

# 1 Introdução

## 1.1 Exemplos de problemas de programação linear

### O problema da dieta

Considere  $n$  diferentes alimentos e  $m$  diferentes nutrientes, e suponha que você possui uma tabela com o conteúdo nutricional de uma unidade ou porção de cada alimento:

	alim. 1	...	alim. n
nutr. 1	$a_{11}$	...	$a_{1n}$
$\vdots$	$\vdots$		$\vdots$
nutr. m	$a_{m1}$	...	$a_{mn}$

Note que a  $j$ -ésima coluna da matriz representa o conteúdo nutricional do  $j$ -ésimo alimento. Seja  $b_i$  o requisito nutricional mínimo do nutriente  $i$  em uma dieta balanceada. Podemos interpretar um vetor  $x \in \mathbb{R}_+^n$  como a especificação de uma dieta que utiliza  $x_j$  unidades/porções do alimento  $j$  para cada  $j = 1, \dots, n$ . A dieta  $x$  será balanceada se satisfizer  $\sum_{j=1}^n a_{ij}x_j \geq b_i$  para cada nutriente  $i$ . Se estivermos interessados numa dieta balanceada com uma ingestão mínima de calorias e além disso conhecermos a quantidade de calorias  $c_j$  de cada unidade/porção do alimento  $j$ , podemos

$$\begin{aligned}
 \min \quad & c_1x_1 + \dots + c_nx_n \\
 \text{sujeito a} \quad & a_{11}x_1 + \dots + a_{1n}x_n \geq b_1 \\
 & a_{21}x_1 + \dots + a_{2n}x_n \geq b_2 \\
 & \vdots \\
 & a_{m1}x_1 + \dots + a_{mn}x_n \geq b_m \\
 & x \geq 0.
 \end{aligned}$$

---

<sup>1</sup>Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.

Seja  $A \in \mathbb{R}^{m \times n}$  a matriz cujas entradas são  $a_{ij}$ . O mesmo problema pode ser escrito em notação matricial como

$$\begin{array}{ll}\min & c'x \\ \text{s.a} & Ax \geq b \\ & x \geq 0.\end{array}$$

**Variante 1:** Se estivermos interessados em obter a dieta balanceada mais barata e soubermos o preço  $r_j$  de cada unidade/porção do alimento  $j$ , então desejaremos minimizar a expressão  $r'x$ .

**Variante 2:** Se  $b$  representar os requisitos exatos de uma dieta “ideal”, então  $x$  será uma dieta ideal se satisfizer  $Ax = b$ ,  $x \geq 0$ . Sob estas restrições poderemos minimizar o conteúdo calórico ou o custo de uma dieta ideal.

### Um problema de produção

Uma empresa produz  $n$  diferentes produtos usando  $m$  diferentes matérias-primas. Seja  $b_i$  a quantidade disponível da  $i$ -ésima matéria-prima,  $a_{ij}$  a quantidade de matéria-prima  $i$  necessária para a produção do produto  $j$ , e  $c_j$  o lucro obtido com a venda do produto  $j$ . Se a variável  $x_j$  representa a quantidade de produto  $j$  a ser produzida, então a empresa terá o máximo lucro resolvendo o problema

$$\begin{array}{ll}\max & c'x \\ \text{s.a} & Ax \leq b \\ & x \geq 0.\end{array}$$

### O problema do plantão

Um hospital quer fazer a programação semanal dos plantões noturnos de seus enfermeiros. A cada dia da semana a demanda por enfermeiros de plantão é diferente, representada por um inteiro  $d_j$ ,  $j = 1, \dots, 7$ . Cada enfermeiro sempre trabalha 5 noites seguidas em plantão. O problema é encontrar o número mínimo de enfermeiros que o hospital precisa contratar.

Se tentássemos criar uma variável  $x_j$  representando o número de enfermeiros de plantão no dia  $j$  não seríamos capazes de escrever a restrição de que cada enfermeiro sempre trabalha 5 noites seguidas (experimentalmente!). Ao invés disso, representamos por  $x_j$  o número de enfermeiros que *começa* a trabalhar no dia  $j$ ; assim os enfermeiros que começarem

a trabalhar no dia 5 trabalharão nos dias 5, 6, 7, 1 e 2. O problema pode então ser formulado como

$$\begin{array}{ll}
\min & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \\
\text{s.a} & x_1 + x_4 + x_5 + x_6 + x_7 \geq d_1 \\
& x_1 + x_2 + x_5 + x_6 + x_7 \geq d_2 \\
& x_1 + x_2 + x_3 + x_6 + x_7 \geq d_3 \\
& x_1 + x_2 + x_3 + x_4 + x_7 \geq d_4 \\
& x_1 + x_2 + x_3 + x_4 + x_5 \geq d_5 \\
& x_2 + x_3 + x_4 + x_5 + x_6 \geq d_6 \\
& x_3 + x_4 + x_5 + x_6 + x_7 \geq d_7 \\
& x_j \geq 0, \quad x_j \in \mathbb{Z}
\end{array}$$

**Exercício 1.1** Verifique que a solução ótima da relaxação contínua deste problema (sem a restrição  $x_j \in \mathbb{Z}$ ) pode ser obtida através da fórmula

$$x = \frac{1}{5} \begin{bmatrix} 3 & -2 & 3 & -2 & 3 & -2 & -2 \\ -2 & 3 & -2 & 3 & -2 & 3 & -2 \\ -2 & -2 & 3 & -2 & 3 & -2 & 3 \\ 3 & -2 & -2 & 3 & -2 & 3 & -2 \\ -2 & 3 & -2 & -2 & 3 & -2 & 3 \\ 3 & -2 & 3 & -2 & -2 & 3 & -2 \\ -2 & 3 & -2 & 3 & -2 & -2 & 3 \end{bmatrix} d.$$

Este é um problema de programação linear *inteira*. Em algumas situações especiais um problema deste tipo pode ser resolvido como um problema de programação linear (sem a restrição  $x_j \in \mathbb{Z}$ ).

## Classificação de padrões

O problema de classificação de padrões corresponde a tentar identificar a classe de um objeto a partir da descrição de algumas de suas propriedades (seu padrão). Consideremos um exemplo bem simples: são dadas várias imagens de maçãs e laranjas, e para cada imagem um vetor  $a \in \mathbb{R}^3$  tal que  $a_1$  é a curvatura do objeto representado,  $a_2$  é o comprimento da haste e  $a_3$  é sua cor. O conjunto  $\{a^i\}_{i \in S}$  contém padrões de maçãs, e  $\{a^i\}_{i \notin S}$  contém padrões de laranjas. Um classificador linear para distinguir as maçãs e laranjas dadas é um par  $(x, y) \in \mathbb{R}^3 \times \mathbb{R}$  que satisfaz

$$\begin{aligned}
(a^i)'x &\geq y, \quad i \in S \\
(a^i)'x &< y, \quad i \notin S.
\end{aligned}$$

Dado um novo padrão  $\bar{a}$  de uma imagem desconhecida, o mesmo será declarado pelo classificador uma maçã se satisfizer  $\bar{a}'x \geq y$ , uma laranja caso contrário.

## Ordenação

Este exemplo ilustra a versatilidade de modelos de programação linear, mas não deve ser tomado como uma aplicação real: considere que se queira ordenar os números dados  $a_1, \dots, a_n$ . O valor ótimo do problema abaixo

$$\begin{array}{ll} \min & a'x \\ \text{s.a} & \sum_{i=1}^n x_i = 1 \\ & x_i \in \{0, 1\} \end{array}$$

é o menor dentre os valores  $a_1, \dots, a_n$ .

Este é um exemplo de problema de programação linear *0-1* que pode ser reformulado como problema de programação linear trocando-se a restrição  $x_i \in \{0, 1\}$  por  $0 \leq x_i \leq 1$ . O método simplex (que veremos em breve) é capaz de encontrar soluções para o problema reformulado que satisfazem  $x_i = 0$  ou  $x_i = 1$ .

**Exercício 1.2** *Verifique que, trocando-se a restrição  $\sum_{i=1}^n x_i = 1$  por  $\sum_{i=1}^n x_i = k$  no problema original obtém-se como solução ótima a soma dos  $k$  menores valores dentre  $a_1, \dots, a_n$ .*

**Exercício 1.3** *Use o programa pLsolve para resolver o problema de planejamento de produção da DEC usando como entrada o arquivo abaixo:*

$$\max: 60x_1 + 40x_2 + 30x_3 + 30x_4 + 15x_5;$$

$$\begin{array}{rcll} x_1 + & x_2 + & x_3 + & x_4 + & x_5 <= 7 & ; \\ 4x_1 + & 2x_2 + & 2x_3 + & 2x_4 + & x_5 <= 8 & ; \\ & x_2 & & + & x_4 & <= 3 & ; \\ x_1 & & & & & <= 1.8; \\ & & x_3 & & & <= 0.3; \\ x_1 + & x_2 + & x_3 & & & <= 3.8; \\ & & & x_4 + & x_5 <= 3.2; \\ & x_2 & & & & >= 0.5; \\ & & & x_4 & & >= 0.5; \\ & & & & x_5 >= 0.4; \end{array}$$

Troque as restrições do problema original pelas restrições alternativas da modelagem e compare as soluções obtidas.

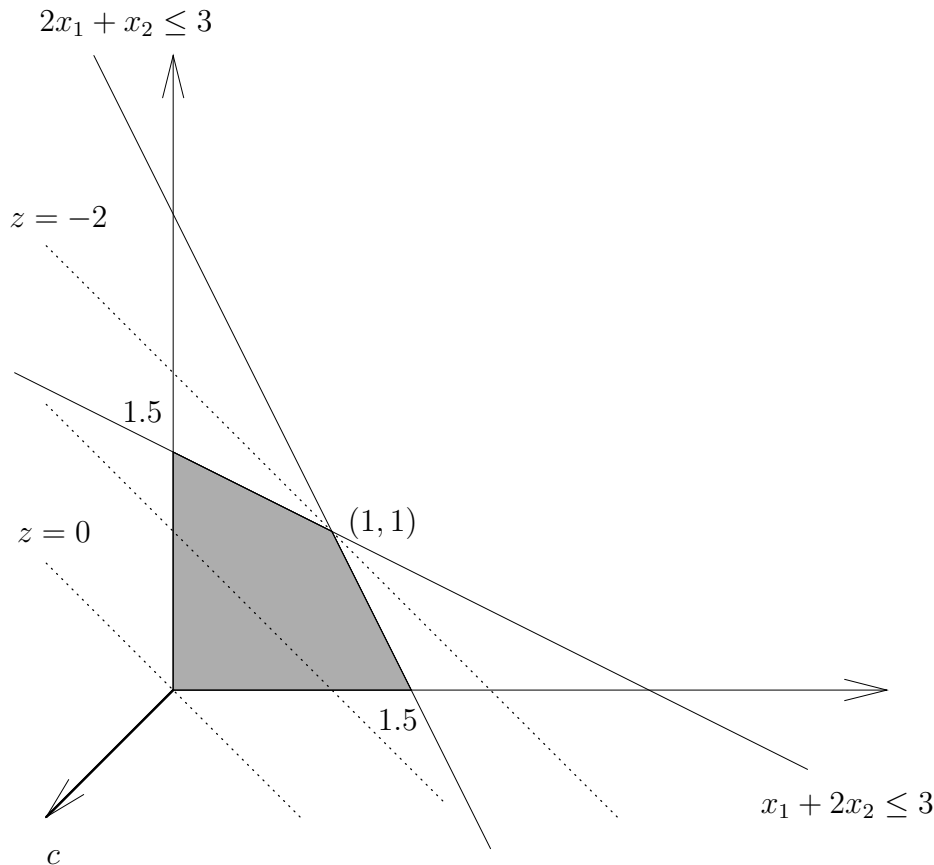
Resolva o problema do plantão usando o `plSolve` (invente os valores de  $d_1, \dots, d_7$ ) e verifique que a solução obtida corresponde à fórmula do exercício 1.1.

## 1.2 Representação e solução gráficas

Considere o problema

$$\begin{array}{llllll} \min & -x_1 & - & x_2 & & \\ \text{s.a} & x_1 & + & 2x_2 & \leq & 3 \\ & 2x_1 & + & x_2 & \leq & 3 \\ & x_1, & x_2 & \geq & 0 & \end{array}$$

O conjunto viável está representado a seguir.



Para encontrar a solução ótima podemos considerar o conjunto dos pontos que têm um mesmo valor de função objetivo, digamos  $z$ . Este é uma linha descrita pela equação  $-x_1 - x_2 = z$  e é perpendicular ao vetor  $c = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ . Para cada valor de  $z$  obtemos uma linha paralela diferente:

aumentando  $z$  seguimos na direção apontada pelo vetor  $c$ , diminuindo  $z$  seguimos na direção  $-c$ . Para minimizar a função objetivo devemos procurar o ponto viável o mais distante na direção  $-c$ : este é o ponto  $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  com valor ótimo  $z = -2$ .

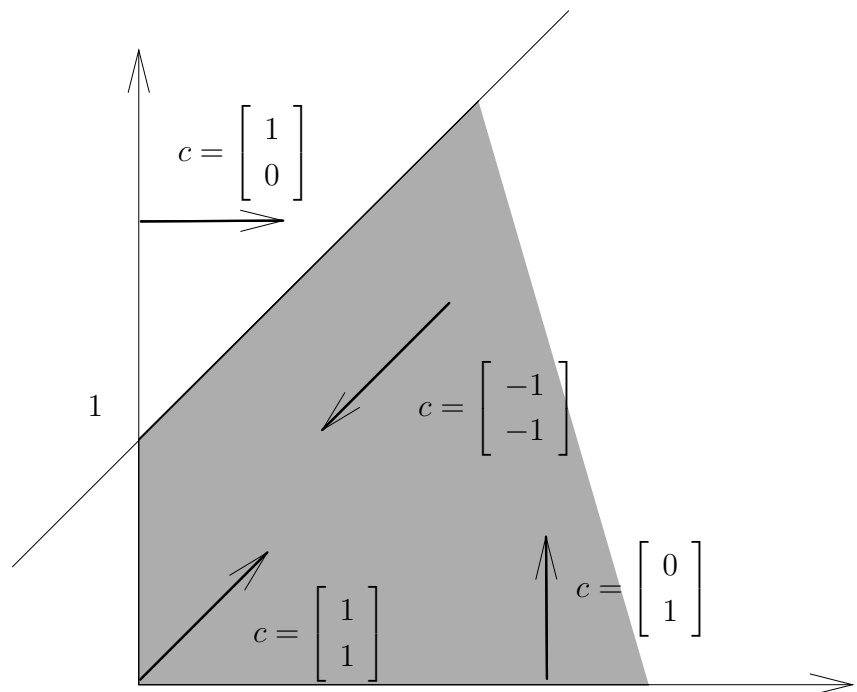
Uma prova algébrica de que esta é uma solução ótima é obtida ao notar-se que toda solução viável satisfaz  $3x_1 + 3x_2 \leq 6$  (soma das duas primeiras restrições), ou equivalentemente,  $-x_1 - x_2 \geq -2$ . Assim  $x$  não só é viável como possui o menor valor possível dentre as soluções viáveis.

**Exercício 1.4** *Obtenha graficamente a solução ótima do problema abaixo e verifique algebricamente que ela é de fato ótima.*

$$\begin{array}{ll} \min & -x_1 + x_2 - x_3 \\ \text{s.a} & 0 \leq x_1 \leq 1 \\ & 0 \leq x_2 \leq 1 \\ & 0 \leq x_3 \leq 1 \end{array}$$

Considere o exemplo a seguir:

$$\begin{array}{llll} \min & c_1x_1 & + & c_2x_2 \\ \text{s.a} & -x_1 & + & x_2 \leq 1 \\ & x_1, & x_2 & \geq 0 \end{array}$$



1. Se  $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ , a única solução ótima é  $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ .

2. Se  $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ , todo vetor da forma  $x = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}$  com  $0 \leq x_2 \leq 1$  é solução ótima (infinitas soluções, conjunto de soluções limitado).
3. Se  $c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ , todo vetor da forma  $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$  com  $x_1 \geq 0$  é solução ótima (infinitas soluções, conjunto de soluções ilimitado).
4. Se  $c = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ , é possível obter uma sequência de soluções viáveis com valores tendendo a  $-\infty$ . Dizemos que o problema é ilimitado e que o valor ótimo é  $-\infty$  (mas nenhuma solução viável atinge este valor!)
5. Se o problema tivesse a restrição adicional  $x_1 + x_2 \leq -2$ , o conjunto viável seria vazio.

### 1.3 Variantes do problema de programação linear

Um problema de programação linear pode ser expresso da maneira mais geral como

$$\begin{aligned}
 &\min / \max \quad c'x \\
 &\text{s.a} \quad a'_i x \geq b_i, \quad i \in M_1, \\
 &\quad \quad a'_i x \leq b_i, \quad i \in M_2, \\
 &\quad \quad a'_i x = b_i, \quad i \in M_3, \\
 &\quad \quad x_j \geq 0, \quad j \in N_1, \\
 &\quad \quad x_j \leq 0, \quad j \in N_2.
 \end{aligned}$$

$x_1, \dots, x_n$  são chamadas *variáveis de decisão* e um vetor  $x$  que satisfaz todas as restrições é chamado de *solução viável* ou *ponto viável*. O conjunto de soluções viáveis é chamado de *conjunto viável* ou *região viável*. Para  $j \notin N_1 \cup N_2$  dizemos que  $x_j$  é uma variável *livre de sinal*. A função  $f(x) = c'x$  é chamada *função objetivo* e uma solução viável  $x^*$  que minimiza a função objetivo (isto é, que satisfaz  $c'x^* \leq c'x$  para todo  $x$  viável) é chamada de *solução ótima* do problema de minimização, com o correspondente *valor ótimo* igual a  $c'x^*$  (define-se analogamente solução ótima do problema de maximização). Se, por outro lado, for possível encontrar soluções viáveis com valores de função objetivo tão baixos quanto se queira, diz-se que o valor ótimo é  $-\infty$  e que o problema é ilimitado.

Um tal grau de generalidade na formulação é inconveniente e desnecessário do ponto de vista da teoria, que pode ser desenvolvida para

moldes mais simples de problemas lineares. Discutimos a seguir algumas reformulações que não tiram expressividade do modelo de programação linear.

$$\begin{array}{l}
x^* \text{ é sol. ótima de } \max c'x \\
\Updownarrow \\
c'x^* \geq c'x \quad \forall x \text{ viável} \\
\Updownarrow \\
-c'x^* \leq -c'x \quad \forall x \text{ viável} \\
\Updownarrow \\
x^* \text{ é sol. ótima de } \min -c'x.
\end{array}$$

Inicialmente note que não é necessário estudar problemas de maximização e de minimização separadamente, pois são equivalentes  $\max c'x$  e  $\min -c'x$ .

Toda restrição de igualdade do tipo  $a'_i x = b_i$  é equivalente às duas desigualdades  $a'_i x \leq b_i$  e  $a'_i x \geq b_i$ ; além disso qualquer desigualdade do tipo  $a'_i x \leq b_i$  pode ser re-escrita como  $(-a_i)'x \geq -b_i$ . Note ainda que as restrições de sinal  $x_j \geq 0$  e  $x_j \leq 0$  são casos particulares das restrições do tipo  $a'_i x \geq b_i$ , onde  $a_i = e_j$  e  $b_i = 0$ .

Através destas equivalências concluímos que qualquer problema de programação linear pode ser escrito na forma compacta que designamos de *forma geral de programação linear*:

$$\begin{array}{ll}
\min & c'x \\
\text{s.a} & Ax \geq b.
\end{array}$$

Outra forma muito importante no desenvolvimento de algoritmos de programação linear é a *forma canônica (ou padrão) de programação linear*:

$$\begin{array}{ll}
\min & c'x \\
\text{s.a} & Ax = b \\
& x \geq 0.
\end{array}$$

Vimos que a forma canônica é um caso particular da forma geral, ou seja, podemos re-escrever qualquer problema em forma canônica na forma geral. O inverso também é verdade: pode-se transformar qualquer problema de programação linear para a forma canônica. Esta transformação depende de duas operações básicas:

**Eliminação de variáveis livres de sinal** Dada uma variável  $x_j$  livre de sinal, substitui-se todas as ocorrências desta por  $x_j^+ - x_j^-$ , onde  $x_j^+, x_j^-$  são novas variáveis não-negativas, isto é, sujeitas às condições  $x_j^+ \geq 0$  e  $x_j^- \geq 0$ . A idéia é que todo número real pode ser escrito como a diferença de dois números não-negativos.

**Eliminação de desigualdades** Dada uma desigualdade da forma  $a'_i x \leq b_i$ , introduz-se uma nova variável  $r_i$  e as restrições (compatíveis com a forma canônica)

$$\begin{array}{l}
a'_i x + r_i = b_i \\
r_i \geq 0.
\end{array}$$



A variável  $r_i$  é chamada de *residual*.

Naturalmente uma variável não-positiva  $x_j \leq 0$  é substituída por  $-x_j$ , e uma desigualdade da forma  $a'_i x \geq b_i$  é substituída por  $a'_i x - r_i = b_i$  e  $r_i \geq 0$ .

Como exemplo, o problema

$$\begin{array}{llllll} \min & 2x_1 & - & x_2 & + & 4x_3 \\ \text{s.a} & x_1 & + & x_2 & & + x_4 \leq 2 \\ & & & 3x_2 & - & x_3 = 5 \\ & & & & & x_3 + x_4 \geq 3 \\ & x_1 & & & & \geq 0 \\ & & & x_3 & & \leq 0 \end{array}$$

pode ser re-escrito na forma geral como

$$\begin{array}{llllll} \min & 2x_1 & - & x_2 & + & 4x_3 \\ \text{s.a} & -x_1 & - & x_2 & & - x_4 \geq -2 \\ & & & 3x_2 & - & x_3 \geq 5 \\ & & & -3x_2 & + & x_3 \geq -5 \\ & & & & & x_3 + x_4 \geq 3 \\ & x_1 & & & & \geq 0 \\ & & & -x_3 & & \geq 0 \end{array}$$

e na forma canônica como

$$\begin{array}{llllllllll} \min & 2x_1 & - & x_2^+ & + & x_2^- & - & 4x_3 & & & \\ \text{s.a} & x_1 & + & x_2^+ & - & x_2^- & & & + & x_4 & + r_1 = 2 \\ & & & 3x_2^+ & - & 3x_2^- & + & x_3 & & & = 5 \\ & & & & & & & -x_3 & + & x_4 & - r_2 = 3 \\ & x_1, & & x_2^+, & & x_2^-, & & x_3, & & x_4, & r_1, & r_2 \geq 0. \end{array}$$

**Exercício 1.5** Passe o problema de planejamento da DEC para a forma canônica.

## 1.4 Pré-requisitos e notação

### Conjuntos

Utiliza-se as seguintes notações:  $|S|$  denota a cardinalidade do conjunto  $S$ ,  $S \setminus T = \{s \in S \mid s \notin T\}$ ,  $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$  e  $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ .

## Vetores e matrizes

$A \in \mathbb{R}^{m \times n}$  denota uma matriz

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

$x \in \mathbb{R}^n$  denota o vetor coluna  $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ ; alguns vetores especiais

utilizados são  $0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$  e  $e^i$  o  $i$ -ésimo vetor da base canônica.

$A'$  denota a *transposta* de  $A$ , que satisfaz  $A'_{ij} = A_{ji}$ ,  $\forall i, j$ . Se  $x, y \in \mathbb{R}^n$ , então  $x'y = \sum_{i=1}^n x_i y_i$ . Lembre-se que  $x \perp y$  ( $x$  é ortogonal a  $y$ )  $\iff x'y = 0$ . A *norma Euclideana* de  $x \in \mathbb{R}^n$  é denotada por  $\|x\| = \sqrt{x'x}$ .

**Exercício 1.6** *Demonstre a desigualdade de Cauchy-Schwarz:*

$$|x'y| \leq \|x\| \|y\|,$$

e verifique que a igualdade vale se, e somente se,  $x$  é múltiplo de  $y$  ou vice-versa. Dica: se  $x \neq 0$ , considere a distância de  $y$  à projeção de  $y$  sobre  $x$ .

Dada uma matriz  $A$  denota-se por  $A^j$  sua  $j$ -ésima coluna, e  $A_i$  sua  $i$ -ésima linha. O *produto* de duas matrizes  $A \in \mathbb{R}^{m \times n}$  e  $B \in \mathbb{R}^n \times k$  é uma matriz em  $\mathbb{R}^{m \times k}$  cujas componentes são dadas por  $[AB]_{ij} = \sum_{l=1}^n a_{il} b_{lj} = A_i B^j$ , ou ainda,

$$AB = \left[ \begin{array}{c|c|c} AB^1 & \cdots & AB^k \end{array} \right] = \left[ \begin{array}{c} A_1 B \\ \vdots \\ A_m B \end{array} \right].$$

O produto de matrizes satisfaz  $(AB)C = A(BC)$  e  $(AB)' = B'A'$ , mas não satisfaz  $AB = BA$  em geral (produza um contra-exemplo!) A matriz *identidade* é denotada por  $I$  e satisfaz  $IA = A$  e  $BI = B$  para quaisquer  $A$  e  $B$  de dimensões compatíveis.

Dados  $A \in \mathbb{R}^{m \times n}$  e  $x \in \mathbb{R}^n$ , valem as identidades  $Ae^i = A^i$ ,  $x = \sum_{i=1}^n x_i e^i$  e

$$Ax = \sum_{i=1}^n A^i x_i = \begin{bmatrix} A_1 x \\ \vdots \\ A_m x \end{bmatrix}.$$

Dado  $x \in \mathbb{R}^n$  utiliza-se as notações  $x \geq 0 \iff x_i \geq 0 \forall i$  e  $x > 0 \iff x_i > 0 \forall i$ ; analogamente define-se as desigualdades  $A \geq 0$  e  $A > 0$ . Estas definições estendem-se naturalmente:  $x \geq y \iff x - y \geq 0$  e  $x > y \iff x - y > 0$ .

## Inversão de matrizes

Uma matriz quadrada  $A$  é dita *não-singular* ou *invertível* se existe uma matriz  $B$  tal que  $AB = BA = I$ ; esta matriz é única e denotada por  $A^{-1}$ . Se  $A, B \in \mathbb{R}^{n \times n}$  são invertíveis então  $AB$  é invertível e  $(AB)^{-1} = B^{-1}A^{-1}$ .

Uma coleção de vetores  $x^1, \dots, x^k \in \mathbb{R}^n$  é dita *linearmente dependente* se existem  $a_1, \dots, a_k$  reais tais que  $a \neq 0$  e  $\sum_{i=1}^k a_i x^i = 0$ ; caso contrário diz-se que os vetores são *linearmente independentes*. Vale o seguinte resultado

**Teorema 1.2** *Seja  $A$  uma matriz quadrada. As afirmações seguintes são equivalentes:*

1.  $A$  é invertível;

2.  $A'$  é invertível;

3.  $\det(A) \neq 0$ ;

4.  $A_1, \dots, A_n$  são linearmente independentes;

5.  $A^1, \dots, A^n$  são linearmente independentes;

6. Existe  $b$  tal que o sistema  $Ax = b$  tem uma única solução.

7. Para todo  $b$  o sistema  $Ax = b$  tem uma única solução;

$$\begin{aligned} A'(A^{-1})' &= (A^{-1}A)' = I' = I, \\ (A^{-1})'A' &= (AA^{-1})' = I' = I. \end{aligned}$$

$$\det(A^{-1}) \det(A) = \det(I) = 1.$$

$Ax = 0$  possui solução única.

**Prova.**

(3  $\implies$  4) Suponha que  $\sum_{i=1}^n \alpha_i A_i = 0$ , e considere

$$B = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Então  $[BA]_1 = \sum_{i=1}^n \alpha_i A_i = 0 \implies 0 = \det(BA) = \det(B) \det(A) \implies \det(B) = \alpha_1 = 0$ . Repita o argumento para  $\alpha_2, \dots, \alpha_n$ .

(4  $\implies$  5) Suponha que  $A\alpha = \sum_{i=1}^n \alpha_i A^i = 0$ . Então  $A_i \alpha = 0$ ,  $\forall i$ , logo  $\alpha = \sum_{i=1}^n \text{proj}_{A_i}(\alpha) = \frac{A_i \alpha}{\|A_i\|^2} A_i = 0$ .

(6  $\implies$  7) Suponha que  $Ax = b$  possui uma única solução  $z$  e que  $Ax = b'$  possui duas soluções  $x$  e  $y$ . Então  $A(z + x - y) = Az + A(x - y) = Az + Ax - Ay = b + b' - b' = b$ , logo  $z + x - y$  é solução de  $Ax = b$  e portanto  $x = y$ .

(7  $\implies$  1) Defina  $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$  por  $T(b) = x \iff Ax = b$ . Então  $AT(b) = b$ ,  $\forall b \in \mathbb{R}^n$  e  $TA(x) = x$ ,  $\forall x \in \mathbb{R}^n$ , ou seja,  $AT \equiv TA \equiv I$  e a matriz que representa  $T$  é a inversa de  $A$ .

## Subespaços e bases

$\emptyset \neq S \subset \mathbb{R}^n$  é um *subespaço* de  $\mathbb{R}^n$  se  $ax + by \in S$  para quaisquer  $x, y \in S$  e  $a, b \in \mathbb{R}$ . O *subespaço gerado* por  $x^1, \dots, x^k$  é o conjunto de vetores  $y$  da forma  $y = \sum_{i=1}^k a_i x^i$ , onde  $a \in \mathbb{R}^k$ ;  $y$  é dito uma *combinação linear* de  $x^1, \dots, x^k$ .

Se  $B$  e  $C$  geram  $S$ , e  $n = |C| >$

$|B| = m$ , então  $C = BA$  onde  $A \in \mathbb{R}^{m \times n}$ . Seja  $x \neq 0 : Ax = 0$ .

Então  $Cx = BAx = B0 = 0$ , ou seja,  $C$  não é LI.

Dado um subespaço  $\emptyset \neq S \subset \mathbb{R}^n$ , uma *base* de  $S$  é uma coleção de vetores linearmente independentes que geram  $S$ . Toda base de  $S$  tem o mesmo número de vetores e este número é chamado de *dimensão* de  $S$ ; em particular  $\dim(\mathbb{R}^n) = n$  e todo subespaço de  $\mathbb{R}^n$  tem dimensão menor ou igual a  $n$ . Por definição,  $\dim(\{0\}) = 0$ .

Se  $S$  é um subespaço *próprio* de  $\mathbb{R}^n$  (isto é, diferente de  $\mathbb{R}^n$ ), então existe um vetor  $a \neq 0$  tal que  $a'x = 0$  para todo  $x \in S$ . Em geral, se  $\dim(S) = m < n$ , então existem  $n - m$  vetores linearmente independentes ortogonais a  $S$ .

**Teorema 1.3** *Suponha que o subespaço  $S$  gerado pelos vetores  $x^1, \dots, x^k$  tem dimensão  $m$ . Então:*

1. *Existe uma base de  $S$  contendo  $m$  vetores dentre  $x^1, \dots, x^k$ ;*
2. *Se  $l \leq m$  e  $x^1, \dots, x^l$  são linearmente independentes, pode-se formar uma base de  $S$  começando com  $x^1, \dots, x^l$  e escolhendo  $m - l$  vetores dentre  $x^{l+1}, \dots, x^k$*

**Exercício 1.7** *Demonstre o teorema acima. Note que a parte 1 é caso particular da parte 2 (com  $l = 0$ ).*

Seja  $A \in \mathbb{R}^{m \times n}$ . O *espaço-coluna* gerado por  $A$  é o subespaço (de  $\mathbb{R}^m$ ) gerado pelas colunas de  $A$ ; analogamente definimos o *espaço-linha* de  $A$ . As dimensões dos espaço-linha e espaço-coluna de  $A$  coincidem e este número é denominado *posto* ou *característica* de  $A$ .  $A$  possui *posto completo* ou *característica plena* se  $\text{posto}(A) = \min\{m, n\}$ . O conjunto  $\{x \in \mathbb{R}^n \mid Ax = 0\}$  é chamado de *espaço-nulo* de  $A$ .

