

CAPÍTULO IV - PROCESSOS ALEATÓRIOS

- Definição, caracterização e classificação.
- Média e autocorrelação.
- Exemplos de processos aleatórios.
- **Processos de Markov e de incrementos independentes.**
- **Processos Gaussianos.**
- **Estacionariedade e ergodicidade**
- **Processos complexos.**
- **Propriedades da correlação para processos estacionários (tempo contínuo e discreto);**

Processos de Markov

- Nós vimos, até agora, alguns exemplos de processos aleatórios, dos quais o processo de Wiener e o de Poisson são fundamentais, pois muitos outros processos podem ser obtidos a partir destes.
- Para esses dois processos, para se obter um conjunto de distribuições de n -ésima ordem foram utilizadas as densidades de primeira ordem e a propriedade de independência entre incrementos.
- Uma outra forma de se obter a densidade de ordem n a partir das densidades de primeira e segunda ordem (ou condicionais), para uma certa classe de processos, é mostrada a seguir:

Processos de Markov

- Seja a densidade de ordem n de um processo $\mathbf{x}(t)$, onde os valores x_i correspondem à variável aleatória $\mathbf{x}(t_i)$, que pode ser escrita como:

$$\begin{aligned} f_{\mathbf{x}}(x_n, x_{n-1}, \dots, x_3, x_2, x_1) &= f(x_n | x_{n-1}, \dots, x_3, x_2, x_1) f(x_{n-1}, \dots, x_3, x_2, x_1) = \\ &= f(x_n | x_{n-1}, \dots, x_3, x_2, x_1) f(x_{n-1} | x_{n-2}, \dots, x_3, x_2, x_1) f(x_{n-2}, \dots, x_3, x_2, x_1) = \dots = \\ &= f(x_n | x_{n-1}, \dots, x_3, x_2, x_1) f(x_{n-1} | x_{n-2}, \dots, x_3, x_2, x_1) \dots f(x_3 | x_2, x_1) f(x_2 | x_1) f(x_1) \end{aligned}$$

com $t_n > t_{n-1} > \dots > t_2 > t_1$.

- Se, para esse processo,

$$f(x_n | x_{n-1}, \dots, x_2, x_1) = f(x_n | x_{n-1})$$

para todos os valores $x_1, \dots, x_n$; todos os instantes $t_1 < \dots < t_n$ e qualquer inteiro $n > 0$, ou seja, a densidade condicional depende apenas do valor mais recente conhecido, esse processo é classificado como processo de Markov.

Processos de Markov

- Portanto:

$$f_{\mathbf{x}}(x_n, x_{n-1}, \dots, x_3, x_2, x_1) = f(x_n | x_{n-1}) f(x_{n-1} | x_{n-2}) \dots f(x_3 | x_2) f(x_2 | x_1) f(x_1)$$

- Assim, para processos dessa classe, a densidade de ordem n pode ser determinada conhecendo-se apenas as densidades de primeira ordem e as densidades condicionais entre dois instantes sucessivos quaisquer (ou as densidades de segunda ordem), já que

$$f(x_i | x_j) = f(x_i, x_j) / f(x_j)$$

No caso de um processo discreto, ele será dito de Markov se

$$P(\mathbf{x}_n = x_n | \mathbf{x}_{n-1} = x_{n-1}, \dots, \mathbf{x}_2 = x_2, \mathbf{x}_1 = x_1) = P(\mathbf{x}_n = x_n | \mathbf{x}_{n-1} = x_{n-1})$$

para todos os valores $x_1, \dots, x_n$; todos os instantes $t_1 < \dots < t_n$ e qualquer inteiro $n > 0$

Processos de Markov

- O valor do processo num instante de tempo t determina, então, as probabilidades condicionais para os valores futuros do processo.
- Os valores do processo são, então, chamados de estado do processo, e as probabilidades condicionais seriam as probabilidades de transição de estado.
- Se apenas um conjunto finito ou contável de valores x_i são possíveis, o processo discreto de Markov é denominado cadeia de Markov.
- Um exemplo de cadeia de Markov é o processo de Poisson visto anteriormente, e o processo de Wiener é um exemplo de processo de Markov contínuo.

Propriedades

- Se $\mathbf{x}(t)$ é Markov, então ,

$$f(x_n | x_{n+1}, \dots, x_{n+k-1}, x_{n+k}) = f(x_n | x_{n+1})$$

ou seja, a densidade condicional depende apenas do valor futuro mais recente conhecido;

- Demonstração:

$$\begin{aligned} f(x_n | x_{n+1}, \dots, x_{n+k-1}, x_{n+k}) &= \frac{f(x_n, x_{n+1}, \dots, x_{n+k-1}, x_{n+k})}{f(x_{n+1}, \dots, x_{n+k-1}, x_{n+k})} = \\ &= \frac{f(x_{n+k} | x_{n+k-1}) f(x_{n+k-1} | x_{n+k-2}) \cdots f(x_{n+1} | x_n) f(x_n)}{f(x_{n+k} | x_{n+k-1}) f(x_{n+k-1} | x_{n+k-2}) \cdots f(x_{n+2} | x_{n+1}) f(x_{n+1})} = \\ &= \frac{f(x_{n+1} | x_n) f(x_n)}{f(x_{n+1})} = \frac{f(x_{n+1}, x_n)}{f(x_{n+1})} = f(x_n | x_{n+1}) \end{aligned}$$

Propriedades

- se $\mathbf{x}(t)$ é Markov, e o estado presente é conhecido, então, o passado é independente do futuro, ou seja

$$f(x_1, x_3 | x_2) = f(x_1 | x_2) f(x_3 | x_2) \quad t_1 < t_2 < t_3;$$

- Demonstração:

$$\begin{aligned} f(x_1, x_3 | x_2) &= \frac{f(x_1, x_2, x_3)}{f(x_2)} = \frac{f(x_3 | x_2) f(x_2 | x_1) f(x_1)}{f(x_2)} = \\ &= \frac{f(x_3 | x_2) f(x_2, x_1)}{f(x_2)} = f(x_3 | x_2) f(x_1 | x_2) \end{aligned}$$

- Se $\mathbf{x}(t)$ é Markov, $E\{\mathbf{x}(t_n) | \mathbf{x}(t_{n-1}), \dots, \mathbf{x}(t_1)\} = E\{\mathbf{x}(t_n) | \mathbf{x}(t_{n-1})\}$.

Propriedades

- Se, com $t_1 < t_2$, $\mathbf{x}(t_2) - \mathbf{x}(t_1)$ é independente de $\mathbf{x}(t)$ para todo $t \leq t_1$, então, $\mathbf{x}(t)$ é Markov;

- Demonstração:

$$\begin{aligned} \begin{cases} \Delta = \mathbf{x}_n - \mathbf{x}_{n-1} \\ \Gamma = \mathbf{x}_{n-1} \end{cases} &\Rightarrow \begin{cases} \mathbf{x}_n = \Delta + \Gamma \\ \mathbf{x}_{n-1} = \Gamma \end{cases} \Rightarrow \\ \Rightarrow f_{\mathbf{x}_n, \mathbf{x}_{n-1}}(x_n, x_{n-1}) &= f_{\Delta, \Gamma}(x_n - x_{n-1}, x_{n-1}) = f_{\Delta}(x_n - x_{n-1}) f_{\mathbf{x}_{n-1}}(x_{n-1}) \end{aligned}$$

$$\begin{aligned} f(x_n | x_{n-1}, \dots, x_2, x_1) &= \frac{f(x_n, x_{n-1} | x_{n-2}, \dots, x_2, x_1)}{f(x_{n-1} | x_{n-2}, \dots, x_2, x_1)} = \\ &= \frac{f_{\Delta}(x_n - x_{n-1} | x_{n-2}, \dots, x_2, x_1) f(x_{n-1} | x_{n-2}, \dots, x_2, x_1)}{f(x_{n-1} | x_{n-2}, \dots, x_2, x_1)} = \\ &= f_{\Delta}(x_n - x_{n-1}) = \frac{f_{\Delta}(x_n - x_{n-1}) f(x_{n-1})}{f(x_{n-1})} = \frac{f(x_n, x_{n-1})}{f(x_{n-1})} = \\ &= f(x_n | x_{n-1}) \end{aligned}$$

