MODELIZAÇÃO ESPECTRAL AUTOREGRESSIVA DE PROCESSOS ESTOCÁSTICOS PONTUAIS:

MÉTODOS, ALGORITMOS E APLICAÇÕES

Ana Paula Rocha

Dezembro 1992

Modelização espectral autoregressiva de processos estocásticos pontuais: métodos, algoritmos e aplicações

Ana Paula de Frias Viegas Proença Rocha

Tese submetida para provas de Doutoramento

Grupo de Matemática Aplicada
Faculdade de Ciências
Universidade do Porto

Dezembro de 1992

Resumo

A teoria de processos estocásticos pontuais é frequentemente utilizada para a descrição de séries de acontecimentos ocorrendo aleatoriamente no tempo. Em particular, a análise estatística de processos pontuais tem sido aplicada na caracterização de sinais biomédicos.

No domínio da frequência, os processos pontuais estacionários são geralmente descritos pela função de densidade espectral das contagens. O espectro das contagens difere significativamente do espectro de uma série temporal, requerendo métodos de estimação específicos.

Neste trabalho discute-se a modelização espectral autoregressiva (AR) do espectro das contagens de processos estocásticos pontuais.

A forma de abordagem adoptada permite desenvolver algoritmos eficientes para a estimação do espectro das contagens, em clara correspondência com as técnicas habitualmente descritas na literatura para processos estocásticos em geral. Em particular, além dos métodos não paramétricos "clássicos", possibilita a modelização espectral paramétrica, bem como o desenvolvimento de métodos iterativos para melhoria da fidelidade na descrição global, por inclusão de informação relativa às características do espectro das contagens.

De salientar ainda que os algoritmos descritos neste trabalho se baseiam na estimação prévia da sucessão densidade de covariância. Assim, de forma análoga ao caso de processos estocásticos em geral e, nomeadamente, na modelização espectral AR, é fundamental garantir a obtenção de sucessões semi-definidas positivas. Uma vez que os estimadores correntes na literatura de processos pontuais não apresentam esta característica, discute-se também a estimação da função densidade de covariância e função de intensidade condicional e apresentam-se algoritmos de implementação eficiente que garantam a verificação das propriedades requeridas.

A aplicação dos métodos e algoritmos desenvolvidos é ilustrada com sinais biomédicos, considerando dois casos particulares: caracterização do miograma de interferência (EMG) e da variabilidade de sinais cardiovasculares.

Contribuições da tese

A contribuição genérica fundamental desta tese consiste na *abordagem adoptada*: mostra-se que alguns problemas de estimação em processos pontuais (nomeadamente da densidade de covariância, da intensidade condicional e do espectro das contagens), podem ser formulados na perspectiva da análise de séries temporais. Em particular, esta orientação permite desenvolver algoritmos eficientes para a estimação do espectro das contagens de um processo estocástico pontual, em clara correspondência com os métodos, quer paramétricos quer não paramétricos, habitualmente descritos na literatura para processos estocásticos em geral.

Neste trabalho é dada particular atenção à *modelização espectral AR do espectro das contagens e ao seu aperfeiçoamento*. As contribuições específicas são as seguintes:

- Desenvolvimento de estimadores semi-definidos positivos, susceptíveis de implementação eficiente, para a a função densidade de covariância e intensidade condicional de processos estocásticos pontuais.
- Modelização espectral AR do espectro das contagens de processos estocásticos pontuais.
- Aperfeiçoamento da modelização espectral AR de processos estocásticos pontuais, a partir de inclusão directa de informação sobre as características do espectro das contagens na região das altas frequências.

Agradecimentos

Agradeço ao Professor Doutor Pedro Lago, que orientou este trabalho, o empenho e entusiasmo dispensados. Gostaria de agradecer as condições de trabalho que me criou e a disponibilidade incondicional nos momentos decisivos.

Neste trabalho usam-se dados experimentais. Agradeço as facilidades de recolha de dados e experimentação concedidas pelo Serviço de Neurofisiologia Clínica do Hospital Geral de Santo António, muito em especial ao Dr. Viana Pinheiro, pelo auxílio indispensável. Agradeço também a colaboração da Faculdade de Medicina da Universidade do Porto e do Centro de Medicina Desportiva do Norte, em especial aos Professores Doutores Falcão de Freitas e Ovídio Costa.

Agradeço ao Instituto Nacional de Investigação Científica o apoio prestado - Bolsa de estudo para Doutoramento no País (Processo 23390).

Finalmente, mas não menos sinceros, os meus agradecimentos para a minha Família, a quem tantas vezes privei do meu apoio.

Ao António e ao Bernardo

Índice

Resumo

Contribuições da tese

1 Introdução 4

1.1 Processos estocásticos pontuais: aplicação na caracterização de sinais biomédicos 4

1.1.1 Caracterização da complexidade dos potenciais no EMG 5

1.1.2 Variabilidade de sinais cardiovasculares 9

1.2 Análise espectral de processos pontuais 14

1.2.1 Introdução: descrição de um processo pontual 14

1.2.2 Análise espectral do processo das contagens 14

1.3 Formulação do problema 18

1.3.1 Motivação 18

1.3.2 Forma de abordagem 18

1.4 Organização do texto e contribuições fundamentais da Tese 20

2 Análise espectral das contagens: métodos e algoritmos de implementação eficiente 22

2.1 Introdução 23

2.1.1 O periodograma como estimador para o espectro das contagens 23

2.1.2 Estimadores espectrais consistentes 24

2.2 Implementação eficiente da estimação do espectro das contagens 27

2.2.1 Introdução 27

2.2.2 Resultados fundamentais 27

2.2.3 O método de Blackman-Tukey 29

2.2.4 Análise espectral AR de processos pontuais 30

2.2.5 Comentários finais 32

2.3 Estimação da função densidade de covariância para a análise espectral de processos pontuais 33

2.3.1 Introdução 33

2.3.2 Métodos de estimação baseados no histograma das contagens . 34

2.3.3 Estimadores semi-definidos positivos 35

2.3.4 Estimação de $\{g_d(n\Delta t)\}$: resultados de simulação 39

2.4 Análise Espectral das contagens: estudo de simulação 41
2.4.1 Método de Blackman-Tukey 41
2.4.2 Modelização Espectral AR 43
2.5 Comentários finais 48

3 Análise espectral AR de processos estocásticos pontuais: inclusão directa de informação sobre as características do espectro 49
3.1 Introdução 50
3.1.1 Modelização espectral AR: estudo de simulação com base num número não mínimo de equações 50
3.1.2 Modelização AR do espectro das contagens a partir da sucessão densidade de covariância teórica 54
3.1.3 Comentários Finais 56
3.2 Aperfeiçoamento da modelização espectral AR 58
3.2.1 Introdução 58
3.2.2 Medida de erro "modificada" 58
3.2.3 Modelo espectral AR equivalente 61
3.3 Determinação do modelo espectral AR por iteração sobre os parâmetros 63
3.3.1 Método de projecção do gradiente 63
3.3.2 Método de Newton modificado 65
3.3.3 Comentários finais e interpretação da implementação do método de Newton modificado 67
3.4 Algoritmos eficientes tipo Levinson baseados numa densidade de covariância "modificada" 68
3.4.1 Introdução 69
3.4.2 Densidade de covariância modificada e modelo espectral correspondente: resultados de simulação 70
3.4.3 Aperfeiçoamento do modelo espectral AR por iteração na densidade de covariância 74
3.5 Aperfeiçoamento da modelização espectral AR: alguns comentários . . . 77
3.5.1 Escolha da ordem 77
3.5.2 Obtenção de indicação da gama de frequências onde o espectro deverá ser plano 80
3.5.3 Fidelidade da descrição espectral na gama das baixas frequências 84
3.6 Comentários finais 86

4 Modelização AR do espectro das contagens: aplicação na caracterização de sinais biomédicos 88
4.1 Modelização espectral AR do processo pontual associado aos intervalos RR no ECG 89
4.1.1 Obtenção de informação sobre a frequência respiratória a partir da análise espectral das contagens 89
4.1.2 Modelização espectral AR e decomposição espectral 91
4.2 Modelização espectral AR de processos pontuais extraidos do EMG . . 94
4.2.1 Caracterização da complexidade dos potenciais 94

4.2.2 Obtenção da frequência média de acendimento das unidades motoras no EMG e sua alteração com a fadiga 102

5 Conclusões Finais **107**

A Sistema de aquisição e processamento do EMG **111**

Referências **113**

Capítulo 1

Introdução

Neste capítulo introduz-se o tópico da dissertação, dando especial ênfase à motivação e formulação do problema em estudo.

1.1 Processos estocásticos pontuais: aplicação na caracterização de sinais biomédicos

A teoria de processos estocásticos pontuais é particularmente adequada para a descrição matemática de séries de acontecimentos ocorrendo aleatoriamente no tempo. As áreas de aplicação são numerosas; nomeadamente, a análise estatística de processos pontuais tem sido usada na caracterização de sinais biomédicos [Cohen 86].

A Neurofisiologia tem constituido uma das principais áreas de aplicação da teoria de processos estocásticos pontuais ao estudo de sinais bioeléctricos: com efeito, um trem de potenciais pode ser completamente caracterizado pela forma do potencial e pela sucessão dos instantes de ocorrência, isto é pelo processo estocástico pontual associado aos instantes de ocorrência dos potenciais.

Neste trabalho discute-se a análise espectral de processos estocásticos pontuais e ilustra-se a sua utilização na caracterização de sinais biomédicos considerando duas aplicações: caracterização do *miograma de interferência* (EMG) e da *variabilidade de sinais cardiovasculares*.

Embora seja corrente a utilização de técnicas de análise espectral (nomeadamente no caso das aplicações consideradas) [Cohen 86], em muitas situações é desejável adaptar o método de análise às características particulares do sinal.

Assim, conforme se descreve sucintamente em seguida, uma abordagem baseada na análise espectral de processos estocásticos pontuais não só é claramente sugerida, como apresenta a vantagem de complementar ou mesmo obviar as dificuldades encontradas na caracterização de um sinal.

1.1.1 Caracterização da complexidade dos potenciais no EMG

A contracção voluntária dos músculos estriados produz, além de actividade mecânica, um sinal eléctrico: o miograma. O sinal que se observa, registado com eléctrodos convencionais, apresenta comportamento aleatório extremamente complexo, como se ilustra na figura 1.1, designando-se por *miograma de interferência* (EMG) [Basmajian 78, Basmajian e De Luca 85, e referências aí incluidas].

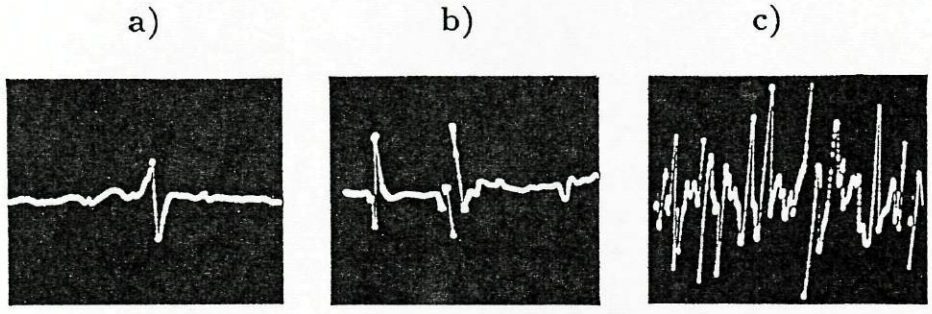


Figura 1.1: Registos electromiográficos para níveis de contracção crescente (de [Basmajian 78]): **a)** corresponde ao potencial associado a uma unidade motora, que se repete aleatoriamente no tempo, como se mostra em **b)**; numa contracção moderada a forte é recrutado um número muito elevado de unidades motoras e o registo obtido corresponde a uma sobreposição temporal de registos do tipo **a)** o que dá origem a um sinal como o representado em **c)**.

Em estados de doença muscular avançada observam-se alterações profundas tanto na actividade eléctrica como na força produzida. Contudo, nos estados menos avançados, é geralmente necessário um exame cuidadoso do miograma o qual inclui a observação dos potenciais individuais das unidades motoras e a sua caracterização pela medição da amplitude, duração e complexidade (número de fases) [Desmedt 83, Basmajian e De Luca 85, Desmedt 89].

Este tipo de observação, além de requerer uma cooperação do doente, apresenta limitações uma vez que só é aplicável em boas condições para contracções fracas e portanto às unidades motoras de limiar baixo (as primeiras a serem recrutadas), ou então em estados de doença mais avançados em que o número de unidades motoras activas é muito reduzido.

A ocorrência de potenciais polifásicos no EMG tem significado clínico [Desmedt 83, Basmajian e De Luca 85, Desmedt 89]. Como o EMG se assemelha a ruído, para a caracterização da complexidade dos potenciais que constituem o miograma de interferência (numa gama variada de contracções musculares) são particularmente indicadas as técnicas de análise espectral. A forma da função de densidade espectral do EMG poderá indicar a existência de potenciais polifásicos uma vez que estes, tendo uma menor duração, dão origem a uma densidade espectral de largura de banda maior do que o normal [Basmajian e De Luca 85, Desmedt 83, Desmedt 89].

Existem contudo limitações na aplicação da análise espectral directa, devido essencialmente a muitos outros factores que influenciam a forma do espectro e interactuam de forma complexa [Jones e Lago 77, Lindstrom e Petersen 83, Basmajian e De Luca 85]. Apesar dessas influências poderem dar informação útil sobre alguns parâmetros, em particular sobre a velocidade média de condução das fibras [Lindstrom e Petersen 83] e sobre a frequência média de acendimento das unidades motoras [Lago e Jones 77, Lago e Jones 81, Paiss e Inbar 87, e referências aí incluidas], torna-se, no entanto, difícil estabelecer um padrão de ausência significativa de potenciais polifásicos em termos de um padrão normal da função de densidade espectral do EMG.

As limitações que a análise espectral directa apresenta tornam necessária a inclusão de informação adicional directamente ligada com a geração do sinal.

Embora não haja uma descrição matemática simples, uma classe de modelos que se tem vindo a revelar bastante útil na descrição do EMG, para contracções isotónicas e isométricas pouco prolongadas, baseia-se na soma de processos estocásticos, possivelmente não independentes, representando a actividade das unidades motoras intervenientes [Jones e Lago 77, Agarwal e Gottlieb 82, e referências aí incluidas]. Como se ilustra na figura 1.2,

$$EMG = \sum_{i=1}^{s} \sum_{j} h_i(t - t_{ij}), \tag{1.1}$$

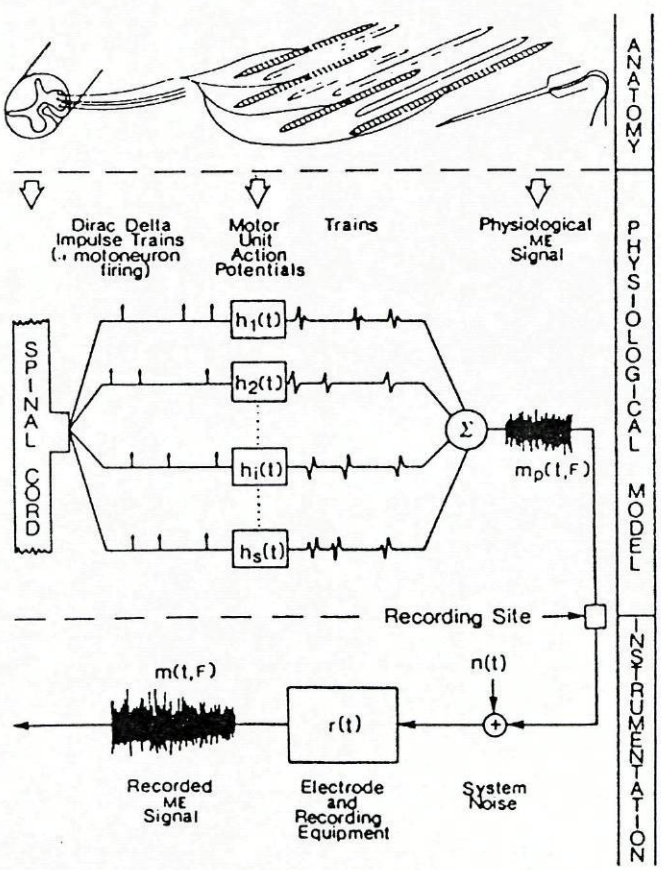


Figura 1.2: Representação esquemática de um modelo de geração do miograma; extraído de [De Luca 79].

onde s representa o número de unidades motoras activas e, $h_i(t)$ e $\{t_{ij}\}$ indicam, respectivamente, a forma dos potenciais e a sucessão dos instantes de acendimento.

Reconhece-se facilmente que a informação sobre a presença de potenciais polifásicos está contida no agrupamento dos instantes de ocorrência dos máximos locais significativos nos trens de potenciais constituintes (figura 1.3). Assim, o processo pontual associado aos máximos locais significativos no EMG traduz características importantes do sinal, representando-se na figura 1.4 as operações envolvidas na sua extracção.

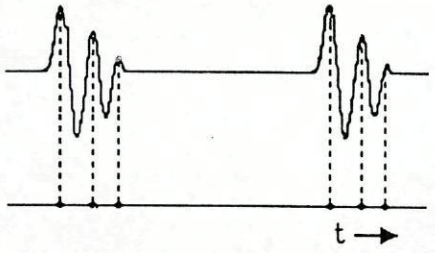


Figura 1.3: Representação esquemática de um trem de potenciais polifásicos e da sucessão dos máximos locais significativos.

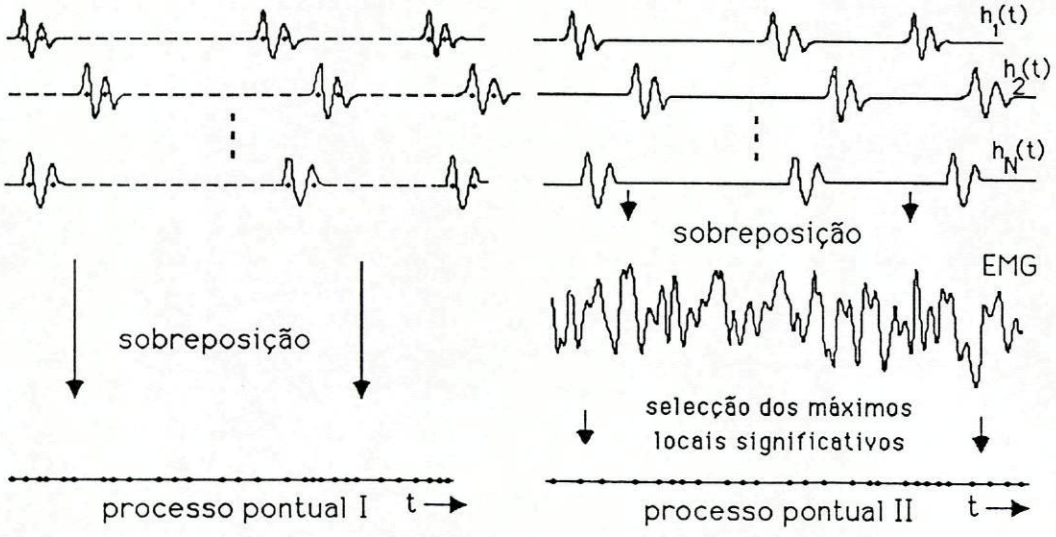


Figura 1.4: Ilustração esquemática da geração do processo pontual obtido por sobreposição dos máximos locais significativos dos trens de potenciais (processo I). Ilustração esquemática da geração do EMG e do processo de extracção dos máximos locais significativos (processo II).

Verifica-se que a sucessão resultante da sobreposição dos instantes de ocorrência dos máximos locais correspondentes aos trens de potenciais constituintes (processo I na figura 1.4) contem informação relevante sobre a forma (número de fases) dos potenciais, das estatísticas de acendimento e da sincronização (relação de dependência) entre diferentes unidades motoras [Lago e Jones 83]. Contudo, o efeito de agrupamento não será facilmente perceptível dado o elevado grau de sobreposição. O mesmo se passa com

a sucessão dos máximos locais significativos extraída do EMG de interferência (processo II na figura 1.4), sendo a aplicação de técnicas de análise espectral particularmente adequada para a sua descrição.

Estas observações constituem a base de um método de análise do EMG a partir da análise espectral de um processo pontual associado à sucessão dos máximos locais significativos no sinal [Lago e Jones 83, Lago et al 82, Lago et al 83]. O método tem uma base teórica bem estabelecida e tem demonstrado potencialidades no estudo das alterações do EMG com a fadiga e como auxiliar de classificação e diagnóstico [Jones et al 87, Jones et al 90, Lago e Jones 83, Lago et al 84].

Em particular, verifica-se que o espectro das contagens do processo estocástico pontual extraido do EMG contem informação sobre o grau de incidência de potenciais polifásicos, conforme se ilustra na figura 1.5 [Lago e Jones 83, Lago et al 84].

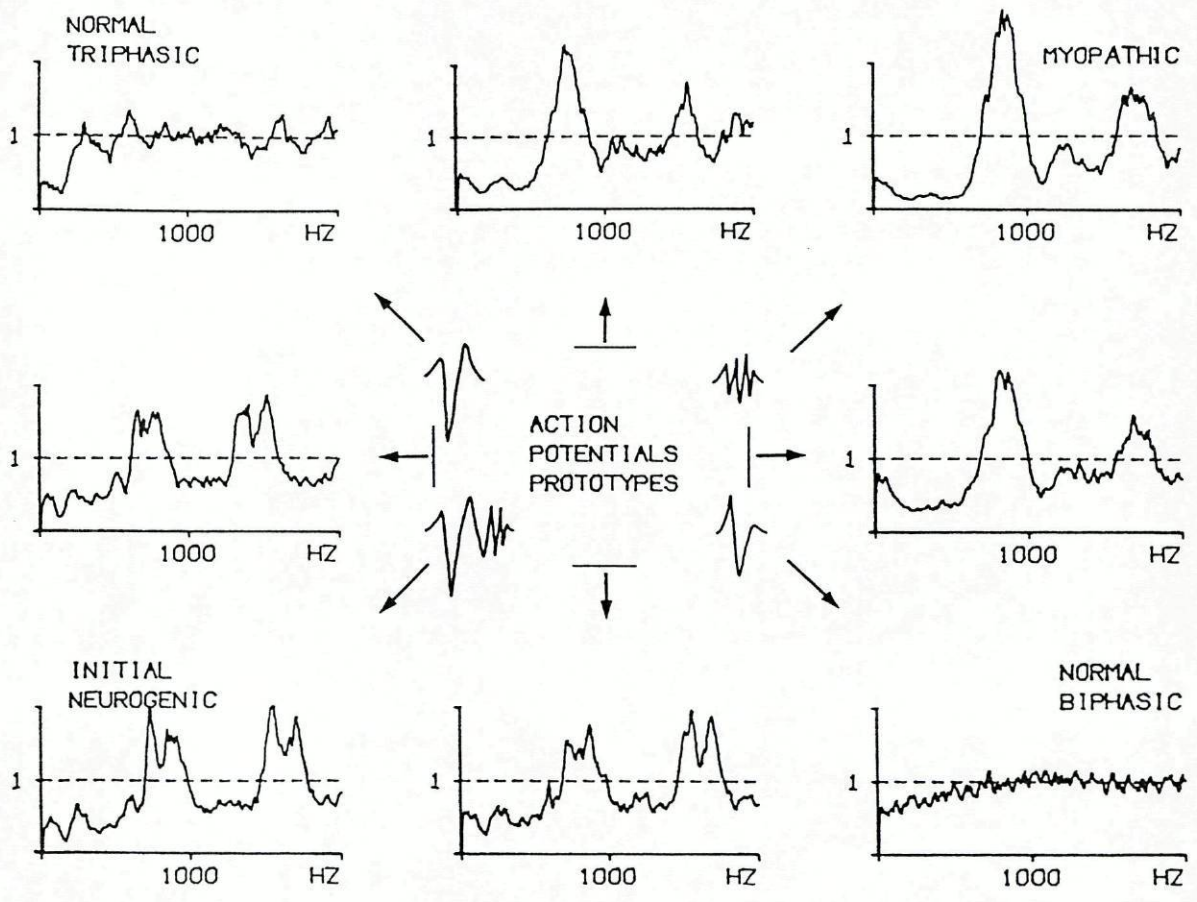


Figura 1.5: Espectros das contagens normalizados típicos (estudo de simulação); extraído de [Lago et al 84].

Na ausência de potenciais polifásicos e de sincronização, a função de densidade espectral das contagens, normalizada pelo número médio de acontecimentos por unidade de tempo, é plana e aproximadamente igual a 1 excepto na região de muito baixas frequências (diagrama do canto superior esquerdo na figura 1.5).

A existência de potenciais polifásicos origina uma estrutura em "clusters" no processo pontual, o que provoca uma amplificação nas baixas frequências e uma atenuação na banda média do espectro das contagens do processo extraido do EMG. Assim, a presença deste tipo de potenciais reflecte-se em desvios significativos da densidade espectral normalizada em relação à unidade numa gama de frequências muito maior.

Neste trabalho serão apresentados métodos eficientes para a análise espectral de processos pontuais os quais possibilitam uma extensão e aperfeiçoamento do método. Note-se que os resultados apresentados na figura 1.5, baseados em estimadores não paramétricos para o espectro das contagens, sugerem claramente vantagens que poderiam advir duma modelização espectral AR do espectro.

1.1.2 Variabilidade de sinais cardiovasculares

Sabe-se que os sinais de natureza cardiovascular (nomeadamente a frequência cardíaca, pressão arterial e mesmo os próprios complexos electrocardiográficos—complexos QRS), são "quase" periódicos, apresentando contudo variabilidade batimento a batimento [Appel et al 89]. A título exemplificativo representam-se na figura 1.6 registos simultâneos do sinal electrocardiográfico (ECG), pressão arterial e variação respiratória.

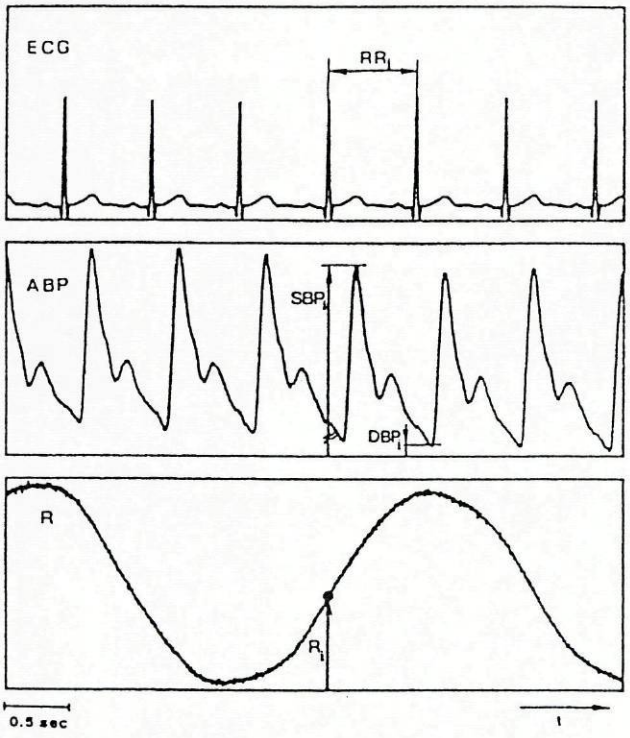


Figura 1.6: Registos simultâneos do ECG, pressão arterial (ABP) e respiração (R) com representação dos valores relevantes no iº ciclo—intervalo RR (RR_i), pressão sistólica (SBP_i), diastólica (DBP_i) e respiração (R_i); extraído de [Baselli et al 86].

A Obstrectrícia foi a primeira área em que se reconheceu o interesse na caracterização dessa variabilidade: a monitorização fetal com diminuição da variabilidade da

frequência cardíaca é um identificador de eventual sofrimento do feto.

Com efeito, a variabilidade, batimento a batimento, das variáveis cardiovasculares reflecte uma correlação entre várias perturbações da função cardiovascular (devidas nomeadamente alterações posturais, alterações da temperatura ambiente, efeitos da variação respiratória na pressão intratorácica) e a resposta correspondente dos sistemas de regulação [Akselrod et al 85, Appel et al 89, Malliani et al 91].

Assim, o estudo dessa variabilidade poderá dar indicação acerca das perturbações a que o sistema cardiovascular está exposto, bem como a resposta reguladora a essas alterações. Este tipo de análise reveste-se de enorme importância, uma vez que poderá permitir quantificar o mecanismo de controlo do sistema cardiovascular, constituindo um meio não invasivo para aceder à sua integridade, em variados estados de doença.

A "frequência" cardíaca no ECG (baseada na identificação dos complexos QRS e posterior medição dos intervalos RR—*tacograma*) é o sinal cardiovascular mais acessível para caracterizar essa variabilidade. Têm sido utilizadas técnicas de análise espectral (não paramétrica e paramétrica) na descrição da variabilidade dos intervalos RR ("frequência cardíaca"), considerando a série discreta (tacograma) associada à ordem do batimento [Akselrod et al 85, Appel et al 89, Malliani et al 91].

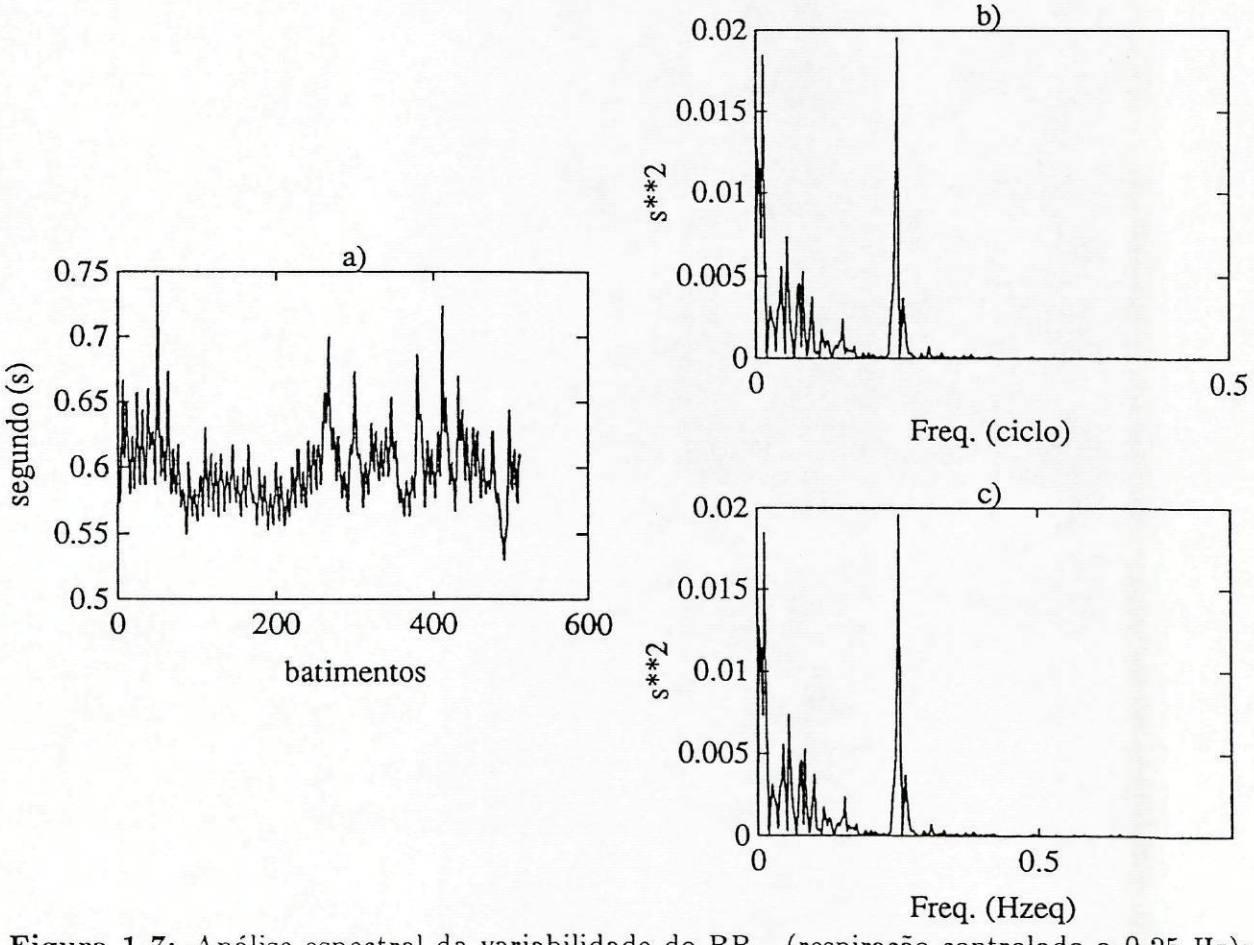


Figura 1.7: Análise espectral da variabilidade do RR (respiração controlada a 0.25 Hz); **a)** tacograma (512 pontos), **b)** e **c)** espectro não paramétrico.

Na figura 1.7 representa-se o espectro não paramétrico da variabilidade do RR para

uma série curta, considerando uma situação de respiração controlada. A escala de frequência em c) ("Hz equivalente", Hzeq), corresponde à normalização da escala em b) pelo intervalo RR médio, e é um artifício necessário pelo facto de na escala horizontal do tacograma não existir informação sobre o tempo.

Tal como a figura 1.7 ilustra, o efeito da actividade respiratória sobre a variabilidade do RR está correctamente traduzida na nova escala em frequência. Esta normalização é uma necessidade imposta por este tipo de análise e não introduz erro apreciável sempre que a variabilidade nos intervalos RR não for significativa.

Estudos farmacológicos realizados em animais não anestesiados, confirmados posteriormente no homem, permitem concluir que a análise espectral da variabilidade da frequência cardíaca, possibilita a identificação de diversas componentes, correspondentes à actividade do sistema nervoso autónomo (simpático e parasimpático), nomeadamente a componente associada à actividade respiratória [Akselrod et al 81, Akselrod et al 85, Pagani et al 86], como se ilustra esquematicamente na figura 1.8.

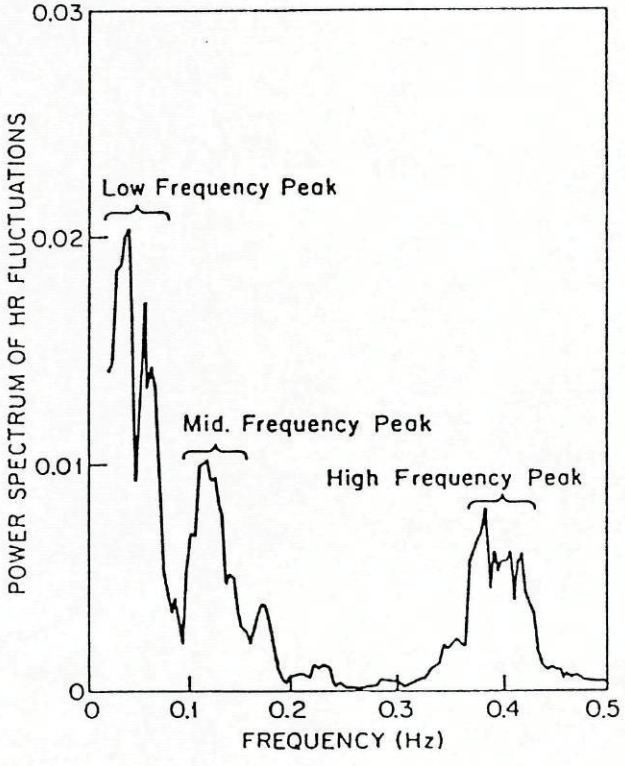


Figura 1.8: Análise espectral da variabilidade da frequência cardíaca evidenciando as principais componentes: baixas frequências (VLF), frequências intermédias (LF) e altas frequências (HF) (respiração $\approx$ 25/minuto); extraído de [Akselrod et al 81].

Observa-se uma componente nas baixas frequências (*VLF*), aproximadamente abaixo de 0.05 Hz, e duas componentes nas frequências intermédias e altas (*LF* e *HF*), centradas respectivamente, na gama aproximada de 0.1 Hz e na frequência respiratória; assim, a componente HF, associada à actividade respiratória, não estará presente no espectro do tacograma se a frequência respiratória for elevada relativamente à frequência cardíaca média [Akselrod et al 81].

Na literatura é referido que HF está apenas relacionada com a actividade do sistema nervoso parassimpático, enquanto que as componentes a baixa frequência estão relacionadas com a actividade, quer do simpático quer do parassimpático [Akselrod et al 81].

Este efeito é notório quando se analiza a variabilidade do RR em séries longas. Conforme se ilustra na figura 1.9, a variabilidade do RR apresenta uma variação circadiana onde se destaca a componente HF durante a noite, correspondendo ao período do sono e predomínio da actividade vagal (sistema parassimpático) [Costa et al 92]. Com efeito, a componente HF é significativa e corresponde a uma gama de frequências bastante estável, nesse período, devido à regularidade da actividade respiratória.

Note-se, que o espectro da variabilidade da frequência cardíaca não apresenta geralmente componentes tão bem diferenciados como os da figura 1.8: devido à frequência respiratória ser geralmente mais baixa, bem como à sua variabilidade, as componentes LF e HF tendem a misturar-se.

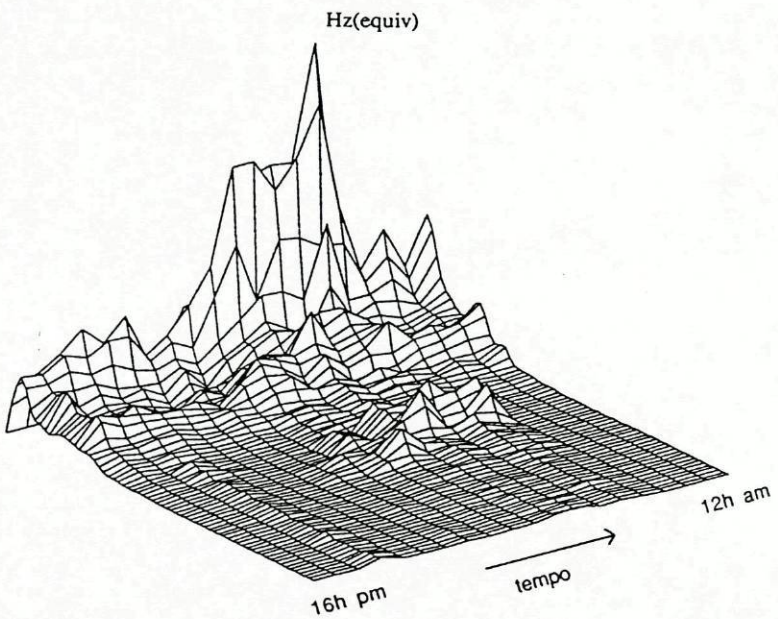


Figura 1.9: Evolução do espectro da variabilidade da frequência cardíaca relativo a um registo longo (das 16h às 12h do dia seguinte).

Referindo de novo o exemplo da figura 1.7, considere-se o espectro normalizado de forma que a área delimitada pela curva seja igual à variância do sinal. Assim, a decomposição da variabilidade do RR poderá ser feita a partir do cálculo de áreas no espectro. No caso da análise não paramétrica as componentes correspondem a bandas fixas (tipicamente: VLF-0 a 0.05 Hzeq, LF-0.05 a 0.15 Hzeq, HF-0.15 a 0.40 Hzeq), enquanto que na paramétrica (figura 1.10) são identificadas automaticamente, relativamente à sua frequência central [Costa et al 93, Pagani et al 86, Zetterberg 69].

A componente *VLF* é dominada essencialmente pela actividade física e alterações posturais, sendo necessários registos muito longos para a analizar. É um índice de interpretação díficil não sendo muito utilizado.

Assim, a avaliação do equilíbrio entre os dois componentes do sistema nervoso autónomo é geralmente feita por determinação de das áreas LF, HF e razão LF/HF.

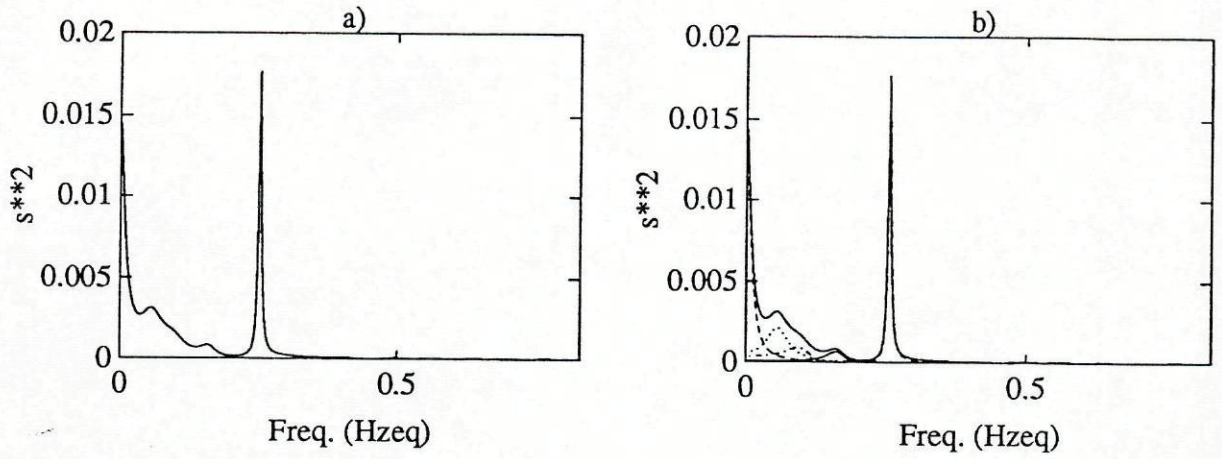


Figura 1.10: Análise espectral paramétrica da variabilidade do RR (mesmos dados da figura 1.7; **a)** espectro AR(27) (ordem do modelo optimizada pelo critério de AIC [Akaike 74]) , **b)** espectro AR(27) e decomposição espectral AR (componentes espectrais superiores a 5% da variância total) [Zetterberg 69].

Tem sido verificado que este tipo de caracterização é susceptível de variadas aplicações clínicas, nomeadamente no estudo de hipertensão, enfarte agudo do miocárdio e comportamento após transplante cardíaco [Malliani et al 91, e referências aí incluidas]. No caso da PAF (Polineuropatia Amiloidotica Familiar) permite mesmo obter indicação da disfunção do sistema nervoso autónomo em portadores assintomáticos [Puig et al 91, Carvalho et al 91].

Finalmente refira-se que têm também sido realizados estudos relativos à análise espectral cruzada da variabilidade dos intervalos RR, nomeadamente com a pressão arterial e variação respiratória [Akselrod et al 85, Appel et al 89, Malliani et al 91].

Apesar da análise da variabilidade dos intervalos RR baseada na série discreta associada a cada batimento ter mostrado potencialidades, também neste caso se poderá pensar em introduzir uma descrição alternativa. Com efeito, o ECG pode ser considerado como um processo estocástico pontual se apenas se estiver interessado, por exemplo, no estudo da variabilidade dos intervalos RR e não no detalhe de todo o percurso de polarização e despolarização do músculo cardíaco. Assim, considerando nomeadamente o estudo da variação na frequência cardíaca, baseada na identificação dos complexos QRS e posterior medição dos intervalos RR, o ECG pode ser representado por um processo pontual associado aos instantes de ocorrência correspondentes.

A abordagem a partir da análise espectral de processos pontuais permite descrições alternativas que poderão ser vantajosas, em particular quando se pretende relacionar a série de acontecimentos a outros sinais de natureza essencialmente contínua, como a respiração ou a temperatura [DeBoer et al 84].

De salientar ainda que os algoritmos eficientes para a estimação do espectro das contagens apresentados neste trabalho permitirão também caracterizações paramétricas e, em particular, a identificação de componentes espectrais relevantes, podendo portanto constituir uma alternativa, ou um complemento, às descrições tradicionais.

1.2 Análise espectral de processos pontuais

1.2.1 Introdução: descrição de um processo pontual

Os processos estocásticos pontuais admitem duas descrições complementares no domínio do tempo, podendo ser caracterizados, quer em termos dos intervalos de tempo entre acontecimentos sucessivos, quer em termos de contagens do número de acontecimentos em intervalos fixos no tempo [Cox e Isham 80, Cox e Lewis 66].

Considerando processos pontuais estacionários, existem no domínio da frequência as duas descrições correspondentes: a função de densidade espectral dos intervalos e a função de densidade espectral das contagens, respectivamente.

A função de densidade espectral dos intervalos é a função de densidade espectral da sucessão discreta dos intervalos de tempo entre ocorrências sucessivas [Cox e Lewis 66]. Assim, a definição e estimação do espectro dos intervalos não difere essencialmente da definição e estimação do espectro de uma série temporal discreta.

A função de densidade espectral das contagens vai ser definida na secção seguinte. Em contraste com a análise espectral dos intervalos, a análise espectral das contagens difere significativamente da análise espectral de uma série temporal.

Neste trabalho consideram-se descrições de processos estocásticos pontuais no domínio da frequência, essencialmente baseadas na análise espectral das contagens. Salvo indicação expressa em contrário, será apenas considerado o caso de processos pontuais estacionários sem ocorrências múltiplas.

1.2.2 Análise espectral do processo das contagens

A função de densidade espectral das contagens [Cox e Isham 80, Cox e Lewis 66], $G(\omega)$, é definida como sendo a transformada de Fourier da função densidade de covariância completa, $g(t)$, a qual é dada, para processos pontuais estacionários e sem ocorrências múltiplas, pela expressão

$$g(t) = \lambda\,\delta(t) + \lambda\,(m(t) - \lambda) \tag{1.2}$$

onde $\delta(t)$ indica a função delta de Dirac, λ é a intensidade média,

$$\lambda = \lim_{\Delta t \to 0+} \text{Prob. } \{\text{acontecimento em }]t, t+\Delta t]\} \;/\; \Delta t \tag{1.3}$$

e $m(t)$ a função de intensidade condicional, função simétrica definida como zero na origem e para $t \neq 0$ pela probabilidade condicional

$$m(t) = \lim_{\Delta t \to 0+} \{\text{Prob. acontec. em }]u + |t|, u + |t| + \Delta t] \mid \text{acontec. em } u\}/\Delta t. \tag{1.4}$$

Tomando a transformada de Fourier de (1.2), verifica-se então que, para um processo estocástico pontual estacionário com probabilidade nula de ocorrências múltiplas, o espectro das contagens, $G(\omega)$,

$$G(\omega) = \lambda + \lambda \int_{-\infty}^{+\infty} (m(t) - \lambda) e^{-j\omega t} dt \tag{1.5}$$

é dado para frequências $w \neq 0$ pela expressão [Cox e Lewis 66]

$$G(\omega) = \lambda(1 + \int_0^{+\infty} m(t)e^{j\omega t}dt + \int_0^{+\infty} m(t)e^{-j\omega t}dt), \tag{1.6}$$

isto é, $G(\omega) = \lambda(1 + M(j\omega))$, onde $M(j\omega)$ designa a transformada de Fourier de $m(t)$.

