

Processamento de Sinal

Conceitos, Métodos e Aplicações

**Texto Tutorial da Disciplina:
APSI - LEEC**

J.P. Marques de Sá – jmsa@fe.up.pt
Faculdade de Engenharia da Universidade do Porto
© 2001 J.P. Marques de Sá

Índice

3	Processos Estocásticos.....	3
3.1	Conceitos Básicos.....	3
3.2	Espectro de Potência	10
3.3	SLITs com entradas estocásticas	11
3.4	Filtragem de Sinais Estocásticos.....	16
3.4.1	Detecção de sinais conhecidos com o filtro adaptado	16
3.4.2	Suavização de sinais desconhecidos	20
3.4.2.1	Suavização de média movente	20
3.4.2.2	Técnica de Médias.....	26
3.4.3	Estimação de processos estocásticos.....	28
3.4.3.1	Filtro de Wiener	28

3 Processos Estocásticos

3.1 Conceitos Básicos

Um processo estocástico $\mathbf{x}(n)$ é uma família de funções (apenas tratamos o caso discreto), reais ou complexas, definidas num espaço de probabilidades.

Uma amostra $x(n)$ de $\mathbf{x}(n)$ é denominada uma *realização* ou *instância* do processo.

1 - Média: $E[\mathbf{x}(n)] = \mu_x(n)$

2 – Correlação cruzada: $E[\mathbf{x}(n_1)\mathbf{y}^*(n_2)] = r_{xy}(n_1, n_2)$

3 - Autocorrelação: $E[\mathbf{x}(n_1)\mathbf{x}^*(n_2)] = r_{xx}(n_1, n_2) \equiv r(n_1, n_2)$

$r(n, n) = E[\mathbf{x}^2(n)]$ é a *intensidade* ou *potência média* do processo.

$$E[(\mathbf{x}(n_1) + \mathbf{x}(n_2))^2] = E\left\{[\mathbf{x}(n_1) + \mathbf{x}(n_2)][\mathbf{x}^*(n_1) + \mathbf{x}^*(n_2)]\right\} = r(n_1, n_1) + r(n_2, n_2) + 2r(n_1, n_2)$$

4 – Variância: $\sigma_x^2 = V[\mathbf{x}(n)] = E[(\mathbf{x}(n) - \mu_x(n))^2]$

5 – Covariância cruzada:

$$E[(\mathbf{x}(n_1) - \mu_x(n_1))(\mathbf{y}^*(n_2) - \mu_y(n_2))] = c_{xy}(n_1, n_2)$$

Logo:

$$c_{xy}(n_1, n_2) = r_{xy}(n_1, n_2) - \mu_x(n_1)\mu_y^*(n_2)$$

6 – Autocovariância: $c(n_1, n_2) = r(n_1, n_2) - \mu_x(n_1)\mu_y^*(n_2)$

Em particular: $c(n, n) = r(n, n) - |\mu_x(n)|^2$

7 – Centragem:

Seja o processo: $\mathbf{x}_c(n) = \mathbf{x}(n) - \mu_x(n)$.

Então o processo $\mathbf{x}_c$ tem média nula e

$$c_{x_c x_c}(n_1, n_2) = r_{x_c x_c}(n_1, n_2) = c_{xx}(n_1, n_2)$$

8 – Estacionaridade:

Em sentido estrito: Todas as propriedades estatísticas são invariantes para uma translação do tempo.

Em sentido lato (*fracamente estacionário*):

$$E[\mathbf{x}(n)] = \mu = \text{constante} \quad \text{estatística de 1ª ordem}$$

$$r(n, n + \tau) = r(\tau) = E\left[\mathbf{x}(n + \tau)\mathbf{x}^*(n)\right] = E\left[\mathbf{x}\left(n + \frac{\tau}{2}\right)\mathbf{x}^*\left(n - \frac{\tau}{2}\right)\right] \quad \text{estatística de 2ª ordem}$$

Em particular: $E[\mathbf{x}^2(n)] = r(0)$

Para um processo normal (todas as fdp's conjuntas de $n_1, \dots, n_N$ são normais) a estacionaridade lato implica a estrita, dado que o processo fica univocamente determinado pela média e pela autocorrelação.

Dois processos são conjuntamente estacionários se ambos são estacionários e a covariância cruzada só depende de τ .

9 – Ergodicidade:

Um processo $\mathbf{x}(n)$ é ergódico se todas as suas propriedades estatísticas se podem determinar a partir de uma sua única realização. Para tal o processo deverá ser estacionário, podendo-se, então, substituir médias de conjuntos por médias temporais.

$$E[\mathbf{g}(n)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N g(n)$$

10 – **Ruído branco** é um processo tal que: $R(\tau) = I\delta(\tau)$.

Exemplos:

a) Números aleatórios com distribuição uniforme em $[-1, 1[$. São possíveis $24 \cdot 3600 \cdot 100$ instâncias equiprováveis.

Trata-se de um processo estacionário e ergódico de média nula.