Processos de incrementos independentes

- Um processo para o qual as variáveis aleatórias y_i dadas por

$$y_1 = x(t_1) - x(t_0); y_2 = x(t_2) - x(t_1); \dots; y_n = x(t_n) - x(t_{n-1}); t_0 < t_1 < t_2 < \dots < t_n,$$

são independentes é denominado um processo de incrementos independentes. É fácil de ver, portanto, que sendo $\Delta x_i = x_i - x_{i-1}$,

$$\begin{aligned} F_x(x_n | x_{n-1}, \dots, x_1) &= P\{x(t_n) \leq x_n | x(t_{n-1}) = x_{n-1}, \dots, x(t_1) = x_1\} = \\ &= P\{y_n \leq \Delta x_n | y_{n-1} = \Delta x_{n-1}, \dots, y_1 = \Delta x_1\} = P\{y_n \leq \Delta x_n\} = \\ &= P\{x(t_n) \leq x_n | x(t_{n-1}) = x_{n-1}\} = F_x(x_n | x_{n-1}) \Rightarrow \\ &\Rightarrow f_x(x_n | x_{n-1}, \dots, x_1) = f_x(x_n | x_{n-1}) \end{aligned}$$

- ou seja, todo processo de incrementos independentes também é Markov. O argumento reverso, porém, não é verdadeiro: Nem todo processo de Markov é de incrementos independentes.

Processos de Markov de tempo discreto

- Um processo contínuo de tempo discreto para o qual

$$f(x_{n+k} | x_{n-1}, \dots, x_1, x_0) = f(x_{n+k} | x_{n-1})$$

para todos os valores $x_0, \dots, x_{n-1}, x_{n+k}$ e qualquer inteiro $n > 0$ e qualquer $k > 0$, é denominado processo de Markov de tempo discreto, ou sequência aleatória de Markov.

- No caso de uma sequência discreta devemos, analogamente, ter

$$P(x_{n+k} | x_{n-1}, \dots, x_1, x_0) = P(x_{n+k} | x_{n-1})$$

- A sequência discreta de Markov, de tempo discreto, é também denominada cadeia de Markov.

Processos de Markov de tempo discreto

- É fácil de mostrar que essas expressões serão verdadeiras para qualquer k se o forem para $k=0$, ou seja:

$$f(x_n | x_{n-1}, \dots, x_1, x_0) = f(x_n | x_{n-1}) \Rightarrow f(x_{n+k} | x_{n-1}, \dots, x_1, x_0) = f(x_{n+k} | x_{n-1})$$

Demonstração:

$$\begin{aligned} f(x_{n+k} | x_{n-1}, \dots, x_1, x_0) &= \frac{f(x_{n+k}, x_{n-1}, \dots, x_1, x_0)}{f(x_{n-1}, \dots, x_1, x_0)} = \\ &= \frac{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_{n+k}, \mathbf{x}_{n+k-1}, \dots, \mathbf{x}_n, x_{n-1}, \dots, x_1, x_0) d\mathbf{x}_{n+k-1} \dots d\mathbf{x}_n}{f(x_{n-1}, \dots, x_1, x_0)} = \\ &= \frac{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_{n+k} | x_{n+k-1}) f(x_{n+k-1} | x_{n+k-2}) \dots f(x_n | x_{n-1}) f(\mathbf{x}_{n-1}, \dots, \mathbf{x}_1, x_0) d\mathbf{x}_{n+k-1} \dots d\mathbf{x}_n}{f(\mathbf{x}_{n-1}, \dots, \mathbf{x}_1, x_0)} = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_{n+k}, \mathbf{x}_{n+k-1}, \dots, \mathbf{x}_n | x_{n-1}) d\mathbf{x}_{n+k-1} \dots d\mathbf{x}_n = f(x_{n+k} | x_{n-1}) \end{aligned}$$

Exemplo: processo de Markov gaussiano

- Seja $\mathbf{x}(n)$ uma sequência randômica definida para $n \geq 0$ com

$$f_{\mathbf{x}}(x; 0) = N(0, \sigma_0)$$

e função densidade de probabilidade de transição dada por

$$f_{\mathbf{x}}(x_n | x_{n-1}; n, n-1) = N(\rho x_{n-1}, \sigma_w)$$

com $|\rho| < 1$.

- Queremos determinar a densidade da variável aleatória $\mathbf{x}(n)$ para um $n > 0$ arbitrário.
- Em geral, poderíamos avançar recursivamente, partindo da densidade inicial, e calculando as integrais

$$f_{\mathbf{x}}(x_n; n) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x_n | x_{n-1}; n, n-1) f_{\mathbf{x}}(x_{n-1}; n-1) dx_{n-1}$$

- para $n = 1, 2, \dots$

Exemplo: processo de Markov gaussiano

- Note porém que:

$$\begin{aligned}\int_{-\infty}^{\infty} f_x(x_n | x_{n-1}; n, n-1) f_x(x_{n-1}; n-1) dx_{n-1} &= \int_{-\infty}^{\infty} g\left(\frac{x_n - \rho x_{n-1}}{\sigma_w}\right) f_x(x_{n-1}; n-1) dx_{n-1} = \\ &= \int_{-\infty}^{\infty} g\left(\frac{\rho}{\sigma_w}\left(\frac{x_n}{\rho} - x_{n-1}\right)\right) f_x(x_{n-1}; n-1) dx_{n-1} = g\left(\frac{x_n}{\sigma_w}\right) * f_x\left(\frac{x_n}{\rho}; n-1\right)\end{aligned}$$

- E, começando-se com uma densidade gaussiana ($n=0$) todas as densidades seguintes serão gaussianas, pois a convolução de funções gaussianas é também gaussiana.

Exemplo: processo de Markov gaussiano

- Assim, precisamos determinar, apenas, os dois primeiros momentos centrais da variável aleatória $\mathbf{x}(n)$.
- A média será dada por:

$$\begin{aligned} m_{\mathbf{x}}(n) &= E\{\mathbf{x}(n)\} = E\{E\{\mathbf{x}(n)|\mathbf{x}(n-1)\}\} = E\{\rho \mathbf{x}(n-1)\} = \rho m_{\mathbf{x}}(n-1) = \\ &= \rho^2 m_{\mathbf{x}}(n-2) = \dots = \rho^n m_{\mathbf{x}}(0) = 0 \end{aligned}$$

- e a variância por

$$\begin{aligned} \sigma_{\mathbf{x}}^2(n) &= E\{\mathbf{x}^2(n)\} = E\{E\{\mathbf{x}^2(n)|\mathbf{x}(n-1)\}\} = E\{\sigma_{\mathbf{w}}^2 + \rho^2 \mathbf{x}^2(n-1)\} = \\ &= \sigma_{\mathbf{w}}^2 + \rho^2 E\{\mathbf{x}^2(n-1)\} = \sigma_{\mathbf{w}}^2 + \rho^2 \sigma_{\mathbf{w}}^2 + \rho^4 E\{\mathbf{x}^2(n-2)\} = \dots = \\ &= \sigma_{\mathbf{w}}^2 \left(\sum_{i=0}^{n-1} \rho^{2i} \right) + \rho^{2n} E\{\mathbf{x}^2(0)\} = \sigma_{\mathbf{w}}^2 \left(\sum_{i=0}^{n-1} \rho^{2i} \right) + \rho^{2n} \sigma_0^2 \underset{n \rightarrow \infty}{=} \frac{1}{1-\rho^2} \sigma_{\mathbf{w}}^2 \end{aligned}$$

