

pid_{vq} : percentagem de instantes de decisão para o teste de variação quadrática.

ET_v^-, ET_v^+ : médias teóricas do tempo necessário para decisão no teste de verosimilhança.

\bar{T}_v^-, \bar{T}_v^+ : médias (empíricas) do tempo necessário para decisão no teste de razão de razão de verosimilhança.

ET_{vq}^-, ET_{vq}^+ : médias teóricas do tempo necessário para decisão no teste de variação quadrática.

$\bar{T}_{vq}^-, \bar{T}_{vq}^+$: médias (empíricas) do tempo necessário para decisão no teste de variação quadrática.

i_0 : instante a partir do qual se começa a contabilizar a estatística do teste de decisão.

$$S^1 = \frac{1}{100 \times 1500} \sum_{i=1}^{100} \sum_{j=0}^{1500} |x_{ij} - \hat{x}_{ij}| .$$

$$S^2 = \frac{1}{100 \times 1500} \sum_{i=1}^{100} \sum_{j=0}^{1500} |x_{ij} - \hat{x}_{ij}|^2 .$$

1. Introdução

Neste trabalho propomo-nos efectuar o estudo da estimação das variáveis de estado para um modelo de Markov escondido não linear em tempo discreto, em que as funções que o descrevem são lineares por bocados. Tomamos, portanto, como ponto de partida um modelo de Markov escondido (HMM - *Hidden Markov Model*), isto é, um processo estocástico, neste caso em tempo discreto, constituído por uma sequência de estados, não observáveis, que representa o processo de Markov e por uma outra sequência de variáveis aleatórias condicionalmente independentes, que representa o processo das observações. A estimação das variáveis de estado, em cada instante, é feita à custa das observações até esse mesmo instante, procedimento este que é habitualmente designado por filtragem. Nos modelos em que as funções intervenientes são funções reais lineares podemos utilizar o filtro de Kalman-Bucy para estimar o estado em cada instante e respectiva covariância de erro de estimação. Nos modelos não lineares utiliza-se geralmente uma solução aproximada para o problema de estimação (Andrews-Grewal, 1993, Harvey, 1989 e Jazwinski, 1970, entre outros). De entre as aproximações possíveis o chamado Filtro de Kalman Estendido é uma das mais usadas pelas suas propriedades (Picard, 1991). Exemplos da aplicação do Filtro de Kalman Estendido em modelos não lineares podem ser encontrados por exemplo em Sorenson (1990). Existem diversos tipos de modelos não lineares, como é o caso dos modelos de limiar (“threshold”), modelos aqui em estudo. Tong-Yeung (1989), Tong (1990) e Turkman (1998) fazem uma abordagem destes modelos numa perspectiva de séries temporais. Nesta tese vamos tratar o caso em que as funções que descrevem o modelo são lineares por bocados. A solução deste problema, como veremos, passa pelo cálculo de filtros em paralelo correspondentes aos ramos das funções respeitantes às regiões possíveis para o estado. Para produzir as estimativas torna-se necessário decidir sobre a região onde se situam as variáveis de estado. Para isso há primeiro que isolar os intervalos de tempo em que o processo de estado não muda de região. Sem perda de generalidade suporemos que estas regiões são \mathcal{R}^- e \mathcal{R}^+ , que estão portanto separadas pelo número zero. Então, a estimação deverá ser precedida por dois testes, um para a detecção de passagens em zero e outro para a decisão sobre o sinal do estado, nos intervalos sem cruzamentos em zero. No que respeita ao primeiro teste, temos à escolha entre dois tipos

de teste: um teste sobre as observações ou um teste sobre o estimador de Kalman-Bucy. Na decisão sobre o sinal do estado podemos também optar entre dois testes, o da razão de verosimilhança e o de variação quadrática. Os dois tipos de testes referidos para a decisão sobre o sinal podem ainda ser aplicados ao processo das observações ou às próprias estimativas de Kalman-Bucy. Quer os testes de passagens em zero quer os testes de decisão sobre o sinal foram estudados em Fleming et al. (1991), Oliveira-Roubaud (1991) e Oliveira (2000) para modelos discretos e em Fleming-Pardoux (1989), Fleming et al. (1988), Pardoux-Roubaud (1991) e Roubaud (1992 e 1995) para os análogos modelos contínuos. Neste trabalho usaremos apenas os testes de detecção de passagens em zero e decisão sobre o sinal baseados nas observações.

Como veremos estes métodos só podem ser aplicados se se verificar uma hipótese, dita hipótese de detectabilidade. Esta subdivide-se em dois casos distintos: (HD1) e (HD2), para os quais os testes assumem aspectos diferentes.

O principal objectivo deste trabalho consiste em analisar e comparar o comportamento destes dois testes de decisão sobre o sinal baseados nas observações (razão de verosimilhança e variação quadrática) na resolução do problema de estimação, face à hipótese de detectabilidade (HD2). Uma breve abordagem à primeira hipótese (hipótese (HD1)) é usada apenas em termos comparativos.

A análise comparativa do desempenho dos testes e os resultados numéricos apresentados baseiam-se na simulação de alguns modelos.

Este trabalho divide-se em 7 capítulos que se distribuem da seguinte forma. Após esta breve introdução formulamos o problema em estudo no capítulo 2. Os filtros de Kalman-Bucy em paralelo, usados na estimação, são apresentados no capítulo 3. Os capítulos 4 e 5 contêm o teste de detecção de passagens em zero e os testes de decisão sobre o sinal, respectivamente. Neste último capítulo são tratados o teste da razão de verosimilhança (secção 5.1) e o teste de variação quadrática (secção 5.2). Apresenta-se ainda uma pequena abordagem à hipótese de detectabilidade (HD1) e ao teste de variação quadrática com esta hipótese, no final da secção 5.2. No capítulo 6 são apresentados e analisados os resultados numéricos obtidos em alguns modelos, em que o método exposto foi aplicado. Por fim são apresentadas conclusões sobre o desempenho dos testes (capítulo 7). Detalhes sobre os resultados numéricos dos testes podem ser analisados com a ajuda das tabelas e figuras que foram incluídas no apêndice A. No apêndice B são incluídas as listagens das rotinas desenvolvidas em linguagem de programação 'Matlab' e que permitiram obter os resultados numéricos referidos ao longo deste trabalho.

2. Formulação do problema

Consideremos o seguinte modelo não linear :

$$\begin{cases} X_{k+1} = X_k + \varepsilon b(X_k) + \sqrt{\varepsilon} \sigma(X_k) u_k, & k = 0, 1, 2, \dots, K \\ Y_k = h(X_k) + \sqrt{\varepsilon} v_k, \end{cases} \quad (2.1)$$

em que $\{X_k, k \geq 0\}$ e $\{Y_k, k \geq 0\}$ são processos estocásticos que tomam valores em \mathfrak{R} e correspondem aos processos de estado e de observação, respectivamente. O parâmetro ε é positivo e suposto pequeno e as sequências $\{u_k\}_k$ e $\{v_k\}_k$ são processos estocásticos que pretendem modelar perturbações, ruídos ou incertezas. No nosso caso, são considerados como ruídos brancos gaussianos normalizados e independentes de X_0 . Considera-se o intervalo de tempo contínuo $[0, T]$ que, para uma discretização com passo de tempo $\Delta t = \varepsilon$, dá origem a $K+1$ instantes em tempo discreto, com $K = \text{int}(T/\varepsilon)$. Os processos $\{X_k\}_k$ e $\{Y_k\}_k$ são assim indexados, nestes instantes, da maneira habitual.

No problema que pretendemos tratar suporemos que as funções b, σ e h são funções reais de variável real lineares por bocados :

- $b(x) = b_- x |_{\{x < 0\}} + b_+ x |_{\{x \geq 0\}}$
- $\sigma(x) = \sigma_- |_{\{x < 0\}} + \sigma_+ |_{\{x \geq 0\}}$
- $h(x) = h_- x |_{\{x < 0\}} + h_+ x |_{\{x \geq 0\}}$

O processo estocástico $\{(X_k, Y_k)\}_k$, modelado pelo sistema (2.1), pode ser visto como um modelo de Markov escondido, em que $\{X_k\}_k$ é um processo de Markov em espaço de estados infinito e o processo $\{Y_k\}_k$ corresponde a uma sequência de variáveis aleatórias, dependentes de $\{X_k\}_k$. A designação de modelo de Markov escondido é motivada pelo facto do processo de

estado $\{X_k\}_k$ não ser observado, mas existir um processo directamente correlacionado com o mesmo, $\{Y_k\}_k$, que representa as observações do sistema em estudo.

O problema consiste em estimar em cada instante k o estado X_k , através das observações $\{Y_n\}_{0 \leq n \leq k}$ recolhidas até esse instante. Como habitualmente, designaremos por filtragem a estimação de um valor de X_k (não observado) baseada num conjunto de observações $\{Y_1, \dots, Y_n\}$, para $k=n$. O estimador será designado por filtro.

Consideremos a σ -álgebra das observações até ao instante k , $\gamma_0^k = \sigma\{y_0, y_1, \dots, y_k\}$. Começemos por introduzir o estimador e respectiva covariância de erro, propostos por Kalman-Bucy para o modelo linear gaussiano multidimensional, que correspondem à média e matriz de covariância da lei condicional de X_k dado γ_0^k ,

$$\hat{X}_{k/k} = E[X_k / \gamma_0^k]$$

e

$$Q_{k/k} = E[(X_k - \hat{X}_{k/k})(X_k - \hat{X}_{k/k})^T / \gamma_0^k],$$

respectivamente. Quando o modelo é linear, o estimador, dito de Kalman-Bucy, é um estimador óptimo no sentido do erro quadrático médio (Jazwinski, 1970), desde que a condição inicial seja gaussiana (a lei condicional é também gaussiana), e é sub-óptimo, na classe de estimadores lineares, se esta condição não se verificar.

Não se verificando a linearidade das funções b, σ e h a determinação da lei condicional é em geral impraticável pelo que se adopta habitualmente um estimador aproximado dito Filtro de Kalman Estendido (Jazwinski, 1970, Picard, 1986 e Roubaud, 1995). No entanto, não nos devemos esquecer que embora o Filtro de Kalman Estendido seja em geral eficiente em muitas situações práticas, a sua eficiência não é garantida, dependendo das condições do modelo. Picard (1986) demonstrou que a eficiência se verifica se a função h for injectiva, para além de outras hipóteses de regularidade bastante gerais sobre as funções que intervêm no modelo. Como estamos a lidar com uma função de observação unidimensional linear por bocados, h não será necessariamente injectiva. Quando h é uma função injectiva (monótona), isto é no nosso caso $h_- h_+ > 0$, o problema de estimação não oferece portanto dificuldades. Podemos dizer que, nessa altura, se tivéssemos $\varepsilon=0$ então $X_k = h^{-1}(Y_k)$, isto é, X_k era conhecido de forma exacta. No caso de ε pequeno, teríamos

$X_k = \frac{Y_k}{h} - \sqrt{\varepsilon} v_k$, onde h tomaria o valor de h_- ou h_+ , consoante o seu sinal. Assim, o estimador $\tilde{X}_k = \frac{Y_k}{h}$, onde h representaria h_- ou h_+ consoante o sinal de Y_k , seria suficientemente bom para estimar X_k . Verifica-se que o estimador de Kalman Estendido, \hat{X}_k^E , com ε suficientemente pequeno, tem uma variância muito reduzida e converge para o estimador óptimo (Picard, 1986).

Quando h não é injectiva (não monótona), ou seja, no nosso caso se $h_- h_+ < 0$, as estimativas para as variáveis de estado não podem ser directamente obtidas pela estimação de $h(X_k)$. Nesta situação, o Filtro de Kalman Estendido não é em geral assintoticamente eficiente, sendo necessária a determinação da região onde $\{X_k\}_k$ se situa ou, se quisermos, do sinal de $\{X_k\}_{k \geq 0}$ em cada instante. Neste caso, dado que as funções no modelo são lineares por bocados, a ideia é a de considerar um estimador para cada um dos sistemas correspondentes aos (dois) ramos de linearidade das funções. Assim, mantêm-se dois filtros em paralelo, optando-se em cada instante pela aplicação de um ou de outro, consoante se toma a decisão de considerar que o processo evolui numa das regiões (\mathcal{R}^- ou \mathcal{R}^+). Esta decisão baseia-se na aplicação de testes sequenciais. É precisamente neste caso que nos situaremos nesta tese, isto é, admitimos a seguinte hipótese:

$$h_- h_+ < 0 \quad (\text{H1})$$

Naturalmente a questão da determinação do sinal do estado passa primeiro pela garantia de se ter uma variância condicional reduzida. Para isso, foram introduzidas duas hipóteses de detectabilidade (Fleming et al., 1991 e Oliveira-Roubaud, 1995). São elas :

$$h_-^2 \sigma_-^2 \neq h_+^2 \sigma_+^2 \quad (\text{HD1})$$

$$h_-^2 \sigma_-^2 = h_+^2 \sigma_+^2 \text{ e } b_+ \neq b_- \quad (\text{HD2})$$

Sob qualquer destas hipóteses é possível, à custa de dois estimadores de Kalman-Bucy em paralelo, obter um estimador convergente.

No caso da hipótese (HD1), o sinal de X_k é essencialmente determinado atendendo à variação quadrática da função $h(X_k)$. No caso da hipótese (HD2), é a função $b(X_k)$ que contém a informação essencial para a determinação do sinal de X_k , sendo observada

indirectamente através de Y_k . Por consequência veremos que, quando se verifica a primeira hipótese, a hipótese (HD1), o processo de estimação é mais rápido do que quando se verifica a segunda, a hipótese (HD2). Verificando-se a hipótese (HD1), a diferença entre as quantidades $\sigma_-^2 h_-^2$ e $\sigma_+^2 h_+^2$ permite ao teste distinguir mais rapidamente a região em que se encontram as variáveis de estado. Quando se verifica a hipótese (HD2), o teste necessita de um maior número de instantes de tempo para detectar as diferenças na estatística provocadas pelos parâmetros b_- e b_+ . Assim, a hipótese (HD2) torna-se mais delicada na obtenção de estimativas para o processo $\{X_k\}_k$ e é sob esta hipótese que se situarão os estudos e resultados apresentados ao longo deste trabalho.

A obtenção das estimativas para as variáveis de estado do modelo (2.1) passa por duas etapas :

1ª etapa : Detectar os intervalos de tempo em que a trajectória de $\{X_k\}_k$ não cruza o nível zero.

2ª etapa : Decidir para cada um desses intervalos se se tem $X_k > 0$ ou $X_k < 0$.

Sob as hipóteses de detectabilidade apresentadas é possível obter intervalos de tempo suficientemente longos, em que o processo $\{X_k\}_k$ toma valores distantes de zero. Nestes intervalos podemos aplicar os filtros de Kalman-Bucy, ora para um ramo ora para outro, de acordo com o sinal das variáveis de estado nesses intervalos.

Às duas etapas acima referidas correspondem testes estatísticos. Na primeira etapa, utiliza-se um teste de detecção de passagens em zero, que permite identificar os intervalos a que, para uma dada probabilidade, correspondem realizações das variáveis de estado que não cruzam o limiar zero. Após a determinação destes intervalos é necessário decidir qual dos filtros de Kalman-Bucy a aplicar em cada um desses intervalos. Para isso é necessário conhecer a região (semirecta) em que o estado se situa, ou por outras palavras, decidir sobre o sinal do estado. A estimativa a reter é então a que provém do sistema correspondente. Na decisão sobre o sinal consideraremos dois tipos de testes: testes de razão de verosimilhança e testes de variação quadrática.

A hipótese de detectabilidade (HD1) foi já objecto de vários estudos incluindo os testes de detecção de passagens em zero e da decisão sobre o sinal (ver Fleming et al., 1991). Tal como já referimos na introdução, neste estudo limitar-nos-emos na análise da hipótese (HD2), sendo a primeira hipótese utilizada apenas em termos comparativos em exemplos de

verificação do comportamento do teste de variação quadrática (ver capítulo 6). No que diz respeito ao teste de detecção de passagens em zero, estudos anteriores mostram que é possível adoptar o teste proposto por Fleming et al (1991) sob a hipótese (HD1) ao problema sob a hipótese (HD2), sem dificuldade (ver Oliveira-Roubaud, 1995). No que diz respeito ao teste de decisão sobre o sinal Oliveira-Roubaud (1991) propõem, sob esta última hipótese, dois testes da razão de verosimilhança e Oliveira (2000) propõe um teste de variação quadrática. O principal objectivo deste estudo consiste em analisar o comportamento do teste de variação quadrática e compará-lo com o teste da razão de verosimilhança, baseados nas observações, face à hipótese de detectabilidade (HD2).

Convém notar que o modelo (2.1) resulta da discretização do modelo contínuo

$$\begin{cases} d\tilde{X}_t = b(\tilde{X}_t)dt + \sigma(\tilde{X}_t)dU_t, & 0 \leq t \leq T \\ d\tilde{Y}_t = h(\tilde{X}_t)dt + \varepsilon dV_t, \end{cases} \quad (2.2)$$

com passo de tempo $\Delta t = \varepsilon$, no intervalo $[0, T]$. Neste modelo contínuo os processos $\{U_k, k \geq 0\}$ e $\{V_k, k \geq 0\}$ são processos de Wiener standard independentes, com valores em \mathcal{R} . O problema de estimação em observação parcial destes modelos contínuos foi tratado por diversos autores (Roubaud, 1995 e 1993, Picard, 1986, Fleming-Pardoux, 1989 e Fleming-Ji-Pardoux, 1988), sendo usado como ponto de partida para a construção dos estimadores e para comparações e verificação de propriedades para os modelos em tempo discreto como o modelo (2.1).

Neste trabalho, usaremos a seguinte notação para indicar que um determinado processo é da ordem de grandeza de ε^q .

Notação 2.1 – Para um dado processo $\{G_k\}$ considera-se que

$$G_k = O(\varepsilon^q),$$

onde $q \in \mathcal{R}_+$, para significar que existem $c_1, c_2, c_3 > 0$ tais que

$$E[G_k^2] \leq c_1 \exp\{-c_2 k \varepsilon\} + c_3 \varepsilon^{2q}, \forall k \in N, \varepsilon > 0.$$

3. Dois estimadores de Kalman-Bucy em paralelo

Relembramos que o problema em questão neste trabalho consiste em estimar o processo de estado não observado $\{X_k\}_k$ em cada instante k , dadas as observações até esse mesmo instante. O estimador óptimo, no sentido da média quadrática, é dado pela esperança condicional $E[X_k | \gamma_0^k]$ (Jazwinski, 1970). No caso em que o modelo é um sistema linear gaussiano, a lei condicional também é gaussiana e a sua média e variância são suficientes para caracterizar essa distribuição, como é bem conhecido, constituindo o filtro proposto por Kalman e Bucy nos anos 60 (Kalman-Bucy, 1961). A necessidade motivada pelas aplicações práticas levou a que a teoria da filtragem linear fosse estendida aos sistemas não lineares dando origem ao filtro de Kalman Estendido (Andrews-Grewal, 1993 e Jazwinski, 1970 entre muitos outros). Assim a ideia da utilização deste estimador em sistemas não lineares consiste em considerar que, nas vizinhanças de uma estimativa, o sistema comporta-se localmente de forma aproximada à do sistema linear, de tal forma que, ao aplicar o filtro de Kalman-Bucy correspondente a esse sistema localmente linearizado produzimos estimativas próximas das óptimas no sentido do erro quadrático médio. No nosso caso temos um sistema unidimensional linear por bocados portanto, em cada um desses intervalos de linearidade, a ideia é aplicar o filtro correspondente ao ramo em causa consoante se decide que o processo de estado está num ramo ou no outro. Para tomar estas decisões, que antecedem a aplicação dos filtros, são usados os testes de decisão sobre o sinal, que iremos tratar no capítulo 5. Para detectar os intervalos de linearidade são usados os testes apresentados no capítulo 4.