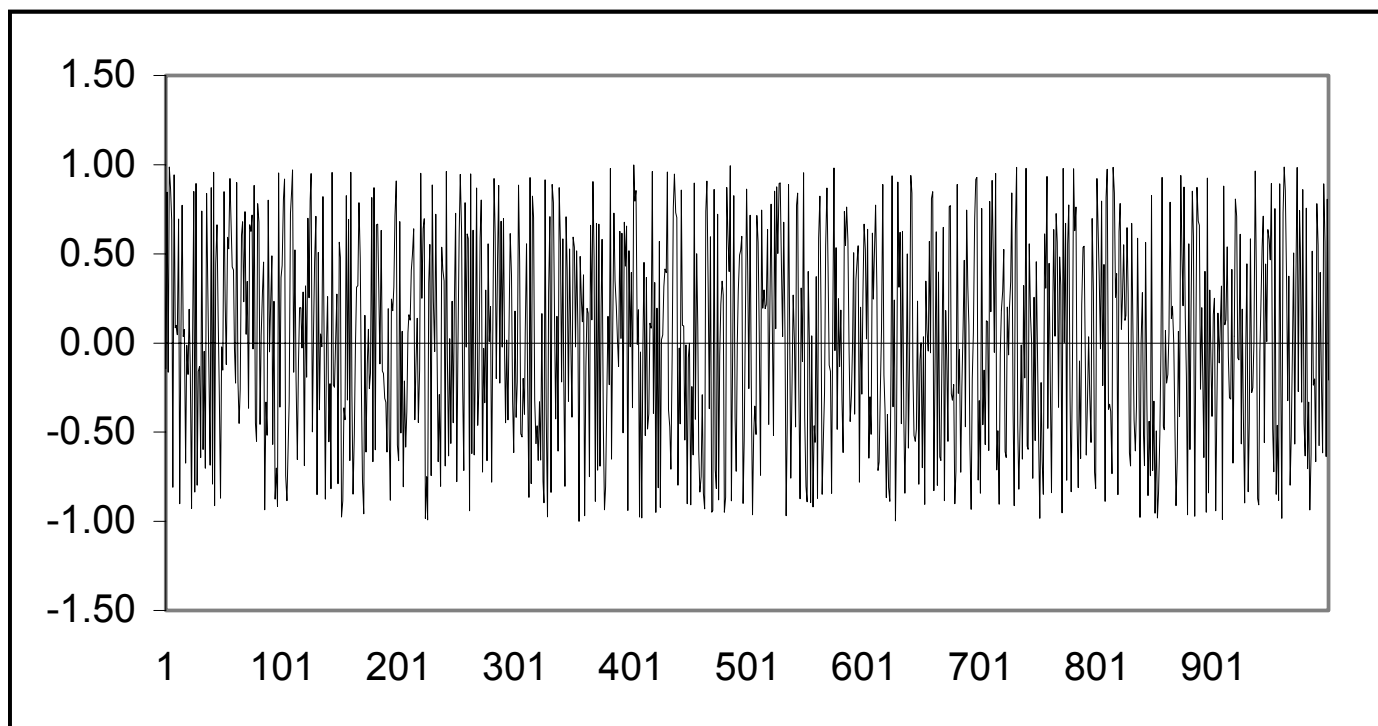


Figura 3.1. Uma realização de $x(n)$.

Experiência de 25 realizações (consecutivas...) de $N = 1000$ amostras.

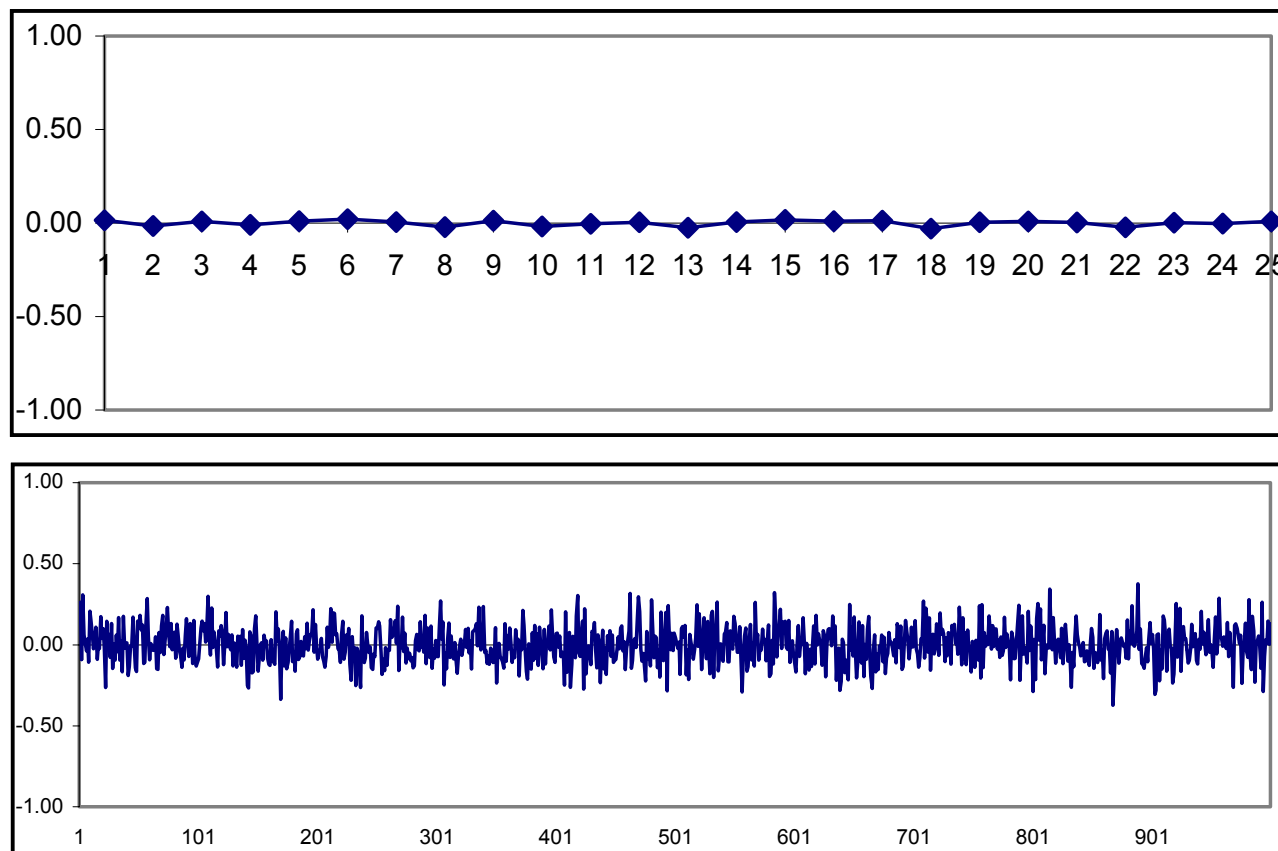


Figura 3.2. Média temporal estimada em 1000 amostras para 25 sinais (em cima); média de conjunto estimada em 25 sinais para 1000 amostras (em baixo).

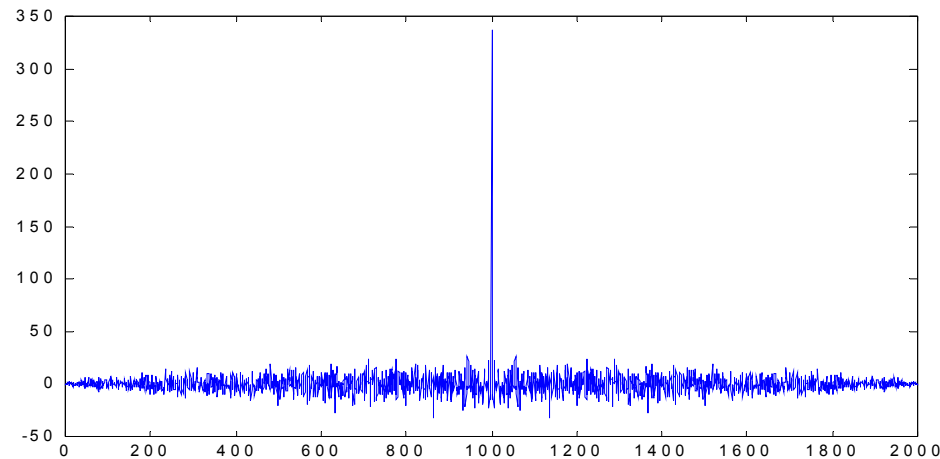
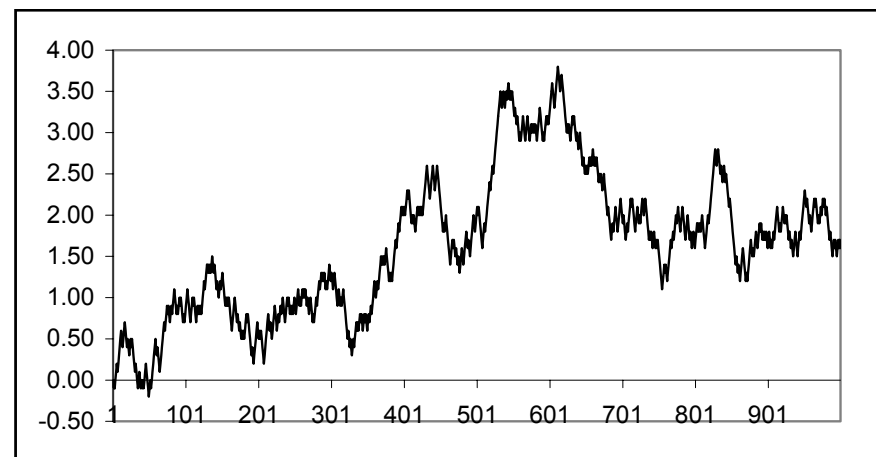


