

Otimização Linear

Nelson Maculan e Marcia H. Costa Fampa
Universidade Federal do Rio de Janeiro
e-mail: {maculan, fampa}@cos.ufrj.br

15 de dezembro de 2004

Sumário

1	Rápido Histórico e Introdução	4
1.1	Histórico e Introdução	4
2	Definições e Propriedades dos Problemas de Programação Linear	7
2.1	Observações	17
2.2	Exercícios	18
3	Método do Simplex	19
3.1	Introdução	19
3.2	Algoritmo do Simplex	20
3.3	Busca de uma Solução Básica Viável	24
3.4	Cálculo de B^{-1}	25
3.5	Interpretação Geométrica da Mudança de Base	27
3.6	Convergência do Método do Simplex	33
3.6.1	Motivação	34
3.6.2	Exemplo de Ciclagem	35
3.6.3	Regra de Bland	36
3.7	Exercícios	38
4	Dualidade em Programação Linear	40
4.1	Uma Interpretação das Variáveis Duais	45
4.2	Método Dual do Simplex	45
4.3	Análise de Sensibilidade	51
4.3.1	Alterando o lado direito das restrições	52
4.3.2	Alterando o vetor custo	53
4.4	Pós-otimização	53
4.4.1	Alterando o lado direito das restrições	54
4.4.2	Alterando o vetor custo	55
4.4.3	Acrescentando mais Restrições	57
4.5	Exercícios	61

5	Obtenção de uma Solução Viável	64
5.1	Método de Eliminação de Fourier	64
5.2	Exemplos	65
5.3	Aplicação do Método de Fourier quando houver Igualdades . .	69
5.4	Um Método Heurístico	70
5.5	Métodos Introduzindo Soluções Artificiais	72
5.6	Viabilidade de Sistemas com Igualdades e Desigualdades Li- neares	73
5.7	Exercícios	75
6	Conjuntos Poliédricos Convexos	76
6.1	Hiperplano, Vértices, Raios e Direções	76
6.2	Conjunto Poliédrico Convexo	78
6.3	Caracterização das Arestas de um Conjunto Poliédrico Convexo	85
6.4	Teorema da representação de um Conjunto Poliédrico Convexo	87
6.5	Exercícios	90
7	Geração de Colunas	91
7.1	Introdução	91
7.2	Problema Auxiliar Limitado	105
7.3	Exercícios	105
8	Tratamento Implícito de Variáveis Limitadas	107
8.1	Método Primal do Simplex para Restrições Canalizadas	107
8.2	Busca de uma Solução Básica Viável	111
8.3	Método Dual do Simplex para Restrições Canalizadas	116
8.4	Exercícios	120
9	Métodos de Pontos Interiores	122
9.1	Introdução	122
9.2	O Método de Elipsóides	124
9.3	O Método Afim-Escala	128
9.4	A Trajetória Central	139
9.4.1	Condições de Karush-Kuhn-Tucker (ou KKT)	141
9.5	Algoritmos de Trajetória Central	146
9.6	Algoritmos de Redução Potencial	178
9.7	Exercícios	190

10 Introdução à Programação Inteira	192
10.1 Esquematizando o Método do Simplex por Operações entre Colunas	194
10.2 Métodos de Planos de Corte	196
10.3 Método de Balas para Otimização Linear 0-1	201
10.3.1 Esquema de Enumeração	202
10.3.2 Convergência do Método de Balas	207
10.3.3 Otimização Não Linear 0-1	208
10.4 Métodos de “Branch-and-Bound”	209
10.4.1 Ilustração dos Métodos de “Branch-and-Bound”	209
10.5 Exercícios	219
11 Dualidade em Programação Inteira	222
11.1 Relaxação Lagrangeana	222
11.2 Solução do Problema Dual	224
11.2.1 Técnica de geração de colunas	224
11.2.2 Método utilizando sub-gradientes	228
11.3 Exercícios	244

Capítulo 1

Rápido Histórico e Introdução

1.1 Histórico e Introdução

A Programação Linear ou Otimização Linear faz parte das disciplinas que compõem a Programação Matemática (Otimização) e é um elemento importantíssimo na Pesquisa Operacional. Nos métodos de solução propostos em Programação Não Linear com restrições há quase sempre a utilização da Programação Linear de maneira iterativa. Os problemas clássicos de otimização em grafos são problemas de Programação Linear específicos. A Programação Linear tem um papel duplo na Programação Matemática, pois os algoritmos utilizados para a sua solução podem ser de natureza combinatória (discreta) ou contínua. Nos problemas de Otimização Combinatória, modelados por problemas de Programação Linear Inteira, suas relaxações também serão problemas de Programação Linear. Três obras gerais sobre a Programação Matemática (Otimização), para que o leitor ou a leitora possa ter uma melhor idéia, são [Mino 86, NeRiTo 89, Maf 2000].

As aplicações da Otimização Linear são comuns em quase todos os setores do cotidiano, por exemplo, nas indústrias, nos transportes, na saúde, na educação, na agricultura, nas finanças, na economia, nas administrações públicas.

O primeiro algoritmo para a solução dos problemas de Programação Linear foi denominado método ou algoritmo do simplex, parece que foi Fourier [Fou 890] em 1826 que apresentou essa idéia visando à solução de sistemas de desigualdades lineares, mas foi Dantzig em 1947 que forneceu os resultados teóricos e computacionais do método do simplex, ver [Dan 51, Dan 63, Dan 91, Sc 86]. O método do simplex é um dos poucos algoritmos que foi implantado, comercialmente, em computador há mais de 35 anos, no início os fabricantes de computadores forneciam esses códigos, nos dias de hoje, os

códigos mais conhecidos possuem diferentes versões, dependendo do sistema operacional utilizado. Em 1939, o matemático Kantorovich [Ka 39] na União Soviética, já havia modelado e resolvido alguns problemas de otimização ligados ao planejamento econômico, graças a esses trabalhos, Leonid V. Kantorovich recebeu o prêmio Nobel de Economia em 1975.

Nos capítulos 2 a 8 apresentamos o método do simplex de maneira algébrica, numérica e geométrica, assim como algumas de suas extensões. Apesar de sua grande aplicação prática, o método do simplex pode ter comportamento exponencial em seu número de iterações. Em 1978, Khachian [Kha 79], utilizando os métodos elipsoidais, apresentou seu algoritmo, no qual o número máximo de iterações para resolver um problema de Programação Linear é limitado por uma função polinomial do tamanho dos dados do problema numa memória de computador. O método de Khachian [Kha 79] ou dos elipsóides, que apresentamos no capítulo 9, é um dos algoritmos mais elegantes em otimização, no entanto, sua performance prática deixa muito a desejar.

Em 1984, Karmarkar [Kar 84] propôs seu algoritmo de pontos interiores com comportamento numérico teórico polinomial e com bons resultados práticos. Após a publicação desse trabalho de Karmarkar [Kar 84], vários algoritmos de pontos interiores foram apresentados. Citamos alguns livros que tratam dos métodos de pontos interiores: [Ja 97, Vand 98, Wr 97]. Dedicamos o capítulo 9 a alguns desses algoritmos, assim como fornecemos alguns programas em MATLAB.

Nos capítulos 10 e 11 introduzimos os problemas de Otimização Inteira ou Programação Inteira ou Otimização Discreta ou Otimização Combinatória, isto é, problemas de Otimização Linear onde algumas ou todas variáveis são restritas a valores inteiros. Gomory [Go 58a, Go 58b], no final dos anos 50, propôs os primeiros algoritmos de planos de corte para solucionar esses problemas, ver [Ma 78, Ma 83, NeWo 88, Sc 86]. Temos a intenção de escrever um outro livro sobre a Otimização Inteira.

Procuramos listar na bibliografia deste livro um número razoável de referências didáticas em Programação Linear publicadas desde os anos 50, inclusive fizemos um esforço em buscar referências em língua portuguesa. A obra mais clássica é o livro de Dantzig [Dan 63], seguem-se várias outras, aqui em ordem alfabética: [Ar 93, BaJa 77, BrOlBo 81, ChCoHe 53, Ch 83, DoSaSo 58, Ga 60, Gar 60, Gas 58, GiMuWr 91, GiEl 71, GoYo 73, Gon 89, Ha 65, He 94, HuCa 87, Ja 97, Las 70, Lu 89, Mac 75, MaPe 80, Mu 76, Or 68, Pu 75, Sak 83, Sc 86, Si 72, Sim 58, St 72, Vand 98, We 90, Wr 97, Zi 74].

Outros livros de Otimização Combinatória trazem sempre uma introdução à Programação Linear, tais como: [GaNe 72, Gr 71, Hu 69, KoFi 69, Ko 80,

Law 76, Ma 78, Ma 83, NeWo 88, PaSt 82, Sa 75, SyDeKo 83, Ta 75, Var 72, Wo 98].

Utilizamos, desde o início, uma notação matricial para apresentar os resultados teóricos e computacionais. No que se refere à modelagem de problemas de decisão pela Programação Linear, solicitamos ao leitor ou à leitora que busque esse enfoque nas referências [BrHaMa 77, JaLa 98, Th 82, GuPrSe 2000, MaPe 80] e em livros de introdução à Pesquisa Operacional.

O conteúdo deste livro pode ser apresentado, parcial ou totalmente, nas disciplinas de Programação Matemática, de Otimização (Programação) Linear, de Pesquisa Operacional, de Otimização Combinatória nos seguintes cursos de graduação e pós-graduação: Engenharias, Matemática, Ciência da Computação, Física, Química, Economia, Administração, Estatística, Atuária.

Capítulo 2

Definições e Propriedades dos Problemas de Programação Linear

Um problema de programação linear pode ser definido sob a seguinte forma:

$$\text{maximizar } z = \sum_{j=1}^p c_j x_j \quad (2.1)$$

sujeito a:

$$\sum_{j=1}^p a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, q \quad (2.2)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, p, \quad (2.3)$$

onde c_j , a_{ij} e b_i são dados (números reais) e x_j representa para $j = 1, 2, \dots, p$, as variáveis de decisão. A função linear a ser maximizada em (2.1) é denominada função objetivo, função econômica ou função critério. As restrições de não negatividade (2.3) são conhecidas como triviais.

Cada restrição i de (2.2) pode ser substituída com o acréscimo de uma variável $x_{p+i} \geq 0$, denominada variável de folga, por uma restrição de igualdade e uma restrição trivial:

$$\sum_{j=1}^p a_{ij} x_j \leq b_i \Leftrightarrow \begin{cases} \sum_{j=1}^p a_{ij} x_j + x_{p+i} = b_i, \\ x_{p+i} \geq 0. \end{cases}$$

ou

$$\sum_{j=1}^p a_{ij} x_j \geq b_i \Leftrightarrow \begin{cases} \sum_{j=1}^p a_{ij} x_j - x_{p+i} = b_i, \\ x_{p+i} \geq 0. \end{cases}$$

O leitor ou a leitora poderá, facilmente, verificar esta afirmação.

Uma restrição de igualdade poderá também ser substituída por duas desigualdades:

$$\sum_{j=1}^p a_{ij}x_j = b_i \Leftrightarrow \begin{cases} \sum_{j=1}^p a_{ij}x_j \leq b_i, \\ \sum_{j=1}^p a_{ij}x_j \geq b_i. \end{cases}$$

Sendo dado um problema de programação linear com restrições de igualdades e desigualdades, poderemos acrescentar variáveis de folga às desigualdades não triviais, passando dessa maneira a trabalharmos com restrições de igualdades e desigualdades triviais.

Assim sendo, um problema de programação linear poderá sempre ser escrito da seguinte maneira:

$$(PPL) : \text{maximizar } z = \sum_{j=1}^n c_j x_j$$

sujeito a:

$$\begin{aligned} \sum_{j=1}^n a_{ij}x_j &= b_i, \quad i = 1, 2, \dots, m \\ x_j &\geq 0, \quad j = 1, 2, \dots, n, \end{aligned}$$

que poderá ser ainda apresentado sob a forma abaixo:

$$(PPL) : \text{maximizar } z = cx \tag{2.4}$$

sujeito a:

$$Ax = b \tag{2.5}$$

$$x \geq 0, \tag{2.6}$$

onde $c = (c_1 \ c_2 \ \dots \ c_n)$, $x^T = (x_1 \ x_2 \ \dots \ x_n)$, $b^T = (b_1 \ b_2 \ \dots \ b_m)$, $A = (a_1 \ a_2 \ \dots \ a_n)$ e $a_j^T = (a_{1j} \ a_{2j} \ \dots \ a_{mj})$, isto é, $c^T \in \mathcal{R}^n$, $x \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $A \in \mathcal{R}^{m \times n}$ e $a_j \in \mathcal{R}^m$.

A desigualdade (2.6) indica que cada componente do vetor x é não negativa. Indicaremos, portanto, que um dado vetor x tem pelo menos uma componente negativa através da notação $x \not\geq 0$.

Definição 2.1 *Seja $X = \{x \in \mathcal{R}^n | Ax = b, x \geq 0\}$. O conjunto X é denominado conjunto ou região viável do (PPL) e se $x \in X$, então x é uma solução viável do mesmo problema. Dado $x^* \in X$, x^* é denominado uma solução ótima do (PPL) se $cx^* \geq cx$, para todo $x \in X$.*

Suporemos, sem perda de generalidade, que a matriz A tenha posto igual a m , isto é, existem m colunas de A linearmente independentes.

Como observação podemos dizer que a presença de uma variável x_j irrestrita em sinal será expressa: $x_j = x_j^+ - x_j^-$, $x_j^+ \geq 0$ e $x_j^- \geq 0$, deixando sempre o problema na forma (PPL) .

Particionaremos a matriz A da seguinte maneira: $A = (B \ N)$, onde B é uma matriz quadrada $m \times m$ e inversível. Analogamente particionaremos os vetores x e c : $x^T = (x_B^T \ x_N^T)$, $c = (c_B \ c_N)$, x_B e c_B possuirão m componentes associadas à matriz B . Dessa maneira o (PPL) poderá ser escrito:

$$(PPL) : \text{maximizar } z = c_B x_B + c_N x_N \quad (2.7)$$

sujeito a:

$$Bx_B + Nx_N = b \quad (2.8)$$

$$x_B \geq 0, \ x_N \geq 0. \quad (2.9)$$

Explicitaremos x_B em função de x_N em (2.8):

$$x_B = B^{-1}b - B^{-1}Nx_N. \quad (2.10)$$

Façamos $x_N = 0$ e $\bar{x}_B = B^{-1}b$.

Definição 2.2 \bar{x} é uma solução básica de (2.5) se $\bar{x}^T = (\bar{x}_B^T \ 0)$. As variáveis associadas às componentes de \bar{x}_B são denominadas básicas e as demais não básicas. Quando \bar{x}_B possuir ao menos uma componente nula diremos que \bar{x} é uma solução básica degenerada.

No caso em que \bar{x}_B for não negativo, isto é, $\bar{x}_B \geq 0$, então \bar{x} satisfará à restrição (2.6). Por força do hábito, diremos que esta solução \bar{x} é uma solução básica primal viável. Primal pois mais adiante introduziremos a noção de dual.

Sejam I_B o conjunto dos índices das colunas de A pertencendo à matriz B e I_N o conjunto dos demais índices de A . Lembremos que $I_B \cap I_N = \emptyset$ e $I_B \cup I_N = \{1, 2, \dots, n\}$.

Levando a expressão de x_B em (2.10) na função objetivo (2.7) teremos uma outra forma do (PPL) :

$$(PPL) : \text{maximizar } z = c_B B^{-1}b - (c_B B^{-1}N - c_N)x_N \quad (2.11)$$

sujeito a:

$$x_B = B^{-1}b - B^{-1}Nx_N \quad (2.12)$$

$$x_B \geq 0, \ x_N \geq 0. \quad (2.13)$$

Por comodidade, definiremos seguindo alguns autores clássicos dos textos de programação linear, por exemplo, Dantzig [Dan 63] e Simonnard [Si 72], novos parâmetros para o último (*PPL*):

$$\begin{aligned} u &= c_B B^{-1}, & u^T &\in \mathcal{R}^m, \\ \bar{x}_B &= B^{-1}b, & \bar{x}_B &\in \mathcal{R}^m, \\ z_j &= ua_j \quad (j \in I_B \cup I_N), & z_j &\in \mathcal{R}, \\ y_j &= B^{-1}a_j \quad (j \in I_B \cup I_N), & y_j &\in \mathcal{R}^m, \\ \bar{z} &= c_B B^{-1}b = ub = c_B \bar{x}_B. \end{aligned}$$

Assim poderemos escrever $(c_B B^{-1}N - c_N)x_N = \sum_{j \in I_N} (z_j - c_j)x_j$ e o (*PPL*) se tornará:

$$(PPL) : \text{ maximizar } z = \bar{z} - \sum_{j \in I_N} (z_j - c_j)x_j \quad (2.14)$$

sujeito a:

$$x_B = \bar{x}_B - \sum_{j \in I_N} y_j x_j \quad (2.15)$$

$$x_B \geq 0, \quad x_j \geq 0, \quad j \in I_N. \quad (2.16)$$

Definindo $y_j^T = (y_{1j} \ y_{2j} \ \dots \ y_{mj})$, $x_B^T = (x_{B(1)} \ x_{B(2)} \ \dots \ x_{B(m)})$ e $\bar{x}_B^T = (\bar{x}_{B(1)} \ \bar{x}_{B(2)} \ \dots \ \bar{x}_{B(m)})$ então (2.15) poderá ainda ser escrito como:

$$x_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_N} y_{ij} x_j, \quad i = 1, \dots, m. \quad (2.17)$$

Proposição 2.1 *Se $\bar{x}_B \geq 0$ e $z_j - c_j \geq 0$, $\forall j \in I_N$ então o vetor $x^* \in \mathcal{R}^n$, onde $x_{B(i)}^* = \bar{x}_{B(i)}$, $i = 1, 2, \dots, m$ e $x_j^* = 0$, $j \in I_N$, será uma solução ótima do (*PPL*).*

Demonstração

Como $z_j - c_j \geq 0$ e $x_j \geq 0$, $\forall j \in I_N$, então de (2.14) temos $z \leq \bar{z} = cx^*$. O máximo de z não ultrapassará $\bar{z} = cx^*$, mas x^* é uma solução viável do (*PPL*), logo x^* é uma solução ótima do (*PPL*). ■

No caso da propriedade 2.1, x^* é uma solução básica de (2.5).

Suponhamos agora que $\hat{x} \in \mathcal{R}^n$ seja uma solução viável de (2.5) e (2.6), logo o será também de (2.15) e (2.16), isto é,

$$\hat{x}_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_N} y_{ij} \hat{x}_j, \quad i = 1, 2, \dots, m \quad (2.18)$$

e $\hat{x} \geq 0$, $j \in I_B \cup I_N$, fornecendo um valor \hat{z} à função objetivo:

$$\hat{z} = \bar{z} - \sum_{j \in I_N} (z_j - c_j) \hat{x}_j = c\hat{x}.$$

Suporemos também que $\hat{x} \in \mathcal{R}^n$ não seja uma solução básica de (2.5), isto quer dizer que haverá ao menos uma componente $\hat{x}_j > 0$, $j \in I_N$.

Será possível passar da solução \hat{x} a uma solução básica viável x^* do (PPL) tal que $cx^* \geq \hat{z} = c\hat{x}$?

Para respondermos esta pergunta começaremos fazendo variar o valor de uma variável x_k , $k \in I_N$ enquanto que o valor das outras variáveis cujos índices pertencem a I_N não se modificam, isto é, $x_j = \hat{x}_j$ para $j \in I_N - \{k\}$. De (2.18):

$$x_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j - y_{ik} x_k, \quad i = 1, 2, \dots, m \quad (2.19)$$

onde x_k poderá variar (aumentar ou diminuir).

Sabemos que $x_k \geq 0$, $x_{B(i)} \geq 0$, $i = 1, 2, \dots, m$, e que os outros valores associados a x_j , $j \in I_N - \{k\}$, não serão modificados. Assim sendo: $x_{B(i)} \geq 0$ implica que

$$\bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j - y_{ik} x_k \geq 0, \quad i = 1, 2, \dots, m. \quad (2.20)$$

Consideremos L_0, L_1, L_2 uma partição de $\{1, 2, \dots, m\}$, tal que

$$L_0 = \{i \mid y_{ik} = 0\}, \quad L_1 = \{i \mid y_{ik} > 0\}, \quad L_2 = \{i \mid y_{ik} < 0\}.$$

Busquemos os limites de variação para x_k pois sabemos que de (2.20):

$$y_{ik} x_k \leq \bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j, \quad i = 1, 2, \dots, m. \quad (2.21)$$

Para $i \in L_0$ basta que o valor de x_k seja não-negativo.

Para $i \in L_1$:

$$x_k \leq \frac{1}{y_{ik}} \left(\bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j \right).$$

Para $i \in L_2$:

$$x_k \geq \frac{1}{y_{ik}} \left(\bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j \right).$$

Sejam

$$\alpha_k = \frac{1}{y_{sk}} \left(\bar{x}_{B(s)} - \sum_{j \in I_N - \{k\}} y_{sj} \hat{x}_j \right) = \min_{i \in L_1} \left\{ \frac{1}{y_{ik}} \left(\bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j \right) \right\},$$

$$\beta_k = \frac{1}{y_{lk}} \left(\bar{x}_{B(l)} - \sum_{j \in I_N - \{k\}} y_{lj} \hat{x}_j \right) = \max_{i \in L_2} \left\{ \frac{1}{y_{ik}} \left(\bar{x}_{B(i)} - \sum_{j \in I_N - \{k\}} y_{ij} \hat{x}_j \right) \right\},$$

e $\gamma_k = \max\{0, \beta_k\}$.

Logo $\gamma_k \leq x_k \leq \alpha_k$.

Quando $L_1 = \phi \Rightarrow \alpha_k = \infty$ e quando $L_2 = \phi \Rightarrow \beta_k = -\infty$.

A matriz $B = (a_{B(1)} \ a_{B(2)} \ \dots \ a_{B(m)})$ extraída de A foi utilizada para chegarmos ao sistema (2.15) a partir de (2.5). Os vetores $a_{B(i)}$, $i = 1, 2, \dots, m$, formam uma base do \mathcal{R}^m , logo existem $\lambda_i \in \mathcal{R}$, $i = 1, 2, \dots, m$, para os quais $a_k = \sum_{i=1}^m \lambda_i a_{B(i)}$. Seja $s \in \{1, 2, \dots, m\} = M$ tal que $\lambda_s \neq 0$, então $a_{B(s)} = \frac{1}{\lambda_s} \left(a_k - \sum_{i \in M - \{s\}} \lambda_i a_{B(i)} \right)$, como $a_{B(s)} \neq 0$, a coluna a_k não pode ser escrita como uma combinação linear das colunas $a_{B(i)}$, $i \in M - \{s\}$; isto quer dizer que $a_{B(1)}, a_{B(2)}, \dots, a_{B(s-1)}, a_k, a_{B(s+1)}, \dots, a_{B(m)}$ formam também uma base do \mathcal{R}^m .

Seja $v^T = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_m)$, assim podemos escrever $a_k = Bv$, logo $v = B^{-1}a_k$, isto é, $v = y_k$.

Basta que $y_{sk} \neq 0$ para que possamos substituir a base formada pelas colunas de B por uma outra base em que o vetor $a_{B(s)}$ é substituído por a_k .

Já estamos aptos a responder nossa pergunta.

Procedimento 1

Tomemos x_k tal que $x_k = \hat{x}_k > 0$ e $k \in I_N$.

1º caso: $z_k - c_k > 0$, decresceremos o valor de x_k até alcançar γ_k ;

se $\gamma_k = 0$, faremos $x_k = 0$ e utilizaremos (2.19) para atualizar os valores de $x_{B(i)}$, $i = 1, 2, \dots, m$;

se $\gamma_k = \beta_k$, faremos $x_k = \beta_k$ que ocasionará $x_{B(l)} = 0$ em (2.19), como $y_{lk} \neq 0$ então poderemos fazer

$$I_B := (I_B - \{B(l)\}) \cup \{k\},$$

$$I_N := (I_N - \{k\}) \cup \{B(l)\},$$

isto é, teremos uma nova matriz B inversível, extraída de A , onde a coluna $a_{B(l)}$ será substituída por a_k ;

2º caso: $z_k - c_k < 0$, aumentaremos o valor de x_k até alcançar α_k ;

se $\alpha_k = +\infty$, a solução do (PPL) será ilimitada, pois $x_k \rightarrow +\infty$ implica $z \rightarrow +\infty$;

se $\alpha_k < \infty$, faremos $x_k = \alpha_k$ que ocasionará $x_{B(s)} = 0$ em (2.19), como $y_{sk} \neq 0$ então poderemos fazer

$$I_B := (I_B - \{B(s)\}) \cup \{k\},$$

$$I_N := (I_N - \{k\}) \cup \{B(s)\},$$

isto é, teremos uma nova matriz B inversível, extraída de A , onde a coluna $a_{B(s)}$ será substituída por a_k ;

3º caso: $z_k - c_k = 0$, aplicaremos o que foi realizado no 1º caso.

Fim do procedimento 1

Para cada $j \in I_N$ tal que $\hat{x}_j > 0$, o procedimento 1 feito para o índice k será repetido até que os valores atribuídos às variáveis x_j , $j \in I_N$, sejam nulos, ou que a solução máxima (ótima) do (PPL) seja ilimitada (2º caso, $\alpha_k = +\infty$).

O procedimento 1 será aplicado r vezes onde $r = |\{j \in I_N \mid \hat{x}_j > 0\}|$.

Com as explicações anteriores poderemos enunciar mais duas propriedades a seguir.

Proposição 2.2 *Se (2.5), (2.6) admitirem uma solução viável, então haverá ao menos uma solução básica de (2.5) satisfazendo (2.6).*

Proposição 2.3 *Se o (PPL) possuir ótimo finito ao menos uma solução ótima será básica viável.*

A propriedade 2.2 poderá ser demonstrada utilizando o 1º caso do procedimento 1 para qualquer $z_k - c_k \in \mathcal{R}$, $k \in I_N$, $\hat{x}_k > 0$.

Aplicaremos r vezes, onde $r = |\{j \in I_N \mid \hat{x}_j > 0\}|$, o procedimento 1 para demonstrarmos a propriedade 3.

Exemplo 2.1 Seja o

$$(PPL) : \text{maximizar } z = 3x_1 + 5x_2$$

sujeito a:

$$\begin{array}{rcl} x_1 & \leq & 4 \\ & x_2 & \leq 6 \\ 3x_1 + 2x_2 & \leq & 18 \\ x_1 & \geq & 0 \\ & x_2 & \geq 0 \end{array}$$

Associaremos às restrições não triviais as variáveis de folga $x_3 \geq 0$, $x_4 \geq 0$, $x_5 \geq 0$ tais que o (PPL) fique sob a seguinte forma.

$$(PPL) : \text{maximizar } z = 3x_1 + 5x_2 + 0x_3 + 0x_4 + 0x_5$$

sujeito a:

$$\begin{array}{rcl} x_1 & + & x_3 = 4 \\ & x_2 & + x_4 = 6 \\ 3x_1 + 2x_2 & & + x_5 = 18 \\ x_j & \geq & 0, \quad j = 1, 2, 3, 4, 5, \end{array}$$

onde

$$A = (a_1 \ a_2 \ a_3 \ a_4 \ a_5) = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix},$$

$$b = \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix}, \ c = (3 \ 5 \ 0 \ 0 \ 0).$$

Tomemos

$$I_B = \{3, \ 2, \ 5\}, \ I_N = \{1, \ 4\},$$

$$B(1) = 3, \ B(2) = 2, \ B(3) = 5,$$

$$B = (a_3 \ a_2 \ a_5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \ \text{logo } B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix},$$

$$c_B = (0 \ 5 \ 0), \ u = c_B B^{-1} = (0 \ 5 \ 0) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} = (0 \ 5 \ 0),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_2 \\ \bar{x}_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 6 \end{pmatrix},$$

$$\bar{z} = c_B B^{-1}b = ub = (0 \ 5 \ 0) \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = 30,$$

$$z_1 = ua_1 = (0 \ 5 \ 0) \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = 0 \Rightarrow z_1 - c_1 = 0 - 3 = -3,$$

$$z_4 = ua_4 = (0 \ 5 \ 0) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 5 \Rightarrow z_4 - c_4 = 5 - 0 = 5,$$

$$y_1 = B^{-1}a_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix},$$

$$y_4 = B^{-1}a_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}.$$

Poderemos escrever:

maximizar z

sujeito a:

$$\begin{aligned} z &= 30 + 3x_1 - 5x_4 \\ x_3 &= 4 - x_1 \\ x_2 &= 6 - x_4 \\ x_5 &= 6 - 3x_1 + 2x_4 \\ x_j &\geq 0, \quad j = 1, 2, 3, 4, 5. \end{aligned}$$

Tomemos $\hat{x}_1 = 1$, $\hat{x}_2 = 4$, $\hat{x}_3 = 3$, $\hat{x}_4 = 2$, $\hat{x}_5 = 7$ uma solução viável deste problema, facilmente verificada, fornecendo $\hat{z} = 30 + 3 \times 1 - 5 \times 2 = 23$. A partir da solução \hat{x} objetivamos encontrar uma solução básica viável que forneça um valor z^* para z , tal que $z^* \geq c\hat{x}$.

A seguir usaremos o procedimento 1.

Como $z_4 - c_4 = 5 > 0$, estamos no 1º caso, portanto faremos x_4 decrescer de valor:

$$4 - \hat{x}_1 \geq 0 \quad \text{qualquer que seja } x_4,$$

$$6 - x_4 \geq 0 \Rightarrow x_4 \leq 6,$$

$$6 - 3\hat{x}_1 + 2x_4 \geq 0 \Rightarrow 2x_4 \geq 3\hat{x}_1 - 6 \Rightarrow x_4 \geq \frac{3 \times 1 - 6}{2} = -\frac{3}{2}, \quad \beta_4 = -\frac{3}{2},$$

$$\text{logo } \gamma_4 = \max \left\{ 0, -\frac{3}{2} \right\} = 0.$$

Basta fazermos $x_4 = 0$ e teremos a nova solução viável: $\hat{x}_1 = 1$, $\hat{x}_4 = 0$ fornecendo $\hat{x}_3 = 3$, $\hat{x}_2 = 6$, $\hat{x}_5 = 3$ e $\hat{z} = 30 + 3 \times 1 = 33$.

Examinaremos agora x_1 . Como $z_1 - c_1 = -3$ estamos no 2º caso, portanto faremos x_1 aumentar de valor:

$$x_3 = 4 - x_1 \geq 0 \Rightarrow x_1 \leq 4,$$

$$x_2 = 6 - \hat{x}_4 \geq 0 \quad \text{qualquer que seja } x_1,$$

$$x_5 = 6 - 3x_1 + 2\hat{x}_4 \geq 0 \Rightarrow 3x_1 \leq 6 + 2\hat{x}_4 \Rightarrow x_1 \leq \frac{6 + 2 \times 0}{3} = 2,$$

$$\text{logo } \alpha_1 = \min \{2, 4\} = 2, \quad s = 3, \quad B(s) = 5.$$

A nova base será definida por

$$I_B = \{3, 2, 1\}, \quad I_N = \{4, 5\},$$

$$B(1) = 3, \quad B(2) = 2, \quad B(3) = 1,$$

$$B = (a_3 \ a_2 \ a_1) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix}, \quad \text{logo } B^{-1} = \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

$$c_B = (0 \ 5 \ 3), \ u = c_B B^{-1} = (0 \ 5 \ 3) \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} = (0 \ 3 \ 1),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix},$$

$$\bar{z} = c_B B^{-1}b = ub = (0 \ 3 \ 1) \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = 36,$$

$$z_4 = ua_4 = (0 \ 3 \ 1) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 3 \Rightarrow z_4 - c_4 = 3 - 0 = 3,$$

$$z_5 = ua_5 = (0 \ 3 \ 1) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1 \Rightarrow z_5 - c_5 = 1 - 0 = 1,$$

$$y_4 = B^{-1}a_4 = \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ 1 \\ -\frac{2}{3} \end{pmatrix},$$

$$y_5 = B^{-1}a_5 = \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ 1 \\ \frac{1}{3} \end{pmatrix}.$$

Novamente escreveremos o (*PPL*) sob a seguinte forma:

maximizar z

sujeito a:

$$\begin{aligned} z &= 36 - 3x_4 - x_5 \\ x_3 &= 2 - \frac{2}{3}x_4 + \frac{1}{3}x_5 \\ x_2 &= 6 - x_4 - x_5 \\ x_1 &= 2 + \frac{2}{3}x_4 - \frac{1}{3}x_5 \\ x_j &\geq 0, \ j = 1, 2, 3, 4, 5. \end{aligned}$$

A solução obtida será $\hat{x}_1 = 2$, $\hat{x}_2 = 6$, $\hat{x}_3 = 2$, $\hat{x}_4 = 0$, $\hat{x}_5 = 0$, que é uma solução básica primal viável. Neste caso obtivemos $z_j - c_j \geq 0$, $\forall j \in I_N$, assim sendo, pela propriedade 2.1, esta última solução é também ótima, fornecendo $z^* = 36$.

Ilustraremos, na figura 2.1, o desenvolvimento deste exemplo no espaço $x_1 \times x_2$ do problema original. Os lados do pentágono formado pelos vértices

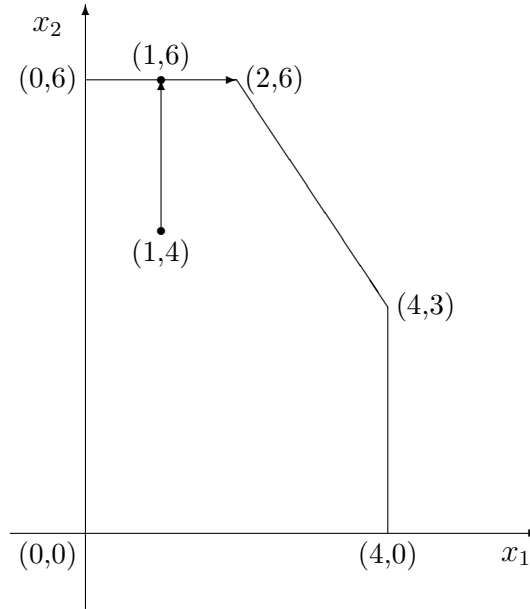


Figura 2.1: Busca de uma solução básica

$(0,0)$, $(4,0)$, $(4,3)$, $(2,6)$ e $(0,6)$ e seu interior representam a região viável do problema. Do ponto inicial $(1,4)$ passamos a $(1,6)$ e deste à solução básica $(2,6)$. ■

2.1 Observações

Pela propriedade 2.3 podemos pensar em só considerar, para efeito de encontrar um ótimo do (PPL) , as soluções básicas viáveis deste problema.

O sistema de equações lineares (2.5) pode possuir até \mathcal{C}_n^m soluções básicas, pois \mathcal{C}_n^m é o número total de matrizes $m \times m$ diferentes extraídas de A , algumas podendo não ser inversíveis e entre as inversíveis poderemos ter $\bar{x}_B \not\geq 0$. Supondo mais uma vez que o problema de programação linear tenha solução finita, poderemos escrever o seguinte problema de otimização:

$$(P) : \begin{array}{ll} \text{maximizar} & z = cx \\ \text{sujeito a:} & x \in V, \end{array}$$

onde $V = \{x^1, x^2, \dots, x^t\}$, e $x^i \in \mathcal{R}^n$, $i = 1, 2, \dots, t$, são soluções básicas viáveis do (PPL) . Lembremos que $|V| = t \leq \mathcal{C}_n^m$.

Definição 2.3 (P) é um problema de otimização combinatória.

2.2 Exercícios

1. Sendo dado o problema de programação linear

$$(P) : \text{maximizar } z = 3x_1 + 5x_2$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & & & + & x_3 & & & = & 4 \\ & & x_2 & & & + & x_4 & = & 6 \\ 3x_1 & + & 2x_2 & & & & & + & x_5 = 18 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

Verificar que as colunas associadas às variáveis x_1 , x_2 , x_3 formam um base ótima do (P) .

2. Seja

$$(P1) : \text{maximizar } z = 10x_1 + 12x_2 + 3x_3 + 4x_4 + 2x_5$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & + & x_2 & + & x_3 & & & = & 4 \\ 2x_1 & + & 3x_2 & & & - & x_4 & = & 6 \\ x_1 & & & & & & & + & x_5 = 3 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

A partir da solução $x_1 = 1$, $x_2 = 2$, $x_3 = 1$, $x_4 = 2$, $x_5 = 2$, fornecendo um valor de $z = 49$. Encontrar uma solução básica de $(P1)$ tal que o valor de z associado à esta solução básica seja maior ou igual a 49.

Capítulo 3

Método do Simplex

3.1 Introdução

Dantzig, em 1947, introduziu o método do simplex para resolver um problema de programação linear (*PPL*).

A idéia do método é partir de uma solução básica de (2.5) satisfazendo (2.6), isto é, uma solução básica primal viável, passar para outra solução básica primal viável sem que o valor da função objetivo diminua (no caso de maximização). Como o número de soluções básicas é finito, o algoritmo, sob algumas condições, convergirá.

Dada a matriz B quadrada e inversível, extraída de A , tal que $\bar{x}_B \geq 0$, colocaremos o problema de programação linear (2.4), (2.5) e (2.6) sob a forma (2.14), (2.15) e (2.16). Utilizando a propriedade 2.1, testaremos se esta solução é ótima, caso não o seja tentaremos aumentar o valor de uma variável x_k , $k \in I_N$, tal que $z_k - c_k < 0$, como já explicado no 2º caso do procedimento 1. Se $\alpha_k = +\infty$ então não haverá ótimo finito, caso contrário procederemos exatamente como foi ilustrado no exemplo 2.1.

No caso em que só iremos trabalhar com soluções básicas viáveis, o cálculo de α_k é mais simplificado:

$$\alpha_k = \frac{\bar{x}_{B(s)}}{y_{sk}} = \min_{i \in L_1} \left\{ \frac{\bar{x}_{B(i)}}{y_{ik}} \right\}$$

e caso $L_1 = \emptyset$ faremos $\alpha_k = +\infty$.

A seguir descreveremos um procedimento que resume o método do simplex.

3.2 Algoritmo do Simplex

Procedimento 2 (maximização)

Dada uma solução básica primal viável para o (PPL).

Se $z_j - c_j \geq 0, \forall j \in I_N$, a solução dada é uma solução ótima. PARE.

Caso contrário, escolhe-se um $k \in I_N$ para o qual $z_k - c_k < 0$;

se $\alpha_k = +\infty$, a solução do (PPL) é ilimitada. PARE.

se $\alpha_k < +\infty$, faremos $x_k = \alpha_k$, acarretando $x_{B(s)} = 0$,
a coluna a_k ocupará o lugar da coluna $a_{B(s)}$ em B .
MUDANÇA DE BASE.

Fim do procedimento 2.

Para cada nova base o procedimento 2 é repetido até que uma regra de parada seja verificada. Este procedimento possui duas saídas: solução ótima encontrada ou solução ilimitada. Para que este método iterativo realmente termine, teremos que considerar as condições de Bland [Bl 77] que serão examinadas mais tarde.

Exemplo 3.1 Tomemos o (PPL) do exemplo 2.1, já com as variáveis de folga adicionadas.

