MCTA017 – Programação Matemática

EPR202 – Métodos de Otimização Aplicados à Eng. de Produção

Aula 4

Revisão de Álgebra Linear (Parte II)

Prof. Dr. Aritanan Gruber

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC

Quadrimestre Suplementar 2020

Outline

Subespaços Ortogonais

Para um conjunto $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subseteq \mathbb{R}^n$, definimos o *complemento ortogonal* a S como

$$S^\perp := \{\boldsymbol{w} \in \mathbb{R}^n : \langle \boldsymbol{w}, \boldsymbol{v}_j \rangle = 0, \forall\, j \in [n]\}.$$

Observe que $S^\perp$ é um subespaço vetorial e que todo vetor em $S^\perp$ é ortogonal a todo vetor em $\operatorname{lin.span}(S)$.

Proposição

Se $S \subseteq \mathbb{R}^n$ é um subespaço, então

(a) $S = \left(S^\perp\right)^\perp$; *($S \subset \left(S^\perp\right)^\perp$ se S não é subespaço)*

(b) $S \cap S^\perp = \{\mathbf{0}\}$;

(c) $\mathbb{R}^n = S \oplus S^\perp$ *e* $\dim(\mathbb{R}^n) = \dim(S) + \dim(S^\perp)$.

Vamos provar os itens (b) e (c). O item (a) é consequência e fica como exercício.

Subespaços Ortogonais

Prova. Seja S um subespaço de $\mathbb{R}^n$. Provamos $\mathbb{R}^n = S + S^\perp$ primeiro. Seja $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ uma base ortonormal para S.

Todo $\boldsymbol{x} \in \mathbb{R}^n$ pode ser escrito como

$$\boldsymbol{x} = \underbrace{\langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \cdots + \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k}_{=:\boldsymbol{u}} + \underbrace{\boldsymbol{x} - \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 - \cdots - \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k}_{=:\boldsymbol{v}}.$$

Claramente, $\boldsymbol{u} \in S$. Como $\langle \boldsymbol{v}, \boldsymbol{e}_j \rangle = \langle \boldsymbol{x}, \boldsymbol{e}_j \rangle - \langle \boldsymbol{x}, \boldsymbol{e}_j \rangle = 0$, pois $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0$ para $i \neq j$ e $\langle \boldsymbol{e}_j, \boldsymbol{e}_j \rangle = 1$, temos que $\boldsymbol{v} \in S^\perp$. Logo, mostramos que todo $\boldsymbol{x} \in S + S^\perp$, ou que $\mathbb{R}^n = S + S^\perp$.

Para concluir, tome $\boldsymbol{x} \in S \cap S^\perp$. Como $\boldsymbol{x} \in S^\perp$, temos que $\boldsymbol{x}$ é ortogonal a todos os vetores em S, incluindo o próprio $\boldsymbol{x}$, pois também está em S. Assim, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$, o que implica que $\boldsymbol{x} = \boldsymbol{0}$.

Portanto, $S \cap S^\perp = \{\boldsymbol{0}\}$ e consequentemente, $\mathbb{R}^n = S \oplus S^\perp$. Segue imediatamente que $\dim(\mathbb{R}^n) = \dim(S) + \dim(S^\perp)$. $\square$

Subespaços Ortogonais

Nota: A prova anterior fornece o princípio de projeções ortogonais.

Parafraseando o item (a), em espaços de dimensão finita, a “ortogonalização” de subespaços é *involutiva*: se S é um subespaço (primal), $S^{\perp}$ é o subespaço ortogonal (dual) a S e o dual do dual é igual ao primal.

As seguintes relações envolvendo subespaços associados a matrizes reais serão úteis:

Proposição

Para $\boldsymbol{A} \in \mathbb{R}^{m \times n}$,

(a) $\big(\mathrm{row.\,span}(\boldsymbol{A})\big)^{\perp} = \mathrm{null}(\boldsymbol{A})$,

(b) $\big(\mathrm{range}(\boldsymbol{A})\big)^{\perp} = \mathrm{null}(\boldsymbol{A}^{\top})$.

Vamos provar o item (b). A prova de (a) fica como exercício.