Figura 3.3. Autocorrelação temporal de uma realização de 1000 números com distribuição uniforme em $[-1, 1[$ (aproximação de ruído branco).

b) Passeio binário simétrico:

$$\mathbf{x}(n) = \sum_{k=0}^n \mathbf{s}(k) \quad \text{com} \quad P(s(k) = s) = P(s(k) = -s) = \frac{1}{2}$$

Figura 3.4. Uma realização do passeio aleatório simétrico com $s = 0.1$. Usou-se a anterior realização uniforme para as probabilidades do incremento.



É possível mostrar que:

$$E[\mathbf{x}(n)] = 0; \quad E[\mathbf{x}^2(n)] = ns^2 \quad \text{logo a variância cresce com } n.$$

O processo não é estacionário e, portanto, não é ergódico.

c) Seja:

$$x(n) = \begin{cases} x_1(n), & \text{sequência de Bernoulli com } P = 0.5; \quad P(x_1(n)) = 0.5 \\ x_2(n), & \text{sequência de Bernoulli com } P = 0.3; \quad P(x_2(n)) = 0.3 \end{cases}$$

Então:

$$E[\mathbf{x}(n)] = 0.4$$
$$R(n_1, n_2) = 0.4\delta(\tau)$$

Logo, o processo é estacionário. Contudo, não é ergódico (p. ex. $\bar{x}_1(n) = 0.5 \neq E[\mathbf{x}(n)]$).

3.2 Espectro de Potência

No caso de sinais estocásticos a transformada discreta de Fourier de uma realização não tem significado útil.

Para sinais estacionários usamos o *espectro de potência*:

$$S(\omega) = \sum_{m=-\infty}^{\infty} r(m)e^{-jm\omega} \Leftrightarrow r(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega)e^{jm\omega} d\omega$$

Veremos mais tarde o problema de estimar $S(\omega)$.

Notar, em particular, que:

$$r(0) = \int_{-\pi}^{\pi} S(\omega) d\omega$$

Para o ruído branco: $S(\omega) = I$ (daí o nome)

O espectro de potência cruzada é definido de forma semelhante:

$$S_{xy}(\omega) = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-jm\omega}$$

Para processos não estacionários define-se um espectro de potência variante no tempo:

$$S(k, \omega) = \sum_{m=-\infty}^{\infty} r(k, m)e^{-jm\omega}$$

3.3 SLITs com entradas estocásticas

$y(n) = x(n) \otimes h(n)$ com:

- $E[y(n)] = E[x(n)] \otimes h(n)$
- $r_{xy}(n_1, n_2) = \sum_{k=-\infty}^{\infty} r_{xx}(n_1, n_2 - k) h^*(k)$
- $r_{yy}(n_1, n_2) = \sum_{k=-\infty}^{\infty} r_{xy}(n_1 - k, n_2) h(k)$

Em particular, para processos estacionários, $y(n)$ é também estacionário, com:

$$r_{xy}(m) = \sum_{k=-\infty}^{\infty} r_{xx}(m+k) h^*(k)$$

$$r_{yy}(m) = \sum_{k=-\infty}^{\infty} r_{xy}(m, k) h(k) \Rightarrow r_{yy}(m) = r_{xx}(m) \otimes \rho(m)$$

e
$$\rho(m) = \sum_{k=-\infty}^{\infty} h(m+k) h^*(k)$$

Usando $H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}$, conclui-se:

- $S_{xy}(\omega) = S_{xx}(\omega)H^*(\omega)$
- $S_{yy}(\omega) = S_{xx}(\omega)|H(\omega)|^2 \Rightarrow E[|\mathbf{y}(n)|^2] = r_{yy}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{xx}(\omega)|H(\omega)|^2 d\omega$

Suponhamos que um processo discreto $\mathbf{y}(n)$ é obtido com uma excitação de entrada $\mathbf{x}(n)$ de um SLIT de ordem p :

$$\mathbf{y}(n) + a_1\mathbf{y}(n-1) + \dots + a_p\mathbf{y}(n-p) = b_0\mathbf{x}(n) + \dots + b_p\mathbf{x}(n-p)$$

Consideremos uma excitação por ruído branco, $S_{ff}(\omega) = 1$. Então:

$$S_{yy}(\omega) = |H(\omega)|^2 = H(\omega)H(-\omega) = H(z)H(z^{-1}) \Big|_{z=e^{j\omega}}$$

Logo:

$$S_{yy}(z) = H(z)H(z^{-1})$$

Exemplo:

Processo estocástico de primeira ordem.

$$\mathbf{y}(n) - a\mathbf{y}(n-1) = \mathbf{x}(n); \quad S_{xx}(\omega) = 1; \quad H(z) = \frac{z}{z-a}$$

Então:

$$S_{yy}(z) = \frac{z}{(z-a)(1-az)} = \frac{-a^{-1}}{(z+z^{-1}) - (a+a^{-1})};$$

$$r_{yy}(m) = \frac{a^{|m|}}{1-a^2};$$

$$S_{yy}(\omega) = \frac{1}{1-2a \cos \omega + a^2}.$$

Por forma a obtermos simplesmente $r_{yy}(m) = a^{|m|}$, basta multiplicar $S_{yy}(z)$ por $(1-a^2)$, correspondendo a:

$$\mathbf{y}(n) - a\mathbf{y}(n-1) = \sqrt{1-a^2} \mathbf{x}(n)$$

Experiência para $\alpha=0.8$ e entrada ruído Gaussiano branco (distribuição normal, $N(0,1)$).