1ª solução básica:

$$I_B = \{3, 4, 5\}, I_N = \{1, 2\},$$

$$B(1) = 3, B(2) = 4, B(3) = 5,$$

$$B = (a_3 \ a_4 \ a_5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

A base B está associada a uma solução básica primal viável, isto é, faremos $x_1 = x_2 = 0$ e teremos $x_3 = 4$, $x_4 = 6$ e $x_5 = 18$. Neste caso $B^{-1} = B = I$. E ainda

$$c_B = (0 \ 0 \ 0), \ u = c_B B^{-1} = (0 \ 0 \ 0)I = (0 \ 0 \ 0),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \\ \bar{x}_{B(3)} \end{pmatrix} = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_4 \\ \bar{x}_5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix},$$

$$\bar{z} = c_B B^{-1}b = ub = (0 \ 0 \ 0)(4 \ 6 \ 18)^T = 0,$$

$$z_1 = ua_1 = (0 \ 0 \ 0) \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = 0 \Rightarrow z_1 - c_1 = 0 - 3 = -3 < 0,$$

$$z_2 = ua_2 = (0 \ 0 \ 0) \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = 0 \Rightarrow z_2 - c_2 = 0 - 5 = -5 < 0,$$

$$y_1 = B^{-1}a_1 = Ia_1 = a_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix},$$

$$y_2 = B^{-1}a_2 = Ia_2 = a_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

O problema ficará sob a seguinte forma:

maximizar z

sujeito a:

$$\begin{aligned} z &= 0 + 3x_1 + 5x_2 \\ x_3 &= 4 - x_1 \\ x_4 &= 6 - x_2 \\ x_5 &= 18 - 3x_1 - 2x_2 \\ x_j &\geq 0, \quad j = 1, 2, 3, 4, 5. \end{aligned}$$

Fazendo $x_1 = x_2 = 0$ teremos $x_3 = 4$, $x_4 = 6$, $x_5 = 18$ fornecendo $z = 0$. Faremos uma das variáveis x_1 ou x_2 crescer de valor, provocando o aumento de z . Tomemos, por exemplo, x_2 para ter seu valor aumentado, isto é, faremos a coluna a_2 entrar na nova base. Como $L_1 = \{2, 3\}$, pois $y_{12} = 0$, $y_{22} = 1$ e $y_{32} = 2$, passaremos a calcular α_2 :

$$\alpha_2 = \min \left\{ \frac{\bar{x}_{B(2)}}{y_{22}}, \frac{\bar{x}_{B(3)}}{y_{32}} \right\} = \min \left\{ \frac{6}{1}, \frac{18}{2} \right\} = 6 = \frac{\bar{x}_{B(2)}}{y_{22}},$$

logo $a_{B(2)}$ deixará a base, sendo substituída pela coluna a_2 .

2ª solução básica:

$$I_B = \{3, 2, 5\}, \quad I_N = \{1, 4\},$$

$$B(1) = 3, \quad B(2) = 2, \quad B(3) = 5,$$

$$B = (a_3 \ a_2 \ a_5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \text{ logo } B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix},$$

$$c_B = (0 \ 5 \ 0), \quad u = c_B B^{-1} = (0 \ 5 \ 0) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} = (0 \ 5 \ 0),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_2 \\ \bar{x}_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 6 \end{pmatrix},$$

$$z_1 = ua_1 = (0 \ 5 \ 0) \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = 0 \Rightarrow z_1 - c_1 = 0 - 3 = -3 < 0,$$

$$z_4 = ua_4 = (0 \ 5 \ 0) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 5 \Rightarrow z_4 - c_4 = 5 - 0 = 5 > 0.$$

Calcularemos o essencial para a passagem à terceira solução básica, isto é, a_1 entrará na nova base; necessitamos obter α_1 para saber qual a coluna de B que será substituída por a_1 .

$$y_1 = B^{-1}a_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} y_{11} \\ y_{21} \\ y_{31} \end{pmatrix},$$

logo $L_1 = \{1, 3\}$ e

$$\alpha_1 = \min_{i \in L_1} \left\{ \frac{\bar{x}_{B(i)}}{y_{i1}} \right\} = \min \left\{ \frac{4}{1}, \frac{6}{3} \right\} = \frac{6}{3} = 2 = \frac{\bar{x}_{B(3)}}{y_{31}} = \frac{\bar{x}_5}{y_{31}}.$$

Assim sendo $a_5 = a_{B(3)}$ deixará a base.

3ª solução básica:

$$I_B = \{3, 2, 1\}, \quad I_N = \{4, 5\},$$

$$B(1) = 3, \quad B(2) = 2, \quad B(3) = 1,$$

$$B = (a_3 \ a_2 \ a_1) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix} \quad \text{logo} \quad B^{-1} = \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix},$$

$$c_B = (0 \ 5 \ 3), \quad u = c_B B^{-1} = (0 \ 5 \ 3) \begin{pmatrix} 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} = (0 \ 3 \ 1),$$

$$\bar{x} = B^{-1}b = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix},$$

$$\bar{z} = ub = (0 \ 3 \ 1) \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = 36,$$

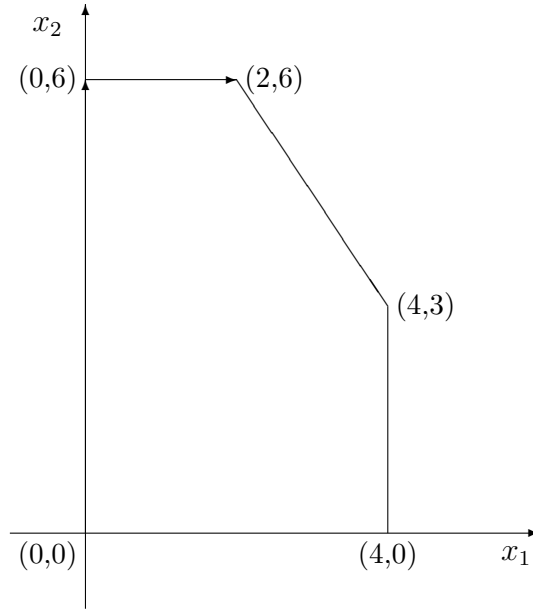


Figura 3.1: Iterações do método primal do simplex

$$z_4 = ua_4 = (0 \ 3 \ 1) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 3 \Rightarrow z_4 - c_4 = 3 - 0 = 3 > 0,$$

$$z_5 = ua_5 = (0 \ 3 \ 1) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1 \Rightarrow z_5 - c_5 = 1 - 0 = 1 > 0.$$

Como $z_j - c_j \geq 0$, $\forall j \in I_N$, esta solução básica (3ª solução) é ótima (propriedade 1).

Então $x_1 = 2$, $x_2 = 6$, $x_3 = 2$, $x_4 = x_5 = 0$ é uma solução ótima, fornecendo $z = 36$.

Ilustraremos, na figura 3.1, o desenvolvimento deste exemplo no espaço $x_1 \times x_2$ do problema original. Os lados do pentágono formado pelos vértices $(0,0)$, $(4,0)$, $(4,3)$, $(2,6)$ e $(0,6)$ e seu interior representam a região viável do problema. Do ponto inicial $(0,0)$ passamos a $(0,6)$ e deste à solução ótima $(2,6)$ ■.

3.3 Busca de uma Solução Básica Viável

Quando não tivermos uma solução básica viável para o (PPL) , poderemos proceder da seguinte maneira.

Acrescentaremos uma variável artificial $g_i \geq 0$ à esquerda de cada restrição $\sum_{j=1}^n a_{ij}x_j = b_i$, $i = 1, 2, \dots, m$. Suporemos $b_i \geq 0$, $i = 1, 2, \dots, m$. Teríamos o seguinte conjunto de restrições:

$$\sum_{j=1}^n a_{ij}x_j + g_i = b_i, \quad i = 1, 2, \dots, m \quad (3.1)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n, \quad (3.2)$$

$$g_i \geq 0, \quad i = 1, 2, \dots, m. \quad (3.3)$$

Construiremos um outro problema de programação linear

$$(PA) : \text{minimizar } \sum_{i=1}^m g_i$$

sujeito a

$$\sum_{j=1}^n a_{ij}x_j + g_i = b_i, \quad i = 1, 2, \dots, m$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n,$$

$$g_i \geq 0, \quad i = 1, 2, \dots, m.$$

Utilizaremos $v(\cdot)$ para representar o valor ótimo da função objetivo do problema de programação linear (\cdot) .

É fácil verificar que as variáveis $x_j = 0$, $j = 1, 2, \dots, n$ e $g_i = b_i \geq 0$ estão associadas a uma solução básica de (3.1) satisfazendo (3.2) e (3.3). Esta solução básica será tomada como solução inicial para a solução de (PA) utilizando o método do simplex (procedimento 2) para o caso de minimização, lembrando que $\min z = -\max(-z)$.

Se $v(PA) > 0$ o conjunto de restrições do (PPL) é vazio.

Se $v(PA) = 0$ a solução ótima obtida para (PA) terá $g_i = 0$, $i = 1, 2, \dots, m$ e $x_j = \bar{x}_j \geq 0$ para $j = 1, 2, \dots, n$ satisfazendo a $\sum_{j=1}^n a_{ij}\bar{x}_j = b_i$, $i = 1, 2, \dots, m$. Se a base final da solução ótima de (PA) não contiver nenhuma coluna associada às variáveis artificiais g_i , $i = 1, 2, \dots, m$ esta será também uma base primal viável do (PPL) original. Caso a solução básica ótima encontrada para o (PA) seja degenerada, isto é, há pelo menos uma coluna associada a g_i na base ótima e $g_i = 0$, poderemos iniciar o método do simplex para o (PPL) com esta base, não permitindo que as variáveis g_i associadas à

base tenham valores diferentes de zero. Mais tarde trataremos com detalhes este assunto.

A solução do (PA) é conhecida como sendo a primeira fase do método do simplex.

3.4 Cálculo de B^{-1}

Lembremos que em cada etapa do método do simplex necessitamos determinar:

$$u = c_B B^{-1}, \quad \bar{x}_B = B^{-1}b \text{ e } y_k = B^{-1}a_k,$$

onde a_k é a coluna escolhida para entrar na base.

Na realidade temos três sistemas de equações lineares simultâneas para serem resolvidos:

$$uB = c_B, \quad B\bar{x}_B = b \text{ e } By_k = a_k,$$

que determinam respectivamente os vetores u , \bar{x}_B e y_k .

Caso só tivéssemos de resolver uma vez esses três sistemas, poderíamos utilizar os métodos numéricos de resolução de sistemas de equações lineares simultâneas sem a inversão explícita de B , no entanto, em cada iteração do simplex buscamos a solução dos três sistemas para os quais B , c_B e a_k variam de iteração em iteração.

Geralmente, como descrito no exemplo 3.1, a primeira base B associada a uma solução básica viável, no método do simplex, é a matriz unitária I . Por outro lado, já verificamos que de uma iteração à seguinte a matriz B se transforma em outra matriz B' trocando-se somente uma coluna de B .

Seja $B_r = (a_1 \ a_2 \ \dots \ a_{r-1} \ a_r \ a_{r+1} \ \dots \ a_m)$ uma matriz quadrada inversível, $m \times m$. Suponhamos conhecida B_r^{-1} .

Dada $B_p = (a_1 \ a_2 \ \dots \ a_{r-1} \ a_p \ a_{r+1} \ \dots \ a_m)$ como se poderá determinar B_p^{-1} , caso exista, utilizando B_r^{-1} ?

Consideremos

$$B_r^{-1}B_p = (B_r^{-1}a_1 \ B_r^{-1}a_2 \ \dots \ B_r^{-1}a_{r-1} \ B_r^{-1}a_p \ B_r^{-1}a_{r+1} \ \dots \ B_r^{-1}a_m),$$

sabemos que $B_r^{-1}a_j = e_j$, $j \neq p$, onde e_j é um vetor com todas as componentes nulas exceto a j -ésima componente que é igual a um.

Seja $v = B_r^{-1}a_p = (v_1 \ v_2 \ \dots \ v_{r-1} \ v_r \ v_{r+1} \ \dots \ v_m)^T$, então

$$B_r^{-1}B_p = (e_1 \ e_2 \ \dots \ e_{r-1} \ v \ e_{r+1} \ \dots \ e_m) = E_r, \quad (3.4)$$

ou ainda

$$E_r = \begin{pmatrix} 1 & 0 & 0 & 0 & v_1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & v_2 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & v_{r-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & v_r & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & v_{r+1} & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & v_{m-1} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & v_m & 0 & 0 & 0 & 1 \end{pmatrix}.$$

De (3.4) teremos que $B_p = B_r E_r$. Verificamos que existirá B_p^{-1} se e somente se existir E_r^{-1} . Notemos que o determinante de E_r é igual a v_r , basta que $v_r \neq 0$ para termos E_r^{-1} , assim sendo:

$$B_p^{-1} = (B_r E_r)^{-1} = E_r^{-1} B_r^{-1}. \quad (3.5)$$

Proposição 3.1 *Se $v_r \neq 0$, então a inversa de E_r será*

$$E_r^{-1} = (e_1 \ e_2 \ \dots \ e_{r-1} \ \bar{v} \ e_{r+1} \ \dots \ e_m),$$

onde

$$\bar{v} = \left(-\frac{v_1}{v_r} \ -\frac{v_2}{v_r} \ \dots \ -\frac{v_{r-1}}{v_r} \ \frac{1}{v_r} \ -\frac{v_{r+1}}{v_r} \ \dots \ -\frac{v_{m-1}}{v_r} \ -\frac{v_m}{v_r} \right)^T.$$

Ou ainda

$$E_r^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & -\frac{v_1}{v_r} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -\frac{v_2}{v_r} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -\frac{v_{r-1}}{v_r} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{v_r} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{v_{r+1}}{v_r} & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -\frac{v_{m-1}}{v_r} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -\frac{v_m}{v_r} & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Demonstração

Basta realizarmos $E_r E_r^{-1}$ e encontraremos I . ■

Suponhamos agora que $B = (a_1 \ a_2 \ \dots \ a_m)$ é uma matriz quadrada $m \times m$ e desejamos encontrar sua inversa, caso exista.

Partiremos da matriz unitária $I = (e_1 \ e_2 \ \dots \ e_m)$, cuja inversa é a própria matriz I . Iniciaremos calculando a inversa de $B_1 = (a_1 \ e_2 \ \dots \ e_m)$ a partir de I tal como foi feito neste capítulo, isto é, $E_1 = B_1$, então $B_1^{-1} = E_1^{-1}$. A partir de B_1 da qual já conhecemos a inversa, passaremos a calcular a inversa de $B_2 = (a_1 \ a_2 \ e_3 \ \dots \ e_m)$. Sabemos que $B_1^{-1}B_2 = (e_1 \ B_1^{-1}a_2 \ e_3 \ \dots \ e_m) = E_2$, a inversa de E_2 é obtida pelo método já visto. Assim sendo $B_2 = B_1E_2 \Rightarrow B_2^{-1} = E_2^{-1}B_1^{-1} = E_2^{-1}E_1^{-1}$. Seja $B_3 = (a_1 \ a_2 \ a_3 \ e_4 \ \dots \ e_m)$ cuja inversa desejamos obter, logo poderemos escrever $B_2^{-1}B_3 = (e_1 \ e_2 \ B_2^{-1}a_3 \ e_4 \ \dots \ e_m) = E_3 \Rightarrow B_3^{-1} = E_3^{-1}B_2^{-1} = E_3^{-1}E_2^{-1}E_1^{-1}$. Continuando o raciocínio obteremos $B^{-1} = E_m^{-1}E_{m-1}^{-1}\dots E_3^{-1}E_2^{-1}E_1^{-1}$. Como saberemos que B não tem inversa?

Exemplo 3.2 Seja

$$E = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}, \det(E) = 3, \quad E^{-1} = \begin{pmatrix} 1 & -\frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{3} & 0 & 1 \end{pmatrix}. \quad \blacksquare$$

Observação: verificamos que $v = y_k$, onde $y_k = B^{-1}a_k$, no caso em que a_k substituirá a r -ésima coluna de B para formar a matriz B' . Logo, se $y_{rk} \neq 0$, então B' será inversível.

3.5 Interpretação Geométrica da Mudança de Base

Um conjunto $C \subseteq \mathcal{R}^n$ é convexo se para quaisquer x^1 e $x^2 \in C$ temos que $x = \lambda x^1 + (1 - \lambda)x^2$ para $0 \leq \lambda \leq 1$ pertencerá também a C .

O conjunto $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\}$ é, por definição um conjunto poliédrico quando $X \neq \emptyset$.

Proposição 3.2 *O conjunto X é convexo.*

Demonstração

Sejam x^1 e $x^2 \in X$ então $x^1 \geq 0$, $Ax^1 = b$ e $x^2 \geq 0$, $Ax^2 = b$. Para $0 \leq \lambda \leq 1$ podemos escrever $\lambda x^1 \geq 0$, $\lambda Ax^1 = \lambda b$ e $(1 - \lambda)x^2 \geq 0$, $(1 - \lambda)Ax^2 = (1 - \lambda)b$ ou $\lambda x^1 + (1 - \lambda)x^2 \geq 0$, $\lambda Ax^1 + (1 - \lambda)Ax^2 = \lambda b + (1 - \lambda)b$, assim temos $\bar{x} = \lambda x^1 + (1 - \lambda)x^2 \geq 0$, $A[\lambda x^1 + (1 - \lambda)x^2] = A\bar{x} = b$, logo $\bar{x} \in X$. \blacksquare

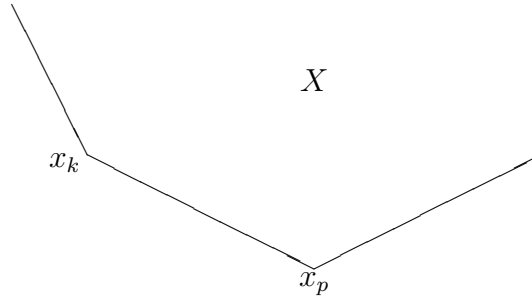


Figura 3.2: Vértices ou pontos extremos

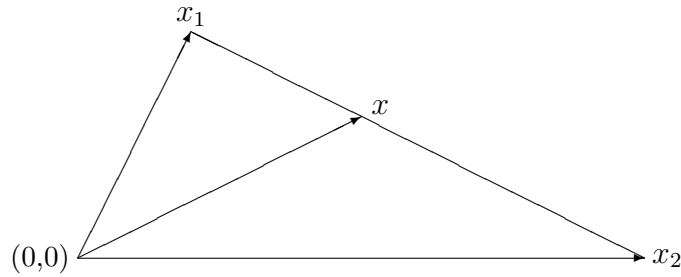


Figura 3.3: Combinação convexa

Definição 3.1 x é um vértice (ponto extremo) de um conjunto poliédrico X se $x \in X$ e não existir $\lambda \in [0, 1]$ tal que $x = \lambda x^1 + (1 - \lambda)x^2$, onde x^1 e $x^2 \in X$, para $x \neq x^1$ e $x \neq x^2$.

Na figura 3.2 ilustramos dois vértices, x^k e x^p de um conjunto poliédrico em \mathcal{R}^2 .

A expressão $x = \lambda x^1 + (1 - \lambda)x^2$, para $\lambda \in [0, 1]$, representa todos os pontos do segmento de reta unindo x^1 e x^2 . Dizemos que x é uma combinação convexa de x^1 e x^2 , ver figura 3.3.

Proposição 3.3 Seja $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\} \neq \emptyset$ e seja A uma matriz $m \times n$, de posto igual a m . Uma solução básica de $Ax = b$, satisfazendo $x \geq 0$ corresponde a um vértice de X .

Demonstração

Consideremos $A = (B \ N)$, e $B \in \mathcal{R}^{m \times m}$ uma matriz quadrada inversível e $N \neq 0$, pois se $N = 0$ haveria um único ponto em X que seria o próprio vértice.

Tomemos $\bar{x} = (\bar{x}_B \ \bar{x}_N)^T$, onde $\bar{x}_N = 0$, uma solução básica viável, isto é, $\bar{x}_B = B^{-1}b \geq 0$.

Sejam x^1 e $x^2 \in X$ diferentes de \bar{x} , vamos supor que exista um $\lambda \in [0, 1]$ tal que $\bar{x} = \lambda x^1 + (1 - \lambda)x^2$, sabemos que $x^1 \geq 0$, $Ax^1 = Bx_B^1 + Nx_N^1 = b$, $x^2 \geq 0$, $Ax^2 = Bx_B^2 + Nx_N^2 = b$, como $N \neq 0$, $x_N^1 \neq 0$ e $x_N^2 \neq 0$, caso contrário, isto é, $x_N^1 = 0$ implicaria $x_B^1 = \bar{x}_B$ e $x_N^2 = 0$ implicaria $x_B^2 = \bar{x}_B$; contrariando a hipótese de que $x^1 \neq \bar{x}$ e $x^2 \neq \bar{x}$. Temos que

$$\bar{x} = \lambda x^1 + (1 - \lambda)x^2$$

implica

$$\bar{x}_B = \lambda x_B^1 + (1 - \lambda)x_B^2 \text{ e } 0 = \lambda x_N^1 + (1 - \lambda)x_N^2. \quad (3.6)$$

Como $x_N^1 \geq 0$ e $x_N^1 \neq 0$, $x_N^2 \geq 0$ e $x_N^2 \neq 0$, $\lambda \geq 0$, $(1 - \lambda) \geq 0$; λ e $(1 - \lambda)$ não podem ser anulados ao mesmo tempo, então (3.6) nunca será verificada. Assim sendo demonstramos, por absurdo, que uma solução básica viável corresponde a um vértice de conjunto poliédrico X . ■

Ilustremos, a seguir, a propriedade 3.3. Seja

$$X = \{x = (x_1 \ x_2 \ x_3)^T \in \mathcal{R}^3 \mid 2x_1 + 3x_2 + 6x_3 = 6, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0\}.$$

As soluções básicas de $2x_1 + 3x_2 + 6x_3 = 6$ são:

$$x_1 = \frac{6}{2} = 3, \ x_2 = x_3 = 0,$$

$$x_2 = \frac{6}{3} = 2, \ x_1 = x_3 = 0,$$

$$x_3 = \frac{6}{6} = 1, \ x_1 = x_2 = 0.$$

Todas satisfazem às restrições de não-negatividade. Na figura 3.4 representamos o conjunto X pelo triângulo unindo os três vértices.

A seguir daremos uma interpretação geométrica do processo de mudança de base (pivoteamento) no método do simplex. Para facilitarmos a notação

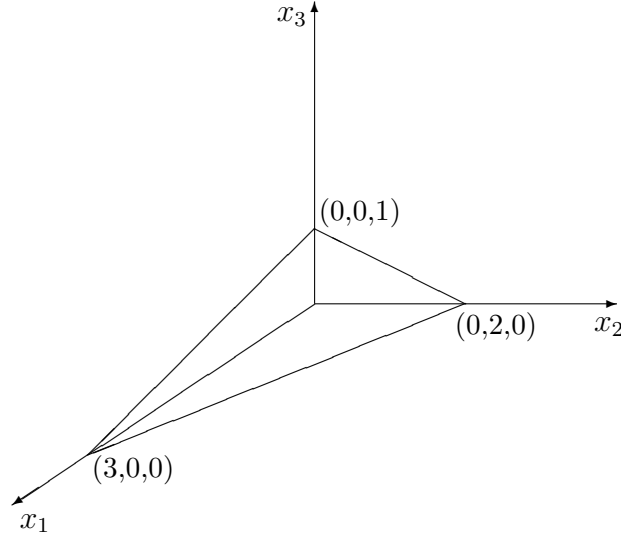


Figura 3.4: Conjunto poliédrico convexo, fechado e limitado

faremos $B(i) = i$, $i = 1, 2, \dots, m$; assim sendo o (PPL) sob a forma (2.14), (2.17) e (2.16) será escrito a seguir.

$$(PPL) : \text{maximizar } z = \bar{z} - \sum_{j \in I_N} (z_j - c_j)x_j \quad (3.7)$$

sujeito a:

$$x_i = \bar{x}_i - \sum_{j \in I_N} y_{ij}x_j, \quad i \in I_B \quad (3.8)$$

$$x_j = 0 + x_j, \quad j \in I_N \quad (3.9)$$

$$x_i \geq 0, \quad i \in I_B \quad (3.10)$$

$$x_j \geq 0, \quad j \in I_N. \quad (3.11)$$

As restrições do tipo (3.9) são colocadas para completar o sistema de equações, de modo que a interpretação geométrica se faça em \mathcal{R}^n .

Estudemos o processo de entrada na base de uma coluna a_k , $k \in I_N$.

$$x_i = \bar{x}_i - y_{ik}x_k, \quad i \in I_B,$$

$$x_j = 0 + \delta_{jk}x_k, \quad j \in I_N,$$

onde

$$\delta_{jk} = 1 \text{ se } j = k \text{ e } \delta_{jk} = 0 \text{ se } j \neq k,$$

ou ainda,

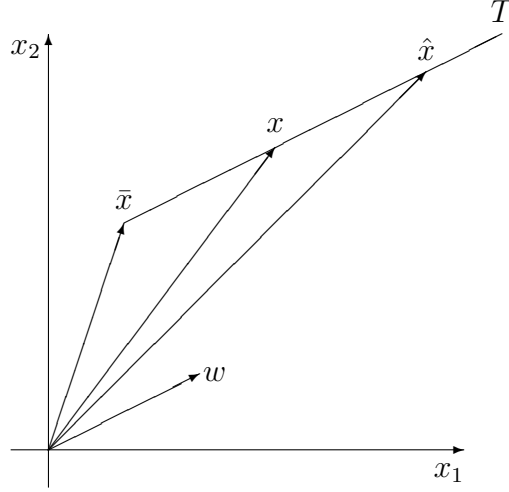


Figura 3.5: Direção de subida do método do simplex

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ x_m \\ x_{m+1} \\ \vdots \\ x_{k-1} \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_{m-1} \\ \bar{x}_m \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} -y_{1k} \\ -y_{2k} \\ \vdots \\ -y_{m-1\ k} \\ -y_{mk} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} x_k,$$

caso denominemos $w^T = (-y_{1k} \ -y_{2k} \ \dots \ -y_{mk} \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)$, poderemos ainda escrever:

$$x = \bar{x} + x_k w, \ x_k \geq 0. \quad (3.12)$$

Quando x_k aumentar, o vetor x será modificado, na figura 3.5 ilustramos este procedimento em \mathcal{R}^2 . Aí observamos que x estará na semi-reta T , paralela ao vetor w .

No caso em que $x_k \leq \alpha_k < \infty$ teremos $\hat{x} = \bar{x} + \alpha_k w$. Verificamos que ao passarmos da solução básica \bar{x} para a solução também básica \hat{x} estaremos caminhando sobre o segmento da semi-reta T entre \bar{x} e \hat{x} . Este segmento é uma aresta do conjunto poliédrico formado pelas restrições do (PPL), como será visto no capítulo 6. Isto quer dizer que w fornece a direção desta aresta. A determinação de w é imediata pelo método do simplex.

Definição 3.2 Quando um conjunto poliédrico convexo for limitado será denominado um politopo.

Exemplo 3.3 Tomemos, novamente, o (PPL) do exemplo 2.1, já com as variáveis de folga consideradas.

Começemos com

$$B = (a_1 \ a_4 \ a_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 2 \end{pmatrix},$$

calculemos B^{-1} da maneira apresentada na seção 3.4.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} = E_3^{-1}; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{3}{2} \\ \frac{3}{2} \end{pmatrix}.$$

Seguindo a notação da seção 3.4 teremos

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ \frac{3}{2} & 0 & 1 \end{pmatrix}, \text{ cuja inversa } \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ \frac{3}{2} & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{3}{2} & 0 & 1 \end{pmatrix},$$

logo $B^{-1} = E_1^{-1}E_3^{-1}$. Finalmente

$$B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{3}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & -\frac{1}{2} \\ -\frac{3}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Lembremos que $I_N = \{3, 5\}$ e passamos a calcular \bar{x}_B , u e $z_j - c_j$, $j \in I_N$.

$$u = c_B B^{-1} = (3 \ 0 \ 5) \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & -\frac{1}{2} \\ -\frac{3}{2} & 0 & \frac{1}{2} \end{pmatrix} = \left(-\frac{9}{2} \ 0 \ \frac{5}{2}\right),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_4 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix},$$

$$z_3 = ua_3 = \left(-\frac{9}{2} \ 0 \ \frac{5}{2}\right) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -\frac{9}{2} \Rightarrow z_3 - c_3 = -\frac{9}{2} - 0 = -\frac{9}{2} < 0,$$

$$z_5 = ua_5 = \begin{pmatrix} -\frac{9}{2} & 0 & \frac{5}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{5}{2} \Rightarrow z_5 - c_5 = \frac{5}{2} - 0 = \frac{5}{2} > 0.$$

Façamos a_3 entrar na base $B^{-1}a_3 = \left(1 \ \frac{3}{2} \ -\frac{3}{2}\right)^T$, assim sendo

$$\begin{pmatrix} x_1 \\ x_4 \\ x_2 \\ x_3 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 3 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ -\frac{3}{2} \\ \frac{3}{2} \\ 1 \\ 0 \end{pmatrix} x_3.$$

É fácil verificar que $\alpha_3 = 2$, logo faremos $x_3 = 2$.

Esquematizaremos no plano $x_1 \times x_2$, ver figura 3.6, a operação a ser finalizada. Para isso observamos que

$$x_1 = 4 - \alpha, \quad x_2 = 3 + \frac{3}{2}\alpha,$$

para $0 \leq \alpha \leq \alpha_3 = 2$, ou ainda

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ \frac{3}{2} \end{pmatrix} \alpha,$$

mas, para $\alpha = 2$,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}.$$

Notemos que a aresta que vai do ponto $(4, 3)$ ao ponto $(2, 6)$ é paralela ao vetor $w = \begin{pmatrix} -1 & \frac{3}{2} \end{pmatrix}^T$. ■

3.6 Convergência do Método do Simplex

Começaremos esta seção motivando a leitora ou o leitor para a possibilidade do método do simplex não convergir.

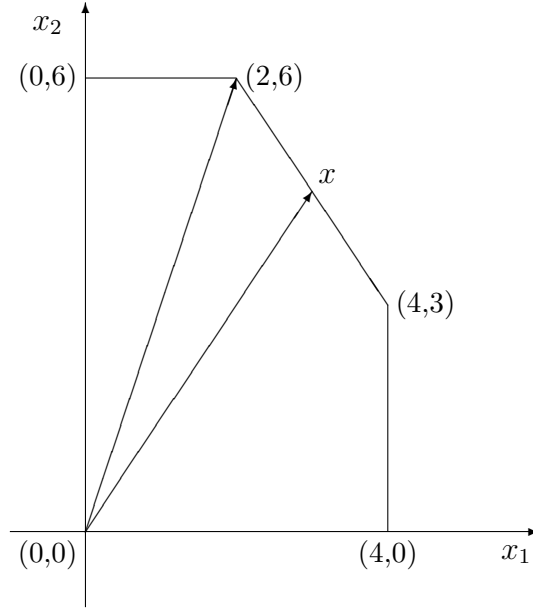


Figura 3.6: Direção de uma aresta do poliedro. O ponto x pode ser escrito como $x = (4 \ 3)^T + (-1 \ \frac{3}{2})^T \alpha$, $\alpha \in [0, 2]$

3.6.1 Motivação

Consideremos novamente o (PPL) em sua forma padrão:

$$(PPL) : \text{maximizar } z = cx \quad (3.13)$$

sujeito a:

$$Ax = b \quad (3.14)$$

$$x \geq 0. \quad (3.15)$$

Onde $c = (c_1 \ c_2 \ \dots c_n)$, $x^T = (x_1 \ x_2 \ \dots x_n)$, $b^T = (b_1 \ b_2 \ \dots b_m)$, $A = (a_1 \ a_2 \ \dots a_n)$ e $a_j^T = (a_{1j} \ a_{2j} \ \dots a_{mj})$, isto é, $c^T \in \mathcal{R}^n$, $x \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $A \in \mathcal{R}^{m \times n}$ e $a_j \in \mathcal{R}^m$.

Suporemos, sem perda de generalidade, que a matriz A tenha posto igual a m . Como já foi feito anteriormente poderemos escrever ainda

$$(PPL) : \text{maximizar } z = \bar{z} - \sum_{j \in I_N} (z_j - c_j) x_j \quad (3.16)$$

sujeito a:

$$x_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_N} y_{ij} x_j, \ i = 1, \dots, m. \quad (3.17)$$

$$x_j \geq 0, \quad j \in I_B \cup I_N. \quad (3.18)$$

O número total de soluções básicas de (3.14) tem por cota superior

$$\mathcal{C}_{n,m} = \frac{n!}{m!(n-m)!}.$$

O método primal do simplex parte de uma solução básica de (3.14) satisfazendo (3.15), isto é, uma solução básica viável do (*PPL*) e, de maneira iterativa (pivoteamento), passa a outra solução básica viável até que tenhamos satisfeito às condições de otimalidade ($z_j - c_j \geq 0, \quad j \in I_N$) ou concluindo que a solução é ilimitada.

Se de uma iteração para outra a variável x_l entrar na base então z terá um novo valor, $z = \hat{z}$, onde $\hat{z} = \bar{z} - (z_l - c_l)x_l$. Lembremos que $z_l - c_l < 0$. Caso x_l tome um valor estritamente positivo então $\hat{z} > \bar{z}$, caso x_l tenha um valor nulo, teremos $\hat{z} = \bar{z}$. Neste último caso z não modifica seu valor de uma iteração para outra. Quando uma variável básica tomar um valor nulo diremos que a solução básica é degenerada.

Quando aplicarmos o método do simplex a um (*PPL*) em que todas as soluções básicas não são degeneradas z crescerá a cada iteração, dessa maneira nunca repetiremos uma solução básica. O método do simplex convergirá sempre neste caso.

No entanto, quando tivermos soluções básicas degeneradas poderemos observar, às vezes, que após algumas iterações, sem que o valor de z se modifique, tenhamos uma solução básica já visitada anteriormente pelo método do simplex. Diremos, neste caso, que o método do simplex cicla.

3.6.2 Exemplo de Ciclagem

Seja o (*PPL*), ver [Ch 83]:

$$(\text{PPL}) : \text{maximizar } z = 10x_1 - 57x_2 - 9x_3 - 24x_4$$

sujeito a:

$$\begin{array}{rcccccccl} 0.5x_1 & - & 5.5 & - & 2.5x_3 & - & 9x_4 & + & x_5 & & = & 0 \\ 0.5x_1 & - & 1.5x_2 & - & 0.5x_3 & + & x_4 & & & + & x_6 & = & 0 \\ x_1 & & & & & & & & & & + & x_7 & = & 1 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, \dots, 7.$$

Caso partamos da solução básica tal que $I_B = \{5, 6, 7\}$ e utilizarmos como critério de entrada o índice associado ao menor $z_j - c_j, \quad j \in I_N$ e, no caso de empate, como critério de saída escolhermos a coluna $a_{B(i)}$ com o menor

$B(i)$ para sair da base, verificaremos que z não variará de valor durante seis iterações e na sexta retornaremos a $I_B = \{5, 6, 7\}$!

Talvez com outro critério possamos contornar o fenômeno de ciclagem que pode apresentar o método do simplex.

3.6.3 Regra de Bland

Bland [Bl 77] enunciou uma regra com um critério de escolha para a coluna que vai entrar na base e um outro critério para a coluna que sairá da base durante o método do simplex que evita a ciclagem do algoritmo. Esta regra é do menor índice ou de Bland. Passemos a apresentar esta regra.

Regra de Bland (Método primal de simplex)

Critério de entrada (problema de maximização):

a_p entra na base se $z_p - c_p < 0$ e p for o menor índice entre todos os $j \in I_N$ tais que $z_j - c_j < 0$.

Critério de saída:

a_s sai da base se $\frac{\bar{x}_s}{y_{lp}} = \min_{y_{ip} > 0, i=1,2,\dots,m} \left\{ \frac{\bar{x}_{B(i)}}{y_{ip}} \right\} = \theta$, onde $s = B(l)$ e s é o menor índice entre todos aqueles para os quais $\frac{\bar{x}_{B(i)}}{y_{ip}} = \theta$.

Proposição 3.4 *Caso aplicarmos a regra de Bland durante a resolução de um problema de programação linear, utilizando o método primal do simplex, nunca haverá ciclo, isto é, o método do simplex convergirá.*

Demonstração

A idéia da demonstração é que aplicando a regra de Bland o método do simplex cycle, dessa maneira a prova dar-se-á por absurdo.

Como haverá ciclagem, isto é, após um certo número de iterações voltaremos a uma solução básica já visitada pelo método do simplex anteriormente. Para que isso seja possível, o valor de z não se modificará nessas iterações e as soluções básicas em questão serão todas degeneradas.

Suponhamos que na iteração q tenhamos encontrado a mesma solução básica da iteração 1 ($q > 1$), utilizando a regra de Bland.

Denominaremos K o conjunto dos índices das colunas que participaram da base e não participaram da base em pelo menos uma dessas q iterações.

Seja $t = \max\{i \mid i \in K\}$. Consideremos que a_t saia da base na iteração \bar{q} ($1 \leq \bar{q} \leq q$), isto é, será não básica na iteração $\bar{q} + 1$ e que a_p entre na base na iteração \bar{q} , isto é, a_p estará na base na iteração $\bar{q} + 1$, logo $p \in K$ por definição de K .

Como supusemos que o procedimento cicle, poderemos continuar as iterações $\bar{q} + 1, \bar{q} + 2, \dots$ até que a_t volte à base. Seja \hat{q} a iteração em que a_t voltou a base.

Sabemos que o valor de z em todas essas iterações permaneceu invariável. Na iteração \bar{q} teremos:

$$x_{\bar{B}(i)} = \bar{x}_{\bar{B}(i)} - \sum_{j \in I_N} \bar{y}_{ij} x_j, \quad i = 1, \dots, m. \quad (3.19)$$

e

$$z = \bar{z} - \sum_{j \in I_N} \bar{y}_{0j} x_j, \quad (3.20)$$

onde \bar{B} é a base na iteração \bar{q} , $\bar{y}_{0j} = \bar{z}_j - c_j$, $\bar{z}_j = c_{\bar{B}} \bar{B}^{-1} a_j$ e $\bar{y}_j = \bar{B}^{-1} a_j$.

Na iteração \hat{q} explicitaremos z da seguinte maneira:

$$z = \bar{z} - \sum_{j \in I_N} \hat{y}_{0j} x_j, \quad (3.21)$$

onde \hat{B} é a base, $\hat{y}_{0j} = \hat{z}_j - c_j$ e $\hat{z}_j = c_{\hat{B}} \hat{B}^{-1} a_j$, lembremos que $\hat{y}_{0j} = 0$ para $j \in I_{\hat{B}}$.

A expressão (3.21) pode também ser considerada sob a seguinte forma:

$$z = \bar{z} - \sum_{j=1}^n \hat{y}_{0j} x_j, \quad (3.22)$$

que deve ser verificada para $x_p = \lambda$, $x_j = 0$, $j \in I_{\bar{N}} - \{p\}$ e $x_{\bar{B}(i)} = \bar{x}_{\bar{B}(i)} - \bar{y}_{ip} \lambda$, $i = 1, 2, \dots, m$ fornecendo

$$z = \bar{z} - \bar{y}_{0p} \lambda. \quad (3.23)$$

Assim sendo, (3.22) tornar-se-á:

$$z = \bar{z} - \hat{y}_{0p} \lambda - \sum_{i=1}^m \hat{y}_{0\bar{B}(i)} (\bar{x}_{\bar{B}(i)} - \bar{y}_{ip} \lambda). \quad (3.24)$$

De (3.23) e (3.24) podemos escrever:

$$\bar{z} - \bar{y}_{0p} \lambda = \bar{z} - \hat{y}_{0p} \lambda - \sum_{i=1}^m \hat{y}_{0\bar{B}(i)} (\bar{x}_{\bar{B}(i)} - \bar{y}_{ip} \lambda). \quad (3.25)$$

A equação (3.25) pode ser também expressa como:

$$(\hat{y}_{0p} - \bar{y}_{0p} - \sum_{i=1}^m \hat{y}_{0\bar{B}(i)} \bar{y}_{ip}) \lambda = - \sum_{i=1}^m \hat{y}_{0\bar{B}(i)} \bar{x}_{\bar{B}(i)} = \text{constante para qualquer } \lambda.$$

Logo esta constante tem que ser nula, assim sendo:

$$\hat{y}_{0p} - \bar{y}_{0p} - \sum_{i=1}^m \hat{y}_{0\bar{B}(i)} \bar{y}_{ip} = 0. \quad (3.26)$$

Como a_p entra na base na iteração \bar{q} então $\bar{y}_{0p} < 0$ e como a_p não está entrando na base na iteração \hat{q} então $\hat{y}_{0p} \geq 0$ (pois $p < t$ e estamos utilizando a regra de Bland).

Para que a expressão (3.26) seja verificada existirá r tal que $\hat{y}_{0\bar{B}(r)} \bar{y}_{rp} > 0$, $r \in \{1, 2, \dots, m\}$. Sabemos que $a_{\bar{B}(r)}$ é básica em \bar{q} e não é básica em \hat{q} , pois $\hat{y}_{0\bar{B}(r)} \neq 0$, portanto $\bar{B}(r) \in K$ e $\bar{B}(r) \leq t$.

Verificaremos que $\bar{B}(r) \neq t$, pois a_t deixa a base na iteração \bar{q} e a_p entra implicando $\bar{y}_{ip} > 0$ para $\bar{B}(i) = t$. Como a_t volta à base na iteração \hat{q} , temos $\hat{y}_{0t} < 0$, logo $\hat{y}_{0t} \bar{y}_{ip} < 0$ para $\bar{B}(i) = t, i \in \{1, 2, \dots, m\}$.