Subespaços Ortogonais

Prova.
Seja $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. Temos que

$$\begin{aligned}
\boldsymbol{x} \in \big(\operatorname{range}(\boldsymbol{A})\big)^{\perp} &\iff \langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle = 0 \qquad\qquad \forall\, \boldsymbol{y} \in \mathbb{R}^n \\
&\iff \langle \boldsymbol{A}^\top \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \qquad\qquad \forall\, \boldsymbol{y} \in \mathbb{R}^n \\
&\iff \boldsymbol{A}^\top \boldsymbol{x} = \boldsymbol{0} \\
&\iff \boldsymbol{x} \in \operatorname{null}(\boldsymbol{A}^\top),
\end{aligned}$$

em que $\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{A}^\top \boldsymbol{x} = \langle \boldsymbol{y}, \boldsymbol{A}^\top \boldsymbol{x} \rangle = \langle \boldsymbol{A}^\top \boldsymbol{x}, \boldsymbol{y} \rangle$. $\square$

De forma similar, é possível provar que $\operatorname{range}(\boldsymbol{A}) = \operatorname{null}(\boldsymbol{A}^\top)^{\perp}$, $\big(\operatorname{range}(\boldsymbol{A}^\top)\big)^{\perp} = \operatorname{null}(\boldsymbol{A})$, e $\operatorname{range}(\boldsymbol{A}^\top) = \operatorname{null}(\boldsymbol{A})^{\perp}$.

Representação de Subespaços

Um *hiperplano (linear)* em $\mathbb{R}^n$ é um conjunto

$$H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} = 0\}$$

para algum vetor $\boldsymbol{a} \neq \mathbf{0}$ em $\mathbb{R}^n$. Claramente, H é um subespaço.

O resultado abaixo resume uma parte do que fizemos até aqui.

Teorema (representação de subespaços)

Um subespaço $S \subseteq \mathbb{R}^n$ é gerado por uma coleção finita de vetores e também é a interseção de um número finito de hiperplanos lineares.

Prova. Como S é um subespaço de $\mathbb{R}^n$, $\dim(S) \leq n < \infty$ e qualquer base para S é finita. O mesmo argumento se aplica a $S^\perp$ e se $\mathcal{B}^\perp = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_\ell\}$ é uma base para $S^\perp$, temos que

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{u}_i^\top \boldsymbol{x} = 0,\, \boldsymbol{u}_i \in \mathcal{B}^\perp\}. \quad \square$$

Representação de Subespaços

Consequentemente, para qualquer subespaço S de $\mathbb{R}^n$ existem matrizes $\boldsymbol{A} \in \mathbb{R}^{\ell \times n}$ e $\boldsymbol{C} \in \mathbb{R}^{n \times n}$ tais que

$$S = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} = \{\boldsymbol{x} : \boldsymbol{x} = \boldsymbol{C}\boldsymbol{y} \text{ para algum } \boldsymbol{y} \in \mathbb{R}^n\},$$

em que $\ell := \dim(S^\perp) = n - \dim(S)$.

Ou seja, S é o kernel (espaço nulo) de $\boldsymbol{A}$ e a imagem (espaço linha) de $\boldsymbol{C}$:

$$S = \text{null}(\boldsymbol{A}) = \text{range}(\boldsymbol{C}).$$

Nota: $\boldsymbol{A}$ e $\boldsymbol{C}$ são, ainda, matrizes de transformações lineares.

Representação de Subespaços

Podemos usar o teorema da representação de subespaços para obter uma nova prova da existência de um sistema fundamental de soluções para um sistema linear homogêneo.

Corolário

Para qualquer matriz $\boldsymbol{A}$, existem vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ tais que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ se e somente se $\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots \lambda_k \boldsymbol{x}_k$ para certos $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Prova. Aplique o teorema de representação em $\operatorname{null}(\boldsymbol{A})$. □

Nota: $\lambda_1, \lambda_2, \ldots, \lambda_k$ acima podem ser feitos racionais.

Outline

Conjuntos Afins

Para $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, o conjunto de pontos

$$(1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y} = \boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}) \qquad \forall\, \lambda \in \mathbb{R}$$

é uma *linha* em $\mathbb{R}^n$ que passa por $\boldsymbol{x}$ e $\boldsymbol{y}$, se $\boldsymbol{x} \neq \boldsymbol{y}$, ou um ponto em caso contrário.

Um conjunto $\emptyset \neq L \subseteq \mathbb{R}^n$ é *afim* se

$$(1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y} \in L \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in L \text{ e } \lambda \in \mathbb{R}^n.$$

$\mathbb{R}^n$ é claramente afim e, por definição, o conjunto vazio $\emptyset$ é afim.