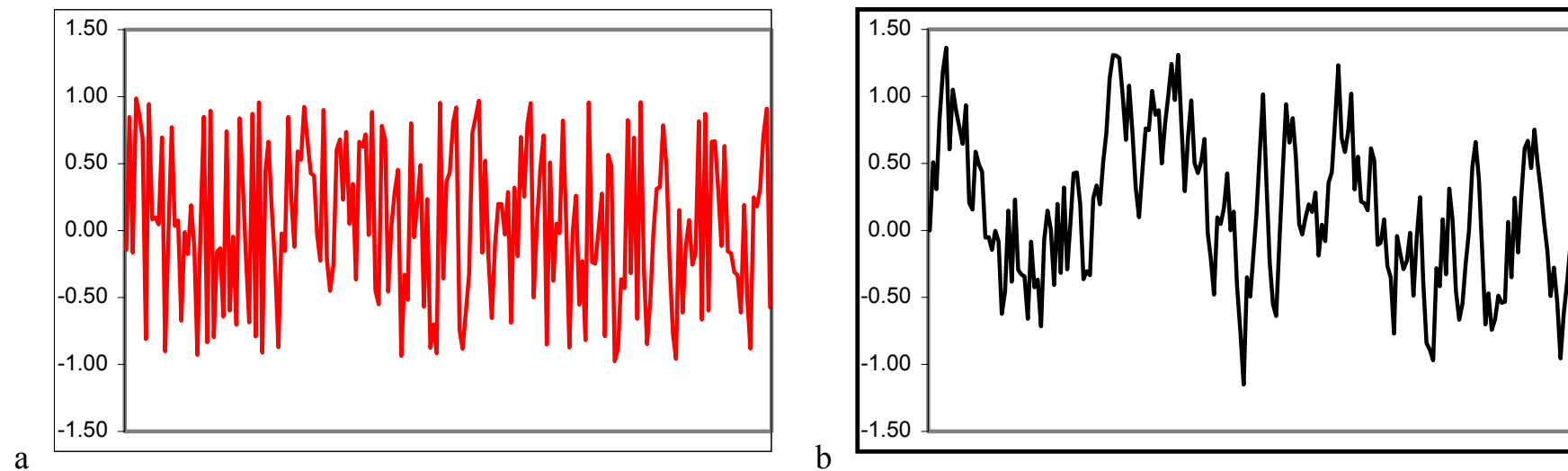


Figura 3.5. a) Uma realização de ruído Gaussiano branco; b) Resposta da anterior realização por um sistema de primeira ordem com $\alpha = 0.8$.

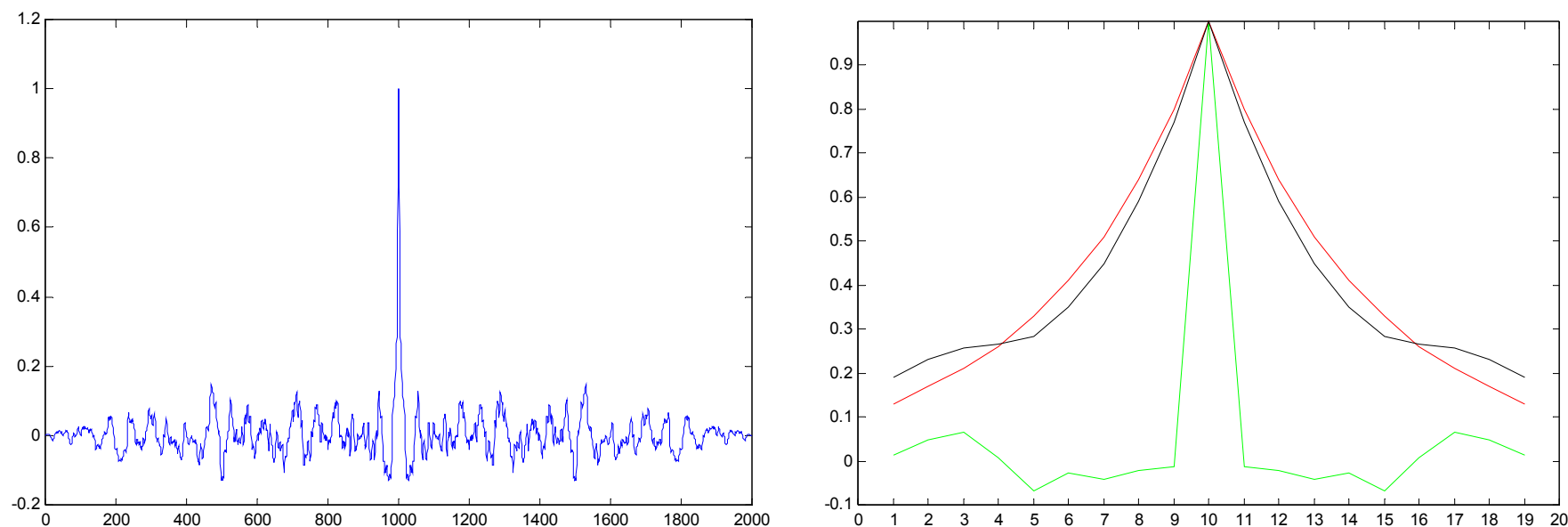


Figura 3.6. à esquerda, autocorrelação (estimativa) da realização do anterior processo de 1ª ordem usando 1000 amostras; à direita, detalhe das seguintes autocorrelações (o ponto de abcissa 10 corresponde ao anterior 1000): ruído Gaussiano branco (verde); processo de 1ª ordem (preto); curva teórica do processo de 1ª ordem, $0.8^{|m-10|}$ (vermelho).

3.4 Filtragem de Sinais Estocásticos

Dado: $\mathbf{x}(n) = f(n) + \mathbf{e}(n)$

Estimar $f(n)$ na presença de ruído aditivo $\mathbf{e}(n)$, com espectro de potência $S(\omega)$.

Três casos:

- O sinal $f(n)$ é conhecido
- O sinal $f(n)$ é desconhecido
- O sinal é um processo estocástico, $\mathbf{f}(n)$, com espectro de potência conhecido

Consideramos apenas sinais reais.

3.4.1 Detecção de sinais conhecidos com o filtro adaptado

Pretende-se estimar se o sinal $f(n)$, conhecido, ocorre ou não em $\mathbf{x}(n)$.

Usamos um estimador linear:

$$\mathbf{y}(n) = \mathbf{x}(n) \otimes h(n) = f(n) \otimes h(n) + \mathbf{e}(n) \otimes h(n) = y_f(n) + \mathbf{y}_e(n)$$

Pretendemos determinar $h(n)$, tal que para $n = n_0$ seja máximo:

$$SNR = \frac{|y_f(n_0)|}{\sqrt{E[y_e^2(n_0)]}}$$

Usando o resultado anterior, temos:

$$E[y_e^2(n)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) |H(\omega)|^2 d\omega$$

Por outro lado: $y_f(n_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega) H(\omega) e^{j\omega n_0} d\omega$

Considere-se a *desigualdade de Schwartz*:

$$\left| \int_a^b z(x) w(x) dx \right|^2 \leq \int_a^b |z(x)|^2 dx \int_a^b |w(x)|^2 dx, \text{ ocorrendo a situação de igualdade para } z(x) = k w^*(x).$$

