MCTA017 – Programação Matemática

EPR202 – Métodos de Otimização Aplicados à Eng. de Produção

Aula 3

Revisão de Álgebra Linear

(Parte I)

Prof. Dr. Aritanan Gruber

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC

Quadrimestre Suplementar 2020

Álgebra Linear

Por um lado, uma extensão do conceito e do ferramental de vetores em $\mathbb{R}^2$ e $\mathbb{R}^3$, os *espaços euclidianos*, para espaços de dimensões maiores (finitas e infinita).

Por outro, um arcabouço abstrato para o tratamento unificado de objetos que possuem uma estrutura linear.

Enfoque em aspectos algébricos, frequentemente com interpretações geométricas.

Álgebra Linear

Neste curso, não precisaremos da teoria de álgebra linear em sua totalidade; nem em sua maior generalidade. Focamos em espaços vetoriais de dimensões finitas sobre $\mathbb{R}$ (às vezes, sobre $\mathbb{Q}$), subespaços lineares e afins, (in)dependência, bases, projeções, transformações e dualidade.

Nota: Este não é um curso de álgebra linear; então assumimos em nossa revisão que você já possui familiaridade com o assunto.

A referida teoria será, depois, estendida com a ajuda de tópicos de análise convexa.

Outline

Espaços e Subespaços Vetoriais Lineares

Combinações Lineares, Independência, Bases

Sistemas Homogêneos e Subespaços de Matrizes

Transformações Lineares

Espaços com Produto Interno

Espaços Vetoriais Lineares

Considere

- um conjunto $V \neq \emptyset$, cujos elementos chamados *vetores*, e
- um corpo $\mathbb{F}$ (como $\mathbb{Q}, \mathbb{R}$), cujos elementos são ditos *escalares*,

e dotados de

(a) uma operação de *adição vetorial*

$$\boldsymbol{x} + \boldsymbol{y} \in V \qquad \text{para todos} \qquad \boldsymbol{x}, \boldsymbol{y} \in V$$

com as propriedades de

- associatividade: $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$, $\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$;
- comutatividade: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$, $\forall\, \boldsymbol{x}, \boldsymbol{y} \in V$;
- identidade aditiva: existe $\mathbf{0} \in V$ tal que $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$, $\forall\, \boldsymbol{x} \in V$;
- inversos aditivos: $\forall\, \boldsymbol{x} \in V$, $\exists\, \boldsymbol{x}' \in V$ tal que $\boldsymbol{x} + \boldsymbol{x}' = \mathbf{0}$;

(b) uma operação de *multiplicação escalar*

$$\lambda \cdot \boldsymbol{x} = \lambda \boldsymbol{x} \in V \qquad \text{para todos} \qquad \lambda \in \mathbb{F},\ \boldsymbol{x} \in V$$

com as propriedades de

- associatividade: $\lambda \cdot (\mu \cdot \boldsymbol{x}) = (\lambda\mu) \cdot \boldsymbol{x}$, $\forall\, \lambda, \mu \in \mathbb{F},\ \boldsymbol{x} \in V$;
- identidade multiplicativa: existe $1 \in \mathbb{F}$ t.q. $1 \cdot \boldsymbol{x} = \boldsymbol{x}$, $\forall\, \boldsymbol{x} \in V$;

Espaços Vetoriais Lineares

e tal que as operações de adição e multiplicação são distributivas:

$$\lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y} \qquad \text{e} \qquad (\lambda + \mu) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \mu \cdot \boldsymbol{x},$$

para todos $\lambda, \mu \in \mathbb{F}$ e $\boldsymbol{x}, \boldsymbol{y} \in V$.

Uma tal estrutura $(V, \mathbb{F}, +, \cdot, \mathbf{0}, 1)$ é *espaço vetorial (linear)*.

Nota: É comum a expressão "V é um espaço vetorial sobre $\mathbb{F}$."

Proposição (de prova simples)

Para um espaço vetorial V sobre $\mathbb{F}$,

(a) *a identidade de adição, $\mathbf{0}$, é única em V;*

(b) *cada elemento $\boldsymbol{x} \in V$ possui um único inverso aditivo $-\boldsymbol{x}$;*

(c) $0 \cdot \boldsymbol{x} = \mathbf{0}$ *e* $\lambda \cdot \mathbf{0} = \mathbf{0}$, $\forall \lambda \in \mathbb{F},\ \boldsymbol{x} \in V$;

(d) $(-1) \cdot \boldsymbol{x} = -\boldsymbol{x}$ *para todo* $\boldsymbol{x} \in V$, *com* $-1 \in \mathbb{F}$.