Conjuntos afins são intimamente ligados a espaços vetoriais lineares, como mostrado pelos dois resultados que seguem.

Proposição (exercício)

Os subespaços vetoriais de $\mathbb{R}^n$ são conjuntos afins que possuem $\mathbf{0}$.

Conjuntos Afins

Dois conjuntos afins L e M são *paralelos* se, para algum $\boldsymbol{a} \in \mathbb{R}^n$,

$$L = M + \boldsymbol{a} := \{\boldsymbol{x} + \boldsymbol{a} : \boldsymbol{x} \in M\}.$$

Dizemos que $\boldsymbol{a}$ é o *translado* de L com relação a M.

A noção de paralelismo entre conjuntos afins é induz um classe de equivalências, pois para L, M, N conjuntos afins, temos:

- *reflexividade:* L é paralelo a L;
- *simetria:* L é paralelo a M se e somente se M é paralelo a L;
- *transitividade:* se L é paralelo a M e M é paralelo a N, então L é paralelo a N.

Não surpreendente, conjuntos afins são paralelos a subespaços vetoriais.

Conjuntos Afins

Especificamente:

Teorema

Para todo subespaço afim $L \neq \emptyset$, $V = L - L := \{\boldsymbol{x} - \boldsymbol{y} : \boldsymbol{x}, \boldsymbol{y} \in L\}$ é o único subespaço vetorial paralelo a L.

Prova.

Sejam V e W dois espaços vetoriais paralelos a L. Temos então que V é paralelo a W. Isto é, existe $\boldsymbol{a} \in \mathbb{R}^n$ tal que $V = W + \boldsymbol{a}$.

Como $\mathbf{0} \in V$, temos que $-\boldsymbol{a} \in W$ e, assim, que $\boldsymbol{a} \in W$ implicando em $W \supseteq W + \boldsymbol{a} = V$. Repetindo o argumento com V e W trocados, obtemos que $V \supseteq W$ e, portanto, $V = W$.

Observe, agora, que para todo $\boldsymbol{x} \in L$, $L - \boldsymbol{x} = L + (-\boldsymbol{x})$ é um translado de L contendo $\mathbf{0}$. Pela proposição anterior, este conjunto afim é um subespaço vetorial V — o único paralelo a L.

Conjuntos Afins

Finalmente, como $V = L - \boldsymbol{x}$ independentemente do vetor $\boldsymbol{x} \in L$ que escolhermos, segue que $V = L - L$. $\square$

O teorema acima possibilita abusarmos da nomenclatura e nos referirmos a conjuntos afins não vazios por *espaços vetoriais afins*.

O fato de um conjunto afim não vazio possuir um único espaço vetorial *associado* (ou *subjacente*) permite importarmos o ferramental desenvolvido para espaços vetoriais — com ligeiras modificações.

Antes da importação, dois exemplos de interesse.

Conjuntos Afins

Neste curso, estamos interessados nos seguintes conjuntos afins.

(1) *Hiperplanos afins:* conjuntos da forma

$$H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} = b\} \qquad (\bullet)$$

para vetores $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \mathbf{0}$, e escalares $b \in \mathbb{R}$.

Se $b = 0$, o hiperplano é homogêneo (subespaço vetorial). Hiperplanos afins em $\mathbb{R}$ são pontos, em $\mathbb{R}^2$, retas e em R^3, planos. Não há nomenclatura especial para hiperplanos em $\mathbb{R}^k$, com $k \geq 4$.

Teorema

(a) *Para todo hiperplano H, temos que $\dim(H) = n - 1$.*

(b) *Todo conjunto afim em $\mathbb{R}^n$ de dimensão $n - 1$ é um hiperplano; a representação deste como em $(\bullet)$ é única módulo um múltiplo escalar comum a $\boldsymbol{a}$ e b.*

Conjuntos Afins

O fato de que $\dim(H) = \dim(\mathbb{R}^n) - 1$ implica que o hiperplano H divide $\mathbb{R}^n$ em dois *subespaços*:

$$H^- = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} < b\} \quad \text{e} \quad H^+ = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} > b\}.$$

(2) Soluções de *sistemas lineares afins* (ou *não-homogêneos*):

$$X = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$$

para matrizes $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ e vetores $\boldsymbol{b} \in \mathbb{R}^m$.