Só poderemos então considerar $\bar{B}(r) < t$ e assim sendo não teremos $\hat{y}_{0\bar{B}(r)} < 0$ (e sim $\hat{y}_{0\bar{B}(r)} > 0$) implicando $\bar{y}_{rp} > 0$ para que possamos satisfazer $\hat{y}_{0\bar{B}(r)} \bar{y}_{rp} > 0$.

Todas as iterações de \bar{q} a \hat{q} estão associadas a bases degeneradas implicando que o valor de $x_{\bar{B}(r)}$ seja o mesmo na base ou fora, isto é, igual a zero. Assim sendo, $\bar{x}_{\bar{B}(r)} = 0$ e como $\bar{y}_{rp} > 0$ então, pela regra de Bland, $a_{\bar{B}(r)}$ deveria deixar a base na iteração \bar{q} , pois $\bar{B}(r) < t$. Esta contradição completa a prova. ■

3.7 Exercícios

1. Seja

$$(P) : \text{maximizar } x_0 = 6x_1 + 4x_2$$

sujeito a:

$$\begin{array}{rcccccccl} 3x_1 & + & 2x_2 & + & x_3 & & & = & 18 \\ x_1 & & & & & + & x_4 & = & 4 \\ & & x_2 & & & & + & x_5 & = & 6 \\ & & & & x_j \geq 0, & j = 1, 2, 3, 4, 5. \end{array}$$

Verificar que a solução não básica viável de (P) , $x_1 = 3, x_2 = \frac{9}{2}, x_3 = 0, x_4 = 1, x_5 = \frac{3}{2}$, fornecendo $x_0 = 36$ é ótima de (P) .

2. Dados os pontos (x_i, y_i) , $i = 1, 2, \dots, p$, onde x_i e y_i são números reais. Gostaríamos de escrever que $y = ax + b$ tal que os parâmetros a e b forneçam o mínimo de $\max_{i=1,2,\dots,p} \{|y_i - (ax_i + b)|\}$. Determinar um problema de programação linear que encontre esses parâmetros a e b . Justificar.

3. Dados os pontos (x_i, y_i) , $i = 1, 2, \dots, p$, onde x_i e y_i são números reais. Gostaríamos de escrever que $y = ax^2 + bx + c$ tal que os parâmetros a , b e c forneçam o mínimo de $\sum_{i=1,2,\dots,p} |y_i - (ax_i^2 + bx_i + c)|$. Fornecer um problema de programação linear que encontre esses parâmetros a , b e c . Justificar.

4. Seja

$$(P) : \text{ maximizar } z = 4x_1 + 3x_2 + 2x_3 + 3x_4 + x_5$$

sujeito a:

$$\begin{array}{ccccccccc} 3x_1 & + & 2x_2 & + & x_3 & + & 2x_4 & + & x_5 & = & 13 \\ 5x_1 & + & 4x_2 & + & 3x_3 & + & 4x_4 & + & x_5 & = & 25 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

Verificar que (P) não é vazio e que todas suas variáveis $x_j \geq 0$, $j = 1, 2, 3, 4, 5$ são limitadas superiormente. Demonstrar que toda solução viável de (P) é também solução ótima de (P) .

Capítulo 4

Dualidade em Programação Linear

Tomemos o problema de programação linear (2.1), (2.2) e (2.3),

$$(P) : \text{ maximizar } z = \sum_{j=1}^p c_j x_j \quad (4.1)$$

sujeito a:

$$\sum_{j=1}^p a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, q \quad (4.2)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, p. \quad (4.3)$$

Associemos a cada restrição de (4.2) a variável $u_i \geq 0$, $i = 1, 2, \dots, q$ e definamos o seguinte problema

$$(D) : \text{ minimizar } d = \sum_{i=1}^q b_i u_i \quad (4.4)$$

sujeito a:

$$\sum_{i=1}^q a_{ij} u_i \geq c_j, \quad j = 1, 2, \dots, p \quad (4.5)$$

$$u_i \geq 0, \quad i = 1, 2, \dots, q. \quad (4.6)$$

Por definição diremos que (P) é o problema primal e (D) o seu dual.

Proposição 4.1 *O dual de (D) é (P) .*

A demonstração desta propriedade é deixada à leitora ou ao leitor, lembrando que $\min d = -\max(-d)$.

Os problemas (P) e (D) podem ser colocados sob forma matricial:

$$(P) : \text{maximizar } z = cx \quad (4.7)$$

sujeito a:

$$Ax \leq b \quad (4.8)$$

$$x \geq 0. \quad (4.9)$$

e

$$(D) : \text{minimizar } d = ub \quad (4.10)$$

sujeito a:

$$uA \geq c \quad (4.11)$$

$$u \geq 0, \quad (4.12)$$

onde c^T e $x \in \mathcal{R}^p$, b e $u^T \in \mathcal{R}^q$ e $A \in \mathcal{R}^{q \times p}$.

O resultado apresentado a seguir é conhecido como **Teorema da Dualidade Fraca**

Teorema 4.1 *Se \bar{x} satisfizer (4.8) e (4.9) e \bar{u} satisfizer (4.11) e (4.12) então teremos $c\bar{x} \leq \bar{u}b$.*

Demonstração

Sabemos que $A\bar{x} \leq b$ e $\bar{u} \geq 0$ então

$$\bar{u}A\bar{x} \leq \bar{u}b. \quad (4.13)$$

Por outro lado temos também que $\bar{u}A \geq c$ e $\bar{x} \geq 0$ então

$$\bar{u}A\bar{x} \geq c\bar{x}. \quad (4.14)$$

De (4.13) e (4.14) teremos:

$$c\bar{x} \leq \bar{u}A\bar{x} \leq \bar{u}b, \quad (4.15)$$

logo $c\bar{x} \leq \bar{u}b$. ■

Proposição 4.2 *Se \bar{x} for uma solução viável de (P) , \bar{u} uma solução viável de (D) e $c\bar{x} = \bar{u}b$ então \bar{x} será um ótimo de (P) e \bar{u} será um ótimo de (D) .*

Demonstração

Da propriedade 4.1 temos que $c\bar{x} \leq ub$ para todo u viável de (D) e como, por hipótese, $c\bar{x} = \bar{u}b$, o que implica $\bar{u}b \leq ub$ para todo u viável de (D) , logo \bar{u} é uma solução ótima (mínimo) de (D) . A demonstração que \bar{x} é uma solução ótima (máximo) de (P) é análoga. ■

Proposição 4.3 *Se (P) tiver uma solução “ ótima ” ilimitada então (D) será vazio.*

Demonstração

Suporemos que (D) não seja vazio e que \bar{u} seja uma solução viável de (D) , então pela propriedade 4.1, $c\bar{x} \leq \bar{u}b$ para todo x viável de (P) , o que é um absurdo pois máximo $\{cx\} \rightarrow +\infty$. Logo (D) será vazio. ■

Quando incorporarmos as variáveis de folga $x_{p+i} \geq 0$, $i = 1, 2, \dots, q$ às restrições (4.2) poderemos redefinir c, x e A de maneira que (P) possa ser escrito de outra forma:

$$(P_1) : \text{maximizar } z = cx \quad (4.16)$$

sujeito a:

$$Ax = b \quad (4.17)$$

$$x \geq 0. \quad (4.18)$$

É simples verificarmos que o dual de (P_1) será

$$(D_1) : \text{minimizar } d = ub \quad (4.19)$$

sujeito a:

$$uA \geq c, \quad (4.20)$$

onde c^T e $x \in \mathcal{R}^n$, b e $u^T \in \mathcal{R}^m$ e $A \in \mathcal{R}^{m \times n}$.

Podemos verificar que em (D_1) , u é um vetor cujas componentes podem ser negativas ou nulas ou positivas. São livres quanto ao sinal.

Tal como foi feito no capítulo 2, seja $A = (B \ N)$, tal que $\det(B) \neq 0$. Se $\bar{u} = c_B B^{-1}$, tal que $\bar{u}A \geq c$, então, \bar{u} é uma solução viável de (D_1) . Neste caso diremos que $\bar{x} = (\bar{x}_B \ 0)^T$, onde $\bar{x}_B = B^{-1}b$, é uma **solução básica dual viável**.

Sabemos que $\bar{u}a_j = z_j$, $j \in I_B \cup I_N$, então $\bar{u}A \geq c$ implica em $z_j \geq c_j$ ou ainda $z_j - c_j \geq 0$, $j \in I_B \cup I_N$. Estas últimas desigualdades representam as condições de otimalidade do método do simplex, como visto anteriormente.

Se $\bar{x} = (\bar{x}_B \ 0)^T$, onde $\bar{x}_B = B^{-1}b$ (primal viável) e $\bar{u} = c_B B^{-1}$, tal que $\bar{u}A \geq c$, então \bar{x} é uma solução ótima de (P_1) .

Poderemos, facilmente, verificar que a propriedade 4.1, deduzida para (P_1) e (D_1) forneceria (4.15) sob a seguinte forma:

$$cx \leq uAx = ub, \quad (4.21)$$

para qualquer x de (P_1) e qualquer u de (D_1) .

O resultado apresentado a seguir é conhecido como **Teorema da Dualidade Forte**.

Teorema 4.2 *Se \hat{x} for uma solução ótima de (P_1) e \hat{u} uma solução ótima de (D_1) então $c\hat{x} = \hat{u}b$.*

Demonstração

Sem perda de generalidade podemos supor que \hat{x} seja uma solução básica viável de (P_1) associada a uma base B formada por m colunas de A e que $\bar{u} = c_B B^{-1}$.

Como \hat{x} é uma solução ótima de (P_1) , isto é, $z_j - c_j \geq 0$, $j \in I_B \cup I_N$ então $\bar{u}A \geq c$, ou ainda \bar{u} é viável de (D_1) .

Por outro lado, $c\hat{x} = c_B B^{-1}b = \bar{u}b$. De (4.21) temos que $c\hat{x} \leq ub$, para todo u viável de (D_1) , ou seja $\bar{u}b \leq ub$ para todo u viável de (D_1) . Assim sendo, \bar{u} é uma solução ótima de (D_1) , implicando $\bar{u}b = \hat{u}b$, completando a prova. ■

Consideremos agora o seguinte exemplo:

Exemplo 4.1

$$\text{maximizar } z = 3x_1 + 4x_2$$

sujeito a:

$$\begin{array}{rcll} x_1 & - & x_2 & \leq -1 \\ - & x_1 & + & x_2 \leq 0 \\ x_1 & \geq & 0 & \\ & & & x_2 \geq 0, \end{array}$$

cujos dual é

$$\text{minimizar } d = -u_1$$

sujeito a:

$$\begin{array}{rcll} u_1 & - & u_2 & \geq 3 \\ - & u_1 & + & u_2 \geq 4 \\ u_1 & \geq & 0 & \\ & & & u_2 \geq 0. \end{array}$$

Verificamos facilmente que ambos os problemas são vazios. ■

As propriedades 4.2, 4.3 e 4.2 mais o exemplo 4.1 permitem que enunciemos o seguinte teorema clássico em programação linear: **Teorema da Dualidade (existência)**.

Teorema 4.3 *Dado um par de problemas (um primal e seu dual) uma e somente uma das três afirmações é verdadeira:*

1. os dois problemas são vazios;
2. um é vazio e o outro é ilimitado;
3. ambos admitem soluções ótimas finitas (as respectivas funções objetivo no ótimo assumem o mesmo valor).

Finalmente, apresentamos a seguir um resultado conhecido como **Teorema das Folgas Complementares**

Teorema 4.4 *Se \hat{x} é ótimo de (P) e \hat{u} é ótimo de (D) então $(\hat{u}A - c)\hat{x} = 0$ e $\hat{u}(A\hat{x} - b) = 0$.*

Demonstração

Temos que $c\hat{x} = \hat{u}b$, assim sendo (4.15) se torna $c\hat{x} = \hat{u}A\hat{x} = \hat{u}b$, ou ainda, $c\hat{x} = \hat{u}A\hat{x}$ e $\hat{u}A\hat{x} = \hat{u}b$, logo $(\hat{u}A - c)\hat{x} = 0$ e $\hat{u}(A\hat{x} - b) = 0$. ■

Da propriedade 4.4 podemos ainda escrever:

$$\sum_{j=1}^p \hat{x}_j \left(\sum_{i=1}^q a_{ij} \hat{u}_i - c_j \right) = 0, \text{ como } \hat{x}_j \geq 0 \text{ e } \sum_{i=1}^q a_{ij} \hat{u}_i - c_j \geq 0, j = 1, 2, \dots, p$$

implica

$$\hat{x}_j \left(\sum_{i=1}^q a_{ij} \hat{u}_i - c_j \right) = 0, j = 1, 2, \dots, p \quad (4.22)$$

e

$$\sum_{i=1}^q \hat{u}_i \left(\sum_{j=1}^p a_{ij} \hat{x}_j - b_i \right) = 0, \text{ como } \hat{u}_i \geq 0 \text{ e } \sum_{j=1}^p a_{ij} \hat{x}_j - b_i \geq 0, i = 1, 2, \dots, q \text{ implica}$$

$$\hat{u}_i \left(\sum_{j=1}^p a_{ij} \hat{x}_j - b_i \right) = 0, i = 1, 2, \dots, q. \quad (4.23)$$

As relações (4.22) e (4.23) são denominadas condições de complementaridade.

A propriedade 4.4 deduzida a partir de (P_1) e (D_1) nos forneceria apenas $(\hat{u}A - c)\hat{x} = 0$, pois $A\hat{x} - b = 0$ em (P_1) .

É interessante notar que quando \bar{x} for uma solução básica (primal viável ou não) de (P_1) e $\bar{u} = c_B B^{-1}$ então $\bar{x}_j (\sum_{i=1}^q a_{ij} \bar{u}_i - c_j) = 0$, $j \in I_B \cup I_N$. As relações de complementaridade são sempre verificadas na aplicação do método do simplex.

4.1 Uma Interpretação das Variáveis Duais

Seja $z(b) = \max_x \{cx \mid Ax = b, x \geq 0\}$, $z : \mathcal{R}^m \rightarrow \mathcal{R}$.

Suponhamos que $\bar{x} = (x_B \ 0)^T$, $x_B = B^{-1}b \geq 0$ seja uma solução básica ótima do primal e $\bar{u} = c_B B^{-1}$ uma solução ótima do dual. Podemos escrever $\bar{u} = (\bar{u}_1 \ \bar{u}_2 \ \dots \ \bar{u}_k \ \dots \ \bar{u}_m)$.

Sejam $e_k^T = (0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)$, onde a componente igual a 1 ocupa a k -ésima linha do vetor e_k , $b' = b + e_k$ e $\bar{x}'_B = B^{-1}b' \geq 0$; esta nova solução continua sendo ótima de $z(b')$, pois $\bar{x}' = (\bar{x}'_B \ 0)^T$ é primal e dual viável.

Sabemos que $z(b) = \bar{u}b$ e $z(b') = \bar{u}b' = \bar{u}b + \bar{u}_k$, logo $z(b') = z(b) + \bar{u}_k$, \bar{u}_k poderá assim ser interpretada como sendo a variação de z quando aumentarmos de uma unidade a k -ésima componente de b , isto é, substituímos b_k por $b_k + 1$.

Podemos portanto, interpretar \bar{u}_k como sendo a derivada parcial de z em relação a b_k :

$$\frac{\partial z}{\partial b_k} = \bar{u}_k. \quad (4.24)$$

4.2 Método Dual do Simplex

Tomemos novamente o problema de programação linear sob a forma (2.7), (2.8) e (2.9), onde $A = (B \ N)$, tal que, B^{-1} exista.

Seja $\bar{x} = (\bar{x}_B \ 0)^T$, onde $\bar{x}_B = B^{-1}b$, é uma solução básica de (4.17) e tal que $\bar{u} = c_B B^{-1}$ satisfaça $\bar{u}A \geq c$ (\bar{u} é uma solução viável de (4.19) e (4.20)). Neste caso dizemos, por definição, que a base B é dual viável. Se $\bar{x}_B = B^{-1}b \geq 0$ dizemos também que B é uma base primal viável. Como $\bar{u}A \geq c$ implica $z_j - c_j \geq 0$, $j \in I_B \cup I_N$, então se B for primal e dual viável $\bar{x} = (\bar{x}_B \ 0)^T$ será uma solução ótima do primal e $\bar{u} = c_B B^{-1}$ uma solução ótima do dual.

Suponhamos que \bar{x} não esteja associada a uma base B primal viável, isto é, $\bar{x}_B = B^{-1}b \not\geq 0$ (existe, pelo menos, uma componente negativa de \bar{x}_B). A idéia do método dual do simplex para resolver (4.16), (4.17) e (4.18) é de partir de uma base B básica dual viável, passar para uma nova base dual

viável, pela substituição de uma coluna de B . Este procedimento é repetido até que atinjamos uma base primal e dual viável.

Com a mesma notação do capítulo 2, consideremos

$$x_{B(k)} = \bar{x}_{B(k)} - \sum_{j \in I_N} y_{kj} x_j, \quad (4.25)$$

onde $\bar{x}_{B(k)} < 0$.

Se $y_{kj} \geq 0$ para todo $j \in I_N$ em (4.25) e como $x_j \geq 0$ para todo $j \in I_N$, $x_{B(k)}$ nunca poderá ser não negativo, implicando que o problema primal será vazio e o dual ilimitado.

Consideremos o conjunto $L_k = \{j \in I_N \mid y_{kj} < 0\} \neq \emptyset$.

Como $\bar{x}_{B(k)} < 0$, escolheremos a coluna $a_{B(k)}$ para deixar a base e tomaremos a_p associada ao índice p , tal que $y_{kp} < 0$ para entrar na nova base. Qual deve ser p tal que a nova base continue dual viável?

A matriz B' formada pela substituição em B da coluna $a_{B(k)}$ por a_p , tem sua inversa calculada, como foi exposto na seção 3.4, da seguinte maneira:

$(B')^{-1} = E_k^{-1} B^{-1}$, onde $E_k = (e_1 \ e_2 \ \dots \ e_{k-1} \ y_p \ e_{k+1} \ \dots \ e_m)$; lembramos que $y_p = B^{-1} a_p$.

Desejamos calcular $y'_j = (B')^{-1} a_j$, $j \in I_N$, ou ainda $y'_j = E_k^{-1} (B^{-1} a_j) = E_k^{-1} y_j$, isto é :

$$y'_{ij} = y_{ij} - \frac{y_{ip}}{y_{kp}} y_{kj}, \quad i \neq k, \quad (4.26)$$

$$y'_{kj} = \frac{y_{kj}}{y_{kp}}.$$

Observação: y_{kp} será o pivô.

A atualização dos $z_j - c_j$ será feita de maneira equivalente à mostrada em (4.26), como será visto ainda neste capítulo.

O problema (2.7), (2.8) e (2.9) poderá ser escrito:

maximizar z

sujeito a:

$$z - c_B x_B - c_N x_N = 0 \quad (4.27)$$

$$B x_B + N x_N = b \quad (4.28)$$

$$x_B \geq 0, \ x_N \geq 0.$$

Solucionaremos o sistema (4.27) e (4.28) obtendo z e x_B em função de x_N .

Seja

$$\hat{B} = \begin{pmatrix} 1 & -c_B \\ 0 & B \end{pmatrix}, \text{ como } \det(B) \neq 0 \Rightarrow \text{existe } \hat{B}^{-1}.$$

$$\hat{B}^{-1} = \begin{pmatrix} 1 & -c_B \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} 1 & c_B B^{-1} \\ 0 & B^{-1} \end{pmatrix} = \begin{pmatrix} 1 & u \\ 0 & B^{-1} \end{pmatrix}.$$

Temos que

$$\begin{pmatrix} z_j - c_j \\ y_j \end{pmatrix} = \begin{pmatrix} 1 & c_B B^{-1} \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} -c_j \\ a_j \end{pmatrix}.$$

Denominemos $y_{0j} = z_j - c_j$. Logo a atualização de $z_j - c_j$ será feita como em (4.26), isto é:

$$y'_{0j} = y_{0j} - \frac{y_{0p}}{y_{kp}} y_{kj}. \quad (4.29)$$

Desejamos que a escolha de p , índice da coluna que substituirá $a_{B(k)}$ na nova base, seja feita tal que a nova base continue dual viável: $y'_{0j} \geq 0$, $j \in I_B \cup I_N$, ou ainda $y_{0j} - \frac{y_{0p}}{y_{kp}} y_{kj} \geq 0$; consideremos dois casos a seguir.

1º caso: $y_{kj} < 0$ então

$$\frac{y_{0j}}{y_{kj}} - \frac{y_{0p}}{y_{kp}} \leq 0 \quad \text{ou} \quad \frac{y_{0p}}{y_{kp}} \geq \frac{y_{0j}}{y_{kj}}, \quad j \in I_N. \quad (4.30)$$

Lembremos que $j \in I_B$ implica $y_{0j} = 0$. De (4.30) temos que

$$\frac{y_{0p}}{y_{kp}} = \max_{j \in L_k} \left\{ \frac{y_{0j}}{y_{kj}} \right\}. \quad (4.31)$$

2º caso: $y_{kj} \geq 0$ então

$$y'_{0j} = y_{0j} - \frac{y_{0p}}{y_{kp}} y_{kj} \geq 0, \text{ pois } y_{0p} \geq 0 \text{ e } y_{kp} < 0, \text{ logo } -\frac{y_{0p}}{y_{kp}} y_{kj} \geq 0.$$

Apenas nos interessará os $y_{kj} < 0$, para os quais (4.31) nos fornecerá o índice p da coluna que entrará na base, ocupando o lugar de $a_{B(k)}$ na próxima iteração.

Verificamos também que a variação do valor da função objetivo em cada iteração: $\bar{z} = c_B B^{-1} b$ e $\bar{z}' = \bar{z} - y_{0p} \frac{\bar{x}_{B(k)}}{y_{kp}}$, como $y_{0p} \geq 0$, $\bar{x}_{B(k)} < 0$, $y_{kp} < 0$, logo $\bar{z}' \leq \bar{z}$, no caso em que $y_{0p} > 0$ teremos $\bar{z}' < \bar{z}$. Estamos minimizando a função objetivo do programa dual (4.19) e (4.20).

Passaremos à descrição de um procedimento para a solução de um (PPL) pelo método dual do simplex no caso de maximização.

Procedimento 3 (maximização, método dual do simplex)

Dada uma base B dual viável para (4.16), (4.17) e (4.18).

Se $\bar{x}_B \geq 0$, a base B está associada a uma solução primal e dual viável. PARE.

Caso contrário ($\bar{x}_B \not\geq 0$) escolhe-se um k para o qual $\bar{x}_{B(k)} < 0$, se $L_k = \emptyset$, o (PPL) é vazio. PARE.

se $L_k \neq \emptyset$, toma-se a coluna a_p , $p \in I_N$, talque

$$\frac{y_{0p}}{y_{kp}} = \max_{j \in L_k} \left\{ \frac{y_{0j}}{y_{kj}} \right\},$$

a coluna a_p ocupará o lugar da coluna $a_{B(k)}$ em B .
(MUDANÇA DE BASE).

Fim do procedimento 3

O procedimento 3 é aplicado, iterativamente, para cada nova base até que uma das regras de parada seja verificada.

Exemplo 4.2 Seja o

$$(PPL) : \text{maximizar } z = -4x_1 - 5x_2$$

sujeito a:

$$\begin{aligned} x_1 + 4x_2 &\geq 5 \\ 3x_1 + 2x_2 &\geq 7 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Associaremos às restrições não triviais as variáveis de folga $x_3 \geq 0$, $x_4 \geq 0$ tais que o (PPL) fique sob a seguinte forma.

$$(PPL) : \text{maximizar } z = -4x_1 - 5x_2 + 0x_3 + 0x_4$$

sujeito a:

$$\begin{aligned} x_1 + 4x_2 - x_3 &= 5 \\ 3x_1 + 2x_2 - x_4 &= 7 \\ x_j &\geq 0, \quad j = 1, 2, 3, 4, \end{aligned}$$

onde

$$A = (a_1 \ a_2 \ a_3 \ a_4) = \begin{pmatrix} 1 & 4 & -1 & 0 \\ 3 & 2 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 7 \end{pmatrix}, \quad c = (-4 \ -5 \ 0 \ 0).$$

Tomemos

$$I_B = \{3, 4\}, \quad I_N = \{1, 2\},$$

$$B = (a_3 \ a_4) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \text{ logo } B^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$u = c_B B^{-1} = (0 \ 0) \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = (0 \ 0),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \end{pmatrix} = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} -5 \\ -7 \end{pmatrix}.$$

Verificamos que $\bar{x} = (\bar{x}_B \ 0)^T \not\geq 0$ (B não é primal viável). No entanto,

$$z_1 = ua_1 = (0 \ 0) \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 0 \Rightarrow z_1 - c_1 = 0 - (-4) = 4 > 0$$

e

$$z_2 = ua_2 = (0 \ 0) \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 0 \Rightarrow z_2 - c_2 = 0 - (-5) = 5 > 0,$$

logo B é dual viável.

Tomemos $B(2) = 4$ pois $\bar{x}_4 < 0$, procedendo assim estamos escolhendo a_4 para deixar a base.

Calculemos y_{2j} , $j \in I_N$, obtendo:

$$y_{21} = (0 \ -1) \begin{pmatrix} 1 \\ 3 \end{pmatrix} = -3 \text{ e } y_{22} = (0 \ -1) \begin{pmatrix} 4 \\ 2 \end{pmatrix} = -2,$$

onde $(0 \ -1)$ é a segunda linha de B^{-1} . Verificamos que $L_2 = \{1, 2\}$.

$$\max \left\{ \frac{4}{-3}, \frac{5}{-2} \right\} = \frac{4}{-3} = \frac{z_1 - c_1}{y_{21}},$$

logo a coluna a_1 substituirá a coluna a_4 na próxima base.

Segunda base:

$$I_B = \{3, 1\}, \ I_N = \{4, 2\},$$

$$B^{-1} = \begin{pmatrix} -1 & 1 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix},$$

$$u = c_B B^{-1} = (0 \ -4) \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} = (0 \ -\frac{4}{3}),$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \end{pmatrix} = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} -\frac{8}{3} \\ \frac{7}{3} \end{pmatrix} \not\geq 0,$$

$$\bar{z} = \left(0 \ -\frac{4}{3}\right) \begin{pmatrix} 5 \\ 7 \end{pmatrix} = -\frac{28}{3}.$$

Como $\bar{x}_{B(1)} = \bar{x}_3 = -\frac{8}{3} < 0$, a_3 sairá da base na próxima iteração.
 Calculemos y_{1j} , $j \in I_N$:

$$y_{14} = \begin{pmatrix} -1 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -\frac{1}{3} \text{ e } y_{12} = \begin{pmatrix} -1 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = -\frac{10}{3},$$

onde $\begin{pmatrix} -1 & \frac{1}{3} \end{pmatrix}$ é a primeira linha de B^{-1} e $L_1 = \{4, 2\}$.

$$z_4 = ua_4 = \begin{pmatrix} 0 & -\frac{4}{3} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{4}{3} \Rightarrow z_4 - c_4 = \frac{4}{3} - 0 = \frac{4}{3} > 0,$$

$$z_2 = ua_2 = \begin{pmatrix} 0 & -\frac{4}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = -\frac{8}{3} \Rightarrow z_2 - c_2 = -\frac{8}{3} - (-5) = \frac{7}{3} > 0,$$

$$\max \left\{ \frac{\frac{4}{3}}{-\frac{1}{3}}, \frac{\frac{7}{3}}{-\frac{10}{3}} \right\} = \frac{\frac{7}{3}}{-\frac{10}{3}} = -\frac{7}{10} = \frac{z_2 - c_2}{y_{12}},$$

assim sendo a_2 entrará na base no lugar de a_3 na próxima etapa.

Terceira base:

$$I_B = \{2, 1\}, \quad I_N = \{4, 3\},$$

$$B^{-1} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix},$$

$$u = c_B B^{-1} = (-5 \quad -4) \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} = \begin{pmatrix} -\frac{7}{10} & -\frac{11}{10} \end{pmatrix},$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \end{pmatrix} = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{8}{10} \\ \frac{18}{10} \end{pmatrix} \geq 0,$$

e

$$\bar{z} = c_B B^{-1}b = c_B \bar{x}_B = (-5 \quad -4) \begin{pmatrix} \frac{8}{10} \\ \frac{18}{10} \end{pmatrix} = -\frac{112}{10}.$$

Podemos então dizer que $x_1^* = \frac{18}{10}$, $x_2^* = \frac{8}{10}$, $x_3^* = x_4^* = 0$, implicando $z^* = -\frac{112}{10}$, é a solução ótima do primal (*PPL*); $u_1^* = -\frac{7}{10}$, $u_2^* = -\frac{11}{10}$ é a solução ótima do dual, a variável dual u_1 está associada à restrição $x_1 + 4x_2 \geq 5$ e u_2 à restrição $3x_1 + 2x_2 \geq 7$.

Na figura 4.1 será ilustrado o método dual do simplex para o exemplo aqui tratado. No espaço $x_1 \times x_2$ notamos que o método sai de uma solução básica dual viável (porém não primal viável) representada pelo ponto $(0, 0)$, prossegue para outra solução básica dual viável (ainda não primal viável) associada ao ponto $(\frac{7}{3}, 0)$ e, finalmente, atinge a solução básica dual e primal viável (ótima) ilustrada pelo ponto $(\frac{18}{10}, \frac{8}{10})$. ■

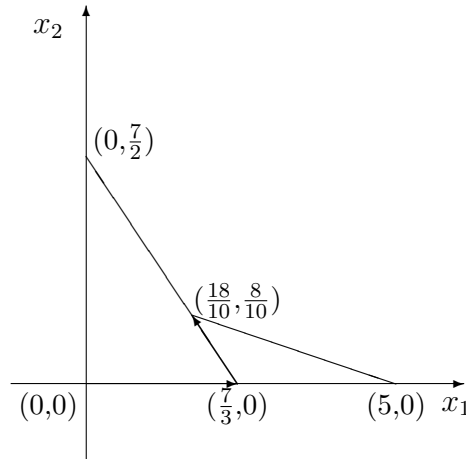


Figura 4.1: Iterações do método dual do simplex

Observação: verificamos facilmente que a expressão (4.31) pode ser substituída por

$$\frac{y_{0p}}{y_{kp}} = \min_{j \in L_k} \left\{ \left| \frac{y_{0j}}{y_{kj}} \right| \right\}, \quad (4.32)$$

seja para o caso de maximização, seja para o caso de minimização da função objetivo.

4.3 Análise de Sensibilidade

Nesta seção faremos uma análise sobre o impacto sofrido pela solução ótima de um problema de programação linear quando a sua estrutura é modificada.

Muitas vezes, devido a imprecisões nos dados de um problema de programação linear, é interessante avaliar quão sensível a solução ótima do problema é com relação a pequenas variações nestes dados, ou seja, é interessante analisar qual o intervalo de oscilação permitido a estes dados sem que a base ótima do problema seja alterada.

A seguir faremos esta análise, conhecida como análise de sensibilidade, com relação a oscilações nos custos associados às variáveis e no lado direito das restrições do problema

(*PPL*) : maximizar $z = cx$

sujeito a:

$$Ax = b \quad (4.33)$$

$$x \geq 0,$$

cuja solução ótima supomos conhecida e associada a uma base B , primal e dual viável.

4.3.1 Alterando o lado direito das restrições

Suponhamos que o lado direito das restrições do (*PPL*) (4.33) seja alterado de b para $b + \delta b$. Neste caso, a base ótima do (*PPL*) não deixa de ser dual viável, mas para que esta base continue sendo ótima, ela deverá manter-se também primal viável.

Desta forma, para que a base B continue sendo ótima para o problema modificado, devemos ter

$$B^{-1}(b + \delta b) \geq 0.$$

Exemplo 4.3 Consideremos o problema de programação linear do exemplo 4.2. Verifiquemos qual o intervalo em que o lado direito da primeira restrição pode se encontrar sem que a base ótima do problema seja alterada.

Para que a base ótima determinada por $I_B = \{2, 1\}$ não seja alterada, devemos ter:

$$B^{-1} \begin{pmatrix} b_1 + \delta b_1 \\ b_2 \end{pmatrix} \geq 0 \Rightarrow \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 5 + \delta b_1 \\ 7 \end{pmatrix} \geq 0 \Rightarrow$$

$$\frac{3}{10}(5 + \delta b_1) - \frac{7}{10} \geq 0 \Rightarrow \delta b_1 \geq -\frac{8}{3},$$

$$-\frac{2}{10}(5 + \delta b_1) + \frac{28}{10} \geq 0 \Rightarrow \delta b_1 \leq 9.$$

Portanto devemos ter $\delta b_1 \in [-\frac{8}{3}, 9]$, ou seja, para que a base ótima do problema não seja alterada, o lado direito da primeira restrição deve satisfazer a

$$b_1 \in \left[\frac{7}{3}, 14 \right]. \quad \blacksquare$$

4.3.2 Alterando o vetor custo

Suponhamos agora que o vetor custo associado às variáveis do (PPL) (4.33) seja alterado de c para $c + \delta c$. Neste caso, a base ótima do (PPL) não deixa de ser primal viável, mas para que esta base continue sendo ótima, ela deverá manter-se também dual viável.

Desta forma, para que a base B continue sendo ótima para o problema modificado, devemos ter

$$(c_B + \delta c_B)B^{-1}N - (c_N + \delta c_N) \geq 0.$$

Exemplo 4.4 Consideremos novamente o problema de programação linear do exemplo 4.2. Verifiquemos qual o intervalo em que o custo associado a variável x_2 pode se encontrar sem que a base ótima do problema seja alterada.

Para que a base ótima determinada por $I_B = \{2, 1\}$ não seja alterada, devemos ter:

$$\begin{aligned} (c_2 + \delta c_2 \quad c_1)B^{-1}N - (c_4 \quad c_3) &\geq 0 \Rightarrow \\ (-5 + \delta c_2 \quad -4) \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} - (0 \quad 0) &\geq 0 \Rightarrow \\ (-5 + \delta c_2 \quad -4) \begin{pmatrix} \frac{1}{10} & -\frac{3}{10} \\ -\frac{4}{10} & \frac{2}{10} \end{pmatrix} - (0 \quad 0) &\geq 0 \Rightarrow \\ \frac{1}{10}(-5 + \delta c_2) + \frac{16}{10} \geq 0 &\Rightarrow \delta c_2 \geq -11, \\ -\frac{3}{10}(-5 + \delta c_2) - \frac{8}{10} \geq 0 &\Rightarrow \delta c_2 \leq \frac{7}{3}. \end{aligned}$$

Portanto devemos ter $\delta c_2 \in [-11, \frac{7}{3}]$, ou seja, para que a base ótima do problema não seja alterada, o custo associado à segunda variável deve satisfazer a

$$c_2 \in \left[-16, -\frac{8}{3}\right]. \quad \blacksquare$$

4.4 Pós-otimização

Consideremos agora, que após a obtenção da solução ótima do (PPL) (4.33), alguma mudança na sua estrutura foi introduzida, levando a um novo problema (\overline{PPL}) , para o qual a base ótima do (PPL) possivelmente não é mais primal ou dual viável.

O processo de pós-otimização consiste em obter a solução ótima do problema modificado (\overline{PPL}) , considerando-se que a solução ótima do problema original (PPL) é conhecida.

O processo de pós-otimização torna-se mais eficiente quando a base B , associada a uma solução ótima do (PPL) pode ser utilizada para inicializar o método do simplex ou o método dual do simplex na resolução do problema modificado. Este procedimento evita o trabalho de obtenção de uma base inicial primal ou dual viável para o (\overline{PPL}) .

Consideraremos a seguir três alterações que são freqüentemente introduzidas sobre a estrutura de um problema de programação linear. Na última alteração, introdução de uma nova restrição ao problema, veremos que apesar da base B deixar de ser primal e dual viável, ela pode ser utilizada na construção de uma nova base dual viável para o problema modificado.

4.4.1 Alterando o lado direito das restrições

Suponhamos que o vetor b que define o lado direito das restrições no problema (PPL) seja alterado para $b + \delta b$ e consideremos

$$(\overline{PPL}) : \text{maximizar } z = cx$$

sujeito a:

$$Ax = b + \delta b$$

$$x \geq 0,$$

Como já visto na seção anterior, a base B associada a uma solução ótima de (PPL) não deixa de ser dual viável para o problema (\overline{PPL}) . Suponhamos, no entanto que $B^{-1}(b + \delta b) \not\geq 0$, ou seja, que B deixa de ser uma base primal viável. Neste caso, para obter a solução ótima do problema modificado, podemos aplicar o método dual do simplex, tomando B como base inicial.

Exemplo 4.5 Seja o (PPL) do exemplo 4.2. Suponhamos que o lado direito da primeira restrição seja alterado de 5 para 15.

Como já verificamos no exemplo 4.3, a base ótima do problema original deixa de ser primal viável com esta alteração, fato este que comprovamos abaixo.

$$\bar{x}_B = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \end{pmatrix} = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = B^{-1}b = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 15 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{38}{10} \\ -\frac{2}{10} \end{pmatrix},$$

ou seja, $\bar{x}_B \not\geq 0$.

Por outro lado, sabemos que B não deixa de ser dual viável. Sendo assim, aplicamos a seguir o método dual do simplex para resolver o problema modificado, tomando a base B como base inicial.

Base inicial:

$$I_B = \{2, 1\}, I_N = \{4, 3\},$$

Como $\bar{x}_{B(2)} = \bar{x}_1 = -\frac{2}{10} < 0$, a_1 sairá da base na próxima iteração.

Calculemos y_{2j} , $j \in I_N$:

$$y_{24} = \begin{pmatrix} -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -\frac{4}{10} \text{ e } y_{23} = \begin{pmatrix} -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \frac{2}{10},$$

onde $\begin{pmatrix} -\frac{2}{10} & \frac{4}{10} \end{pmatrix}$ é a segunda linha de B^{-1} e $L_2 = \{4\}$.

Como L_2 só tem um elemento, ele define a coluna que entrará na base, ou seja, na próxima iteração a_4 entrará na base no lugar de a_1 .

Segunda base:

$$I_B = \{2, 4\}, I_N = \{1, 3\},$$

$$B^{-1} = \begin{pmatrix} 4 & 0 \\ 2 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix},$$

$$u = c_B B^{-1} = (-5 \ 0) \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix} = \begin{pmatrix} -\frac{5}{4} & 0 \end{pmatrix},$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \end{pmatrix} = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix} \begin{pmatrix} 15 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{15}{4} \\ \frac{2}{4} \end{pmatrix} \geq 0,$$

e

$$\bar{z} = c_B B^{-1}b = c_B \bar{x}_B = (-5 \ 0) \begin{pmatrix} \frac{15}{4} \\ \frac{2}{4} \end{pmatrix} = -\frac{75}{4}.$$

Como $x_B \geq 0$, esta é uma solução ótima para o problema. Podemos então dizer que $x_1^* = 0$, $x_2^* = \frac{15}{4}$, $x_3^* = 0$, $x_4^* = \frac{2}{4}$, implicando $z^* = -\frac{75}{4}$, é a solução ótima do (PPL) modificado. ■

4.4.2 Alterando o vetor custo

Suponhamos agora que o vetor c , que define o custo associado às variáveis do problema (PPL), seja alterado para $c + \delta c$ e consideremos e que a base B deixa de ser dual viável para o problema modificado:

$$(\overline{PPL}) : \text{ maximizar } z = (c + \delta c)x$$

sujeito a:

$$Ax = b$$

$$x \geq 0,$$

ou seja, $(c_B + \delta c_B)B^{-1}N - (c_N + \delta c_N) \not\geq 0$.

Neste caso, para obter a solução ótima de (\overline{PPL}) , podemos aplicar o método do simplex, tomando B como base inicial.

Exemplo 4.6 Consideremos novamente o (PPL) do exemplo 4.2. Suponhamos agora que o custo associado à variável x_2 seja alterado de -5 para -1.

Verificamos no exemplo 4.4 que a base ótima do problema original deixa de ser dual viável com esta alteração, como comprovamos abaixo.

$$\begin{aligned} & (c_2 \ c_1)B^{-1}N - (c_4 \ c_3) = \\ & (-1 \ -4) \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} - (0 \ 0) = \\ & (-1 \ -4) \begin{pmatrix} \frac{1}{10} & -\frac{3}{10} \\ -\frac{4}{10} & \frac{2}{10} \end{pmatrix} - (0 \ 0) = \begin{pmatrix} \frac{15}{10} & -\frac{5}{10} \end{pmatrix} \not\geq 0. \end{aligned}$$

Por outro lado, sabemos que B não deixa de ser primal viável. Sendo assim, aplicamos a seguir o método do simplex para resolver o problema modificado, tomando a base B como base inicial.