No caso particular importante de um processo estocástico pontual *"renewal"*, isto é, das variáveis aleatórias intervalos de tempo entre acontecimentos sucessivos serem mutuamente independentes e terem a mesma distribuição marginal, $f(t)$, o espectro das contagens tem uma expressão simples. Com efeito, designando por $F(j\omega)$ a transformada de Fourier da função $f(t)$, conclui-se que [Cox e Lewis 66],

$$G(\omega) = \lambda \ (1 + \frac{F(j\omega)}{1 - F(j\omega)} + \frac{F(-j\omega)}{1 - F(-j\omega)}), \omega \neq 0. \tag{1.7}$$

A análise espectral das contagens de um processo estocástico pontual difere significativamente da análise espectral de um processo estocástico em geral, $\{x(t)\}$.

Contudo, existe uma perfeita analogia das operações envolvidas, conforme se descreve comparativamente na tabela 1.1 (de salientar o facto de neste trabalho terem sido adoptadas as designações correntes em Processamento de Sinal [Kay 88, Marple 87], diferindo ligeiramente das utilizadas em Séries Temporais).

Em particular, designando por $N(t)$ a variável aleatória "número de acontecimentos no intervalo $]0, t]$", as funções de intensidade condicional e densidade de covariância, exprimem-se, respectivamente, a partir das funções de autocorrelação e autocovariância do processo diferencial $\{\Delta N(t)\}$ [Cox e Lewis 66],

$$\Delta N(t) = \lim_{\Delta t \to 0^+} \frac{N(t + \Delta t) - N(t)}{\Delta t}.$$

No caso de processos estocásticos pontuais a dependência entre os acontecimentos é descrita, no domínio do tempo, a partir da funções $m(t)$ e $g(t)$.

Verifica-se que $\lim_{t \to \infty} m(t) = \lambda$ e a independência entre os acontecimentos (processo de Poisson) traduz-se numa intensidade condicional constante, isto é, $m(t) = \lambda$ [Cox e Lewis 66].

A característica fundamental do espectro das contagens, $G(\omega)$, consiste no facto de, em contraste com o espectro de uma série temporal, $\phi(\omega)$, $G(\omega)$ não tender para zero com ω crescente, como se ilustra na figura 1.11.

Com efeito, a partir de (1.5), conclui-se que

$$\lim_{\omega \to +\infty} G(\omega) = \lambda. \tag{1.8}$$

A independência total entre os acontecimentos (processo de Poisson) traduz-se, no domínio da frequência, por um espectro das contagens constante, isto é $G(\omega) = \lambda$.

PROCESSO ESTOCÁSTICO	PROCESSO PONTUAL
Autocorrelação $\boldsymbol{R}(\tau) = \mathrm{E}[x(t)x(t+\tau)]$	**Função Intensidade Condicional** $\boldsymbol{m}(t) = \lim_{\Delta t \to 0+} \frac{\Pr.(\{\text{ac.em}]t, t+\Delta t]\} \mid \text{ac.em } t=0)}{\Delta t}$ $t > 0$ $(m(t) = m(-t), t < 0, m(0) = 0)$
Valor médio $\boldsymbol{\mu} = \mathrm{E}[x(t)]$	**Intensidade média** $\boldsymbol{\lambda} = \lim_{\Delta t \to 0+} \frac{\Pr.(\{\text{ac.em}]t, t+\Delta t]\})}{\Delta t}$
Autocovariância $\boldsymbol{C}(\tau) = \mathrm{R}(\tau) - \mu^2$	**Função Densidade de Covariância** $\boldsymbol{g}(t) = \lambda\,\delta(t) + \lambda\,(m(t) - \lambda)$
Função de densidade espectral $\boldsymbol{C}(\tau) \longleftrightarrow \boldsymbol{\phi}(\omega)$ $\phi(\omega) = \int_{-\infty}^{+\infty} C(\tau) e^{-j\omega\tau} d\tau$ $\lim_{\omega \to +\infty} \phi(\omega) = 0$	**Espectro das contagens** $\boldsymbol{g}(t) \longleftrightarrow \boldsymbol{G}(\omega)$ $G(\omega) = \int_{-\infty}^{+\infty} g(t) e^{-j\omega t} dt$ $\lim_{\omega \to +\infty} \boldsymbol{G}(\omega) = \lambda$

Tabela 1.1: Paralelismo entre os conceitos e operações envolvidas na análise espectral de processos estocásticos estacionários e processos estocásticos pontuais estacionários sem ocorrências múltiplas.

A título exemplificativo das características e diferenças fundamentais entre a função de densidade espectral das contagens de um processo pontual e o espectro de um processo estocástico em geral, representam-se na figura 1.11 o espectro do EMG, obtido com eléctrodos de superfície, e o espectro das contagens do processo pontual associado à variabilidade da frequência cardíaca, medida pelos instantes de ocorrência das ondas R no ECG.

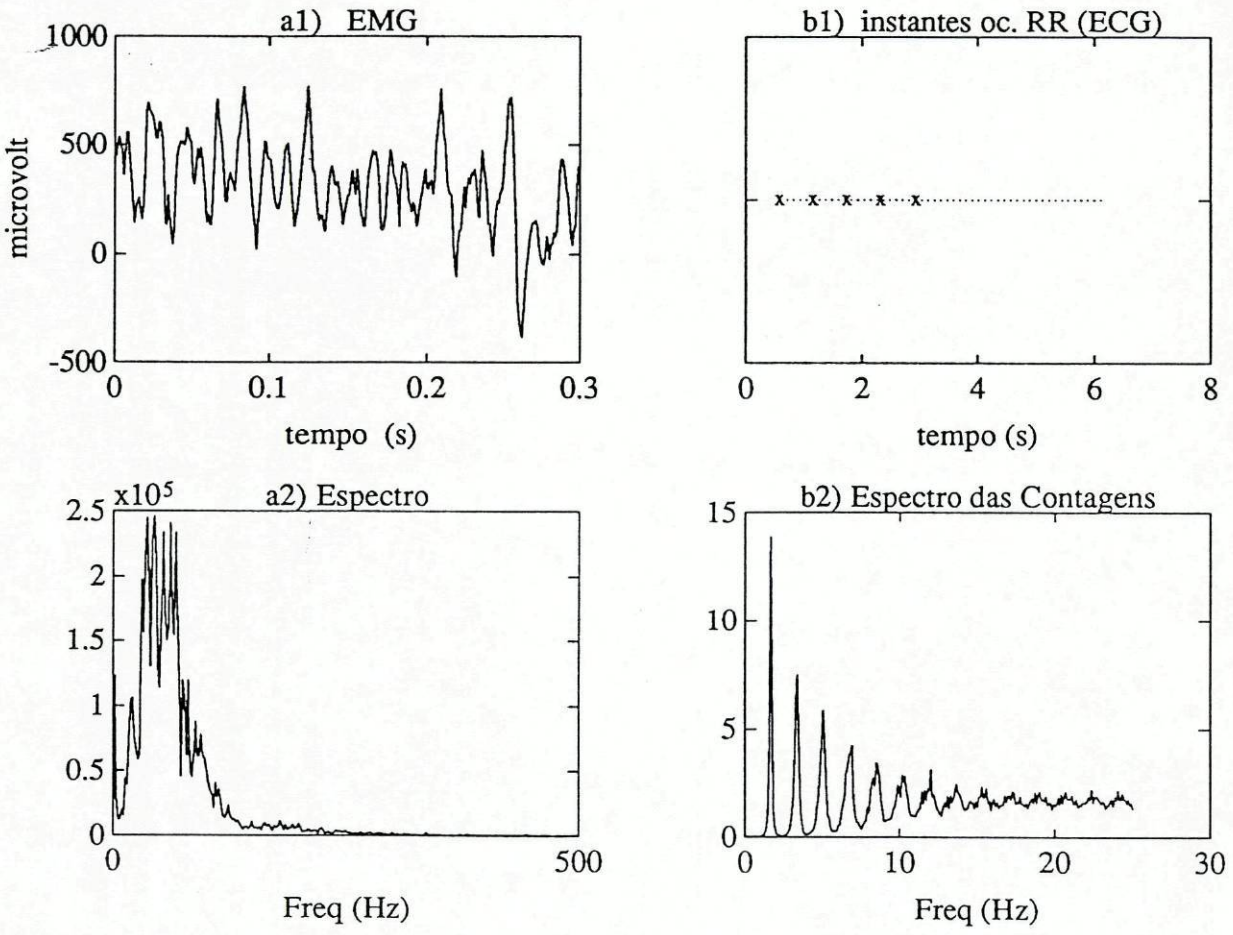


Figura 1.11: EMG - a1) sinal obtido com eléctrodo de superfície e **a2)** espectro; **ECG - b1)** processo estocástico pontual associado à variabilidade da frequência cardíaca e **b2)** função de densidade espectral das contagens.

1.3 Formulação do problema

1.3.1 Motivação

Tendo em vista a aplicação da análise espectral de processos pontuais, torna-se fundamental o desenvolvimento de algoritmos eficientes para a estimação do espectro.

Conforme já foi referido, os processos pontuais estacionários podem ser descritos, no domínio da frequência, quer a partir do espectro dos intervalos, quer a partir do espectro das contagens.

A estimação do espectro dos intervalos não difere essencialmente da estimação do espectro de uma série temporal discreta, podendo utilizar-se os métodos (paramétricos e não paramétricos), correntemente descritos na literatura [Kay 88, Marple 87]. Não será, portanto, discutida neste trabalho.

No domínio da frequência, os processos pontuais estacionários são geralmente descritos pela função de densidade espectral das contagens.

No entanto, o espectro das contagens de um processo pontual difere significativamente do espectro de uma série temporal, como se ilustrou na secção anterior. A análise espectral do processo das contagens requer métodos específicos, que garantam uma descrição correcta na zona das altas frequências.

Uma das dificuldades da estimação corrente do espectro das contagens, a partir do periodograma da série de acontecimentos observada [Cox e Lewis 66], decorre do facto de não ser susceptível de implementação eficiente. Este aspecto é particularmente limitativo na caracterização de sinais biológicos, uma vez que os registos podem ser bastante longos.

Além disso, seria extremamente vantajoso transpor para a análise espectral de processos pontuais os métodos e algoritmos eficientes e amplamente divulgados na literatura de Processamento de Sinal.

1.3.2 Forma de abordagem

Apesar das diferenças entre o espectro das contagens e o espectro de uma série temporal, o paralelismo das operações envolvidas (descrito na tabela 1.1) sugere que alguns dos problemas de estimação em processos pontuais possam ser formulados na perspectiva de análise de séries temporais.

Esta abordagem permite o desenvolvimento de algoritmos eficientes em clara correspondência com os métodos correntes em Processamento e Sinal. Inicialmente são apresentados os métodos não paramétricos "clássicos": *periodograma* e *método de Blackman-Tukey*.

A modelização espectral AR tem sido largamente utilizada na análise de séries temporais e pode, em princípio, ser aplicada na análise espectral do processo das

contagens com vantagens semelhantes sobre a abordagem não paramétrica.

Um dos objectivos principais deste trabalho consiste no desenvolvimento de métodos paramétricos para a análise espectral de processos pontuais, discutindo-se detalhadamente a *modelização espectral AR do espectro das contagens* .

É dada especial atenção ao aperfeiçoamento da fidelidade na modelização espectral AR, por inclusão de informação relativa às características do espectro das contagens.

De salientar ainda que os algoritmos descritos neste trabalho se baseiam numa estimação prévia da sucessão densidade de covariância. Assim, é fundamental, de forma análoga ao caso de processos estocásticos em geral, e em particular para a modelização espectral AR, que esteja garantida a obtenção de sucessões semi-definidas positivas.

Uma vez que os estimadores correntes na literatura de processos pontuais não apresentam esta característica, discute-se também a estimação da função densidade de covariância e apresentam-se algoritmos de implementação eficiente que preservem aquela propriedade.

Neste trabalho apresenta-se um estudo comparativo de métodos (paramétricos e não paramétricos) para a estimação do espectro das contagens de um processo estocástico pontual.

Na comparação de métodos de estimação espectral é essencial definir as características a avaliar.

Neste trabalho a preocupação essencial refere-se ao estudo da potência dos métodos e ao aperfeiçoamento dos algoritmos relativamente à fidelidade na descrição global do espectro. A fidelidade na descrição da forma global do espectro garante a descrição das suas características fundamentais, e revela-se particularmente adequada para as aplicações estudadas.

De forma análoga aos processos estocásticos em geral, existem dificuldades na obtenção de expressões analíticas para as medidas estatísticas correntes (erro de centricidade, variância) do desempenho de alguns dos estimadores espectrais aqui estudados. Em particular, muitas vezes apenas se dispõe de resultados assimptóticos.

Uma forma corrente de obviar estas dificuldades consiste em ilustrar e avaliar as propriedades dos estimadores espectrais pelo método de Monte Carlo [Kay 88, Marple 87]. Neste trabalho será adoptada uma abordagem análoga: a comparação será feita a partir da dispersão da sobreposição dos espectros estimados para diversas realizações e do espectro médio correspondente. Obtem-se assim indicação sobre a variabilidade dos estimadores e o erro de centricidade associado, em função da frequência.

A implementação dos métodos e algoritmos desenvolvidos foi feita no ambiente de programação MATLAB, utilizando o "MATLAB-Signal Processing Toolbox".

1.4 Organização do texto e contribuições fundamentais da Tese

O objectivo deste trabalho consiste no desenvolvimento de métodos e algoritmos para a modelização espectral autoregressiva (AR) de processos estocásticos pontuais e na sua aplicação na caracterização de sinais biomédicos.

De referir que a motivação para o desenvolvimento de métodos paramétricos para a análise espectral de processos estocásticos pontuais surgiu no decorrer do estudo da aplicação em electromiografia descrita na Introdução. A modelização espectral é claramente sugerida pela necessidade de descrever, de forma global e sucinta, os espectros das contagens do processo pontual associado à sucessão dos máximos locais significativos no EMG.

A exposição está organizada em capítulos, cujos objectivos fundamentais e insersão no esquema geral se descrevem em seguida:

- Introdução
  - Motivação e formulação do problema
  - Caracterização de sinais biomédicos utilizando técnicas de análise espectral: apresentação resumida de duas aplicações particulares
    - caracterização da complexidade no EMG
    - caracterização da variabilidade de sinais cardiovasculares
- Análise espectral das contagens: métodos e algoritmos de implementação eficiente
  - Estudo comparativo de diversos métodos (paramétricos e não paramétricos) para a estimação do espectro das contagens de um processo pontual
  - Estimadores semi-definidos positivos para a função densidade de covariância de um processo pontual
- Modelização espectral AR de processos pontuais: inclusão directa de informação sobre as características do espectro das contagens
  - Aperfeiçoamento da modelização espectral AR.
  - Desenvolvimento de algoritmos iterativos para a determinação do modelo espectral "melhorado"
- Análise espectral de processos estocásticos pontuais: aplicação na caracterização de sinais biomédicos.
  - Ilustração da modelização espectral AR de processos pontuais na caracterização do EMG e da variabilidade em sinais cardiovasculares
- Conclusões finais

De salientar que a aplicação da modelização espectral AR do processo das contagens, na caracterização do EMG e da variabilidade em sinais cardiovasculares, é ilustrada utilizando *dados experimentais* .

- No caso do EMG os sinais utilizados foram recolhidos com um sistema para aquisição, armazenamento e quantificação do EMG desenvolvido no âmbito num projecto de investigação em colaboração com o Serviço de Neurofisiologia Clínica do Hospital Geral de Santo António.

  Este processo constituiu uma fase extremamente morosa, embora *essencial* para a continuidade do trabalho, não sendo descrita nesta tese por não conter aspectos inovadores relativamente a outros sistemas de aquisição [Stålberg e Antoni 83].

  O sistema desenvolvido baseia-se num computador PDP 11/23 e descreve-se sucintamente em Apêndice.

- No caso da variabilidade de sinais cardiovasculares os dados foram cedidos pela Faculdade de Medicina da Universidade do Porto (Medicina 3) e pelo Centro de Medicina Desportiva do Norte, no âmbito de um projecto de investigação em colaboração com essas entidades.

A contribuição genérica fundamental desta tese consiste na *abordagem adoptada*: mostra-se que alguns problemas de estimação em processos pontuais (nomeadamente da densidade de covariância, da intensidade condicional e do espectro das contagens), podem ser (re)formulados na perspectiva da análise de séries temporais. Em particular, esta orientação permite desenvolver algoritmos eficientes para a estimação do espectro das contagens de um processo estocástico pontual, em clara correspondência com os métodos, quer paramétricos quer não paramétricos, habitualmente descritos na literatura para processos estocásticos em geral.

Neste trabalho é dada particular atenção à *modelização espectral AR do espectro das contagens e ao seu aperfeiçoamento*, atendendo às características do espectro das contagens de um processo estocástico pontual. As contribuições específicas são as seguintes:

- Desenvolvimento de estimadores semi-definidos positivos, susceptiveis de implementação eficiente, para a a função densidade de covariância e intensidade condicional de processos estocásticos pontuais.

- Modelização espectral AR do espectro das contagens de processos estocásticos pontuais.

- Aperfeiçoamento da modelização espectral AR de processos estocásticos pontuais, a partir de inclusão directa de informação sobre as características do espectro das contagens na região das altas frequências.

Capítulo 2

Análise espectral das contagens: métodos e algoritmos de implementação eficiente

Neste capítulo apresenta-se um estudo comparativo de diversos métodos de implementação eficiente para a estimação do espectro das contagens.

A abordagem adoptada e os resultados fundamentais aqui apresentados, permitem desenvolver algoritmos eficientes, em clara correspondência com os métodos correntes para séries temporais. Em particular, para além da análise espectral não paramétrica "clássica"(periodograma e método de Blackman-Tukey), possibilitam o desenvolvimento de métodos paramétricos alternativos, nomeadamente a modelização espectral AR.

Na aplicação dos algoritmos desenvolvidos neste trabalho é essencial a utilização de métodos adequados para a estimação das funções intensidade condicional e densidade de covariância de um processo pontual. Para além dos métodos correntes, descrevem-se também neste capítulo algoritmos de implementação eficiente, conduzindo sempre a estimadores semi-definidos positivos, estudando-se as suas propriedades assimpóticas e quantificando a distorsão no espectro estimado correspondente.

Na comparação de métodos de análise espectral é essencial definir as características a avaliar. Neste capítulo procura-se, utilizando dados de simulação, ilustrar e discutir a potência dos métodos na descrição do espectro das contagens. A estrutura do capítulo é a seguinte:

- Periodograma da série de acontecimentos observada: propriedades e limitações.
- Abordagem geral para a implementação eficiente da estimação do espectro das contagens. Discussão de dois casos particulares: método de Blackman-Tukey e modelização espectral AR.
- Estimação da sucessão densidade de covariância: desenvolvimento de estimadores que garantam a obtenção de sucessões semi-definidas positivas.
- Estudo de simulação ilustrativo dos estimadores descritos e comentários finais.

2.1 Introdução

Considere-se uma realização de um processo estocástico pontual consistindo em N ocorrências nos instantes $t_1, \ldots, t_N$ num intervalo $[0, T]$.

2.1.1 O periodograma como estimador para o espectro das contagens

O ponto de partida natural para a estimação do espectro das contagens, $G(\omega)$, consiste na estimação da intensidade média, λ, e da função de intensidade condicional, $m(t)$, e substituição na expressão (1.6).

Os métodos correntes para a estimação da função de intensidade condicional, $m(t)$, baseiam-se no histograma formado pelos $N(N-1)/2$ intervalos de tempo que se podem definir entre os instantes de ocorrência, $t_1, t_2, \ldots, t_N$, dos N acontecimentos no intervalo $[0, T]$ considerado [Cox e Lewis 66, Lewis 70].

Este processo pode ser representado analiticamente pela estatística $\tilde{m}(t)$

$$\tilde{m}(t) = N^{-1} \sum_{n=1}^{N-1} \sum_{k=1}^{N-n} \delta\left(t_{n+k} - t_n - |t|\right); \; 0 < |t| \leq T. \tag{2.1}$$

Saliente-se desde já que a implementação prática deste estimador torna necessária a introdução de algum amaciamento, o que será objecto de análise numa das secções seguintes.

O estimador espectral obtido [Lewis 70]

$$\hat{\lambda}\left(1 + \int_0^{+\infty} \tilde{m}(t) e^{j\omega t} dt + \int_0^{+\infty} \tilde{m}(t) e^{-j\omega t} dt\right), \tag{2.2}$$

onde $\hat{\lambda}$ indica a intensidade média estimada por

$$\hat{\lambda} = N/T \tag{2.3}$$

e $\tilde{m}(t)$ é dada por, (2.1) em $[0, T]$ e 0 fora desse intervalo, pode ser escrito na forma

$$I_T(\omega) = T^{-1} \, |X(j\omega)|^2, \text{onde } X(j\omega) = \sum_{n=1}^{N} e^{-j\omega t_n}, \tag{2.4}$$

sendo designado por *periodograma da série de acontecimentos observada.*

Verifica-se uma completa analogia com o periodograma convencional [Marple 87, Kay 88], quer na forma de obtenção, quer na expressão (2.4).

Com efeito, o periodograma da série de acontecimentos observada, $I_T(\omega)$, tem propriedades análogas ao periodograma convencional, em particular é não cêntrico e não consistente, sendo o desvio padrão da ordem de grandeza dos valores estimados [Cox e Lewis 66, Lewis 70].

No entanto, existem algumas diferenças importantes, que dificultam a estimação do espectro das contagens e a obtenção das propriedades do estimador (2.4). Nomeadamente, quer a função de densidade espectral das contagens, quer o periodograma da sucessão observada não são funções periódicas.

No que respeita a questões de implementação a principal dificuldade consiste no facto de a utilização de (2.4) poder levantar sérias limitações, particularmente quando o número de ocorrências é elevado. Em contraste com o periodograma convencional, não é possível a aplicação da Transformada Rápida de Fourier (FFT).

O efeito da duração finita do registo reflete-se de forma análoga ao caso de processos estocásticos em geral. Com efeito, para um registo de duração T, conclui-se a partir de (2.2) e da expressão para o erro de centricidade em $\tilde{m}(t)$ [Cox e Lewis 66, Lewis 70]

$$E[I_T(\omega)] = \lambda + \lambda^2 W_T(\omega) + \frac{1}{2\pi}(G(\omega) - \lambda) * W_T(\omega) \tag{2.5}$$

onde $W_T(\omega) = T((\sin(\omega T/2))/(\omega T/2))^2$. Apesar do estimador não ser cêntrico, é assimptoticamente cêntrico com exclusão da frequência $w = 0$, facto que decorre de em (1.6) e (2.2) não se ter levado em conta a intensidade média. No entanto, esta dificuldade pode ser facilmente obviada considerando o cálculo do periodograma em frequências da forma $\omega_p = \frac{2\pi p}{T}, p = 1, 2, \ldots$.

No caso de processos estocásticos pontuais, o processo de Poisson desempenha o papel do ruído branco Gausseano para processos estocásticos em geral.

Para analizar a variância do estimador considere-se o caso de um processo de Poisson com intensidade média λ. Considerando frequências ω_p, da forma indicada, conclui-se que [Cox e Lewis 66]:

- $2\lambda^{-1} I_T(\omega_p) \sim \chi^2_2 \quad \omega_p = \frac{2\pi p}{T}, p = 1, 2, \ldots$
- $I_T(\omega_p)$ independente de $I_T(\omega_m)$ para $p \neq m$.

Verifica-se assim que $\mathrm{E}[I_T(\omega_p)] = \lambda$ e $\mathrm{Var}(I_T(\omega_p)) = \lambda^2$. As flutuações em $I_T(\omega)$ tendem para a sua própria ordem de grandeza e a dispersão do periodograma não diminui quando $T \to \infty$.

Contrariamente ao que se passa para o caso do periodograma convencional, é dificil generalizar as conclusões para uma classe mais vasta de processos e o número de frequências independentes ω_p é infinito [Cox e Lewis 66].

2.1.2 Estimadores espectrais consistentes

Para obviar a dificuldade do periodograma ser um estimador não consistente poderá, da forma habitual e analogamente ao periodograma convencional, considerar-se quer o amaciamento quer a segmentação [Koopmans 74, Jenkins e Watts 68].

No primeiro caso deverá tomar-se o cuidado de não propagar o erro de centricidade para $w = 0$, quando se procede ao amaciamento; deverá também escolher-se uma função de peso truncada de forma adequada, uma vez que $I_T(\omega)$ não é periódico.

Assim, poderá considerar-se alternativamente:

i) o amaciamento do periodograma pelo cálculo de médias pesadas dos pontos do periodograma em torno das frequências ω_p. Isto é,

$$\overline{I}_T(\omega_p) = \sum_{m=-\frac{1}{2}K+\frac{1}{2}}^{\frac{1}{2}K-\frac{1}{2}} h_m I_T(\omega_{p+m}), \tag{2.6}$$

onde a sucessão de pesos deverá satisfazer à condição adicional $\sum_m h_m = 1$, K (neste caso particular suposto ímpar) indica o número de pontos em que é feito o amaciamento e p é um inteiro [Cox e Lewis 66, e referências aí incluidas].

ii) a segmentação, considerando o registo inicial dividido em K segmentos independentes e de igual duração $T' = T/K$, e determinando o periodograma médio

$$\overline{I}(\omega) = \frac{1}{K}\sum_{i=1}^{K}\frac{1}{T'}|X^i(j\omega)|^2, \text{ onde } X^i(j\omega) = \sum_{(i-1)T'<t_n<iT'} e^{-j\omega t_n}. \tag{2.7}$$

Considerando em **i)** o caso particular da janela de amacimento de Daniell, isto é pesos idênticos para todas as frequências, verifica-se, para as duas situações **i)** e **ii)**, que se obtem uma redução de variância de um factor de K, uma vez que a distribuição passa a ser proporcional a uma distribuição χ^2_{2K}.

Na figura 2.1 ilustra-se a inconsistência do periodograma e a redução na variância pela utilização de um periodograma médio $\overline{I}(\omega)$.

De uma forma geral verifica-se para qualquer das alternativas, **i)** e **ii)**, em analogia com o que se passa para processos estocásticos em geral, que reduzir o erro de centricidade e a variância são dois requisitos contraditórios, devendo estabelecer-se um compromisso em face do registo a analizar; a única forma de sair deste dilema consiste em aumentar a duração do registo [Koopmans 74, Jenkins e Watts 68].

No primeiro caso, o controlo do compromisso erro de centricidade/variância é feito a partir da janela do amaciamento: o erro de centricidade será tanto mais pequeno quanto menor for K mas, de uma forma geral, a variância diminui quando K aumenta.

No segundo caso deverá tomar-se em atenção que, embora a variância seja inversamente proporcional ao número de segmentos considerados, K, o erro de centricidade aumenta. Com efeito, a partir da expressão (2.5) conclui-se que a centricidade é tanto melhor quanto maior for a duração do segmento. Uma vez que $T = KT'$ é constante, à medida que o número de periodogramas aumenta a variância diminui. Assim o valor de K deverá ser escolhido, atendendo à informação à priori sobre o processo e as necessidades específicas de cada situação.

Além das limitações intrínsecas a qualquer estimador não paramétrico, a principal limitação no uso do peridograma, particularmente quando o número N de ocorrências é elevado, consiste na impossibilidade de usar a FFT, sendo a segmentação dos dados a única forma de reduzir o tempo de cálculo.

A utilização do periodograma médio permite reduzir o número de operações, proporcional a N^2 em (2.4), para N^2/K.

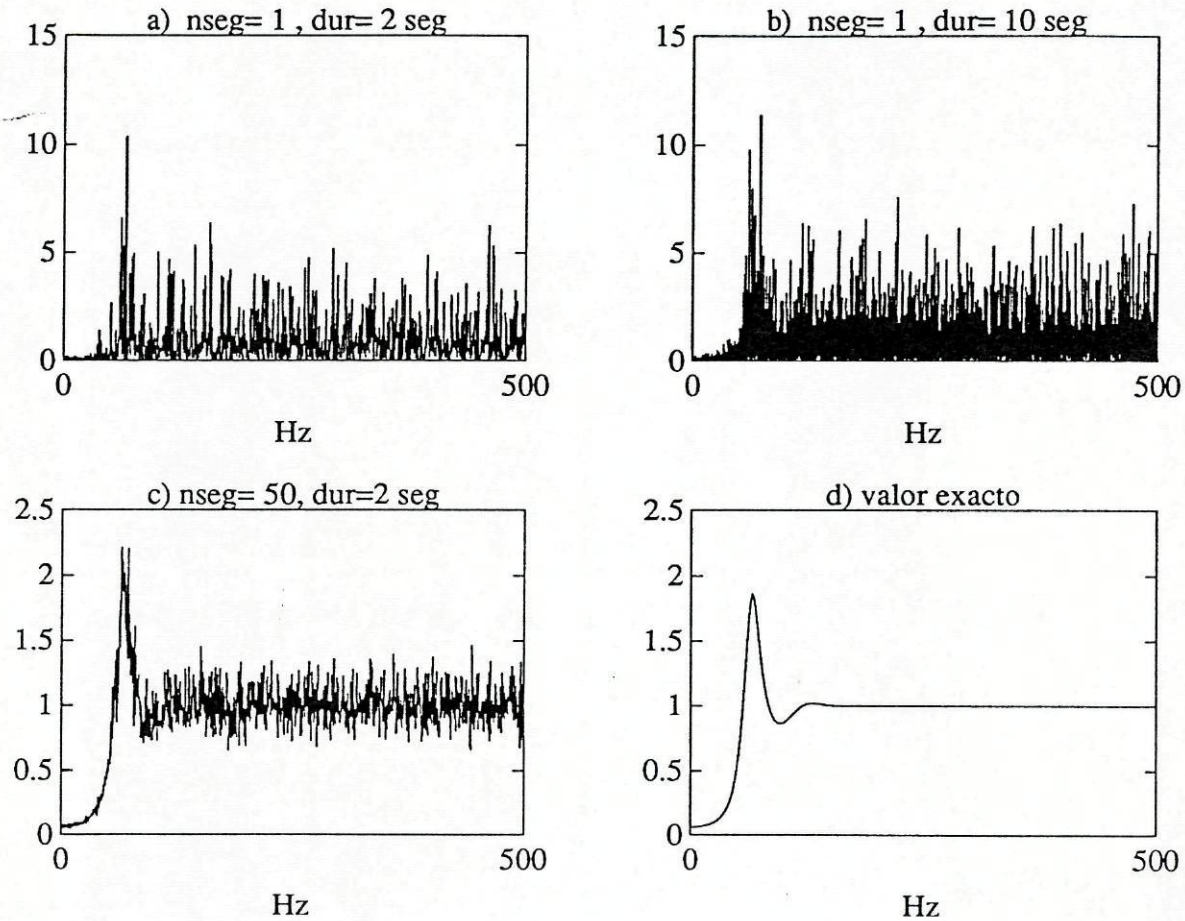


Figura 2.1: Ilustração da inconsistência do periodograma e da redução da variância pela utilização de um periodograma médio, para um processo estocástico pontual renewal Gausseano ($\lambda = 70s^{-1}, \sigma/\mu = 0.25$). Representação dos espectros das contagens estimados normalizados, $G(\omega)/\hat{\lambda}$, e do valor exacto correspondente; **nseg** e **dur** indicam, respectivamente, o número de segmentos e a sua duração.

2.2 Implementação eficiente da estimação do espectro das contagens

2.2.1 Introdução

Como alternativa à utilização do periodograma, têm sido sugeridos métodos de implementação eficiente, nomeadamente pela redução da estimação do espectro das contagens à estimação do espectro de um sinal amostrado regularmente [French e Holden 70, Lago e Jones 82].

No primeiro caso, o algoritmo de French-Holden permite obter, à custa do trem de funções δ associado ao processo pontual, um sinal contínuo o qual pode ser amostrado regularmente. Atendendo à característica fundamental dos processos pontuais, (expressão 1.8), os autores sugerem a aplicação de um filtro ideal truncado passa-baixo, o que possibilita uma amostragem sem sobreposição de bandas [French e Holden 70].

No segundo caso é aproveitado o facto de, na prática, os instantes observados em que ocorrem os acontecimentos, serem múltiplos de um determinado intervalo Δt, o qual caracteriza a resolução da técnica de medição. Apesar da estimação da função de densidade espectral de processos estocásticos pontuais e de sinais discretos amostrados regularmente serem problemas distintos, requerendo diferentes abordagens, na prática as diferenças são atenuadas em consequência da quantificação na medição dos instantes de ocorrência dos acontecimentos. Verifica-se que a estimação da função de densidade espectral de um processo pontual pode reduzir-se, sob certas condições, à estimação da função de densidade espectral de um sinal aleatório discreto de 0 e 1, associados à marca temporal mais próxima [Lago e Jones 82].

A definição de espectro das contagens e a analogia com a análise espectral de séries temporais (tabela 1.1) sugere naturalmente que o desenvolvimento de métodos eficientes (permitindo em particular a utilização da FFT,) poderá consistir em avaliar a função densidade de covariância em múltiplos de um intervalo de tempo Δt.

A desvantagem dos algoritmos que foram propostos com este objectivo [Brillinger 75, Brillinger 78], consiste no facto de não estar garantida a obtenção de estimativas não negativas para o espectro das contagens.

2.2.2 Resultados fundamentais

Dado um processo estocástico pontual com função densidade de covariância $g(t)$ definida em (1.2), considere-se a sequência $\{g_d(n\Delta t)\}$,

$$g_d(n\Delta t) = \lambda\,(\Delta t)^{-1}\,\delta_{0n} + \lambda\,(m(n\Delta t) - \lambda);\; n = 0, \pm 1, \ldots \tag{2.8}$$

onde δ_{ij} indica o símbolo delta de Kronecker. Considerando o espectro da sucessão densidade de covariância $\{g_d(n\Delta t)\}$, $G_d(\omega)$, definido por

$$G_d(\omega) = \Delta t \sum_{n=-\infty}^{n=+\infty} g_d(n\Delta t)\, e^{-jn\omega\Delta t} \tag{2.9}$$

conclui-se a partir de (1.2), (2.8) e (2.9) que $G_d(\omega)$ e $G(\omega)$ estão relacionados por

$$G_d(\omega) = \lambda + (G(\omega) - \lambda) * \sum_{n=-\infty}^{n=+\infty} \delta(\omega + 2\pi n/\Delta t) \tag{2.10}$$

onde $*$ indica a operação de convolução.

A equação (2.10) mostra que na prática $G_d(\omega)$ é uma boa aproximação de $G(\omega)$ para uma escolha adequada de Δt. Com efeito, uma vez que $|G(\omega) - \lambda|$ tende para zero rapidamente para valores crescentes de $|\omega|$, conclui-se facilmente de (2.10) que basta escolher Δt suficientemente pequeno, de forma que $|G(\omega) - \lambda| \simeq 0$ para frequências $|\omega| \geq \pi/\Delta t$. Nestas condições não surgem problemas com a sobreposição de bandas, podendo aceitar-se a aproximação $G_d(\omega) \simeq G(\omega)$ para $|\omega| \leq \pi/\Delta t$.

Conclui-se então, para uma escolha adequada de Δt e considerando frequências $|\omega| \leq \pi/\Delta t$, que:

- Uma estimativa para o espectro das contagens pode ser obtida a partir da estimação da sucessão discreta $\{g_d(n\Delta t)\}$, seguida por, em alternativa,
  - cálculo da transformada de Fourier, corrigida pelo factor multiplicativo Δt, equação (2.9); analogamente ao caso de séries temporais a estimação de $G(\omega)$, é reduzida à estimação de $\{g_d(n\Delta t)\}$ (*método de Blackman-Tukey*).
  - modelização do espectro da sucessão $\{g_d(n\Delta t)\}$, $G_d(\omega)$, definido por (2.9)
- De uma forma geral, a análise espectral de um processo pontual pode ser reduzida, a partir da equação (2.9), à análise espectral de uma sucessão discreta com função de covariância $C(n) = \Delta t \cdot g_d(n\Delta t)$.

Na formulação do problema aqui apresentada, é dada especial atenção à característica fundamental de um processo pontual (expressão 1.8), evitando os problemas inerentes à sobreposição de bandas. Assim, com exclusão da origem, a sucessão (2.8) pode ser considerada como obtida por amostragem de $g(t)$. A separação da origem é essencial para a descrição adequada do espectro na região das altas frequências e comparação das sequências estimadas para diferentes frequências de amostragem.

A analogia com a análise espectral de processos estocásticos verifica-se, na condição imposta a Δt, uma vez que esta é equivalente ao teorema de amostragem [Koopmans 74]. Também a relação entre $G(\omega)$ e a função de densidade espectral, $\Phi(\omega)$, de uma sucessão discreta com covariância $C(n)$, corresponde à relação entre o espectro de um processo estocástico contínuo e o espectro da sucessão discreta obtida por amostragem (intervalo de amostragem Δt) [Koopmans 74]; assim, $\hat{G}(\omega) = \hat{\Phi}(\omega\Delta t)$, $|\omega| \leq \pi/\Delta t$.

Finalmente será importante salientar que, ao reduzir o problema da análise espectral de processos pontuais ao cálculo do espectro de uma sucessão discreta torna-se possível, de uma forma geral, a utilização e adaptação dos métodos eficientes (paramétricos e não paramétricos) correntes na literatura de Processamento de Sinal, como alternativa aos métodos específicos da literatura de processos pontuais.

2.2.3 O método de Blackman-Tukey

Supondo que Δt foi escolhido de modo a satisfazer as condições formuladas na secção 2.2.2, poderá usar-se a equação (2.9) para a estimação do espectro das contagens a partir da sucessão densidade de covariância estimada, $\{\hat{g}_d(n\Delta t)\}$.

Dado um registo consistindo num número finito de instantes de ocorrência, $t_1, \ldots, t_N$, num intervalo $[0, T]$, o problema que se põe na prática consiste no facto de se pretender estimar um número infinito de valores da sucessão densidade de covariância. Analogamente ao que se passa para séries temporais em geral, a forma natural de ultrapassar esta dificuldade consiste em substituir a sucessão densidade de covariância por uma sequência finita, estimada com base na amostra.

Restringindo a soma em (2.9) ao domínio finito apropriado, $n = -M, \ldots M$, onde $M = int(T/\Delta t)$, o espectro das contagens poderá então ser estimado, a partir de

$$\hat{G}(\omega) = \Delta t \sum_{n=-M}^{M} \hat{g}_d(n\Delta t)\, e^{-jn\omega\Delta t}, \quad |\omega| \leq \pi/\Delta t. \tag{2.11}$$

O problema da estimação da sucessão densidade de covariância será estudado detalhadamente na secção 2.3. Conforme será ilustrado, as sucessões densidade de covariância estimadas, $\{\hat{g}_d(n\Delta t)\}$, apresentam grande variabilidade (particularmente para os valores elevados de n) [Lago et al 89].

Assim, o estimador espectral (2.11) apresenta enorme variabilidade, reflectindo a variabilidade inerente aos valores estimados para a densidade de covariância. Este efeito é inteiramente análogo ao que se passa para séries temporais em geral [Priestley 81] e a forma corrente de obviar este problema consiste dar um menor peso aos valores $\{\hat{g}_d(n\Delta t)\}$, para n elevado.

O *método de Blackman-Tukey* para processos pontuais consistirá então em,

$$\hat{G}_{BT}(\omega) = \Delta t \sum_{n=-M}^{n=+M} w(n)\hat{g}_d(n\Delta t)\, e^{-jn\omega\Delta t}, \tag{2.12}$$

onde, tomando $L \leq M-1$, $w(n)$ indica uma janela na sucessão densidade de covariância com as seguintes propriedades:

$$\begin{aligned} &\text{a)} \quad 0 \leq \omega(k) \leq w(0) = 1 \\ &\text{b)} \quad \omega(-k) = \omega(k) \\ &\text{c)} \quad \omega(k) = 0, |k| > L. \end{aligned} \tag{2.13}$$

A janela $w(n)$ deverá ser escolhida de forma que a sua inclusão não altere o caracter não negativo do espectro das contagens, isto é deverá ser semi-definida positiva. A janela mais simples que satisfaz a esta última condição é a janela de Bartlett,

$$w(n) = \begin{cases} 1 - |i|/m & \text{para } |n| < L \\ 0 & \text{restantes valores de } n. \end{cases}$$

Finalmente, saliente-se que o método de Blackman-Tukey para processos estocásticos pontuais se reduz ao método correspondente para processos estocásticos discretos (processo discreto com função de covariância $C(n) = \Delta t.g_d(n\Delta t)$). Isto é,

$$\hat{G}(\omega/\Delta t) = \sum_{n=-M}^{n=+M} w(n)\Delta t\hat{g}_d(n\Delta t)\, e^{-jn\omega}, |\omega| < \pi/\Delta t.$$

Assim, poderão utilizar-se todos os resultados na literatura de séries temporais relativamente às suas propriedades e distribuição [Jenkins e Watts 68, Koopmans 74, Priestley 81]. Em particular, também aqui não se poderá reduzir simultaneamente o erro de centricidade e a variância, e a escolha de janelas adequadas permite obter estimadores espectrais consistentes.

O compromisso *"erro de centricidade/variância"* é controlado, da forma habitual, por uma escolha adequada da janela na densidade de covariância, em particular do valor de L ($L \ll M$) a considerar. De forma análoga ao que se passa para processos estocásticos em geral, a diminuição na variância (considerando valores de L pequenos) é acompanhada de uma diminuição na resolução e de um aumento no erro de centricidade, conforme será ilustrado na secção 2.4 [Priestley 81].

2.2.4 Análise espectral AR de processos pontuais

Os resultados fundamentais apresentados na secção 2.2.2, permitem, para escolhas adequadas de Δt, reduzir a análise espectral de um processo estocástico pontual à análise espectral de uma série temporal discreta com covariância $\{\Delta t.g_d(n\Delta t)\}$.

Com efeito, $G_d(\omega)$ (determinado pela equação (2.9) a partir do espectro da sucessão $\{\Delta t.g_d(n\Delta t)\}$), é uma boa aproximação para $G(\omega)$, para $|\omega| \leq \pi/\Delta t$, se Δt for escolhido de forma que $|G(\omega) - \lambda| \approx 0$ para ω fora da gama $|\omega| \leq \pi/\Delta t$.

Suposta satisfeita esta condição, o problema da *modelização do espectro* $G_d(w)$ pode ser formulado na forma seguinte:

Formulação 1 *Dado* $G_d(\omega)$ *e um determinado número de polos* p, *pretende-se aproximar* $G_d(\omega)$ *por um espectro autoregressivo* $G_a(\omega)$,

$$G_a(\omega) = \frac{V}{|1 + \sum_{n=1}^{p} a_n e^{-jnw\Delta t}|^2} \tag{2.14}$$

usando para medida de erro, $E(p)$

$$E(p) = \frac{V\Delta t}{2\pi} \int_{-\pi/\Delta t}^{+\pi/\Delta t} \frac{G_d(\omega)}{G_a(\omega)}\, d\omega. \tag{2.15}$$

Os parâmetros $\{a_n\}$ são a solução do sistema de p equações lineares [Makhoul 75]

$$\sum_{k=1}^{p} a_k g_d(|n-k|\Delta t) = -g_d(n\Delta t), \;\; n = 1, \ldots, p \tag{2.16}$$

e pode mostrar-se que V é o erro médio quadrático mínimo, o qual é dado por

$$E(p) = \Delta t \, (g_d(0\Delta t) + \sum_{k=1}^{p} a_k g_d(k\Delta t)). \tag{2.17}$$

As equações (2.16), para a determinação dos parâmetros autoregressivos, podem ser escritas na forma matricial

$$\boldsymbol{C}_{pp} \, \boldsymbol{a}_p = -\boldsymbol{c}_p, \tag{2.18}$$

onde $\boldsymbol{C}_{pp}$ indica a matriz densidade de covariância, matriz simétrica Toeplitz de dimensão $(p \times p)$, cujos elementos são dados por

$$C_{pp}(i,j) = g_d(|i-j|\Delta t), \;\; 1 \leq i \leq p, \; 1 \leq j \leq p,$$

$\boldsymbol{a}_p$ é o vector de parâmetors autoregressivos e $\boldsymbol{c}_p$ um vector com elementos

$$c_p(i) = g_d(i\Delta t), \;\; 1 \leq i \leq p.$$

Os parâmetros autoregressivos podem ser obtidos por resolução directa do sistema (2.18). No entanto, uma vez que a matriz dos coeficientes $\boldsymbol{C}_{pp}$ é simétrica Toeplitz, aquele sistema pode ser resolvido recursivamente sem inversão matricial para valores crescentes de p usando o algoritmo de Levinson [Marple 87].

O erro quadrático mínimo para o modelo de ordem p é determinado a partir dos parâmetros autoregressivos estimados. Alternativamente, a aplicação do algoritmo de Levinson também permite a determinação recursiva do erro quadrático mínimo em cada passo. Note-se, no entanto, que nesse caso será necessário incluir o factor de escala Δt na definição da matriz $\boldsymbol{C}_{pp}$ e no vector $\boldsymbol{c}_p$.

A função densidade de covariância é semi-definida positiva. Assim, a sucessão densidade de covariância $\{g_d(n\Delta t)\}$ definida por (2.9) é também semi-definida positiva. Pode-se assim concluir que a solução de (2.16) e de (2.18) corresponde a um modelo espectral $G_a(\omega)$ com polos no interior do círculo unitário e a uma medida de erro $E(p)$ não negativa.

Note-se que os parâmetros $\{a_n\}$ e o erro médio quadrático mínimo V são os mesmos que se obteriam pela modelização autoregressiva de uma série temporal discreta com função de covariância $\{\Delta t.g_d(n\Delta t)\}$ a partir das equações de Yule-Walker.

Analogamente ao caso do método de Blackman-Tukey, esta analogia permite utilizar, na análise espectral AR de processos pontuais, os resultados correntes na literatura de Processamento de Sinal [Kay 88, Marple 87]. Em particular, sugere que os critérios para a selecção da ordem dos modelos autoregressivos para uma série temporal discreta também possam ser aplicados para a selecção do número de coeficientes no modelo $G_a(\omega)$ para o espectro das contagens.

Nos exemplos apresentados neste trabalho serão usados os critérios de AIC e CAT [Akaike 74, Parzen 74]. É importante referir que a aplicação desses critérios para a selecção da ordem, p, requer que a propriedade de $g(t)$ ser uma função semi-definida positiva seja preservada nas sucessões estimadas. Este aspecto será retomado na secção 2.4.

2.2.5 Comentários finais

Os resultados fundamentais apresentados na secção 2.2.2 sugerem a utilização de todas as técnicas (correntes e de implementação eficiente) em Processamento de Sinal na estimação do espectro das contagens, como alternativa aos métodos específicos descritos na literatura de processos estocásticos pontuais.