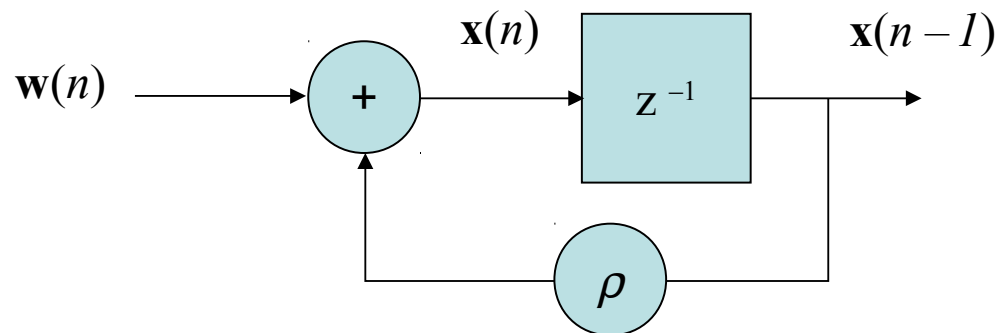
Exemplo: sistema discreto de 1a. ordem

- Considere a equação de diferenças
 - $\mathbf{x}(n) = \rho \mathbf{x}(n-1) + \mathbf{w}(n); \mathbf{x}(0) = \mathbf{w}(0)$
- onde $\mathbf{w}(n)$ é uma sequência randômica independente e identicamente distribuída (i.i.d).
- Claramente $\mathbf{x}(n)$ é uma sequência de Markov.
- Para $n > 0$ a função densidade de probabilidade conjunta das variáveis aleatórias $\mathbf{x}(i)$, $i = 0, 1, \dots, n$ será dada por

$$\begin{aligned} f_{\mathbf{x}}(x_n, x_{n-1}, \dots, x_0) &= f_{\mathbf{x}}(x_n | x_{n-1}, \dots, x_0) \cdot f_{\mathbf{x}}(x_{n-1} | x_{n-2}, \dots, x_0) \dots f_{\mathbf{x}}(x_1 | x_0) f_{\mathbf{x}}(x_0) = \\ &= f_{\mathbf{x}}(x_n | x_{n-1}) f_{\mathbf{x}}(x_{n-1} | x_{n-2}) \dots f_{\mathbf{x}}(x_1 | x_0) f_{\mathbf{w}}(x_0) = f_{\mathbf{w}}(x_0) \prod_{k=1}^n f_{\mathbf{w}}(x_k - \rho x_{k-1}) \end{aligned}$$

- onde x_k é o valor tomado pela variável aleatória $\mathbf{x}(k)$

Exemplo: sistema discreto de 1a. ordem



Seqüência randômica de Markov de ordem p

- Podemos generalizar a propriedade de Markov para cobrir equações de diferença de ordem superior estendendo o conceito de dependência direta para amostras distando mais do que apenas uma unidade, como se segue.
- Seja o inteiro positivo p denominado ordem da sequência Markov- p . Uma sequência aleatória contínua será denominada Markov- p se

$$f_{\mathbf{x}}(x_{n+k} | x_{n-1}, x_{n-2}, \dots, x_0) = f_{\mathbf{x}}(x_{n+k} | x_{n-1}, x_{n-2}, \dots, x_{n-p})$$

- para todo $k \geq 0$ e todo $n \geq p$.

Seqüência randômica de Markov de ordem p

- Vemos, então, que a medida que a ordem p do modelo de Markov aumenta, o erro cometido ao se aproximar uma seqüência genérica (não Markov) por uma seqüência de Markov diminui, desde que a dependência mais forte seja sempre com as amostras mais próximas ($\mathbf{x}(n-1)$ e $\mathbf{x}(n-2)$ por exemplo) e a dependência seja desprezível com as amostras mais distantes, o que parece razoável para a maior parte dos casos práticos.
- O modelo Markov- p é muito útil em técnicas modernas de estimação espectral e em estimação linear recursiva (filtro de Kalman).

Equação de Chapman-Kolmogorov

- Dado um processo de Markov, ele estará totalmente caracterizado se forem conhecidas sua densidade de primeira ordem, $f_{\mathbf{x}}(x;t)$ e sua densidade condicional $f_{\mathbf{x}}(x_2|x_1;t_2|t_1)$.
- Porém essa densidade condicional não pode ser qualquer, ela deve ser consistente de forma que sempre possamos obter a densidade de primeira ordem para um instante $t_2 > t_1$ a partir dela, ou seja

$$f_{\mathbf{x}}(x_2;t_2) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x_2|x_1;t_2|t_1) f_{\mathbf{x}}(x_1;t_1) dx_1$$

Equação de Chapman-Kolmogorov

- Que condição geral deve então ela satisfazer?
- A resposta a essa pergunta é dada pela equação de Chapman-Kolmogorov, demonstrada a seguir.
- Considere três instantes de tempo, $t_3 > t_2 > t_1$, e as variáveis aleatórias do processo de Markov nesses instantes, $\mathbf{x}(t_1)$, $\mathbf{x}(t_2)$ e $\mathbf{x}(t_3)$. Primeiramente, vamos escrever a seguinte densidade conjunta,

$$f_{\mathbf{x}}(x_3, x_1) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x_3 | x_2, x_1) f_{\mathbf{x}}(x_2, x_1) dx_2$$

- onde suprimimos os instantes t_i por simplicidade de notação.
- Dividindo-se ambos os lados dessa equação por $f_{\mathbf{x}}(x_1)$, obtemos

$$f_{\mathbf{x}}(x_3 | x_1) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x_3 | x_2, x_1) f_{\mathbf{x}}(x_2 | x_1) dx_2$$

Equação de Chapman-Kolmogorov

- Utilizando-se, agora, a propriedade de Markov, obtemos

$$f_x(x_3|x_1) = \int_{-\infty}^{\infty} f_x(x_3|x_2) f_x(x_2|x_1) dx_2$$

- que é conhecida como equação de Chapman-Kolmogorov para a densidade de transição de um processo de Markov

Processos gaussianos

- Um processo estocástico é denominado gaussiano quando as variáveis aleatórias $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_n)$ são conjuntamente gaussianas para qualquer n , e quaisquer $t_1, t_2, \dots, t_n$.
- Nesse caso o processo estará completamente caracterizado se eu conhecer apenas $E\{\mathbf{x}(t)\}$ e $C_{\mathbf{x}}(t_1, t_2)$:

$$C_{\mathbf{x}} = E\left\{\left(\mathbf{x} - \underline{m}_{\mathbf{x}}\right)\left(\mathbf{x} - \underline{m}_{\mathbf{x}}\right)^T\right\} = E\left\{\begin{bmatrix} \mathbf{x}(t_1) - m_{\mathbf{x}}(t_1) \\ \vdots \\ \mathbf{x}(t_n) - m_{\mathbf{x}}(t_n) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t_1) - m_{\mathbf{x}}(t_1) & \cdots & \mathbf{x}(t_n) - m_{\mathbf{x}}(t_n) \end{bmatrix}\right\} =$$
$$= \begin{bmatrix} C_{\mathbf{x}}(t_1, t_1) & \cdots & C_{\mathbf{x}}(t_1, t_n) \\ \vdots & \ddots & \vdots \\ C_{\mathbf{x}}(t_n, t_1) & \cdots & C_{\mathbf{x}}(t_n, t_n) \end{bmatrix}$$

- O processo de Wiener, visto anteriormente, é um exemplo de processo gaussiano.

Processos estacionários

- Dizemos que um processo estocástico, $\mathbf{x}(t)$, é estacionário, no sentido restrito, se suas estatísticas não são afetadas por um deslocamento da origem dos tempos, isto é, se os processos $\mathbf{x}(t)$ e $\mathbf{x}(t+\varepsilon)$ tiverem as mesmas estatísticas para qualquer ε .
- Dizemos que os processos $\mathbf{x}(t)$ e $\mathbf{y}(t)$ são conjuntamente estacionários se as estatísticas conjuntas de $\mathbf{x}(t)$ e $\mathbf{y}(t)$ forem as mesmas de $\mathbf{x}(t+\varepsilon)$ e $\mathbf{y}(t+\varepsilon)$ para qualquer ε .
- Dizemos que o processo complexo $\mathbf{z}(t) = \mathbf{x}(t) + j \mathbf{y}(t)$ é estacionário se $\mathbf{x}(t)$ e $\mathbf{y}(t)$ são conjuntamente estacionários.
- Exemplo:

$$\mathbf{x}(t) = A \cos(2 \pi f_0 t + \boldsymbol{\theta})$$

$\boldsymbol{\theta}$ = variável aleatória com densidade de probabilidade uniforme em $[0, 2 \pi]$.