Base inicial:

$$I_B = \{2, 1\}, \ I_N = \{4, 3\},$$

Como visto acima, $z_4 - c_4 = \frac{15}{10} > 0$ e $z_3 - c_3 = -\frac{5}{10} < 0$, logo a coluna a_3 entrará na base na próxima iteração do algoritmo.

$$y_3 = B^{-1}a_3 = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{3}{10} \\ \frac{2}{10} \end{pmatrix} = \begin{pmatrix} y_{13} \\ y_{23} \end{pmatrix},$$

logo $L_1 = \{2\}$. Como L_1 só tem um elemento, este elemento determinará a coluna que sairá da base, ou seja, $a_{B(2)} = a_1$ sairá da base, sendo substituída pela coluna a_3 .

Segunda base:

$$I_B = \{2, 3\}, \quad I_N = \{4, 1\},$$

$$B = (a_2 \ a_3) = \begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix} \text{ logo } B^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix},$$

$$c_B = (-1 \ 0), \quad u = c_B B^{-1} = (-1 \ 0) \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix} = \left(0 \ -\frac{1}{2}\right),$$

$$\bar{x} = B^{-1}b = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ 9 \end{pmatrix},$$

$$\bar{z} = ub = \left(0 \ -\frac{1}{2}\right) \begin{pmatrix} 5 \\ 7 \end{pmatrix} = -\frac{7}{2},$$

$$z_1 = ua_1 = \left(0 \ -\frac{1}{2}\right) \begin{pmatrix} 1 \\ 3 \end{pmatrix} = -\frac{3}{2} \Rightarrow z_1 - c_1 = -\frac{3}{2} - (-4) = \frac{5}{2} > 0,$$

$$z_4 = ua_4 = \left(0 \ -\frac{1}{2}\right) \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{2} \Rightarrow z_4 - c_4 = \frac{1}{2} - 0 = \frac{1}{2} > 0.$$

Como $z_j - c_j \geq 0, \forall j \in I_N$, esta solução básica é ótima. Logo $x_1 = 0, x_2 = \frac{7}{2}, x_3 = 9, x_4 = 0$ é uma solução ótima, fornecendo $z = -\frac{7}{2}$. ■

4.4.3 Acrescentando mais Restrições

Suponhamos agora que acrescentemos mais uma restrição ao (PPL) (4.33) da forma $sx \geq b_{m+1}$, onde $s = (s_1 \ s_2 \ \dots \ s_n)$ é um vetor linha dado e b_{m+1} um real também conhecido.

Sem inicializarmos novamente todo o processo de solução do (PPL) com mais uma restrição poderemos pensar em, a partir da solução ótima obtida para o (PPL) original, reotimizar o novo problema utilizando os resultados já obtidos.

Introduziremos a variável de folga $x_{n+1} \geq 0$ a essa nova restrição: $sx - x_{n+1} = b_{m+1}$ e o novo (PPL) será escrito:

$$(\overline{PPL}) : z = cx + 0x_{n+1}$$

sujeito a:

$$Ax - 0x_{n+1} = b$$

$$sx - x_{n+1} = b_{m+1}$$

$$x \geq 0, \quad x_{n+1} \geq 0,$$

em termos matriciais as restrições ficariam da seguinte forma:

$$\begin{pmatrix} A & 0 \\ s & -1 \end{pmatrix} \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} b \\ b_{m+1} \end{pmatrix}, \text{ lembrando que } A \in \mathcal{R}^{m \times n}.$$

Consideremos $s = (s_B \ s_N)$, onde s_B está associado a B . Tomemos a matriz

$$\bar{B} = \begin{pmatrix} B & 0 \\ s_B & -1 \end{pmatrix}.$$

Como B^{-1} existe então \bar{B}^{-1} também existe, pois $\det(\bar{B}) = \det(B) \times (-1)$.

Proposição 4.4 \bar{B} está associada a uma solução dual viável de (\overline{PPL}) .

Demonstração

Sabemos que

$$\bar{B}^{-1} = \begin{pmatrix} B^{-1} & 0 \\ s_B B^{-1} & -1 \end{pmatrix} \text{ e } \bar{u} = \bar{c}_B \bar{B}^{-1}, \text{ onde } \bar{c}_B = (c_B \ 0),$$

logo

$$\bar{c}_B \bar{B}^{-1} = (c_B \ 0) \begin{pmatrix} B^{-1} & 0 \\ s_B B^{-1} & -1 \end{pmatrix} = (c_B B^{-1} \ 0) = (u \ 0).$$

Teremos ainda

$$\bar{z}_j = \bar{u} \bar{a}_j = (u \ 0) \begin{pmatrix} a_j \\ s_j \end{pmatrix} = u a_j = z_j, \quad j = 1, 2, \dots, n$$

e

$$\bar{z}_{n+1} = \bar{u} \bar{a}_{n+1} = (u \ 0) \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0.$$

Assim sendo temos que $\bar{z}_j - c_j = z_j - c_j \geq 0$, $j = 1, 2, \dots, n$; pois a solução associada a B é ótima do (PPL) . Por outro lado $\bar{z}_{n+1} - c_{n+1} = 0 - 0 = 0$. Logo \bar{B} está associada a uma solução \bar{u} viável do dual de (\overline{PPL}) . ■

Exemplo 4.7 Voltemos ao (PPL) do exemplo 4.2, acrescentamos a este a restrição $4x_1 + 5x_2 \geq 20$ à qual associaremos a variável de folga $x_5 \geq 0$: $4x_1 + 5x_2 - x_5 = 20$.

Teremos que

$$\bar{A} = (\bar{a}_1 \ \bar{a}_2 \ \bar{a}_3 \ \bar{a}_4 \ \bar{a}_5) = \begin{pmatrix} 1 & 4 & -1 & 0 & 0 \\ 3 & 2 & 0 & -1 & 0 \\ 4 & 5 & 0 & 0 & -1 \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 5 \\ 7 \\ 20 \end{pmatrix},$$

$$\bar{c} = (-4 \ -5 \ 0 \ 0 \ 0).$$

Sabemos que $\bar{B} = (\bar{a}_2 \ \bar{a}_1 \ \bar{a}_5)$ nos fornece uma base dual viável e que

$$\bar{B} = \left(\begin{array}{cc|c} 4 & 1 & 0 \\ 2 & 3 & 0 \\ 5 & 4 & -1 \end{array} \right) = \left(\begin{array}{c|c} B & 0 \\ s_B & -1 \end{array} \right), \quad \bar{B}^{-1} = \left(\begin{array}{c|c} B^{-1} & 0 \\ s_B B^{-1} & -1 \end{array} \right),$$

e

$$B^{-1} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix}.$$

Então

$$s_B B^{-1} = (5 \ 4) \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} = \begin{pmatrix} 7 & 11 \\ 10 & 10 \end{pmatrix},$$

logo,

$$\bar{B}^{-1} = \left(\begin{array}{cc|c} \frac{3}{10} & -\frac{1}{10} & 0 \\ -\frac{2}{10} & \frac{4}{10} & 0 \\ \frac{7}{10} & \frac{11}{10} & -1 \end{array} \right),$$

verificamos que

$$\bar{x}_B = \bar{B}^{-1} \bar{b} = \bar{B}^{-1} \begin{pmatrix} 5 \\ 7 \\ 20 \end{pmatrix} = \begin{pmatrix} \frac{8}{10} \\ \frac{18}{10} \\ -\frac{88}{10} \end{pmatrix} \not\geq 0, \quad \bar{u} = \begin{pmatrix} -\frac{7}{10} & -\frac{11}{10} & 0 \end{pmatrix},$$

e

$$\bar{u} \bar{b} = (u \ 0) \begin{pmatrix} b \\ b_{n+1} \end{pmatrix} = ub = -\frac{112}{10}.$$

Como $\bar{x}_B(3) = \bar{x}_5 = -\frac{88}{10} < 0$ faremos a_5 deixar a base. Para escolhermos uma coluna não básica para entrar na base devemos calcular os y_{3j} , $j \in \{3, 4\} = I_N$. Para isso tomaremos a terceira linha de \bar{B}^{-1} e as colunas \bar{a}_3 e \bar{a}_4 , tal como segue:

$$y_{33} = \begin{pmatrix} \frac{7}{10} & \frac{11}{10} & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = -\frac{7}{10} \text{ e } y_{34} = \begin{pmatrix} \frac{7}{10} & \frac{11}{10} & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = -\frac{11}{10}.$$

Lembremos que $z_3 - c_3 = \frac{7}{10}$ e que $z_4 - c_4 = \frac{11}{10}$, calculemos agora

$$\min \left\{ \left| \frac{\frac{7}{10}}{-\frac{7}{10}} \right|, \left| \frac{\frac{11}{10}}{-\frac{11}{10}} \right| \right\},$$

há empate. Escolheremos \bar{a}_4 para entrar na base no lugar de \bar{a}_5 .

Nova base:

$$\bar{B} = (\bar{a}_2 \ \bar{a}_1 \ \bar{a}_4), \quad \bar{B}^{-1} = \begin{pmatrix} \frac{4}{11} & 0 & -\frac{1}{11} \\ -\frac{5}{11} & 0 & 0 \\ -\frac{7}{11} & -1 & \frac{10}{11} \end{pmatrix}$$

e

$$\bar{x}_B = \bar{B}^{-1}\bar{b} = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_1 \\ \bar{x}_4 \end{pmatrix} = \bar{B}^{-1} \begin{pmatrix} 5 \\ 7 \\ 20 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 8 \end{pmatrix} \geq 0.$$

A última base \bar{B} encontrada é primal e dual viável, assim sendo a solução $x_1 = 5$, $x_2 = 0$, $x_3 = 0$, $x_4 = 8$, $x_5 = 0$, fornecendo $z = -20$, é ótima para o (PPL) do exemplo da última seção ao qual foi acrescentado mais uma restrição.

É interessante observar que esta solução básica ótima é degenerada, pois a variável associada à coluna \bar{a}_2 tem valor nulo, isto é, $x_2 = 0$. A leitora ou o leitor poderia ter uma explicação para este fato?

Por curiosidade calculemos $\bar{z}_3 - \bar{c}_3$. Para isso observemos que $\bar{u} = \bar{c}_B \bar{B}^{-1} = (0 \ 0 \ -1)$ logo $\bar{z}_3 = \bar{u} \bar{a}_3 = 0$. Temos também que $\bar{c}_3 = 0$. Logo $\bar{z}_3 - \bar{c}_3 = 0$; caso façamos \bar{a}_3 entrar na base, usando a método primal do simplex, o valor de z não será modificado. Por que?

Para saber qual coluna sairá da base para dar entrada a coluna \bar{a}_3 , calcularemos o vetor

$$y_3 = \bar{B}^{-1}a_3 = \begin{pmatrix} y_{13} \\ y_{23} \\ y_{33} \end{pmatrix} = \begin{pmatrix} \frac{4}{11} & 0 & -\frac{1}{11} \\ -\frac{5}{11} & 0 & 0 \\ -\frac{7}{11} & -1 & \frac{10}{11} \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{4}{11} \\ \frac{5}{11} \\ \frac{7}{11} \end{pmatrix}.$$

Já temos que $\bar{x}_B(2) = \bar{x}_1 = 5$ e que $\bar{x}_B(3) = \bar{x}_4 = 8$. Consideraremos agora $\min \left\{ \frac{5}{\frac{5}{11}}, \frac{8}{\frac{7}{11}} \right\} = \frac{5}{\frac{5}{11}} = \frac{\bar{x}_1}{y_{23}}$. A coluna \bar{a}_1 deixará a base. A nova base será agora

$$\bar{B} = (\bar{a}_2 \ \bar{a}_3 \ \bar{a}_4), \quad \bar{B}^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{5} \\ -1 & 0 & \frac{1}{5} \\ 0 & -1 & \frac{2}{5} \end{pmatrix},$$

$$\bar{x}_B = \bar{B}^{-1}\bar{b} = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \bar{B}^{-1} \begin{pmatrix} 5 \\ 7 \\ 20 \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \\ 1 \end{pmatrix}.$$

Facilmente verificamos que esta última base também é dual viável. Obtemos $x_1 = 0$, $x_2 = 4$, $x_3 = 11$, $x_4 = 1$, $x_5 = 0$, fornecendo $z = -20$.

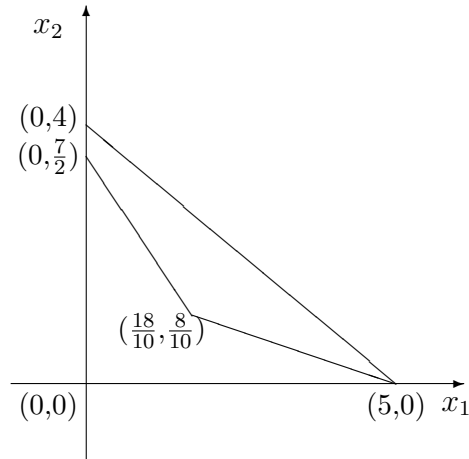


Figura 4.2: Pós-otimização

Todas as soluções ótimas do (\overline{PPL}) podem ser escritas da seguinte maneira:

$$x^* = \lambda \begin{pmatrix} 5 \\ 0 \\ 0 \\ 8 \\ 0 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} 0 \\ 4 \\ 11 \\ 1 \\ 0 \end{pmatrix}, \text{ para } \lambda \in [0, 1].$$

Ilustraremos todo o procedimento de pós-otimização e a busca de outra solução básica ótima na figura 4.2. Partimos do ponto $(\frac{18}{10}, \frac{8}{10})$ passamos para o ponto $(5, 0)$ (primal e dual viável) e, finalmente, atingimos o outro ponto ótimo $(0, 4)$. ■

4.5 Exercícios

1. Seja o problema de programação linear

$$(P) : \text{maximizar } x_0 = 5x_1 + 8x_2$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & + & x_2 & + & x_3 & & & = & 2 \\ x_1 & - & 2x_2 & & & + & x_4 & = & 0 \\ -x_1 & + & 4x_2 & & & & & + & x_5 = 1 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

Utilizando os métodos do primal e do dual do simplex, tratar dos seguintes itens.

- (a) Verificar que as colunas associadas às variáveis x_1 , x_2 e x_3 formam uma base ótima de (P) . Esta verificação deve ser feita invertendo a matriz básica fornecida e verificando sua primal e sua dual viabilidades.
- (b) Seja (\bar{P}) o problema de programação linear formado por (P) e pela restrição $x_1 + x_2 = 1$. Qual será a solução ótima de (\bar{P}) ? Utilizar o método dual do simplex para reotimizar.

2. Seja o problema de programação linear

$$(P) : \text{minimizar } z = 3x_1 + 2x_2$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & + & 3x_2 & - & x_3 & & & = & 9 \\ 5x_1 & + & 6x_2 & & & - & x_4 & = & 30 \\ x_1 & + & 2x_2 & & & & & - & x_5 = 8 \end{array}$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

Utilizando os métodos do primal e do dual do simplex, tratar dos seguintes itens.

- (a) Verificar que as colunas associadas às variáveis x_2 , x_3 e x_5 formam uma base ótima de (P) . Esta verificação deve ser feita invertendo a matriz básica fornecida e verificando sua primal e sua dual viabilidades.
- (b) Seja (\bar{P}) o problema de programação linear formado por (P) e pela restrição $3x_1 + x_2 \geq 6$. Qual será a solução ótima de (\bar{P}) ? Utilizar o método dual do simplex para reotimizar.
- (c) Verificar em que intervalo deve estar o custo associado à variável (i) x_1 , (ii) x_2 , para que a base ótima de (P) não seja alterada.

- (d) Se em (P) substituirmos na segunda restrição 30 por α , verificar que para $24 \leq \alpha < \infty$ a base ótima de (P) será sempre associada às variáveis x_2 , x_3 e x_5 .
- (e) Seja (\bar{P}) o problema de programação linear formado por (P) , ao se substituir o lado direito da segunda restrição por 20. Qual será a solução ótima de (\bar{P}) ? Utilizar o método dual do simplex para reotimizar.
- (f) Seja agora (\bar{P}) formado por (P) , ao se substituir o custo associado à variável x_2 por 10. Qual será a solução ótima de (\bar{P}) ? Utilizar o método do simplex para reotimizar.

3. Seja

$$(P) : \text{minimizar } z = 6x_1 + 9x_2 + 42x_3 + 36x_4$$

sujeito a:

$$\begin{aligned} x_1 & \quad + 3x_3 + 5x_4 \geq 2 \\ x_2 & + 4x_3 + 2x_4 \geq 3 \\ x_j & \geq 0, \quad j = 1, 2, 3, 4. \end{aligned}$$

Escrever (D) o problema dual de (P) . Resolver graficamente (D) . A partir da solução ótima de (D) encontrar a solução ótima de (P) utilizando as relações das folgas complementares.

- 4. Estudar os valores de α, β e λ pertencendo ao conjunto dos números reais para que o problema maximizar $z = \alpha x_1 + \beta x_2$, sujeito a $x_1 - x_2 = \lambda$, $x_1 \geq 0$, $x_2 \geq 0$, possua uma solução ótima limitada. Neste caso fornecer uma solução ótima quando $\lambda \geq 0$ e uma outra quando $\lambda \leq 0$.
- 5. Supondo que para qualquer sistema de desigualdades lineares, sob forma matricial, $My \leq d$ exista um método para encontrar um \bar{y} , tal que $M\bar{y} \leq d$; como usar este método para resolver o problema: maximizar $z = cx$, sujeito a, $Ax \leq b$, $x \geq 0$? Isto é, como seriam representados M e d em função de c , A e b ?

Capítulo 5

Obtenção de uma Solução Viável

Na seção 3.3 fizemos uma apresentação de um método para encontrar uma solução básica inicial, conhecido também como método das duas fases. Neste capítulo faremos um estudo mais geral.

Começemos tratando do problema de encontrar uma solução viável de um sistema de desigualdades lineares.

Seja $X = \{x \in \mathcal{R}^n \mid Ax \leq b\}$, onde $A \in \mathcal{R}^{m \times n}$ e $b \in \mathcal{R}^m$. Desejamos determinar um vetor $\bar{x} \in X$ ou ainda tal que $A\bar{x} \leq b$. Quando $b \geq 0$, $\bar{x} = 0$ é uma solução. Lembremos que X pode ser vazio.

5.1 Método de Eliminação de Fourier

Passaremos a desenvolver o método de eliminação de Fourier [Fou 890] para determinar soluções de sistemas de desigualdades lineares.

Consideremos $\sum_{j=1}^n a_{ij}x_j \leq b_i$ a i -ésima desigualdade de $Ax \leq b$, suponhamos que $a_{i1} \neq 0$, assim podemos escrever: $a_{i1}x_1 \leq b_i - \sum_{j=2}^n a_{ij}x_j$.

Se $a_{i1} > 0$ então

$$x_1 \leq \frac{b_i}{a_{i1}} - \sum_{j=2}^n \frac{a_{ij}}{a_{i1}}x_j;$$

e se $a_{i1} < 0$ então

$$x_1 \geq \frac{b_i}{a_{i1}} - \sum_{j=2}^n \frac{a_{ij}}{a_{i1}}x_j.$$

O conjunto dos índices das linhas $I = \{1, 2, 3, \dots, m\}$ pode ser particionado em I_0 , I_1 e I_2 da seguinte maneira:

$$I_0 = \{i \in I \mid a_{i1} = 0\}, \quad I_1 = \{i \in I \mid a_{i1} > 0\} \text{ e } I_2 = \{i \in I \mid a_{i1} < 0\}.$$

Desta maneira $Ax \leq b$ será escrito sob a seguinte forma:

$$\begin{aligned} x_1 &\leq \frac{b_i}{a_{i1}} - \sum_{j=2}^n \frac{a_{ij}}{a_{i1}} x_j, & i \in I_1, \\ \frac{b_i}{a_{i1}} - \sum_{j=2}^n \frac{a_{ij}}{a_{i1}} x_j &\leq x_1, & i \in I_2, \end{aligned}$$

e

$$\sum_{j=2}^n a_{ij} x_j \leq b_i, \quad i \in I_0.$$

Primeiramente consideremos $I_1 \neq \emptyset$ e $I_2 \neq \emptyset$, então para eliminarmos x_1 teremos que considerar o novo sistemas de desigualdades:

$$\frac{b_k}{a_{k1}} - \sum_{j=2}^n \frac{a_{kj}}{a_{k1}} x_j \leq \frac{b_l}{a_{l1}} - \sum_{j=2}^n \frac{a_{lj}}{a_{l1}} x_j, \quad k \in I_2, \quad l \in I_1$$

e

$$\sum_{j=2}^n a_{ij} x_j \leq b_i, \quad i \in I_0.$$

No sistema de desigualdades acima, onde eliminamos x_1 , o número de desigualdades é igual a $|I_1| \times |I_2| + |I_0|$. O conjunto das soluções que satisfazem este sistema é a projeção de X no espaço $x_2 \times x_3 \times \dots \times x_n$. Poderíamos então repetir o procedimento visando à eliminação de x_2 no espaço projetado. Continuando faríamos o mesmo para x_3 e sucessivamente até obtermos um sistema de desigualdades lineares com uma única variável x_n , que será da forma $\alpha \leq x_n \leq \beta$, onde α pode ser igual a $-\infty$ e β a $+\infty$. Caso $\alpha > \beta$, X é vazio.

Supondo o caso em que $X \neq \emptyset$, basta tomarmos um $\bar{x}_n \in [\alpha, \beta]$, este valor será dado a x_n no sistema contendo apenas as variáveis x_{n-1} e x_n , teremos analogamente $\alpha(\bar{x}_n) \leq x_{n-1} \leq \beta(\bar{x}_n)$ e tomaremos $x_{n-1} = \bar{x}_{n-1}$ tal que $\bar{x}_{n-1} \in [\alpha(\bar{x}_n), \beta(\bar{x}_n)]$. Passaremos a seguir com os valores \bar{x}_n e \bar{x}_{n-1} ao sistema de desigualdades lineares contendo somente as variáveis x_{n-2} , x_{n-1} e x_n e obteremos $\bar{x}_{n-2} \in [\alpha(\bar{x}_{n-1}, \bar{x}_n), \beta(\bar{x}_{n-1}, \bar{x}_n)]$. O procedimento é continuado até obtermos um valor de $x_1 = \bar{x}_1$ tal que $\bar{x}_1 \in [\alpha(\bar{x}_2, \bar{x}_3, \dots, \bar{x}_{n-1}, \bar{x}_n), \beta(\bar{x}_2, \bar{x}_3, \dots, \bar{x}_{n-1}, \bar{x}_n)]$.

5.2 Exemplos

Seja

$$X = \{(x_1 \ x_2)^T \mid 3x_1 + 2x_2 \geq 18, \ x_1 \leq 4, \ x_2 \leq 6, \ x_1 \geq 0, \ x_2 \geq 0\},$$

desejamos encontrar um $\bar{x} \in X$. Verificamos, facilmente, que $(0 \ 0)^T \notin X$. Lembrando que $3x_1 + 2x_2 \geq 18$ implica $x_1 \geq 6 - \frac{2}{3}x_2$, poderemos escrever o sistema de desigualdades lineares sob uma outra forma:

$$\begin{array}{rcl} & x_1 & \leq 4 \\ & 0 & \leq x_1 \\ 6 - \frac{2}{3}x_2 & \leq & x_1 \\ & x_2 & \leq 6 \\ 0 & \leq & x_2. \end{array}$$

Eliminando x_1 do sistema acima teremos que $0 \leq 4$, $6 - \frac{2}{3}x_2 \leq 4$, $x_2 \leq 6$ e $x_2 \geq 0$. Observamos que $0 \leq 4$ não traz nenhuma contradição, no entanto, poderíamos ter obtido um resultado do tipo $9 \leq 4$, isto é, $X = \phi$.

Como $6 - \frac{2}{3}x_2 \leq 4$ implica $x_2 \geq 3$, teremos:

$$\begin{array}{rcl} 3 & \leq & x_2 \\ 0 & \leq & x_2 \\ & x_2 & \leq 6, \end{array}$$

ou seja $3 \leq x_2 \leq 6$, basta tomarmos um $\bar{x}_2 \in [3, 6]$, por exemplo, $\bar{x}_2 = \frac{9}{2}$. Levando este valor no sistema com x_1 e x_2 , obteremos:

$$\begin{array}{rcl} & x_1 & \leq 4 \\ & 0 & \leq x_1 \\ 6 - \frac{2}{3} \times \frac{9}{2} & \leq & x_1, \end{array}$$

logo $3 \leq x_1 \leq 4$, isto é, $\alpha(\bar{x}_1) = 3$ e $\beta(\bar{x}_1) = 4$. Tomemos $\bar{x}_1 = \frac{7}{2}$. O ponto $(\frac{7}{2} \ \frac{9}{2})^T$ é uma solução de X . A figura 5.1 ilustra o exemplo.

Consideremos agora um outro exemplo, onde $X = \{x = (x_1 \ x_2)^T \mid 3x_1 + 2x_2 \geq 18, \ 2x_1 + x_2 \geq 10, \ x_1 + 8x_2 \geq 8\}$ e desejamos novamente encontrar um $\bar{x} \in X$. A origem $(0 \ 0)^T$ não pertence a X . Assim sendo:

$$\begin{array}{rcl} 3x_1 + 2x_2 \geq 18 & \Rightarrow & x_1 \geq 6 - \frac{2}{3}x_2, \\ 2x_1 + x_2 \geq 10 & \Rightarrow & x_1 \geq 5 - \frac{1}{2}x_2, \\ x_1 + 8x_2 \geq 8 & \Rightarrow & x_1 \geq 8 - 8x_2. \end{array}$$

logo, neste caso, $I_1 = \phi$ e $I_0 = \phi$, assim x_2 poderá tomar qualquer valor real. Tomemos $\bar{x}_2 = 2$ implicando que

$$\begin{array}{rcl} x_1 \geq 6 - \frac{2}{3} \times 2 & \Rightarrow & x_1 \geq \frac{14}{3}, \\ x_1 \geq 5 - \frac{1}{2} \times 2 & \Rightarrow & x_1 \geq 4, \\ x_1 \geq 8 - 8 \times 2 & \Rightarrow & x_1 \geq -8, \end{array}$$

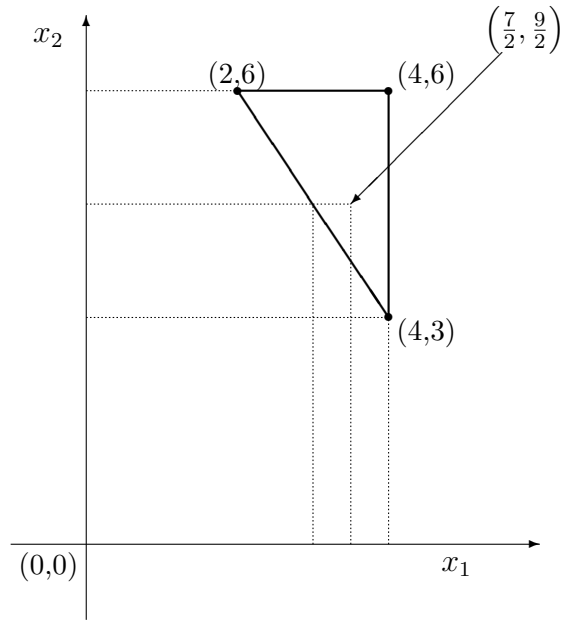


Figura 5.1: Obtenção de uma solução para $X = \{(x_1 \ x_2)^T \mid 3x_1 + 2x_2 \geq 18, \ x_1 \leq 4, \ x_2 \leq 6, \ x_1 \geq 0, \ x_2 \geq 0\}$ pelo método de eliminação de Fourier

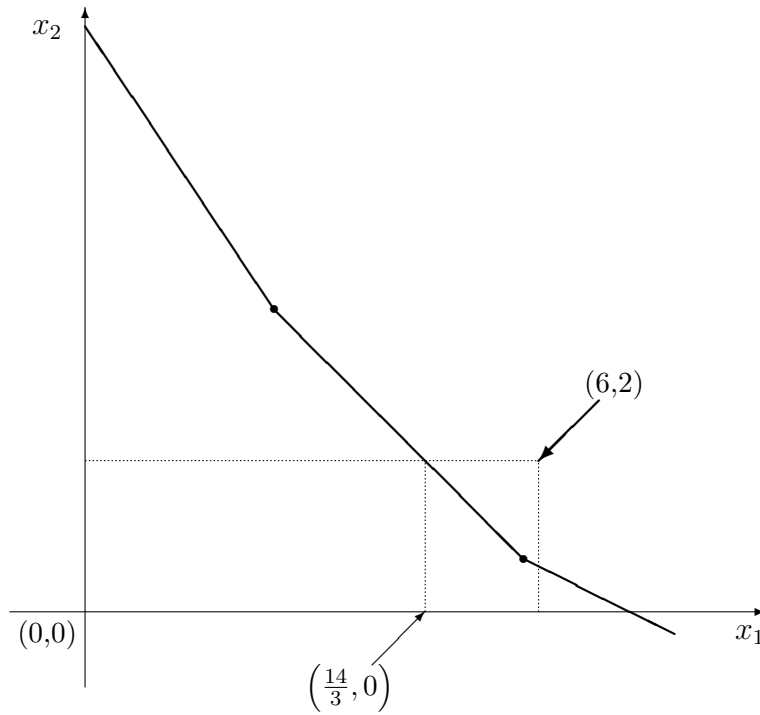


Figura 5.2: Obtenção de uma solução para $X = \{x = (x_1 \ x_2)^T \mid 3x_1 + 2x_2 \geq 18, \ 2x_1 + x_2 \geq 10, \ x_1 + 8x_2 \geq 8\}$ pelo método de eliminação de Fourier

logo $\alpha(\bar{x}_2) = \frac{14}{3}$ e $\beta(\bar{x}_2) = +\infty$; basta escolhermos $\bar{x}_1 \in \left[\frac{14}{3}, +\infty\right)$, tomemos $\bar{x}_1 = 6$. O ponto $(6 \ 2)^T \in X$. A figura 5.2 esclarece o procedimento.

Podemos também utilizar o método de eliminação de Fourier para resolvermos um problema de programação linear.

Seja

$$\begin{aligned} &\text{maximizar } z = 3x_1 + 5x_2 \\ &\text{sujeito a} \\ &3x_1 + 2x_2 \geq 18, \ x_1 \leq 4, \ x_2 \leq 6, \ x_1 \geq 0, \ x_2 \geq 0. \end{aligned}$$

Este problema é equivalente a

$$\begin{aligned} &\text{maximizar } z \\ &\text{sujeito a} \\ &z \leq 3x_1 + 5x_2, \ 3x_1 + 2x_2 \geq 18, \ x_1 \leq 4, \ x_2 \leq 6, \ x_1 \geq 0, \ x_2 \geq 0. \end{aligned}$$

Iniciaremos buscando uma solução viável $\bar{x} = (\bar{x}_1 \ \bar{x}_2 \ \bar{z})^T$ para as restrições do problema acima. Para isso começaremos eliminando x_1 :

$$\frac{1}{3}z - \frac{5}{3}x_2 \leq x_1, \ 6 - \frac{2}{3}x_2 \leq x_1, \ 0 \leq x_1, \ x_1 \leq 4, \ x_2 \leq 6, \ 0 \leq x_2.$$

Fornecendo

$$\frac{1}{5}z - \frac{12}{5} \leq x_2, \ 3 \leq x_2, \ 0 \leq x_2, \ x_2 \leq 6.$$

A variável x_2 será agora eliminada e ficaremos com um sistema só com a variável z :

$$\frac{1}{5}z - \frac{12}{5} \leq 6, \ 3 \leq 6, \ 0 \leq 6, \text{ logo } z \in (-\infty, 42].$$

Como desejamos o máximo de z , tomaremos $z = \bar{z} = 42$. Fazendo $z = 42$ na projeção do sistema no espaço $x_2 \times z$ teremos $6 \leq \bar{x}_2, \ 3 \leq \bar{x}_2, \ 0 \leq \bar{x}_2, \ \bar{x}_2 \leq 6$, assim sendo $6 \leq \bar{x}_2 \leq 6$, implicando $\bar{x}_2 = 6$.

Levando os valores $z = 42$ e $x_2 = 6$ nas restrições originais obteremos $4 \leq \bar{x}_1, \ 2 \leq \bar{x}_1, \ 0 \leq \bar{x}_1, \ \bar{x}_1 \leq 4$, implicando $4 \leq \bar{x}_1 \leq 4$, então $\bar{x}_1 = 4$. A solução $\bar{x}_1 = 4, \ \bar{x}_2 = 6$, fornecendo $\bar{z} = 42$ é ótima. Neste caso ela é única. Por que?

O conjunto das restrições é esboçado na figura 5.3.

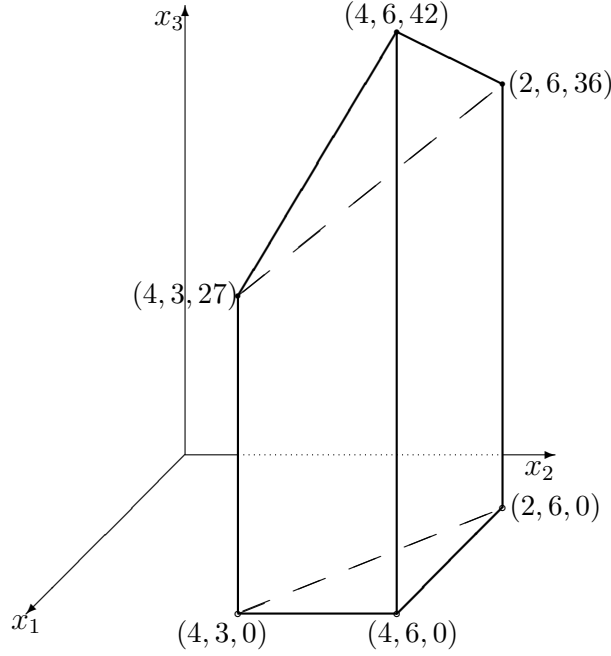


Figura 5.3: Aplicação do método de eliminação de Fourier para resolver um problema de programação linear

5.3 Aplicação do Método de Fourier quando houver Igualdades

Trataremos de determinar uma solução viável de $X = \{x \in \mathcal{R}^n \mid Ax = b, Dx \leq d, x \geq 0\}$, onde $A \in \mathcal{R}^{m \times n}$, $D \in \mathcal{R}^{p \times n}$, $d \in \mathcal{R}^p$, suporemos que o posto de A seja igual a m . Como o posto de A é m , poderemos tomar uma submatriz quadrada $B \in \mathcal{R}^{m \times m}$ de A inversível.

Particionaremos A e x da maneira usual: $A = (B \ N)$ e $x = (x_B \ x_N)^T$, assim escreveremos $Bx_B + Nx_N = b$ ou ainda

$$x_B = B^{-1}b - B^{-1}Nx_N. \quad (5.1)$$

A matriz D também será particionada, $D = (D_B \ D_N)$, o sistema $Dx \leq d$ será escrito:

$$D_Bx_B + D_Nx_N \leq d. \quad (5.2)$$

Levando o valor de x_B de (5.1) em (5.2):

$$D_BB^{-1}b - D_BB^{-1}Nx_N + D_Nx_N \leq d$$

ou ainda

$$(D_N - D_BB^{-1}N)x_N \leq d - D_BB^{-1}b. \quad (5.3)$$

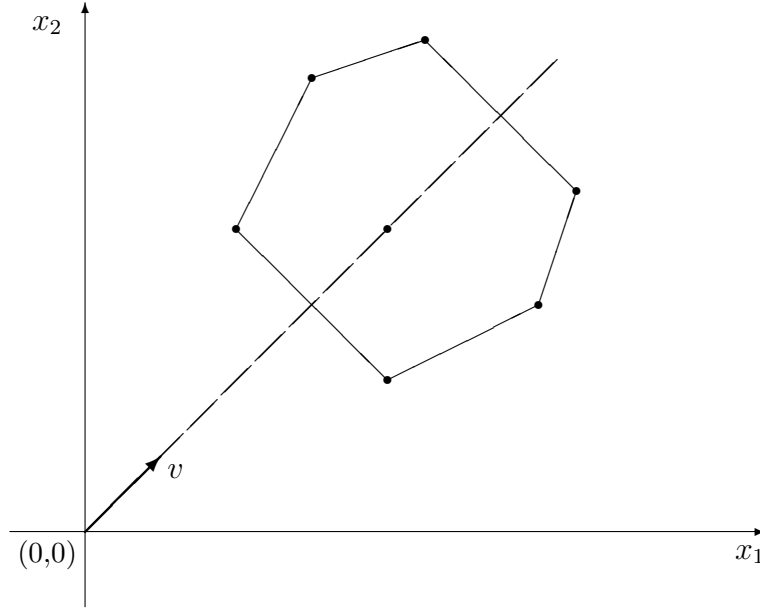


Figura 5.4: Método heurístico para obtenção de uma solução para $X = \{x \in \mathcal{R}^n \mid Ax \leq b\}$

Como $x \geq 0$, implicando $x_B \geq 0$ e $x_N \geq 0$, temos

$$B^{-1}b - B^{-1}Nx_N \geq 0 \quad (5.4)$$

e

$$x_N \geq 0. \quad (5.5)$$

Utilizaremos o método de eliminação de Fourier para encontrarmos uma solução viável de (5.3), (5.4) e (5.5). Uma vez encontrada uma solução $x_N = \bar{x}_N$, esta é levada em (5.1) para obtermos $\bar{x}_B = B^{-1}b - B^{-1}N\bar{x}_N$.

5.4 Um Método Heurístico

Consideremos novamente $X = \{x \in \mathcal{R}^n \mid Ax \leq b\}$ e queremos determinar um $\bar{x} \in X$. Suporemos que o interior (relativo) de X não seja vazio. A idéia é a de testar se um ponto $x(\lambda) = \lambda v$, onde v é um vetor dado e $\lambda \in \mathcal{R}$, pode para um certo λ pertencer a X . Em outras palavras, existe $\bar{\lambda} \in \mathcal{R}$ tal que $x(\bar{\lambda}) \in X$? Esta idéia é ilustrada na figura 5.4.

Buscaremos um $\bar{\lambda}$ tal que $Ax(\bar{\lambda}) \leq b$, isto é, $\bar{\lambda}Av \leq b$, que vem a ser a determinação de um λ que satisfaça a $\alpha \leq \lambda \leq \beta$. Se $\alpha > \beta$ então para a direção v não haverá λ tal que $A(\lambda v) \leq b$.

O método é de tentativas, começaremos com $v = e = (1 \ 1 \ \dots \ 1)^T$ e seja a_{ij} um elemento genérico da matriz A . O sistema $\lambda Ae \leq b$ ficaria

$$\lambda \sum_{j=1}^n a_{ij} \leq b_i, \quad i = 1, 2, \dots, m.$$

Denominemos $s_i = \sum_{j=1}^n a_{ij}$, $i = 1, 2, \dots, m$ e consideraremos as seguintes possibilidades:

- caso 1: $s_i = 0$ e $b_i < 0$, não existirá λ para a direção e ;
- caso 2: $s_i = 0$ e $b_i \geq 0$, qualquer λ satisfará à i -ésima desigualdade;
- caso 3: $s_i > 0$, $\lambda \leq \frac{b_i}{s_i}$ e
- caso 4: $s_i < 0$, $\lambda \geq \frac{b_i}{s_i}$.

Se não acontecer o caso 1, tomaremos:

$$\alpha = \max_{i \text{ tal que } s_i < 0} \left\{ \frac{b_i}{s_i} \right\} \quad \beta = \min_{i \text{ tal que } s_i > 0} \left\{ \frac{b_i}{s_i} \right\},$$

logo $\alpha \leq \lambda \leq \beta$.

Observação: se $\alpha > \beta$ não haverá λ para a direção e , caso contrário poderemos, por exemplo, tomar $\bar{\lambda} = \frac{\alpha + \beta}{2}$ e $x(\bar{\lambda}) = \bar{\lambda}e$ será uma solução. A seguir apresentaremos dois exemplos utilizando a direção

$$e = (1 \ 1 \ \dots \ 1)^T.$$

Primeiro exemplo:

Seja $X = \{x = (x_1 \ x_2)^T \mid 3x_1 + 2x_2 \geq 18, \ x_1 \geq 0, \ x_2 \geq 0\}$, teremos que $3x_1 + 2x_2 \geq 18 \Rightarrow (3 + 2)\lambda \geq 18 \Rightarrow \lambda \geq \frac{18}{5}$; $x_1 \geq 0 \Rightarrow \lambda \geq 0$ e $x_2 \geq 0 \Rightarrow \lambda \geq 0$. Basta tomarmos $\lambda \in [\frac{18}{5}, +\infty)$, por exemplo, $\bar{\lambda} = 4$, fornecendo um ponto viável $\bar{x}_1 = \bar{x}_2 = 4$ para X .