**Exercício 1.8** *Prove que*

Observe que 1 e 2 implicam  $\dim(\text{col}(A)) = \dim(\text{lin}(A))$ .

1.  $\dim(\text{nulo}(A)) = n - \dim(\text{linhas}(A))$ .
2.  $\dim(\text{colunas}(A)) + \dim(\text{nulo}(A)) = n$ .

## Subespaços afins

Dado um subespaço linear  $S_0 \subset \mathbb{R}^n$  e um vetor  $x_0 \in \mathbb{R}^n$ , o conjunto  $S = S_0 + x_0 = \{x + x_0 \mid x \in S_0\}$  corresponde a uma translação do subespaço  $S_0$  e é denominado *subespaço afim*.  $S$  tem a mesma dimensão de  $S_0$ , por definição.

Como um primeiro exemplo, considere  $k + 1$  vetores em  $\mathbb{R}^n$ , e o conjunto  $S = \{x^0 + \lambda_1 x^1 + \dots + \lambda_k x^k \mid \lambda \in \mathbb{R}^k\}$ .  $S$  é um subespaço afim obtido a partir da translação do subespaço  $S_0$  gerado pelos vetores  $x^1, \dots, x^k$ . Se estes últimos vetores forem linearmente independentes, então tanto  $S_0$  quanto  $S$  terão dimensão  $k$ .

Outro exemplo é dado pelo conjunto  $S = \{x \in \mathbb{R}^n \mid Ax = b\}$  onde  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  e supomos  $S \neq \emptyset$ . Seja  $x^0 \in \mathbb{R}^n$  tal que  $Ax^0 = b$ ; então  $x \in S \iff Ax = Ax^0 = b$ , ou seja,  $x \in S \iff A(x - x^0) = 0 \iff x - x^0 \in S_0 = \{y \mid Ay = 0\}$ . Ou seja,  $S$  é a translação do espaço-nulo de  $A$  pelo vetor  $x^0$ . Cada restrição  $a'_i x = b_i$  define um hiperplano de  $\mathbb{R}^n$  (um subespaço de dimensão  $n - 1$ ), e em princípio diminui em 1 a dimensão do conjunto  $S$ : se  $A$  possui  $m$  linhas linearmente independentes, então  $\dim(S) = n - m$ .

## 2 Geometria e Programação Linear

### 2.1 Poliedros e conjuntos convexos

Alguns conceitos geométricos importantes no estudo de programação linear são introduzidos a seguir.

**Definição 2.1**  $P \subset \mathbb{R}^n$  é um **poliedro** se existem  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$  tais que  $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ .

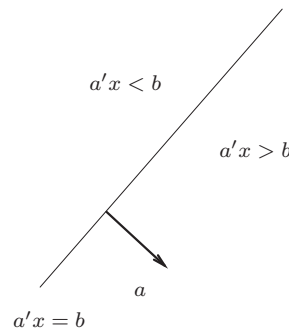
Como já vimos, qualquer problema de programação linear pode ser escrito na *forma geral*, onde o conjunto de pontos viáveis é expresso por  $Ax \geq b$ ; assim o conjunto viável de qualquer PL é um poliedro. O conjunto  $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  define um tipo especial de poliedro que chamaremos de *poliedro na forma canônica*.

**Definição 2.2** Um conjunto  $S \subset \mathbb{R}^n$  é **limitado** se existe  $K \in \mathbb{R}_+$  tal que  $|x_i| < K, \forall x \in S$ .

**Exercício 2.1** Mostre que  $S \subset \mathbb{R}^n$  é limitado se, e somente se, existe  $K \in \mathbb{R}_+$  tal que  $\|x\| < K, \forall x \in S$ .

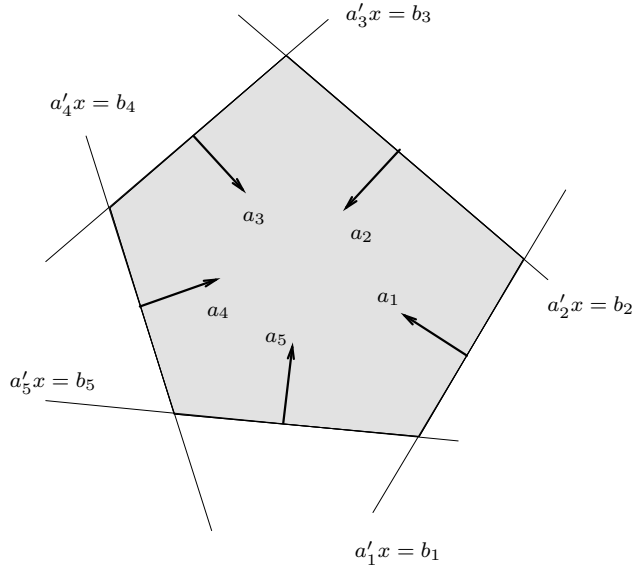
**Definição 2.3** Um **hiperplano** é um conjunto da forma  $\{x \in \mathbb{R}^n \mid a'x = b\}$ , e um **semi-espço** é um conjunto da forma  $\{x \in \mathbb{R}^n \mid a'x \geq b\}$ , onde  $a \in \mathbb{R}^n$  e  $b \in \mathbb{R}$ .

**Exercício 2.2** Mostre que  $a$  é ortogonal a qualquer direção interna ao hiperplano  $\{x \mid a'x = b\}$ .

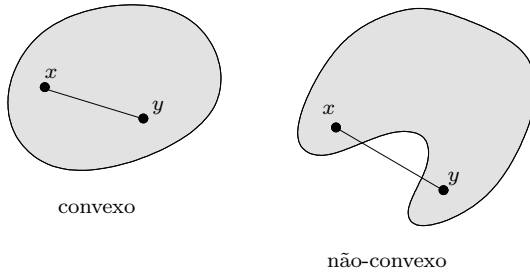


<sup>1</sup>Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.

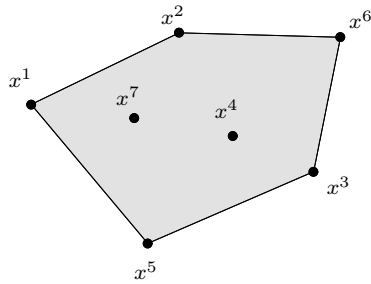
Note que hiperplanos e semi-espacos s3o casos particulares de poliedros, e que todo poliedro 3 uma intersec33o finita de semi-espacos.



**Defini33o 2.4**  $S \subset \mathbb{R}^n$  3 **convexo** se  $\forall x, y \in S, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in S$ .



**Defini33o 2.5** Sejam  $x^1, \dots, x^k \in \mathbb{R}^n$  e  $\lambda_1, \dots, \lambda_k \in \mathbb{R}_+$  tais que  $\lambda_1 + \dots + \lambda_k = 1$ . O vetor  $\sum_{i=1}^k \lambda_i x^i$  3 dito uma **combina33o convexa** dos vetores  $x^1, \dots, x^k$ . O **casco convexo** dos vetores  $x^1, \dots, x^k$  3 o conjunto de todas as combina333es convexas destes vetores.



## Teorema 2.1

1. A intersec33o de conjuntos convexas 3 convexa.

2. *Todo poliedro é convexo.*
3. *Um conjunto convexo é fechado por combinações convexas.*
4. *O casco convexo de um conjunto finito de vetores é convexo.*

**Prova.**

1. Sejam  $x, y \in S = \bigcap_{i \in I} S_i \subset \mathbb{R}^n$  e  $\lambda \in (0, 1)$ . Então  $x, y \in S_i, \forall i \in I$  e, como  $S_i$  é convexo,  $\lambda x + (1 - \lambda)y \in S_i, \forall i \in I$ , ou seja,  $\lambda x + (1 - \lambda)y \in S$ .
2. Vamos mostrar que todo semi-espaco é convexo: seja  $S = \{x \in \mathbb{R}^n \mid a'x \geq b\}$  um semi-espaco,  $x, y \in S$  e  $\lambda \in (0, 1)$ . Então  $a'(\lambda x + (1 - \lambda)y) = \lambda a'x + (1 - \lambda)a'y \geq \lambda b + (1 - \lambda)b = b$  (usando  $a'x \geq b, a'y \geq b, \lambda, (1 - \lambda) > 0$ ) e assim  $\lambda x + (1 - \lambda)y \in S$ ; logo  $S$  é convexo. Pela parte (1) toda intersecção finita de semi-espacos (ou seja, todo poliedro) é convexa(o).
3. A afirmação é que se  $S$  é convexo,  $x^1, \dots, x^k \in S$  e  $\lambda_1, \dots, \lambda_k \in \mathbb{R}_+$  satisfazem  $\lambda_1 + \dots + \lambda_k = 1$ , então  $\sum_{i=1}^k \lambda_i x^i \in S$ . O caso  $k = 2$  é a própria definição de convexidade; considere então que o resultado vale para  $k$  e vamos provar que vale também para  $k + 1$ , completando a prova por indução. Sejam  $x^1, \dots, x^{k+1} \in S$  e  $\lambda_1, \dots, \lambda_{k+1} \in \mathbb{R}_+$  tais que  $\lambda_1 + \dots + \lambda_{k+1} = 1$ . Suponha que  $\lambda_{k+1} \neq 1$  (senão teríamos  $\sum_{i=1}^{k+1} \lambda_i x^i = x^{k+1} \in S$ ); então

$$\sum_{i=1}^{k+1} \lambda_i x^i = \lambda_{k+1} x^{k+1} + (1 - \lambda_{k+1}) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i.$$

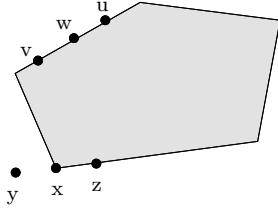
Como  $\frac{\lambda_i}{1 - \lambda_{k+1}} \geq 0$  e  $\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = \frac{\sum_{i=1}^k \lambda_i}{1 - \lambda_{k+1}} = \frac{1 - \lambda_{k+1}}{1 - \lambda_{k+1}} = 1$ , temos que  $y = \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \in S$ , e por convexidade,  $\sum_{i=1}^{k+1} \lambda_i x^i = \lambda_{k+1} x^{k+1} + (1 - \lambda_{k+1})y \in S$ .

4. Seja  $S$  o casco convexo de  $x^1, \dots, x^k$  e  $y, z \in S$ . Então  $y = \sum_{i=1}^k \mu_i x^i$  e  $z = \sum_{i=1}^k \nu_i x^i$ , com  $\mu, \nu \geq 0$  e  $\sum_{i=1}^k \mu_i = \sum_{i=1}^k \nu_i = 1$ . Seja  $\lambda \in (0, 1)$ ; então  $\lambda y + (1 - \lambda)z = \lambda(\sum_{i=1}^k \mu_i x^i) + (1 - \lambda)(\sum_{i=1}^k \nu_i x^i) = \sum_{i=1}^k (\lambda \mu_i + (1 - \lambda)\nu_i) x^i$ . Como  $\lambda \mu_i + (1 - \lambda)\nu_i \geq 0, \forall i$  e  $\sum_{i=1}^k \lambda \mu_i + (1 - \lambda)\nu_i = \lambda \sum_{i=1}^k \mu_i + (1 - \lambda) \sum_{i=1}^k \nu_i = \lambda + (1 - \lambda) = 1$ , segue que  $\lambda y + (1 - \lambda)z \in S$ . Logo  $S$  é convexo.



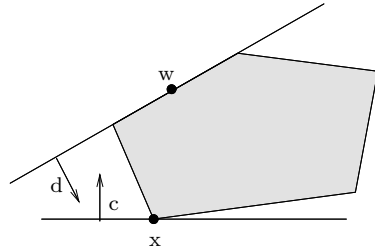
## 2.2 Pontos extremos, vértices e soluções básicas viáveis

**Definição 2.6** Seja  $P$  um poliedro.  $x \in P$  é um **ponto extremo** de  $P$  se não existem  $y, z \in P$ ,  $y \neq z$  e  $\lambda \in (0, 1)$  tais que  $x = \lambda y + (1 - \lambda)z$ .



**Exercício 2.3** Mostre que a condição  $\left[ \nexists y, z \in P, y \neq z \text{ t.q. } x = \frac{1}{2}(y + z) \right]$  é equivalente à da definição acima.

**Definição 2.7** Seja  $P \subset \mathbb{R}^n$  um poliedro.  $x \in P$  é um **vértice** de  $P$  se existe  $c \in \mathbb{R}^n$  t.q.  $c'x < c'y, \forall y \in P \setminus \{x\}$ .

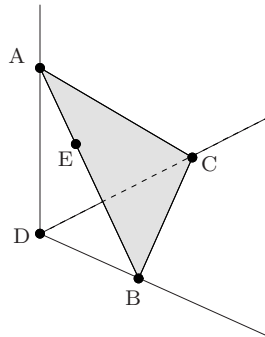


Procuramos traduzir os conceitos expressos pelas definições 2.6 e 2.7 para um contexto algébrico, que utilize a representação do poliedro e que permita uma verificação computacional. Para isso considere um poliedro  $P \subset \mathbb{R}^n$  descrito pelas igualdades e desigualdades a seguir:

$$\begin{aligned} a'_i x &\geq b_i, & i \in M_1 \\ a'_i x &\leq b_i, & i \in M_2 \\ a'_i x &= b_i, & i \in M_3, \end{aligned}$$

onde  $M_1$ ,  $M_2$  e  $M_3$  são conjuntos (finitos) de índices.

**Definição 2.8** Se  $x^*$  satisfaz  $a'_i x^* = b_i$  para  $i \in \bigcup_{k=1}^3 M_k$  dizemos que esta restrição é **ativa** em  $x^*$ .



$$P = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1, x \geq 0\}$$

Note que podemos associar a cada ponto do poliedro um conjunto de restrições ativas naquele ponto. Estamos interessados em considerar conjuntos de restrições ativas que estão associados a um único ponto do poliedro. Esta unicidade está associada à independência linear dos vetores que definem as restrições, como veremos a seguir.

**Teorema 2.2** *Seja  $x^* \in \mathbb{R}^n$  e  $I = \{i \in \bigcup_{k=1}^3 M_k \mid a'_i x^* = b_i\}$ . Então são equivalentes:*

1. *Existem  $n$  vetores linearmente independentes dentre os  $a_i$ ,  $i \in I$ ;*
2. *O espaço gerado por  $\{a_i \mid i \in I\}$  é o  $\mathbb{R}^n$ ;*
3. *O sistema de equações  $a'_i x = b_i$ ,  $i \in I$  tem solução única.*

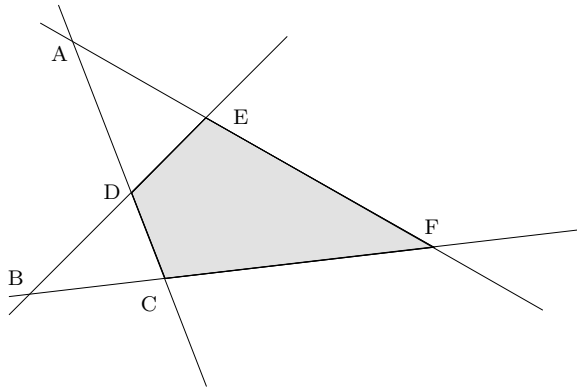
**Prova.**

- $(1 \Rightarrow 2)$  Como existem  $n$  vetores  $a_i$  linearmente independentes, o espaço gerado por eles tem dimensão  $\geq n$ . Como este espaço gerado é subconjunto do  $\mathbb{R}^n$  e o único subespaço do  $\mathbb{R}^n$  de dimensão  $n$  é o próprio  $\mathbb{R}^n$ , segue a implicação.
- $(2 \Rightarrow 3)$  Suponha por contradição que exista outra solução  $y^* \neq x^*$  para o sistema  $a'_i x = b_i$ ,  $i \in I$ . Então  $x^* - y^*$  é ortogonal a todos os  $a_i$ ,  $i \in I$ , e portanto não pode ser gerado por eles. Isso contradiz (2).
- $(3 \Rightarrow 1)$  Suponha por contradição que não existam  $n$  vetores linearmente independentes dentre os  $a_i$ ,  $i \in I$ ; então existe  $d \neq 0$  ortogonal a todos os  $a_i$ ,  $i \in I$  e portanto se  $x^*$  é solução de  $a'_i x = b_i$ ,  $i \in I$ ,  $x^* + d$  também é solução deste sistema, contrariando (3).

Freqüentemente diremos que certas restrições são *linearmente independentes* querendo dizer que os vetores  $a_i$  que as definem são linearmente independentes. Note que se  $x^*$  é solução do sistema  $a'_i x = b_i$ ,  $i \in I$  não necessariamente é verdade que  $x^*$  seja viável; outras restrições (com índices fora do conjunto  $I$ ) poderão ser violadas. Isso nos leva à seguinte definição.

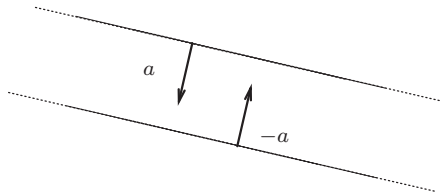
**Definição 2.9** *Seja  $P$  um poliedro e  $x^* \in \mathbb{R}^n$ .  $x^*$  é dito **solução básica** se todas as restrições de igualdade são satisfeitas e além disso dentre as restrições ativas existem  $n$  linearmente independentes. Uma **solução básica viável** é uma solução básica que satisfaz todas as restrições, inclusive as inativas.*

Esta distinção entre igualdades e desigualdades faz com que a definição acima seja *dependente* da representação do poliedro: se trocássemos todas as igualdades por pares de desigualdades, novas soluções básicas (inviáveis) surgiriam.



Soluções básicas (A e B são inviáveis)

Observe que se um poliedro possui menos restrições do que variáveis ( $m < n$ ), então o número de restrições ativas em qualquer ponto é também menor do que  $n$  e portanto não existem soluções básicas. Isso também ocorre quando  $m \geq n$  mas existem menos do que  $n$  restrições linearmente independentes:



O teorema a seguir relaciona as definições anteriores.

**Teorema 2.3** *Seja  $P$  um poliedro e  $x^* \in P$ . São equivalentes:*

1.  $x^*$  é um vértice;
2.  $x^*$  é um ponto extremo;
3.  $x^*$  é uma solução básica viável.

**Prova.**

Sem perda de generalidade considere que  $P$  é descrito pelas desigualdades  $a'_i x \geq b_i$  e igualdades  $a'_i x = b_i$ .

- $(1 \Rightarrow 2)$  Suponha que  $x^* \in P$  é vértice, ou seja, que existe  $c \in \mathbb{R}^n$  tal que  $c'x^* < c'y$ ,  $\forall y \in P \setminus \{x^*\}$ . Para quaisquer  $y, z \in P$ ,  $y \neq z$  e  $\lambda \in (0, 1)$  temos que

$c'(\lambda y + (1 - \lambda)z) = \lambda c'y + (1 - \lambda)c'z > \lambda c'x^* + (1 - \lambda)c'x^* = c'x^*$  (a desigualdade se deve a ou  $x^* \neq y$  ou  $x^* \neq z$ ), assim  $\lambda y + (1 - \lambda)z \neq x^*$ ; ou seja,  $x$  não pode ser expresso como combinação convexa de outros pontos de  $P$  (definição de ponto extremo)

- (2  $\Rightarrow$  3) Suponha por contradição que  $x^* \in P$  não seja uma solução básica viável; vamos mostrar que  $x^*$  não pode ser ponto extremo. Seja  $I = \{i \mid a'_i x^* = b_i\}$ . Como  $x^*$  não é solução básica viável, existem menos de  $n$  dentre os  $a_i$ ,  $i \in I$  linearmente independentes, e portanto existe um  $d \neq 0$  tal que  $a'_i d = 0$ ,  $\forall i \in I$ . Como todas as restrições inativas (com índices fora de  $I$ ) satisfazem  $a'_i x^* > b_i$ , existe um  $\epsilon > 0$  tal que o intervalo  $[x^* - \epsilon d, x^* + \epsilon d] \subset P$  e, em particular,  $x^* = \frac{1}{2}(x^* - \epsilon d) + \frac{1}{2}(x^* + \epsilon d)$ , com  $x^* - \epsilon d \neq x^* + \epsilon d$ . Isso contradiz a extremalidade de  $x^*$ .
- (3  $\Rightarrow$  1) Seja  $c = \sum_{i \in I} a_i$ , onde  $I = \{i \mid a'_i x^* = b_i\}$ . Então para qualquer  $y \in P$ ,

$$c'y = \sum_{i \in I} a'_i y \geq \sum_{i \in I} b_i = \sum_{i \in I} a'_i x^* = c'x^*.$$

Além disso,  $x^*$  é a única solução do sistema  $a'_i x = b_i$ ,  $i \in I$ , ou seja,  $\forall y \in P \setminus \{x^*\}$  vale  $c'y > c'x^*$ , e portanto  $x^*$  é vértice.

Note que a definição de vértice e de ponto extremo independe da representação do poliedro; pelo teorema anterior concluímos que a definição de solução básica viável também não depende da representação, muito embora a definição de solução básica dependa. Note também que a definição algébrica (solução básica viável) pode ser implementada computacionalmente: pode-se testar se um conjunto de vetores é linearmente independente utilizando o método da triangularização. Outra propriedade importante é:

**Corolário 2.1** *Dado um número finito de desigualdades e igualdades lineares, o número de soluções básicas (e consequentemente de vértices) é finito.*

**Prova.**

Basta notar que o número de subconjuntos de restrições linearmente independentes é finito, e cada um está associado

a no máximo uma solução básica  $x^*$  (no sentido de corresponder exatamente ao conjunto de restrições ativas em  $x^*$ ).

Apesar de finito, o número de vértices de um poliedro pode ser muito alto mesmo que a descrição do poliedro seja “pequena”: o hipercubo unitário em  $\mathbb{R}^n$  é descrito por  $2n$  desigualdades ( $x_i \geq 0$  e  $x_i \leq 1$ ,  $i = 1, \dots, n$ ) e no entanto todos os  $2^n$  pontos da forma  $(x_1, \dots, x_n)'$  com  $x_i \in \{0, 1\}$  são vértices do hipercubo.

Outro conceito importante associado aos vértices é o de vizinhança ou adjacência:

**Definição** *Duas soluções básicas são ditas **adjacentes** se existem  $n - 1$  restrições linearmente independentes que sejam ativas em ambas as soluções. Se duas soluções básicas adjacentes são viáveis, o segmento de reta que as une é chamado de **aresta** do poliedro.*

**Exercício 2.4** *Mostre que todos os pontos da forma  $(x_1, \dots, x_n)'$  com  $x_i \in \{0, 1\}$  são vértices do hipercubo unitário em  $\mathbb{R}^n$ . Mostre que dois vértices  $(x_1, \dots, x_n)'$  e  $(y_1, \dots, y_n)'$  são adjacentes se, e somente se, existe um  $j \in \{1, \dots, n\}$  tal que  $x_i = y_i$ ,  $\forall i \neq j$  e  $x_j = 1 - y_j$ .*

## 2.3 Poliedros na forma canônica

Vamos especializar as definições e resultados da seção anterior para poliedros da forma  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ ; estes resultados serão muito importantes para o desenvolvimento do método simplex.

Freqüentemente faremos a hipótese fundamental de que a matriz  $A \in \mathbb{R}^{m \times n}$  possui linhas linearmente independentes e, em particular, que  $m \leq n$ . Veremos mais tarde que isso não acarreta perda de generalidade na discussão, pois se  $P \neq \emptyset$  então linhas linearmente dependentes de  $A$  correspondem a restrições supérfluas e que podem ser removidas.

Note que toda solução básica de  $P$  precisa satisfazer as restrições  $Ax = b$  por definição; isso fornece  $m$  restrições linearmente independentes, de acordo com a hipótese fundamental sobre  $A$ . Para obter um vértice precisamos obter mais  $n - m$  restrições ativas, ou seja, variáveis  $x_i = 0$  (tais que a restrição  $x_i \geq 0$  fique ativa). A necessidade de que o conjunto de restrições ativas resultante seja linearmente independente nos fornece a seguinte caracterização:

**Teorema 2.4** *Considere  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  e que  $A \in \mathbb{R}^{m \times n}$  possui linhas linearmente independentes. Então  $x \in \mathbb{R}^n$  é solução básica se e somente se  $Ax = b$  e existem índices  $b_1, \dots, b_m$*

tais que as colunas  $A^{b_1}, \dots, A^{b_m}$  são linearmente independentes e  $x_i = 0, \forall i \notin \{b_1, \dots, b_m\}$ .

**Prova.**

Inicialmente suponha que  $x \in \mathbb{R}^n$  satisfaz  $Ax = b$  e existem índices  $b_1, \dots, b_m$  tais que as colunas  $A^{b_1}, \dots, A^{b_m}$  são linearmente independentes e  $x_i = 0, \forall i \notin \{b_1, \dots, b_m\}$ . O sistema linear  $\sum_{i=1}^m A^{b_i} x_{b_i} = b$  possui solução única, pois a matriz  $[A^{b_1} \dots A^{b_m}]$  é inversível. Então o sistema  $Ax = b, x_i = 0, \forall i \notin \{b_1, \dots, b_m\}$  possui solução única, visto que substituindo  $x_i = 0$  em  $Ax = b$  obtemos  $Ax = \sum_{i=1}^n A^i x_i = \sum_{i=1}^m A^{b_i} x_{b_i} = b$ . Logo  $x^*$  é solução básica.

Por outro lado, considere que  $x^*$  é solução básica e seja  $I = \{i \mid x_i \neq 0\} = \{b_1, \dots, b_k\}$ . Por definição,  $x^*$  é a única solução do sistema  $Ax = b, x_i = 0, \forall i \notin I$ , e o sistema equivalente  $\sum_{i=1}^k A^{b_i} x_{b_i} = b$  possui solução única. Então as colunas  $A^{b_1}, \dots, A^{b_k}$  são linearmente independentes; em particular  $k \leq m$ . Como  $\text{posto}(A) = \dim(\text{col}(A)) = \dim(\text{lin}(A)) = m$ , podemos obter  $m - k$  colunas  $A^{b_{k+1}}, \dots, A^{b_m}$  de tal forma que  $A^{b_1}, \dots, A^{b_m}$  são linearmente independentes. Além disso  $\forall i \notin \{b_1, \dots, b_m\}$  temos que  $i \notin \{b_1, \dots, b_k\} = I$ , e portanto  $x_i = 0$ .

**Exercício 2.5** *Mostre que  $x^* \in P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  é um vértice de  $P$  se e somente se as colunas  $\{A^i \mid x_i^* \neq 0\}$  são linearmente independentes. Dica: separe em 2 casos; 1 -  $A$  possui linhas linearmente independentes e 2 -  $A$  possui linhas linearmente dependentes.*

Com o resultado acima podemos construir todos os vértices de um poliedro usando o algoritmo abaixo.

#### Algoritmo para construir vértices

1. Escolha  $m$  índices  $I = \{b_1, \dots, b_m\}$  tais que as colunas  $A^{b_1}, \dots, A^{b_m}$  sejam linearmente independentes;
2. Faça  $x_i = 0, \forall i \notin I$ ;
3. Resolva o sistema  $\sum_{i \in I} A^i x_i = b$ ;
4. Teste se a solução  $x^*$  satisfaz  $x^* \geq 0$ .

Veja que no passo 4 a viabilidade da solução se reduz ao teste  $x^* \geq 0$ , pois o fato dela satisfazer  $\sum_{i \in I} A^i x_i^* = b$  e  $x_i^* = 0, \forall i \notin I$  garantem

que  $Ax^* = b$ . Se  $x^*$  é uma solução básica, as variáveis  $x_{b_1}, \dots, x_{b_m}$  são chamadas **variáveis básicas** e as restantes são chamadas **não-básicas**. As colunas  $A^{b_1}, \dots, A^{b_m}$  são chamadas **colunas básicas** e como são linearmente independentes formam uma **base** do  $\mathbb{R}^m$ . Chamamos de **matriz básica** à matriz

$$B = [A^{b_1} \dots A^{b_m}]$$

e o correspondente vetor básico associado é  $x_B = (x_{b_1}, \dots, x_{b_m})'$ . O valor das variáveis básicas é obtido resolvendo-se o sistema  $Bx_B = b$ , cuja única solução é  $x_B = B^{-1}b$ .

**Exemplo 2.1** *Seja o poliedro  $P = \{x \in \mathbb{R}^7 \mid Ax = b, x \geq 0\}$  onde*

$$A = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ 12 \\ 4 \\ 6 \end{bmatrix}.$$

*Note que a matriz  $A$  possui uma submatriz identidade, cujas colunas são linearmente independentes. Considerando as colunas  $A^4, A^5, A^6, A^7$  como básicas temos a solução associada  $x = (0, 0, 0, 8, 12, 4, 16)$  que é um vértice (por ser não-negativa). Tomando as colunas  $A^3, A^5, A^6, A^7$  e resolvendo o sistema correspondente teremos a solução  $x = (0, 0, 4, 0, -12, 4, 6)$  que é uma solução básica inviável.*

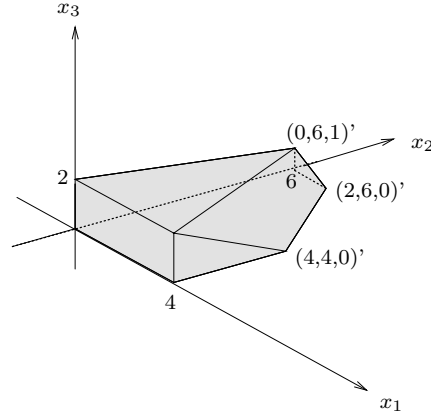
*Suponha que a matriz possuisse uma coluna  $A^8$  idêntica à  $A^7$ ; então os conjuntos  $\{A^3, A^5, A^6, A^7\}$  e  $\{A^3, A^5, A^6, A^8\}$  seriam idênticos. Porém as duas bases (associadas aos conjuntos de índices  $\{3, 5, 6, 7\}$  e  $\{3, 5, 6, 8\}$ ) são distintas. Note que as soluções associadas seriam  $x = (0, 0, 4, 0, -12, 4, 6, 0)$  e  $y = (0, 0, 4, 0, -12, 4, 0, 6)$ , que são distintas.*

*Por outro lado, ainda no problema original, considere a solução  $x = (4, 0, 2, 0, 0, 0, 6)$ . Verifique que ela está associada a quatro bases distintas, formadas pelos conjuntos  $\{A^1, A^2, A^3, A^7\}$ ,  $\{A^1, A^3, A^4, A^7\}$ ,  $\{A^1, A^3, A^5, A^7\}$  e  $\{A^1, A^3, A^6, A^7\}$ . Repare ainda que esta solução está associada a 8 restrições ativas, de onde se podem extrair também quatro subconjuntos de 7 restrições linearmente independentes (garantindo solução única do sistema de equações associado).*

*O poliedro acima pode ser visto como uma representação alternativa para o poliedro em  $\mathbb{R}^3$  abaixo, e toda a discussão acima (traduzindo a idéia de “base” para “subconjuntos de 3 restrições ativas”) pode ser*

adaptada para este contexto com a ajuda do desenho que segue.

$$Q = \left\{ x \in \mathbb{R}^3 \left| \begin{array}{rclcl} x_1 & + & x_2 & + & 2x_3 & \leq & 8 \\ & & x_2 & + & 6x_3 & \leq & 12 \\ x_1 & & & & & \leq & 4 \\ & & x_2 & & & \leq & 6 \\ x_1, x_2, x_3 & \geq & 0 & & & & \end{array} \right. \right\}.$$



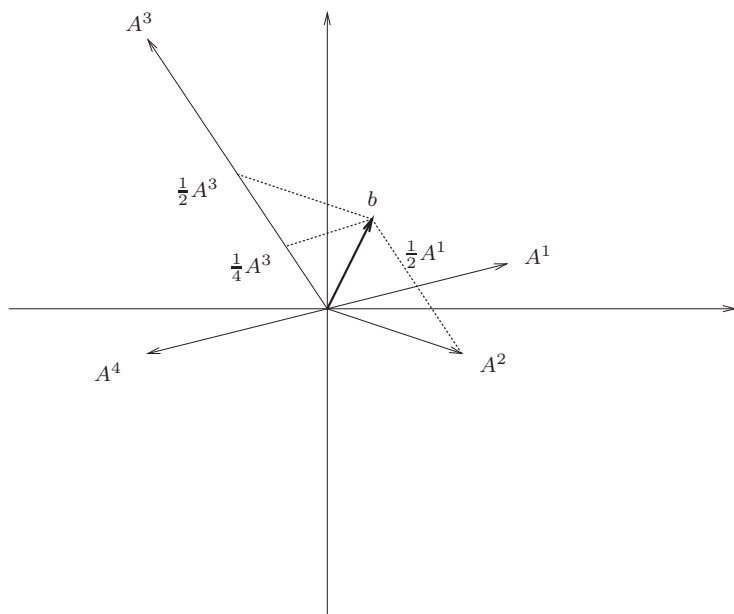
Para ter uma intuição geométrica de soluções básicas devemos utilizar uma outra interpretação do problema, que corresponde a encontrar coeficientes  $x_i$  que mostrem que  $b$  é combinação linear dos vetores  $A^i$  ( $b = \sum_{i=1}^n A^i x_i = Ax$ ). Numa solução básica isso é atingido utilizando apenas  $m$  vetores linearmente independentes; para a solução básica ser viável, todos os coeficientes  $x_i$  têm que ser positivos. Lembrando de nossos problemas-exemplo do capítulo 1, as colunas de  $A$  estão normalmente associadas ao modo como as variáveis de decisão se relacionam com as exigências do problema, e  $b$  representa um atendimento ideal (ou mínimo ou máximo) destas exigências. Assim  $A^i$  representa como o computador  $i$  utiliza as matérias-primas UCP, memória, leitor de discos, e  $b$  representa as quantidades máximas disponíveis destas matérias primas (problemas da DEC e problema de produção);  $A^i$  representa quanto o alimento  $i$  fornece de cada nutriente e  $b$  a quantidade de cada nutriente numa dieta balanceada (ou ideal) (problema da dieta);  $A^i$  representa a rotina semanal de trabalho dos enfermeiros que começam no dia  $i$  e  $b$  é o quadro funcional mínimo do hospital ao longo da semana (problema do plantão). Representando as colunas  $A^i$  e o vetor  $b$  graficamente teremos um esquema dos recursos e exigências do problema e poderemos obter soluções básicas utilizando  $m$  (ou menos) recursos.

