Instituto Superior Técnico
Departamento de Matemática
Secção de Álgebra e Análise

Última actualização: 18/Nov/2003

ÁLGEBRA LINEAR A

REVISÃO DA PARTE III

Parte III - (a) Ortogonalidade

Conceitos:
produto interno, ortogonalidade, norma, vectores ortonormais (o.n.), ângulo, complemento ortogonal, **projecção ortogonal**, decomposição ortogonal, soma directa, espaço euclideano real, espaço euclideano complexo, **transformação ortogonal**, matriz ortogonal, matriz transposta, matriz simética, matriz anti-simétrica

O **produto interno usual** de dois vectores $v = (x_1, \ldots, x_n)$ e $w = (y_1, \ldots, y_n)$ em $\mathbb{R}^n$ é o número real

$$v \cdot w = x_1y_1 + x_2y_2 + \ldots + x_ny_n = v^tw \ .$$

v e w dizem-se **ortogonais** (ou perpendiculares), e representa-se $v \perp w$, se $v \cdot w = 0$.
A **norma** (ou comprimento) de v é o número não negativo

$$|v| = \sqrt{v \cdot v} = \sqrt{x_1^2 + \ldots + x_n^2} \ .$$

Um vector $v \in \mathbb{R}^n$ diz-se **unitário** se $|v| = 1$.
Vectores $v_1, \ldots, v_m \in \mathbb{R}^n$ dizem-se **ortonormais (o.n.)** quando

$$v_i \cdot v_j = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

ou seja, quando são ortogonais entre si e unitários.

Propriedades do produto interno (caracterizam um produto interno, não necessariamente o usual em $\mathbb{R}^n$): para quaisquer $u, v, w \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, tem-se

- $v \cdot w = w \cdot v$ (*simetria*)
- $(\lambda u + v) \cdot w = \lambda(u \cdot w) + v \cdot w$ (*linearidade*)
- $v \cdot v > 0 \ , \quad \forall v \neq 0$ (*positividade*)

Factos sobre a **ortogonalidade** de vectores e espaços (para o produto interno usual em $\mathbb{R}^n$ ou qualquer outro espaço vectorial V com um produto interno):

- Vectores o.n. são linearmente independentes.
- Dado um qualquer subespaço vectorial W de $\mathbb{R}^n$, o seu **complemento ortogonal**
$$W^\perp = \{v \in \mathbb{R}^n : v \cdot w = 0 \ \forall w \in W\}$$
é um subespaço vectorial de $\mathbb{R}^n$. Tem-se $(W^\perp)^\perp = W$ e $W \cap W^\perp = \{0\}$.
- Se W é um subespaço de $\mathbb{R}^n$ e se v é um qualquer vector de $\mathbb{R}^n$, então existe um *único* vector $w \in W$ com a propriedade da diferença $v-w$ pertencer ao complemento ortogonal $W^\perp$:
$$v = \underbrace{w}_{\in W} + \underbrace{v-w}_{\in W^\perp} .$$
A este vector w chama-se a **projecção ortogonal de** v **em** W e representa-se $w = \mathrm{proj}_W v$.
- Relativamente a uma base o.n. $w_1, \ldots, w_m$ de W, o vector projecção ortogonal de v em W é dado pela fórmula
$$\boxed{\mathrm{proj}_W v = (v \cdot w_1)w_1 + \ldots + (v \cdot w_m)w_m} .$$
- A **projecção ortogonal em** W é a transformação linear definida por
$$\mathrm{proj}_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n , \quad v \mapsto \mathrm{proj}_W v ,$$
cuja imagem é $\mathrm{Im}\ \mathrm{proj}_W = W$ e cujo núcleo é $\mathrm{Ker}\ \mathrm{proj}_W = W^\perp$.
- Um subespaço vectorial W de V e o seu complemento ortogonal $W^\perp$ formam uma **decomposição ortogonal** de V:
$$V = W \oplus W^\perp$$
i.e., cada vector $v \in V$ pode ser escrito de uma forma *única* como a soma de vectores $w = \mathrm{proj}_W v \in W$ e $v - w \in W^\perp$.
Portanto, tem-se a fórmula para a dimensão: $\boxed{\dim V = \dim W + \dim W^\perp}$.