Consideremos dois modelos lineares correspondentes aos ramos em que se separam as funções intervenientes no modelo (2.1). Referenciêmo-los como modelo (+) e modelo (-):

$$\begin{aligned}
(+)\quad & \begin{cases} X_{k+1} = (1 + \varepsilon b_+) X_k + \sigma_+ \sqrt{\varepsilon} u_k \\ Y_k = h_+ X_k + \sqrt{\varepsilon} v_k \end{cases} \\
(-)\quad & \begin{cases} X_{k+1} = (1 + \varepsilon b_-) X_k + \sigma_- \sqrt{\varepsilon} u_k \\ Y_k = h_- X_k + \sqrt{\varepsilon} v_k \end{cases}
\end{aligned}$$

Podemos dizer que se trata dos modelos resultantes da permanência do processo de estado em valores positivos ou valores negativos, respectivamente. A separação do modelo original nestes dois modelos dá origem ao aparecimento de dois filtros que terão de ser calculados em simultâneo, um para cada modelo. Daí a designação de filtros de Kalman-Bucy em paralelo.

Sejam (\hat{X}_k^+, Q^+) e (\hat{X}_k^-, Q^-) os filtros associados aos sistemas lineares (+) e (-):

$$\begin{cases} \hat{X}_{k/k}^+ = \hat{X}_{k/k-1}^+ + K_k^+ (Y_k - h_+ \hat{X}_{k/k-1}^+) \\ Q_{k/k}^+ = Q_{k/k-1}^+ - Q_{k/k-1}^+ h_+ (h_+ Q_{k/k-1}^+ h_+ + \varepsilon)^{-1} h_+ Q_{k/k-1}^+, \end{cases}$$

com $\hat{X}_{k/k-1}^+ = (1 + \varepsilon b_+) \hat{X}_{k-1/k-1}^+$, $K_k^+ = Q_{k/k-1}^+ h_+ (h_+ Q_{k/k-1}^+ h_+ + \varepsilon)^{-1}$ e

$$Q_{k/k-1}^+ = (1 + \varepsilon b_+)^2 Q_{k-1/k-1}^+ + \sigma_+^2 \varepsilon$$

e

$$\begin{cases} \hat{X}_{k/k}^- = \hat{X}_{k/k-1}^- + K_k^- (Y_k - h_- \hat{X}_{k/k-1}^-) \\ Q_{k/k}^- = Q_{k/k-1}^- - Q_{k/k-1}^- h_- (h_- Q_{k/k-1}^- h_- + \varepsilon)^{-1} h_- Q_{k/k-1}^-, \end{cases}$$

com $\hat{X}_{k/k-1}^- = (1 + \varepsilon b_-) \hat{X}_{k-1/k-1}^-$, $K_k^- = Q_{k/k-1}^- h_- (h_- Q_{k/k-1}^- h_- + \varepsilon)^{-1}$ e

$$Q_{k/k-1}^- = (1 + \varepsilon b_-)^2 Q_{k-1/k-1}^- + \sigma_-^2 \varepsilon.$$

As condições iniciais gaussianas que se consideram, são as seguintes :

- $\hat{X}_0^+ = \hat{X}_0^- = E(X_0)$
- $Q_0^+ = Q^+$ e $Q_0^- = Q^-$, onde Q^+ e Q^- são as variâncias estacionárias das leis condicionais.

As condições iniciais são importantes no sentido em que influenciam o comportamento inicial do modelo, o que poderá não ser acompanhado pelo estimador, mas o efeito de dependência da condição inicial tende a desaparecer rapidamente no tempo (Oliveira-Roubaud, 1991 e Picard, 1991).

Naturalmente, os resultados da estimação usando os filtros em paralelo aqui apresentados, serão tanto melhores quanto mais longos forem os intervalos de linearidade, ou seja, quanto mais prolongadas forem as permanências do processo de estado numa das semirectas \mathfrak{R}^- ou \mathfrak{R}^+ .

Podemos resumir o processo de estimação em três passos essenciais: a selecção dos intervalos de tempo em que um dos filtros conduz a uma boa estimativa (teste de detecção de passagens em zero, capítulo 4), a decisão de qual dos dois filtros se deve aplicar (teste de decisão sobre o sinal do processo não observado, capítulo 5) e por último, a estimação propriamente dita usando o estimador \hat{X}_k^+ ou \hat{X}_k^- , consoante o caso.

4. Teste de detecção de passagens em zero

Como já tivemos oportunidade de referir o procedimento para a obtenção de estimativas das variáveis de estado envolve dois tipos de testes. Inicialmente, é necessário seleccionar os intervalos de tempo, ditos de linearidade, em que, para uma probabilidade próxima de um, o processo de estado não cruza o zero. Após a selecção dos intervalos procede-se à decisão sobre o sinal que permite escolher o estimador a aplicar em cada um desses intervalos. Neste capítulo trataremos o primeiro tipo de testes, os testes de detecção de passagens em zero, sendo os de decisão sobre o sinal objecto de estudo do capítulo seguinte (capítulo 5).

O teste de detecção de passagens em zero pode ser aplicado às observações ou ao próprio estimador de $\{X_k\}_k$, $\{\hat{X}_k\}_k$, tendo-se neste caso optado por um teste baseado nas observações. No caso em que o teste de detecção é baseado no próprio estimador, o procedimento é basicamente o mesmo, excepto naturalmente no que diz respeito à estatística do teste. Para mais detalhes poderá ser consultado o trabalho desenvolvido por Fleming et al. (1991) e Oliveira-Roubaud (1991). Este teste pode ser aplicado quer sob a hipótese (HD2) quer sob a hipótese (HD1).

Nesta tese pretendemos essencialmente analisar o comportamento dos testes de decisão sobre o sinal, mas a aplicação destes últimos passa obrigatoriamente pelo teste de detecção de passagens em zero, que abordaremos neste capítulo e que passamos a introduzir.

Fixemos uma trajectória do processo $\{X_k\}_k$. No intervalo de tempo contínuo $[0, T]$, consideremos um intervalo $[a, b]$ em que essa trajectória não cruza o zero. A não existência de cruzamentos em zero corresponde à ocorrência de um dos seguintes acontecimentos:

$$A_- = \{ X_k < 0; k = i_0, i_0 + 1, \dots, i_0 + m \},$$
$$A_+ = \{ X_k > 0; k = i_0, i_0 + 1, \dots, i_0 + m \},$$

com $i_0 = \lceil a/\varepsilon \rceil$ e $m = \lceil (b-a)/\varepsilon \rceil$. Num abuso de linguagem, por uma questão de simplificação do texto, designaremos um intervalo de tempo correspondente aos índices i_0 a i_0+m por intervalo de tempo discreto $[i_0, i_0 + m]$. Detectar um destes intervalos corresponde a testar se um dos dois acontecimentos ocorre ou não. Perante uma dada trajectória o nosso objectivo será determinar um procedimento que permita detectar os intervalos nestas condições.

Este procedimento, a que chamaremos teste de detecção de cruzamentos em zero, usará uma estatística de teste que passaremos a construir. Ele consistirá na adopção de uma constante como valor crítico, a que chamaremos c_{obs} , e na verificação perante os dados se o valor da estatística está ou não na região crítica.

Se analisarmos o modelo em tempo discreto (2.1) constatamos que $h(X)=0$ sse $X=0$ e que $h(X_k)$ “pequeno” implica, por continuidade de h , X_k próximo de zero. No entanto $h(X_k)$ não é observável, mas para ε “pequeno”, aproxima-se de Y_k . Assim, Y_k é um candidato à partida como estatística do teste na detecção de passagens em zero para o processo $\{X_k\}_k$.

Desejavelmente a escolha de c_{obs} deveria resultar de um compromisso entre a manutenção de intervalos de monotonia suficientemente grandes e a redução da probabilidade de erro do teste. No entanto conhecemos as limitações que os testes de hipóteses por vezes nos impõem. Para além disso, os testes que aplicamos numa situação de observação parcial têm a dificuldade acrescida de serem baseados num processo observado ($\{Y_k\}_k$) quando se pretende decidir sobre um processo não observado ($\{X_k\}_k$).

A nossa preocupação, agora, na construção do teste será precisamente a determinação de uma fórmula para esta constante $c_{obs} > 0$, designada por alguns autores como “cutoff”. O teste baseia-se portanto na análise da ocorrência do seguinte acontecimento :

$$C = \left\{ |Y_k| \geq c_{obs} ; k = i_0, \dots, i_0 + m \right\} .$$

O Teorema que se segue (Fleming et al., 1991) permite-nos formar um primeiro juízo sobre a qualidade assintótica do teste.

Teorema 4.1 – Para $c_{obs} > 0$ e $\varepsilon_0 > 0$ dados, existe $\beta > 0$ tal que $\forall \varepsilon \in (0, \varepsilon_0]$ se tem

$$P \left(\left\{ (A_- \cup A_+)^c \mid C \right\} \right) \leq e^{-\beta/\varepsilon}.$$

Na obtenção de uma fórmula para a constante c_{obs} pretendemos minimizar a probabilidade de erro de que um cruzamento não seja detectado, a qual deverá corresponder ao valor máximo para a probabilidade condicional

$$P \left\{ \left\{ X_k \cdot X_{k+1} < 0 \right\} \mid |Y_k| \geq c_{obs}, |Y_{k+1}| \geq c_{obs} \right\},$$

nos instantes $k = i_0, \dots, i_0 + m$.

Dado que, sob a hipótese (H1) se tem $h_- h_+ < 0$ suponhamos, por um momento, que $h_- < 0$ e $h_+ > 0$, isto é, $h(x) \geq 0, \forall x$. O outro caso, $h_- > 0$ e $h_+ < 0$, pode ser tratado da mesma maneira. Dado que, se $h_- < 0$ e $h_+ > 0$, se tem

$$P \left(\left\{ Y_k < -c_{obs} \right\} \right) \leq P \left(\left\{ v_k < -c_{obs} / \sqrt{\varepsilon} \right\} \right) \leq \frac{1}{2} e^{-c_{obs}^2 / (2\varepsilon)},$$

e, tendo em conta o resultado do Teorema 4.1, preocupar-nos-emos apenas com o estudo do condicionamento pelo acontecimento $\{Y_k \geq c_{obs}, Y_{k+1} \geq c_{obs}\}$.

Procuremos avaliar a probabilidade condicionada começando por analisar as intersecções possíveis com o acontecimento $\{X_k \cdot X_{k+1} < 0\}$.

Definimos os acontecimentos :

$$D_c^1 = \{X_k < 0, X_{k+1} > 0\} \cap \{Y_k \geq c, Y_{k+1} \geq c\},$$

$$D_c^2 = \{X_k > 0, X_{k+1} < 0\} \cap \{Y_k \geq c, Y_{k+1} \geq c\},$$

$\forall c > 0$. Então

$$D_{c_{obs}}^1 \cup D_{c_{obs}}^2 = \{X_k \cdot X_{k+1} < 0\} \cap \{Y_k \geq c_{obs}, Y_{k+1} \geq c_{obs}\}$$

e pretendemos avaliar a probabilidade deste acontecimento.

Condicionando o modelo (2.1) pelo acontecimento D_c^1 , obtém-se :

$$\begin{cases} X_{k+1} = (1 + \varepsilon b_-) X_k + \sigma_- \sqrt{\varepsilon} u_k \\ Y_k = h_- X_k + \sqrt{\varepsilon} v_k, \quad Y_{k+1} = h_+ X_{k+1} + \sqrt{\varepsilon} v_{k+1}. \end{cases}$$

Então:

$$\begin{aligned} & h_- Y_{k+1} - h_+ (1 + \varepsilon b_-) Y_k = \\ & = h_- (h_+ X_{k+1} + \sqrt{\varepsilon} v_{k+1}) - h_+ (1 + \varepsilon b_-) (h_- X_k + \sqrt{\varepsilon} v_k) \\ & = h_- h_+ [(1 + \varepsilon b_-) X_k + \sigma_- \sqrt{\varepsilon} u_k] + h_- \sqrt{\varepsilon} v_{k+1} - h_- h_+ (1 + \varepsilon b_-) X_k - h_+ (1 + \varepsilon b_-) \sqrt{\varepsilon} v_k \\ & = h_+ h_- \sigma_- \sqrt{\varepsilon} u_k + h_- \sqrt{\varepsilon} v_{k+1} - h_+ (1 + \varepsilon b_-) \sqrt{\varepsilon} v_k, \end{aligned}$$

isto é,

$$\frac{1}{\sqrt{\varepsilon}} (h_- Y_{k+1} - h_+ (1 + \varepsilon b_-) Y_k) = h_+ h_- \sigma_- u_k + h_- v_{k+1} - h_+ (1 + \varepsilon b_-) v_k$$

e $Z_k = h_+ h_- \sigma_- u_k + h_- v_{k+1} - h_+ (1 + \varepsilon b_-) v_k$ segue uma lei gaussiana $N(0, h_+^2 h_-^2 \sigma_-^2 + h_-^2 + h_+^2 (1 + \varepsilon b_-)^2)$.

Se D_c^1 ocorre então podemos estabelecer a seguinte majoração:

$$\frac{1}{\sqrt{\varepsilon}} |h_- - h_+ (1 + \varepsilon b_-)| c \leq Z_k.$$

Atendendo à distribuição de Z_k , para uma dada probabilidade de risco α_d ⁽¹⁾ é elementar obter $\lambda > 0$ tal que:

$$P(Z_k^0 \geq \lambda) = \alpha_d,$$

onde $\{Z_k^0\}$ representa um processo normalizado associado a $\{Z_k\}$ e $\phi^{-1}(\lambda) = 1 - \alpha_d$, com ϕ^{-1} para representar a inversa da função distribuição da lei normal reduzida. Daí obtemos a expressão de c . Repetimos o mesmo raciocínio supondo agora que $h_- > 0$ e $h_+ < 0$. Finalmente deparamos com a seguinte escolha para o valor de c :

⁽¹⁾ Nas aplicações tomaremos $\alpha_d = 5\%$.

$$c_1 = \lambda \sqrt{\varepsilon} \frac{\sqrt{h_+^2 h_-^2 \sigma_-^2 + h_-^2 + h_+^2 (1 + \varepsilon b_-)^2}}{|h_- - h_+ (1 + \varepsilon b_-)|}$$

e então $P(D_{c_1}^1) \leq \alpha_d$.

Se seguirmos um procedimento semelhante para o acontecimento D_c^2 , obtemos para c o valor :

$$c_2 = \lambda \sqrt{\varepsilon} \frac{\sqrt{h_+^2 h_-^2 \sigma_+^2 + h_+^2 + h_-^2 (1 + \varepsilon b_+)^2}}{|h_+ - h_- (1 + \varepsilon b_+)|}$$

e portanto, para uma probabilidade de erro α_d , a constante do teste de detecção de passagens em zero tem a seguinte expressão :

$$c_{obs} = \max(c_1, c_2) \quad (4.3)$$

Após a determinação da constante c_{obs} podemos agora seleccionar os intervalos $[i_0, i_0 + m]$ em que ocorre apenas um dos acontecimentos alternativos, A_- ou A_+ . Os intervalos detectados por este teste determinam os instantes para uma possível estimação pelo filtro de Kalman-Bucy, sendo para isso necessário decidir primeiro sobre o sinal das variáveis de estado em cada um desses intervalos de linearidade, procedimento este que abordaremos no capítulo que se segue (capítulo 5).

5. Testes de decisão sobre o sinal do processo não observado

Após a selecção dos intervalos em que o processo de estado, para uma dada probabilidade, não tem passagens em zero, é necessário determinar o sinal do processo não observado ou, mais concretamente, é necessário decidir sobre qual dos filtros (\hat{X}_k^+ ou \hat{X}_k^- dados no capítulo 3) deve ser usado nesses mesmos intervalos. Na decisão sobre o sinal podemos considerar dois tipos de testes, o da razão de verosimilhança e o de variação quadrática. As estatísticas para estes testes assumem aspectos diferentes consoante a hipótese de detectabilidade que se considera ((HD1) ou (HD2)). Tal como já tivemos oportunidade de referir anteriormente, neste trabalho pretende-se, para o caso da hipótese de detectabilidade (HD2), analisar o teste de variação quadrática e compará-lo com o teste da razão de verosimilhança. Considerámos no entanto, brevemente, a hipótese (HD1) para termos de comparação na discussão do comportamento dos testes (fim da secção 5.2). Nas secções 5.1 e 5.2 apresentam-se os testes da razão de verosimilhança e de variação quadrática segundo a hipótese (HD2), que relembramos a seguir .

$$(HD2): \quad h_-^2 \sigma_-^2 = h_+^2 \sigma_+^2 \quad \text{e} \quad b_+ \neq b_-$$

Os dois tipos de testes que apresentamos neste capítulo permitem decidir sobre o sinal do processo de estado optando entre duas alternativas, que correspondem à ocorrência do acontecimento A_- ou à ocorrência do acontecimento A_+ , onde

$$A_- = \{ X_k < 0; k = i_0, i_0 + 1, \dots, i_0 + m \},$$
$$A_+ = \{ X_k > 0; k = i_0, i_0 + 1, \dots, i_0 + m \}.$$

Os intervalos de tempo discreto $[i_0, i_0 + m]$ considerados provêm do primeiro teste, o teste de detecção, apresentado no capítulo anterior. O instante discreto i_0 corresponde ao primeiro instante a partir do qual se inicia a contabilização da estatística a qual, como veremos, será do tipo incremental.

5.1 – Teste da razão de verosimilhança

O teste da razão de verosimilhança, tal como o de variação quadrática, é desenvolvido para determinar em qual das duas regiões é que o estado X_k se encontra, em cada um dos intervalos detectados pelo teste de detecção de passagens em zero. O teste da razão de verosimilhança que vamos considerar baseia-se no processo das observações mas, como já referimos, é possível também descrever um teste de razão de verosimilhança sobre o estimador de Kalman Estendido (Fleming et al., 1991, Oliveira-Roubaud, 1995).

Consideraremos então o processo das observações correspondentes a um intervalo sem passagens em zero, $\{Y_k, i_0 + 1 \leq k \leq i_0 + m\}$, em que i_0 corresponde ao primeiro instante a partir do qual se inicia a contabilização da estatística.

De acordo com o modelo (2.1) os valores observados para o processo $\{Y_k\}_k$ seguem a equação

$$Y_k = h(X_k) + \sqrt{\varepsilon} v_k,$$

que pode ser reescrita sob as hipóteses A_- e A_+ da seguinte maneira :

$$\begin{aligned} \text{Sob } A_+ : & Y_k = (1 + \varepsilon b_+) X_k + \sqrt{\varepsilon} \dot{\Gamma} u_k^+ + \sqrt{\varepsilon} h_+ v_k^+, \\ \text{Sob } A_- : & Y_k = (1 + \varepsilon b_-) X_k + \sqrt{\varepsilon} \dot{\Gamma} u_k^- + \sqrt{\varepsilon} h_- v_k^-, \end{aligned}$$

com $\dot{\Gamma}^2 = (h_- \sigma_-)^2 = (h_+ \sigma_+)^2$.

Consideremos a σ -álgebra $F_{i_0}^k = \sigma \{x_{i_0}, u_{i_0}, \dots, u_k\}$, $i_0 \leq k \leq i_0 + m$ e as medidas de probabilidade P^+ e P^- , tais que :

$$P^+(D) = P(D), \forall D \subset A_+,$$

$$P^-(D) = P(D), \forall D \subset A_-.$$

O processo $\{Y_k; k \geq i_0 + 1\}$ é $F_{i_0}^k$ - mensurável (i.e. previsível) e $\sum_{k=i_0}^{i_0+m} Y_k^2 < +\infty$ P - q.c.

Nas condições do modelo que estamos a tratar e em particular lembrando que $b_- \neq b_+$, podemos recorrer a uma versão discreta do Teorema de Girsanov (Oliveira, 1994) para determinar a estatística de teste, usando o logaritmo da razão de verosimilhança.