**Exemplo** Considere o poliedro  $P = \{x \in \mathbb{R}^4 \mid Ax = b, x \geq 0\}$  onde

$$A = \begin{bmatrix} 4 & 3 & -4 & -4 \\ 1 & -1 & 6 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$



As colunas  $A^i$  e o vetor  $b$  podem ser representados como segue:



Note que  $A^1$  e  $A^2$  são linearmente independentes, mas a solução básica correspondente é inviável pois será necessária uma quantidade  $x_2$  negativa para escrever  $b = x_1 A^1 + x_2 A^2$ ; também serão inviáveis as bases associadas a  $\{A^2, A^4\}$  e  $\{A^3, A^4\}$ . Os vetores  $A^1$  e  $A^3$  são linearmente independentes e produzem uma solução básica viável  $x = (\frac{1}{2}, 0, \frac{1}{4}, 0)'$ ; também será uma solução básica viável  $x = (0, 1, \frac{1}{2}, 0)'$  associada a  $\{A^2, A^3\}$ .  $A^1$  e  $A^4$  não formam uma base pois são linearmente dependentes.

### Correspondência entre bases e soluções básicas

Como vimos anteriormente, soluções básicas distintas correspondem a bases distintas, pois uma base está associada a um sistema de equações que possui solução única. No entanto duas bases diferentes podem corresponder à mesma solução básica (que satisfaz os dois sistemas de equações); além do exemplo que já vimos, considere o caso extremo em que  $b = 0$  e todas as bases estão associadas à mesma solução viável ( $x = 0$ ). A consideração deste fenômeno é muito importante do ponto de vista computacional, e está associado ao fenômeno de *degenerescência* que veremos na próxima seção.

### Soluções básicas adjacentes e bases adjacentes

Lembremos que duas soluções básicas distintas são ditas adjacentes se elas possuem  $n - 1$  restrições ativas linearmente independentes em comum. Em problemas na forma canônica, dizemos que duas bases são adjacentes se elas possuem  $m - 1$  colunas em comum (diferem em apenas uma coluna).

**Exercício 2.6** Prove que duas soluções básicas adjacentes tem sempre associadas duas bases adjacentes. Prove que se duas bases adjacentes produzem soluções básicas distintas, então estas soluções são adjacentes.

**Exemplo 2.2** No exemplo 2.1 as bases  $\{A^4, A^5, A^6, A^7\}$  e  $\{A^3, A^5, A^6, A^7\}$  são adjacentes pois diferem em apenas uma coluna. As soluções básicas correspondentes  $x = (0, 0, 0, 8, 12, 4, 6)'$  e  $y = (0, 0, 4, 0, -12, 4, 6)'$  também são adjacentes: o problema está em  $\mathbb{R}^7$  e elas possuem 6 restrições ativas linearmente independentes em comum ( $x_1 \geq 0, x_2 \geq 0$  e as 4 restrições de igualdade). No desenho em  $\mathbb{R}^3$  estas soluções correspondem a  $x = (0, 0, 0)'$  e  $y = (0, 0, 4)'$  (esta última é inviável e não aparece no desenho). Verifique as colunas básicas e restrições ativas em comum para outros pares de soluções adjacentes daquele desenho.

**A hipótese posto**( $A$ ) =  $\dim(\text{linha}(A)) = m$

Havíamos mencionado que a hipótese de que  $A$  possui linhas linearmente independentes não restringia em nada a discussão desta seção. Como exemplo, imagine que uma das linhas de  $A$  seja combinação linear das outras, digamos  $A_m = \sum_{i < m} \alpha_i A_i$ . Então existem duas possibilidades:

1.  $b_m = \sum_{i < m} \alpha_i b_i$ : neste caso qualquer solução  $x$  que satisfaça  $A_i x = b_i$  para  $i = 1, \dots, m-1$  irá satisfazer

$$A_m x = \left( \sum_{i < m} \alpha_i A_i \right) x = \sum_{i < m} \alpha_i A_i x = \sum_{i < m} \alpha_i b_i = b_m,$$

ou seja, a restrição  $A_m = b_m$  é desnecessária;

2.  $b_m \neq \sum_{i < m} \alpha_i b_i$ : pelo mesmo argumento concluiremos que qualquer  $x$  que satisfaça  $A_i x = b_i$  para  $i = 1, \dots, m-1$  NÃO irá satisfazer  $A_m = b_m$ , e portanto o poliedro correspondente é vazio.

Restringindo a discussão para poliedros não-vazios, vemos então que restrições que se escrevem como combinação linear das demais podem ser descartadas. Isso justifica o seguinte teorema:

**Teorema 2.5** Seja  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} \neq \emptyset$ , onde  $A \in \mathbb{R}^{m \times n}$ . Suponha que  $\text{posto}(A) = k < m$  e que as linhas  $A_1, \dots, A_k$  sejam linearmente independentes. Então  $P = \{x \in \mathbb{R}^n \mid A_i x = b_i, i = 1, \dots, k, x \geq 0\}$ .

**Prova.**

Como  $\text{posto}(A) = k$  e  $A_1, \dots, A_k$  são linearmente independentes, estas linhas formam uma base do espaço-linha associado a  $A$ . Assim as linhas  $A_j$ ,  $j = k + 1, \dots, m$  são geradas (como combinação linear) pelas linhas  $A_1, \dots, A_k$ . Utilizando o argumento acima, sempre caímos no caso 1 (pois o poliedro é não-vazio), e com isso provamos que  $\{x \in \mathbb{R}^n \mid A_i x = b_i, i = 1, \dots, k, x \geq 0\} \subset P$ . A inclusão contrária é imediata, pois qualquer  $x$  que satisfaz  $Ax = b$  satisfaz  $A_i x = b_i$ ,  $i = 1, \dots, k, k + 1, \dots, m$ .

Deste teorema concluímos que qualquer poliedro não-vazio na forma canônica pode ser re-escrito como  $P = \{x \in \mathbb{R}^n \mid Dx = f, x \geq 0\}$ , onde  $D \in \mathbb{R}^{k \times n}$  e  $f \in \mathbb{R}^k$  podem ser escritos como

$$D = \left[ \begin{array}{c} A_{i_1} \\ A_{i_2} \\ \vdots \\ A_{i_k} \end{array} \right] \quad \text{e} \quad f = \left[ \begin{array}{c} b_{i_1} \\ b_{i_2} \\ \vdots \\ b_{i_k} \end{array} \right]$$

e as linhas de  $D$  são linearmente independentes.

**Exemplo 2.3** Considere o poliedro definido por

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 2 \\ x_1 + x_2 &= 1 \\ x_1 + x_3 &= 1 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

A primeira restrição corresponde à soma das outras duas, que por sua vez são linearmente independentes. Assim o poliedro acima é equivalente ao definido por

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_3 &= 1 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

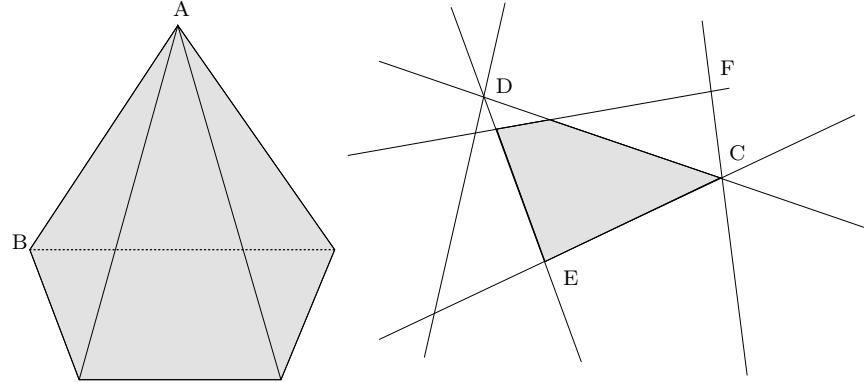
### 3 Degenerescência

O fenômeno conhecido como degenerescência está associado a termos em uma solução básica mais do que o número mínimo necessário de restrições ativas:

**Definição 2.10** Uma solução básica  $x \in \mathbb{R}^n$  é **degenerada** se mais do que  $n$  restrições são ativas em  $x$ .

Assim, se o poliedro está em  $\mathbb{R}^2$  uma solução básica degenerada está na intersecção de 3 ou mais retas; em  $\mathbb{R}^3$  a intersecção de 4 ou mais planos define uma solução básica degenerada. No caso de poliedros canônicos, pode-se traduzir a definição acima como:

**Definição 2.11** *Seja  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  com  $A \in \mathbb{R}^{m \times n}$  e  $x$  uma solução básica. Dizemos que  $x$  é **degenerada** se  $x$  possui mais de  $n - m$  componentes nulas.*



Soluções básicas viáveis não-degeneradas: B,E  
 Soluções básicas viáveis degeneradas: A,C  
 Solução básica (inviável) não-degenerada: F  
 Solução básica (inviável) degenerada: D

**Exemplo** Voltando ao exemplo 2.1 da página 23, podemos observar que a solução básica  $x = (2, 6, 0)'$  da formulação em  $\mathbb{R}^3$  é não-degenerada, pois possui exatamente 3 restrições ativas:  $x_1 + x_2 + 2x_3 = 8$ ,  $x_2 = 6$  e  $x_3 = 0$ ; a solução correspondente  $x = (2, 6, 0, 0, 6, 2, 0)'$  em  $\mathbb{R}^7$  também é não-degenerada, pois possui exatamente  $n - m = 7 - 4 = 3$  componentes nulas. Por outro lado, a solução  $x = (4, 0, 2)'$  em  $\mathbb{R}^3$  é degenerada, pois são ativas 4 restrições:  $x_1 + x_2 + 2x_3 = 8$ ,  $x_2 + 6x_3 = 12$ ,  $x_1 = 4$  e  $x_2 = 0$ ; respectivamente, a solução  $x = (4, 0, 2, 0, 0, 0, 6)'$  em  $\mathbb{R}^7$  também é degenerada, pois possui 4 componentes nulas.

**Degenerescência não é uma propriedade exclusivamente geométrica**

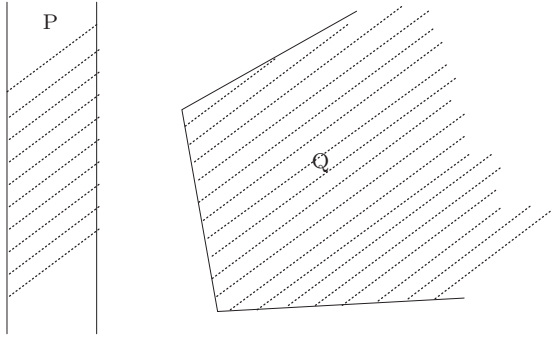
Note que a propriedade de degenerescência está ligada à representação do poliedro. Por exemplo, no poliedro canônico  $P = \{x \in \mathbb{R}^3 \mid x_1 - x_2 = 0, x_1 + x_2 + 2x_3 = 2, x \geq 0\}$ , a solução  $(1, 1, 0)'$  é não-degenerada (pois possui apenas uma componente nula) enquanto a solução  $(0, 0, 1)'$  é degenerada (possui duas componentes nulas); mas o mesmo conjunto  $P \subset \mathbb{R}^3$  pode ser descrito como  $\{x \in \mathbb{R}^3 \mid x_1 - x_2 = 0, x_1 + x_2 + 2x_3 = 2, x_1 + x_2 \geq 0, x_1 \leq 1, x_2 \leq 1\}$  (poliedro não-canônico) e nessa representação a solução  $(1, 1, 0)'$  é degenerada e  $(0, 0, 1)'$  é não-degenerada. (verifique!)

Outro exemplo: seja  $x^*$  uma solução básica viável não-degenerada de um poliedro canônico  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ , onde  $A \in \mathbb{R}^{m \times n}$ ; como  $x^*$  é básica e não-degenerada,  $x^*$  possui exatamente  $n - m$  zeros. Mas  $P$  pode ser colocado na forma geral  $P = \{x \in \mathbb{R}^n \mid Ax \geq b, -Ax \geq -b, x \geq 0\}$ ; nesta nova representação  $x^*$  continuará possuindo  $n - m$  zeros, e além destes mais  $2m$  restrições ativas ( $Ax \geq b$  e  $-Ax \geq -b$ ), num total de  $n + m$  restrições ativas, o que mostra que  $x^*$  é degenerada em relação a esta representação.

## 4 Existência de vértices

Note que nem todo poliedro possui vértices. Por exemplo, um semi-espaço em  $\mathbb{R}^2$  não possui vértices; se  $A \in \mathbb{R}^{m \times n}$  e  $m < n$  então o poliedro  $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$  não possui solução básica viável.

**Definição 2.12** Dizemos que um poliedro  $P \subset \mathbb{R}^n$  contém uma reta se existem  $x \in P$  e  $d \in \mathbb{R}^n \setminus \{0\}$  tais que  $x + \lambda d \in P$  para qualquer  $\lambda \in \mathbb{R}$ .



P contém uma reta e não possui vértices  
Q possui vértices e não contém uma reta

O teorema a seguir mostra que é a propriedade acima caracteriza a (não) existência de vértices.

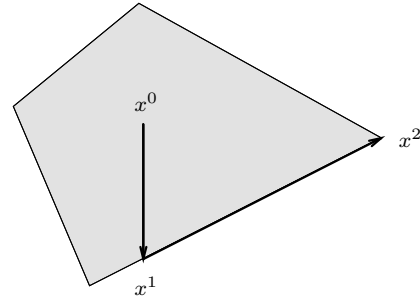
**Teorema 2.6** Seja  $P = \{x \in \mathbb{R}^n \mid Ax \geq b\} \neq \emptyset$ . São equivalentes:

1.  $P$  possui (pelo menos) um vértice;
2.  $A$  possui  $n$  linhas linearmente independentes;
3.  $P$  não contém uma reta.

**Prova.**

- $(1 \implies 2)$  Pelo teorema 2.3, se  $x \in P$  é vértice, então  $x$  é solução básica viável e, em particular, possui  $n$  restrições ativas linearmente independentes.

- (2  $\implies$  3) Suponha por contradição que  $P$  contém uma reta, ou seja, que existem  $x \in P$  e  $d \in \mathbb{R}^n \setminus \{0\}$  tais que  $A(x + \lambda d) \geq b$  para qualquer  $\lambda \in \mathbb{R}$ . Afirmamos que  $Ad = 0$ : se  $A_i d$  não fosse 0 existiriam valores de  $\lambda$  que violam  $A_i(x + \lambda d) \geq b_i$  (por exemplo,  $\lambda = \frac{b_i - A_i x - 1}{A_i d}$ ). Então  $d$  é ortogonal a todas as linhas da matriz, e como existem  $n$  dentre elas linearmente independentes temos  $d = 0$ , uma contradição.
- (3  $\implies$  1) Seja  $x^0 \in P$  e  $I = \{i \mid A_i x^0 = b_i\}$ . Se  $\{A_i \mid i \in I\}$  contém  $n$  linhas linearmente independentes,  $x^0$  é um vértice e temos o resultado. Do contrário, existe um vetor  $d \in \mathbb{R}^n \setminus \{0\}$  ortogonal ao espaço gerado pelos  $A_i$ ,  $i \in I$ . Por hipótese, a reta  $x^0 + \lambda d$ ,  $\lambda \in \mathbb{R}$  não está contida em  $P$ , e portanto existe um  $\lambda^* \neq 0$  tal que  $x^0 + \lambda^* d \in P$  e alguma restrição  $A_j x \geq b_j$  inativa em  $x^0$  ( $A_j x^0 > b_j$ ) torna-se ativa em  $x^0 + \lambda^* d$  ( $A_j(x^0 + \lambda^* d) = b_j$ ). Note que  $A_j$  não pode ser gerado por  $\{A_i \mid i \in I\}$ , visto que  $A_j d = \frac{1}{\lambda^*}(b_j - A_j x^0) \neq 0$ . Além disso, todas as restrições ativas em  $x^0$  permanecem ativas em  $x^0 + \lambda^* d$ , pois para  $i \in I$  temos  $A_i(x^0 + \lambda^* d) = A_i x^0 + \lambda^* A_i d = A_i x^0 + \lambda^* \cdot 0 = b_i$ .



Com isso temos um ponto  $x^1 = x^0 + \lambda^* d \in P$  que possui pelo menos uma restrição ativa a mais do que  $x^0$ , linearmente independente em relação às anteriores. Se  $x^1$  possui  $n$  restrições ativas l.i.,  $x^1$  é vértice e temos o resultado. Do contrário, podemos repetir o argumento acima e construir  $x^2 \in P$  com pelo menos uma restrição ativa a mais do que  $x^1$ , l.i. em relação às anteriores. No máximo depois de  $n$  iterações, o método descrito acima produzirá um vértice de  $P$ , o que mostra o resultado.

Um corolário direto do teorema acima é:

**Corolário 2.2** *Todo poliedro limitado não-vazio possui (pelo menos) um vértice. Todo poliedro canônico não-vazio possui (pelo menos um) vértice.*

**Prova.**

Para a primeira afirmação: um poliedro limitado não contém uma reta. Para a segunda: todo poliedro canônico está contido em  $\{x \in \mathbb{R}^n \mid x \geq 0\}$ , e este último conjunto não contém uma reta.

## 5 Otimalidade de vértices

**Teorema 2.7** *Seja o problema de programação linear*

$$(PL) \begin{cases} \min & c'x \\ \text{s.a} & x \in P, \end{cases}$$

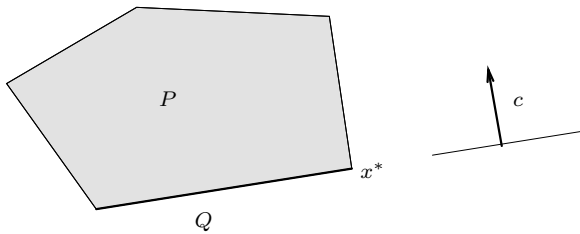
*e suponha que este problema admite uma solução ótima e que  $P$  possui pelo menos um vértice. Então existe um vértice de  $P$  que é solução ótima de  $(PL)$ .*

**Prova.**

Seja  $Q$  o conjunto de soluções ótimas de  $(PL)$ , ou seja,

$$Q = \{x \in P \mid c'x = v\},$$

onde  $v = \min\{c'x \mid x \in P\} \in \mathbb{R}$  e  $Q \neq \emptyset$  (pela hipótese de existência da solução ótima). Então  $Q$  também é um poliedro, não vazio, e  $Q$  não pode conter uma reta, visto que  $Q \subset P$ . Pelo teorema 2.6  $Q$  possui um vértice  $x^*$ ; vamos mostrar que  $x^*$  também é vértice de  $P$ , o que juntamente com  $c'x^* = v$  concluirá a demonstração.



Suponha por contradição que  $x^*$  não é vértice de  $P$ . Então existem  $y, z \in P$ ,  $y \neq z$  tais que  $x^* = \frac{1}{2}y + \frac{1}{2}z$ . Como  $y$  e  $z$

são viáveis,  $c'y \geq v$  e  $c'z \geq v$ ; além disso  $c'x^* = \frac{1}{2}c'y + \frac{1}{2}c'z = v$ , o que só é possível com  $c'y = c'z = v$ . Mas isso mostra que  $y, z \in Q$ , e portanto  $x^*$  não seria vértice de  $Q$ , uma contradição. Logo  $x^*$  é vértice de  $P$  e é solução ótima, o que conclui a prova.

Uma das hipóteses do teorema 2.7 é que exista uma solução ótima, para então concluir que existe um vértice ótimo. O próximo resultado completa o anterior, ainda no caso em que  $P$  possui vértices (não contém retas): ele afirma que se não houver um vértice ótimo então o problema é ilimitado (não possui solução ótima).

**Teorema 2.8** *Seja o problema de programação linear*

$$(PL) \begin{cases} \min & c'x \\ \text{s.a} & x \in P, \end{cases}$$

*e suponha que  $P$  possui pelo menos um vértice. Então ou o valor ótimo de  $(PL)$  é  $-\infty$ , ou existe um vértice de  $P$  que é solução ótima de  $(PL)$ .*

**Prova.**

O núcleo da demonstração é a seguinte propriedade: a menos que o problema seja ilimitado, a partir de qualquer ponto viável podemos obter um vértice com o valor da função objetivo melhor do que o ponto inicial; como o número de vértices é finito, o melhor deles será a solução ótima do problema.

Considere  $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$  (poliedro na forma geral) e seja  $x \in P$  qualquer. Se  $x$  é vértice, tomamos  $v = x$ . Do contrário, podemos escolher  $d \in \mathbb{R}^n \setminus \{0\}$  tal que  $A_i d = 0$ , para  $i \in I = \{i \mid A_i x = b_i\}$ , e podemos supor sem perda de generalidade que  $c'd \leq 0$  (do contrário, usaríamos  $-d$ ).

Agora existem duas possibilidades:  $c'd < 0$  ou  $c'd = 0$ .

No primeiro caso, ou a semi-reta  $\{x + \lambda d \mid \lambda \geq 0\}$  está contida em  $P$ , assumindo valores  $c'(x + \lambda d) = \underbrace{c'x}_{\text{cte}} + \lambda \underbrace{c'd}_{<0}$  e portanto o valor ótimo do problema é  $-\infty$ , ou existe um valor  $\lambda^* > 0$  tal que pelo menos mais uma restrição é ativa em  $x^1 = x + \lambda^* d \in P$  (como na demonstração do teorema 2.6) e além disso  $c'x^1 = c'x + \lambda^* c'd \leq c'x$ .

No segundo caso ( $c'd = 0$ ), como a reta  $\{x + \lambda d \mid \lambda \in \mathbb{R}\}$  não está contida em  $P$ , existe um  $\lambda^* \in \mathbb{R}$  tal que pelo menos

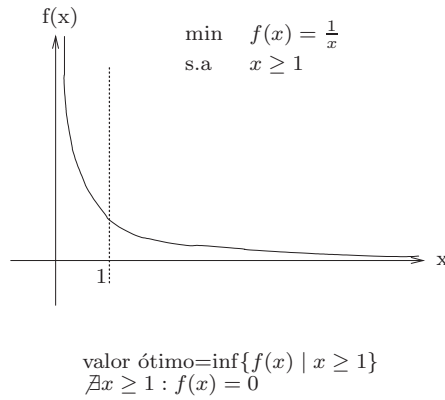


mais uma restrição é ativa em  $x^1 = x + \lambda^* d \in P$  e além disso  $c'x^1 = c'x + \lambda c'd = c'x$ .

Repetindo o argumento no máximo  $n$  vezes, ou teremos uma direção de ilimitação da função objetivo, ou teremos um vértice  $v^x$  tal que  $c'v^x \leq c'x$ .

Como o número de vértices é finito (corolário 2.1), existe um vértice  $v^*$  tal que  $c'v^* \leq c'v$  para qualquer outro vértice  $v$ . Então se o problema não é ilimitado,  $c'v^* \leq c'v^x \leq c'x, \forall x \in P$ , o que mostra que  $v^*$  é solução ótima do problema.

Um exemplo de programação não-linear mostra que a existência de valor ótimo real não implica na existência de solução ótima em geral:



Esta é uma propriedade forte que vale para o caso de programação linear. Com o auxílio dos teoremas anteriores, podemos generalizar o resultado e caracterizar todas as possibilidades de um problema de programação linear:

**Corolário 2.3 (modificado)** *Seja o problema de programação linear*

$$(PL) \begin{cases} \min & c'x \\ \text{s.a} & x \in P, \end{cases}$$

*Então exatamente uma das seguintes afirmações é verdadeira sobre o problema (PL):*

1.  $P = \emptyset$  e o valor ótimo é  $+\infty$ ;
2. Existe uma direção de ilimitação da função objetivo, e o valor ótimo é  $-\infty$ ;
3. Existe um vértice ótimo.

**Exercício 5.1** *Demonstre o teorema acima, formalizando o seguinte argumento e citando os teoremas utilizados: qualquer problema de programação linear (PL) pode ser transformado em um problema na forma canônica (PLC) que é “equivalente” no sentido de existir uma função injetora do conjunto viável de (PLC) para o conjunto viável de (PL) que preserva os valores das funções objetivo de (PL) e (PLC). Mostre que as conclusões do teorema 2.8 para o (PLC) podem ser transportadas para o (PL).*

**Exercícios sugeridos para o capítulo 2:** 2.1, 2.3-2.10, 2.12-2.17, 2.19 e 2.22.

### 3 O Método Simplex

Neste capítulo consideraremos sempre o problema de programação linear na forma canônica

$$(PLC) \begin{cases} \min & c'x \\ \text{s.a} & x \in P \end{cases}$$

onde

$$P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\},$$

$b \in \mathbb{R}^m$  e  $A \in \mathbb{R}^{m \times n}$  possui linhas linearmente independentes.

#### 3.1 Condições de otimalidade

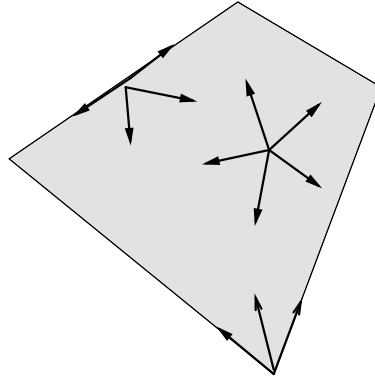
Uma estratégia comum a vários algoritmos de otimização consiste em construir uma sequência de pontos viáveis  $x^0, x^1, \dots, x^k$  que correspondam a melhorias sucessivas do valor da função objetivo, com a propriedade de que  $x^{i+1}$  está “próximo” de  $x^i$ ; se nenhum ponto nas proximidades de  $x^k$  melhoram o valor da função objetivo o método pára e a solução obtida é *localmente ótima*.

**Exercício 3.1** *Mostre que se a solução  $x^*$  é localmente ótima para o PLC, ou seja, se existe um  $\varepsilon > 0$  tal que  $c'x^* \leq c'(x^* + d)$  para qualquer  $d \in \{d \in \mathbb{R}^n \mid \|d\| < \varepsilon \text{ e } x^* + d \in P\}$ , então  $x^*$  é solução ótima do (PLC).*

A partir de um ponto  $x \in P$ , estamos interessados em considerar direções que não saem imediatamente do conjunto viável, como na figura abaixo:

---

<sup>1</sup>Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.



**Definição 3.1**  $d \in \mathbb{R}^n$  é chamada de **direção viável** a partir de  $x \in P$  se existe  $\theta > 0$  tal que  $x + \theta d \in P$ .

Observe que na definição acima  $x + \alpha d \in P$ ,  $\forall \alpha \in [0, \theta]$ , por convexidade. Se  $x$  é uma solução básica viável e  $\{B_1, \dots, B_m\}$  são os índices de uma base associada, temos

$$x_B = (x_{B_1}, \dots, x_{B_m})' = B^{-1}b = \left[ \begin{array}{c|c|c} A^{B_1} & \dots & A^{B_m} \end{array} \right]^{-1} b$$

e  $x_j = 0$  para todas as variáveis não básicas. Estamos interessados em restringir a busca ao conjunto das soluções básicas, pois este é um conjunto finito (corolário 2.1) que contém a solução ótima do PLC sempre que uma solução ótima existir (teorema 2.8).

Usaremos como noção de “proximidade” a definição de adjacência e buscaremos uma solução básica adjacente restringindo as direções  $d$  consideradas a manter ativas  $n-1$  dentre as restrições ativas em  $x$ : além das  $m$  restrições que formam  $Ax = b$ , manteremos ativas  $n - m - 1$  das restrições  $x_i \geq 0$  ativas ( $x_i = 0$ ); como são  $n - m$  as variáveis não-básicas, podemos atingir este objetivo exigindo que  $d_j = 1$  para alguma variável  $x_j$  não-básica, e  $d_i = 0$  para todas as demais variáveis não-básicas. O novo ponto  $y = x + \theta d$  terá que satisfazer

$$\begin{aligned} y_B &= x_B + \theta d_B \\ y_j &= x_j + \theta d_j = 0 + \theta = \theta \\ y_i &= x_i + \theta d_i = 0 + \theta \cdot 0 = 0, \forall i \neq j, B_1, \dots, B_m. \end{aligned}$$

Como estamos interessados em soluções viáveis, exigiremos que  $y \in P$  para algum valor de  $\theta > 0$ . Por um lado, isso força  $y$  a satisfazer  $Ay = b$ , ou seja

$$b = Ay = A(x + \theta d) = Ax + \theta Ad = b + \theta Ad \implies Ad = 0;$$

lembrando que  $d_j = 1$  e  $d_i = 0, \forall i \neq j, B_1, \dots, B_m$ , temos  $0 = Ad = \left(\sum_{i=B_1, \dots, B_m} A^i d_i\right) + d_j A^j + \left(\sum_{i \neq j, B_1, \dots, B_m} A^i d_i\right) = Bd_B + A^j$  e portanto

$$d_B = -B^{-1}A^j.$$

Por outro lado, precisamos garantir que  $y \geq 0$ , ou seja, que  $x + \theta d \geq 0$ . Esta análise será feita na próxima seção.

A direção  $d$  acima construída, obtida fixando-se  $d_j = 1$  e  $d_i = 0, \forall i \neq j, B_1, \dots, B_m$ , é chamada de  $j$ -ésima direção básica. Com respeito a esta direção podemos distinguir duas situações:

1.  $x$  é uma solução não-degenerada: neste caso  $x_B > 0$ , e conseqüentemente para  $\theta > 0$  pequeno teremos  $x + \theta d \geq 0$ . Em particular,  $d$  é uma direção viável de acordo com a definição 3.1;
2.  $x$  é degenerada: então existe um índice  $B_i$  tal que  $x_{B_i} = 0$ , e neste caso a direção construída pode ou não ser viável, dependendo do sinal de  $d_{B_i}$ .

