

Processos Estocásticos

Paulo Lopes dos Santos
T-P Perdicoulis

22 de setembro de 2024

Conteúdo

1	Processos Estocásticos	1
1.1	Introdução	1
1.2	Distribuição dum processo estocástico	3
1.3	Média e Covariância	3
1.4	Processos estocásticos estacionários	5
1.5	Processos ergódicos	9
1.6	Densidade espectral	10
1.7	Previsão linear em PE	13
1.8	Modelos de entrada-saída de processos estocásticos	17
1.8.1	Modelo de Média móvel (MA)	17
1.8.2	Modelo Auto-Regressivo (AR)	18
1.8.3	Modelo Auto-Regressivo de Média Móvel (ARMA)	18
1.8.4	Filtro branqueador	19
1.9	Modelos de PE no espaço de estados	19

Resumo

Capítulo 1

Processos Estocásticos

1.1 Introdução

No mundo real, existem muitas variáveis aleatórias (v.a.s) que evoluem no tempo segundo leis probabilísticas. Como exemplo dessas variáveis, podemos referir a evolução da temperatura num determinado local (ver Figura 1.1), o ruído térmico em circuitos

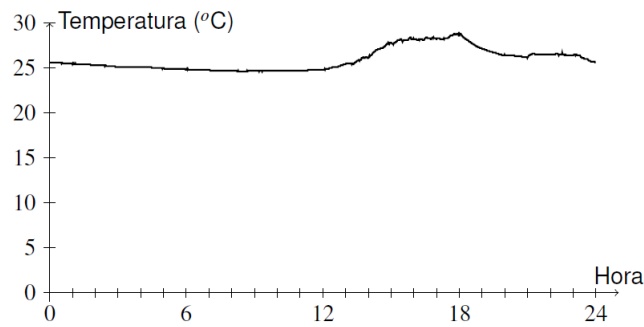


Figura 1.1: Evolução da temperatura num determinado local

eléctricos, o ruído de sensores (Figura 1.1), os sinais de radar, a cotação das acções na bolsa, o consumo de água numa cidade, a taxa de ocupação hoteleira de uma região, etc.. Chamamos *processo estocástico* (PE), ou série temporal, ao conjunto de todas as variações possíveis destas variáveis. Os PEs são, pois, famílias de funções (reais ou complexas) do tempo, ou seja, um PE é um conjunto de variáveis aleatórias (v.a), indexadas por um conjunto T . Sendo Ω o espaço amostral associado a essas v.a.s, podemos expressar um processo estocástico na forma

$$\{x(t, \omega), t \in T, \omega \in \Omega\}.$$

Se fixarmos $t = t_1$, teremos uma v.a. $x(t_1, \omega)$ no espaço amostral Ω . Se fixarmos o evento $\omega = \omega_1$ obteremos uma função do tempo $x(t, \omega_1)$ designada por *realização*.

Exemplo 1 *Variável aleatória e realização dum PE:*

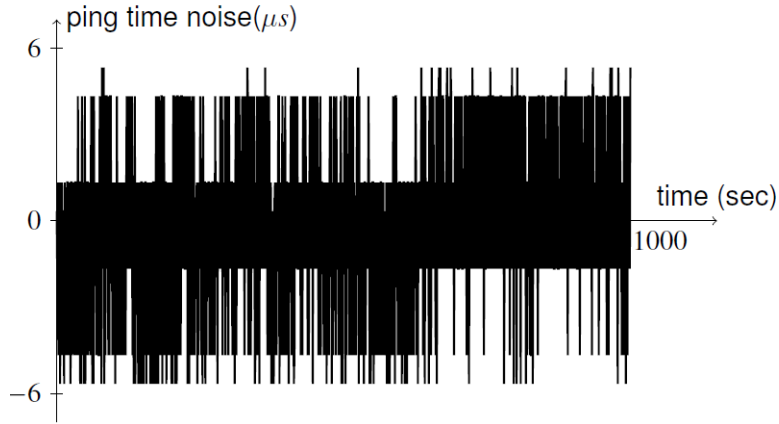


Figura 1.2: Ruído de um sensor ultrassônico

Se $x(t, \omega)$ for a evolução da temperatura ao longo de um dia, então, se $t = 12h$, $x(12, \omega)$ é uma variável aleatória que descreve a temperatura ao meio-dia. Se $\omega = 15$ de outubro de 2021, $x(t, 15 \text{ de outubro de } 2021)$ é a função que mapeia a temperatura em função do tempo, isto é, é a realização de $x(t, \omega)$ no dia 15 de outubro de 2021. $\square$

Para simplificar a notação, representaremos o PE $x(t, \omega)$ apenas por $x(t)$ ou $\{x(t)\}$, deixando cair o evento na lista de argumentos como é usual na representação das variáveis aleatórias.

Consideremos o conjunto dos índices, T , como $\mathbb{R} =]-\infty, \infty[$ ou ainda um intervalo finito $]a, b[\in \mathbb{R}$. Dizemos, neste caso, que o processo estocástico é de parâmetro contínuo, ou simplesmente, *contínuo*. Por outro lado, dizemos que o PE é de parâmetro discreto (ou simplesmente, *discreto*) quando $T \equiv \mathbb{Z} = \{t = 0, \pm 1, \pm 2, \dots\}$ ou ainda $T \equiv \mathbb{Z} = \{t = t_1, t_1 + 1, \dots, t_1 + n\} \subset \mathbb{Z}$.

Exemplo 2 - PE de parâmetro contínuo ou simplesmente PE contínuo:

Consideremos, de novo, a evolução da temperatura ao longo de um dia, já com a nova notação $x(t)$. Como varia continuamente no intervalo $t \in [0, 24]$, $x(t)$ é um PE de parâmetro contínuo. $\square$

Exemplo 3 - PE de parâmetro discreto ou simplesmente PE discreto:

Seja $x^*(t)$ um PE que resulta da amostragem de $x(t)$ de hora a hora, isto é, $t \in \{0, 1, \dots, 23\}$. Então $x^*(t)$ é PE de parâmetro discreto. $\square$

1.2 Distribuição dum processo estocástico

Tomemos os instantes $t_1, t_2, \dots, t_k$ do PE $x(t)$. Nestes instantes, o PE pode ser visto

como uma v.a. k -dimensional $X_{t_1:t_k} = \begin{pmatrix} x(t_1) \\ x(t_2) \\ \vdots \\ x(t_k) \end{pmatrix}$ cuja função de distribuição é

definida por

$$P\{x(t_1) \leq a_1, x(t_2) \leq a_2, \dots, x(t_k) \leq a_k\} \\ = \int_{-\infty}^{a_1} \int_{-\infty}^{a_2} \int_{-\infty}^{a_k} p_{t_1, t_2, \dots, t_k}(x_1, \dots, x_k) dx_1 dx_2, \dots, dx_k$$

onde $a_1, a_2, \dots, a_k \in \mathbb{R}$ e $p_{t_1, t_2, \dots, t_k}(x_1, \dots, x_k)$ é a *função de densidade de probabilidade conjunta* de $X_{t_1:t_k}$. Esta distribuição conjunta diz-se uma distribuição finita do PE nos instantes $t_1, t_2, \dots, t_k$. Tem-se, então, que a distribuição de um PE pode ser determinada por todas as suas distribuições finitas. Em particular, se todas as distribuições de x forem gaussianas, o PE diz-se *gaussiano*.