Usando este resultado, obtém-se:

$$\left| \int_{-\pi}^{\pi} F(\omega) H(\omega) e^{j\omega n_0} d\omega \right|^2 \leq \int_{-\pi}^{\pi} \frac{|F(\omega)|^2}{S(\omega)} d\omega \int_{-\pi}^{\pi} S(\omega) |H(\omega)|^2 d\omega$$

Donde:

$$SNR^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|F(\omega)|^2}{S(\omega)} d\omega$$

O máximo de SNR ocorre para a situação de igualdade que corresponde a:

$$\sqrt{S(\omega)}H(\omega) = k \frac{F^*(\omega)}{\sqrt{S(\omega)}} e^{-j\omega n_0}$$

Obtém-se, então, o *filtro adaptado*:

$$H(\omega) = k \frac{F^*(\omega)}{S(\omega)} e^{-j\omega n_0}$$

No caso de ruído branco, $S(\omega) = N_0$, temos:

$$h(n) = \frac{k}{N_0} f(n_0 - n) \quad \text{e} \quad SNR_{\max} = \sqrt{E / N_0}, \text{ sendo } E \text{ a energia do sinal.}$$

Exemplo:

$$f(n) = [0.25 \ 0.75 \ 2.00 \ 4.00 \ 5.00 \ 4.50 \ 3.00]$$

$e(n)$: ruído branco (números aleatórios com distribuição uniforme em $[-1, 1]$)

$$x(n) = 0.35 f(n) + e(n)$$

$$h(n) = [3.00 \ 4.50 \ 5.00 \ 4.00 \ 2.00 \ 0.75 \ 0.25]$$

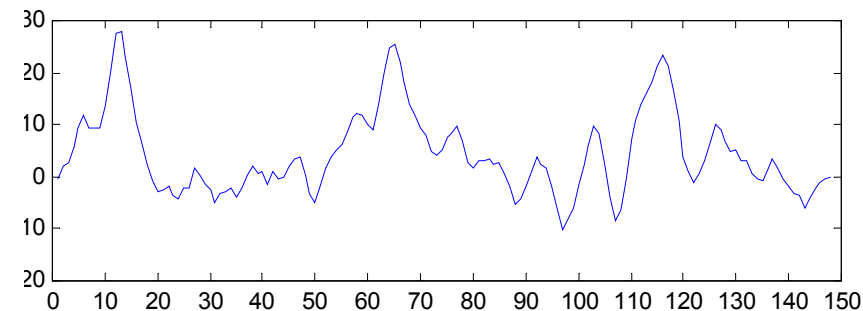
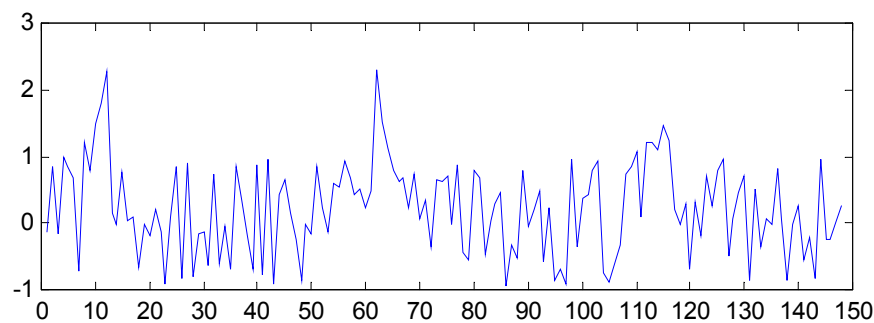
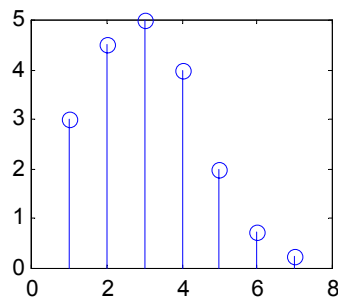
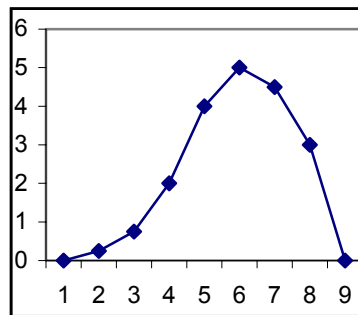


Figura 3.7. Em cima: Um sinal (à esquerda) mergulhado em ruído branco (à direita). As ocorrências do sinal em $n_0=13, 65$ são visíveis; a ocorrência em $n_0=116$ não o é. Em baixo: filtro adaptado (à esquerda) e sinal filtrado (à direita). Os três máximos locais proeminentes deste sinal ocorrem precisamente nas anteriores posições.

3.4.2 Suavização de sinais desconhecidos

3.4.2.1 Suavização de média movente

Assumimos que $f(n)$ é desconhecido e $\mathbf{e}(n)$ é um processo estocástico de média nula e autocorrelação $r(n_1, n_2)$.

Temos: $E[\mathbf{x}(n)] = f(n)$; $\sigma_x^2 = E[\mathbf{e}^2(n)] = r(n, n)$

Suavizamos o sinal por média movente de $2N + 1$ amostras:

$$\mathbf{y}(n) = \frac{1}{2N} \sum_{k=-N}^N \mathbf{x}(n-k)$$

Logo:

$$\mathbf{y}(n) = \frac{1}{2N} \sum_{k=-N}^N f(n-k) + \frac{1}{2N} \sum_{k=-N}^N \mathbf{e}(n-k) = y_f(n) + \mathbf{y}_e(n)$$

Queremos determinar N por forma a minimizar o valor médio do quadrado do erro $\mathbf{y}(n) - f(n)$.

Mas $\mathbf{y}(n)$ é um processo estocástico de média $y_f(n)$ e variância:

$$\sigma_y^2 = E[y_e^2(n)] = \frac{1}{4N^2} \sum_{k_1=n-N}^{n+N} \sum_{k_2=n-N}^{n+N} r(k_1, k_2)$$

Suponhamos que o ruído é branco: $r(n_1, n_2) = I(n_1)\delta(n_1 - n_2)$

Então, prova-se que (Papoulis, 1977):

- Variância: $\sigma_y^2 = \frac{1}{2N} I(n + n_0)$ com $|n_0| < N$
- Viés: $b = E[y(n)] - f(n) = \frac{N^2}{6} f''(n + n_1), \quad |n_1| < N$

Dilema:

Para N largo a variância é pequena e o viés é grande; para N pequeno a variância é grande e o viés pequeno.

Note-se que quando $f''(n) = 0$ (e.g. intervalos onde o sinal é nulo) o viés é nulo.

Na situação em que $f''(n)$ e $I(n)$ são suficientemente suaves, temos:

$$I(n + n_0) \approx I(n); \quad f''(n + n_1) \approx f''(n)$$

Então:

$$\varepsilon = E[(y(n) - f(n))^2] = b^2 + \sigma_y^2 \cong \frac{N^4}{36} [f''(n)]^2 + \frac{I(n)}{2N}$$

O mínimo de ε ocorre para:

$$N_m = \sqrt[5]{\frac{9I(n)}{2[f''(n)]^2}}$$

Note-se que depende de $f''(n)$ e $I(n)$.

