

EE881 – Princípios de Comunicações I

Resumo de Ferramentas Matemáticas

Henrique Miyamoto

1º semestre 2022

NOTAÇÃO

x, \mathbf{v} e \mathbf{M} representam escalares, vetores e matrizes determinísticas, respectivamente. X e \mathbf{X} representam variáveis aleatórias e vetores aleatórios, respectivamente. A unidade imaginária é denotada $\mathbf{j} = \sqrt{-1}$. O conjugado complexo de z é denotado z^* . A função indicadora $\mathbb{1}_{\mathcal{A}}(x)$ vale 1 se $x \in \mathcal{A}$, e 0 se $x \notin \mathcal{A}$. A abreviação *sse* significa “se, e somente se”.

1. ÁLGEBRA LINEAR E MATRIZES

A. Álgebra Linear

Definição 1.1. Um *espaço vetorial* é formado por:

- um corpo \mathbb{F} (e.g., \mathbb{R} ou \mathbb{C});
- um conjunto \mathcal{V} de objetos chamados *vetores*;
- uma operação *soma*, que associa $u, v \in \mathcal{V}$ ao vetor $u + v$, satisfazendo, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$:
 - 1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
 - 2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
 - 3) $\exists! \mathbf{0} \in \mathcal{V}$ t.q. $\mathbf{u} + \mathbf{0} = \mathbf{u}$;
 - 4) $\forall \mathbf{u} \in \mathcal{V}, \exists \mathbf{v}$ t.q. $\mathbf{u} + \mathbf{v} = \mathbf{0}$.
- uma operação *produto*, que associa $\alpha \in \mathbb{F}$ e $u \in \mathcal{V}$ ao vetor αu , satisfazendo, para todos $\alpha, \beta \in \mathbb{F}$ e $\mathbf{u}, \mathbf{v} \in \mathcal{V}$:
 - 1) $1\mathbf{u} = \mathbf{u}$;
 - 2) $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u})$;
 - 3) $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$;
 - 4) $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$.

São exemplos o espaço euclidiano \mathbb{R}^n , o espaço \mathbb{C}^n , o espaço $C^0([a, b])$ das funções contínuas complexas $f: [a, b] \rightarrow \mathbb{C}$.

Definição 1.2. Seja \mathcal{V} um espaço vetorial sobre \mathbb{C} . Um *produto interno* é uma operação $\langle \cdot, \cdot \rangle: \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$, que associa $(\mathbf{u}, \mathbf{v}) \mapsto \langle \mathbf{u}, \mathbf{v} \rangle$, satisfazendo:

- 1) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$;
- 2) $\langle \alpha\mathbf{u}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$;
- 3) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle^*$;
- 4) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, com igualdade sse $\mathbf{u} = \mathbf{0}$.

Um espaço vetorial munido de um produto interno é dito *espaço com produto interno*.

Proposição 1.1. *Valem as propriedades adicionais:*

- 1) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$;
- 2) $\langle \mathbf{u}, \alpha\mathbf{v} \rangle = \alpha^*\langle \mathbf{u}, \mathbf{v} \rangle$.

Concretamente, interessam-nos o produto interno canônico do espaço \mathbb{C}^n ,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n u_i v_i^*, \quad (1)$$

e o produto interno do espaço de funções $C^0([a, b])$,

$$\langle f, g \rangle = \int_a^b f(t)[g(t)]^* dt. \quad (2)$$

Um produto interno induz uma *norma* $\|\mathbf{u}\| := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ para os vetores $\mathbf{u} \in \mathcal{V}$ do espaço vetorial.

Teorema 1.1. *Seja \mathcal{V} espaço com produto interno, $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ e $\alpha \in \mathbb{C}$. Valem:*

- 1) $\|\mathbf{u}\| \geq 0$, com igualdade sse $\|\mathbf{u}\| = 0$;
- 2) $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$;
- 3) $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$ (*Cauchy-Schwarz*);
- 4) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (*desigualdade triangular*);
- 5) $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\Re\{\langle \mathbf{u}, \mathbf{v} \rangle\}$;
- 6) $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$.