Em particular, permitem desenvolver uma análise espectral não paramétrica "clássica" em analogia com os métodos correntes para séries temporais. A abordagem adoptada permite reduzir a estimação do espectro das contagens à estimação da sucessão densidade de covariância em múltiplos de um determinado intervalo de tempo Δt e formular o método de Blackman-Tukey para processos estocásticos pontuais.

A escolha de janelas adequadas permite obter estimadores espectrais consistentes. Para garantir a obtenção de valores estimados não negativos, deverão utilizar-se janelas na densidade de covariância bem como sucessões densidade de covariância estimadas semi-definidas positivas.

Tem havido um grande desenvolvimento de novas técnicas para estimação espectral [Kay 88, Marple 87]. Tem sido dada uma grande ênfase no desenvolvimento da análise espectral paramétrica como alternativa aos métodos não paramétricos clássicos. Em particular, a modelização espectral autoregressiva (AR) tem adquirido uma grande importância na análise de séries temporais.

A análise espectral AR pode ser também aplicada à análise espectral de processos pontuais com vantagens semelhantes sobre a abordagem não paramétrica. A análise espectral autoregressiva de processos estocásticos pontuais pode ser motivada de diversas formas, e neste trabalho foi formulada a partir de modelização espectral.

Na aplicação da modelização espectral AR e na utilização dos critérios correntes para a selecção da ordem também será fundamental utilizar sucessões densidade de covariância estimadas garantidamente semi-definidas positivas. Se esta condição for verificada o modelo espectral AR, $G_a(\omega)$, tem polos no interior do círculo unitário e a medida de erro $E(p)$ é não negativa.

Finalmente, note-se que a aplicação do método de Blackman-Tukey permite ultrapassar o problema prático de, dada uma realização num intervalo $[0, T]$, pretender estimar um número infinito de valores da sucessão densidade de covariância.

Alternativamente na modelização espectral AR, os modelos espectrais são baseados num número reduzido de valores da densidade de covariância e (tal como é inerente aos modelos espectrais AR) a sucessão densidade de covariância é extrapolada para as ordens superiores.

2.3 Estimação da função densidade de covariância para a análise espectral de processos pontuais

O primeiro passo para a implementação eficiente da análise espectral de processos estocásticos pontuais consiste na estimação da função densidade de covariância avaliada em múltiplos de um determinado intervalo de tempo Δt.

Apresentam-se nesta secção métodos eficientes para a estimação da densidade de covariância, avaliada em múltiplos de um determinado intervalo de tempo Δt, adequados para a análise espectral de processos pontuais. Assim, para uma descrição espectral correcta na região das altas frequências, será dada particular atenção à estimação daquelas funções na origem.

Tendo em vista a aplicação dos métodos descritos neste trabalho, será fundamental o desenvolvimento de estimadores que garantam a obtenção de sucessões estimadas semi-definidas positivas.

2.3.1 Introdução

Os métodos tradicionalmente usados para estimação da função densidade de covariância, baseiam-se na estimação prévia da intensidade média e da intensidade condicional e substituição na definição (1.2) [Cox e Lewis 66].

Conforme já foi referido na Introdução (definição do periodograma), os métodos correntes para a estimação da função de intensidade condicional, $m(t)$, baseiam-se no histograma formado pelos $N(N-1)/2$ intervalos de tempo que se podem definir entre os instantes de ocorrência dos N acontecimentos no intervalo $[0,T]$ [Cox e Lewis 66, Brillinger 75]. Assim, considerando $[0,T]$ dividido em intervalos de amplitude Δt e o histograma correspondente, obtem-se a estimativa

$$\hat{m}(n\Delta t + \Delta t/2) = (N\Delta t)^{-1}(\#[(j,k); n\Delta t < t_j - t_k < n\Delta t + \Delta t]), n = 0, 1, \ldots . \quad (2.19)$$

Este processo pode ser representado analiticamente como correspondendo ao amaciamento da estimativa não amaciada $\tilde{m}(t)$, dada por (2.1), pela função peso de Daniell [Cox e Lewis 66]. Com efeito, esta função traduz o agrupamento em intervalos de amplitude Δt, podendo escrever-se

$$\hat{m}(n\Delta t + \Delta t/2) = (\Delta t)^{-1} \int_{n\Delta t}^{(n+1)\Delta t} \tilde{m}(u)\, du, \;\; n = 0, 1, \ldots \quad (2.20)$$

Note-se que embora de implementação extremamente eficiente, uma vez que envolve apenas operações de classificação e contagem, o estimador da intensidade condicional (2.20) só é adequado se $|n\Delta t| \ll T$ e apresenta nos outros casos um grande erro de

centricidade. Pode no entanto ser introduzida uma correcção de centricidade em (2.20) quer multiplicando pelo factor correctivo [Cox e Lewis 66]

$$c_1(t) = T\,(T - |t|)^{-1}, \tag{2.21}$$

quer adicionando o factor correctivo linear [Brillinger 75]

$$c_2(t) = N\,|t|\,T^{-2}. \tag{2.22}$$

A estimativa amaciada corrente para $g(t)$ $(t \neq 0)$ correspondente a (2.20) é

$$\hat{g}(n\Delta t + \Delta t/2) \;=\; \hat{\lambda}(\hat{m}(n\Delta t + \Delta t/2) - \hat{\lambda}), (n \neq 0) \tag{2.23}$$

onde $\hat{\lambda}$ indica a intensidade média estimada por (2.3).

2.3.2 Métodos de estimação baseados no histograma das contagens

Para a obtenção de estimadores para as funções de intensidade condicional e densidade de covariância avaliadas em múltiplos "inteiros" de Δt, deverá tomar-se em especial atenção a definição da origem (ver (1.2) e (2.8)).

Por conveniência introduz-se a função intensidade condicional completa, $v(t)$, definida por [Lago et al 89]

$$v(t) = \delta(t) + m(t). \tag{2.24}$$

A função simétrica $\tilde{v}(t)$ definida, no intervalo $[-T, T]$, por

$$\tilde{v}(t) = \delta(t) + \tilde{m}(t) \tag{2.25}$$

com $\tilde{m}(t)$ dada por (2.1), é uma estimativa não amaciada para $v(t)$.

Estimativas amaciadas definidas em múltiplos inteiros de Δt podem ser obtidas nomeadamente de

$$\hat{v}_d(n\Delta t) = (\Delta t)^{-1}\;(\;\delta_{0n} + \int_{n\Delta t - \Delta t/2}^{n\Delta t + \Delta t/2} \tilde{m}(u)\,du\;), \tag{2.26}$$

e duma forma geral pode escrever-se

$$\hat{v}_d(n\Delta t) = \Delta t^{-1}\,\delta_{0n} + \;(\tilde{m}(t) * w(t))_{t=n\Delta t} \tag{2.27}$$

onde $\omega(t)$ designa uma janela de área unitária [Lago et al 89].

O caso particular do estimador (2.26) corresponde a um amaciamento por $\omega_0(t)$, janela rectangular unitária de duração Δt.

Analogamente aos métodos correntes, pode ser introduzida correcção do erro de centricidade em (2.26). A implementação do método é eficiente envolvendo apenas

operações simples de classificação e contagem. Assim, o estimador definido por (2.26) pode ser obtido a partir de

$$\hat{v}_d(n\Delta t) = \Delta t^{-1}(\delta_{0n} + \#[(j,k); n\Delta t - \Delta t/2 < t_j - t_k < n\Delta t + \Delta t/2]/N). \quad (2.28)$$

Estimada a sucessão $\{\hat{v}_d(n\Delta t)\}$, $\{\hat{g}_d(n\Delta t)\}$ poderá ser obtida a partir de

$$\hat{g}_d(n\Delta t) = \hat{\lambda}\,(\hat{v}_d(n\Delta t) - \hat{\lambda}) \quad (2.29)$$

onde $\hat{\lambda}$ indica a intensidade média estimada por (2.3).

A sucessão densidade de covariância estimada a partir de (2.26) e (2.29), com janela rectangular $\omega_0(t)$, será designada neste trabalho por $\{\hat{g}_{0d}(n\Delta t)\}$.

O isolamento da origem em (2.26) garante uma descrição correcta na zona das altas frequências. Com efeito, escolhendo Δt de forma adequada (ver secção 2.2.2) e indicando por $G_{0d}(\omega)$ o espectro correspondente a $\{g_{0d}(n\Delta t)\}$, obtido a partir de (2.9), verifica-se que o comportamento assimptótico pode ser descrito por

$$\lim_{T\to\infty} E[G_{0d}(\omega)] \simeq \lambda + W_0(\omega)\,(G(\omega) - \lambda), |\omega| < \pi/\Delta t$$

onde $W_0(\omega) = (\sin(\omega\Delta t/2)/(\omega\Delta t/2))$ designa a transformada de Fourier de $\omega_0(t)$.

Este comportamento assimptótico será ilustrado na secção 2.3.3 e pode ser generalizado a qualquer janela de amaciamento, $w(t)$, desde que na expressão anterior se substitua $W_0(\omega)$ por $W(\omega)$.

As funções de intensidade condicional completa e densidade de covariância são semi-definidas positivas, muito embora esta característica não seja preservada pelos estimadores (2.27) e (2.29).

Assim, o espectro associado, $\hat{G}_d(\omega)$, calculado a partir de (2.9) não é garantidamente não negativo. Relembra-se também que, para a aplicação da modelização espectral autoregressiva do espectro das contagens e em particular para a utilização dos critérios correntes para a selecção da ordem, será necessário que a sucessão densidade de covariância estimada seja garantidamente semi-definida positiva.

Finalmente será de salientar que os algoritmos para estimação da função densidade de covariância em múltiplos de um intervalo Δt referidos na literatura [Brillinger 75, Brillinger 78] correspondem a uma abordagem deste tipo, onde a função densidade de covariância é estimada a partir da intensidade média e intensidade condicional. Não fica, portanto, também garantido que as sucessões densidade de covariância estimadas sejam semi-definidas positivas.

2.3.3 Estimadores semi-definidos positivos

A estimação da sucessão $\{\hat{v}_d(n\Delta t)\}$ a partir de (2.27) conduz a sucessões semi-definidas positivas, desde que no amaciamento de $\tilde{m}(t)$ se considere uma função peso semi-

-definida positiva. No entanto a sucessão densidade de covariância calculada a partir da expressão (2.29) poderá não ser semi-definida positiva, mesmo que a sucessão $\{\hat{v}_d(n\Delta t)\}$ o seja.

Assim, para garantir aquela propriedade, parece preferivel estimar directamente a função densidade de covariância e não partir da estimação prévia da função de intensidade condicional. Métodos alternativos deste tipo serão descritos nesta secção.

A analogia com a análise espectral de séries temporais, sugere naturalmente que uma forma de obviar as dificuldades encontradas poderá consistir em abordar o problema da estimação da sucessão densidade de covariância de forma semelhante ao da estimação da função de covariância $C(t)$ de uma série temporal $\{x(t)\}$ com média desconhecida μ. Por outro lado, a função $\tilde{v}(t)$ pode exprimir-se como função de autocorrelação de um trem de funções δ, isto é [Lago et al 89]

$$\tilde{v}(t) = N^{-1} \int_{0-}^{T-|t|} \theta(u)\, \theta(u + |t|)\, du, \tag{2.30}$$

onde

$$\theta(t) = \sum_{k=1}^{N} \delta(t_k - t). \tag{2.31}$$

No caso de processos estocásticos pontuais é assim sugerida a introdução da estimativa não amaciada para a função densidade de covariância, $\tilde{g}_1(t)$ [Lago et al 89]

$$\tilde{g}_1(t) = T^{-1} \int_{0-}^{T-|t|} (\theta(u) - \hat{\lambda})\, (\theta(u + |t|) - \hat{\lambda})\, du \tag{2.32}$$

com $\hat{\lambda}$ e $\theta(t)$ dados por (2.3) e (2.31), respectivamente.

Na tabela 2.1 descrevem-se comparativamente as operações envolvidas para processos estocásticos em geral (P.E.) e processos estocásticos pontuais (P.P.), dada uma realização num intervalo de amplitude T. Como é sabido, no caso dos processos estocásticos os estimadores associados, $\hat{C}(t)$, são garantidamente semi-definidos positivos [Priestley 81], o mesmo acontecendo com $\tilde{g}_1(t)$.

P.E.	$\hat{\mu} = T^{-1} \int_0^T x(u)du$	$\hat{C}(t) = T^{-1} \int_0^{T-\lvert t\rvert} (x(u) - \hat{\mu})(x(u + \lvert t\rvert) - \hat{\mu})du$
P.P.	$\hat{\lambda} = NT^{-1}$	$\tilde{g}_1(t) = T^{-1} \int_{0-}^{T-\lvert t\rvert} (\theta(u) - \hat{\lambda})(\theta(u + \lvert t\rvert) - \hat{\lambda})du$

Tabela 2.1:

Adoptando esta abordagem poderão desenvolver-se métodos para obter uma sucessão $\{\hat{g}_d(n\Delta t)\}$ com as propriedades pretendidas, nomeadamente [Lago et al 89]:

a) O método consistirá em:

i) Cálculo de $\tilde{g}_1(t)$

ii) Amaciamento com uma função semidefinida positiva $w_1(t)$

iii) Amostragem a Δt

o que conduz à estimativa

$$\hat{g}_{1d}(n\Delta t) = \hat{\lambda}\,\Delta t^{-1}\,\delta_{0n} + \hat{q}_1(n\Delta t) \tag{2.33}$$

onde

$$\hat{q}_1(t) = T^{-1}\,w_1(t) * \int_{0+}^{T-|t|} (\theta(u) - \hat{\lambda})\,(\theta(u+|t|) - \hat{\lambda})\,du \tag{2.34}$$

b) Alternativamente, as expressões (2.32) e (2.33) sugerem a troca das fases de amaciamento e amostragem, isto é considerar nomeadamente [Lago et al 89]:

i) amaciamento de $\theta(t) - \hat{\lambda}$ com uma função $w_2(t)$

ii) amostragem a Δt

iii) cálculo da sucessão densidade de covariância

obtendo-se a estimativa

$$\hat{g}_{2d}(n\Delta t) = m^{-1} \sum_{k=0}^{m-|n|} \phi(k)\phi(k+|n|),\ \ n = 0, \pm 1, \ldots, \pm m - 1 \tag{2.35}$$

onde $m = int(T/\Delta t)$, $\phi(n) = (w_2(t) * (\theta(t) - \hat{\lambda}))_{t=n\Delta t+\Delta t/2}$ e $(\theta(t) - \hat{\lambda})$ definido por (2.31) em $[0, T]$ e zero fora desse intervalo.

No primeiro caso, o isolamento da origem permite obter um comportamento assimptótico adequado, bastando escolher uma janela $w_1(t)$ semi-definida positiva para garantir a obtenção de sucessões semi-definidas positivas. No segundo caso, embora seja preservada aquela propriedade, terá de se escolher a janela $w_2(t)$ e a sua amplitude de forma a satisfazer a determinadas condições que garantam um comportamento assimptótico correcto na região das altas frequências[Lago et al 89]. As janelas mais simples que satisfazem às condições exigidas são, respectivamente,

a) $w_1(t)$, janela triangular unitária em $[-\Delta t, \Delta t]$

b) $w_2(t)$, janela rectangular unitária em $[-\Delta t/2, \Delta t/2]$

as quais conduzem a comportamentos assimptóticos idênticos.

Com efeito, nestes dois casos, indicando por $G_{id}(\omega)$, $i = 1, 2$, os espectros estimados correspondentes, obtidos a partir de (2.9) e escolhendo Δt satisfazendo à condição indicada na secção 2.2.2, conclui-se que [Lago et al 89]

$$\lim_{T\to\infty} E[G_{id}(\omega)] \simeq \lambda + W(\omega)\,(G(\omega) - \lambda), |\omega| < \pi/\Delta t, i = 1, 2. \qquad (2.36)$$

Nesta expressão $W(\omega) = W_1(\omega) = |W_2(\omega)|^2$, onde $W_1(\omega)$ e $W_2(\omega)$ designam, respectivamente as transformadas de Fourier das funções $w_1(t)$ e $w_2(t)$.

A equação (2.36) traduz a distorsão provocada pelo amaciamento, e poderá ser utilizada para a escolha da duração da janela, sendo idêntica para as janelas consideradas e dada por $W(\omega) = (\sin(\omega\Delta t/2)/(\omega\Delta t/2))^2, i = 1, 2$.

Conforme se ilustra na figura 2.2, o comportamento assimptótico (quando $T \to \infty$) do espectro estimado a partir de (2.33) ou (2.35), em **b**), é idêntico ao obtido com (2.26) e (2.29), em **a**), (escala de frequências normalizada pela multiplicação por Δt).

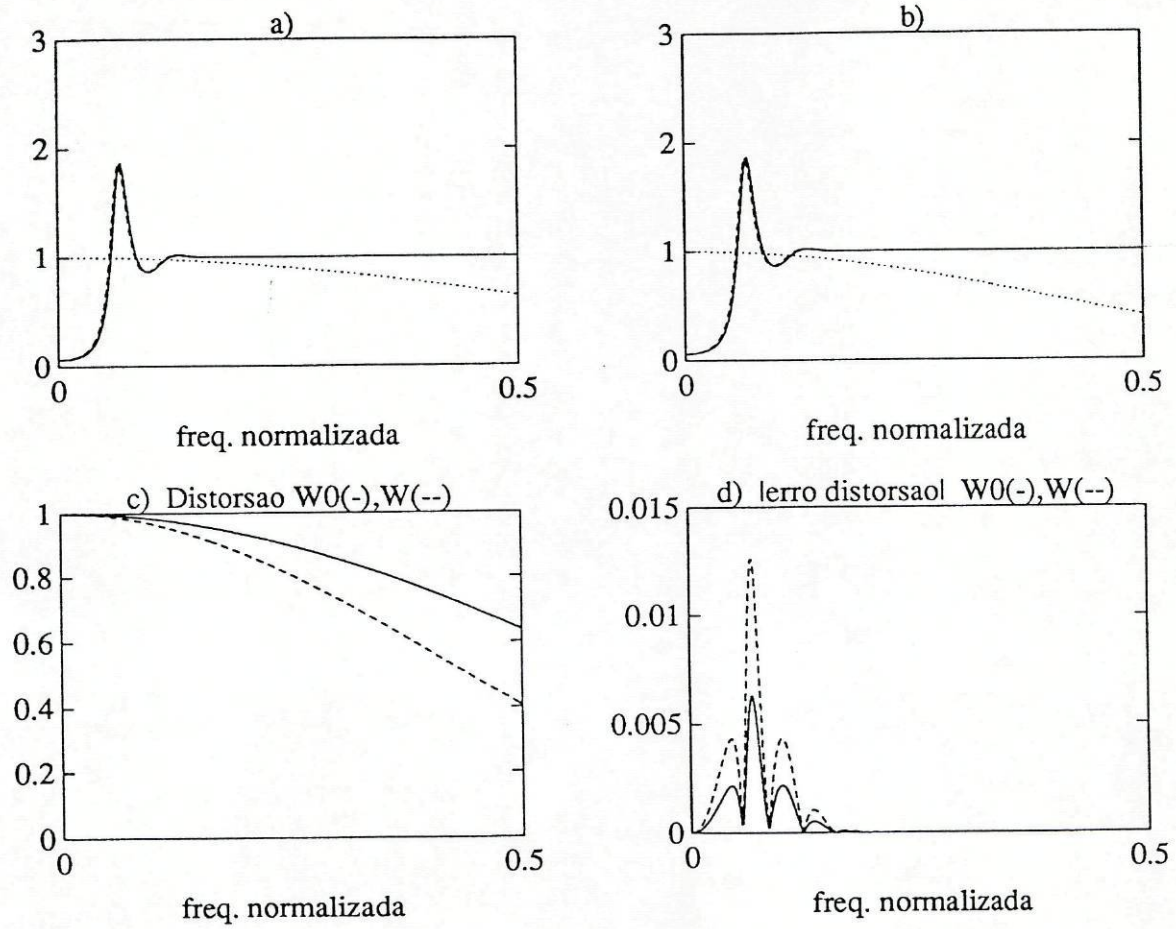


Figura 2.2: Ilustração do efeito de distorsão ($T \to \infty$) na estimação do espectro das contagens de um processo renewal Gausseano ($\lambda = 70s^{-1}, \sigma/\mu = 0.25$). Escala de frequências normalizada, $\Delta t = 1$ ms. Em **a**) e **b**) representam-se o valor exacto (-) e o valor assimptótico (- -) do espectro estimado a partir de ((2.26) e (2.29)) e ((2.33) ou (2.35)), respectivamente, bem como a distorsão (. . .), dada por $W_0(\omega)$ em **a**) e $W(\omega)$ em **b**). Em **c**) e **d**) representam-se sobrepostos a distorsão e o erro assimptótico, para $W_0(\omega)$ (-) e $W(\omega)$ (- -).

Para a resolução temporal empregada, o efeito de distorsão (...) não é significativo, observando-se a sobreposição do valor exacto(-) e do valor assimptótico do espectro estimado (- -), dado por (2.36), tanto para $W_0(\omega)$ (a), como para $W(\omega)$ (b).

Finalmente na figura 2.2 representa-se, respectivamente em **c)** e **d)**, o efeito de distorsão e o comportamento assimptótico do erro de distorsão ($T \rightarrow \infty$), para as duas situações, $W_0(\omega)$ (-) e $W(\omega)$ (- -).

Conclui-se assim, em qualquer dos casos, que o espectro das contagens definido por (2.9), é uma boa aproximação de $G(\omega)$, desde que Δt seja suficientemente pequeno para que se verifique que

- $|G(w) - \lambda| \simeq 0$ para frequências $|w| \geq \pi/(\Delta t)$
- o efeito de distorsão da janela não seja significativo na zona onde o espectro difere significativamente de λ.

Finalmente será importante salientar que a aplicação dos métodos de estimação desenvolvidos nesta secção, também requer apenas a utilização de operações de classificação e contagem baseadas na sequência de acontecimentos observada, conduzindo portanto a algoritmos simples e de implementação extremamente eficiente.

De notar ainda que, se a sucessão densidade de covariância estimada for semi-definida positiva, a estimativa para a função de intensidade condicional completa obtida a partir de

$$\hat{v}_d(n\Delta t) = \hat{\lambda}^{-1}\, \hat{g}_d(n\Delta t) + \hat{\lambda} \tag{2.37}$$

é garantidamente semi-definida positiva.

2.3.4 Estimação de $\{g_d(n\Delta t)\}$: resultados de simulação

Na figura 2.3 apresenta-se um exemplo de estimação da função densidade de covariância de um processo pontual renewal Gausseano.

A partir destes exemplos e de muitos outros, pode-se concluir que as estimativas obtidas com técnicas de construção de histogramas (secção 2.3.2) e com os algoritmos apresentados neste trabalho (secção 2.3.3) são muito semelhantes no que respeita ao valor médio e à variância. Contudo, poderão surgir diferenças significativas quando se utilizam as estimativas da densidade de covariância na estimação do espectro das contagens, a partir de (2.9).

Com efeito, podem observar-se valores negativos nos espectros, como é ilustrado na figura 2.3 para o método de Blackman-Tukey. Este problema apresenta-se frequentemente na análise de dados experimentais, devido à enorme variabilidade presente, e será ilustrado no capítulo das aplicações (caracterização do EMG). Note-se que numa situação como esta não se poderia empregar a densidade de covariância na modelização AR do espectro e escolha da ordem.

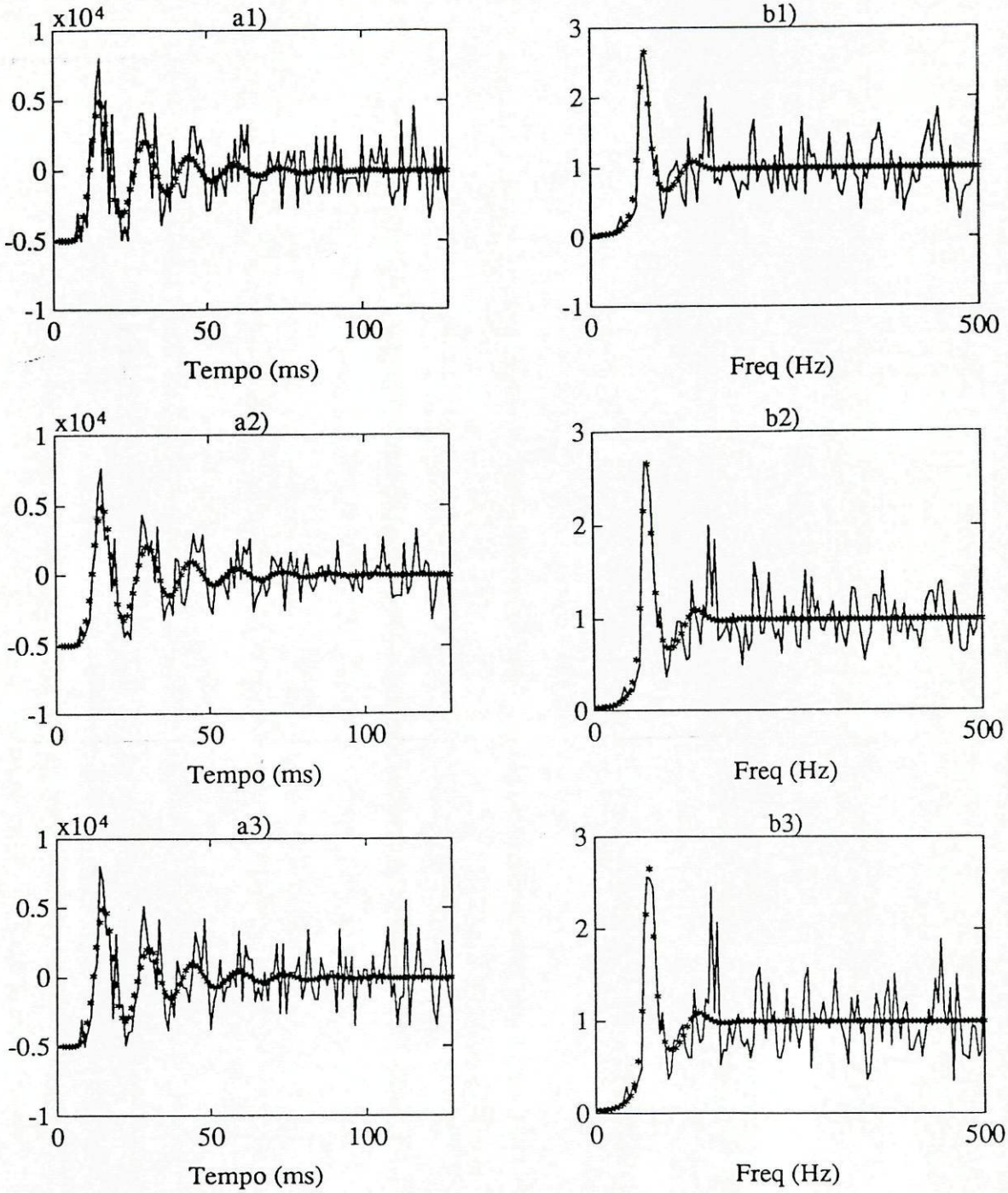


Figura 2.3: a) Função densidade de covariância estimada ($T = 1s, \Delta t = 1ms$) para um processo pontual renewal Gausseano ($\lambda = 70s^{-1}, \sigma/\mu = 0.2$), a partir de **a1)** (2.26) e (2.29) (correcção aditiva do erro de centricidade), **a2)** (2.33), e **a3)** (2.35) (o valor na origem não foi incluido nos diagramas); b) espectros normalizados (não paramétricos) das contagens, $G(\omega)/\hat{\lambda}$, correspondentes, obtidos pelo método de Blackman-Tukey (expressão (2.12), janela de Bartlett, L=256). Sobrepostos nos gráficos (*) representa-se a função densidade de covariância teórica, obtida a partir do valor exacto do espectro, bem como o valor exacto normalizado do espectro. Este foi determinado a partir de (1.7), com $F(j\omega) = e^{-j\mu\omega+(\sigma j\omega)^2/2}$.

2.4 Análise Espectral das contagens: estudo de simulação

Apresentam-se alguns exemplos de aplicação dos estimadores (paramétricos e não paramétricos) do espectro das contagens descritos nas secções anteriores. A implementacão dos estimadores espectrais baseia-se numa estimação prévia da sucessão densidade de covariância utilizando os métodos descritos na secção 2.3.

Neste estudo são usados dados de simulação, considerando dois exemplos que correspondem a espectros das contagens com características bastante diferentes, conforme se ilustra na figura 2.4. Nomeadamente, para $\lambda = 1/\mu = 70s^{-1}$

- *processo pontual renewal Gausseano*, de parâmetros μ e σ^2 $(1/\mu = \lambda, \sigma/\mu = 0.25)$, apresentando um pico pronunciado na intensidade média
- *processo pontual renewal Gama* de parâmetros n e a $(1/\mu = \lambda = a/n,\ n = 3)$, com um espectro das contagens essencialmente plano.

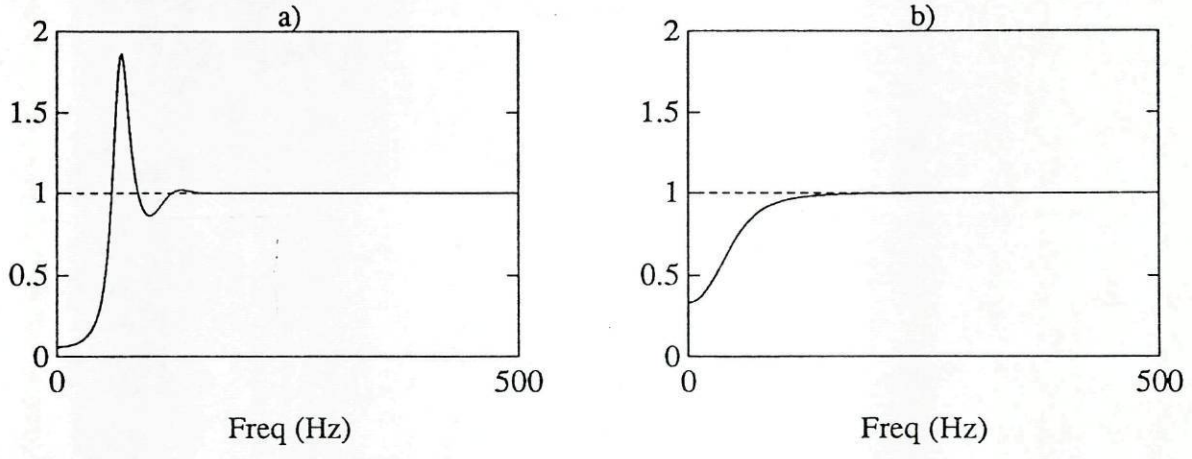


Figura 2.4: Espectros das contagens normalizados, $G(\omega)/\lambda$, **a)** processo renewal Gausseano $(\lambda = 70s^{-1}, \sigma/\mu = 0.25)$; **b)** processo renewal Gama $(\lambda = 70s^{-1}, n = 3)$.

Para o estudo do erro de centricidade e da variância dos diversos estimadores espectrais, utilizam-se sobreposições de 50 realizações: uma representação deste tipo traduz a variabilidade do estimador e o valor médio das realizações permite obter indicação sobre o erro de centricidade. Neste estudo de simulação é também representado o valor exacto do espectro das contagens.

2.4.1 Método de Blackman-Tukey

O método de Blackman-Tukey foi implementado a partir da equação (2.12) considerando a janela de Bartlett.

Conforme foi ilustrado na figura 2.3, para L fixo, a variabilidade (particularmente nas altas frequências) depende do estimador usado para a função densidade de covariância. Verifica-se sistematicamente que a estimação da densidade de covariância a partir de (2.33) é a que conduz a uma menor variabilidade, sendo considerado apenas este caso nos exemplos seguintes.

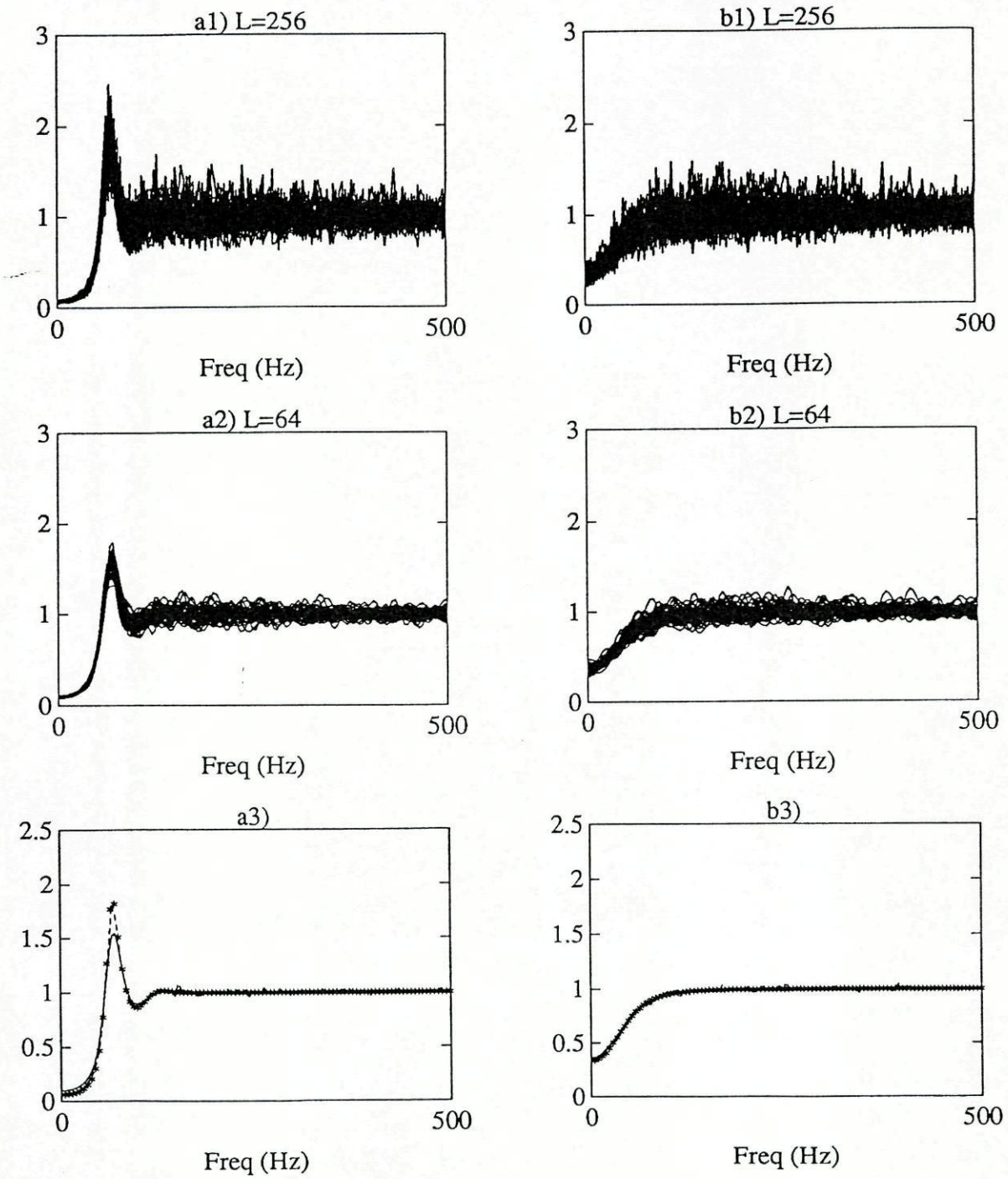


Figura 2.5: Método de Blackman-Tukey: espectros da contagens normalizados, sobreposição de 50 realizações ($T = 5s, \Delta t = 1ms, \lambda = 70s^{-1}$) **a)** processo renewal Gausseano ($\sigma/\mu = 0.25$); **b)** processo renewal Gama ($n = 3$); **1)** L=256,**2)** L=64. Em **3)** estão sobrepostos a média dos espectros em **1)** (L=256,- -) e **2)** (L=64,-), e o valor exacto (*).

Pretende-se ilustrar na figura 2.5 o "compromisso erro de centricidade/variância", o qual é controlado a partir de L, uma vez que a janela está previamente escolhida.

Verifica-se, tal como nos estimadores espectrais correntes, que no caso de não haver picos pronunciados (distribuição Gama), a variância pode ser tornada pequena (tomando um valor baixo para L), sem incorrer num forte erro de centricidade. Este efeito deve-se ao facto de a densidade de covariância decrescer rapidamente para zero.

No caso do processo Gausseano já se terá de escolher um valor suficientemente elevado de L para permitir uma descrição adequada do pico.

2.4.2 Modelização Espectral AR

Na modelização espectral AR, o sistema (2.16) foi resolvido utilizando as equações de Yule-Walker e o algoritmo de Levinson. A ordem óptima, p, foi determinada por minimização dos critérios $AIC(p)$ e $CAT(p)$,

$$AIC(p) = m \log E(p) + 2(p+1) \tag{2.38}$$

$$CAT(p) = m^{-1} \sum_{j=1}^{p} \left(\frac{1}{e(j)} - \frac{1}{e(p)}\right) \tag{2.39}$$

onde $m = int(T/\Delta t)$, $E(p)$ indica o erro mínimo de predição para o modelo de ordem p e $e(j) = (m/(m-j))\, E(j)$, para $p = 1, \ldots, 50$. O espectro AR, $G_a(\omega)$ (2.14), foi calculado pela transformada de Fourier da sucessão $\{1, a(1), \ldots, a(p)\}$.

A sucessão densidade de covariância foi obtida usando os estimadores indicados na secção 2.3.3, equações (2.33) e (2.35). Os métodos baseados em construção de histogramas, equações (2.27) e (2.29) não podem ser sempre aplicados—violação da condição da sucessão estimada ser semi-definida positiva. No entanto verifica-se que, nos casos em que tal é possível, os modelos espectrais AR correspondentes apresentam características idênticas.

Nas figuras 2.6 e 2.7 representa-se o espectro das contagens normalizado para as duas situações anteriormente consideradas. Como seria de esperar, devido às propriedades do modelo espectral $G_a(\omega)$ utilizado, o declive do espectro estimado é zero nas frequências $\omega = 0$ e $\omega = \pi/\Delta t$ [Makhoul 75]. Representam-se ainda os diagramas de polos associados ao modelo espectral; os polos são reais ou complexos conjugados e estão no interior do círculo unitário, uma vez que foram utilizados estimadores semi-definidos positivos para a sucessão densidade de covariância. O espectro não paramétrico, obtido a partir do método de Blackman-Tuckey ($L = 256$), é também representado para comparação.

Conclui-se, a partir destes exemplos e de muitos outros, que a modelização espectral AR conduz a boas descrições do espectro das contagens com uma redução bastante significativa dos dados. Como seria de esperar, devido ao amaciamento inerente aos estimadores espectrais AR, o espectro AR corresponde aproximadamente a uma curva média do espectro não paramétrico. Não foram encontradas diferenças significativas na aplicação dos critérios AIC ou CAT mínimos para a selecção da ordem. Com a ordem seleccionada por aqueles critérios a qualidade da descrição é uniforme, quer nos

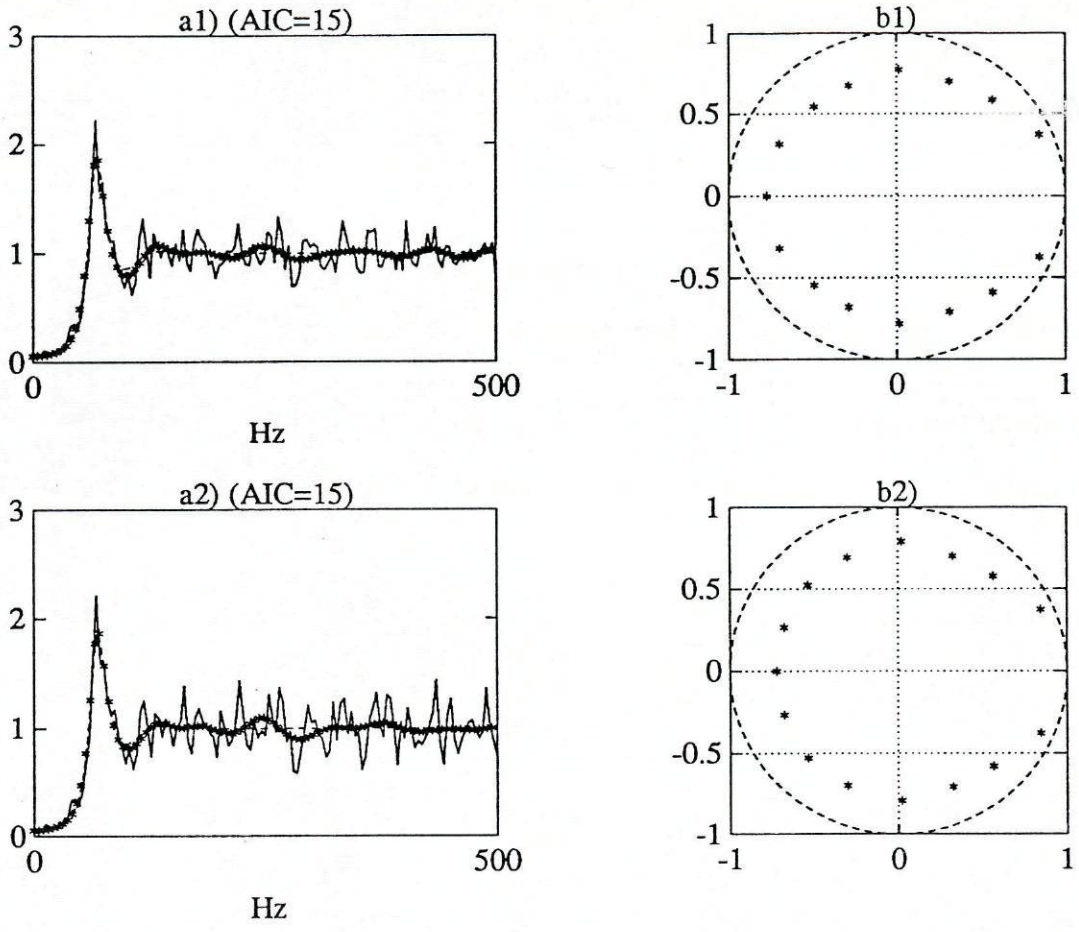


Figura 2.6: a) Espectros das contagens normalizados AR e não paramétrico para um processo renewal Gausseano (mesmos dados da figura 2.5) obtidos com a função densidade de covariância estimada a partir de (2.33) (**a1**)) e (2.35) (**a2**)); b) diagramas de polos correspondentes; valor exacto do espectro (- -).

picos quer nos vales [Lago et al 89, Lago et al 90, Rocha et al 87, Rocha e Lago 88, Rocha e Lago 90].

A comparação entre a modelização espectral AR baseada nas equações de Yule-Walker e no método dos mínimos quadrados, permite verificar que os espectros obtidos são idênticos [Rocha et al 87]. Assim será preferível utilizar as equações de Yule-Walker, de implementação computacional muito mais eficiente [Kay 88, Marple 87].

O espectro das contagens de um processo estocástico pontual difere bastante do espectro de uma série temporal uma vez que não tende para zero com ω crescente. Este comportamento introduz algumas limitações na modelização da região de altas frequências. A experiência parece indicar que se obtêm melhores resultados utizando a sucessão densidade de covariância estimada a partir da expressão (2.33) do que a partir de (2.35), para as janelas anteriormente indicadas. Na parte plana do espectro das contagens, região de altas frequências, a descrição baseada na função densidade de covariância obtida utilizando a janela triangular, é superior. Excluindo a região de altas frequências a descrição com base em qualquer dos métodos é semelhante.

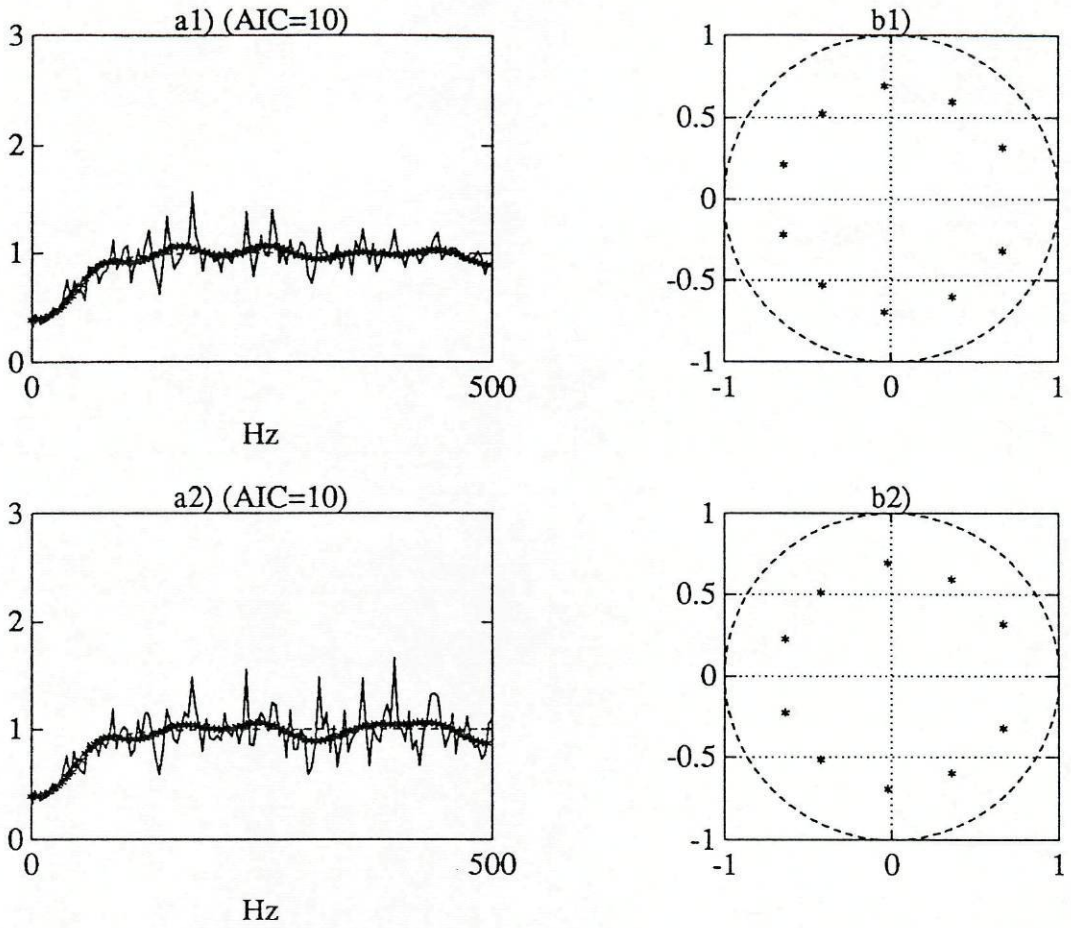


Figura 2.7: **a)** Espectros das contagens normalizados AR e não paramétrico para um processo renewal Gama (mesmos dados da figura 2.5) obtidos com a função densidade de covariância estimada a partir de (2.33) (**a1**)) e (2.35) (**a2**)); **b)** diagramas de polos correspondentes; valor exacto do espectro (- -).