Diz-se que V é a **soma directa** dos seus subespaços vectoriais W_1 e W_2, e escreve-se $V = W_1 \oplus W_2$, quando $V = W_1 + W_2$ (i.e., estes subespaços juntos geram V) e $W_1 \cap W_2 = \emptyset$ (i.e., a intersecção destes subespaços é vazia.

Teorema de Pitágoras: $\boxed{|v+w|^2 = |v|^2 + |w|^2 \iff v \perp w}$.
Consequências: Se W é um subespaço de $\mathbb{R}^n$,

- $|\mathrm{proj}_W v| \leq |v|$ e
- $|\mathrm{proj}_W v| = |v|$ sse $v \in W$.

Desigualdade de Cauchy-Schwarz: $\boxed{|v \cdot w| \leq |v| \cdot |w|}$ e
tem-se a igualdade $|v \cdot w| = |v| \cdot |w|$ sse v e w são proporcionais (ou colineares).
Consequência: Define-se o **ângulo** θ entre dois vectores v e w não nulos como sendo

$$\theta = \arccos \frac{v \cdot w}{|v||w|} \in [0, \pi] \ .$$

Assim se obtém a fórmula $v \cdot w = |v||w| \cos\theta$.

Desigualdade Triangular: $\boxed{|v + w| \leq |v| + |w|}$ e
tem-se a igualdade $|v + w| = |v| + |w|$ sse v e w são proporcionais e têm o mesmo sentido.

Ortogonalização de Gram-Schmidt: processo iterativo para obter uma base o.n. $w_1, \ldots, w_m$ de um espaço vectorial V com produto interno a partir de uma base qualquer $v_1, \ldots, v_m$ de V.

(1) $w_1 = \frac{v_1}{|v_1|}$
(2) $w_2 = \frac{v_2 - (v_2 \cdot w_1) w_1}{|v_2 - (v_2 \cdot w_1) w_1|}$
(3) $w_3 = \frac{v_3 - (v_3 \cdot w_1) w_1 - (v_3 \cdot w_2) w_2}{|v_3 - (v_3 \cdot w_1) w_1 - (v_3 \cdot w_2) w_2|}$
(4) etc.

Porque é que as **bases ortonormais são importantes**?

- Uma base o.n. simplifica os cálculos devido aos produtos internos nulos $w_i \cdot w_j = 0$. Em particular, a fórmula para uma projecção é especialmente simples quando se tem uma base o.n.
- Em Sistemas Dinâmicos e em Mecânica Quântica, muitos problemas lidam com matrizes simétricas (ou hermiteanas, no caso complexo) em que existem bases o.n. naturais (formadas por vectores próprios).
- O processo de Gram-Schmidt pode ser usado para definir classes de polinómios importantes em aplicações, tais como os polinómios de Legendre, os polinómios de Chebyshev e os polinómios de Hermite.
- O processo de Gram-Schmidt conduz a factorização de matrizes permitindo resolução numérica de sistemas muito grandes de equações lineares, como aqueles que ocorrem em Astrofísica.
- Etc.