Segundo exemplo:

Seja $X = \{x = (x_1 \ x_2 \ x_3)^T \mid 6x_1 + 8x_2 + 2x_3 = 24, \ x_1 + x_2 \geq 0, \ x_1 \geq 0, \ x_2 \geq 0, \ x_3 \geq 0\}$, teremos que $6x_1 + 8x_2 + 2x_3 = 24 \Rightarrow x_3 = 12 - 3x_1 - 4x_2 \geq 0 \Rightarrow 3x_1 + 4x_2 \leq 12$, assim sendo $(3 + 4)\lambda \leq 12 \Rightarrow \lambda \leq \frac{12}{7}$; $x_1 + x_2 \geq 0 \Rightarrow 2\lambda \geq 0 \Rightarrow \lambda \geq 0$, as outras duas restrições implicam também que $\lambda \geq 0$. Tomemos um $\lambda \in [0, \frac{12}{7}]$, por exemplo, $\bar{\lambda} = 1$, logo $\bar{x}_1 = \bar{x}_2 = 1$ e $\bar{x}_3 = 12 - 3 - 4 = 5$ é um ponto de X .

Este método heurístico ingênuo apresentado nem sempre consegue determinar uma solução viável de X , no entanto, pela sua simplicidade poderá ser utilizado em uma primeira etapa. Caso não se encontre uma solução viável de X após testar algumas direções v , passar-se-á para um outro método.

5.5 Métodos Introduzindo Soluções Artificiais

Como foi dito no início deste capítulo, a busca de uma solução viável para um sistema de desigualdades e igualdades lineares introduzindo uma solução artificial foi descrita rapidamente na seção 3.3. Nesta seção apresentaremos duas maneiras de encontrar soluções viáveis de sistemas lineares.

Primeiro Método

Seja $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\}$, tomemos um vetor não negativo $\bar{x} \geq 0$, definimos $\bar{b} = b - A\bar{x}$ e seja $\lambda \in \mathcal{R}$. Consideraremos o seguinte problema de otimização em λ e x :

$$\text{minimizar } \lambda \tag{5.6}$$

sujeito a:

$$Ax + \bar{b}\lambda = b, \tag{5.7}$$

$$x \geq 0 \text{ e } \lambda \geq 0. \tag{5.8}$$

Verificamos facilmente que $x = \bar{x}$ e $\lambda = 1$ é uma solução de (5.7) e (5.8). Se o mínimo de λ em (5.6), (5.7) e (5.8) for igual a zero para $x = x^*$ então $x^* \in X$. Caso contrário, isto é, $\lambda > 0$ no ótimo, teremos $X = \emptyset$.

Como foi explicado no capítulo 2, poder-se-á, a partir de uma solução inicial viável de (5.7) e (5.8), buscar uma solução básica de (5.7) satisfazendo (5.8) sem aumentar o valor de λ , visando à utilização do método do simplex.

Este tipo de enfoque que acabamos de apresentar pode ser útil nos métodos de pontos interiores para a solução de problemas de programação linear.

Segundo Método

Podemos também acrescentar uma variável $v_i \geq 0$ artificial a cada equação do sistema $Ax = b$, isto é, consideraremos as seguintes equações $\sum_{j=1}^n a_{ij}x_j + v_i = b_i$, $i = 1, 2, \dots, m$. Suporemos, sem perda de generalidade, que $b \geq 0$.

O problema de otimização a ser considerado terá a seguinte forma:

$$\text{minimizar } e^T v \tag{5.9}$$

sujeito a:

$$v + Ax = b, \tag{5.10}$$

$$x \geq 0 \text{ e } v \geq 0, \tag{5.11}$$

onde $e^T = (1 \ 1 \ \dots \ 1)$ e $v^T = (v_1 \ v_2 \ \dots \ v_m)$.

A solução $v = b$ e $x = 0$ é básica de (5.10) satisfazendo (5.11). Se a solução ótima de (5.9), (5.10) e (5.11) fornecer $e^T v = 0$, que corresponde a $v = 0$ e $x = \bar{x}$. Desejamos que \bar{x} seja uma solução básica de $Ax = b$.

Utilizando o algoritmo primal de simplex para solucionar (5.9), (5.10) e (5.11), dois casos serão considerados no ótimo:

- $v = 0$ está associado a colunas não básicas de (5.10), então \bar{x} é uma solução básica de $Ax = b$;
- existe $v_i = 0$ associada a uma coluna básica de (5.10), se houver um $y_{ij} \neq 0$, para $j = 1, 2, \dots, n$, basta colocarmos na base a coluna a_j no lugar da coluna e_i associada à variável v_i ; se $y_{ij} = 0, j = 1, 2, \dots, n$, eliminaremos a linha i .

No caso em que $e^T v > 0$ no ótimo então $X = \phi$.

5.6 Viabilidade de Sistemas com Igualdades e Desigualdades Lineares

Consideraremos agora $X = \{x \in \mathcal{R}^n \mid A_1 x \leq b_1, A_2 x = b_2\}$ e desejamos saber se $X \neq \phi$, onde $A_1 \in \mathcal{R}^{p \times n}$ e $A_2 \in \mathcal{R}^{q \times n}$.

Sejam $0 \leq w_1^T \in \mathcal{R}^p$, $w_2^T \in \mathcal{R}^q$. Poderemos escrever $w_1 A_1 x \leq w_1 b_1$ e $w_2 A_2 x = w_2 b_2$. Consideraremos então:

$$(w_1 A_1 + w_2 A_2)x \leq w_1 b_1 + w_2 b_2. \quad (5.12)$$

É fácil verificar que se existirem $w_1 \geq 0$ e w_2 tais que $w_1 A_1 + w_2 A_2 = 0$ e $w_1 b_1 + w_2 b_2 < 0$ então $X = \phi$.

Tomemos um exemplo para ilustrar o resultado acima, onde $X = \{x = (x_1 \ x_2 \ x_3)^T \mid 3x_1 + 2x_2 + 4x_3 \leq 5, x_1 + 2x_2 + x_3 \leq 4, -x_1 \leq 0, -x_2 \leq 0, -x_3 \leq 0, x_1 + x_2 + x_3 = 3\}$, e sejam $w_1 = (1 \ 3 \ 0 \ 2 \ 1)$ e $w_2 = -6$. Escreveremos para estas desigualdades e igualdade a expressão (5.12): $(1 \times 3 + 3 \times 1 - 6 \times 1)x_1 + (1 \times 2 + 3 \times 2 - 2 \times 1 - 6 \times 1)x_2 + (1 \times 4 + 3 \times 1 - 1 \times 1 - 6 \times 1)x_3 \leq (1 \times 5 + 3 \times 4 - 6 \times 3)$, isto é, $(0 \ 0 \ 0)(x_1 \ x_2 \ x_3)^T \leq -1 \Rightarrow 0 \leq -1$ (!), logo $X = \phi$.

Definição 5.1 Segundo H. W. Kuhn [Ku 56], o sistema $A_1 x \leq b_1$, $A_2 x = b_2$ é inconsistente se há $w_1 \geq 0$ e w_2 tais que $w_1 A_1 + w_2 A_2 = 0$ e $w_1 b_1 + w_2 b_2 < 0$.

Kuhn [Ku 56] enunciou o teorema seguinte.

Teorema 5.1 *Um sistema de igualdades e desigualdades lineares será vazio se e somente se for inconsistente.*

Demonstração

(\Rightarrow) Já demonstramos acima que um sistema inconsistente é vazio.

(\Leftarrow) Consideremos novamente $A_1x \leq b_1$, $A_2x = b_2$ e suporemos $b_2 \geq 0$. Particionaremos $b_1 = (b_\alpha \ b_\beta)^T$, onde $b_\alpha \geq 0$ e $b_\beta < 0$. Particionaremos também as linhas de $A_1 = \begin{pmatrix} A_\alpha \\ A_\beta \end{pmatrix}$ para que possamos construir o seguinte problema de programação linear visando a encontrar uma solução viável para o sistema de igualdades e desigualdades lineares em questão:

(PA) : maximizar $-(\xi_\alpha^T v_\alpha + \xi_\beta^T v_\beta + \xi_2^T v_2)$
sujeito a:

$$A_\alpha x + v_\alpha \leq b_\alpha,$$

$$A_\beta x - v_\beta \leq b_\beta,$$

$$A_2 x + v_2 = b_2,$$

$$v_\alpha \geq 0, \ v_\beta \geq 0, \ v_2 \geq 0,$$

onde $\xi_k = (1 \ 1 \ \dots \ 1)^T$, pertencendo ao mesmo espaço que b_k , para $k = \alpha, \beta, 2$.

Como a função objetivo de (PA) é limitada para o máximo, pois $-(\xi_\alpha^T v_\alpha + \xi_\beta^T v_\beta + \xi_2^T v_2) \leq 0$ e $x = 0$, $v_\alpha = b_\alpha$, $v_\beta = -b_\beta$ e $v_2 = b_2$ é uma solução viável de (PA), então teremos que o ótimo de (PA) será limitado.

Denominaremos $w_1 = (w_\alpha \ w_\beta) \geq 0$ e w_2 os vetores cujas componentes são as variáveis do seguinte dual de (PA) :

(DA) : minimizar $w_\alpha b_\alpha + w_\beta b_\beta + w_2 b_2$
sujeito a:

$$w_\alpha A_\alpha + w_\beta A_\beta + w_2 A_2 = 0,$$

$$w_\alpha \geq -\xi_\alpha^T, \ -w_\beta \geq -\xi_\beta^T, \ w_2 \geq -\xi_2^T,$$

$$w_\alpha \geq 0, \ w_\beta \geq 0.$$

Pelas propriedades da dualidade sabemos $\text{val}(PA) = \text{val}(DA)$, pois (PA) não é vazio e possui ótimo limitado. Denominamos $\text{val}(\cdot)$ o valor da função objetivo no ótimo de (\cdot).

Sendo o sistema $A_1x \leq b_1$, $A_2x = b_2$ vazio por hipótese, sabemos que $\text{val}(PA) < 0$ e assim $\text{val}(PA) = \text{val}(DA) < 0$, implicando que

$$w_\alpha b_\alpha + w_\beta b_\beta + w_2 b_2 = w_1 b_1 + w_2 b_2 < 0, \text{ onde } w_1 = (w_\alpha \ w_\beta) \geq 0,$$

e

$$w_\alpha A_\alpha + w_\beta A_\beta + w_2 A_2 = w_1 A_1 + w_2 A_2 = 0,$$

logo o sistema é inconsistente. ■

5.7 Exercícios

1. Teorema de Farkas [Fa 02]. Demonstrar que o conjunto $\{(x, w) \in \mathcal{R}^n \times \mathcal{R}^m \mid Ax \leq 0, cx > 0, w^T A = c, w \geq 0\}$ é vazio, onde $A \in \mathcal{R}^{m \times n}$ e $c^T \in \mathcal{R}^n$ são dados.
2. Resolver pelo método do simplex partindo de uma solução artificial: maximizar $z = 3x_1 + 5x_2$, sujeito a: $x_1 \leq 4, x_2 \leq 6, 3x_1 + 2x_2 \geq 18, x_1 \geq 0, x_2 \geq 0$.
3. Solucionar o mesmo problema acima utilizando o teorema das folgas complementares.

Capítulo 6

Conjuntos Poliédricos Convexos

A definição de um conjunto convexo foi feita nos capítulos anteriores. A seguir daremos as definições de hiperplano, de vértice, de raios e direções, assim como algumas de suas propriedades inerentes. Introduziremos também o estudo dos conjuntos poliédricos convexos.

6.1 Hiperplano, Vértices, Raios e Direções

Definição 6.1 *Hiperplano*

Sejam $p \in \mathcal{R}^n$, tal que $p \neq 0$ e $\alpha \in \mathcal{R}$ dados. O conjunto $H = \{x \in \mathcal{R}^n \mid p^T x = \alpha\}$ é denominado um hiperplano.

O vetor p é ortogonal ao hiperplano H , pois consideremos $x^0 \in H$, isto é, $p^T x^0 = \alpha$ e para todo $x \in H$ teremos também $p^T x = \alpha$, logo $p^T (x - x^0) = 0$. A figura 6.1 ilustra esta propriedade.

Proposição 6.1 *Um hiperplano é um conjunto convexo.*

Demonstração

Sejam x^1 e $x^2 \in H$, isto é: $p^T x^1 = \alpha$ e $p^T x^2 = \alpha$; tomemos $\lambda \in [0, 1]$ e consideremos $\lambda p^T x^1 = \lambda \alpha$ e $(1 - \lambda) p^T x^2 = (1 - \lambda) \alpha$ ou ainda $p^T (\lambda x^1) + p^T [(1 - \lambda) x^2] = \lambda \alpha + (1 - \lambda) \alpha$ que implica em $p^T [\lambda x^1 + (1 - \lambda) x^2] = \alpha$. Assim sendo, qualquer combinação convexa de dois pontos de H também pertence a H . ■

Um hiperplano divide \mathcal{R}^n em duas regiões, denominadas de semi-espacos.

Um semi-espaco é um conjunto $\{x \in \mathcal{R}^n \mid p^T x \leq \alpha\}$, quando $p \neq 0$. Na realidade podemos considerar um semi espaco fechado $\{x \in \mathcal{R}^n \mid p^T x \leq \alpha\}$ ou aberto $\{x \in \mathcal{R}^n \mid p^T x < \alpha\}$.

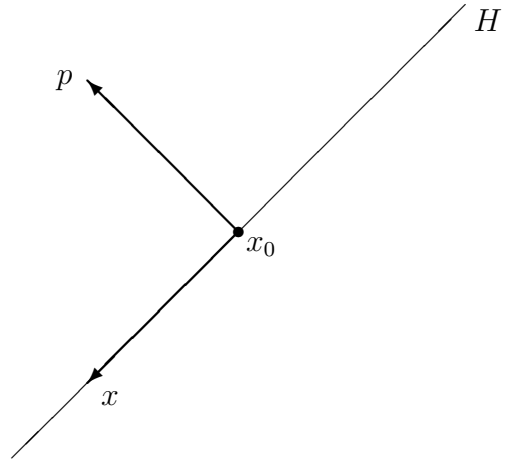


Figura 6.1: O vetor p é ortogonal a $x - x_0$

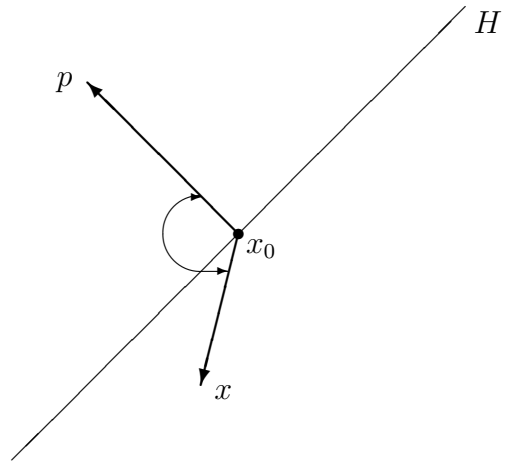


Figura 6.2: O ângulo entre p e x é maior do que 90°

Tomemos $x^0 \in \{x \in \mathcal{R}^n \mid p^T x = \alpha\}$ e $x \in \{x \in \mathcal{R}^n \mid p^T x \leq \alpha\}$, ou seja $p^T x^0 = \alpha$ e $p^T x \leq \alpha$, logo $p^T x \leq p^T x^0$ implicando $p^T(x - x^0) \leq 0$ como ilustramos na figura 6.2.

Definição 6.2 *Vértice ou ponto extremo de um conjunto*

Seja $X \subseteq \mathcal{R}^n$ um conjunto, $v \in X$ é um vértice ou um ponto extremo de X se não existirem x^1 e $x^2 \in X$, $x^1 \neq x^2$, tais que $v = \frac{1}{2}x^1 + \frac{1}{2}x^2$.

Na figura 6.3, dois esboços ilustram a definição de vértice.

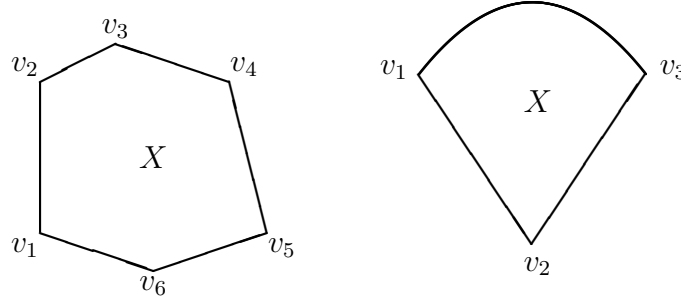


Figura 6.3: Vértices: No conjunto a direita, além de v_2 , todos os pontos no arco v_1v_3 são vértices

Definição 6.3 Raios e direções

Um raio é um conjunto da forma $\{x \in \mathcal{R}^n \mid x = x^0 + \lambda d, \lambda \geq 0\}$, onde x^0 é um ponto dado e a direção $d \in \mathcal{R}^n$ também o é.

A figura 6.4 ilustra a definição de raio.

6.2 Conjunto Poliédrico Convexo

Consideremos $H_i = \{x \in \mathcal{R}^n \mid (p^i)^T x \leq \alpha_i\}$, $p^i \neq 0$ e α_i dados para $i = 1, 2, \dots, m$. $X = \bigcap_{i=1}^m H_i$ é um conjunto poliédrico convexo por definição se $X \neq \emptyset$.

Um conjunto poliédrico convexo é a interseção de um número finito de semi-espacos fechados. Esta interseção deve ser não vazia. Daremos na figura 6.5 alguns exemplos de conjuntos poliédricos convexos em \mathcal{R}^2 :

As restrições dos problemas de programação linear formam um conjunto poliédrico convexo. Por que?

Quando um conjunto poliédrico convexo for limitado, diremos que este conjunto é um hiper-poliedro ou um politopo.

É interessante verificarmos que no caso de um politopo X , caso conheçamos todos os seus vértices v^1, v^2, \dots, v^q , podemos expressar qualquer $x \in X$ como sendo uma combinação convexa dos vértices de X , isto é, $x \in X$ se e somente se, $x = \sum_{j=1}^q \lambda_j v^j$, $\sum_{j=1}^q \lambda_j = 1$ e $\lambda_j \geq 0$, $j = 1, 2, \dots, q$. Este resultado será demonstrado ainda neste capítulo.

Na figura 6.6 temos seis vértices e $x \in X$ poderá ser expresso como $x = \sum_{j=1}^6 \lambda_j v^j$, $\sum_{j=1}^6 \lambda_j = 1$ e $\lambda_j \geq 0$, $j = 1, 2, \dots, 6$.

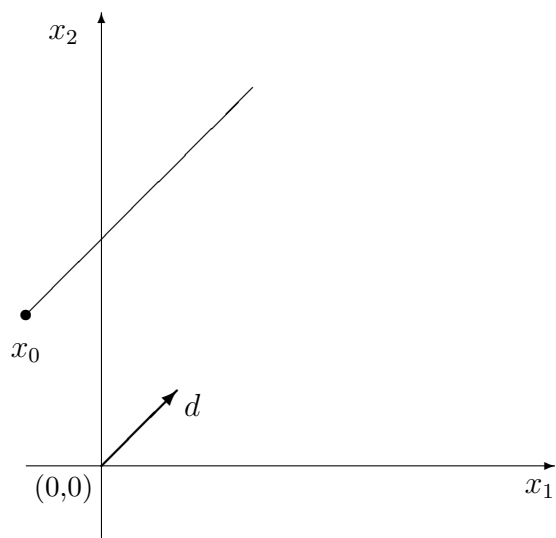


Figura 6.4: Raio de direção d e vértice x_0

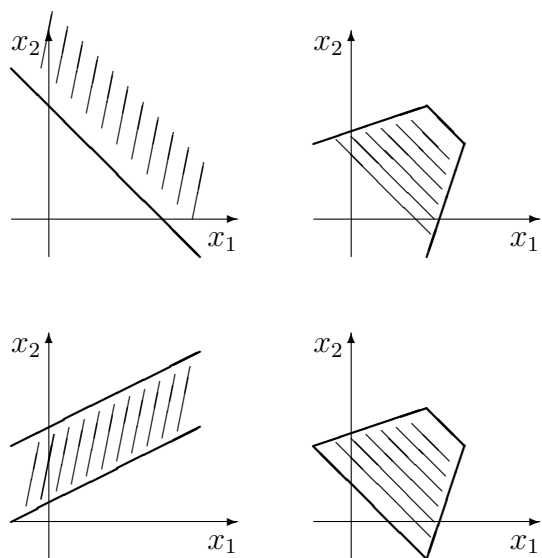


Figura 6.5: Conjuntos Poliédricos Convexos

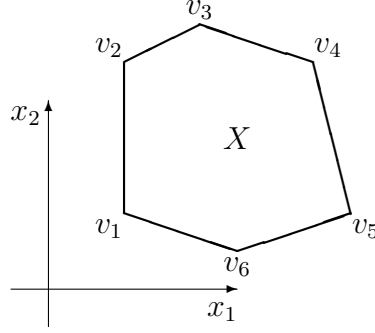


Figura 6.6: Combinação convexa dos vértices do politopo

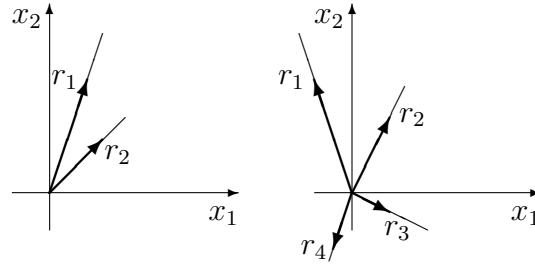


Figura 6.7: Cones poliédricos convexos

No caso particular de $\alpha_i = 0$, $i = 1, 2, \dots, m$ teremos

$$H_i = \{x \in \mathcal{R}^n \mid (p^i)^T x \leq 0\}, \quad p^i \neq 0, \quad C = \bigcap_{i=1}^m H_i.$$

O conjunto poliédrico convexo C é denominado de cone poliédrico convexo. Na figura 6.7 ilustramos cones poliédricos convexos em \mathcal{R}^2 e \mathcal{R}^3 .

Os vetores r^i ilustrados na figura são denominados raios extremos de C e serão definidos da seguinte maneira: r é um raio extremo de C se não existirem d^1 e $d^2 \in C$, $d^1 \neq \beta d^2$ para qualquer $\beta \in \mathcal{R}_+$ tais que $r = \frac{1}{2}d^1 + \frac{1}{2}d^2$.

Sendo $A \in \mathcal{R}^{m \times n}$ então $C = \{x \in \mathcal{R}^n \mid Ax \leq 0\}$, é um cone poliédrico convexo. Supondo que r^1, r^2, \dots, r^q sejam os raios extremos de C mostraremos mais à frente que para $\forall x \in C$ teremos $x = \sum_{i=1}^q \mu_i r^i$, $\mu_i \geq 0$, $i = 1, 2, \dots, q$.

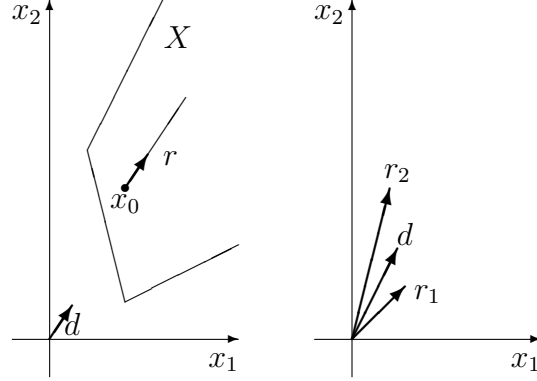


Figura 6.8: Conjunto poliédrico não limitado

Voltemos ao conjunto poliédrico convexo $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\}$, onde $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, e o posto de A igual a m .

Se X contiver um raio então X não será limitado. Para verificar isto tomaremos $x^0 \in X$ e $d \neq 0$ um vetor e consideraremos $r = x^0 + \lambda d$, $\lambda \geq 0$ um raio de X . Para que $r \in X$ teremos que $Ar = b$, $r \geq 0$, ou ainda que $A(x^0 + \lambda d) = b$ e $x^0 + \lambda d \geq 0$, para $\forall \lambda \geq 0$.

De $A(x^0 + \lambda d) = b$ teremos que $Ax^0 + \lambda Ad = b$, como $Ax^0 = b$ então $\lambda Ad = 0$, para $\forall \lambda \geq 0$, isto é, $Ad = 0$. Como $x^0 + \lambda d \geq 0$, $\forall \lambda \geq 0$, logo $d \geq 0$.

Definição 6.4 Diremos que $d \neq 0$, $d \geq 0$ é uma direção de $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\}$, se $Ad = 0$.

Se não existir $d \neq 0$ tal que $Ad = 0$ e $d \geq 0$ então X será um conjunto limitado.

Na figura 6.8 apresentamos um conjunto poliédrico X em \mathcal{R}^2 , não limitado e contendo o raio $r = x^0 + \lambda d$, $\lambda \geq 0$. Apresentamos também um cone poliédrico C associado às arestas de X não limitadas.

É importante notarmos que se d é uma direção de X então αd para $\alpha > 0$ também o será. De maneira geral, se d^1 e d^2 forem direções de X , então $d \neq 0$, $d = \alpha_1 d^1 + \alpha_2 d^2$, para $\alpha_1 \geq 0$ e $\alpha_2 \geq 0$ também o será.

Denominaremos de direções extremas de X as direções associadas às arestas não limitadas de X . Caso X seja limitado poderemos dizer que o conjunto das direções extremas de X é vazio.

Como uma direção extrema d de X é paralela a uma aresta não limitada de X , então não existem d^1 e d^2 direções de X , $d^1 \neq \beta d^2$ para todo $\beta > 0$, tal que $d = \frac{1}{2}d^1 + \frac{1}{2}d^2$. Isto é, o vetor d não poderá ser escrito como uma combinação estritamente convexa de d^1 e d^2 .

A seguir caracterizaremos as direções extremas de $X = \{x \in \mathcal{R}^n \mid Ax = b, x \geq 0\}$.

Suponhamos $A = (B \ N)$, onde $B \in \mathcal{R}^{m \times m}$ é inversível, o que é sempre possível, pois supusemos que o posto de A fosse igual a m . Consideremos d uma direção de X , isto é, $d \neq 0$, $Ad = 0$ e $d \geq 0$; tomemos d da seguinte forma:

$$d^T = (\underbrace{d_1 \ d_2 \ \dots \ d_k \ 0 \ 0 \ \dots \ 0}_m \ \underbrace{0 \ \dots \ 0 \ d_l \ 0 \ \dots \ 0}_{n-m}),$$

onde $d_j > 0$, $j = 1, 2, \dots, k$ e $d_l > 0$ para $l > m$.

Os m primeiros índices estão associados às colunas $a_1, a_2, \dots, a_k, \dots, a_m$ de B .

Verificaremos que se as colunas a_1, a_2, \dots, a_k não fossem linearmente independentes, a direção d poderia ser escrita como uma combinação estritamente convexa de duas outras direções de X .

Para realizarmos essa verificação, suporemos que as colunas a_1, a_2, \dots, a_k são linearmente dependentes. Neste caso, existem escalares nem todos nulos λ_j , $j = 1, 2, \dots, k$ tais que $\sum_{j=1}^k \lambda_j a_j = 0$. Como $d_j > 0$, $j = 1, 2, \dots, k$, é fácil notar que existe um $\alpha > 0$ tal que

$$d_j - \alpha \lambda_j > 0 \text{ e } d_j + \alpha \lambda_j > 0 \text{ para } j = 1, 2, \dots, k. \quad (6.1)$$

Construiremos os dois seguintes vetores:

$$d^1 = \begin{pmatrix} d_1 - \alpha \lambda_1 \\ d_2 - \alpha \lambda_2 \\ \vdots \\ d_k - \alpha \lambda_k \\ 0 \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ e } d^2 = \begin{pmatrix} d_1 + \alpha \lambda_1 \\ d_2 + \alpha \lambda_2 \\ \vdots \\ d_k + \alpha \lambda_k \\ 0 \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Por construção $d^1 \geq 0$ e $d^2 \geq 0$. Verificaremos a seguir que d^1 e d^2 são também direções de X :

$$\begin{aligned} Ad^1 &= \sum_{j=1}^k (d_j - \alpha \lambda_j) a_j + d_l a_l = Ad - \alpha \sum_{j=1}^k \lambda_j a_j \Rightarrow \\ Ad^1 &= Ad = 0, \text{ pois } \sum_{j=1}^k \lambda_j a_j = 0. \end{aligned}$$

De maneira análoga mostra-se que $Ad^2 = 0$.

Como nem todos os λ_j são nulos então $d^1 \neq d^2$ e também $d^1 \neq \beta d^2$ para $\beta > 0$, assim sendo d^1 e d^2 são direções distintas. Mas, neste caso, $d = \frac{1}{2}d^1 + \frac{1}{2}d^2$.

Assim sendo para que d possa ser uma direção extrema de X , as colunas a_1, a_2, \dots, a_k terão que ser linearmente independentes.

Podemos ainda escrever:

$$0 = Ad = \sum_{j=1}^m d_j a_j + d_l a_l = Bd_B + d_l a_l, \text{ onde } d_B^T = (\underbrace{d_1 \ d_2 \ \dots \ d_k}_m \ 0 \ \dots \ 0).$$

Teremos que $Bd_B = -d_l a_l$ ou ainda $d_B = -d_l B^{-1} a_l$ logo:

$$d = \begin{pmatrix} d_B \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -d_l B^{-1} a_l \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} = d_l \begin{pmatrix} -B^{-1} a_l \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

lembramos que $d_l > 0$. Assim sendo d será uma direção extrema de X e o vetor

$$\begin{pmatrix} -B^{-1} a_l \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

também o será.

Suporemos agora que uma direção \bar{d} de X possa ser representada da seguinte forma:

$$\bar{d}^T = (\underbrace{d_1 \ d_2 \ \dots \ d_k \ 0 \ 0 \ \dots \ 0}_m \ \underbrace{0 \ \dots \ 0 \ d_l \ 0 \ \dots \ 0 \ d_p \ 0 \ \dots \ 0}_{n-m}).$$

Aqui também suporemos que as m primeiras componentes de \bar{d} estejam associadas às colunas da matriz inversível B . Poderemos então escrever:

$$0 = A\bar{d} = \sum_{j=1}^m d_j a_j + d_l a_l + d_p a_p = b d_B + d_l a_l + d_p a_p,$$

implicando

$$B d_B = -d_l a_l - d_p a_p \quad \text{e} \quad d_B = -d_l B^{-1} a_l - d_p B^{-1} a_p.$$

Poderemos ainda representar o vetor \bar{d} como

$$\begin{pmatrix} d_B \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \\ d_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -d_l B^{-1} a_l - d_p B^{-1} a_p \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \\ d_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -d_l B^{-1} a_l \\ 0 \\ \vdots \\ 0 \\ d_l \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} -d_p B^{-1} a_p \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ d_p \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

assim sendo

$$\bar{d} = d_l \begin{pmatrix} -B^{-1} a_l \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + d_p \begin{pmatrix} -B^{-1} a_p \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = d_l \bar{d}^1 + d_p \bar{d}^2.$$

Calculemos $A\bar{d}^1$ e $A\bar{d}^2$, $A\bar{d}^1 = B(-B^{-1}a_l) + a_l = -a_l + a_l = 0$ e $A\bar{d}^2 = B(-B^{-1}a_p) + a_p = -a_p + a_p = 0$, logo \bar{d}^1 e \bar{d}^2 são também direções de X , implicando que \bar{d} seja uma combinação linear não negativa de \bar{d}^1 e \bar{d}^2 .

Analogamente poderemos considerar uma direção da forma de \bar{d} , onde haja mais de duas componentes positivas entre as $n - m$ última componentes de \bar{d} .

E quando todas as $n - m$ últimas componentes de \bar{d} forem nulas, isto é,

$$\bar{d}^T = (\underbrace{d_1 \ d_2 \ \dots \ d_k \ 0 \ 0 \ \dots \ 0}_m \ \underbrace{0 \ \dots \ 0}_{n-m}).$$

Verificaremos que $0 = A\bar{d} = Bd_B$ implicando $d_B = 0$ pois B é inversível. Logo $\bar{d} = 0$ não poderá ser uma direção de X .

Agora estamos aptos a caracterizar as arestas não limitadas dos conjuntos poliédricos convexos do tipo $X = \{x \in \mathcal{R}^n \mid Ax = b, \ x \geq 0\}$. Basta que durante a execução do método do simplex estejamos em uma solução básica B e exista um vetor da forma $y_s = -B^{-1}a_s \geq 0$, onde a_s é uma coluna de A . A solução básica nos fornece um vértice de X e a partir de y_s construiremos uma direção d , teremos desta maneira um raio que definirá uma aresta não limitada de X . Para cada base viável de X teremos no máximo $n - m$ direções extremas.

Na próxima seção caracterizaremos as arestas limitadas de X .

6.3 Caracterização das Arestas de um Conjunto Poliédrico Convexo

Quando apresentamos o método do simplex no capítulo 3, ilustramos que a trajetória para irmos de um vértice para o seguinte estava em uma aresta do conjunto poliédrico convexo do exemplo. Poderíamos pensar que uma trajetória passando pelo interior relativo do conjunto poliédrico convexo fosse possível também durante a execução do método do simplex. Verificaremos a seguir que, utilizando o método do simplex, nunca caminharemos pelo interior relativo do conjunto das restrições.

Consideremos novamente $X = \{x \in \mathcal{R}^n \mid Ax = b, \ x \geq 0\}$, tal que o posto de A seja igual ao seu número de linhas m . E que a matriz B formada pelas primeiras m colunas de A seja uma base primal viável de A . Seja \bar{x} a solução básica associada a B , isto é, $\bar{x}^T = (\bar{x}_B^T \ 0)$, onde $\bar{x}_B = B^{-1}b$. Tomemos uma coluna a_k de A não estando em B que substituirá uma coluna de B para

criar uma nova base. Escreveremos que

$$x^\alpha = \bar{x} + \begin{pmatrix} -B^{-1}a_k \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \alpha.$$

Esta escolha de a_k está condicionada a que $-B^{-1}a_k \not\geq 0$, assim sendo existirá $\bar{\alpha} > 0$ (caso não degenerado) associado à próxima solução básica \hat{x} , tal que $0 \leq \alpha \leq \bar{\alpha}$. Esta solução \hat{x} definirá um novo vértice de X . Ou ainda que

$$\hat{x} = \bar{x} + \begin{pmatrix} -B^{-1}a_k \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \bar{\alpha}.$$

Ou seja $x^\alpha = \lambda \bar{x} + (1 - \lambda) \hat{x}$, $0 \leq \lambda \leq 1$. Todo vetor da forma x^α , $0 \leq \alpha \leq 1$ estará sobre o segmento de reta entre \bar{x} e \hat{x} . Para mostrar que este segmento é uma aresta de X basta verificar que não existem x^1 e $x^2 \in X$ diferentes de βx^α ($\beta > 0$), para todo $0 \leq \alpha \leq 1$, tais que $x^\alpha = \frac{1}{2}x^1 + \frac{1}{2}x^2$.

Consideraremos outra vez $A = (B \ N)$ e suporemos que existam x^1 e $x^2 \in X$ diferentes de βx^α ($\beta > 0$), tais que $x^\alpha = \frac{1}{2}x^1 + \frac{1}{2}x^2$ e mostraremos a afirmação acima por absurdo.

Sabemos que $Ax^i = b$, $i = 1, 2$ e que $x^i \geq 0$, $i = 1, 2$; pois $x^i \in X$, $i = 1, 2$. Particionaremos $x^i = (x_B^i \ x_N^i)^T$, $i = 1, 2$ para que possamos escrever

$$Bx_B^i + Nx_N^i = b, \ i = 1, 2. \quad (6.2)$$

Verificamos facilmente que $x_N^i \neq 0$, $i = 1, 2$; pois caso não fosse assim teríamos em (6.2) $x_B^i = B^{-1}b$, $i = 1, 2$ implicando que $x^i = \bar{x}$, $i = 1, 2$, algo em contradição, pois $\bar{x} = x^\alpha$ para $\alpha = 0$.

O vetor x^α será também particionado: $x^\alpha = (x_B^\alpha \ x_N^\alpha)^T$.

Lembremos que x_N^α é um vetor com uma só componente diferente de zero, representando a k -ésima componente de x^α que é igual a α .

Passamos a considerar

$$x_B^\alpha = \frac{1}{2}x_B^1 + \frac{1}{2}x_B^2, \quad (6.3)$$

$$x_N^\alpha = \frac{1}{2}x_N^1 + \frac{1}{2}x_N^2. \quad (6.4)$$

Como $x_N^i \geq 0$ e $x_N^i \neq 0$, $i = 1, 2$, teremos em (6.4) que x_N^i , $i = 1, 2$ terá apenas uma componente diferente de zero, correspondendo à componente de x_N^α diferente de zero. Ou ainda podemos dizer que a k -ésima componente de x^i , $i = 1, 2$ será positiva e igual a x_k^i , $i = 1, 2$. Logo (6.2) será escrita

$$Bx_B^i + x_k^i a_k = b, \quad i = 1, 2. \quad (6.5)$$

De (6.5) temos que

$$x_B^i = B^{-1}b + (-B^{-1}a_k)x_k^i, \quad i = 1, 2. \quad (6.6)$$

para que $x_B^i \geq 0$, $i = 1, 2$, temos que ter $0 \leq x_k^i \leq \bar{\alpha}$. Assim sendo os pontos x^1 e x^2 estão no segmento definido por x^α , para $0 \leq \alpha \leq \bar{\alpha}$, contrariando a hipótese de que x^1 e x^2 sejam diferentes de x^α , para $0 \leq \alpha \leq \bar{\alpha}$.

Nesta seção verificamos que, realmente, o método do simplex em cada iteração caminha sobre uma aresta do conjunto poliédrico convexo que forma o conjunto de restrições do problema de programação linear.

Na seção que se segue apresentaremos uma versão do teorema principal da representação de um conjunto poliédrico convexo em função de seus vértices e de suas direções extremas.

6.4 Teorema da representação de um Conjunto Poliédrico Convexo

O teorema que será enunciado e demonstrado a seguir foi proposto por Minkowski [Min 11], ver também o capítulo 7 de [Sc 86].

Consideremos mais uma vez $X = \{x \in \mathcal{R}^n \mid Ax = b, \ x \geq 0\}$, $A \in \mathcal{R}^{m \times n}$ com posto igual a m . Suporemos que v^1, v^2, \dots, v^p sejam os vértices de X e que r^1, r^2, \dots, r^q as direções extremas de X .

Teorema 6.1 *Um ponto $x \in X$ se e somente se existirem $\lambda_j \geq 0$, $j = 1, 2, \dots, p$, tais que $\sum_{j=1}^p \lambda_j = 1$ e $\mu_i \geq 0$, $i = 1, 2, \dots, q$ para os quais*

$$x = \sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i.$$

Demonstração

Sabemos que $Av^j = b$ e $v^j \geq 0$, $j = 1, 2, \dots, p$ e que $Ar^i = 0$ e $r^i \geq 0$, $i = 1, 2, \dots, q$.

