Introdução à Programação Linear (MAC - 315 / 2000): versão 3.1

Leônidas de Oliveira Brandão
http://www.ime.usp.br/~leo
http://www.matematica.br

8 de agosto de 2001

Sumário

Conceitos básicos

0.1 Convenções

0.1.1 Representando matrizes e vetores

Para representar um VETOR, utilizaremos sempre um tipo diferente de fonte: $\boldsymbol{v}$, por exemplo, é um vetor e $\mathbb{0}$ corresponde ao vetor nulo. Sempre os interpretaremos como sendo vetores-coluna. Já MATRIZES, escreveremos da seguinte forma: $\mathbf{A}$ é um exemplo de matriz e $\mathbf{O}$ corresponde à matriz nula.

Sendo assim, cada coluna de uma matriz pode ser entendida como um vetor-coluna que denotaremos como a^i, e cada linha será interpretada como o tranposto de um vetor-coluna e será escrita como a_i.

Dados dois vetores $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, dizemos que:

- $\boldsymbol{x} \leqq \boldsymbol{y} \iff \langle \forall i,\ x_i < y_i$ ou $x_i = y_i \rangle$ (ou equivalentemente: $\langle \nexists i \in \{1, 2, \ldots, n\} : x_i > y_i \rangle$)
- $\boldsymbol{x} \leq \boldsymbol{y} \iff \boldsymbol{x} \leqq \boldsymbol{y}$ e $\boldsymbol{x} \neq \boldsymbol{y}$ (se n=1, $x \leq y \Longrightarrow x < y$)

0.2 Demonstrações

Nesta seção apresentaremos uma breve discussão sobre teoremas: demonstrações e contra-exemplos.

Uma sentença matemática pode estar errada ou correta. No primeiro deve existir um CONTRA-EXEMPLO para a mesma (um exemplo no qual a sentança não seja válida), enquanto que no segundo deve ser possível demonstrar a inexistência de contra-exemplos.

Estas sentenças propõem uma TESE a partir da suposição de algumas HIPÓTESES

$$\text{Teorema: conjunto de hipóteses} \Longrightarrow \text{tese.}$$

e o processo de demonstração, consiste em encadear argumentos lógicos que comprovem a correção da tese.

A seguir apresentaremos algumas das técnicas de demonstrações mais úteis.

0.2.1 Inspeção

Consiste no exame direto da tese, normalmente empregado quando existe um número finito de possibilidades e que podem ser facilmente examinados.

Exemplo 0.2.1 *Resolver qualquer problema do tipo:*

$$\max \boldsymbol{c}'\boldsymbol{x} := \max \sum_{i=1}^{n} c_i x_i, \ \boldsymbol{x} \in \{\boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3, ..., \boldsymbol{x}^n\}.$$

0.2.2 Lógica

Nesta técnica devemos utilizar argumentos lógicos baseados em resultados já conhecidos, definições ou axiomas. Para isso são úteis resultados de lógica como os seguintes:

$$A \Longrightarrow B \equiv \neg B \Longrightarrow \neg A.$$

Leis de Morgan:

$\neg(A \text{ } e \text{ } B) \equiv \neg A \text{ } ou \text{ } \neg B$
$\neg(A \text{ } ou \text{ } B) \equiv \neg A \text{ } e \text{ } \neg B$

Cuidados com os quantificadores

Os quantificadores `existe` ($\exists$) e qualquer ($\forall$) costumam causar grandes problemas em sentenças matemáticas. Um tipo de erro muito frequente diz respeito às suas negações. Veremos alguns exemplos:

Exemplo: Considere as seguintes frase e sua negativa

	Sentenças	Negativa ERRADA	Negativas corretas
1	*todos os alunos reprovaram*	~~*todos os alunos foram aprovados*~~	*nem todos os alunos foram reprovados*, ou *alguns alunos foram aprovados*
2	*todos os números são pares*	~~*todos os números não são pares*~~	*nem todos os números são pares*, ou *alguns números não são pares*
3	$\forall a \in A \quad \Rightarrow \quad a \in B$	~~$a \in A \quad \Rightarrow \quad a \notin B$~~	"$\forall a \in A \not\Rightarrow a \in B$" ou "$\exists a \in A : a \notin B$"

Note que na linha 3, a primeira sentença equivale a "$A \subseteq B$", a segunda a "$A \cap B = \emptyset$" e a terceira a "$A \not\subseteq B$" (ou "$B \setminus A \neq \emptyset$").

Exemplo: $A = \{x \in \mathbb{N} : \ n \ \text{ múltiplo de } 4\}$, que colocado na forma de expressões lógica equivale a $A \quad \Rightarrow \quad B$
$B = \{x \in \mathbb{N} : \ n \ \text{ par}\}$
(i.é, se $x \in A$, então $x \in B$).

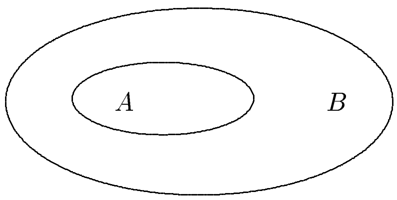


Figura 1: A está contido em B

0.2.3 Contradição

Demonstrações por contradição estão baseadas no PRINCÍPIO DO TERCEIRO EXCLUIDO, isto é, para qualquer sentença matemática A, existe apenas duas possibilidades: A é verdadeira ou A é falsa.

A estrutura de uma tal demonstração é supor o contrário e, através de argumentos lógicos válidos, chegar a um absurdo.

Teorema 0.1 (PRINCÍPIO DA INDUÇÃO) *Seja $T(n)$ qualquer propriedade sobre os naturais ($\mathbb{N}$).*

$$\text{Se} \left\langle \begin{array}{l} H_b: \ T(n_0) \text{ vale, e} \\ H_p: T(n) \Longrightarrow T(n+1), \forall n \geqq n_0, \end{array} \right\rangle, \ \text{então} \ \langle \ T(n) \text{ vale para todo } n \geqq n_0 \ (n \in \mathbb{N}) \ \rangle$$

Demonstração Supor $\exists n \in \mathbb{N}$ tal que $T(n)$ não vale e que n seja o primeiro para o qual isto ocorre. Sabemos que $n \geqq n_0 + 1$, pois por hipótese $T(n_0)$ vale (hipótese H_b). Sendo assim, $n - 1 \in \mathbb{N}$ e $T(n-1)$ vale (pois n é o primeiro contra-exemplo à propriedade $T(.)$). Daí, usando a hipótese H_b, $T(n)$ vale (contradição !!!!), o que é uma contradição com o fato de n ser o primeiro natural no qual $T(n)$ não é válido.

Logo, não pode existir um tal n, e daí conclui-se a tese. ∎

0.2.4 Indução

Uma demonstração por indução é baseada no teorema anterior, Princípio da Indução (PI), e por isso só possível utilizá-la quando a tese puder ser indexada pelos naturais. Nestas circunstâncias, devemos provar que valem as hipótese H_b e H_p do citado teorema, respectivamente, denominadas BASE DA INDUÇÃO e PASSO DA INDUÇÃO.

Assim, uma demonstração por indução tem a seguinte estrutura:

(Base da indução) Mostrar que $T(n_0)$ vale (para um $n_0 \in \mathbb{N}$, é a hipótese H_b do teorema PI);

(Passo da indução) Mostrar que $T(n) \quad \Longrightarrow \quad T(n+1)$, $\forall n \geqq n_0$ (correspondente a hipótese H_p de PI)

Daí pode-se concluir, usando o teorema PI:

$$\text{Se valem } (H_b) \text{ e } (H_p), \text{ então } T(n) \text{ vale para todo } n \geqq n_0.$$

Capítulo 0

Introdução

Resumo

Os conceitos iniciais de Programação Linear apareceram por volta de 1940, motivados por questões economicas e belicistas. Entre os primeiros trabalhos que levaram a esta grande área, podemos destacar John von Neumann que em 1928 publicou um artigo sobre TEORIA DOS JOGOS[1] e George B. Dantzig que de fato formalizou a teoria, que depois veio a chamar Programação Linear ("Linear Programming"), em 1951 um artigo tratando da MAXIMIZAÇÃO de funções lineares sujeitas a restrições também lineares[2]. Apesar de não existir um premio Nobel para a Matemática, o premio de 1975 foi recebido Koopmans (economista) e Kantorovich (matemático da então União Soviética), basicamente por trabalhar sobre a teoria formalizada por Dantzig.

0.1 Introdução

Para ilustrar um modelo de problema linear, vamos considerar uma hipotética fábrica que produz n diferentes tipos de produtos, utilizando m tipos de insumos. O que poderia interessar aos administradores/planejadores ?

Primeiro é necessário estabelecer qual o "problema". Num tal exemplo podemos estar interessados em maximizar o lucro pela venda dos produtos ou minimizar o gasto na produção (com insumo e eventualmente com estocagem, dentre outros custos). Resumidamente, podemos ter as seguintes variáveis norteando a produção:

x_i	quantidade de produtos tipo i para venda ($i \in \{1, 2, \ldots, n\}$);
u_j	valor unitário para compra de insumo tipo j ($j \in \{1, 2, \ldots, m\}$);
c_i	valor unitário para venda do produto i;
a_{ji}	quantidade de insumos tipo j, numa unidade do produto i;
b_j	limite para insumo tipo j (total disponível ou limitação de estocagem).

A partir destes dados para produção devemos MODELAR o problema. A partir das descrições acima podemos concluir que a_{ij} é determinado pela tecnologia/recursos e portanto devem ser pré-fixados. Quanto aos valores de compra dos insumos ou de venda dos produtos, depende das "forças de mercado" e por isso, podemos supor pré-determinados. Portanto o que sobra como variável determinante para a produção é a quantidade de cada produto a ser produzidos.

[1] *Zur Theorie der Gessellschaftsspiele*, Mathematische Annalen, 100:295-320.

[2] *Maximization of a Linear Function of Variables Subject to Linear Inequalities*, em T.C. Koopmans (editor), Activity Analysis of Production and Allocation, pp. 339-347, John Wiley & Sons, 1951.

Sendo c_i o valor unitário de venda do produto i e x_i sua variável de produção, poderíamos tentar a seguinte modelagem: buscamos maximizar as vendas, dada por $c_1x_1 + c_2x_2 + \cdots + c_nx_n$, respeitando as restrições de estocagem/insumo: $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqq b_i$, para cada tipo de insumo i. O que pode ser re-escrito na forma matricial,

$$
(\mathsf{P}) \qquad \begin{array}{ll}
\max \sum_{i:=1}^{n} c_i x_i & (\text{FUNÇÃO OBJETIVO}) \\
\text{sujeito a:} & \begin{array}{rcl}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & \leqq & b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & \leqq & b_2 \\
\vdots & & \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & \leqq & b_m \\
(x_1; x_2; \ldots; x_n) \geqq \mathbf{0}. & &
\end{array}
\end{array}
$$

Assim, o objetivo numa tal fábrica é encontrar uma quantidade não negativa de produção, $\overline{x}_i$, para cada produto i, respeitando as restrições acima e que maximize o valor a ser obtido com suas vendas.

0.1.1 Um mercado de dois agentes

Mas antes de discutirmos como resolver o problema acima, vamos pensar que existam apenas "dois agentes" no mercado, a fábrica e seus fornecedores. Neste modelo podemos obter uma interpretação "dual" à (P).

Vamos admitir que os fornecedores desejem determinar o "melhor" preço $\{u_j\}_{j=1}^m$ para seus insumos. Como os fornecedores compram toda a produção da fábrica, precisam levar em consideração as características dessa.

