DCO1013 - Probabilidade e processos estocásticos

Levy Gabriel da S. G.
Engenharia elétrica - UFRN

Lathi, B. P., Ding, Z., "Modern Digital and Analog Communication Systems", 4a edição, Editora Oxford University Press, 2009.

Para o cálculo da **probabilidade**, sendo n eventos X_i mutuamente exclusivos ($X_j \cap X_k = \varnothing$, $\forall\, j \neq k$), com $i = 1, 2, ...n$ e considerando $N(X_i)$ o número de ocorrências do evento X_i em um espaço amostral S de tamanho N, a probabilidade da união de todos os eventos será:

$$P\left(\bigcup_{i=1}^{n} X_i\right) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{n} N(X_i) = \sum_{i=1}^{n} P(X_i) \tag{1}$$

A **probabilidade condicional** $P(B|A)$ denota a probabilidade de ocorrência do evento B dado que o evento A ocorreu. Considerando um experimento de N repetições, no qual o evento A ocorre n_1 vezes, e destas vezes que A ocorreu, B ocorreu n_2 vezes, portanto sendo n_2 o número de vezes que o evento $A \cap B$ ocorreu. Assim:

$$P(A \cap B) = \lim_{N\to\infty}\left(\frac{n_2}{N}\right) = \lim_{N\to\infty}\left(\frac{n_1}{N}\right)\left(\frac{n_2}{n_1}\right) = P(A)P(B|A) \tag{2}$$

Se o evento B for independente de A: $P(A \cap B) = P(A)P(B)$, logo $P(B|A) = P(B)$.

O **teorema da probabilidade total** facilita o cálculo da probabilidade de ocorrência de um evento Y diante de um grande número de resultados a considerar. Assim, para n eventos disjuntos X_i ($X_j \cap X_k = \varnothing$, $\forall\, j \neq k$), com $i = 1, 2, ...n$, e que formam uma partição do espaço amostral S tal que: $\bigcup_{i=1}^{n} X_i = S$, a probabilidade de ocorrência de Y será:

$$P(Y) = \sum_{i=1}^{n} P(X_i)P(Y|X_i) \tag{3}$$

O **teorema de Bayes** permite determinar a possibilidade de ocorrência de uma causa particular de um evento dentre muitas causas disjuntas possíveis. Para as mesmas considerações do teorema da probabilidade total para X_i e $P(Y) > 0$, com $j = 1, 2, ...n$, tem-se:

$$P(X_j|Y) = \frac{P(Y|X_j)P(X_j)}{\sum_{i=1}^{n} P(Y|X_i)P(X_i)} \tag{4}$$

A **correlação** é uma medida da relação entre duas variáveis aleatórias. Ela depende da covariância ($\sigma_{xy} = \overline{(x-\bar{x})(y-\bar{y})}$) e variância ($\sigma_x = \overline{(x-\bar{x})^2}$), assim o coeficiente de correlação $\rho_{xy} \in [-1, 1]$ será:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{5}$$

Variáveis aleatórias independentes são descorrelacionadas, mas não necessariamente o contrário é válido. "Existe apenas um caso especial para o qual independência e descorrelação são equivalentes quando as variáveis x e y são conjuntamente gaussianas. Reparemos que, quando x e y são conjuntamente gaussianas, x e y são individualmente gaussianas."

O **valor quadrático médio da soma de variáveis descorrelacionadas** x e y, com $\sigma_{xy} = 0$ será:

$$\sigma_z^2 = \sigma_z^2 + \sigma_y^2 + \sigma_{xy} = \sigma_z^2 + \sigma_y^2 \tag{6}$$

Haykin, S., Moher, M., "Sistemas de Comunicação", 5a Edição, Editora Bookman, 2011.

A média de um processo aleatório $X(t)$ é a esperança da variável aleatória obtida pela observação do processo ao longo de certo tempo t com fdp $f_{X(t)}(x)$.

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x)dx \tag{7}$$

Um processo estocástico é dito como **estacionário de primeira ordem** se sua função distribuição ou densidade de probabilidade não varia com o tempo, satisfazendo: $f_{X(t_1)}(x) = f_{X(t_2)}(x)$ para qualquer t. Isso também implica que seu valor esperado e variância mantém-se constantes.