Exemplo 4 - Processo Wiener $W(t)$:

O processo Wiener $W(t)$, que modela o movimento Browniano, tem as seguintes propriedades:

1. $W(0) = 0$.
2. Para qualquer $t_1 < t_2$, a diferença $W(t_2) - W(t_1)$ tem distribuição normal $\mathcal{N}(0, t_2 - t_1)$.
3. Os incrementos $W(t_2) - W(t_1)$ são independentes para intervalos não sobrepostos.
4. $W(t)$ tem caminhos contínuos e não diferenciáveis em quase todos os pontos.

Se $\{W(t)\}$ for um processo de Wiener, então para $t_1 < t_2$, o incremento $W(t_2) - W(t_1)$ é uma variável aleatória com distribuição normal $\mathcal{N}(0, t_2 - t_1)$.

Se quisermos modelar o movimento de uma partícula sujeita a pequenas colisões aleatórias, podemos usar o processo de Wiener para descrever a posição da partícula em cada instante de tempo t □

1.3 Média e Covariância

Acabamos de ver que os PE são conjuntos de v.a.. Dado estas últimas serem frequentemente caracterizadas pelos seus momentos de primeira e segunda ordem, é natural que o mesmo aconteça com os PE. Definimos então:

$$M(t_1, t_2, \dots, t_k) = E\{x(t_1), x(t_2), \dots, x(t_k)\} = \quad (1.1) \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdots x_k p_{t_1, t_2, \dots, t_k}(x_1, \dots, x_k) dx_1 dx_2 \cdots dx_k$$

como *momento de ordem k* do PE $\{x(t)\}$, onde $E\{\cdot\}$ é a esperança matemática. No que se segue, como aliás referimos atrás, estamos sobretudo interessados nos momentos de primeira e segunda ordem, que são respectivamente designados por *função média* e *função (auto-)covariância*. A média é representada por:

$$\mu_x(t) = E\{x(t)\} = M_x(t) \quad (1.2)$$

e a covariância por:

$$\Lambda_{xx}(t, s) = E\{(x(t) - \mu_x(t))(x(s) - \mu_x(s))\} = M_x(t, s) - M_x(t)M_x(s). \quad (1.3)$$

Outra representação para a covariância é $cov\{x(t), x(s)\}$. Em particular:

$$\sigma_x^2(t) = cov\{x(t), x(t)\} = \Lambda_{xx}(t, t) \quad (1.4)$$

que designamos por *variância* de $\{x(t)\}$.

PE com variância finita dizem-se PE de segunda ordem.

Exemplo 5 - Função de densidade de probabilidade dum PE gaussiano:

Seja $\{x(t), t = 0, \pm 1, \pm 2, \dots\}$ um PE gaussiano. Sejam $\mu_x(t)$ e $\sigma(t, s)$ a média e a covariância, respectivamente. Então a função de densidade de probabilidade conjunta de $x(t_1), \dots, x(t_k)$ é expressa como :

$$p_{t_1, t_2, \dots, t_k}(x_1, \dots, x_k) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^k \Sigma_{i,j}^{-1} (x_i - \mu_x(t_i))(x_j - \mu_x(t_j)) \right\}$$

onde $\Sigma = (\sigma(t_i, t_j)) \in \mathbb{R}^{k \times k}$ é a matriz de covariâncias, e $\Sigma_{i,j}^{-1}$ designa o elemento (i, j) da matriz inversa Σ^{-1} . Se $x(t)$ for ruído branco temos então que Σ é uma matriz diagonal. $\square$

Consideremos agora $\{x(t), t = 0, \pm 1, \pm 2, \dots\}$ um vector de PE de dimensão n , i.e.,

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \in \mathbb{R}^n$$

onde $x_i(t), i = 1, \dots, n$ são PE escalares. Posto isto, definimos então a média vectorial (ou média de $x(t)$) como:

$$\mu_x(t) = Ex(t) = \begin{pmatrix} Ex_1(t) \\ \vdots \\ Ex_n(t) \end{pmatrix} = \begin{pmatrix} \mu_{x_1}(t) \\ \vdots \\ \mu_{x_n}(t) \end{pmatrix} \quad (1.5)$$

bem como a covariância vectorial (ou covariância de $x(t)$):

$$\begin{aligned} \Lambda_{xx}(t, s) &= E\{(x(t) - \mu_x(t))(x(s) - \mu_x(s))^T\} = \\ &= \begin{pmatrix} E\{\tilde{x}_1(t)\tilde{x}_1(s)\} & E\{\tilde{x}_1(t)\tilde{x}_2(s)\} & \cdots & E\{\tilde{x}_1(t)\tilde{x}_n(s)\} \\ E\{\tilde{x}_2(t)\tilde{x}_1(s)\} & E\{\tilde{x}_2(t)\tilde{x}_2(s)\} & \cdots & E\{\tilde{x}_2(t)\tilde{x}_n(s)\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{\tilde{x}_n(t)\tilde{x}_1(s)\} & E\{\tilde{x}_n(t)\tilde{x}_2(s)\} & \cdots & E\{\tilde{x}_n(t)\tilde{x}_n(s)\} \end{pmatrix} \end{aligned} \quad (1.6)$$

sendo $\tilde{x}(t) = x(t) - \mu_x(t)$. Os elementos da diagonal de $\Lambda_{xx}(t, s)$ são as funções de covariância de $\{x_i(t), t = 0, \pm 1, \pm 2, \dots\}$, sendo os restantes as funções de covariância cruzada entre $\{x_i(t)\}$ e $\{x_j(t)\}, i \neq j$.

1.4 Processos estocásticos estacionários

Um PE diz-se *estacionário* se não varia com o tempo. Isto significa que a função de densidade de probabilidade conjunta só depende da diferença temporal entre as amostras, sendo independente do intervalo de tempo entre os instantes em que são consideradas. Estatisticamente, isto equivale a afirmar que o futuro é igual ao passado, e pode ser expresso através da função de densidade de probabilidade conjunta:

$$p_{t_1, t_2, \dots, t_k}(x_1, x_2, \dots, x_k) = p_{t_1 + \tau, t_2 + \tau, \dots, t_k + \tau}(x_1, x_2, \dots, x_k), \quad \tau = 0, \pm 1, \pm 2, \dots \quad (1.7)$$

Se esta condição se verificar para qualquer valor de τ , dizemos ser o PE *fortemente estacionário*. Num PE fortemente estacionário, com momento de segunda ordem, a média é constante e a covariância depende da diferença entre 2 instantes, i.e.,

$$\mu_x(t) = E\{x(t)\} = \mu_x(0) = \mu_x \quad (1.8)$$

$$\Lambda_{xx}(t, s) = \Lambda_{xx}(t - s, 0) = \Lambda_{xx}(t - s) \quad (1.9)$$

Exemplo 6 - Ruído branco:

Um PE $\{v(t), t = 0, 1, \dots\}$ diz-se *ruído branco* se as v.a. $v(t)$ e $v(s)$ forem igualmente distribuídas e independentes entre si para qualquer $t \neq s$, i.e.,

$$p_{t,s}(v(t), v(s)) = p_t(v(t)) p_s(v(s)), \quad t \neq s$$

Como $v(t)$ e $v(s)$ são igualmente distribuídas, então $\mu_v(t) = E\{v(s)\} = \mu_v$,

$$M_v(t, t) = E\{v(t)^2\} = \mu_v^2 + \sigma_v^2$$

$$M_v(t, s) = E\{v(t)v(s)\} = E\{v(t)\} E\{v(s)\} = \mu_v^2$$

Temos ainda que

$$\begin{aligned} \Lambda_{xx}(t, s) &= E\{(v(t) - \mu_v)(v(s) - \mu_v)\} \\ &= E\{v(t)v(s) - \mu_v v(s) - \mu_v v(t) + \mu_v^2\} \\ &= E\{v(t)v(s)\} - \mu_v E\{v(t)\} - \mu_v E\{v(s)\} + \mu_v \mu_v \\ &= E\{v(t)v(s)\} - \mu_v^2 \\ &= \begin{cases} 0 & , \quad t \neq s \\ E\{v(t)^2\} - \mu_v^2 & , \quad t = s \end{cases} \end{aligned}$$

Sendo só dependente de $t - s$, Logo, o ruído branco é um PE estacionário.