Espaços Vetoriais Lineares

Exemplos

Espaços vetoriais sobre $\mathbb{R}$ importantes neste curso: $\mathbb{R}^n$ e $\mathbb{R}^{m\times n}$.

O primeiro, $\mathbb{R}^n$, diz respeito aos vetores com n coordenadas, cada qual, um valor real: $\mathbb{R}^n := \big\{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}\big\}$. Adição e multiplicação são definidas por componentes (coordenadas):

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

e $\lambda \cdot (x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$.

Apesar de termos escrito os vetores em linhas (às vezes, usaremos o símbolo de transposição $\boldsymbol{x}^\top$), eles, na verdade, são colunas:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Espaços Vetoriais Lineares

Exemplos

O segundo, $\mathcal{M}(m,n) := \mathbb{R}^{m\times n}$, é o espaço das matrizes com m linhas e n colunas com entradas reais. Adição e multiplicação são definidas por componentes (coordenadas):

$$\boldsymbol{A}+\boldsymbol{B} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \cdots & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{1,2} & \cdots & a_{2,n}+b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}+b_{m,1} & a_{m,2}+b_{m,2} & \cdots & a_{m,n}+b_{m,n} \end{pmatrix}.$$

Espaços Vetoriais Lineares

Exemplos

$$\lambda \cdot \boldsymbol{A} = \lambda \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \cdots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \cdots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \cdots & \lambda a_{m,n} \end{pmatrix}.$$

Claramente, cada linha de $\boldsymbol{A}$ é um vetor em $\mathbb{R}^n$ transposto, e cada coluna, um vetor em $\mathbb{R}^m$.

Vamos denotar a linha i de $\boldsymbol{A}$ por $\boldsymbol{A}_{i,*}$ ou por $\boldsymbol{a}_i^\top$, e a coluna j por $\boldsymbol{A}_{*,j}$ ou por $\boldsymbol{A}_j$.

Exercício

Prove que, de fato, $\mathbb{R}^n$ e $\mathbb{R}^{m\times n}$ são espaços vetoriais sobre $\mathbb{R}$.

Subespaços Vetoriais

Para um espaço vetorial V sobre $\mathbb{F}$, um subconjunto $\emptyset \neq S \subseteq V$ é um *subespaço vetorial (linear)* de V se, para todos $\boldsymbol{x}, \boldsymbol{y} \in S$ e $\lambda, \mu \in \mathbb{F}$,

$$\lambda \boldsymbol{x} + \mu \boldsymbol{y} \in S,$$

com as operações de adição e multiplicação como em V. Isto é, S é, por sí mesmo, um espaço vetorial sobre $\mathbb{F}$.

Exemplos: (verifique as afirmações!)

- $\{(0, x_2) : x_2 \in \mathbb{R}\}$ é um subespaço de $\mathbb{R}^2$.
- Qualquer reta em $\mathbb{R}^2$ que passe pela origem $(0, 0)$.
- $\{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$ é um subespaço de $\mathbb{R}^3$.
- $\left\{ \begin{pmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix} : a_{1,1}, a_{1,2}, a_{2,2} \in \mathbb{R} \right\}$, o conjunto das matrizes 2×2 *triangulares superiores* é um supespaço de $\mathbb{R}^{2\times 2}$.
- $\mathbb{S}^n := \{\boldsymbol{A} \in \mathbb{R}^{n\times n} : \boldsymbol{A} = \boldsymbol{A}^\top\}$, o conjunto das matrizes quadradas ($n \times n$) e *simétricas* é um subespaço de $\mathbb{R}^{n\times n}$.

Operações em Espaços Vetoriais

Sejam V_1, V_2 espaços vetoriais sobre $\mathbb{F}$.

Proposição

A intersecção de V_1 e V_2 é um espaço vetorial sobre $\mathbb{F}$.

Prova. Sejam $\boldsymbol{x}, \boldsymbol{y} \in X = V_1 \cap V_2$ e $\lambda, \mu \in \mathbb{F}$. Como $\boldsymbol{x}, \boldsymbol{y} \in V_i$ e V_i é um espaço vetorial, $\lambda\boldsymbol{x}, \mu\boldsymbol{y} \in V_i$ e, assim, $\lambda\boldsymbol{x} + \mu\boldsymbol{y} \in V_i$. Logo, $\lambda\boldsymbol{x} + \mu\boldsymbol{y} \in X$ completando a prova. $\square$

A *soma* de V_1 e V_2 é dada por

$$V_1 + V_2 := \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x} \in V_1,\ \boldsymbol{y} \in V_2\}.$$

A soma é *direta*, denotada por $V_1 \oplus V_2$, se $V_1 \cap V_2 = \{\mathbf{0}\}$.