Exemplo:

Determinar o N óptimo para a filtragem de sinais com $f''(n) = 0.2$ e $I(n) = 1$.

Aplicando a fórmula, obtém-se: $N_m = 2.57 \approx 3$. Logo, um comprimento total do filtro de $L = 2N + 1 = 7$.

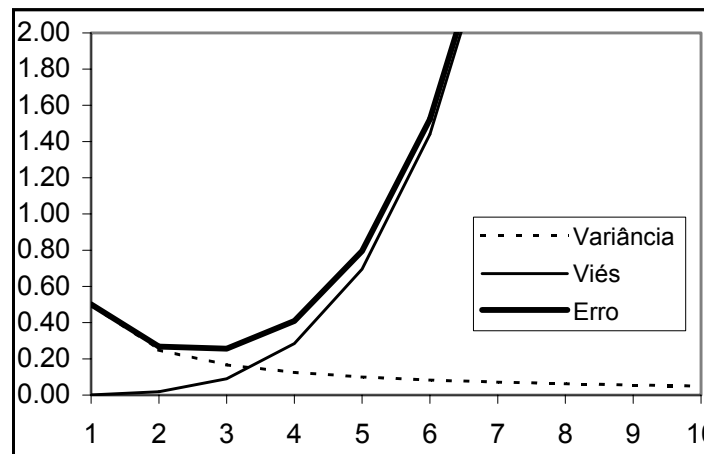


Figura 3.8. Curvas de variância, viés e erro para $f'(n) = 0.2$ e $I(n) = 1$.

Seja $f(n)$ constituído por várias réplicas de sinais parabólicos, do tipo $s(n) = G[49 - (n-7)^2]$, com ganho $G = 0.1$, embebido em ruído branco com distribuição Gaussiana de variância 1. Estamos nas condições de aplicação do resultado anterior.

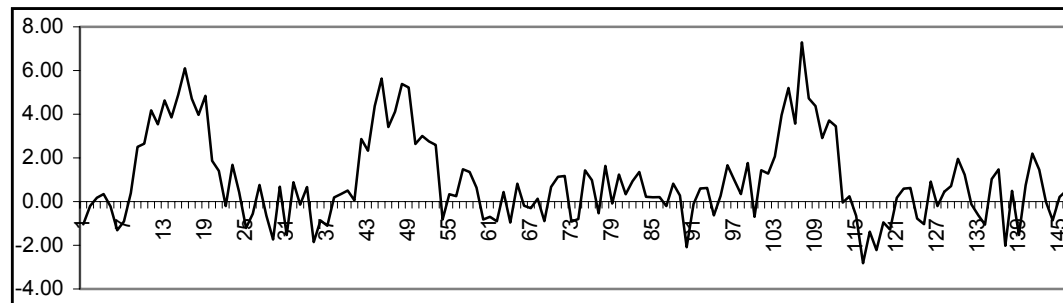
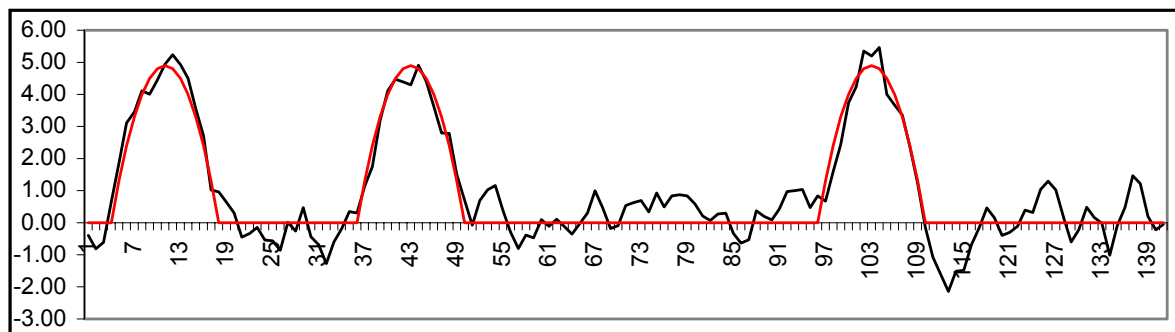


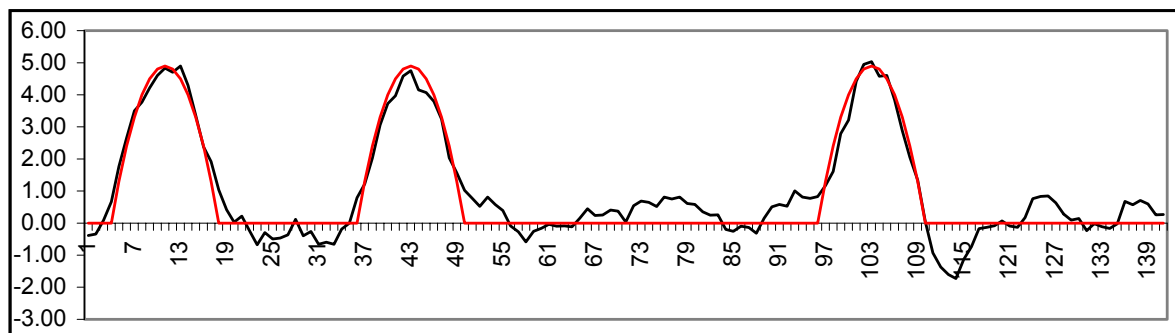
Figura 3.9. Uma realização do processo $x(n)$, com $f(n)$ constituído por "impulsos parabólicos" com $f'(n) = 0.2$ e ruído branco com $I(n) = 1$.

