Capítulo 1

Variáveis Aleatórias e Processos Estocásticos

1.1 Variáveis aleatórias

Ao tratar de variáveis aleatórias, estaremos referindo a resultados de experimentos. Um exemplo muito simples e muito comum é jogar uma moeda. Em princípio, de acordo com as leis da mecânica, seria possível prever o resultado. Porém, as condições iniciais do arremesso, a interação com o ar durante o movimento e a interação com a mesa durante o choque inelástico na queda deveriam ser completamente determinadas. Como isso é praticamente impossível, é mais conveniente dizer que o resultado é imprevisível e o máximo que podemos fazer é estimar a probabilidade de sair cara ou coroa. Se a moeda for equilibrada e os lançamentos não forem maliciosos, depois de um grande número de realizações teremos números muito próximos de ocorrências de cara e coroa. Concluímos então que a probabilidade de se obter um dos dois possíveis resultados é $\frac{1}{2}$. Podemos estender o exemplo acima para casos em que o número de fatores incontroláveis que podem influenciar o resultado de um experimento é extremamente grande, como no caso do valor da tensão elétrica na saída de um medidor de temperatura ambiente em uma determinada hora do dia, o número de pulsos produzidos por um contador de fótons em um certo intervalo, etc. As variáveis aleatórias podem ser contínuas, como um sinal elétrico que pode assumir qualquer valor entre $V_{\text{mín}}$ e $V_{\text{máx}}$, ou discretas, como por exemplo o número de pulsos por tempo de amostragem de um contador de fótons. Sempre que possível, daremos exemplos com variáveis discretas, passando ao limite do contínuo quando necessário.

O conceito de variável aleatória não se limita a variáveis unidimensionais e nem a variáveis reais. Neste curso utilizaremos amplamente as variáveis aleatórias complexas.

Podemos modelar uma variável aleatória discreta como um sorteio feito sobre o conjunto Ω de todos os resultados possíveis de um experimento (espaço amostral)[1], sendo que a cada resultado ω_i (evento elementar) está associada uma chance de ocorrência p_i (probabilidade), tal que $\sum_i p_i = 1$. Não nos deteremos aqui nos detalhes da formulação da teoria de probabilidades. O leitor interessado pode consultar a referência [1].

1.2 Sequências aleatórias e processos estocásticos

Entendemos por função estocástica uma função qualquer cujos valores assumidos são variáveis aleatórias. As funções aleatórias podem ser de argumento discreto, caso em que recebem o nome de *sequências aleatórias*, ou de argumento contínuo. Exemplos de sequências aleatórias: números de fótons detectados em intervalos de amostragem subsequentes, valores de temperatura ambiente registrados de hora em hora, etc. No caso funções de variável contínua, a variável é em geral o tempo, caso em que essas funções são chamadas de *processos estocásticos*, ou às vezes a posição. Exemplo de processo estocástico: Sinal elétrico de um microfone registrando o ruído

[1] No caso de variáveis complexas, podemos considerar que tanto a parte real como a parte imaginária são resultados de experimentos.

ambiente. Por outro lado, se este ruído for registrado por um gravador digital, o tempo será discretizado por uma determinada taxa de amostragem e o valor da tensão elétrica será discretizado por um conversor analógico-digital. Teremos aí uma sequência aleatória (tempo discretizado) de uma variável discreta (bytes). Devido às características do ouvido e de todo o processamento de sinal no corpo, percebemos essa sequência aleatória como um processo estocástico. Neste curso, não faremos distinção entre sequências e processos e chamaremos as funções aleatórias genericamente de processos estocásticos.

1.2.1 Caracterização de processos estocásticos

Seja $\xi(t)$ um processo estocástico e seja $F_t(x) = \text{Prob}\{\xi(t) < x\}$ a função de distribuição acumulada de ξ [1, 2]. Estendendo a definição de $F_t(x)$, temos

$$F_t(x) = \text{Prob}\{\xi(t) < x\}, \tag{1.1}$$

$$F_{t_1,t_2}(x_1, x_2) = \text{Prob}\{\xi(t_1) < x_1 \,\&\, \xi(t_2) < x_2\}, \tag{1.2}$$

$$F_{t_1,t_2,...,t_n}(x_1, x_2, ..., x_n) = \text{Prob}\{\xi(t_1) < x_1 \,\&\, \xi(t_2) < x_2 \,\&...\&\, \xi(t_n) < x_n\}. \tag{1.3}$$

Dizemos que um processo estocástico está completamente caracterizado se conhecemos todas as suas distribuições acumuladas $F_{t_1,...,t_n}$ para qualquer n tal que todos os t_n pertençam ao domínio de $\xi(t)$.