Para fazer uma unidade do produto do i, a referida fábrica gastaria em insumos $a_{1i}u_1 + a_{2i}u_2 + \cdots + a_{mi}u_m$ (reais), sendo que cada unidade deste seria vendida por c_i (reais). Logo os fornecedores podem desejar calcular o "valor agregado" de seus produtos (insumos), exigindo que o valor gasto pela fábrica numa unidade do produto i, com insumos, seja ao menos c_i, ou seja, os fornecedores desejam que o preço total dos insumos usados na manufatura de uma unidade do produto i seja tal que

$$
a_{1i}u_1 + a_{2i}u_2 + \cdots + a_{mi}u_m \geqq c_i. \tag{1}
$$

Além disso, a fábrica compra no máximo b_i unidades de insumo i, assim o gasto total da fábrica com insumos é limitado por

$$
b_1u_1 + b_2u_2 + \cdots + b_nu_n. \tag{2}
$$

Por outro lado, os fornecedores não desejam "quebrar" a fábrica, pois esta compra seus insumos e lhes vende seus produtos. Assim, estabelecem como meta buscar o menor preço a ser cobrado pelos seus insumos ($\{u_j\}_{j=1}^m$), de modo a garantir o equilibrio com os preços dos produtos que depois irão comprar. Deste modo, como $\boldsymbol{u} \in \mathbb{R}_+^m$, o problema, sob o ponto de vista dos fornecedores, pode ser assim resumido

$$
(\mathsf{D}) \qquad \begin{array}{ll}
\min \sum_{i:=1}^{m} b_i u_i & (\text{FUNÇÃO OBJETIVO DUAL}) \\
\text{sujeito a:} & \begin{array}{rcl}
a_{11}u_1 + a_{21}u_2 + \cdots + a_{m1}u_n & \geqq & c_1 \\
a_{12}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n & \geqq & c_2 \\
\vdots & & \\
a_{1n}u_1 + a_{2n}u_2 + \cdots + a_{mn}u_n & \geqq & c_m \\
(u_1; u_2; \ldots; u_n) \geqq \mathbf{0}. & &
\end{array}
\end{array}
$$

Assim, temos um novo problema para ser resolvido e desde já fica a indagação das possíveis relações existentes entre ambos os problemas. Nesta formulação, (P) é denominado problema PRIMAL e (D) seu DUAL.

Para entendermos melhor o que ocorre podemos fazer o que é sempre recomendável numa tal situação: analisarmos alguns casos particulares ou instâncias dos problemas.

0.1.2 Exemplos

Nesta seção examinaremos dois exemplos de problemas lineares. O primeiro deles é muito simples, permitindo uma resolução gráfica. O segundo é baseado em um caso concreto vivido pela Digital Equipament Corporation (DEC) - que hoje faz parte da Compaq (vide `Bertsimas&Tsitsiklis-1997`).

Exemplo 0.1.1 *Seja* $P = \{\boldsymbol{x} \in \mathbb{R}^2 : \mathbf{A}\boldsymbol{x} \leqq \boldsymbol{b}, \boldsymbol{x} \geqq \mathbf{0}\}$ *e* $D = \{\boldsymbol{u} \in \mathbb{R}^2 : \mathbf{A}'\boldsymbol{u} \geqq \boldsymbol{c}, \boldsymbol{u} \geqq \mathbf{0}\}$*, sendo*

$$\mathbf{A} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ \boldsymbol{b} := \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad e \quad \boldsymbol{c} := \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Examinando a figura abaixo podemos concluir que o ótimo em P *é* $\overline{\boldsymbol{x}} = (3,1)'$*, com valor ótimo igual a* $\boldsymbol{c}'\overline{\boldsymbol{x}} = 7$*, enquanto o ponto ótimo em* D *é* $\overline{\boldsymbol{u}} = (3/2, 1/2)'$*, com valor ótimo* $\boldsymbol{b}'\overline{\boldsymbol{u}} = 7$.

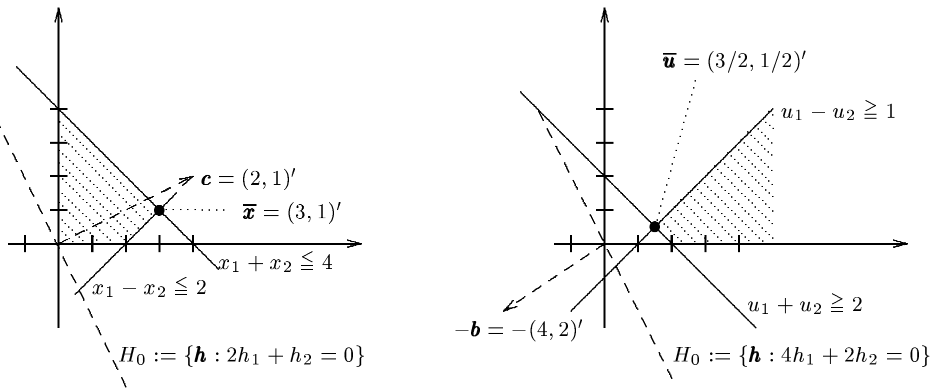


Figura 1: O poliedro é uma intersecção finita de semi-espaços e hiperplanos

Do exemplo acima, podemos conjecturar que exista uma forte relação entre ambos os problemas, parecida com uma situação de equilíbrio. De fato, veremos na seção ?? que ambos os problemas mantém um relação estreita e que quando ambos forem VIÁVEIS[3] o valor objetivo dos mesmos coincidirá.

O exemplo a seguir consta do livro `Bertsimas&Tsitsiklis-1997`, páginas 6 a 10.

Exemplo 0.1.2 *No quarto dradrimestre de 1988, a DEC introduziu uma nova família de computadores, denominados* GP-1, GP-2, GP-3, WS-1 *e* WS-2. *A tabela abaixo descreve as características dos sistemas.*

Sistema	*Preço*	*# "disk drives"*	*# placas de 256K*
GP-1	*\$60 000*	*0.3*	*4*
GP-2	*\$40 000*	*1.7*	*2*
GP-3	*\$30 000*	*0.0*	*2*
WS-1	*\$60 000*	*1.4*	*2*
WS-2	*\$60 000*	*0.0*	*1*

[3] Um problema (P) (ou (D)), com variável $\boldsymbol{x} \in \mathbb{R}^n$ ($\boldsymbol{y} \in \mathbb{R}^m$), é viável se existir um vetor $\overline{\boldsymbol{x}} \in \mathbb{R}^n$ ($\overline{\boldsymbol{y}} \in \mathbb{R}^m$) satisfazendo as retrições de (P) (de ((D))).

A terceira coluna ("disk drives" - leitores de disquetes), é o número de leitores por unidades vendidas e quarta o número de placas de memórias de $256Kb$*. Por exemplo, o sistema GP-2 usa* 2 *placas de memória e, em média, cada* 10 *unidades saem com* 17 *unidades leitoras de disquetes.*

Foram levantadas as seguintes dificuldades para o referido quadrimestre:

O fornecedor de UCP pode prover no máximo 7000 *unidades;*

O fornecedor de leitores de disquetes acreditava poder entregar entre 3000 *e* 7000 *unidades;*

O suprimento de placas de memória (de $256Kb$*) era limitado entre* 8000 *e* 16000 *unidades.*

Quanto à demanda, o departamento de "marketing" determinou que a demanda no primeiro quadrimestre de 1989 seria de 1800 *GP-1,* 300 *de GP-3,* 3800 *da família GP e* 3200 *da família WS. Esta projeção incluia os pedidos já feitos de* 500 *unidades de GP-2,* 500 *de WS-1 e* 400 *de WS-2.*

No quadrimestre anterior, para reduzir o tempo de produção, a DEC produziu GP-1, GP-3 e WS-2 sem leitor de disquete e unidades de GP-2 e WS-1 com leitor.

A DEC pode substituir unidades de memória de $256Kb$ *por dois unidades de* $128Kb$ *no GP-1, podendo produzir até* 4000 *unidades destas placas "menores" no próximo quadrimestre.*

Resumindo os objetivos dos administradores, podemos dizer que desejavam encontrar a melhor produção para o próximo quadrimestre, considerando estes dados.

Denotarmos por x_i*,* $i \in \{1, 2, \ldots, 5\}$ *o número de unidades, em milhares, a serem produzidas de GP-1, GP-2, GP-3, WS-1 e WS-2, respectivamente. Deste modo, como* $1000x_i$ *representa o número de unidades, deve ser um número inteiro. Por simplicidade, podemos truncar* x_i *a partir da terceira casa decimal, sem a necessidade de considerar um problema de programação inteira (aqueles cujas soluções devem ser números inteiros).*

Usando um modelo simplificado, onde não é considerado as memórias alternativas, deseja-se resolver o seguinte problema linear:

$$
\begin{array}{llllllllllllll}
\max & \multicolumn{9}{l}{60x_1 + 40x_2 + 30x_3 + 30x_4 + 15x_5} & & & \textit{(total de venda em milhões de dólares)} \\
\textit{sujeito a:} & x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & \leqq & 7 & \textit{(UCP disponíveis)} \\
 & 4x_1 & + & 2x_2 & + & 2x_3 & + & 2x_4 & + & x_5 & \leqq & 8 & \textit{(memórias disponíveis de } 256Kb\textit{)} \\
 & & & x_2 & & & + & x_4 & & & \leqq & 3 & \textit{(leitores disponíveis)} \\
 & x_1 & & & & & & & & & \leqq & 1.8 & \textit{(demanda máxima por GP-1)} \\
 & & & & & x_3 & & & & & \leqq & 0.3 & \textit{(demanda máxima por GP-3)} \\
 & x_1 & + & x_2 & + & x_3 & & & & & \leqq & 3.8 & \textit{(demanda máxima por GP)} \\
 & & & & & & & x_4 & + & x_5 & \leqq & 3.2 & \textit{(demanda máxima por WS)} \\
 & & & x_2 & & & & & & & \leqq & 0.5 & \textit{(demanda mínima por GP-2)} \\
 & & & & & & & x_4 & & & \leqq & 0.5 & \textit{(demanda mínima por WS-1)} \\
 & & & & & & & & & x_5 & \leqq & 0.4 & \textit{(demanda mínima por WS-2)} \\
 & x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & \leqq & 0. &
\end{array}
$$

0.1.3 Formas de um problema de programação linear

Um problema de programação linear pode ser apresentado em sua forma normal ou em sua forma canônica. Temos, abaixo, um exemplo para cada tipo de apresentação:

FORMA GERAL:

$$\max/\min \; \mathbf{c'x} \quad \text{(FUNÇÃO OBJETIVO)}$$

$$\begin{aligned} \text{sujeito a:} \quad & \mathbf{A}\boldsymbol{x} \geqq \boldsymbol{b}^1 \\ & \mathbf{B}\boldsymbol{x} = \boldsymbol{b}^2 \\ & \mathbf{C}\boldsymbol{x} \leqq \boldsymbol{b}^3 \end{aligned}$$

$$\boldsymbol{x} \in \{\mathbb{R}^n, \mathbb{R}^n_+, \mathbb{R}^n_-\} \textit{ ou variantes}$$

FORMA CANÔNICA:

$$\textit{max } \boldsymbol{c}'\boldsymbol{x} \quad \text{(FUNÇÃO OBJETIVO)}$$

$$\begin{aligned} \textit{sujeito a}: \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geqq \mathbb{0} \end{aligned}$$

Para podermos elaborar alguns argumentos geométricos sobre problemas lineares, precisaremos de algumas definições, apresentadas na próxima subseção.