O ruído branco é um sinal imprevisível, pois, o seu valor, num dado instante, é independente de todo o passado e futuro. A aleatoriedade do ruído branco é ilustrada nas figuras 1.3 e 1.4 que mostram realizações de ruído branco gaussiano e ruído branco binário, respetivamente $\square$

No entanto, existem PEs com distribuições diferentes em instantes de tempo diferentes mas com momentos de primeira e segunda ordem estacionários¹. Dizemos que

¹Um momento de primeira ordem é estacionário se não variar com o tempo. Um momento de segunda ordem de segunda ordem é estacionário se só depender do espaçamento entre os instantes de tempo

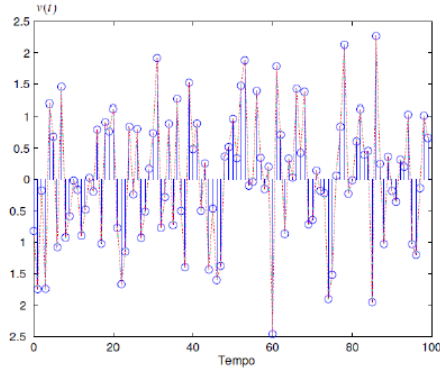


Figura 1.3: Realização de ruído branco gaussiano de média nula.

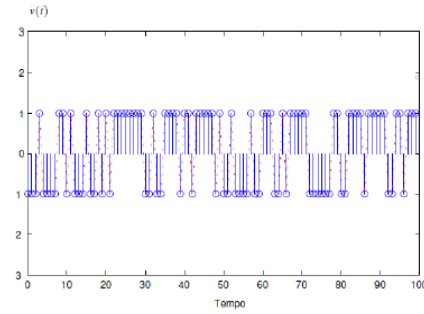


Figura 1.4: Realização de ruído branco binário de média nula.

estes PEs são *fracamente estacionários*. É óbvio que os PEs fortemente estacionários também são fracamente estacionários, não sendo no entanto o recíproco verdadeiro.

Exemplo 7 - Caminho aleatório:

O ruído branco é usado, de forma adequada, para gerar diferentes processos com diferentes propriedades estocásticas, como por exemplo, o caminho aleatório $x(t)$, que é expresso como uma soma de ruídos brancos:

$$x(t) = v(1) + v(2) + \dots + v(t), \quad t = 1, 2, \dots \quad (1.10)$$

com $x(0) = 0$. Podemos ainda escrever (1.10) recursivamente como:

$$x(t) = x(t-1) + v(t), \quad \text{com } x(0) = 0 \quad (1.11)$$

Calculemos a sua média e variância.

Considerando $\{v(t)\}$ um PE gaussiano de média zero e variância σ_v^2 , temos que $E\{v(t)\} = 0$ e $E\{v(t)v(s)\} = \delta(t-s)\sigma_v^2$, em que $\delta(t)$ é o impulso de Dirac. Podemos então calcular a média do passeio aleatório como:

$$\mu_x(t) = E\{v(1) + v(2) + \dots + v(t)\} = 0.$$

No que diz respeito à covariância, considerando $t > s$, $x(t) - x(s) = \sum_{i=s+1}^t v(i)$ e

$x(s) = \sum_{i=1}^s v(i)$ são independentes, temos então:

$$\begin{aligned} \Lambda_{xx}(t, s) &= E\{x(t)x(s)\} = E\{[x(t) - x(s)]x(s) + x(s)^2\} = E\{x(s)^2\} = \\ &= \sum_{i=1}^s E\{v(i)^2\} = s\sigma_v^2 \end{aligned}$$

De igual forma, quando $t < s$, temos que $\Lambda_{xx}(t, s) = t\sigma_v^2$. Donde se conclui que a covariância de um caminho aleatório é $\Lambda_{xx}(t, s) = \min(t, s)\sigma_v^2$, que não é de forma alguma uma função de $t - s$; donde se conclui não ser o caminho aleatório um PE estacionário.

A Figura 1.5 mostra uma realização dum passeio aleatório. □

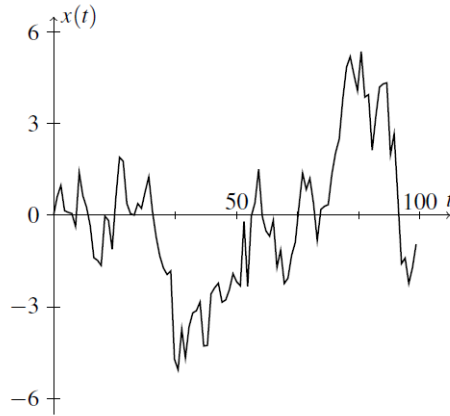


Figura 1.5: Realização dum passeio aleatório.

Exemplo 8 - Relação entre passeio aleatório e o processo de Wiener:

Consideremos um que uma partícula faz caminho aleatório num reta horizontal com n passos em intervalos de tempo $\Delta t = \frac{T}{n}$, onde T é o tempo total. Em cada passo, a partícula move-se da direita para a esquerda ou da esquerda para a direita com uma distribuição normal $\mathcal{N}(0, \Delta t)$ tal que $n\Delta t = T$, ou seja, a magnitude do passo é ajustada para que a variância do deslocamento total num intervalo de tempo com duração T seja finita. A posição ao fim de n passos (ou ao fim do intervalo de tempo T) é

$$X_n = \sum_{i=0}^n \chi_i, \quad \chi_i \sim \mathcal{N}\left(0, \frac{T}{n}\right). \quad (1.12)$$

À medida que o número de passos n aumenta, a variância da distância percorrida em cada passo $\Delta t = \frac{T}{n}$ diminui. Quando $n \rightarrow \infty$ e, consequentemente o $\Delta t \rightarrow 0$, o

caminho aleatório normalizado aproxima-se do movimento Browniano (processo de Wiener) $W(t) = \lim_{n \rightarrow \infty} X_n$. $\square$

Sendo $\{x(t)\}$ um processo estacionário de segunda ordem, sua média é constante, igual a μ_x . Temos sem perda de generalidade que $x(t) - \mu_x$, é um PE de média nula pelo que iremos sempre considerar, daqui para a frente, que $E\{x(t)\} = 0$. Por outro lado, sendo a sua função de covariância unicamente dependente de $t - s$, iremos expressá-la na forma

$$\Lambda_{xx}(\tau) = E\{x(t + \tau)x(t)\}, \tau = 0, \pm 1, \pm 2, \dots \quad (1.13)$$

Lema 1. *A função covariância de um processo estacionário $\{x(t)\}$ tem as seguintes propriedades:*

- (i) (Limite) $|\Lambda_{xx}(\tau)| \leq \Lambda_{xx}(0) = \sigma_x^2, \quad \tau = 0, \pm 1, \pm 2, \dots$
- (ii) (Simetria) $\Lambda_{xx}(\tau) = \Lambda_{xx}(-\tau)$.
- (iii) (Não negatividade) $\sum_{i,k=1}^n a_i a_k \Lambda_{xx}(\tau_i - \tau_k) \geq 0$

Estas propriedades permitem-nos testar se um processo é estacionário. Caso não se verifiquem temos a certeza de estar perante um processo não estacionário.