Definição 1.3. Seja \mathcal{V} espaço com produto interno. Dizemos que dois vetores $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ são *ortogonais* se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Definição 1.4. Uma *base* \mathcal{B} do espaço vetorial \mathcal{V} é um subconjunto de vetores $\mathcal{B} \subset \mathcal{V}$ tais que eles são linearmente independentes e geram o espaço \mathcal{V} .

Se $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ é uma base do espaço vetorial \mathcal{V} de dimensão finita, então qualquer vetor $\mathbf{u} \in \mathcal{V}$ pode ser escrito de forma única como $\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{b}_i$ (i.e., os coeficientes α_i são únicos).

Definição 1.5. Uma base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ para o espaço vetorial \mathcal{V} é dita *ortonormal* se satisfaz $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$, se $i \neq j$, e $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 1$, se $i = j$.

O procedimento de *ortonormalização de Gram-Schmidt* permite encontrar uma base ortonormal $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ de \mathcal{V} , dado um conjunto de vetores linearmente independentes $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathcal{V}$:

- comece com $\mathbf{b}_1 := \mathbf{u}_1 / \|\mathbf{u}_1\|$;
- para $i = 2, \dots, n$, faça:

$$\mathbf{v}_i := \mathbf{u}_i - \sum_{j=1}^{i-1} \langle \mathbf{u}_i, \mathbf{b}_j \rangle \mathbf{b}_j \quad \text{e} \quad \mathbf{b}_i := \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}.$$

B. Matrizes

A *matriz adjunta* ou *transposto hermitiano* de uma matriz \mathbf{H} é a matriz $\mathbf{H}^\dagger := (\mathbf{H}^\top)^*$.

Definição 1.6. Uma matriz $U \in \mathbb{C}^{n \times n}$ é dita *unitária* se $U^\dagger U = \mathbf{I}$. Se, além disso, ela tem entradas reais, então é dita *ortogonal*.

Teorema 1.2. Seja $U \in \mathbb{C}^{n \times n}$. São equivalentes:

- U é unitária;
- U é não-singular e $U^\dagger = U^{-1}$;
- $UU^\dagger = \mathbf{I}$;
- U^\dagger é unitária;
- as colunas de U formam conjunto ortonormal;
- as linhas de U formam conjunto ortonormal;
- para todo $\mathbf{x} \in \mathbb{C}^n$, defina $\mathbf{y} := U\mathbf{x}$; então $\mathbf{y}^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{x}$.

Teorema 1.3. Qualquer matriz quadrada A pode ser escrita como $A = URU^\dagger$, em que U é unitária e R é triangular superior, cujas entradas diagonais são autovalores de A .

Definição 1.7. Uma matriz $H \in \mathbb{C}^{n \times n}$ é dita *hermitiana* se $H = H^\dagger$, e *anti-hermitiana* se $H = -H^\dagger$. Se H é hermitiana (resp., anti-hermitiana) e tem entradas reais, então é dita *simétrica* (resp., *antissimétrica*).

Teorema 1.4. Uma matriz hermitiana $H \in \mathbb{C}^{n \times n}$ pode ser escrita como

$$H = U\Lambda U^\dagger = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger, \quad (3)$$

em que U é unitária e $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ é composta pelos autovalores λ_i de H . Ademais, os autovalores são reais e a i -ésima coluna de U , \mathbf{u}_i , é um autovetor associado a λ_i .

Definição 1.8. Uma matriz hermitiana $H \in \mathbb{C}^{n \times n}$ é dita *definida positiva* (resp., *semi-definida positiva*) se

$$\mathbf{u}^\dagger H \mathbf{u} > 0 \quad (\text{resp., } \mathbf{u}^\dagger H \mathbf{u} \geq 0)$$

para todo $\mathbf{u} \in \mathbb{C}^n$, com $\mathbf{u} \neq \mathbf{0}$.

Os autovalores de uma matriz definida positiva (resp., semi-definida positiva) são positivos (resp., não-negativos).