Proposição (exercício)

$V_1 + V_2$ e $V_1 \oplus V_2$ são espaços vetoriais sobre $\mathbb{F}$.

Operações em Espaços Vetoriais

Exemplo:

$$\{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\} \oplus \{(0, 0, x_3) : x_3 \in \mathbb{R}\}$$
$$= \{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \mathbb{R}\} = \mathbb{R}^3.$$

A união de espaços vetoriais, em geral, não é um espaço vetorial. Por exemplo,

$$X := \{(x, x) : x \in \mathbb{R}\} \cup \{(x, -x) : x \in \mathbb{R}\}$$

não é um espaço vetorial, pois $(z, z) + (z, -z) = (2z, 0) \notin X$ para $z \neq 0$. Claramente, $\{(x, x) : x \in \mathbb{R}\}$ e $\{(x, -x) : x \in \mathbb{R}\}$ são subespaços de $\mathbb{R}^2$.

Nota: $\{(x, x) : x \in \mathbb{R}\} \oplus \{(x, -x) : x \in \mathbb{R}\} = \mathbb{R}^2$.

Outline

Combinações e Fechos Lineares

(V um espaço vetorial sobre $\mathbb{F}$)

Um vetor $\boldsymbol{y} \in V$ é uma *combinação linear* de vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in V$ se existem escalares $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$ tais que

$$\boldsymbol{y} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_n \boldsymbol{x}_n.$$

Exemplos:

- $(3, 2, 1) = 3(1, 0, 0) + 2(0, 1, 0) + (0, 0, 1)$.
- $(3, 2, 1) = 2(2, 1, 3) - (1, 0, 5)$.
- $\begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} = 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 5 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Fechos e Geradores Lineares

(V um espaço vetorial sobre $\mathbb{F}$)

Para $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\} \subseteq V$, não necessariamente um subespaço, o *fecho linear* (ou *linear span*) de X é

$$\operatorname{lin.span}(X) := \{\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_n \boldsymbol{x}_n : \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}\},$$

o conjunto de todas as combinações lineares dos elementos em X.

Nota: $\operatorname{lin.span}(\emptyset) = \{\mathbf{0}\}$.

Proposição

Para todo $X \subseteq V$, $\operatorname{lin.span}(X)$ é um subespaço de V.

X é um *gerador* de/para V (ou X *gera* V) se $\operatorname{lin.span}(X) = V$.

Fechos e Geradores Lineares

Exemplo: Para $i \in [n] := \{1, 2, \ldots, n\}$, defina $\boldsymbol{e}_i \in \mathbb{R}^n$ como o vetor que possui a coordenada i igual a 1 e as demais iguais a 0:

$$\boldsymbol{e}_i := (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i+1}).$$

O conjunto $E = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ é um gerador de $\mathbb{R}^n$.

De fato, para qualquer $\boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$, temos que $\boldsymbol{z} \in \operatorname{lin.span}(E)$, pois

$$\boldsymbol{z} = z_1 \cdot \boldsymbol{e}_1 + z_2 \cdot \boldsymbol{e}_2 + \cdots + z_n \cdot \boldsymbol{e}_n.$$

A inclusão $\operatorname{lin.span}(E) \subseteq \mathbb{R}^n$ é imediata. $\square$

Bases

Seja V um espaço vetorial e $X^* \subseteq V$ um conjunto gerador *minimal* de V, isto é,

- $V = \operatorname{lin.span}(X^*)$ e
- para todo $X \subseteq X^*$ tal que $\operatorname{lin.span}(X) = V$, $X^* = X$.

X^* é uma *base* de/para V.

Nota: Em geral, um espaço vetorial admite várias bases.

Pergunta: O que podemos dizer sobre os elementos de uma base para V? Isto é, como uma base difere de um conjunto gerador não mínimal?

Dependência e Independência Linear

Os vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in V \setminus \{\mathbf{0}\}$ são *linearmente independentes* se

$$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_n \boldsymbol{x}_n = \mathbf{0}$$

implica que $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Em caso contrário, eles são *linearmente dependentes*.

De outra forma, se $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in V \setminus \{\mathbf{0}\}$ são linearmente dependentes, existe $\lambda_j \neq 0$ tal que

$$\boldsymbol{x}_j = -\frac{\lambda_1}{\lambda_j}\boldsymbol{x}_1 - \frac{\lambda_2}{\lambda_j}\boldsymbol{x}_2 - \cdots - \frac{\lambda_{j-1}}{\lambda_j}\boldsymbol{x}_{j-1} - \frac{\lambda_{j+1}}{\lambda_j}\boldsymbol{x}_{j+1} - \cdots - \frac{\lambda_n}{\lambda_j}\boldsymbol{x}_n,$$

com o lado direito diferente de $\mathbf{0}$.