Para testar o desempenho da modelização espectral autoregressiva foram realizados estudos de simulação com a ordem seleccionada pelo critério de AIC (CAT). Mostram-se na figura 2.8 os espectros AR correspondentes a 50 realizações dos processos Gausseano e Gama, bem como os espectros teóricos correspondentes. A variância é aproximadamente a mesma para os dois casos (e menor do que nos métodos não paramétricos para descrições equivalentes), observando-se em qualquer dos casos um pequeno erro de centricidade.

O erro de centricidade observado poderá ser consequência (em parte) de não se estar a considerar a ordem adequada na modelização espectral AR. Existem diversos critérios, correntemente referidos na literatura, para a selecção da ordem [Kay 88, Marple 87]. Nomeadamente, poderá utilizar-se a sequência dos coeficientes de reflexão, k_i: é sabido que, para um processo estocástico AR de ordem p_0, $k_i = 0$, para $i > p_0$ e também que as suas estimativas de máxima verosimilhança têm uma distribuição $\mathrm{N}(0, 1/N_p)$ (N_p, número de pontos da amostra) [Kay 88, e referências aí incluidas]. Assim, conforme se ilustra na figura 2.9, poderá utilizar-se em alternativa a sucessão dos coeficientes de reflexão, $\{k_i\}$, para obter um indicador da ordem do modelo AR.

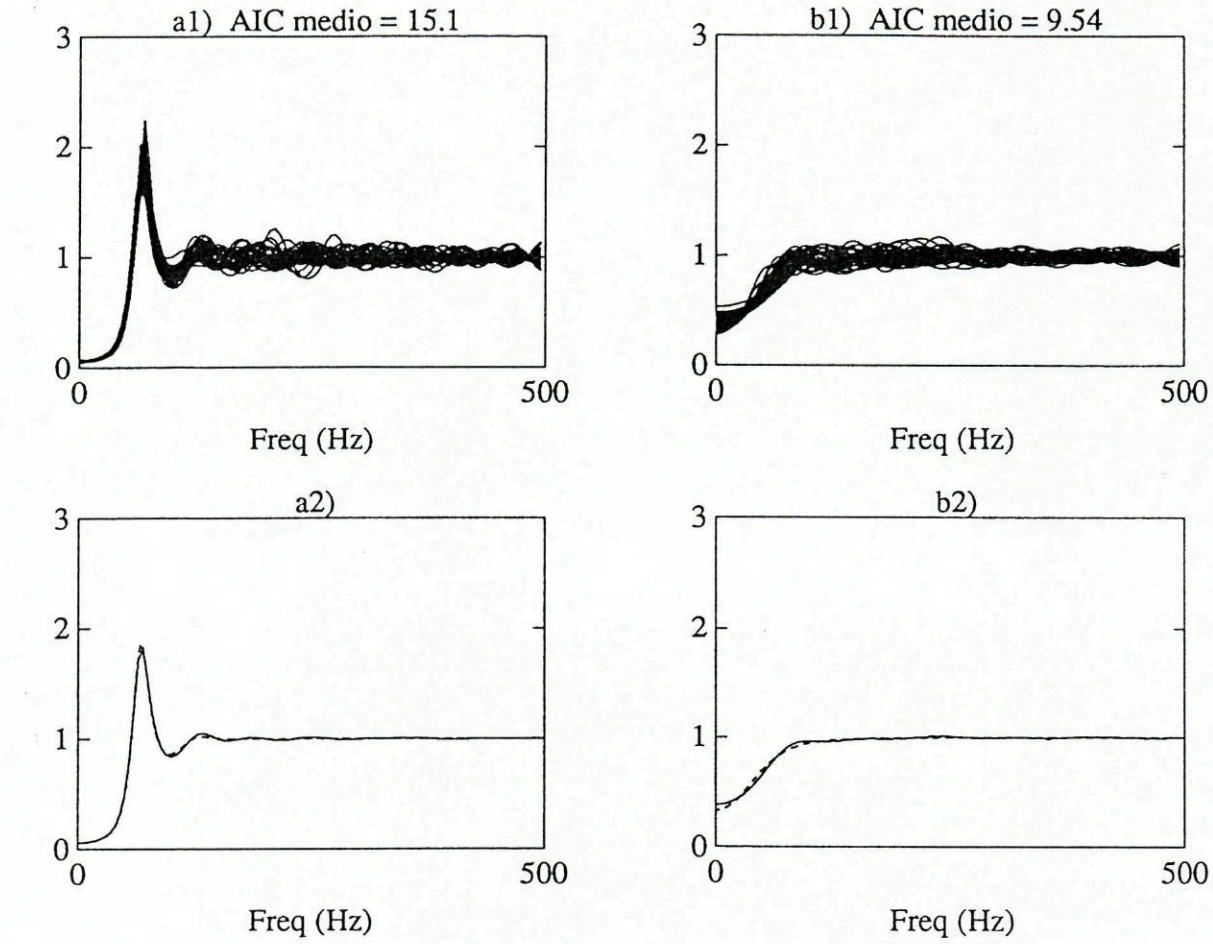


Figura 2.8: Avaliação do desempenho da modelização espectral AR do espectro das contagens utilizando as equações de Yule-Walker e a ordem seleccionada pelo critério de AIC (mesmos dados da figura 2.5); **a)** processo renewal Gausseano, **b)** processo renewal Gama; em qualquer dos casos, **1)** sobreposição de 50 realizações, **2)** espectro paramétrico normalizado médio (-) e valor exacto (- -).

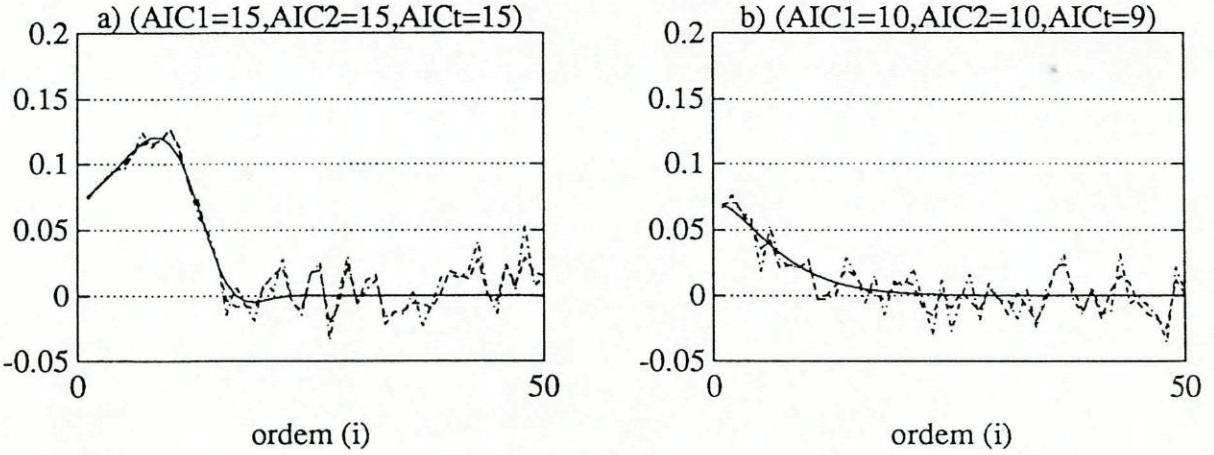


Figura 2.9: Sequência dos coeficientes de reflexão $k_i, i = 1, \ldots, 50$, para a densidade de de covariância estimada por (2.33) (1,- -) e (2.35) (2,- .) e o valor teórico correspondente ao valor exacto do espectro (t,-); em qualquer dos casos, processo renewal **a)** Gausseano e **b)** Gama (mesmos dados da figura 2.5), indicando-se a ordem do modelo AR correspondente seleccionada pela minimização do critério de AIC.

Verifica-se efectivamente que a utilização de uma ordem ligeiramente superior à indicada pelo critério de AIC permite atenuar o erro de centricidade. Na figura 2.10 utilizou-se a ordem indicada pelos coeficientes de reflexão correspondentes à modelização dos espectros teóricos.

No entanto, em contraste com a modelização AR do espectro teórico, a utilização de uma ordem demasiado elevada por vezes acentua as flutuações na região das altas frequências, onde o espectro já deveria ser plano. Estes aspectos serão objecto de análise mais cuidada no estudo do aperfeiçoamento da modelização espectral AR.

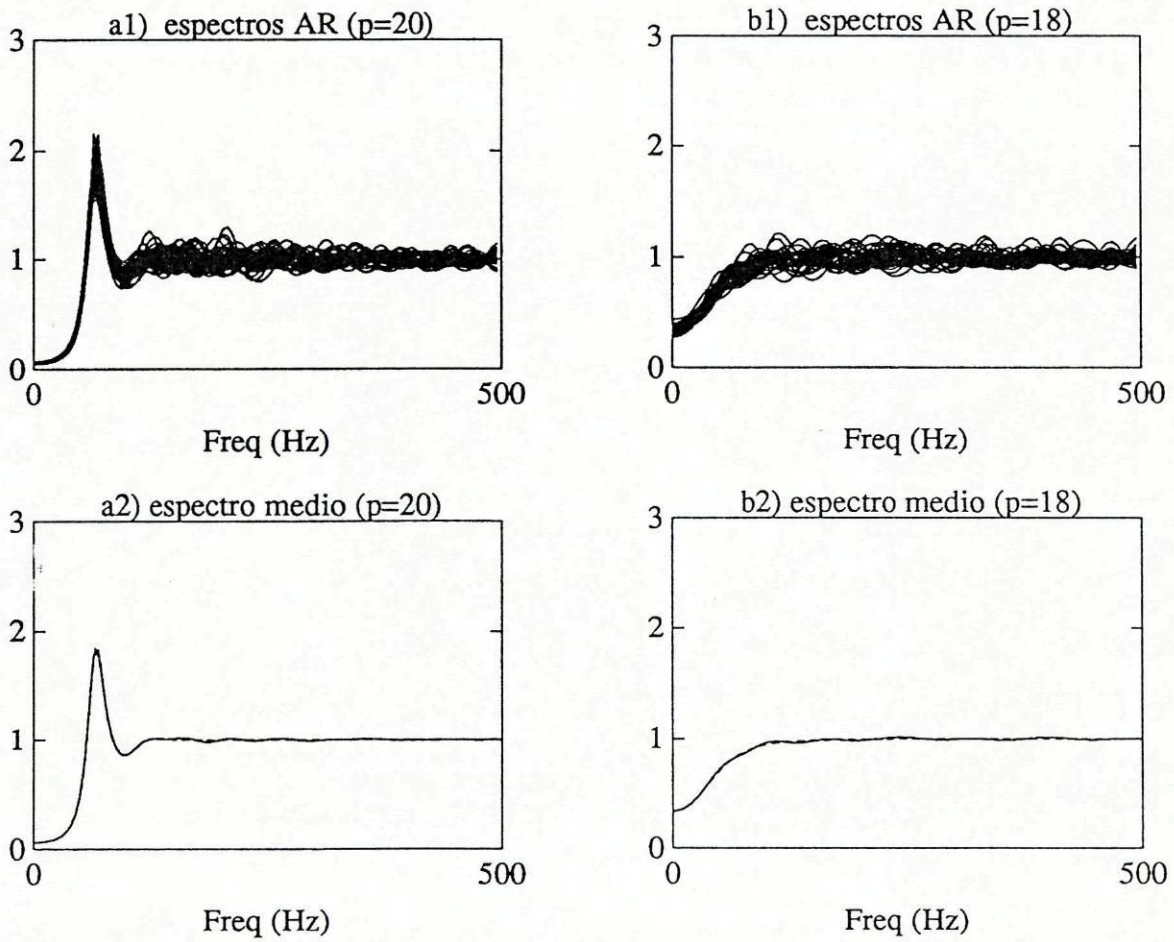


Figura 2.10: Avaliação do desempenho da modelização espectral AR do processo das contagens utilizando as equações de Yule-Walker e a ordem p fixa, obtida à custa dos coeficientes de reflexão teóricos (mesmos dados da figura 2.5); **a)** processo renewal Gausseano, **b)** processo renewal Gama ; em qualquer dos casos, **1)** sobreposição de 50 processos, **2)** espectro paramétrico normalizado médio (-) e valor exacto (- -). A comparação com a figura 2.8 mostra a redução obtida no erro de centricidade.

Note-se deste já que a curva correspondente à sequência dos coeficientes de reflexão apresenta uma forma muito distinta no caso do espectro ser essencialmente plano, como para o processo renewal Gama (figura 2.9). Este facto sugere a exploração desta característica na obtenção de um indicador para a forma do espectro das contagens.

2.5 Comentários finais

A abordagem adoptada nesta secção permite o desenvolvimento de métodos (quer paramétricos, quer não paramétricos), susceptíveis de implementação eficiente, para a análise espectral de processos estocásticos pontuais, em perfeita analogia com os métodos correntes para processos estocásticos em geral.

O estudo de simulação realizado indica que a modelização espectral AR de processos estocásticos pontuais, baseada em estimadores semi-definidos positivos para a sucessão densidade de covariância, apresenta um certo número de vantagens sobre a abordagem não paramétrica.

A análise espectral AR conduz a boas descrições do espectro das contagens de um processo pontual, bem como a uma redução significativa nos dados. Com a ordem seleccionada pelos critérios AIC ou CAT a qualidade da descrição pela modelização AR é uniforme em picos e vales, correspondendo aproximadamente a uma curva média do espectro não paramétrico.

O espectro das contagens de um processo estocástico pontual difere significativamente do espectro de uma série temporal, uma vez que tende para a intensidade média, λ, com ω crescente. Este comportamento introduz alguma dificuldade na modelização na região das altas frequências, como consequência daquela propriedade não ter sido tomada em linha de conta na modelização.

No capítulo seguinte será discutido o aperfeiçoamento da modelização espectral AR de processos estocásticos pontuais, incorporando directamente informação sobre as características do espectro.

Capítulo 3

Análise espectral AR de processos estocásticos pontuais: inclusão directa de informação sobre as características do espectro

As características do espectro das contagens introduzem algumas dificuldades na modelização espectral AR, em particular na região das altas frequências.

Os estudos de simulação realizados (descritos neste capítulo) indicam que não parece possivel aperfeiçoar significativamente a modelização espectral AR, para a medida de erro adoptada. A simplicidade dos modelos AR, sobre outros modelos mais complexos, nomeadamente modelos ARMA, justifica plenamente o estudo de medidas de erro alternativas.

Neste capítulo discute-se o aperfeiçoamento da modelização AR do espectro das contagens de um processo pontual, incorporando na medida de erro informação sobre as características do espectro. O estudo, ilustrado por simulação, incidirá nos seguintes tópicos:

- Modelização AR baseada num número não mínimo de valores da densidade de covariância.
- Aperfeiçoamento da modelização espectral AR: formulação do problema e adaptação às características do espectro das contagens. Introdução de uma medida de erro modificada.
- Desenvolvimento de algoritmos iterativos, para a obtenção dos parâmetros AR.
- Interpretação do aperfeiçoamento da modelização espectral AR com base numa densidade de covariância modificada: algoritmos simplificados correspondentes.
- Alguns comentários à escolha da ordem e obtenção de indicação sobre a gama de frequências onde o espectro deverá ser aproximadamente plano.

3.1 Introdução

A modelização espectral AR de processos pontuais, baseada em estimadores semi-definidos positivos para a sucessão densidade de covariância, conduz a boas descrições do espectro das contagens. No entanto, os resultados até aqui apresentados indicam a necessidade de introduzir alguns aperfeiçoamentos.

As dificuldades na modelização da região das altas frequências do espectro das contagens (a qual deverá ser aproximadamente plana) são uma consequência desta propriedade não ter sido explicitamente considerada na formulação do problema. Com efeito, os modelos AR são baseados num número reduzido de valores da densidade de covariância e (tal como é inerente à modelização espectral autoregressiva) a sucessão densidade de covariância é implicitamente extrapolada para os valores superiores.

Para além disso, outra das limitações na qualidade dos estimadores espectrais decorre do facto de em geral não ser conhecido o valor exacto da densidade de covariância, a qual deverá ser estimada a partir de uma realização truncada do processo.

O ponto de partida natural para melhorar a modelização espectral AR consiste em usar um número de termos na sucessão densidade de covariância superior ao mínimo estritamente necessário, de forma a traduzir melhor as características do processo.

A analogia com a modelização AR de processos estocásticos em geral, sugere a utilização de métodos correntemente referidos na literatura, que consideram um número não mínimo de equações e exploram as propriedades algébricas das matrizes densidade de covariância associadas. Assim, de forma análoga ao que se passa para processos estocásticos em geral, seria de esperar obter um melhor desempenho (pela redução da sensibilidade dos estimadores espectrais AR relativamente ao método utilizado na estimação da sucessão densidade de covariância) ou mesmo indicação sobre a ordem no modelo espectral AR [Cadzow 82, Gardner 88, Kay 88].

Nas secções iniciais deste capítulo apresentam-se alguns resultados de um extenso estudo de simulação, com diversos métodos correntemente utilizados (com algum sucesso, no caso de processos estocásticos em geral) para melhorar a qualidade da modelização espectral. Pretende-se essencialmente ilustrar as dificuldades encontradas na modelização do espectro das contagens.

O aperfeiçoamento da modelização espectral AR, pela inclusão directa de informação relativa às características do espectro das contagens, será discutido na secção 3.2.

3.1.1 Modelização espectral AR: estudo de simulação com base num número não mínimo de equações

Usando apenas os $p+1$ primeiros valores da densidade de covariância, os parâmetros AR, $\{a_n\}$, são a solução <u>única</u> do sistema "bem determinado" (2.18).

O paralelismo entre a análise espectral de processos estocásticos pontuais e a análise espectral de sinais discretos, sugere a obtenção dos parâmetros AR a partir de um

número de equações $t > p$ [Cadzow 82, Lago et al 90]. Neste caso o sistema (2.18) é substituido pelo sistema, que deve ser resolvido no sentido do método dos mínimos quadrados,

$$C_{tp}\, a_p = -c_t, \tag{3.1}$$

onde C_{tp} indica a matriz Toeplitz de dimensão $(t \times p)$, com elementos

$$C_{tp}(i,j) = g_d(|i-j|\Delta t),\ 1 \leq i \leq t,\ 1 \leq j \leq p,$$

e $c_t(i) = g_d(i\Delta t),\ 1 \leq i \leq t$.

Conforme se ilustra na figura 3.1 o espectro obtido a partir de (3.1) apresenta uma qualidade de descrição inferior ao obtido com (2.18), facto que reflecte o aumento na variância devido ao grande número de equações utilizado [Kay 88].

Este comportamento foi observado sistematicamente no estudo de simulação efectuado. De referir que quando t não é muito maior do que p, a determinação do espectro a partir de (3.1) também não melhora a descrição.

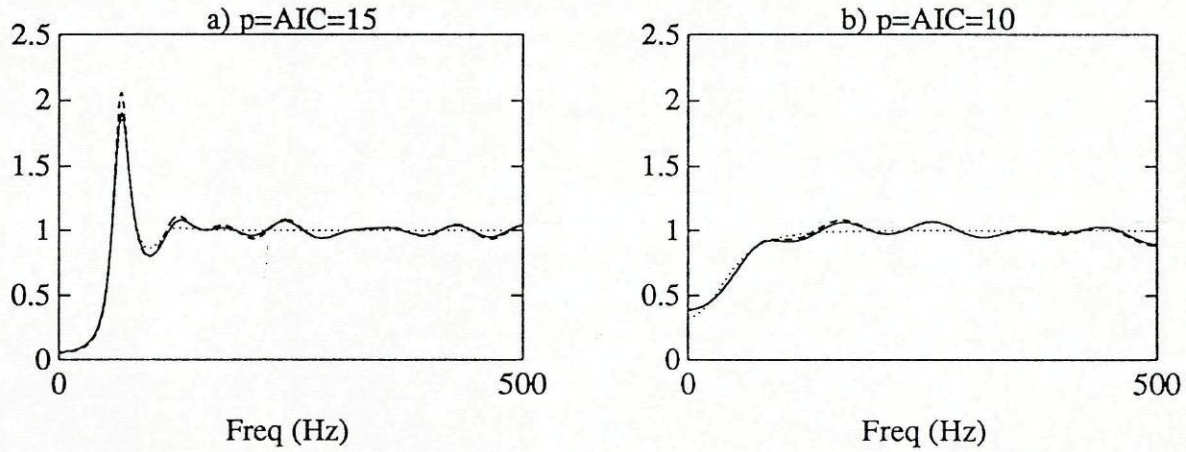


Figura 3.1: Espectro das contagens AR normalizado obtido a partir de (2.18), considerando a sucessão densidade de covariância estimada por (2.33) e o valor de p de acordo com os critérios de AIC e CAT para um processo renewal Gausseano e Gama, respectivamente a) e b). A tracejado representa-se o espectro obtido pelos mínimos quadrados considerando as primeiras $t = 200$ equações de Yule-Walker modificadas dadas por (3.1). Dados da figura 2.5, valor exacto normalizado do espectro (...).

Uma outra possibilidade poderá consistir na selecção de um modelo de ordem estendida $q < t$, muito superior à ordem p indicada pelos critéros AIC/CAT e usar as propriedades algébricas da matriz densidade de covariância associada, C_{tq} (de dimensão $(t \times q)$) [Cadzow 82, Gardner 88, Noble e Daniel 77].

Se a matriz C_{tq} tivesse uma característica "virtual" baixa, a decomposição em valores singulares (**SVD**) de C_{tq} poderia ser utilizada para obter uma indicação para a ordem a utilizar na modelização AR. Neste caso, um determinado número d de valores singulares dominaria os valores restantes $q - d$ valores e os parâmetros AR para o modelo melhorado poderiam ser determinados a partir de

$$C_{tq}^{d}\, a_q = -c_t, \tag{3.2}$$

onde C_{tq}^{d} indica a matriz de dimensão $(t \times q)$ que melhor aproxima a matriz C_{tq} no sentido da norma de Frobenius ($\| \ \|_F$) [Noble e Daniel 77]. Apesar do modelo ter ordem máxima q, deveria comportar-se como um modelo de ordem mais baixa, $d < q$.

Este método foi utilizado com algum sucesso para processos estocásticos em geral [Cadzow 82].

Contudo para processos estocásticos pontuais $g(0)$ é dominante podendo, a partir da equação (2.8), decompor-se a matriz C_{tq} na forma

$$C_{tq} = \lambda(\Delta t)^{-1} I_{tq} + A_{tq}, \tag{3.3}$$

onde I_{tq} indica a matriz "identidade" rectangular, de dimensão $(t \times q)$, com elementos

$$I_{tq}(i,j) = \delta_{ij}, \ 1 \leq i \leq t, \ 1 \leq j \leq q,$$

e A_{tq} é obtida a partir de C_{tq} subtraindo $\lambda(\Delta t)^{-1}$ à diagonal principal. Conclui-se assim que a matriz densidade de covariância C_{tq} tem característica virtual máxima, q.

No entanto, poderão obter-se matrizes associadas para processos pontuais com característica virtual baixa.

Na figura 3.2 representa-se o comportamento das sequências de valores singulares associados com as equações de Yule-Walker deslocadas, isto é as equações (2.16) para $n = i, i+1, \ldots, i+t-1$ onde $i > 1$). Verifica-se que a característica virtual da matriz densidade de covariância C_{tq} associada às equações deslocadas diminui quando o deslocamento i aumenta. Este comportamento é notório no índice normalizado $\nu(k) = \|C_{tq}^{k}\|_F / \|C_{tq}\|_F$, o qual mede a qualidade da aproximação de C_{tq} por C_{tq}^{k} [Cadzow 82].

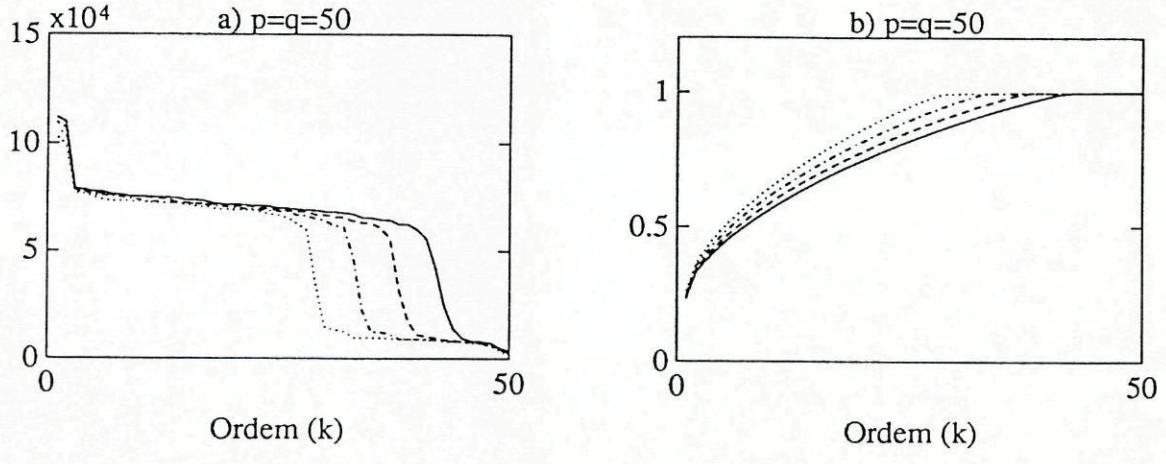


Figura 3.2: a) Sequências dos valores singulares associados com as equações de Yule-Walker modificadas para $p = q = 50$, $t = 200$ e deslocamentos $i = 5$ (-), 10 (- -), 15 (-.), 20 (..) e b) coeficientes de aproximação normalizados, $\nu(k)$, correspondentes para um processo renewal Gausseano (mesmos dados da figura 3.1).

A razão normalizada atinge o valor máximo de um à medida que k se aproxima de q. Para matrizes de característica virtual baixa, $\nu(k)$ é aproximadamente um para valores de k significativamente menores do que q.

Assim, a sucessão de valores singulares associada à matriz densidade de covariância pode ter as propriedades pretendidas, para uma escolha de deslocamento suficientemente grande. Contudo, conforme é indicado pela experiência, as primeiras equações são essenciais para a determinação dos parâmetros AR. Os resultados obtidos a partir da SVD são fracos e não parecem merecer grande confiança.

Na literatura é também correntemente referida a utilização da decomposição da matriz de covariância na identificação de espaços ortogonais associados às diversas componentes espectrais. Nomeadamente, é referida a identificação de sinusoides em ruído, pela separação num espaço associado ao sinal e outro ao ruído [Kay 88, Marple 87, e referências aí incluidas].

Na figura 3.3 mostra-se a diferença no comportamento das sequências de valores singulares das matrizes, A_{tq} e C_{tq}, e correspondentes coeficientes de aproximação normalizados. Este exemplo ilustra o facto da matriz A_{tq} apresentar uma característica virtual baixa.

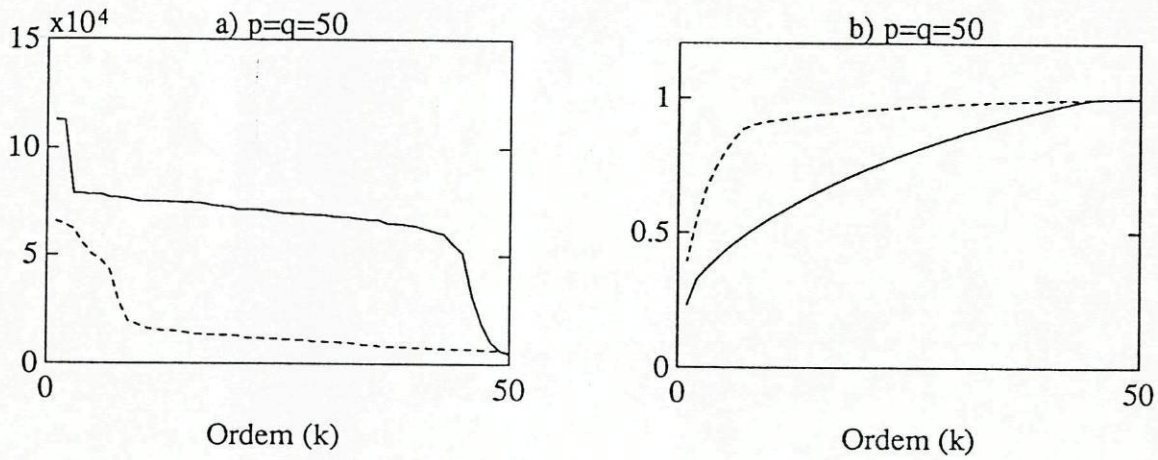


Figura 3.3: **a)** Sequências dos valores singulares associados às matrizes $\mathbf{A}_{tq}$ (- -) e C_{tq} (-) e **b)** coeficientes de aproximação normalizados correspondentes para $p = q = 50$, $t = 200$; processo renewal Gausseano, mesmos dados das figuras 3.1e 3.2.

Assim, no caso da modelização espectral AR de processos estocásticos pontuais, é natural pensar em utilizar a decomposição da matriz densidade de covariância (3.3) na melhoria da fidelidade espectral.

No entanto, apesar da matriz A_{tq} poder dar indicação sobre a ordem do modelo, ela não é adequada para a resolução do problema. Com efeito, a sucessão associada não é garantidamente semidefinida positiva, não correspondendo portanto a uma densidade de covariância.

Pode concluir-se, a partir dos exemplos apresentados e muitos outros, que existem grandes dificuldades em melhorar a modelização espectral AR de processos pontuais, para os métodos correntemente utilizados no caso de processos estocásticos em geral.

Parece assim ser necessário utilizar modelos alternativos mais complexos, que permitam descrever melhor as características do espectro das contagens, nomeadamente modelos ARMA(p,p). A modelização espectral ARMA permite de facto obter boas descrições na zona das altas frequências. Contudo a escolha da ordem do modelo espectral é bastante crítica [Lago et al 90], o que suscita algumas interrogações sobre a aplicabilidade geral do método.

A simplicidade dos modelos AR, justifica plenamente que se analise o problema com maior detalhe. Nomeadamente, as dificuldades na descrição da zona das altas frequências poderão advir dos métodos de estimação da densidade de covariância, em particular da existência de lobos laterais nas janelas espectrais associadas aos estimadores utilizados.

Assim, será fundamental começar por verificar se o modelo AR é adequado para a modelização do espectro das contagens, considerando o caso de conhecimento exacto da sucessão densidade de covariância.

3.1.2 Modelização AR do espectro das contagens a partir da sucessão densidade de covariância teórica

Nesta secção discute-se o desempenho da modelização espectral AR supondo conhecido o valor teórico da sucessão densidade de covariância.

Em qualquer dos casos considerados, processos pontuais renewal Gama e Gausseano, não é fácil obter uma expressão analítica para a sucessão densidade de covariância. Assim, os valores teóricos foram calculados directamente a partir do valor exacto dos espectros, considerando uma resolução temporal suficiente e utilizando a relação inversa de (2.9),

$$g_d(n\Delta t) = \frac{1}{2\pi} \int_{-\pi/\Delta t}^{+\pi/\Delta t} G_d(\omega)\, e^{j\omega n\Delta t}\, d\omega, \quad n = 0, \pm 1, \ldots \tag{3.4}$$

A implementação dos cálculos foi feita a partir da transformada inversa de Fourier.

Verifica-se, conforme se ilustra na figura 3.4, que no caso de se conhecer a sucessão densidade de covariância, o modelo espectral AR permite descrever adequadamente o espectro das contagens. É contudo essencial considerar um modelo de ordem suficientemente alta, a qual deverá ser superior à ordem indicada pelo critério de AIC (na figura a ordem AIC corresponde, tal como no capítulo anterior, a registos com duração de 5s e a uma resolução temporal $\Delta t = 1ms$).

Para uma ordem suficientemente elevada, o modelo espectral AR não se distingue do valor exacto e a parte plana do espectro corresponde, em qualquer dos casos, a polos igualmente significativos.

Um aumento da ordem reflecte-se sempre numa melhoria na fidelidade da descrição espectral, com uma diminuição da amplitude das flutuações na região das altas frequências. Verifica-se no entanto que é suficiente utilizar a ordem indicada pelo comportamento dos coeficientes de reflexão teóricos (ver figura 2.9).

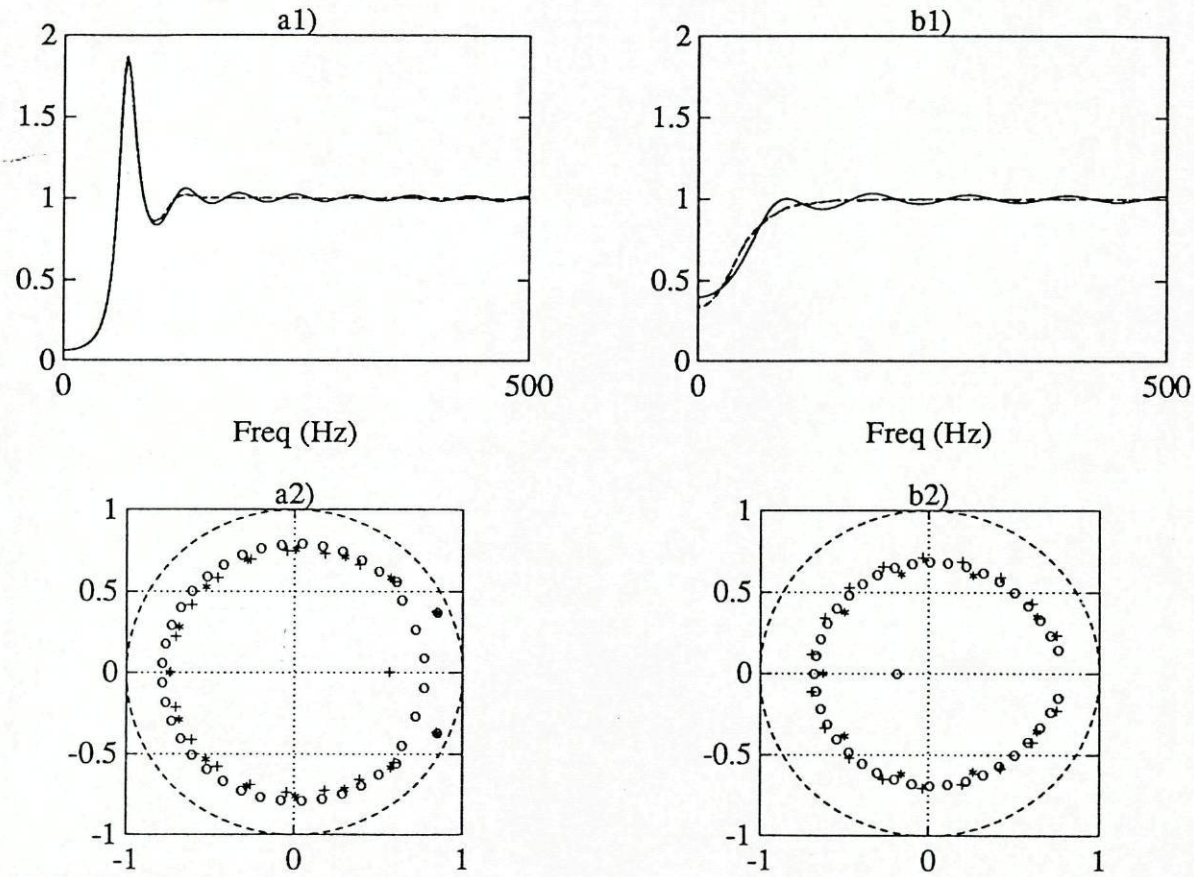


Figura 3.4: A linha contínua representa-se o espectro das contagens AR normalizado para um processo renewal Gausseano (**a1**) e um processo renewal Gama (**b1**), obtidos a partir do valor exacto da densidade de covariância, para $p = p_0 = AIC$ (15 e 9, respectivamente). Para valores da ordem mais elevados (p_1 e p_2, respectivamente, 20 e 40 em (**a1**) e 18 e 40 em (**b1**)) o espectro AR não se distingue do teórico (a linha tracejada). **a2)** e **b2)** diagramas de polos correspondentes a p_0, p_1 e p_2, respectivamente (*), (+) e (o). Mesmos dados da figura 2.4, considerando uma resolução temporal $\Delta t = 1ms$.

Em contraste, no caso de se utilizarem sucessões densidade de covariância estimadas, será fundamental não utilizar uma ordem demasiado elevada.

A figura 3.5 ilustra o efeito decorrente do aumento da ordem do modelo. Observa-se neste caso um detalhe crescente, reflectindo uma maior fidelidade na descrição de todas as flutuações no espectro não paramétrico correspondente. A parte plana do espectro deixa também de corresponder a uma distribuição regular dos polos, os quais já não se apresentam igualmente significativos.

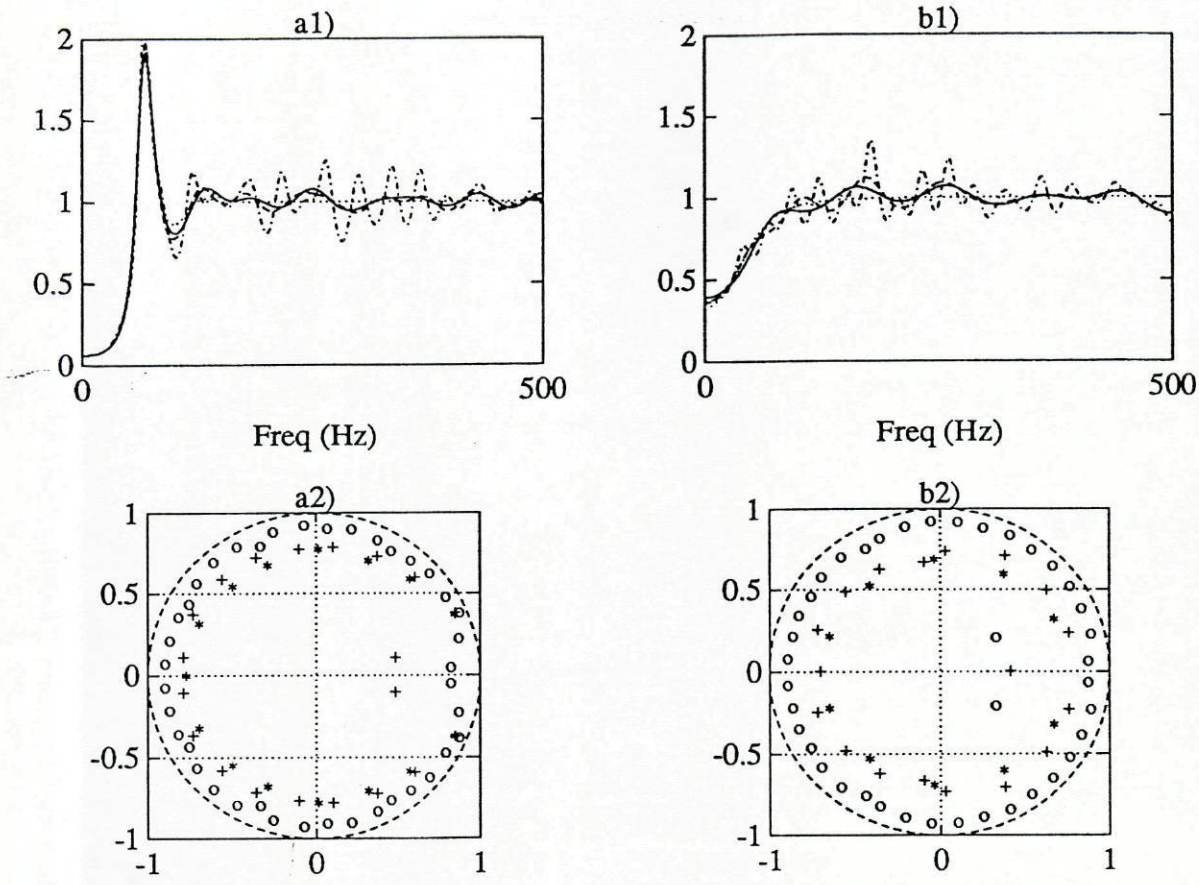


Figura 3.5: a1) e b1) Espectros das contagens AR normalizados para, respectivamente, processo renewal Gausseano (p = 15(AIC,-), 20 (- -) e 40 (- .)) e processo renewal Gama (p = 10(AIC,-), 18 (- -) e 40 (- .)), obtidos com a densidade de covariância estimada a partir de (2.33) e diagramas de polos correspondentes, respectivamente (*), (+) e (o), em (**a2**) e (**b2**); valor exacto do espectro (...). Mesmos dados da figura 2.5 (T=5s, resolução temporal $\Delta t = 1ms$).

3.1.3 Comentários Finais

Os estudos de simulação efectuados permitem concluir que, apesar das dificuldades encontradas na descrição da região das altas frequências, o modelo AR é adequado para a modelização do espectro das contagens de um processo pontual. Com efeito, com base no valor exacto da sucessão densidade de covariância obtêm-se boas descrições espectrais, desde que se considere uma ordem suficientemente elevada.

Sendo assim, torna-se necessário verificar se as dificuldades encontradas na modelização da parte plana do espectro poderão ser justificados pela variabilidade nos valores estimados da sucessão densidade de covariância ou se, nomeadamente, reflectem a presença de lobos laterais nas janelas espectrais associadas aos estimadores semi-definidos positivos desenvolvidos (secção 2.3.3).

Perturbando ligeiramente a sucessão densidade de covariância teórica pode obter-se indicação do efeito da variabilidade dos valores estimados para a sucessão densidade de covariância na modelização espectral AR.

Conforme se ilustra na figura 3.6 (para o caso de um processo pontual renewal Gausseano) a sucessão perturbada resultante tem um aspecto semelhante ao da sucessão densidade de covariância estimada no que respeita ao valor médio e variância, conduzindo a um modelo espectral AR que apresenta (tal como o obtido com as sucessões densidade de covariância estimadas) flutuações acentuadas na região na região das altas frequências.

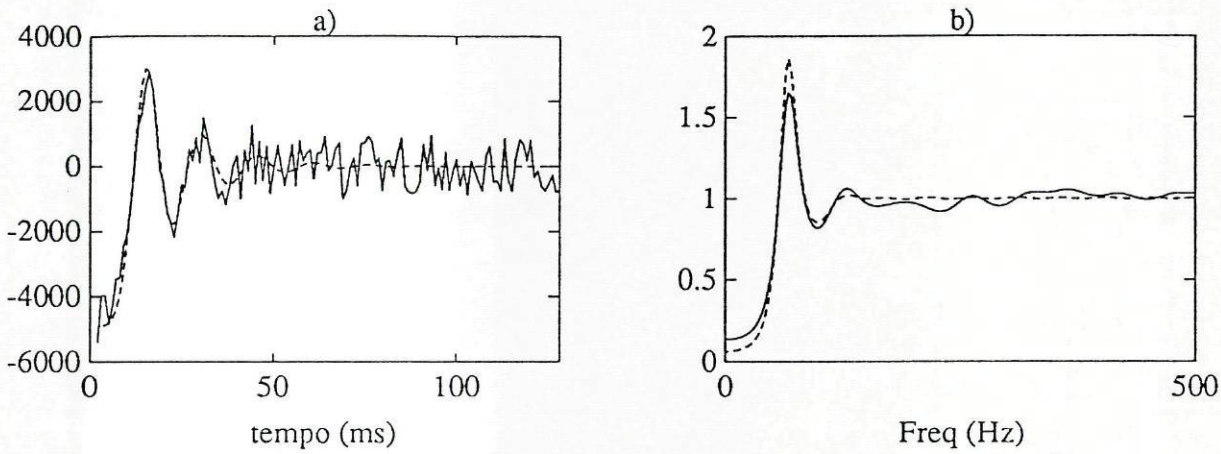


Figura 3.6: Ilustração do efeito da perturbação do valor teórico da densidade de covariância **(a)** no espectro das contagens AR normalizado **(b)**. Mesmos dados da figura 2.5 para $p = 20$; valores teóricos correspondentes a tracejado.

Parece poder concluir-se que as dificuldades na descrição nas altas frequências não decorrem do uso dos estimadores semi-definidos positivos para a sucessão densidade de covariância apresentados na secção 2.3.3. Com efeito, as dificuldades encontradas na descrição da região das altas frequências poderão apenas traduzir a variabilidade nos valores estimados para a sucessão densidade de covariância.

Do extenso trabalho de experimentação numérica efectuado (ilustrado nas secções 3.1.1 a 3.1.3) conclui-se que não parece fácil melhorar a qualidade dos estimadores espectrais AR, para a medida de erro utilizada, expressão (2.15).

Uma possibilidade em aberto, poderá consistir em utilizar medidas de erro alternativas, eventualmente mais adequadas para a modelização do espectro das contagens de um processo estocástico pontual.

3.2 Aperfeiçoamento da modelização espectral AR

3.2.1 Introdução

A inclusão da condição de que o modelo espectral AR para o espectro das contagens deve ser aproximadamente constante na região das altas frequências, é facilmente incorporada na medida de erro a partir da quantificação dos desvios para λ.

O critério de erro de Itakura-Saito corresponde a uma medida de distorsão que tem sido utilizada com frequência em diversas aplicações [Itakura e Saito 68]. Em particular, este critério foi usado com sucesso no problema da modelização espectral AR de processos não pontuais, e apresenta vantagens sobre os métodos baseados no preditor linear no caso dos espectros a modelar serem discretos [El-Jaroudi e Makhoul 91].

A avaliação da distorsão entre duas funções de densidade espectral, $\Phi_1(\omega)$ e $\Phi_2(\omega)$, a partir do critério de erro de Itakura-Saito, é feita considerando a função

$$d(\omega) = \frac{\Phi_1(\omega)}{\Phi_2(\omega)} - \log \frac{\Phi_1(\omega)}{\Phi_2(\omega)} - 1. \tag{3.5}$$

Note-se que $d(\omega) = 0$ se e só se $\Phi_1(\omega) = \Phi_2(\omega)$. A função $d(\omega)$ é não negativa e apresenta mínimos locais nos pontos que verificam a condição $\Phi_1(\omega) = \Phi_2(\omega)$.

Assim, *na região das altas frequências* (onde o espectro deverá ser suficientemente plano), a avaliação da distorsão do modelo espectral AR, $G_a(\omega)$, relativamente a λ, poderá ser feita a partir do critério de erro de Itakura-Saito. Isto é, considerando a função $d(\omega)$ dada por (3.5), para $\Phi_1(\omega) = \lambda$ e $\Phi_2(\omega) = G_a(\omega)$.

3.2.2 Medida de erro "modificada"

Para valores de Δt suficientemente pequenos, o espectro da sucessão $\{g_d(n\Delta t)\}$, $G_d(\omega)$ (definido por 2.9), é uma boa aproximação para $G(\omega)$, para $|\omega| \leq \pi/\Delta t$, conforme foi mostrado na secção 2.2.2.