Queremos observar o comportamento da função objetivo nas direções básicas. A variação do valor da função objetivo andando-se na  $j$ -ésima direção básica  $d$  será

$$c'(x + \theta d) - c'x = \theta c'd = \theta(c_j + c'_B d_B) = \theta(c_j - c'_B B^{-1}A^j).$$

**Definição 3.2** *Seja  $x$  uma solução básica e  $B$  a matriz básica associada. Para cada índice  $j$  definimos o  $j$ -ésimo custo reduzido como*

$$\bar{c}_j = c_j - c'_B B^{-1}A^j.$$

**Exemplo 3.1** *Considere o PLC*

$$\begin{array}{llllll} \min & c_1 x_1 & + & c_2 x_2 & + & c_3 x_3 & + & c_4 x_4 \\ \text{s.a} & x_1 & + & x_2 & + & x_3 & + & x_4 & = & 2 \\ & 2x_1 & & & + & 3x_3 & + & 4x_4 & = & 2 \\ & & & & & x_1, \dots, x_4 & \geq & 0 & & . \end{array}$$

*Considere a base associada às variáveis  $\{x_1, x_2\}$ , com  $B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$  e a solução correspondente  $x = (1, 1, 0, 0)'$ , que é não-degenerada. A direção básica associada a  $j = 3$  é obtida fazendo-se  $d_3 = 1, d_4 = 0$  e calculando-se*

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = -B^{-1}A^3 = -\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix};$$

O custo reduzido  $\bar{c}_3$  corresponde à variação da função objetivo entre os pontos  $x$  e  $x + d$ , cujo valor é

$$\bar{c}_3 = c_3 - c'_B B^{-1} A^3 = c_3 + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}' \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} = -\frac{3}{2}c_1 + \frac{1}{2}c_2 + c_3.$$

Note que na definição de custo reduzido, se  $j$  é um índice básico, então  $\bar{c}_j = c_j - c'_B B^{-1} A^j = c_j - c'_B e^j = c_j - c_j = 0$ . Se existe uma variável não-básica  $x_j$  tal que  $\bar{c}_j < 0$  e é possível andar na  $j$ -ésima direção básica sem sair do poliedro, então é possível passar para um ponto melhor. Do contrário, a solução considerada é ótima, como veremos a seguir.

**Teorema 3.1** *Seja  $x$  uma solução básica viável e  $\bar{c}$  o vetor de custos reduzidos. Então*

1. *Se  $\bar{c} \geq 0$ ,  $x$  é solução ótima do PLC;*
2. *Se  $x$  é ótima e não-degenerada, então  $\bar{c} \geq 0$ .*

**Prova.**

Seja  $B$  a matriz básica associada a  $x$ . Seja  $y \in P$  uma solução viável qualquer, e considere  $d = y - x$ . Temos  $Ax = Ay = b \implies Ad = 0$ , e conseqüentemente  $\sum_{i=1}^n A^i d_i = Bd_B + \sum_{i \in N} A^i d_i = 0$ , onde  $N$  é o conjunto de índices das variáveis não-básicas. Como  $B$  é inversível, podemos escrever

$$d_B = - \sum_{i \in N} B^{-1} A^i d_i$$

e

$$c'd = c'_B d_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c'_B B^{-1} A^i) d_i = \sum_{i \in N} \bar{c}_i d_i.$$

1. Note que  $d_i = y_i - x_i \geq 0$ , pois  $x_i = 0$  (variável não-básica) e  $y_i \geq 0$  (viabilidade); além disso  $\bar{c}_i \geq 0$  por hipótese. Assim  $c'd = \sum_{i \in N} \bar{c}_i d_i \geq 0$ , ou seja  $c'y \geq c'x$ , o que mostra que  $x$  é ótimo.
2. Suponha por contradição que existe um índice (não-básico)  $j$  tal que  $\bar{c}_j < 0$  e seja  $d$  a  $j$ -ésima direção básica a partir de  $x$ . Como  $x$  é não-degenerada,  $x_B > 0$  e portanto existe um  $\theta > 0$  tal que  $x + \theta d \in P$ . Mas

$$c'(x + \theta d) = c'x + \theta c'd = c'x + \theta \left( \sum_{i \in N} \bar{c}_i d_i \right) = c'x + \theta \bar{c}_j < c'x,$$

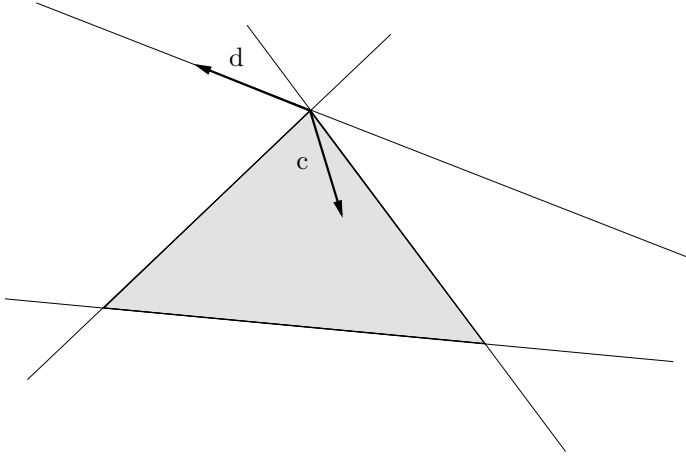
onde a última igualdade é obtida lembrando-se que  $d_j = 1$  e  $d_i = 0, \forall i \in N \setminus \{j\}$ ; mas a desigualdade obtida contradiz a otimalidade de  $x$ . Concluimos, portanto, que  $\bar{c} \geq 0$ .

De acordo com o teorema acima, para concluir que uma solução básica não-degenerada é ótima basta calcular o custo reduzido para as  $n - m$  direções básicas e verificar o sinal do mesmo. Isso motiva a seguinte

**Definição 3.3** Uma matriz básica é dita **ótima** se  $B^{-1}b \geq 0$  e  $\bar{c}' = c' - c'_B B^{-1}A \geq 0'$ .

No entanto, no caso de uma solução básica degenerada, não podemos afirmar a partir do teorema que se existe  $\bar{c}_j < 0$  então a solução não é ótima: a  $j$ -ésima direção básica poderia não ser viável. Em outras palavras, custos reduzidos não-negativos são uma condição *suficiente* de otimalidade, mas não *necessária*. O caso degenerado será tratado detalhadamente na seção seguinte.

**Exercício 3.2** Construa um exemplo de uma solução ótima degenerada e uma base que não satisfaz  $\bar{c} \geq 0$ , traduzindo o desenho a seguir para uma formulação como PLC;



## 3.2 Desenvolvimento do método simplex

Consideraremos inicialmente o caso em que todas as soluções básicas do poliedro são não-degeneradas. Seja  $x$  uma solução básica e  $d$  a  $j$ -ésima direção básica a partir de  $x$ , satisfazendo  $d_B = -B^{-1}A^j$ ,  $d_j = 1$  e  $d_i = 0, \forall i \neq j, B_1, \dots, B_m$ . Estamos interessados em analisar a viabilidade da solução  $x + \theta d$ : lembrando que  $Ad = 0$  sabemos que  $A(x + \theta d) = b, \forall \theta \in \mathbb{R}$ ; restaria verificar que  $x + \theta d \geq 0$ . Temos duas situações possíveis:

1.  $d \geq 0$ : neste caso  $x + \theta d \geq 0, \forall \theta \geq 0$ ;
2.  $d$  possui alguma coordenada negativa: neste caso, existe um índice  $B_i$  tal que  $d_{B_i} < 0$ , e o valor máximo de  $\theta$  que satisfaz  $x + \theta d \geq 0$  será

$$\theta^* = \min_{B_i | d_{B_i} < 0} \left( -\frac{x_{B_i}}{d_{B_i}} \right);$$

**Exemplo 3.2** Voltando ao exemplo 3.1, tínhamos a solução básica  $x = (1, 1, 0, 0)'$  associada à base  $\{x_1, x_2\}$  e a direção básica correspondente a  $j = 3$  era  $d = (-\frac{3}{2}, \frac{1}{2}, 1, 0)'$ . Supondo que  $c = (2, 0, 0, 0)'$  teremos  $\bar{c}_3 = -3 < 0$  e conseqüentemente a função objetivo decresce ao longo da direção  $d$ . Teremos

$$\theta^* = \min_{B_i | d_{B_i} < 0} \left( -\frac{x_{B_i}}{d_{B_i}} \right) = -\frac{x_1}{d_1} = \frac{2}{3},$$

e conseqüentemente

$$y = x + \theta^* d = x + \frac{2}{3} d = (0, \frac{4}{3}, \frac{2}{3}, 0)'.$$

Note que as colunas  $A^2$  e  $A^3$  associadas a  $y_i \neq 0$  são linearmente independentes, e  $y$  é a solução básica associada à base  $\{x_2, x_3\}$ . Em relação à base associada a  $x$  temos a entrada de  $x_3$  na base e a saída de  $x_1$ : a entrada determinada pela escolha da direção básica associada a  $j = 3$  e a saída determinada pelo índice que define  $\theta^*$ , ou seja, pela primeira restrição a se tornar ativa no caminho de  $x$  para  $x + \theta d$  (a restrição  $x_1 \geq 0$ ). Note que  $y$  e  $x$  são adjacentes, pois possuem  $m - 1 = 1$  variáveis básicas em comum (a variável  $x_2$ ).

Supondo que  $\theta^* = -\frac{x_{B_k}}{d_{B_k}} > 0$ , teremos o novo ponto  $y = x + \theta^* d$  que satisfaz

$$\begin{aligned} y_B &= x_B + \theta^* d_B = x_B + \frac{x_{B_k}}{d_{B_k}} B^{-1} A^j \geq 0 \\ y_{B_k} &= x_{B_k} + \theta^* d_{B_k} = x_{B_k} - \frac{x_{B_k}}{d_{B_k}} d_{B_k} = 0 \\ y_j &= x_j + \theta^* d_j = -\frac{x_{B_k}}{d_{B_k}} > 0 \\ y_i &= x_i + \theta^* d_i = 0 + \theta^* 0 = 0, \forall i \neq j, B_1, \dots, B_m. \end{aligned}$$

Assim, substituindo a variável  $x_{B_k}$  pela variável  $x_j$  na base, obtemos a nova base  $\{B_1, \dots, B_{k-1}, j, B_{k+1}, \dots, B_m\}$  associada a  $y$ , com a matriz básica

$$\bar{B} = \left[ \begin{array}{c|c|c|c|c|c|c} A^{B_1} & \dots & A^{B_{k-1}} & A^j & A^{B_{k+1}} & \dots & A^{B_m} \end{array} \right].$$



### Teorema 3.2

1. As colunas  $A^{B_i}$  e  $A^j$  são linearmente independentes e  $\bar{B}$  é uma matriz básica;
2. O vetor  $y = x + \theta^* d$  é uma solução básica viável associada à matriz básica  $\bar{B}$ .

**Prova.**

1. Suponha por contradição que  $A^{B_i}$ ,  $i = 1, \dots, k-1, k+1, \dots, m$  e  $A^j$  sejam linearmente dependentes, ou seja, que exista  $\lambda \neq 0$  tal que  $\bar{B}\lambda = 0$ . Então  $B^{-1}\bar{B}\lambda = 0$ , ou seja, os vetores  $B^{-1}A^{B_i} = e^i$ ,  $i = 1, \dots, k-1, k+1, \dots, m$  e  $B^{-1}A^j = -d_B$  são linearmente dependentes. Mas  $(B^{-1}A^j)_k = -d_{B_k} > 0$  por construção, e portanto  $\lambda$  satisfaz

$$\left\{ \begin{array}{ccccccc} \lambda_1 & & -\lambda_k d_{B_1} & & & & = 0 \\ & \ddots & \vdots & & & & \vdots \\ & & \lambda_{k-1} & -\lambda_k d_{B_{k-1}} & & & = 0 \\ & & & -\lambda_k d_{B_k} & & & = 0 \\ & & & -\lambda_k d_{B_{k+1}} & \lambda_{k+1} & & = 0 \\ & & & \vdots & & \ddots & \vdots \\ & & & -\lambda_k d_{B_m} & & & \lambda_m = 0. \end{array} \right.$$

Da linha  $k$  concluímos que  $\lambda_k = 0$  e das outras  $\lambda = 0$ . Esta contradição estabelece o item 1.

2. Por  $Ad = 0$  temos  $Ay = b$  e pela definição de  $\theta^*$  temos  $y \geq 0$ . Além disso  $y_i = 0$ ,  $\forall i \neq \bar{B}_1, \dots, \bar{B}_m$  e as colunas  $A^{\bar{B}_1}, \dots, A^{\bar{B}_m}$  são linearmente independentes pela parte 1. Isso estabelece a afirmação de 2.

Além disto, para o caso não-degenerado que estamos considerando temos que  $y \neq x$  (pois  $d \neq 0$  e  $\theta^* > 0$ ),  $y$  é adjacente a  $x$  (pois possui  $m-1$  variáveis básicas em comum) e  $c'y < c'x$  (pois  $c'd < 0$ ).

Podemos resumir **uma iteração típica do simplex** como segue:

1. No início da iteração conhecemos uma base formada pelas colunas  $A^{B_1}, \dots, A^{B_m}$  e o vértice associado  $x$ .
2. Calcule os custos reduzidos  $\bar{c}_j = c_j - c'_B B^{-1}A^j$  para todos os índices não-básicos  $j$ . Se  $\bar{c} \geq 0$ ,  $x$  é uma solução ótima e o algoritmo pára. Do contrário, escolha algum  $j$  tal que  $\bar{c}_j < 0$ .

3. Calcule  $u = B^{-1}A^j$ . Se  $u \leq 0$  tem-se  $\theta^* = \infty$ , logo o custo ótimo é  $-\infty$  e o algoritmo pára.
4. Se existe algum componente  $u_i > 0$ , define-se

$$\theta^* = \min_{u_i > 0} \frac{x_{B_i}}{u_i}.$$

5. Seja  $k$  um índice tal que  $\theta^* = \frac{x_{B_k}}{u_k}$ . Forme uma nova base substituindo  $A^{B_k}$  por  $A^j$ . Os valores da nova solução básica  $y$  são dados por  $y_j = \theta^*$  e  $y_{B_i} = x_{B_i} - \theta^* u_i$ ,  $\forall i \neq k$  (coordenadas básicas), além de  $y_{B_k} = 0$  (não-básica).

A inicialização do simplex será discutida na seção 3.5. Por enquanto consideramos dada uma base viável inicial (lembre-se que para problemas canônicos viáveis em que  $A$  possui linhas l.i. sempre existe alguma base viável).

**Teorema 3.3** *Suponha que o conjunto viável é não-vazio e que toda solução básica é não-degenerada. Nestas condições o método simplex termina em um número finito de iterações, com uma das seguintes propriedades:*

1.  $B$  é uma base ótima e  $x$  é uma solução ótima; ou
2. Existe  $d$  tal que  $Ad = 0$ ,  $d \geq 0$  e  $c'd < 0$  e o valor ótimo do problema é  $-\infty$ .

**Prova.**

A cada iteração intermediária o método passa de uma solução básica viável  $x$  para uma solução básica viável  $y$  que satisfaz  $c'y < c'x$ . Como o número de soluções básicas é finito (corolário 2.1) o método pára obrigatoriamente depois de um número finito de passos. Há dois critérios de parada: No passo 2, se todos os custos reduzidos são não-negativos, então  $B$  é uma base ótima (teorema 3.1 (a)). No passo 3, se  $u \leq 0$  a  $j$ -ésima direção básica satisfaz  $Ad = 0$  (por definição),  $d \geq 0$  (porque  $u \leq 0$ ) e  $c'd = \bar{c}_j < 0$  (pelo passo 2); assim  $x + \theta d \in P$  para qualquer  $\theta \geq 0$  e  $\lim_{\theta \rightarrow \infty} c'(x + \theta d) = -\infty$ , o que mostra que o problema é ilimitado.

Note que este teorema é uma versão construtiva do teorema 2.8.

## O método simplex para problemas degenerados

Considere a aplicação do método simplex descrito acima em problemas degenerados. Duas situações novas podem aparecer durante a execução do algoritmo:

1. Se o vértice  $x$  do início da iteração é degenerado, pode existir algum  $u_i > 0$  com correspondente  $x_{B_i} = 0$ , e portanto  $\theta^* = 0$ ; isto acontece quando a  $j$ -ésima direção básica é inviável, e neste caso a “nova” solução básica é  $y = x$ . A base associada a  $y$  no entanto é diferente da de  $x$ , pois foi obtida desta última através da substituição da coluna  $A^{B_k}$  por  $A^j$ ; em particular a conclusão do teorema 3.2 ainda é válida.
2. Mesmo que tenhamos  $\theta^* > 0$ , pode acontecer de uma ou mais componentes básicas se anularem no novo ponto  $x + \theta^* d$ ; isso acontece quando há empates do tipo  $\frac{x_{B_l}}{u_l} = \frac{x_{B_k}}{u_k} = \theta^*$ . Como apenas uma variável pode sair da base, as outras permanecerão na base com valor 0, e portanto o novo vértice  $y$  será degenerado.

Mudanças de bases associadas ao mesmo vértice podem levar o algoritmo a descobrir eventualmente uma direção básica viável com custo reduzido negativo. Por outro lado, pode-se obter uma repetição cíclica de bases associadas a um mesmo vértice, e o algoritmo não terminar nunca. Este fenômeno é chamado de *ciclagem* e pode ser evitado através de escolhas cuidadosas das variáveis que entram e saem da base nos casos de empate.

### Seleção do pivô

Nos passos 2 e 5 do simplex não estão determinadas as escolhas da variável que entra na base (índice  $j$  tal que  $\bar{c}_j < 0$ ) e da variável que sai da base (índice  $k$  tal que  $\theta^* = \frac{x_{B_k}}{u_k}$ ). Regras determinísticas de escolha destas variáveis são chamadas de *regras de pivotação*.

Com relação à variável que *entra* na base, algumas escolhas são naturais:

1. Escolha o índice  $j$  associado ao custo reduzido  $\bar{c}_j$  mais negativo. Esta escolha está associada à maior *taxa de decréscimo* da função objetivo; o decréscimo real depende do tamanho de passo  $\theta^*$  que daremos naquela direção. Outra possibilidade é:

2. Escolha o índice  $j$  com o valor de  $\theta_j^* \bar{c}_j$  mais negativo (aqui  $\theta_j^*$  é o valor de  $\theta^*$  associado à  $j$ -ésima direção básica). Esta escolha nos levará ao vizinho  $y$  de menor custo.

Estas duas regras são computacionalmente caras, pois é necessário calcular o custo reduzido associado a todas as variáveis não-básicas. Uma regra muito simples e que também é muito eficiente na prática é a **regra de Bland**, ou regra do menor índice, que escolhe o primeiro  $j$  (na ordem natural dentre os índices não-básicos) que satisfaz  $\bar{c}_j < 0$ .

A regra de Bland também pode ser aplicada na escolha da variável que sai da base: basta escolher o menor índice  $B_k$  que satisfaz  $\theta^* = \frac{x_{B_k}}{u_k}$ .

Veremos na seção 3.4 que o uso da regra de Bland garante que não ocorre ciclagem em problemas degenerados.

### 3.3 Implementações do método simplex

Vamos discutir algumas alternativas de implementação do método simplex que diferem essencialmente no modo como as contas (por exemplo custos reduzidos e direções básicas) são realizadas e na quantidade de informação que é aproveitada de uma iteração para a iteração seguinte.

Vamos contabilizar o número aproximado de operações em cada iteração usando a notação  $\mathcal{O}(f(n))$ , que pode ser entendida – *informalmente* – como a ordem ou taxa de crescimento da função; assim  $\mathcal{O}(n)$  representa uma função qualquer que cresce (no máximo) linearmente e  $\mathcal{O}(n^2)$  representa uma função com taxa de crescimento (no máximo) quadrático. Por exemplo, para calcular o produto interno  $p'b$  onde  $p, b \in \mathbb{R}^m$  são necessárias  $\mathcal{O}(m)$  operações aritméticas; o produto matriz-vetor  $Bb$  requer  $\mathcal{O}(m^2)$  operações; calcular a inversa de uma matriz  $B$  ou resolver um sistema linear da forma  $Bx = b$  requer  $\mathcal{O}(m^3)$  operações aritméticas.

#### Implementação ingênua

A implementação mais direta do método simplex como formulado na seção anterior leva em consideração que apenas os índices  $B_1, \dots, B_m$  das variáveis básicas são conhecidos no início de cada iteração. A matriz  $B$  é formada a partir das colunas  $A^{B_1}, \dots, A^{B_m}$  e calcula-se o vetor auxiliar

$$p' = c'_B B^{-1}$$

resolvendo-se o sistema linear  $p'B = c'_B$ ; este vetor é chamado de vetor de *multiplicadores* associados à base  $B$ . O custo reduzido  $\bar{c}_j = c_j -$

$c'_B B^{-1} A^j$  associado à variável  $x_j$  é então calculado como

$$\bar{c}_j = c_j - p' A^j.$$

Dependendo da regra de pivotação utilizada pode ser necessário calcular todos os custos reduzidos, ou então calcular um por vez até encontrar o primeiro  $\bar{c}_j$  negativo. Supondo que a variável  $x^j$  seja escolhida para entrar na base, calculamos  $u = B^{-1} A^j$  resolvendo o sistema  $Bu = A^j$ , e determinamos a  $j$ -ésima direção básica. Finalmente calculamos  $\theta^*$  e a variável a sair da base, para determinar a nova solução básica viável.

Note que para calcular  $p$  e  $u$  são necessárias  $\mathcal{O}(m^3)$  operações aritméticas. Calcular o custo reduzido de uma variável custa  $\mathcal{O}(m)$  operações, e como no pior caso teremos que calcular todos os custos reduzidos, teríamos um gasto de  $\mathcal{O}(m(n - m)) = \mathcal{O}(mn)$ . No total, teremos um total de operações aritméticas de  $\mathcal{O}(m^3 + mn)$ .

## Método simplex revisado

A parte computacionalmente mais custosa da implementação ingênua do simplex consiste em resolver dois sistemas de equações. Alternativamente, poderíamos considerar que a matriz  $B^{-1}$  é conhecida no início de uma iteração qualquer: isso tornaria os cálculos de  $p$  e  $u$  simples produtos matriz-vetor com um custo de  $\mathcal{O}(m^2)$  operações aritméticas. Para esta alternativa ser viável, precisamos de um método eficiente de atualização de  $B^{-1}$  nas mudanças de base ao final de cada iteração.

Seja  $B = \left[ \begin{array}{c|c|c} A^{B_1} & \dots & A^{B_m} \end{array} \right]$  a matriz básica disponível no início de uma iteração e  $\bar{B} = \left[ \begin{array}{c|c|c|c|c|c|c} A^{B_1} & \dots & A^{B_{k-1}} & A^j & A^{B_{k+1}} & \dots & A^{B_m} \end{array} \right]$

a matriz que queremos calcular para a próxima iteração. Estas duas matrizes possuem  $n - 1$  colunas em comum; é de se esperar que as matrizes  $B^{-1}$  e  $\bar{B}^{-1}$  também possuam algum tipo de informação em

comum. Em particular,

$$B^{-1}\bar{B} = \begin{bmatrix} 1 & 0 & \cdots & u_1 & \cdots & 0 \\ 0 & 1 & \cdots & u_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & & u_k & & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_m & \cdots & 1 \end{bmatrix};$$

portanto se conseguirmos descobrir uma matriz  $Q$  tal que  $QB^{-1}\bar{B} = I$ , então poderemos calcular  $\bar{B}^{-1} = QB^{-1}$  usando apenas um produto de matrizes ( $\mathcal{O}(m^2)$  operações).

**Definição 3.4** *Dada uma matriz qualquer, a operação que consiste em somar um múltiplo de uma linha qualquer a uma outra linha (ou à mesma linha) é chamada de **operação elementar de linhas**.*

**Exemplo 3.3** *Seja  $Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$  e  $C = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 5 & 6 \end{bmatrix}$  e considere o pro-*

*duto  $QC = \begin{bmatrix} 11 & 14 \\ 2 & 4 \\ 5 & 6 \end{bmatrix}$ . Note que a matriz  $Q$  representa a operação*

*elementar de linhas que corresponde a somar à primeira linha de  $C$  o dobro da terceira linha.*

Em geral, somar à  $i$ -ésima linha de uma matriz  $C$  o produto da  $j$ -ésima linha da mesma matriz por  $\beta$  é equivalente a multiplicar  $C$  à esquerda pela matriz  $Q = I + D_{ij}$ , onde  $D_{ij}$  é uma matriz com  $\beta$  na posição  $(i, j)$  e zero em todas as outras posições. Com exceção do caso (degenerado)  $i = j$  e  $\beta = -1$ , que corresponde a apagar completamente a informação de uma das linhas não será utilizado, temos em todos os outros casos  $\det(Q) \neq 0$  e portanto  $Q$  é inversível. Note ainda que para calcular o produto  $QC$  são necessárias apenas  $\mathcal{O}(m)$  operações.

Considere agora que aplicamos uma seqüência de operações elementares de linhas representadas pelas matrizes  $Q_1, \dots, Q_K$ ; o resultado final será equivalente a multiplicar à esquerda pela matriz  $Q = Q_K \cdots Q_1$ .

Examinando a estrutura da matriz  $B^{-1}\bar{B}$  não é difícil encontrar uma seqüência de operações elementares de linhas que transformam a matriz  $B^{-1}\bar{B}$  na identidade:

1. Divida a  $k$ -ésima linha por  $u_k$  (lembre-se que  $u_k = -d_{B_k} > 0$ );
2. para cada  $i \neq k$  subtraia a  $k$ -ésima linha multiplicada por  $u_i$ .

A seqüência destas operações equivale a multiplicar  $B^{-1}\bar{B}$  à esquerda por uma matriz  $Q$  e isso garante, como vimos antes, que  $\bar{B}^{-1} = QB^{-1}$ ; isso mostra que para calcular  $\bar{B}^{-1}$  basta aplicar a seqüência de operações acima à matriz  $B^{-1}$ .

**Exemplo 3.4** Seja  $B^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{bmatrix}$ ,  $u = \begin{bmatrix} -4 \\ 2 \\ 2 \end{bmatrix}$  e suponha

que  $k = 3$ . Queremos então transformar o vetor  $u$  no vetor  $e^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ .

Para isso, dividimos a terceira linha por 2, somamos à primeira linha o produto da terceira linha por 4 e subtraímos da segunda linha o produto da terceira linha por 2. Com isso, obtemos

$$\bar{B}^{-1} = \begin{bmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -1.5 & -1 \end{bmatrix}.$$

Atualizando a inversa da matriz básica desta maneira obtemos uma implementação do simplex conhecida como *método simplex revisado*:

0. No início da iteração conhecemos uma base formada pelas colunas  $A^{B_1}, \dots, A^{B_m}$  e a matriz inversa de base  $B^{-1}$ , além do vértice associado  $x$ .
1. Calcule o vetor  $p' = c'_B B^{-1}$  os custos reduzidos  $\bar{c}_j = c_j - p' A^j$  para todos os índices não-básicos  $j$ . Se  $\bar{c} \geq 0$ ,  $x$  é uma solução ótima e o algoritmo pára. Do contrário, escolha algum  $j$  tal que  $\bar{c}_j < 0$ .
2. Calcule  $u = B^{-1} A^j$ . Se  $u \leq 0$  tem-se  $\theta^* = \infty$ , logo o custo ótimo é  $-\infty$  e o algoritmo pára.
3. Se existe algum componente  $u_i > 0$ , define-se

$$\theta^* = \min_{u_i > 0} \frac{x_{B_i}}{u_i}.$$

4. Seja  $k$  um índice tal que  $\theta^* = \frac{x_{B_k}}{u_k}$ . Forme uma nova base substituindo  $A^{B_k}$  por  $A^j$ . Os valores da nova solução básica  $y$  são dados por  $y_j = \theta^*$  e  $y_{B_i} = x_{B_i} - \theta^* u_i$ ,  $\forall i \neq k$  (coordenadas básicas), além de  $y_{B_k} = 0$  (não-básica).

5. Monte a matriz  $[B^{-1} \mid u] \in \mathbb{R}^{m \times (m+1)}$ . Aplique a seqüência de  $m$  operações elementares de linhas de modo a tornar a última coluna igual ao vetor  $e^k$  ( $k$ -ésimo vetor da base canônica); as primeiras  $m$  colunas da matriz correspondem à próxima inversa de base  $\bar{B}^{-1}$ .

Note que nesta versão o cálculo de  $p$  e  $u$  é feito por produtos matriz-vetor, que custam  $\mathcal{O}(m^2)$  operações aritméticas. A atualização de  $B^{-1}$  corresponde a uma seqüência de  $m$  operações elementares de linhas, cada uma das quais custando  $\mathcal{O}(m)$ , num total de  $\mathcal{O}(m^2)$ . Finalmente o cálculo dos custos reduzidos pode levar (no pior caso)  $\mathcal{O}(mn)$ . Cada iteração do simplex revisado custa portanto  $\mathcal{O}(m^2 + mn)$  operações aritméticas.

**Exemplo** *Considere o problema*

$$\begin{cases} \min & -x_1 & -x_2 & (+0x_3 & +0x_4) \\ \text{s.a} & x_1 & & +x_3 & = 1 \\ & x_1 & +x_2 & & +x_4 = 2 \\ & & & x \geq 0 \end{cases}$$

e considere a base inicial formada pelas variáveis  $x_3$  e  $x_4$ .

1. iteração:  $B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $x = (0, 0, 1, 2)'$ . Calculamos

$$p' = c'_B B^{-1} = (0, 0) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (0, 0),$$

$$\bar{c}_1 = c_1 - p' A^1 = -1 - (0, 0) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1$$

$$\text{e } \bar{c}_2 = c_2 - p' A^2 = -1 - (0, 0) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1.$$

Por Bland, escolhemos  $x_1$  para entrar na base e calculamos

$$u = B^{-1} A^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\text{e } \theta^* = \min\left\{\frac{x_{B_1}}{u_1}, \frac{x_{B_2}}{u_2}\right\} = \min\left\{\frac{1}{1}, \frac{2}{1}\right\} = 1,$$

de onde  $k = 1$  e  $x_3$  sai da base. Fazemos

$$x \longleftarrow x + \theta^*(1, 0, -1, -1)' = (1, 0, 0, 1)'$$

e calculamos a nova inversa da base  $\{x_1, x_4\}$  aplicando as operações elementares de linhas que transformam  $u$  em  $e^k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ :

$$[B^{-1} \mid u] = \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \right] \rightsquigarrow [\bar{B}^{-1} \mid e^k] = \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ -1 & 1 & 0 \end{array} \right].$$



2. iteração:  $B^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ ,  $x = (1, 0, 0, 1)'$ . Calculamos

$$p' = c'_B B^{-1} = (-1, 0) \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = (-1, 0),$$

$$\bar{c}_2 = c_2 - p' A^2 = -1 - (-1, 0) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1$$

$$e \bar{c}_3 = c_3 - p' A^3 = 0 - (-1, 0) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1.$$

Como  $\bar{c}_2 < 0$ ,  $x_2$  entra na base; calculamos

$$u = B^{-1} A^2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$e \theta^* = \frac{x_{B_2}}{u_2} = 1,$$

de onde  $k = 2$  e  $x_4$  sai da base. Fazemos

$$x \leftarrow x + \theta^*(0, 1, 0, -1)' = (1, 1, 0, 0)'$$

e calculamos a nova inversa da base  $\{x_1, x_2\}$  aplicando as operações elementares de linhas que transformam  $u$  em  $e^k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ .

Como  $u = e^k$  neste caso teremos  $\bar{B}^{-1} = B^{-1}$ : a base mudou (porque os índices mudaram), mas a matriz básica permaneceu igual (porque as colunas  $A^2$  e  $A^4$  são iguais).

3. iteração:  $B^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ ,  $x = (1, 1, 0, 0)'$ . Calculamos

$$p' = c'_B B^{-1} = (-1, -1) \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = (0, -1),$$

$$\bar{c}_3 = c_3 - p' A^3 = 0 - (0, -1) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

$$e \bar{c}_4 = c_4 - p' A^4 = 0 - (0, -1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1.$$

Como  $\bar{c} \geq 0$ , esta base é ótima. A solução ótima do problema é  $x^* = (1, 1, 0, 0)'$  com valor ótimo  $c'x^* = -2$ . Observe que como  $\bar{c}_3 = 0$  e a direção básica associada  $(-1, 1, 1, 0)'$  é viável, todos os pontos do segmento que une  $(1, 1, 0, 0)'$  até  $(0, 2, 1, 0)'$  são ótimos.