Logo, $\boldsymbol{x}_j \in \operatorname{lin.span}(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_{j+1}, \ldots, \boldsymbol{x}_n\})$ e $\boldsymbol{x}_j$ é "supérfluo".

Bases e Dimensões

Mostramos com isso que:

Proposição

Um conjunto gerador X é uma base de V se $\operatorname{lin.span}(X) = V$ e os vetores em X são linearmente independentes.

A mesma ideia pode ser usada para mostrar que (verifique!):

Teorema

Todas as bases de V tem o mesmo número de elementos.

O que implica, em espaços vetoriais, que todo conjunto gerador minimal é *minimo* (com respeito à independência linear).

A *dimensão* de V é

$$\dim(V) = |X^*|,$$

com X^* é uma base de V.

V tem dimensão *finita* se $|X^*| < \infty$ e *infinita* em caso contrário.

Bases e Dimensões

Exemplo: Vimos que $E = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ é um conjunto gerador de $\mathbb{R}^n$. Claramente, que os vetores em E são linearmente independentes. Logo, E é uma base para $\mathbb{R}^n$ e com isso, que $\dim(\mathbb{R}^n) = n$.

Sejam $S \subseteq T \subseteq V$, com S e T subespaços de V. Não é difícil ver que

$$\dim(S) \leq \dim(T) \leq \dim(V).$$

Além disso, se $S \subset T$, temos que toda base de S pode ser estendida a uma base de T (prove!).

Seja $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ uma base para V, com os vetores $\boldsymbol{v}_i$ listados em alguma ordem fixada. Qualquer vetor $\boldsymbol{x} \in V$ pode ser representado com relação à $\mathcal{B}$ (sob a ordem fixada) como $\boldsymbol{x} = (\lambda_1, \lambda_2, \ldots, \lambda_n)_{\mathcal{B}}$, em que

$$\boldsymbol{x} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_n \boldsymbol{v}_n.$$

Outline

Sistemas Lineares Homogêneos

Um *sistema linear homogêneo* com m equações e n variáveis reais, $x_1, x_2, \ldots, x_n$, costuma ser representado como

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & 0 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & 0 \end{array}$$

em que $a_{i,j} \in \mathbb{R}$ são ditos *coeficientes*.

Claramente, $x_1 = x_2 = \cdots = x_n = 0$ é uma solução, chamada *trivial*. Logo, todo sistema homogêneo é *consistente*.

O foco torna-se, então: o sistema admite soluções não nulas? em caso afirmativo, o número de soluções não nulas é finito ou infinito? em sendo infinito, é possível descrevê-las de forma finita? e, finalmente, como encontrar tais soluções?

Sistemas Lineares Homogêneos

A forma como foi escrito produz um enfoque natural na organização do sistema por linhas. É conveniente abreviar esta forma para:

$$\begin{array}{rcl} \boldsymbol{a}_1^\top \boldsymbol{x} & = & 0 \\ \boldsymbol{a}_2^\top \boldsymbol{x} & = & 0 \\ \vdots & & \vdots \\ \boldsymbol{a}_m^\top \boldsymbol{x} & = & 0 \end{array}$$

com $\boldsymbol{a}_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,n})$ e $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

Sistemas Lineares Homogêneos

Uma variação consiste em organizar o sistema por colunas:

$$x_1 \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + x_2 \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Fica claro que qualquer solução para o sistema é uma combinação linear dos vetores coluna formados pelos coeficientes.

Finalmente, definindo a matriz $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ como $\boldsymbol{A}_{i,j} = a_{i,j}$ e $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, temos a notação matricial

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}.$$

Claramente, todas as notações são equivalentes. É somente uma questão de usarmos a notação mais apropriada (mais fácil de trabalhar com) em cada circunstância.

Subespaços Associados a Matrizes

Seja $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ uma matriz com m linhas e n colunas.

O *espaço coluna* (ou *range*) de $\boldsymbol{A}$ é o fecho linear das colunas de $\boldsymbol{A}$:

$$\operatorname{range}(\boldsymbol{A}) = \operatorname{lin.span}\left(\{\boldsymbol{A}_{*,j} : j \in [n]\}\right).$$

O *espaço linha* de $\boldsymbol{A}$ é o fecho linear das linhas de $\boldsymbol{A}$:

$$\operatorname{row.span}(\boldsymbol{A}) = \operatorname{lin.span}\left(\{\boldsymbol{A}_{i,*} : i \in [m]\}\right).$$

Proposição

As dimensões de $\operatorname{range}(\boldsymbol{A})$ *e* $\operatorname{row.span}(\boldsymbol{A})$ *são sempre iguais.*

Tal valor é dito o *posto* (ou *rank*) de $\boldsymbol{A}$ e, claramente,

$$\operatorname{rank}(\boldsymbol{A}) \leq \min\{m, n\}.$$

$\boldsymbol{A}$ tem *posto completo* (é *full dimensional*) se $\operatorname{rank}(\boldsymbol{A}) = \min\{m, n\}$.