Para um modelo com passo de tempo Δt e sem ruído de observação

$$\begin{cases} X_{k+1} = X_k + \Delta t b(X_k) + \sqrt{\Delta t} \sigma(X_k) u_k, \\ Y_k = h(X_k), \end{cases}$$

definimos a estatística

$$L_n = \ln \left. \frac{dP^+}{dP^-} \right|_{F_{i_0}^k},$$

com $1 \leq n \leq m - 1$. Oliveira-Roubaud (1995) deduzem com base numa versão discreta do Teorema de Girsanov (ver Oliveira, 1994) que

$$L_n = \frac{1}{\dot{\Gamma}^2} \left[(b_+ - b_-) \sum_{k=i_0+1}^{i_0+n} Y_k (Y_{k+1} - Y_k) - \frac{1}{2} \Delta t (b_+^2 - b_-^2) \sum_{k=i_0+1}^{i_0+n} Y_k^2 \right].$$

Tendo em conta a igualdade

$$Y_k (Y_{k+1} - Y_k) = -\frac{1}{2} \left((Y_{k+1} - Y_k)^2 - Y_{k+1}^2 + Y_k^2 \right),$$

podemos reescrever a estatística L_n na seguinte forma :

$$L_n = \frac{b_+ - b_-}{2\dot{\Gamma}^2} (Y_{i_0+n+1}^2 - Y_{i_0+1}^2) - \frac{b_+ - b_-}{2\dot{\Gamma}^2} \sum_{k=i_0+1}^{i_0+n} (Y_{k+1} - Y_k)^2 - \Delta t \frac{b_+^2 - b_-^2}{2\dot{\Gamma}^2} \sum_{k=i_0+1}^{i_0+n} Y_k^2,$$

onde, sob a hipótese $A_+ \cup A_-$ se tem que, $\sum_{k=i_0+1}^{i_0+n} (Y_{k+1} - Y_k)^2 - \dot{\Gamma}^2 n \Delta t = O(\sqrt{\Delta t})$.

Para o modelo (2.1), em estudo, usamos uma aproximação L_n^ε desta estatística com $\Delta t = \varepsilon$:

$$L_n^\varepsilon = \frac{b_+ - b_-}{2\hat{\Gamma}^2} (Y_{i_0+n+1}^2 - Y_{i_0+1}^2) - \frac{1}{2}n\varepsilon(b_+ - b_-) - \varepsilon \frac{b_+^2 - b_-^2}{2\hat{\Gamma}^2} \sum_{k=i_0+1}^{i_0+n} Y_k^2, \quad (5.4)$$

que tem a seguinte propriedade (Oliveira-Roubaud, 1995):

Teorema 5.1 – Considerando o acontecimento C definido no capítulo anterior, tem-se que, sob C ,

$$L_n^\varepsilon - L_n = O(\sqrt{\varepsilon}).$$

É esta estatística L_n^ε que tomaremos como estatística de teste. A regra de decisão do teste é a seguinte:

- Se $L_{N^*} \geq l_2$, decide-se pelo acontecimento A_+ ;
- Se $L_{N^*} \leq -l_1$, decide-se pelo acontecimento A_- ;
- Senão, não se decide,

onde $l_1 > 0$ e $l_2 > 0$ são duas constantes que forneceremos de seguida e N^* é um tempo de paragem definido por:

$$N^* = \inf \{n : L_n \geq l_2 \text{ ou } L_n \leq -l_1\} \wedge m,$$

correspondente ao número mínimo de instantes necessário para decidir sobre o sinal do estado num dado intervalo detectado.

O passo seguinte será então a determinação dos valores dos parâmetros l_1 e l_2 que nos permitem escrever a regra de decisão. Consideremos as probabilidades de erro do tipo I, p_- e p_+ ⁽²⁾, respeitantes à decisão sobre as hipóteses alternativas A_- e A_+ :

$$\begin{cases} p_+ = P(\text{'rejeitar } A_+ \text{'} | A_+) \\ p_- = P(\text{'rejeitar } A_- \text{'} | A_-) \end{cases} \quad (5.5)$$

⁽²⁾ Em geral considera-se $p_+ = p_- = 0.05$.

Fixando as probabilidades de erro p_+ e p_- , os valores de l_1 e l_2 resultam da resolução do sistema (não linear) seguinte (ver Oliveira-Roubaud, 1995):

$$\begin{cases} p_+ = \frac{1 - e^{-l_2}}{e^{l_1} - e^{-l_2}} \\ p_- = \frac{1 - e^{-l_1}}{e^{l_2} - e^{-l_1}} \end{cases} \quad (5.6)$$

Como podemos constatar a solução deste sistema é independente dos parâmetros do modelo e portanto os valores de l_1 e l_2 são constantes, ou seja, dependem apenas dos valores de erro admitidos para o teste.

Dado que pretendemos comparar o comportamento dos testes, torna-se interessante caracterizar a performance do teste em termos do tempo necessário para atingir uma decisão, razão pela qual medimos aquilo a que chamaremos o tempo médio de espera para uma decisão. É possível obter os valores esperados do tempo de espera, teoricamente, usando uma aproximação em tempo contínuo (Fleming et al., 1991 e Oliveira-Roubaud, 1995). Considerando os tempos de paragem N^+ e N^- que representam a variável aleatória N^* sob as hipóteses A_+ e A_- respectivamente, definimos os correspondentes tempos de espera $T_v^+ = N^+ \varepsilon$ e $T_v^- = N^- \varepsilon$ contabilizados no intervalo de tempo contínuo $[0, T]$. Os valores esperados dos tempos de espera são dados por :

$$E(T_v^+) = \frac{-2b_+ (2 + \varepsilon b_+)}{(b_+ - b_-)^2} [l_2 - p_+ (l_1 + l_2)],$$

$$E(T_v^-) = \frac{-2b_- (2 + \varepsilon b_-)}{(b_+ - b_-)^2} [l_1 - p_- (l_1 + l_2)].$$

Torna-se também interessante comparar os valores médios teóricos do tempo de espera, assim obtidos, com os seus valores amostrais. Esta comparação será feita no capítulo 6 sobre alguns exemplos de aplicação. Os tempos de espera (ou de paragem) permitem fazer comparações entre testes ou até mesmo entre hipóteses, dentro do mesmo teste, de forma a se analisarem as alterações no que diz respeito à rapidez de decisão sobre o sinal.

O sistema de equações (5.6) por si só pode ter várias soluções e terá de ser resolvido numericamente. Do ponto de vista prático, para $p_+ = p_- = \alpha$, a dificuldade de obtenção de

uma solução adequada do sistema (5.6) pode ser ultrapassada procurando uma solução em que $l_1 = l_2$. Nessa altura facilmente se verifica que :

$$l_1 = l_2 = \ln \frac{1 - \alpha_s}{\alpha_s}.$$

Esta solução foi adoptada em Oliveira-Roubaud (1995) com sucesso e será também adoptada nesta tese. Por exemplo, para $\alpha_s = 0.05$ obtem-se $l_1 = l_2 = 2.94$ em qualquer um dos casos de aplicação.

Veremos também que, do ponto de vista prático, é recomendável considerar como primeiro instante de cálculo da estatística de teste não o índice i_0 mas sim um índice ligeiramente posterior ⁽³⁾. Esta pequena alteração em nada altera o espírito do teste. O seu efeito é apenas o de evitar pequenas instabilidades que possam ocorrer quando o sistema ainda esteja a “entrar num intervalo de linearidade”. Esta observação também é válida para o teste de variação quadrática, cujo procedimento analisamos na secção seguinte.

5.2 – Teste de variação quadrática

Com o intuito de introduzir a estatística deste teste começemos por considerar as seguintes notações :

$$\Delta_k^+ = Y_{k+1} - (1 + b_+ \varepsilon) Y_k,$$

$$\Delta_k^- = Y_{k+1} - (1 + b_- \varepsilon) Y_k,$$

e vejamos qual a ideia que preside ao teste de variação quadrática.

No caso de se ter $X_k > 0$ e $X_{k+1} > 0$ (acontecimento A_+), estes processos são dados pelas seguintes expressões :

⁽³⁾ Nos exemplos de aplicação que apresentamos no capítulo 6 consideramos, de um modo geral, i_0 incrementado de 6 unidades. No entanto alguns ensaios foram realizados com i_0 incrementado de 11 unidades, de modo a estudar a sensibilidade do teste relativamente ao parâmetro.

$$\begin{aligned}
\Delta_k^+ &= Y_{k+1} - (1 + b_+ \varepsilon) Y_k \\
&= h_+(X_{k+1}) + \sqrt{\varepsilon} v_{k+1} - (1 + b_+ \varepsilon)(h_+(X_k) + \sqrt{\varepsilon} v_k) \\
&= (1 + \varepsilon b_+) h(X_k) + h_+ \sigma_+ \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_+ \varepsilon)(h_+(X_k) + \sqrt{\varepsilon} v_k) \\
&= h_+ \sigma_+ \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_+ \varepsilon) \sqrt{\varepsilon} v_k,
\end{aligned}$$

$$\begin{aligned}
\Delta_k^- &= Y_{k+1} - (1 + b_- \varepsilon) Y_k \\
&= h_+(X_{k+1}) + \sqrt{\varepsilon} v_{k+1} - (1 + b_- \varepsilon)(h_+(X_k) + \sqrt{\varepsilon} v_k) \\
&= (1 + \varepsilon b_+) h(X_k) + h_+ \sigma_+ \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_- \varepsilon)(h_+(X_k) + \sqrt{\varepsilon} v_k) \\
&= h_+(b_+ - b_-) \varepsilon X_k + h_+ \sigma_+ \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_- \varepsilon) \sqrt{\varepsilon} v_k,
\end{aligned}$$

expressões que decorrem do próprio modelo (2.1). No caso de ocorrer $X_k < 0$ e $X_{k+1} < 0$ (acontecimento A_-), obtém-se após um cálculo análogo :

$$\begin{aligned}
\Delta_k^+ &= h_-(b_- - b_+) \varepsilon X_k + h_- \sigma_- \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_+ \varepsilon) \sqrt{\varepsilon} v_k, \\
\Delta_k^- &= h_- \sigma_- \sqrt{\varepsilon} u_k + \sqrt{\varepsilon} v_{k+1} - (1 + b_- \varepsilon) \sqrt{\varepsilon} v_k.
\end{aligned}$$

Analisando estas expressões facilmente concluímos que, se o acontecimento A_+ ocorre, a variável aleatória Δ_k^+ tem distribuição $\mathcal{N}(0, \Gamma_+^2 \varepsilon)$ e se o acontecimento A_- ocorre, a variável aleatória Δ_k^- tem distribuição $\mathcal{N}(0, \Gamma_-^2 \varepsilon)$, sendo

$$\begin{aligned}
\Gamma_+^2 &= h_+^2 \sigma_+^2 + 1 + (1 + b_+ \varepsilon)^2, \\
\Gamma_-^2 &= h_-^2 \sigma_-^2 + 1 + (1 + b_- \varepsilon)^2.
\end{aligned}$$

Podemos então dizer que a decisão entre as alternativas A_- e A_+ nos conduz naturalmente a um teste de hipóteses sobre as variâncias das variáveis aleatórias Δ_k^- e Δ_k^+ .

Note-se que $\Gamma_+^2 - \Gamma_-^2 = O(\varepsilon)$ (ver notação 2.1). Mais precisamente $\Gamma_+^2 - \Gamma_-^2 = \varepsilon(b_+ - b_-)(2 + (b_+ + b_-)\varepsilon)$, logo não será surpreendente se chegarmos a um teste que necessita de um tempo longo para tomar uma decisão.

Com o intuito de construir este teste considere-se a razão das respectivas funções de densidade (gaussianas) :

$$Z_k = \ln \left(\frac{f_N(\Delta_k^+, 0, \Gamma_+^2 \varepsilon)}{f_N(\Delta_k^-, 0, \Gamma_-^2 \varepsilon)} \right).$$

Ela permite definir a seguinte estatística :

$$S_n = \sum_{k=i_0}^{i_0+n} Z_k,$$

com $0 \leq n \leq m$. Desenvolvendo a razão Z_k obtém-se :

$$Z_k = \ln \left(\frac{\Gamma_-}{\Gamma_+} \right) + \frac{1}{2\varepsilon} \left(\frac{\Delta_k^-^2}{\Gamma_-^2} - \frac{\Delta_k^+^2}{\Gamma_+^2} \right)$$

e portanto :

$$S_n = (n+1) \ln \left(\frac{\Gamma_-}{\Gamma_+} \right) + \frac{1}{2\varepsilon} \sum_{k=i_0}^{i_0+n} \left(\frac{\Delta_k^-^2}{\Gamma_-^2} - \frac{\Delta_k^+^2}{\Gamma_+^2} \right) \quad (5.7)$$

Construída a estatística de teste, S_n , falta agora determinar a região crítica para a estatística, que é calculada atendendo às probabilidades de erro consideradas, p_- e p_+ ⁽⁴⁾, definidas como na secção anterior (ver (5.5)).

Seja $[-l_1, l_2]$, com $l_1 > 0$ e $l_2 > 0$, a região crítica para a estatística S_n , que é construída para um nível de significância α_s . Os valores para l_1 e l_2 podem ser determinados aproximadamente se em (5.5) substituirmos os segundos membros por expressões aproximadas. Temos então o seguinte sistema de equações não linear:

$$\begin{cases} p_+ = \frac{1 - e^{-(\mu_+ / \gamma_+^2) l_2}}{e^{(\mu_+ / \gamma_+^2) l_1} - e^{-(\mu_+ / \gamma_+^2) l_2}} \\ p_- = \frac{1 - e^{-(\mu_- / \gamma_-^2) l_1}}{e^{-(\mu_- / \gamma_-^2) l_2} - e^{-(\mu_- / \gamma_-^2) l_1}}, \end{cases} \quad (5.8)$$

em que

⁽⁴⁾ Em geral considera-se $p_+ = p_- = 0.05$.

$$\begin{aligned}\mu_+ &= \ln\left(\frac{\Gamma_-}{\Gamma_+}\right) + \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\Gamma_+^2, \\ \mu_- &= \ln\left(\frac{\Gamma_-}{\Gamma_+}\right) + \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\Gamma_-^2, \\ \gamma_+ &= \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\sqrt{\Gamma_+^4 + 2(1+b_+\varepsilon)^2}, \\ \gamma_- &= \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\sqrt{\Gamma_-^4 + 2(1+b_-\varepsilon)^2}.\end{aligned}$$

A maneira como estas expressões são deduzidas assim como detalhes sobre a construção do teste são apresentados em Oliveira (2000).

Definida a estatística S_n e calculados os valores de l_1 e l_2 segundo as probabilidades de erro p_- e p_+ , podemos agora estabelecer a regra de decisão para o teste (de variação quadrática):

- Se $S_{N^*} \leq -l_1$, decidir pelo acontecimento A_- ;
- Se $S_{N^*} \geq l_2$, decidir pelo acontecimento A_+ ;
- Caso contrário, não decidir,

onde N^* é um tempo de paragem definido da mesma forma que no teste anterior (secção 5.1), agora para a estatística S_n . O tempo necessário para o teste decidir em cada intervalo, N^* , corresponde ao número mínimo de instantes que a estatística necessita para atingir os valores $-l_1$ ou l_2 , sempre que se verifica uma decisão. São os já referidos tempos de espera para a decisão, que se podem contabilizar separadamente para as decisões sob a hipótese A_+ ou sob a hipótese A_- , N^+ e N^- , respectivamente. Supondo que o modelo (2.1) resulta da discretização no tempo de um modelo contínuo num intervalo de tempo $[0, T]$, os tempos de espera $T_{vq}^+ = N^+ \varepsilon$ e $T_{vq}^- = N^- \varepsilon$ podem ser caracterizados teoricamente pelas suas esperanças matemáticas. Têm-se então as seguintes expressões (Oliveira-Roubaud, 1991 e Fleming et al., 1991):

$$\begin{aligned}E(T_{vq}^+) &= \left(\frac{l_2}{\mu_+} - p_+ \frac{l_2 + l_1}{\mu_+}\right) \varepsilon, \\ E(T_{vq}^-) &= \left(-\frac{l_2}{\mu_-} + p_- \frac{l_2 + l_1}{\mu_-}\right) \varepsilon.\end{aligned}$$

Voltando ao sistema de equações não linear (5.8), convém notar que a determinação da sua solução, de um ponto de vista prático, não é imediata. De facto esta tem geralmente de ser obtida por métodos numéricos, de entre os métodos habitualmente usados para resolução de sistemas não lineares com os inconvenientes que tão bem conhecemos. Existe em geral um conjunto de soluções possíveis e aproximadas, ficando o sucesso do teste muito dependente da escolha de uma solução inicial para $-l_1$ e l_2 para arranque do método numérico. Por outro lado, não temos garantia do sistema ter sempre solução e por vezes, quando existe mais do que uma solução, torna-se necessário optar pelo par de valores (l_1, l_2) mais “adequado”. Podemos verificar que os resultados do teste são muito sensíveis aos parâmetros l_1 e l_2 e por isso tem de ser estudado algum procedimento a seguir para a sua determinação na prática. No caso da hipótese (HD1) Fleming et al. (1991) contornaram esta dificuldade usando a seguinte aproximação:

$$\begin{aligned}
 l_1 &= - \left| \frac{1}{\theta_+} \right| \log(\alpha_s), \\
 l_2 &= - \left| \frac{1}{\theta_-} \right| \log(\alpha_s), \\
 \text{com } \theta_+ &= \frac{\mu_+}{\gamma_+} \text{ e } \theta_- = \frac{\mu_-}{\gamma_-}.
 \end{aligned} \tag{5.9}$$

Estas aproximações foram testadas em alguns exemplos de aplicação e revelaram-se eficazes no caso da hipótese (HD1). Infelizmente, como veremos mais à frente (secção 6.2), o mesmo não se pode dizer para a hipótese (HD2). Os resultados numéricos que o demonstram são apresentados de forma mais pormenorizada no capítulo 6, nos exemplos de aplicação.

Embora neste trabalho estejamos interessados em estudar o problema sob a hipótese de detectabilidade (HD2), é interessante introduzir aqui o teste de variação quadrática sob a hipótese (HD1), $h_-^2 \sigma_-^2 \neq h_+^2 \sigma_+^2$. Fleming et al. (1991) desenvolveram este teste e verificaram que a condição $h_-^2 \sigma_-^2 \neq h_+^2 \sigma_+^2$ permite ao teste ganhar uma maior sensibilidade na determinação da região em que se encontra o processo de estado, relativamente à condição $h_-^2 \sigma_-^2 = h_+^2 \sigma_+^2$ da hipótese (HD2) e que, quando se verifica (HD1), os valores dos parâmetros b_+ e b_- se tornam irrelevantes no que diz respeito aos resultados de decisão do teste. Na hipótese (HD2) tinha-se $h_-^2 \sigma_-^2 = h_+^2 \sigma_+^2$ e o teste via-se obrigado a recorrer às

diferenças entre os parâmetros b_+ e b_- para poder decidir sobre o sinal. Assim, quando consideramos a hipótese (HD1) podemos omitir os termos em b_+ e b_- considerando $b_+ = b_- = 0$ sem alterar os resultados da análise deste teste.

O teste pode ser descrito do seguinte modo (Fleming et al., 1991). Para as alternativas A_- e A_+ a testar considere-se a estatística

$$S_n = (n + 1) \ln \left(\frac{\Gamma_-}{\Gamma_+} \right) + \frac{1}{2} \left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2} \right) \sum_{k=i_0}^{i_0+n} \frac{\Delta_k^2}{\varepsilon}, \quad (5.10)$$

com $0 \leq n \leq m$, onde

$$\begin{aligned} \Delta_k &= Y_{k+1} - Y_k, \\ \Gamma_+^2 &= h_+^2 \sigma_+^2 + 2, \\ \Gamma_-^2 &= h_-^2 \sigma_-^2 + 2. \end{aligned}$$

Esta estatística será calculada em cada um dos intervalos detectados, $[i_0, i_0 + m]$.