Desenhe o poliedro em  $\mathbb{R}^2$  associado a este problema (considerando  $x^3$  e  $x^4$  variáveis residuais) bem como o trajeto do simplex (sempre considerando as duas primeiras coordenadas de pontos e direções).

## Implementação do simplex tabular

A próxima implementação a ser discutida é conhecida por *simplex tabular*. Ao invés de supor que se conhece no início de uma iteração qualquer apenas a matriz  $B^{-1}$ , esta implementação mantém atualizados os vetores  $B^{-1}b$  (valores das variáveis básicas) e  $B^{-1}A^1, \dots, B^{-1}A^n$ , armazenados no formato matricial

$$B^{-1} [b \mid A]$$

conhecido como *tableau* do simplex. A coluna  $B^{-1}A^j = u$  associada à variável que entra na base e a  $k$ -ésima linha associada à variável  $x_{B_k}$  que sai da base são chamadas de *coluna* e *linha pivô*, respectivamente; o elemento  $(B^{-1}A^j)_k = u_k$  é o *elemento pivô*, e é sempre positivo (a menos que o algoritmo pare por causa do critério de ilimitação  $u \leq 0$ ).

A informação contida no tableau do simplex pode ser interpretada lembrando-se que, se  $B$  é uma matriz inversível qualquer, então

$$b = Ax \iff B^{-1}b = B^{-1}Ax,$$

para qualquer  $x \in \mathbb{R}^n$ . Os coeficientes que aparecem no tableau correspondem à representação do sistema linear acima.

Como no caso do simplex revisado, precisamos garantir que a informação transmitida para o proxima iteração, neste caso o tableau  $B^{-1} [b \mid A]$ , seja atualizada ao final de cada iteração. O tableau ao final de uma iteração qualquer deverá conter  $\bar{B}^{-1} [b \mid A]$ , o que pode ser obtido multiplicando-se o tableau original à esquerda pela matriz  $Q$  que satisfaz  $QB^{-1} = \bar{B}^{-1}$ . Ainda como no simplex revisado, isso é feito aplicando-se operações elementares de linhas que visam transformar a coluna  $B^{-1}A^j$  na  $k$ -ésima coluna da identidade, ou ainda, transformando o elemento pivô em 1 e as demais coordenadas daquela coluna em 0.

A escolha da variável que sai da base, assim como o cálculo de  $\theta^*$ , corresponde a considerar as frações  $\frac{(B^{-1}b)_k}{(B^{-1}A^j)_k} \left( = \frac{x_{B_k}}{u_k} \right)$  para  $(B^{-1}A^j)_k > 0$ . A determinação de ilimitação está associada à condição  $B^{-1}A^j \leq 0$ .

Para prover o tableau do simplex com a informação necessária para a decisão da variável que entra na base, acrescentamos uma linha extra (linha 0) contendo as informações  $-c'x = -c'_B x_B = -c'_B B^{-1}b$  (onde  $x$

é a solução associada à base atual) e o vetor  $\bar{c}' = c' - c'_B B^{-1}A$ :

$-c'_B x_B$	$\bar{c}_1$	$\cdots$	$\bar{c}_n$
$x_{B_1}$			
$\vdots$	$B^{-1}A^1$	$\cdots$	$B^{-1}A^n$
$x_{B_m}$			

e com isso temos no tableau toda a informação que o método simplex precisa. Veremos a seguir que a regra de atualização da linha 0 é idêntica à das outras linhas: adicionar à linha 0 um múltiplo da linha pivô de modo a zerar a entrada  $\bar{c}_j$ , que corresponde à variável que entra na base e fica justamente acima da coluna  $B^{-1}A^j = u$  no tableau.

No início de uma iteração qualquer do simplex a linha 0 contém

$$[-c'_B x_B \mid \bar{c}'] = [0 \mid c'] - p' [b \mid A]$$

onde  $p' = c'_B B^{-1}$ . Suponha que apliquemos a atualização proposta à linha 0, que corresponde a somar à linha 0 o produto da linha pivô por um múltiplo positivo (de modo a zerar o próximo  $\bar{c}_j$ ). A linha 0 teria então a nova forma

$$[0 \mid c'] - \bar{p}' [b \mid A]$$

onde  $\bar{p}'$  corresponde a uma nova combinação linear das linhas de  $[b \mid A]$ . Lembrando que a variável  $j$  ocupa o lugar da variável  $B_k$  na próxima base, teremos

$$c_j - p' A^j = c_{\bar{B}_k} - \bar{p}' A^{\bar{B}_k} = 0.$$

Considere os outros índices básicos  $B_i$ ,  $i \neq k$ . A  $B_i$ -ésima coluna do tableau é  $B^{-1}A^{B_i} = e^i$ , assim o elemento da linha  $k$  e coluna  $B_i$  é 0 e portanto a atualização proposta não vai alterar o elemento  $\bar{c}_{B_i}$ , que vai continuar sendo 0:

$$c_{\bar{B}_i} - \bar{p}' A^{\bar{B}_i} = 0, \quad i \neq k.$$

Com isso teremos  $c_{\bar{B}_i} - \bar{p}' A^{\bar{B}_i} = 0$ ,  $i = 1, \dots, m$  ou ainda  $c'_{\bar{B}} = \bar{p}' \bar{B}$ . Multiplicando os dois lados desta equação por  $\bar{B}^{-1}$  concluímos que  $\bar{p}' = c'_{\bar{B}} \bar{B}^{-1}$ , ou seja, a nova linha 0 corresponde exatamente ao tableau associado à nova base  $\bar{B}$ .

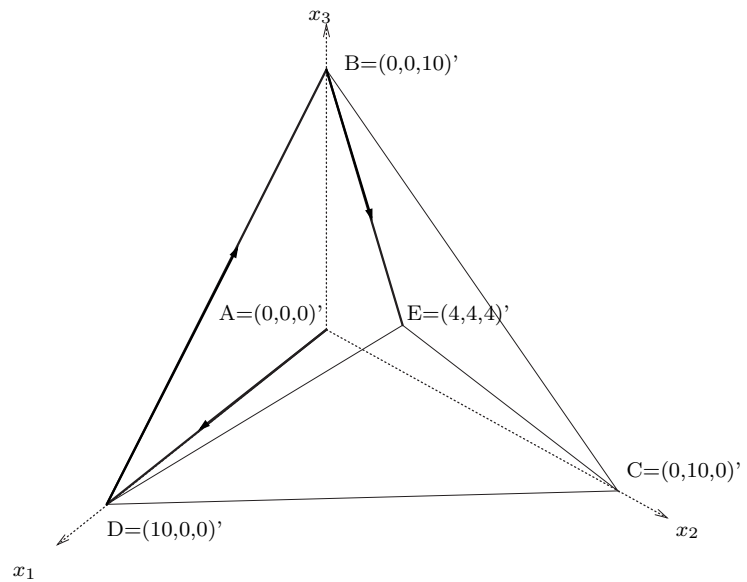
Uma iteração do simplex tabular fica assim:

0. É conhecido o tableau associado a uma base  $B$ .
1. Examinamos a linha 0 buscando um custo reduzido negativo. Se tal custo não existir, temos uma base ótima; do contrário, escolhemos  $j$  tal que  $\bar{c}_j < 0$ .

2. Se a  $j$ -ésima coluna  $B^{-1}A^j = u$  é não-positiva, o problema é ilimitado; do contrário, escolha  $k$  tal que  $\frac{x_{B_k}}{u_k} = \min_{u_i > 0} \{\frac{x_{B_i}}{u_i}\}$ . A variável  $x_{B_k}$  sai da base e a variável  $x_j$  entra em seu lugar.
3. Some a cada linha do tableau um múltiplo positivo da linha pivô (linha  $k$ ) de modo a tornar o elemento pivô ( $u_k$ ) igual a 1 e todo o resto da coluna pivô (coluna  $j$ ) igual a 0.

**Exemplo 3.5** Considere o problema

$$\left\{ \begin{array}{llll} \min & -10x_1 & - & 12x_2 & - & 12x_3 \\ s.a & x_1 & + & 2x_2 & + & 2x_3 \leq 20 \\ & 2x_1 & + & x_2 & + & 2x_3 \leq 20 \\ & 2x_1 & + & 2x_2 & + & x_3 \leq 20 \\ & & & & & x \geq 0 \end{array} \right.$$



que na forma canônica fica

$$\left\{ \begin{array}{ll} \min & (-10, -12, -12, 0, 0, 0)x \\ s.a & \begin{bmatrix} 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 20 \\ 20 \\ 20 \end{bmatrix} \\ & x \geq 0 \end{array} \right.$$

Usaremos como solução básica viável inicial  $x = (0, 0, 0, 20, 20, 20)'$  associada à base  $\{x_4, x_5, x_6\}$ . Como  $c_B = 0$  teremos  $[-c_B x_B \mid c -$

$c'_B B^{-1} A] = [0 \mid c]$ , e como  $B^{-1} = I$  teremos o tableau inicial

		$\downarrow$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	0		-10	-12	-12	0	0	0
$x_4 =$	20		1	2	2	1	0	0
$\leftarrow x_5 =$	20		2	1	2	0	1	0
$x_6 =$	20		2	2	1	0	0	1

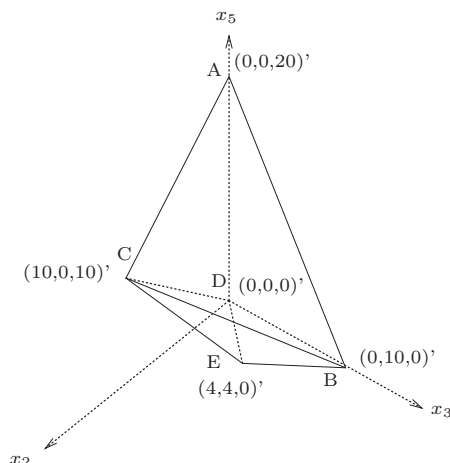
**1. iteração:** Escolhemos  $x_1$  para entrar na base ( $\bar{c}_1 = -10 < 0$ ), e notamos que o  $\min\{\frac{x_4}{u_1}, \frac{x_5}{u_2}, \frac{x_6}{u_3}\}$  é atingido para  $k = 2$  e  $k = 3$ . Escolhendo  $k = 2$  ( $x_5$  sai da base), teremos que pivotar sobre o elemento da linha 2, coluna 1 do tableau (indicado pelas setas). A nova base  $\{x_4, x_1, x_6\}$  está associada ao tableau

		$\downarrow$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	100		0	-7	-2	0	5	0
$\leftarrow x_4 =$	10		0	1.5	1	1	-0.5	0
$x_1 =$	10		1	0.5	1	0	0.5	0
$x_6 =$	0		0	1	-1	0	-1	1

e à solução  $x = (10, 0, 0, 10, 0, 0)'$ . Note que os números do tableau correspondem ao problema  $\min -7x_2 - 2x_3 + 5x_5$  sujeito a

$$\begin{aligned} 10 &= 1.5x_2 + x_3 + x_4 - 0.5x_5 \\ 10 &= x_1 + 0.5x_2 + x_3 + 0.5x_5 \\ 0 &= x_2 - x_3 - x_5 + x_6 \end{aligned}$$

que pode ser visualizado considerando-se as variáveis  $x_1$ ,  $x_4$  e  $x_6$  como residuais:



**2. iteração:** Voltando ao simplex, vamos escolher  $\bar{c}_3 = -2 < 0$  para entrar na base (por Bland escolheríamos  $\bar{c}_2 = -7 < 0$ ). Há um empate entre  $\frac{x_4}{u_1}$  e  $\frac{x_1}{u_2}$ ; escolhemos  $k = 1$  como o índice que realiza o mínimo ( $x_4$  sai da base). Pivotando sobre o elemento da linha 1, coluna 3, temos

			$\downarrow$				
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	120	0	-4	0	2	4	0
$x_3 =$	10	0	1.5	1	1	-0.5	0
$x_1 =$	0	1	-1	0	-1	1	0
$\leftarrow x_6 =$	10	0	2.5	0	1	-1.5	1

associado à solução  $x = (0, 0, 10, 0, 0, 10)'$ .

**3. iteração:** Neste último tableau teremos  $j = 2$  (o único índice tal que  $\bar{c}_j < 0$ ) e  $k = 3$  (pois  $\frac{10}{2.5} < \frac{10}{1.5}$ ); assim  $x_2$  entra na base e  $x_6$  sai da base. O tableau final é

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	136	0	0	0	3.6	1.6	1.6
$x_3 =$	4	0	0	1	0.4	0.4	-0.6
$x_1 =$	4	1	0	0	-0.6	0.4	0.4
$x_2 =$	4	0	1	0	0.4	-0.6	0.4

que corresponde à solução ótima  $x^* = (4, 4, 4, 0, 0, 0)'$  (pois  $\bar{c} \geq 0$ ). O caminho traçado pelo simplex pode ser acompanhado na figura do poliedro original como a sequência de pontos A-D-B-E. Esta sequência depende das escolhas das variáveis que entram e saem da base: se na segunda iteração tivéssemos usado Bland e escolhido  $x_2$  para entrar na base, por exemplo, teríamos uma pivotação degenerada (linha 3, coluna 2) e passaríamos da base  $\{x_4, x_1, x_6\}$  para a base  $\{x_4, x_1, x_2\}$ , associadas ao mesmo ponto  $D = (10, 0, 0, 10, 0, 0)'$ . O caminho seria A-D-E, mas o número de iterações continuaria igual. Como o simplex só modifica um índice básico por iteração, é impossível passar da base inicial  $\{x_4, x_5, x_6\}$  para a base ótima  $\{x_1, x_2, x_3\}$  em menos de 3 iterações.

## Complexidade do simplex tabular e comparação com o revisado

O simplex tabular mantém atualizado o tableau que tem tamanho  $\mathcal{O}(mn)$ . A atualização de cada entrada do tableau usa um número constante de operações (uma soma/subtração e um produto), portanto

a atualização do tableau inteiro custa  $\mathcal{O}(mn)$  operações aritméticas. A escolha da variável que entra ( $\mathcal{O}(n)$ ) e da variável que sai ( $\mathcal{O}(m)$ ) não afeta a ordem de grandeza da expressão  $mn$ , assim uma iteração do simplex tabular tem complexidade  $\mathcal{O}(mn)$ .

Lembramos que a complexidade do simplex revisado era  $\mathcal{O}(m^2+mn)$  que representa exatamente a mesma ordem de grandeza que  $\mathcal{O}(mn)$  (pois  $m \leq n \implies m^2 \leq mn$ ). Esta complexidade do simplex revisado foi calculada supondo-se que todos os custos reduzidos eram computados: utilizando-se Bland poderíamos ter uma situação em que o primeiro custo reduzido já é negativo e aquela iteração do simplex teria complexidade  $\mathcal{O}(m^2)$  (correspondente a atualizar  $B^{-1}$  e calcular  $p'$  e  $u$ ). Já no simplex tabular não é possível deixar de atualizar todo o tableau e melhorar a complexidade  $\mathcal{O}(mn)$ . Assim concluímos que uma iteração do simplex revisado não pode ser mais lenta que uma iteração do simplex tabular, e deve inclusive ser muito mais rápida na maioria das iterações.

Em relação ao uso da memória, as iterações do simplex revisado só armazenam a inversa da base, que ocupa espaço  $\mathcal{O}(m^2)$ . Já o tableau ocupa  $\mathcal{O}(mn)$ . Apesar da descrição do problema em princípio ter tamanho  $\mathcal{O}(mn)$  (por causa da matriz  $A$ ), na maioria dos problemas práticos (e grandes) a estrutura de  $A$  é esparsa e permite seu armazenamento eficiente; em contrapartida o tableau não será esparsa em geral, mesmo que  $A$  e  $B$  o sejam, e o uso da memória no simplex tabular será sempre  $\mathcal{O}(mn)$ .

Resumindo esta discussão, temos as seguintes avaliações da complexidade destas duas implementações do simplex:

**Revisado:** Memória  $\mathcal{O}(m^2)$ , Tempo (melhor caso)  $\mathcal{O}(m^2)$ , Tempo (pior caso)  $\mathcal{O}(mn)$ ;

**Tabular:** Memória  $\mathcal{O}(mn)$ , Tempo (melhor caso)  $\mathcal{O}(mn)$ , Tempo (pior caso)  $\mathcal{O}(mn)$

## Aspectos práticos de implementação

Algumas idéias específicas de melhoria da implementação do simplex estão associadas com aumentar a estabilidade numérica do método. Tanto o simplex revisado quanto o tabular partem de uma informação inicial obtida diretamente da descrição do problema, e passam a carregar informação de uma iteração para a outra acumulando erros, devido aos arredondamentos e truncamentos em função da representação finita dos valores de ponto flutuante.

Uma idéia neste sentido é a *reinversão*. Depois de um certo número de iterações do simplex revisado em que a inversa de base já acumulou uma certa quantidade de erros, a matriz  $B^{-1}$  é recalculada a partir dos valores originais das colunas  $A^{B_1}, \dots, A^{B_m}$ . No caso do simplex tabular pode-se recomputar o tableau a partir da nova  $B^{-1}$  pela definição do tableau  $(-c'_B x_B, \bar{c}, B^{-1}b, B^{-1}A^1, \dots, B^{-1}A^n)$ .

Outra idéia relacionada no caso do simplex revisado é, ao invés de atualizar a matriz  $B^{-1}$ , armazenar as  $m$  operações elementares de linhas que transformam  $B^{-1}$  em  $\bar{B}^{-1}$ , ou seja, os múltiplos da linha pivô de  $B^{-1}$  que foram somados a cada linha (de  $B^{-1}$ ). Para calcular  $u = \bar{B}^{-1}A^j$ , por exemplo, poderíamos aplicar a mesma seqüência de operações elementares ao vetor  $B^{-1}A^j$ , a um custo de  $\mathcal{O}(m)$  operações aritméticas (supondo  $B^{-1}A^j$  conhecido). Esta idéia pode ser utilizada ao longo de várias iterações, guardando a cada iteração o índice da linha pivô e os  $m$  coeficientes correspondentes à atualização da inversa de base. Supondo que se armazene todos os vetores  $B^{-1}A^j$  onde  $B$  é a base inicial, o custo de se calcular  $p'$  e  $u$  após  $l$  iterações será  $\mathcal{O}(lm)$ ; utilizando-se esta estratégia em seqüências de  $L$  iterações (onde  $L$  é constante), depois das quais a matriz  $B^{-1}$  é computada explicitamente, os custos de computar  $p'$  e  $u$  serão  $\mathcal{O}(m)$  em todas as iterações do método, o que representa uma grande economia de tempo.

## Ciclagem e a regra de Bland

**Exemplo 3.6** *Neste exemplo veremos que o simplex pode realmente ciclar dependendo das escolhas das variáveis que entram e saem da base em cada iteração. Considere o tableau inicial*

		↓	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	$-\frac{3}{4}$	20	$-\frac{1}{2}$	6	0	0	0	0
← $x_5 =$	0	$\frac{1}{4}$	-8	-1	9	1	0	0	0
$x_6 =$	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0	1	0	0
$x_7 =$	1	0	0	1	0	0	0	1	1

e suponha que adotemos as seguintes regras: **1)** a variável a entrar da base é sempre aquela com o menor custo reduzido (mais negativo); **2)** dentre as variáveis que podem sair da base escolhemos sempre aquela de menor índice. Então o primeiro pivô será o elemento na posição



$(1,1)$  e a sequência de tableaux obtida é

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	0	-4	$-\frac{7}{2}$	33	4	0	0
$x_1 =$	0	1	-32	-4	36	3	0	0
$\leftarrow x_6 =$	0	0	4	$\frac{3}{2}$	-15	-2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	0	0	-2	18	1	1	0
$\leftarrow x_1 =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	$\frac{3}{8}$	$-\frac{15}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	0
$x_7 =$	1	0	0	1	0	0	0	1

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	$\frac{1}{4}$	0	0	-3	-2	3	0
$x_3 =$	0	$\frac{1}{8}$	0	1	$-\frac{21}{2}$	$-\frac{3}{2}$	1	0
$\leftarrow x_2 =$	0	$-\frac{3}{64}$	1	0	$\frac{3}{16}$	$\frac{1}{16}$	$-\frac{1}{8}$	0
$x_7 =$	1	$-\frac{1}{8}$	0	0	$\frac{21}{2}$	$\frac{3}{2}$	-1	1

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	$-\frac{1}{2}$	16	0	0	-1	1	0
$\leftarrow x_3 =$	0	$-\frac{5}{2}$	56	1	0	2	-6	0
$x_4 =$	0	$-\frac{1}{4}$	$\frac{16}{3}$	0	1	$\frac{1}{3}$	$-\frac{2}{3}$	0
$x_7 =$	1	$\frac{5}{2}$	-56	0	0	-2	6	1

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	$-\frac{7}{4}$	44	$\frac{1}{2}$	0	0	-2	0
$x_5 =$	0	$-\frac{5}{4}$	28	$\frac{1}{2}$	0	1	-3	0
$\leftarrow x_4 =$	0	$\frac{1}{6}$	-4	$-\frac{1}{6}$	1	0	$\frac{1}{3}$	0
$x_7 =$	1	0	0	1	0	0	0	1

		$\downarrow$						
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
	3	$-\frac{3}{4}$	20	$-\frac{1}{2}$	6	0	0	0
$\leftarrow x_5 =$	0	$\frac{1}{4}$	-8	-1	9	1	0	0
$x_6 =$	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

onde o último tableau é idêntico ao inicial e o simplex repetirá a mesma seqüência de tableaux infinitamente. Observe que  $\theta^* = 0$  em todas as iterações e que a solução viável foi sempre  $x = (0, 0, 0, 0, 0, 0, 1)'$  com custo 3. Apenas as bases que representam esta solução básica degenerada foram alternando-se ao longo da execução do algoritmo.

Veremos a seguir que a regra de Bland garante que o método simplex não apresenta ciclagem.

**Teorema** *O método simplex utilizando a **regra de Bland**:*

1. *dentre as variáveis candidatas a entrar na base, escolha a de índice  $j = \min\{i \mid \bar{c}_i < 0\}$ ;*
2. *dentre as variáveis candidatas a sair da base, escolha a de índice  $B_k = \min\{B_l \mid u_l > 0 \text{ e } \frac{x_{B_l}}{u_l} = \theta^*\}$ ;*

*não apresenta ciclagem.*

**Prova.**

A prova é feita por contradição: suponha que possa existir um ciclo usando a regra de Bland, formado pelas bases

$$B^{(1)}, B^{(2)}, \dots, B^{(l)}, B^{(1)}, B^{(2)}, \dots, B^{(l)}, \dots$$

onde todas as bases estão associadas a uma mesma solução degenerada  $\bar{x}$ .

**constatação:** toda variável que entra na base durante a ciclagem terá que sair antes de  $l$  iterações (altura em que entrará de novo na base).

**conseqüência:** dentre as variáveis que entram e saem de alguma base durante a ciclagem existe uma variável  $x_q$  de índice máximo.

Vamos explorar a estrutura do problema no momento em que  $x_q$  entrará na base. Seja  $B$  a matriz básica nesta iteração. Se  $x_q$  entrará na base pela regra de Bland, é porque

$$\bar{c}_q < 0 \quad \text{e} \quad \bar{c}_j \geq 0, \quad \forall j < q.$$

Considere o problema equivalente

$$\begin{cases} \min & \bar{c}'x \\ \text{s.a} & \bar{A}x = \bar{b} \\ & x \geq 0 \end{cases}$$

onde  $\bar{A} = B^{-1}A$ ,  $\bar{b} = B^{-1}b$  e  $\bar{c}' = c' - c'_B B^{-1}A$ . Note que

$$\bar{c}'x = \overset{=0}{\bar{c}'_B} x_B + \sum_{\substack{j \in N \\ j < q}}^{\geq 0} \bar{c}'_j x_j + \bar{c}'_q x_q + \sum_{\substack{j \in N \\ j > q}}^{< 0} \bar{c}'_j x_j,$$

e que as variáveis  $\{x_j \mid j \in N \text{ e } j > q\}$  nunca entram nas bases do ciclo (porque  $x_q$  é a variável de índice máximo que entra e sai de alguma base).

Considere agora o momento em que  $x_q$  sairá da base, e seja  $\bar{B}$  a matriz básica nesta iteração e  $x_p$  a variável que entrará no lugar de  $x_q$ . Podemos calcular os custos reduzidos com relação à formulação anterior: se  $x_p$  entrará na base, é porque

$$\bar{\bar{c}}_p = \bar{c}_p - \bar{c}'_{\bar{B}} \bar{B}^{-1} \bar{A}^p = \bar{c}_p - \bar{c}'_{\bar{B}} \bar{u} < 0,$$

onde  $\bar{u} = \bar{B}^{-1} \bar{A}^p$ . Como  $x_p$  é uma das variáveis que entram e saem da base,  $p < q$  e conseqüentemente  $\bar{c}_p \geq 0$ ; isso mostra que

$$\bar{c}'_{\bar{B}} \bar{u} = \bar{c}_{\bar{B}_1} \bar{u}_1 + \bar{c}_{\bar{B}_2} \bar{u}_2 + \cdots + \bar{c}_{\bar{B}_m} \bar{u}_m > 0.$$

Como  $x_q$  está na base  $\bar{B}$  temos  $q = \bar{B}_s$  (para algum  $s$ ) e um dos termos da somatória acima é  $\bar{c}_q \bar{u}_s$ ; lembrando que  $\bar{c}_q < 0$  e que  $\bar{u}_s > 0$  (pois este é o pivô), temos  $\bar{c}_q \bar{u}_s < 0$ . Como  $\bar{c}_{\bar{B}_1} \bar{u}_1 + \bar{c}_{\bar{B}_2} \bar{u}_2 + \cdots + \bar{c}_{\bar{B}_m} \bar{u}_m > 0$  precisa existir um outro índice  $\bar{B}_r$  tal que  $\bar{c}_{\bar{B}_r} \bar{u}_r > 0$ ; como  $x_{\bar{B}_r}$  participa de uma base do ciclo,  $\bar{B}_r \notin \{j \in N \mid j > q\}$  e conseqüentemente  $\bar{c}_{\bar{B}_r} \geq 0$ , de onde

$$\bar{c}_{\bar{B}_r} > 0 \text{ e } \bar{u}_r > 0.$$

Em particular,  $x_{\bar{B}_r}$  não pertencia à base original  $B$  (pois  $\bar{c}_B = 0$ ) e assim  $\bar{x}_{\bar{B}_r} = 0$  na solução degenerada  $\bar{x}$  associada ao ciclo. Mais ainda,  $\bar{B}_r < q$ , porque  $x_{\bar{B}_r}$  entrou na base em algum momento do ciclo. Nos termos do simplex tabular teríamos a seguinte situação:

		↓		
		$x_{\bar{B}_r}$	$x_p$	$x_q$
		0	$\bar{\bar{c}}_p < 0$	0
← $x_q =$	0	0		0
		0	$\bar{u}_s > 0$	1
		0		0
← $x_{\bar{B}_r} =$	0	1	$\bar{u}_r > 0$	0
		0		0

Mas as condições

$$\bar{u}_r > 0, \quad \frac{\bar{x}_{\bar{B}_r}}{\bar{u}_r} = 0 \quad \text{e} \quad \bar{B}_r < q$$

contradizem o fato de  $x_q$  ter sido escolhida pela regra de Bland para sair da base. Esta contradição mostra que a hipótese da existência do ciclo sob a regra de Bland não pode ser verdadeira.

**conclusão:** este argumento mostra que a regra de Bland garante a seguinte propriedade: *uma variável  $x_q$  só pode sair da base depois que entre na base alguma variável de índice maior do que  $q$ , que era não-básica quando  $x_q$  entrou na base.* Esta propriedade não é compatível com a existência de um ciclo, onde existe uma variável  $x_q$  com índice máximo dentre as variáveis que entram e saem da base.

### 3.4 Inicialização do simplex

Até agora consideramos a existência de uma base inicial dada. Em alguns casos uma tal base pode ser facilmente construída; se o problema foi obtido a partir de um problema da forma  $Ax \leq b$  com  $b \geq 0$ , na forma canônica as colunas associadas às variáveis residuais formam uma matriz identidade, e a solução básica associada será  $x_B = b \geq 0$ ,  $x_N = 0$  que é viável.

No caso geral a matriz  $A$  não possui necessariamente uma base tal que  $B = I$ , e encontrar uma base viável inicial para o simplex é tão difícil quanto resolver um problema de programação linear. Considere o problema

$$\begin{cases} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0, \end{cases}$$

e suponha que  $b \geq 0$  (como as restrições são de igualdade, podemos multiplicar algumas das linhas por  $-1$  se necessário).

A técnica usual para produzir uma base viável inicial consiste em acrescentar uma matriz identidade à matriz original, identidade esta associada a  $m$  variáveis artificiais, e resolver o problema auxiliar:

$$\begin{cases} \min & y_1 + y_2 + \cdots + y_m \\ \text{s.a} & Ax + Iy = b \\ & x, y \geq 0. \end{cases}$$

Note que  $x$  é solução do problema original se e somente se  $\begin{bmatrix} x \\ y=0 \end{bmatrix}$  é solução do problema auxiliar. A base associada às variáveis  $y_1, \dots, y_m$  produz a solução  $x = 0$  e  $y = b \geq 0$ , que é viável no problema auxiliar.

Note ainda que  $y \geq 0$  garante que

$$y_1 + y_2 + \dots + y_m = 0 \iff y_1 = y_2 = \dots = y_m = 0,$$

de onde podemos afirmar que

*o problema original é viável se e somente se o valor ótimo do problema auxiliar é zero.*

Podemos então aplicar o simplex ao problema auxiliar a partir da base inicial  $\{y_1, \dots, y_m\}$  e observar o resultado: se a solução ótima  $\begin{bmatrix} x^* \\ y^* \end{bmatrix}$  tem valor ótimo zero então  $y^* = 0$  e  $x^*$  é viável no problema original; se o valor ótimo é positivo, o problema original é inviável (o que não deixa de ser uma resposta ao problema).

Neste ponto ainda não resolvemos *necessariamente* o problema de produzir uma base para o problema original. Suponha que o problema original é viável e que o método simplex produz uma solução ótima  $\begin{bmatrix} x^* \\ y^*=0 \end{bmatrix}$ . Se a base ótima produzida contiver apenas variáveis  $x_i$  originais, então ela será uma base viável para o problema original (pois será formada por  $m$  colunas de  $A$  linearmente independentes, associadas à solução  $x^*$  viável). Por outro lado se a base ótima contiver variáveis artificiais  $y_i$  (com  $y_i^* = 0$ ), as colunas correspondentes na matriz básica associada não pertencem à matriz  $A$  e sim à identidade acrescentada ao problema. Se não pudermos forçar a saída das variáveis artificiais da base não poderemos eliminá-las do problema e voltar ao contexto do problema original.

### Retirando as variáveis artificiais da base

Suponha que a solução ótima do problema auxiliar seja  $\begin{bmatrix} x^* \\ y^*=0 \end{bmatrix}$  e que a base ótima obtida é formada pelas variáveis  $x_{B_1}, \dots, x_{B_k}$  e  $y_{B_{k+1}}, \dots, y_{B_m}$ , onde  $k < m$  (se necessário pode-se reordenar os índices básicos). Temos duas possibilidades:

**A matriz  $A$  possui linhas L.I.** Neste caso as colunas  $A^{B_1}, \dots, A^{B_k}$  podem ser completadas com outras colunas da matriz  $A$  até formar uma base; as variáveis associadas às outras colunas possuem valor 0 na solução  $x^*$  porque são não-básicas, e entrarão na base também com valor 0;

**A matriz  $A$  possui linhas L.D.** Neste caso não existem bases, mas é possível identificar linhas redundantes (combinações lineares de outras linhas) e reduzir o número de linhas até obter um conjunto linearmente independente.

O procedimento utilizado para obter uma base inicial do simplex após resolver o problema auxiliar consiste em tentar retirar as variáveis artificiais da base. A cada passo ocorrerá OU a substituição de uma variável artificial por uma variável original OU a eliminação de uma variável artificial juntamente com uma linha redundante. Esse procedimento será explicado fazendo-se referência ao tableau do simplex.