Subespaços Associados a Matrizes

O conjunto

$$\mathrm{null}(\boldsymbol{A}) := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$$

é chamado *kernel* ou *espaço nulo* (*null space*) de $\boldsymbol{A}$.

$\mathrm{null}(\mathbb{A})$ é um subespaço de $\mathbb{R}^n$ e $\mathrm{range}(\boldsymbol{A})$ é um subespaço de $\mathbb{R}^m$.

Teorema

$\dim(\mathbb{R}^n) = \dim(\mathrm{null}(\boldsymbol{A})) + \dim(\mathrm{range}(\boldsymbol{A}))$.

Prova. Se $\boldsymbol{A} = \boldsymbol{0}$, então $\mathrm{null}(\boldsymbol{A}) = \mathbb{R}^n$, $\mathrm{range}(\boldsymbol{A}) = \{\boldsymbol{0}\}$ e o resultado é válido. Suponha $\boldsymbol{A} \neq \boldsymbol{0}$ e seja $\mathcal{K} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ uma base para $\mathrm{null}(\boldsymbol{A})$.

$\mathcal{K}$ pode ser estendida a uma base para $\mathbb{R}^n$. Sejam $\mathcal{J} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_\ell\}$ os vetores que realizam esta extensão ($\ell = n - k$).

Subespaços Associados a Matrizes

Para $\boldsymbol{x} \in \mathbb{R}^n$, temos então que

$$\boldsymbol{x} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \cdots + \lambda_k \boldsymbol{u}_k + \mu_1 \boldsymbol{v}_1 + \mu_2 \boldsymbol{v}_2 + \cdots + \mu_\ell \boldsymbol{v}_\ell.$$

Aplicando $\boldsymbol{A}$ acima e notando que $\boldsymbol{A}\boldsymbol{u}_i = \boldsymbol{0}$, temos

$$\begin{aligned}\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{A}(\lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \cdots + \lambda_k \boldsymbol{u}_k) + \boldsymbol{A}(\mu_1 \boldsymbol{v}_1 + \mu_2 \boldsymbol{v}_2 + \cdots + \mu_\ell \boldsymbol{v}_\ell) \\ &= \mu_1 \boldsymbol{A}\boldsymbol{v}_1 + \mu_2 \boldsymbol{A}\boldsymbol{v}_2 + \cdots + \mu_\ell \boldsymbol{A}\boldsymbol{v}_\ell.\end{aligned}$$

Isso implica que $\mathcal{L} = \{\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_\ell\}$ é um conjunto gerador para $\mathrm{range}(\boldsymbol{A})$. Resta provarmos que $\mathcal{L}$ é linearmente independente.

Temos que $\phi_1 \boldsymbol{A}\boldsymbol{v}_1 + \phi_2 \boldsymbol{A}\boldsymbol{v}_2 + \cdots + \phi_\ell \boldsymbol{A}\boldsymbol{v}_\ell = \boldsymbol{0}$ se e somente se $\boldsymbol{A}(\phi_1 \boldsymbol{v}_1 + \phi_2 \boldsymbol{v}_2 + \cdots + \phi_\ell \boldsymbol{v}_\ell) = \boldsymbol{0} \iff \phi_1 \boldsymbol{v}_1 + \phi_2 \boldsymbol{v}_2 + \cdots + \phi_\ell \boldsymbol{v}_\ell = \boldsymbol{0}$.

Como $\mathcal{J}$ é linearmente independente, temos que $\phi_j = 0$ e assim, $\mathcal{L}$ é uma base para $\mathrm{range}(\boldsymbol{A})$ de dimensão ℓ. $\square$

Subespaços Associados a Matrizes

Este é um bom momento para resumirmos o que fizemos acima...