Tomando, como é habitual, p_+ como sendo a probabilidade de rejeitar erradamente o acontecimento A_+ e p_- como sendo a probabilidade de rejeitar erradamente o acontecimento A_- , os valores dos limites da região crítica, l_1 e l_2 , podem ser obtidos resolvendo o sistema

$$\begin{cases} p_+ = \frac{1 - e^{-(\mu_+ / \gamma_+^2) l_2}}{e^{(\mu_+ / \gamma_+^2) l_1} - e^{-(\mu_+ / \gamma_+^2) l_2}} \\ p_- = \frac{1 - e^{-(\mu_- / \gamma_-^2) l_1}}{e^{-(\mu_- / \gamma_-^2) l_2} - e^{-(\mu_- / \gamma_-^2) l_1}}, \end{cases} \quad (5.11)$$

em que

$$\begin{aligned}\mu_+ &= \ln\left(\frac{\Gamma_-}{\Gamma_+}\right) + \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\Gamma_+^2, \\ \mu_- &= \ln\left(\frac{\Gamma_-}{\Gamma_+}\right) + \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\Gamma_-^2, \\ \gamma_+ &= \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\sqrt{\Gamma_+^4 + 2(1 + b_+ \varepsilon)^2}, \\ \gamma_- &= \frac{1}{2}\left(\frac{1}{\Gamma_-^2} - \frac{1}{\Gamma_+^2}\right)\sqrt{\Gamma_-^4 + 2(1 + b_- \varepsilon)^2}.\end{aligned}$$

Tal como nas situações anteriores, a resolução do sistema (5.11) não é imediata, pelo facto de se ter um sistema de equações não linear. No entanto, no caso da hipótese (HD1), verifica-se, na prática, ser mais fácil encontrar uma solução aproximada que satisfaça os níveis de erro considerados, $p_+ = p_- = \alpha_s$ ⁽⁵⁾. A aproximação (5.9) sugerida pelos autores Fleming et al. (1991), apresentada acima, permite encontrar boas soluções para l_1 e l_2 , tal como teremos oportunidade de verificar nas aplicações práticas (capítulo 6).

Quanto à regra de decisão ela é análoga à do teste sob a hipótese (HD2), figurando naturalmente a estatística S_n apresentada em (5.10) no lugar da estatística anteriormente apresentada em (5.7).

Os valores médios teóricos dos tempos de paragem são agora dados por :

$$\begin{aligned}E(T_{vq}^+) &= \left(\frac{l_2}{\mu_+} - p_+ \frac{l_2 + l_1}{\mu_+}\right)\varepsilon, \\ E(T_{vq}^-) &= \left(-\frac{l_2}{\mu_-} + p_- \frac{l_2 + l_1}{\mu_-}\right)\varepsilon.\end{aligned}$$

Apresentados os procedimentos a seguir na aplicação dos testes da razão de verosimilhança e de variação quadrática, falta-nos agora comparar os seus resultados na prática, o que faremos no capítulo seguinte (capítulo 6).

⁽⁵⁾ Na prática, consideraremos $\alpha_s = 0.05$.

6. Exemplos de Aplicação

Este capítulo é dedicado ao estudo do comportamento dos testes de detecção de passagens em zero e de decisão sobre o sinal. Para isso recorreremos à simulação numérica de alguns exemplos de modelos. Começaremos por analisar brevemente o comportamento do teste de detecção de passagens em zero na selecção dos intervalos em que, segundo uma dada probabilidade, se podem obter boas estimativas para as variáveis de estado X_k . Mas o principal objectivo consiste em analisar o desempenho do teste de variação quadrática, na decisão sobre o sinal do estimador a aplicar, e compará-lo com o teste da razão de verosimilhança tendo em conta a hipótese (HD2).

A análise que se segue é baseada em 100 trajectórias geradas, num intervalo de tempo contínuo $[0,15]$ o que, para $\mathcal{E}=0.01$, corresponde a considerar uma partição em 1500 instantes de tempo discreto ($K=1500$). A simulação dos processos de estado e observações foi feita partindo da condição inicial $X_0 \sim N(-5, 0.1)$.

A performance dos testes é medida através das percentagens de instantes em que os testes actuam e das percentagens de erros na detecção e decisão sobre o sinal para o processo não observado. As probabilidades de erro são comparadas com as frequências de decisões erradas encontradas nas simulações, pe_v e pe_{vq} . Para o cálculo destas percentagens consideram-se apenas os intervalos em que efectivamente não houve cruzamentos em zero, independentemente de estarem ou não bem detectados pelo primeiro teste.

Os testes de decisão sobre o sinal são ainda analisados tendo em conta os tempos médios de paragem, que correspondem aos tempos (mínimos) de que, em média, um teste precisa para decidir. A análise dos tempos de paragem permite-nos comparar a rapidez de decisão em ambos os testes (razão de verosimilhança e variação quadrática) ou até mesmo analisar o efeito dos parâmetros do modelo na rapidez de decisão para cada um dos testes em particular. Neste capítulo apresentaremos alguns dos resultados mais importantes, sob a forma de tabela ou gráfico, sendo os restantes resultados apresentados em apêndice de forma mais detalhada.

Neste capítulo serão consideradas as seguintes notações:

TD : teste de detecção.

TV : teste da razão de verosimilhança .

TVQ : teste de variação quadrática.

α_d : nível de significância para o teste de detecção de passagens em zero.

α_s : nível de significância para o teste de decisão sobre o sinal.

$-l_1, l_2$: valores críticos para a estatística do teste de decisão sobre o sinal.

l_1^v, l_2^v : valores de l_1 e l_2 usados no teste da razão de verosimilhança.

l_1^{vq}, l_2^{vq} : valores de l_1 e l_2 usados no teste de variação quadrática.

pe_d : percentagem de erro na detecção de passagens em zero.

pe_v : percentagem de erro no teste da razão de verosimilhança, na decisão sobre o sinal.

pe_{vq} : percentagem de erro no teste de variação quadrática, na decisão sobre o sinal.

pi_d : percentagem de instantes onde é aplicado o teste de detecção.

pid_v : percentagem de instantes de decisão para o teste da razão de verosimilhança.

pid_{vq} : percentagem de instantes de decisão para o teste de variação quadrática.

ET_v^-, ET_v^+ : médias teóricas do tempo necessário para decisão no teste de verosimilhança.

\bar{T}_v^-, \bar{T}_v^+ : médias (empíricas) do tempo necessário para decisão no teste da razão de verosimilhança.

ET_{vq}^-, ET_{vq}^+ : médias teóricas do tempo necessário para decisão no teste de variação quadrática.

$\bar{T}_{vq}^-, \bar{T}_{vq}^+$: médias (empíricas) do tempo necessário para decisão no teste de variação quadrática.

i_0 : instante a partir do qual se começa a contabilizar a estatística do teste de decisão.

$$S^1 = \frac{1}{100 \times 1500} \sum_{i=1}^{100} \sum_{j=0}^{1500} |x_{ij} - \hat{x}_{ij}|.$$

$$S^2 = \frac{1}{100 \times 1500} \sum_{i=1}^{100} \sum_{j=0}^{1500} |x_{ij} - \hat{x}_{ij}|^2.$$

6.1. Teste de detecção de passagens em zero

Começamos por analisar o teste de passagens em zero baseado nas observações, verificando o efeito dos parâmetros do modelo sobre a performance deste teste. Tal como já foi explicado no capítulo 4, este teste processa-se de acordo com o seguinte procedimento :

- cálculo de uma constante dita de “cutoff” dada pela expressão (4.3) sobre a sequência de observações;
- determinação dos intervalos em que a observação é superior, em valor absoluto, à constante calculada, isto é, determinação dos intervalos de tempo em que o acontecimento $C = \{ |y_k| \geq c_{obs} \}$ ocorre.

A escolha de c_{obs} deve ser um compromisso entre a manutenção de intervalos de monotonia suficientemente grandes e a redução da probabilidade de erro do teste. Nos exemplos que se seguem fixámos uma probabilidade de erro $\alpha_d = 5\%$. Note-se que a constante c_{obs} varia consoante os parâmetros do modelo, sendo muito sensível principalmente ao valor de $|h_- - h_+|$. Quanto mais elevada for esta diferença, maior é o valor desta constante e portanto mais selectivo se torna o teste.

Consideremos o modelo (2.1) com $b_- = 1$ e $b_+ = -0.05$ e diferentes valores de $\sigma_-^2 h_-^2 = \sigma_+^2 h_+^2$. Os resultados da aplicação do teste em alguns exemplos constam da tabela 6.1.

b_-	b_+	h_-	h_+	σ_-	σ_+	c	pi_d	pe_d
1	-0.05	-1	1	1	1	0.1429	61.04%	4.07%
1	-0.05	-1.5	1.5	1	1	0.1701	61.64%	4.46%
1	-0.05	-2	4	1	0.5	0.252	62.36%	1.68%

Tabela 6.1 – Resultados para o teste de detecção de passagens em zero.

Para $|h_- - h_+| = 2$ e $|h_- - h_+| = 6$ obtiveram-se percentagens de erro na decisão de cerca de 4% e 1,68%, respectivamente.

Em geral, a percentagem de instantes que verificam o acontecimento C (pi_d) ronda os 60% apesar de se notar uma ligeira subida nessa percentagem quando aumenta o valor de

$|h_- - h_+|$, que corresponde ao aumento da amplitude dos intervalos detectados. Tal como seria desejável, este erro de detecção de passagens em zero não ultrapassa o nível $\alpha_d = 5\%$.

Vejam agora o comportamento de dois testes de decisão sobre o sinal, o teste da razão de verosimilhança (TV) e o teste de variação quadrática (TVQ), face aos diferentes parâmetros do modelo.

6.2 – Testes de decisão sobre o sinal

O teste de detecção da secção anterior permite seleccionar os intervalos em que, para um dado nível de erro α , as variáveis de estado X_k não cruzam o valor zero. Falta agora decidir se a estes intervalos correspondem ocorrências no semiplano positivo ou no semiplano negativo. Nesta etapa actuam os testes de decisão sobre o sinal.

São apresentados de seguida alguns resultados para dois tipos diferentes de teste de decisão. Em ambos i_0 foi incrementado de 6 unidades. Quer isto dizer que, em cada intervalo de decisão são ignorados os primeiros 6 instantes, de forma a que a decisão não seja precipitada nem tomada em intervalos demasiadamente pequenos. Como fizemos notar no capítulo anterior, após a detecção de uma passagem em zero deve ser aguardado um certo tempo antes de se começarem a calcular as estatísticas, pois nestes testes a decisão do sinal assenta sobre condições de estacionaridade (Fleming et al., 1991 e Oliveira-Roubaud, 1991).

Como vimos, o primeiro passo para a aplicação de um teste de decisão sobre o sinal é o cálculo dos limites l_1 e l_2 . No caso do teste da razão de verosimilhança é possível encontrar a solução $l_1 = l_2 = \ln\left(\frac{1-\alpha_s}{\alpha_s}\right)$ para o sistema de equações não lineares apresentado no capítulo 5 (sistema 5.6) que se verifica ser adequada ao problema (Oliveira-Roubaud, 1991). Assim, independentemente dos parâmetros do modelo, obtém-se $l_1 = l_2 = 2.94$ para uma probabilidade de erro $\alpha_s = 0.05$. No caso do teste de variação quadrática já será necessário calcular novos valores de cada vez que se alteram os parâmetros do modelo. Para além disso, veremos a seguir que nos deparamos com algumas dificuldades quando pretendemos optar por uma boa solução para o par (l_1, l_2) . Convém notar também que a dificuldade em encontrar soluções exactas leva a recorrer a métodos de aproximação (como por exemplo, a

função ‘fsolve’ da linguagem de programação Matlab que permite a resolução numérica de sistemas não lineares).

Incluimos de seguida três figuras (Figuras 6.1 a 6.3) que ilustram bem a importância de uma boa escolha para l_1 e l_2 como soluções de (5.8). Todas elas dizem respeito a uma mesma trajectória para o processo de estado, estimada por diferentes testes e/ou soluções de l_1 e l_2 .

Na Figura 6.1 estão representados os resultados da filtragem usando o teste da razão de verosimilhança, em que os parâmetros l_1 e l_2 são dados por $l_1 = l_2 = \ln\left(\frac{1-\alpha_s}{\alpha_s}\right)$ ($l_1 = l_2 = 2.94$). Nas Figuras 6.2 e 6.3 estão representados os resultados para o teste de variação quadrática com duas soluções diferentes para os parâmetros l_1 e l_2 . Estes valores foram obtidos numericamente através da função ‘fsolve’ do Matlab.

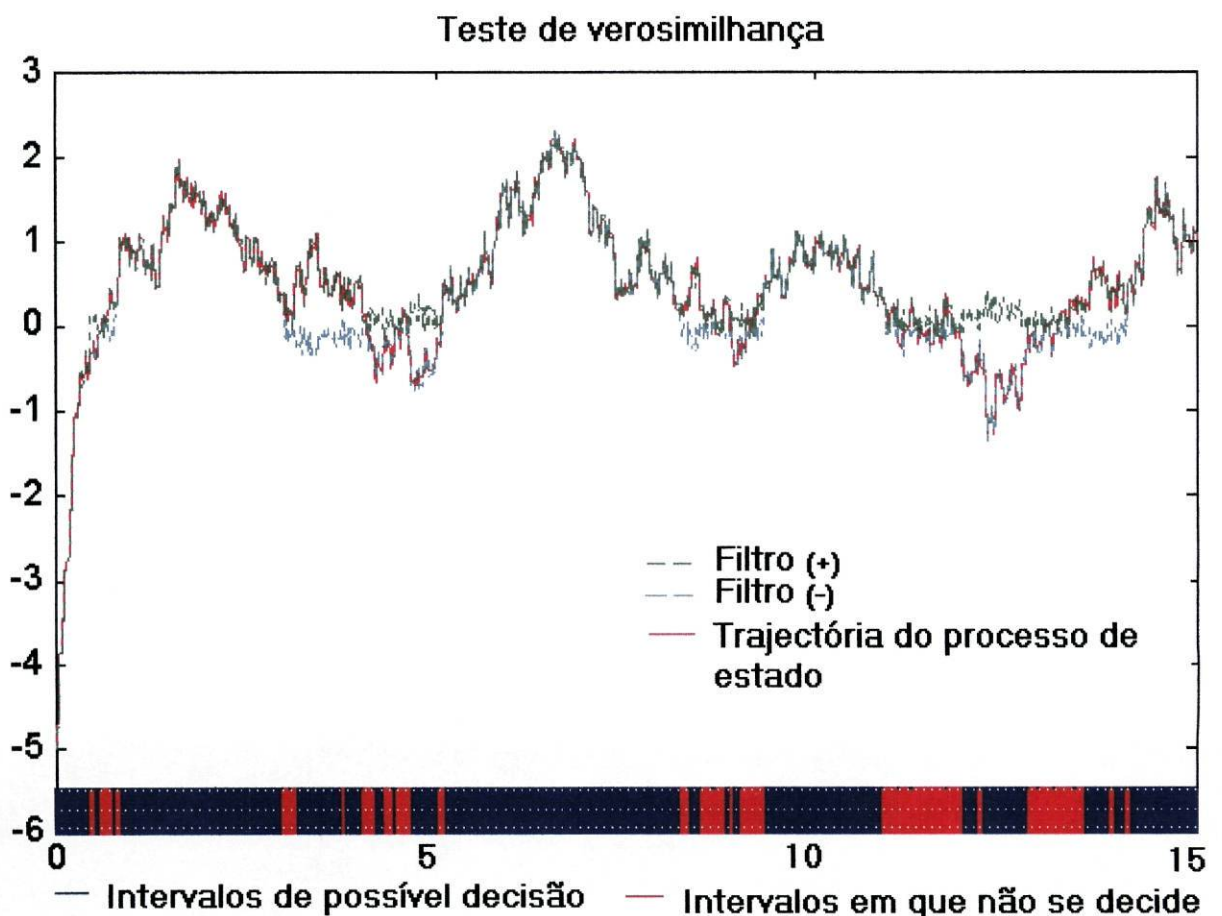


Figura 6.1 – Resultado da filtragem aplicando o teste de verosimilhança, em que l_1 e l_2 são dados por $l_1 = l_2 = \ln\left(\frac{1-\alpha_s}{\alpha_s}\right)$.

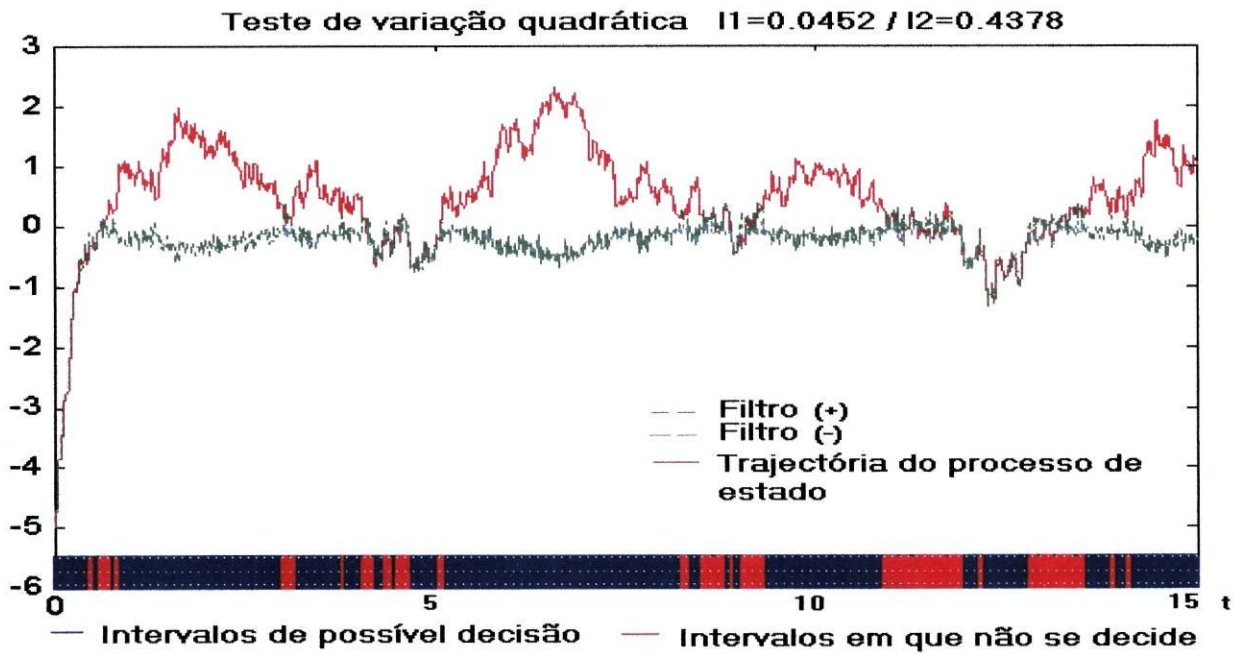


Figura 6.2 - Resultado da filtragem aplicando o teste de variação quadrática, em que $l_1 = 0.0452$, $l_2 = 0.4378$.