0.1.4 Hiperplanos e Semi-espaços

Para melhor entendermos os problemas envolvidos num programa linear, serão necessários conhecimentos de Álgebra Linear, Vetores e Geometria e Cálculo. Para começar precisamos definir o conjunto de restrições lineares, que é um POLIEDRO: uma intersecção (finita) de semi-espaços e hiperplanos.

$$\text{Hiperplano } H_{\boldsymbol{c},z} = \{\boldsymbol{x} : \boldsymbol{c}'\boldsymbol{x} = z\}$$
$$\text{Semi-espaço } S_{\boldsymbol{c},z} = \{\boldsymbol{x} : \boldsymbol{c}'\boldsymbol{x} \leqq z\}$$

Chamamos de HIPERPLANO a generalização no $\mathbb{R}^n$ da noção que temos de uma reta no $\mathbb{R}^2$ e de um plano no $\mathbb{R}^3$, em que $\boldsymbol{c}'$ é um vetor não nulo pertencente ao $\mathbb{R}^n$ e z é um escalar. Um hiperplano divide uma região em dois SEMI-ESPAÇOS, cuja união formaria todo o $\mathbb{R}^n$.

Na figura abaixo, temos um poliedro, ou seja, a intersecção finita de semi-espaços e de hiperplanos:

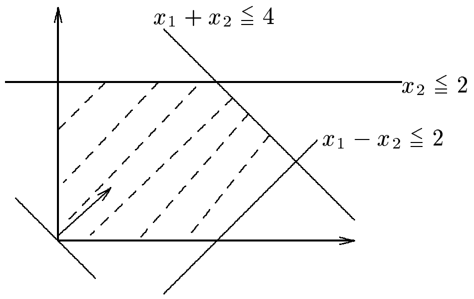


Figura 2: O poliedro é uma intersecção finita de semi-espaços e hiperplanos

Como exercício, represente os conjuntos $H_{\boldsymbol{c},1}$ e $H_{\boldsymbol{c},2}$ para $\boldsymbol{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Quais os pontos $\boldsymbol{x}^1 \in H_{\boldsymbol{c},1}$ e $\boldsymbol{x}^2 \in H_{\boldsymbol{c},2}$ que minimizam a distância destes hiperplanos à origem ?

Na subseção 1.4 apresentaremos algumas propriedades úteis à Programação Linear.

0.2 Exemplos

0.2.1 Exemplo 1: Transporte

Neste exemplo, temos n pontos de distribuição e n consumidores. Chamaremos de c_{ij} o custo relativo a levar o produto do ponto de distribuição i até o consumidor j. Como exemplo, vamos utilizar o de uma companhia que processa café em m fábricas e distribui o produto para n mercados. Vamos supor, neste caso, que o custo para transportar o café da fábrica i para o mercado j seja c_{ij}. Chamaremos de a_i a capacidade de produção da fábrica i, e de b_j a demanda do produto no mercado j. Temos que encontrar um x_{ij} que minimize o custo de transporte, sendo $i = 1, ..., m$ e $j = 1, ..., n$.

Temos, abaixo, uma representação gráfica deste problema:

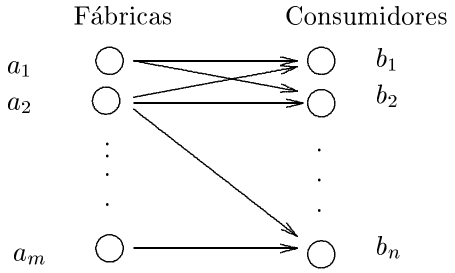


Figura 3: O problema do transporte

Como exercício sugerimos que o leitor tente modelar matematicamente este problema!!

0.2.2 Exemplo 2: Problema de produção

Neste exemplo, temos 2 tipos de produtos e de insumos. Chamaremos de a_{ij} a quantidade de insumo j no produto i. Os CUSTOS estarão relacionados ao problema se minimizar $\boldsymbol{c'x}$, e o LUCRO ao de maximizar $\boldsymbol{c'x}$.

Qual seria o lucro ? $\alpha+\beta$? E se trocarmos o problema para maximizar x_1+2x_2 ? Qual seria uma função objetivo $\boldsymbol{c}$ que resulte numa solução que não produza o produto tipo 2 ?

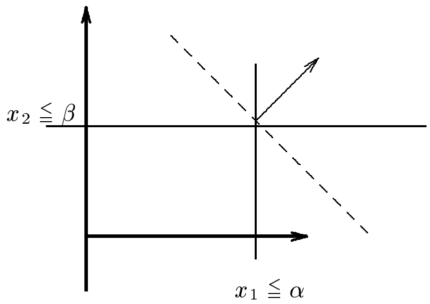


Figura 4: O problema de produção

0.2.3 Exemplo 3: Encontrando o vértice ótimo

Neste caso retomaremos o exemplo da figura 2, $\max 2x_1 + x_2$, sujeito à $\boldsymbol{x} \in P$, onde

$$P = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 \geqq 0,\ x_2 \geqq 0, x_2 \leqq 2,\ x_1 + x_2 \leqq 4, x_1 - x_2 \leqq 2\}.$$

Qual a solução para $\boldsymbol{c} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$? E para $\boldsymbol{c} = \alpha \begin{pmatrix} 2 \\ 1 \end{pmatrix}$?

Sejam $\boldsymbol{x}^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\boldsymbol{x}^2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$, $\boldsymbol{x}^3 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, $\boldsymbol{x}^4 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\boldsymbol{x}^5 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

Temos que $\boldsymbol{c}'\boldsymbol{x}^4 > \boldsymbol{c}'\boldsymbol{x}$, $\forall \boldsymbol{x} \in P$ e $\boldsymbol{x} \neq \boldsymbol{x}^4$! Por quê?

Se colocarmos $\boldsymbol{c} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$?

Temos $\boldsymbol{c}'\boldsymbol{x}^3 = 6 > \boldsymbol{c}'\boldsymbol{x}$, $\forall \boldsymbol{x} \in P$. Verifique!

O ponto de ótimo está sempre num vértice ? O que é um vértice ?

Pode existir ilimitação ? (apague as restrições $x_1 + x_2 \leqq 4$ e $x_1 - x_2 \leqq 2$)

0.3 Perguntas a serem respondidas

A partir do que foi examinado nos exemplos acima, podemos destacar várias questões a serem discutidas num curso de Programação Linear, a saber:

1. Como provar que vértices são os principais candidatos a ponto ótimo? *(estudo de convexidade)*
2. Como reconhecer um ponto ótimo ? E uma ilimitação ? *(condições de otimalidade)*
3. Como gerar vértices ? *(caracterização de vértices)*
4. Como automatizar a resolução ? *(algoritmo Simplex - depende de 2 e 3)*
5. O que ocorre quando mudamos “levemente” $\boldsymbol{c}$ ou $\boldsymbol{b}$? *(dual e análise de sensibilidade)*

Capítulo 1

Matrizes e Vetores

Neste texto faremos uma rápida revisão de conceitos fundamentais de matrizes e vetores. Recomendamos àqueles que tiverem dúvidas que consultem, por exemplo, a seção 1.5 de `Bertsekas-1997` ou o capítulo 2 de `Hadley-1962` (ou `Hadley-1982`, em Português). Para um estudo mais completo do assunto sugerimos textos de Álgebra Linear, como `Monteiro-1969` ou `Hoffman&Kunze-1970`.

1.1 Matrizes

Uma MATRIZ $\mathbf{A}_{m,n}$ com m linhas e n colunas (sobre o corpo dos reais) é uma tabela retangular de escalares $[a_{i,j}]$. Representando por a^i e por a_j, respectivamente, a i-ésima coluna e j-ésima linha de $\mathbf{A}$, temos

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \left[\ a^1 \mid a^2 \mid \cdots \mid a^n\ \right] = \begin{bmatrix} a_1 \\ \hline a_2 \\ \hline \vdots \\ \hline a_n \end{bmatrix}$$

1.1.1 Propriedades e conceitos

1. Sejam $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ matrizes do $\mathbb{R}^{m\times n}$ e $\lambda \in \mathbb{R}$, então:

   - $\mathbf{A} = \mathbf{B} \iff \langle a_{ij} = b_{ij}, \quad (i,j) \in \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}\rangle$
   - $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (COMUTATIVIDADE da soma)
   - $\overline{\mathbf{A}} = \lambda\mathbf{A} \iff \quad \langle \overline{a}_{ij} = \lambda a_{ij}, \quad (i,j) \in \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}\rangle$
   - MATRIZ NULA: é a matriz $\mathbf{O}$ cujas entradas são todas nulas
     $\mathbf{A} + \mathbf{O} = \mathbf{A} = \mathbf{O} + \mathbf{A}$ (ELEMENTO NEUTRO da soma)
   - TRANSPOSTA: a matriz transposta de $\mathbf{A}$ é $\mathbf{A}'$, cuja i-ésima linha (coluna) é a i-coluna (linha) de $\mathbf{A}$. Definindo $\overline{\mathbf{A}} := \mathbf{A}'$, tem-se
     $$\overline{a}_{ij} = a_{ji} \quad (\mathbf{A}_{m\times n} \quad \Longrightarrow \quad \overline{\mathbf{A}}_{n\times m})$$
     Usando os vetores linha ou colunas, temos que,

$$\mathbf{A}_{m\times n} = \left[a^1|\cdots|a^n\right] = \left[\begin{array}{c} a_1 \\ \hline \vdots \\ \hline a_m \end{array}\right] \iff \mathbf{A}' = \left[\begin{array}{c} (a^1)' \\ \hline \vdots \\ \hline (a^n)' \end{array}\right] = [(a_1)'|\cdots|(a_m)']$$

- Matriz identidade (de ordem n, $\mathbf{I}_n$):

$$\mathbf{I}_{n\times n} = \left[e^1|\cdots|e^n\right] = [\delta_{ij}], \text{ onde } \left\langle \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases} \right\rangle \quad (\delta \text{ de Kronecker})$$

2. Produto de matrizes

Sejam $\mathbf{A} \in \mathbb{R}^{m\times n}$ e $\mathbf{B} \in \mathbb{R}^{n\times p}$, se $\mathbf{C} := \mathbf{A}_{m\times n}\mathbf{B}_{n\times p}$, então

$$\mathbf{C} \in \mathbb{R}^{m\times p}, \quad \text{com } c_{ij} := \sum_{k=1}^{n} a_{ik}b_{kj}.$$

- sendo $\mathbf{O} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$, $\mathbf{O}_{t\times m}\mathbf{A} = \mathbf{O}_{t\times n}$ e $\mathbf{A}\mathbf{O}_{n\times t} = \mathbf{O}_{m\times t}$

- $\mathbf{AI} = \mathbf{A} = \mathbf{IA}$ (elemento neutro do produto)

Exercício 1 *Dados* $\mathbf{A} = \begin{bmatrix} 2 & 3 & \frac{1}{2} \\ -1 & 0 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ *e* $\mathbf{C} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix}$, *compute* $\mathbf{A}'\mathbf{B}$, $\mathbf{A}'\mathbf{B}$, $\mathbf{BC}$, $\mathbf{CB}$.

Note que o produto matricial não é comutativo: com matrizes quadradas, $\mathbf{AB}$ não necessariamente equivale a $\mathbf{BA}$.