A **função autocorrelação** de um processo é definido como o produto de duas variáveis aleatórias obtidas pela observação de um processo estocástico em dois instantes de tempo diferentes, com $f_{X(t_1),X(t_2)}(x_1, x_2)$ representando a fdp conjunta de $X(t_1)$ e $X(t_2)$:

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1),X(t_2)}(x_1, x_2)dx_1 dx_2 \tag{8}$$

Um processo estocástico é dito como **estacionário de segunda ordem** se sua fdp conjunta depende somente da diferença entre os tempos observados, implicando que a função de autocorrelação também possui essa dependência, ou seja para todo t: $R_X(t_1, t_2) = R_X(t_2 - t_1)$. Assim, a função auto covariância também dependente apenas da diferença entre os tempos será:

$$C_X(t_1, t_2) = E[(X(t_1) - \mu_X)(X(t_2) - \mu_X)] = R_X(t_2 - t_1) - \mu_X^2 \tag{9}$$

Propriedades da função de autocorrelação de um processo estacionário redefinida como $R_X(\tau) = E[X(t+\tau)X(t)]$:

- O valor médio quadrático do processo pode ser obtido da função autocorrelação para $\tau = 0$: $R_X(\tau) = E[X^2(t)]$.
- A função autocorrelação é uma função par de τ: $R_X(\tau) = R_X(-\tau) = E[X(t)X(t-\tau)]$.
- O valor máximo da função autocorrelação é dado em $\tau = 0$: $|R_X(\tau)| \leq R_X(0)$.

A **função autocorrelação** possui o sentido físico de descrever a "interdependência"de duas variáveis aleatórias observadas em um processo aleatório e espaçadas de τ segundos. Assim, quanto mais o processo estocástico varia com o tempo, mais rapidamente a função autocorrelação decresce com o aumento de τ. Considerando um caso mais geral de dois processos estocásticos $X(t)$ e $Y(t)$, com autocorrelações, respectivamente, $R_X(t, u)$ e $R_Y(t, u)$, a **função correlação cruzada** desses dois processos será:

$$R_{XY}(t, u) = E[X(t)Y(t)] \tag{10}$$

Essa função em geral não é par em τ e não possui máximo na origem. Porém obedece certa simetria de forma que (problema 5.12): $R_{XY}(\tau) = R_{YX}(-\tau)$.

Vale destacar se os processos X e Y forem estacionário no sentido amplo e conjunto, a correlação cruzada é re-escrita como: $R_{XY}(t, u) = R_{XY}(\tau)$.

Uma propriedade conhecida como **ergodicidade** permite que a média do conjunto seja estimada pela média temporal da função amostra.

No que diz respeito à **densidade espectral de energia** de um processo estocástico $X(t)$ como $S_X(f)$, esta pode ser encontrada a partir da transformada de Furier da função autocorrelação do processo, assim:

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f \tau} d\tau \tag{11}$$

Juntamente com a inversa de 11, estas equações são chamadas relações Einstein-Wiener-Khintchine. De forma que mostra que, se a autocorrelação ou densidade espectral de potência de um processo são conhecidos, o outro pode ser encontrado.

Propriedades da densidade espectral de potência (PSD) considerando um processo estocástico estacionário em sentido amplo:

- O valor DC da PSD do processo descrito é a área sob o gráfico da função autocorrelação: $S_X(0) = \int_{-\infty}^{\infty} R_X(\tau)d\tau$.
- O valor quadrático médio do processo descrito é a área sob o gráfico da PSD: $E[X^2(t)] = R_X(0) = \int_{-\infty}^{\infty} S_X(f)df$.
- A PSD é sempre positiva, pois $S_X(f) \approx E[|P(f)|^2]$, com $P(f)$ sendo a transformada de Fourier do processo $X(t)$.
- A PSD de um processo de valores reais é uma função par da frequência.

Ao passar um processo aleatório por um filtro estável, linear, invariante no tempo, a PSD de saída será modificada de: $S_Y(f) = H(f)H^*(f)S_X(f) = |H(f)|^2 S_X(f)$.

Albuquerque, J. P. A., Fortes, J. M. P., Finamore, W. A, "Probabilidade, Variáveis Aleatórias e Processos Estocásticos", Ed. Interciência e Ed. PUC-Rio, 2008.