1.2.2 Representação de *ensemble*

Podemos entender um processo estocástico como uma sequência temporal de experimentos cujos resultados são sorteados no conjunto Ω de possibilidades para cada valor de t. Tratando-se de sorteios, se tivermos dois processos derivados de sistemas físicos idênticos submetidos às mesmas condições externas, os valores sorteados nos dois experimentos ao mesmo tempo serão em geral diferentes. Como exemplo, podemos imaginar o sinal elétrico de dois detectores idênticos captando a luz de dois corpos negros à mesma temperatura.

Uma outra maneira de representar um processo estocástico é por meio de um *ensemble*, isto é, o conjunto hipotético infinito de todas as possíveis realizações de $\xi(t)$. Assim, o valor de $\xi(t_1)$ para um determinado processo pode ser visto como um sorteio sobre todas as possíveis realizações temporais de $\xi(t)$ em $t = t_1$, conforme representado na Fig. 1. As funções de distribuição $F_{t_1,t_2,...,t_n}(x_1, x_2, ..., x_n)$ poderiam, em princípio, ser obtidas do *ensemble*.

1.2.3 Médias e correlações

Na prática, é em geral impossível de se obterem todas as funções de distribuição de probabilidade $F_{t_1,t_2,...,t_n}(x_1, x_2, ..., x_n)$ para um certo n de interesse. Em vez disso, os processos estocásticos são estudados em termos das médias e das correlações de ξ em diferentes tempos, mensuráveis e definidas como

$$\left\langle \xi^k(t) \right\rangle = \int x^k \, dF_t(x), \tag{1.4}$$

$$\left\langle \xi^{k_1}(t_1)\xi^{k_2}(t_2) \right\rangle = \iint x_1^{k_1} x_2^{k_2} \, dF_{t_1,t_2}(x_1, x_2), \tag{1.5}$$

$$\left\langle \xi^{k_1}(t_1)\xi^{k_2}(t_2)...\xi^{k_n}(t_n) \right\rangle = \int ... \int x_1^{k_1} x_2^{k_2} ... x_n^{k_n} \, dF_{t_1,t_2,...t_n}(x_1, x_2, ..., x_n), \tag{1.6}$$

Onde todas as integrais vão de $-\infty$ a $+\infty$. Aqui foi usado o conceito de integral de Stiltjes [1]:

$$\int_a^b g(x) \, dF(x) = \lim_{n\to\infty} \sum_{k=1}^{n} g(x_k)[F(x_k) - F(x_{k-1})], \tag{1.7}$$

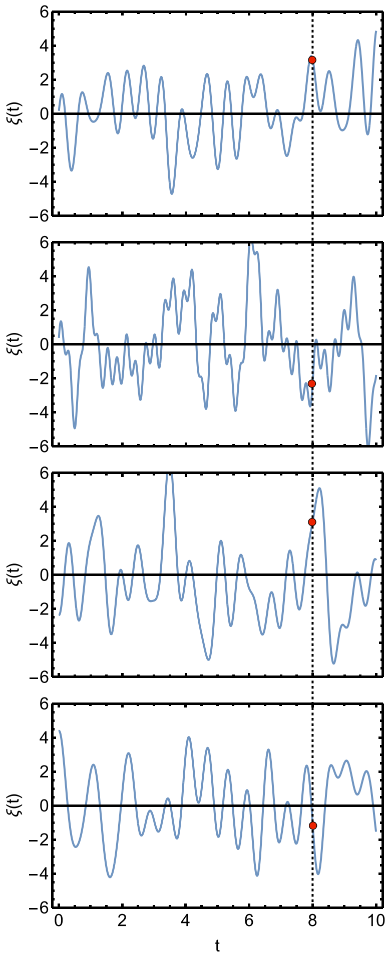

Figura 1.1
Representação de 4 possíveis realizações de um processo estocástico $\xi(t)$. Os pontos vermelhos mostram quatro possíveis valores de ξ que podem ser obtidos em um único experimento realizado em $t = 8$.