Considere um sistema linear homogêneo $\boldsymbol{Ax} = \mathbf{0}$, para $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

- O sistema sempre admite uma solução trivial: $\boldsymbol{x} = \mathbf{0}$.
- A solução trivial é a única solução do sistema se $\mathrm{null}(\boldsymbol{A}) = \{\mathbf{0}\}$; de outra forma, se $\dim(\mathrm{null}(\boldsymbol{A})) = 0$.
- Caso $\dim(\mathrm{null}(\boldsymbol{A})) > 0$, uma base para $\mathrm{null}(\boldsymbol{A})$ pode ser encontrada "dentro" de uma base para $\mathbb{R}^n$.
- Sendo $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ é uma base para $\mathrm{null}(\boldsymbol{A})$, todo vetor

$$\boldsymbol{z} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \cdots + \lambda_k \boldsymbol{u}_k$$

  é uma solução para o sistema, pois $\boldsymbol{Az} = 0$.

Os vetores $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ acima formam um *sistema fundamental de soluções* de $\boldsymbol{Ax} = \mathbf{0}$.

Subespaços Associados a Matrizes

As relações abaixo serão úteis.

Proposição

Seja $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ uma matriz quadrada. São equivalentes:

(a) $\boldsymbol{A}$ *é invertível:* $\exists\, \boldsymbol{A}^{-1} \in \mathbb{R}^{n\times n}$ *tal que* $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$;

(b) $\boldsymbol{A}^\top$ *é invertível;*

(c) $\det(\boldsymbol{A}) := \sum_{\sigma\in\mathfrak{S}_n} \mathit{sgn}(\sigma) \prod_i \boldsymbol{A}_{i,\sigma(i)} \neq 0$; *(determinante de* $\boldsymbol{A}$*)*

(d) *As linhas de* $\boldsymbol{A}$ *são linearmente independentes;*

(e) *As colunas de* $\boldsymbol{A}$ *são linearmente independentes;*

(f) $\mathrm{rank}(A) = n$; *(*$\boldsymbol{A}$ *tem posto completo)*

(g) *O sistema* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ *tem uma única solução;*

(i) $\mathrm{null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$. *(*$\boldsymbol{0}$ *é o único elemento no kernel de* $\boldsymbol{A}$*)*

Nota: num futuro próximo, vamos estender o maquinário desenvolvido até aqui para sistemas afins e sistemas com desigualdades.

Outline

Transformações Lineares

Sejam V e W dois espaços vetoriais sobre $\mathbb{F}$.

Um mapeamento $T : V \to W$ é uma *transformação linear* se

- $T(\boldsymbol{x} + \boldsymbol{y}) = T(\boldsymbol{x}) + T(\boldsymbol{y})$, para $\boldsymbol{x}, \boldsymbol{y} \in V$;
- $T(\lambda \cdot \boldsymbol{x}) = \lambda T(\boldsymbol{x})$, para $\lambda \in \mathbb{F}$ e $\boldsymbol{x} \in V$.

A transformação linear T é dita um *operador linear* se $W = V$; caso $W = \mathbb{F}$, é dita um *funcional linear*.

Quando $V = \mathbb{R}^n$ e $W = \mathbb{R}$, o funcional linear T é referido simplesmente como uma *função linear* real.

Para $\boldsymbol{a} \in \mathbb{R}^n$ e $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ variáveis reais, temos que

$$f(\boldsymbol{x}) = \boldsymbol{a}^\top \boldsymbol{x}$$

é uma função linear; e toda função linear tem a forma acima (requer uma prova).

Transformações Lineares

Se V e W tiverem dimensão finita n e m, a transformação linear $T : V \to W$ pode ser representada por uma matriz com m linhas e n colunas. Vamos ilustrar este resultado para $V = \mathbb{R}^n$ e $W = \mathbb{R}^m$.

Sejam $\mathcal{B}_n = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ e $\mathcal{B}_m = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ bases para $\mathbb{R}^n$ e $\mathbb{R}^m$. Defina $\boldsymbol{u}_i = T(\boldsymbol{v}_i)$, para $\boldsymbol{v}_i \in \mathcal{B}_n$. Temos que $\boldsymbol{u}_i \in \mathbb{R}^m$. Como $\mathcal{B}_m$ é base,

$$\boldsymbol{u}_i = a_{1,i}\boldsymbol{w}_1 + a_{2,i}\boldsymbol{w}_2 + \cdots + a_{m,i}\boldsymbol{w}_m.$$

Para $\boldsymbol{x} \in \mathbb{R}^n$, temos que $\boldsymbol{x} = \lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_n\boldsymbol{v}_n$. Assim,

$$\begin{aligned} T(\boldsymbol{x}) &= T(\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_n\boldsymbol{v}_n) \\ &= \lambda_1 T(\boldsymbol{v}_1) + \lambda_2 T(\boldsymbol{v}_2) + \cdots + \lambda_n T(\boldsymbol{v}_n) \\ &= \lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \cdots + \lambda_n\boldsymbol{u}_n. \end{aligned}$$

Substituindo $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ acima...