Exemplo 9 - Funções de covariância válidas e inválidas:

Das três sequências discretas representadas na Figura 1.6 só a de cima é que obedece às três propriedades do Lema 1 e, consequentemente, só esta é que pode ser a função de covariância dum processo estocástico estacionário. $\square$

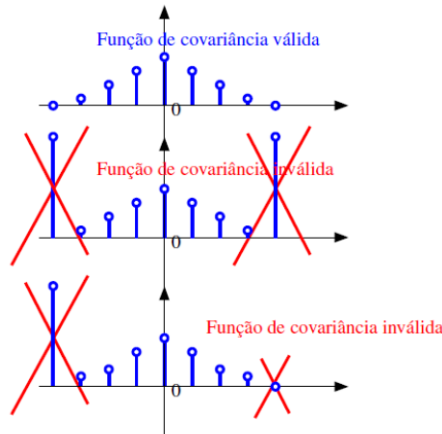


Figura 1.6: Sequências candidatas a função de covariância

Consideremos o PE conjunto $\{x(t), y(t), t = 0, \pm 1, \pm 2, \dots\}$ de média nula.

Os processos $\{x(t)\}$ e $\{y(t)\}$ dizem-se *juntamente estacionários* se, considerando o vector PE $w(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, este for um PE estacionário. Dada a matriz de covariância de w :

$$\begin{aligned} \Lambda_{ww}(t + \tau, t) &= \begin{pmatrix} \Lambda_{xx}(t + \tau, t) & \Lambda_{xy}(t + \tau, t) \\ \Lambda_{yx}(t + \tau, t) & \Lambda_{yy}(t + \tau, t) \end{pmatrix} = E \left\{ \begin{pmatrix} x(t + \tau) \\ y(t + \tau) \end{pmatrix} (x^T(t) y^T(t)) \right\} \\ &= \text{cov} \{w(t + \tau)w(t)\}, \end{aligned} \quad (1.14)$$

a estacionaridade implica que as covariâncias de (1.14) sejam somente função da diferença τ .

A $\Lambda_{xy}(\tau) = E \{x(t + \tau)y(t)\}$ chama-se *covariância cruzada* entre x e y .

Lema 2. A covariância cruzada, $\Lambda_{xy}(\tau)$, tem as seguintes propriedades:

- (i) (*Anti-simetria*) $\Lambda_{xy}(\tau) = \Lambda_{yx}(-\tau)$
- (ii) (*Limite*) $|\Lambda_{xy}(\tau)| \leq \Lambda_{xx}(0)\Lambda_{yy}(0), \quad \tau = \pm 1, \pm 2, \dots, 9$

1.5 Processos ergódicos

Um dos grandes desafios da análise de PEs é estimar os seus parâmetros estocásticos através da observação dos dados ao longo do tempo. Dado os parâmetros estarem relacionados com os valores esperados de alguma função de um PE, estudamos então o problema da estimação da média e covariância de um PE estacionário.

Suponhamos $\{x(t), t = 0, \pm 1, \pm 2, \dots\}$ um PE estacionário de segunda ordem. Definimos a média temporal deste PE como

$$m_x(N) = \frac{1}{2N + 1} \sum_{t=-N}^N x(t), \quad (1.15)$$

e correlação temporal como

$$r_{xx}(\tau, N) = \frac{1}{2N + 1} \sum_{t=-N}^N x(t + \tau)x(t), \quad \tau = 0, \pm 1, \pm 2, \dots \quad (1.16)$$

Se $\{x(t), t = 0, \pm 1, \pm 2, \dots\}$ tiver média nula, esta quantidade é também designada *função de (auto-)covariância*. As funções média (1.8) e covariância (1.9) são definidas como médias de um conjunto, i.e, como valores esperados. Mas (1.16) e (1.15) são definidas como médias temporais de uma realização.

$$x = (\dots, x(-1), x(0), x(1), \dots)$$

Aquando da análise de dados, nós lidamos antes com uma realização, gerada por uma experiência, e não com um conjunto de dados. Posto isto, sob o ponto de vista prático, a definição de momentos em termos de média temporal é preferível à média definida por um conjunto de dados. Falta agora saber se a média temporal, $r_{xx}(\tau)$, é ou não igual à média do conjunto $\Lambda_{xx}(\tau)$.

Todo o PE cujos parâmetros estatísticos são estimados de forma precisa a partir de uma simples realização, diz-se *processo ergódico*. Ou seja, num processo ergódico as médias temporais convergem para as esperanças matemáticas e são utilizadas como estimativas destas. No que se segue enunciamos teoremas ergódicos para a esperança e a covariância.

Definição 1. Diz-se que um processo estacionário tem média ergódica se

$$\lim_{N \rightarrow \infty} E \left\{ [m_x(N) - \mu_x]^2 \right\} = 0,$$

em que

$$\begin{aligned} \mu_x &= E \{x(t)\} \\ m_x(N) &= \frac{1}{2N+1} \sum_{t=-N}^N x(t), \end{aligned}$$

ou seja se $m_x(N)$ convergir em média quadrática para μ_x .

Definição 2. Um PE ergódico, $x(t)$, de média nula tem covariância ergódica se

$$E \{r_{xx}(\tau) - \Lambda_{xx}(\tau)\} = 0$$

em que

$$\begin{aligned} \lambda_{xx}(\tau) &= E \{x(t)x(t-\tau)\} \\ r_{xx}(\tau) &= \frac{1}{2N+1} \sum_{t=-N}^N E \{x(t+\tau)x(t)\}. \end{aligned}$$

Maior detalhe sobre as condições de ergodicidade da média e da covariância podem ser vistas em T. Katayama. A partir deste momento só serão considerados PE com média e covariância ergódicas.

1.6 Densidade espectral

Se a função de covariância dum processo estacionário $x(t)$ for absolutamente somável, isto é, se

$$\sum_{\tau=-\infty}^{\infty} |\Lambda_{xx}(\tau)| < \infty,$$

então terá uma transformada de Fourier. Essa transformada é designada como *densidade espectral* de $x(t)$.

Definição 3. Se $x(t)$ for um PE estacionário com função de covariância absolutamente somável, então

$$\Phi_{xx}(\omega) = \sum_{\tau=-\infty}^{\infty} \Lambda_{xx}(\tau) e^{-j\omega\tau}$$

é a sua densidade espectral.

Como $\Phi_{xx}(\omega)$ é a transformada de Fourier de $\Lambda_{xx}(\tau)$, $\Lambda_{xx}(\tau)$ pode ser determinado através da transformada inversa de Fourier:

$$\Lambda_{xx}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{xx}(\omega) e^{j\omega\tau} d\omega$$

Vemos, assim, que existe uma relação de um para um entre $\Lambda_{xx}(\tau)$ e $\Phi_{xx}(\omega)$, pelo que ambas as funções transportam a mesma informação. Então, para $\tau = 0$, temos:

$$\Lambda_{xx}(0) = \sigma_x^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{xx}(\omega) d\omega.$$

Considerando que $x(t)$ é ergódico podemos escrever

$$\Lambda_{xx}(0) = r_{xx}(0) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{t=-N}^N x(t)^2 = P_x = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{xx}(\omega) d\omega,$$

em que P_x significa potência de $x(t)$. Vemos, deste modo, que Φ_{xx} representa a distribuição da potência de $x(t)$ pelo espectro $\omega \in]-\pi, \pi]$ e daí o nome de densidade espectral.