(\Rightarrow)

$$A(\sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i) = \sum_{j=1}^p \lambda_j (Av^j) + \sum_{i=1}^q \mu_i (Ar^i) = b(\sum_{j=1}^p \lambda_j) + 0 = b.$$

(\Leftarrow)

Tomemos $x \in X$ e desejamos verificar a existência de λ_j e μ_i tais que:

$$\sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i = x, \quad (6.7)$$

$$\sum_{j=1}^p \lambda_j = 1, \quad (6.8)$$

$$-\lambda_j \leq 0, \quad j = 1, 2, \dots, p, \quad (6.9)$$

$$-\mu_i \leq 0, \quad i = 1, 2, \dots, q. \quad (6.10)$$

Se o sistema de desigualdades e igualdades (6.7)-(6.10) for vazio existirão números u_s , $s = 1, 2, \dots, n, n+1$, $w_j \geq 0$, $j = 1, 2, \dots, p$ e $t_i \geq 0$, $i = 1, 2, \dots, q$, tais que (ver teorema 5.1):

$$\sum_{s=1}^n u_s v_s^j + u_{n+1} - w_j = 0, \quad j = 1, 2, \dots, p, \quad (6.11)$$

$$\sum_{s=1}^n u_s r_s^i - t_i = 0, \quad i = 1, 2, \dots, q, \quad (6.12)$$

$$\sum_{s=1}^n u_s x_s + u_{n+1} < 0. \quad (6.13)$$

Como $w_j \geq 0$, $j = 1, 2, \dots, p$ e $t_i \geq 0$, $i = 1, 2, \dots, q$ poderemos escrever também

$$\sum_{s=1}^n u_s v_s^j + u_{n+1} \geq 0, \quad j = 1, 2, \dots, p, \quad (6.14)$$

$$\sum_{s=1}^n u_s r_s^i \geq 0, \quad i = 1, 2, \dots, q, \quad (6.15)$$

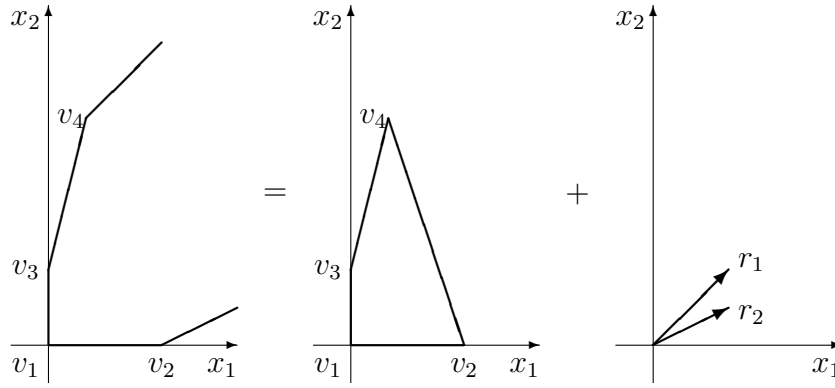


Figura 6.9: $X =$ envoltória convexa dos vértices de $X +$ cone assintótico

$$\sum_{s=1}^n u_s x_s + u_{n+1} < 0. \quad (6.16)$$

Seja $u = (u_1 \ u_2 \ \dots \ u_n)$ logo $ux = \sum_{s=1}^n u_s x_s$. Consideremos agora o seguinte problema de programação linear:

(PPL) : minimizar $z = ux$, sujeito a $x \in X$.

Este (PPL) não é vazio pois supusemos a existência de $x \in X$. As relações representadas por (6.15) podem ser escritas também $ur^i \geq 0$, $i = 1, 2, \dots, q$.

Já vimos que as direções extremas de X podem ser obtidas a partir de uma solução básica primal viável de X e que uma solução ilimitada de um problema de programação linear é sempre verificada através do método do simplex na presença de uma direção extrema de X . Seja \bar{x} uma solução básica primal viável de X e r uma direção extrema de X obtida a partir de \bar{x} pelo método do simplex, sabemos que os pontos da forma $x = \bar{x} + \lambda r$, para $\lambda \geq 0$ pertencem a X , levemos esses pontos à função objetivo do (PPL) e teremos: $z = u\bar{x} + \lambda(ur)$, como $u\bar{x}$ é um valor fixo então se $ur < 0$ implica que $z \rightarrow -\infty$ quando $\lambda \rightarrow \infty$. Caso tivéssemos que $ur^i \geq 0$, $i = 1, 2, \dots, q$ saberíamos, pelo método do simplex, que z teria um ponto de mínimo finito.

Podemos então dizer que o (PPL) em questão além de não ser vazio possui solução ótima limitada que é um vértice de X , por exemplo, v^k , $k \in \{1, 2, \dots, p\}$, logo $uv^k = \sum_{s=1}^n u_s v_s^k \leq \sum_{s=1}^n u_s x_s$ para todo $x \in X$. De (6.14) teremos que $\sum_{s=1}^n u_s v_s^k \geq -u_{n+1}$. Dessas últimas duas expressões obteremos $\sum_{s=1}^n u_s x_s \geq -u_{n+1}$ que é contraditória com a expressão (6.16).

Logo não existem u_s , $s = 1, 2, \dots, n, n+1$; w_j , $j = 1, 2, \dots, p$ e t_i , $i = 1, 2, \dots, q$ satisfazendo a (6.11), (6.12) e (6.13) implicando a existência de λ_j , $j = 1, 2, \dots, p$ e μ_i , $i = 1, 2, \dots, q$. ■

O resultado deste teorema pode ser ilustrado na figura 6.9.

6.5 Exercícios

1. Seja o (PPL) : maximizar $z = cx$ sujeito a $Ax = b$, $x \geq 0$. Se o máximo de z não for limitado, mostrar que existe uma direção d do conjunto poliédrico convexo que representa os pontos das restrições do (PPL) tal que $cd > 0$.
2. Mostrar que existe $\alpha \in \mathcal{R}$ satisfazendo as condições em (6.1).

Capítulo 7

Geração de Colunas

7.1 Introdução

Como foi visto, o teorema 6.1 serve para a representação de um conjunto poliédrico convexo, uma das aplicações desse teorema será apresentada neste capítulo visando à solução de problemas de programação linear por geração de colunas. A seguir consideraremos o seguinte problema de programação linear:

$$(P) : \text{ minimizar } z = cx$$

sujeito a:

$$Ax = b$$

$$x \geq 0,$$

onde $c^T, x \in \mathcal{R}^n$, $A \in \mathcal{R}^{m \times n}$, e $b \in \mathcal{R}^m$.

Particionaremos A e b da seguinte maneira

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

onde $A_1 \in \mathcal{R}^{m_1 \times n}$, $A_2 \in \mathcal{R}^{m_2 \times n}$, $b_1 \in \mathcal{R}^{m_1}$ e $b_2 \in \mathcal{R}^{m_2}$.

Assim (P) poderá ser colocado sob a seguinte forma:

$$(P) : \text{ minimizar } z = cx$$

sujeito a:

$$A_1x = b_1$$

$$A_2x = b_2$$

$$x \geq 0.$$

Denominemos $X = \{x \in \mathcal{R}^n \mid A_2x = b_2, x \geq 0\} \neq \emptyset$. Sejam $V(X) = \{v^1, v^2, \dots, v^p\}$ o conjunto dos v rtices de X e $R(X) = \{r^1, r^2, \dots, r^q\}$ o conjunto dos raios extremos de X . Pelo teorema estudado na se  o 6.4, qualquer $x \in X$ poder  ser escrito como uma combina  o convexa dos elementos de $V(X)$ mais uma combina  o n o negativa dos elementos de $R(X)$, isto  , existem $\lambda_j \in \mathcal{R}$ e $\mu_i \in \mathcal{R}$, tais que

$$x \in X \Rightarrow x = \sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i,$$

para

$$\sum_{j=1}^p \lambda_j = 1, \lambda_j \geq 0, j = 1, 2, \dots, p, \mu_i \geq 0, i = 1, 2, \dots, q.$$

Levando esta  ltima express o de $x \in X$ a $A_1x = b_1$ teremos (P) escrito em fun  o dos $\lambda_j, j = 1, 2, \dots, p$ e $\mu_i, i = 1, 2, \dots, q$ como segue.

$$(P) : \text{ minimizar } z = c \left(\sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i \right)$$

sujeito a:

$$A_1 \left(\sum_{j=1}^p \lambda_j v^j + \sum_{i=1}^q \mu_i r^i \right) = b_1$$

$$\sum_{j=1}^p \lambda_j = 1$$

$$\lambda_j \geq 0, j = 1, 2, \dots, p, \mu_i \geq 0, i = 1, 2, \dots, q,$$

ou ainda

$$(P) : \text{ minimizar } z = \sum_{j=1}^p (cv^j) \lambda_j + \sum_{i=1}^q (cr^i) \mu_i \quad (7.1)$$

$$\sum_{j=1}^p (A_1 v^j) \lambda_j + \sum_{i=1}^q (A_1 r^i) \mu_i = b_1 \quad (7.2)$$

$$\sum_{j=1}^p \lambda_j = 1 \quad (7.3)$$

$$\lambda_j \geq 0, j = 1, 2, \dots, p, \mu_i \geq 0, i = 1, 2, \dots, q. \quad (7.4)$$

A matriz dos coeficientes dos λ_j e μ_i poder  ser esquematizada:

$$M = \begin{pmatrix} A_1 v^1 & A_1 v^2 & \dots & A_1 v^p & A_1 r^1 & A_1 r^2 & \dots & A_1 r^q \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

O problema (P) sob a forma (7.1)-(7.4) é denominado o *problema mestre*.

Tomemos uma matriz quadrada B formada por $m_1 + 1$ colunas de M , tal que B seja inversível. A matriz B terá colunas dos seguintes tipos:

$$\begin{pmatrix} A_1 v^j \\ 1 \end{pmatrix} \quad \text{e/ou} \quad \begin{pmatrix} A_1 r^i \\ 0 \end{pmatrix}.$$

Escreveremos esta matriz ainda como

$$B = \begin{pmatrix} A_1 v^1 & A_1 v^2 & \dots & A_1 v^{p_1} & A_1 r^1 & A_1 r^2 & \dots & A_1 r^{q_1} \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \end{pmatrix},$$

onde $p_1 + q_1 = m_1 + 1$.

Definiremos também $c_B = (cv^1 \ cv^2 \ \dots \ cv^{p_1} \ cr^1 \ cr^2 \ \dots \ cr^{q_1})$ e $u = c_B B^{-1}$.

Suporemos que

$$B^{-1} \begin{pmatrix} b_1 \\ 1 \end{pmatrix} \geq 0,$$

isto é, a matriz B está associada a uma solução básica primal viável de (7.1)-(7.4). Teremos que verificar se esta mesma base B está também associada a uma solução dual viável de (7.1)-(7.4). Para isso calcularemos:

$$z_j = u \begin{pmatrix} A_1 v^j \\ 1 \end{pmatrix}, \quad z_j - cv^j \quad \text{e} \quad z_i = u \begin{pmatrix} A_1 r^i \\ 0 \end{pmatrix}, \quad z_i - cr^i.$$

Se $z_j - cv^j \leq 0$, $j = 1, 2, \dots, p$ e $z_i - cr^i \leq 0$, $i = 1, 2, \dots, q$ então B também definirá uma solução básica ótima de (7.1)-(7.4). É claro que o número de vértices e de raios extremos de X pode ser muito grande impossibilitando os cálculos de todos os $z_j - cv^j$ e $z_i - cr^i$, assim sendo poderemos proceder da seguinte maneira:

Seja $u = (u_1 \ u_0)$, onde $u_0 \in \mathcal{R}$, então

$$z_j - cv^j = u_1 A_1 v^j + u_0 - cv^j = (u_1 A_1 - c)v^j + u_0, \quad j = 1, 2, \dots, p; \quad (7.5)$$

e

$$z_i - cr^i = u_1 A_1 r^i - cr^i = (u_1 A_1 - c)r^i, \quad i = 1, 2, \dots, q. \quad (7.6)$$

Poderemos pensar em calcular o máximo dos $z_j - cv^j$ em (7.5) e o máximo dos $z_i - cr^i$ em (7.6); se ambos os máximos forem não positivos a matriz B estará associada a uma solução ótima de (7.1)-(7.4).

Sabemos que

$$\max_{j=1,2,\dots,p} \{z_j - cv^j\} = u_0 + \max_{j=1,2,\dots,p} \{(u_1 A_1 - c)v^j\} = u_0 + \max_{v \in V(X)} (u_1 A_1 - c)v,$$

ainda podemos considerar o problema de programação linear

$$(PA) : \text{ maximizar } (u_1 A_1 - c)v + u_0$$

sujeito a:

$$A_2 v = b_2$$

$$v \geq 0,$$

para o qual só estaremos interessados nas soluções básicas de $A_2 v = b_2$, satisfazendo $v \geq 0$, pois assim verificaríamos os vértices de X que otimizam (PA) . O método do simplex será utilizado na solução de (PA) .

Três casos devem ser considerados.

- Na solução de (PA) podemos durante o desenrolar do método do simplex verificar a existência de uma solução ilimitada, isto é, encontramos uma solução da forma $v + \alpha r$, para $\alpha \geq 0$ que pertence a X , onde $v \in V(X)$ e $r \in R(X)$, tal que $(u_1 A_1 - c)r > 0$, neste caso faremos a coluna $\begin{pmatrix} A_1 r \\ 0 \end{pmatrix}$ entrar na base, como já foi visto no capítulo 3.
- Ao resolvermos (PA) obtivemos um vértice ótimo \bar{v} , tal que $(u_1 A_1 - c)\bar{v} + u_0 > 0$, a coluna $\begin{pmatrix} A_1 \bar{v} \\ 1 \end{pmatrix}$ entrará na base, como já foi visto no capítulo 3.
- Se o vértice ótimo v^* obtido para (PA) fornecer $(u_1 A_1 - c)v^* + u_0 \leq 0$, logo a base B estará associada a um ótimo de (7.1)-(7.4). Com os λ_j e μ_i associados à base B e seus vértices e raios extremos correspondentes calcularemos o x ótimo de (P) . Denominaremos de $I_{B(\lambda)}$ os índices j dos λ_j em B e de $I_{B(\mu)}$ os índices i dos μ_i em B , assim sendo a solução ótima ficaria $x = \sum_{j \in I_{B(\lambda)}} \lambda_j v^j + \sum_{i \in I_{B(\mu)}} \mu_i r^i$.

Exemplo 7.1

$$(P) : \text{ minimizar } z = x_1 + 2x_2 + x_3 + 2x_4 + 6x_5$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & + & x_2 & - & x_3 & & & = & 5 \\ 4x_1 & + & x_2 & & & & - & x_5 & = & 8 \\ x_1 & - & 2x_2 & & & + & x_4 & & = & 2 \\ & & & & & & & & & x_j \geq 0, \quad j = 1, 2, 3, 4, 5. \end{array}$$

Definiremos $X = \{(x_1 \ x_2 \ x_3 \ x_4 \ x_5)^T \mid x_1 - 2x_2 + x_4 = 2, \ x_j \geq 0, \ j = 1, 2, 3, 4, 5\}$. Os vetores v^j , $j = 1, 2, \dots, p$ representarão os vértices de X e r^i , $i = 1, 2, \dots, q$ os raios extremos de X . Poderemos ainda escrever que $v^j = (v_1^j \ v_2^j \ v_3^j \ v_4^j \ v_5^j)^T$ e $r^i = (r_1^i \ r_2^i \ r_3^i \ r_4^i \ r_5^i)^T$. Assim sendo para todo $x = (x_1 \ x_2 \ x_3 \ x_4 \ x_5)^T$ escreveremos

$$x_k = \sum_{j=1}^p v_k^j \lambda_j + \sum_{i=1}^q r_k^i \mu_i, \quad k = 1, 2, 3, 4, 5,$$

onde

$$\sum_{j=1}^p \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, 2, \dots, p, \quad \mu_i \geq 0, \quad i = 1, 2, \dots, q.$$

Levando o valor de x_k , $k = 1, 2, 3, 4, 5$, em função de λ_j e μ_i na função objetivo e nas duas primeiras restrições de (P) , teremos o seguinte problema mestre.

(P) : minimizar $z =$

$$\sum_{j=1}^p (v_1^j + 2v_2^j + v_3^j + 2v_4^j + 6v_5^j) \lambda_j + \sum_{i=1}^q (r_1^i + 2r_2^i + r_3^i + 2r_4^i + 6r_5^i) \mu_i \quad (7.7)$$

sujeito a:

$$\sum_{j=1}^p (v_1^j + v_2^j - v_3^j) \lambda_j + \sum_{i=1}^q (r_1^i + r_2^i - r_3^i) \mu_i = 5 \quad (7.8)$$

$$\sum_{j=1}^p (4v_1^j + v_2^j - v_5^j) \lambda_j + \sum_{i=1}^q (4r_1^i + r_2^i - r_5^i) \mu_i = 8 \quad (7.9)$$

$$\sum_{j=1}^p \lambda_j = 1 \quad (7.10)$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, p, \quad \mu_i \geq 0, \quad i = 1, 2, \dots, q. \quad (7.11)$$

Como não conhecemos uma solução básica viável para o problema mestre (7.7)-(7.11), utilizaremos o método das duas fases do simplex. Assim sendo definiremos o seguinte problema artificial.

$$(ART) : \text{ minimizar } \xi = g_1 + g_2 + g_3 \quad (7.12)$$

sujeito a:

$$\sum_{j=1}^p (v_1^j + v_2^j - v_3^j) \lambda_j + \sum_{i=1}^q (r_1^i + r_2^i - r_3^i) \mu_i + g_1 = 5 \quad (7.13)$$

$$\sum_{j=1}^p (4v_1^j + v_2^j - v_5^j) \lambda_j + \sum_{i=1}^q (4r_1^i + r_2^i - r_5^i) \mu_i + g_2 = 8 \quad (7.14)$$

$$\sum_{j=1}^p \lambda_j + g_3 = 1 \quad (7.15)$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, p, \quad \mu_i \geq 0, \quad i = 1, 2, \dots, q, \quad g_s \geq 0, \quad s = 1, 2, 3. \quad (7.16)$$

As colunas associadas às variáveis artificiais g_1, g_2 e g_3 formam uma base $B = I_3$ primal viável de (ART) . Podemos em função dessa escolha de base definir:

$$c_B = (1 \ 1 \ 1), \quad u = c_B B^{-1} = (1 \ 1 \ 1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (1 \ 1 \ 1);$$

$$z_j = ua_j = (1 \ 1 \ 1) \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix}$$

ou

$$z_i = ua_i = (1 \ 1 \ 1) \begin{pmatrix} r_1^i + r_2^i - r_3^i \\ 4r_1^i + r_2^i - r_5^i \\ 0 \end{pmatrix},$$

para os j associados a λ_j e os i associados a μ_i temos que $c_j = c_i = 0$ em (ART) .

Consideremos primeiramente os $z_j - c_j = z_j$ visando ao estudo da viabilidade dual de B :

$$z_j = ua_j = (1 \ 1 \ 1) \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix} = 5v_1^j + 2v_2^j - v_3^j - v_5^j + 1.$$

Busquemos o máximo de $z_j - c_j$, onde $j = 1, 2, \dots, p$, para isto formemos o problema auxiliar:

$$(PA) : \text{ maximizar } t = 5v_1 + 2v_2 - v_3 - v_5 + 1 \quad (7.17)$$

sujeito a:

$$v_1 - 2v_2 + v_4 = 2 \quad (7.18)$$

$$v_k \geq 0, \quad k = 1, 2, 3, 4, 5. \quad (7.19)$$

Devemos resolver (PA) utilizando o método do simplex. Na definição de (7.17)-(7.19) o índice j está implícito.

De (7.18): $v_1 = 2 + 2v_2 - v_4$ e $t = 5(2 + 2v_2 - v_4) + 2v_2 - v_3 - v_5 + 1$. Verificamos que para $v_2 = v_3 = v_4 = v_5 = 0$ teremos $v_1 = 2$, logo $v^1 = (2 \ 0 \ 0 \ 0 \ 0)^T$ é um vértice de X . Este vértice fornece $t = 11 > 0$, isto é, se a coluna

$$\begin{pmatrix} 2 + 0 - 0 \\ 4 \times 2 + 0 - 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 1 \end{pmatrix}$$

associada à variável λ_1 entrasse na base no problema (ART), o valor da função objetivo ξ diminuiria. Não precisaríamos resolver (PA) até a otimalidade, basta encontrarmos um vértice de X que forneça um $t > 0$. No exemplo presente resolveremos (PA) visando ao seu ótimo ou determinando que sua solução é ilimitada. Colocando v_1 em função de v_2 e v_4 teremos $t = 11 + 12v_2 - v_3 - 5v_4 - v_5$. Podemos notar que se o valor de v_2 crescer a partir de zero, mantendo-se $v_3 = v_4 = v_5 = 0$, o valor de t aumentará também. Neste caso $v_1 = 2 + 2v_2 \geq 0$ o que implica $v_2 \geq -1$, logo a componente v_2 não é limitada superiormente. Assim sendo, se $v_2 \rightarrow \infty$ implica $t \rightarrow \infty$.

Passaremos a determinar o raio extremo associado:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2v_2 \\ 1v_2 \\ 0v_2 \\ 0v_2 \\ 0v_2 \end{pmatrix},$$

logo $r^1 = (2 \ 1 \ 0 \ 0 \ 0)^T$ é um raio extremo de X .

Calculemos então

$$z_i - c_i = z_i = (1 \ 1 \ 1) \begin{pmatrix} r_1^i + r_2^i - r_3^i \\ 4r_1^i + r_2^i - r_5^i \\ 0 \end{pmatrix} = 5r_1^i + 2r_2^i - r_3^i - r_5^i,$$

Neste caso $i = 1$, $r_1^1 = 2$, $r_2^1 = 1$, $r_3^1 = r_4^1 = r_5^1 = 0$, $z_{\mu_i} - c_{\mu_i} = z_{\mu_i} = 5 \times 2 + 2 \times 1 - 0 - 0 = 12$. A coluna

$$\begin{pmatrix} 2 + 1 - 0 \\ 4 \times 2 + 1 - 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix}$$

associada à variável μ_1 entrará na base no problema (ART). Para sabermos qual coluna da atual base B sairá faremos os seguintes cálculos:

$$\begin{pmatrix} \bar{g}_1 \\ \bar{g}_2 \\ \bar{g}_3 \end{pmatrix} = B^{-1} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix},$$

$$\bar{\xi} = (1 \ 1 \ 1) \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = 14,$$

$$B^{-1} \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix}.$$

Poderemos esquematizar a determinação da coluna que sairá da base da seguinte maneira:

	1	μ_1	
g_1	5	3	
g_2	8	9*	
g_3	1	0	

pois $\frac{8}{9} < \frac{5}{3}$.

Entrará a coluna associada à variável μ_1 e sairá a coluna associada à variável artificial g_2 . A nova base B será:

$$B = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ logo } B^{-1} = \begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ } c_B = (1 \ 0 \ 1),$$

$$u = c_B B^{-1} = (1 \ 0 \ 1) \begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \left(1 \ -\frac{1}{3} \ 1\right),$$

$$\bar{\xi} = \left(1 \ -\frac{1}{3} \ 1\right) \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \frac{10}{3},$$

$$z_j - c_j = z_j = \left(1 \ -\frac{1}{3} \ 1\right) \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix} = -\frac{1}{3}v_1^j + \frac{2}{3}v_2^j - v_3^j + \frac{1}{3}v_5^j + 1.$$

Consideremos novamente:

$$(PA) : \text{ maximizar } t = -\frac{1}{3}v_1 + \frac{2}{3}v_2 - v_3 + \frac{1}{3}v_5 + 1 \quad (7.20)$$

sujeito a:

$$v_1 - 2v_2 + v_4 = 2 \quad (7.21)$$

$$v_k \geq 0, \ k = 1, 2, 3, 4, 5. \quad (7.22)$$

Escrevendo novamente v_1 em função de v_2 e v_4 , isto é, $v_1 = 2 + 2v_2 - v_4$, fazendo com que $t = -\frac{1}{3}(2 + 2v_2 - v_4) + \frac{2}{3}v_2 - v_3 + \frac{1}{3}v_5 + 1$, ou ainda $t = \frac{1}{3} - v_3 + \frac{1}{3}v_4 + \frac{1}{3}v_5$. Se v_4 entrar na base no lugar de v_2 obteremos o vértice

$v^2 = (0 \ 0 \ 0 \ 2 \ 0)^T$ associado à variável λ_2 e $t = \frac{1}{3} + \frac{1}{3} \times 2 = 1$. Se v_5 entrar na base fará também com que t aumente de valor, mas v_5 não é limitado superiormente, logo se $v_5 \rightarrow \infty$ implica $t \rightarrow \infty$. Determinemos o raio extremo de X associado à variável μ_2 .

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0v_5 \\ 0v_5 \\ 0v_5 \\ 0v_5 \\ 1v_5 \end{pmatrix},$$

logo $r^2 = (0 \ 0 \ 0 \ 0 \ 1)^T$. Calculemos agora

$$z_{\mu_2} - c_{\mu_2} = z_{\mu_2} = \begin{pmatrix} 1 & -\frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} r_1^i + r_2^i - r_3^i \\ 4r_1^i + r_2^i - r_5^i \\ 0 \end{pmatrix} = -\frac{1}{3}r_1 + \frac{2}{3}r_2 - r_3 + \frac{1}{3}r_5.$$

Neste caso $i = 2$ e $r_1^2 = r_2^2 = r_3^2 = r_4^2 = 0$, $r_5^2 = 1$, $\Rightarrow z_{\mu_2} - c_{\mu_2} = \frac{1}{3}$. A coluna

$$\begin{pmatrix} 0 + 0 - 0 \\ 0 + 0 - 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

associada à variável μ_2 entrará na base no problema (*ART*). Para saber qual coluna deixará a base repetiremos o processo anterior:

$$\begin{pmatrix} \bar{g}_1 \\ \bar{\mu}_1 \\ \bar{g}_3 \end{pmatrix} = B^{-1} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ \frac{8}{9} \\ 1 \end{pmatrix},$$

$$B^{-1} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{9} \\ 0 \end{pmatrix}.$$

Determinaremos a coluna que sairá da base utilizando o seguinte esquema:

	1	μ_2
g_1	$\frac{7}{3}$	$\frac{1}{3}^*$
μ_1	$\frac{8}{9}$	$-\frac{1}{9}$
g_3	1	0

Assim sendo a coluna associada à variável artificial g_1 sairá da base e teremos a nova base:

$$B = \begin{pmatrix} 0 & 3 & 0 \\ -1 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Para calcularmos B^{-1} sabemos que

$$\begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{9} \\ 0 \end{pmatrix}.$$

Seja

$$E = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{9} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ e } E^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

logo

$$B^{-1} = \begin{pmatrix} 0 & 3 & 0 \\ -1 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$c_B = (0 \ 0 \ 1), \ u = (0 \ 0 \ 1) \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (0 \ 0 \ 1);$$

por outro lado

$$z_j - c_j = z_j = (0 \ 0 \ 1) \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix} = 1, \ t = 1$$

em (PA) para qualquer vértice de X .

Tomemos $v^2 = (0 \ 0 \ 0 \ 2 \ 0)^T$ associado à variável λ_2 , cuja coluna associada em (ART) será:

$$\begin{pmatrix} 0 + 0 - 0 \\ 0 + 0 - 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Esta coluna entrará na base em (ART) . Teremos, mais uma vez, que saber qual coluna sairá da base.

$$\begin{pmatrix} \bar{\mu}_2 \\ \bar{\mu}_1 \\ \bar{g}_3 \end{pmatrix} = B^{-1} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ \frac{5}{3} \\ 1 \end{pmatrix},$$

$$B^{-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Utilizando o mesmo esquema para saber qual coluna sairá da base:

	1	λ_2
μ_2	7	0
μ_1	$\frac{5}{3}$	0
g_3	1	1^*

A coluna associada à variável artificial g_3 deixará a base. Neste caso particular a matriz básica é a mesma que a anterior, no entanto, sua terceira coluna está associada agora à variável λ_2 . A solução básica que minimiza (*ART*) não possui variáveis artificiais na base, logo é uma solução básica primal viável para (*P*). Isto é,

$$\begin{pmatrix} \bar{\mu}_2 \\ \bar{\mu}_1 \\ \bar{\lambda}_2 \end{pmatrix} = B^{-1} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ \frac{5}{3} \\ 1 \end{pmatrix},$$

que é uma solução básica primal viável de (*P*).

Passamos então à segunda fase do método do simplex. Temos agora que $c_B = (cr^2 \ cr^1 \ cv^2)$, onde $c = (1 \ 2 \ 1 \ 2 \ 6)$ e

$$cr^2 = (1 \ 2 \ 1 \ 2 \ 6) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 6, \quad cr^1 = (1 \ 2 \ 1 \ 2 \ 6) \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 4,$$

$$cv^2 = (1 \ 2 \ 1 \ 2 \ 6) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} = 4.$$

Assim $c_B = (6 \ 4 \ 4)$, logo

$$u = (6 \ 4 \ 4) \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \left(\frac{58}{3} \ -6 \ 4 \right),$$

$$\bar{z} = \left(\frac{58}{3} \ -6 \ 4 \right) \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \frac{158}{3}.$$

Verifiquemos a dual viabilidade da base atual com relação ao problema (P) .

$$\begin{aligned} z_j - c_j = ua_j - c_j &= \begin{pmatrix} \frac{58}{3} & -6 & 4 \end{pmatrix} \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix} - cv^j \\ &= -\frac{17}{3}v_1^j + \frac{34}{3}v_2^j - \frac{61}{3}v_3^j - 2v_4^j + 4. \end{aligned}$$

definiremos novamente o problema auxiliar (PA) :

$$(PA) : \text{ maximizar } t = -\frac{17}{3}v_1 + \frac{34}{3}v_2 - \frac{61}{3}v_3 - 2v_4 + 4 \quad (7.23)$$

sujeito a:

$$v_1 - 2v_2 + v_4 = 2 \quad (7.24)$$

$$v_k \geq 0, \quad k = 1, 2, 3, 4, 5. \quad (7.25)$$

Expressando novamente $v_1 = 2 + 2v_2 - v_4$, teremos $t = -\frac{22}{3} - \frac{61}{3}v_3 + \frac{11}{3}v_4$. Faremos v_4 entrar na base no lugar de v_1 . Sabemos que $v_4 = 2 - v_1 + 2v_2$ e $t = 0 - \frac{11}{3}v_1 + \frac{22}{3}v_2 - \frac{61}{3}v_3$ e como v_2 não é limitado superiormente, se $v_2 \rightarrow \infty$ então $t \rightarrow \infty$. Determinemos o raio extremo associado:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0v_2 \\ 1v_2 \\ 0v_2 \\ 2v_2 \\ 0v_2 \end{pmatrix},$$

logo $r^3 = (0 \ 1 \ 0 \ 2 \ 0)^T$. A coluna $(r_1^3 + r_2^3 - r_3^3 \ 4r_1^3 + r_2^3 - r_5^3 \ 0)^T$ associada à variável μ_3 entrará na nova base de (P) . Determinemos esta coluna:

$$\begin{pmatrix} r_1^i + r_2^i - r_3^i \\ 4r_1^i + r_2^i - r_5^i \\ 0 \end{pmatrix} = \begin{pmatrix} 0 + 1 - 0 \\ 4 \times 0 + 1 - 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$B^{-1} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{1}{3} \\ 0 \end{pmatrix}.$$

Repetindo a mesma esquematização para saber a coluna que sairá da base temos:

	1	μ_3
μ_2	7	2^*
μ_1	$\frac{5}{3}$	$\frac{1}{3}$
λ_2	1	0

Sai a coluna associada à variável μ_2 . Calculemo a inversa da nova base

$$B = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E = \begin{pmatrix} 2 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ logo } E^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{6} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$B^{-1} = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{6} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\begin{pmatrix} \bar{\mu}_3 \\ \bar{\mu}_1 \\ \bar{\lambda}_2 \end{pmatrix} = B^{-1} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}.$$

Sabemos que $cr^3 = (1 \ 2 \ 1 \ 2 \ 6) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} = 6$, $c_B = (cr^3 \ cr^1 \ cv^2) = (6 \ 4 \ 4)$.

Determinemos u :

$$u = (6 \ 4 \ 4) \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \left(\frac{25}{3} \ -\frac{7}{3} \ 4 \right), \bar{z} = \left(\frac{25}{3} \ -\frac{7}{3} \ 4 \right) \begin{pmatrix} 5 \\ 8 \\ 1 \end{pmatrix} = 27.$$

Escreveremos novamente

$$\begin{aligned} z_j - c_j &= \left(\frac{25}{3} \ -\frac{7}{3} \ 4 \right) \begin{pmatrix} v_1^j + v_2^j - v_3^j \\ 4v_1^j + v_2^j - v_5^j \\ 1 \end{pmatrix} - (v_1^j + 2v_2^j + v_3^j + 2v_4^j + 6v_5^j) \\ &= -2v_1^j + 4v_2^j - \frac{28}{3}v_3^j - 2v_4^j - \frac{11}{3}v_5^j + 4. \end{aligned}$$

O problema auxiliar ficará:

$$(PA) : \text{ maximizar } t = -2v_1 + 4v_2 - \frac{28}{3}v_3 - 2v_4 - \frac{11}{3}v_5 + 4 \quad (7.26)$$

sujeito a:

$$v_1 - 2v_2 + v_4 = 2 \quad (7.27)$$

$$v_k \geq 0, \ k = 1, 2, 3, 4, 5. \quad (7.28)$$

Fazendo, mais uma vez, $v_1 = 2 + 2v_2 - v_4$ temos $t = 0 - \frac{28}{3}v_3 - \frac{11}{3}v_5$.

Assim sendo, o vértice $v^1 = (2 \ 0 \ 0 \ 0 \ 0)^T$ é ótimo de (PA) , fornecendo $t = 0$. Logo a última base B é ótima de (P) . A solução ótima do problema mestre será:

$$\mu_3 = \frac{7}{2}, \text{ associada a } r^3 = (0 \ 1 \ 0 \ 2 \ 0)^T,$$

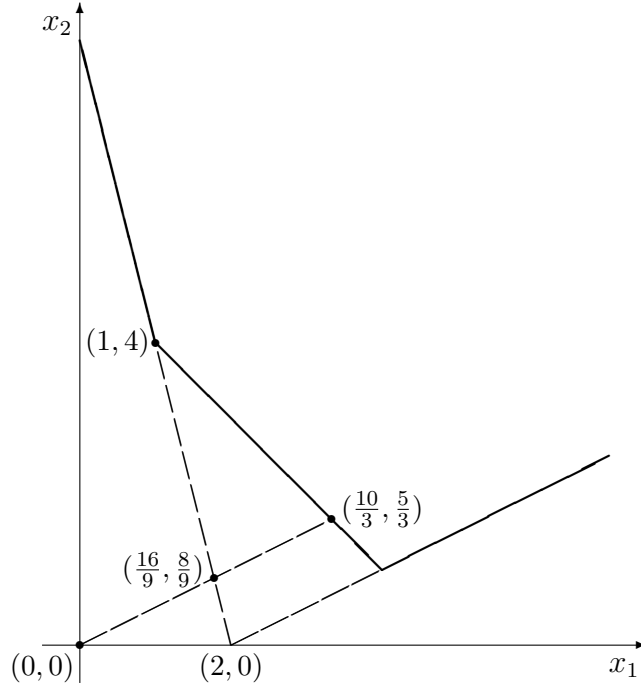


Figura 7.1: Solução do problema mestre (7.7)-(7.11) por geração de colunas

$$\mu_1 = \frac{1}{2}, \text{ associada a } r^1 = (2 \ 1 \ 0 \ 0 \ 0)^T,$$

$$\lambda_2 = 1, \text{ associada a } v^2 = (0 \ 0 \ 0 \ 2 \ 0)^T,$$

forneendo $z = 27$. Retornemos ao problema nas variáveis x_j , $j = 1, 2, 3, 4, 5$.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \frac{7}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 9 \\ 0 \end{pmatrix},$$

ou seja $x_1 = 1$, $x_2 = 4$, $x_3 = 0$, $x_4 = 9$, $x_5 = 0$, fornecendo $z = 27$.

Caso projetemos as restrições do problema original no plano $x_1 \times x_2$ poderemos ilustrar a evolução da solução do problema mestre neste plano, como se pode observar na figura 7.1. ■

7.2 Problema Auxiliar Limitado

Se o conjunto $X = \{x \in \mathcal{R}^n \mid A_2x = b_2, x \geq 0\} \neq \emptyset$ for limitado $R(X) = \emptyset$, só teremos no problema mestre colunas do tipo

$$\begin{pmatrix} A_1v^j \\ 1 \end{pmatrix}.$$

Neste caso, o problema auxiliar (PA) terá sempre um vértice por solução ótima, e conseqüentemente o valor da função objetivo de (PA) , no ótimo, será igual a $(u_1A_1 - c)v^j + u_0$, onde $v^j \in V(X)$. Suponha ainda que $(u_1A_1 - c)v^j + u_0 > 0$, e que λ_j entrará na nova base B do problema mestre fazendo com que

$$z = \bar{z} - [(u_1A_1 - c)v^j + u_0]\bar{\lambda}_j, \text{ onde } 0 \leq \bar{\lambda}_j \leq 1.$$

Lembremos que \bar{z} é o valor da função objetivo do problema mestre antes da entrada de λ_j na base. Logo

$$z \geq \bar{z} - [(u_1A_1 - c)v^j + u_0].$$

Então $\bar{z} - [(u_1A_1 - c)v^j + u_0]$ é uma cota inferior para a solução ótima de (P) . A partir de uma solução básica viável do problema mestre associada à matriz B , a cada iteração do simplex (a cada nova coluna gerada) encontraremos uma cota inferior para o mínimo de (P) . Não se pode garantir que a cota inferior calculada na iteração j será menor do que a calculada na iteração $j + 1$. Isto é, não se pode garantir um comportamento monótono para a cota inferior em função do número de iterações.

Na figura 7.2 ilustramos um possível comportamento dessas cotas inferiores, onde z_j é o valor de z na etapa j e t_j o valor da função objetivo de (PA) na etapa j associada à geração de uma nova coluna para a etapa $j + 1$. Assumimos que z^* é o valor ótimo de z em (P) .

7.3 Exercícios

1. Resolver o seguinte problema de programação linear por geração de colunas:

$$\text{minimizar } z = -3x_1 - 5x_2$$

sujeito a:

$$\begin{array}{rcl} x_1 & \leq & 4 \\ & x_2 & \leq 6 \\ 3x_1 + 2x_2 & \leq & 18 \\ x_1 & \geq & 0 \\ & x_2 & \geq 0, \end{array}$$

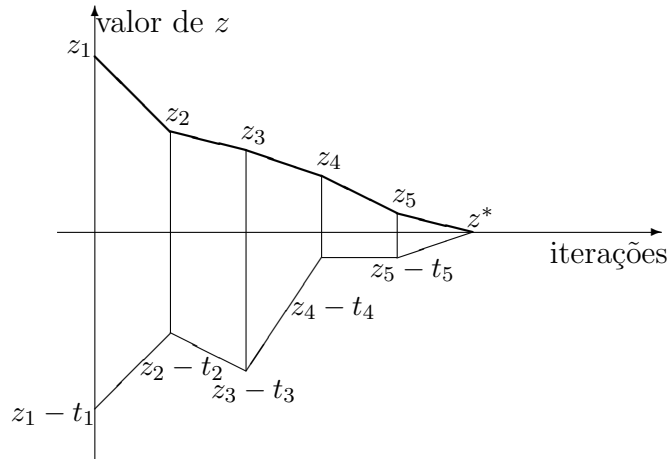


Figura 7.2: Um possível comportamento de cotas inferiores para a solução ótima de (P) , obtidas a partir da solução ótima de (PA)

onde $X = \{(x_1 \ x_2)^T \mid 0 \leq x_1 \leq 4 \text{ e } 0 \leq x_2 \leq 6\}$. Calcular as cotas inferiores em cada geração de coluna. Esquematizar no plano $x_1 \times x_2$ as iterações realizadas de geração de colunas para a obtenção da solução ótima do problema mestre.

2. Resolver pelo método de geração de colunas o problema, maximizar $2x_1 + 4x_2 + x_3$, sujeito a: $2x_1 + x_2 + x_3 \leq 10$, $x_1 + x_2 - x_3 \leq 4$, $0 \leq x_1 \leq 4$, $0 \leq x_2 \leq 6$, $1 \leq x_3 \leq 6$, onde o problema auxiliar para a geração de colunas terá como restrições $0 \leq x_1 \leq 4$, $0 \leq x_2 \leq 6$, $1 \leq x_3 \leq 6$. Fornecer a cada iteração uma cota superior para o ótimo da função objetivo, utilizando a solução do problema auxiliar. Tentar esquematizar o método em um esboço no espaço a três dimensões $x_1 \times x_2 \times x_3$.

Capítulo 8

Tratamento Implícito de Variáveis Limitadas

Muitas vezes temos restrições num problema de programação linear do tipo $\alpha_j \leq x_j \leq \beta_j$, onde j pertence a um subconjunto de $\{1, 2, \dots, n\}$. Denominaremos estas restrições de canalizadas. Neste capítulo trataremos de considerar as restrições canalizadas de maneira implícita.