O problema da modelização de $G_d(\omega)$, incorporando a informação de que o espectro deverá ser plano na região das altas frequências, poderá ser formulado de forma inteiramente análoga à apresentada no capítulo anterior. Suposto escolhido o número de pólos p, bastará substituir, na formulação 1, a medida de erro $E(p)$, dada por (2.15), por uma medida de erro que traduza aquela característica fundamental do espectro das contagens.

Assim, para uma escolha adequada de Δt, o problema da *modelização* do espectro $G_d(w)$, incorporando a informação de que o espectro deverá ser *suficientemente plano na região das altas frequências* pode ser formulado da seguinte forma:

Formulação 2 *Dado $G_d(\omega)$ e um determinado número de polos p, pretende-se aproximar $G_d(\omega)$ por um espectro autoregressivo $G_a(\omega)$, dado por*

$$G_a(\omega) = \frac{V}{|1 + \sum_{n=1}^{p} a_n e^{-jnw\Delta t}|^2} \tag{3.6}$$

usando a medida de erro "modificada", $E_{ma}(p)$,

$$E_{ma}(p) = \frac{V\Delta t}{2\pi} \int_{-\pi/\Delta t}^{+\pi/\Delta t} (\frac{G_d(\omega)}{G_a(\omega)} + \beta [\frac{\lambda}{G_a(\omega)} - \log \frac{\lambda}{G_a(\omega)} - 1] W(\omega)) \, d\omega, \tag{3.7}$$

onde β designa um peso constante positivo e $W(\omega)$ indica uma janela em frequência, simétrica e não negativa, a qual deverá anular-se fora da zona, $|\omega| \geq \omega_{lim}$, onde o espectro já é suficientemente próximo de λ.

Nesta formulação será de salientar que:

- O funcional de custo adoptado, E_{ma}, dado por (3.7), não é mais do que o funcional inicialmente escolhido, ao qual se adicionou uma parcela que traduz a medida em que o espectro estimado deverá ser suficientemente próximo de λ.
- $W(\omega)$ corresponde ao módulo da resposta em frequência de um filtro passa-alto com frequência de corte $\omega_{lim} \leq \pi/\Delta t$. A escolha de β e da janela $W(\omega)$ permite penalizar adequadamente as diferenças para λ e indicar a zona onde o espectro estimado deverá ser aproximadamente plano.
- Para a aplicação do funcional de custo E_{ma} será necessário ter indicação da zona em que o espectro já é aproximadamente constante, nomeadamente a partir de um espectro não paramétrico estimado previamente. Este ponto será analizado numa secção posterior.
- Finalmente, refira-se que uma modificação da medida de erro, considerando em (3.7) apenas a inclusão da razão $\lambda/G_a(\omega)$, em alternativa a (3.5), não seria adequada. Para além desta razão não dar uma contribuição nula para o erro (mesmo no caso do modelo espectral coincidir com λ), a solução corresponderia à modelização de $G_d(\omega) + \lambda\beta W(\omega)$ para $|\omega| < \pi/\Delta t$, na formulação 1. Não seria portanto possível obter um espectro das contagens plano e aproximadamente igual a λ na região das altas frequências.

Verifica-se facilmente que a minimização de E_{ma} (dado por (3.7)), relativamente aos parâmetros AR, $\{a_n\}_{n=1,\cdots,p}$, conduz ao sistema de equações

$$g_d(n\Delta t) + \sum_{k=1}^{p} a_k g_d(|n-k|\Delta t) = g_{ma}(n\Delta t) + \sum_{k=1}^{p} a_k g_{ma}(|n-k|\Delta t), \; n = 1, \ldots, p \tag{3.8}$$

onde $g_d(n\Delta t)$, dada por (3.4), designa a sucessão densidade de covariância correspondente ao espectro $G_d(\omega)$ (expressão (2.9)) e

$$g_{ma}(n\Delta t) = \beta\,(\varphi_a(n\Delta t) - \lambda\, w(n\Delta t)),\ n = 0, \pm 1, \ldots \tag{3.9}$$

Nesta expressão, utilizam-se as notações $w(n\Delta t)$ e $\varphi_a(n\Delta t)$ para designar as sucessões cujo espectro, definido por (2.9), corresponde a $W(\omega)$ e $G_a(\omega)W(\omega)$, respectivamente. Isto é, de forma análoga a (3.4),

$$w(n\Delta t) = \frac{1}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} W(\omega)\, e^{j\omega n\Delta t}\, d\omega \tag{3.10}$$

e

$$\varphi_a(n\Delta t) = \frac{1}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} (G_a(\omega)W(\omega))\, e^{j\omega n\Delta t}\, d\omega. \tag{3.11}$$

O valor de V em (3.6) poderia ser calculado a partir da conservação de energia entre o espectro a modelar, $G_d(\omega)$, e o modelo espectral $G_a(\omega)$ determinado a partir de (3.8). Alternativamente, será utilizada a condição

$$\frac{\Delta t}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} \frac{G_d(\omega)}{G_a(\omega)}\, d\omega = 1. \tag{3.12}$$

Esta condição permite traduzir a propriedade da solução corresponder a uma boa descrição do espectro não paramétrico, para uma ordem suficientemente elevada, e é verificada para a medida de erro inicialmente considerada (formulação 1), uma vez que o erro mínimo é, nesse caso, igual a V.

Designando por $\boldsymbol{a}$ o vector, de dimensão $p+1$, dos parâmetros AR incluindo uma primeira componente $a_0 = 1$, isto é $\boldsymbol{a} = (a_0, a_1, \cdots, a_p)'$, verifica-se facilmente, a partir da relação (3.4), que (3.12) pode escrever-se na forma equivalente

$$\Delta t\ \boldsymbol{a}'\, \boldsymbol{M}_d\, \boldsymbol{a} = V. \tag{3.13}$$

Nesta expressão $\boldsymbol{M}_d$ designa a matriz quadrada simétrica toeplitz $\boldsymbol{C}_{PP}$, de dimensão $P = p+1$, definida em (2.18).

Note-se que valor de V calculado a partir da expressão (3.13) coincide exactamente com o determinado a partir de (2.17), na formulação 1 para a medida de erro (2.15), e também corresponde ao erro médio quadrático mínimo para a medida de erro modificada E_{ma}.

A principal dificuldade na utilização da medida de erro modificada consiste no facto do sistema de equações (3.8) ser *não linear*, uma vez que $\varphi_a(n\Delta t)$ depende dos próprios parâmetros AR. Para a sua resolução deverão ser utilizados métodos iterativos nos parâmetros AR, o que será objecto de estudo neste capítulo.

Assim, o vector dos parâmetros AR passará a designar-se por $\boldsymbol{b} = (b_0, b_1, \cdots, b_p)'$, onde b_0 poderá ser diferente de 1. Para este modelo espectral ser equivalente ao da primeira formulação deverá no entanto fazer-se a minimização de E_{mb} relativamente aos parâmetros AR, $\{b_n\}_{n=0,\cdots,p}$, restringida por uma condição que traduza a relação (3.12).

3.2.3 Modelo espectral AR equivalente

Para uma escolha adequada de Δt, o problema da modelização de $G_d(\omega)$ poderá ser formulado a partir do modelo espectral equivalente $G_b(\omega)$:

Formulação 3 *Dado $G_d(\omega)$ e um determinado número de polos p, pretende-se aproximar $G_d(\omega)$ por um espectro autoregressivo $G_b(\omega)$*

$$G_b(\omega) = \frac{1}{|\sum_{n=0}^{p} b_n e^{-jnw\Delta t}|^2} \tag{3.14}$$

usando a medida de erro "modificada" $E_{mb}(p)$,

$$E_{mb}(p) = \frac{\Delta t}{2\pi} \int_{-\pi/\Delta t}^{+\pi/\Delta t} (\frac{G_d(\omega)}{G_b(\omega)} + \beta [\frac{\lambda}{G_b(\omega)} - \log \frac{\lambda}{G_b(\omega)} - 1] W(\omega)) \, d\omega, \tag{3.15}$$

e impondo a restrição

$$\Delta t \ \boldsymbol{b}' \boldsymbol{M}_d \boldsymbol{b} = 1. \tag{3.16}$$

Nesta formulação, onde β e $W(\omega)$ têm o significado atrás referido, a restrição imposta na minimização corresponde à condição (3.12) e equivale à expressão (3.13) para a determinação do valor de V, na formulação 2.

O lagrangeano associado com o problema de minimização restringida tem uma forma muito simples [Luenberger 73]

$$\mathcal{L}(\boldsymbol{b}, \theta) = E_{mb}(\boldsymbol{b}, p) + \theta h(\boldsymbol{b}) \tag{3.17}$$

onde

$$h(\boldsymbol{b}) = \Delta t \ \boldsymbol{b}' \boldsymbol{M}_d \boldsymbol{b} - 1. \tag{3.18}$$

A solução é dada por,

$$\begin{aligned} \nabla_{\boldsymbol{b}} \mathcal{L}(\boldsymbol{b}, \theta) &= \nabla E_{mb}(\boldsymbol{b}, p) = 0 \\ \nabla_{\theta} \mathcal{L}(\boldsymbol{b}, \theta) &= h(\boldsymbol{b}) = 0 \end{aligned} \tag{3.19}$$

isto é,

$$\begin{aligned} \nabla E_{mb} = 2\Delta t \, (\boldsymbol{M}_d - \beta \, \boldsymbol{M}_b) \, \boldsymbol{b} &= 0 \\ \Delta t \ \boldsymbol{b}' \boldsymbol{M}_d \boldsymbol{b} &= 1. \end{aligned} \tag{3.20}$$

Nesta expressão $\boldsymbol{M}_b$ designa também uma matriz quadrada, simétrica Toeplitz de dimensão $P = p + 1$ com elementos

$$M_b(i,j) = g_{mb}(|i-j|\Delta t),\ 1 \leq i,\ j\ \leq P. \tag{3.21}$$

De forma análoga à formulação 2, $g_{mb}(n\Delta t)$ é definida a partir da sucessão $\varphi_b(n\Delta t)$, correspondente ao produto da janela espectral $W(\omega)$ pelo modelo $G_b(\omega)$. Isto é

$$g_{mb}(n\Delta t) = \beta\,(\varphi_b(n\Delta t) - \lambda\, w(n\Delta t))),\ n = 0, \pm 1, \ldots \tag{3.22}$$

onde $w(n\Delta t)$ é dada pela expressão (3.10) e

$$\varphi_b(n\Delta t) = \frac{1}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} (G_b(\omega)W(\omega))\, e^{j\omega n\Delta t}\, d\omega, \tag{3.23}$$

expressão correspondente a (3.11) para o modelo espectral $G_b(\omega)$.

Conclui-se assim que a determinação dos parâmetros AR pela minimização de (3.15), relativamente aos parâmetros AR, $\{b_n\}_{n=0,\cdots,p}$, conduz ao sistema de equações normais

$$\sum_{k=0}^{p} b_k g_d(|n-k|\Delta t) = \beta\ (\sum_{k=0}^{p} b_k g_{mb}(|n-k|\Delta t)),\ n = 0, \ldots, p \tag{3.24}$$

o qual se pode escrever na forma matricial

$$\boldsymbol{M}_d\, \boldsymbol{b} = \beta\ \boldsymbol{M}_b \boldsymbol{b}. \tag{3.25}$$

Será finalmente de salientar que:

- De forma análoga a (3.8) o sistema de equações normais (3.24) é não linear, tornando-se necessário o desenvolvimento de algoritmos iterativos para a determinação da solução.

  Contudo, esta tarefa é bastante simplificada pela expressão analítica do lagrangeano. A restrição, apesar de ser não linear, corresponde a uma forma quadrática bastante simples não havendo dificuldade em implementar métodos de minimização correntemente referidos na literatura [Luenberger 73].

- Para a formulação 3, o erro mínimo é igual a <u>um</u>.

  Com efeito, na solução do problema o termo correspondente à descrição da parte plana do espectro anula-se e a primeira parcela tem o valor <u>um</u>, devido à restrição imposta.

3.3 Determinação do modelo espectral AR por iteração sobre os parâmetros

Nesta secção descrevem-se algoritmos iterativos para a resolução do problema de minimização formulado na secção anterior.

Os algoritmos serão ilustrados com processos pontuais simulados, considerando os casos particulares de processos renewal Gausseano e Gama, cujos espectros exactos se representam na figura 2.4. A janela $W(\omega)$ será implementada a partir da resposta em frequência de um filtro passa-alto de Butterworth (ordem $n = 5$), considerando frequências de corte adequadas, determinadas a partir dos espectros teóricos.

Será utilizada como aproximação inicial $\boldsymbol{b}^{(0)}$, solução correspondente à modelização a partir da medida de erro inicial $E(p)$ dada por (2.15). Isto é, $\boldsymbol{b}^{(0)} = (b_0^{(0)}, \cdots, b_p^{(0)})'$,

$$\begin{aligned} b_0^{(0)} &= 1/\sqrt{V} \\ b_i^{(0)} &= a_i/\sqrt{V}, i = 1, \cdots, p \end{aligned} \tag{3.26}$$

onde $\{a_n\}$, designa a solução do sistema de equações (2.16) e V o erro médio quadrático mínimo correspondente à formulação inicial, determinado a partir de (2.17) (ou de forma equivalente pela equação (3.13)).

Note-se que o valor inicial $\boldsymbol{b}^{(0)}$ satizfaz à restrição (3.16), podendo mesmo constituir uma boa descrição do espectro não paramétrico, para uma ordem p suficientemente elevada.

3.3.1 Método de projecção do gradiente

O ponto de partida natural para a determinação da solução do problema de minimização restringida é o *método de projecção do gradiente*.

Em cada passo é feita uma iteração a partir do método de descida mais rápida, para uma direcção de descida correspondente à projecção do gradiente no subespaço tangente às restrições activas [Luenberger 73]. Isto é,

$$\boldsymbol{b}^{(k+1)} = \boldsymbol{b}^{(k)} - \alpha_k \boldsymbol{P}_k \bigtriangledown E_{mb}(\boldsymbol{b}^{(k)}) \tag{3.27}$$

onde α_k designa o passo na iteração, o qual deverá ser escolhido de forma a minimizar $E_{mb}(\boldsymbol{b}^{(k+1)})$. Designa-se por $\boldsymbol{P}_k$ a matriz de projecção calculada no ponto $\boldsymbol{b}^{(k)}$, a qual é dada, para o caso de uma restrição não linear $h(\boldsymbol{b})$, por

$$\boldsymbol{P}_k = \boldsymbol{I} - \bigtriangledown h(\boldsymbol{b}^{(k)}) \bigtriangledown h(\boldsymbol{b}^{(k)})'/(\bigtriangledown h(\boldsymbol{b}^{(k)})' \bigtriangledown h(\boldsymbol{b}^{(k)})). \tag{3.28}$$

A satisfação da restrição pode ainda ser melhorada iterativamente em cada passo [Luenberger 73].

No problema em estudo, $\bigtriangledown E_{mb}(\boldsymbol{b})$ é dada por (3.20) e, atendendo à expressão (3.18) para $h(\boldsymbol{b})$,

$$\bigtriangledown h(\boldsymbol{b}) = 2\,\Delta t\ \boldsymbol{M}_d\,\boldsymbol{b}. \tag{3.29}$$

Neste caso particular a satisfação da restrição não é crítica. A condição de restrição é muito satisfatoriamente verificada apenas para uma iteração no valor obtido por projecção do gradiente em cada passo, conforme se detalha em seguida.

Cada passo do algoritmo de projecção do gradiente utilizado consiste assim em:

$$\begin{aligned} \boldsymbol{b}_0^{(k+1)} &= \boldsymbol{b}^{(k)} - \alpha_k \boldsymbol{P}_k \bigtriangledown E_{mb}(\boldsymbol{b}^{(k)}) \\ \boldsymbol{b}^{(k+1)} &= \boldsymbol{b}_0^{(k+1)} - \bigtriangledown h(\boldsymbol{b}^{(k)}) h(\boldsymbol{b}_0^{(k+1)}) / (\bigtriangledown h(\boldsymbol{b}^{(k)})' \bigtriangledown h(\boldsymbol{b}^{(k)})). \end{aligned} \tag{3.30}$$

O método de projecção do gradiente conduz a uma solução simples para o problema da determinação da direcção de descida adequada à restrição. No entanto, conforme é referido na literatura [Luenberger 73], o algoritmo correspondente pode não convergir. Além disso, como é inerente ao método de descida mais rápida, a escolha do passo é bastante crítica.

Na figura 3.7 ilustra-se a aplicação do algoritmo descrito em (3.30) para um processo renewal Gausseano, considerando um passo, α, constante. Os estudos de simulação realizados, tanto neste caso como para processos renewal Gama, indicam que é necessário tomar um valor do passo pequeno, não se observando convergência para valores de $\alpha\beta$ superiores a 10^{-3}.

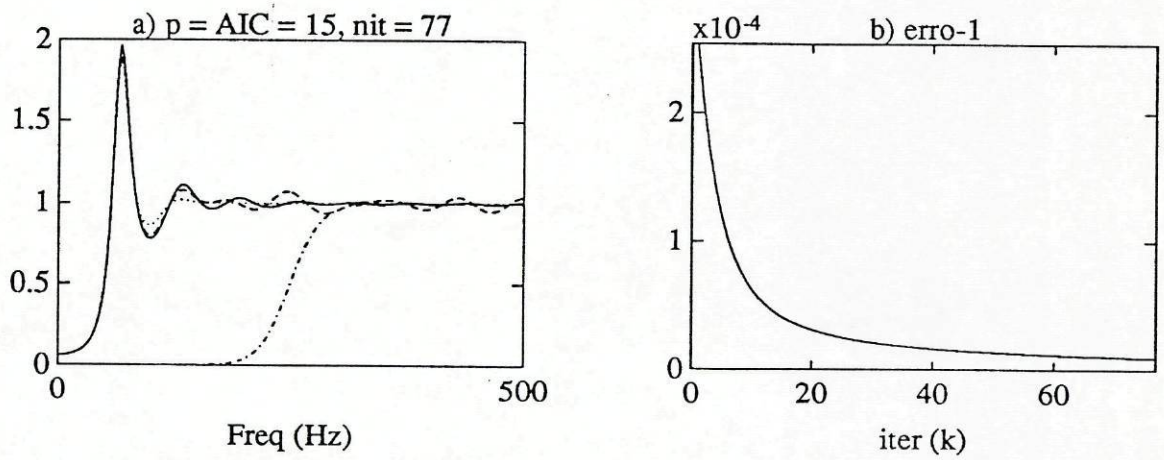


Figura 3.7: **a)** Modelo espectral AR determinado a partir de (3.30) ($p = AIC$, $\alpha\beta = 10^{-3}$, $w_{lim} = 250$Hz), para um processo renewal Gausseano ($T = 5s$,$\lambda = 70s^{-1}$, $\sigma/\mu = 0.25$) após **nit** iterações; densidade de covariância estimada por (2.33) para $\Delta t = 1ms$. Sobrepostos no gráfico representam-se, o modelo espectral inicial (- -), a janela $W(\omega)$ (-.) e o valor exacto do espectro (...). **b)** Diferença da medida de erro E_{mb} para a unidade em cada iteração.

A convergência é lenta, reflectindo o facto do método de descida mais rápida ser um método de primeira ordem. Este efeito é agravado para espectros essencialmente planos, como por exemplo para processos renewal Gama, e este aspecto será ilustrado posteriormente. Torna-se assim essencial desenvolver algoritmos iterativos de ordem superior, que possam ser aplicados de forma eficiente a uma gama variada de situações.

3.3.2 Método de Newton modificado

Um algoritmo iterativo básico para a resolução do problema de minimização da função de custo E_{mb} utiliza a relação recursiva [Luenberger 73]

$$\boldsymbol{b}^{(k+1)} = \boldsymbol{b}^{(k)} - \alpha_k \boldsymbol{S}_k \bigtriangledown E_{mb}(\boldsymbol{b}^{(k)}) \tag{3.31}$$

onde $\boldsymbol{S}_k$ indica uma matriz simétrica quadrada de dimensão $P = p + 1$ e onde, como habitualmente, α_k é escolhido de forma a minimizar $E_{mb}(\boldsymbol{b}^{(k+1)})$.

Tomar para $\boldsymbol{S}_k$ a inversa da matriz Hessiana de E_{mb} corresponde ao *método de Newton clássico*, algoritmo iterativo de segunda ordem. Para outras escolhas de $\boldsymbol{S}_k$ obtêm-se algoritmos alternativos, nomeadamente, o método da descida mais rápida, para $\boldsymbol{S}_k = \boldsymbol{I}$. O *método de Newton modificado* corresponde a considerar $\boldsymbol{S}_k$ uma aproximação da inversa da matriz hessiana, sendo bastante utilizado na prática.

Para o processo iterativo definido por (3.31) ser garantidamente um processo de descida para valores pequenos de α, é necessário em geral requerer que $\boldsymbol{S}_k$ seja uma matriz semi-definida positiva [Luenberger 73].

Na resolução do problema em estudo será utilizado o *método de Newton modificado*, considerando $\boldsymbol{S}_k$ uma aproximação da inversa da matriz hessiana. Nomeadamente, atendendo à expressão (3.20) para $\bigtriangledown E_{mb}$

$$\boldsymbol{S}_k = \boldsymbol{S} = (2\,\Delta t \boldsymbol{M}_d)^{-1}. \tag{3.32}$$

Note-se que esta escolha de $\boldsymbol{S}_k$ é adequada uma vez que garante a convergência do algoritmo, mesmo quando se usam valores estimados para a sucessão densidade de covariância, desde que se utilizem os estimadores semi-definidos positivos descritos na secção 2.3.3.

Para a aplicação do método de Newton modificado na resolução deste problema será ainda necessário, em cada iteração, normalizar os coeficientes de forma a verificar a condição de restrição (3.18).

Cada passo do algoritmo de Newton modificado utilizado, corresponde assim a:

$$\begin{aligned} \boldsymbol{b}_0^{(k+1)} &= \boldsymbol{b}^{(k)} - \alpha\, \boldsymbol{S} \bigtriangledown E_{mb}(\boldsymbol{b}^{(k)}) \\ \boldsymbol{b}^{(k+1)} &= \frac{\boldsymbol{b}_0^{(k+1)}}{\sqrt{\Delta t\ \boldsymbol{b}_0^{(k+1)\,\prime}\, \boldsymbol{M}_d\, \boldsymbol{b}_0^{(k+1)}}} \end{aligned} \tag{3.33}$$

onde $\boldsymbol{S}$ e $E_{mb}(\boldsymbol{b}^{(k)})$ são dadas, respectivamente, por (3.32) e (3.20). O passo α_k, em cada iteração, será mantido fixo ao longo do processo iterativo e igual a α.

A aplicação deste algoritmo ao exemplo apresentado na figura 3.7 permite obter, como seria de esperar, um modelo espectral de características idênticas com um muito menor número de iterações (figura 3.8). A utilização do algoritmo (3.33) é mesmo fundamental no caso de espectros essencialmente planos, conforme se ilustra na figura 3.9.

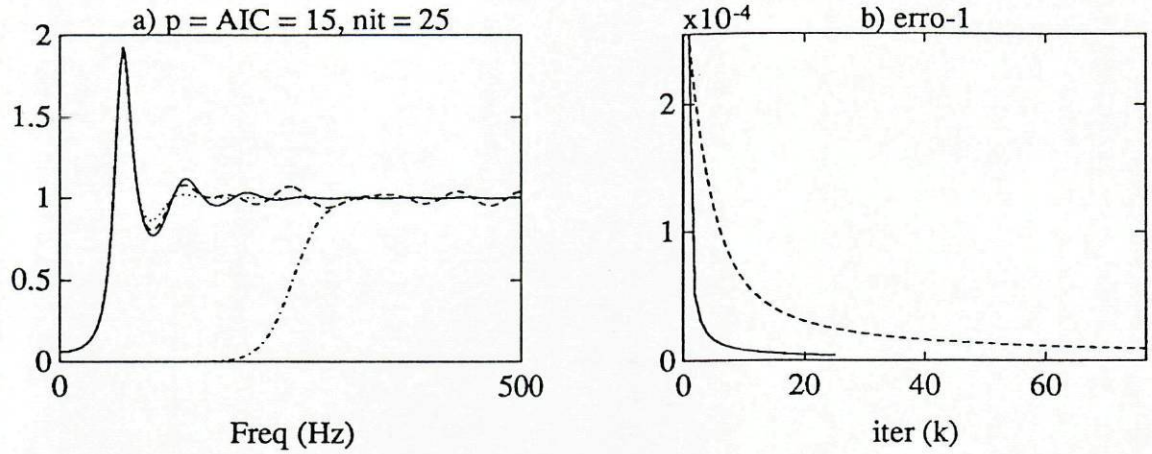


Figura 3.8: **a)** Modelo espectral AR obtido com o algoritmo (3.33) e $\alpha\beta = 0.5$. **b)** Diferença da medida de erro E_{mb} para a unidade em cada iteração; para facilidade de comparação representa-se a tracejado o valor correspondente do erro para o algoritmo (3.30). Mesmos dados da figura 3.7.

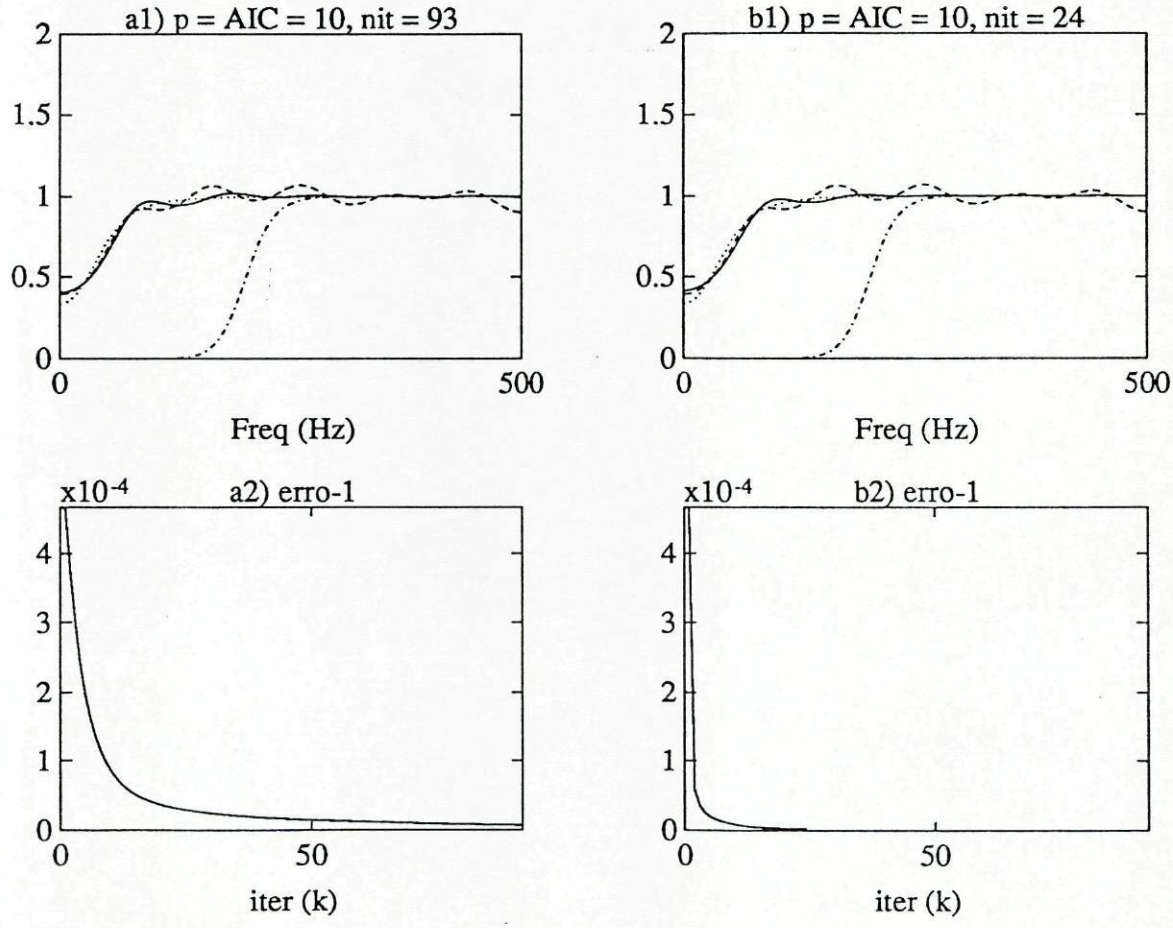


Figura 3.9: Modelo espectral AR obtido a partir de (3.30) (**a1**) e (3.33) (**b1**), ($\alpha\beta = 10^{-3}$ e 0.5, respectivamente; $p = AIC$, $w_{lim} = 200$Hz), para um processo renewal Gama ($n = 3, \lambda = 70s^{-1}$) após **nit** iterações; densidade de covariância estimada por (2.33) ($\Delta t = 1ms, T = 5s$). Sobrepostos nos gráficos representam-se, o modelo espectral inicial (- -), a janela $W(\omega)$ (-.) e o valor exacto do espectro (. . .). **a2)** e **b2)** Diferença da medida espectral E_{mb} para a unidade, em cada iteração, para cada um dos algoritmos.

Uma questão importante na aplicação do método de Newton modificado, é a escolha do passo, α, na iteração. Uma escolha adequada de α permite garantir uma diminuição do erro em cada iteração e conduzir a uma convergência rápida.

Este comportamento é típico deste tipo de algoritmos, verificando-se que para valores de α diferentes do valor óptimo poderá ocorrer uma convergência mais ou menos lenta, conforme se ilustra na figura 3.10 para $\alpha\beta = 0.3, 0.5, 0.65$.

As experiências realizadas por simulação, indicam que o processo é convergente para $0 < \alpha\beta \leq 0.65$, e que o valor 0.5 assegura a convergência mais rápida.

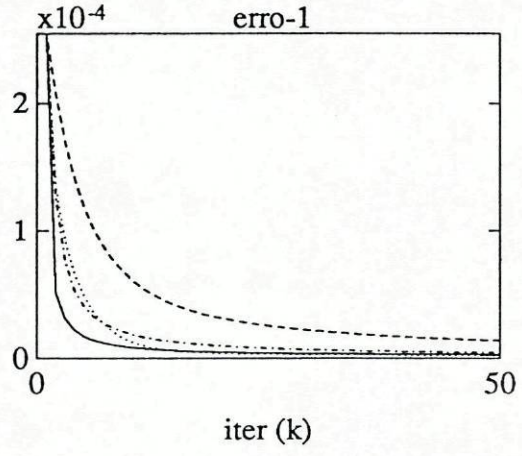


Figura 3.10: Diferença do erro espectral E_{mb} para a unidade, como função da ordem da iteração, utilizando (3.33) para um processo renewal Gausseano; $\alpha\beta = 0.3(- .), 0.5(-), 0.65(\ldots)$. Para facilidade de comparação representa-se também (- -) o valor correspondente do erro para o algoritmo (3.30). Mesmos dados da figura 3.7.

3.3.3 Comentários finais e interpretação da implementação do método de Newton modificado

Os resultados obtidos permitem concluir que a utilização do método de Newton modificado (a partir das equações (3.33) e para uma escolha adequada do passo de iteração), conduz a um algoritmo convergente e susceptível de implementação eficiente para o aperfeiçoamento da modelização espectral AR de processos pontuais.

Nomeadamente, será de salientar que:

- A escolha de $\boldsymbol{S}_k$, dada por (3.32), não só garante a convergência processo iterativo, como permite o desenvolvimento de algoritmos de implementação extremamente eficiente. Uma vez que a matriz $\boldsymbol{S}_k$ se mantem fixa, o seu cálculo pode ser feito no início do processo (uma única inversão matricial).

- Escolhida a frequência de corte ω_{lim} e a janela $W(\omega)$, a sequência $\{\omega(n\Delta t)\}$ também se mantem fixa ao longo do processo. Em cada iteração é apenas necessário actualizar $\{\varphi_b(n\Delta t)\}$, em função do espectro paramétrico correspondente.

Finalmente, será de salientar que o método de Newton modificado, tal como foi implementado, é susceptível de interpretação como resolução directa do sistema de equações (3.25).

Com efeito, substituindo em (3.33) as expressões para $\nabla E_{mb}(\boldsymbol{b}^{(k)})$ e $\boldsymbol{S} = \boldsymbol{S}_k$, respectivamente, (3.20) e (3.32), obtem-se o algoritmo equivalente:

$$\begin{aligned} \boldsymbol{b}_0^{(k+1)} &= (\boldsymbol{M}_d)^{-1}(\beta\, \boldsymbol{M}_b(\boldsymbol{b}^{(k)})\, \boldsymbol{b}^{(k)}) \\ \boldsymbol{b}_1^{(k+1)} &= (1-\alpha)\boldsymbol{b}^{(k)} + \alpha\, \boldsymbol{b}_0^{(k+1)} \\ \boldsymbol{b}^{(k+1)} &= \frac{\boldsymbol{b}_1^{(k+1)}}{\sqrt{\Delta t\ \boldsymbol{b}_1^{(k+1)\,\prime}\, \boldsymbol{M}_d\, \boldsymbol{b}_1^{(k+1)}}} \end{aligned} \tag{3.34}$$

Em cada iteração, o valor de $\boldsymbol{b}_0^{(k+1)}$ corresponde à resolução directa do sistema (3.25), agora "linearizado", pelo cálculo do termo independente para $\boldsymbol{b} = \boldsymbol{b}^{(k)}$.

Note-se que para garantir a convergência deste algoritmo (necessariamente para um passo de iteração $\alpha \neq 1$), a normalização devida à restrição (3.16), é precedida de uma correcção do valor obtido para $\boldsymbol{b}_0^{(k+1)}$.

Conforme já foi referido, a matriz densidade de covariância $\boldsymbol{M}_d$ mantem-se fixa ao longo de todo o processo iterativo, podendo ser invertida no início do processo.

Será também importante salientar que, pelo facto da matriz $\boldsymbol{M}_d$ ser simétrica Toeplitz, podem ser utilizados algoritmos eficientes para a resolução do sistema linear envolvido na determinação de $\boldsymbol{b}_0^{(k+1)}$ em (3.34).

3.4 Algoritmos eficientes tipo Levinson baseados numa densidade de covariância "modificada"

O aperfeiçoamento da modelização espectral AR de processos pontuais é susceptível de implementação eficiente, utilizando o método de Newton modificado, que conduz à resolução de um sistema linear em cada iteração. Apesar de poder ser aproveitado o facto da matriz dos coeficientes ser simétrica Toeplitz, em contraste com a medida de erro (2.17) para a formulação 1, não é possível a aplicação do algoritmo de Levinson.

Seria vantajoso utilizar do algoritmo de Levinson em cada iteração. Com efeito, o número de operações por iteração seria ainda mais reduzido.

Conforme será mostrado nesta secção, o aperfeiçoamento da modelização espectral AR considerando a medida de erro modificada, admite uma interpretação simples para o caso de $\beta = 1$. Para este valor de β obtêm-se algoritmos simplificados, susceptíveis de implementação extremamente eficiente e, em particular, torna-se possível utilizar o algoritmo de Levinson em cada iteração.

3.4.1 Introdução

Retomando a formulação 2, em termos do modelo espectral $G_a(\omega)$, reconhece-se facilmente que o sistema de equações normais (3.8), decorrente da minimização de E_{ma} (dada por (3.7)) relativamente aos parâmetros AR, $\{a_n\}_{n=1,\cdots p}$, pode ser escrito numa forma equivalente.

Com efeito, definindo a sucessão $\{g_m(n\Delta t)\}$ por

$$g_m(n\Delta t) = g_d(n\Delta t) + \beta\,(\lambda\, w(n\Delta t) - \varphi_a(n\Delta t)),\ n = 0, \pm 1, \ldots \tag{3.35}$$

onde $g_d(n\Delta t)$ designa a sucessão densidade de covariância (3.4), correspondente ao espectro $G_d(\omega)$, e as sucessões $w(n\Delta t)$ e $\varphi_a(n\Delta t)$ são dadas pelas expressões (3.10) e (3.11), respectivamente, o sistema de equações (3.8) pode escrever-se na forma

$$\sum_{k=1}^{p} a_k g_m(|n-k|\Delta t) = -g_m(n\Delta t),\ n = 1, \ldots, p. \tag{3.36}$$

O sistema (3.36) correspondente à medida de erro modificada, $E_{ma}(p)$, é inteiramente análogo ao sistema (2.16), para a medida de erro inicial, $E(p)$. Esta observação sugere que a interpretação da incorporação na medida de erro da informação sobre as propriedades do espectro das contagens, pode ser interpretada como correspondendo à modificação da sucessão densidade de covariância. Esta modificação tem como objectivo obter modelos espectrais com um comportamento mais adequado na região das altas frequências.

Interessa assim verificar em que condições é que a sucessão, $\{g_m(n\Delta t)\}$ dada por (3.35), a qual será designada por *sucessão densidade de covariância modificada*, é semi-definida positiva. Neste caso o problema poderá (alternativamente) ser reduzido ao problema inicialmente estudado (formulação 1, secção 2.2.4).

Designando por $G_m(\omega)$, $W(\omega)$ e $G_a(\omega)W(\omega)$, os espectros das sucessões $\{g_m(n\Delta t)\}$, $\{w(n\Delta t)\}$, e $\{\varphi_a(n\Delta t)\}$, respectivamente, definidos de forma análoga a $G_d(\omega)$, a partir da expressão (2.9), conclui-se que

$$G_m(\omega) = G_d(\omega) + \beta W(\omega)\,(\lambda - G_a(\omega)). \tag{3.37}$$

Assim, supondo que nesta expressão $G_a(\omega)$ representa um modelo espectral AR previamente determinado (nomeadamente o obtido a partir da medida de erro inicial $E(p)$) verifica-se facilmente que $G_m(\omega)$ poderá ser garantidamente não negativo.

Com efeito, na região das baixas frequências ($W(\omega) = 0$) o espectro "modificado", $G_m(\omega)$, é idêntico a $G_d(\omega)$ e portanto não negativo. Nas altas frequências ($W(\omega) \neq 0$), o espectro $G_m(\omega)$ apresenta-se como um espectro $G_d(\omega)$ corrigido das diferenças do modelo espectral $G_a(\omega)$ para λ. Devido ao amaciamento inerente aos estimadores espectrais AR, o modelo espectral $G_a(\omega)$ corresponde aproximadamente a uma curva média do espectro não paramétrico e, como nessa região o espectro já é aproximadamente plano e igual a λ, aquela diferença é muito inferior, em valor absoluto, ao valor de $G_d(\omega)$.

Conclui-se assim que, *tomando* $\beta = 1$ *e escolhendo uma frequência* ω_{lim} *adequada* , fica assegurado que, partindo de uma sucessão densidade de covariância semi-definida positiva, o espectro $G_m(\omega)$ é garantidamente não negativo e a sucessão densidade de covariância modificada correspondente, $\{g_m(n\Delta t)\}$, obtida a partir de (3.35), preserva aquela propriedade.

3.4.2 Densidade de covariância modificada e modelo espectral correspondente: resultados de simulação

Nesta secção discutem-se os efeitos decorrentes da substituição da sucessão densidade de covariância por uma sucessão densidade de covariância modificada. Os resultados apresentados nas figuras 3.11 e 3.12 foram obtidos por simulação considerando o caso de processos estocásticos pontuais renewal Gausseanos e Gama, respectivamente.

Pretende-se ilustrar que a utilização da sucessão densidade de covariância modificada permite obter um modelo espectral AR com um comportamento mais adequado na região das altas frequências. Para uma melhor interpretação das alterações observadas comparam-se também a sucessão densidade de covariância estimada, a sucessão modificada e o valor teórico correspondente.

Na figuras 3.11 e 3.12 representam-se em (a1) e (b1), respectivamente, a sucessão densidade de covariância estimada a partir de (2.33) e os respectivos espectros das contagens (paramétrico e não paramétrico) normalizados. O espectro paramétrico (+) corresponde ao modelo espectral AR obtido por minimização do critério de AIC para a medida de erro (2.15) na formulação 1.

Sobrepostos nos gráficos representa-se a tracejado o valor da densidade de covariância teórica e o valor exacto normalizado do espectro. A janela espectral $W(\omega)$ (-.), utilizada nos cálculos, foi obtida de forma análoga à das secções anteriores, escolhendo a frequência limite ω_{lim} com base no valor exacto do espectro.

Para a determinação do modelo espectral correspondente à sucessão densidade de covariância modificada deverá começar-se por resolver o sistema (3.36), o qual é não linear, uma vez que $\varphi_a(n\Delta t)$ envolve os próprios parâmetros AR.

No entanto, esta dificuldade poderá ser facilmente obviada. A forma natural de "linearizar" aquele sistema, consiste em utilizar nos cálculos uma sucessão $\varphi_a(n\Delta t)$, calculada a partir do modelo espectral $G_a(\omega)$ previamente determinado com base na medida de erro inicial $E(p)$, dada por (2.15).

Note-se que, atendendo às conclusões da secção anterior e, uma vez que se utilizam estimadores semi-definidos positivos para a sucessão densidade de covariância, uma escolha adequada de w_{lim} garante a obtenção de sucessões $\{g_m(n\Delta t)\}$ semi-definidas positivas.

A determinação de um modelo espectral AR melhorado, a partir da sucessão densidade de covariância modificada, pode reduzir-se assim ao problema inicialmente formulado, para a medida $E(p)$.

Os parâmetros AR, $\{a_n\}_{n=1,\cdots,p}$, e o valor de V para o modelo espectral relativo à densidade de covariância modificada, são determinados a partir das equações (2.16) e (2.17), respectivamente, substituindo $g_d(n\Delta t)$ por $g_m(n\Delta t)$. Em particular, poderão utilizar-se algoritmos eficientes tipo Levinson.

Na figuras 3.11 e 3.12 representa-se em (a2) e (b2), respectivamente, a sucessão densidade de covariância modificada (calculada com base no modelo espectral $G_a(\omega)$ em (b1)) e os correspondentes espectros, paramétrico e não paramétrico, normalizados.

As diferenças entre os espectros relativos à sucessão densidade de covariância inicial (it=0) e à sucessão modificada correspondente (it=1) são notórias nas altas frequências, essencialmente nos modelos espectrais autoregressivos e em particular para o processo Gama.

Verifica-se em (b2) que o modelo espectral AR relativo à sucessão densidade de covariância modificada, representada em (a2) (it=1) é muito mais plano na região das altas frequências do que a aproximação inicial $G_a(\omega)$ representada em (b1) (it=0). Esta característica observa-se mesmo nas zonas onde não são muito evidentes diferenças nos espectros não paramétricos correspondentes ($G_d(\omega)$ e $G_m(\omega)$, respectivamente) e reflecte a modificação nas sucessões densidade de covariância. Verifica-se que na sucessão densidade de covariância modificada (a2) são corrigidas algumas das diferenças da densidade de covariância inicial (a1) relativamente ao valor teórico.

Para maior facilidade de comparação representam-se, sobrepostos em (c1), os modelos espectrais relativos à sucessão densidade de covariância inicial e à sucessão densidade de covariância modificada.

Nos exemplos apresentados, foi escolhida uma ordem p para o modelo espectral autoregressivo suficientemente elevada. Com efeito, considerando como aproximação inicial para o espectro paramétrico o modelo espectral AR correspondente à minimização do critério de AIC para a medida de erro $E(p)$, obtêm-se modelos espectrais $G_a(\omega)$ que constituem boas descrições do espectro não paramétrico $G_d(\omega)$ conforme se representa em (b1).

Verifica-se, neste caso, que o espectro $G_m(\omega)$, representado em (b2), corresponde mesmo a um melhor estimador espectral do que $G_d(\omega)$. A amplitude das flutuações na região das altas frequências diminui, devido ao facto de o espectro não paramétrico $G_m(\omega)$ estar (em parte) corrigido das diferenças para λ, nessa gama de frequências.

Este facto é bastante notório nos diagramas de polos correspondentes (c2).

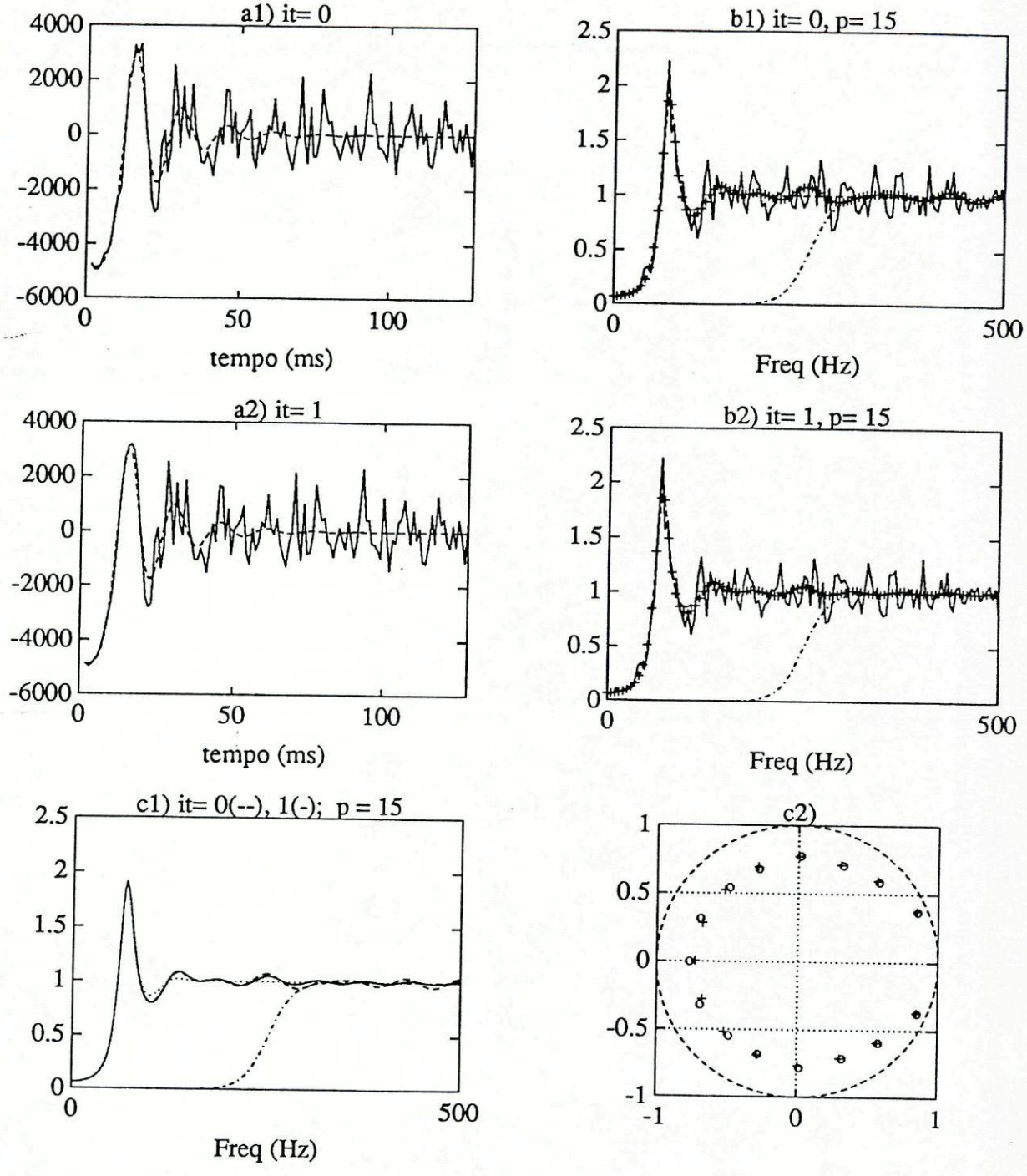


Figura 3.11: a1) Densidade de covariância estimada por (2.33), ($T = 5s$,$\Delta t = 1ms$), para um processo renewal Gausseano ($\lambda = 70s^{-1}$, $\sigma/\mu = 0.25$) e **b1)** espectros normalizados, não paramétrico (método de Blackman-Tuckey ($L = 256$),-) e modelo espectral AR (p=AIC,+) respectivos; **a2)** e **b2)** representam a sucessão densidade de covariância modificada e os espectros normalizados (paramétrico e não paramétrico) correspondentes. Representa-se também a janela espectral utilizada (-.), bem como a densidade de covariância teórica e o valor exacto normalizado do espectro (- -). O valor da densidade de covariância na origem não está representado nos diagramas. Em **c1)** (- - ,o) e **c2)** (-,+) representam-se os espectros AR em **b1)** e **b2)** (espectro teórico (...)) e os diagramas de polos correspondentes.