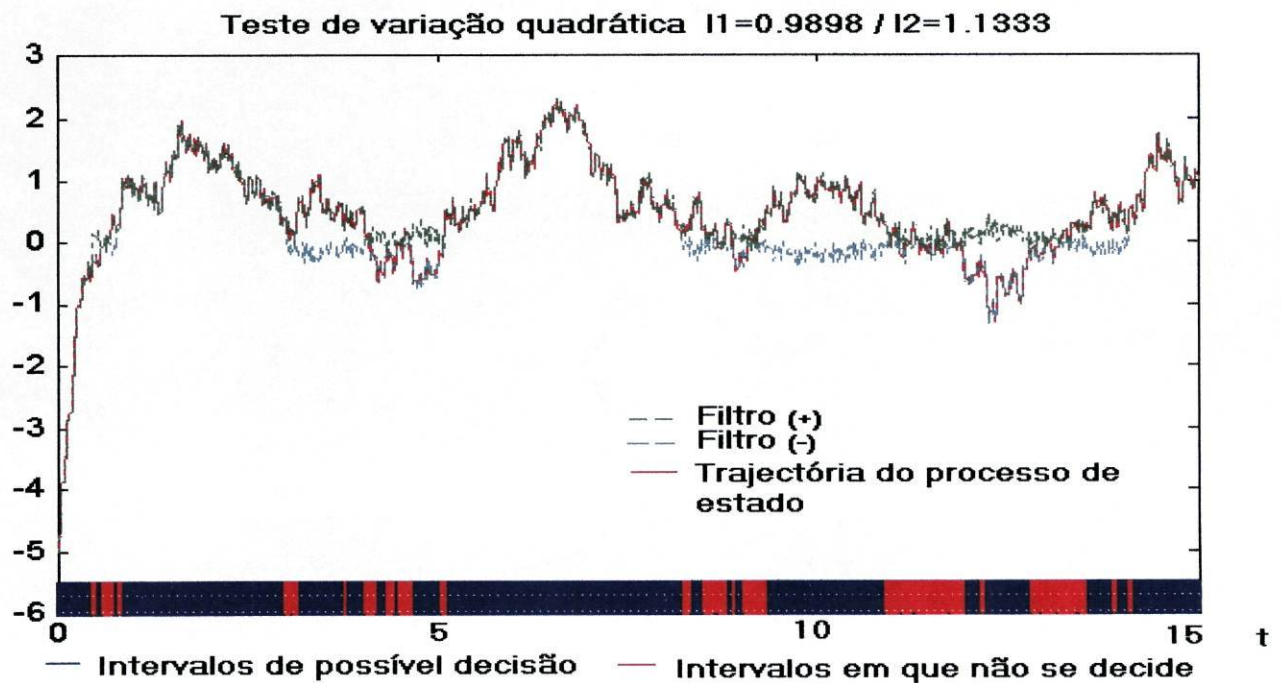


Figura 6.3 - Resultado da filtragem aplicando o teste de variação quadrática, em que $l_1 = 0.9898$, $l_2 = 1.1333$.

A zona identificada com barra azul corresponde aos intervalos seleccionados pelo teste de detecção de passagens em zero. A trajetória do processo de estado está representada a vermelho, sendo os filtros positivo e negativo representados a tracejado em verde e azul, respectivamente. Nos instantes em que os dois traços coincidem, significa que foi tomada

uma decisão e, se eles estiverem do mesmo lado que a trajectória do processo de estado, significa que a decisão foi correcta.

Numa primeira análise da Figura 6.1, podemos verificar que o teste da razão de verosimilhança funcionou bem, decidindo em praticamente todos os intervalos de detecção e de uma forma correcta. Já na Figura 6.2, verificamos que o teste de variação quadrática decidiu em praticamente todos os intervalos de detecção e quase sempre de forma errada. O facto do valor l_1 estar muito próximo de zero faz com que o teste decida muitas vezes pelo lado negativo, sem que corresponda à realidade. Para este modelo, na simulação de 100 trajectórias, incluindo a da figura, obteve-se um erro na decisão muito superior a 5%. Analisando os valores de l_1 e l_2 e os respectivos resultados de estimação, podemos concluir que os valores $l_1 = 0.0452$ e $l_2 = 0.04378$ originam um mau funcionamento do teste. Vejamos agora o que se passa se escolhermos um outro par de valores (l_1, l_2) (ver Figura 6.3). Neste exemplo, uma abordagem do sistema de equações (5.8) análoga à sugerida em Fleming et al. (1991) para os modelos sob a hipótese (HD1) (ver expressão (5.9)) permite encontrar valores de l_1 e l_2 mais equilibrados e que dão origem a probabilidades de erro do tipo I inferiores a 5% (a saber $p_+ = 0$ e $p_- = 0.0413$). A simples alteração dos parâmetros l_1 e l_2 proporcionou um funcionamento adequado do teste, muito semelhante ao do teste da razão de verosimilhança (Figura 6.1).

Este exemplo ilustra bem a importância da escolha dos parâmetros l_1 e l_2 , que pode ter grande influência na qualidade do teste. Do ponto de vista prático, nem sempre é fácil encontrar soluções para os parâmetros, que permitem obter probabilidades de erro adequadas. Há que ensaiar diferentes alternativas, partindo do que foi sugerido por Fleming et al. (1991).

Passamos agora ao estudo do comportamento dos testes em alguns exemplos (casos 1 a 4). Começemos por considerar o modelo (2.1) com os seguintes parâmetros:

$$\begin{cases} b_+ = -0.01, b_- = -0.5 \\ h_+ = 1, h_- = -1 \\ \sigma_+ = 1, \sigma_- = 1 \end{cases} \quad (\text{Caso 1})$$

Os resultados de aplicação dos testes estão resumidos na Tabela 6.2. Resultados mais detalhados podem ser consultados no apêndice A.

Neste caso obteve-se uma percentagem de erro na decisão do teste da razão de verosimilhança muito semelhante à do teste de variação quadrática (6.56% e 6.82%,

respectivamente). As percentagens de instantes de tempo com decisões também foram muito semelhantes (44.46% e 41.79%, respectivamente). Vejamos o que se passa se neste exemplo se aumentar a diferença entre os parâmetros b_- e b_+ . Consideremos dois exemplos em que em relação ao Caso 1 se diminui essencialmente o valor do parâmetro b_- :

$$b_- = -1 \quad (\text{Caso 2})$$

$$b_- = -4 \quad (\text{Caso 3})$$

Da análise da Tabela 6.2 retiramos que, ao aumentar $|b_+ - b_-|$ se obteve uma maior percentagem de instantes de decisão (pid_v e pid_{vq}) e, por outro lado, uma diminuição no erro da decisão (pe_v e pe_{vq}), para ambos os testes. O facto da decisão ser mais rápida nestes últimos casos pode também ser constatado através dos valores médios empíricos \bar{T}_v^-, \bar{T}_v^+ e $\bar{T}_{vq}^-, \bar{T}_{vq}^+$ que diminuíram consideravelmente. Note-se que é natural esperar que o aumento da distância entre os parâmetros b_- e b_+ facilite a decisão sobre o sinal nesses intervalos, uma vez que quanto maior é essa distância mais rapidamente os valores das variáveis de estado X_k diferem no seu comportamento, consoante se situam na região \mathcal{R}^+ ou \mathcal{R}^- .

Casos	b_-	b_+	pid_v	pid_{vq}	pe_v	pe_{vq}	\bar{T}_v^-	\bar{T}_v^+	\bar{T}_{vq}^-	\bar{T}_{vq}^+
1	-0.5	-0.01	41.46%	41.79%	6.56%	6.82%	1.2	2.68	1.46	6.66
2	-1	-0.05	47.21%	34.74%	3.95%	4.41%	0.42	1.44	0.67	3.86
3	-4	-0.05	61.36%	58.77%	4.77%	4.59%	0.09	0.37	0.17	0.5

Tabela 6.2 – Resultados de decisão nos testes da razão de verosimilhança e de variação quadrática para diferentes valores de b_- e b_+ .

Também de notar que, relativamente à detecção de intervalos sem passagens em zero, o aumento de b_- e b_+ em valor absoluto facilita esta detecção uma vez que esse aumento provoca um afastamento mais rápido do estado X_k do limiar 0 (zero).

Comparando agora os dois testes de decisão sobre o sinal, entre si, verificamos que, embora as percentagens de erro sejam semelhantes em ambos os testes, os tempos médios de decisão para o teste de variação quadrática são ligeiramente superiores.

Tentemos agora analisar o efeito da quantidade $\dot{\Gamma}^2 = (\sigma_- h_-)^2 = (\sigma_+ h_+)^2$ no comportamento dos testes. Se analisarmos a estatística utilizada no teste de razão de

verossimilhança (expressão (5.4)) podemos constatar que as alterações na quantidade $\dot{\Gamma}^2 = (\sigma_- h_-)^2 = (\sigma_+ h_+)^2$ deverão influenciar muito os valores da estatística.

Retomando o caso 2, em que $(\sigma_- h_-)^2 = (\sigma_+ h_+)^2 = 1$, vejamos o que sucede se diminuirmos esta quantidade. Consideremos um novo exemplo em que apenas se diminui o valor $(\sigma_- h_-)^2 = (\sigma_+ h_+)^2$:

$$\begin{cases} \sigma_- = 1, h_- = -0.8 \\ \sigma_+ = 1, h_+ = 0.8 \end{cases} \quad \text{(Caso 4)}$$

Os resultados são apresentados na tabela seguinte:

Casos	h_-	h_+	σ_-	σ_+	pid_v	pid_{vq}	pe_v	pe_{vq}	\bar{T}_v^-	\bar{T}_v^+	\bar{T}_{vq}^-	\bar{T}_{vq}^+
2	-1	1	1	1	47.21%	34.74%	3.95%	4.41%	0.42	1.44	0.67	3.86
4	-0.8	0.8	1	1	45.06%	25.98%	8.56%	1.72%	0.32	1.37	1.23	5.3

Tabela 6.3 – Resultados de decisão nos testes da razão de verossimilhança e de variação quadrática, para diferentes funções $\sigma h(x_k)$ ($b_- = -1, b_+ = -0.05$).

A diminuição da quantidade $|h(x_k)|$ faz diminuir os valores do processo das observações reduzindo ligeiramente a percentagem de instantes que verificam o acontecimento C do teste de detecção. Por outro lado, a estatística do teste da razão de verossimilhança necessita de um menor número de instantes de tempo para atingir os valores $-l_1$ ou l_2 , mas vem acompanhada de um aumento do erro de decisão pe_v . Enquanto que para $(\sigma_- h_-)^2 = (\sigma_+ h_+)^2 = 1$ (caso 2) se obtiveram decisões em 47.21% dos instantes e erro em 3,95% das decisões, para $(\sigma_- h_-)^2 = (\sigma_+ h_+)^2 = 0.64$ (caso 4) obtiveram-se decisões em $pid_v = 45.06\%$ dos instantes e $pe_v = 8.56\%$ de decisões erradas.

Quanto ao teste de variação quadrática, verificou-se uma diminuição na frequência de decisões erradas (4.4% e 1.7% respectivamente). Verificámos ainda (ver Tabela A1 no apêndice A) que, se $\dot{\Gamma}^2 = (\sigma_- h_-)^2 = (\sigma_+ h_+)^2$ aumenta consideravelmente este teste toma decisões erradas com maior frequência revelando um mau comportamento na prática. Uma razão para este comportamento poderá estar em que quando $\dot{\Gamma}^2$ aumenta, o termo principal da estatística deste teste (ver (5.7)) diminui de peso. No entanto, o primeiro termo, que nestes casos é sempre negativo, assume maior importância levando por vezes a decisões precipitadas. O facto da estatística atingir o valor $-l_1$, quando as variáveis estão do lado

positivo, leva a concluir que, nestes casos, o teste poderá não ser tão eficaz. Como verificámos, no entanto, que a performance do teste está muito relacionada com a maneira como são determinados os valores de l_1 e l_2 , enveredámos esforços para obter outros valores possíveis para estas constantes o que se revelou infrutífero. Não quer isto dizer, evidentemente, que eles não existam mas sim que seria necessário uma investigação especificamente com este propósito do âmbito da análise numérica. Esta situação mostra que é necessário ter em atenção as características dos modelos e verificar, por via da simulação, se o teste de facto permite obter bons resultados quando aplicado na prática. Pelos exemplos que foram testados constata-se que, aparentemente, quando $|\sigma_{-}h_{-}| = |\sigma_{+}h_{+}|$ é elevado (ultrapassa a unidade) torna-se problemática a aplicação do teste de variação quadrática, sendo nesses casos aconselhável dar preferência ao outro teste de decisão.

Se por um lado o teste de variação quadrática tem algumas limitações em termos de aplicação, por outro lado este teste revela-se mais eficaz nos intervalos próximos de zero mas que ainda correspondem a intervalos detectados. Em geral o teste de variação quadrática toma mais frequentemente decisões nessas zonas, comparativamente ao teste da razão de verosimilhança. Este comportamento é ilustrado nas Figuras 6.4 e 6.5 relativas ao caso 3.

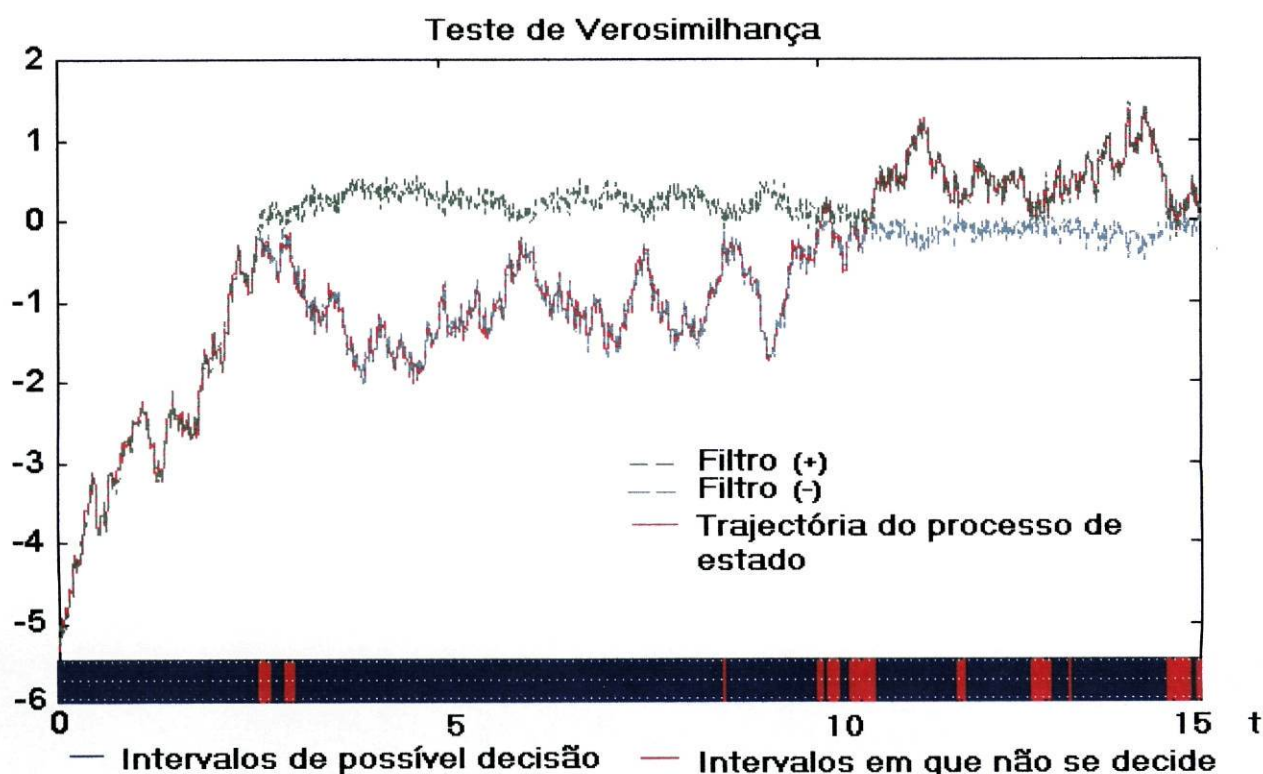


Figura 6.4 – Resultado da filtragem utilizando o teste de verosimilhança para uma trajetória do caso 3.

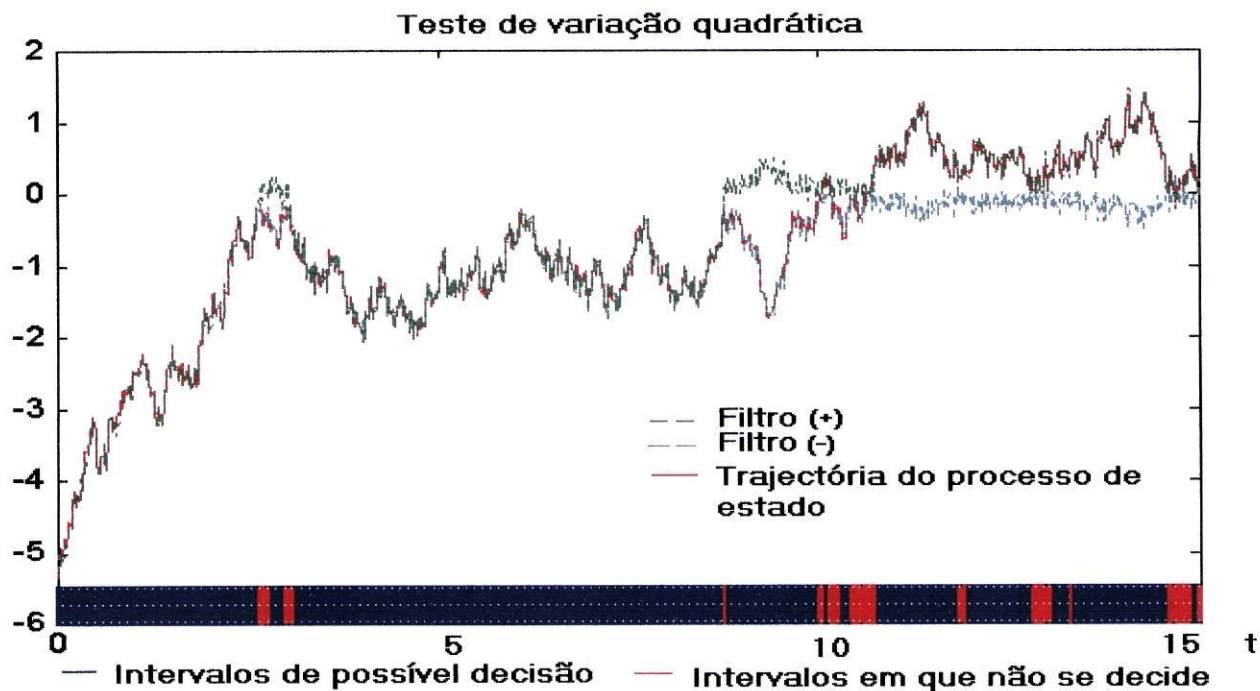


Figura 6.5 – Resultado da filtragem utilizando o teste de variação quadrática para uma trajetória do caso 3.

Como podemos ver, em ambos os testes, existem zonas no intervalo em tempo contínuo $[10,15]$ onde o teste de decisão sobre o sinal é aplicado, no entanto nenhum deles permite tomar uma decisão (nesses intervalos a estatística não chega a atingir os limites $-l_1$ e l_2). Para os instantes anteriores a $t=10$, pode-se destacar na Figura 6.5 um intervalo longo, identificado com barra azul, onde o teste de variação quadrática é aplicado enquanto que o teste da razão de verosimilhança (Figura 6.4) não conduz a uma decisão. O teste de variação quadrática opta correctamente pelo filtro do lado negativo nesse intervalo. Este é apenas um exemplo dos muitos em que situações como esta ocorrem. De facto, em zonas de detecção mais próximas de zero o teste de variação quadrática, se tiver um intervalo suficientemente longo para poder decidir, parece funcionar melhor do que o teste da razão de verosimilhança. Podemos dizer que este último tem geralmente melhor performance do que o teste de variação quadrática por estar menos sujeito a decisões erradas, mas perde performance no que respeita à percentagem de decisão.

Até ao momento todos os exemplos que foram analisados verificam a hipótese de detectabilidade (HD2). Para melhor estudarmos o comportamento dos testes vamos analisar de seguida alguns exemplos de aplicação que verificam a hipótese (HD1) e compará-los com os anteriores, sabendo de antemão que as percentagens de decisão terão de ser maiores e, sobretudo, os tempos de espera para a decisão mais curtos. Nesse sentido considerámos três exemplos identificados na Tabela 6.4 como casos 1' a 3'.

Casos	b_-	b_+	h_-	h_+	σ_-	σ_+	l_1	l_2	pid_{vq}	pe_{vq}	\bar{T}_{vq}^-	\bar{T}_{vq}^+
1'	-1	-0.05	-1	2	1	1	5.1772	2.1526	75.3%	4.33%	0.51	0.2
2'	-1	-0.01	-2	4	1	1	6.7053	1.1628	88.11%	4.99%	0.34	0.12
3'	-1	-0.01	-1.8	2	1	1	3.4734	2.9298	49.18%	1.16%	1.06	1.52

Tabela 6.4 – Resultados do teste de variação quadrática segundo a hipótese (HD1).

Dado que, para os modelos que verificam a hipótese (HD1), os dois testes foram já comparados entre si em Fleming et al (1991), incluímos aqui apenas os resultados para o teste de variação quadrática.