Claramente, X é equivalente a

$$\bigcap_{i=1}^{m} \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\},$$

a intersecção de m hiperplanos afins.

Conjuntos Afins

Note que X pode ser vazio se o sistema afim for inconsistente, e que $X = \mathbb{R}^n$ se $\boldsymbol{A} = \boldsymbol{0}$ e $\boldsymbol{b} = \boldsymbol{0}$.

De forma similar ao que ocorre com subespaços lineares, temos um teorema de representação de conjuntos afins.

Todo conjunto afim pode ser representado *internamente*, via bases afins ou *externamente*, como o conjunto solução de um sistema afim.

Antes de fornecermos uma prova, precisamos prosseguir com a importação e adaptação dos conceitos requeridos.

Outline

Combinações, Fechos e Bases Afins

No que segue, seja $\emptyset \neq L \subseteq \mathbb{R}^n$ um conjunto afim, V o subespaço vetorial associado a L e $\boldsymbol{a} \in \mathbb{R}^n$ o translado de L com relação a V.

Dados vetores $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in L$, suponha que queremos exprimir $\boldsymbol{x}$ como uma combinação dos demais, i.é,

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k,$$

para $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$.

Da relação $L = V + \boldsymbol{a}$, temos que existem vetores $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in V$ tais que $\boldsymbol{x}_i = \boldsymbol{v}_i + \boldsymbol{a}$. Logo, a combinação deve satisfazer

$$\begin{aligned} \boldsymbol{x} &= \lambda_1(\boldsymbol{v}_1 + \boldsymbol{a}) + \lambda_2(\boldsymbol{v}_2 + \boldsymbol{a}) + \cdots + \lambda_k(\boldsymbol{v}_k + \boldsymbol{a}) \\ &= \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k + (\lambda_1 + \lambda_2 + \cdots + \lambda_k)\boldsymbol{a}. \end{aligned}$$

Como $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{a}$ para algum $\boldsymbol{v} \in V$, temos que é necessário que $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$.

Combinações, Fechos e Bases Afins

Assim, dizemos que $\boldsymbol{x}$ é uma *combinação afim* de $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ se existem escalares $\lambda_1, \lambda_2, \ldots, \lambda_k$ tais que

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_n \boldsymbol{x}_n \quad \text{e} \quad \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1.$$

O *fecho afim* (ou *affine span*) de $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ é

$$\operatorname{aff.span}(X) := \left\{ \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i : \sum_{i=1}^{k} \lambda_i = 1,\ \lambda_i \in \mathbb{R} \right\},$$

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

X é um conjunto *gerador afim* de/para L se $\operatorname{aff.span}(X) = L$.

Combinações, Fechos e Bases Afins

Os vetores em X são *afim independentes* se a única solução do sistema

$$\begin{aligned} \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k &= \mathbf{0} \\ \lambda_1 + \lambda_2 + \cdots + \lambda_k &= 1 \end{aligned}$$

é $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. Em caso contrário, eles são ditos *afim dependentes*.

De forma semelhante à dependência linear, alguns vetores são "supérfluos" caso os vetores em X sejam afim dependentes.

Se X é gerador para L e seus elementos afim independentes, X é uma *base afim* de/para L.

Um argumento semelhante ao feito no caso linear mostra todas as bases afins de L têm o mesmo número ℓ de elementos. Tal número é definido como a *dimensão* de L: $\dim(L) = \ell$.

Combinações, Fechos e Bases Afins

Pergunta: Qual a relação entre $\dim(L)$ e $\dim(V)$?

Oras, como $L = V + \boldsymbol{a}$, segue que

$$\dim(L) = \dim(V) + \dim(\{\boldsymbol{a}\}) = \dim(V),$$

pois qualquer conjunto unitário ($\{\boldsymbol{a}\}$) tem dimensão zero.

Isso implica que o maior número de vetores afim independentes em $\mathbb{R}^n$ é $n+1$: n vetores linearmente independentes mais o vetor $\mathbf{0}$.

Tecnicalidade: E se tivéssemos $L = \emptyset$? Neste caso, L não tem um subespaço vetorial associado e definimos $\dim(\emptyset) = -1$.

Combinações, Fechos e Bases Afins

Revisitando a construção feita para combinações afins sob a ótica de independência, obtemos o resultado que segue.