Seja  $l > k$  e considere a  $l$ -ésima variável básica, que é uma variável artificial. Suponha que a  $l$ -ésima linha de  $B^{-1}A$  possui um elemento diferente de zero, na coluna  $B^{-1}A^j$ .

			$\downarrow$	
		$x_{B_i}$	$x_j$	$y_{B_l}$
		0	?	0
$x_{B_i} =$	?	0	?	0
		1	?	0
		0	?	0
$\leftarrow y_{B_l} =$	0	0	$\neq 0$	1
		0	?	0

Então a coluna  $A^j$  é não-básica e além disso é linearmente independente de  $A^{B_1}, \dots, A^{B_k}$ : basta ver que os vetores  $B^{-1}A^{B_i} = e^i$ ,  $i < k < l$  possuem a  $l$ -ésima coordenada igual a zero, e portanto  $B^{-1}A^j$  não pode ser combinação linear de  $B^{-1}A^{B_i}$ ,  $i < k < l$  (equivalentemente  $A^j$  não pode ser combinação linear de  $A^{B_i}$ ,  $i < k < l$ ). Podemos introduzir  $x_j$  na base e retirar  $y_{B_l}$  fazendo uma pivotação sobre o elemento  $B_l^{-1}A^j \neq 0$ . Como  $y_{B_l} = 0$ , a pivotação será degenerada, isso é, não modificará nem a solução  $x^*$  nem o valor ótimo do problema auxiliar, e poderá ser efetuada mesmo que o pivô seja negativo.

Suponha por outro lado que a  $l$ -ésima linha de  $B^{-1}A$  é nula. Então  $B_l^{-1}A = \sum_{i=1}^m B_{li}^{-1}A_i = 0'$ , o que mostra que as linhas de  $A$  são linearmente dependentes. Como  $B_l^{-1}b = y_{B_l} = 0$ , temos que a  $l$ -ésima linha do tableau  $B_l^{-1}A = B_l^{-1}b$  é redundante e pode ser eliminada. Se estivermos usando o simplex revisado, podemos concluir que a linha de índice  $B_l$  do problema original é redundante: como  $B^{-1}I^{B_l} = e^l \implies B_l^{-1}I^{B_l} = 1 \implies B_{lB_l}^{-1} = 1$ , teremos  $\sum_{i=1}^m B_{li}^{-1}A_i = 0' \implies A_{B_l} = \sum_{\substack{i=1 \\ i \neq B_l}}^m B_{li}^{-1}A_i$ .

**Exemplo** Considere o problema

$$\left\{ \begin{array}{ll} \min & x_1 - x_2 - x_3 \\ \text{s.a} & x_1 + x_2 + x_3 = 2 \\ & x_1 - 2x_2 + x_3 = 2 \\ & 2x_1 + 2x_2 + 2x_3 = 4 \\ & x_1, x_2, x_3 \geq 0 \end{array} \right.$$

e o problema auxiliar

$$\left\{ \begin{array}{ll} \min & y_1 + y_2 + y_3 \\ \text{s.a} & x_1 + x_2 + x_3 + y_1 = 2 \\ & x_1 - 2x_2 + x_3 + y_2 = 2 \\ & 2x_1 + 2x_2 + 2x_3 + y_3 = 4 \\ & x_1, x_2, x_3, y_1, y_2, y_3 \geq 0 \end{array} \right.$$

Temos um vértice inicial fazendo  $x_1 = x_2 = x_3 = 0$  e  $(y_1, y_2, y_3)' = b = (2, 2, 4)'$ . A matriz básica correspondente é a identidade e  $c_B = (1, 1, 1)'$ . Os custos reduzidos são  $\bar{c}'_N = c'_N - c'_B B^{-1} A^N = (0, 0, 0) - (1, 1, 1) [A^1 \mid A^2 \mid A^3] = (-4, -1, -4)$ . A aplicação do simplex tabular para o problema auxiliar usando a regra de Bland gera os seguintes tableaux:

		$\downarrow$	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
			$-4$	$-1$	$-4$	$0$	$0$	$0$
$\leftarrow y_1 =$	$2$	$1$	$1$	$1$	$1$	$0$	$0$	
$y_2 =$	$2$	$1$	$-2$	$1$	$0$	$1$	$0$	
$y_3 =$	$4$	$2$	$2$	$2$	$0$	$0$	$1$	

		$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
	$0$	$0$	$3$	$0$	$4$	$0$	$0$
$x_1 =$	$2$	$1$	$1$	$1$	$1$	$0$	$0$
$y_2 =$	$0$	$0$	$-3$	$0$	$-1$	$1$	$0$
$y_3 =$	$0$	$0$	$0$	$0$	$-2$	$0$	$1$

Este último corresponde a uma solução ótima do problema auxiliar com valor ótimo  $0$ . Assim sabemos que  $(x_1, x_2, x_3)' = (2, 0, 0)'$  é viável no problema original. Porém a base obtida contém variáveis artificiais.

Observando a segunda linha do tableau (associada à variável  $y_2$ ) nas posições correspondentes às colunas  $x_1, x_2, x_3$  encontramos um elemento não-nulo (associado à variável  $x_2$ ). Podemos então introduzir  $x_2$  na base, eliminando  $y_2$  através de uma pivotação degenerada:

			$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
$x_1 =$	2		1	1	1	1	0	0
$\leftarrow y_2 =$	0		0	-3	0	-1	1	0
$y_3 =$	0		0	0	0	-2	0	1

		$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
$x_1 =$	2	1	0	1	$\frac{2}{3}$	$\frac{1}{3}$	0
$x_2 =$	0	0	1	0	$\frac{1}{3}$	$-\frac{1}{3}$	0
$y_3 =$	0	0	0	0	-2	0	1

Note que a linha zero já não tem importância nesta pivotação: apenas trocamos um índice da base, sem sair do ponto  $(2, 0, 0, 0, 0, 0)'$ . Observando a terceira linha do tableau (associada à variável  $y_3$ ) nas posições correspondentes às colunas  $x_1, x_2, x_3$  temos apenas zeros, o que permite concluir que esta linha do tableau é redundante no contexto do problema original. Eliminamos a linha e a coluna associadas a  $y_3$  e chegamos ao tableau:

		$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
$x_1 =$	2	1	0	1	$\frac{2}{3}$	$\frac{1}{3}$
$x_2 =$	0	0	1	0	$\frac{1}{3}$	$-\frac{1}{3}$

Agora temos apenas variáveis originais na base, e com isso podemos voltar para o problema original. Note que nas colunas artificiais do tableau a informação contida é  $B^{-1}I = B^{-1}$  (a identidade corresponde às colunas associadas às variáveis  $y_1$  e  $y_2$ ). Para preencher a linha zero do problema original associada a esta base, calculamos

$$-c'_B B^{-1}b = -(1, -1) \begin{bmatrix} 2 \\ 0 \end{bmatrix} = -2$$

$$\bar{c}_3 = c_3 - c'_B B^{-1}A^3 = -1 - (1, -1) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -2$$

e lembramos que  $\bar{c}_1 = \bar{c}_2 = 0$  (porque  $x_1$  e  $x_2$  são básicas). Teremos então o tableau a seguir



			$x_1$	$x_2$	$x_3$	↓
		-2	0	0	-2	
←	$x_1 =$	2	1	0	1	
	$x_2 =$	0	0	1	0	

que nos conduz a solução ótima do problema original:

		$x_1$	$x_2$	$x_3$
	2	2	0	0
$x_3 =$	2	1	0	1
$x_2 =$	0	0	1	0

A solução ótima é  $(0, 0, 2)'$  com valor ótimo  $-2$ .

## O método simplex de duas fases

Resumimos então o comportamento do método simplex completo. Note que não são mais necessários nem o conhecimento de uma base inicial nem a hipótese de que as linhas de  $A$  sejam L.I.:

### Fase 1: inicialização

1. Se necessário, multiplique as linhas originais por -1 de modo a obter  $b \geq 0$ ;
2. Crie variáveis artificiais  $y_1, \dots, y_m$ , se necessário, e aplique o método simplex ao problema auxiliar com o objetivo de minimizar  $\sum_{i=1}^m y_i$ ;
3. Se o valor ótimo do problema auxiliar é positivo, o problema original é inviável e o método pára;
4. Se o valor ótimo do problema original é zero e todas as variáveis básicas são originais, elimine as variáveis artificiais e vá para a fase 2;
5. Se a  $l$ -ésima variável básica é artificial, observe a linha  $B_l^{-1}A$  do tableau. Se esta linha é nula, elimine a  $l$ -ésima linha do tableau (no caso do simplex revisado, elimine a  $B_l$ -ésima linha da matriz  $A$  e a  $l$ -ésima linha e coluna da inversa de base); elimine também a variável artificial  $y_{B_l}$ . Se ao contrário existir  $j$  tal que  $B_l^{-1}A^j \neq 0$ , faça uma mudança de base usando  $B_l^{-1}A^j$  como pivô, e elimine a variável artificial  $y_{B_l}$ . Repita esta operação até que a base possua apenas variáveis originais.

## Fase 2: solução

1. Considere a inversa de base ou o tableau do final da fase 1;
2. Calcule os custos reduzidos das variáveis usando esta base inicial e os coeficientes da função objetivo original;
3. Aplique o método simplex para o problema original.

Note que, utilizando-se uma regra anti-ciclagem qualquer, o método simplex de duas fases terminará em tempo finito para qualquer problema de programação linear canônico, cobrindo todas as possibilidades a seguir:

1. Se o problema original é inviável, isso é detectado no passo 3 da fase 1;
2. Se o problema é viável mas a matriz  $A$  possui linhas redundantes, isso é detectado e corrigido no final da fase 1;
3. Se o valor ótimo é  $-\infty$ , isso é detectado durante o passo 3 da fase 2;
4. Se existe uma solução ótima, o simplex produzirá uma base ótima no final da fase 2.

## 3.5 Eficiência computacional do método simplex

Na seção 3.3 fizemos uma análise preliminar do custo computacional de UMA iteração do método simplex, para cada uma das implementações propostas. Estamos interessados em analisar o número de iterações necessárias para resolver um problema de programação linear.

### O número de iterações no pior caso

Embora possamos exprimir a complexidade computacional de uma iteração do simplex como uma função polinomial das dimensões  $m$  e  $n$ , o número de iterações do simplex no pior caso não admite uma tal limitação a priori. Um exemplo de comportamento não polinomial do simplex pode ser construído como segue.

**Teorema 3.5** *Seja  $\varepsilon \in (0, \frac{1}{2})$  e considere o problema de programação linear*

$$\left\{ \begin{array}{ll} \min & -x_n \\ \text{s. a} & \varepsilon \leq x_1 \leq 1 \\ & \varepsilon x_{i-1} \leq x_i \leq 1 - \varepsilon x_{i-1}, \quad i = 2, \dots, n. \end{array} \right.$$

Então são verdadeiras as seguintes afirmações:

1. O conjunto viável possui  $2^n$  vértices;
2. Os vértices podem ser ordenados numa seqüência  $v^1, \dots, v^{2^n}$  de modo que  $v^{i+1}$  é adjacente a  $v^i$  e  $c'v^{i+1} < c'v^i$ ;
3. Existe uma regra de pivotação sob a qual o simplex percorre a seqüência de vértices  $v^1, \dots, v^{2^n}$  nesta ordem, levando  $2^n - 1$  iterações.

Este exemplo não responde a seguinte pergunta (em aberto): *é verdade que exemplos como o anterior existem para quaisquer regras de pivotação possíveis?*

### O número de iterações no caso médio

Apesar da existência comprovada de exemplos especiais cuja solução exige um número exponencial de iterações do método simplex, a observação do comportamento do método simplex em situações típicas sugere que o número de iterações é em média  $\mathcal{O}(m)$ . A dificuldade de produzir resultados gerais no caso médio reside na dificuldade de estabelecer uma distribuição de probabilidade razoável sobre o universo de problemas possíveis.

Um exemplo de resultado particular para o caso médio do simplex segue. Suponha que são dados  $m$  vetores-linha  $A_1, \dots, A_m \in \mathbb{R}^n$ , um vetor  $c \in \mathbb{R}^n$  e um vetor  $b \in \mathbb{R}^m$  e considere que para cada índice  $i = 1, \dots, m$  podemos escolher uma das desigualdades  $A_i x \leq b_i$  ou  $A_i x \geq b_i$  com probabilidade igual. Poderão ser gerados deste modo  $2^m$  problemas lineares diferentes, dos quais um certo número  $L$  de problemas serão viáveis. Pode-se mostrar então que, usando uma regra especial de pivotação, no caso médio dentre os  $L$  problemas viáveis o método simplex terminará em no máximo  $\frac{n}{2}$  iterações (vide Schrijver: “Theory of Linear and Integer Programming”; o resultado é de M. Haimovich). Esta estimativa linear está em concordância com uma “intuição empírica” do comportamento do simplex.

**Exercícios sugeridos para o capítulo 3: 3.2-3.7, 3.9, 3.12, 3.17-3.20, 3.22-3.25, 3.28 e 3.33.**

## 4 Dualidade

### 4.1 Motivação

O estudo de dualidade em programação linear tem suas origens no método de multiplicadores de Lagrange, que foi desenvolvido para minimização de funções diferenciáveis com restrições de igualdade, visando recair em problemas sem restrições. Considere o problema

$$\begin{cases} \min & x^2 + y^2 + z^2 \\ \text{s.a} & x + y + z = 1 \end{cases}$$

A este problema associamos a função Lagrangeano  $L(x, y, z, p)$ :

$$L(x, y, z, p) = x^2 + y^2 + z^2 + p(1 - x - y - z),$$

onde a variável  $p$  é interpretada como uma penalização (ou preço) que deve ser paga se a restrição  $x + y + z = 1$  for violada. A função Lagrangeano não possui restrições e assim pode ser minimizada procurando pontos de derivada nula em relação a  $x$ ,  $y$  e  $z$ :

$$\begin{aligned} \frac{\partial L}{\partial x} &= 2x - p = 0 \\ \frac{\partial L}{\partial y} &= 2y - p = 0 \\ \frac{\partial L}{\partial z} &= 2z - p = 0 \end{aligned}$$

o que fornece a solução paramétrica  $x = y = z = \frac{p}{2}$ . Considerando novamente a restrição original  $x + y + z = 1$  descobrimos que a solução é  $x = y = z = \frac{1}{3}$ . O Lagrangeano usando a penalização  $p = \frac{2}{3}$  (associada à solução ótima) pode ser interpretado como uma combinação conveniente da função objetivo e da restrição original, combinação esta que não *força* a restrição  $x + y + z = 1$ , mas torna desinteressante (do ponto de vista da minimização) a opção de violar aquela restrição pagando o preço  $\frac{2}{3}(1 - x - y - z)$  correspondente.

---

<sup>1</sup>Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.

Considere o problema de programação linear na forma canônica

$$(PLC) \begin{cases} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0 \end{cases}$$

onde  $b \in \mathbb{R}^m$  e  $A \in \mathbb{R}^{m \times n}$ , e suponha que este problema possua uma solução ótima  $x^*$ . Associamos a cada restrição  $A_i x = b_i$  uma penalização  $p_i$  e criamos o seguinte problema relaxado:

$$\begin{cases} \min & c'x + p'(b - Ax) \\ \text{s.a} & x \geq 0. \end{cases}$$

A relaxação do problema consiste em permitir vetores  $x$  que não satisfaçam  $Ax = b$ , desde que o preço  $p'(b - Ax)$  seja adicionado à função objetivo. Chamando de  $g(p)$  o valor ótimo deste problema, para um vetor de preços fixado, temos a seguinte propriedade:

$$g(p) = \min_{x \geq 0} [c'x + p'(b - Ax)] \leq c'x^* + \overbrace{p'(b - Ax^*)}^{=0} = c'x^*,$$

e isso explica o uso do termo “problema relaxado”: como o último problema tem mais opções (pontos viáveis) do que o primeiro, ele tem mais possibilidades para a minimização e a chance de encontrar um mínimo mais baixo que o ótimo do problema original.

Outro modo de entender o resultado acima é que cada vetor de preços  $p$  produz um limitante inferior  $g(p)$  para o valor ótimo do (PLC) original. Poderíamos procurar o maior limitante inferior para o valor ótimo do (PLC) de acordo com esta estratégia resolvendo o problema

$$\begin{cases} \max & g(p) \\ \text{s.a} & p \in \mathbb{R}^m \end{cases}$$

ou equivalentemente

$$\max_{p \in \mathbb{R}^m} \min_{x \geq 0} c'x + p'(b - Ax).$$

Este problema é chamado de dual do problema original (chamado de *primal*, neste contexto). Note que

$$\min_{x \geq 0} [c'x + p'(b - Ax)] = p'b + \min_{x \geq 0} (c' - p'A)x$$

e que

$$\min_{x \geq 0} (c' - p'A)x = \begin{cases} 0 & \text{se } c' - p'A \geq 0 \\ -\infty & \text{caso contrário} \end{cases}$$

Como queremos maximizar  $g(p)$ , os vetores  $p$  tais que  $g(p) = -\infty$  não precisam ser considerados. Com isso vemos que o problema dual pode ser expresso equivalentemente como

$$(DLC) \begin{cases} \max & p'b \\ \text{s.a} & A'p \leq c \\ & p \in \mathbb{R}^m \end{cases}$$

onde  $A'p \leq c \iff (A'p)' \leq c' \iff p'A \leq c'$ . Observe que o (PLC) original é equivalente a

$$\min_{x \geq 0} \max_{p \in \mathbb{R}^m} c'x + p'(b - Ax)$$

pois

$$\max_{p \in \mathbb{R}^m} c'x + p'(b - Ax) = \begin{cases} c'x & \text{se } Ax = b \\ +\infty & \text{caso contrário.} \end{cases}$$

Outro exemplo: se o problema primal fosse

$$(PL) \begin{cases} \min & c'x \\ \text{s.a} & Ax \geq b \end{cases}$$

teríamos o problema equivalente

$$\min_{x \in \mathbb{R}^n} \max_{p \geq 0} c'x + p'(b - Ax).$$

Note que neste contexto queremos aplicar outro tipo de penalização: queremos garantir que não vale a pena violar a restrição  $Ax \geq b$  na formulação acima. Se existir uma restrição violada  $A_i x < b_i$ , impomos uma multa positiva  $p_i$  muito alta que torne o produto  $p'(b - Ax)$  muito grande (positivo). Isto tornará desinteressante a violação daquela restrição do ponto de vista da minimização. Não podemos permitir preços negativos pois estaríamos desencorajando também a possibilidade de  $A_i x > b_i$ , que é possível no problema original. Note finalmente que

$$\max_{p \geq 0} c'x + p'(b - Ax) = \begin{cases} c'x & \text{se } Ax \geq b \\ +\infty & \text{caso contrário} \end{cases}$$

e que para  $p \geq 0$  conseguimos manter a propriedade

$$g(p) = \min_{x \in \mathbb{R}^n} c'x + p'(b - Ax) \leq c'x^* + \overbrace{p'(b - Ax^*)}^{\leq 0} \leq c'x^*,$$

ou seja,  $g(p)$  é um limitante inferior para o valor ótimo do (PL). O dual do (PL) corresponde a  $\max_{p \geq 0} g(p)$ , ou equivalentemente,

$$\max_{p \geq 0} \min_{x \in \mathbb{R}^n} c'x + p'(b - Ax) = \max_{p \geq 0} \min_{x \in \mathbb{R}^n} p'b + (c' - p'A)x.$$

Mais uma vez, observando que

$$\min_{x \in \mathbb{R}^n} p'b + (c' - p'A)x = \begin{cases} p'b & \text{se } c' - p'A = 0 \\ -\infty & \text{caso contrário,} \end{cases}$$

concluiremos que o dual do (PL) acima é equivalente a

$$(DL) \begin{cases} \max & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0 \end{cases}$$

Assim, o sinal dos preços duais deve ser designado de modo a penalizar a violação das restrições primais, isto é, de modo a permitir que a expressão  $p'(b - Ax)$  fique (muito) positiva sempre que a restrição original não é satisfeita; equivalentemente, para todos os pontos viáveis devemos ter  $p'(b - Ax) \leq 0$ , de modo a anular esta parcela ao calcularmos  $\max_p p'(b - Ax)$ . Assim,  $Ax \geq b$  dá origem a preços  $p \geq 0$ , e  $Ax = b$  dá origem a preços  $p \in \mathbb{R}^m$ .

Finalmente, observe que este tipo de transformação também pode ser aplicado a problemas de maximização, seja transformando-os em um problema de minimização equivalentes, seja reformulando as penalizações visando tornar  $p'(b - Ax)$  (muito) negativo, redefinindo  $g(p)$  como um limitante superior para o valor ótimo e entendendo o dual como o problema de encontrar o menor limitante superior para o valor ótimo primal. Como exemplo,

$$\begin{cases} \max & c'x \\ \text{s.a} & Ax \leq b \end{cases}$$

é equivalente a

$$\max_{x \in \mathbb{R}^n} \min_{p \geq 0} \overbrace{c'x + p'(b - Ax)}^{g(p)}$$

cujo dual é

$$\min_{p \geq 0} \max_{x \in \mathbb{R}^n} p'b + (c' + p'A)x$$

ou, equivalentemente,

$$\begin{cases} \min & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0 \end{cases}$$

**Exercício 4.1** Calcule os duais dos problemas a seguir:

$$(P_1) \begin{cases} \max & p'b \\ \text{s.a} & A'p \leq c \\ & p \in \mathbb{R}^m \end{cases}$$

$$(P_2) \begin{cases} \max & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0 \end{cases}$$

$$(P_3) \begin{cases} \min & p'b \\ \text{s.a} & A'p = c \\ & p \leq 0 \end{cases}$$

## 4.2 O problema dual

Note da discussão anterior que as restrições primais e o sinal dos preços duais estão estreitamente relacionados, bem como o sinal das variáveis primais e a forma das restrições duais. Podemos resumir todas estas relações considerando o problema primal de programação linear a seguir:

$$(P) \begin{cases} \min & c'x \\ \text{s.a} & A_i x \geq b_i, \quad i \in M_1, \\ & A_i x \leq b_i, \quad i \in M_2, \\ & A_i x = b_i, \quad i \in M_3, \\ & x_j \geq 0, \quad j \in N_1, \\ & x_j \leq 0, \quad j \in N_2, \\ & x_j \in \mathbb{R}, \quad j \in N_3. \end{cases}$$

Seguindo o mesmo tipo de argumento da seção anterior, concluiremos que o seu dual é

$$(D) \begin{cases} \max & p'b \\ \text{s.a} & p_i \geq 0, \quad i \in M_1, \\ & p_i \leq 0, \quad i \in M_2, \\ & p_i \in \mathbb{R}, \quad i \in M_3, \\ & (A^j)'p \leq c_j, \quad j \in N_1, \\ & (A^j)'p \geq c_j, \quad j \in N_2, \\ & (A^j)'p = c_j, \quad j \in N_3. \end{cases}$$

**Exercício 4.2** *Mostre que (D) é o dual de (P) repetindo os argumentos da seção anterior.*

Note que a cada restrição primal temos uma variável dual associada, e a cada variável primal temos uma restrição dual associada; note que restrições de igualdade estão associadas a variáveis livres de sinal, e restrições de desigualdades estão associadas a variáveis com algum sinal. A função objetivo primal faz parte das restrições duais e o lado direito primal aparece como função objetivo dual.



**Exemplo 4.1** Considere os problemas primal e dual a seguir:

$$\left\{ \begin{array}{ll} \min & x_1 + 2x_2 + 3x_3 \\ \text{s.a} & -x_1 + 3x_2 = 5 \\ & 2x_1 - x_2 + 3x_3 \geq 6 \\ & x_3 \leq 4 \\ & x_1 \geq 0, x_2 \leq 0, x_3 \in \mathbb{R} \end{array} \right. \quad \left\{ \begin{array}{ll} \max & 5p_1 + 6p_2 + 4p_3 \\ \text{s.a} & -p_1 + 2p_2 \leq 1 \\ & 3p_1 - p_2 \geq 2 \\ & 3p_2 + p_3 = 3 \\ & p_1 \in \mathbb{R}, p_2 \geq 0, p_3 \leq 0 \end{array} \right.$$

Vamos transformar o problema dual em um problema equivalente de minimização e calcular o seu dual:

$$\left\{ \begin{array}{ll} \min & -5p_1 - 6p_2 - 4p_3 \\ \text{s.a} & -p_1 + 2p_2 \leq 1 \\ & 3p_1 - p_2 \geq 2 \\ & 3p_2 + p_3 = 3 \\ & p_1 \in \mathbb{R}, p_2 \geq 0, p_3 \leq 0 \end{array} \right. \quad \left\{ \begin{array}{ll} \max & z_1 + 2z_2 + 3z_3 \\ \text{s.a} & -z_1 + 3z_2 = -5 \\ & 2z_1 - z_2 + 3z_3 \leq -6 \\ & z_3 \geq -4 \\ & z_1 \leq 0, z_2 \geq 0, z_3 \in \mathbb{R} \end{array} \right.$$

Este último problema é equivalente ao problema original, com  $z = -x$ .

O princípio ilustrado pelo teorema acima (o dual do dual é o primal) vale em geral:

**Teorema 4.1** Seja  $(P)$  um problema (primal) de programação linear e  $(D)$  o seu dual. O problema dual  $(D_D)$  obtido a partir de  $(D)$  é equivalente a  $(P)$ .

**Exercício 4.3** Prove o teorema acima para a formulação geral do exemplo 4.1.

**Exemplo 4.2** Considere o par primal-dual a seguir:

$$\left\{ \begin{array}{ll} \min & c'x \\ \text{s.a} & Ax \geq b \\ & x \in \mathbb{R}^n \end{array} \right. \quad \left\{ \begin{array}{ll} \max & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0 \end{array} \right.$$

Substituindo as desigualdades do problema original por variáveis residuais e depois calculando o dual, teremos:

$$\left\{ \begin{array}{ll} \min & c'x \\ \text{s.a} & Ax - x^r = b \\ & x \in \mathbb{R}^n \\ & x^r \geq 0 \end{array} \right. \quad \left\{ \begin{array}{ll} \max & p'b \\ \text{s.a} & A'p = c \\ & -p \leq 0 \\ & p \in \mathbb{R}^m \end{array} \right.$$

Por outro lado, substituindo as variáveis do problema original por pares de variáveis não negativas e depois calculando o dual, teremos:

$$\left\{ \begin{array}{ll} \min & c'x^+ - c'x^- \\ \text{s.a} & Ax^+ - Ax^- \geq b \\ & x^+ \geq 0 \\ & x^- \geq 0 \end{array} \right. \quad \left\{ \begin{array}{ll} \max & p'b \\ \text{s.a} & A'p \leq c \\ & -A'p \leq -c \\ & p \geq 0 \end{array} \right.$$

Observe que são equivalentes os três problemas primais, bem como os respectivos duais.

**Exemplo 4.3** Considere o (PLC) e seu dual

$$(PLC) \left\{ \begin{array}{ll} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0 \end{array} \right. \quad (DLC) \left\{ \begin{array}{ll} \max & b'p \\ \text{s.a} & A'p \leq c \\ & p \in \mathbb{R}^m \end{array} \right.$$

e suponha que  $A_m = \sum_{i=1}^{m-1} \gamma_i A_i$  e  $b_m = \sum_{i=1}^{m-1} \gamma_i b_i$ , ou seja, que a última restrição de igualdade do (PLC) é supérflua. Note que

$$b'p = \left( \sum_{i=1}^{m-1} b_i p_i \right) + b_m p_m = \sum_{i=1}^{m-1} b_i p_i + \sum_{i=1}^{m-1} \gamma_i b_i p_m = \sum_{i=1}^{m-1} b_i (p_i + \gamma_i p_m),$$

$$A'p = \left( \sum_{i=1}^{m-1} (A_i)' p_i \right) + (A_m)' p_m = \sum_{i=1}^{m-1} (A_i)' (p_i + \gamma_i p_m).$$

Ou seja, o problema dual é equivalente a

$$\left\{ \begin{array}{ll} \max & \sum_{i=1}^{m-1} b_i q_i \\ \text{s.a} & \sum_{i=1}^{m-1} (A_i)' q_i \leq c \\ & q \in \mathbb{R}^{m-1} \end{array} \right.$$

onde  $q_i = p_i + \gamma_i p_m$ . Este é exatamente o dual que seria obtido se eliminássemos a restrição supérflua do problema original.

Os dois exemplos acima fornecem o seguinte resultado:

**Teorema 4.2** Suponha que o problema de programação linear  $(P_2)$  foi obtido de  $(P_1)$  através de uma sequência de operações do tipo:

1. substituição de uma variável  $x_i \in \mathbb{R}$  por  $x_i^+, x_i^- \geq 0$ ;
2. substituição de uma restrição  $A_i x \leq b$  por  $A_i x + x^r = b$ ,  $x^r \geq 0$ ;
3. eliminação de restrições de igualdade supérfluas.

Então os duais de  $(P_1)$  e  $(P_2)$  são equivalentes.

### 4.3 O teorema de dualidade

**Teorema 4.3 (Dualidade fraca)** *Se  $x$  é viável primal e  $p$  é viável dual, então*

$$p'b \leq c'x.$$

**Prova.**

Note que no caso canônico o resultado acima é trivial:

$$p'b = \overbrace{p'A}^{\leq c} \bar{x} \stackrel{\geq 0}{\leq} c'x.$$

No caso geral, considere os valores que aparecem nas funções Lagrangeano primal e dual:

$$\begin{aligned} u_i &= p_i(b_i - A_i x) \\ v_j &= (c_j - p'A^j)x_j. \end{aligned}$$

Pela viabilidade de  $x$  e  $p$ , o sinal de  $p_i$  tem que ser contrário ao de  $b_i - A_i x$ :

$$\begin{aligned} A_i x \geq b_i &\implies p_i \geq 0 \\ A_i x \leq b_i &\implies p_i \leq 0 \\ A_i x = b_i &\implies p_i \in \mathbb{R}. \end{aligned}$$

Isso mostra que  $u_i \leq 0$ . Analogamente mostra-se que  $v_i \geq 0$ .

Portanto

$$\begin{aligned} p'b - p'A x &= \sum_{i=1}^m u_i \leq 0 \leq \sum_{j=1}^n v_j = c'x - p'A x \\ &\Downarrow \\ p'b &\leq c'x. \end{aligned}$$

Apesar de bastante simples, o resultado acima já permite algumas conclusões sobre a relação entre os problemas primal e dual:

**Corolário 4.1** *Se o primal é ilimitado (valor ótimo  $-\infty$ ), o dual é inviável. Analogamente, se o dual é ilimitado (valor ótimo  $+\infty$ ), o primal é inviável.*

**Prova.**

Suponha que o primal é ilimitado, e suponha por contradição que  $p$  é viável dual. Pelo teorema fraco de dualidade,  $p'b \leq c'x$ ,  $\forall x$  viável. Pela ilimitação primal, existe um  $\bar{x}$  viável tal que  $c'\bar{x} < p'b$ , o que fornece a contradição  $p'b \leq c'x < p'b$ . A outra implicação é análoga.

**Corolário 4.2** *Se  $x$  viável primal e  $p$  viável dual satisfazem  $p'b = c'x$ , então  $x$  e  $p$  são soluções ótimas de seus respectivos problemas.*

**Prova.**

Seja  $y$  uma solução viável primal qualquer. Pelo teorema fraco de dualidade,  $p'b \leq c'y$ ; assim  $c'x = p'b \leq c'y$ ,  $\forall y$  viável. A prova de que  $p$  é ótimo do problema dual é análoga.

**Teorema 4.4 (Dualidade forte)** *Se um problema de programação linear possui uma solução ótima então seu dual também possui solução ótima e os valores ótimos dos dois problemas são iguais.*

**Prova.**

A prova será dividida em dois casos: (PLC) e caso geral.