8.1 Método Primal do Simplex para Restrições Canalizadas

Poderemos considerar um problema de programação linear sob a seguinte forma:

$$(PPL) : \text{maximizar } z = cx \quad (8.1)$$

sujeito a:

$$Ax = b \quad (8.2)$$

$$\alpha_j \leq x_j \leq \beta_j, \quad j = 1, 2, \dots, n; \quad (8.3)$$

onde $c = (c_1 \ c_2 \ \dots c_n)$, $x^T = (x_1 \ x_2 \ \dots x_n)$, $b^T = (b_1 \ b_2 \ \dots b_m)$, $A = (a_1 \ a_2 \ \dots a_n)$ e $a_j^T = (a_{1j} \ a_{2j} \ \dots a_{mj})$, isto é, $c^T \in \mathcal{R}^n$, $x \in \mathcal{R}^n$, $b \in \mathcal{R}^m$, $A \in \mathcal{R}^{m \times n}$ e $a_j \in \mathcal{R}^m$, $\alpha_j < \beta_j$, $j = 1, 2, \dots, n$.

Lembremos que x é o vetor cujas componentes são as variáveis de decisão do (PPL).

Quando x_j não for limitada superiormente consideraremos $\beta_j = +\infty$ e da mesma maneira x_j não for limitada inferiormente tomaremos $\alpha_j = -\infty$. A variável x_j será dita livre quando $\beta_j = +\infty$ e $\alpha_j = -\infty$; neste caso faremos $x_j = x_j^+ - x_j^-$, $x_j^+ \geq 0$ e $x_j^- \geq 0$, ou seja $\alpha_j^+ = \alpha_j^- = 0$ e $\beta_j^+ = \beta_j^- = +\infty$.

Ao longo de todo este capítulo suporemos que se $\alpha_j = -\infty$ então β_j será finito e se $\beta_j = +\infty$ então α_j será finito.

Suporemos que a característica da matriz A seja igual a m . Como já foi realizado no capítulo 2, particionaremos a matriz A da seguinte maneira: $A = (B \ N)$, onde B é uma matriz quadrada $m \times m$ e inversível. Analogamente particionaremos os vetores x e c : $x^T = (x_B^T \ x_N^T)$, $c = (c_B \ c_N)$, x_B e c_B possuirão m componentes associadas à matriz B . Suporemos ainda que I_B seja o conjunto dos índices das colunas de A pertencendo à matriz B e I_N seja o conjunto dos demais índices de A . Lembremos que $I_B \cap I_N = \emptyset$ e $I_B \cup I_N = \{1, 2, \dots, n\}$.

Outros parâmetros serão utilizados: $u = c_B B^{-1}$, $u^T \in \mathcal{R}^m$, $\bar{x}_B = B^{-1}b$, $\bar{x}_B \in \mathcal{R}^m$, $z_j = ua_j$ ($j \in I_B \cup I_N$), $z_j \in \mathcal{R}$, $y_j = B^{-1}a_j$ ($j \in I_B \cup I_N$), $y_j \in \mathcal{R}^m$, $\bar{z} = c_B B^{-1}b = ub = c_B \bar{x}_B$.

Desta maneira o (*PPL*) poderá ser escrito:

$$(PPL) : \text{maximizar } z = \bar{z} - \sum_{j \in I_N} (z_j - c_j)x_j \quad (8.4)$$

sujeito a:

$$x_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_N} y_{ij}x_j, \quad i = 1, \dots, m. \quad (8.5)$$

$$\alpha_{B(i)} \leq x_{B(i)} \leq \beta_{B(i)}, \quad i = 1, 2, \dots, m; \quad (8.6)$$

$$\alpha_j \leq x_j \leq \beta_j, \quad j \in I_N. \quad (8.7)$$

Definiremos uma solução básica de (8.2) ou de (8.5) quando fizermos $x_j = \alpha_j$ ou $x_j = \beta_j$, para $j \in I_N$; claro que α_j ou β_j tem que ser finito pelas hipóteses feitas acima.

Particionaremos também $I_N = I_\alpha \cup I_\beta$, onde $I_\alpha = \{j \in I_N \mid x_j = \alpha_j\}$ e $I_\beta = \{j \in I_N \mid x_j = \beta_j\}$. Podemos assim escrever:

$$\hat{x}_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_\alpha} y_{ij}\alpha_j - \sum_{j \in I_\beta} y_{ij}\beta_j, \quad i = 1, 2, \dots, m. \quad (8.8)$$

Quando $\alpha_{B(i)} \leq \hat{x}_{B(i)} \leq \beta_{B(i)}$, $i = 1, 2, \dots, m$, diremos que estamos diante de uma solução básica primal viável.

É interessante notarmos que a mesma base B pode fornecer mais de uma solução básica viável dependendo da partição de I_N .

A leitora ou o leitor poderá verificar que esta solução básica primal viável corresponde a um vértice do poliedro formado pelas restrições (8.2) e (8.3).

O valor de z correspondendo a esta solução básica será:

$$\hat{z} = \bar{z} - \sum_{j \in I_\alpha} (z_j - c_j)\alpha_j - \sum_{j \in I_\beta} (z_j - c_j)\beta_j. \quad (8.9)$$

Observando a expressão (8.9) é fácil deduzir as seguintes condições de otimalidade para o caso de maximização:

$$z_j - c_j \geq 0, \forall j \in I_\alpha \text{ e } z_j - c_j \leq 0, \forall j \in I_\beta. \quad (8.10)$$

O algoritmo do simplex neste caso partirá de uma solução básica primal viável. Verificamos se esta solução básica satisfaz às condições de otimalidade (8.10). Se a resposta for afirmativa esta solução básica será ótima do (PPL). Caso contrário partiremos à busca de uma nova solução básica primal viável que possa aumentar o valor de z .

1º caso: $k \in I_\alpha$ e $z_k - c_k < 0$, o valor de $x_k = \alpha_k$ poderá passar a $x_k = \alpha_k + \lambda$ ($\lambda \geq 0$), podendo aumentar o valor de z . Como a nova solução básica deverá ser viável, isto é,

$$\alpha_k + \lambda \leq \beta_k \quad (8.11)$$

e

$$\alpha_{B(i)} \leq \bar{x}_{B(i)} - \sum_{j \in I_\alpha} y_{ij} \alpha_j - \sum_{j \in I_\beta} y_{ij} \beta_j - y_{ik} \lambda \leq \beta_{B(i)}, \quad i = 1, 2, \dots, m. \quad (8.12)$$

Podemos expressar (8.12) da seguinte maneira:

$$\alpha_{B(i)} \leq \hat{x}_{B(i)} - y_{ik} \lambda \leq \beta_{B(i)}, \quad i = 1, 2, \dots, m. \quad (8.13)$$

Consideraremos

$$L_k^+ = \{i \mid y_{ik} > 0\}, \quad L_k^- = \{i \mid y_{ik} < 0\} \text{ e } L_k^0 = \{i \mid y_{ik} = 0\}.$$

Nas restrições de (8.13) para as quais $i \in L_k^0$ nada será modificado com a variação do valor de λ , isto é, os valores de $x_{B(i)}$, $i \in L_k^0$ ficarão inalterados.

Teremos então, a partir de (8.11) e (8.12), que

$$\lambda \leq \beta_k - \alpha_k, \quad \lambda \leq \frac{\hat{x}_{B(i)} - \alpha_{B(i)}}{y_{ik}}, \quad i \in L_k^+, \quad \lambda \leq \frac{\hat{x}_{B(i)} - \beta_{B(i)}}{y_{ik}}, \quad i \in L_k^-.$$

Sejam

$$\lambda_{B(p)} = \frac{\hat{x}_{B(p)} - \alpha_{B(p)}}{y_{pk}} = \min_{i \in L_k^+} \left\{ \frac{\hat{x}_{B(i)} - \alpha_{B(i)}}{y_{ik}} \right\}$$

e

$$\lambda_{B(q)} = \frac{\hat{x}_{B(q)} - \beta_{B(q)}}{y_{qk}} = \min_{i \in L_k^-} \left\{ \frac{\hat{x}_{B(i)} - \beta_{B(i)}}{y_{ik}} \right\}.$$

Consideraremos agora:

$$\theta = \min\{ \beta_k - \alpha_k, \lambda_{B(p)}, \lambda_{B(q)} \}. \quad (8.14)$$

Antes de continuarmos faremos as seguintes observações:

(i) quando para $\forall i \in L_k^+$, $\alpha_i = -\infty$ ou quando $L_k^+ = \phi$ faremos $\lambda_{B(p)} = +\infty$;

(ii) quando para $\forall i \in L_k^-$, $\beta_i = \infty$ ou $L_k^- = \phi$ faremos $\lambda_{B(q)} = +\infty$.

Se $\beta_k = +\infty$, e tivermos as condições (i) e (ii) acima, então $\theta = +\infty$. Isto é, λ não será limitado superiormente, logo neste caso quando $\lambda \rightarrow +\infty$ implica $z \rightarrow +\infty$, determinando que o (PPL) é ilimitado.

Supomos agora que θ seja finito e consideraremos a possibilidades seguintes:

1. $\theta = \beta_k - \alpha_k$, a base B será mesma na próxima iteração, no entanto, os valores de $\hat{x}_{B(i)}$ poderão ser modificados, pois $x_k = \beta_k$, $I_\alpha := I_\alpha - \{k\}$ e $I_\beta := I_\beta \cup \{k\}$; I_B não se modifica;
2. $\theta = \lambda_{B(p)}$, a coluna a_k entrará na base substituindo $a_{B(p)}$ e $x_{B(p)} = \alpha_{B(p)}$, teremos ainda que $I_B := (I_B - \{B(p)\}) \cup \{k\}$, $I_\alpha := (I_\alpha - \{k\}) \cup \{B(p)\}$; I_β não se modifica;
3. $\theta = \lambda_{B(q)}$, a coluna a_k entrará na base substituindo $a_{B(q)}$ e $x_{B(q)} = \beta_{B(q)}$, teremos ainda que $I_B := (I_B - \{B(q)\}) \cup \{k\}$, $I_\beta := I_\beta \cup \{B(q)\}$; $I_\alpha := I_\alpha - \{k\}$.

2º caso: $k \in I_\beta$ e $z_k - c_k > 0$, logo $x_k = \beta_k$ poderá passar ao valor $x_k = \beta_k - \lambda$ ($\lambda \geq 0$), podendo aumentar o valor de z . Seguiremos o mesmo procedimento desenvolvido para o 1º caso. Assim sendo, devemos ter:

$$\beta_k - \lambda \geq \alpha_k \quad (8.15)$$

e

$$\alpha_{B(i)} \leq \hat{x}_{B(i)} + y_{ik}\lambda \leq \beta_{B(i)}, \quad i = 1, 2, \dots, m. \quad (8.16)$$

Consideraremos novamente $L_k^+ = \{i \mid y_{ik} > 0\}$, $L_k^- = \{i \mid y_{ik} < 0\}$ e $L_k^0 = \{i \mid y_{ik} = 0\}$. Nas restrições de (8.16) para as quais $i \in L_k^0$ nada será modificado com a variação do valor de λ , isto é, os valores de $x_{B(i)}$, $i \in L_k^0$ ficarão inalterados.

Sabemos que λ terá que satisfazer

$$\lambda \leq \beta_k - \alpha_k, \quad \lambda \leq \frac{\beta_{B(i)} - \hat{x}_{B(i)}}{y_{ik}}, \quad i \in L_k^+, \quad \lambda \leq \frac{\alpha_{B(i)} - \hat{x}_{B(i)}}{y_{ik}}, \quad i \in L_k^-.$$

Sejam

$$\lambda_{B(p)} = \frac{\beta_{B(p)} - \hat{x}_{B(p)}}{y_{pk}} = \min_{i \in L_k^+} \left\{ \frac{\beta_{B(i)} - \hat{x}_{B(i)}}{y_{ik}} \right\}$$

e

$$\lambda_{B(q)} = \frac{\alpha_{B(q)} - \hat{x}_{B(q)}}{y_{qk}} = \min_{i \in L_k^-} \left\{ \frac{\alpha_{B(i)} - \hat{x}_{B(i)}}{y_{ik}} \right\}.$$

Passaremos a considerar

$$\theta = \min\{ \beta_k - \alpha_k, \lambda_{B(p)}, \lambda_{B(q)} \}. \quad (8.17)$$

Supondo em (8.17) que θ seja finito:

1. $\theta = \beta_k - \alpha_k$, teremos $I_\alpha := I_\alpha \cup \{k\}$ e $I_\beta := I_\beta - \{k\}$ e I_B será o mesmo na próxima iteração;
2. $\theta = \lambda_{B(p)}$, a coluna a_k substituirá $a_{B(p)}$, teremos ainda que $I_B := (I_B - \{B(p)\}) \cup \{k\}$, $I_\beta := (I_\beta - \{k\}) \cup \{B(p)\}$; I_α não se modifica;
3. $\theta = \lambda_{B(q)}$, a coluna a_k entrará na base substituindo $a_{B(q)}$ e $I_B := (I_B - \{B(q)\}) \cup \{k\}$, $I_\alpha := I_\alpha \cup \{B(q)\}$; $I_\beta := I_\beta - \{k\}$.

Adaptando o mesmo raciocínio feito para o 1º caso, teremos no 2º caso, quando λ não for limitado superiormente, isto é, em (8.17) $\theta = +\infty$, que o (PPL) será ilimitado.

8.2 Busca de uma Solução Básica Viável

Trataremos de encontrar uma solução básica de (8.2) satisfazendo (8.3). Para isso acrescentaremos uma variável artificial v_i , $i = 1, 2, \dots, m$ às linhas de (8.2), isto é,

$$\sum_{j=1}^n a_{ij}x_j + v_i = b_i, \quad i = 1, 2, \dots, m. \quad (8.18)$$

Suporemos que $I_N = \{1, 2, \dots, n\}$ seja o conjunto das variáveis não básicas, isto é, todas as variáveis x_j , $j = 1, 2, \dots, n$ serão consideradas não básicas inicialmente. Tomemos I_α e I_β , como fizemos anteriormente, uma partição de I_N , faremos com que toda a variável x_j tome valor α_j (finito) ou β_j (finito).

Determinemos

$$s_i = b_i - \sum_{j \in I_\alpha} a_{ij}\alpha_j - \sum_{j \in I_\beta} a_{ij}\beta_j, \quad i = 1, 2, \dots, m.$$

Se $s_i \geq 0$ faremos $v_i \geq 0$ e terá seu coeficiente na função objetivo artificial igual a +1. No caso em que $s_i < 0$ faremos $v_i \leq 0$ e seu coeficiente na função objetivo artificial será -1.

Sejam $I_1 = \{ i \mid v_i \geq 0 \}$ e $I_2 = \{ i \mid v_i \leq 0 \}$, assim sendo nosso problema auxiliar que constituirá a primeira fase do método do simples será

$$(PA) : \text{ minimizar } \xi = \sum_{i \in I_1} v_i - \sum_{i \in I_2} v_i$$

sujeito a

$$v_i + \sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, 2, \dots, m,$$

$$\alpha_j \leq x_j \leq \beta_j, \quad j = 1, 2, \dots, n,$$

$$0 \leq v_i, \quad i \in I_1,$$

$$v_i \leq 0, \quad i \in I_2.$$

Verificamos facilmente que $v_i = s_i$, $i = 1, 2, \dots, m$; $x_j = \alpha_j$, $j \in I_\alpha$ e $x_j = \beta_j$, $j \in I_\beta$ é uma solução primal viável de (PA) .

Caso o mínimo de ξ em (PA) seja estritamente positivo o (PPL) será vazio.

Quando no ótimo de (PA) tivermos $\xi = 0$ e todas as variáveis v_i , $i = 1, 2, \dots, m$ forem não básicas, então a solução obtida eliminando todas as variáveis v_i será uma solução básica primal viável de (PPL) . Se houver $v_i = 0$ na base, procederemos como foi apresentado no capítulo 5.

Exemplo 8.1

$$(PPL) : \text{ maximizar } z = 3x_1 + 5x_2$$

sujeito a:

$$\begin{aligned} 3x_1 + 2x_2 &\geq 18 \\ 2x_1 + x_2 &\leq 12 \\ 0 &\leq x_1 \leq 4 \\ 0 &\leq x_2 \leq 6, \end{aligned} \tag{8.19}$$

Introduzindo as variáveis de folga $x_3 \geq 0$ e $x_4 \geq 0$, o sistema de restrições passa a ser equivalente a:

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 18 \\ 2x_1 + x_2 + x_4 &= 12 \\ 0 &\leq x_1 \leq 4 \\ 0 &\leq x_2 \leq 6 \\ 0 &\leq x_3 \\ 0 &\leq x_4, \end{aligned}$$

daremos, por exemplo, os seguintes valores às variáveis x_j , $j = 1, 2, 3, 4$: $x_1 = 4, x_2 = x_3 = x_4 = 0$, ou seja $I_\alpha = \{2, 3, 4\}$ e $I_\beta = \{1\}$.

Calculemos

$$s_1 = 18 - 3 \times 4 - 2 \times 0 + 1 \times 0 = 6 \geq 0, \quad s_2 = 12 - 2 \times 4 - 1 \times 0 - 1 \times 0 = 4 \geq 0,$$

ambos são não negativos.

O problema auxiliar (primeira fase) será escrito como se segue

$$(PA) : \text{minimizar } \xi = v_1 + v_2$$

sujeito a

$$\begin{array}{rcccccccl} v_1 & & + & 3x_1 & + & 2x_2 & - & x_3 & & = & 18 \\ & v_2 & + & 2x_1 & + & x_2 & & & + & x_4 & = & 12 \\ 0 & \leq & v_1 & & & & & & & & \\ & 0 & \leq & v_2 & & & & & & & \\ & & 0 & \leq & x_1 & \leq & 4 & & & & \\ & & & 0 & \leq & x_2 & \leq & 6 & & & \\ & & & & & 0 & \leq & x_3 & & & \\ & & & & & & 0 & \leq & x_4. & & \end{array}$$

A base $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = B^{-1}$. I_α e I_β já foram definidos anteriormente.

Determinemos $u = c_B B^{-1} = (1 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (1 \ 1)$.

Calculemos:

$$z_1 - c_1 = (1 \ 1) \begin{pmatrix} 3 \\ 2 \end{pmatrix} - 0 = 3 + 2 = 5 \text{ e } x_1 = \beta_1 = 4,$$

satisfazendo a condição de otimalidade (minimização);

$$z_2 - c_2 = (1 \ 1) \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 0 = 2 + 1 = 3 \text{ e } x_2 = \alpha_2 = 0,$$

não satisfazendo a condição de otimalidade (minimização).

Se x_2 tomar um valor estritamente positivo o valor de ξ diminuirá na próxima iteração.

Sabemos que

$$\begin{pmatrix} \bar{v}_1 \\ \bar{v}_2 \end{pmatrix} = B^{-1}b = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 18 \\ 12 \end{pmatrix} = \begin{pmatrix} 18 \\ 12 \end{pmatrix}$$

e que, como

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad y_j = a_j, \quad j = 1, 2, 3, 4.$$

Temos que:

$$\begin{aligned} v_1 &= \bar{v}_1 - 3 \times 4 - 2 \times 0 + 1 \times 0 - 2\lambda = 18 - 12 - 2\lambda \geq 0 \Rightarrow \lambda \leq 3, \\ v_2 &= \bar{v}_2 - 2 - 1 \times 0 - 1 \times 0 - \lambda = 12 - 8 - \lambda \geq 0 \Rightarrow \lambda \leq 4, \\ \text{e } \beta_2 - \alpha_2 &= 6 - 0 = 6. \end{aligned}$$

Logo $\theta = \min\{6, 4, 3\} = 3$, assim sendo a coluna $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, associada à variável x_2 , entrará na base substituindo $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, associada à variável artificial v_1 . Neste caso podemos eliminar v_1 e $I_\alpha = \{3, 4\}$. Consideremos agora a nova iteração:

$$\begin{aligned} B &= \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \Rightarrow B^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix}, \\ c_B &= (0 \ 1), \ u = (0 \ 1) \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} = (-\frac{1}{2} \ 1) \end{aligned}$$

e

$$z_1 - c_1 = (-\frac{1}{2} \ 1) \begin{pmatrix} 3 \\ 2 \end{pmatrix} - 0 = -\frac{3}{2} + 2 - 0 = \frac{1}{2} > 0, \ x_1 = \beta_1 = 4,$$

satisfazendo a condição de otimalidade;

$$z_3 - c_3 = (-\frac{1}{2} \ 1) \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 0 = \frac{1}{2} + 0 - 0 = \frac{1}{2} > 0, \ x_3 = \alpha_3 = 0,$$

não satisfazendo a condição de otimalidade.

Se a variável x_3 assumir um valor estritamente positivo o valor de ξ crescerá. Passaremos a calcular

$$\begin{pmatrix} \bar{x}_2 \\ \bar{v}_2 \end{pmatrix} = B^{-1}b = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 18 \\ 12 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix},$$

$$y_1 = B^{-1}a_1 = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix},$$

$$\hat{x}_2 = 9 - \frac{3}{2} \times 4 = 9 - 6 = 3, \ \hat{v}_2 = 3 - \frac{1}{2} \times 4 = 3 - 2 = 1,$$

$$y_3 = B^{-1}a_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix},$$

assim podemos escrever

$$\begin{aligned} 0 \leq x_2 = 3 + \frac{1}{2}\lambda \leq 6 & \Rightarrow \lambda \leq 6, \\ 0 \leq v_2 = 1 - \frac{1}{2}\lambda & \Rightarrow \lambda \leq 2. \end{aligned}$$

A coluna $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ associada à variável x_3 substituirá a coluna $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ associada à variável artificial v_2 em B . Agora não há mais variável artificial na base e poderemos começar a segunda fase do simplex.

Verificamos que $I_B = \{2, 3\}$, $I_\alpha = \{4\}$ e $I_\beta = \{1\}$ assim sendo,

$$\begin{aligned} B = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \Rightarrow B^{-1} &= \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}. \end{aligned}$$

Sabemos que

$$c_B = (5 \ 0), \quad u = c_B B^{-1} = (5 \ 0) \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = (0 \ 5),$$

logo

$$z_1 - c_1 = u a_1 - c_1 = (0 \ 5) \begin{pmatrix} 3 \\ 2 \end{pmatrix} - 3 = 10 - 3 = 7 > 0,$$

não satisfazendo a condição de otimalidade (maximização).

Se x_1 tomar um valor estritamente menor do que $\beta_4 = 4$, o valor de z aumentará. Assim sendo, passaremos a calcular

$$\begin{pmatrix} \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = B^{-1}b = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 18 \\ 12 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \end{pmatrix},$$

$$y_1 = B^{-1}a_1 = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

$$y_4 = B^{-1}a_4 = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

$$\hat{x}_2 = 12 - 2 \times 4 - 1 \times 0 = 4, \quad \hat{x}_3 = 6 - 1 \times 4 - 2 \times 0 = 2$$

assim teremos que

$$\begin{aligned} 0 \leq x_2 = 4 + 2\lambda \leq 6 &\Rightarrow \lambda \leq 1, \\ 0 \leq x_3 = 2 + 1\lambda &\Rightarrow \lambda \geq -2 \text{ (mas } \lambda \geq 0), \\ 0 \leq x_1 = 4 - \lambda \leq 4 &\Rightarrow \lambda \leq 4. \end{aligned}$$

Dessa maneira a_1 entrará na base no lugar de a_2 . teremos então $I_B = \{1, 3\}$, $I_\alpha = \{4\}$ e $I_\beta = \{2\}$.

Consideremos

$$\begin{aligned} B &= \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} \Rightarrow \\ B^{-1} &= \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{pmatrix}. \end{aligned}$$

Como $c_B = (3 \ 0)$, teremos que $u = (3 \ 0) \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{pmatrix} = (0 \ \frac{3}{2})$.

Calculemos agora

$$z_4 - c_4 = \left(0 \ \frac{3}{2}\right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 0 = \frac{3}{2} > 0,$$

como $4 \in I_\alpha$, satisfazendo a condição de otimalidade.

$$z_2 - c_2 = \left(0 \ \frac{3}{2}\right) \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 5 = -\frac{7}{2} < 0,$$

como $2 \in I_\beta$, satisfazendo a condição de otimalidade.

Finalmente teremos que $x_1 = 3$, $x_3 = 3$, $x_2 = 6$ e $x_4 = 0$ é uma solução básica ótima do (PPL), fornecendo $z = \left(0 \ \frac{3}{2}\right) \begin{pmatrix} 18 \\ 12 \end{pmatrix} - \left(-\frac{7}{2}\right) \times 6 = 18 + 21 = 39$.

Na figura 8.1 ilustraremos os passos do método do simplex executados no exemplo. ■

8.3 Método Dual do Simplex para Restrições Canalizadas

Nesta seção apresentaremos o método dual do simplex adaptado para tratar implicitamente as restrições canalizadas.

Apresentaremos novamente o problema de programação linear (8.1), (8.2) e (8.3) sob a seguinte forma:

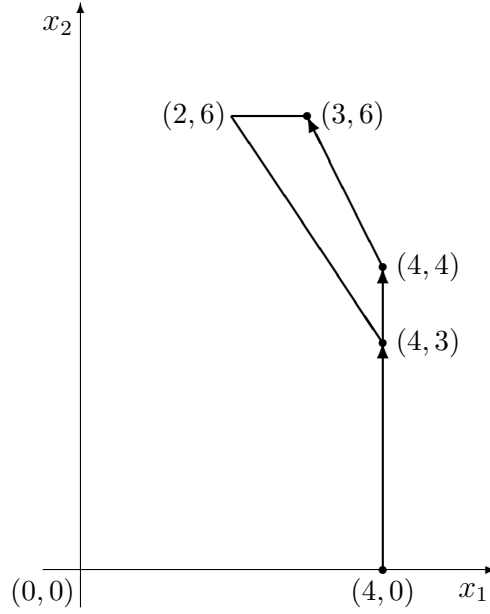


Figura 8.1: Passos do método do simplex executados no exemplo 8.19

$$(PPL) : \text{maximizar } z = cx \quad (8.20)$$

sujeito a:

$$\sum_{j=1}^n a_j x_j = b \quad (8.21)$$

$$x_j \leq \beta_j, \quad j = 1, 2, \dots, n; \quad (8.22)$$

$$-x_j \leq -\alpha_j, \quad j = 1, 2, \dots, n. \quad (8.23)$$

Suporemos, sem perda de generalidade, que α_j e β_j , $j = 1, 2, \dots, n$ sejam finitos. Associamos o vetor $u = (u_1 \ u_2 \ \dots \ u_m)$ às restrições de (8.21), as variáveis $w_j \geq 0$, $j = 1, 2, \dots, n$ às restrições de (8.22) e, finalmente, $t_j \geq 0$, $j = 1, 2, \dots, n$ às restrições de (8.23).

O dual do (PPL) será escrito:

$$(DPL) : \text{minimizar } d = ub + \sum_{j=1}^n \beta_j w_j - \sum_{j=1}^n \alpha_j t_j \quad (8.24)$$

sujeito a:

$$ua_j + w_j - t_j = c_j, \quad j = 1, 2, \dots, n, \quad (8.25)$$

$$w_j \geq 0, \quad t_j \geq 0, \quad j = 1, 2, \dots, n. \quad (8.26)$$

Lembrando que $ua_j = z_j$. Assim sendo as restrições (8.25) poderão ser representadas por

$$z_j - c_j = t_j - w_j, \quad j = 1, 2, \dots, n, \quad (8.27)$$

Como foi comentado no capítulo 4, as relações de complementaridade são sempre verificadas durante a aplicação do método do simplex, seja primal ou seja dual. Vamos supor que B esteja associada à seguinte solução básica de (8.21), isto é,

$$\hat{x}_{B(i)} = \bar{x}_{B(i)} - \sum_{j \in I_\alpha} y_{ij} \alpha_j - \sum_{j \in I_\beta} y_{ij} \beta_j, \quad i = 1, 2, \dots, m,$$

sem que as restrições $\alpha_{B(i)} \leq \hat{x}_{B(i)} \leq \beta_{B(i)}$, $i = 1, 2, \dots, m$, sejam todas verificadas.

Supondo que $\alpha_j < \beta_j$, $j = 1, 2, \dots, n$ e lembrando que $u(Ax - b) = 0$, pois a solução x que estamos considerando é básica, logo satisfaz $Ax = b$. Verificaremos a seguir os valores de w_j e t_j . Para que as relações de complementaridade $w_j(x_j - \beta_j) = 0$ e $t_j(x_j - \alpha_j) = 0$, $j = 1, 2, \dots, n$ possam ser verificadas:

$$j \in I_\alpha, \quad x_j = \alpha_j \Rightarrow x_j < \beta_j \Rightarrow w_j = 0 \Rightarrow z_j - c_j = t_j \geq 0; \quad (8.28)$$

$$j \in I_\beta, \quad x_j = \beta_j \Rightarrow x_j > \alpha_j \Rightarrow t_j = 0 \Rightarrow z_j - c_j = -w_j \leq 0. \quad (8.29)$$

Quando $j \notin I_\alpha \cup I_\beta$, isto é, $j \in I_B$, teremos $z_j - c_j = 0$ que implica por (8.27), $t_j = w_j$; mas por hipótese $\alpha_j < \beta_j$ que fornece ou $t_j = 0$ ou $w_j = 0$, para que tenhamos as relações de complementaridade, neste caso, $t_j = w_j = 0$.

Dado o exposto dizemos que uma solução básica de (8.21) associada à matriz quadrada inversível B , é dual viável de (PPL) se $z_j - c_j \geq 0$, $\forall j \in I_\alpha$ e $z_j - c_j \leq 0$, $\forall j \in I_\beta$.

A idéia é de partir de uma solução básica dual viável do (PPL) , passar a outra, também dual viável, até encontrar uma solução primal e dual viável.

Para iniciar a exposição do método dual do simplex no contexto deste capítulo suporemos conhecida uma solução básica do (PPL) dual viável e que exista $k \in \{1, 2, \dots, m\}$ tal que $\hat{x}_{B(k)} < \alpha_{B(k)}$. Isto é a solução considerada não é primal viável.

O valor de $x_{B(k)}$ deve ser aumentado para que atinja, pelo menos, $\alpha_{B(k)}$, assim temos que ao menos um $y_{kj} < 0$, $j \in I_\alpha$ ou $y_{kj} > 0$, $j \in I_\beta$. Suponhamos que a coluna a_p , onde $p \in I_\alpha \cup I_\beta$, seja escolhida para entrar na base substituindo $a_{B(k)}$. O pivô será y_{kp} e como já foi visto em (4.29),

$y'_{0j} = y_{0j} - \frac{y_{0p}}{y_{kp}} y_{kj}$. Para que a nova solução básica continue dual viável teremos que ter:

$$y'_{0j} \geq 0, \quad j \in I_\alpha \text{ e } y'_{0j} \leq 0, \quad j \in I_\beta. \quad (8.30)$$

Consideremos $L_\alpha = \{j \in I_\alpha \mid y_{kj} < 0\}$ e $L_\beta = \{j \in I_\beta \mid y_{kj} > 0\}$. Deveremos verificar que

- $\frac{y_{0p}}{y_{kp}} \geq \frac{y_{0j}}{y_{kj}}, \quad j \in L_\alpha$, pois $y'_{0j} \geq 0, \quad j \in I_\alpha$, e
- $\frac{y_{0p}}{y_{kp}} \geq \frac{y_{0j}}{y_{kj}}, \quad j \in L_\beta$, pois $y'_{0j} \leq 0, \quad j \in I_\beta$.

A coluna a_p a entrar na base estará associada ao p da expressão abaixo:

$$\frac{y_{0p}}{y_{kp}} = \max_{j \in L_\alpha \cup L_\beta} \left\{ \frac{y_{0j}}{y_{kj}} \right\}.$$

Após o pivoteamento teremos que $x_{B(k)} = \alpha_{B(k)}$ ou seja $B(k) \in I_\alpha$.

Observemos que se $L_\alpha = L_\beta = \phi$ o (PPL) não possuirá solução primal viável.

A título de ilustração tomemos o exemplo 8.1, cuja solução básica ótima é definida por $I_B = \{1, 3\}$, $I_\alpha = \{4\}$ e $I_\beta = \{2\}$. Consideremos $\frac{7}{2} \leq x_1 \leq 4$ no lugar de $0 \leq x_1 \leq 4$. A solução ótima obtida no exemplo 8.1 não será mais primal viável, pois agora temos que $x_1 \geq \frac{7}{2}$, mas continuará dual viável.

A coluna a_1 deixará a base, para sabermos qual coluna entrará na base, teremos que calcular y_{1j} , $j \in I_\alpha \cup I_\beta = \{4, 2\}$.

Lembremos que

$$B = (a_1 \ a_3) = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{pmatrix},$$

$$a_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad a_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Logo

$$y_{12} = \begin{pmatrix} 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{2} \geq 0, \quad y_{14} = \begin{pmatrix} 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} \geq 0.$$

Verificamos que $L_\alpha = \phi$ e $L_\beta = \{2\}$. Caso em que só há um elemento a ser considerado, assim sendo a_2 entrará na base substituindo a_1 .

Agora teremos que

$$B = (a_2 \ a_3) = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

Poderemos escrever

$$\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 18 \\ 12 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} x_1 - \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} x_4,$$

ou ainda

$$\begin{aligned} x_2 &= 12 - 2x_1 - x_4, \\ x_3 &= 6 - x_1 - 2x_4. \end{aligned}$$

Sabemos que $x_1 = \frac{7}{2}$ e $x_4 = 0$, fornecendo $x_2 = 5, x_3 = \frac{5}{2}$, solução primal e dual viável, logo ótima. Solicitamos à leitora ou ao leitor que ilustre o procedimento deste ultimo exemplo sobre a figura 25.

De maneira análoga, vamos supor agora a existência de uma solução básica do (PPL) dual viável e que exista $k \in \{1, 2, \dots, m\}$ tal que $\hat{x}_{B(k)} > \beta_{B(k)}$. Isto é a solução considerada não é primal viável.

O valor de $x_{B(k)}$ deve ser diminuído para que atinja, pelo menos, $\beta_{B(k)}$, assim temos que ao menos um $y_{kj} > 0$, $j \in I_\alpha$ ou $y_{kj} < 0$, $j \in I_\beta$.

Suponhamos que a coluna a_p , onde $p \in I_\alpha \cup I_\beta$, seja escolhida para entrar na base substituindo $a_{B(k)}$. O pivô será y_{kp} e como já foi visto em (4.29), $y'_{0j} = y_{0j} - \frac{y_{0p}}{y_{kp}} y_{kj}$. Para que a nova solução básica continue dual viável teremos que ter:

$$y'_{0j} \geq 0, j \in I_\alpha \text{ e } y'_{0j} \leq 0, j \in I_\beta. \quad (8.31)$$

Consideremos $L_\alpha = \{j \in I_\alpha \mid y_{kj} > 0\}$ e $L_\beta = \{j \in I_\beta \mid y_{kj} < 0\}$. Deveremos verificar que

- $\frac{y_{0p}}{y_{kp}} \leq \frac{y_{0j}}{y_{kj}}$, $j \in L_\alpha$, pois $y'_{0j} \geq 0$, $j \in I_\alpha$, e
- $\frac{y_{0p}}{y_{kp}} \leq \frac{y_{0j}}{y_{kj}}$, $j \in L_\beta$, pois $y'_{0j} \leq 0$, $j \in I_\beta$.

A coluna a_p a entrar na base estará associada ao p da expressão abaixo:

$$\frac{y_{0p}}{y_{kp}} = \min_{j \in L_\alpha \cup L_\beta} \left\{ \frac{y_{0j}}{y_{kj}} \right\}.$$

Após o pivoteamento teremos que $x_{B(k)} = \beta_{B(k)}$ ou seja $B(k) \in I_\beta$.

Observemos que se $L_\alpha = L_\beta = \emptyset$ o (PPL) não possuirá solução primal viável.

8.4 Exercícios

1. Utilizando o método dual apresentado neste capítulo, como seriam determinadas as variáveis duais ótimas associadas às restrições canalizadas?

2. Sendo dado o seguinte problema de programação linear com variáveis canalizadas:

maximizar $x_0 = \sum_{j=1}^n c_j x_j$, sujeito a: $\sum_{j=1}^n a_j x_j \leq b$, $0 \leq x_j \leq d_j$, onde $c_j \geq 0, a_j > 0, d_j > 0, j = 1, 2, \dots, n, b > 0$.

Demonstrar que se $\frac{c_j}{a_j} \geq \frac{c_{j+1}}{a_{j+1}}, j = 1, 2, \dots, n-1$ e k tal que $\sum_{j=1}^{k-1} a_j d_j \leq b < \sum_{j=1}^k a_j d_j$ então uma solução ótima do problema será:

$$\begin{aligned} x_j &= d_j, & j &= 1, 2, \dots, k-1; \\ x_j &= 0, & j &= k+1, k+2, \dots, n; \\ x_k &= \frac{b - \sum_{j=1}^{k-1} a_j d_j}{a_k}. \end{aligned}$$

Capítulo 9

Métodos de Pontos Interiores

9.1 Introdução

O método do Simplex visto nos capítulos anteriores tem se mostrado bastante eficaz na resolução de problemas de programação linear aplicados. Em 1972, no entanto, Klee e Minty [KlMi 72] apresentaram um problema teórico com n restrições e $2n$ variáveis para o qual o método executa $2^n - 1$ iterações até encontrar a solução ótima.

Ficou então aberta a questão quanto a existência de um método eficiente para resolver o problema de programação linear. Um método é dito eficiente se ele tem complexidade polinomial, ou seja, se o número de instruções por ele requerido para solucionar um problema é limitado por um polinômio no tamanho do problema.

Em 1979, Khachian [Kha 79] respondeu a esta questão com a publicação do primeiro algoritmo polinomial para resolver o problema de programação linear, o método de elipsóides. Apesar de sua grande importância teórica, no entanto, o método de elipsóides se mostrou ineficaz na prática.

Em 1984, Karmarkar [Kar 84] revolucionou a área da programação linear com a publicação de um algoritmo com complexidade polinomial e bom desempenho quando aplicado a problemas práticos. Esta publicação deu origem a um novo campo de pesquisa chamado de métodos de pontos interiores. Ao contrário do método do Simplex, que utilizando a estrutura combinatória do problema, caminha pelos vértices de sua região viável, os métodos de pontos interiores caminham pelo interior da região viável.

Comparando-se os métodos de pontos interiores com o método do Simplex, os primeiros claramente serão os melhores, se critérios teóricos forem considerados, como por exemplo a complexidade de pior caso. No entanto, na prática os métodos competem até hoje. Uma análise comparativa entre os

melhores métodos de pontos interiores, sob o ponto de vista computacional, e o método do Simplex foi realizada por Illés e Terlaky em [ITe 2002], onde concluem que, de uma forma geral, não há método vencedor. Na prática, a chave para o sucesso dos métodos é a utilização da estrutura dos problemas, da esparsidade e da arquitetura dos computadores.

Os métodos de pontos interiores serão tema deste capítulo. Devido a sua grande importância histórica, iniciaremos o capítulo descrevendo o método de elipsoides. Em seguida apresentaremos o algoritmo afim-escala. Este algoritmo foi descrito pela primeira vez por Dikin em 1967 [Di 67], mas apenas recentemente foi reconhecido, após reaparecer como uma simplificação do algoritmo de Karmarkar. Resultados quanto a convergência global deste método já foram demonstrados por diversos pesquisadores, incluindo o próprio Dikin [Di 74], que publicou em 1974 a primeira análise de convergência. No entanto, até hoje não se conseguiu demonstrar a polinomialidade do método. Conjectura-se inclusive que ele não é polinomial por ter um comportamento bastante parecido com o método do Simplex quando aplicado a alguns problemas. Entretanto, a direção percorrida pelo algoritmo afim-escala em cada iteração, é uma importante ferramenta que será utilizada em praticamente todos os algoritmos de pontos interiores. O seu entendimento portanto, ajudará bastante no aprendizado destes algoritmos.

Em seguida apresentaremos a trajetória central e os algoritmos de pontos interiores primais-duais de trajetória central, de redução potencial e os algoritmos inviáveis.

Consideraremos neste capítulo o seguinte problema de programação linear:

$$(P) : \begin{array}{ll} \text{maximizar} & c^T x \\ \text{sujeito a:} & Ax = b \\ & x \geq 0, \end{array}$$

onde c e $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$ é uma matriz com posto completo.

O conjunto de soluções viáveis de (P) e o conjunto de pontos interiores associado são dados por:

$$\mathcal{X} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$$

e

$$\mathcal{X}^0 = \{x \in \mathcal{X} | x > 0\}.$$

O problema dual de (P) é

$$(D) : \begin{array}{ll} \text{minimizar} & b^T u \\ \text{sujeito a:} & A^T u - s = c \\ & s \geq 0, \end{array}$$

onde $s \in \mathbb{R}^n$ e $u \in \mathbb{R}^m$.