N

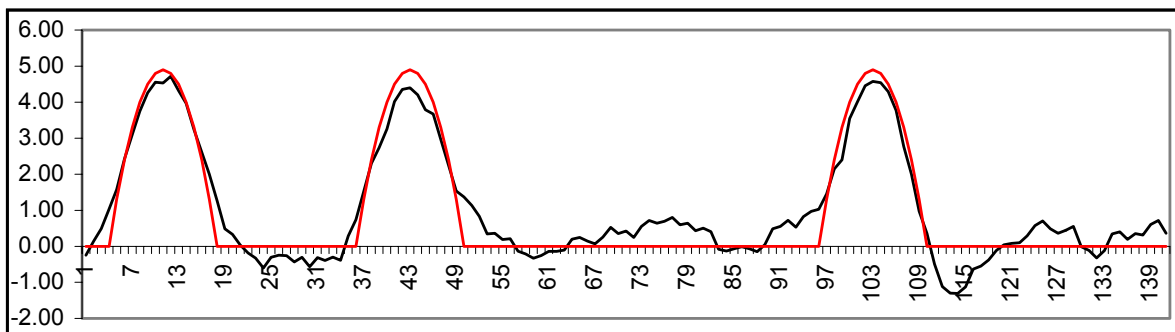
1



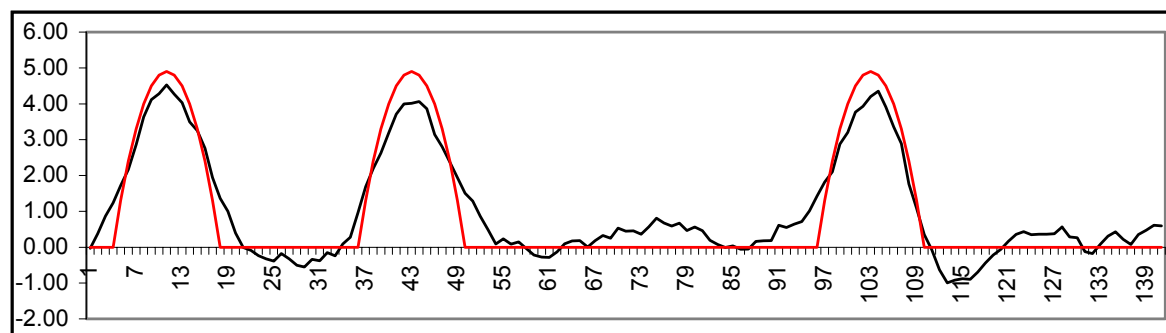
2



3



4



5

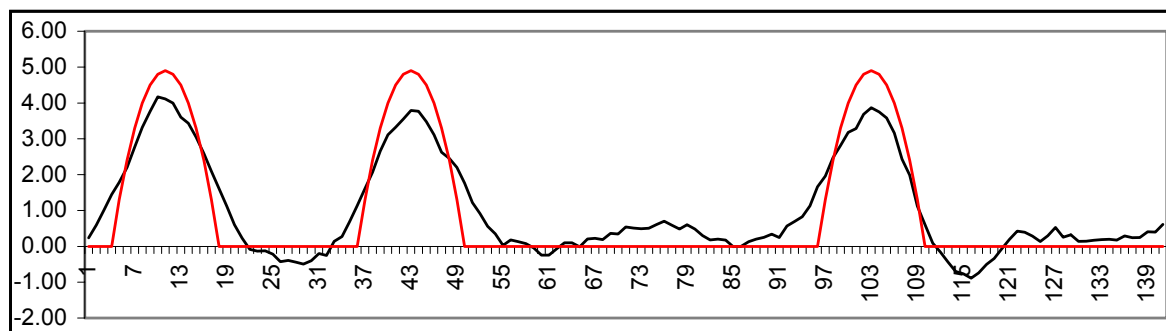


Figura 3.10. Resultado da filtragem da anterior realização de $x(n)$ para vários valores de N . Notar o efeito do viés e da variância.

Erros obtidos nas filtrações anteriores (realização com 148 amostras):

N	$L = 2N + 1$	$\sum \varepsilon^2(n)$
1	3	69.61
2	5	56.62
3	7	57.61
4	9	63.31
5	11	70.90

3.4.2.2 Técnica de Médias

Condições:

- Processo ergódico.
- Uma realização do processo consiste na repetição espaçada de várias réplicas de um sinal.
- Ruído de média nula, $E[\mathbf{e}(n)] = 0$.

Podemos obter uma versão suavizada do sinal da seguinte forma:

- Alinhar temporalmente as diversas réplicas do sinal, segmentando-o e usando uma medida de correlação cruzada.
- Determinar a média das réplicas do sinal:

$$\sum_{i=1}^M \mathbf{x}_i(n) \xrightarrow{M \rightarrow \infty} E[\mathbf{x}(n)] = f(n) + E[\mathbf{e}(n)] = f(n)$$

Exemplo:

Determinação de um complexo médio de ondas do ECG.

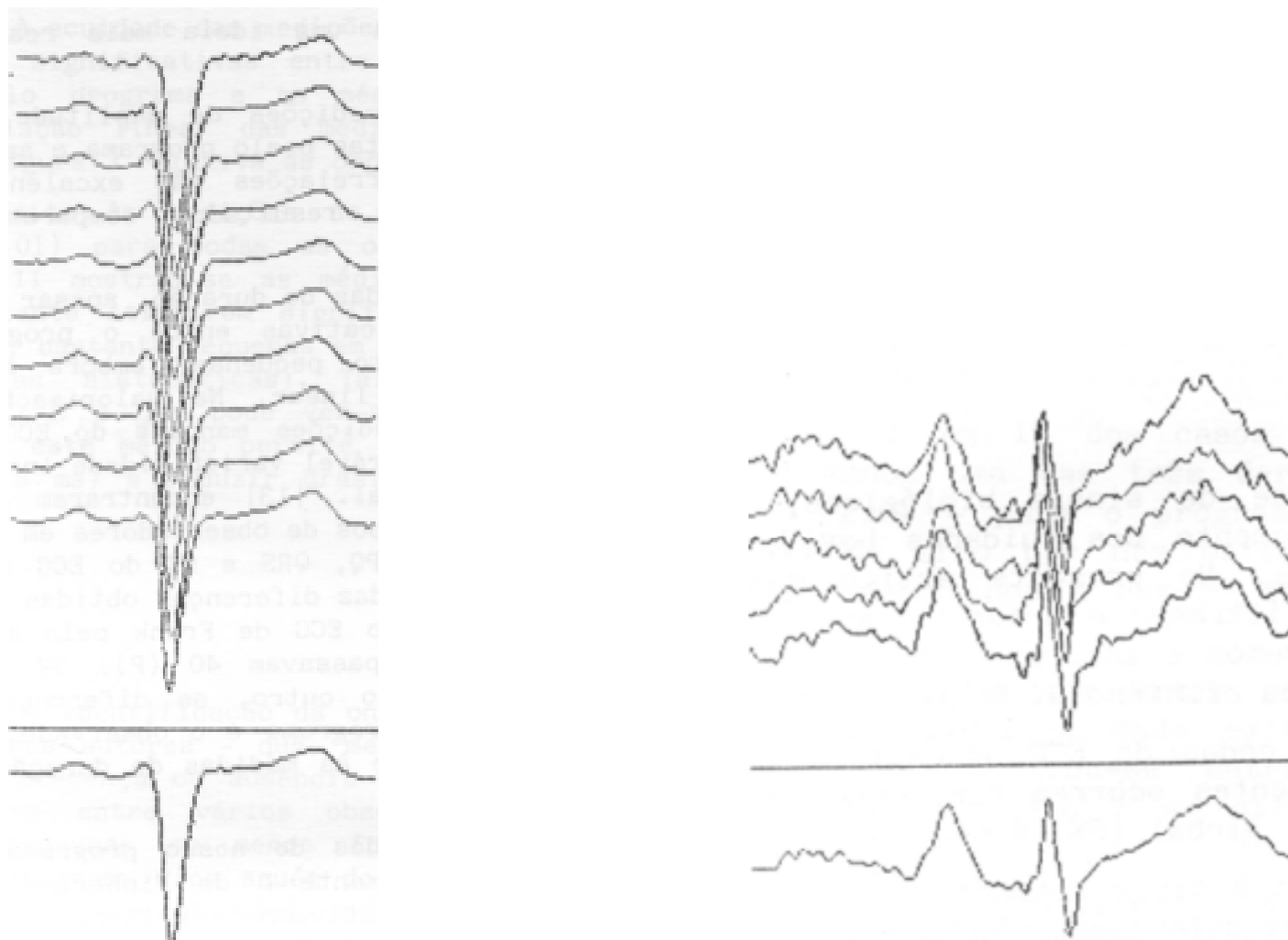


Figura 3.11. Exemplos de obtenção de complexos médios de ondas de ECGs, usando a técnica de médias. Note-se como se obtém uma suavização adequada mesmo para um número pequeno de réplicas do sinal.