Pela *decomposição em valores singulares (SVD)*, qualquer matriz $A \in \mathbb{C}^{m \times n}$ pode ser escrita na forma $A = UDV^\dagger$, onde $U \in \mathbb{C}^{m \times m}$ e $V \in \mathbb{C}^{n \times n}$ são matrizes unitárias e $D \in \mathbb{R}^{m \times n}$ é uma matriz diagonal, com entradas não negativas.

2. ANÁLISE DE SINAIS

Estamos interessados nos sinais $g : \mathbb{R} \rightarrow \mathbb{C}$ com energia finita¹. A *transformada de Fourier* $g_{\mathcal{F}}$ de uma função $g : \mathbb{R} \rightarrow \mathbb{C}$ é dada por

$$g_{\mathcal{F}}(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt \quad (4)$$

e a *transformada inversa* de $g_{\mathcal{F}} : \mathbb{R} \rightarrow \mathbb{C}$ é

$$g(t) = \int_{-\infty}^{\infty} g_{\mathcal{F}}(f) e^{j2\pi ft} df. \quad (5)$$

¹A rigor, nas funções \mathcal{L}_2 , i.e., Lebesgue-mensuráveis e com integral (de Lebesgue) $\int_{-\infty}^{\infty} |g(t)|^2 dt < \infty$. A integral de Lebesgue é uma generalização da integral de Riemann.

Teorema 2.1 (Parseval). Sejam $f(t)$ e $g(t)$ duas funções com energia finita e $f_{\mathcal{F}}(f)$ e $g_{\mathcal{F}}(f)$ suas respectivas transformadas de Fourier. Então

$$\langle f(t), g(t) \rangle = \langle f_{\mathcal{F}}(f), g_{\mathcal{F}}(f) \rangle \quad e \quad \|f(t)\|^2 = \|f_{\mathcal{F}}(f)\|^2.$$

Tabela I

TRANSFORMADA DE FOURIER DE ALGUMAS FUNÇÕES.

$f(t)$	$f_{\mathcal{F}}(f)$
$e^{-a t }$	$\frac{2a}{a^2 + (2\pi f)^2}$
$e^{-at}, t \geq 0$	$\frac{1}{a + j2\pi f}$
$e^{-\pi t^2}$	$e^{-\pi f^2}$
$\mathbb{1}_{[-a, a]}(t)$	$2a \text{ sinc}(2af)$
$\text{sinc}(x) := \sin(\pi x)/\pi x$	

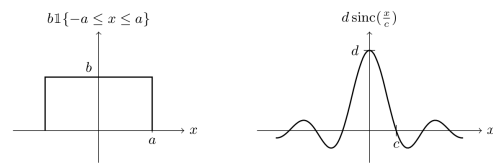


Figura 1. O truque para calcular a transformada de Fourier da função retângulo consiste em igualar as áreas do retângulo, $2ab$, com a do seno cardinal, cd (área do triângulo inscrito no maior lobo).

A *série de Fourier* de uma função $f : \mathbb{R} \rightarrow \mathbb{C}$, de período fundamental T , pode escrita

$$f(x) = \sum_{i \in \mathbb{Z}} A_i e^{j2\pi xi/T}, \quad (6)$$

com coeficientes

$$A_i = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-j2\pi xi/T} dx. \quad (7)$$

Trata-se da representação do sinal fundamental $f : [-\frac{T}{2}, \frac{T}{2}] \rightarrow \mathbb{C}$ na base ortogonal $\{\phi_i(x)\}_{i \in \mathbb{Z}}$, com

$$\phi_i(x) := e^{j2\pi xi/T} \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}(x).$$

Note que $TA_i = \langle f, \phi_i \rangle$.

3. PROBABILIDADE E ESTATÍSTICA

A. Variáveis aleatórias reais

Um *espaço de probabilidade* (Ω, \mathcal{F}, P) é formado pelo *espaço amostral* Ω (conjunto de todos resultados possíveis), pelo conjunto de *eventos* \mathcal{F} (subconjuntos de Ω aos quais uma probabilidade é associada), e pela *medida de probabilidade* P .