Corolários

- Da definição segue-se que a função densidade de ordem n de um processo estacionário deve ser tal que

$$f_{x_1 \dots x_n}(x_1, \dots, x_n; t_1, \dots, t_n) = f_{x_1 \dots x_n}(x_1, \dots, x_n; t_1 + \varepsilon, \dots, t_n + \varepsilon)$$

para qualquer ε .

- Em particular, deve-se ter

$$f_x(x; t) = f_x(x; t + \varepsilon),$$

ou seja, a densidade de primeira ordem deve ser independente de t ,

$$f_x(x; t) = f_x(x),$$

e portanto,

$$E\{\mathbf{x}(t)\} = m_x = \text{constante}.$$

Corolários

- Também a densidade de segunda ordem deve ser tal que

$$f_{\mathbf{x}_1\mathbf{x}_2}(x_1, x_2; t_1, t_2) = f_{\mathbf{x}_1\mathbf{x}_2}(x_1, x_2; t_1 + \varepsilon, t_2 + \varepsilon) = f_{\mathbf{x}_1\mathbf{x}_2}(x_1, x_2; \tau)$$

com $\tau = t_2 - t_1$.

- Temos, então, que

$$R_{\mathbf{x}}(t_1, t_2) = R_{\mathbf{x}}(\tau) = E\{\mathbf{x}(t) \mathbf{x}(t + \tau)\} = E\{\mathbf{x}(t - \tau) \mathbf{x}(t)\} = R_{\mathbf{x}}(-\tau)$$

para processos reais.

Estacionariedade de ordem finita

- Dizemos que um processo é estacionário de ordem k se

$$f_{\mathbf{x}}(x_1, \dots, x_n; t_1, \dots, t_n) = f_{\mathbf{x}}(x_1, \dots, x_n; t_1 + \varepsilon, \dots, t_n + \varepsilon)$$

- para qualquer $n \leq k$.

Estacionariedade no sentido amplo

- Dizemos que um processo é estacionário no sentido amplo se seu valor esperado é uma constante e sua autocorrelação depende apenas de $t_1 - t_2$:

$$E\{\mathbf{x}(t)\} = m_{\mathbf{x}} \quad \text{e} \quad E\{\mathbf{x}(t+\tau) \mathbf{x}(t)\} = R_{\mathbf{x}}(\tau)$$

Exemplos

- O processo $x(t) = A \cos(2 \pi f_0 t + \theta)$, como visto, tem

$$m_x = 0 \text{ e } R_x(t, t + \tau) = A^2/2 \cos(2 \pi f_0 \tau).$$

- Portanto ele é estacionário no sentido amplo (obviamente pois é também estacionário no sentido restrito).
- Os processo de Poisson e de Wiener não são estacionários nem mesmo no sentido amplo pois suas funções de correlação são dadas por:

$$C_w(t_1, t_2) = \alpha \min(t_1, t_2)$$

Exemplos

- O processo $x(t) = A \cos(2\pi v t)$ com

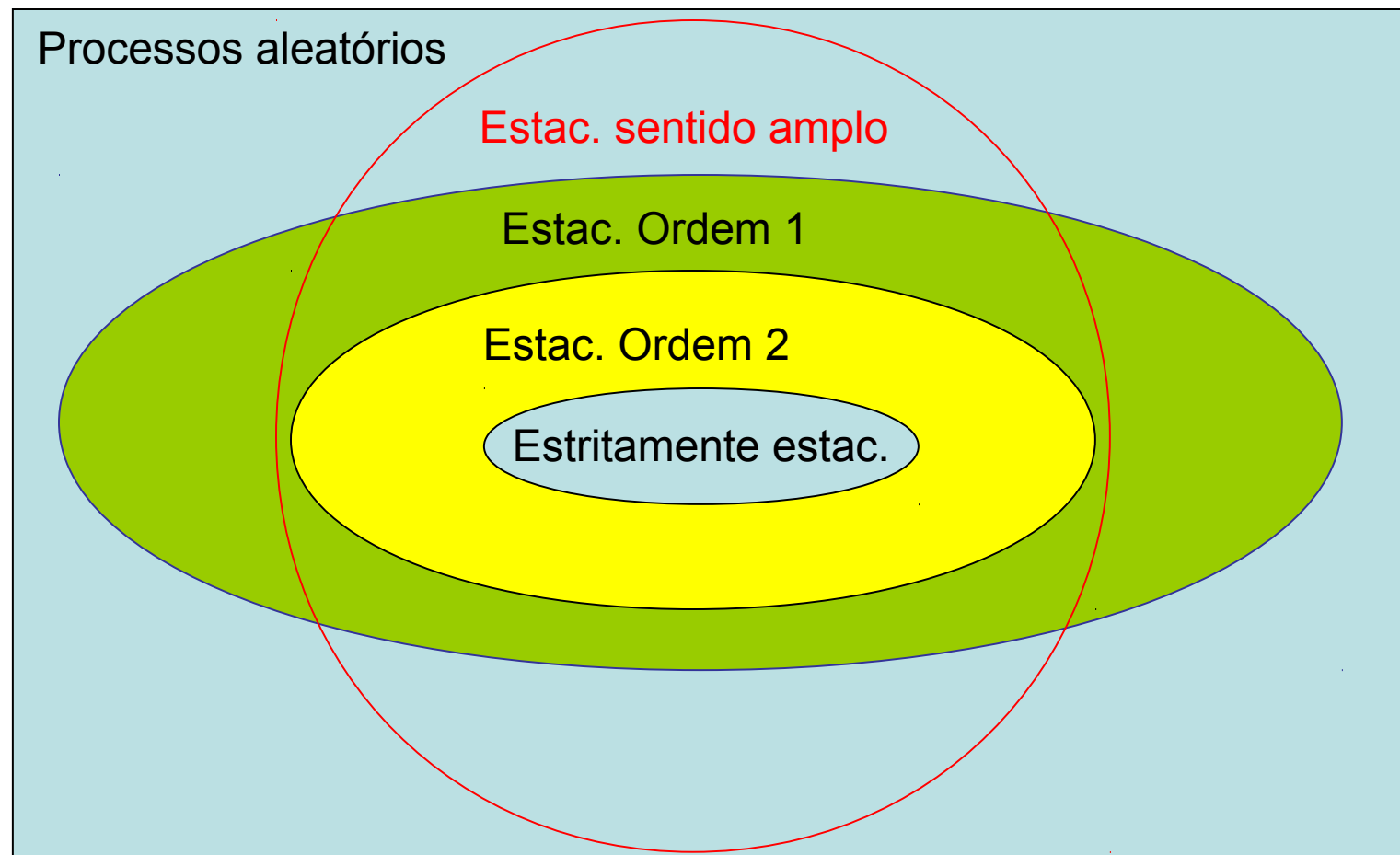
$$f_v(v) = \frac{1}{f_2 - f_1} \quad f_1 \leq v \leq f_2$$

tem média dada por

$$\begin{aligned} m_x(t) &= \int_{f_1}^{f_2} A \cos(2\pi v t) \frac{1}{f_2 - f_1} dv = \frac{A}{f_2 - f_1} \frac{\sin(2\pi f_2 t) - \sin(2\pi f_1 t)}{2\pi t} = \\ &= A \frac{\sin(\pi(f_2 - f_1)t)}{\pi(f_2 - f_1)t} \cos(\pi(f_2 + f_1)t) \end{aligned}$$

- Portanto, ele não é estacionário, nem mesmo no sentido amplo.