A **matriz transposta** de uma matriz $m \times n$ A é a matriz $n \times m$ A^t cuja linha j é a coluna j de A. Propriedades da transposta:

- $(A^t)^t = A$
- $(A+B)^t = A^t + B^t$
- $(AB)^t = B^t A^t$
- Quando A é invertível, A^t também é e $(A^t)^{-1} = (A^{-1})^t$.
- $(\mathrm{Im}\ A)^\perp = \mathrm{Ker}\ A^t$

Num subespaço V de $\mathbb{R}^n$ (ou num qualquer espaço vectorial V com produto interno), uma transformação linear $T : V \to V$ diz-se uma **transformação ortogonal** quando preserva a norma, i.e., $|T(v)| = |v|$, $\forall v \in V$, ou equivalentemente quando preserva o produto interno, i.e., $T(v) \cdot T(w) = v \cdot w$, $\forall v, w \in V$. Em particular, transformações ortogonais preservam ortogonalidade, i.e., se T é ortogonal e $v \perp w$, então $T(v) \perp T(w)$.

Quando uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é representada por uma matriz quadrada A relativamente à base canónica, tem-se:

$$\begin{aligned} A \text{ é } \textbf{matriz ortogonal} \quad &\iff \quad T \text{ é ortogonal} \\ &\iff \quad T(e_1), \ldots, T(e_n) \text{ é base o.n. de } \mathbb{R}^n \\ &\iff \quad \text{as colunas de } A \text{ são o.n.} \\ &\iff \quad \boxed{A^t A = \mathrm{Id}} \end{aligned}$$

ou seja, A é matriz ortogonal sse é invertível com inversa igual à sua matriz transposta.

Para um subespaço W de $\mathbb{R}^n$ com base o.n. $w_1, \ldots, w_m$, a **projecção ortogonal** $\mathrm{proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ é representada (relativamente à base canónica de $\mathbb{R}^n$) pela matriz quadrada $\boxed{AA^t}$ onde A é a matriz $n \times m$ com colunas dadas por $w_1, \ldots, w_m$.

Porque é que as **transformações ortogonais são importantes**?

- Em Mecânica Clássica, as leis são invariantes pelas ditas transformações de Galileu – composições de translações com transformações ortogonais.
- Em Mecânica Quântica, as evoluções de um sistema, quando escritas como matrizes reais, são transformações ortogonais.
- As chamadas transformadas de Fourier são transformações ortogonais. Em aplicações essas transformadas são muito úteis em computação gráfica (por exemplo, ficheiros JPG) e em compressão de som (por exemplo, ficheiros MP3).
- Muitas mudanças de coordenadas são transformações ortogonais; em particular, essas mudanças preservam volume (ver "Determinantes").
- Etc.

Parte III - (b) Determinantes

Conceitos:
determinante, permutação, menor, cofactor, matriz dos cofactores, volume,
valor próprio, **vector próprio**, polinómio característico, traço

O **determinante** de uma matriz $n \times n$ é o número

$$\det A = \sum_{\sigma \in P_n} (-1)^\sigma a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

onde P_n é o conjunto de todas as permutações de n elementos e $(-1)^\sigma = \pm 1$ é o *sinal* da permutação. O determinante de A também se pode representar por $|A|$.

Caso 2×2

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

A primeira parcela corresponde à permutação $(\sigma(1), \sigma(2)) = (1, 2)$ (que não troca nenhum índice) e a segunda parcela corresponde à permutação $(\sigma(1), \sigma(2)) = (2, 1)$ (que troca os dois índices).

Caso 3×3

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} \end{aligned}$$

Propriedades dos determinantes

- $$\det \begin{bmatrix} | & & | & & | \\ v_1 & \ldots & (v+w) & \ldots & v_n \\ | & & | & & | \end{bmatrix} = \det \begin{bmatrix} | & & | & & | \\ v_1 & \ldots & v & \ldots & v_n \\ | & & | & & | \end{bmatrix} + \det \begin{bmatrix} | & & | & & | \\ v_1 & \ldots & w & \ldots & v_n \\ | & & | & & | \end{bmatrix}$$
- $$\det \begin{bmatrix} | & & | & & | \\ v_1 & \ldots & (kv) & \ldots & v_n \\ | & & | & & | \end{bmatrix} = k \det \begin{bmatrix} | & & | & & | \\ v_1 & \ldots & v & \ldots & v_n \\ | & & | & & | \end{bmatrix} \quad (k \in \mathbb{R})$$