Lema 3. A densidade espectral $\Phi_{xx}(\omega)$ tem as seguintes propriedades:

(i) $\Phi_{xx}(\omega) = \Phi_{xx}(-\omega)$ (simetria par)

(ii) $\Phi_{xx}(\omega) \geq 0$ (não negativa)

Demonstração. A propriedade (i) resulta do facto de $\lambda_{xx}(\tau)$ ser uma sequência par e (ii) é consequência de que a variância de $x(t)$, $\sigma_x(t) = \int_{-\pi}^{\pi} \Phi_{xx}(\omega) d\omega$ ser não negativa e do Lema 4. Se Φ_{xx} fosse negativo numa banda de frequências $\omega_1 < \omega < \omega_2$, seria possível obter, através duma filtragem adequada², um sinal com variância negativa, o que é uma impossibilidade. $\square$

Exemplo 10 - Densidade espectral do ruído branco

Seja $e(t)$ ruído branco. Então, a sua sequência de covariância é

$$\lambda_{ee}(\tau) = \begin{cases} \sigma_e^2, & \tau = 0 \\ 0, & \tau \neq 0, \end{cases}$$

sendo a sua densidade espectral

$$\Phi_{ee}(\omega) = \sum_{\tau=-\infty}^{\infty} \lambda_{ee}(\tau) e^{-j\omega\tau} = \lambda_{ee}(0) = \sigma_e^2. \quad (1.17)$$

Logo, o ruído branco tem uma densidade espectral constante, podendo-se afirmar que, tal como a luz branca, tem um espectro plano (ver Figura 1.6). Daí o nome de ruído branco. $\square$

É muito frequente os PE propagarem-se através de sistemas dinâmicos. Veremos, em seguida, o que acontece a um processo quando se propaga através de um sistema linear invariante no tempo e estável.

²Por exemplo, um filtro passa-banda ideal com banda de passagem $\omega_1 < \omega < \omega_2$.

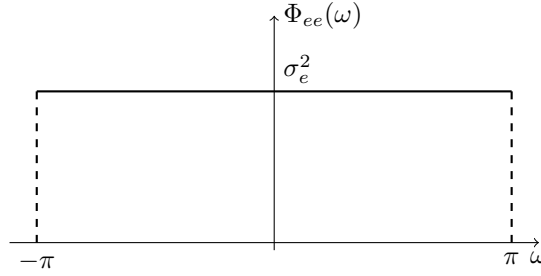


Figura 1.7: Densidade espectral do ruído branco.

Lema 4. *Se a entrada de um sistema linear e invariante no tempo com função de transferência $G(z)$ estável for um PE estacionário de 2ª ordem e média nula $\{u(t)\}$, a saída $\{y(t)\}$ também é um PE estacionário de 2ª ordem e média nula com densidade espectral*

$$\Phi_{yy}(\omega) = |G(e^{j\omega})|^2 \Phi_{uu}(\omega).$$

Demonstração. Ver T. Katayama. □

É comum modelar um PE estacionário como a saída dum sistema causal linear e invariante no tempo, estável e de fase mínima, excitado por ruído branco. Este procedimento é fundamentado pelo seguinte teorema:

Teorema 1 (Factorização da densidade espectral). *A função de densidade espectral $\Phi_{yy}(\omega)$ pode ser factorizada na forma*

$$\Phi_{yy}(\omega) = H(e^{j\omega}) H(e^{-j\omega}) \sigma^2$$

em que

$$H(z) = \sum_{k=0}^{\infty} h(k) z^{-k}$$

é uma função de transferência estável com $h(0) = 1$ e com todos os zeros no interior do círculo unitário.

Demonstração. Ver T. Katayama. □

Este teorema e o Lema 4 permitem modelar qualquer PE como a saída de um sistema excitado por ruído branco. Veremos alguns desses modelos na secção seguinte.

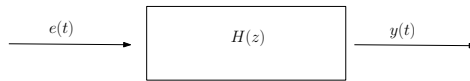


Figura 1.8: Processo estocástico estacionário, com $e(t)$ ruído branco.

Antes, porém, convém referir que o termo densidade espectral também é utilizado para designar a transformada z da sequência de autocorrelação, ou seja

$$\Phi_{xx}(z) = \sum_{\tau=-\infty}^{\infty} A_{xx}(\tau) z^{-\tau}.$$

1.7 Previsão linear em PE

Seja $\{y(t)\}$ um PE de média nula, estacionário e de segunda ordem. O nosso problema é prever $y(t+m)$, $m = 1, 2, \dots$ a partir de $y(t), y(t-1), \dots$. De acordo com o Teorema 1, podemos considerar que $y(t)$ é a saída dum sistema linear e invariante no tempo com função de transferência

$$H(z) = \sum_{k=0}^{\infty} h(k)z^{-k} \quad (1.18)$$

com $h(0) = 1$. Tem-se, então

$$y(t) = \sum_{k=0}^{\infty} h(k)e(t-k) = e(t) + h(1)e(t-1) + \dots \quad (1.19)$$

e

$$\begin{aligned} y(t+m) &= e(t+m) + h(1)e(t+m-1) + \dots + h(m-1)e(t+1) \\ &\quad + h(m)e(t) + h(m+1)e(t-m) + \dots \\ &= \sum_{k=0}^{m-1} h(s)e(t+m-k) + \sum_{k=0}^{\infty} h(m+k)e(t-k). \end{aligned} \quad (1.20)$$

Se definirmos o operador q tal que

$$qx(t) = x(t+1) \quad (1.21)$$

podemos reescrever as três últimas equações na forma

$$H(q) = \sum_{k=0}^{\infty} h(k)q^{-k} \quad (1.22)$$

$$y(t) = \sum_{k=0}^{\infty} h(k)q^{-k}e(t) = H(q)e(t) \quad (1.23)$$

$$\begin{aligned} y(t+m) &= H(q)q^m e(t) \\ &= \sum_{k=0}^{m-1} h(k)q^{m-k}e(t) + \sum_{k=0}^{\infty} h(m+k)q^{-k}e(t) \end{aligned} \quad (1.24)$$

Como $e(t)$ é ruído branco, a melhor previsão de $e(t+1), \dots, e(t+n)$ são os seus valores médios que são iguais a zero. Logo, $\hat{y}(t+1)$, a melhor previsão linear de $y(t+m)$ no instante t é

$$\hat{y}(t+m) = \sum_{k=0}^{\infty} h(m+k)q^{-k}e(t). \quad (1.25)$$

De (1.24),

$$\sum_{k=0}^{\infty} h(m+k)q^{-k}e(t) = H(q)q^m e(t) - \sum_{k=0}^{m-1} h(k)q^{m-k}e(t) \quad (1.26)$$

e, consequentemente,

$$\hat{y}(t+m) = \left[H(q)q^m - \sum_{s=0}^{m-1} h(s)q^{m-s} \right] e(t) \quad (1.27)$$

Como, segundo o Teorema 1, $H(q)$ tem todos os zeros no interior do círculo unitário, então é inversamente estável e, dada a sequência $\dots, y(t-1), y(t)$, a sequência $\dots, e(t-1), e(t)$ pode ser calculada através de

$$e(t) = H^{-1}(q)y(t) = \sum_{k=0}^{\infty} \bar{h}(k)q^{-k}y(t), \quad (1.28)$$

em que $\bar{h}(k)$, $k = 0, \dots, \infty$ é a resposta impulsional de $H(z)^{-1}$. Logo

$$\hat{y}(t+m) = \left[H(q)q^m - \sum_{s=0}^{m-1} h(s)q^{m-s} \right] H^{-1}(q)y(t). \quad (1.29)$$

De (1.24) e (1.25), o erro de previsão é

$$\begin{aligned} \hat{e}(t+m) &= y(t+m) - \hat{y}(t+m) \\ &= \sum_{k=0}^{m-1} h(k)q^{m-k}e(t) + \sum_{k=0}^{\infty} h(m+k)q^{-k}e(t) - \sum_{k=0}^{\infty} h(m+k)q^{-k}e(t) \\ &= \sum_{k=0}^{m-1} h(k)q^{m-k}e(t) \\ &= e(t+m) + h(1)e(t+m-1) + \dots + h(m-1)e(t+1) \end{aligned}$$

Na perspectiva da identificação de sistemas, a previsão mais interessante acontece para $m = 1$ que pode ser obtida através de (1.29) fazendo $m = 1$ e recordando que $h(0) = 1$:

$$\hat{y}(t+1) = [H(q)q - 1]e(t) = [H(q) - 1]qe(t) = [H(q) - 1]e(t+1)$$

como

$$e(t) = H^{-1}(q)y(t)$$

então

$$\begin{aligned} \hat{y}(t+1) &= [H(q) - 1]H^{-1}(q)y(t+1) \\ \Leftrightarrow \hat{y}(t) &= [H(q) - 1]H^{-1}(q)y(t) \end{aligned}$$

sendo o erro de previsão $y(t) - \hat{y}(t) = e(t)$.