Tal como sob a hipótese (HD2), é necessário determinar os valores para l_1 e l_2 resolvendo um sistema de equações não linear, o sistema (5.11). A aproximação convenientemente sugerida por Fleming et al. (expressão (5.9)) permite encontrar soluções de uma forma rápida e que respeitam aproximadamente a condição de que as probabilidades de erro sejam de 5%.

Podemos ver pelos casos apresentados na Tabela 6.4 que, tal como era de esperar, nesta situação o teste apresenta tempos de espera para a decisão mais reduzidos, comparativamente às situações sob a hipótese (HD2) (Tabela 6.2). Na Figura 6.6 podemos ver o resultado da aplicação do teste de variação quadrática a uma trajectória do modelo apresentado no caso 1', identificado na Tabela 6.4.

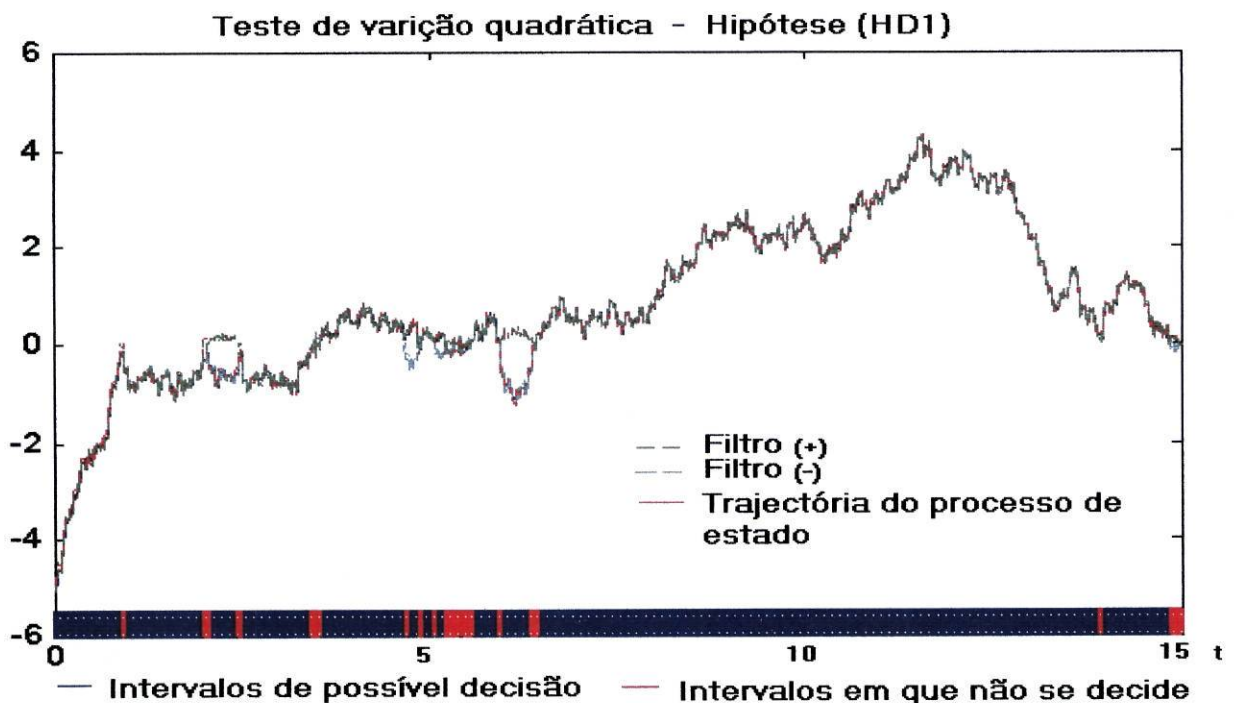


Figura 6.6 – Resultado da filtragem utilizando o teste de variação quadrática (hipótese (HD1), caso 1').

Note-se que, com a proximidade entre as quantidades $(\sigma_-h_-)^2$ e $(\sigma_+h_+)^2$, o teste necessita de mais tempo para poder decidir, o que vem de encontro ao que seria de esperar pela forma como foi construído.

Em qualquer um dos três casos de aplicação, obteve-se uma percentagem de erro inferior aos 5% (nível fixado). O teste de variação quadrática, que apresentara algumas limitações para a hipótese (HD2), nos modelos que verificam a hipótese (HD1) mostrou-se eficaz respondendo bem aos diferentes parâmetros do modelo (2.1). Podemos ver pelo gráfico da figura anterior que, para esta hipótese, o teste de variação quadrática permite tomar decisões sobre o sinal mesmo em intervalos de detecção (detectados pelo primeiro teste) com poucos instantes e até nos intervalos em que o processo de estado se encontra bastante próximo de zero, o que não acontecia para os modelos sob a hipótese (HD2). Enquanto que na hipótese (HD2) se tinha $(\sigma_-h_-)^2 = (\sigma_+h_+)^2$ e o teste de decisão sobre o sinal recorria aos valores de $b(x_k)$ para poder decidir, na hipótese (HD1) tem-se $(\sigma_-h_-)^2 \neq (\sigma_+h_+)^2$, o que por si só já é suficiente para distinguir os diferentes ramos de linearidade do estado a estimar. Neste caso, os valores de $b(x_k)$ assumem um papel menos importante, sendo os resultados de qualidade de estimação indiferentes às alterações dos mesmos (Fleming et al., 1991).

Em alguns dos casos apresentados, os resultados que obtivemos poderiam eventualmente vir a ser melhorados se considerássemos um passo de tempo menor, especialmente as situações em que o número de intervalos detectados é reduzido, o que aconteceu sobretudo sob a hipótese (HD2) (ver por exemplo os Casos 4, 7, 8 e 9 da Tabela A1, apêndice A). A diminuição do passo de tempo $\Delta t = \varepsilon$ acarretaria, no entanto, uma grande sobrecarga em termos computacionais, dificultando a análise de um número tão grande de casos de aplicação.

7. Conclusões

Ao longo deste trabalho pretendeu-se analisar os modelos de Markov escondidos unidimensionais lineares por bocados em tempo discreto, numa perspectiva de estimação do processo de estado não observado. Mediante uma hipótese de detectabilidade sobre o modelo, os estimadores tradicionais (Filtro de Kalman-Bucy) permitem estimar o estado especificamente, desde que precedidos da aplicação de uma sequência de dois testes estatísticos. Foram estudados o teste de detecção de passagens em zero, para o processo não observado, e dois testes de decisão sobre o sinal (teste da razão de verosimilhança e teste de variação quadrática). Estes testes podem ser desenhados para serem aplicados ao processo das observações ou ao próprio estimador, tendo-se neste trabalho optado pela primeira abordagem.

O desempenho do teste de variação quadrática, na decisão sobre o sinal foi comparado com o teste da razão de verosimilhança, sob a hipótese de detectabilidade (HD2). Para a análise da performance dos testes recorremos às percentagens de instantes de detecção e decisão e respectivas frequências de detecções e decisões erradas obtidas nas simulações. Foram ainda considerados os tempos médios de paragem que nos permitem analisar a rapidez de decisão de cada teste.

Confirmámos os resultados obtidos por Oliveira-Roubaud (1995) segundo os quais o teste de razão de verosimilhança fornecia bons resultados no procedimento de estimação do estado. Verificámos que o teste de variação quadrática não é geralmente melhor do que o anterior mas permite-nos por vezes obter estimativas em intervalos de tempo em que o teste da razão de verosimilhança não toma uma decisão, como é o caso de intervalos com poucos instantes ou em que o processo de estado toma valores muito próximos de zero. Verificámos no entanto existirem exemplos em que o teste de variação quadrática não fornece bons resultados e dificilmente cumpre os níveis de erro exigidos. Não se aconselha portanto a aplicação do teste de variação quadrática sem prévia verificação do comportamento sob simulação. Aparentemente isto acontece quando estão em causa valores elevados para $(\sigma h)^2(x)$. Nem sempre as probabilidades de erro de decisão foram iguais ou inferiores aos 5% (nível fixado). Um estudo mais demorado, com um número muito maior de simulações e um passo de tempo inferior seria necessário para averiguar das razões do aparecimento destes

valores de erro e verificar se são significativamente diferentes de 5%. O comportamento dos testes está muito dependente dos valores encontrados para os parâmetros l_1 e l_2 (valores críticos da estatística do teste) e seria necessário realizar um estudo numérico aprofundado para melhorar a determinação destes parâmetros (que resultam da resolução de um sistema de equações não linear). A aproximação sugerida por Fleming et al (1991) (ver expressão 5.9) nem sempre se revelou adequada para os modelos que verificam a hipótese de detectabilidade (HD2), embora se considere que a expressão é adequada para uso nos modelos em que se verifica a hipótese (HD1). A qualidade dos testes deveria melhorar consideravelmente se utilizássemos um passo de tempo inferior ao utilizado para a discretização do modelo (2.2), o que acarretaria no entanto uma grande sobrecarga em termos computacionais, pelo que nos limitámos a um estudo mais curto.

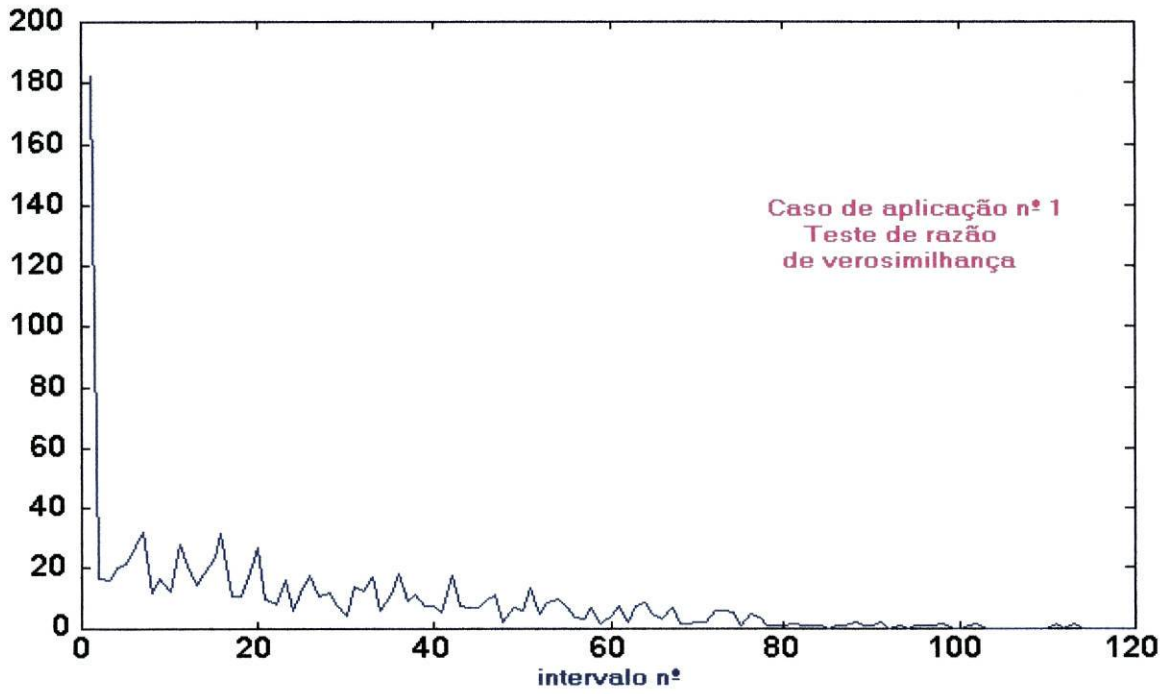
Nos modelos em estudo considerou-se o valor zero como o valor que separa as regiões de linearidade para o processo de estado (valor limiar), mas poder-se-ia considerar um outro valor diferente de zero e seguir um procedimento semelhante ao adoptado, mediante mudanças de variável convenientes. O modelo em tempo discreto considerado é unidimensional mas este trabalho poderia ser futuramente estendido, com as alterações devidas, a modelos em que os processos das observações e de estado têm dimensão superior a um. Grande parte dos resultados e propriedades que se aplicam aos problemas em causa foram já deduzidos teoricamente para modelos em tempo contínuo e generalizados para uma dimensão superior a um (Fleming-Pardoux, 1989, Roubaud, 1992 e Roubaud, 1995). No entanto não se conhecem extensões análogas para modelos em tempo discreto.

Seria interessante completar este trabalho com um estudo realizado sobre a aplicação destes testes num caso real.

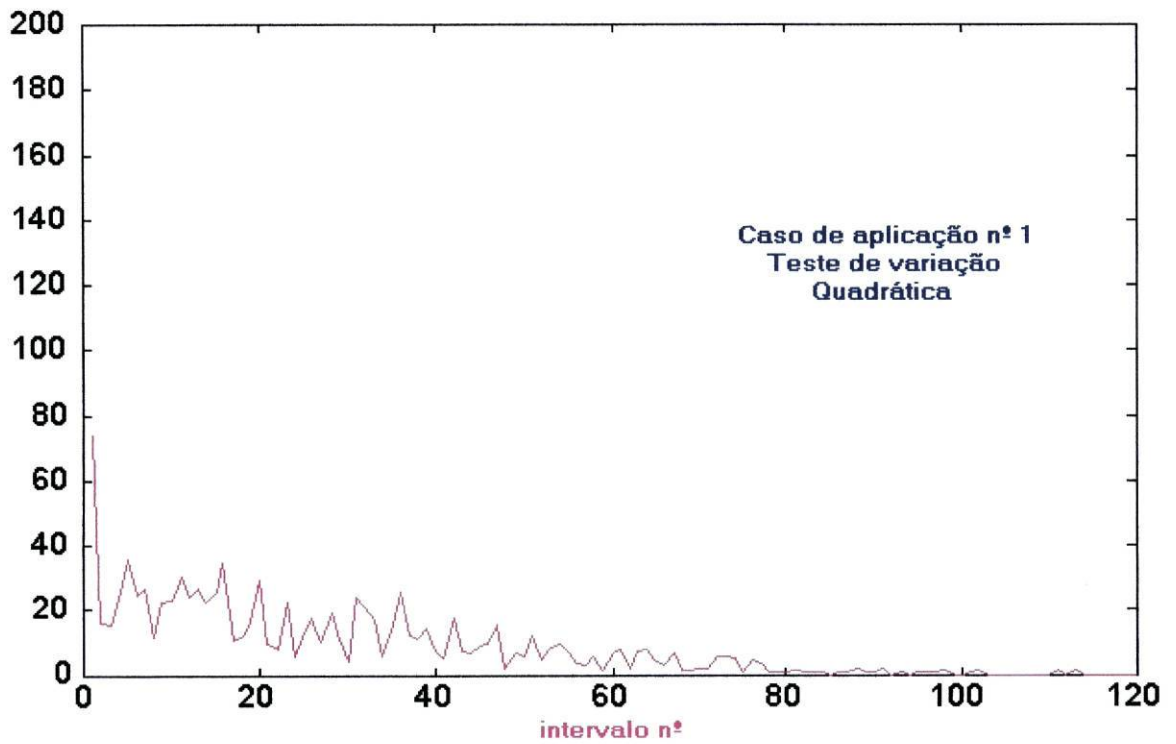
Apêndice A

**(Tabelas e Gráficos
de Resultados)**

Média do número de intervalos necessários para decidir em cada intervalo



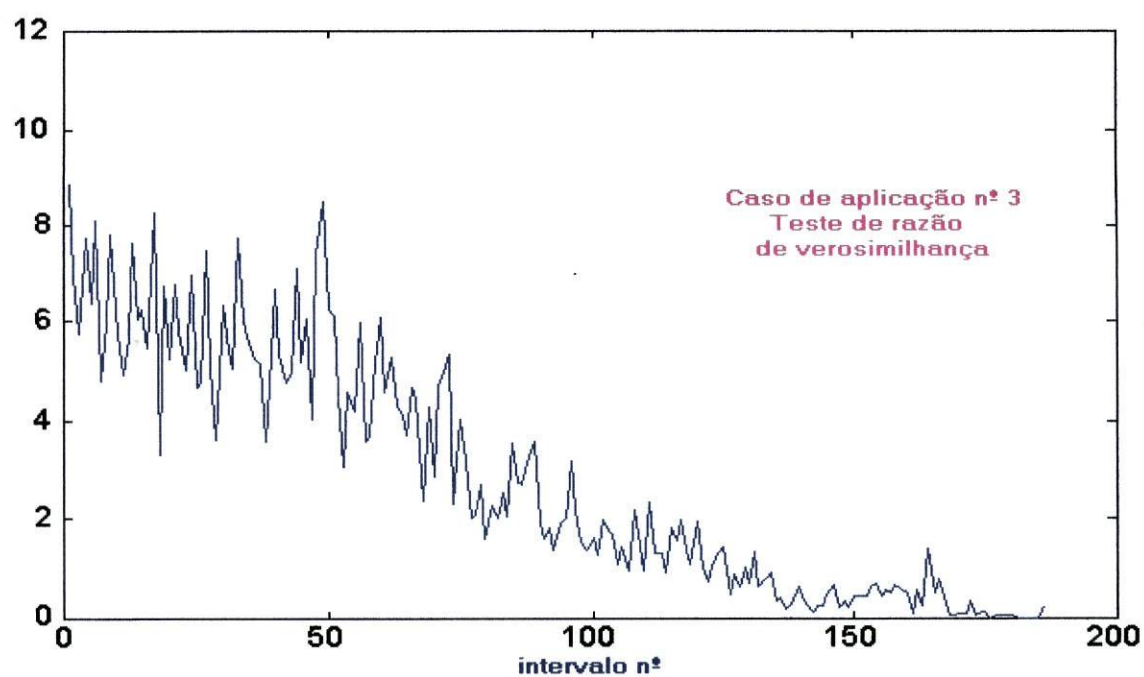
(a)



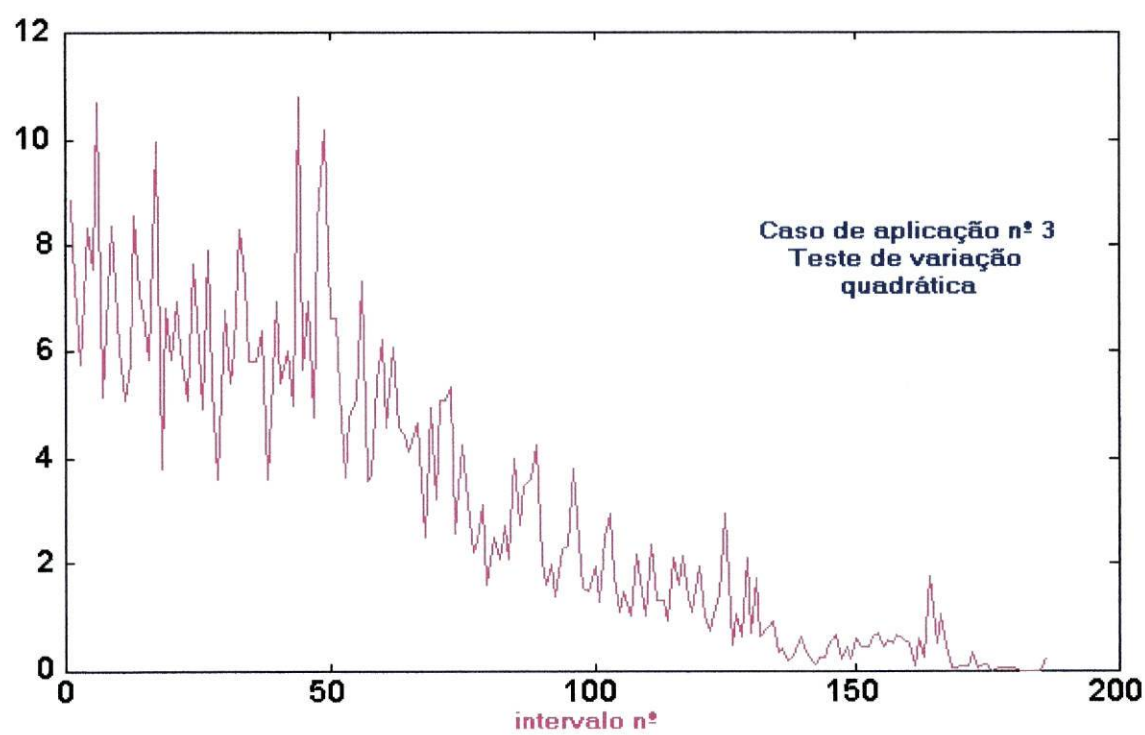
(b)

Figura A1 - Índice do intervalo em que se aplica o teste de decisão sobre o sinal num intervalo em tempo discreto $[0,1500]$ para o caso de aplicação nº 1: (a) teste de razão de verossimilhança; (b) teste de variação quadrática.