Estacionariedade



Processos assintoticamente estacionários

- Processos assintoticamente estacionários são aqueles que se tornam estacionários no limite quando t tende para infinito.
- Exemplo:
 - O processo do exemplo anterior tem média

$$m_x(t) = A \operatorname{sinc}[(f_2 - f_1)t] \cos[\pi(f_2 + f_1)t] \underset{t \rightarrow \infty}{=} 0$$

e autocorrelação

$$R_x(t, t + \tau) = \frac{A^2}{2(f_2 - f_1)} \cdot$$

$$\begin{aligned} & \cdot \{f_2 \operatorname{sinc}[2f_2(2t + \tau)] - f_1 \operatorname{sinc}[2f_1(2t + \tau)] + f_2 \operatorname{sinc}(2f_2\tau) - f_1 \operatorname{sinc}(2f_1\tau)\} = \\ & \underset{t \rightarrow \infty}{=} \frac{A^2}{2(f_2 - f_1)} [f_2 \operatorname{sinc}(2f_2\tau) - f_1 \operatorname{sinc}(2f_1\tau)] \end{aligned}$$

portanto ele é assintoticamente estacionário no sentido amplo.

Processos cicloestacionários

- Dado um processo não estacionário $\mathbf{x}(t)$, para o qual existe um período T tal que $\mathbf{x}(t)$ e $\mathbf{x}(t+nT)$ têm as mesmas estatísticas para qualquer n , esse processo é dito cicloestacionário e, portanto,

$$m_{\mathbf{x}}(t) = m_{\mathbf{x}}(t+T) \text{ e } R_{\mathbf{x}}(t_1, t_2) = R_{\mathbf{x}}(t_1+T, t_2+T).$$

- Se apenas a média e a autocorrelação obedecem à propriedade de cicloestacionariedade, o processo é dito cicloestacionário no sentido amplo.
- Exemplos de processos cicloestacionários incluem sinais modulados, sensores de varredura (por exemplo, sinal de televisão) e sinais de comunicação digital.
- Embora processos cicloestacionários não sejam estacionários, exceto em casos triviais, é muitas vezes apropriado converter um processo cicloestacionário em um processo estacionário como será mostrado no exemplo a seguir.

Exemplo: Modulação PSK

- PSK = phase-shift keying
- Na comunicação entre sistemas digitais, sequências binárias (dados) são geradas e então moduladas (movidas para uma região do espectro eletromagnético) para que possam ser transmitidas através de um canal de comunicação (p. ex. linha telefônica).
- Um método básico de modulação binária é o PSK.
- Nesse método, dígitos binários, modelados por uma sequência binária $\mathbf{b}(n)$, são mapeados bit a bit em uma sequência de ângulos $\theta(n)$ que é usada para modular a fase de um sinal senoidal $\cos(2 \pi f_c t)$.

Exemplo: Modulação PSK

- Especificamente, seja $\mathbf{b}(n)$ uma sequência de Bernoulli, que toma os valores 0 e 1 com igual probabilidade. E seja:

$$\theta(n) = \begin{cases} +\pi/2 & \text{se } \mathbf{b}(n)=1 \\ -\pi/2 & \text{se } \mathbf{b}(n)=0 \end{cases}$$

- Sendo

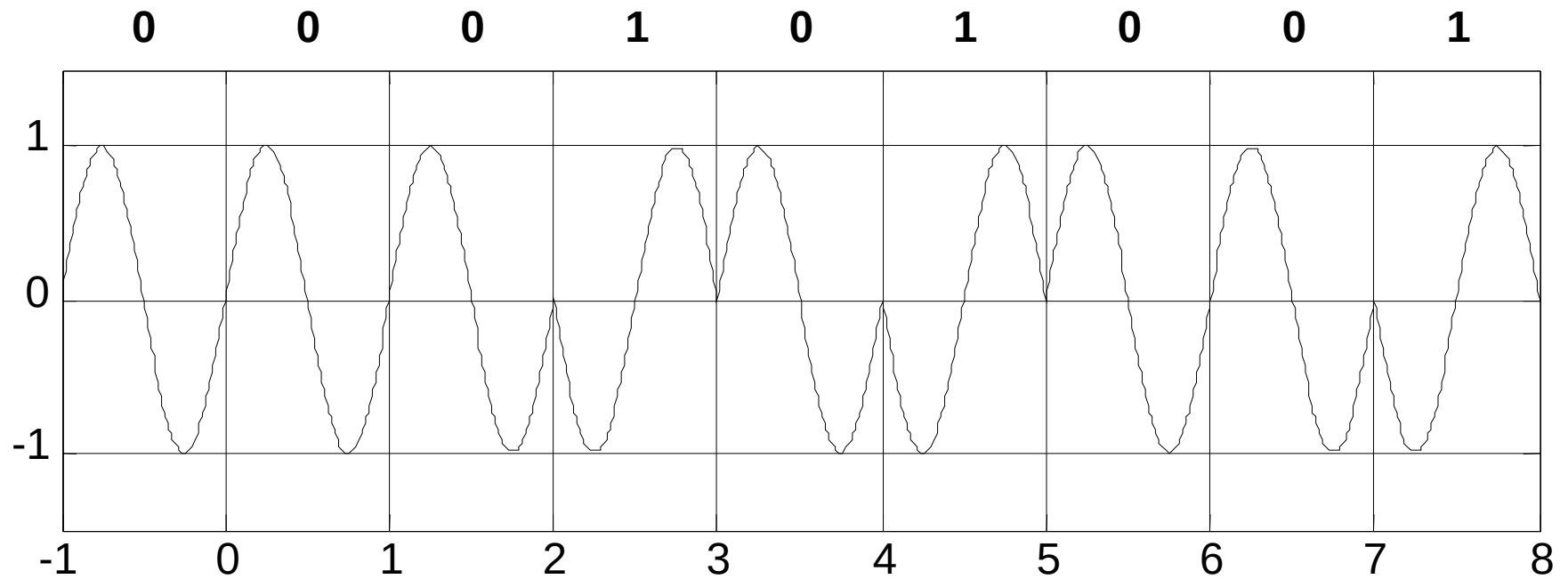
$$\theta_a(t) = \theta(n) \quad \text{para} \quad nT \leq t < (n+1)T,$$

construímos o sinal modulado como

$$\mathbf{x}(t) = \cos[2\pi f_c t + \theta_a(t)]$$

- onde T é a constante de tempo de transição de um bit.
- Normalmente T é escolhido para ser um múltiplo de $1/f_c$ de forma a haver um número inteiro de ciclos por bit.

Exemplo: Modulação PSK



Exemplo: Modulação PSK

- Vamos determinar a média e a autocorrelação desse processo. Vamos, inicialmente definir duas funções básicas,

$$s_I(t) = \begin{cases} \cos(2\pi f_c t) & 0 \leq t < T \\ 0 & \text{caso contrario} \end{cases}$$

- e

$$s_Q(t) = \begin{cases} \sin(2\pi f_c t) & 0 \leq t < T \\ 0 & \text{caso contrario} \end{cases}$$

Exemplo: Modulação PSK

- Assim,

$$\begin{aligned}\mathbf{x}(t) &= \cos[\boldsymbol{\theta}_a(t)] \cos(2\pi f_c t) - \sin[\boldsymbol{\theta}_a(t)] \sin(2\pi f_c t) = \\ &= \sum_{n=-\infty}^{\infty} \cos[\boldsymbol{\theta}(n)] s_I(t-nT) - \sum_{n=-\infty}^{\infty} \sin[\boldsymbol{\theta}(n)] s_Q(t-nT)\end{aligned}$$

- Portanto,

$$E\{\mathbf{x}(t)\} = \sum_{n=-\infty}^{\infty} E\{\cos[\boldsymbol{\theta}(n)]\} s_I(t-nT) - \sum_{n=-\infty}^{\infty} E\{\sin[\boldsymbol{\theta}(n)]\} s_Q(t-nT) = 0$$

- pois $\cos[\boldsymbol{\theta}(n)] = 0$ e $\sin[\boldsymbol{\theta}(n)] = \pm 1$ com igual probabilidade neste caso.