Proposição

São equivalentes:

(a) $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$ *são afim independentes;*

(b) $(\boldsymbol{x}_2 - \boldsymbol{x}_1), (\boldsymbol{x}_3 - \boldsymbol{x}_1), \ldots, (\boldsymbol{x}_k - \boldsymbol{x}_1) \in \mathbb{R}^n$ *são linearmente independentes;*

(c) $(\boldsymbol{x}_1, -1), (\boldsymbol{x}_2, -1), \ldots, (\boldsymbol{x}_k, -1) \in \mathbb{R}^{n+1}$ *são linearmente independentes.*

Note que independência linear implica em independência afim, e que a recíproca não é verdadeira.

Conjuntos Ortogonais e Transformações Afins

O produto interno canônico de $\mathbb{R}^n$ tem a mesma funcionalidade sobre subespaços vetoriais e conjuntos afins em $\mathbb{R}^n$.

Além disso, nossa construção de subespaços ortogonais foi suficientemente geral para acomodar conjuntos afins. Assim,

$$L^{\perp} := \{\boldsymbol{y} \in \mathbb{R}^n : \langle \boldsymbol{y}, \boldsymbol{x} \rangle = 0,\, \boldsymbol{x} \in L\}$$

é o subespaço ortogonal a L.

Toda *transformação afim* $T : (L \subseteq \mathbb{R}^n) \to (M \subseteq \mathbb{R}^m)$ tem a forma

$$T(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{a}) + \boldsymbol{b},$$

em que $f : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear e $\boldsymbol{b} \in \mathbb{R}^m$ é o translado de M.

A representação de f por uma matriz $\boldsymbol{A}_f$ pode ser claramente adaptada a T: $\boldsymbol{A}_T = (\boldsymbol{A}_f, \boldsymbol{b}) = (\boldsymbol{A}_f \,|\, \boldsymbol{b})$.

Outline

Representações de Conjuntos Afins

A prova da representação interna de conjuntos afins foi feita, sem alarde, quando falamos de bases afins. A da representação externa, foi indicada quando mencionamos ortogonalidade em conjuntos afins.

Resumimos as construções no seguinte resultado.

Corolário

Todo conjunto afim não vazio é gerado por uma coleção finita de vetores e também pela intersecção de um número finito de hiperplanos afins.

Prova. exercício.

Teorema Fundamental da Álgebra Linear

($\boldsymbol{A} \in \mathbb{R}^{m \times n}$ e $\boldsymbol{b} \in \mathbb{R}^m$)

Teorema (fundamental)

O sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ tem solução se e somente se $\boldsymbol{y}^\top \boldsymbol{b} = 0$ para cada vetor $\boldsymbol{y} \in \mathbb{R}^m$ tal que $\boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{0}$.

Prova. A necessidade é imediata, já que $\boldsymbol{y}^\top \boldsymbol{b} = \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} = 0$ sempre que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ e $\boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{0}$.

Para suficiência, seja $L := \{\boldsymbol{z} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z} \text{ para algum } \boldsymbol{x}\}$. Pelo Teorema de representação de espaços vetoriais, existe uma matriz $\boldsymbol{C}$ tal que $L = \{\boldsymbol{z} : C\boldsymbol{z} = \boldsymbol{0}\}$.

Por ortogonalidade, temos que $\boldsymbol{C}\boldsymbol{A} = \boldsymbol{0}$. Isso implica que $\boldsymbol{C}\boldsymbol{b} = \boldsymbol{0}$ (estamos usando as linhas de $\boldsymbol{C}$ como $\boldsymbol{y}$ e, por hipótese, $\boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{0}$ implica em $\boldsymbol{y}^\top \boldsymbol{b} = 0$), que por ventura implica em $\boldsymbol{b} \in L$.

Portanto, existe um $\boldsymbol{x}$ tal que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. $\square$

Teorema Fundamental da Álgebra Linear

Teorema (sistema fundamental de soluções)

Seja $\boldsymbol{x}_0$ uma solução do sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Existem vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ tais que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ se e somente se

$$\boldsymbol{x} = \boldsymbol{x}_0 + \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k$$

para certos escalares $\lambda_1, \lambda_2, \ldots, \lambda_k$ (podem ser tomados em $\mathbb{Q}$).

Prova.