Média do número de intervalos necessários para decidir em cada intervalo



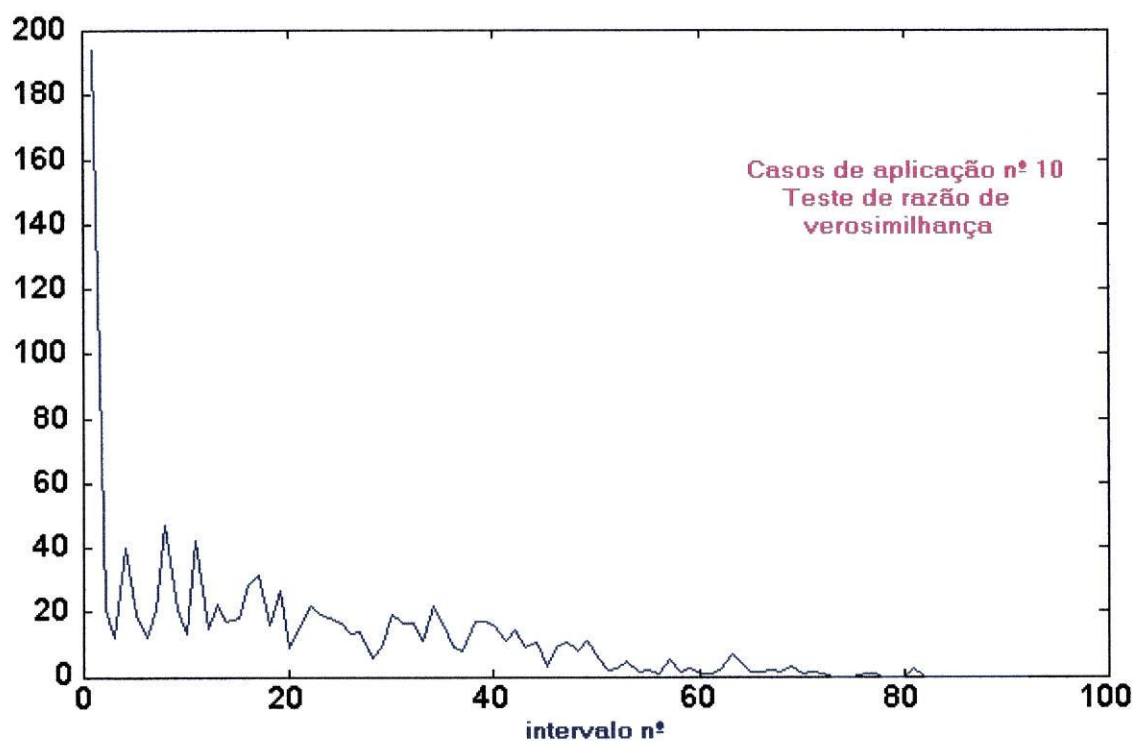
(a)



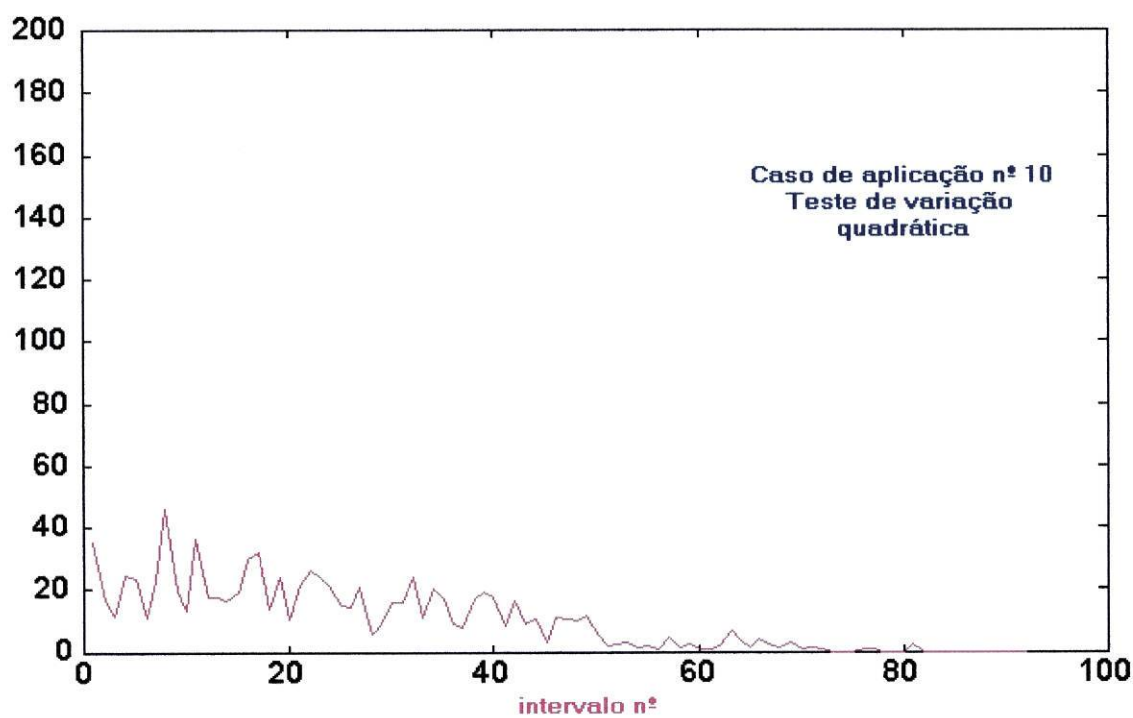
(b)

Figura A2 - Índice do intervalo em que se aplica o teste de decisão sobre o sinal num intervalo em tempo discreto $[0,1500]$ para o caso de aplicação n° 3: (a) teste de razão de verossimilhança; (b) teste de variação quadrática.

Média do número de intervalos necessários para decidir em cada intervalo



(a)



(b)

Figura A3 - Índice do intervalo em que se aplica o teste de decisão sobre o sinal num intervalo em tempo discreto [0,1500] para o caso de aplicação n° 10 : (a) teste de razão de verossimilhança; (b) teste de variação quadrática.

Tabela de Resultados para os Testes de Verosimilhança e Variação Quadrática

Hipótese (HD2)

Casos	b_-	b_+	h_-	h_+	σ_-	σ_+	c	θ_-	θ_+	l_1^{vq}	l_2^{vq}	S_v^1	S_v^2	S_{vq}^1	S_{vq}^2	pi_d	pe_d	pe_v	pe_{vq}	pid_v	pid_{vq}	\bar{T}_v^-	\bar{T}_v^+	ET_v^-	ET_v^+	\bar{T}_{vq}^-	\bar{T}_{vq}^+	ET_{vq}^-	ET_{vq}^+
	1	-0.5	-0.01	-1	1	1	1	0.1429	-0.3271	4.8948	0.612	4.5621	0.7791	1.7366	0.8577	2.072	62.53%	3.35%	6.56%	6.82%	41.46%	41.79%	1.2	2.68	21.99	0.44	1.46	6.66	2.94
2	-1	-0.05	-1	1	1	1	0.1429	-0.6384	3.1717	0.9445	3.5435	0.4434	0.5787	0.6842	1.1048	61.04%	4.07%	3.95%	4.41%	47.21%	34.74%	0.42	1.44	0.12	0.59	0.67	3.86	3.18	0.74
3	-4	-0.05	-1	1	1	1	0.1429	-2.7554	66.275	0.9898	1.1333	0.1953	0.1266	0.237	0.2153	58.05%	3.76%	4.77%	4.59%	61.36%	58.77%	0.09	0.37	2.66	0.03	0.17	0.5	0.89	0.01
4	-1	-0.05	-0.8	0.8	1	1	0.134	-0.7556	1.8396	1.6284	3.9646	0.4208	0.48	0.7337	1.1878	59.27%	2.75%	8.56%	1.72%	45.06%	25.98%	0.32	1.37	11.67	0.59	1.23	5.3	5.95	0.82
5	-1	-0.05	-2	4	1	0.5	0.252	-0.027	10.5298	0.2845	5.0276	0.3615	0.4502	0.6167	1.1387	62.36%	1.68%	1.98%	32.45%	55.83%	58.62%	0.44	1.83	11.67	0.59	0.71	2.94	0.08	1.06
6	-1	-0.05	-1.5	1.5	1	1	0.1701	-0.3246	6.8297	0.4386	3.8782	0.52	0.91	0.6692	1.4978	61.64%	4.46%	6.45%	15.31%	49.52%	49.28%	0.47	1.42	11.67	0.59	0.8	2.84	0.98	0.81
7	-1	-0.05	-0.6	0.6	1	1	0.1267	-0.8583	0.6801	4.4050	3.4904	0.3756	0.3935	1.0047	1.8857	57.63%	2.7%	1075%	0	44.2%	5.76%	0.35	1.08	11.67	0.59	--	6.67	17.69	0.69
8	-1	-0.05	-0.5	0.5	1	1	0.1238	-0.9018	0.1902	15.7497	3.3222	0.3348	0.2911	0.937	1.5474	57.2%	2.29%	9.64%	0	46.83%	5.52%	0.28	0.99	11.67	0.59	--	5.91	65.25	0.52
9	-1	-0.25	-1.5	1.5	1	1	0.1701	-0.3179	0.3866	7.6765	9.1517	0.8133	1.7087	1.3985	3.7536	61.63%	4.66%	10.16%	0	29.36%	3.92%	0.57	1.53	18.72	4.7	--	56.3	48.36	14.76
10	-0.5	-0.01	-1.5	1.5	1	1	0.1701	-0.1621	9.6449	0.3106	4.0657	0.983	2.6975	0.7758	2.0932	63.11%	5.05%	6.56%	16.67%	39.68%	59.18%	1.35	2.98	21.99	0.44	1.48	5.54	0.76	0.64

Tabela A1 – Resultados para casos de aplicação em que foram aplicados ambos os testes de decisão sobre o sinal, o da razão de verosimilhança e o de variação quadrática.

**Tabela de Resultados para o Teste de Variação Quadrática
Hipótese (HD1)**

Casos	b_-	b_+	h_-	h_+	σ_-	σ_+	c	θ_-	θ_+	l_1^{vq}	l_2^{vq}	S_{vq}^1	S_{vq}^2	pi_d	pe_d	pe_{vq}	pid_{vq}	\bar{T}_{vq}^-	\bar{T}_{vq}^+	ET_{vq}^-	ET_{vq}^+
	1'	-1	-0,05	-1	2	1	1	0,165	-1,2706	0,5786	5,1772	2,1526	0,1667	1,1897	60,98%	4,37%	4,33%	75,3%	0,51	0,2	0,49
2'	-1	-0,01	-2	4	1	1	0,252	-1,8461	0,4468	6,7053	1,6228	0,14	0,39	63,04%	6,08%	4,99%	88,11%	0,34	0,12	0,29	0,03
3'	-1	-0,01	-1,8	2	1	1	0,1952	-1,0225	0,8625	3,4734	2,9298	0,5576	1,1057	62,24%	5,08%	1,16%	49,18%	1,06	1,52	6,81	5,14

Tabela A2 – Resultados do teste de variação quadrática (hipótese (HD1)) .

Tabela de Resultados para o Teste de Razão de Verossimilhança

Hipótese (HD2)

Casos	b_-	b_+	h_-	h_+	σ_-	σ_+	c	S_v^1	S_v^2	pi_d	pe_d	pe_v	pid_v	\bar{T}_v^-	\bar{T}_v^+	ET_v^-	ET_v^+
	-1	0.25	-1	1	1	1								0.6	1.58	18.72	4.7
11	-1	0.25	-1	1	1	1	0.1429	0.633	1.019	60.02%	3.86%	5%	29.68%	0.6	1.58	18.72	4.7
12	-1	-0.25	-1	1	1	1	0.1429	0.9395	1.9011	59.31%	4.06%	10.53%	8.53%	0.69	2.28	42.12	21.12
13	-1	-0.25	-0.5	0.5	1	1	0.1253	0.4391	0.4419	55.34%	2.02%	11.33%	26.71%	0.43	1.24	18.72	4.7
14	-1	-0.25	-0.5	0.5	1.5	1.5	0.1321	0.7703	1.1388	58.71%	2.71%	12.86%	22.7%	0.68	1.61	18.72	4.7
15	-1	-0.25	-1	1	0.5	0.5	0.2238	0.3702	0.4484	56.62%	2.01%	17.9%	32.84%	0.13	1.08	18.72	4.7

Tabela A3 – Resultados do teste da razão de verossimilhança (hipótese (HD2)). Nestes casos não foi possível aplicar o teste de variação quadrática.

Nota: Considerámos $I_1^v = I_2^v = 2.94$.

Apêndice B

**(Rotinas em Matlab
para a obtenção
de resultados práticos)**

```
clc
clear
```

```
% simulação dos processos de estado e observação sob indicação dos valores a
% considerar para os parâmetros do modelo
```

```
epsilon=0.01;
b_mais=0;sigma_mais=1;h_mais=1;
b_menos=0;sigma_menos=1;h_menos=-2;
n=input('Quantas trajetórias quer simular ? ');
k=input('E quantos instantes deverá ter cada trajetória ? ');
aux1=sqrt((h_mais^2)*(h_menos^2)*(sigma_mais^2)+(h_menos^2)+(h_mais^2));
aux2=sqrt((h_mais^2)*(h_menos^2)*(sigma_menos^2)+(h_menos^2)+(h_mais^2));
c=((1.65*sqrt(epsilon))/abs(h_menos-h_mais))*max(aux1,aux2)
for j=1:n
    x(j,1)=sqrt(0.1)*randn-1;
    if x(j,1)>0
        y(j,1)=h_mais*x(j,1)+sqrt(epsilon)*randn;
    else
        y(j,1)=h_menos*x(j,1)+sqrt(epsilon)*randn;
    end
    for i=2:k+1
        if x(j,i-1)>0
            b=b_mais;sigma=sigma_mais;h=h_mais;
        else
            b=b_menos;sigma=sigma_menos;h=h_menos;
        end
        x(j,i)=(1+epsilon*b)*x(j,i-1)+sqrt(epsilon)*sigma*randn;
        y(j,i)=h*x(j,i)+sqrt(epsilon)*randn;
    end
end
end
```

```
clc
clear
load c:\tesel\dados.mat
```

% Divisão por intervalos em que o processo de estado não cruza o valor zero
% Calcula o erro na detecção de cruzamentos e calcula a percentagem de
% instantes para os intervalos de detecção

```
intervalos=[];
erros=[];
soma=0;
[k1 k2]=size(y);
for j=1:k1
    if abs(y(j,1))>=c
        k=1;
        t1=1;
    else
        k=0;
        t1=1;
    end
    interv=0;
    er=0;
    for i=2:1001
        if k==1
            if abs(y(j,i))<c
                t2=i-1;
                k=0;
                interv=interv+1;
                valor=sum(sign(x(j,t1:t2)));
                soma=soma+t2-t1+1;
                if abs(valor)~=t2-t1+1
                    er=er+1;
                end
                t1=i;
            end
        elseif abs(y(j,i))>=c
            t2=i-1;
            %interv=interv+1;
            k=1;
            %valor=sum(sign(x(j,t1:t2)));
            %if abs(valor)~=t2-t1+1;
            % er=er+1;
            %end
            t1=i;
        end
    end
    if i==1001 & k==1
        t2=i;
        interv=interv+1;
    end
end
```

```
soma=soma+t2-t1+1;
valor=sum(sign(x(j,t1:t2)));
if abs(valor)~=t2-t1+1;
    er=er+1;
end
end
end
intervalos=[intervalos interv];
erros=[erros er];
end
disp('A frequência de erro é :')
freq_erro=sum(erros)/sum(intervalos)
percentagem=soma/(k1*k2)
```

```
function F=solucao(l)
```

```
% Função que permite calcular as probabilidades de erro do  
% tipo I, dados os valores l1,l2 e theta_mais,theta_menos.
```

```
theta_mais=input('Qual o valor de theta_mais a considerar ?');  
theta_menos=input('Qual o valor de theta_menos a considerar ?');  
p_mais=0.05;  
p_menos=0.05;  
F=[((1-exp(-theta_mais*l(2)))/(exp(theta_mais*l(1))-exp(-theta_mais*l(2))))];  
((1-exp(theta_menos*l(1)))/(exp(-theta_menos*l(2))-exp(theta_menos*l(1))))];
```

```
function F=solucao1(l)
```

```
% Função utilizada como parâmetro de entrada na função 'fsolve'  
% para calcular a solução ótima para o par (l1,l2)
```

```
theta_mais=input('Qual o valor de theta_mais a considerar ?');  
theta_menos=input('Qual o valor de theta_menos a considerar ?');  
p_mais=0.05;  
p_menos=0.05;  
F=[(((1-exp(-theta_mais*l(2)))/(exp(theta_mais*l(1))-exp(-theta_mais*l(2))))-p_mais;  
((1-exp(theta_menos*l(1)))/(exp(-theta_menos*l(2))-exp(theta_menos*l(1))))-  
p_menos];
```

```
Function [theta_mais,theta_menos,l1,l2,F] = var_quad (b_mais, b_menos, h_mais, h_menos, sigma_mais, sigma_menos, epsilon)
```

```
% função que permite calcular o valor das probabilidades de erro de tipo I,  
% atendendo à aproximação sugerida para a determinação dos valores de l1 e l2.  
% Teste de variação quadrática, sob a hipótese (HD2)
```

```
clc  
gama_mais=sqrt((h_mais^2)*(sigma_mais^2)+1+(1+b_mais*epsilon)^2)  
gama_menos=sqrt((h_menos^2)*(sigma_menos^2)+1+(1+b_menos*epsilon)^2)  
  
sigma=sqrt(4*(1+(sigma_mais*h_mais)^2)/((gama_mais*gama_menos)^2));  
  
fi_mais=sqrt((sigma/4)*sqrt(2+4*((gama_mais^4)+2)/((gama_menos*gama_mais)^4))  
);  
fi_menos=sqrt((sigma/4)*sqrt(2+4*((gama_menos^4)+2)/((gama_menos*gama_mais)^4))  
^4));  
  
theta_mais=(log(gama_menos/gama_mais)+((h_mais^2)*(sigma_mais^2)*((b_mais-  
b_menos)^2)*epsilon)/(4*(gama_menos^2)*abs(b_mais)))/((fi_mais^2)*epsilon);  
  
theta_menos=(log(gama_menos/gama_mais)+((h_menos^2)*(sigma_menos^2)*((b_  
mais-  
b_menos)^2)*epsilon)/(4*(gama_mais^2)*abs(b_menos)))/((fi_menos^2)*epsilon);
```

```
l1=-(1/abs(theta_mais))*log(0.05);  
l2=-(1/abs(theta_menos))*log(0.05);
```

```
[F]=solucao([l1 l2]);
```

```
function  
[theta_mais,theta_menos,l1,l2,F]=v_qa(b_mais,b_menos,h_mais,h_menos,sigma_m  
ais,sigma_menos,epsilon)
```

```
% função que permite calcular o valor das probabilidades de erro de tipo I,  
% atendendo à aproximação sugerida para a determinação dos valores de l1 e l2.  
% Teste de variação quadrática, sob a hipótese (HD1)
```

```
clc  
gama_mais=sqrt(((h_mais*sigma_mais)^2)+1+((1+b_mais*epsilon)^2))  
gama_menos=sqrt(((h_menos*sigma_menos)^2)+1+((1+b_menos*epsilon)^2))  
  
aux=0.5*((1/(gama_menos^2))-(1/(gama_mais^2)))  
miu_mais=log(gama_menos/gama_mais)+(aux*(gama_mais^2))
```

```
miu_menos=log(gama_menos/gama_mais)+(aux*(gama_menos^2))

fi_mais=aux*sqrt((gama_mais^4)+2*((1+b_mais*epsilon)^2))
fi_menos=aux*sqrt((gama_menos^4)+2*((1+b_menos*epsilon)^2))

theta_mais=miu_mais/(fi_mais^2);
theta_menos=miu_menos/(fi_menos^2);

l1=-(1/abs(theta_mais))*log(0.05);
l2=-(1/abs(theta_menos))*log(0.05);

[F]=solucao([l1 l2]);
```

```
clear
clc
load c:\tese\dados.mat
```

```
% Calcula o estimador com os dois filtros de Kalman-Bucy em paralelo de acordo
% com a escolha do sinal nos intervalos de detecção
% Teste de razão de verosimilhança
```

```
gama=sigma_mais*h_mais;
F_mais=(1+b_mais*epsilon);
F_menos=(1+b_menos*epsilon);
aux1=(b_mais-b_menos);
l=2.94;
n0=5;
```

```
% Divisão por intervalos sem cruzamentos em zero
```

```
sigma_m=sigma;
sigma_M=sigma;
[k1,k2]=size(y);
nInt1=0;
nTempos1=0;
nInt2=0;
nTempos2=0;
nE=0;
for i=1:k1
    ntempos=0;
    t1=0;
    t2=0;
    x_e_m(i,1)=-5;
    x_e_M(i,1)=-5;
    instantes=find(abs(y(i,:))>=c);
    if instantes(1)==1
        t1=1;
    end
    j=2;
    while j<=k2
        if t1==0 & find(instantes==j)~=[]
            t1=j;
        else
            x_p_m=F_menos*x(i,j-1);
            x_m=F_menos*x(i,j-1)+sqrt(epsilon)*randn;
            y_m=h_menos*x(i,j)+sqrt(epsilon)*randn;
            sigma_p_m=F_menos*sigma_m*F_menos'+1;
            k_m=sigma_p_m*h_menos*inv(h_menos*sigma_p_m*h_menos'+1);
            sigma_m=sigma_p_m-k_m*h_menos*sigma_p_m;
            x_e_m(i,j)=x_p_m+k_m*(y(i,j)-h_menos*x_p_m);
            x_p_M=F_mais*x(i,j-1);
            x_M=F_mais*x(i,j-1)+sqrt(epsilon)*randn;
```

```

y_M=h_mais*x(i,j)+sqrt(epsilon)*randn;
sigma_p_M=F_mais*sigma_M*F_mais'+1;
k_M=sigma_p_M*h_mais'*inv(h_mais*sigma_p_M*h_mais'+1);
sigma_M=sigma_p_M-k_M*h_mais*sigma_p_M;
x_e_M(i,j)=x_p_M+k_M*(y(i,j)-h_mais*x_p_M);
end
if t1~=0
k=0;
while k==0
    if find(instantes==j+1)==[]
        k=1;
        t2=j;
    else
        k=0;
        j=j+1;
    end
end

%
if t1==1
    t1=2;
end
s=0;
for t=t1:t2
    x_p_M=F_mais*x(i,t-1);
    x_M=F_mais*x(i,t-1)+sqrt(epsilon)*randn;
    y_M=h_mais*x(i,t)+sqrt(epsilon)*randn;
    sigma_p_M=F_mais*sigma_M*F_mais'+1;
    k_M=sigma_p_M*h_mais'*inv(h_mais*sigma_p_M*h_mais'+1);
    sigma_M=sigma_p_M-k_M*h_mais*sigma_p_M;
    x_e_M(i,t)=x_p_M+k_M*(y(i,t)-h_mais*x_p_M);
    x_p_m=F_menos*x(i,t-1);
    x_m=F_menos*x(i,t-1)+sqrt(epsilon)*randn;
    y_m=h_menos*x(i,t)+sqrt(epsilon)*randn;
    sigma_p_m=F_menos*sigma_m*F_menos'+1;
    k_m=sigma_p_m*h_menos'*inv(h_menos*sigma_p_m*h_menos'+1);
    sigma_m=sigma_p_m-k_m*h_menos*sigma_p_m;
    x_e_m(i,t)=x_p_m+k_m*(y(i,t)-h_menos*x_p_m);
end
if instantes(1)==1 & t1==2
    t1=1;
end
ntempos=ntempos+1;
tempos1(i,ntempos)=t2-t1+1;
tempos2(i,ntempos)=t2-t1+1;
for t=t1+n0:t2
    n=t-t1-n0;
    aux2=y(i,t1+n0+n)^2-y(i,t1+n0)^2;
L_aux=(aux1/(2*(gama^2)))*aux2-0.5*n*epsilon*aux1-epsilon*((b_mais^2-b_menos^2)/(2*(gama^2)))*sum(y(i,t1+n0:t1+n0+n-1).^2);
if (L_aux>=l | L_aux<=-l) & s==0 & t>=t1