onde o intervalo $[a, b]$ é dividido em n partes e $|x_k - x_{k-1}| \to 0$ quando $n \to \infty$. A função $F(x)$ é não decrescente e é chamada de *medida*. A razão de usarmos a integral de Stieltjes é que ela permite tratar simultaneamente dos casos contínuo e discreto com as mesmas expressões. Por exemplo, se a função distribuição $F_t(x)$ é contínua, a Eq.(1.4) se transforma em uma integral de Riemann:

$$\left\langle \xi^k(t) \right\rangle = \int x^k \, dF_t(x) = \int x^k \, \frac{dF_t(x)}{dx} dx = \int x^k \, P_t(x) dx, \tag{1.8}$$

onde $P_t(x)$ é a densidade de probabilidade da variável $\xi(t)$. Se $F_t(x)$ tiver a forma de uma escada, tal que $\frac{dF_t(x)}{dx} = 0$ para a maioria dos valores de x e $\frac{dF_t(x)}{dx} \propto \delta(x - x_i)$ para um conjunto de pontos $\{x_1, x_2, ...\}$, a Eq.(1.4) transforma-se em

$$\left\langle \xi^k(t) \right\rangle = \int x^k \, dF_t(x) = \int x^k \, \frac{dF_t(x)}{dx} dx = \sum_i x_i^k \, P_t(x_i). \tag{1.9}$$

As médias e correlações são medidas na prática como

$$\left\langle \xi^k(t) \right\rangle_t = \frac{1}{T} \int_0^T \xi^k(t) \, dt, \tag{1.10}$$

$$\left\langle \xi^{k_1}(t) \xi^{k_2}(t+\tau) \right\rangle_t = \frac{1}{T} \int_0^T \xi^{k_1}(t) \xi^{k_2}(t+\tau) \, dt, \tag{1.11}$$

$$\left\langle \xi^{k_1}(t) \xi^{k_2}(t+\tau_1) ... \xi^{k_n}(t+\tau_n) \right\rangle_t = \frac{1}{T} \int_0^T \xi^{k_1}(t) \xi^{k_2}(t+\tau_1) ... \xi^{k_n}(t+\tau_n) \, dt. \tag{1.12}$$

para T muito grande. A rigor, deveríamos tomar o limite $T \to \infty$. Em princípio, para um processo qualquer $\xi(t)$, não há nenhuma garantia de que $\langle ... \rangle$ definidas nas Eqs. (1.4–1.6) e $\langle ... \rangle_t$ definidas nas Eqs. (1.10–1.12) sejam equivalentes.

1.3 Processos estocásticos estacionários

Dizemos que um processo estocástico no tempo $\xi(t)$ é estacionário se as funções de distribuição $F_{t_1, t_2, ..., t_n}(x_1, x_2, ..., x_n)$ são invariantes por translações no tempo, isto é, se

$$F_{t_1+\tau, t_2+\tau, ..., t_n+\tau}(x_1, x_2, ..., x_n) = F_{t_1, t_2, ..., t_n}(x_1, x_2, ..., x_n). \tag{1.13}$$

Neste caso, nem as Eqs. (1.4–1.6) nem as Eqs. (1.10–1.12) dependem de t. Para processos estacionários, o teorema ergódico garante que as médias de *ensemble* $\langle ... \rangle$ e as médias temporais $\langle ... \rangle_t$ são equivalentes.

O conceito de processo estacionário tem um papel importantíssimo no desenvolvimento da teoria de processos estocásticos, como será comprovado nos capítulos seguintes. Na grande maioria dos casos, nosso interesse se limitará à média $\langle \xi(t) \rangle$ e às propriedades de funções de correlação do tipo $\langle \xi(t) \xi(t+\tau) \rangle$. Neste contexto, o conceito de processo estacionário é relaxado para o de *processo estacionário no sentido amplo*, no qual assumimos apenas que $\langle \xi(t) \rangle$ e $\langle \xi(t) \xi(t+\tau) \rangle$ são independentes de t.