Considere o problema linear canônico

$$(PLC) \begin{cases} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0 \end{cases}$$

e suponha que as linhas de  $A$  são linearmente independentes. Se o (PLC) possui solução ótima, o método simplex aplicado a este problema nos fornece uma base viável ótima  $B$ . Pelo critério de otimalidade,

$$\bar{c}' = c' - c'_B B^{-1} A = c' - p' A \geq 0,$$

onde  $p' = c'_B B^{-1}$  é o vetor de multiplicadores do simplex. Assim  $p$  é viável para o problema dual

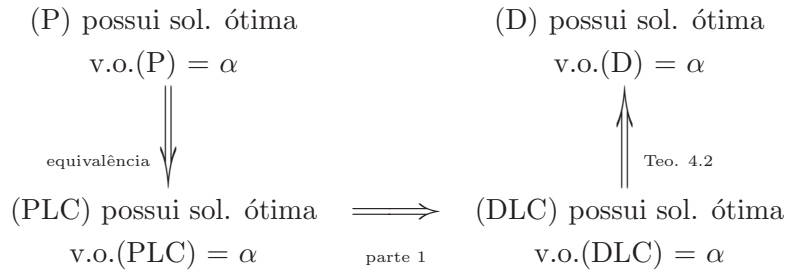
$$(PLC) \begin{cases} \max & p'b \\ \text{s.a} & A'p \leq c. \end{cases}$$

Além disto,

$$p'b = c'_B B^{-1} b = c'_B x_B = c'x,$$

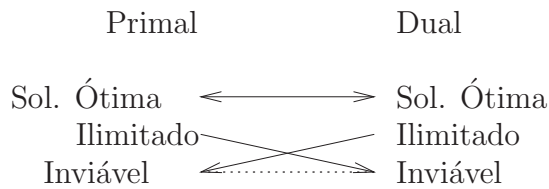
onde  $x$  é a solução ótima associada à base  $B$ . Pelo corolário 4.2  $p$  é uma solução ótima dual e o valor ótimo dual é igual ao valor ótimo primal.

Considere agora um problema de programação linear (P) qualquer e suponha que (P) possui uma solução ótima. Sabemos que podemos transformar este problema em um problema linear canônico (PLC) equivalente onde  $A$  possui linhas linearmente independentes, que possui uma solução ótima com o mesmo valor ótimo de (P).



Pela primeira parte desta demonstração, o dual do (PLC) possui solução ótima e o valor ótimo deste dual é igual ao valor ótimo do (PLC). Pelo teorema 4.2 os duais de (P) e do (PLC) são equivalentes, e portanto ambos possuem soluções ótimas e o mesmo valor ótimo.

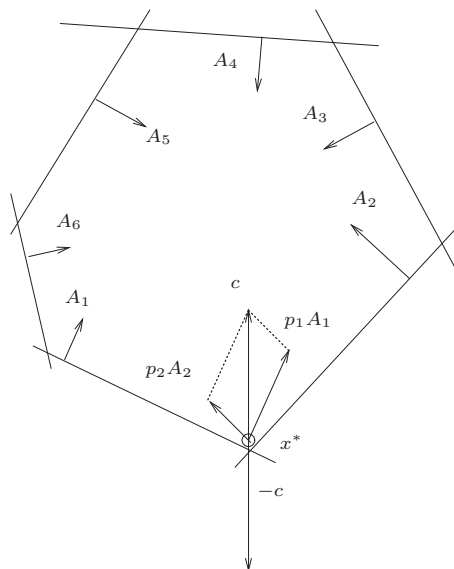
Lembramos que a solução de qualquer problema de programação linear pode levar a três conclusões possíveis: otimalidade, ilimitação e inviabilidade. Estas conclusões no caso de um problema qualquer de programação linear e seu dual podem ser expressas através de um diagrama:



Note que os casos que não estão associados no diagrama não são possíveis, de acordo com os teoremas de dualidade fraca e forte. A situação em que primal e dual são ambos inviáveis é possível, e pode ser ilustrada com este exemplo:

$$\left\{ \begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s.a} & x_1 + x_2 = 1 \\ & 2x_1 + 2x_2 = 3 \end{array} \right. \quad \left\{ \begin{array}{ll} \max & p_1 + 3p_2 \\ \text{s.a} & p_1 + 2p_2 = 1 \\ & p_1 + 2p_2 = 2 \end{array} \right.$$

**Exemplo 4.4** Considere um poliedro na forma  $Ax \geq b$  e suponha que coloquemos uma bola de dimensões desprezíveis em seu interior. Impulsionada pela gravidade, esta bola atingirá um estado de equilíbrio na quina mais baixa do poliedro.



Este ponto  $x^*$  é solução ótima do problema

$$\begin{cases} \min & c'x \\ \text{s.a} & Ax \geq b, \end{cases}$$

onde  $c$  é um vetor vertical. O equilíbrio da bola naquele ponto é caracterizado pelo fato de que a gravidade (direção  $-c$ ) é contrabalançada pelas forças normais  $A_i$  exercidas pelas paredes do poliedro. Assim existirão coeficientes  $p_i \geq 0$  tais que

$$-c + \sum_i p_i A_i = 0 \iff c = \sum_i p_i A_i;$$

em particular,  $p$  é viável no problema dual

$$\begin{cases} \max & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0. \end{cases}$$

Além disto, os coeficientes  $p_i$  associados a paredes que não encostam na bola são nulos, pois não há reação daquelas paredes. Isso mostra que

$$p_i(b_i - A_i x^*) = 0, \forall i,$$

pois nas paredes que tocam a bola teremos  $b_i = A_i x^*$  e nas demais teremos  $p_i = 0$ . Conseqüentemente,

$$p'b = \sum_i p_i b_i = \sum_i p_i A_i x^* = (\sum_i p_i A_i) x^* = p' A x^* = c' x^*,$$

ou seja,  $p$  é viável dual e possui o mesmo valor da função objetivo dual que  $c'x^*$ . Assim  $p$  é ótimo dual e o valor ótimo dual é igual ao valor ótimo primal.

## Folgas complementares

A condição  $p_i(b_i - A_i x) = 0$ ,  $\forall i$ , no exemplo anterior é chamada de *folgas complementares*, pois diz que as restrições  $p_i \geq 0$  e  $A_i x \geq b_i$  não podem apresentar folgas (desigualdades estritas) simultaneamente.

Outra condição de folgas complementares associadas ao mesmo par primal-dual é  $(c_j - p' A^j)x_j = 0$ ,  $\forall j$ , que no exemplo anterior é satisfeita trivialmente (pois  $(A^j)'p = c_j$ ,  $\forall j$ ). Quando  $x$  e  $p$  são viáveis (primal e dual, respectivamente), estas condições equivalem a  $c'x = p'b$ .

**Teorema 4.5 (Folgas complementares)** *Seja  $x$  uma solução viável primal e  $p$  uma solução viável dual. Então  $x$  e  $p$  são soluções ótimas (dos problemas primal e dual, respectivamente) se e somente se*

$$\begin{cases} p_i(b_i - A_i x) = 0, & \forall i, \\ (c_j - p' A^j)x_j = 0, & \forall j. \end{cases}$$

### Prova.

Na demonstração do teorema de dualidade fraca (4.3) verificamos que  $p_i(b_i - A_i x) \leq 0$ ,  $\forall i$  e  $(c_j - p' A^j)x_j \geq 0$ ,  $\forall j$ .

Assim,

$$p'b - p'Ax = \sum_i p_i(b_i - A_i x) \leq 0 \leq \sum_j (c_j - p' A^j)x_j = c'x - p'Ax.$$

Se, por um lado,  $x$  e  $p$  são soluções ótimas, temos  $c'x = p'b$  (teorema de dualidade forte) e assim todos os termos da desigualdade acima são iguais. Em particular,  $\sum_i p_i(b_i - A_i x) = 0$  e  $\sum_j (c_j - p' A^j)x_j = 0$ . Como os termos de cada uma destas somatórias têm todos o mesmo sinal ( $\leq 0$  na primeira,  $\geq 0$  na segunda), segue que  $p_i(b_i - A_i x) = 0$ ,  $\forall i$ , e  $(c_j - p' A^j)x_j = 0$ ,  $\forall j$ .

Por outro lado, se vale folgas complementares, então da desigualdade anterior segue que  $p'b - p'Ax = c'x - p'Ax$ , ou seja  $c'x = p'b$ . Pelo corolário 4.2 (do teorema de dualidade fraca) temos que  $x$  é solução ótima primal e  $p$  é solução ótima dual.

As condições de folgas complementares dizem que uma variável (primal ou dual) só pode ser diferente de 0 se a restrição correspondente (no dual ou no primal) estiver ativa ou, equivalentemente, que uma restrição (primal ou dual) só pode estar inativa se a variável correspondente (dual ou primal) possuir valor 0.

Quando um dos problemas (primal ou dual) possui restrições de igualdade, algumas das condições de folgas complementares são automaticamente satisfeitas; por exemplo quando o primal é o (PLC), as condições  $p_i(b_i - A_i x) = 0, \forall i$ , são satisfeitas para todos os pontos viáveis, pois  $A_i x = b_i$ . Um exemplo onde todas as folgas são possíveis e nenhuma das condições acima é automaticamente satisfeita é dado pelo par

$$(P) \begin{cases} \min & c'x \\ \text{s.a} & Ax \geq b \\ & x \geq 0 \end{cases} \quad (D) \begin{cases} \max & p'b \\ \text{s.a} & A'p \leq c \\ & p \geq 0 \end{cases}$$

**Exemplo 4.6** Considere o par de problemas

$$(P) \begin{cases} \min & 13x_1 + 10x_2 + 6x_3 \\ \text{s.a} & 5x_1 + x_2 + 3x_3 = 8 \\ & 3x_1 + x_2 = 3 \\ & x \geq 0 \end{cases} \quad (D) \begin{cases} \max & 8p_1 + 3p_2 \\ \text{s.a} & 5p_1 + 3p_2 \leq 13 \\ & p_1 + p_2 \leq 10 \\ & 3p_1 \leq 6. \end{cases}$$

A solução  $x^* = (1, 0, 1)'$  é ótima, como verificaremos a seguir. É fácil ver que  $x^*$  é básica (associada à base  $\{x_1, x_3\}$ ) e não-degenerada (pois  $x_B^* > 0$ ). As condições  $p_i(b_i - A_i x^*) = 0, \forall i$ , são satisfeitas automaticamente, bem como  $(c_2 - p'A^2)x_2^* = 0$  (pois  $x_2^* = 0$ ). As demais condições de folgas complementares são  $(c_1 - p'A^1) \cdot 1 = 0$  e  $(c_3 - p'A^3) \cdot 1 = 0$ , o que fornece o sistema

$$\begin{cases} 5p_1 + 3p_2 = 13 \\ 3p_1 = 6 \end{cases}$$

que possui como única solução  $p_1^* = 2$  e  $p_2^* = 1$ . Note que esta solução é dual viável e que  $(p^*)'b = 19 = c'x^*$ , o que mostra que  $p^*$  é ótima no dual e confirma o fato de que  $x^*$  é ótimo no primal.

Este exemplo pode ser generalizado: em uma solução ótima não-degenerada temos  $x_{B_i}^* > 0$  para cada variável básica  $x_{B_i}$ . Assim, por folgas complementares, teremos  $p'A^j = c_j$  e portanto  $p'[A^{B_1} | \dots | A^{B_m}] = (c_{B_1}, \dots, c_{B_m})$  ou, equivalentemente,  $p'B = c'_B$ . Este sistema possui a solução única  $p' = c'_B B^{-1}$  que será viável no dual se e somente se a base  $B$  é ótima primal.

## Uma visão geométrica

Considere o par primal-dual

$$\begin{cases} \min & c'x \\ \text{s.a} & A_i x \geq b_i \quad i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{cases} \quad \begin{cases} \max & b'p \\ \text{s.a} & \sum_{i=1}^m p_i A_i = c' \\ & p \geq 0. \end{cases}$$

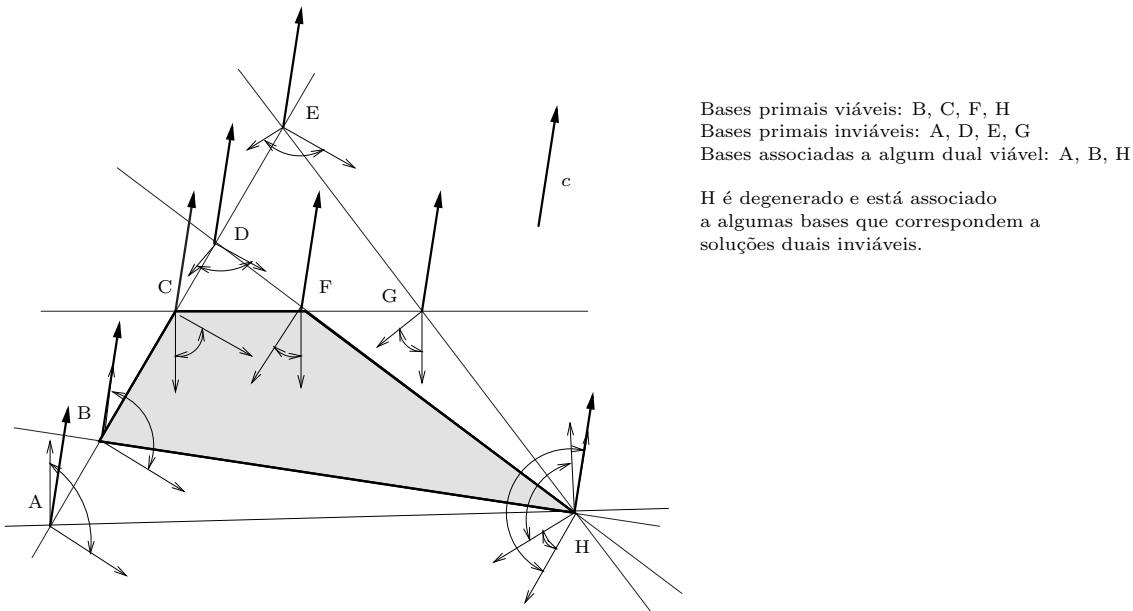


Suponha que os vetores  $A_i$  geram  $\mathbb{R}^n$  (em particular,  $m \geq n$ ) e considere uma base de  $\mathbb{R}^n$  formada por  $A_i$ ,  $i \in I$ . O sistema  $A_I x = b_I$  possui uma única solução que vamos chamar de  $x^I$ .  $x^I$  é uma solução básica por construção; suponha além disso que  $x^I$  é não-degenerada, ou seja, que  $A_i x^I \neq b_i$ ,  $\forall i \notin I$ .

Seja  $p \in \mathbb{R}^m$  qualquer. Para que  $x^I$  seja ótimo primal e  $p$  seja ótimo dual, precisam satisfazer as seguintes condições:

- 1  $A_i x^I \geq b_i$ ,  $\forall i$ , (viabilidade primal)
- 2  $p_i = 0$ ,  $\forall i \notin I$  (folgas complementares)
- 3  $\sum_{i=1}^m p_i A_i = c'$  (viabilidade dual I)
- 4  $p \geq 0$  (viabilidade dual II)

Substituindo (2) em (3) temos  $\sum_{i \in I} p_i A_i = c'$ . Pela independência linear dos  $A_i$  este sistema possui uma única solução que chamaremos de  $p^I$ . Note que os vetores  $A_i$  correspondem a colunas L.I. do dual que está na forma canônica, assim estes vetores formam uma base do dual associada a  $p^I$ . Concluimos portanto que, construindo  $p^I$  desta maneira (satisfazendo folgas complementares), a viabilidade de  $p^I$  depende tão somente da restrição  $p^I \geq 0$ , ou seja, do fato de  $c$  ser uma combinação linear dos vetores  $A_i$  com coeficientes não-negativos.



Se  $\bar{x}$  é uma solução básica degenerada então podem existir vários subconjuntos  $I$  distintos que correspondem a bases de  $\mathbb{R}^n$  formadas pelos vetores  $A_i$ . Para cada base teremos uma solução dual  $p^I$  diferente que satisfaz folgas complementares; se uma destas soluções for viável

dual, e  $\bar{x}$  for viável primal, então o par  $(\bar{x}, p^I)$  será um par primal-dual ótimo (pelo teorema 4.5),

#### 4.4 Variáveis duais ótimas como custos marginais

Outra interpretação para as variáveis duais no ponto de ótimo, além da interpretação mecânica (equilíbrio), é a de associar as variáveis  $p_i$  a preços. Suponha que o problema primal é um problema linear canônico associado a um processo de produção e que o vetor  $b$  corresponde ao estoque de matérias-primas utilizadas (vide capítulo 1). Ao resolver-se o problema, uma pergunta que surge naturalmente é: *pode-se melhorar a solução ótima alterando-se o estoque?* Poderíamos vender parte do estoque de algumas matérias-primas menos essenciais e comprar outro tipo de matéria-prima.

Suponha que  $x^*$  é uma solução ótima não-degenerada e que  $B$  é a base ótima associada. Pela não-degenerescência,  $x_B = B^{-1}b > 0$ . Suponha que o estoque  $b$  seja trocado por  $b + d$ . Por continuidade,  $B^{-1}(b + d) > 0$  desde que  $d$  seja suficientemente pequeno. Assim a mesma base será viável para o problema alterado. Além disso,  $\bar{c}' = c' - c'_B B^{-1}A \geq 0$  independentemente do lado direito do problema ser  $b$  ou  $b + d$ . Assim, se  $B^{-1}(b + d) \geq 0$ , então esta base é ótima também no novo problema. O valor ótimo deste problema será

$$c'_B B^{-1}(b + d) = p'(b + d) = p'b + p'd = c'x^* + p'd,$$

ou seja, a variação do valor ótimo será  $p'd$ . Cada variável  $p_i$  representa o quanto se ganha ou perde no valor ótimo para uma variação  $d_i$  no lado direito do problema. Se for possível modificar o problema através de uma variação  $d$  do estoque que satisfaça  $p'd < 0$ , o valor ótimo do problema modificado será menor do que o do problema original.

Se modificarmos o lado direito do problema original de tal modo que a base ótima deixa de ser viável ( $B^{-1}(b + d) \not\geq 0$ ), podemos ainda aproveitar a informação de  $p$  ser viável dual para aplicarmos o simplex dual usando  $B$  como base inicial.

#### 4.5 O Simplex Dual

Na prova do teorema forte de dualidade utilizamos o fato de que se um problema na forma canônica possui solução ótima então o simplex aplicado a ele consegue encontrar uma base viável ótima. Vimos que a condição de otimalidade  $\bar{c}' = c' - c'_B B^{-1}A \geq 0$  era equivalente à

condição de viabilidade dual  $A'p \leq c$  para o vetor de multiplicadores do simplex  $p' = c'_B B^{-1}$ . Observe que este vetor está definido para todas as bases do problema (sejam elas viáveis ou inviáveis), e que

$$p'b = c'_B B^{-1}b = c'_B x_B = c'x.$$

Podemos então pensar no simplex como um método que percorre bases do problema primal associadas ao par  $(x, p)$ , mantendo a condição de viabilidade primal ( $x_B = B^{-1}b \geq 0$ ) e folgas complementares (que correspondem a  $p'b = c'x$  neste caso), buscando atingir viabilidade dual ( $c' - p'A \geq 0$ ).

Uma alternativa é explorar as bases do poliedro na forma canônica mantendo a cada iteração as condições de viabilidade *dual* e folgas complementares, buscando obter viabilidade primal. Um tal método passaria por uma sequência de soluções *inviáveis* do primal, e obteria otimalidade (de  $x$  e  $p$ ) quando verificasse a condição de viabilidade primal.

Considere o tableau associado a uma base  $B$

$-c'_B x_B$	$\bar{c}_1$	$\cdots$	$\bar{c}_n$
$x_{B_1}$			
$\vdots$	$B^{-1}A^1$	$\cdots$	$B^{-1}A^n$
$x_{B_m}$			

e suponha que  $\bar{c} \geq 0$  (ou equivalentemente  $A'p \leq c$ ) mas não temos necessariamente que  $x_B \geq 0$  (se esta última condição fosse verdadeira a base seria ótima, como vimos anteriormente). Considere um índice  $l$  tal que  $x_{B_l} < 0$  e considere a  $l$ -ésima linha do tableau,  $v = B_l^{-1}A$ , que definiremos como a linha pivô. A coluna pivô será qualquer coluna  $j$  que satisfaça  $v_j < 0$  e

$$\frac{\bar{c}_j}{|v_j|} = \min_{v_i < 0} \frac{\bar{c}_i}{|v_i|}.$$

Observe que  $x_j$  é não-básico (pois a coluna  $j$  do tableau não é uma coluna da identidade!) e podemos forçar a entrada de  $x_j$  na base substituindo  $x_{B_l}$  fazendo uma pivotação sobre o elemento  $v_j$ . Observe que o novo valor de  $\bar{c}'$  será  $\bar{c}' + \frac{\bar{c}_j}{|v_j|} \geq 0$  (devido à escolha de  $j$ ) e que o novo valor de  $-c'_B x_B$  será  $-c'_B x_B + \frac{\bar{c}_j}{|v_j|} x_{B_l} \leq -c'_B x_B$ . Isso mostra que a nova base é viável dual e que o valor da solução dual não decresce.

**Exemplo 4.7** Considere o tableau abaixo.

			$\downarrow$			
		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
	0	2	6	10	0	0
$x_4 =$	2	-2	4	1	1	0
$\leftarrow x_5 =$	-1	4	-2	-3	0	1

Como  $x_{B_2} < 0$  escolhemos esta linha como linha pivô. Os elementos  $v_j < 0$  desta linha são  $v_2 = -2$  e  $v_3 = -3$ ; como  $\frac{6}{|-2|} < \frac{10}{|-3|}$  escolhemos  $j = 2$ . Pivotando sobre o elemento  $v_2$  temos o novo tableau

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
	-3	14	0	1	0	3
$x_4 =$	0	6	0	-5	1	2
$x_2 =$	$\frac{1}{2}$	-2	1	$\frac{3}{2}$	0	$-\frac{1}{2}$

Observe que o custo aumentou para 3 e que  $x_B = B^{-1}b \geq 0$ , o que juntamente com  $\bar{c} \geq 0$  mostra que a solução  $x = (0, \frac{1}{2}, 0, 0, 0)'$  é ótima.

Duas observações merecem ser feitas antes de enunciarmos o método simplex dual.

A primeira é que sempre que  $\bar{c}_j > 0$  na coluna pivô teremos que o novo valor de  $-c'_B x_B$  será  $-c'_B x_B + \frac{\bar{c}_j}{|v_j|} x_{B_l} < -c'_B x_B$  e portanto o custo da solução dual cresce estritamente. Se isto for verdade em todas as iterações, o número de iterações será finito (pois o número de bases é finito). Se numa iteração  $\bar{c}_j = 0$  na coluna pivô então a solução dual não muda e a pivotação é dual-degenerada; será necessário utilizar alguma regra anti-ciclagem (como Bland, por exemplo) para garantir a terminação do simplex dual em um número finito de iterações.

A segunda observação é que se  $x_{B_l} < 0$  e a linha correspondente do tableau  $v = B_l^{-1}A$  é não-negativa, então o problema dual é ilimitado (exercício 4.22), e conseqüentemente o primal é inviável.

### Uma iteração do simplex dual

0. A iteração começa com o tableau associado a uma base  $B$  tal que  $\bar{c} \geq 0$ ;
1. Se  $x_B \geq 0$  pare, a solução é ótima; do contrário, escolha um índice  $l$  tal que  $x_{B_l} < 0$ ;
2. Se  $v = B_l^{-1}A \geq 0$  pare, o primal é inviável; do contrário, escolha um índice  $j$  tal que  $\frac{\bar{c}_j}{|v_j|} = \min_{v_i < 0} \frac{\bar{c}_i}{|v_i|}$ ;

3. Faça a pivotação sobre o elemento  $v_j$ , que corresponde a fazer  $x_{B_l}$  sair da base e  $x_j$  entrar em seu lugar.

### Quando usar o simplex dual

Um caso bastante particular onde é interessante utilizar o simplex dual é quando o problema original é da forma  $\min c'x$  s.a  $Ax \leq b$ , onde  $c \geq 0$ . Passando este problema para a forma canônica temos que a base formada pelas variáveis residuais, embora não necessariamente viável primal, é viável dual pois

$$\bar{c}' = c' - \overset{=0}{c'_B} B^{-1}A = c' \geq 0.$$

Assim podemos formar o tableau associado a esta base e aplicar o método simplex dual.

Outra situação ainda mais relevante é quando conhecemos uma solução ótima para um (PLC) e queremos resolver o mesmo problema com um lado direito  $b$  diferente. A mudança do vetor  $b$  pode prejudicar a viabilidade primal da base ótima conhecida, mas não afeta o vetor de custos reduzidos (pois este não depende de  $b$ ) e assim a base conhecida é viável dual, o que permite a utilização do simplex dual. Voltaremos a este exemplo no capítulo 5.

### A geometria do simplex dual

Observe que a solução dual construída a partir de uma base primal ( $p' = c'_B B^{-1}$ ) satisfaz  $p'B = p' [A^{B_1} | \dots | A^{B_m}] = c'_B$ , ou equivalentemente

$$(A^{B_i})'p = c_{B_i}, \quad i = 1, 2, \dots, m$$

o que corresponde a  $m$  restrições do problema dual, ativas em  $p$ , e linearmente independentes (pois as colunas da base são linearmente independentes). Como as variáveis duais estão em  $\mathbb{R}^m$ , temos que  $p' = c'_B B^{-1}$  é na realidade uma solução básica do dual. Assim, uma base do (PLC) está associada tanto a uma solução básica primal quanto a uma solução básica dual. Estas soluções (primal e dual) podem ser viáveis ou inviáveis; a base em questão será ótima se e somente se as duas soluções forem viáveis (primal e dual respectivamente).

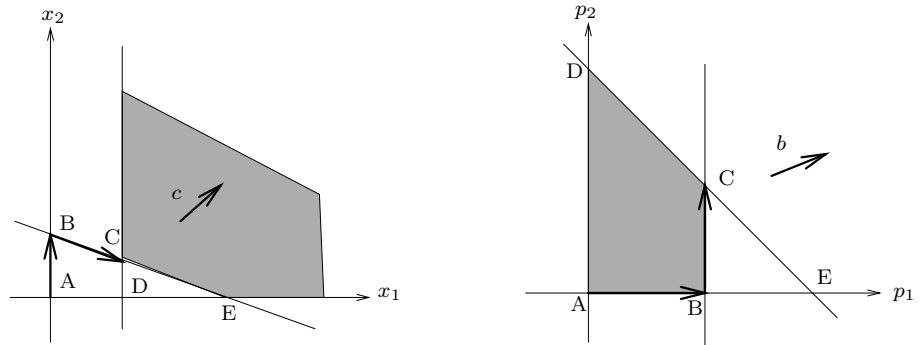
Mais ainda, ao passar de uma base para a outra o simplex (dual) preserva  $m - 1$  colunas l.i. que são comuns às duas bases, o que corresponde no dual, de acordo com o argumento anterior, a  $m - 1$  restrições ativas linearmente independentes em comum entre as duas soluções

básicas duais correspondentes. Assim o simplex dual percorre um caminho de soluções básicas viáveis que são adjacentes no poliedro associado ao problema dual, até atingir otimalidade dual (que corresponde à viabilidade primal da solução  $x$  associada). No entanto, aplicar o simplex dual não corresponde necessariamente a aplicar o simplex primal ao problema dual passado para a forma canônica (um exemplo pode ser encontrado no exercício 4.25 do livro).

**Exemplo 4.8** Considere o par de problemas primal-dual

$$\begin{array}{llll} \min & x_1 + x_2 & & \max & 2p_1 + 2p_2 \\ \text{s.a} & x_1 + 2x_2 - x_3 & = & 2 & \text{s.a} & p_1 + p_2 \leq 1 \\ & x_1 & & -x_4 = 1 & & 2p_1 \leq 1 \\ & x \geq 0 & & & & p \geq 0 \end{array}$$

Os conjuntos viáveis dos dois problemas são ilustrados a seguir (no primal as variáveis  $x_3$  e  $x_4$  são interpretadas como variáveis de folga).



Há um total de 5 bases no primeiro problema, associadas a 5 soluções básicas distintas, identificadas na figura como A, B, C, D e E. As soluções básicas correspondentes no dual foram identificadas com as mesmas letras. Como exemplo, a base formada pelas colunas  $A^3$  e  $A^4$  determinam a solução básica primal  $x = (0, 0, -2, -1)'$  (ponto A). A solução dual correspondente é obtida fazendo-se  $p'A^3 = c_3$  e  $p'A^4 = c_4$ , cuja solução é  $p = (0, 0)'$ . Esta é uma solução básica viável no dual, e pode ser utilizada para inicializar o simplex dual:

			$\downarrow$		
		$x_1$	$x_2$	$x_3$	$x_4$
	0	1	1	0	0
$\leftarrow x_3 =$	-2	-1	-2	1	0
$x_4 =$	-1	-1	0	0	1

Aplicando o simplex dual, teremos

		$\downarrow$	$x_1$	$x_2$	$x_3$	$x_4$
	$-1$	$\frac{1}{2}$	$0$	$\frac{1}{2}$	$0$	
$x_2 =$	$1$	$\frac{1}{2}$	$1$	$-\frac{1}{2}$	$0$	
$\leftarrow x_4 =$	$-1$	$-1$	$0$	$0$	$1$	

		$x_1$	$x_2$	$x_3$	$x_4$
	$-\frac{3}{2}$	$0$	$0$	$\frac{1}{2}$	$\frac{1}{2}$
$x_2 =$	$\frac{1}{2}$	$0$	$1$	$-\frac{1}{2}$	$\frac{1}{2}$
$x_1 =$	$1$	$1$	$0$	$0$	$-1$

Esta seqüência de tableaus corresponde ao caminho  $A$ - $B$ - $C$  tanto no problema primal quanto no dual. No primal o caminho percorre uma seqüência de bases inviáveis até encontrar uma base viável que também é ótima. No dual, o caminho tem as mesmas propriedades do simplex primal: as bases são todas viáveis e a função objetivo (dual) melhora (aumenta) a cada passo.

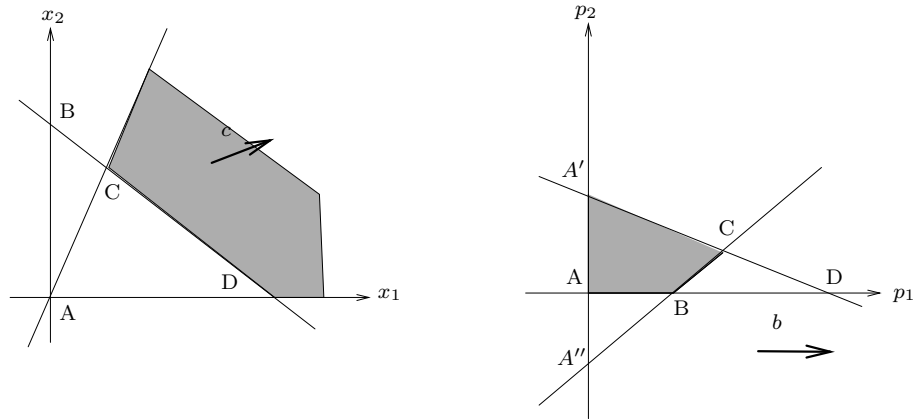
## Dualidade e degenerescência

Dada uma base  $B$  do (PLC), considere a solução básica dual  $p' = c'_B B^{-1}$  associada. Se esta solução é degenerada, então além das restrições  $p' A^{B_i} = c_{B_i}$ ,  $i = 1, \dots, m$  teremos alguma outra restrição  $p' A^j = c_j$  ativa, correspondente a uma variável  $x_j$  primal não-básica; conseqüentemente teremos  $\bar{c}_j = c_j - p' A^j = 0$ . Isso mostra que degenerescência da solução dual associada à base  $B$  corresponde a termos algum custo reduzido nulo associado a uma variável não-básica.

**Exemplo 4.9** Considere o par primal-dual

$$\begin{array}{ll}
 \min & 3x_1 + x_2 \\
 \text{s.a.} & x_1 + x_2 \geq 2 \\
 & 2x_1 - x_2 \geq 0 \\
 & x \geq 0
 \end{array}
 \qquad
 \begin{array}{ll}
 \max & 2p_1 \\
 \text{s.a.} & p_1 + 2p_2 \leq 3 \\
 & p_1 - p_2 \leq 1 \\
 & p \geq 0
 \end{array}$$

Os conjuntos viáveis dos dois problemas são ilustrados a seguir



Observe que há um total de seis bases no problema primal, mas apenas quatro soluções básicas (pontos A, B, C e D). No dual as seis bases correspondem a seis soluções básicas distintas (pontos A, A', A'', B, C e D).