Transformações Lineares

obtemos

$$\begin{aligned} T(\boldsymbol{x}) &= \lambda_1(a_{1,1}\boldsymbol{w}_1 + a_{2,1}\boldsymbol{w}_2 + \cdots + a_{m,1}\boldsymbol{w}_m) \\ &+ \lambda_2(a_{1,2}\boldsymbol{w}_1 + a_{2,2}\boldsymbol{w}_2 + \cdots + a_{m,2}\boldsymbol{w}_m) \\ &\vdots \\ &+ \lambda_n(a_{1,n}\boldsymbol{w}_1 + a_{2,n}\boldsymbol{w}_2 + \cdots + a_{m,n}\boldsymbol{w}_m) \\ &= (\lambda_1 a_{1,1} + \lambda_2 a_{1,2} + \cdots + \lambda_n a_{1,n})\boldsymbol{w}_1 \\ &+ (\lambda_1 a_{2,1} + \lambda_2 a_{2,2} + \cdots + \lambda_n a_{2,n})\boldsymbol{w}_2 \\ &\vdots \\ &+ (\lambda_1 a_{m,1} + \lambda_2 a_{m,2} + \cdots + \lambda_n a_{m,n})\boldsymbol{w}_m \\ &= \boldsymbol{\lambda}^\top \boldsymbol{a}_1 \boldsymbol{w}_1 + \boldsymbol{\lambda}^\top \boldsymbol{a}_2 \boldsymbol{w}_2 + \cdots + \boldsymbol{\lambda}^\top \boldsymbol{a}_m \boldsymbol{w}_m, \end{aligned}$$

com $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ e $\boldsymbol{a}_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,m})$.

Transformações Lineares

Considerando a representação de $\boldsymbol{x}$ em $\mathcal{B}_n$,

$$\boldsymbol{x} = (\lambda_1, \lambda_2, \ldots, \lambda_n)_{\mathcal{B}_n},$$

e a representação de $T(\boldsymbol{x})$ em $\mathcal{B}_m$,

$$T(\boldsymbol{x}) = (\boldsymbol{\lambda}^\top \boldsymbol{a}_1, \boldsymbol{\lambda}^\top \boldsymbol{a}_2, \ldots, \boldsymbol{\lambda}^\top \boldsymbol{a}_m)_{\mathcal{B}_m},$$

e definindo a matriz

$$\boldsymbol{A} = \begin{pmatrix} -- & \boldsymbol{a}_1^\top & -- \\ -- & \boldsymbol{a}_2^\top & -- \\ \vdots & \vdots & \vdots \\ -- & \boldsymbol{a}_m^\top & -- \end{pmatrix}$$

vem que

$$T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}.$$

Transformações Lineares

Podemos, então, abusar da notação e usar o mesmo símbolo para denotar uma transformação linear e a matriz que a representa.

Os resultados envolvendo subespaços associados a matrizes são válidos para transformações lineares.

Por exemplo,

Teorema

Para $\boldsymbol{T} : V \to W$, $\dim(V) = \dim(\mathrm{null}(\boldsymbol{T})) + \dim(\mathrm{range}(\boldsymbol{T}))$.

Um resultado interessante que será indiretamente útil neste curso:

Teorema

$\boldsymbol{T}$ *é injetora sse* $\mathrm{null}(\boldsymbol{T}) = \{\mathbf{0}\}$ *e sobrejetora sse* $\mathrm{range}(\boldsymbol{T}) = W$.

Mudanças de Base

Sejam $\mathcal{B}$ e $\mathcal{B}'$ bases distintas de $\mathbb{R}^n$.

Problema

Dado um vetor $\boldsymbol{x} \in \mathbb{R}^n$ representado na base $\mathcal{B}$, encontrar a representação de $\boldsymbol{x}$ na base $\mathcal{B}'$.

O desenvolvimento anterior fornece a resposta: mudanças de base são realizadas por operadores lineares.

Se $B : \mathbb{R}^n \to \mathbb{R}^n$ faz a mudança $\mathcal{B} \mapsto \mathcal{B}'$ e $\boldsymbol{A}_B$ é a matriz associada a B, temos que $\boldsymbol{A}_B \boldsymbol{x}$ é a representação de $\boldsymbol{x}$ em $\mathcal{B}$".

Nota: $\boldsymbol{A}_B$ tem posto completo; logo, é invertível.