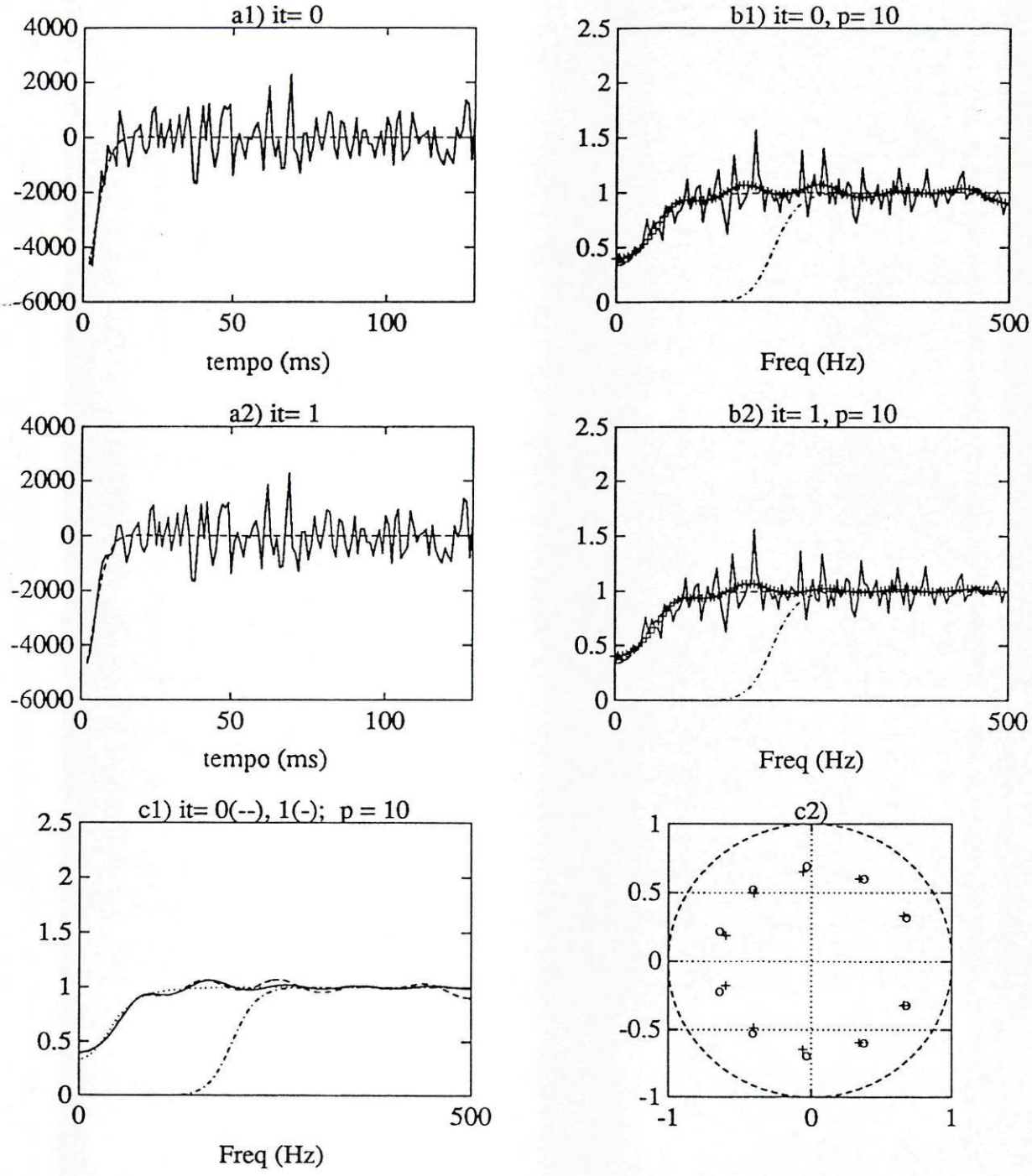


Figura 3.12: Função densidade de covariância estimada a partir de (2.33), para um processo renewal Gama ($\lambda = 70s^{-1}$, $n = 3$), densidade de covariância modificada, e espectros não paramétricos, paramétricos e diagramas de polos correspondentes; ver legenda da figura 3.11.

3.4.3 Aperfeiçoamento do modelo espectral AR por iteração na densidade de covariância

Os resultados apresentados sugerem naturalmente o desenvolvimento de algoritmos iterativos (simplificados e de implementação extremamente eficiente) para o aperfeiçoamento do modelo espectral AR, por modificações sucessivas da densidade de covariância.

Seja p a ordem escolhida para o modelo espectral AR, $G_a(\omega)$, e valores iniciais para a densidade de covariância e para os parâmetros espectrais, respectivamente,

$$\{g_m^{(0)}(n\Delta t)\},\ \boldsymbol{a}^{(0)} = (1, a_1^{(0)}, \cdots, a_p^{(0)})' \text{ e } V^{(0)}.$$

Cada passo do algoritmo para a determinação do modelo espectral AR por iteração na sucessão densidade de covariância, consistirá assim em:

1. Actualização da densidade de covariância

$$g_m^{(k+1)}(n\Delta t) = g_m^{(k)}(n\Delta t) + (\lambda\, w(n\Delta t) - \varphi_a^{(k)}(n\Delta t)),\ n = 0, \pm 1, \ldots, \tag{3.38}$$

onde

$$w(n\Delta t) = \frac{1}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} W(\omega)\, e^{j\omega n\Delta t}\, d\omega$$

e

$$\varphi_a^{(k)}(n\Delta t) = \frac{1}{2\pi}\int_{-\pi/\Delta t}^{+\pi/\Delta t} (G_a^{(k)}(\omega)W(\omega))\, e^{j\omega n\Delta t}\, d\omega$$

é calculado para o modelo espectral $G_a^{(k)}(\omega)$ actual, isto é

$$G_a^{(k)}(\omega) = \frac{V^{(k)}}{|1 + \sum_{n=1}^{p} a_n^{(k)} e^{-jnw\Delta t}|^2}.$$

2. Determinação dos parâmetros do modelo espectral correspondente, $\boldsymbol{a}^{(k+1)}$ e $V^{(k+1)}$,

$$\begin{aligned} \boldsymbol{a}^{(k+1)} &: \text{ solução do sistema linear de p equações} \\ &\sum_{i=1}^{p} a_i^{(k+1)} g_m^{(k+1)}(|n-i|\Delta t) = -g_m^{(k+1)}(n\Delta t), \\ V^{(k+1)} &= \Delta t\,(g_m^{(k+1)}(0\Delta t) + \sum_{i=1}^{p} a_i^{(k+1)} g_m^{(k+1)}(i\Delta t)). \end{aligned} \tag{3.39}$$

Para a inicialização do algoritmo considera-se para a sucessão densidade de covariância modificada o valor inicial,

$$g_m^{(0)}(n\Delta t) = g_d(n\Delta t),\ n = 0, \pm 1, \ldots. \tag{3.40}$$

Os parâmetros espectrais iniciais $\boldsymbol{a}^{(0)} = (1, a_1^{(0)}, \cdots, a_p^{(0)})'$ e $V^{(0)}$ são obtidos considerando como aproximação inicial a solução do sistema (2.16) correspondente à modelização a partir da medida de erro inicial $E(p)$ e o erro médio quadrático mínimo respectivo, (2.17). Isto é,

$$\begin{aligned} \boldsymbol{a}^{(0)} \quad &: \quad \text{solução do sistema linear de p equações} \\ &\sum_{i=1}^{p} a_i^{(0)} g_m^{(0)}(|n-i|\Delta t) = -g_m^{(0)}(n\Delta t), \\ V^{(0)} \quad &= \quad \Delta t \, (g_m^{(0)}(0\Delta t) + \sum_{i=1}^{p} a_i^{(0)} g_m^{(0)}(i\Delta t)). \end{aligned} \tag{3.41}$$

Em cada iteração, a determinação do modelo espectral para a medida de erro modificada, reduz-se completamente, ao problema estudado para a medida de erro inicial, considerando a densidade de covariância actualizada. Nomeadamente, os resultados apresentados na secção anterior, correspondem à inicialização deste algoritmo, a partir de (3.41), seguida de uma iteração para a densidade de covariância modificada, com base em (3.38) e (3.39), utilizando o algoritmo de Levinson.

Na figura 3.13 ilustra-se a aplicação deste algoritmo. Como seria de esperar, no modelo espectral final a parte constante do espectro corresponde a polos uniformemente distribuidos e a igual distância do círculo unitário.

Este e outros exemplos permitem concluir que a aplicação do algoritmo baseado na iteração na densidade de covariância, conduz a resultados semelhantes ao método de Newton modificado. Conforme se ilustra na figura 3.13, o algoritmo simplificado poderá mesmo ser mais vantajoso, conduzindo a um menor número de iterações, em particular no caso de espectros essencialmente planos.

No entanto, a vantagem essencial da utilização da iteração na densidade de covariância sobre o método de Newton, consiste na redução do número de operações aritméticas, em particular, quando se aplica o algoritmo de Levinson em cada iteração.

Note-se que a actualização da densidade de covariância, a partir da expressão (3.38) é fundamental para garantir a convergência rápida do processo. Se, em cada iteração, se actualizassem apenas os parâmetros AR, considerando a densidade de covariância "fixa", isto é, se se substituisse a relação (3.38) por

$$g_m^{(k+1)}(n\Delta t) = g_d(n\Delta t) + (\lambda \, w(n\Delta t) - \varphi_a^{(k)}(n\Delta t)), \; n = 0, \pm 1, \ldots$$

viria

$$G_m^{(k+1)}(\omega) = G_d(\omega) + W(\omega) \, (\lambda - G_a^{(k)}(\omega)).$$

A análise desta expressão permite, em geral, concluir que o processo seria muito lentamente convergente, com uma diminuição do erro apenas de duas em duas iterações.

Com efeito, a última parcela traduziria as diferenças de $G_a^{(k)}(\omega)$ para λ. Supondo inicializado o processo e feita a primeira iteração, na iteração seguinte, a correcção

seria muito menos significativa relativamente a $G(\omega)$ do que na primeira. Assim, seria obtido um modelo espectral ligeiramente mais plano do que na aproximação inicial, embora pior do que na primeira. Este efeito seria observado de duas em duas iterações.

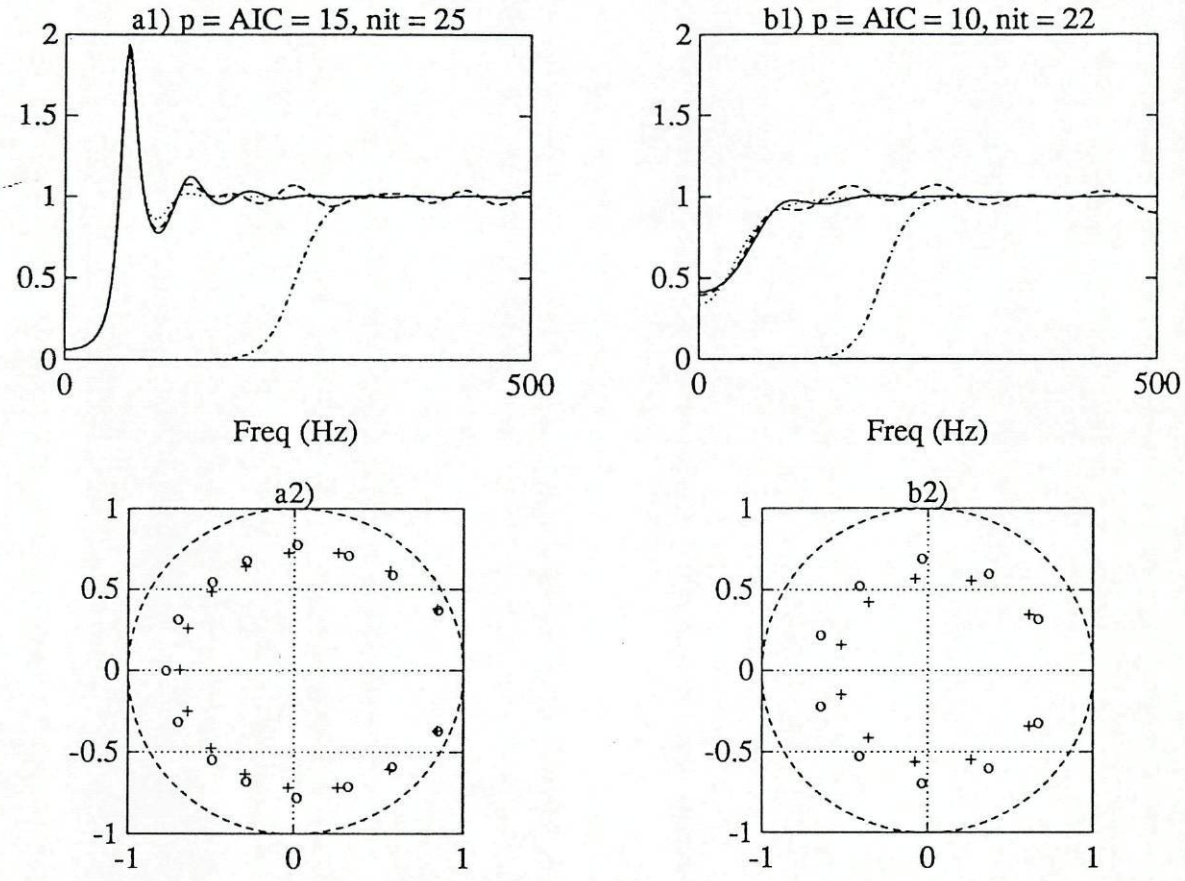


Figura 3.13: Modelo espectral AR determinado a partir de iteração na densidade de covariância no caso de um processo renewal Gausseano (**a1**) e de um processo Gama (**b1**), após **nit** iterações, e diagramas de polos correspondentes (+). Sobrepostos nos gráficos representa-se a aproximação inicial (- -), o diagrama de polos correspondente (o) e a janela espectral utilizada (-.). Mesmos dados das figuras 3.11 e 3.12.

Será finalmente de salientar que para garantir a convergência do processo é também essencial manter a ordem do modelo a partir da primeira iteração.

A iteração na densidade de covariância a partir da expressão (3.38) garante que em cada passo, é determinado o modelo espectral correspondente a um espectro cada vez mais plano na região das altas frequências. Fica assim assegurada a *diminuição* do erro em cada iteração.

3.5 Aperfeiçoamento da modelização espectral AR: alguns comentários

No aperfeiçoamento da modelização espectral autoregressiva do processo das contagens e aplicação dos algoritmos apresentados há essencialmente duas variáveis "livres", sobre as quais se poderá actuar, nomeadamente:

- p: ordem do modelo espectral AR
- w_{lim}: frequência limite na janela espectral, $W(\omega)$.

Conclui-se, a partir dos resultados até aqui apresentados, que o bom desempenho espectral depende fortemente de uma escolha adequada daqueles valores.

Embora não seja objecto deste trabalho um estudo exaustivo dessas variáveis, apresenta-se nesta secção uma discussão preliminar, sugerindo possíveis direcções para uma análise mais aprofundada.

Finalmente fazem-se alguns comentários à fidelidade da descrição espectral na gama das baixas frequências.

3.5.1 Escolha da ordem

Os resultados apresentados (quer com base no conhecimento exacto da densidade de covariância, quer utilizando uma sucessão densidade de covariância estimada a partir dos próprios dados) indicam que a utilização do critério de AIC conduz a valores conservativos para a ordem do modelo espectral AR.

No caso do conhecimento da densidade de covariância teórica o problema que se põe consiste, essencialmente, em escolher um valor mínimo para a ordem do modelo. Conforme se ilustrou neste capítulo (ver figura 3.4), obtêm-se boas descrições para a modelizacão espectral AR do espectro das contagens teórico, desde que se utilize uma ordem suficientemente elevada, superior à indicada por minimização do critério de AIC. Verificou-se também que a qualidade do modelo espectral melhora com o aumento da ordem.

Neste caso, os coeficientes de reflexão K_i podem ser utilizados para a obtenção de indicação sobre a ordem mínima a usar na modelização espectral AR. Com efeito, para um modelo AR de ordem p_0, $K_i = 0$, $p > p_0$ [Kay 88].

Em contraste, no caso da modelização do espectro das contagens a partir de uma sucessão densidade de covariância estimada a escolha da ordem é bastante crítica. A ordem do modelo espectral deverá ser superior à indicada pelo critério de AIC, mas não demasiadamente alta, conforme se ilustrou na figura 3.5.

É correntemente referido na literatura que os coeficientes de reflexão estimados podem ser utilizados para a determinação da ordem do modelo AR [Kay 88, e referências

aí incluidas]. Com efeito, para um processo discreto AR de ordem p_0, as estimativas de máxima verosimilhança para os coeficientes de reflexão admitem a distribuição

$$\hat{K}_i \sim N(0, \frac{1}{N_p}), \; i \geq p_0 + 1, \tag{3.42}$$

onde N_p representa o número de pontos na amostra.

Atendendo à analogia entre a análise espectral de processos estocásticos pontuais e de sucessões discretas (apresentada na secção 2.2.2) poderá utilizar-se este resultado no estabelecimento de intervalos de confiança para os coeficientes reflexão e obtenção de indicação sobre a ordem a adoptar. Para isso, na modelização espectral AR de processos pontuais deverá, considerar-se $N_p = T/\Delta t$ em (3.42).

Na figura 3.14 representam-se o modelo espectral AR de ordem $p = 40$ e a sucessão dos coeficientes de reflexão, $\{\hat{K}_i\}_{i=1,...40}$, correspondente, para o caso do processo renewal Gausseano apresentado no exemplo da figura 3.5.

A aplicação do resultado (3.42) permite obter um intervalo de confiança para os coeficientes de reflexão estimados de ordem superior a p_0, $\{\hat{K}_i\}_{i \geq p_0}$. A título exemplificativo representa-se na figura o intervalo de confiança a 90% correspondente, podendo adoptar-se uma ordem escolhida com base no valor i_0, a partir do qual os coeficientes de reflexão estão contidos naquele intervalo.

Este critério conduz à escolha de um modelo de ordem $p = 15$, valor coincidente com o decorrente da aplicação do critério de AIC.

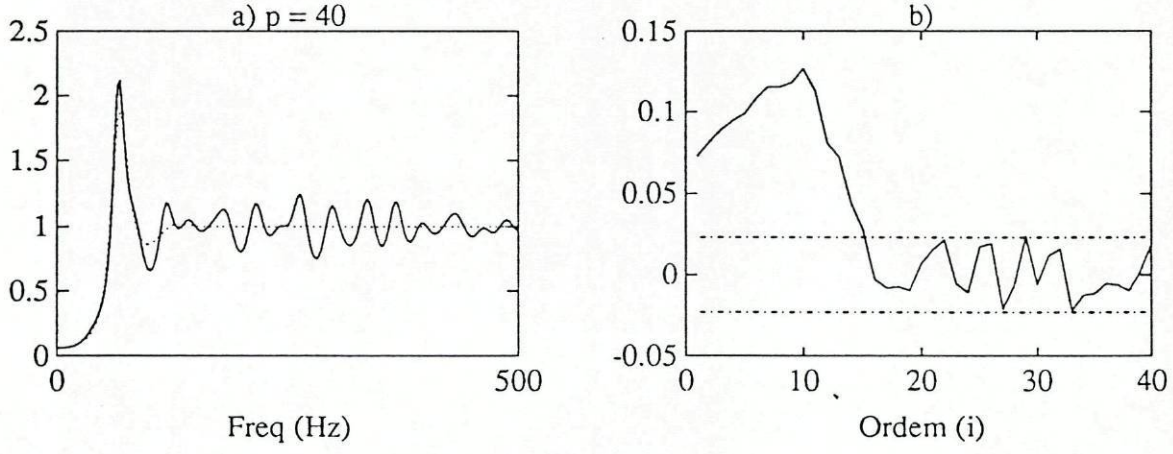


Figura 3.14: **a)** Modelo espectral AR ($p = 40$); **b)** sucessão dos coeficientes de reflexão, $\{K_i\}_{i=1,...,40}$, e intervalo de confiança a 90% correspondente (-.). Processo pontual renewal Gausseano, mesmos dados da figura 3.5.

Conclui-se assim, que a utilização de (3.42) na determinação da ordem do modelo espectral AR, a partir da sucessão dos coeficientes de reflexão estimada com base dos próprios dados, conduz essencialmente a valores semelhantes aos indicados pelo critério de AIC. Não permite, portanto, em princípio, melhorar a informação sobre a escolha da ordem.

Os resultados obtidos a partir do aperfeiçoamento da modelização espectral AR sugerem que se possa utilizar a sucessão dos coeficientes de reflexão correspondentes ao modelo espectral AR aperfeiçoado, para uma ordem relativamente elevada bastante superior à indicada pelo critério de AIC, na obtenção de uma melhor informação relativamente à ordem a adoptar.

Na figura 3.15 representam-se o modelo espectral AR melhorado (ordem $p = 40$) obtido por iteração na densidade de covariância e a sucessão dos coeficientes de reflexão, $\{K_i\}_{i=1,...40}$, correspondente. Para facilidade de comparação apresentam-se sobrepostos os valores teóricos e a aproximação inicial.

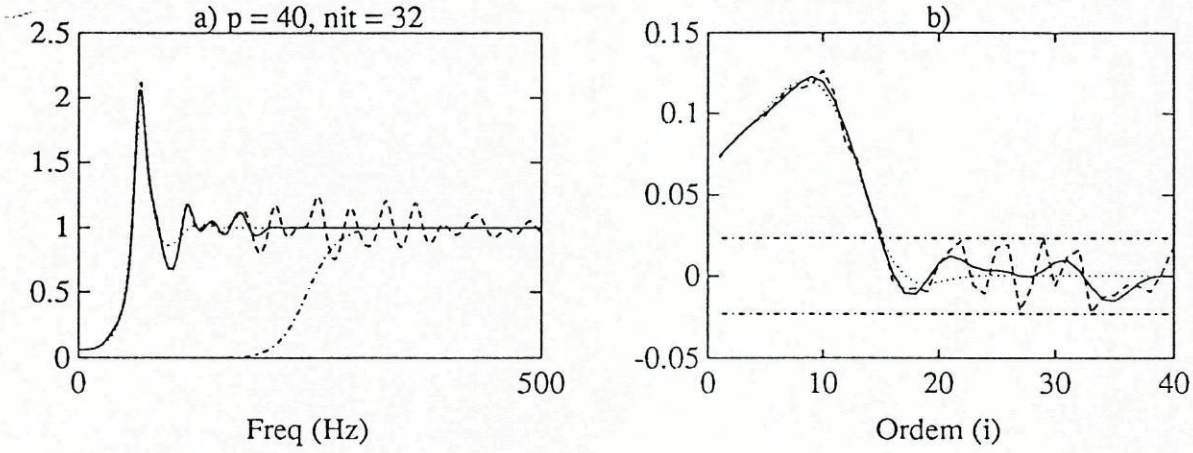


Figura 3.15: **a)** Modelo espectral AR "melhorado"($p = 40$) obtido a partir de iteração na densidade de covariância e **b)** sucessão dos coeficientes de reflexão, $\{K_i\}_{i=1,...,40}$ correspondente. Sobrepostos nos gráficos representam-se os valores teóricos respectivos (...), bem como o modelo espectral inicial, a sucessão dos coeficientes de reflexão e o intervalo de confiança a 90% correspondentes (- -). Processo pontual renewal Gausseano, dados da figura 3.5.

A análise da figura sugere que a sucessão dos coeficientes de reflexão "melhorada" permite definir um intervalo mais estreito que o intervalo de confiança correspondente ao modelo espectral inicial. Neste caso particular, em analogia com a escolha anterior, seria natural adoptar um valor superior a 15 para a ordem do modelo.

Na figura 3.16 representam-se os modelos espectrais aperfeiçoados para as ordens 15 e 16. Verifica-se que a escolha da ordem do modelo é realmente bastante crítica. Por comparação dos modelos espectrais correspondentes àquelas duas ordens (e mesmo dos modelos de ordem superior) conclui-se que a escolha do modelo de ordem 16 permite obter um melhor de desempenho na modelização do espectro do processo considerado, observando-se uma redução bastante significativa do número de iterações.

Foram sistemáticamente observados resultados análogos noutras situações.

Estes exemplos ilustram as dificuldades envolvidas na escolha da ordem na modelização espectral AR de processos estocásticos pontuais. Os resultados obtidos sugerem que se poderá revelar bastante útil, a incorporação de informação sobre as características do espectro, nomeadamente a partir de modelos espectrais "melhorados".

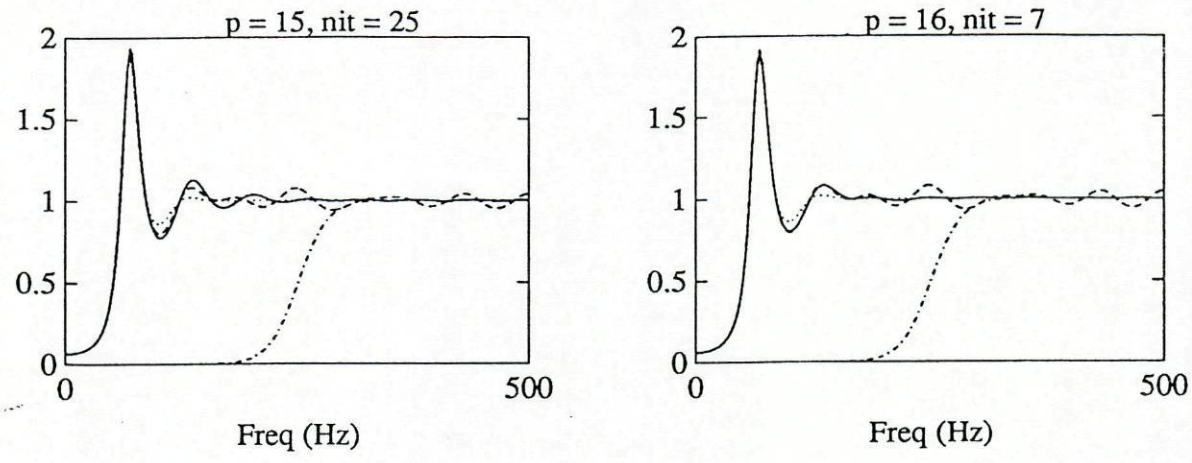


Figura 3.16: Modelos espectrais AR "melhorados",para $p = 15$ e $p = 16$, obtidos por iteração na densidade de covariância. Representa-se também o valor exacto do espectro (...), bem como o modelo espectral inicial (- -) e a janela espectral utilizada (-.). Processo pontual renewal Gausseano, dados da figura 3.5.

3.5.2 Obtenção de indicação da gama de frequências onde o espectro deverá ser plano

A aplicação dos algoritmos desenvolvidos para o aperfeiçoamento da modelização espectral AR do processo das contagens pressupõe que é conhecida a frequência a partir da qual o espectro já é aproximadamente constante. Nos exemplos até aqui apresentados, foram utilizados os valores exactos dos espectros na determinação da frequência limite, w_{lim} (a utilizar nas janela espectral $W(\omega)$), uma vez que aqueles eram conhecidos.

O problema que se põe na prática é a estimação de w_{lim} a partir dos próprios dados, o que será objecto de estudo nesta secção.

O espectro das contagens é do tipo "banda larga" ($G(\omega) \simeq \lambda$, $|w| > \omega_{lim}$).

A informação sobre w_{lim} pode ser obtida a partir de uma descrição global e de baixa resolução do espectro das contagens. Isto é, pretende-se reter apenas informação sobre a *forma dominante*: a parte constante do espectro.

Os modelos espectrais AR (independentemente da ordem utilizada) são métodos de alta resolução. Como tal, permitem obter boas descricões (de alta resolução) da *forma global* do espectro. Não são portanto adequados para descrever apenas a forma dominante, isto é, a parte constante do espectro.

Por outro lado é conhecido que, para processos estocásticos não pontuais, o método não paramétrico de *variância mínima* (muitas vezes designado por método de Capon, de máxima verosimilhança) [Kay 88, Marple 87, e referências aí incluidas] conduz a uma descrição global de resolução mais baixa do que os espectros AR [Burg 72]. Assim, poderá permitir obter informação sobre a forma dominante do espectro.

A aplicação do método de variância mínima na análise espectral de processos estocásticos pontuais é imediata a partir dos resultados fundamentais apresentados na secção 2.2.2. De uma forma geral, a análise espectral de um processo estocástico pontual pode ser reduzida, para frequências $|\omega| < \pi/\Delta t$, à análise espectral de uma sucessão discreta com função de covariância $C(n) = \Delta t.g_d(n\Delta t)$.

Assim, o estimador espectral de variância mínima para do espectro das contagens poderá ser definido por [Kay 88]

$$G_{VM}(\omega) = \frac{p\Delta t}{\boldsymbol{e}^H(\omega)\, R_p^{-1} \boldsymbol{e}(\omega)}, \quad |\omega| < \pi/\Delta t, \tag{3.43}$$

onde $\boldsymbol{e}(\omega)$ designa o vector de frequências complexas

$$\boldsymbol{e}(\omega) = (1,\ e^{j\omega\Delta t},\ e^{j\,2\omega\Delta t},\ \ldots,\ e^{j\,p\omega\Delta t})', \tag{3.44}$$

e $\mathbf{H}$ o operador transposta hermítica. Indica-se por $\boldsymbol{R}_p$ a matriz quadrada, simétrica Toeplitz de dimensão $P = p + 1$ da densidade de covariância, corrigida do factor de escala Δt, isto é,

$$R_p = \Delta t.C_{PP} \tag{3.45}$$

O valor de p foi incluido na expressão (3.43), para garantir uma escala correcta na estimação espectral [Kay 88].

As propriedades do estimador de variância mínima no caso de processos estocásticos em geral são conhecidas [Kay 88, Marple 87, e referências aí incluidas]. Em particular, verifica-se que o desempenho do método de variância mínima é geralmente bom, observando-se um erro de centricidade se existirem picos aguçados.

A escolha do valor de p, dimensão da matriz densidade de covariância utilizada, é bastante crítica. O valor escolhido para p representa o compromisso tradicional "erro de centricidade/variância" dos métodos não paramétricos: deverá escolher-se p suficientemente grande para ter uma resolução adequada, mas não demasiado elevado de forma a induzir uma grande variabilidade espectral.

É geralmente referido que o método de variância mínima apresenta uma resolução superior aos métodos clássicos de estimação espectral, Blackman-Tukey e periodograma. No entanto (conforme já foi referido) a resolução é inferior à modelização espectral AR, reflectindo a relação entre o método de variância mínima e a modelização espectral AR, no caso da covariância ser conhecida [Burg 72].

Finalmente será de salientar que, em contraste com o método de Blackman-Tukey e os estimadores espectrais AR, a transformada inversa de Fourier do espectro determinado pelo método de variância mínima, não dá os valores da covariância utilizados para estimar o espectro.

Na figura 3.17 ilustra-se a aplicação do método de variância mínima na estimação do espectro das contagens de um processo estocástico pontual, para $p = AIC$.

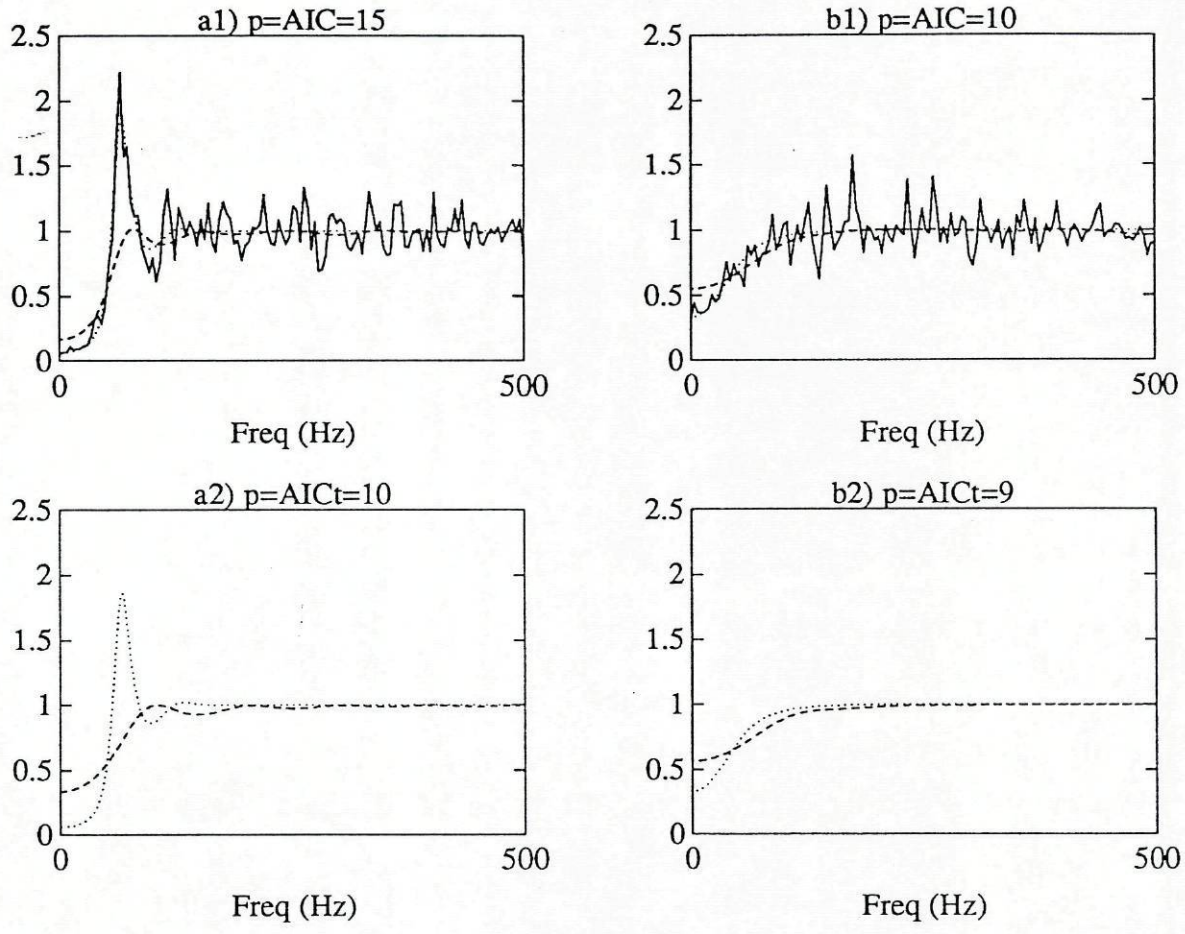


Figura 3.17: Modelo espectral determinado a partir do método de variância mínima (- -, $p = AIC$) para um processo renewal Gausseano (**a1** e **a2**) e um processo renewal Gama (**b1** e **b2**), no caso da sucessão densidade de covariância ser estimada a partir dos dados e de ser conhecido o seu valor exacto, respectivamente. Sobrepostos nos gráficos representa-se o valor exacto normalizado dos espectros (. . .), bem como o espectro não paramétrico correspondente às sucessões densidade de covariância estimadas (-, método de Blackman-Tuckey, L=256) em (**a1**) e (**b1**). Mesmos dados das figuras 3.11 e 3.12.

Verifica-se, quer no caso do conhecimento exacto da densidade de covariância, quer no caso de se utilizarem valores estimados, que o método de máxima verosimilhança permite obter boas descrições espectrais para espectros essencialmente planos, como o do processo pontual gama. No caso do processo Gausseano, devido à existência de um pico acentuado, observa-se um forte erro de centricidade, o que reflecte a escolha um valor demasiado baixo para p.

No entanto, quer para espectros com picos acentuados, quer para espectros essencialmente planos, conclui-se que a aplicação do método de variância mínima para $p = AIC$ pode ser utilizado para obter indicação da gama de frequências onde o espectro é aproximadamente plano.

A análise da figura 3.17 permite concluir, quer no caso do conhecimento exacto da densidade de covariância, quer para valores estimados a partir dos dados, que o espectro deverá ser aproximadamente plano para frequências da ordem de

- 200–250 Hz, para o processo renewal Gausseano.
- 200 Hz, para o processo renewal Gama.

A escolha destes valores para w_{lim} permite obter modelos espectrais melhorados com um bom desempenho espectral. Nomeadamente, nos exemplos que têm vindo a ser apresentados, têm sido utilizadas as frequências 250Hz e 200 Hz, para os processsos estocásticos renewal Gausseanos e Gama, respectivamente (ver nomeadamente figura 3.13).

Note-se para uma escolha $w_{lim} = 200Hz$, no caso do processo renewal Gausseano, se obtêm resultados equivalentes. Neste caso, deverá utilizar-se uma janela espectral com uma região de transição menos extensa (correspondendo na implementação a um filtro de Butterworth de ordem mais alta).

A vantagem essencial da aplicação do método de variância mínima consiste no facto de poder ser obtida informação sobre a frequência limite, w_{lim}, quer no caso de se conhecer a densidade de covariância, quer no caso de os valores serem estimados a partir dos próprios dados, como acontece na prática.

Saliente-se finalmente que para a obtenção de indicação sobre a frequência w_{lim} poderiam utilizar-se outros métodos. Nomeadamente poder-se-ia pensar em utilizar intervalos de confiança para o espectro não paramétrico, ou mesmo modelização espectral MA.

A vantagem de utilizar um indicador baseado no método de variância mínima, consiste no facto de poder ser implementado de uma forma extremamente simples. Com efeito o estimador de variância mínima baseia-se apenas, tal como os algoritmos para o aperfeiçoamento da modelização espectral AR, na matriz densidade de covariância.

3.5.3 Fidelidade da descrição espectral na gama das baixas frequências

O aperfeiçoamento da modelização espectral AR, descrito neste capítulo, permite apenas actuar na zona das altas frequências, onde o espectro deverá ser aproximadamente plano.

O desempenho do estimadores espectrais AR na zona das baixas frequências reflecte a variabilidade na densidade de covariância estimada. Assim, para obter melhores descrições espectrais nas baixas frequências será necessário utilizar registos longos e/ou segmentação para a determinação da densidade de covariância.

Verifica-se também, mesmo nestes casos, que a descrição na gama das baixas frequências é particularmente sensível à ordem do modelo espectral e que o critério de AIC conduz a valores conservativos para a ordem a utilizar.

Na figura 3.18 ilustra-se esta situação, para um processo renewal Gausseano, considerando a covariância média correspondente a 50 realizações independentes.

Representa-se em **a)** e **b)**, respectivamente, o espectro não paramétrico e o respectivo modelo espectral AR melhorado considerando a ordem indicada pela aplicação do critério de AIC). Em **c)** representa-se a sucessão dos coeficientes de reflexão correspondente ao um modelo espectral melhorado de ordem $p = 40$, muito superior à indicada pelo critério de AIC ($p = 15$).

Verifica-se, que se pode obter uma indicação sobre a ordem a utilizar a partir da sucessão dos coeficientes de reflexão correspondentes a um modelo espectral melhorado, para uma ordem p bastante superior à indicada pelo critério de AIC.

Neste exemplo particular seria natural substituir o modelo de ordem 15 (**b**) por um modelo espectral de ordem 17, o qual se representa em **c)**, e que apresenta um melhor desempenho global na região das baixas frequências.

Note-se que a descrição do pico apesar de aparentemente inferior corresponde a uma descrição bastante boa do espectro não paramétrico correspondente.

Conclui-se assim, que a escolha da ordem é bastante crítica, mesmo em situações em que a variabilidade é reduzida.

No entanto, conforme se ilustra na figura, para uma escolha adequada da ordem, podem obter-se modelos espectrais melhorados para o espectro médio, correspondendo a uma descrição bastante fiel na região das baixas frequências.

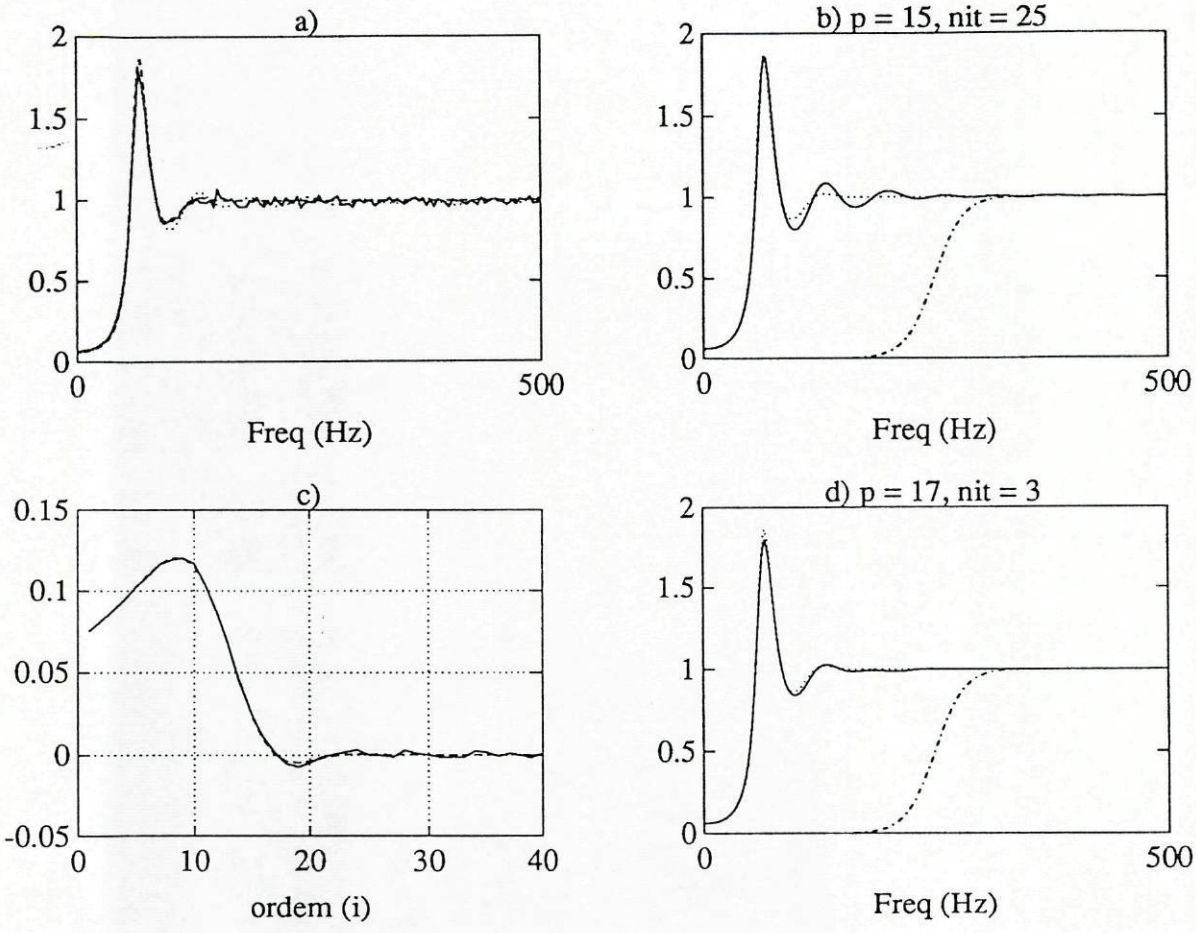


Figura 3.18: a) Espectro não paramétrico médio e **b)** modelo espectral autoregressivo melhorado corrrespondente a 50 realizações independentes de um processo renewal Gausseano; **c)** sucessão dos coeficientes de reflexão correspondentes ao modelo espectral melhorado de ordem 40 ; **d)** modelo espectral melhorado de ordem 17. Sobrepostos nos gráficos representam-se os valores teóricos (- -) correspondentes. Modelos espectrais melhorados obtidos por iteração na densidade de covariância; mesmos dados da figura 3.11.

3.6 Comentários finais

Apesar da modelização espectral autoregressiva de processos estocásticos pontuais (baseada em estimadores semi-definidos positivos para a densidade de covariância) conduzir a boas descrições do espectro das contagens, existem algumas dificuldades na descrição adequada da região das altas frequências, onde o espectro deverá ser aproximadamente plano.

Os estudos de simulação descritos neste capítulo, com base na utilização de um número não mínimo de valores da densidade de covariância e nas propriedades algébricas das matrizes associadas, indicam que não parece ser possível que se consiga aperfeiçoar significativamente a modelização espectral AR, com as medidas de erro correntemente utilizadas para processos não pontuais. No entanto, conforme também foi aqui ilustrado, a modelização espectral AR baseada no conhecimento exacto da função densidade de covariância permite obter boas descrições para o espectro das contagens, desde que se considere uma ordem suficientemente elevada.

A simplicidade dos modelos espectrais AR, sobre outros modelos que permitam traduzir melhor as características do espectro das contagens, justifica plenamente o aperfeiçoamento da modelização espectral AR, pela adaptação às propriedades do espectro. Neste capítulo formulou-se o problema de modelização AR do espectro das contagens a partir de uma medida de erro modificada, incorporando directamente a informação de que o espectro deverá ser plano na região das altas frequências.

A utilização das medidas de erro modificadas conduz ao desenvolvimento de algoritmos iterativos nos parâmetros, para a determinação da solução, os quais são detalhadamente descritos. Verifica-se que os modelos espectrais AR aperfeiçoados, obtidos a partir de estimadores semi-definidos positivos para a densidade de covariância e do conhecimento da gama de frequências onde o espectro deverá ser aproximadamente plano, conduz a boas descrições do espectro das contagens, com um comportamento adequado na região das altas frequências.