Mais propriedades dos determinantes

- $\det(kA) = k^n \det A$.
- Matrizes transpostas têm o mesmo determinante: $\det(A^t) = \det A$.
- Se B se obtém de A trocando duas linhas, então $\det B = -\det A$.
- Se duas linhas de A são iguais, então $\det A = 0$.
- Se B se obtém de A adicionando uma linha a outra, então $\det B = \det A$.
- O determinante de uma matriz triangular (superior ou inferior) é o produto das entradas da diagonal.

Uma matriz A é invertível sse o seu determinante não é zero:

$$\boxed{A^{-1} \text{ existe} \quad \Longleftrightarrow \quad \det A \neq 0}$$

Porquê?
Aplique-se o método de Gauss a A:

$$[A] \longrightarrow \dots \longrightarrow [\text{forma escalonada de } A] \ ,$$

fazendo

- um número total N de trocas de linhas,
- multiplicações de linhas por escalares não-nulos $k_1, \dots, k_p$ e
- substituições de linhas por essas linhas somadas com outras.

Então, pelas propriedades anteriores dos determinantes,

$$\det A = (-1)^N k_1 \cdots k_p \det(\text{forma escalonada de } A) \ .$$

Logo,

$$\begin{aligned} A \text{ é invertível} \quad &\Longleftrightarrow \quad \text{a forma escalonada de } A \text{ é a identidade} \\ &\Longleftrightarrow \quad \det(\text{forma escalonada de } A) \neq 0 \\ &\Longleftrightarrow \quad \det A \neq 0 \ . \end{aligned}$$

O determinante de um produto de matrizes é o produto dos determinantes:

$$\boxed{\det(AB) = \det A \det B}$$

Porquê?
Se A não é invertível, então AB não é invertível e a fórmula reduz-se a $0 = 0$.
Se A é invertível, aplique-se o método de Gauss a A e AB justapostas:

$$[A|AB] \longrightarrow \dots \longrightarrow [\mathrm{Id}|B] \ ,$$

fazendo

- um número total N de trocas de linhas,
- multiplicações de linhas por escalares não-nulos $k_1, \dots, k_p$ e
- substituições de linhas por essas linhas somadas com outras.

Então, pelas propriedades anteriores (argumento semelhante ao acima),

$$\det A = (-1)^N k_1 \cdots k_p \det \mathrm{Id} \quad \text{e} \quad \det AB = (-1)^N k_1 \cdots k_p \det B$$

donde se conclui (por ser $\det \mathrm{Id} = 1$) que $\det AB = \det A \det B$.

Mais propriedades dos determinantes

- $\det(A^k) = (\det A)^k$ $(k = 1, 2, 3, \ldots)$.
- Quando A é invertível, $\det\left(A^{-1}\right) = (\det A)^{-1}$.
- Matrizes ortogonais têm determinante ± 1.

Cálculo de determinantes pela fórmula de Laplace

Para uma matriz $n \times n$ A, o **menor** ij de A é a matriz $(n-1) \times (n-1)$ A_{ij} que se obtém de A eliminando a linha i e a coluna j.

O determinante de uma matriz $A = (a_{ij})$ pode ser calculado em termos dos determinantes dos menores de A ao longo de uma qualquer linha ou de uma qualquer coluna:

- Fórmula de Laplace ao longo da linha i

$$\boxed{\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}}$$

- Fórmula de Laplace ao longo da coluna j

$$\boxed{\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}}$$

Por exemplo, para uma matriz 3×3 a fórmula ao longo da primeira linha dá

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} .$$

Resolução de sistemas pela regra de Cramer

Dada uma matriz $n \times n$ invertível A e um vector $b \in \mathbb{R}^n$, o sistema de equações lineares $\boxed{Av = b}$ tem uma e uma só solução dada por

$$v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{com} \qquad \boxed{x_i = \frac{\det A_i}{\det A}}$$

onde A_i é a matriz obtida de A substituindo a coluna i pelo vector b.