Uma *medida de probabilidade* é uma função $P : \mathcal{F} \rightarrow [0, 1]$ tal que $P(\emptyset) = 0$, $P(\Omega) = 1$ e $P(A_1 \cup A_2) = P(A_1) + P(A_2)$, para A_1, A_2 eventos disjuntos.

Definição 3.1. Uma *variável aleatória (v.a.) real* X sobre o espaço de probabilidade (Ω, \mathcal{F}, P) é uma função $X : \Omega \rightarrow \mathbb{R}$ tal que $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$, $\forall x \in \mathbb{R}$. A *função distribuição acumulada* é dada por $F_X(x) := P(\{\omega : X(\omega) \leq x\})$.

No caso de uma v.a. discreta, a *função massa de probabilidade* indica $P(x) = P(\{X = x\})$. No caso

contínuo, a *função densidade de probabilidade* é definida $f(x) := F'_X(x)$ e permite calcular a probabilidade $P(\{a \leq X \leq b\}) = \int_a^b f(x) dx$.

A *esperança matemática* de uma v.a. discreta é $\mathbb{E}[X] := \sum_i x_i P(x_i)$, e de uma v.a. contínua é $\mathbb{E}[X] := \int x f(x) dx$. A *variância* é dada por $\text{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Exemplo 3.1 (V.a. gaussiana). Uma v.a. gaussiana $X \sim \mathcal{N}(\mu, \sigma^2)$ de média μ e variância σ^2 tem função densidade de probabilidade (pdf) dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \quad (8)$$

Um *vetor aleatório* n -dimensional sobre o espaço de probabilidade (Ω, \mathcal{F}, P) é uma função $\mathbf{X} : \Omega \rightarrow \mathbb{R}^n$, tal que cada componente é uma v.a. sobre (Ω, \mathcal{F}, P) . Pode ser visto como uma coleção de n v.a.s: $\mathbf{X} = (X_1, \dots, X_n)$. A pdf do vetor aleatório é a pdf conjunta de X_1, \dots, X_n . A *esperança* é $\mathbb{E}[\mathbf{X}] := (\mathbb{E}[X_1], \dots, \mathbb{E}[X_n])$ e a *matriz de covariância* é $K_{\mathbf{X}} := \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top]$.

Exemplo 3.2 (Vetor aleatório gaussiano). Um vetor aleatório $\mathbf{X} = (X_1, \dots, X_n)$ formado por componentes gaussianas independentes e identicamente distribuídas (i.i.d.) $X_i \sim \mathcal{N}(0, \sigma^2)$ tem pdf

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\mathbf{x}^\top \mathbf{x}}{2\sigma^2}\right). \quad (9)$$

Em geral, se $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, K_{\mathbf{X}})$ é um vetor gaussiano com média $\boldsymbol{\mu}$ e matriz de covariância (não-singular) $K_{\mathbf{X}}$, sua pdf é

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det K_{\mathbf{X}}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top K_{\mathbf{X}}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right). \quad (10)$$

Exemplo 3.3 (Mudança de variável). Seja X uma v.a. de pdf f_X e $Y := g(X)$ uma v.a. definida pela função bijetiva e diferenciável $g : \mathbb{R} \rightarrow \mathbb{R}$. A pdf f_Y de Y é dada por

$$f_Y(y) = \frac{f_X(x)}{|g'(x)|}, \quad (11)$$

com $x := g^{-1}(y)$. No caso geral, de vetores aleatórios \mathbf{X} e $\mathbf{Y} := g(\mathbf{X})$, com $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bijetiva e diferenciável, temos

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}(\mathbf{x})}{|\det J(\mathbf{x})|}, \quad (12)$$

com $J(\mathbf{x}) = \left[\frac{\partial g_i}{\partial x_j} \right]_{(i,j)}$ matriz jacobiana.

B. Processos estocásticos

Definição 3.2. Um *processo estocástico* é uma coleção de variáveis aleatórias sobre o espaço de probabilidade (Ω, \mathcal{F}, P) , indexadas por um conjunto de instantes T . Se $T = \mathbb{Z}$, o processo é dito a *tempo discreto*, como $\{X[i] : i \in \mathbb{Z}\}$; se $T = \mathbb{R}$, o processo é dito a *tempo contínuo*, como $\{X(t) : t \in \mathbb{R}\}$.