Apesar de se ter perdido, em parte, a simplicidade da formulação para a medida de erro inicial, podem ser desenvolvidos algoritmos eficientes, para a obtenção dos modelos melhorados. Em particular a utilização do método de Newton modificado, considerando uma aproximação da matriz hessiana baseada na matriz (simétrica e Toeplitz) densidade de covariância, conduz a algoritmos rapidamente convergentes e susceptíveis de implementação eficiente.

A interpretação da incorporação directa de informação sobre as características do espectro das contagens em termos de uma densidade de covariância modificada, permite desenvolver algoritmos simplificados por iteração na sucessão densidade de covariância. Neste caso o problema é reduzido à formulação baseada na medida de erro inicial, em particular é possível a utilização de algoritmos eficientes, nomeadamente do algoritmo de Levinson para a determinação dos parâmetros. Os resultados obtidos são equivalentes aos do método de Newton modificado.

Na implementação dos métodos e algoritmos desenvolvidos, particularmente no caso em que a densidade de covariância é estimada a partir dos próprios dados, é fundamental escolher adequadamente a ordem do modelo espectral e determinar a gama de frequências onde o espectro deverá ser aproximadamente constante.

Verifica-se que a escolha da ordem é bastante crítica e que o critério de AIC conduz a valores conservativos para a ordem a utilizar. É claramente sugerida a utilização da sucesssão dos coeficientes de reflexão relativa a um modelo espectral "melhorado", de ordem elevada, na obtenção de informação sobre a ordem a considerar na modelização AR.

Quanto à escolha da frequência w_{lim}, os resultados apresentados na parte final deste capítulo, permitem concluir que a aplicação do método de variância mínima, para uma ordem $p = AIC$, dá indicação sobre a frequência limite, w_{lim}, a utilizar nas janelas espectrais.

Capítulo 4

Modelização AR do espectro das contagens: aplicação na caracterização de sinais biomédicos

Neste capítulo ilustra-se a aplicação dos métodos e algoritmos desenvolvidos para a modelização espectral AR de processos pontuais, na caracterização de sinais biomédicos.

Serão utilizados dados experimentais relativos a dois casos particulares: sinais cardiovasculares (ECG) e miograma de interferência (EMG).

Com os resultados apresentados pretende-se ilustrar que a utilização da modelização espectral AR do espectro das contagens poderá (vantajosamente) permitir o desenvolvimento de métodos alternativos de descrição do sinal nas duas situações consideradas. A estrutura do capítulo é a seguinte:

- Modelização espectral AR do processo pontual associado aos intervalos RR no ECG:
  - Obtenção de informação sobre a frequência respiratória a partir do espectro das contagens. Modelização espectral AR e decomposição espectral.
- Modelização espectral AR de processos pontuais extraidos do EMG:
  - Caracterização da complexidade dos potenciais no EMG.
  - Obtenção de informação sobre a frequência média de acendimento das unidades motoras e sua alteração com a fadiga.

4.1 Modelização espectral AR do processo pontual associado aos intervalos RR no ECG

Conforme foi referido no capítulo 1, têm sido utilizadas técnicas de análise espectral (paramétrica e não paramétrica) na descrição da variabilidade dos intervalos RR, a partir da série temporal discreta (tacograma) associada à ordem do batimento.

Nesta secção pretende-se ilustrar que a utilização de uma descrição alternativa, a partir do processo estocástico pontual associado aos instantes de ocorrência das ondas R no ECG, permite traduzir mais correctamente a actividade respiratória.

Apresentam-se também alguns resultados relativos à aplicação da modelização espectral AR do processo das contagens, na caracterização da variabilidade da frequência cardíaca.

4.1.1 Obtenção de informação sobre a frequência respiratória a partir da análise espectral das contagens

Tendo por objectivo o estudo da variabilidade da frequência cardíaca, o ECG pode ser representado por um processo estocástico pontual associado aos instantes de ocorrência das ondas R.

Assim, no domínio da frequência, aquele processo pontual admite as duas descrições complementares referidas no capítulo 1: a função de densidade espectral dos intervalos e a função de densidade espectral das contagens.

A título exemplificativo representam-se na figura 4.1 o espectro do tacograma (correspondente ao espectro dos intervalos no processo pontual) e o espectro das contagens, numa situação de respiração controlada a 0.25 Hz.

Note-se que no caso do espectro dos intervalos em **a2)** se utilizou a escala em frequências corrente, em Hzeq, correspondente a uma normalização pelo intervalo médio. Este artifício permite obviar o facto de no tacograma não existir informação sobre o tempo. A actividade respiratória fica correctamente traduzida na nova escala em frequência.

O espectro das contagens foi obtido a partir do método de Blackman-Tukey.

Em contraste com o espectro dos intervalos, o espectro das contagens retem directamente informação sobre a escala temporal, através da intensidade média λ (o valor limite do espectro, expressão (1.8)).

A gama de frequências no espectro das contagens é apenas limitada pela resolução temporal Δt escolhida. Para a resolução temporal utilizada na figura, $\Delta t = 25\ ms$, os picos no espectro indicam apenas o batimento cardíaco, quase regular.

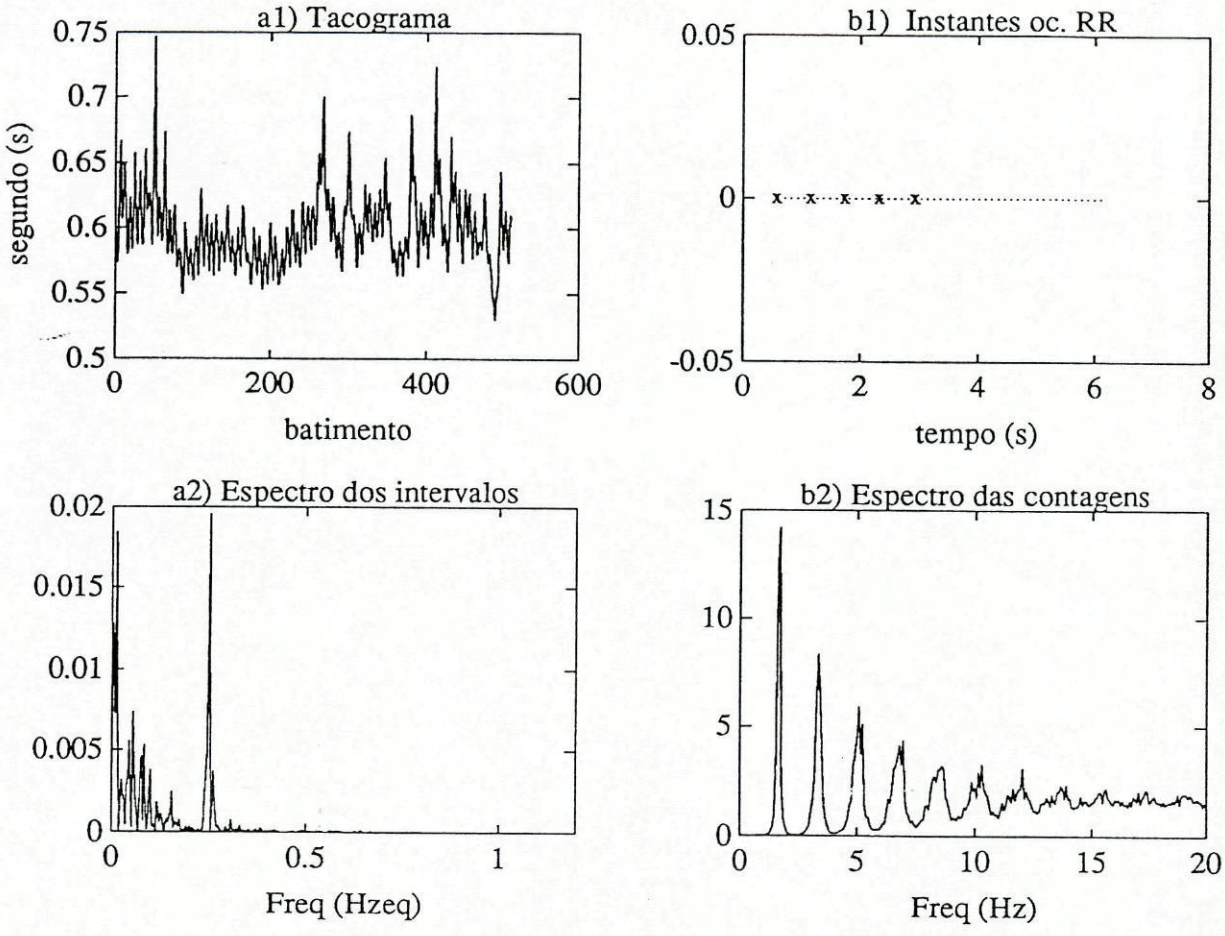


Figura 4.1: Análise espectral do processo pontual associado à variabilidade do RR: **a1)** sucessão dos intervalos RR — tacograma (512 pontos), **b1)** processo pontual associado à ocorrência das ondas R no ECG, **a2)** espectro não paramétrico do tacograma, **b2)** espectro não paramétrico das contagens do processo pontual em **a1)**, determinado a partir dos primeiros 512 valores da estimativa da densidade de covariância (determinada a partir de (2.33), resolução temporal $\Delta t = 25\ ms$). Mesmos dados da figura 1.7.

Embora não se represente na figura, a análise da gama das baixas frequências do espectro não paramétrico das contagens em **b2)**, não evidencia a componente correspondente à frequência respiratória.

Verifica-se, no entanto, conforme se ilustra na figura 4.2, que o espectro não paramétrico determinado a partir do método de Blackman-Tukey, utilizando um número suficientemente elevado de valores da densidade de covariância, permite a identificacão da componente associada à frequência respiratória.

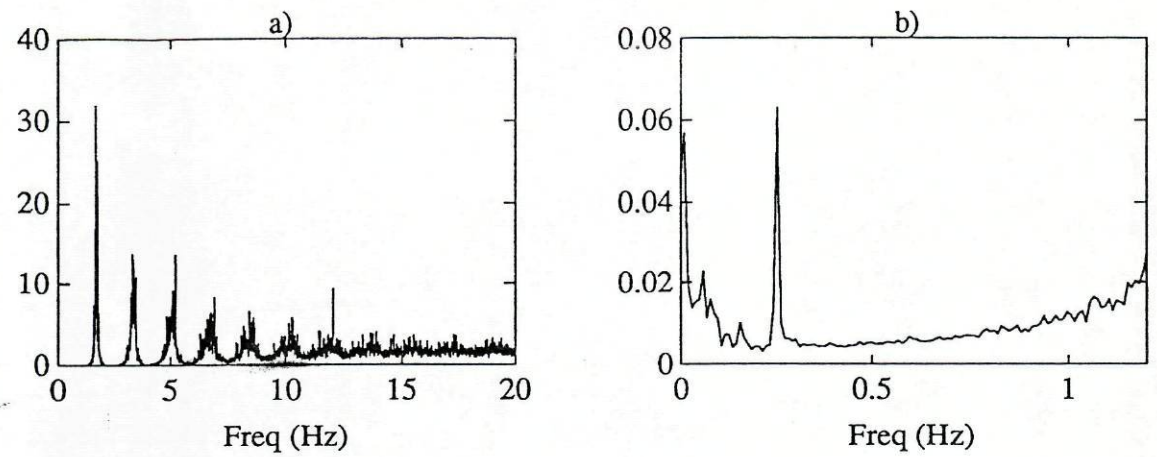


Figura 4.2: Espectro não paramétrico das contagens do processo pontual associado à variabilidade do RR , obtido a partir do método de Blackman-Tuckey para uma resolução temporal $\Delta t = 25\ ms$, com os primeiros 4096 valores da densidade de covariância, considerando: **a)** a gama total de frequências, $[0, \frac{1}{2\Delta t}]$, e **b)** a gama das muito baixas frequências, abaixo de 1.2 Hz. Mesmos dados da figura 4.1.

4.1.2 Modelização espectral AR e decomposição espectral

A modelização espectral AR poderá revelar-se particularmente útil na identificação da componente associada à actividade respiratória.

Para a aplicação da modelização espectral e obtenção de boas descrições, mantendo detalhe na região das baixas frequências, terá de se escolher uma ordem bastante elevada, conforme se ilustra na figura 4.3, tomando $p = 200$. Para garantir uma ordem baixa seria necessário efectuar uma modelização selectiva na zona das baixas frequências. Este tipo de método será ilustrado numa das secções seguintes, numa das aplicações em electromiografia.

Nesta secção considera-se, alternativamente, a decomposição do modelo espectral, com a identificação automática das componentes a partir da sua frequência central.

Este processo descrito na literatura, para o caso de processos estocásticos em geral, baseia-se numa decomposição da função de autocorrelação (e de forma correspondente, do próprio espectro) numa soma de componentes simples, aplicando o teorema dos resíduos [Zetterberg 69].

No caso de processos estocásticos pontuais procede-se, de forma inteiramente análoga à decomposição da densidade de covariância, e à decomposição correspondente do espectro das contagens.

Na figura 4.4 ilustra-se a decomposição espectral AR do espectro das contagens e a identificação da componente associada à actividade respiratória (frequência 0.256 Hz).

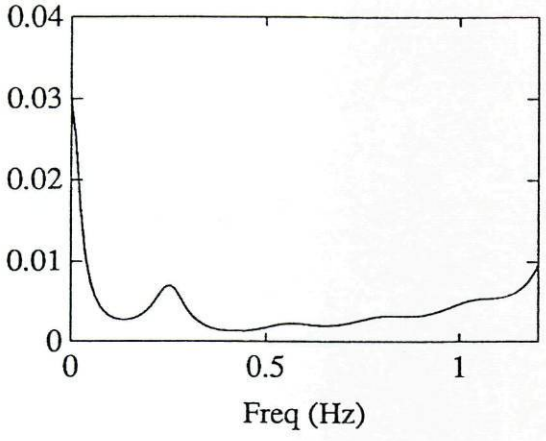


Figura 4.3: Espectro AR ($p = 200$) das contagens do processo pontual associado à variabilidade do RR. Densidade de covariância estimada a partir de (2.33) para uma resolução temporal $\Delta t = 25\,ms$. Representa-se apenas a gama das muito baixas frequências, abaixo de 1.2 Hz. Mesmos dados da figura 4.1.

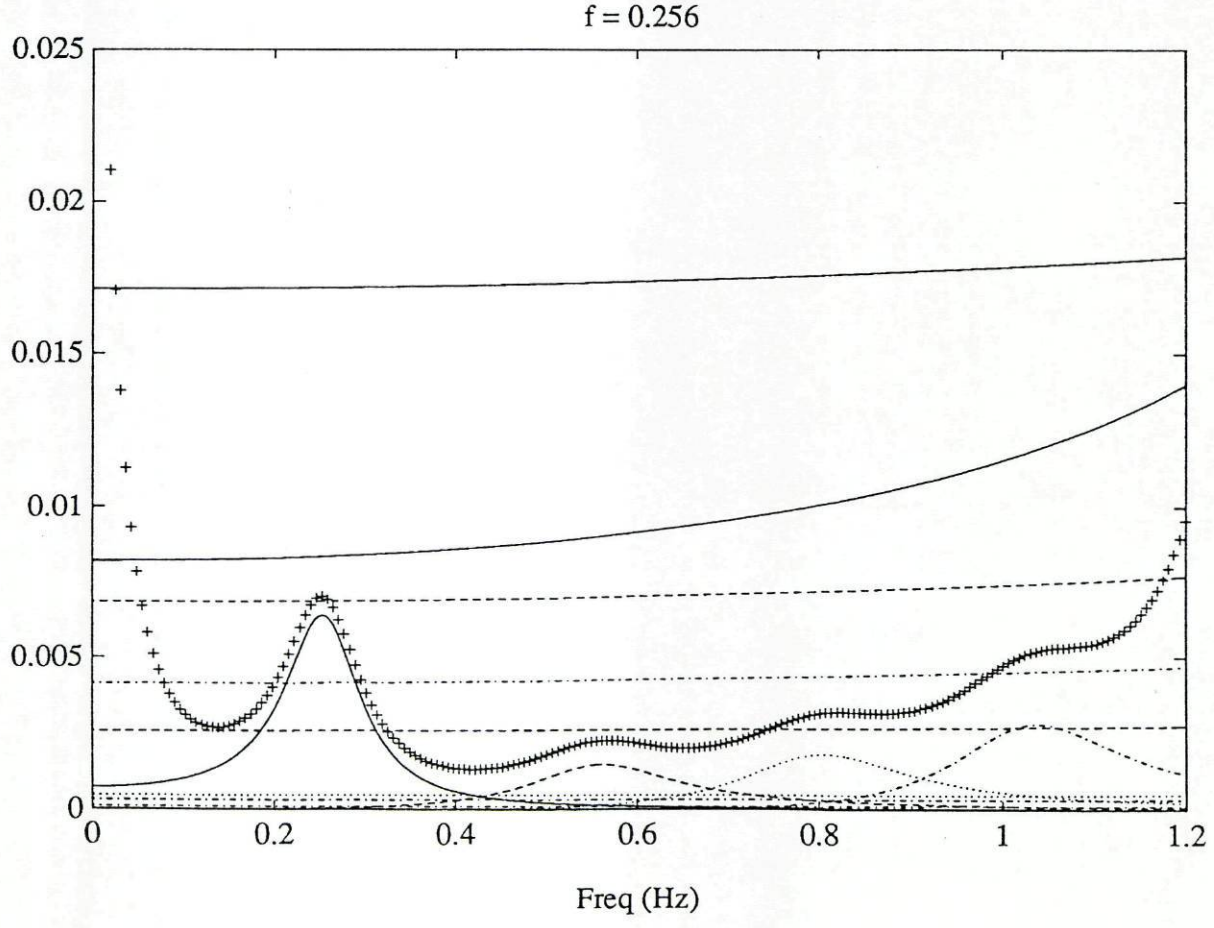


Figura 4.4: Decomposição espectral AR ($p = 200$) das contagens do processo pontual associado à variabilidade do RR. Densidade de covariância estimada a partir de (2.33) para uma resolução temporal $\Delta t = 25\ ms$, considerando a gama das muito baixas frequências, abaixo de 1.2 Hz. Mesmos dados da figura 4.1.

Na figura 4.4, representa-se a contribuição de todas as componentes, na gama de frequências de 0 a 1.2 Hz.

No entanto poderão apenas ser determinadas algumas das componentes, nomeadamente, poderá apenas determinar-se a correspondente à actividade respiratória.

Nestes casos o processo de identificação poderá ser tornado mais eficiente, devendo começar-se por isolar os zeros correspondentes às componentes a analizar. Para isso deverão utilizar-se os algoritmos específicos para a determinação de zeros de polinómios: em particular o algoritmo de Lehmer-Schur (eventualmente seguido do método de Newton) [Ralston 65].

Em resumo, os resultados apresentados, ilustram as seguintes conclusões:

- A análise espectral das contagens permite obter informação semelhante à do espectro dos intervalos, com a vantagem de não estar limitada na gama de frequência, desde que na estimação espectral se utilize uma resolução temporal adequada.

  É referido na literatura que a componente correspondente à actividade respiratória poderá não estar presente se a frequência respiratória for elevada relativamente à frequência cardíaca média [Akselrod et al 81]. Com a utilização da análise espectral das contagens será sempre possível analizar aquela componente.

- As limitações impostas pelos métodos correntes para a estimação do espectro das contagens, é ultrapassada com os algoritmos desenvolvidos neste trabalho.

  Em particular, poderá utilizar-se a modelização espectral AR e fazer a identificação de componentes espectrais a partir da decomposição espectral AR do espectro das contagens, de forma inteiramente análoga aos métodos baseados na análise espectral do tacograma.

4.2 Modelização espectral AR de processos pontuais extraidos do EMG

Conforme já foi referido no capítulo 1, o processo estocástico pontual associado aos máximos locais significativos no miograma (EMG) traduz características importantes do sinal.

Em particular, a caracterização do miograma de interferência, a partir da análise espectral de processos estocásticos pontuais associados, tem demonstrado potencialidades no estudo das alterações do EMG com a fadiga e como auxiliar de classificação e diagnóstico [Jones et al 87, Jones et al 90, Lago e Jones 83, Lago et al 84]. Estas aplicações, baseadas em estimadores não paramétricos para o espectro das contagens, sugerem claramente a vantagem que poderia advir da utilização de uma modelização espectral AR.

Neste capítulo são consideradas duas situações particulares em que poderá haver vantagem na utilização da modelização espectral autoregressiva do processo estocástico pontual associado aos máximos locais significativos. Nomeadamente: caracterização da complexidade dos potenciais e obtenção de indicação sobre a frequência de acendimento das unidades motoras, com eventual monitorização da sua alteração por efeito da fadiga.

4.2.1 Caracterização da complexidade dos potenciais

Nesta secção apresentam-se alguns resultados da aplicação da modelização espectral AR do processo pontual associado ao miograma de interferência na caracterização da complexidade dos potenciais [Rocha et al 87]. Pretende-se indicar uma possível extensão, com base na modelização espectral autoregressiva do espectro das contagens, aos métodos (não paramétricos) referidos no capítulo 1.

Os registos foram obtidos sob contracção voluntária máxima, para o músculo deltoide de um indivíduo de normal, e em dois doentes, um com atrofia neurogénia e outro com distrofia muscular de Duchenne. Na figura 4.5 considera-se também um registo com potenciais miopáticos.

Os sinais foram obtidos com o sistema descrito em Apêndice, com uma frequência de amostragem de 12500 Hz. O processo pontual associado foi extraído com um factor de discriminação de picos de 50 $\mu volt$.

A função densidade de covariância foi estimada utilizando os métodos descritos, nas secções 2.3.2 e 2.3.3, e a modelização espectral AR foi realizada como descrito na secção 2.2.4.

A determinação dos parâmetros AR e do erro médio quadrático mínimo foi feita recursivamente com o algoritmo de Levinson e o número de parâmetros determinado pela minimização dos critérios de AIC(p) e CAT(p), (2.38) e (2.39), respectivamente, para $p = 1, \cdots, 50$. Não foram encontradas diferenças na aplicação dos dois critérios.

Os espectros não paramétricos foram determinados pelo método de Blackman--Tukey (expressão (2.12), janela de Bartlett, $L = 256$).

O processo estocástico pontual associado à sucessão dos máximos locais significativos no EMG registado com eléctrodos de agulha concêntrica, apresenta tipicamente uma intensidade média de 100 a 700 acontecimentos por segundo e um espectro significativamente diferente do seu valor limite para frequências até 2 a 8 KHz e próximo da intensidade média acima dessa gama [Lago e Jones 83, Lago et al 84].

Apresentam-se alguns exemplos nas figuras 4.5, 4.6 e 4.7.

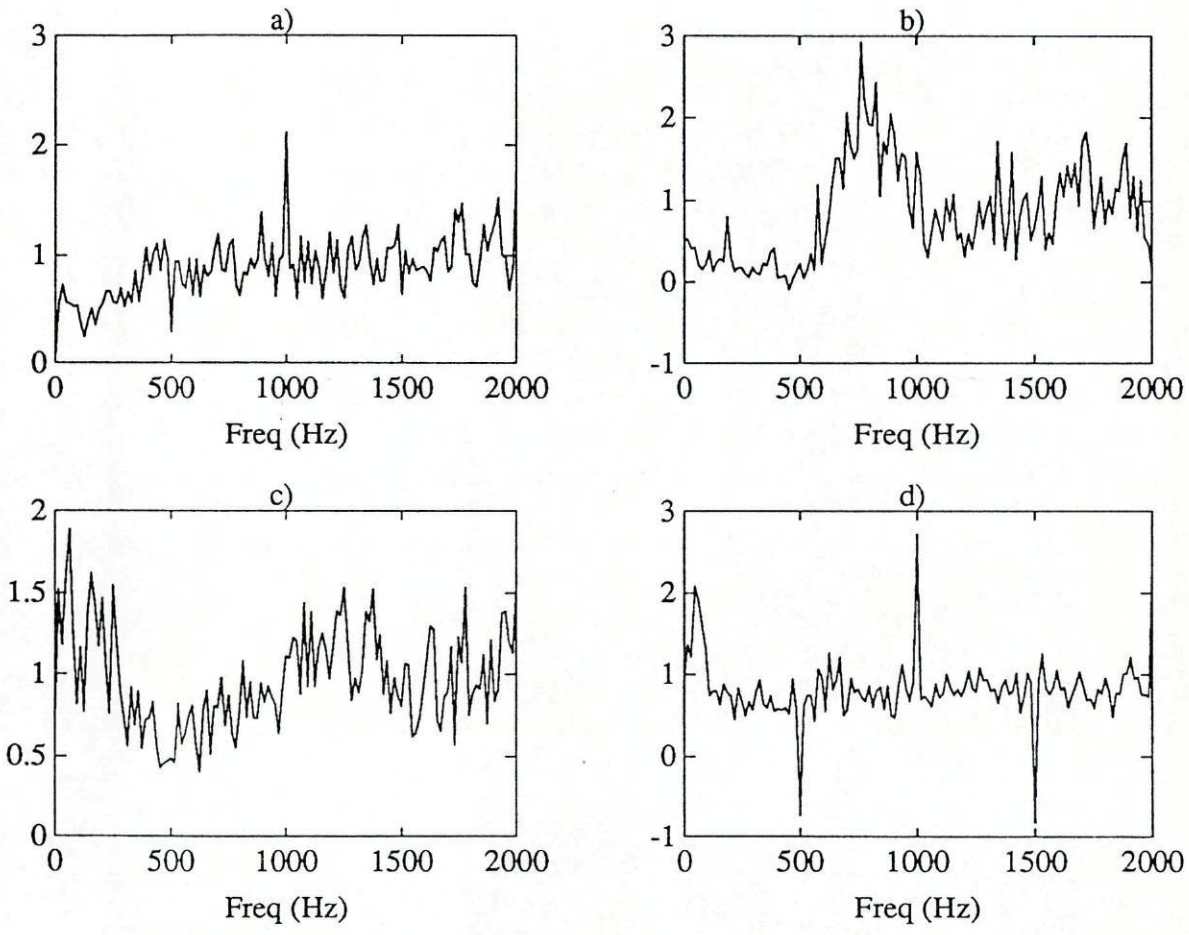


Figura 4.5: Espectro não paramétrico normalizado das contagens, $G(\omega)/\hat{\lambda}$, obtido pelo método de Blackman-Tukey, para o processo estocástico pontual associado aos máximos locais significativos no EMG. Densidade de covariância estimada a partir de (2.26) e (2.29) com correcção aditiva do erro de centricidade ($\Delta t = 0.25ms$). **a)** indivíduo normal de controlo, **b)** potenciais miopáticos, **c)** doente com atrofia neurogénia, **d)** doente com distrofia muscular de Duchenne. Registo **b)** com duração $T = 500$ ms, restantes registos com duração 800 ms.

No capítulo 2 foi referido que a estimação da função densidade de covariância a partir dos métodos referidos na secção 2.3.2 (isto é nomeadamente, a partir de (2.26) e (2.29)) pode conduzir a valores espectrais estimados negativos, não sendo portanto adequada para a modelização espectral AR do espectro das contagens.

Esta situação pode ocorrer com grande frequência no caso de dados experimentais, devido à enorme variabilidade presente, como é ilustrado na figura 4.5, verificando-se que a utilização dos estimadores semi-definidos positivos, descritos na secção 2.3.3, é essencial para a modelização espectral autoregressiva do espectro das contagens.

Apresentam-se, nas figuras 4.6 e 4.7, alguns exemplos relativos à análise espectral do processo pontual associado ao EMG, com função densidade de covariância estimada a partir de (2.33) e (2.35), respectivamente.

Para facilidade de comparação representam-se também os espectros não paramétricos correspondentes obtidos pelo método de Blackman-Tukey.

Estes exemplos e muitos outros mostram que:

- A modelização espectral AR dá consistentemente uma boa descrição do espectro das contagens do processo estocástico pontual associado aos máximos locais significativos no EMG.
- Os resultados obtidos com dados experimentais são concordantes com os apresentado no capítulo 1, com dados de simulação:
  - Como era de esperar, devido ao amaciamento inerente à modelização espectral AR, o espectro AR corresponde aproximadamente a uma curva média do espectro não paramétrico.
  - Com a ordem escolhida a partir da minimização dos critérios AIC e CAT o desempenho da modelização espectral AR é bom e semelhante quer nos picos, quer nos vales. Não foram encontradas diferenças significativas para os critérios AIC e CAT.
  - A experiência com dados experimentais também permite concluir que se obtem um melhor desempenho com o estimador baseado em (2.33) do que a partir de (2.35), conforme se ilustra nas figuras 4.6 e 4.7;

    Na parte plana do espectro das contagens (região das altas frequências) o espectro AR baseado na estimativa da densidade de covariância obtida a partir de (2.33) apresenta geralmente menos flutuações.

    Excluindo a região das altas frequências, os dois métodos conduzem a resultados idênticos.

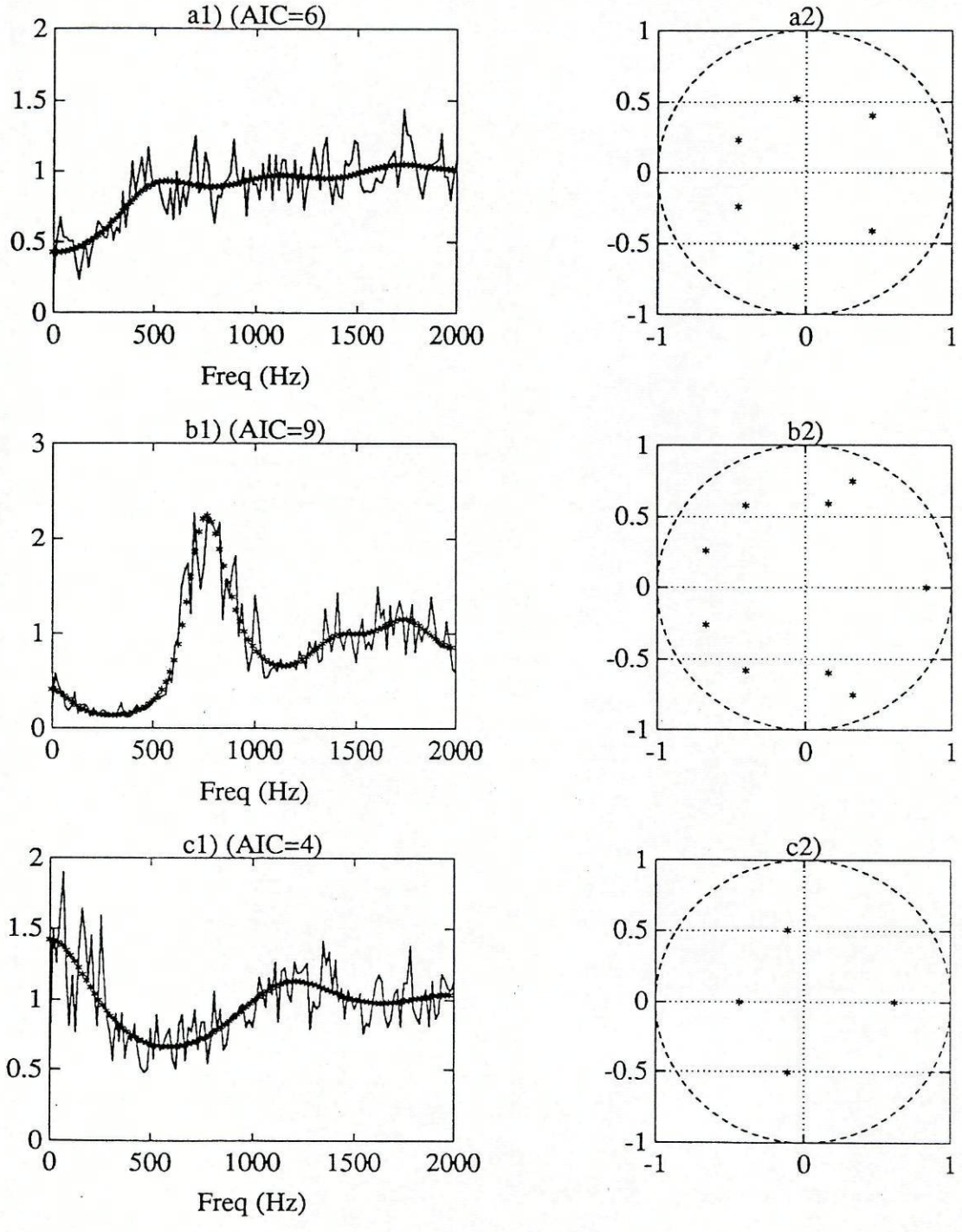


Figura 4.6: Espectro AR e não paramétrico normalizados, $G(\omega)/\hat{\lambda}$, para o processo estocástico pontual associado aos máximos locais significativos no EMG, obtidos com a densidade de covariância estimada a partir de (2.33)($\Delta t = .25ms$) e diagramas de polos correspondentes; **a)** indivíduo normal de controlo, **b)** potenciais miopáticos, **c)** doente com atrofia neurogénia. Mesmos dados da figura 4.5.

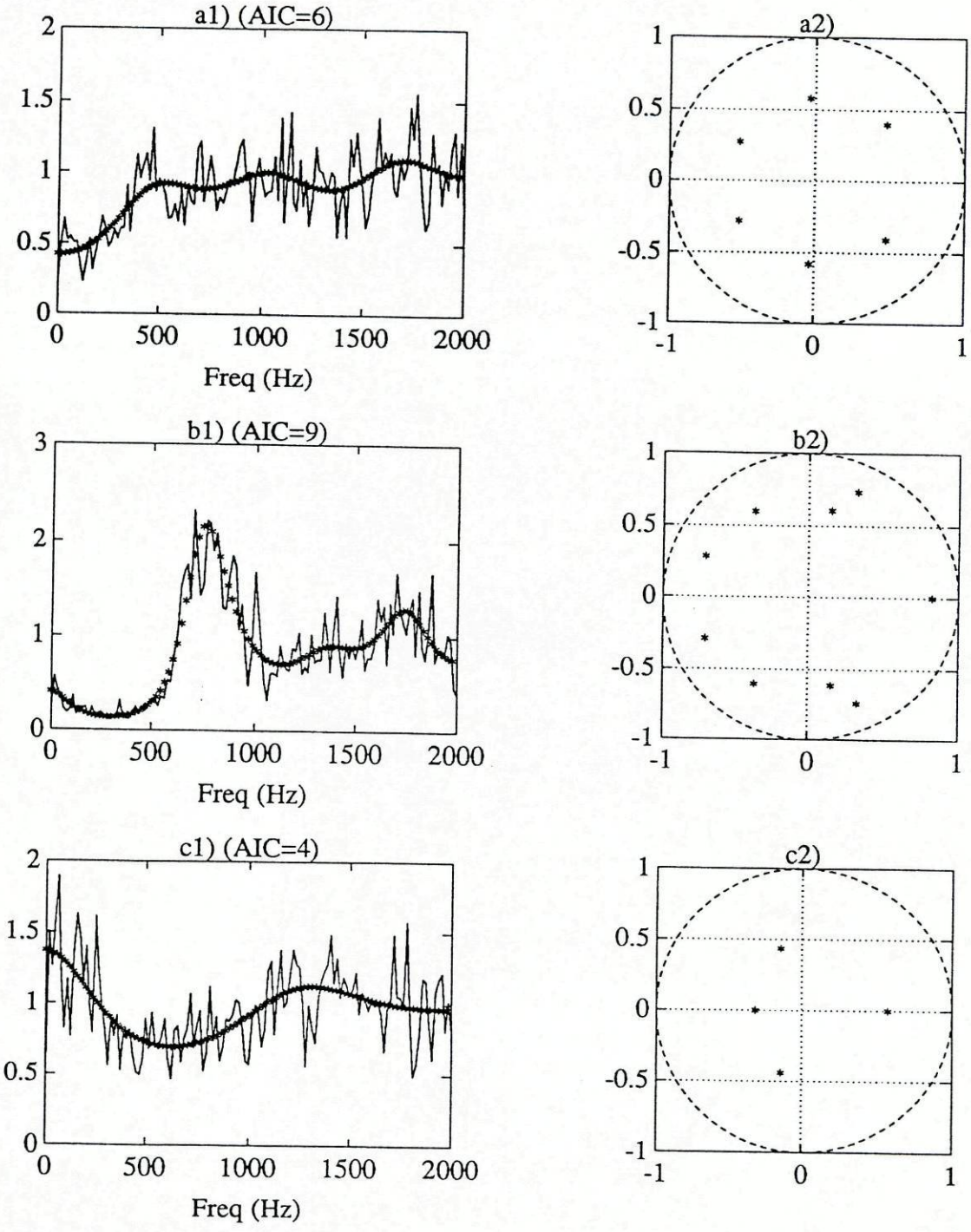


Figura 4.7: Espectro AR e não paramétrico normalizados das contagens, $G(\omega)/\hat{\lambda}$, para o processo estocástico pontual associado aos máximos locais significativos no EMG, obtidos com a densidade de covariância estimada a partir de $(2.35)(\Delta t = .25ms)$ e diagramas de polos correspondentes; **a)** indivíduo normal de controlo, **b)** potenciais miopáticos, **c)** doente com atrofia neurogénia. Mesmos dados da figura 4.5.

Um problema que se põe na prática é a escolha da resolução temporal (Δt) a empregar, de forma a garantir uma descrição espectral adequada.

Conforme se viu no capítulo anterior, o método de variância mínima permite obter informação sobre a gama de frequências onde o espectro deverá ser aproximadamente plano.

Assim, o método de variância mínima poderá ser utilizado como auxiliar na escolha e validação do valor de Δt a utilizar na análise espectral de processos pontuais.

Nos exemplos apresentados nesta secção utilizou-se uma resolução temporal de $0.25\ ms$. Conforme se ilustra na figura 4.8, considerando dois dos casos particulares ilustrados nesta secção, a resolução temporal é suficientemente alta garantindo que o espectro normalizado já é aproximadamente constante (e igual a 1) na gama $|\omega| > \pi/\Delta t$.

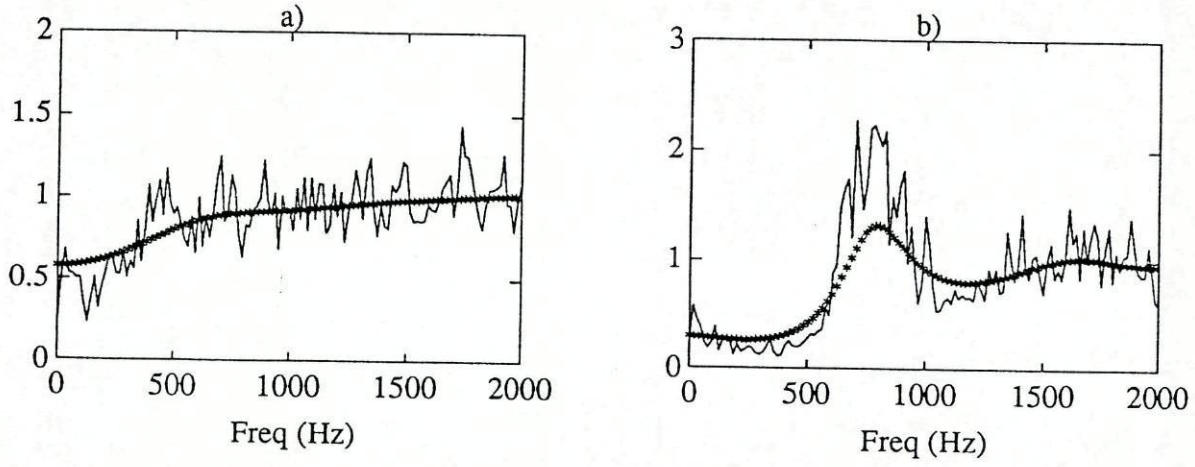


Figura 4.8: Espectro das contagens normalizado obtido a partir do método de variância mínima ($p = AIC$) para um indivíduo normal de controlo (a) e na presença de potenciais miopáticos (b). Sucessão densidade de covariância estimada a partir de (2.33), para $\Delta t = 0.25\ ms$. Mesmos dados das figuras anteriores.

Apesar da modelização espectral AR permitir obter boas descrições do espectro das contagens, existem algumas dificuldades na descrição das altas frequências, em particular para valores elevados da resolução temporal.

A descrição na região das altas frequências pode ser melhorada, utilizando os algoritmos para aperfeiçoamento da modelização espectral AR descritos no capítulo 3, conforme se ilustra na figura 4.9, considerando os casos particulares representados na figura 4.8. Os modelos espectrais melhorados (obtidos a partir de $\omega_{lim} = 1000$ Hz em a) e $\omega_{lim} = 1500$ Hz em b)) apresentam um comportamento mais adequado nas altas frequências do que os modelos espectrais iniciais correspondentes, representados na figura 4.5.

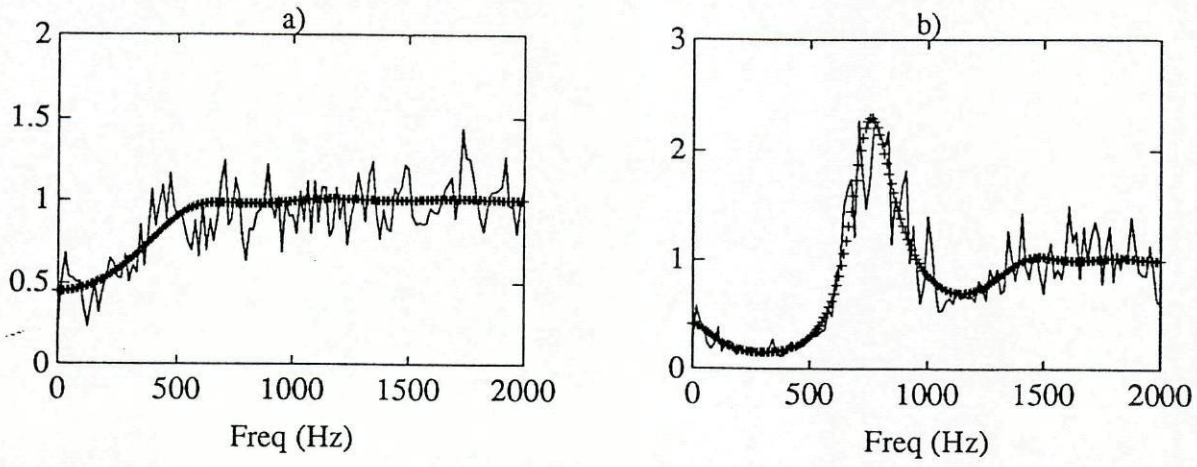


Figura 4.9: Modelo espectral melhorado ($p = AIC$) para um indivíduo normal de controlo (**a**) e na presença de potenciais miopáticos (**b**). Sucessão densidade de covariância estimada a partir de (2.33), para $\Delta t = 0.25\, ms$. Mesmos dados das figuras anteriores.

Saliente-se finalmente que para além da modelização espectral AR dar consistentemente uma boa descrição do espectro das contagens, permite uma *redução bastante significativa* nos dados.

A informação contida no espectro não paramétrico é sintetizada de forma simples pelos parâmetros do modelo AR, quer por outras descrições equivalentes, nomeadamente diagrama de polos ou coeficientes de reflexão.

Na escolha de parâmetros para a descrição e quantificação dos modelos espectrais AR, tendo em vista aplicações em classificação ou diagnóstico, será desejável que haja uma ordenação natural desses parâmetros.

Os parâmetros AR, embora naturalmente ordenados, dependem da ordem do modelo espectral, o que conduziria à utilização de ordens fixas. Este aspecto seria bastante limitativo, em particular na caracterização da complexidade do EMG: a ordem do modelo espectral AR, depende fortemente das características do espectro das contagens e da resolução temporal, Δt, utilizada.

Estas observações sugerem que a descrição poderá ser feita com vantagem a partir dos *coeficientes de reflexão*, isoladamente ou combinados com outros parâmetros.

Para além de poderem dar indicação da forma do espectro (ver figura 4.10), e de corresponderem a uma ordenação natural, os valores de k_i, $i < p$, não se alteram quando p aumenta.

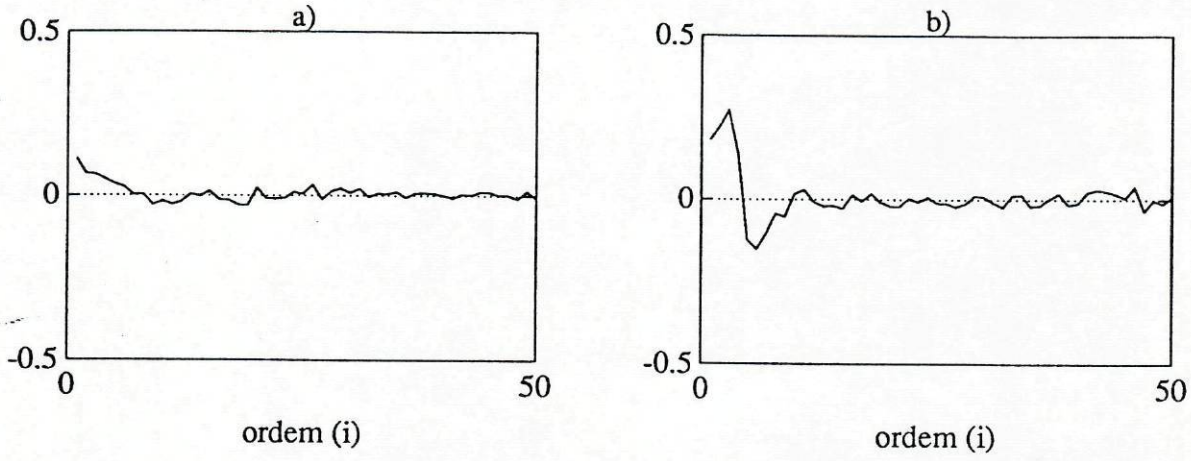


Figura 4.10: Sequência dos coeficientes de reflexão $\{k_i\}_{i=1,\cdots,50}$ correspondentes à modelização espectral AR do processo pontual extraido do EMG de interferência para um indivíduo normal de controlo (**a**) e na presença de potenciais miopáticos (**b**). Sucessão densidade de covariância estimada a partir de (2.33), para $\Delta t = 0.25\ ms$. Mesmos dados das figuras anteriores.

Não é objecto de estudo, neste trabalho, o aperfeiçoamento do método de caracterização da complexidade dos potenciais no miograma de interferência apresentado no capítulo 1.

No entanto, os resultados preliminares descritos nesta secção, sugerem claramente que a modelização espectral AR do espectro das contagens poderá apresentar vantagens, sobre os métodos não paramétricos na caracterização da complexidade dos potenciais.

A redução bastante significativa nos dados poderá, nomeadamente, ser útil no estabelecimento de padrões de normalidade.