Decomposição e produto: se tormarmos matrizes $\mathbf{A}_{m\times n}$ e $\mathbf{B}_{n\times t}$, tais que $m = r+s$, $n = p+q$, $t = u+v$ e

$$\mathbf{A}_{m\times n} := \left[\begin{array}{c|c} \mathbf{C}_{r\times p} & \mathbf{D}_{r\times q} \\ \hline \mathbf{E}_{s\times p} & \mathbf{F}_{s\times q} \end{array}\right] \begin{array}{l} \updownarrow r \\ \updownarrow s \end{array} \qquad \mathbf{B}_{n\times t} := \left[\begin{array}{c|c} \overline{\mathbf{C}}_{p\times u} & \overline{\mathbf{D}}_{p\times v} \\ \hline \overline{\mathbf{E}}_{q\times u} & \overline{\mathbf{F}}_{q\times v} \end{array}\right]$$

$$\longleftarrow p \longrightarrow\longleftarrow q \longrightarrow$$

então o produto de $\mathbf{AB}$ pode ser escrito como

$$\mathbf{AB} = \left[\begin{array}{c|c} \mathbf{C}\overline{\mathbf{C}} + \mathbf{D}\overline{\mathbf{E}} & \mathbf{C}\overline{\mathbf{D}} + \mathbf{D}\overline{\mathbf{F}} \\ \hline \mathbf{E}\overline{\mathbf{C}} + \mathbf{F}\overline{\mathbf{E}} & \mathbf{E}\overline{\mathbf{D}} + \mathbf{F}\overline{\mathbf{F}} \end{array}\right]_{m\times t}$$

Exercício 2 *Considerando as matrizes* $\mathbf{A}$ *e* $\mathbf{B}$ *acima definidas, tomando-se* $\overline{\mathbf{C}} := \mathbf{F}$, $\overline{\mathbf{D}} := -\mathbf{D}$, $\overline{\mathbf{E}} := -\mathbf{E}$ *e* $\overline{\mathbf{D}} := \mathbf{C}$:

(a) quais devem ser as dimensões r, p, s e q para que faça sentido o produto $\mathbf{AB}$ *?*

(b) supondo as matrizes $\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$ *e* $\mathbf{F}$ *simétricas, dê a forma reduzida (mais simplificada) de* $\mathbf{AB}$ *e* $\mathbf{BA}$.

1.1.2 Determinantes

Seja $\mathcal{P} = \{\text{conjunto de permutações de } (1,2,\ldots,n)\}$ e $s_{(p)}$ o número de inversões de $p \in \mathcal{P}$, sendo que

$$s_{(p)} = \#\{(p_i, p_j) : \quad i < j \quad \text{e} \quad p_i > p_j\}.$$

DETERMINANTE é uma função real definida sobre matrizes quadradas, $\det(\mathbf{A}) : \mathbb{R}^{n\times n} \to \mathbb{R}$, definido por

$$\det(\mathbf{A}) = |\mathbf{A}| := \sum_{p\in\mathcal{P}} (-1)^{s_{(p)}} a_{1p_1} a_{2p_2} \ldots a_{np_n}, \text{ sendo } \mathbf{A} \in \mathbb{R}^{n\times n}.$$

Algumas propriedades de determinantes: fixando-se $\mathbf{A} \in \mathbb{R}^{n\times n}$

- $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$, quaisquer que sejam $\mathbf{A}$ e $\mathbf{B}$;
- $\det(\alpha\mathbf{A}) = \alpha \det(\mathbf{A})$, qualquer que seja $\alpha \in \mathbb{R}$;
- Trocar pares de linhas ou de colunas inverte o sinal do determinante, por exemplo, se $\mathbf{B}$ é obtida de uma matriz $\mathbf{A}$ trocando-se a linha 1 com linha 2, então $\det(\mathbf{B}) = -\det(\mathbf{A})$;
- Somar multiplos de linhas (ou de colunas) não altera o determinante, isto é, se construirmos $\overline{\mathbf{A}}$ a partir de $\mathbf{A}$ de tal forma que
$$\overline{\mathbf{a}}^i := \begin{cases} \boldsymbol{a}^k + \alpha a^j, & i = k \\ a^i & i \neq k \end{cases}, \text{ então } \det(\overline{\mathbf{A}}) = \det(\mathbf{A});$$
- Definindo $\boldsymbol{A}_{ij}$ como o COFATOR de a_{ij}, sendo
$$\boldsymbol{A}_{ij} = (-1)^{i+j} \det \mathbf{C}, \text{ sendo } \mathbf{C} \in \mathbb{R}^{n-1\times n-1} \text{ obtida de } \mathbf{A} \text{ eliminando-se a linha } \boldsymbol{a}_i \text{ e a coluna } a^j,$$

  Podemos demonstrar, por indução na ordem n da matriz, que

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}\boldsymbol{A}_{ij} = \sum_{i=1}^{n} a_{ji}\boldsymbol{A}_{ji}, \quad \forall i \in \{1,2,\ldots,n\}. \tag{1.1}$$

1.1.3 Matriz Adjunta de A

Teorema 1.1 *Se uma matriz* $\mathbf{D}$ *contém duas linhas (colunas) iguais, então* $\det(\mathbf{D}) = 0$.

Demonstração Como $\mathbf{D}$ tem uma linha (coluna) repetida, sem perda de generalidade, podemos usar uma matriz auxiliar $\mathbf{A}$ para construir $\mathbf{D}$. Sendo d_i e d_k as linhas repetidas, podemos escrever $\mathbf{D}$ como no exemplo ao lado, ou seja, $d_j := a_j, \quad j \neq k$ e $d_k := a_i$. Deste modo, a k-ésima linha da matriz de cofatores de $\mathbf{D}$ coincide a corresponde linha de cofatores de $\mathbf{A}$,

$$\mathbf{D} := \begin{bmatrix} \boldsymbol{a}_1 \\ \hline \vdots \\ \hline \boldsymbol{a}_i \\ \hline \vdots \\ \hline \boldsymbol{a}_i \\ \hline \vdots \\ \hline \boldsymbol{a}_n \end{bmatrix}$$

$$\mathbf{D}_{kj} = \boldsymbol{A}_{kj} \tag{1.2}$$

Daí podemos concluir que

$$\det(\mathbf{D}) = \sum_{j=1}^{n} d_{kj}\mathbf{D}_{kj} \overset{d_k=a_i}{=} \sum_{j=1}^{n} a_{ij}\mathbf{D}_{kj} \overset{(1.2)}{=} \sum_{j=1}^{n} a_{ij}\boldsymbol{A}_{kj} = 0.$$

finalizando a demonstração. ∎

Este teorema leva a um resultado interessante, que induz a definição de uma nova matriz que é construida a partir de uma matriz quadrada qualquer: dada $\mathbf{A} \in \mathbb{R}^{n\times n}$, a matriz $\mathrm{Adj}(\mathbf{A})$ obtida transpondo-se os cofatores de $\mathbf{A}$ é a MATRIZ ADJUNTA DE $\mathbf{A}$, ou seja, se $\mathbf{B} := \mathrm{Adj}(\mathbf{A})$, então $b_{ij} = \boldsymbol{A}_{ji}$.

Teorema 1.2 *Seja* $\mathbf{A} \in \mathbb{R}^{n\times n}$, *então*

$$\mathbf{A}\mathrm{Adj}(\mathbf{A}) = \det(\mathbf{A})\mathbf{I}.$$

Demonstração Fixando-se qualquer linha (coluna) i de $\mathbf{A}$, podemos construir uma matriz $\mathbf{D}$ repetindo-se a linha (coluna) i como no teorema anterior. Deste modo,

$$\sum_{j=1}^{n} a_{ij}\boldsymbol{A}_{kj} = \sum_{j=1}^{n} a_{ji}\boldsymbol{A}_{jk} = 0, \qquad \text{para } i \neq k. \tag{1.3}$$

Além disso, da identidade (1.1), temos que

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}\boldsymbol{A}_{ij} = \sum_{j=1}^{n} a_{ji}\boldsymbol{A}_{ji},$$

e usando a equação (1.3) acima, segue que

$$\sum_{j=1}^{n} a_{ij}\boldsymbol{A}_{kj} = \sum_{j=1}^{n} a_{ji}\boldsymbol{A}_{jk} = |\mathbf{A}|\delta_{ki} \quad (\delta_{ki} = 0 \text{ quando } i \neq k). \tag{$\star$}$$

Definindo a matriz $\mathbf{B} := \mathrm{Adj}(\mathbf{A})$ $(b_{ij} = a_{ji})$, segue que

$$a_i\boldsymbol{b}k{=}\sum_{j=1}^{n} a_{ij}b_{jk} = \sum_{j=1}^{n} a_{ij}\boldsymbol{A}_{kj} \overset{(\star)}{=} |\mathbf{A}|\delta_{ki},$$

ou seja, $\mathbf{A}\mathrm{Adj}(\mathbf{A}) = |\mathbf{A}|\mathbf{I}_n$. ∎

1.1.4 Inversa

$\mathbf{B} = \mathbf{A}^{-1} \iff \mathbf{AB} = \mathbf{I} = \mathbf{BA}$

- $\mathbf{A}$ é INVERSÍVEL (NÃO SINGULAR) $\iff \det(\mathbf{A}) \neq 0$
- $\mathbf{B} = \mathbf{A}^{-1} = \dfrac{1}{|\mathbf{A}|}\mathrm{Adj}(\mathbf{A})$
- Se $|\mathbf{A}| \neq 0$, então $\mathbf{A}^{-1}\boldsymbol{b}$ resolve $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, qualquer que seja $\boldsymbol{b} \in \mathbb{R}^n$.

1.2 Vetores

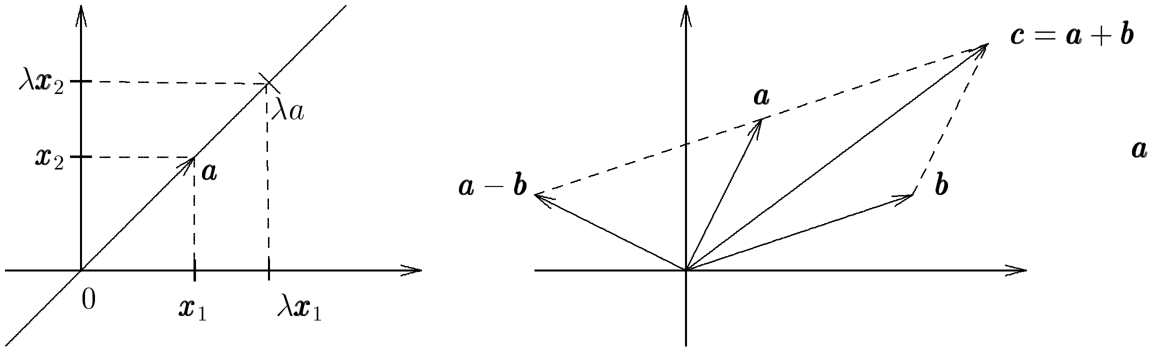
- Espaço gerado
  O espaço gerado por $\boldsymbol{a}$ é $S = \{\boldsymbol{x} \in \mathbb{R}^n : \quad \boldsymbol{x} = \lambda \boldsymbol{a}, \lambda \in \mathbb{R}\}$.