Observe que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ se e somente se $\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{0}$. Aplicando o teorema de representação de espaços vetoriais em

$$\{\boldsymbol{x} : \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{0}\}$$

fornece o resultado. $\square$

Teorema Fundamental da Álgebra Linear

É conveniente reinterpretarmos os dois teoremas anteriores à luz dos subespaços associados a matrizes.

Essencialmente:

- O sistema $\boldsymbol{Ax} = \boldsymbol{b}$ tem solução se e somente se $\boldsymbol{b} \in \operatorname{range}(\boldsymbol{A})$. Condição necessária: $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}((\boldsymbol{A} \,|\, \boldsymbol{b}))$.
- Toda solução $\boldsymbol{x}$ tem a forma $\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{y}$, com $\boldsymbol{z} \in \operatorname{null}(\boldsymbol{A})$ e $\boldsymbol{y} \in \operatorname{row.span}(\boldsymbol{A})$ tal que $\boldsymbol{Ay} = \boldsymbol{b}$.
- Se $\boldsymbol{Ax} = \boldsymbol{b}$ não tem solução, existe $\boldsymbol{y} \in \operatorname{null}(\boldsymbol{A}^\top)$ com $\boldsymbol{y}^\top \boldsymbol{b} \neq 0$.

Certifique-se de que você compreende o que está escrito acima!

Outline

Escalonamento de Sistemas Lineares

Considere o sistema linear $\boldsymbol{Ax} = \boldsymbol{b}$ escrito explicitamente como:

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

O *escalonamento* de tal sistema consiste na transformação dele em um sistema equivalente $\boldsymbol{A'x} = \boldsymbol{b'}$ tal que $\boldsymbol{A'}$ é triangular superior:

$$\begin{array}{ccccccccc} a'_{1,1}x_1 & + & a'_{1,2}x_2 & + & \cdots & + & a'_{1,n}x_n & = & b'_1 \\ & & a'_{2,2}x_2 & + & \cdots & + & a'_{2,n}x_n & = & b'_2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & a'_{m,n}x_n & = & b'_m \end{array}$$

Os sistemas são *equivalentes* se possuem o mesmo conjunto de soluções: $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{Ax} = \boldsymbol{b}\} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A'x} = \boldsymbol{b'}\}$.

Escalonamento de Sistemas Lineares

O escalonamento é obtido pela aplicação em $(\boldsymbol{A} \,|\, \boldsymbol{b})$, um número finito de vezes, das operações:

- trocar as posições de duas linhas:
  $(\boldsymbol{A} \,|\, \boldsymbol{b})_i^\top \leftrightarrow (\boldsymbol{A} \,|\, \boldsymbol{b})_j^\top$;
- multiplicar uma linha por um escalar $c \in \mathbb{R}$ não nulo:
  $(\boldsymbol{A} \,|\, \boldsymbol{b})_i^\top \leftarrow c \cdot (\boldsymbol{A} \,|\, \boldsymbol{b})_i^\top$;
- somar o múltiplo de uma linha a uma outra linha:
  $(\boldsymbol{A} \,|\, \boldsymbol{b})_i^\top \leftarrow (\boldsymbol{A} \,|\, \boldsymbol{b})_i^\top + c \cdot (\boldsymbol{A} \,|\, \boldsymbol{b})_j^\top$.

As três operações acima equivalem à multiplicação de $(\boldsymbol{A} \,|\, \boldsymbol{b})$ por matrizes apropriadas (tente exibi-las; são simples). Logo, $(\boldsymbol{A} \,|\, \boldsymbol{b})$ é escalonado em $(\boldsymbol{A}' \,|\, \boldsymbol{b}')$ por uma sequência de composições de transformações lineares.

A equivalência entre $(\boldsymbol{A} \,|\, \boldsymbol{b})$ e $(\boldsymbol{A}' \,|\, \boldsymbol{b}')$ advém das transformações realizarem combinações lineares em $\operatorname{row.span}\big((\boldsymbol{A} \,|\, \boldsymbol{b})\big)$.