Exemplo 11

O PE gerado por

$$y(t) = u(t) - 2u(t-1)$$

onde $u(t)$ é ruído branco com média nula e variância unitária, tem um espectro generalizado

$$\Phi_{yy}(q) = G(q)G(q^{-1})\sigma_u^2$$

onde

$$G(q) = \frac{Y(q)}{U(q)} = 1 - 2q^{-1} = \frac{q-2}{q}$$

Se suposermos que $u(t)$ é desconhecido, podemos tentar prever $y(t)$ no instante $t-1$ através de

$$\hat{y}(t) = (G(q) - 1)G^{-1}(q)y(t) = -\frac{2q^{-1}}{1-2q^{-1}}y(t) = -\frac{2}{q-2}y(t) = G_p(z)y(t).$$

No entanto este previsor não funciona, pois não é estável ($G_p(z)$ tem um zero fora do círculo unitário). A Figura 1.9 ilustra inutilidade deste previsor, onde se pode constatar que o erro de previsão cresce exponencialmente. Como podemos então obter o melhor

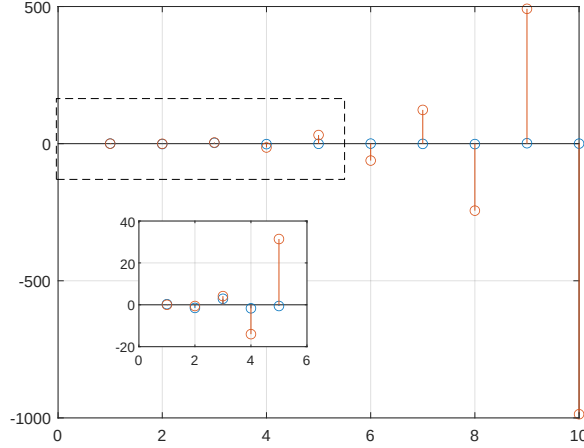


Figura 1.9: Realização de $y(t) = (1 - 2q^{-1})u(t)$ (azul) e previsão $\hat{y}(t) = -2q^{-1}/(1-2q-1)u(t)$ (vermelho). Pode-se ver na ampliação das seis primeiras amostras que, a partir de $t=4$, o erro de previsão cresce exponencialmente.

previsor linear de $y(t)$?

Para responder a esta pergunta, começamos por tentar factorizar o espectro na forma

$$\Phi_{yy}(q) = H(q)H(q^{-1})\sigma_e^2$$

em que $H(q)$ é uma função de transferência com todos os pólos e zeros no interior do círculo unitário e $h(0) = 1$. Os pólos e zeros de $H(q^{-1})$ são os inversos dos de $H(q)$ e, como tal, estão todos no exterior do círculo unitário. Uma vez obtida esta factorização, assumimos ser $y(t)$ gerado pelo modelo

$$y(t) = H(z)e(t)$$

a partir do qual podemos dimensionar o previsor:

$$\hat{y}(t) = (H(q) - 1)H^{-1}(q)y(t),$$

pois agora $H^{-1}(q)$ é estável. Notemos que a estabilidade de $H^{-1}(q)$ é fundamental para que possamos prever $e(t)$ a partir de $y(t)$. O processo que temos vindo a considerar tem um espectro

$$\Phi_{yy}(q) = G(q)G(q^{-1}) = (1 - 2q^{-1})(1 - 2q) = H(q)H(q^{-1})\sigma^2.$$

Como o zero de $G(q)$ está no exterior do círculo unitário, o de $H(q)$ tem que ser o seu inverso, isto é,

$$H(q) = 1 - 0.5q^{-1},$$

sendo

$$\begin{aligned}\Phi_{yy}(\omega) &= (1 - 0.5q^{-1})(1 - 0.5q)\sigma^2 = (1 - 2q^{-1})(1 - 2q) \\ &= (1 - 0.5q)(1 - 0.5q^{-1})(-2q^{-1})(2q) \\ &= (1 - 0.5q)(1 - 0.5q^{-1})4,\end{aligned}$$

pelo que $\sigma^2 = 4$. Temos então o seguinte modelo conceptual

$$y(t) = e(t) - 0.5e(t-1)$$

em que $e(t)$ é ruído branco de média nula e variância $\sigma^2 = 4$. O melhor previsor linear a um passo é, finalmente,

$$\hat{y}(t) = [H(q) - 1]H(q^{-1})y(t) = -\frac{0.5q^{-1}}{1 - 0.5q^{-1}}y(t) = H_p(q)y(t).$$

Como o erro de previsão é igual a $e(t)$, a sua variância é $\sigma^2 = 4$ sendo, por isso, maior do que a que, erradamente, estávamos à espera quando tentámos prever $y(t)$ através de $G_p(q)y(t)$. Notemos que a variância de $e(t)$ é igual à de $u(t)$, multiplicada por um factor igual ao quadrado do zero de $G(q)$ que reflectimos para o interior do círculo unitário quando formamos $H(q)$. Se tivéssemos reflectido mais do que um zero, o factor que multiplicaria σ_u^2 seria o quadrado do produto do módulo dos zeros reflectidos. A Figura 1.10 compara uma realização de $y(t)$ com a previsão $\hat{y}(t) = H_p(q)y(t)$ onde se pode constatar que $\hat{y}(t)$ consegue prever $y(t)$, embora com algum erro. A variância

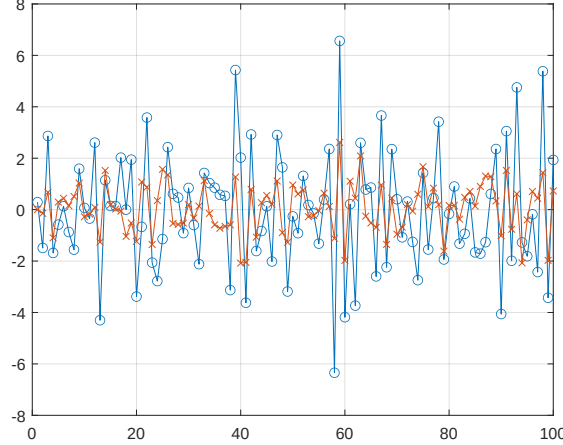


Figura 1.10: Realização de $y(t) = (1 - 2q^{-1})u(t)$ (azul) e previsão $\hat{y}(t) = -0.5q^{-1}/(1 - 0.5q^{-1})y(t)$ (vermelho).

de $y(t)$ é

$$E\{y^2(t)\} = E\{u^2(t) - 2u^2(t-1)\} = E\{u^2(t)\} + 4E\{u^2(t-1)\} = 1 + 4 = 5.$$