1.3.1 Propriedades da função de correlação

Vamos trabalhar de agora em diante com variáveis aleatórias complexas. Como consequências da sua definição, a função de correlação de processos estacionários

$$\Gamma(\tau) = \langle \xi^*(t) \xi(t+\tau) \rangle \tag{1.14}$$

tem as seguintes propriedades, que utilizaremos sem demonstração [2, 3]:

1. $\Gamma(0) > 0$,

2. $\Gamma(-\tau) = \Gamma(\tau)$,

3. $|\Gamma(\tau)| \leq \Gamma(0)$.

4. $\Gamma(\tau)$ é contínua.

Se tomarmos uma sequência de tempos $\{t_1, t_2, ..., t_n\}$ e a cada um deles associarmos um número real qualquer $\{(t_1, a_1), (t_2, a_2), ..., (t_n, a_n)\}$, temos que

5. $= \sum_{j,k=1}^{n} \Gamma(t_j - t_k) a_j a_k \geq 0$.

A propriedade 5. quer dizer que $\Gamma(\tau)$ é positiva semi-definida.

O conceito de função de correlação pode ser estendido para o de correlação cruzada de dois processos estocásticos $\xi_1(t)$ e $\xi_2(t)$ como

$$\Gamma_{12}(\tau) = \langle \xi_1^*(t) \xi_2(t+\tau) \rangle . \tag{1.15}$$

1.3.2 Espectro de potência e o teorema de Wiener-Khinchin

Com base no teorema ergódico, o valor médio de um processo estocástico pode ser sempre obtido experimentalmente com boa precisão por meio da média temporal de $\xi(t)$ tomada em tempos suficientemente longos[2]. Assim, nos casos em que $\langle \xi(t) \rangle \neq 0$, é muito mais conveniente trabalhar com as flutuações de $\Delta\xi(t) = \xi(t) - \langle \xi(t) \rangle$. Em óptica, se estivermos tratando dos campos oscilantes, estes já têm naturalmente média nula. Se estivermos tratando de intensidades, trabalharemos com as suas flutuações sempre que for conveniente. Portanto, no restante deste capítulo, consideraremos apenas processos estocásticos de média nula.

A abordagem padrão da óptica estatística clássica é tratar as flutuações temporais do campo eletromagnético como um processo estocástico estacionário. Isto é obviamente uma idealização, mas que dá excelentes resultados. Da mesma maneira, as flutuações de tensão e corrente em circuitos elétricos são frequentemente tratadas como processos estacionários. Existem outros inúmeros exemplos semelhantes.

Uma das técnicas experimentais mais difundidas na análise de sinais, seja qual for a origem do sinal, é a análise espectral. Por exemplo, se um raio de luz proveniente de uma fonte térmica incide sobre uma rede de difração, a luz espalhada se separa em diferentes frequências de acordo com o ângulo de observação. Se colocarmos um detector a uma certa distância da rede de difração e fizermos uma varredura angular, teremos como resultado, após as correções de eficiência da rede e resposta espectral do detector, o espectro de potência da fonte de luz. O espectro de potência nos fornece a potência relativa $S(\omega)$ da fonte para cada intervalo infinitesimal de frequência $[\omega, \omega + d\omega]$. Em princípio, poderíamos supor que uma vez conhecido o sinal $\xi(t)$, seu espectro de potência pode ser obtido por meio da sua transformada de Fourier $\tilde{\xi}(\omega)$ e vice-versa. Entretanto, um processo estocástico estacionário não pode estar limitado no tempo e consequentemente não satisfaz as condições básicas para a existência de sua transformada de Fourier [2, 4]. Mesmo assim, é possível mostrar que um processo estocástico estacionário de média nula pode ser sempre escrito como uma superposição infinita de funções harmônicas do tempo com amplitudes aleatórias descorrelacionadas[2]:

$$\xi(t) = \int e^{i\omega t} dZ(\omega), \tag{1.16}$$

onde $Z(\omega)$ é uma função estocástica para a qual $\langle dZ^*(\omega) dZ(\omega') \rangle = S(\omega)\delta(\omega - \omega') d\omega d\omega'$.

[2] Por ser um procedimento experimental, a determinação de $\langle \xi(t) \rangle$ terá sempre uma incerteza, ainda que possa ser muito pequena. O tempo de amostragem T deve ser grande, porém dentro do limite para o qual a aproximação do processo como estacionário é válida. Por exemplo, as características macroscópicas da fonte de sinal podem variar ao longo de minutos, horas, dias, anos, etc.