Se $C : \mathbb{R}^n \to \mathbb{R}^n$ for outro operador de mudança de base $\mathcal{B}' \mapsto \mathcal{B}''$, com matriz $\boldsymbol{A}_C$ associada, temos que uma mudança $\mathcal{B} \mapsto \mathcal{B}''$ é dada por:

$$C \circ B(\boldsymbol{x}) = C(B(\boldsymbol{x})) = \boldsymbol{A}_C \boldsymbol{A}_B \boldsymbol{x}.$$

Outline

Produto Interno

Um produto interno em um espaço vetorial V sobre $\mathbb{F}$ é uma operação $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ positiva, conjugado-simétrica e bilinear. A conjugação depende do corpo $\mathbb{F}$.

Em $\mathbb{R}$, a conjugação é a operação de transposição. Assim, um *produto interno* em um espaço V sobre $\mathbb{R}$ satisfaz, para todos $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ e $\lambda, \mu \in \mathbb{R}$:

- *positividade*: $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, com igualdade somente se $\boldsymbol{x} = \boldsymbol{0}$;
- *simetria*: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$;
- *bilinearidade*: $\langle \lambda \boldsymbol{x} + \mu \boldsymbol{y}, \boldsymbol{z} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \mu \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.

O produto interno *canônico* em $\mathbb{R}^n$ é dado por:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Normas

Produtos internos induzem normas. Em particular, o produto interno canônico em $\mathbb{R}^n$ induz a *norma euclidiana* ou norma ℓ_2:

$$\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_2 := \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{\frac{1}{2}}, \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^n,$$

e $\|\boldsymbol{x}\|$ é a *magnitude* ou *comprimento* de $\boldsymbol{x}$. Se $\|\boldsymbol{x}\| = 1$, dizemos que $\boldsymbol{x}$ é *unitário*.

Desigualdade triangular: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Teorema de Pitágoras: $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.

Lema (Desigualdade de Cauchy-Schwarz)

Para quaisquer vetores $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, *vale que* $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$.

Desigualdade de Cauchy-Schwarz

Prova. Podemos supor que $\boldsymbol{x} \neq \mathbf{0}$. Defina $\boldsymbol{u} = \lambda \boldsymbol{x} - \boldsymbol{y}$, para algum $\lambda \in \mathbb{R}$. Temos que

$$\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \langle \lambda \boldsymbol{x} - \boldsymbol{y}, \lambda \boldsymbol{x} - \boldsymbol{y} \rangle = \lambda^2 \langle \boldsymbol{x}, \boldsymbol{x} \rangle - 2\lambda \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle \geq 0,$$

em que a desigualdade final advém do fato que $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \|\boldsymbol{u}\|^2 \geq 0$. Tomando $\lambda = \langle \boldsymbol{x}, \boldsymbol{y} \rangle / \langle \boldsymbol{x}, \boldsymbol{x} \rangle$ e substituindo acima,

$$\frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} - 2\frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} + \langle \boldsymbol{y}, \boldsymbol{y} \rangle \geq 0 \quad \iff \quad \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle,$$

ou $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} \langle \boldsymbol{y}, \boldsymbol{y} \rangle^{1/2} = \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$. $\square$

A igualdade ocorre se e somente se: um dos vetores é igual a zero, ou quando $\boldsymbol{x} = \boldsymbol{y}$.

Distâncias, Ângulos e Projeções

A norma euclidiana induz uma *distância* entre vetores em $\mathbb{R}^n$:

$$d(\boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{y} - \boldsymbol{x}\|, \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Também fornece o ângulo θ entre vetores $\boldsymbol{x}$ e $\boldsymbol{y}$:

$$\theta = \arccos\left(\frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}\right).$$

Assim, $\boldsymbol{x}$ e $\boldsymbol{y}$ são *ortogonais* se $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

A *projeção escalar* de $\boldsymbol{x}$ na direção de $\boldsymbol{y}$ é dada por

$$x_y = \|\boldsymbol{x}\| \cos\theta = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|},$$

em que θ é o ângulo entre $\boldsymbol{x}$ e $\boldsymbol{y}$.

Bases Ortonormais

Uma base $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ de $\mathbb{R}^n$ é *ortogonal* se para $i \neq j$,

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0.$$

$\mathcal{B}$ é *ortonormal* se é ortogonal e $\|\boldsymbol{v}_i\| = 1$, para todo i.

Exemplo: $E = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ é uma base ortonormal de $\mathbb{R}^n$.

Temos ainda, para todo $\boldsymbol{x} \in \mathbb{R}^n$, que:

$$\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \langle \boldsymbol{x}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 + \cdots + \langle \boldsymbol{x}, \boldsymbol{e}_n \rangle \boldsymbol{e}_n.$$

Teorema

Todo espaço vetorial de dimensão finita com produto interno admite uma base ortonormal.

Uma prova é fornecida pelo algoritmo de ortonormalização de Gram-Schmidt.