Inversão de matrizes pela regra de Cramer

Se A é uma matriz invertível, então a sua inversa é dada por

$$A^{-1} = \begin{bmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \ldots & c_{nn} \end{bmatrix} \qquad \text{com} \qquad c_{ij} = \frac{(-1)^{i+j} \det A_{ji}}{\det A}$$

onde A_{ji} é o menor ji de A. Alternativamente,

$$\boxed{A^{-1} = \frac{1}{\det A} (\operatorname{cof} A)^t}$$

onde $\operatorname{cof} A$ é a matriz $n \times n$ cuja entrada ij é dada por $(-1)^{i+j} \det A_{ij}$, chamada a **matriz dos cofactores** de A.

Para dois vectores v_1 e v_2 em $\mathbb{R}^2$,

$$\det \begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix} = \pm \textbf{ área do paralelogramo } \text{definido por } v_1 \text{ e } v_2 \ .$$

Para três vectores v_1, v_2 e v_3 em $\mathbb{R}^3$,

$$\det \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \pm \textbf{ volume do paralelepípedo } \text{definido por } v_1,\ v_2 \text{ e } v_3 \ .$$

Para n vectores $v_1, \ldots, v_n$ em $\mathbb{R}^n$,

$$\det \begin{bmatrix} | & & | \\ v_1 & \ldots & v_n \\ | & & | \end{bmatrix} = \pm \textbf{ volume do paralelepípedo } \text{em } \mathbb{R}^n \text{ definido por } v_1, \ldots v_n \ .$$

Ao **calcular um determinante**:

- Será que a matriz é de um tipo especial, 2×2, 3×3, diagonal, triangular, ortogonal?
- Será que tem colunas ou linhas repetidas ou nulas? Ou que é uma potência?
- Será que há alguma coluna ou linha especialmente adequada à fórmula de Laplace?
- Será que argumentos geométricos ou eliminação de Gauss dão a resposta?

Dada uma matriz $n \times n$ A e um vector não-nulo $v \in \mathbb{R}^n$, diz-se que v é um **vector próprio** de A quando Av é múltiplo escalar de v, i.e., quando $Av = \lambda v$ para algum $\lambda \in \mathbb{R}$. Nesse caso, λ diz-se o **valor próprio** de A associado ao vector próprio v.

$$\begin{aligned} \lambda \text{ é valor próprio de } A \quad &\Longleftrightarrow \quad \mathrm{Ker}\ (A - \lambda \mathrm{Id}) \neq \{0\} \\ &\Longleftrightarrow \quad A - \lambda \mathrm{Id} \text{ não é invertível} \\ &\Longleftrightarrow \quad \boxed{\det(A - \lambda \mathrm{Id}) = 0}\ . \end{aligned}$$

Ao polinómio $p_A(\lambda) = \det(A - \lambda \mathrm{Id})$ chama-se o **polinómio característico** da matriz A. Portanto, um número λ é valor próprio de A sse λ é raiz do polinómio característico de A.

$$\begin{aligned} v \neq 0 \text{ é vector próprio de } A \text{ associado ao valor próprio } \lambda \quad &\Longleftrightarrow \quad (A - \lambda \mathrm{Id})v = 0 \\ &\Longleftrightarrow \quad \boxed{v \in \mathrm{Ker}\ (A - \lambda \mathrm{Id})}\ . \end{aligned}$$

De acordo com a definição de determinante, o polinómio característico de A é um polinómio de grau n em λ com

$$p_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} (\mathrm{tr} A) \lambda^{n-1} + \ldots + (\det A)$$

onde $\mathrm{tr}A$, o **traço** de A, é a soma das entradas da diagonal de A.