Logo, Prevendo $y(t)$ através de $E\{y(t)\} = 0$ obtemos um erro com variância 5 (desvio padrão) $\square$

1.8 Modelos de entrada-saída de processos estocásticos

1.8.1 Modelo de Média móvel (MA)

O Teorema 1 permite-nos encarar um PE como a saída de um sistema causal, linear e invariante no tempo, estável e de fase mínima (sem zeros fora do círculo unitário). O Lema 4 mostra-nos ser a condição de estabilidade absolutamente essencial para a estacionaridade do processo. Veremos adiante que a condição de fase mínima é fundamental para se poder fazer a sua previsão. Neste contexto, um PE $\{y(t)\}$ pode ser descrito pelo modelo:

$$y(t) = \sum_{k=0}^{\infty} h(k)e(t-k) \quad \text{e} \quad h(0) = 1$$

em que os pólos e zeros de

$$H(z) = \sum_{k=0}^{\infty} h(k)z^{-k}$$

estão todos no interior do círculo unitário e $e(t)$ é ruído branco com média e variância

$$E\{e(t)^2\} = \sigma^2.$$

Se $h(k) = 0$ para $k > n$ e se designarmos $h(k)$ por c_k , o modelo será

$$y(t) = \sum_{k=0}^n c_k e(t-k)$$

Definindo z^{-1} como o operador atraso, ou seja $z^{-1}x(t) = x(t-1)$, podemos re-escrever esta equação na forma

$$y(t) = C(z^{-1})e(t)$$

em que

$$C(z^{-1}) = 1 + c_1 z^{-1} + \dots + c_n z^{-n} = H(z)$$

é a função de transferência. Como neste modelo há uma regressão da entrada $e(t)$, é conhecido como modelo de média móvel (MA, *moving average*). Dado que função de transferência dum processo de média móvel é $H(z) = C(z^{-1})$, então o melhor previsor linear a um passo é

$$\hat{y}(t) = [C(q^{-1}) - 1] C^{-1}(q^{-1}) y(t) = \frac{c_1 q^{-1} + \dots + c_n q^{-n}}{1 + c_1 q^{-1} + \dots + c_n q^{-n}} y(t), \quad (1.30)$$

ou, equivalentemente,

$$\hat{y}(t) = c_1 y(t-1) + \dots + c_n y(t-n) - c_1 \hat{y}(t-1) + \dots + c_n \hat{y}(t-n). \quad (1.31)$$

Dado que num processo estocástico estacionário a função $H(z)$ é estável, então, se a resposta impulsional de $H(z)$ tiver duração infinita,

$$\lim_{t \rightarrow \infty} = 0. \quad (1.32)$$

Assim, qualquer PE pode ser aproximado por um processo de MA de ordem n , e o erro da aproximação diminui quando a ordem aumenta.

1.8.2 Modelo Auto-Regressivo (AR)

Os modelos MA são muito simples e permitem descrever qualquer PE. No entanto, para que tenham uma precisão aceitável, necessitam frequentemente de ter uma ordem elevada (n muito grande). Nestas condições é preferível descrever $y(t)$ por um modelo do tipo:

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = e(t)$$

que com

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_n q^{-n}$$

pode ser re-escrita como

$$A(q^{-1})y(t) = e(t).$$

Nestes modelos existe uma regressão em $y(t)$, pelo que são conhecidos como modelos auto-regressivos (AR, *auto-regressive*). A sua função de transferência é

$$H(z) = \frac{1}{A(z^{-1})}$$

e a sua ordem n costuma ser significativamente inferior à dos modelos MA.

O melhor previsor linear a um passo dum modelo AR é

$$\hat{y}(t) = \left[\frac{1}{A(q^{-1})} - 1 \right] A(q^{-1}) y(t) = [1 - A(q^{-1})] y(t), \quad (1.33)$$

ou seja

$$\hat{y}(t) = -a_1 y(t-1) - \dots - a_n y(t-n). \quad (1.34)$$

Este previsor é uma função linear dos parâmetros, o que explica em grande parte a sua popularidade, já que os parâmetros podem ser estimados de forma simples utilizando o método dos mínimos quadrados.

1.8.3 Modelo Auto-Regressivo de Média Móvel (ARMA)

Se utilizarmos regressão na entrada, $e(t)$, e na saída, $y(t)$, conseguimos modelos e menor ordem. Estes modelos são do tipo:

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = e(t) + c_1 e(t-1) + \dots + c_n e(t-n)$$

e podem ser escritos na forma compacta

$$A(q^{-1}) y(t) = C(q^{-1}) e(t).$$

São conhecidos modelos auto-regressivos de média móvel (ARMA). A sua função de transferência é:

$$H(z) = \frac{C(z^{-1})}{A(z^{-1})}, \quad (1.35)$$

sendo o melhor previsor linear dado por

$$\hat{y}(t) = \left[\frac{C(q^{-1})}{A(q^{-1})} - 1 \right] \frac{A(q^{-1})}{C(q^{-1})} y(t) = \frac{C(q^{-1}) - A(q^{-1})}{C(q^{-1})} y(t), \quad (1.36)$$

ou, equivalentemente,

$$\hat{y}(t) = (c_1 - a_1) y(t-1) + \dots + (c_n - a_n) y(t-n) - c_1 \hat{y}(t-1) - \dots - c_n \hat{y}(t-n). \quad (1.37)$$

1.8.4 Filtro branqueador

Os PE são conhecidos pelo nome dos modelos que os geram. Assim, os processos gerados por modelos ARMA são PE ARMA, os gerados por modelos AR são PE AR e os gerados por modelos MA são PE MA.

A invertibilidade de $H(z)$ requer que $\frac{1}{H(z)}$ seja uma função estável. Isto permite que se possa gerar ruído branco alimentando um sistema linear e invariante no tempo com função de transferência $\frac{1}{H(z)}$, com o processo $y(t)$. Por este motivo, $\frac{1}{H(z)}$ é

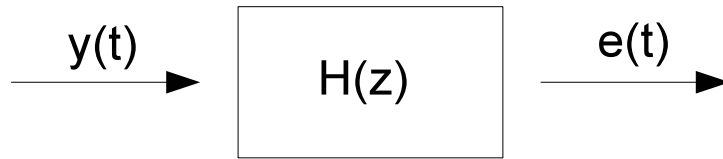


Figura 1.11: Processo estocástico estacionário, com $e(t)$ ruído branco.

também conhecido como filtro branqueador.

1.9 Modelos de PE no espaço de estados

Os PE estacionários de média nula também podem ser descritos por modelos do tipo:

$$\begin{aligned} x(t+1) &= Ax(t) + Ke(t) \\ y(t) &= Cx(t) + e(t) \end{aligned}$$

Estes modelos são conhecidos como modelo de inovação e têm a seguinte função de transferência:

$$H(z) = C(zI_n - A)^{-1}K + 1 = \frac{C(z^{-1})}{A(z^{-1})}. \quad (1.38)$$

O denominador desta função de transferência é a equação característica de A e é dado por

$$\det\{H(z)\} = \det(zI_n - A) = z^n A(z^{-1}) = z^n + a_1 z^{n-1} + \dots + a_n. \quad (1.39)$$

Se pelo menos um dos coeficientes de $C(z^{-1})$ for diferente de zero e todos os coeficientes de $A(z^{-1})$ forem nulos ($a_1 = \dots = a_n = 0$), então $y(t)$ é um processo MA. Por outro lado, se todos os coeficientes de $C(z^{-1})$ forem nulos ($c_1 = \dots = c_n = 0$) e pelo menos um coeficiente de $A(z^{-1})$ for não nulo, $y(t)$ caracteriza-se como um processo autorregressivo. Caso nenhuma dessas condições seja satisfeita, $y(t)$ será um processo ARMA.