O espectro de potência $S(\omega)$ de um processo estocástico estacionário de média nula é dado pelo teorema de Wiener-Khinchin [2]:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle \xi^*(t) \xi(t+\tau) \rangle \, e^{i\omega\tau} \, d\tau. \tag{1.17}$$

Capítulo 2

Teoria Clássica da Coerência Óptica

Na teoria clássica da coerência, lidamos com variados tipos de fontes de luz, incluindo fontes térmicas, lasers, diodos emissores, etc. Os campos eletromagnéticos gerados por essas fontes dependem de uma enorme quantidade de fatores macroscópicos e microscópicos, como por exemplo, as flutuações térmicas dos átomos emissores e as flutuações mecânicas da cavidade de um laser. Devido às diversas fontes de flutuações na produção dos campos, estes só podem ser tratados corretamente se forem considerados como quantidades que flutuam aleatoriamente. Portanto, trataremos os campos como processos estocásticos. Para que possamos tirar máximo proveito da teoria, aproximaremos esses processos por estacionários. Esta aproximação fornece excelentes resultados, desde que as características macroscópicas das fontes (temperatura, dimensões, composição química, etc) sejam suficientemente estáveis dentro do intervalo de tempo estudado.

Em sua definição operacional, os campos elétrico e magnético de uma onda eletromagnética são funções reais da posição e do tempo. Na maioria dos casos, é mais vantajoso fazer os cálculos envolvendo esses campos tratando-os como funções complexas. A seguir, apresentamos uma maneira bastante utilizada para associar os campos a funções complexas.

2.1 O sinal analítico

Seja uma função real $U(t)$, que representa uma componente cartesiana do campo, escrita em termos de sua transformada de Fourier,

$$U(t) = \int_{-\infty}^{\infty} u(\omega)\, e^{-i\omega t} d\omega. \tag{2.1}$$

Como $U(t)$ é real, $U^*(t) = U(t)$ e portanto

$$u^*(\omega) = u(-\omega). \tag{2.2}$$

Já que $u(-\omega)$ pode ser obtido de $u(\omega)$, podemos utilizar somente frequências positivas e representar o campo pela função complexa

$$V(t) = \int_{0}^{\infty} u(\omega)\, e^{-i\omega t} d\omega. \tag{2.3}$$

A função $V(t)$ é chamada de *sinal analítico*, por razões que não discutiremos aqui, e tem algumas propriedades interessantes. É possível mostrar que [3]

$$V(t) = \frac{1}{2}[U(t) + i\tilde{U}(t)], \tag{2.4}$$

onde a função real $\tilde{U}(t)$ é a transformada de Hilbert de $U(t)$. Logo, $U(t) = 2\mathrm{Re}\{V(t)\}$.

Se $U(t)$ é um processo estocástico estacionário, então $V(t)$ também o será[1], e [3],

$$\left\langle U^2(t) \right\rangle = \frac{1}{2} \left\langle V^*(t)V(t) \right\rangle . \tag{2.5}$$

Neste curso, faremos uso constante de campos estocásticos quase monocromáticos, isto é, processos para os quais o espectro de potência $S(\omega)$ está centrado em torno de uma frequência $\bar{\omega}$, com largura $\Delta\omega \ll \bar{\omega}$. Isto quer dizer, com base no teorema de Wiener-Khinchin (1.17), que o tempo de decaimento da função de correlação $\Gamma(\tau) = \langle V^*(t)V(t+\tau) \rangle$ é muito maior que $2\pi/\bar{\omega}$.

[1] Pode parecer contraditório identificar $V(t)$ com um processo estocástico estacionário, uma vez que a primeira definição dada foi através de uma transformada de Fourier (2.3), que não se aplica a processos estocásticos estacionários. Entretanto, a definição (2.4) utilizando a transformada de Hilbert é consistente [5].

Referências Bibliográficas

[1] Barry L. James, *Probabilidade: um curso em nível intermediário*, IMPA (2004).

[2] A. M. Yaglom, *Correlation theory of stationary and related random functions - Volume I: Basic results*, Springer-Verlag (1987).

[3] Leonard Mandel and Emil Wolf, *Optical coherence and quantum optics*, Cambridge (1995).

[4] Ronald N. Bracewell, *The Fourier transform and its applications*, McGraw-Hill (2000).

[5] Georg Lindgren, *Stationary stochastic processes - theory and applications*, CRC (2012).