Exemplo: Modulação PSK

- Para a autocorrelação teremos,

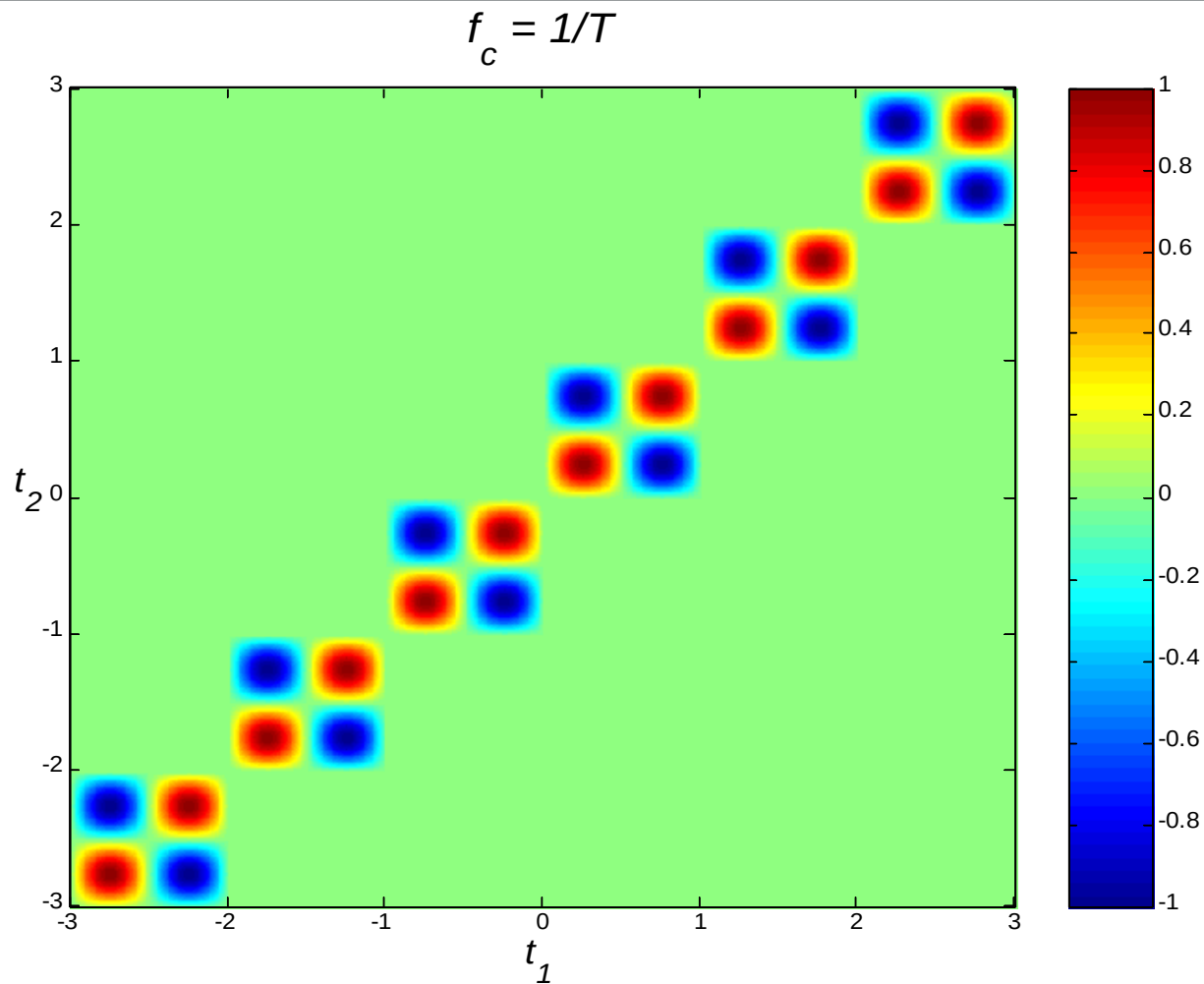
$$\begin{aligned} R_x(t_1, t_2) &= \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} E\{\sin[\theta(n)] \sin[\theta(k)]\} s_Q(t_1 - nT) s_Q(t_2 - kT) = \\ &= \sum_{n=-\infty}^{\infty} s_Q(t_1 - nT) s_Q(t_2 - nT) = \\ &= \begin{cases} s_Q((t_1 \bmod T)) s_Q((t_2 \bmod T)) & \text{se } \text{int}\left(\frac{t_1}{T}\right) = \text{int}\left(\frac{t_2}{T}\right) \\ 0 & \text{caso contrario} \end{cases} \end{aligned}$$

- Pois

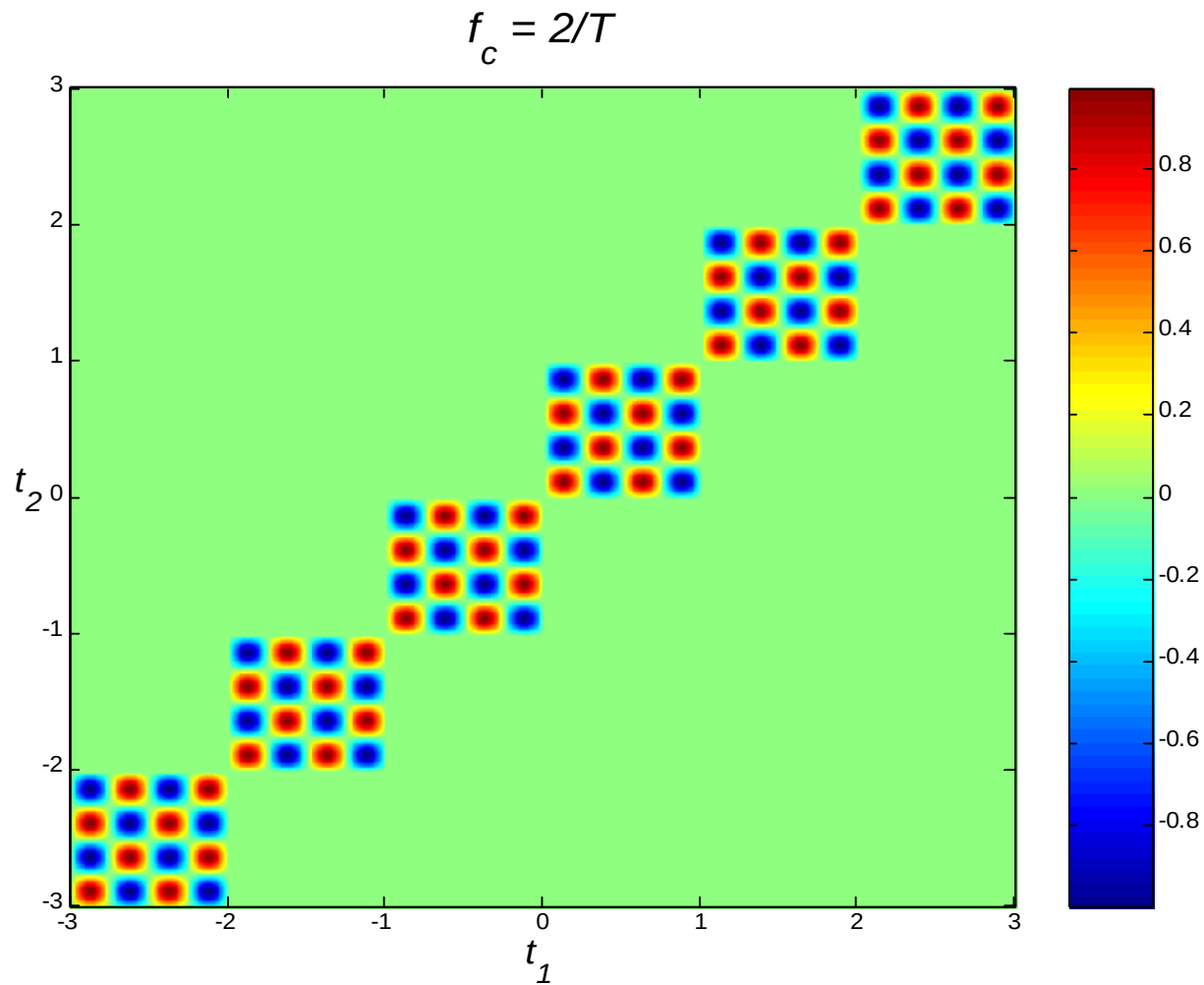
$$E\{\sin[\theta(n)] \sin[\theta(k)]\} = \begin{cases} E\{\sin^2[\theta(n)]\} = E\{1\} = 1 & n=k \\ E\{\sin[\theta(n)]\} E\{\sin[\theta(k)]\} = 0 & n \neq k \end{cases}$$

- Na expressão acima “int(x)” denota a parte inteira do argumento x, e “t mod T” denota $t - T \text{int}(t/T)$.