  Exercício 3 *Prove que S é um espaço vetorial (i.é., valem $\boldsymbol{x}+\boldsymbol{y} \in S$ e $\lambda\boldsymbol{x} \in S$ para todos $\boldsymbol{x}$ e $\boldsymbol{y}$ no $\mathbb{R}^n$ e $\lambda \in \mathbb{R}$).*

  Em geral, se $V := \{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k\}$ são vetores do $\mathbb{R}^n$, o espaço gerado por V é

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : \quad \boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}^i, \boldsymbol{\lambda} \in \mathbb{R}^n\}.$$

- Soma de vetores $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} \iff c_i = a_i + b_i$.



$$\boldsymbol{a} = \sum_{i=1}^{n} a_i \boldsymbol{e}^i$$

Figura 1.1: Vetores, produto por escalar e soma

- Produto escalar
  $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}'\boldsymbol{b} = \sum_{i=1}^{n} a_i b_i$

  Propriedades:

  - $\langle \boldsymbol{a}, \boldsymbol{a} \rangle \geqq 0$ e $\big\langle\ \langle \boldsymbol{a}, \boldsymbol{a} \rangle = 0 \iff \boldsymbol{a} = \mathbb{0}\ \big\rangle$ $\qquad \boxed{\boldsymbol{a}'\boldsymbol{a} \geqq \mathbb{0}}$
  - $\langle \boldsymbol{a}, \boldsymbol{b}+\boldsymbol{c} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} \rangle + \langle \boldsymbol{a}, \boldsymbol{c} \rangle$ $\qquad \boxed{\boldsymbol{a}'(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a}'\boldsymbol{b} + \boldsymbol{a}'\boldsymbol{c}}$
  - $\langle \lambda\boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{a}, \lambda\boldsymbol{b} \rangle = \lambda \langle \boldsymbol{a}, \boldsymbol{b} \rangle$ $\qquad \boxed{(\lambda\boldsymbol{a}')\boldsymbol{b} = \boldsymbol{a}'(\lambda\boldsymbol{b})}$
  - $\langle \mathbf{A}\boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{a}, \mathbf{A}'\boldsymbol{b} \rangle$ $\qquad \boxed{(\mathbf{A}\boldsymbol{a})'\boldsymbol{b} = \boldsymbol{a}'\mathbf{A}'\boldsymbol{b})}$
  - Distância euclidiana: $d(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a}-\boldsymbol{b}\| = \langle \boldsymbol{a}-\boldsymbol{b}, \boldsymbol{a}-\boldsymbol{b} \rangle^{\frac{1}{2}}$
  - Desigualdade de Cauchy-Schwarz: $|\langle \boldsymbol{a}, \boldsymbol{b} \rangle| \leqq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$ $\qquad \boxed{\langle \boldsymbol{a}-\lambda\boldsymbol{b}, \boldsymbol{a}-\lambda\boldsymbol{b} \rangle \geqq 0}$

- Dependência linear
  Um conjunto de vetores do $\mathbb{R}^n$, $V := \{a^1, a^2, \ldots, a^n\}$, é linearmente dependente (l.d.) se, e somente se, $\exists \boldsymbol{\lambda} \in \mathbb{R}^n : \quad \boldsymbol{\lambda} \neq \mathbb{0}$ e $\sum_{i=1}^{n} \lambda_i a^i = \mathbb{0}$. Ou seja, V é l.d. se, e somente se o vetor nulo pode ser escrito como uma combinação linear (a coeficiente não nulos) de V.

  Se o conjunto $\{a^i\}_{i=1}^{n}$ não é l.d., é linearmente independente (l.i.).

Exercício 4 *Prove que* $\{a^1, a^2, \lambda a^1\}$ *é l.d. se* $\lambda \neq 0$.

Solução: $0 = \alpha_1 a^1 + \alpha_2 a^2 + \alpha_3 \lambda a^1 = (\alpha_1 + \lambda\alpha_3)a^1 + \alpha_2 a^2$

Seja $\alpha_2 = 0$ e $\alpha_3 = -\dfrac{\alpha_1}{\lambda}$. □

Exercício 5 *Dado* $V := \{a^1, a^2, \ldots, a^n\}$*, seja* $\mathbf{A}$ *a matriz cuja i-ésima coluna é* a^i*. Prove (pela definição) que, se* V *é l.i., então a transformação* $T : \mathbb{R}^n \to \mathbb{R}^n$ *definada por* $T(\boldsymbol{x}) := \mathbf{A}\boldsymbol{x}$ *é injetora* $(\boldsymbol{x} \neq \boldsymbol{y} \quad \Rightarrow \quad T(\boldsymbol{x}) \neq T(\boldsymbol{y}))$.

Alguns resultados importantes: considerando o espaço $\mathbb{R}^n$, seja $V := \{a^1, a^2, \ldots, a^n\}$ e $\mathbf{A} := [a^1|a^2|\ldots|a^n]$.

1. Se V é l.i., então $|\mathbf{A}| \neq 0$.
2. Se $|\mathbf{A}| \neq 0$, então V é l.i.
3. Se V é l.i., então V gera o $\mathbb{R}^n$.

Ou seja, se V tem n vetores doe $\mathbb{R}^n$ e é l.i. então forma uma base do $\mathbb{R}^n$ como discutido a seguir.

1.3 Bases

Sendo E um espaço vetorial qualquer:

$\{a^1, a^2, \ldots, a^n\}$ gera $E^n \iff \forall \boldsymbol{x} \in E^n, \exists \boldsymbol{\alpha} \in \mathbb{R}^n : \quad \alpha_1 a^1 + \ldots + \alpha_n a^n = \boldsymbol{x}$;

$\{a^1, a^2, \ldots, a^n\}$ é BASE de $E^n \iff$ gera E^n e são l.i.

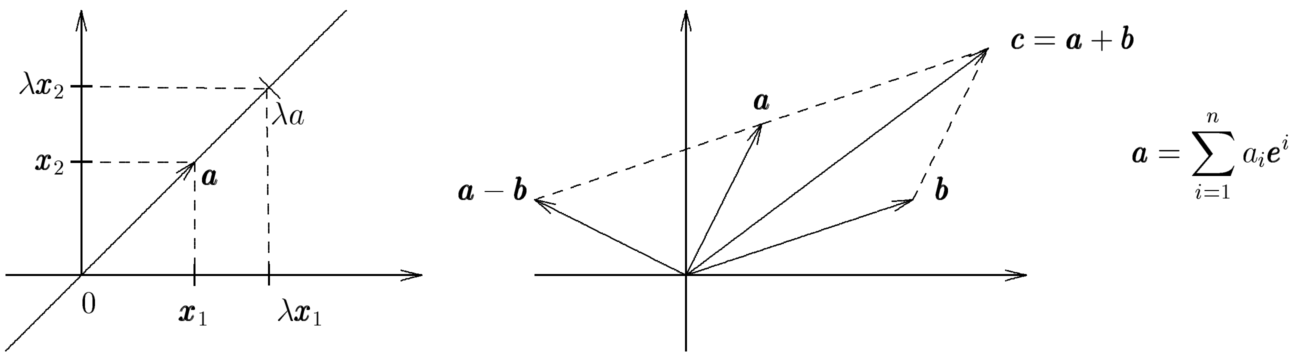


Figura 1.2: Exemplo de vetores que geram $\mathbb{R}^2$

1.3.1 Atualização de base

Se $\{a^1, \ldots, a^n\}$ é base do $\mathbb{R}^n$, então para qualquer que seja o vetor $\boldsymbol{x}$ do $\mathbb{R}^n$ existe uma única representação do vetor $\boldsymbol{x}$ nesta base:

$$\forall \boldsymbol{x} \in E^n, \exists! \boldsymbol{\lambda} \in \mathbb{R}^n : \quad \boldsymbol{x} = \sum_{i=1}^{n} \lambda_i a^i$$

(por contradição, suponha $\exists\overline{\boldsymbol{\lambda}} \neq \boldsymbol{\lambda}$, então $\mathbb{0} = \boldsymbol{x} - \boldsymbol{x} = \sum_{i=1}^{n}(\overline{\lambda}_i - \lambda_i)a^i$, com $\overline{\boldsymbol{\lambda}} - \boldsymbol{\lambda} \neq \mathbb{0} \rightarrow|\leftarrow$).

Teorema 1.3 *Seja* $V := \{a^1, a^2, \ldots, a^n\}$ *uma base de* $\mathbb{R}^n$ *e* $\boldsymbol{\beta} = \sum_{j=1}^{n} \alpha_j a^j$. *Se* $\alpha_i \neq 0$, *então* $\{a^j\}_{j\neq i} \bigcup \{\boldsymbol{\beta}\}$ *é l.i.*

Demonstração Por simplicidade, suporemos que seja $i = n$, então qualquer que seja $\boldsymbol{\lambda} \in \mathbb{R}^n$, como V é l.i., se

$$\mathbb{0} = \sum_{j=1}^{n-1} \lambda_j a^j + \lambda_n \boldsymbol{\beta} = \sum_{j=1}^{n-1} \lambda_j a^j + \lambda_n \left(\sum_{j=1}^{n} \alpha_j a^j \right) = \sum_{j=1}^{n-1} (\lambda_j + \lambda_n \alpha_j) a^j + \lambda_n \alpha_n a^n,$$

então,

$$\lambda_j + \lambda_n \alpha_j = 0, \quad j \in \{1, 2, \ldots, n-1\} \text{ e} \qquad (\star)$$
$$\lambda_n \alpha_n = 0. \qquad (\star\star)$$

Como $\alpha_n \neq 0$, por hipótese, segue que $\lambda_n = 0$ em $(\star\star)$, que substituído em $(\star)$, resulta em $\lambda_j = 0$, $j \in \{1, 2, \ldots, n-1\}$. Ou seja, $\boldsymbol{\lambda} = \mathbb{0}$ e portanto $\{a^j\}_{j\neq i} \bigcup \{\boldsymbol{\beta}\}$ é l.i. como queríamos demonstrar. ■

Como veremos mais tarde, este é um resultado importante para o método SIMPLEX, principalmente por responder a questão seguinte:

sendo B um conjunto de índices tais que $\{a^1, \ldots, \boldsymbol{a}^m\}$ é um base do $\mathbb{R}^m$ e $\boldsymbol{\beta} := a^k$ $(k \notin B)$, como escrever $\boldsymbol{b} \in \mathbb{R}^m$ na base $\{a^1, \ldots, \boldsymbol{a}^{m-1}, \boldsymbol{\beta}\}$?

Vamos aplicar o teorema acima em um caso em que desejamos "trocar" de base: sejam $B := \{1, 2, \ldots, m-1, m\}$ e $\overline{B} := \{1, 2, \ldots, m-1, k\}$ índices de colunas de $\mathbf{A}$ que formam base, sendo $a^k = \sum_{i\in B} \alpha_i a^i$ e $\alpha_m \neq 0$,

conhecemos uma solução $\boldsymbol{x} = \boldsymbol{\lambda}$ para $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, com $\lambda_i = 0$, para todo $i \notin B$, e
desejamos uma solução $\boldsymbol{x} = \boldsymbol{\beta}$, com $\beta_i = 0$, para todo $i \notin \overline{B}$.