Escalonamento de Sistemas Lineares

É mais fácil ilustrarmos as operações em um exemplo:

$$(\boldsymbol{A}\,|\,\boldsymbol{b}) = \left(\begin{array}{rrr|r} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{array}\right)$$

Somando a segunda linha com $3/2$ da primeira e a terceira com a primeira:

$$\simeq \left(\begin{array}{ccr|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 2 & 1 & 5 \end{array}\right)$$

Somando a terceira linha com -4 vezes a segunda:

$$\simeq \left(\begin{array}{ccr|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 1 \end{array}\right)$$

$$= (\boldsymbol{A}'\,|\,\boldsymbol{b}')$$

Escalonamento de Sistemas Lineares

Caso a matriz $\boldsymbol{A}$ não tenha posto completo ($\mathrm{rank}(\boldsymbol{A}) < \min\{m, n\}$), $\boldsymbol{A}'$ não será triangular superior; terá a seguinte estrutura:

$$\boldsymbol{A}' = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{D} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

em que $\boldsymbol{C}$ é triangular superior, $\boldsymbol{D}$ é qualquer, e $\boldsymbol{0}$ é a matriz nula.

O vetor $\boldsymbol{b}$ será transformado em

$$\boldsymbol{b}' = \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{pmatrix}.$$

Claramente, o sistema é inconsistente se $\boldsymbol{d} \neq \boldsymbol{0}$.

Em outras palavras, o sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ tem solução se

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}\left((\boldsymbol{A} \,|\, \boldsymbol{b})\right).$$

Escalonamento de Sistemas Lineares

Por exemplo,

$$(\boldsymbol{A}\,|\,\boldsymbol{b}) = \left(\begin{array}{rrr|r} 1 & 3 & 1 & 9 \\ 1 & 1 & -1 & 1 \\ 3 & 11 & 5 & 36 \end{array}\right) \simeq \left(\begin{array}{rrr|r} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 2 & 2 & 9 \end{array}\right)$$

$$\simeq \left(\begin{array}{rrr|r} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 0 & 0 & 1 \end{array}\right) = (\boldsymbol{A}'\,|\,\boldsymbol{b}')$$

Note que $\boldsymbol{A}'$ tem uma linha nula, pois

$$(3,11,5) - 3(1,3,1) + ((1,1,-1) - (1,3,1)) = 0,$$

e que a entrada correspondente em $\boldsymbol{b}'$ é diferente de 0.

Claramente, $2 = \operatorname{rank}(\boldsymbol{A}) < \operatorname{rank}\big((\boldsymbol{A}\,|\,\boldsymbol{b})\big) = 3$

Eliminação Gaussiana

para cada coluna $k = 0, 1, \ldots, n-2$

seja $\ell = \mathrm{argmax}\{|A[i][k]| : i = k, k+1, \ldots, n-1\}$

se $A[\ell][k] = 0$, ignore a coluna k e continue com o próximo passo do laço

troque as linhas $(A[k][\cdot], b[k]) \leftrightarrow (A[\ell][\cdot], b[\ell])$

seja pivot $= A[k][k]$

para cada linha $i = k+1, k+2, \ldots, n-1$

seja $c = -A[i][k]/\text{pivot}$

faça $A[i][k] = 0$

para cada coluna $j = k+1, k+2, \ldots, n-1$

faça $A[i][j] = A[i][j] + c \cdot A[k][j]$

faça $b[i] = b[i] + c \cdot b[k]$

$\stackrel{*}{\Longrightarrow}$ **Alerta: mudar notação e incluir explicações**

Eliminação Gaussiana

Código em Python

```
def gaussian (A: [[float]], b: [float]):
  n = len (A)
  for k in range (n-1):
    l = max (range (k, n), \
               key = lambda i: abs (A[i][k]))
    if A[l][k] == 0: continue
    A[l], A[k] = A[k], A[l]
    b[l], b[k] = b[k], b[l]

    pivot = A[k][k]
    for i in range (k+1, n):
      c = -A[i][k] / pivot
      A[i][k] = 0
      for j in range (k+1, n):
        A[i][j] += c * A[k][j]
      b[i] += c * b[k]
```

Outline

Conjuntos Convexos

Sejam $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Uma *combinação convexa* de $\boldsymbol{x}$ e $\boldsymbol{y}$ é um vetor da forma

$$\boldsymbol{z} = (1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y} = \boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}) \quad \text{para algum} \quad \lambda \in [0,1].$$

O conjunto de todas as combinações convexas de $\boldsymbol{x}$ e $\boldsymbol{y}$,

$$\operatorname{cvx.hull}(\{\boldsymbol{x}, \boldsymbol{y}\}) := \{(1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y} : \lambda \in [0,1]\},$$

é o segmento de reta em $\mathbb{R}^n$ com extremidades $\boldsymbol{x}$ e $\boldsymbol{y}$.