Exemplo: Modulação PSK



Exemplo: Modulação PSK



Exemplo: Modulação PSK

- Note a diferença entre a função de autocorrelação de um processo cicloestacionário e a de um processo periódico.
- No processo periódico $R_x(t_1, t_2)$ é periódica em duas direções: t_1 e t_2 , enquanto que num processo cicloestacionário ela é periódica apenas ao longo da direção que forma 45° com o eixo t_1 .
- É fácil de ver que a cicloestacionariedade, neste exemplo, provém do fato de o processo $\theta_a(t)$ ser constante por um período e mudar de estado apenas nos instantes $t=nT$, para n inteiro.
- Em muitas situações reais o processo de modulação começa num instante arbitrário t , que pode, portanto, ser modelado como aleatório.
- Assim, o processo anterior pode ser convertido para

$$\mathbf{x}'(t) = \cos \left[2\pi f_c t + \theta_a(t) + 2\pi f_c \alpha \right]$$

- pela adição de uma variável aleatória α , com distribuição uniforme no intervalo $(0, T)$, e independente do processo $\theta_a(t)$.

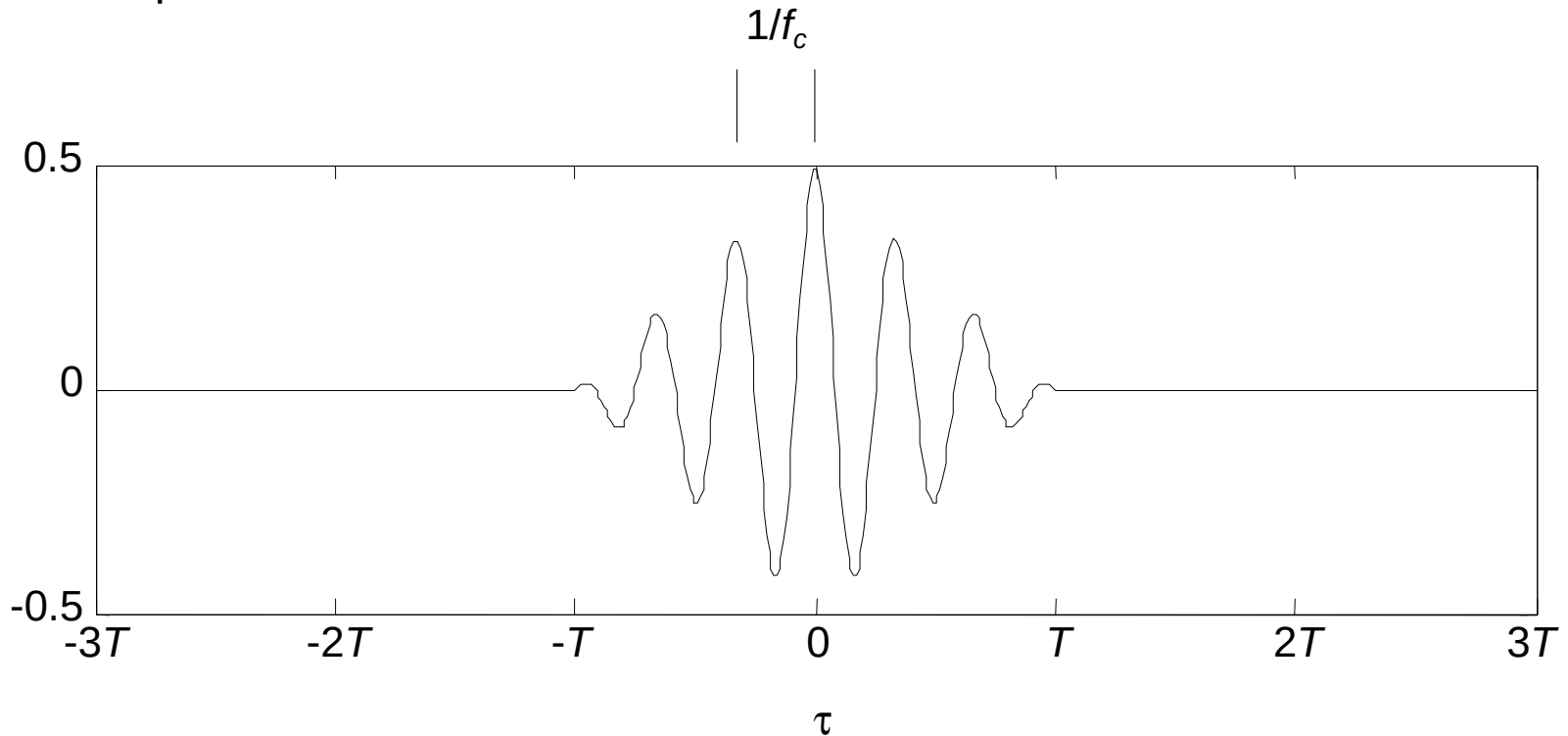
Exemplo: Modulação PSK

- Assim, no cálculo da nova função de autocorrelação, basta substituir t_1 por $t_1 + \alpha$ e t_2 por $t_2 + \alpha$ e integrar em α de 0 a T , dividindo-se por T :

$$\begin{aligned}
 R_x(t_1, t_1 + \tau) &= \frac{1}{T} \int_0^T R_x(t_1 + \alpha, t_1 + \alpha + \tau) d\alpha = \frac{1}{T} \int_{t_1}^{t_1+T} R_x(\varepsilon, \varepsilon + \tau) d\varepsilon = \\
 &= \frac{1}{T} \left\{ \int_{t_1}^{nT} R_x(\varepsilon, \varepsilon + \tau) d\varepsilon + \int_{nT}^{t_1+T} R_x(\varepsilon, \varepsilon + \tau) d\varepsilon \right\} = \\
 &= \frac{1}{T} \left\{ \int_{t_1}^{nT} R_x(\varepsilon, \varepsilon + \tau) d\varepsilon + \int_{(n-1)T}^{t_1} R_x(\beta + T, \beta + T + \tau) d\beta \right\} = \\
 &= \frac{1}{T} \left\{ \int_{t_1}^{nT} R_x(\beta, \beta + \tau) d\beta + \int_{(n-1)T}^{t_1} R_x(\beta, \beta + \tau) d\beta \right\} = \frac{1}{T} \left\{ \int_{(n-1)T}^{t_1} R_x(\beta, \beta + \tau) d\beta + \int_{t_1}^{nT} R_x(\beta, \beta + \tau) d\beta \right\} = \\
 &= \frac{1}{T} \left\{ \int_{(n-1)T}^{nT} R_x(\beta, \beta + \tau) d\beta \right\} = \frac{1}{T} \left\{ \int_0^T R_x(\delta, \delta + \tau) d\delta \right\} = \\
 &= \frac{1}{T} \int_0^T s_Q(\delta) s_Q(\delta + \tau) d\delta = \frac{1}{T} s_Q(\tau) * s_Q(-\tau) = \frac{\cos(2\pi f_c \tau)}{2} \Lambda\left(\frac{|\tau|}{T}\right)
 \end{aligned}$$

Exemplo: Modulação PSK

- Que é função apenas de τ . Portanto, $\mathbf{x}'(t)$ é estacionário no sentido amplo.



Ergodicidade

- Ergodicidade está relacionada com o problema da determinação das estatísticas de um processo $\mathbf{x}(t)$ a partir de uma única amostra desse processo.
- Assim, um processo $\mathbf{x}(t)$ é ergódico, na forma mais geral, se (com probabilidade 1) todas as suas estatísticas podem ser determinadas a partir de uma única função amostra $x(t)$ do processo.
- Em outras palavras, $\mathbf{x}(t)$ é ergódico se as médias no tempo são iguais às medias nas amostras (valor esperado).
- Claramente, para um processo ser ergódico ele precisa ser estacionário.
- Será admitido no resto desta seção que o processo $\mathbf{x}(t)$ é estacionário, pelo menos no sentido amplo.