Em particular poderá permitir a extensão dos métodos de classificação e diagnóstico baseados na análise espectral de processos estocásticos pontuais associados ao miograma de interferência [Jones et al 87, Jones et al 90, Lago e Jones 83, Lago et al 84].

4.2.2 Obtenção da frequência média de acendimento das unidades motoras no EMG e sua alteração com a fadiga

A determinação das estatísticas de acendimento das unidades motoras baseia-se na possibilidade de obter registos electromiográficos altamente selectivos (com um pequeno número de unidades motoras) e a separação desses registos em sequências de potenciais associados a cada uma das unidades, recorrendo a diversas técnicas para decomposição e análise automática do miograma de interferência [Basmajian e De Luca 85, Desmedt 89, e referências aí incluidas].

Este processo apresenta algumas limitações e, em particular, é muito pesado em tempo de cálculo.

A análise espectral do EMG pode ser utilizada, para a extracção de informação útil, e de uma forma simples, sobre a frequência de acendimento das unidades motoras [Lago e Jones 77, Lago e Jones 81, VanBoxtel e Schomaker 83, Paiss e Inbar 87, Blinowska e Verroust 87, Pan et al 89].

Na figura 4.11, representa-se em a) um espectro do EMG, típico, obtido com eléctrodos de superfície. Verifica-se a existência de uma componente significativa na região de frequência de 20 Hz, a qual é susceptível de interpretação simples.

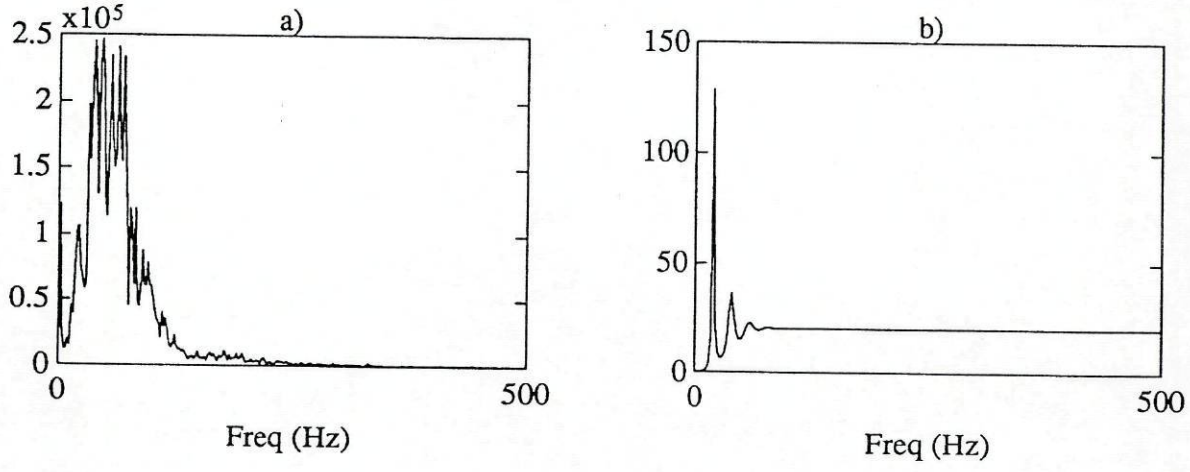


Figura 4.11: a) Densidade espectral do EMG típica para eléctrodos de superfície. b) Espectro das contagens correspondente a um processo pontual renewal Gausseano ($\mu = 50$ms, $\sigma = 6.3$ms), $G(\omega)$ calculado a partir de (1.7), $F(j\omega) = e^{-j\mu\omega+(\sigma j\omega)^2/2)}$.

O EMG pode ser interpretado em termos de um modelo simplificado, considerando na expressão (1.1) um potencial médio, $h(t)$, e supondo os processos pontuais associados aos instantes de acendimento independentes e com características idênticas. Conclui-se assim que, $\phi(\omega) \sim |H(\omega)|^2\phi_{\delta\delta}(\omega), w \neq 0$, onde $H(\omega)$ designa a transformada de Fourier de $h(t)$ e $\phi_{\delta\delta}(\omega)$ a função de densidade espectral do processo pontual associado aos instantes de acendimento de uma unidade motora.

Num grande número de situações, as estatísticas de acendimento podem ser descritas por processos renewal com distribuição próxima da Gausseana para a qual a função $\phi_{\delta\delta}(\omega)$ tem a forma representada na figura 4.11 em b) [Clamann 69, Lago e Jones 77].

O primeiro pico na figura 4.11, em b), corresponde à frequência média de acendimento ($\lambda = \frac{1}{\mu}$) e pode ser observado na função de densidade espectral do EMG ($\lambda \simeq 20$Hz).

A aplicação da análise espectral do EMG na obtenção da frequência média de acendimento das unidades motoras, apresenta contudo algumas limitações. Uma das principais limitações consiste no facto de ser geralmente necessário utilizar registos longos (com duração de dezenas de segundos) e segmentação [Paiss e Inbar 87].

Pretende-se mostrar nesta secção, que com a modelização espectral AR de processos estocásticos pontuais extraidos do EMG, é possível extrair essa informação para registos curtos.

Pretende-se também mostrar que o método é suficientemente selectivo para permitir a detecção das alterações da frequência de acendimento na presença de fadiga: quando é mantida uma contracção muscular constante, a frequência de acendimento diminui independentemente da força produzida [Basmajian e De Luca 85].

Na recolha dos dados, efectuada com eléctrodos de agulha, foi utilizado um protocolo semelhante ao referido em [Paiss e Inbar 87] (para eléctrodos de superfície).

O sinal foi registado no biceps brachii, durante 6 segundos e a uma frequência de amostragem de 5000 Hz, com o antebraço estendido e um peso de 2 Kg na palma da mão, numa situação de contracção isométrica e isotónica.

Para estudar a alteração na frequência média de acendimento foram realizados dois registos, nas condições acima referidas, antes e imediatamente depois de uma contracção voluntária máxima.

Na figura 4.12 representa-se o espectro das contagens do processo pontual associado aos máximos locais significativos (obtido para um factor de discriminação de picos de 50 $\mu volt$), considerando um registo realizado antes de contracção voluntária máxima (registo inicial) e imediatamente depois de contracção voluntária máxima (registo final), a) e b), respectivamente.

A análise da gama das baixas frequências permite identificar no espectro não paramétrico do registo inicial (figura 4.12 a)) uma componente que poderá estar associada à frequência de acendimento das unidades motoras (frequência $\simeq 20Hz$).

No registo final (figura 4.12 b)) observa-se a mesma componente, com um deslocamento para frequências mais baixas, o que poderá ser explicado pelo efeito da fadiga.

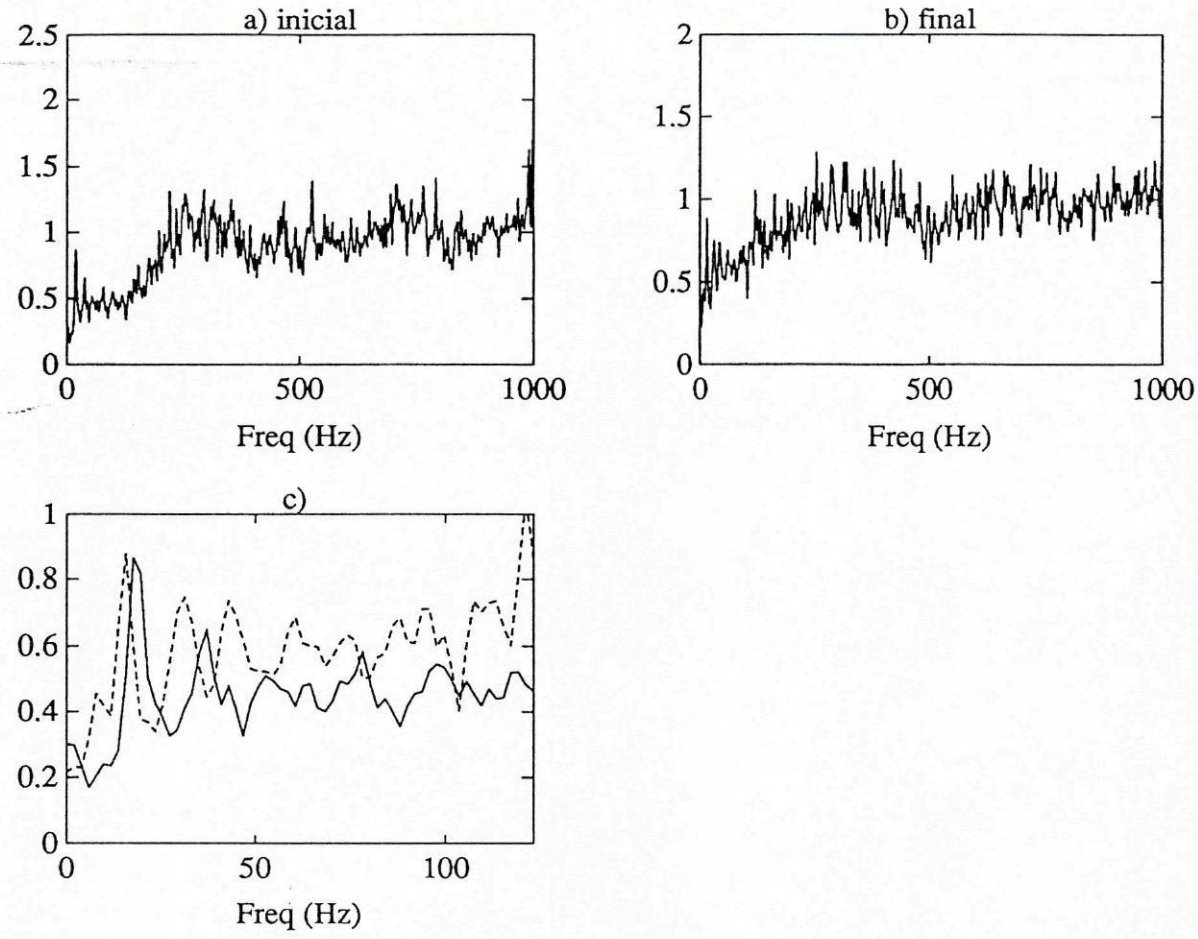


Figura 4.12: Espectros não paramétrico normalizado das contagens para o processo pontual associado aos máximos locais significativos no EMG, obtidos com o método de Blackman--Tuckey a partir da densidade de covariância estimada por (2.33) para $\Delta t = 0.5ms$. **a)** registo inicial (antes de contracção voluntária máxima), **b)** registo final (imediatamente após contracção voluntária máxima). Em **c)** representam-se sobrepostos o espectros inicial (-) e final (- -) na gama das baixas frequências.

A modelização espectral AR poderá revelar-se particularmente útil na identificação da componente associada às estatísticas de acendimento e às alterações na presença de fadiga, conforme será analizado em seguida.

Na figura 4.13 representam-se em **a)** e **b)** o espectros não paramétricos e o modelo espectral AR ($p = AIC$) das contagens, para o processo pontual associado aos máximos locais significativos relativos aos registos inicial e final.

Conforme era de esperar, pelos resultados da secção anterior, obtêm-se boas descrições para o espectro não paramétrico, com a ordem indicada pelos critérios de AIC e CAT. Contudo, não é possível identificar qualquer componente associada à frequência de acendimento das unidades motoras. Essa informação só é patente se se utilizar um modelo com uma ordem muito elevada ($p \simeq 70, 100$) de forma a descrever o detalhe na região das baixas frequências.

Alternativamente poderá utilizar-se a modelização espectral selectiva da gama das baixas frequências [Makhoul 75].

De forma análoga a processos estocásticos em geral, a aplicação deste método consiste na tradução da porção do espectro a modelar $[\omega_\alpha, \omega_\beta]$ na gama fundamental $[0, \pi/\Delta t]$, a partir da mudança de variável

$$\omega' = \frac{\pi}{\Delta t} \frac{\omega - \omega_\alpha}{\omega_\beta - \omega_\alpha},$$

com as frequências negativas definidas da forma habitual, por simetria.

Os parâmetros do modelo AR relativo à gama seleccionada são obtidos a partir de (2.16) e (2.17) para a sucessão densidade de covariância correspondente ao espectro a modelar (calculada pela transformada inversa de Fourier).

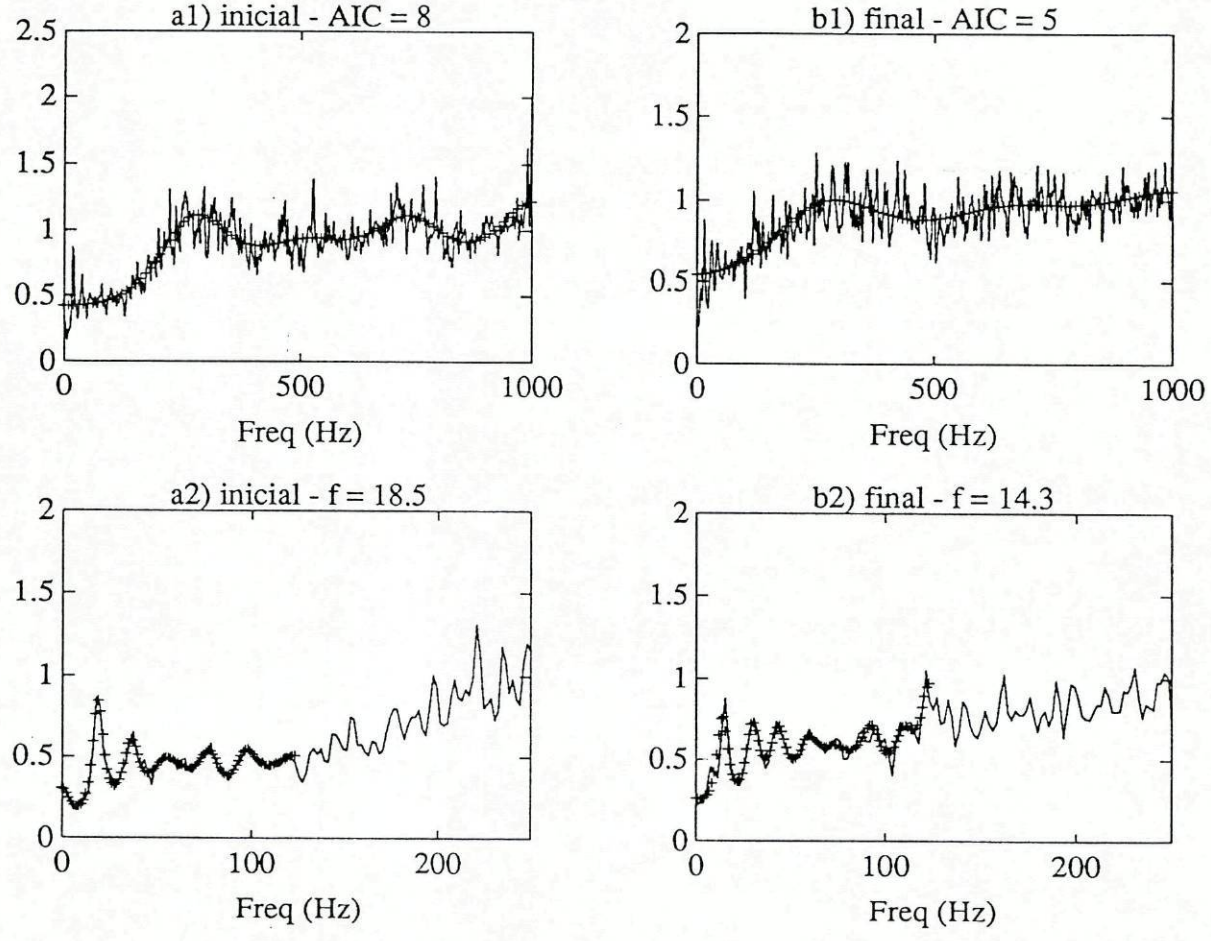


Figura 4.13: Modelização espectral AR do espectro das contagens do processo pontual associado aos máximos locais significativos no EMG, considerando **1)** a gama total de frequências e $p = AIC$ e **2)** modelização selectiva da gama $0 - 120$ Hz ($p = 17$), para **a)** registo inicial (antes de contracção voluntária máxima), **b)** registo final (imediatamente após uma contracção voluntária máxima). Em **a2)** e **b2)** indica-se também a frequência de acendimento identificada pelo argumento dos polos no modelo espectral AR correspondente (f = 18.5 e 14.3 Hz, respectivamente).

Verifica-se que a modelização selectiva da gama de baixas, permite diminuir bastante a ordem do modelo, mantendo a descrição de todo o detalhe, quer no registo inicial, quer no final.

Nomeadamente na figura 4.13 representam-se em **b)** e **c)**, os modelos espectrais selectivos de ordem $p = 17$, para a gama 0 a 120 Hz.

Conclui-se assim que a modelização selectiva da gama de baixas frequências do espectro das contagens permite, para modelos de ordem relativamente baixa, a identificação da componente associada à frequência de acendimento.

Neste caso, a frequência de acendimento poderá ser facilmente caracterizada a partir dos argumentos dos polos no modelo espectral AR correspondente (frequências 18.5 e 14.3 Hz).

Este método, poderá mesmo revelar-se bastante útil no estudo das alterações com a fadiga, conforme se ilustra na figura 4.14.

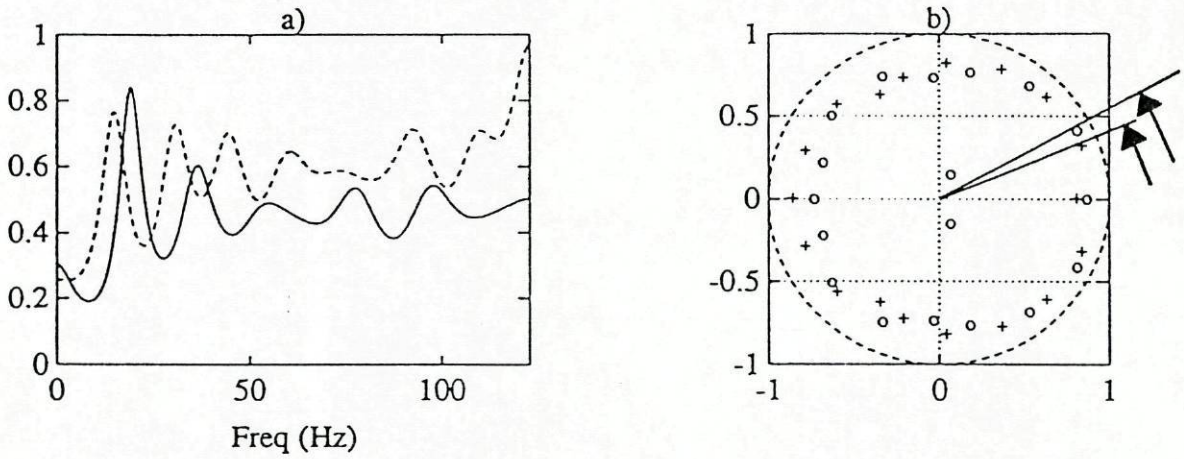


Figura 4.14: Modelos espectrais AR selectivos representados na figura 4.13 (zona das baixas frequências) para o registo inicial (-) e final (- -) e diagramas de polos correspondentes, respectivamente, (o) e (+).

Em resumo, verifica-se que:

- A análise espectral de processo estocástico pontual associado aos máximos locais significativos do EMG permite a identificação da frequência de acendimento e o estudo das suas alterações com a fadiga, para registos curtos.
- A utilização da modelização espectral selectiva, possibilita a identificação da componente associada à frequência de acendimento de uma forma simples a partir dos argumentos dos polos correspondentes.

  O método poderá revelar-se particularmente útil no estudo das alterações da frequência média de acendimento com a fadiga.

Capítulo 5

Conclusões Finais

Neste trabalho discute-se a modelização espectral autoregressiva (AR) do espectro das contagens de um processo estocástico pontual.

A forma de abordagem adoptada explora o paralelismo entre os conceitos e operações envolvidas na análise espectral de processos estocásticos e de processos estocásticos pontuais.

Este ponto de vista permite formular diversos problemas de estimação em processos pontuais, em particular a estimação da densidade de covariância e do espectro das contagens, na perspectiva da análise de séries temporais. Assim, possibilita o desenvolvimento de algoritmos eficientes para a análise de processos pontuais em clara correspondência com os métodos correntes, e amplamente divulgados, na literatura de processos estocásticos em geral.

Em particular, no que respeita à estimação do espectro das contagens, além dos métodos não paramétricos "clássicos" (periodograma e método de Blackman-Tukey) permite o desenvolvimento de métodos paramétricos alternativos para a análise espectral de processos estocásticos pontuais.

Neste trabalho, discute-se detalhadamente a *modelização espectral AR do espectro das contagens* baseada em estimadores semi-definidos positivos para a sucessão densidade de covariância de processos pontuais. É dada especial atenção ao *aperfeiçoamento da fidelidade* na modelização espectral AR, por inclusão de informação relativa às características do espectro das contagens.

Um objectivo importante deste trabalho consiste no estudo comparativo de diversos métodos (paramétricos e não paramétricos) para a estimação do espectro das contagens.

Como é habitual, a aplicação de técnicas de análise espectral a partir de uma amostra finita envolve, na prática, a necessidade de satisfazer variados compromissos, sendo fundamental atender às características que se pretendem avaliar. Também, de forma análoga ao caso de processos estocásticos em geral, a análise estatística dos estimadores espectrais envolve hipóteses restritivas acerca do processo e geralmente apenas aborda as distribuições assimptóticas.

Para ilustrar as características dos métodos de estimação apresentados foram realizados estudos de simulação muito extensos, cujos resultados poderão auxiliar na escolha do método a utilizar em cada situação.

Na avaliação do desempenho de um estimador espectral é essencial definir as características que se pretendem avaliar.

Neste trabalho a preocupação fundamental refere-se ao estudo da potência dos métodos e do desempenho dos algoritmos relativamente à fidelidade da descrição global. A fidelidade da descrição global da forma do espectro das contagens, garante a descrição das suas características fundamentais, e é particularmente adequada para as aplicações estudadas.

A figura 5.1 resume (no caso particular de um processo renewal Gausseano) o tipo de resultados que se obtêm com os métodos descritos neste trabalho, nomeadamente: análise espectral não paramétrica (Blackman-Tukey), modelização espectral AR e o seu aperfeiçoamento com base na introdução de informação sobre as características do espectro. Para facilidade de comparação representa-se também o valor teórico correspondente.

Conforme se ilustra na figura:

- A aplicação do Blackman-Tukey para processos estocásticos pontuais, baseado em estimadores semi-definidos positivos para a sucessão densidade de covariância, garante a obtenção de estimadores espectrais consistentes e não negativos.

- A modelização espectral AR de processos estocásticos pontuais, com base em estimadores semi-definidos positivos para a sucessão densidade de covariância e com a ordem seleccionada pelos critérios de AIC e CAT, conduz a boas descrições do espectro das contagens.

  A qualidade da descrição espectral é uniforme em picos e vales, correspondendo aproximadamente a uma curva média do espectro não paramétrico. Existem algumas dificuldades na descrição adequada da gama das altas frequências, onde o espectro deverá ser aproximadamente plano.

- O aperfeiçoamento da modelização espectral AR, a partir do conhecimento da gama onde o espectro das contagens deverá ser aproximadamente plano, conduz a descrições melhoradas do espectro das contagens, com um comportamento adequado na zona das altas frequências. O método de variância mínima permite obter informação sobre a gama de frequências onde o espectro deverá ser aproximadamente plano.

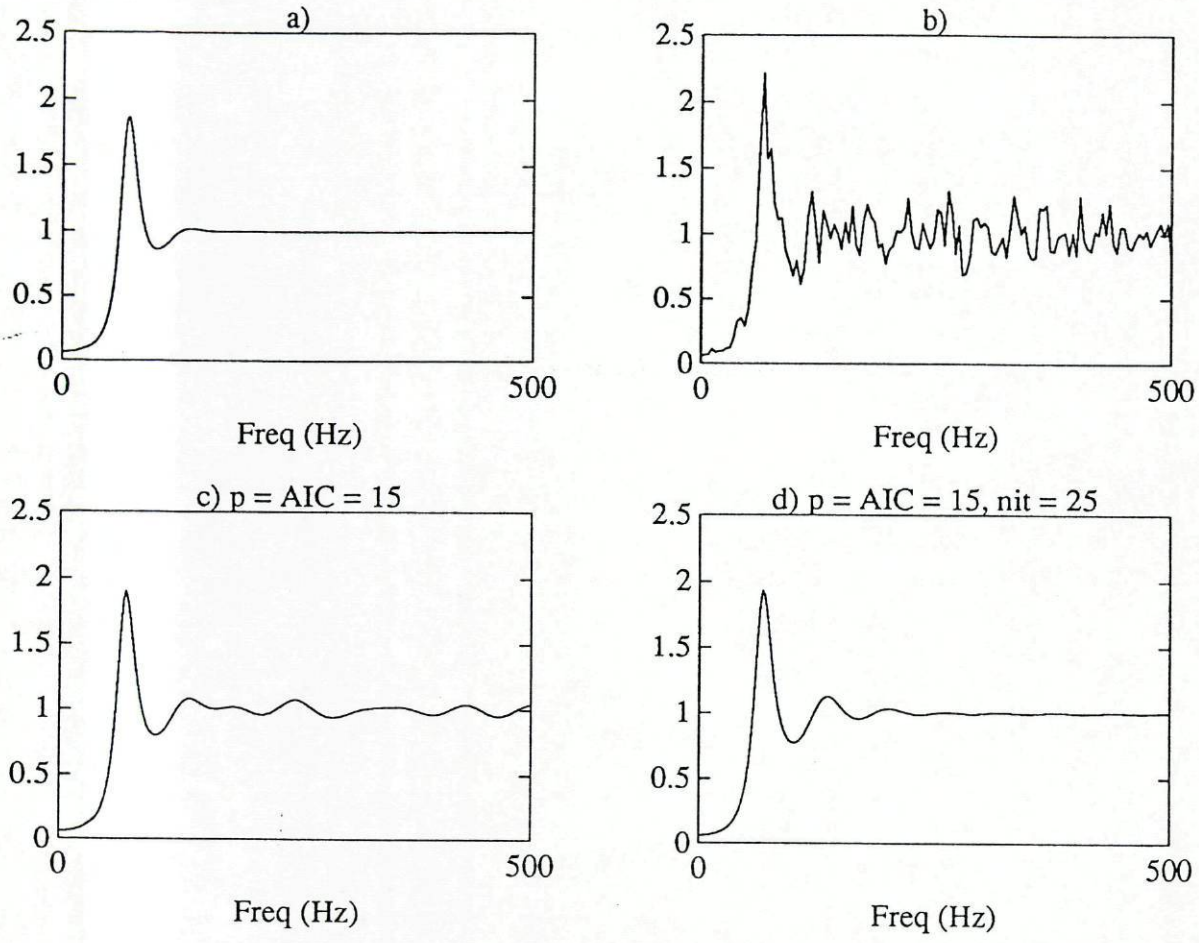


Figura 5.1: Estimação do espectro das contagens de um processo pontual renewal Gausseano (normalização pela intensidade média, $\hat{\lambda}$): **a)** valor teórico, **b)** espectro não paramétrico (método de Blackman-Tukey), **c)** e **d)** modelos espectrais AR, respectivamente, modelo inicial para ordem $p = AIC$ e correspondente modelo aperfeiçoado.

Uma das principais dificuldades da modelização espectral AR consiste na escolha da ordem, p. Uma escolha desajustada para o valor da ordem pode conduzir a um mau desempenho espectral, reflectindo-se num grande erro de centricidade ou em demasiada variabilidade.

Uma das conclusões deste trabalho é que a escolha da ordem é bastante crítica e que o critério de AIC conduz geralmente a valores conservativos para a ordem a utilizar.

Os resultados apresentados neste trabalho, sugerem claramente o desenvolvimento de critérios para a escolha da ordem, baseados num modelo espectral melhorado de ordem elevada.

Finalmente será de salientar que os métodos e algoritmos descritos neste trabalho se baseiam numa estimação directa da sucessão densidade de covariância em alternativa aos métodos descritos na literatura com base na função de intensidade condicional. Os estimadores desenvolvidos garantem que a sucessão densidade de covariância estimada é semi-definida positiva. A utilização destes estimadores revela-se fundamental, uma vez que garantem que o espectro correspondente, determinado a partir de (2.9) é sempre não negativo.

Em particular, no que respeita à modelização espectral AR do espectro das contagens e ao seu aperfeiçoamento, verifica-se que:

- Para a medida de erro inicial (formulação 1), a utilização de sucessões densidade de covariância semi-definidas positivas possibilita a aplicação da modelização espectral AR com os critérios correntes para a escolha da ordem.

- No caso da medida de erro modificada, é também essencial a preservação daquela propriedade. Permite o desenvolvimento de algoritmos iterativos convergentes, uma vez que as matrizes densidade de covariância associadas são garantidamente semi-definidas positivas. Possibilita também a obtenção de sucessões densidade de covariância modificadas, que preservam aquela propriedade, e são mais adequadas para a descrição da parte plana do espectro das contagens.

- A utilização dos estimadores semi-definidos positivos para a sucessão densidade de covariância permite também a aplicação do método de variância mínima (nomeadamente para a obtenção da gama de frequências onde o espectro é plano e na escolha e validação da frequência de amostragem a utilizar). A matriz densidade de covariância associada é garantidamente semi-definida positiva e portanto não singular.

A aplicação dos métodos e algoritmos desenvolvidos é ilustrada com sinais biomédicos, considerando dois casos particulares: caracterização do miograma de interferência (EMG) e da variabilidade de sinais cardiovasculares.

Os resultados obtidos sugerem que a modelização espectral AR do processo das contagens poderá permitir o desenvolvimento de métodos alternativos da descrição do sinal nas duas situações estudadas.

Apêndice A

Sistema de aquisição e processamento do EMG

Foi desenvolvido um sistema para aquisição, armazenamento e quantificação do EMG, suficientemente flexível para a implementação dos métodos de caracterização do sinal correntemente referidos na literatura e de forma a permitir a investigação de novos métodos e variantes.

O sistema, descrito sucintamente neste apêndice, foi desenvolvido no âmbito de um projecto de investigação em colaboração com o Serviço de Neurofisiologia Clínica do Hospital Geral de Santo António, no Porto.

O sistema baseia-se num computador PDP 11/23 de 256 Kbytes de memória, conforme se descreve esquemáticamente na figura A.1. Além das unidades periféricas usuais (unidades de disco, diskette, impressora e terminal) o sistema dispõe de conversores A/D de 12 bits, portas de entrada/saída digital, relógio programável e facilidades gráficas (traçador de gráficos e terminal gráfico interactivo).

As frequências de amostragem necessárias para algumas aplicações em electromiografia, nomeadamente para os métodos baseados em contagens de máximos locais tornam necessária a utilização de frequências de amostragem elevadas.

O sistema desenvolvido permite a aquisição e registo de 32000 amostras à frequência máxima de 25000 amostras por segundo. O programa de aquisição é interactivo, permite a inclusão de identificadores nos ficheiros de dados, sendo a aquisição desencadeada pelo clínico por um interruptor de pedal. Dispõe-se ainda de equipamento para gravação analógica.

O sistema engloba um programa interactivo que permite a escolha do segmento do EMG a processar mediante a introdução de marcas. Foram implementados métodos correntes para o tratamento do EMG: nomeadamente, a análise espectral directa e métodos baseados na contagem dos máximos locais significativos e passagens por zero. Além desses métodos também se encontram implementados métodos baseados na análise espectral de processos estocásticos pontuais associados ao EMG.

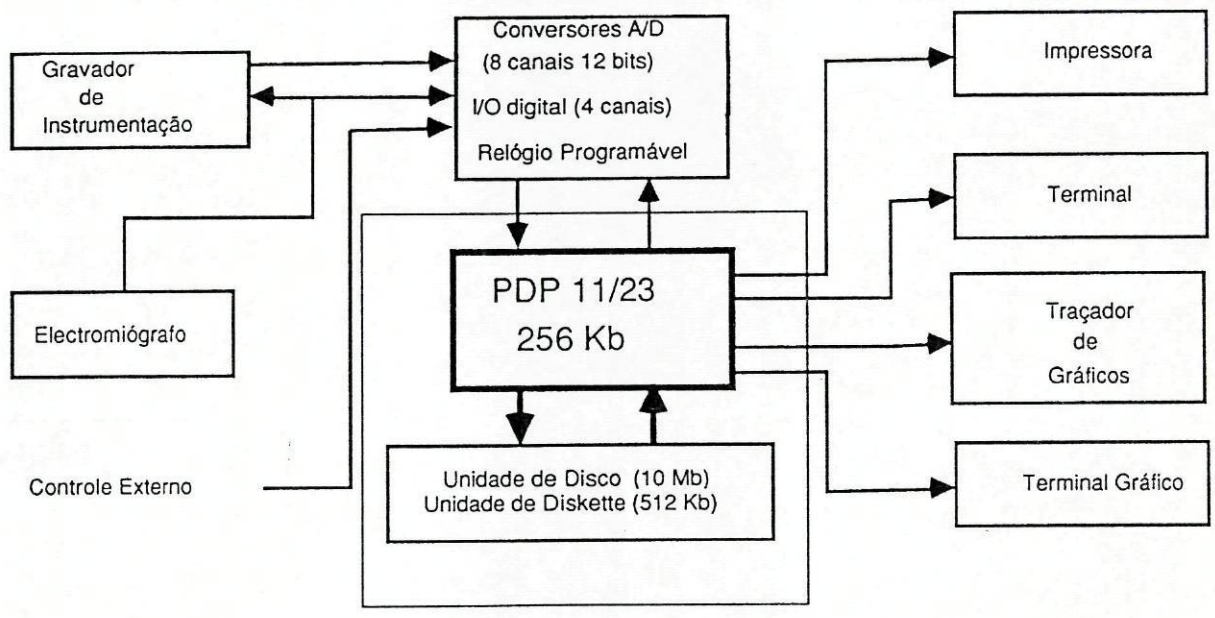


Figura A.1: Sistema de aquisição e processamento do EMG.

Referências

[Agarwal e Gottlieb 82] Agarwal, G.C., Gottlieb, C.L. (1982), "Mathematical modelling and simulation of the postural control loop: part I", *CRC Critical Reviews in Biomedical Engineering*, **8**, 93–134.

[Akselrod et al 81] Akselrod, S., Gordon, D., Ubel, F.A., Shannon, D.C., Barger, A.C., Cohen, R.J. (1981), "Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control", *Science*, **213**, 220–222.

[Akselrod et al 85] Akselrod, S., Gordon, D., Jeffrey, B.M., Snidman, N.C., Shannon, D.C., Cohen, R.J. (1985), "Hemodynamic regulation: investigation by spectral analysis", *Am. Phys. Society*, **H**, 867–875.

[Akaike 74] Akaike, A. (1974), "A new look at the statistical model identification", *IEEE Trans. Automatic Control*, **19**, 716–723.

[Appel et al 89] Appel, M.L., Berger, R.D., Saul, J.P., Smith, J.M., Cohen, R.J. (1989), "Beat to beat variability in cardiovascular variables: Noise or Music?", *J. Am. Coll. Cardiol.*, **14**, 1139–1148.

[Baselli et al 86] Baselli, G., Cerutti, S., Civardi, S., Liberati, D., Lombardi, F., Malliani, A., Pagani, M. (1986), "Spectral and cross-spectral analysis of heart rate and arterial blood pressure variability signals", *Computers and Biomedical Research*, **19** , 520–534.

[Basmajian 78] Basmajian, J.V. (1978), *Muscles alive. Their functions revealed by electromyography* (4th edition), Williams and Wilkins, Baltimore.

[Basmajian e De Luca 85] Basmajian, J.V., De Luca, C.J. (1985), *Muscles alive. Their functions revealed by electromyography* (5th edition), Williams and Wilkins, Baltimore.

[Blinowska e Verroust 87] Blinowska, A., Verroust, J. (1987), "Low frequency power spectrum of the EMG signal", *Electromyography Clin. Neurophysiol.*, **27**, 349–353.

[Brillinger 75] Brillinger, D.R. (1975), "Statistical inference for stationary point processes", *Stochastic processes and related topics*, Puri, M.L. (ed.), Academic Press, New York.

[Brillinger 78] Brillinger, D.R. (1978), "Comparative aspects of the study of ordinary time series and of point processes", *Developments in statistics*, vol. 1, Krishnaiah (ed.), Academic Press.

[Burg 72] Burg, J.P. (1972), "The relation between maximum entropy spectra and maximum likelihood spectra", *Geophysics*, **37**, 375–376.

[Cadzow 82] Cadzow, J.A. (1982), "Spectral estimation: An overdetermined rational model equation approach", *Proc. of the IEEE*, **70**, 907–939.

[Carvalho et al 91] Carvalho, M.J., Costa, O., Freitas, J., Puig, J., Coelho, T., Rocha, A., Lago, P., Man in't Veld, Falcão de Freitas (1991), "Assessement of autonomic function by spectral analysis of heart rate", *Proceedings of 11th Congress of the European Association of Internal Medicine*, Novembro 1991, Lisboa.

[Clamann 69] Clamann, H.P. (1969), "Statistical analysis of motor unit firing patterns in human squeletal muscle", *Biophys. J.*, 9, 1233–1251.

[Cohen 86] Cohen, A.N. (1986), *Biomedical signal processing*, CRC Press, New York.

[Costa et al 92] Costa, O., Lago, P., Rocha, A.P., Carvalho, M.J., Freitas, J., Puig, J., Brandão, A., Falcão de Freitas (1992), "Circadian variation of heart rate variability: A comparative study between short term and long term spectral analysis", *XIV Congress of the European Society of Cardiology*, 30 Agosto - 3 Setembro 1992, Barcelona.

[Costa et al 93] Costa, O., Lago, P., Rocha, A.P., Freitas, A., Puig, J., Carvalho, M.J., Falcão de Freitas (1993), "Heart rate variability: A comparative study between non parametric and parametric spectral analysis", aceite para apresentação em *Société Française de Cardiologie, III Journées Européennes* , Janeiro 1993, Paris.

[Cox e Isham 80] Cox, D.R., Isham, V. (1980), *Point processes*, Chapman and Hall.

[Cox e Lewis 66] Cox, D.R., Lewis, P.A.W. (1966), *The statistical analysis of series of events*, Chapman and Hall.

[DeBoer et al 84] DeBoer, R.W., Karemaker, J.M., Strackee, J. (1984), "Comparing spectra of a series of point events particularly for heart rate variability data", *IEEE Trans. on Biomedical Engineering*, **31**, 385–387.

[De Luca 79] De Luca, C.J. (1979), "Physiology and mathematics of myoelectric signals", *IEEE Trans. Biomedical Engineering*, **26**, 313–325.

[Desmedt 83] Desmedt, J.E. (ed.) (1983), *Computer-aided electromyography*, Karger.

[Desmedt 89] Desmedt, J.E. (ed.) (1989), *Computer-aided electromyography and expert systems*, Elsevier.

[El-Jaroudi e Makhoul 91] El-Jaroudi, A., Makhoul, J. (1991), "Discrete all-pole modelling", *IEEE Trans. on Signal Processing*, **39**, 411–423.

[French e Holden 70] French, S.A., Holden, A.V. (1970), "Alias free sampling of neuronal spike trains", *Kybernetic*, **8**, 165–175.

[Gardner 88] Gardner, W.A. (1988), *Statistical spectral analysis: A nonprobabilistic theory*, Prentice-Hall.

[Itakura e Saito 68] Itakura, F., Saito, S. (1968), "Analysis synthesis telephony based on the maximum likelihood method", *Proc. Sixth Intern. Congr. Acoust.*, Tokyo, C17–20.

[Jenkins e Watts 68] Jenkins, J.M., Watts, D.G. (1968), *Spectral analysis and its applications*, Holden-Day.

[Jones e Lago 77] Jones, N.B, Lago, P.J. (1977), "Factors affecting the interpretation of EMG spectra", *IEE Conf. Public.* **159**, 65–81.

[Jones et al 87] Jones, N.B., Lago, P.J., Parekh, A. (1987), "Principal component analysis of the spectra of point processes — an application to electromyography", *Mathematics in signal processing*, Durrani, T.S. et al (ed.s), Clarendon Press, 147–164.

[Jones et al 90] Jones, N.B., Lago, P.J.A., Parekh, A. (1990), "Classification of point processes using principal components analysis", *Mathematics in signal processing*, McWhirter, R.G. (ed.), Clarendon Press, 369–384.

[Kay 88] Kay, S.M. (1988), *Modern spectral estimation: Theory and application*, Prentice-Hall.

[Koopmans 74] Koopmans, L.H. (1974), *The spectral analysis of time series*, Academic Press.

[Lago e Jones 77] Lago, P.J., Jones, N.B. (1977), "Effect of motor unit firing time statistics on EMG spectra", *Med. & Biol. Eng. & Comput.*, **15**, 648–655.

[Lago e Jones 81] Lago, P.J., Jones, N.B. (1981), "Low frequency spectral analysis of the EMG", *Med. & Biol. Eng. & Comput.*, **19**, 779–782.

[Lago e Jones 82] Lago, P.J.A., Jones, N.B. (1982), "Note on the spectral analysis of neural spike trains", *Med. & Biol. Eng. & Comput.*, **20**, 44–48.

[Lago et al 82] Lago, P.J., Jones, N.B., Rocha, A.P. (1982), "Detection of polyphasic potentials by spectral analysis of a point process derived from the EMG: A comparison of theorectical and simulation results", *Proc. 1982 Portugal Workshop on Signal Processing and its Applications*, B2/9/1–8.

[Lago e Jones 83] Lago, P.J., Jones, N.B. (1983), "Turning points spectral analysis of the interference myoelectric activity", *Med. & Biol. Eng. & Comput.*, **21**, 333–342.

[Lago et al 83] Lago, P.J., Jones, N.B., Rocha, A.P. (1983), "Point process, spectral analysis and the interference EMG", *Proc. 1983 Spain Workshop on Signal Processing and its Applications*, C1/4/315–324.

[Lago et al 84] Lago, P.J., Jones, N.B., Rocha, A.P. (1984), "Characterization of the complexity of the action potential through spectral analysis", *Digital signal processing 1984*, V. Cappelini and G. Constantinides (ed.s), Elsevier Science Publishers B.V., Amsterdam, 684–690.

[Lago et al 89] Lago, P.J.A., Rocha, A.P., Jones N.B. (1989), "Covariance density estimation for autoregressive spectral modelling of point processes", *Biological Cybernetics*, **61**, 3, 195–203.

[Lago et al 90] Lago, P.J.A., Rocha, A.P., Jones, N.B. (1990), "Autoregressive spectral analysis of point processes", *Mathematics in signal processing*, McWhirter, R.G. (ed.), Clarendon Press, 385–396.

[Lewis 70] Lewis, P.A.W. (1970), "Remarks on the theory, computation and application of the spectral analysis of series of events", *J. Sound Vib.*, **12**, 355–375.

[Lindstrom e Petersen 83] Lindstrom, L., Petersen, I. (1983), "Power spectrum analysis of EMG signals and its applications", *Computer aided-electromyography*, Desmedt, J.E. (ed.), Karger, 1–51.

[Luenberger 73] Luenberger, D.G. (1973), *Introduction to linear and nonlinear programming*, Addison-Wesley.

[Makhoul 75] Makhoul, J. (1975), "Spectral linear prediction: properties and applications", *IEEE Trans. Acoust., Speech and Signal Processing*, **23**, 283–296.

[Malliani et al 91] Malliani, A., Pagani, M., Lombardi, F., Cerutti, S. (1991), "Cardiovascular neural regulation explored in the frequency domain", *Circulation*, **84** , 2, 482–492.

[Marple 87] Marple, S.L.Jr. (1987), *Digital spectral analysis with applications*, Prentice-Hall, Englewood Cliffs.

[Noble e Daniel 77] Noble, B., Daniel, J.W. (1977), *Applied linear algebra* (2nd edition), Prentice-Hall.

[Pagani et al 86] Pagani, M., Lombardi, F., Guzzetti, S., Rimoldi, O., Furlan, R., Pizzinelli, P., Sandrone, G., Malfatto, G., Dell'Orto, S., Piccaluga, E., Turiel, M., Baselli, G., Cerutti, S., Malliani, A. (1986), "Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog", *Circulation Research*, **59** , 2, 178–193.

[Paiss e Inbar 87] Paiss, O., Inbar, G.F. (1987), "Autoregressive modeling of surface EMG and its spectrum with application to fatigue", *IEEE Trans. on Biomedical Engineering*, **34**, 761–770.

[Pan et al 89] Pan, Z.S., Zhang, Y., Parker, P.A. (1989), "Motor unit power spectrum and firing rate", *Med. & Biol. Eng. & Comput.*, **29**, 14–18.

[Parzen 74] Parzen, E. (1974), "Some recent advances in time series modeling", *IEEE Trans. Automatic Control*, **19**, 723–730.

[Priestley 81] Priestley, M.B. (1981), *Spectral analysis and time series, volume 1: Univariate series*, Academic Press.

[Puig et al 91] Puig, J., Freitas, J., Carvalho, M., Brandão, A., Rocha, A., Lago, P., Costa, O., Falcão de Freitas (1991),"Effect of transient sino-atrial and atrio-ventricular Wenckeback blocks in heart rate variability assessed by spectral analysis in patients with familial amyloidotic polyneuropathy", *Proceedings of III International Congress of Cardiology*, Novembro 1991, Marrocos.

[Ralston 65] Ralston, A. (1965), *A first course in numerical analysis*, McGraw-Hill.

[Rocha et al 87] Rocha, A.P., Lago, P.J.A., Jones, N.B., Pinheiro A.V. (1987), "Characterization of the turning points spectral density of the EMG", *Digital signal processing 1987*, Cappelini, V., Constantinides, A.G. (ed.s), Elsevier Science Publishers B.V., Amsterdam, 573–578.

[Rocha e Lago 88] Rocha, A.P., Lago, P.J.A. (1988), "Análise espectral de processos estocásticos pontuais: métodos e algoritmos", *Actas I Congresso Português de Engenharia Biomédica*, 19–21.

[Rocha e Lago 90] Rocha, A.P., Lago, P.J.A. (1990), "Estimation of the spectrum of the counts of a stochastic point processes", *Actas XIV Jornadas Hispano Lusas de Matematicas*, **III**, Univ. de La Laguna, Secretariado de Publicaciones, 1423–1427.

[Stålberg e Antoni 83] Stålberg, E., Antoni, L. (1983), "Computer-aided EMG analysis", *Computer-aided electromyography*, Karger, 186–234.

[VanBoxtel e Schomaker 83] Van Boxtel, A., Schomaker, R.B. (1983), "Motor unit firing rate during static contraction indicated by the surface EMG power spectrum", *IEEE Trans. on Biomedical Engineering*, **30**, 601–609.

[Zetterberg 69] Zetterberg, L. H. (1969), "Estimation of the parameters for a linear difference equation with application to EEG analysis", *Math. Biosci.*, **5**, 227–275.