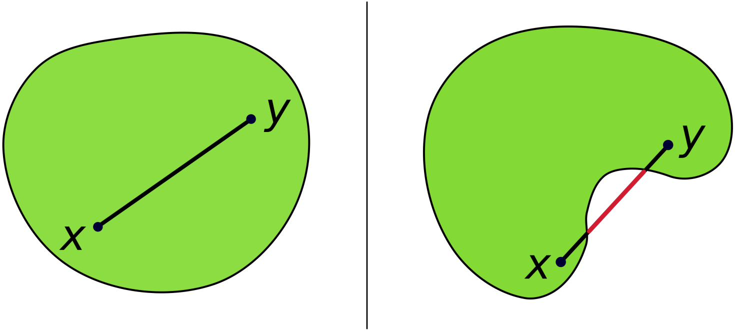
Um conjunto $\emptyset \neq C \subseteq \mathbb{R}^n$ é *convexo* se para quaisquer dois pontos $\boldsymbol{x}, \boldsymbol{y} \in C$, todas as combinações convexas de $\boldsymbol{x}$ e $\boldsymbol{y}$ estão em C; i.é,

$$\operatorname{cvx.hull}(\{\boldsymbol{x}, \boldsymbol{y}\}) \subseteq C \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in C.$$

Por definição, $\emptyset$ é convexo.

Conjuntos Convexos

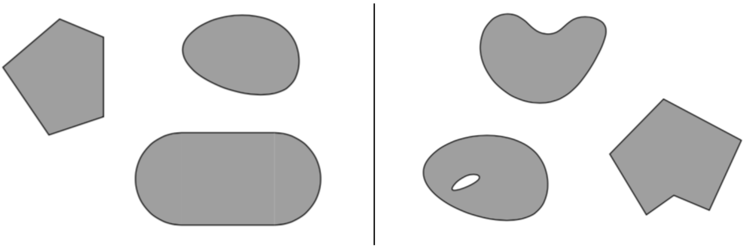
Exemplos



À esquerda um conjuntos convexo; à direita, um não convexo.
(Extraído de Wikipedia)

Conjuntos Convexos

Outros exemplos



À esquerda, conjuntos convexos e à direita, não convexos.
(Extraído de Matoušek-Gartner, pg. 49)

Dimensão e Combinações Convexas

A *dimensão* de um conjunto convexo C é igual à do conjunto afim de menor dimensão que o contém. Em outras palavras,

$$\dim(C) := \min \left\{ \dim(L) : C \subseteq L, \text{ com } L \text{ afim} \right\}.$$

($\dim(\emptyset) = -1$)

Combinações convexas podem envolver mais de dois vetores.

Um vetor $\boldsymbol{x}$ é uma *combinação convexa* de $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$ se

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k,$$

com $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$ e $\lambda_i \geq 0$.

Fechos Convexos

O *fecho convexo* de um conjunto $X \subseteq \mathbb{R}^n$ é o conjunto de todas as combinações convexas dos vetores em X. Logo, X é convexo se e somente se $X = \operatorname{cvx.hull}(X)$.

Se $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\} \subset \mathbb{R}^n$ é finito,

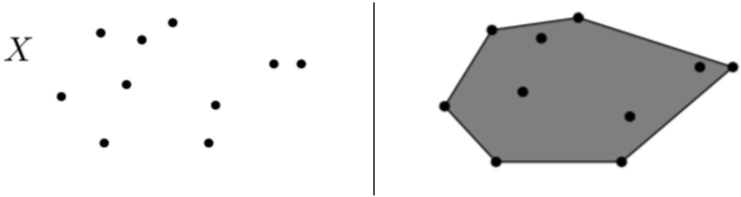
$$\operatorname{cvx.hull}(X) := \left\{ \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i : \sum_{i=1}^{k} \lambda_i = 1,\ \lambda_i \geq 0 \right\}.$$

Neste caso, $\operatorname{cvx.hull}(X)$ é *finitamente gerado*.

Um conjunto convexo finitamente gerado é chamado de *politopo*.

Fechos Convexos

Exemplo



À esquerda, um conjunto finito X; à direita, $\mathrm{cvx.\,hull}(X)$.
(Extraído de Matoušek-Gartner, pg. 49)