Ergodicidade da média

- Introduzindo-se a média no tempo,

$$\frac{1}{2T} \int_{-T}^T \mathbf{x}(t) dt = \mathbf{m}_T$$

vemos que $\mathbf{m}_T$ é uma variável aleatória, e como $E\{\mathbf{x}(t)\} = \text{constante}$,

$$E\{\mathbf{m}_T\} = E\{\mathbf{x}(t)\} = m_x$$

- e a variância de $\mathbf{m}_T$ é dada por (demonstração fica como exercício)

$$\sigma_{\mathbf{m}_T}^2 = \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R_x(\tau) - m_x^2] d\tau$$

- Se, quando T tende a infinito, essa variância tende a zero, então

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \mathbf{x}(t) dt = E\{\mathbf{x}(t)\} = m_x$$

Ergodicidade da média

- e, portanto, a média no tempo de qualquer função amostra é igual a m_x (no sentido quadrático médio):

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt = m_x$$

Ergodicidade da correlação

- Analogamente, pode ser provado que

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t+\tau) x(t) dt = E\{\mathbf{x}(t+\tau) \mathbf{x}(t)\} = R_{\mathbf{x}}(\tau)$$

- se e somente se

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\lambda}{2T}\right) \left[E\{\mathbf{x}(t+\tau+\lambda) \mathbf{x}(t+\lambda) \mathbf{x}(t+\tau) \mathbf{x}(t)\} - R_{\mathbf{x}}^2(\tau) \right] d\lambda = 0$$

- Nesse caso, então, a autocorrelação no tempo iguala a autocorrelação estatística.
- Note que para se testar a ergodicidade da média é suficiente conhecer $m_{\mathbf{x}}$ e $R_{\mathbf{x}}(\tau)$, porém para a ergodicidade da autocorrelação é necessário o conhecimento de momentos de quarta ordem.

Exemplo: senóide com fase aleatória

- $x(t) = A \cos(2 \pi f_0 t + \theta)$,
- $m_x=0$ e $R_x(t,t+\tau)=A^2/2 \cos(2 \pi f_0 \tau)$.

- Vamos calcular a média no tempo de uma amostra genérica:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T A \cos(2 \pi f_0 t + \theta) dt = \lim_{T \rightarrow \infty} \frac{A}{2T} \frac{2 \sin(2 \pi f_0 T + \theta)}{2 \pi f_0} = 0$$

- Portanto o processo é ergódico na média.
- Para a autocorrelação:

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t+\tau) x(t) dt &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T A \cos[2 \pi f_0(t+\tau) + \theta] A \cos[2 \pi f_0 t + \theta] dt = \\ &= \lim_{T \rightarrow \infty} \frac{A^2}{4T} \int_{-T}^T \cos[2 \pi f_0(2t+\tau) + 2\theta] + \cos[2 \pi f_0 \tau] dt = \lim_{T \rightarrow \infty} \frac{A^2}{2} \cos(2 \pi f_0 \tau) = R_x(\tau) \end{aligned}$$

- Portanto ele é ergódico no sentido amplo.

Processos complexos

- Consideraremos, nesta seção, apenas processos que sejam, pelo menos, estacionários no sentido amplo, podendo, eventualmente, ser processos complexos.
- Como definimos anteriormente, a média de um processo aleatório estacionário é dada por $m_x = E\{\mathbf{x}(t)\}$
- Sua autocorrelação, no caso geral de processos complexos, é dada por

$$R_x(\tau) = E\{\mathbf{x}(t+\tau) \mathbf{x}^*(t)\} = E\{\mathbf{x}(t) \mathbf{x}^*(t-\tau)\} = E\{\mathbf{x}^*(t) \mathbf{x}(t-\tau)\}^* = R_x^*(-\tau)$$

- Se $\mathbf{x}(t)$ for real, então $R_x(\tau) = R_x(-\tau)$.
- A correlação cruzada de dois processos aleatórios é dada por

$$R_{xy}(\tau) = E\{\mathbf{x}(t+\tau) \mathbf{y}^*(t)\} = R_{yx}^*(-\tau).$$

Processos complexos

- Definiu-se, também, a autocovariância,

$$C_x(\tau) = E\{[\mathbf{x}(t+\tau) - m_x][\mathbf{x}^*(t) - m_x^*]\} = R_x(\tau) - |m_x|^2$$

- e a covariância cruzada,

$$C_{xy}(\tau) = E\{[\mathbf{x}(t+\tau) - m_x][\mathbf{y}^*(t) - m_y^*]\} = R_{xy}(\tau) - m_x m_y^*$$

- Algumas propriedades simples podem ser derivadas dessas definições:
 - se $\mathbf{z}(t) = \mathbf{x}(t) + \mathbf{y}(t)$, então $R_z(\tau) = R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{yx}(\tau)$
 - se $\mathbf{x}(t)$ e $\mathbf{y}(t)$ são processos independentes e se $\mathbf{w}(t) = \mathbf{x}(t) \mathbf{y}(t)$, então $R_w(\tau) = R_x(\tau) R_y(\tau)$
 - $R_x(0) = E\{|\mathbf{x}(t)|^2\} \geq 0$

Processos complexos

$$- |R_x(\tau)| \leq |R_x(0)| :$$

$$E \{ |\mathbf{x}(t) - a \mathbf{y}(t)|^2 \} = E \{ |\mathbf{x}(t)|^2 \} + |a|^2 E \{ |\mathbf{y}(t)|^2 \} - a E \{ \mathbf{x}^*(t) \mathbf{y}(t) \} - a^* E \{ \mathbf{x}(t) \mathbf{y}^*(t) \} \geq 0$$

Escolhendo-se
$$a = \frac{E \{ \mathbf{x}(t) \mathbf{y}^*(t) \}}{E \{ |\mathbf{y}(t)|^2 \}}$$

temos

$$\begin{aligned} & E \{ |\mathbf{x}(t)|^2 \} + \frac{|E \{ \mathbf{x}(t) \mathbf{y}^*(t) \}|^2}{E^2 \{ |\mathbf{y}(t)|^2 \}} E \{ |\mathbf{y}(t)|^2 \} - \\ & - \frac{E \{ \mathbf{x}(t) \mathbf{y}^*(t) \}}{E \{ |\mathbf{y}(t)|^2 \}} E \{ \mathbf{x}^*(t) \mathbf{y}(t) \} - \frac{E \{ \mathbf{x}^*(t) \mathbf{y}(t) \}}{E \{ |\mathbf{y}(t)|^2 \}} E \{ \mathbf{x}(t) \mathbf{y}^*(t) \} = \\ & = E \{ |\mathbf{x}(t)|^2 \} + \frac{|E \{ \mathbf{x}(t) \mathbf{y}^*(t) \}|^2}{E \{ |\mathbf{y}(t)|^2 \}} - \frac{|E \{ \mathbf{x}(t) \mathbf{y}^*(t) \}|^2}{E \{ |\mathbf{y}(t)|^2 \}} - \frac{|E \{ \mathbf{x}^*(t) \mathbf{y}(t) \}|^2}{E \{ |\mathbf{y}(t)|^2 \}} \geq 0 \Rightarrow \\ & \Rightarrow E \{ |\mathbf{x}(t)|^2 \} E \{ |\mathbf{y}(t)|^2 \} \geq |E \{ \mathbf{x}^*(t) \mathbf{y}(t) \}|^2 \end{aligned}$$

Processos complexos

Assim, fazendo-se $y(t) = x(t + \tau)$, temos

$$|R_x(0)|^2 \geq |R_x(\tau)|^2$$

No caso de processos de tempo discreto, vale, de forma similar,

$$|R_x(0)| \geq |R_x(n)|$$

$$- \quad |R_{xy}(\tau)|^2 \leq R_x(0) R_y(0)$$

Prova: Similar à anterior, substituindo-se $y(t)$ por $y(t + \tau)$.

No caso de processos de tempo discreto, vale, de forma similar,

$$|R_{xy}(n)|^2 \leq R_x(0) R_y(0)$$

- Note que, neste caso, a igualdade só ocorre quando, para algum ε ,

$$a \quad y(t + \varepsilon) = x(t)$$