Seja $\boldsymbol{\lambda} \in \mathbb{R}^n$, tal que $\boldsymbol{b} = \sum_{i\in B} \lambda_i a^i = \sum_{i=1}^{m} \lambda_i a^i = \sum_{i=1}^{m-1} \lambda_i a^i + \lambda_m a^m$. Como $a^k = \sum_{i\in B} \alpha_i a^i = \sum_{i=1}^{m} \alpha_i a^i$, com $\alpha_m \neq 0$, segue que $a^m = \frac{1}{\alpha_m} a^k - \frac{1}{\alpha_m} \sum_{i=1}^{m-1} \alpha_i a^i$. Logo,

$$\boldsymbol{b} = \sum_{i=1}^{m-1} \lambda_i \alpha_i + \lambda_m \left(\frac{1}{\alpha_m} a^k - \frac{1}{\alpha_m} \sum_{i=1}^{m-1} \alpha_i a^i \right) = \sum_{i=1}^{m-1} \left(\lambda_i - \frac{\lambda_m}{\alpha_m} \alpha_i \right) a^i + \frac{\lambda_m}{\alpha_m} a^k. \tag{1.4}$$

Note que $\langle a^k = \sum_{i\in B} \alpha_i a^i = \mathbf{A}_B \boldsymbol{\alpha}_B \rangle \Leftrightarrow \langle \boldsymbol{\alpha}_B = \mathbf{A}_B^{-1} a^k \rangle$ (hipótese: $\{a^i\}_{i\in B}$ base). Deste modo, uma vez conhecida a solução $\boldsymbol{x} = \boldsymbol{\lambda}$ para o sistema $\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \mathbf{A}_B \boldsymbol{x}_B$, podemos encontrar a solução $\boldsymbol{x} = \boldsymbol{\beta}$ para o sistema $\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \mathbf{A}_{\overline{B}} \boldsymbol{x}_{\overline{B}}$ de modo "rápido": basta definir $\boldsymbol{\beta}$ segundo a equação (1.4),

$$\beta_i := \lambda_i - \tfrac{\lambda_m}{\alpha_m} \alpha_i, \ i \in B \setminus \{m\},$$
$$\beta_m := \tfrac{\lambda_m}{\alpha_m}.$$

Esta operação, se aplicada na matriz todo, corresponde a um PIVOTEAMENTO (ou ESCALONAMENTO). Na figura abaixo, representamos graficamente a operação de escalonamento acima: vamos escalonar a matriz $\mathbb{R}^{n\times n+2}$,

$[\mathbf{A}_B|\cdots|\mathbf{a}k|\cdots|\boldsymbol{b}]$ para obtermos $[e^1|\cdots|e^m|\cdots|\mathbf{A}_B^{-1}a^k|\cdots|\mathbf{A}_B^{-1}\boldsymbol{b}]$ e depois $[e^1|\cdots|e^{m-1}|\cdots|e^m|\cdots|\mathbf{A}_{\overline{B}}^{-1}\boldsymbol{b}]$, sendo

$$\left\langle \begin{array}{l} \mathbf{A}_B\boldsymbol{\lambda} = \boldsymbol{b} \iff \boldsymbol{\lambda} = \mathbf{A}_B^{-1}\boldsymbol{b} \\ \mathbf{A}_B\boldsymbol{\alpha} = a^k \iff \boldsymbol{\alpha} = \mathbf{A}_B^{-1}a^k \\ \mathbf{A}_{\overline{B}}\boldsymbol{\beta} = \boldsymbol{b} \iff \boldsymbol{\beta} = \mathbf{A}_{\overline{B}}^{-1}\boldsymbol{b} \end{array} \right\rangle.$$

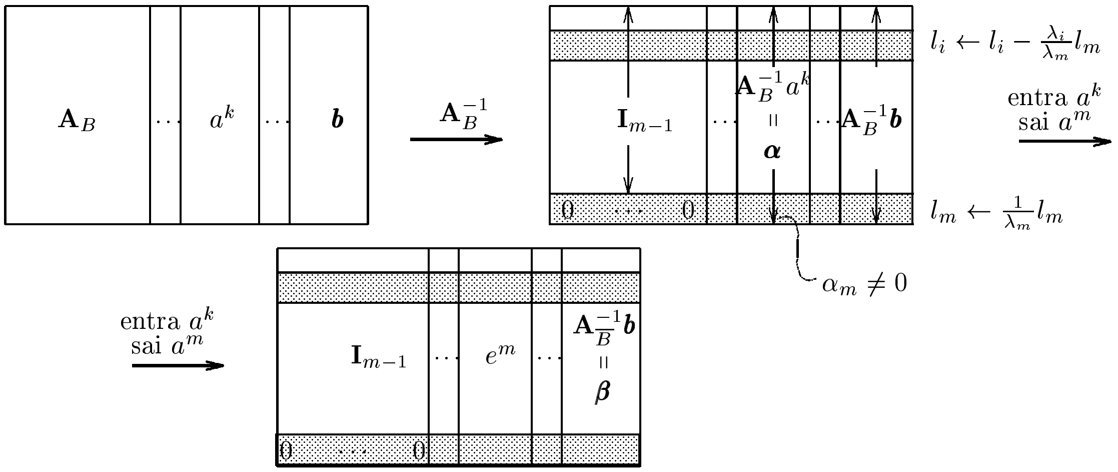


Figura 1.3: Duas "etapas" de escalonamento, para mudança de base

1.3.2 Soluções básicas e degeneradas

O POSTO de uma matriz $\mathbf{A}_{m\times n}$ ($r(\mathbf{A})$ "rank") é o número maximal de colunas l.i. da matriz ($\{a^i\}_{i=1}^n$).

Quando $r(\mathbf{A}) = m$, dizemos que o posto é MÁXIMO ou PLENO.

Se $\mathbf{A}_{m\times n}$ e $m < n$, quantos subconjuntos l.i. de $\{a^i\}_{i=1}^n$ podemos conseguir?

$$\boxed{\begin{array}{ll} \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n-1} + \binom{n}{n} = 2^n : & (\alpha_1, \alpha_1, \ldots, \alpha_n) \quad \alpha_i \in \{0,1\} \\ & \alpha_i = 1 \quad \Longrightarrow \quad \text{coluna } i \text{ está no subconjunto} \end{array}}$$

Usando os resultados da seção anterior é fácil mostrar que: se $\mathbf{A}_{m\times n}$ tem posto máximo, então para todo $\boldsymbol{b} \in \mathbb{R}^n$, existe $B \subseteq \{1, 2, \ldots, n\}$ e um único $\boldsymbol{x} \in \mathbb{R}^m$ tais que $\#B = m$, $\{a^i\}_{i\in B}$ é l.i. e $\sum_{i\in B} x_i a^i = \boldsymbol{b}$.

Neste caso pode-se re-escrever o sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ na forma

$$\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = [\mathbf{A}_B|\mathbf{A}_N] \left[\frac{\boldsymbol{x}_B}{\boldsymbol{x}_N}\right] = \mathbf{A}_B\boldsymbol{x}_B + \mathbf{A}_N\boldsymbol{x}_N,$$

que equivale a forma: $\displaystyle \boldsymbol{b} = \sum_{i=1}^n x_i a^i = \sum_{i\in B} x_i a^i + \sum_{i\in N} x_i a^i, \quad B \cup N = \{1, 2, \ldots, n\}.$

A partir desta identidade devemos notar que: $\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{x}_B = \mathbf{A}_B^{-1}\boldsymbol{b} + \mathbf{A}_B^{-1}\mathbf{A}_N\boldsymbol{x}_N$.

Exemplo 1.3.1 *Supondo $\{a^i\}_{i\in B}$, mostre que: qualquer que seja $\boldsymbol{x} \in \mathbb{R}^n$*

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{x}_B = \mathbf{A}_B^{-1}\boldsymbol{b} + \mathbf{A}_B^{-1}\mathbf{A}_N\boldsymbol{x}_N \tag{1.5}$$

Da identidade (1.5), é possível notar a maior facilidade de obtenção de solução para $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ a partir da expressão à direita: basta "chutar" $\boldsymbol{x}_N = \mathbb{0}$ e pegar o $\boldsymbol{x}_B$ dado pela expressão.

Para qualquer sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, se $\{a^i\}_{i\in B}$ é l.i., então $\boldsymbol{x}$, definido por $\boldsymbol{x}_B = \mathbf{A}_B^{-1}\boldsymbol{b}$ e $\boldsymbol{x}_N = \mathbb{0}$, é denominada SOLUÇÃO BÁSICA.

Se, para algum $i \in B$, $x_i = 0$ então a solução básica é DEGENERADA.

Uma consequência relevante sobre soluções degeneradas, junto com o teorema (1.3), é a possibilidade de existência de diferentes bases para uma mesma solucão. Isto é, podem existirem bases B e $\overline{B}$ tais que $B \neq \overline{B}$, mas $\boldsymbol{x} = \overline{\boldsymbol{x}}$, sendo $\langle \boldsymbol{x}_B := \mathbf{A}_B^{-1}\boldsymbol{b} \wedge \boldsymbol{x}_N := \mathbb{0}\rangle$ e $\langle \boldsymbol{x}_{\overline{B}} := \mathbf{A}_{\overline{B}}^{-1}\boldsymbol{b} \wedge \boldsymbol{x}_{\overline{N}} := \mathbb{0}\rangle$.

Podemos apresentar um exemplo disso usando o exemplo discutido na seção 1.3.1.

1.3.3 Espaços e transformações vetoriais

Dizemos que $\mathbf{S}$ é ESPAÇO VETORIAL se, e somente se $\forall(x,y) \in \mathbf{S}^2$ e $\lambda \in \mathbb{R}$, tivermos

$$\lambda\boldsymbol{x} \in \mathbf{S} \quad \text{e} \quad \boldsymbol{x} + \boldsymbol{y} \in \mathbf{S}. \tag{1.6}$$

Um SUBESPAÇO de $\mathbf{S}$ é qualquer subconjunto $\overline{\mathbf{S}}$ de $\mathbf{S}$ fechado em relação à soma e ao produto por escalar, i.é, se para todo elemento de $\overline{\mathbf{S}}$ vale a equação (1.6) - trocando-se $\mathbf{S}$ por $\overline{\mathbf{S}}$.

Teorema 1.4 *Se $\mathbf{S}$ é um subspaço qualquer, então o vetor nulo pertence a $\mathbf{S}$.* $\boxed{\lambda = 0}$

Exemplo 1.3.2 *Uma reta passando pela origem é subespaço vetorial do $\mathbb{R}^n$.*

Sejam $\boldsymbol{x}$ e $\boldsymbol{y}$ da reta r e $\lambda \in \mathbb{R}$, como a reta tem dimensão 1, $\boldsymbol{y} = \alpha\boldsymbol{x}$, para algum $\boldsymbol{x} \in \mathbb{R}^n$, logo $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \alpha\boldsymbol{x} = (1+\alpha)\boldsymbol{x} \in r$ e $\lambda\boldsymbol{x} \in r$

Exercício 6 *Como gerar o subespaço do $\mathbb{R}^n$ tal que $x_2 = x_3$?*

Solução: Basta tomar $\boldsymbol{x} := x_1e^1 + x_2(e^2 + e^3) + x_4e^4 + \ldots + x_ne^n$, para todo $(x_1, x_2, x_4, x_5, \ldots, x_n) \in \mathbb{R}^{n-1}$. □

Sejam E e F espaços vetoriais. Então

$$T: E \to F \ \text{ tal que } \ \begin{array}{ll} T(\boldsymbol{x}+\boldsymbol{y}) = T(\boldsymbol{x}) + T(\boldsymbol{y}), & \forall(x,y) \in E^2 \\ T(\lambda\boldsymbol{x}) = \lambda T(\boldsymbol{x}), & \forall\lambda \in \mathbb{R} \end{array}$$

Assim, uma matriz $\mathbf{A}_{m\times n}$ pode ser encarada como uma transformação linear do $\mathbb{R}^n$ em $\mathbb{R}^m$. Mais que isso, qualquer transformação linear do $\mathbb{R}^n$ em $\mathbb{R}^m$ é dada por uma matriz m por n.