Tomando $y(t)$ como entrada e $u(t)$ como saída obtém-se o modelo inverso que é dado por

$$\begin{aligned}x(t+1) &= Ax(t) + K(y(t) - Cx(t)) = (A - KC)x(t) + Ky(t) \\e(t) &= -Cx(t) + y(t)\end{aligned}$$

Daqui podemos determinar a função de transferência inversa:

$$H^{-1}(z) = -C(zI_n - (A - KC))^{-1}K + 1 \quad (1.40)$$

e também o melhor previsor linear a um passo:

$$x(t+1) = (A - KC)x(t) + Ky(t) \quad (1.41)$$

$$\hat{y}(t) = Cx(t). \quad (1.42)$$

Exercícios

1. **Definição de Processos Estocásticos:** Defina um processo estocástico e dê dois exemplos do mundo real onde ele pode ser aplicado. Descreva esses exemplos usando a notação de um processo estocástico.
2. **Continuidade do Parâmetro:** Explique a diferença entre um processo estocástico de parâmetro contínuo e um de parâmetro discreto. Dê um exemplo de cada.
3. **Realização de um Processo:** Para um dado processo estocástico $x(t, \omega)$, diga o que é uma *realização*. Considerando que t é o tempo e ω o dia da semana, explique como interpretaria uma realização.
4. **Distribuição de Probabilidade:** Considere um processo estocástico $x(t)$ nos instantes t_1, t_2, t_3 . Defina a função de distribuição conjunta de probabilidade $P(x(t_1), x(t_2), x(t_3))$.

Resolução:

Considere um processo estocástico $x(t)$ nos instantes t_1, t_2, t_3 . A função de distribuição conjunta de probabilidade $P(x(t_1), x(t_2), x(t_3))$ descreve a probabilidade de que o processo $x(t)$ tome valores específicos nos tempos t_1, t_2, t_3 . Ou seja, a função representa a probabilidade de que:

$$P(x(t_1) \leq x_1, x(t_2) \leq x_2, x(t_3) \leq x_3),$$

isto é, a probabilidade de que $x(t_1) = x_1, x(t_2) = x_2$, e $x(t_3) = x_3$.

5. **Distribuição de Probabilidade:** Explique a importância da função de distribuição conjunta de probabilidade.

Resolução:

A função de distribuição conjunta de probabilidade é fundamental para descrever o comportamento de um processo estocástico em múltiplos pontos no tempo. Sua importância pode ser vista nos seguintes aspectos:

- *Dependência temporal:* A distribuição conjunta revela a dependência entre os valores de $x(t)$ em diferentes instantes de tempo. Se os valores em t_1 , t_2 , e t_3 forem altamente correlacionados, essa correlação será refletida na função de distribuição conjunta.
- *Propriedades estatísticas:* A partir da distribuição conjunta, podemos calcular momentos conjuntos, como a covariância entre $x(t_1)$ e $x(t_2)$, permitindo a análise de variabilidade do processo.
- *Modelação de incerteza:* Muitos fenómenos físicos, económicos e de engenharia são modelados por processos estocásticos. A função de distribuição conjunta permite modelar o comportamento incerto desses sistemas ao longo do tempo.
- *Previsão futura:* A distribuição conjunta possibilita inferir probabilidades de eventos futuros, dado o comportamento passado ou presente, o que é crucial em diversas áreas, como controle de sistemas e processamento de sinais

6. **Processos Gaussianos:** Se um processo estocástico tem distribuições Gaussianas para todos os intervalos de tempo, ele é chamado de processo Gaussiano. Dê um exemplo de um processo Gaussiano e explique suas propriedades.

Resolução:

O processo Ornstein-Uhlenbeck (OU) é um processo gaussiano utilizado em diversas áreas, como finanças, física e biologia. É um estacionário e tem uma tendência a retornar a um valor médio ao longo do tempo. É descrito pela seguinte equação diferencial estocástica: $dX(t) = \theta[\mu - X(t)] + \sigma dW(t)$ em que $W(t)$ é o processo de Wiener padrão, isto é, um processo de Wiener em que os incrementos têm variância unitária. Como os incrementos do processo de Wiener padrão são independentes, esta equação é equivalente a $\frac{dX(t)}{dt} = -\theta X(t) + \theta\mu + \sigma\eta(t)$ em que $\eta(t)$ é ruído branco e $\eta(t) \sim \mathcal{N}(0, 1)$.

Os parâmetros de $X(t)$ são:

- $\theta > 0$ - Taxa de reversão da média;
- μ média para que o processo reverte
- σ volatilidade do processo

O processo OU pode ser discretizado no tempo com um período de amostragem T_s sendo descrito por

$$X(k+1) = (1 - \theta T_s)X(k) + \theta\mu T_s + \sigma T_s \eta(k)$$

. em que $X(k) = X(kT_s)$ e $\eta(k) = \eta(kT_s)$.

7. **Média:** Para um processo estocástico $x(t)$, explique como calcular sua função média $\mu_x(t)$. Dado um processo específico, calcule sua média ao longo de um intervalo de tempo.
8. **Função de Covariância:** Defina a função de covariância de um processo estocástico. Para um processo $x(t)$.
9. **Processo Estacionário:** Explique o que significa que um processo estocástico seja estacionário. Forneça um exemplo de um processo estacionário e discuta como suas propriedades estatísticas permanecem inalteradas ao longo do tempo.

10. **Ruído Branco:** Descreva as características do ruído branco. Calcule a covariância do ruído branco em dois instantes de tempo diferentes t_1 e t_2 .
11. **Passeio Aleatório:** Defina um processo de passeio aleatório e derive sua média e variância. Explique por que um caminho aleatório não é estacionário.
12. **Processo Ergódico:** Explique o conceito de um processo ergódico. Como ele difere de um processo estocástico geral? Dê um exemplo de um processo ergódico.
13. **Média dum processo ergódico:** Seja $\{x(t)\}$ um processo ergódico. Explique como é que se estima a média de $\{x(t)\}$ a partir duma realização.
14. **Covariância dum processo ergódico:** Seja $\{x(t)\}$ um processo ergódico. Explique como é que se estima a covariância de $\{x(t)\}$ a partir duma realização.
15. **Densidade Espectral:** O que é a densidade espectral de um processo? Calcule a densidade espectral para o ruído branco.
16. **Previsão em Processos Estocásticos:** Dado um processo estocástico $x(t)$, como é que preveria $x(t + 1)$ usando previsão linear?
17. **Modelo de Média Móvel (MA):** Descreva o modelo de média móvel (MA). Dado um conjunto de dados, ajuste um modelo MA de ordem 1 e interprete os resultados.
18. **Modelo Autorregressivo (AR):** Explique o modelo autorregressivo (AR). Ajuste um modelo AR(1) a uma série temporal e interprete os parâmetros.
19. **Modelo ARMA:** Combine as ideias dos modelos AR e MA para explicar um modelo ARMA. Forneça um exemplo onde um modelo ARMA(1,1) seria adequado.
20. **Covariância de um Passeio Aleatório:** Mostre que a covariância de um passeio aleatório nos instantes t e s é dada por $\min(t, s)\sigma^2$.
21. **Filtro Branqueador:** Explique o conceito de um filtro branqueador e como ele pode ser usado para converter um processo em ruído branco. Aplique este conceito a um processo dado.
22. **Modelo no Espaço de Estados:** Defina um modelo no espaço de estados para um processo estocástico. Dê um exemplo de tal modelo e descreva como ele pode ser usado para previsão.