```

```

tempos1(i,ntempos)=t-t1+1;
tempos2(i,ntempos)=t2-t1+1;
if (t2-t1)>=n0+1
    if L_aux>=l
        x_e_m(i,t1:t2)=x_e_M(i,t1:t2);
        sigma_m=sigma_M;
        nInt1=nInt1+1;
        nTempos1=nTempos1+(t-t1+1);
        if sum(sign(x(i,t1:t2)))~=(t2-t1+1) & abs(sum(sign(x(i,t1:t2
)))) == abs(t2-t1+1)
            nE=nE+1;
        end
    elseif L_aux<=-l
        x_e_M(i,t1:t2)=x_e_m(i,t1:t2);
        sigma_M=sigma_m;
        nInt2=nInt2+1;
        nTempos2=nTempos2+(t-t1+1);
        if sum(sign(x(i,t1:t2)))~=- (t2-t1+1) & abs(sum( sign( x( i,t1:
t2 )))) == abs(t2-t1+1)
            nE=nE+1;
        end
    end
end
s=1;
elseif t==t2 & s==0
    tempos1(i,ntempos)=t-t1+1;
    tempos2(i,ntempos)=t2-t1+1;
end
end
t1=0;
j=t2+1;
else
    j=j+1;
end
end
end
media_instantes1=mean(tempos1(1:k1,1:50));
media_instantes2=mean(tempos1(1:k1,1:100));
percentagem_media=sum(sum(tempos1(1:k1,:)))/sum(sum(tempos2(1:k1,:)));
media_tempos_mais=nTempos1/nInt1
media_tempos_menos=nTempos2/nInt2
esperado_mais=((-2*b_mais*(2+epsilon*b_mais))/((b_mais-b_menos)^2))*(l-0.05*2*l)
esperado_menos = ((-2*b_menos*(2+epsilon*b_menos)) / ((b_mais-b_menos)^2)) *
(l-0.05*2*l)
soma_intervalos_mais=nInt1
soma_intervalos_menos=nInt2
total_intervalos=nInt1+nInt2
total_erro=nE
percentagem_erro=nE/(nInt1+nInt2)

```

```
clear
clc
load c:\tese\dados.mat

% Calcula o estimador com os dois filtros de Kalman-Bucy em paralelo de acordo
% com a escolha do sinal nos intervalos de detecção
% Teste de variação quadrática

gama=sigma_mais*h_mais;
F_mais=(1+b_mais*epsilon);
F_menos=(1+b_menos*epsilon);
aux1=(b_mais-b_menos);
l1=input('Qual é o valor de l1? ');
l2=input('Qual é o valor de l2? ');
gama_mais=sqrt((h_mais^2)*(sigma_mais^2)+1+(1+b_mais*epsilon)^2);
gama_menos=sqrt((h_menos^2)*(sigma_menos^2)+1+(1+b_menos*epsilon)^2);
n0=5;

% Divisão por intervalos sem cruzamentos em zero

sigma_m=sigma;
sigma_M=sigma;
[k1,k2]=size(y);
nInt1=0;
nTempos1=0;
nInt2=0;
nTempos2=0;
nE=0;
for i=1:k1
    ntempos=0;
    t1=0;
    t2=0;
    x_e_m(i,1)=-5;
    x_e_M(i,1)=-5;
    instantes=find(abs(y(i,:))>=c);
    if instantes(1)==1
        t1=1;
    end
    j=2;
    while j<=k2
        if t1==0 & find(instantes==j)~=[]
            t1=j;
        else
            x_p_m=F_menos*x(i,j-1);
            x_m=F_menos*x(i,j-1)+sqrt(epsilon)*randn;
            y_m=h_menos*x(i,j)+sqrt(epsilon)*randn;
            sigma_p_m=F_menos*sigma_m*F_menos'+1;
            k_m=sigma_p_m*h_menos'*inv(h_menos*sigma_p_m*h_menos'+1);
```

```

sigma_m=sigma_p_m-k_m*h_menos*sigma_p_m;
x_e_m(i,j)=x_p_m+k_m*(y(i,j)-h_menos*x_p_m);

x_p_M=F_mais*x(i,j-1);
x_M=F_mais*x(i,j-1)+sqrt(epsilon)*randn;
y_M=h_mais*x(i,j)+sqrt(epsilon)*randn;
sigma_p_M=F_mais*sigma_M*F_mais'+1;
k_M=sigma_p_M*h_mais'*inv(h_mais*sigma_p_M*h_mais'+1);
sigma_M=sigma_p_M-k_M*h_mais*sigma_p_M;
x_e_M(i,j)=x_p_M+k_M*(y(i,j)-h_mais*x_p_M);
end
if t1~=0
    k=0;
    while k==0
        if find(instantes==j+1)==[]
            k=1;
            t2=j;
        else
            k=0;
            j=j+1;
        end
    end
end
if t1==1
    t1=2;
end
s=0;
for t=t1:t2
    x_p_M=F_mais*x(i,t-1);
    x_M=F_mais*x(i,t-1)+sqrt(epsilon)*randn;
    y_M=h_mais*x(i,t)+sqrt(epsilon)*randn;
    sigma_p_M=F_mais*sigma_M*F_mais'+1;
    k_M=sigma_p_M*h_mais'*inv(h_mais*sigma_p_M*h_mais'+1);
    sigma_M=sigma_p_M-k_M*h_mais*sigma_p_M;
    x_e_M(i,t)=x_p_M+k_M*(y(i,t)-h_mais*x_p_M);

    x_p_m=F_menos*x(i,t-1);
    x_m=F_menos*x(i,t-1)+sqrt(epsilon)*randn;
    y_m=h_menos*x(i,t)+sqrt(epsilon)*randn;
    sigma_p_m=F_menos*sigma_m*F_menos'+1;
    k_m=sigma_p_m*h_menos'*inv(h_menos*sigma_p_m*h_menos'+1);
    sigma_m=sigma_p_m-k_m*h_menos*sigma_p_m;
    x_e_m(i,t)=x_p_m+k_m*(y(i,t)-h_menos*x_p_m);

end
if instantes(1)==1 & t1==2
    t1=1;
end
ntempos=ntempos+1;
tempos1(i,ntempos)=t2-t1+1;
tempos2(i,ntempos)=t2-t1+1;
for t=t1+n0:t2-1

```

```

n=t-t1-n0;
soma1=sum((y(i,t1+n0+1:t+1)-(1+b_menos*epsilon)*y(i,t1+n0:t)).^2);
soma2=sum((y(i,t1+n0+1:t+1)-(1+b_mais*epsilon)*y(i,t1+n0:t)).^2);
S_aux=((n+1)*log(gama_menos/gama_mais))+((2*epsilon)^(-1))*((gama_menos^(-2))*soma1-(gama_mais^(-2))*soma2);

if (S_aux>=l2 | S_aux<=-l1) & s==0 & t>=t1
    tempos1(i,n tempos)=t-t1+2;
    tempos2(i,n tempos)=t2-t1+1;
    if (t2-t1)>=n0+1
        if S_aux>=l2
            x_e_m(i,t1:t2)=x_e_M(i,t1:t2);
            sigma_m=sigma_M;
            nInt1=nInt1+1;
            nTempos1=nTempos1+(t-t1+2);
            if sum(sign(x(i,t1:t2)))~=(t2-t1+1) & abs(sum(sign(
x(i,t1:t2))))==abs(t2-t1+1)
                nE=nE+1;
            end
        elseif S_aux<=-l1
            x_e_M(i,t1:t2)=x_e_m(i,t1:t2);
            sigma_M=sigma_m;
            nInt2=nInt2+1;
            nTempos2=nTempos2+(t-t1+2);
            if sum(sign(x(i,t1:t2)))~=- (t2-t1+1) & abs(sum(sign(
x(i,t1:t2))))==abs(t2-t1+1)
                nE=nE+1;
            end
        end
    end
end
end
s=1;
elseif t==t2-1 & s==0
    tempos1(i,n tempos)=t2-t1+1;
    tempos2(i,n tempos)=t2-t1+1;
end
end
t1=0;
j=t2+1;
else
    j=j+1;
end
end
end
media_instantes1=mean(tempos1(1:k1,1:50));
media_instantes2=mean(tempos1(1:k1,1:100));
percentagem_media=sum(sum(tempos1(1:k1,:)))/sum(sum(tempos2(1:k1,:)));
media_tempos_mais=nTempos1/nInt1
media_tempos_menos=nTempos2/nInt2
esperado_mais=((-2*b_mais*(2+epsilon*b_mais))/((b_mais-b_menos)^2))*(l2-0.05*(l1+l2))

```

```

esperado_menos=((-2*b_menos*(2+epsilon*b_menos))/((b_mais-b_menos)^2))*(11-
0.05*(l1+l2))
soma_intervalos_mais=nInt1
soma_intervalos_menos=nInt2
total_intervalos=nInt1+nInt2
total_erros=nE
percentagem_erro=nE/(nInt1+nInt2)

```

```

clc
clear
load dadosa.mat

```

```

% Calcula o erro de estimação do processo de estado após a utilização
% dos testes na decisão sobre o sinal

```

```

[k1 k2]=size(y);
soma=0;
for i=1:k1
    for j=1:k2
        linha(i,j)=max(abs(x(i,j)-x_e_m(i,j))^2,abs(x(i,j)-x_e_M(i,j))^2);
        linha_abs(i,j)=max(abs(x(i,j)-x_e_m(i,j)),abs(x(i,j)-x_e_M(i,j)));
    end
end
coluna_media=mean(linha);
figure(1)
plot((0:k2-1),coluna_media)
title('Erro médio por instante de estimação')
linha_media=mean(linha');
figure(2)
plot((1:k1),linha_media)
title('Erro médio por trajectória')
erro_medio=mean(linha_media)
erro_medio_absoluto=mean(mean(linha_abs))

```

```

*****

clear
clc
load c:\tese\dadosa.mat

% Calcula a percentagem de instantes em que os testes permitem
% decidir sobre o sinal

nTemposDecisao=0;
[k1 k2]=size(y);
for i=1:k1
    k=length(find(x_e_m(i,:)==x_e_M(i,:)));
    nTemposDecisao=nTemposDecisao+k;
end
disp('A percentagem de instantes de decisão é :')
percentagem=nTemposDecisao/(k1*k2)

```

```

*****

load dadosa.mat

% permite representar o processo de estado e respectivos resultados de estimação
% constrói uma barra a duas cores que distingue os intervalos seleccionados pelo
% teste de detecção e nos quais se pode decidir sobre o sinal

i=input('Qual é o nº da trajectória pretendida?');
[t1 t2]=size(y(i,:));
instantes=find(x_e_m(i,:)~=x_e_M(i,:));
%plot(instantes,x_e_m(i,instantes),'r')
k=min(min(x_e_m(i,:),x_e_M(i,:)),x(i,:));

k1=[];k2=[];
for s=1:t2
    if abs(y(i,s))>=c
        k1=[k1 s];
    else
        k2=[k2 s];
    end
end
fig1=(-6:0.01:-5.5)*ones(1,length(k1)); fig2=(-6:0.01:-5.5)*ones(1,length(k2));
figure(1)
plot((0:t2-1),x(i,:), 'c',(0:t2-1),x_e_m(i,:), 'r',(0:t2-1),x_e_M(i,:), 'm:')
hold on
plot(k1,fig1,'y.')
hold on
plot(k2,fig2,'c.')

```

Bibliografia

- Alpuim, T. (1999). *Modelos em espaço de estados, filtro de Kalman e algumas aplicações*, Ciclo de Seminários em Estatística e Aplicações, F.E.U.P.
- Andrews, Angus P. e Grewal, Mohinder S. (1993). *Kalman Filtering, Theory and Practice*, Prentice Hall, New Jersey.
- Fleming, W. H. e Pardoux, E. (1989). *PieceWise monotone filtering with small observation noise*, SIAM J. Control and Optimization, vol. 27, No. 5, pp. 1156-1181.
- Fleming, W. H, Ji, D. e Pardoux, E. (1988). *PieceWise linear filtering with small observation noise*, Proc. 8th Int. Conf. on Analysis and Optim. Of Systems, Antibes 1988, Lect, No. 111, pp. 725-739, Springer, 1988.
- Fleming, W. H., Ji, D., Salame, P. e Zhang, Q. (1991). *Piecewise monotone filtering in discrete-Time with small observation noise*, IEEE Transactions on Automatic Control, Vol. 10, No. 10, pp 1181-1186.
- Harvey, Andrew C. (1989). *Forecasting, Structural time series models and Kalman filter*, Department of Statistical and mathematical Sciences, London.
- Jazwinski, Andrew H. (1970). *Stochastic processes and filtering theory*, Academic Press.

- Kalman, R. E. e Bucy, R. S. (1961). *New results in linear filtering and prediction theory*, J. Basic Eng., vol. 83, pp. 95-108.
- Oliveira, P. Milheiro (1994). *Uma Versão discreta do Teorema de Girsanov*, II Congresso Anual da Sociedade Portuguesa de Estatística, Luso, pp. 211-217.
- Oliveira, P. Milheiro e Roubaud, Marie-Christine (1991). *Filtrage lineaire par morceaux d'un systeme en temps discret avec petiti bruit d'observation*, Unité de Recherche, INRIA-Sophia Antipolis, No. 1451.
- Oliveira, P. Milheiro e Roubaud, M. C. (1995). *Discrete-Time piecewise linear filtering with small observation noise*, IEEE Transactions on Automatic Control, Vol. 40, No. 12, pp 2149-2152.
- Oliveira, P. Milheiro (2000). *A sequential method for the estimation of a partially observed piecewise AR(1) process*, pp 1-5 (preprint).
- Pardoux, E. e Roubaud, M.C. (1991), *Finite dimensional approximate filters in the case of high signal-to-noise ratio*, Conference on the Memory of Zakai, Haifa.
- Picard, J. (1986). *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*, SIAM J. APPL. MATH., vol. 46, No. 6, pp. 1098-1125.
- Picard, J. (1991). *Efficiency of the extended Kalman filter for non linear systems with small noise*, , SIAM J. APPL. MATH, (51), 3, pp. 843-855.

- Roubaud, M. C. (1992). *An approximate filter for a partially observed piecewise linear system with small observation noise*, Systems & Control Letters, North-Holland, No. 20, pp. 239-247.
- Roubaud, M. C. (1995). *Filtrage linéaire par morceaux avec petit bruit d'observation*, Applied Mathematics Optimization, Springer-Verlag, New York, No. 32, pp. 163-194.
- Sorenson, H. W. (1985). *Kalman Filtering: Theory and application*, Department of Applied Mechanics and Engineering Sciences, University of California, San Diego.
- Sowers, Richard B. e Makowski, Armand M. (1992). *Discrete-time filtering for linear systems with non-Gaussian initial conditions : Asymptotic behavior of the difference between the MMSE and LMSE estimates*, IEEE Transactions on Automatic Control, vol. 37, No. 1, pp. 114-120.
- Tong, Howell e Yeung, Iris (1989). *On tests for self-exciting threshold autoregressive-type non-linearity in partially observed time series*, Reviste.
- Tong, Howell (1990). *Non-linear Time Series, A Dynamical system approach*, Clarendon Press Oxford.
- Turkman, K. F. (1998). *Nonlinear processes and models*, Centro Internacinal de Matemática, pp 30-38.