3.4.3 Estimação de processos estocásticos

Temos:

$$\mathbf{x}(n) = \mathbf{f}(n) + \mathbf{e}(n)$$

Assume-se que todos os processos são estacionários e reais.

3.4.3.1 Filtro de Wiener

Pretendemos obter uma estimativa de $\mathbf{f}(n)$, usando uma combinação linear:

$$\hat{\mathbf{f}}(n) = \sum_{k=-\infty}^{\infty} \mathbf{x}(n-k)h(k),$$

por forma a minimizar o erro quadrático médio:

$$\varepsilon = \sum_{n=-\infty}^{\infty} (\mathbf{f}(n) - \hat{\mathbf{f}}(n))^2$$

Para tal, o erro tem de ser ortogonal aos dados:

$$E \left\{ \left[\mathbf{f}(n) - \sum_{k=-\infty}^{\infty} \mathbf{x}(n-k)h(k) \right] \mathbf{x}(n-m) \right\} = 0, \quad \forall m$$

Donde:

$$r_{fx}(m) = \sum_{k=-\infty}^{\infty} r_{xx}(m-k)h(k), \quad \forall m$$

Logo, tomando a transformada em z: $S_{fx}(z) = S_{xx}(z)H(z)$

O filtro óptimo de Wiener, é:

$$H(z) = \frac{S_{fx}(z)}{S_{xx}(z)}$$

com erro

$$\varepsilon = r_{ff}(0) - \sum_{k=-\infty}^{\infty} r_{fx}(k)h(k)$$

Casos especiais:

1 – A correlação cruzada do sinal e do ruído é nula:

$$r_{fe}(m) = 0$$

Então:

$$S_{fx}(z) = S_{ff}(z), \quad S_{xx}(z) = S_{ff}(z) + S_{ee}(z)$$

$$H(z) = \frac{S_{ff}(z)}{S_{ff}(z) + S_{ee}(z)}$$

2 – Verifica-se a condição anterior com ruído branco, $r_{ee}(m) = I\delta(m)$

Então:

$$H(z) = \frac{S_{ff}(z)}{S_{ff}(z) + I}$$

Exemplo:

Processo estocástico de primeira ordem (ver atrás) com ruído branco.

$$r_{ff}(m) = a^{|m|} \Rightarrow S_{ff}(z) = (1 - a^2) \frac{-a^{-1}}{(z + z^{-1}) - (a + a^{-1})}, \quad 0 < |a| < 1$$

Substituindo no anterior $H(z)$:

$$H(z) = \frac{(a - a^{-1}) / I}{(z + z^{-1}) - (b + b^{-1})}$$

$$\text{com } b + b^{-1} = a + a^{-1} + (a^{-1} - a) / I$$

Logo:

$$h(n) = \frac{a^{-1} - a}{I(b^{-1} - b)} b^{|n|}$$

É possível provar que o erro é, neste caso: $\varepsilon = \frac{a^{-1} - a}{b^{-1} - b}$

Para o exemplo da secção 3.3, com $a = 0.8$ e $I = 1$, obtém-se:

$$b = 0.5; \quad h(n) = 0.3 \times 2^{-|n|}; \quad \varepsilon = 0.3$$

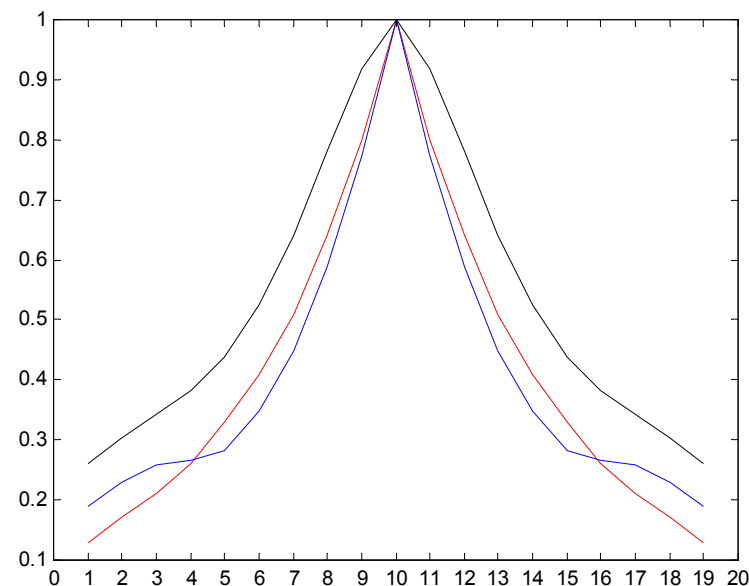


Figura 3.12. Detalhe das seguintes autocorrelações: processo da secção 3.3 (azul), estimada numa realização de 1000 amostras; curva teórica (vermelho); resultado da filtragem pelo anterior filtro de Wiener (preto).

Bibliografia

Kuc R (1988) Introduction to Digital Signal Processing. Mc Graw-Hill, Inc.

Lindner DK (1999) Introduction to Signals and Systems. Mc Graw-Hill, Inc.

Lyons RG (1997) Understanding Digital Signal Processing. Addison Wesley Longman, Inc.

Oppenheim AV, Schafer RW, Buck JR (1999) Discrete-Time Signal Processing. Prentice-Hall Inc.

Papoulis A (1984) Signal Analysis. McGraw-Hill, Inc.

Proakis JG, Manolakis DG (1996) Digital Signal Processing. Principles, Algorithms and Applications. Prentice Hall Int., Inc.

Rabiner LR, Gold B (1975) Theory and Application of Digital Signal Processing. Prentice-Hall, Inc.

Schwarz M, Shaw L (1975) Signal Processing: Discrete Spectral Analysis, Detection, and Estimation. McGraw-Hill.

Steiglitz K (1974) An Introduction to Discrete Systems, J. Wiley & Sons, Inc.

Tretter SA (1976) Introduction to Discrete Time Signal Processing J. Wiley & Sons, Inc.