Exercício 7 *Quando uma tranformação linear tem inversa ? (sob quais condições)*

1.4 Retas e hiperplanos

Para um dado $\boldsymbol{c} \in \mathbb{R}^n$ e um $\alpha \in \mathbb{R}$,

$$H_{\boldsymbol{c},\alpha} = \{\boldsymbol{x} \in \mathbb{R}^n : \quad \boldsymbol{c}'\boldsymbol{x} = \alpha\} \ \text{ é um HIPERPLANO do } \mathbb{R}^n.$$

No $\mathbb{R}^2$, dois vetores l.i. definem unicamente um hiperplano. Podemos escrever este hiperplano de modos, diretamente como os $\boldsymbol{x}$ tais que $\boldsymbol{c}'\boldsymbol{x} = \alpha$ ou como a soma vetorial de um ponto com $n-1$ vetores l.i., como indicada a seguir.

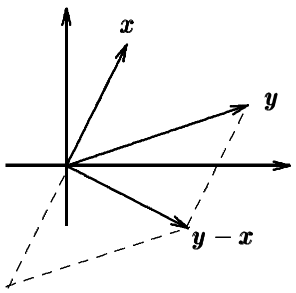


Figura 1.4: Linhas e hiperplanos.

Seja $\boldsymbol{x}_\lambda := \boldsymbol{x} + \lambda(\boldsymbol{y}-\boldsymbol{x})$. Se tomarmos um $\boldsymbol{c} \in \mathbb{R}^2$ tal que $\boldsymbol{c}'(\boldsymbol{y}-\boldsymbol{x}) = 0$, então $\exists \alpha \in \mathbb{R} : \alpha = \boldsymbol{c}'\boldsymbol{x}_\lambda,\ \forall \lambda \in \mathbb{R}$.

$$\boxed{\begin{aligned} \alpha &= \boldsymbol{c}'\boldsymbol{x} \\ \boldsymbol{c}'\boldsymbol{x}_\lambda = \boldsymbol{c}'\boldsymbol{x} + \lambda \boldsymbol{c}'(\boldsymbol{y}-\boldsymbol{x}) &= \boldsymbol{c}'\boldsymbol{x} \end{aligned}}$$

O argumento acima pode ser generalizado para o $\mathbb{R}^n$, tomando-se n vetores l.i. $\{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n\}$ e definindo, por exemplo, $\boldsymbol{x}_\lambda := \boldsymbol{x}^1 + \lambda(\boldsymbol{x}^1 - \boldsymbol{x}^{i+1})$, para cada $i \in \{1, 2, \ldots, n-1\}$.

Assim, definindo os vetores $\boldsymbol{h}^1 := \boldsymbol{x}^1 - \boldsymbol{x}^2$, $\boldsymbol{h}^2 := \boldsymbol{x}^1 - \boldsymbol{x}^3$, até $\boldsymbol{h}^{n-1} := \boldsymbol{x}^1 - \boldsymbol{x}^n$, é fácil ver[1]que existe um $\boldsymbol{c} \neq \mathbb{0}$, para o qual, $\boldsymbol{c}'\boldsymbol{h}^i = 0$, para todo $i \in \{1, 2, \ldots, n-1\}$. Então tomando $\alpha := \boldsymbol{c}'\boldsymbol{x}^1$, tem-se que

$$H_{\boldsymbol{c},\alpha} = \boldsymbol{x}^1 + \lambda_1 \boldsymbol{h}^1 + \lambda_2 \boldsymbol{h}^2 + \cdots + \lambda_{n-1}\boldsymbol{h}^{n-1}, \quad \text{para todo } \boldsymbol{\lambda} \in \mathbb{R}^{n-1}.$$

Exercício 8 *Mostre que, de fato, é possível construir um tal* $\boldsymbol{c}$ *(*$\boldsymbol{c}'\boldsymbol{h}^i = 0,\ \forall i \in \{1, 2, \ldots, n-1\}$*).*

Dica: Note que $\langle \boldsymbol{c}'\boldsymbol{h}^i = 0,\ \forall i \in \{1, 2, \ldots, n-1\} \rangle$ equivale ao sistema linear $\mathbf{H}'\boldsymbol{c} = \mathbb{0}$, sendo que $\mathbf{H}' \in \mathbb{R}^{n-1 \times n}$ e $\boldsymbol{h}^i := \boldsymbol{x}^1 - \boldsymbol{x}^{i+1}$.

Seja $H_{\boldsymbol{c},\alpha} = \{\boldsymbol{x} \in R^n : \boldsymbol{c}'\boldsymbol{x} = \alpha\}$ um hiperplano do $\mathbb{R}^n$. Vamos comparar H_α e $H_{2\alpha}$ para $\boldsymbol{c}$ fixo. Qual a relação ?

$H_{2\alpha} = \{2\boldsymbol{x} : \boldsymbol{x} \in H_\alpha\}$, pois $\boldsymbol{c}'(2x) = 2\boldsymbol{c}'\boldsymbol{x} = 2\alpha$
$H_{2\alpha} = \{\boldsymbol{x} + (\boldsymbol{h}^2 - \boldsymbol{h}^1) : \boldsymbol{x} \in H_\alpha$, pois $\boldsymbol{c}'(\boldsymbol{x} + \boldsymbol{h}^2 - \boldsymbol{h}^1) = \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{c}'\boldsymbol{h}^2 - \boldsymbol{c}'\boldsymbol{h}^1 = \alpha + 2\alpha - \alpha = 2\alpha$

$\boldsymbol{z} \in H_0 : \|\boldsymbol{z}\| = 1$

$H_\alpha = \{\boldsymbol{h}^1 + \boldsymbol{z} : \boldsymbol{z} \in H_0\} = \{\boldsymbol{h}^1 + \beta\boldsymbol{z},\ \beta \in \mathbb{R}\}$
$H_{2\alpha} = \{\boldsymbol{h}^2 + \boldsymbol{z} : \boldsymbol{z} \in H_0\} = \{\boldsymbol{h}^2 + \beta\boldsymbol{z}, \beta \in \mathbb{R}\}$.

Exercício 9 *Se* $\boldsymbol{c}$ *e* $\overline{\boldsymbol{c}}$ *são l.d. (e portanto* $\overline{\boldsymbol{c}} = \lambda\boldsymbol{c}$*, com* $\lambda \neq 0$*), então*

$$\textit{para todo } \alpha \in \mathbb{R}, \quad \textit{existe } \overline{\alpha} \in \mathbb{R} \textit{ para o qual } \{\boldsymbol{x} \in \mathbb{R}^n : \quad \boldsymbol{c}'\boldsymbol{x} = \alpha\} = \{\boldsymbol{x} \in \mathbb{R}^n : \quad \overline{\boldsymbol{c}}'\boldsymbol{x} = \overline{\alpha}\}.$$

Solução: $\overline{\alpha} = \overline{\boldsymbol{c}}'\boldsymbol{x} = (\lambda\boldsymbol{c})'\boldsymbol{x} = \lambda\boldsymbol{c}'\boldsymbol{x} = \lambda\alpha$. □

Exercício 10 *Se* $\boldsymbol{x} \in H_{c,\alpha}$ *então* $\lambda\boldsymbol{x} \in H_{\lambda c, \lambda\alpha}$.

Solução: $\boldsymbol{x} \in H_{c,\alpha} \iff \boldsymbol{c}'\boldsymbol{x} = \alpha \iff (\lambda\boldsymbol{c}')\boldsymbol{x} = \lambda\alpha \iff \boldsymbol{x} \in H_{\lambda c, \lambda\alpha}$. □

1.5 Poliedros e sistemas lineares

Nossa maior preocupação nesta seção é providenciar uma interpretação gráfica de poliedros e sistemas lineares.

[1] $\mathbf{H}'\boldsymbol{c} = \mathbf{0}$ tem solução $\boldsymbol{c} \neq \mathbb{0}$, pois $\mathbf{H} \in \mathbb{R}^{n \times n-1}$, logo $\mathbf{H}' \in \mathbb{R}^{n-1 \times n}$, ou seja, as colunas de $\mathbf{H}'$ devem ser l.d. (n colunas no $\mathbb{R}^{n-1}$).

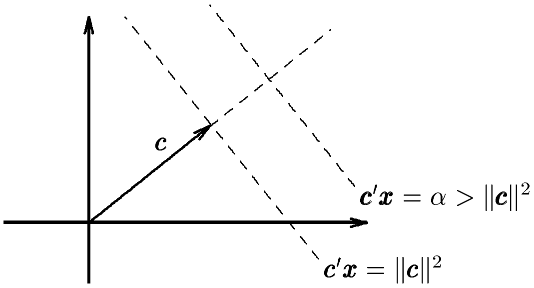


Figura 1.5: Hiperplanos como curvas de nível

Nossa primeira discução foi sobre as diferenças em interpretar a matriz $\mathbf{A}$ do poliedro canônico $X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$. É possível olhar como um vetor (de dimensão n) de produtos escalares com as linhas de $\mathbf{A}$, e "inclinação" definida por $\boldsymbol{b}$. Mas também é possível examinar $\boldsymbol{b}$ como combinação a coeficientes das colunas de $\mathbf{A}$ (e portanto de dimensão m).

Assim, se $X := \{\boldsymbol{x} \in \mathbb{R}^n : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \leqq \mathbb{0}\}$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, podemos:

- *Por linhas*: espaço $\mathbb{R}^n$

  Olhar X como o conjunto dos pontos do $\mathbb{R}^n$ nos m hiperplanos $\boldsymbol{a}_i\boldsymbol{x} = b_i$ ($i \in \{1, 2, \ldots, n\}$) interceptado pelo ortante positivo.

$$\mathbf{A}\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \hline \boldsymbol{a}_2 \\ \hline \cdots \\ \hline \boldsymbol{a}_m \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

- *Por coluna*: espaço $\mathbb{R}^m$

  Olhar X como os multiplicadores das semi-retas (cones) $\lambda \boldsymbol{a}^i$ ($\lambda \in \mathbb{R}_+$) que geram $\boldsymbol{v}$.

$$\mathbf{A}\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}^1 + x_2\boldsymbol{a}^2 + \cdots + x_n\boldsymbol{a}^n = \boldsymbol{b}$$

Para ilustrar a diferença entre as representações, vamos examinar um exemplo no $\mathbb{R}^2$. Seja X definido por

$$\begin{aligned} ax_1 + bx_2 &= 1 \\ cx_1 + dx_2 &= 0. \end{aligned}$$

Assim, a matriz $\mathbf{A}$ ficaria na forma

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

e o sistema de igualdade assume a forma matricial

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Se olharmos para essa matriz da maneira "usual", por linhas, verificamos que ela representa um sistema de equações. Como estamos em $\mathbb{R}^2$, esse sistema de equações representa a intersecção de dois hiperplanos, caracterizando um poliedro. Mas se olharmos a matriz como colunas, verificamos o sequinte:

$$x_1 \begin{pmatrix} a \\ c \end{pmatrix} + x_2 \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Sendo então uma combinação linear, e mais que isso: como $x_i \geqq 0$, $\forall i \in \{1, 2\}$, podemos olhar cada uma das parcelas como geradores de cones.

Generalizando, nessa interpretação de matriz por coluna, para qualquer poliedro canônico na forma $X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$, deve-se notar que $x_1\boldsymbol{a}^1 + \cdots + x_n\boldsymbol{a}^n$ gera um cone de dimensão igual a quantidade de colunos l.i. em $\mathbf{A}$. Logo, para que um vetor $\boldsymbol{x}$ pertença ao poliedro canônico X, é necessário e suficiente que $\boldsymbol{b}$ esteja dentro do cone gerado pelos vetores $\{\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^n\}$.