A *distribuição temporal* (de ordem k) de um processo estocástico $X_t = X(t)$ é dada por

$$F_{X_{t_1}, \dots, X_{t_k}}(x_1, \dots, x_k) := P(X_{t_1} \leq x_1, \dots, X_{t_k} \leq x_k). \quad (13)$$

Para um processo $X(t)$, definimos a *média*

$$m_X(t) := \mathbb{E}[X_t], \quad (14)$$

a *(auto-)correlação*

$$R_X(s, t) := \mathbb{E}[X_s X_t^*], \quad (15)$$

e a *(auto-)covariância*

$$\begin{aligned} K_X(s, t) &:= \mathbb{E}[(X_s - \mathbb{E}[X_s])(X_t - \mathbb{E}[X_t])^*] \\ &= R_X(s, t) - m_X(s)m_X^*(t). \end{aligned} \quad (16)$$

Um processo estocástico é dito *estacionário* em sentido lato (*wide-sense stationary – WSS*) se sua média $m_X(t)$ é constante e sua auto-correlação $R_X(s, t)$ só depende de $\tau := s - t$; escrevemos $R_X(\tau) := R_X(t + \tau, t)$.

A *densidade espectral de potência* (PSD) $S_X(f)$ de um processo $X(t)$, de média nula e WSS, é a transformada de Fourier da sua auto-covariância $K_X(\tau)$, i.e.,

$$S_X(f) = \int_{-\infty}^{\infty} K_X(\tau) e^{-j2\pi f\tau} d\tau. \quad (17)$$

A potência do sinal confinada em uma banda $[f_1, f_2]$ pode ser determinada integrando sua PSD $S_X(f)$.

C. Variáveis aleatórias complexas

Definição 3.3. Uma *v.a. complexa* Z sobre (Ω, \mathcal{F}, P) é um objeto da forma $Z = X + jY$, em que $X = \Re\{Z\}$ e $Y = \Im\{Z\}$ são v.a.s reais sobre (Ω, \mathcal{F}, P) .

A distribuição de Z é dada pela distribuição conjunta de (X, Y) . Um vetor aleatório complexo n -dimensional $\mathbf{Z} = (Z_1, \dots, Z_n)$ tem distribuição especificada pelo vetor aleatório real $2n$ -dimensional $(X_1, \dots, X_n, Y_1, \dots, Y_n)$, com $Z_i = X_i + jY_i$, $1 \leq i \leq n$.

Exemplo 3.4 (Caso gaussiano). Um vetor aleatório $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$ é gaussiano se \mathbf{X} e \mathbf{Y} são vetores aleatórios conjuntamente gaussianos.

Exemplo 3.5 (Caso gaussiano circular). Um caso de especial interesse ocorre quando o vetor aleatório $\mathbf{Z} \in \mathbb{C}^n$ tem *distribuição gaussiana com simetria circular*² $\mathbf{Z} \sim \mathcal{CN}(\mathbf{0}, K_{\mathbf{Z}})$ de média nula e matriz de covariância (não-singular) $K_{\mathbf{Z}}$. Nesse caso, a pdf é dada por

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{\pi^n \det(K_{\mathbf{Z}})} \exp(-\mathbf{z}^\dagger K_{\mathbf{Z}}^{-1} \mathbf{z}). \quad (18)$$

Se $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$ for tal que \mathbf{X} e \mathbf{Y} são vetores reais gaussianos independentes, cada um com componentes i.i.d. $X_i, Y_i \sim \mathcal{N}(0, \frac{1}{2})$, então a pdf de \mathbf{Z} será

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{\pi^n} e^{-\|\mathbf{z}\|^2}. \quad (19)$$

REFERÊNCIAS

- [1] B. Rimoldi, *Principles of Digital Communication: A Top-Down Approach*. Cambridge University Press, 2016.

²Isto é, para todo $\theta \in [0, 2\pi[$, a distribuição de \mathbf{Z} é a mesma de $\mathbf{Z}e^{j\theta}$.