Lembrando que o dual do dual é o primal, poderíamos pensar por um instante no problema da direita como primal. A solução básica associada ao ponto A (da figura da direita) tem como solução dual básica associada o ponto  $(0,0)'$  na figura da esquerda. A degenerescência desta última solução mostra que existem arestas saindo de A (na figura da direita) que não alteram o valor da função objetivo; estas são as arestas que vão para A' e A''. Isto é análogo a existirem custos reduzidos iguais a zero associados a direções básicas, o que observaríamos se transformássemos o problema da direita em um (PLC) (lembre-se que bases, custos reduzidos e direções básicas só estão definidos para o (PLC)).

## 4.6 Sistemas Lineares e o Lema de Farkas

Vimos na seção anterior que resolver um problema de programação linear não é mais difícil do que resolver sistemas de equações e desigualdades lineares em geral. Veremos agora alguns resultados provenientes da teoria de dualidade em programação linear para decidir se um dado sistema linear possui ou não solução.

**Teorema 4.6 (Lema de Farkas)** *Seja  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ . Então exatamente uma das afirmativas seguintes é verdadeira:*

1. Existe  $x \geq 0$  tal que  $Ax = b$ ;
2. Existe  $p$  tal que  $A'p \geq 0$  e  $p'b < 0$ .



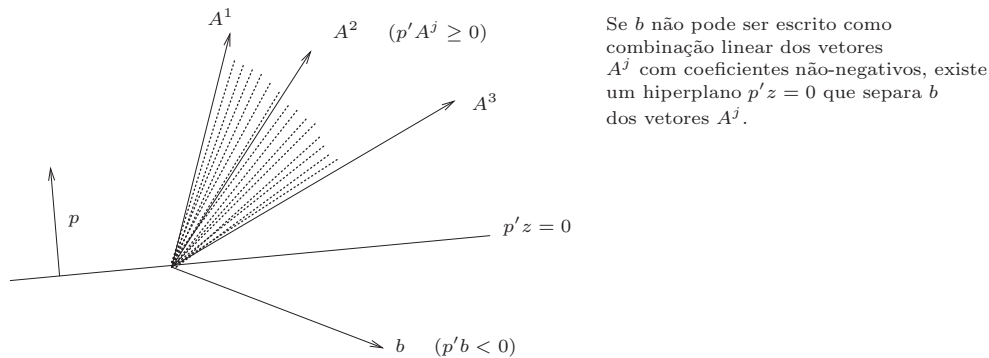
## Prova.

Suponha que vale (1). Então (2) não pode ser verdadeira, pois  $A'p \geq 0 \implies p'b = p'Ax = (A'p)'x \geq 0$ .

Por outro lado, suponha que não vale (1). Considere o par primal-dual

$$\begin{cases} \min & p'b \\ \text{s.a} & A'p \geq 0 \end{cases} \quad \begin{cases} \max & 0'x \\ \text{s.a} & Ax = b \\ & x \geq 0. \end{cases}$$

Observe que o primal é viável (pois  $A'0 \geq 0$ ) e que o dual é inviável por hipótese. A única possibilidade para o primal é ser ilimitado (pois o primal é viável, e se possuísse solução ótima o dual também possuiria, pelo teorema de dualidade forte). Logo o valor ótimo do primal é  $-\infty$  e, em particular, existe uma solução  $p$  viável ( $A'p \geq 0$ ) que satisfaz  $p'b < 0$ .



Observe que este lema permite transformar uma proposição do tipo  $\nexists x : \dots \equiv \forall x : \neg \dots$  (universal, em princípio mais difícil) em uma proposição do tipo  $\exists p : \dots$  (existencial, em princípio mais fácil).

**Corolário 4.3** *Sejam  $A^1, \dots, A^n, b \in \mathbb{R}^m$  e suponha que qualquer vetor  $p$  que satisfaz  $p'A^j \geq 0, \forall j$  também satisfaz  $p'b \geq 0$ . Então  $b$  pode ser escrito como combinação linear de  $A^1, \dots, A^n$  com coeficientes não-negativos.*

## 4.7 De Hiperplanos Separadores a Dualidade

Lembremos que nosso desenvolvimento de dualidade, em particular do teorema forte de dualidade, esteve baseado no fato de que o simplex, munido de uma regra anti-ciclagem para escolha do pivô, termina em

um número finito de passos para qualquer problema de programação linear na forma canônica. O método simplex foi utilizado na demonstração daquele teorema para construir uma solução ótima dual com o mesmo valor da solução ótima primal (cuja existência era hipótese do teorema forte). Usando o teorema forte de dualidade pudemos demonstrar o lema de Farkas, que pode ser interpretado em termos da existência de um hiperplano que separa o vetor  $b$  das colunas de  $A$ .

Nesta seção vamos mostrar que é possível chegar ao teorema forte de dualidade por um caminho puramente geométrico, independente do método simplex. Esta derivação começa a partir de um resultado de hiperplanos separadores para conjuntos convexos, estabelece o lema de Farkas a partir deste resultado e finalmente demonstraremos o teorema forte de dualidade a partir do lema de Farkas. Esta linha de argumentação pode ser generalizada para problemas não-lineares.

Alguns conceitos e resultados básicos serão apresentados a seguir.

**Definição** Um conjunto  $S \subset \mathbb{R}^n$  é **fechado** se para qualquer seqüência  $\{x^k\}_{k \in \mathbb{N}} \subset S$  convergente temos que  $(\lim_{k \rightarrow \infty} x^k) \in S$ .

**Teorema 4.9** *Todo poliedro é fechado.*

**Prova.**

Todo poliedro pode ser representado na forma geral  $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ . Suponha que  $\{x^k\}_{k \in \mathbb{N}} \subset P$  é uma seqüência tal que  $\lim_{k \rightarrow \infty} x^k = x^*$ . Como  $x^k \in P$  temos  $Ax^k \geq b, \forall k \in \mathbb{N}$ . Como  $f(x) = Ax$  é uma função contínua, temos que

$$Ax^* = A \left( \lim_{k \rightarrow \infty} x^k \right) = \left( \lim_{k \rightarrow \infty} Ax^k \right) \geq b,$$

onde a última desigualdade é consequência do fato que a componente  $i$  do vetor  $(\lim_{k \rightarrow \infty} Ax^k)$  é igual a  $\lim_{k \rightarrow \infty} A_i x^k$  onde  $A_i x^k \geq b_i, \forall k \in \mathbb{N}$ .

**Teorema 4.10 (Teorema de Weierstrass)** *Se  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é contínua e  $S \subset \mathbb{R}^n$  é não-vazio, fechado e limitado, então  $f$  admite máximo e mínimo em  $S$ , ou seja, existem  $\bar{x}, \hat{x} \in S$  tais que*

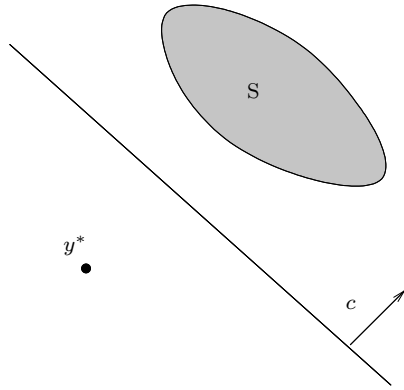
$$f(\hat{x}) \leq f(x) \leq f(\bar{x}), \forall x \in S.$$

**Prova.**

A prova deste resultado foge ao escopo deste curso, mas pode ser encontrada no livro *Real Analysis* de W. Rudin.

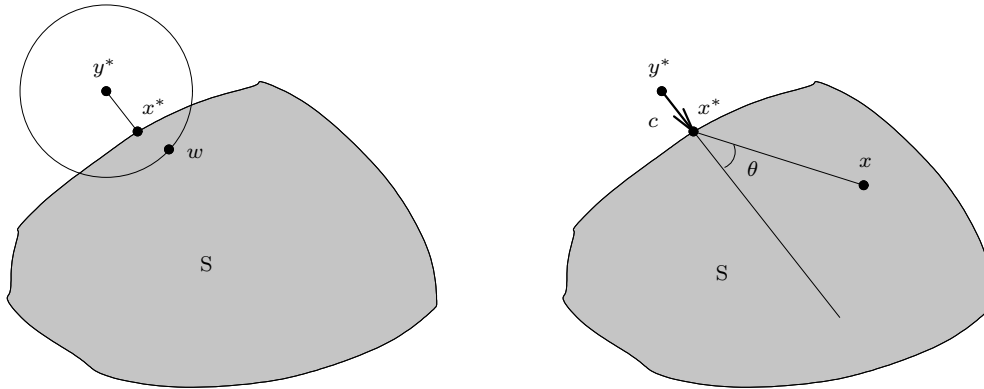
Observe que o teorema de Weierstrass depende fundamentalmente do fato de  $S$  ser fechado e limitado:  $f(x) = x$  não admite nem mínimo nem máximo em  $S = \{x \in \mathbb{R} \mid x > 0\}$ .

**Teorema 4.11 (Teorema do Hiperplano Separador)** *Seja  $S \subset \mathbb{R}^n$  convexo, fechado e não-vazio, e seja  $y^* \in \mathbb{R}^n$  tal que  $y^* \notin S$ . Então existe um vetor  $c \in \mathbb{R}^n$  tal que  $c'y^* < c'x, \forall x \in S$ .*



**Prova.**

Como  $S \neq \emptyset$ , considere um  $w \in S$  qualquer. Seja  $B = \{x \in \mathbb{R}^n \mid \|x - y^*\| \leq \|w - y^*\|\}$  e  $D = S \cap B$ .



Observe que  $D$  é não-vazio ( $w \in D$ ), fechado, limitado e convexo (pois  $S$  e  $B$  são fechados e convexos), e que  $f(x) = \|x - y^*\|$  é uma função contínua. Pelo teorema de Weierstrass existe um  $x^* \in D$  tal que  $\|x^* - y^*\| \leq \|x - y^*\|, \forall x \in D$ . Para  $x \in S \setminus D$  temos  $\|x - y^*\| > \|w - y^*\| \geq \|x^* - y^*\|$ , o que mostra que

$$\|x^* - y^*\| \leq \|x - y^*\|, \forall x \in S.$$

Seja  $x \in S$ . Como  $S$  é convexo, para qualquer  $\lambda \in (0, 1)$  temos que  $x^* + \lambda(x - x^*) = \lambda x + (1 - \lambda)x^* \in S$ . Usando a desigualdade anterior, temos

$$\begin{aligned}\|x^* - y^*\|^2 &\leq \|x^* + \lambda(x - x^*) - y^*\|^2 \\ &= \|x^* - y^*\|^2 + 2\lambda(x^* - y^*)'(x - x^*) + \lambda^2\|x - x^*\|^2,\end{aligned}$$

e assim  $(x^* - y^*)'(x - x^*) \geq -\frac{\lambda}{2}\|x - x^*\|$ . Tomando o limite para  $\lambda \rightarrow 0$  temos

$$(x^* - y^*)'(x - x^*) \geq 0.$$

Isso mostra que o ângulo  $\theta$  representado na figura anterior é agudo. Logo

$$\begin{aligned}(x^* - y^*)'x &\geq (x^* - y^*)'x^* \\ &= (x^* - y^*)'y^* + (x^* - y^*)'(x^* - y^*) \\ &> (x^* - y^*)'y^*.\end{aligned}$$

Fazendo-se  $c = x^* - y^*$  temos o resultado.

### Lema de Farkas como consequência do Teorema do Hiperplano Separador

Lembremos do enunciado do Lema de Farkas: *Seja  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ . Então exatamente uma das afirmativas seguintes é verdadeira:*

1. *Existe  $x \geq 0$  tal que  $Ax = b$ ;*
2. *Existe  $p$  tal que  $A'p \geq 0$  e  $p'b < 0$ .*

A primeira parte da demonstração é idêntica à da primeira prova: Supondo que vale (1), então  $A'p \geq 0 \implies p'b = p'Ax = (A'p)'x \geq 0$ , o que mostra que (2) não pode ser verdadeira.

Por outro lado, suponha que não vale (1). Considere o cone gerado pelas colunas de  $A$

$$S = \{Ax \mid x \geq 0\} = \{y \in \mathbb{R}^n \mid \exists x \geq 0 : Ax = y\}.$$

Como (1) é falso,  $b \notin S$ . Observe que  $S$  é não-vazio (pois  $0 \in S$ ), fechado ( $S$  é um poliedro, tente provar) e convexo ( $x^1, x^2 \geq 0 \implies \lambda x^1 + (1 - \lambda)x^2 \geq 0$ . Portanto  $Ax^1, Ax^2 \in S \implies \lambda Ax^1 + (1 - \lambda)Ax^2 = A(\lambda x^1 + (1 - \lambda)x^2) \in S$ ). Pelo teorema do hiperplano separador, existe um  $p \in \mathbb{R}^m$  tal que  $p'b < p'y$ ,  $\forall y \in S$ . Como  $0 \in S$ ,  $p'b < 0$ . Como  $\lambda A^i \in S$ ,  $\forall \lambda > 0$  (pois  $\lambda A^i = A(\lambda e^i)$ ) temos  $p'b < \lambda p'A^i$  e conseqüentemente  $p'A^i > \frac{p'b}{\lambda}$ . Fazendo  $\lambda \rightarrow +\infty$  temos  $p'A^i \geq 0$ . Como isso vale para  $i = 1, \dots, m$  temos  $A'p \geq 0$ , completando a prova.

## O Teorema Forte de Dualidade como consequência do Lema de Farkas

Considere o par primal-dual

$$\begin{array}{ll} \min & c'x \\ \text{s.a} & Ax \geq b \end{array} \qquad \begin{array}{ll} \max & p'b \\ \text{s.a} & A'p = c \\ & p \geq 0. \end{array}$$

Vamos provar que se o primal possui uma solução ótima  $x^*$  então o dual possui uma solução ótima com o mesmo valor da função objetivo. A extensão para qualquer problema de programação linear segue a mesma linha da prova original deste teorema.

Considere  $I = \{i \mid A_i x^* = b_i\}$  o conjunto das restrições ativas em  $x^*$ . Observe que o conjunto das direções viáveis a partir de  $x^*$  é

$$\begin{aligned} D &= \{d \in \mathbb{R}^n \mid \exists \varepsilon > 0 : A(x^* + \varepsilon d) \geq b\} \\ &= \left\{ d \in \mathbb{R}^n \mid \exists \varepsilon > 0 : \begin{cases} \varepsilon A_i d \geq b_i - A_i x^* = 0, & i \in I \\ \varepsilon A_i d \geq b_i - A_i x^* < 0, & i \notin I \end{cases} \right\} \\ &= \{d \in \mathbb{R}^n \mid A_i d \geq 0, i \in I\}. \end{aligned}$$

Como  $x^*$  é ótimo, temos  $c'd \geq 0$ ,  $\forall d \in D$  ou, analogamente, não existe  $d$  tal que  $A_i d \geq 0$ ,  $i \in I$  e  $c'd < 0$ . Usando o lema de Farkas, temos que  $c'$  pode ser escrito como combinação linear dos  $A_i$ ,  $i \in I$  com coeficientes não-negativos. Equivalentemente, existem  $p_i \geq 0$  tais que  $\sum_{i \in I} p_i A_i = c'$ . Fazendo  $p_i = 0$ ,  $i \notin I$  teremos  $p \in \mathbb{R}^m$ ,  $p \geq 0$  e  $p'A = c'$  ( $A'p = c$ ), o que mostra que  $p$  é viável no problema dual. Além disso,

$$p'b = \sum_{i \in I} p_i b_i = \sum_{i \in I} p_i A_i x^* = \left( \sum_{i \in I} p_i A_i \right) x^* = c'x^*,$$

o que, juntamente com o teorema fraco de dualidade, mostra que  $p$  é ótimo dual e o valor ótimo dual é igual ao valor ótimo primal.

**Exercícios sugeridos para o capítulo 4:** 4.1-4.5, 4.8, 4.10-4.16, 4.18, 4.22, 4.25-4.27, 4.31, 4.35, 4.43(a).

## 5 Análise de Sensibilidade

Neste capítulo consideramos o problema de programação linear na forma canônica e seu dual

$$(PLC) \begin{cases} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0 \end{cases} \quad (DLC) \begin{cases} \max & p'b \\ \text{s.a} & A'p \leq c \\ & p \in \mathbb{R}^m \end{cases}$$

onde  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^n$ . Vamos estudar a dependência do valor ótimo e da solução ótima do problema primal em relação aos dados do problema:  $A$ ,  $b$  e  $c$ . Em particular, vamos determinar condições sob as quais uma variação em um destes dados não altera a solução ótima do problema original, bem como buscar mecanismos de solução quando isto não ocorre, aproveitando a informação da solução ótima do problema original.

Considere que  $x^*$  é uma solução ótima para o (PLC), associada à matriz básica  $B$ . As propriedades

$$\begin{aligned} B^{-1}b &\geq 0 && \text{viabilidade} \\ c' - c'_B B^{-1}A &\geq 0 && \text{otimalidade} \end{aligned}$$

são válidas nesta solução. Veremos a seguir como uma alteração nos parâmetros do problema afeta estas propriedades.

### Acréscimo de uma nova variável

Suponha que o problema passe a ter uma nova variável  $x_{n+1}$  tornando-se

$$(PLC) \begin{cases} \min & c'x + c_{n+1}x_{n+1} \\ \text{s.a} & Ax + A^{n+1}x_{n+1} = b \\ & x, x_{n+1} \geq 0. \end{cases}$$

---

<sup>1</sup>Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.

Observe que a solução  $x = x^*$ ,  $x_{n+1} = 0$  é uma solução básica viável associada à base original  $B$ . Assim a condição de viabilidade não é alterada. Os custos reduzidos das variáveis originais também permanecem os mesmos. O custo reduzido da nova variável será

$$\bar{c}_{n+1} = c_{n+1} - c'_B B^{-1} A_{n+1}.$$

Temos então duas situações possíveis:

1.  $\bar{c}_{n+1} \geq 0$ . Neste caso a solução  $x = x^*$ ,  $x_{n+1} = 0$  é ótima para o problema modificado;
2.  $\bar{c}_{n+1} < 0$ . Neste caso podemos obter a nova solução ótima aplicando o método simplex usando a solução básica  $x = x^*$ ,  $x_{n+1} = 0$  como solução inicial. No caso do simplex tabular, podemos calcular a nova coluna do tableau como

$$\left[ \frac{\bar{c}_{n+1}}{B^{-1} A^{n+1}} \right].$$

Empiricamente podemos esperar que o simplex a partir desta solução necessite de um número de iterações muito menor do que se o problema modificado fosse resolvido sem informação nenhuma.

**Exemplo 5.1** Considere o problema

$$\begin{aligned} \min \quad & -5x_1 - x_2 + 12x_3 \\ \text{s.a} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & 5x_1 + 3x_2 + x_4 = 16 \\ & x \geq 0. \end{aligned}$$

e a solução ótima  $x^* = (2, 2, 0, 0)'$  com o tableau ótimo

		$x_1$	$x_2$	$x_3$	$x_4$
	12	0	0	2	7
$x_1 =$	2	1	0	-3	2
$x_2 =$	2	0	1	5	-3

Observe que a matriz  $B^{-1}$  aparece nas duas últimas colunas do tableau.

Considere agora o problema com uma nova variável  $x_5$ :

$$\begin{aligned} \min \quad & -5x_1 - x_2 + 12x_3 - x_5 \\ \text{s.a} \quad & 3x_1 + 2x_2 + x_3 + x_5 = 10 \\ & 5x_1 + 3x_2 + x_4 + x_5 = 16 \\ & x \geq 0. \end{aligned}$$

Partindo da mesma solução básica  $x = (2, 2, 0, 0)'$  teremos que o custo reduzido em relação à nova variável é

$$\bar{c}_5 = c_5 - c'_B B^{-1} A^5 = -1 - [-5 \quad -1] \begin{bmatrix} -3 & 2 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -4 < 0,$$

o que mostra que esta base não é ótima para o novo problema. Aumentamos o tableau incluindo a coluna  $B^{-1}A^5 = (-1, -2)'$ , e aplicamos o simplex:

			$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	↓
		12	0	0	2	7	-4	
$x_1 =$	2		1	0	-3	2	-1	
← $x_2 =$	2		0	1	5	-3	2	

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
	16	0	2	12	1	0
$x_1 =$	3	1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	0
$x_5 =$	1	0	$\frac{1}{2}$	$\frac{5}{2}$	$-\frac{3}{2}$	1

Assim encontramos a nova solução ótima  $x^* = (3, 0, 0, 0, 1)'$ .

## Acréscimo de uma nova restrição de desigualdade

Suponha que acrescentemos ao problema original uma restrição da forma  $A_{m+1}x \geq b_{m+1}$ . Se a solução ótima do problema original satisfaz  $A_{m+1}x^* \geq b_{m+1}$ , então esta solução também é ótima para o novo problema (prove!). Do contrário, passamos esta restrição para a forma canônica introduzindo uma variável de folga  $x_{n+1}$ , e obtemos um novo problema na forma canônica com a matriz

$$\bar{A} = \begin{bmatrix} A & 0 \\ A_{m+1} & -1 \end{bmatrix}.$$

Seja  $B$  uma base ótima do problema original. Podemos contruir uma solução  $(x^*, x_{n+1})$  definindo  $x_{n+1} = A_{m+1}x^* - b_{m+1}$ . Esta solução é inviável, mas está associada à base formada pelas variáveis  $x_{B_1}, \dots, x_{B_m}, x_{n+1}$ , com a matriz básica dada por

$$\bar{B} = \begin{bmatrix} B & 0 \\ a' & -1 \end{bmatrix},$$

onde  $a$  contém as componentes de  $A_{m+1}$  que estão na base original. Note que esta matriz é de fato inversível, com a inversa dada por

$$\bar{B}^{-1} = \begin{bmatrix} B^{-1} & 0 \\ a'B^{-1} & -1 \end{bmatrix}.$$



Temos assim uma solução primal básica inviável. Note que os custos reduzido em relação a esta solução básica são dados por

$$\begin{aligned} c'_{\text{nov}} - c'_B \bar{B}^{-1} \bar{A} &= [c' \ 0] - [c'_B \ 0] \begin{bmatrix} B^{-1} & 0 \\ a'B^{-1} & -1 \end{bmatrix} \begin{bmatrix} A & 0 \\ A_{m+1} & -1 \end{bmatrix} \\ &= [c' - c'_B B^{-1} A \ 0] \geq 0, \end{aligned}$$

e são não-negativos porque  $B$  é ótima para o problema original. Assim  $\bar{B}$  é uma base viável dual, e podemos aplicar o simplex dual para o novo problema, a partir do tableau

		$x_1 \quad \dots \quad x_n \quad x_{n+1}$
		$c' - c'_B B^{-1} A \quad 0$
$x_B =$	$x_B^*$	$B^{-1} A \quad 0$
$x_{n+1} =$	$A_{m+1} x^* - b_{m+1}$	$a' B^{-1} A - A_{m+1} \quad 1$

**Exemplo 5.2** Considere outra vez o problema

$$\begin{aligned} \min \quad & -5x_1 - x_2 + 12x_3 \\ \text{s.a} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & 5x_1 + 3x_2 + x_4 = 16 \\ & x \geq 0. \end{aligned}$$

com o tableau ótimo

		$x_1 \quad x_2 \quad x_3 \quad x_4$
	12	0 \quad 0 \quad 2 \quad 7
$x_1 =$	2	1 \quad 0 \quad -3 \quad 2
$x_2 =$	2	0 \quad 1 \quad 5 \quad -3

e suponha que seja acrescida a restrição  $x_1 + x_2 \geq 5$ , que é violada pela antiga solução ótima  $x^* = (2, 2, 0, 0)'$ . A última linha do tableau aumentado será dada por

$$a' B^{-1} A - A_{m+1} = [1 \ 1] \begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & 5 & -3 \end{bmatrix} - [1 \ 1 \ 0 \ 0] = [0 \ 0 \ 2 \ -1]$$

e aplicando o simplex dual a partir do tableau aumentado teremos:

					$\downarrow$	
		$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$				
	12	0 \quad 0 \quad 2 \quad 7 \quad 0				
$x_1 =$	2	1 \quad 0 \quad -3 \quad 2 \quad 0				
$x_2 =$	2	0 \quad 1 \quad 5 \quad -3 \quad 0				
$\leftarrow x_5 =$	-1	0 \quad 0 \quad 2 \quad -1 \quad 1				

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
	5	0	0	16	0	7
$x_1 =$	0	1	0	1	0	2
$x_2 =$	5	0	1	-1	0	-3
$x_4 =$	1	0	0	-2	1	-1

## Mudanças no vetor $b$

Suponha que a  $i$ -ésima componente do vetor  $b$  torne-se  $b + \delta e^i$ . Gostaríamos de determinar os valores de  $\delta$  para os quais a base ótima do problema original permanece ótima. Note que a condição de otimalidade primal não depende do vetor  $b$ , assim apenas a viabilidade pode ser afetada. Para manter viabilidade precisamos garantir que a condição  $B^{-1}(b + \delta e^i) \geq 0$  continue valendo. Seja  $g = B^{-1}e^i$  a  $i$ -ésima coluna da matriz  $B^{-1}$ . A condição de viabilidade é equivalente a

$$x_B + \delta g \geq 0$$

ou ainda

$$x_{B_j} + \delta g_j \geq 0, \quad j = 1, \dots, m.$$

Esta condição é válida se e somente se  $\delta \geq -\frac{x_{B_j}}{g_j}$ ,  $\forall g_j > 0$  e  $\delta \leq -\frac{x_{B_j}}{g_j}$ ,  $\forall g_j < 0$ , ou equivalentemente

$$\max_{g_j > 0} -\frac{x_{B_j}}{g_j} \leq \delta \leq \min_{g_j < 0} -\frac{x_{B_j}}{g_j}.$$

Se  $\delta$  está neste intervalo, a base ótima do problema original permanece ótima e o valor ótimo pode ser expresso como função de  $\delta$ :

$$c'_B B^{-1}(b + \delta e^i) = p'b + \delta p_i,$$

onde  $p' = c'_B B^{-1}$  é a solução dual ótima associada à base  $B$ .

Se  $\delta$  está fora do intervalo acima, a solução antiga é inviável mas satisfaz a condição de otimalidade, equivalente à viabilidade dual, e assim podemos usar o simplex dual para encontrar a solução do novo problema.

**Exemplo 5.3** Considere o tableau ótimo

		$x_1$	$x_2$	$x_3$	$x_4$
	12	0	0	2	7
$x_1 =$	2	1	0	-3	2
$x_2 =$	2	0	1	5	-3

e suponha que a primeira componente do vetor  $b$  passe a ser  $10 + \delta$ . A primeira coluna de  $B^{-1}$  é  $(-3, 5)'$ , assim a condição de viabilidade passa a ser  $x_1 = 2 - 3\delta \geq 0$  e  $x_2 = 2 + 5\delta \geq 0$ , ou ainda,  $-\frac{2}{5} \leq \delta \leq \frac{2}{3}$ . O valor ótimo em função de  $\delta$  será  $-12 + 10\delta$ , pois  $p_1 = c'_B B^{-1} e^1 = [-5 \quad -1] \begin{bmatrix} -3 \\ 5 \end{bmatrix} = 10$ .

Se  $\delta \notin [-\frac{2}{5}, \frac{2}{3}]$  então uma das duas variáveis básicas fica negativa, e podemos aplicar o simplex dual a partir desta base.

## Mudanças no vetor $c$

Suponha agora que o vetor  $c$  passe a ser  $c + \delta e^j$ . A viabilidade primal não é afetada (não depende da função objetivo), mas a condição de otimalidade  $c'_B B^{-1} A \leq c'$  pode ser afetada.

Inicialmente suponha que  $j$  é uma variável não-básica. Como o vetor  $c_B$  não muda, a única desigualdade afetada é a do custo reduzido associado a  $x_j$ :

$$c'_B B^{-1} A^j \leq c_j + \delta \iff \delta \geq -\bar{c}_j.$$

Se esta condição é satisfeita, então a base ótima do problema original permanece ótima no problema modificado. Do contrário, podemos usar a base ótima do problema original como base inicial para o método simplex aplicado ao novo problema. Note que, no tableau, apenas a coluna associada a  $x_j$  é modificada.

Suponha, por outro lado, que  $x_j = x_{B_l}$ . Neste caso o vetor  $c_B$  passa a ser  $c_B + \delta e^l$  e todas as desigualdades da condição de otimalidade serão afetadas:

$$\begin{aligned} (c_B + \delta e^l)' B^{-1} A^i &\leq c_i, & \forall i \neq B_l \\ (c_B + \delta e^l)' B^{-1} A^{B_l} &\leq c_{B_l} + \delta \end{aligned}$$

Observe que a última desigualdade é trivialmente satisfeita, pois  $(c_B + \delta e^l)' B^{-1} A^{B_l} = (c_B + \delta e^l)' e^l = c_{B_l} + \delta$ . Observe ainda que para os demais índices da base,  $(c_B + \delta e^l)' B^{-1} A^{B_k} = (c_B + \delta e^l)' e^k = c_{B_k}$ , e a desigualdade correspondente também é trivialmente satisfeita.

Ficamos assim com

$$\delta q_{li} \leq \bar{c}_i, \quad \forall i \notin \{B_1, \dots, B_m\},$$

onde  $q_{li} = (e^l)' B^{-1} A^i = B_l^{-1} A^i$  é o valor da linha  $l$  e coluna  $i$  do tableau. Estas desigualdades determinam uma faixa de valores de  $\delta$  para os quais a base anterior permanece ótima.

**Exemplo 5.4** Voltando ao problema dos exemplos anteriores, podemos calcular, para as variáveis não-básicas  $x_3$  e  $x_4$ , os valores de  $c_3 + \delta_3$  e  $c_4 + \delta_4$  que não perturbam a otimalidade da solução  $x^* = (2, 2, 0, 0)'$ :

$$\begin{aligned}\delta_3 &\geq -\bar{c}_3 = -2, \\ \delta_4 &\geq -\bar{c}_4 = -7.\end{aligned}$$

Uma variação  $c_1 + \delta_1$  no custo associado à variável básica  $x_1$  terá que satisfazer

$$\begin{aligned}\delta_1 q_{13} \leq \bar{c}_3 &\iff \delta_1 \geq -\frac{2}{3}, \\ \delta_1 q_{14} \leq \bar{c}_4 &\iff \delta_1 \leq \frac{7}{2},\end{aligned}$$

ou ainda  $\delta \in [-\frac{2}{3}, \frac{7}{2}]$ , para não perturbar a otimalidade de  $x^*$ .

## Mudanças numa coluna não-básica de $A$

Suponha que a entrada  $A_{ij}$  da matriz passe a ser  $A_{ij} + \delta$ , onde  $x_j$  é uma variável não-básica. Como neste caso a matriz  $B$  não muda, a condição de viabilidade primal permanece inalterada. Além disso apenas o custo reduzido em relação à variável  $x_j$  é alterado, passando a ser

$$c_j - p'(A^j + \delta e^i),$$

onde  $p' = c'_B B^{-1}$ . Assim a condição de otimalidade a ser verificada é

$$c_j - p'(A^j + \delta e^i) \geq 0 \iff \bar{c}_j \geq \delta p_i.$$

Se esta condição é desrespeitada, aplicamos o simplex ao tableau modificado (somente a coluna de  $x_j$  é alterada), e a variável  $x_j$  entrará na base.

**Exemplo** No exemplo anterior, se a entrada  $A_{13}$  da matriz tornar-se  $1 + \delta$ , a otimalidade será preservada somente quando  $\delta$  satisfizer  $\bar{c}_3 \geq \delta p_1$  ou, equivalentemente,

$$\delta \leq \frac{2}{10}.$$

**Exercícios sugeridos para o capítulo 5:** 5.4-5.8, 5.10, 5.13(a,b).