Assim, sendo $I(\boldsymbol{x}) := \{i \in \{1, 2, \ldots, n\} : x_i \neq 0\}$, fixando a matrix $\mathbf{A}_{m\times n}$ e um vetor $\boldsymbol{x}$ em X, teremos

$$\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \sum_{i\in\{1,2,\ldots,n\}} x_i\boldsymbol{a}^i = \sum_{i\in I(x)} x_i\boldsymbol{a}^i + \sum_{i\notin I(x)} x_i\boldsymbol{a}^i.$$

Como $x_i = 0$, para todo $i \in \{1, 2, \ldots, n\} \setminus I(x)$ $(\equiv i \notin I(x))$, então $\sum_{i\notin I(x)} x_i\boldsymbol{a}^i = 0$ e portanto

$$\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \sum_{i\in I(x)} x_i\boldsymbol{a}^i.$$

Usando uma idéia parecida com a decomposição acima, podemos separar as colunas de $\mathbf{A}$ segundo qualquer subconjunto de $\{1, 2, \ldots, n\}$. Assim, tomando qualquer $B \subseteq \{1, 2, \ldots, n\}$, podemos escrever o sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ na forma separada

$$\boldsymbol{b} = \mathbf{A}_B\boldsymbol{x}_B + \mathbf{A}_N\boldsymbol{x}_N = \mathbf{A}_B\boldsymbol{x}_B.$$

Mais que isso, se tivermos um subconjunto $I \subseteq \{1, 2, \ldots, n\}$ tal que $\#I < m$ e $\{\boldsymbol{a}^i\}_{i\in I}$ é l.i., podemos computar um $\boldsymbol{x}$ que resolve o sistema do seguinte modo: sejam[2]

$B := I(x) \cup J$
$N := \{1, 2, \ldots, n\} \setminus (I(x) \cup J)$
$J := \{i_1, i_2, \ldots, i_k\} : \{\boldsymbol{a}^i\}_{i\in B}$ seja l.i. e $k + \#I(x) = m$ (ou seja, $\{\boldsymbol{a}^i\}_{i\in B}$ é base do $\mathbb{R}^m$)
$\overline{\mathbf{A}} := \mathbf{A}_B^{-1}$,

basta definir $\boldsymbol{x}$ como

$$\boldsymbol{x}_B = \overline{\mathbf{A}}\boldsymbol{b} \ \text{ e } \ \boldsymbol{x}_N = \mathbb{0}, \quad \text{isto é,} \quad x_i = \begin{cases} \overline{\boldsymbol{a}}^i\boldsymbol{b}, & i \in B \\ 0, & i \in N. \end{cases}$$

Sugestão de aprendizado: Para quem não ficou satisfeito com as equivalências acima, sugerimos que "coloque a mão na massa". Por exemplo, pegue o exemplo inicial, e substitua a, b, c, d por $1, 1, -1, 1$. Trace a intersecção dos hiperplanos em $\mathbb{R}^2$. Verá que isto resulta em um único ponto, que é um poliedro unitário (o ponto encontrado é um vértice do poliedro - como veremos futuramente). Tente interpretar a matriz como coluna, trace os cones e teste numericamente o último resultado.

[2]Note que, trabalhando com colunas de $\mathbf{A}$ estamos considerando o espaço onde o vetor lado-direito $\boldsymbol{b}$ se encontra, o $\mathbb{R}^m$.

1.5.1 Transformação de poliedros genéricos em poliedros canônicos

Um resultado bastante útil de PL é que qualquer poliedro (intersecção finita de hiperplanos e semi-espaços) pode ser transformada na forma canônica. A técnica de transformação se resume a dois casos, abaixo expostos.

- Transformar restrições de desigualdade em restrições de igualdade introduzindo uma **variável de folga** não negativa:
$$\langle \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i\boldsymbol{x} \leqq b_i \rangle \iff \langle (\boldsymbol{x}; x_i^r) \in \mathbb{R}^n \times \mathbb{R}_+ : \boldsymbol{a}_i\boldsymbol{x} + x_i^r = b_i \rangle$$
$$\langle \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i\boldsymbol{x} \geqq b_i \rangle \iff \langle (\boldsymbol{x}; x_i^r) \in \mathbb{R}^n \times \mathbb{R}_+ : \boldsymbol{a}_i\boldsymbol{x} - x_i^r = b_i \rangle$$

- Transformar variáveis livres de sinal[3] em variáveis não negativas, trocando cada uma por um par de **variáveis artificiais** : sejam $a^+ := max(a, 0)$ e $a^- := -min(x, 0)$ (então $a = a^+ - a^-$),

$$\left\langle \begin{array}{c} x_i \in \mathbb{R} \text{ e } x_j \in \Gamma, j \in \{1, 2, \ldots, n\}\backslash\{i\} \\ x_1\boldsymbol{a}^1 + \cdots + x_i\boldsymbol{a}^i + \cdots + x_n\boldsymbol{a}^n x_2 = \boldsymbol{b} \end{array} \right\rangle \iff \left\langle \begin{array}{c} (x_i^+, x_i^-) \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ e } x_j \in \Gamma, j \in \{1, 2, \ldots, n\}\backslash\{i\} \\ x_1\boldsymbol{a}^1 + \cdots + x_i^+\boldsymbol{a}^i - x_i^-\boldsymbol{a}^i + \cdots + x_n\boldsymbol{a}^n x_2 = \boldsymbol{b} \end{array} \right\rangle$$

Para ilustrar a técnica, consideremos um poliedro na forma $\overline{\mathbf{A}}\overline{\boldsymbol{x}} \leqq \boldsymbol{b}$. Queremos transformá-lo em um poliedro canônico ($X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$). A primeira coisa a fazer, é criar uma variável de folga para produzir a identidade ($= \boldsymbol{b}$),
$$x_i^r = b_i - \overline{\boldsymbol{a}}_i\,\overline{\boldsymbol{x}}.$$
Note que $x_i^r \geqq 0$, desta forma, acertamos o problema da desigualdade, pois
$$\overline{\boldsymbol{a}}_i\,\overline{\boldsymbol{x}} + x_i^r = b_i, \quad \text{com} \quad x_i \geqq 0,$$
introduzindo uma nova variável (que ao menos está necessariamente no ortante positivo como desejamos).

A seguir, devemos eliminar as variáveis livres de sinal: para cada $\overline{x}_i \in \mathbb{R}$ devemos associar dois diferentes pares de variáveis artificiais, x_i^+ e x_i^- ($x_i^+ \geqq 0$ e $x_i^- \geqq 0$). Desta forma, aplicando a técnica de criação de variáveis artificiais: definindo
$$x_i^+ := max(\overline{x}_i, 0) \quad \text{e} \quad x_i^- := -min(\overline{x}_i, 0),$$
teremos $\overline{x}_i = x_i^+ - x_i^-$.

Para um melhor entendimento das duas técnicas expostas, é conveniente aplicá-las a um exemplo. Considere o poliedro $Y = \{\boldsymbol{y} \in \mathbb{R}^2 : 0 \leqq y_1 + y_2 \leqq 1\}$. Devemos inicialmente "eliminar" as desigualdade: $y_1 + y_2 \geqq 0$ e $y_1 + y_2 \leqq 1$. Aplicando a primeira técnica de transformação, obtemos as variáveis artificiais,

$$x_1 := b_1 - \boldsymbol{a}_1\boldsymbol{y} = 1 - y_1 - y_2 \quad \Longrightarrow \quad \langle \boldsymbol{a}_1\boldsymbol{y} + x_1 = b_1,\ x_1 \geqq 0 \iff \boldsymbol{a}_1\boldsymbol{y} \leqq b_1 \rangle, \quad \text{e}$$
$$x_2 := -b_2 + \boldsymbol{a}_2\boldsymbol{y} = y_1 + y_2, \quad \Longrightarrow \quad \langle \boldsymbol{a}_2\boldsymbol{y} - x_2 = b_2,\ x_2 \geqq 0 \iff \boldsymbol{a}_2\boldsymbol{y} \geqq b_2 \rangle.$$

Como y_1 e y_2 são livres de sinal, é necessário definirmos variáveis artificiais correspondentes para cada uma delas (no ortante positivo, $\boldsymbol{z} \geqq \mathbb{0}$),
$$\langle x_3 := y_1^+ \ \text{ e } \ x_4 := y_1^- \rangle \implies y_1 = x_3 - x_4$$
$$\langle x_5 := y_2^+ \ \text{ e } \ x_6 := y_2^- \rangle \implies y_2 = x_5 - x_6$$

Deste modo, definindo

$$\overline{\mathbf{A}} := \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \ \mathbf{A} := \begin{bmatrix} 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \ \text{ e } \ \boldsymbol{b} := \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

[3]Com sinal invertido é simples, basta trocar $x_i \leqq 0$ por $\overline{x}_i \geqq 0$, fazendo $\overline{x}_i = -x_i$.

$$\left.\begin{array}{rl} Y & = \left\{ \boldsymbol{y} \in \mathbb{R}^2 : \begin{bmatrix} y_1 + y_2 \leqq 1 \\ -y_1 - y_2 \leqq 0 \end{bmatrix} \right\} \\ & = \left\{ \boldsymbol{y} \in \mathbb{R}^2 : \overline{\mathbf{A}}\boldsymbol{y} \leqq \boldsymbol{b} \right\} \end{array}\right\} \longleftrightarrow \left\{\begin{array}{rl} X & = \left\{ \boldsymbol{x} \in \mathbb{R}^6_+ : \begin{bmatrix} x_1 + x_3 - x_4 + x_5 - x_6 = 1 \\ x_2 - x_3 + x_4 - x_5 + x_6 = 0 \end{bmatrix} \right\} \\ & = \left\{ \boldsymbol{x} \in \mathbb{R}^6_+ : \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \right\} \end{array}\right.$$

Com estas definições, gere pontos em cada um dos poliedros e, aplicando as regras, obtenha o ponto correspondente no outro.

1.6 Referências bibliográficas

`Hadley-1962` George Hadley; Linear programming. Reading, Mass., Addison-Wesley Pub. Co., 1962.

`Hoffman&Kunze-1970` Kenneth Hoffman e Ray Kunze; Álgebra linear, São Paulo, Edusp, 1970.

`Hoffman&Kunze-1979` Kenneth Hoffman e Ray Kunze; Álgebra linear, seg. ed., Rio de Janeiro, LTC Editora, 1979.

`Monteiro-1969` Luis Henrique Jacy Monteiro; Álgebra linear, sexta ed., São Paulo, Nobel, 1969.

`Hoffman&Kunze-1971` Kenneth Hoffman e Ray Kunze; Linear algebra, seg. ed., Englewood Cliffs, N.J., Prentice-Hall, 1971.

`Hadley-1982` George Hadley; Programação linear, Rio de Janeiro, Ed. Guanabara Dois, 1982.

`Dorfman,et.al.-1987` Robert Dorfman, Paul A. Samuelson e Robert M. Solow, Linear Programming and Economic Analysis, (prim. ed. McGraw-Hill, 1958), Dover Publ. Inc., Nova Yorque, 1987.

`Bertsimas&Tsitsiklis-1997` Dimitris Bertsimas e John N. Tsitsiklis, Introduction do Linear Optimization, Belmont, Massachusetts, Athena Scientific, 1997.