Um processo estocástico é dito **estacionário de ordem** n quando sua função densidade de probabilidade de ordem n não varia com deslocamentos no tempo. Ele também é estacionário para qualquer $k < n$.

Um processo estocástico é dito **estacionário no sentido restrito** quando ele é estacionário de ordem n para qualquer $n \in \mathbb{Z}$.

Um processo estocástico $X(t)$ é dito **estacionário no sentido amplo** se sua média for constante e sua função de autocorrelação depender da diferença $t_2 - t_1$:

$$\mu_X(t) = \mu_X \; ; \forall t \tag{12}$$

$$R_X(t_1, t_2) = R_X(\tau), ; \tau = t_2 - t_1 \tag{13}$$

Processos ergódicos são aqueles cujas estatísticas de processos como a média e a função de autocorrelação podem ser determinadas por uma função-amostra do processo. Nesses processos os valores médios e momentos podem ser determinados por médias temporais.

A **função densidade espectral de potência** (PSD) para um sinal determinístico $x(t)$ será: $S_x(f) = \lim_{T\to\infty} |X_T(f)|^2/\text{T}$. Com a finalidade de generalizar para processos estocásticos, calcula-se a média estatística ao longo do ensemble do processo, resultando em:

$$S_X(f) = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t_1, t_2) e^{-j2\pi f(t_1-t_2) dt_1 dt_2} \tag{14}$$

No caso de um processo estocástico estacionário no sentido amplo ($R_X(t_1, t_2) = R_X(t_1 - t_2)$), assim:

$$S_X(f) = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t_1 - t_2) e^{-j2\pi f(t_1-t_2) dt_1 dt_2} = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau \tag{15}$$

A **potência média** em um intervalo de frequências $[f_1, f_2]$, considerando componentes frequenciais negativas, pode ser encontrada como:

$$P_{X_{[f_1,f_2]}} = \int_{-f_2}^{-f_1} S_X(f) df + \int_{f_1}^{f_2} S_X(f) df \tag{16}$$

Para a potência total basta integrar com $f_1 = 0$ e $f_2 = \infty$. Se o processo estocástico for estacionário no sentido amplo a potência pode ser escrita como: $P_X = R_X(0) = E[X(t)^2]$.

O ruído branco pode ser considerado como um processo estocástico estacionário no sentido amplo, PSD constante. Dessa forma, se $S_X(f) = C$, logo $R_X(\tau) = C\delta(\tau)$.

Problema temático 5.12 - HAYKIN

> Consider a pair of wide-sense stationary random processes $X(t)$ and $Y(t)$. Show that the cross-correlation $R_{XY}(\tau)$ and $R_{YX}(\tau)$ of these processes have the following properties.
>
> - $R_{XY}(\tau) = R_{YX}(-\tau)$;
> - $|R_{XY}(\tau)| \leq \frac{1}{2}[R_X(0) + R_Y(0)]$

Considerando a correlação cruzada para ambos os processos estacionários no sentido amplo:

$$R_{XY}(\tau) = E[X(t)Y(t-\tau)] \tag{17}$$

$$R_{YX}(\tau) = E[Y(t)X(t-\tau)] \tag{18}$$

Tomando $R_{XY}(\tau)$ avaliado em $\tau = -\tau$ e invertendo $t = t + \tau$ tem-se que seu valor será:

$$R_{XY}(-\tau) = E[X(t)Y(t+\tau)] = E[X(t+\tau)Y(t)] = E[Y(t)X(t+\tau)] = R_{XY}(\tau) = R_{YX}(-\tau) \tag{19}$$

No segundo caso, considerando $E[(X(t) \pm Y(t-\tau))^2] \geq 0$ e $R_X(0) = E[X^2(t)]$ e $R_Y(0) = E[Y^2(t)]$, desenvolvendo tem-se:

$$\begin{aligned} E[X^2(t)] \pm 2E[X(t)Y(t-\tau)] + E[Y^2(t-\tau)] \geq 0 \\ E[X(t)Y(t-\tau)] \leq \frac{1}{2}(E[X^2(t)] + E[Y^2(t-\tau)]) \\ |R_{XY}(\tau)| \leq \frac{1}{2}(R_X(0) + R_Y(0)) \end{aligned}$$