O conjunto de soluções viáveis de (D) e o conjunto de pontos interiores associado são dados por:

$$\mathcal{S} = \{s \in \mathbb{R}^n | A^T u - s = c \text{ para algum } u \in \mathbb{R}^m, s \geq 0\}$$

e

$$\mathcal{S}^0 = \{s \in \mathcal{S} | s > 0\}.$$

Notação

Durante todo o capítulo, quando dado um vetor $x = (x_1, x_2, \dots, x_N)^T \in \mathbb{R}^N$, denotaremos por X a matriz diagonal em $\mathbb{R}^{N \times N}$, cujos elementos da diagonal são as componentes do vetor x , ou seja,

$$X = \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_N \end{bmatrix}.$$

Denotaremos por e , o vetor cujos elementos são todos iguais a um e cuja dimensão é indicada pelo contexto, ou seja,

$$e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

e denotaremos por $\|\cdot\|$ a norma Euclideana.

9.2 O Método de Elipsóides

A questão quanto a existência de um método eficiente para resolver o problema de programação linear ficou respondida com a publicação do método de elipsóides pelo soviético Khachian em 1979. O método de elipsóides não fornece um algoritmo prático para resolver o PPL, já que sua convergência é lenta quando aplicado a grande parte das instâncias do problema e comparado ao método do Simplex. No entanto, sua contribuição para a área da programação matemática é grande, uma vez que ele mostra que o problema de programação linear pode ser resolvido de forma eficiente sob uma abordagem teórica. Este resultado motiva a busca de outros algoritmos para resolver o problema que sejam tanto teoricamente eficientes, como também eficazes na prática. De fato, depois da publicação do método de elipsóides,

uma nova classe de algoritmos para o PPL surgiu. Esta classe recebeu o nome de métodos de pontos interiores e será o tema das seções seguintes.

Nesta seção daremos uma noção geométrica de como o método de elipsóides funciona. Uma apresentação mais formal do algoritmo e as demonstrações dos resultados de convergência, podem ser encontrados em [BeTs 97].

O método será inicialmente apresentado para resolver o problema de decidir se o poliedro definido por

$$P = \{x \in \mathbb{R}^n | Ax \geq b\}$$

é vazio ou não. Em seguida descreveremos como ele pode ser utilizado para resolver o PPL.

Primeiramente definimos um elipsóide com centro $y \in \mathbb{R}^n$ como o seguinte conjunto E de vetores de \mathbb{R}^n :

$$E = E(y, D) = \{x \in \mathbb{R}^n | (x - y)^T D^{-1} (x - y) \leq 1\},$$

onde D é uma matriz $n \times n$ simétrica e semidefinida positiva.

A cada iteração k do método de elipsóides é gerado um elipsóide E_k com centro em um dado ponto x_k , que contém o poliedro P , como exemplificado na figura 9.1.

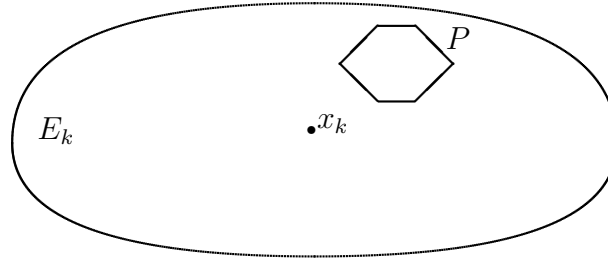


Figura 9.1: Método de Elipsóides: Construção do elipsóide E_k , com centro x_k , que contém o poliedro P .

Se $x_k \in P$ então P não é vazio e o algoritmo pára. Se $x_k \notin P$, então existe i tal que $(a^i)^T x_k < b_i$, onde a^i é a i -ésima linha da matriz A e b_i é o i -ésimo componente do vetor b .

Notemos que se $x \in P$ então $(a^i)^T x \geq b_i$, consequentemente o poliedro P pertence ao semi-espço $\{x \in \mathbb{R}^n | (a^i)^T x \geq (a^i)^T x_k\}$. Sendo assim, se

$x_k \notin P$, sabemos que P pertence a interseção do elipsóide E_k com um semi-espaco que passa pelo centro do elipsóide. Esta região está representada na figura 9.2.

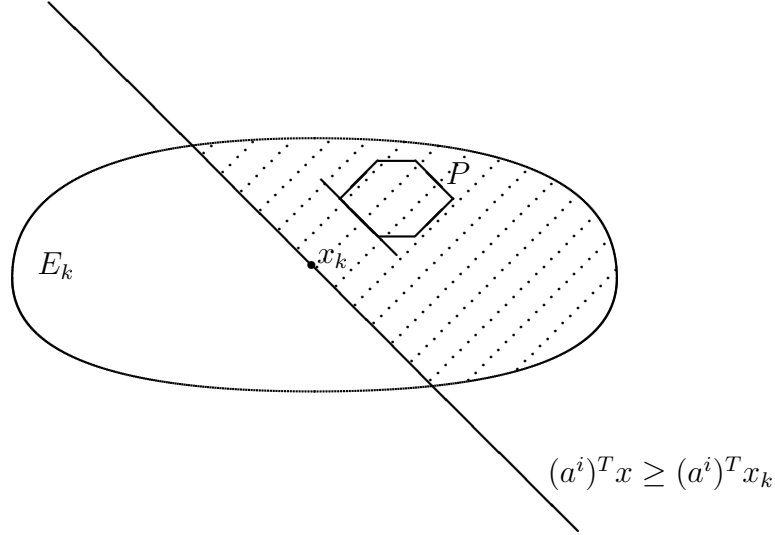


Figura 9.2: Método de Elipsóides: A área pontilhada corresponde à interseção do elipsóide E_k com o semi-espaco $\{x \in \mathbb{R}^n | (a^i)^T x \geq (a^i)^T x_k\}$.

Utilizando então o resultado do teorema 9.1, enunciado a seguir, construímos um elipsóide E_{k+1} que contém esta interseção e tem volume menor que o volume de E_k . Este processo é então repetido até que se obtenha um ponto, centro de um elipsóide, que pertença à P , ou até que se conclua que P é vazio, quando o volume do elipsóide gerado é menor do que um dado número v . A figura 9.3 mostra a última iteração do algoritmo.

O seguinte teorema comprova que é possível construir analiticamente o elipsóide E_{k+1} com volume menor do que o volume de E_k . Sua demonstração pode ser encontrada em [BeTs 97].

Teorema 9.1 *Seja $E = E(y, D)$ um elipsóide em \mathbb{R}^n , $a \in \mathbb{R}^n$ um vetor não nulo e $H = \{x \in \mathbb{R}^n | a^T x \geq a^T y\}$ um semi-espaco em \mathbb{R}^n . Seja:*

$$\begin{aligned} \bar{y} &= y + \frac{1}{n+1} \frac{Da}{\sqrt{a^T Da}} e \\ \bar{D} &= \frac{n^2}{n^2-1} \left(D - \frac{2}{n+1} \frac{Da a^T D}{a^T Da} \right). \end{aligned}$$

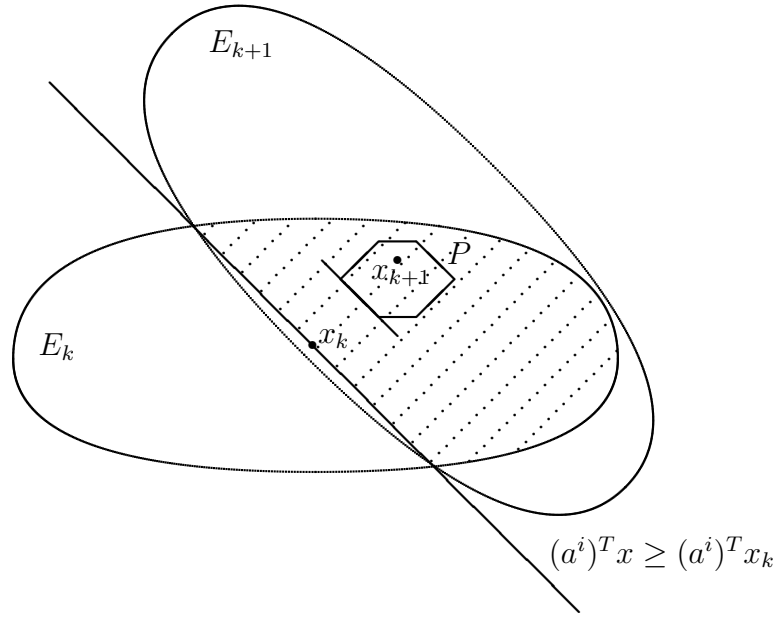


Figura 9.3: Método de Elipsóides: O ponto x_{k+1} , centro do elipsóide E_{k+1} , pertence ao poliedro P .

Os seguintes resultados são verdadeiros:

1. A matriz \bar{D} é simétrica e positiva definida. Consequentemente, $\bar{E} = E(\bar{y}, \bar{D})$ é um elipsóide.
2. $E \cap H \subset \bar{E}$.
3. $\text{Vol}(\bar{E}) < e^{-1/(2(n+1))} \text{Vol}(E)$,
onde $\text{Vol}(S)$ denota o volume do do conjunto $S \in \mathbb{R}^n$.

Da maneira como foi apresentado o método de elipsóides, para que ele seja aplicado com sucesso, é preciso que o poliedro P seja limitado e tenha volume positivo quando não for vazio. Assume-se que é conhecido a priori um elipsóide de volume V que contém P e um número v tal que, se P não for vazio, então $\text{Vol}(P) > v$. Estas hipóteses, no entanto, podem ser relaxadas uma vez que é sempre possível construir um outro poliedro P' que satisfaz a estas hipóteses e será vazio se e somente se P também o for (ver [BeTs 97]).

Finalmente, observamos que o método de elipsóides para decidir se um dado poliedro é vazio, pode ser aplicado para resolver o problema de pro-

gramação linear. Basta considerarmos o problema (P) e seu dual (D) :

$$\begin{array}{ll} (P) : & \text{maximizar } c^T x \\ & \text{sujeito a: } Ax = b \\ & \quad x \geq 0, \\ (D) : & \text{minimizar } b^T u \\ & \text{sujeito a: } A^T u - s = c \\ & \quad s \geq 0, \end{array}$$

e lembrarmos que, pela teoria de dualidade (P) tem solução ótima se e só se (D) também tem. Os vetores x e (u, s) são soluções ótimas de (P) e (D) respectivamente, se o seguinte sistema de inequações lineares é satisfeito:

$$\begin{array}{rcl} c^T x & = & b^T u \\ Ax & = & b \\ x & \geq & 0 \\ A^T u - s & = & c \\ s & \geq & 0 \end{array}$$

Aplicando o método de elipsóides para decidir se este sistema de inequações tem solução, estaremos portanto, resolvendo o problema de programação linear.

9.3 O Método Afim-Escala

Nesta seção apresentaremos o algoritmo afim-escala, descrito primeiramente pelo matemático soviético I. I. Dikin em 1967 [Di 67]. Este algoritmo não é mais utilizado em implementações práticas. A sua importância está no fato de que a direção de busca nele utilizada é também utilizada na maioria dos algoritmos de pontos interiores, como veremos mais tarde.

A Direção Afim-Escala

No algoritmo afim-escala nos é dado um ponto $x^0 \in \mathcal{X}^0$, e o nosso objetivo final é caminhar do ponto x^0 para a solução ótima do problema. A curto prazo, no entanto, nosso objetivo é caminhar numa direção Δx que melhore ao máximo o valor da função objetivo. A direção de máximo aclave de uma dada função, descoberta por Cauchy na primeira metade do século passado, é dada pelo gradiente da função. Por definição, a direção de máximo aclave de uma dada função linear $c^T h$, onde $c \neq 0$, é a solução do problema de maximização da função, em uma bola de raio unitário, ou seja é a solução de

$$\begin{array}{ll} \text{maximizar} & c^T h \\ \text{sujeito a} & \|h\| \leq 1. \end{array} \quad (9.1)$$

A solução de (9.1) pode ser obtida da desigualdade de Cauchy-Schwarz,

$$|c^T h| \leq \|c\| \|h\|,$$

onde a igualdade é verificada quando c e h são colineares e tem o mesmo sentido. Concluimos assim, que a solução de (9.1) é dada por

$$h = \frac{c}{\|c\|}. \quad (9.2)$$

Comprovamos portanto, que a direção de máximo acríve é realmente dada pelo gradiente da função linear $c^T h$.

A direção (9.2) é certamente uma direção de acríscimo para a função objetivo de (P) . No entanto, não necessariamente, ela é uma direção viável para o problema. Ao caminharmos em (9.2), o novo ponto obtido provavelmente não satisfaz a restrição $Ax = b$. Para garantirmos a viabilidade da nova solução, devemos caminhar numa direção d que satisfaça

$$Ad = 0, \quad (9.3)$$

e conseqüentemente teremos $A(x^0 + d) = b$.

Antes de apresentarmos esta direção, no entanto, devemos recordar alguns conceitos de álgebra linear. Lembramos que dada uma transformação linear $A \in \mathbb{R}^{m \times n}$, associa-se a A dois espaços:

o espaço nulo de A ,

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = 0\},$$

e o espaço imagem de A ,

$$\mathcal{I}(A) = \{y \in \mathbb{R}^m | y = Ax, x \in \mathbb{R}^n\}.$$

Existe uma relação interessante entre $\mathcal{N}(A)$ e o espaço imagem de A^T , $\mathcal{I}(A^T)$. Esses dois espaços são sub-espaços ortogonais de \mathbb{R}^n e geram todo o espaço. De fato, é sabido que \mathbb{R}^n é a soma direta desses dois sub-espaços, ou seja, dado qualquer vetor $v \in \mathbb{R}^n$, ele pode ser escrito como

$$v = v_{\mathcal{N}} + v_{\mathcal{I}}, \quad (9.4)$$

onde $v_{\mathcal{N}} \in \mathcal{N}(A)$ e $v_{\mathcal{I}} \in \mathcal{I}(A^T)$. O vetor $v_{\mathcal{N}}$ é a projeção de v sobre $\mathcal{N}(A)$ e $v_{\mathcal{I}}$ é o complemento ortogonal de v em relação a $\mathcal{N}(A)$.

Voltemos agora a análise da direção em que devemos caminhar a partir do ponto x^0 , de forma a aumentar o valor da função objetivo de (P) , sem

sair da região viável do problema. Notemos que a imposição (9.3) implica na pertinência de d ao espaço nulo de A . A direção que passaremos a considerar agora então, é a projeção da direção de máximo aclave (9.2) em $\mathcal{N}(A)$.

Para calcularmos esta projeção, consideraremos inicialmente c escrito como em (9.4), ou seja

$$c = c_{\mathcal{N}} + c_{\mathcal{I}}. \quad (9.5)$$

onde $c_{\mathcal{N}} \in \mathcal{N}(A)$ e $c_{\mathcal{I}} \in \mathcal{I}(A^T)$.

Sabemos que

$$c_{\mathcal{I}} = A^T \lambda, \quad (9.6)$$

para algum $\lambda \in \mathbb{R}^m$.

Substituindo-se (9.6) em (9.5) e multiplicando-se a expressão resultante, a esquerda, por A , temos

$$Ac = Ac_{\mathcal{N}} + AA^T \lambda.$$

Como $Ac_{\mathcal{N}} = 0$ e AA^T é não singular,

$$\lambda = (AA^T)^{-1} Ac. \quad (9.7)$$

Utilizando agora a expressão para λ acima, podemos reescrever (9.5), como

$$c_{\mathcal{N}} = c - c_{\mathcal{I}} = c - A^T(AA^T)^{-1}Ac = (I - A^T(AA^T)^{-1}A)c = P_A c,$$

onde $P_A = [I - A^T(AA^T)^{-1}A]$ é a matriz de projeção sobre o espaço nulo de A .

Verificamos facilmente que a direção $c_{\mathcal{N}}$ é uma direção de acréscimo da função objetivo já que o produto escalar do gradiente da função objetivo por $c_{\mathcal{N}}$ é um número positivo sempre que $c_{\mathcal{N}}$ é um vetor não nulo. Abaixo demonstramos esta afirmação.

$$c^T c_{\mathcal{N}} = (c + c_{\mathcal{N}} - c_{\mathcal{N}})^T c_{\mathcal{N}} = (c_{\mathcal{N}} + c_{\mathcal{I}})^T c_{\mathcal{N}} = \|c_{\mathcal{N}}\|^2.$$

De fato, considerando-se as condições de otimalidade que apresentaremos na próxima seção, é possível mostrar que $c_{\mathcal{N}}$ é solução do problema obtido ao incorporar-se a restrição (9.3) ao problema (9.1). Ou seja, $c_{\mathcal{N}}$ é solução de

$$\begin{aligned} &\text{maximizar} && c^T h \\ &\text{sujeito a} && Ah = 0 \\ &&& \|h\| \leq 1. \end{aligned} \quad (9.8)$$

Na figura 9.4 está representada a direção $c_{\mathcal{N}}$ para o problema de maximizar $-x_1 - x_2$ sujeito apenas as restrições de não negatividade das variáveis

x_1 e x_2 . Neste caso, $P_A = I$ e a direção c_N é dada pelo próprio vetor custo $c = (-1, -1)^T$. Partindo do ponto $(5, 0.5)$, observamos que pouco podemos caminhar ao longo da direção c_N , antes de atingirmos a fronteira da região viável do problema. O progresso obtido com este passo, é portanto pequeno. Uma vez que esta situação é gerada pela proximidade do ponto $(5, 0.5)$ a fronteira da região viável, uma forma de evitá-la seria a de aplicar uma mudança de escala sobre as variáveis do problema, que mapeasse o ponto $(5, 0.5)$ num outro ponto mais distante da fronteira. Devemos portanto, manter este propósito inicial em mente.

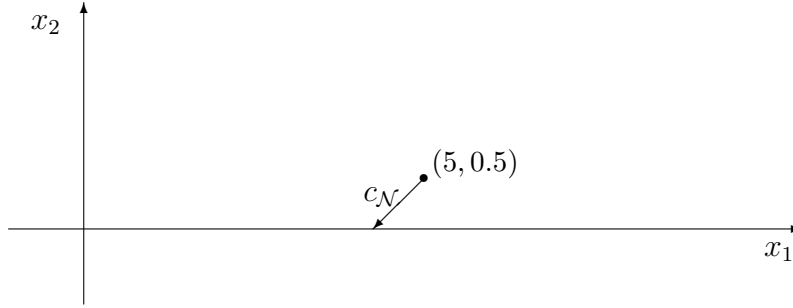


Figura 9.4: Direção de Cauchy

No cálculo da direção de Cauchy (9.8), utilizamos uma região de confiança circular, ou seja, maximizamos a função objetivo do problema numa bola de raio unitário. Através de uma mudança de escala nas variáveis, podemos trabalhar com uma região de confiança elipsoidal.

A idéia do algoritmo afim-escala é de calcular a direção de máximo aclave da função objetivo de (P) , a partir de um dado ponto $x^k \in \mathcal{F}^0$, utilizando como região de confiança um elipsóide, e não mais um círculo. O ideal seria utilizar o maior elipsóide contido dentro da região viável do problema. No entanto, para simplificar o cálculo da direção, os eixos do elipsóide escolhido são paralelos aos eixos coordenados. Neste caso, trabalhamos com o maior elipsóide com centro em x^k e contido no primeiro ortante, sem considerarmos as restrições $Ax = b$. Este elipsóide é chamado de elipsóide de Dikin com centro em x^k e é definido por

$$E = \{x \in \mathbb{R}^n | Ax = 0, x^T (X^k)^{-2} x \leq 1\}.$$

A direção de busca do algoritmo afim-escala para o problema (P) é então definida como a solução de

$$\begin{aligned} & \text{maximizar} && c^T h \\ & \text{sujeito a} && Ah = 0 \\ & && h^T (X^k)^{-2} h \leq 1. \end{aligned} \quad (9.9)$$

Representamos a direção d do algoritmo afim-escala na figura 9.5. Na figura também está representado o elipsóide de Dikin e as curvas de nível da função objetivo do problema. O problema considerado é o mesmo da figura 9.4 e podemos observar com clareza a melhoria obtida ao se substituir a direção de Cauchy pela direção do algoritmo afim-escala.

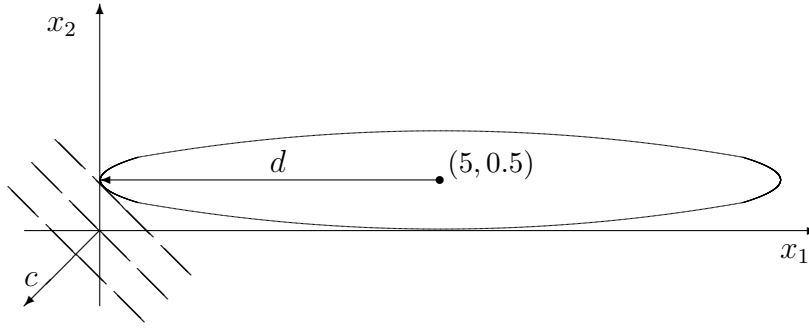


Figura 9.5: Elipsóide de Dikin e Direção Afim-Escala

Finalmente, devemos observar que a direção do afim-escala dada pela solução do problema (9.9) é equivalente a direção de Cauchy projetada no nulo de A , dada por c_N , após uma mudança de escala definida por

$$x = X^k \bar{x}.$$

Esta mudança de escala mapeia o ponto x^k no ponto $e = (1, 1, \dots, 1)^T$, uma vez que $x^k = X^k e$; e transforma o elipsóide de Dikin numa bola de raio unitário centrada em e .

Após a mudança de escala, o problema (P) passa a ser representado por

$$\begin{aligned} (\bar{P}) : & \text{maximizar} && \bar{c}^T \bar{x} \\ & \text{sujeito a:} && \bar{A} \bar{x} = b \\ & && \bar{x} \geq 0, \end{aligned}$$

onde $\bar{c} = X^k c$ e $\bar{A} = AX^k$.

De fato, considerando-se a mudança de escala $h = X^k \bar{h}$, o problema (9.9) passa a ser representado por

$$\begin{aligned} & \text{maximizar} && \bar{c}^T \bar{h} \\ & \text{sujeito a} && \bar{A} \bar{h} = 0 \\ & && \bar{h}^T \bar{h} \leq 1, \end{aligned} \tag{9.10}$$

no qual calculamos a direção viável de máximo aclave da função objetivo de (\bar{P}) , utilizando novamente como região de confiança, uma bola de raio unitário.

Considerando a equivalência dos problemas (9.8) e (9.10), podemos agora apresentar a solução do último deles, dada por:

$$\bar{d} = P_{\bar{A}} \bar{c},$$

onde $P_{\bar{A}} = [I - \bar{A}^T (\bar{A} \bar{A}^T)^{-1} \bar{A}]$ é a matriz de projeção no espaço nulo de \bar{A} .

Retornando a escala original, temos então a direção de busca do algoritmo afim-escala, dada por:

$$d = X^k \bar{d}. \tag{9.11}$$

Finalmente observamos que ao utilizarmos o elipsóide de Dikin como região de confiança no cálculo da direção de máximo aclave, estamos de forma equivalente buscando o nosso propósito inicial: mapear o dado ponto x^k , que pode estar próximo de uma fronteira da região viável, no ponto e , que dista de pelo menos uma unidade de cada fronteira.

Uma vez apresentada a direção de busca do algoritmo afim-escala, estamos agora preparados para a apresentação do algoritmo completo para resolver (P) .

Algoritmo 9.1 (Afim-Escala) *Dados:* $x^0 \in \mathcal{X}^0$ e $\alpha \in (0, 1)$;

$k := 0$;

Repita

Mudança de escala:

$$\bar{A} = AX^k, \quad \bar{c} = X^k c.$$

Projeção:

$$P_{\bar{A}} = I - \bar{A}^T (\bar{A} \bar{A}^T)^{-1} \bar{A}$$

Direção de busca:

$$\bar{d} = P_{\bar{A}} \bar{c}$$

Retorno a escala original:

$$d = X^k \bar{d}$$

Passo:

$$x^{k+1} = x^k + \alpha \lambda d \quad (9.12)$$

Fim (Repita)

No passo (9.12), λ é dado pelo teste da razão, ou seja,

$$\lambda = \min_{j=1, \dots, n} \{-x_j^k / d_j \mid d_j < 0\};$$

e α , que em geral é aproximadamente 0.995, garante que o novo ponto seja estritamente positivo.

Observamos que nenhum critério de parada foi especificado no algoritmo. Para implementá-lo, é preciso que algum critério seja escolhido. Normalmente o critério de parada de algoritmos depende do problema considerado. No exemplo abaixo, no entanto, utilizamos um critério bem geral para algoritmos iterativos. Nele, o método é interrompido quando não há mudança significativa entre as soluções de duas iterações consecutivas.

Exemplo 9.1 Resolver o problema de programação linear abaixo utilizando o algoritmo afim-escala.

$$\begin{aligned} \text{maximizar} \quad & 3x_1 + 5x_2 \\ \text{sujeito a:} \quad & x_1 \leq 4 \\ & x_2 \leq 6 \\ & 3x_1 + 2x_2 \leq 18 \\ & x_1, x_2 \geq 0 \end{aligned} \quad (9.13)$$

É dada uma solução inicial interior viável para o problema: $x_1 = 1$, $x_2 = 1$. Em seguida, reescrevemos o problema na forma padrão, com a adição de variáveis de folga:

$$\begin{aligned} \text{maximizar} \quad & 3x_1 + 5x_2 \\ \text{sujeito a:} \quad & x_1 + x_3 = 4 \\ & x_2 + x_4 = 6 \\ & 3x_1 + 2x_2 + x_5 = 18 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned}$$

A solução inicial interior viável correspondente é

$$x^0 = (1 \ 1 \ 3 \ 5 \ 13)^T,$$

e o valor da função objetivo em x^0 é 8.

Na primeira iteração a mudança de escala é definida pela matriz diagonal

$$X^0 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 3 & & \\ & & & 5 & \\ & & & & 13 \end{pmatrix}.$$

A matriz \bar{A} e o custo \bar{c} obtidos após a mudança de escala são dados por:

$$\bar{A} = AX^0 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3 & 2 & & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 & 5 & 13 \end{pmatrix},$$

logo

$$\begin{aligned} \bar{A} &= AX^0 = \begin{pmatrix} 1 & 3 & 5 & 13 \\ 3 & 2 & & 1 \end{pmatrix}, \\ \bar{c} &= X^0 c = \begin{pmatrix} 1 & 1 & 3 & 5 & 13 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Em seguida calculamos a matriz de projeção no nulo da \bar{A} :

$$\begin{aligned} P_{\bar{A}} &= I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A} \\ &= \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \\ &\quad - \begin{pmatrix} 1 & 3 \\ 3 & 5 \\ & 13 \end{pmatrix} \left[\begin{pmatrix} 1 & 3 & 5 & 13 \\ 3 & 2 & & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 5 \\ & 13 \end{pmatrix} \right]^{-1} \end{aligned}$$

$$\begin{aligned}
& \times \begin{pmatrix} 1 & & 3 & & \\ & 1 & & 5 & \\ 3 & 2 & & & 13 \end{pmatrix} \\
& = \begin{pmatrix} 0.8597 & -0.0287 & -0.2866 & 0.0057 & -0.1940 \\ -0.0287 & 0.9411 & 0.0096 & -0.1882 & -0.1382 \\ -0.2866 & 0.0096 & 0.0955 & -0.0019 & 0.0647 \\ 0.0057 & -0.1882 & -0.0019 & 0.0376 & 0.0276 \\ -0.1940 & -0.1382 & 0.0647 & 0.0276 & 0.0660 \end{pmatrix}.
\end{aligned}$$

A direção de busca será então dada por:

$$\bar{d} = P_{\bar{A}} \bar{c}$$

logo

$$\begin{aligned}
\bar{d} &= \begin{pmatrix} 0.8597 & -0.0287 & -0.2866 & 0.0057 & -0.1940 \\ -0.0287 & 0.9411 & 0.0096 & -0.1882 & -0.1382 \\ -0.2866 & 0.0096 & 0.0955 & -0.0019 & 0.0647 \\ 0.0057 & -0.1882 & -0.0019 & 0.0376 & 0.0276 \\ -0.1940 & -0.1382 & 0.0647 & 0.0276 & 0.0660 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2.4357 \\ 4.6194 \\ -0.8119 \\ -0.9239 \\ -1.2728 \end{pmatrix}.
\end{aligned}$$

Retornando a escala original, temos:

$$\begin{aligned}
d &= X^0 \bar{d} \\
&= \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 3 & & \\ & & & 5 & \\ & & & & 13 \end{pmatrix} \begin{pmatrix} 2.4357 \\ 4.6194 \\ -0.8119 \\ -0.9239 \\ -1.2728 \end{pmatrix} = \begin{pmatrix} 2.4357 \\ 4.6194 \\ -2.4357 \\ -4.6194 \\ -16.5458 \end{pmatrix}
\end{aligned}$$

Determinamos em seguida,

$$\lambda = \min_{j=1,\dots,n} \{-x_j^0/d_j | d_j < 0\} = 0.7857.$$

Finalmente, o novo ponto é dado por:

$$\begin{aligned}
x^1 &= x^0 + \alpha \lambda d \\
&= \begin{pmatrix} 1 \\ 1 \\ 3 \\ 5 \\ 13 \end{pmatrix} + 0.995 * 0.7857 * \begin{pmatrix} 2.4357 \\ 4.6194 \\ -2.4357 \\ -4.6194 \\ -16.5458 \end{pmatrix} = \begin{pmatrix} 2.9041 \\ 4.6113 \\ 1.0959 \\ 1.3887 \\ 0.06501 \end{pmatrix}.
\end{aligned}$$

O valor da função objetivo em x^1 é 31.7688. A primeira iteração do algoritmo está completa. Como houve uma mudança significativa entre as soluções x^0 e x^1 , seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo afim-escala para resolver o exemplo. Consideramos a precisão $\epsilon = 10^{-4}$ e o parâmetro $\alpha = 0.995$.

```

%Algoritmo Afim-Escala
n=5;
k=0;
epsilon=10^(-4);
alfa = 0.995;
A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];
c=[3 5 0 0 0]';
x=[1 1 3 5 13]';
custo=c'*x;
dif = 10;
while (dif>= epsilon)
    Abar = A*diag(x);
    cbar = x.*c;
    aux=(Abar*Abar')\ (Abar*cbar);
    dbar = cbar - Abar'*aux;
    d = x.*dbar;
    aux2=[];
    for i=1:n
        if(d(i)<0)
            aux2 = [aux2; -x(i)/d(i)];
        end
    end
    lambda = min(aux2);
    x=x+alfa*lambda*d;
    custoant=custo;

```

k	0	1	2	3	4	5	6
x_1	1.0000	2.9041	1.9840	2.0030	1.9999	2.0000	2.0000
x_2	1.0000	4.6113	5.9931	5.9954	6.0000	6.0000	6.0000

Tabela 9.1: Iterações geradas pelo Algoritmo Afim-Escala

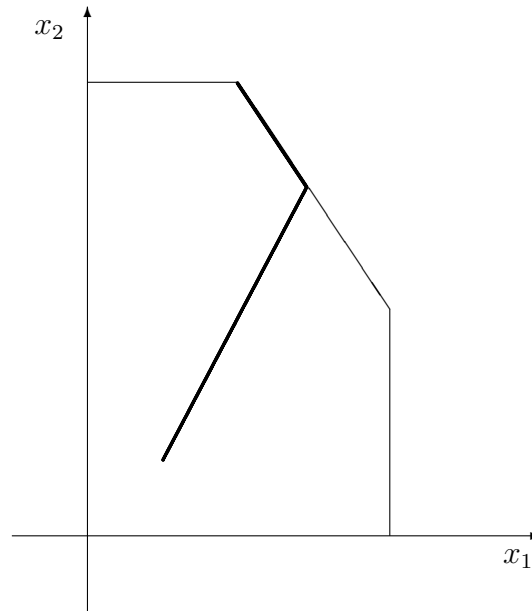


Figura 9.6: Algoritmo Afim-Escala

```

custo=c'*x;
dif=abs(custo-custoant);
k=k+1
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo a cada iteração k estão representados na tabela 9.1.

A trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema, está representada na figura 9.6. ■

9.4 A Trajetória Central

Na seção anterior verificamos que a cada iteração do algoritmo afim-escala, o valor da função objetivo aumenta. Mesmo com a mudança de escala, este fato faz com que, a sequência de iterações gerada pelo método, se aproxime aos poucos da fronteira da região viável de (P) . Com a proximidade da fronteira então, a elipse na qual a função objetivo é minimizada se torna cada vez menor, levando também a passos cada vez menores realizados pelo algoritmo em direção ao ótimo. Para evitar tal dificuldade, os algoritmos de pontos interiores visam, a cada iteração, aumentar o valor da função objetivo de (P) , mantendo-se no entanto, distantes da fronteira, ou seja, no interior da região viável. Pela motivação destes dois objetivos conflitantes, a seguinte função barreira logaritmica, definida para todo $x > 0$, foi empregada:

$$f(x, \mu) := c^T x + \mu \sum_{j=1}^n \log x_j,$$

onde μ é um parâmetro maior que zero.

A função barreira foi utilizada pela primeira vez em otimização por Frisch [Fr 55]. O algoritmo de programação não linear resultante, conhecido como método de barreiras, foi estudado no livro de Fiacco e McConnick [FiMc 55] e é um caso particular de uma família de algoritmos conhecidos como métodos de penalidade.

Definamos agora a seguinte família de problemas barreira associados a (P) e parametrizados por μ :

$$\begin{aligned} &\text{maximizar} && f(x, \mu) \\ &\text{sujeito a:} && Ax = b. \end{aligned} \tag{9.14}$$

Verificamos claramente que o primeiro termo de $f(x, \mu)$ mede o valor da função objetivo de (P) enquanto o segundo termo funciona como uma penalização aos pontos que se aproximam da fronteira da região viável do problema.

Os problemas (9.14) aproximam (P) tanto melhor, quanto menor for o valor de μ . Mostramos na figura 9.7, as soluções ótimas do problema barreira associado ao problema (9.13) para diferentes valores de μ . O ponto x^* corresponde à solução ótima de (9.13). Observamos que para cada valor de μ , a solução de (9.14), a qual é chamada de ponto central, está no interior da região viável de (9.13). Este fato pode ser melhor compreendido se notarmos que em cada uma das faces do poliedro, o valor de uma das variáveis

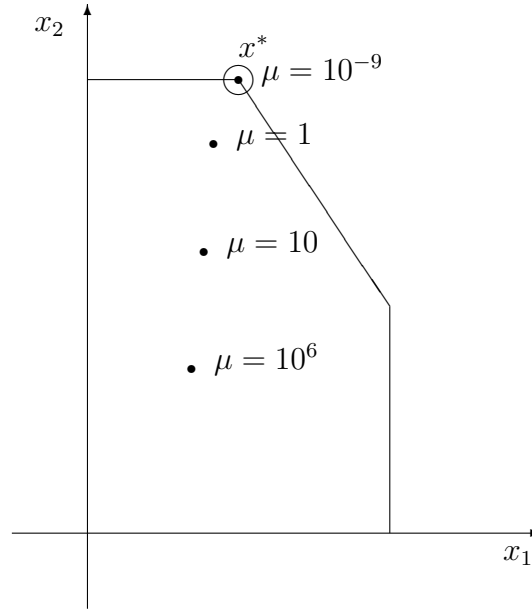


Figura 9.7: Pontos Centrais

do problema se anula. Conseqüentemente, conforme nos aproximamos das faces, o valor da função barreira tende a menos infinito.

O conjunto de pontos centrais de um problema descreve uma curva suave chamada de trajetória central do problema. Pode-se mostrar que, se (P) e (D) têm solução viável com x e s estritamente positivos, então o problema barreira admite solução única para cada valor positivo de μ , já que, neste caso, a função barreira é estritamente côncava [GuRoTeVi 95]. Este fato torna a trajetória central uma curva bem definida. A figura 9.8 ilustra a trajetória central para o problema (9.13). A trajetória central desempenha um importante papel na apresentação dos algoritmos de pontos interiores. de fato, na próxima seção estudaremos uma classe destes algoritmos que se propõem a seguir esta curva em cada iteração. Sendo assim, reservamos o restante desta seção para estudarmos algumas propriedades da trajetória central.

Como já mencionado, cada ponto pertencente a trajetória central de um problema soluciona o problema barreira associado, para um determinado valor de μ . Apresentaremos a seguir as condições de otimalidade que devem ser satisfeitas pela solução de um problema geral de otimização restrita, as quais são conhecidas como condições de Karush-Kuhn-Tucker (ou KKT).

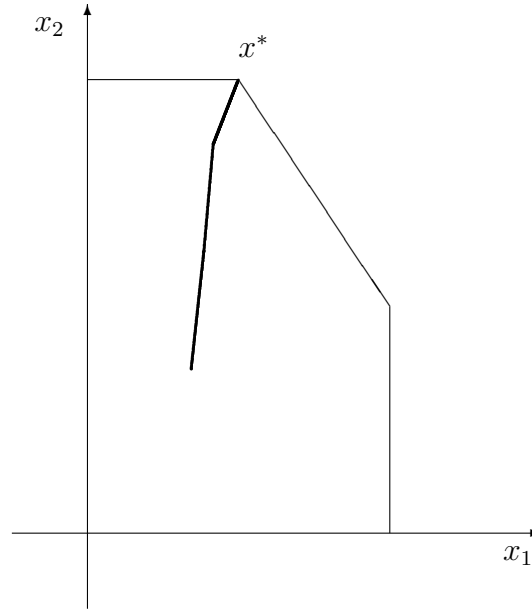


Figura 9.8: Trajetória Central

9.4.1 Condições de Karush-Kuhn-Tucker (ou KKT)

Consideremos o seguinte problema de programação não linear restrito

$$\begin{aligned}
 (NLP) \quad & \text{maximizar} && f(x) \\
 & \text{sujeito a} && g_1(x) = 0 \\
 & && g_2(x) = 0 \\
 & && \vdots \\
 & && g_m(x) = 0,
 \end{aligned}$$

onde $x \in \mathbb{R}^n$ e $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ são funções diferenciáveis.

As condições de KKT para este problema são descritas a seguir:

Se \bar{x} é uma solução ótima de NLP , então existem multiplicadores $y_i \in \mathbb{R}$, $i = 1, \dots, m$, tais que

$$\begin{aligned}
 g_i(\bar{x}) &= 0, & i &= 1, \dots, m \\
 \nabla f(\bar{x}) &= \sum_{i=1}^m y_i \nabla g_i(\bar{x}).
 \end{aligned}$$

Particularizando estas condições para o problema barreira (9.14), definido para um determinado μ , verificamos que a sua solução ótima deve satisfazer ao seguinte sistema de equações não lineares:

$$\begin{aligned} c_j + \mu \frac{1}{x_j} - \sum_{i=1}^q u_i a_{ij} &= 0, \quad j = 1, 2, \dots, n \\ b_i - \sum_{j=1}^p a_{ij} x_j &= 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

onde os multiplicadores $u_i \in \mathbb{R}$, $i = 1, \dots, m$.

Utilizando a notação matricial, podemos reescrever este sistema de equações como:

$$\begin{aligned} A^T u - \mu X^{-1} e &= c \\ Ax &= b. \end{aligned}$$

onde $u \in \mathbb{R}^m$.

Definindo agora $s = \mu X^{-1} e$, $s \in \mathbb{R}^n$, reescrevemos o sistema como

$$\begin{aligned} A^T u - s &= c \\ Ax &= b \\ s &= \mu X^{-1} e. \end{aligned}$$

Finalmente, multiplicando a terceira equação por X , chegamos às condições de otimalidade para o problema da barreira na forma primal-dual:

$$\begin{aligned} A^T u - s &= c \\ Ax &= b \\ XSe &= \mu e \end{aligned} \tag{9.15}$$

Denominaremos a solução deste sistema para um dado μ de (x_μ, u_μ, s_μ) .

Verificamos que, se $\mu = 0$, as equações acima juntamente com as restrições $x, s \geq 0$, são exatamente as condições de otimalidade dos problemas (P) e (D) . Neste caso, as duas primeiras equações garantem a viabilidade primal e dual, enquanto a terceira corresponde às condições de complementaridade. Se $\mu > 0$, a terceira condição é denominada condição de μ -complementaridade. Ela impõe que o produto das variáveis primais e de folga duais $x_i s_i$ tenham o mesmo valor para todo $i = 1, \dots, n$.

O sistema (9.15) é um sistema não linear que envolve $2n + m$ restrições e $2n + m$ variáveis. A utilização do método de Newton para resolver o sistema, dá origem a família de métodos de pontos interiores primais-duais. A cada iteração destes métodos, parte-se de um ponto (x, u, s) tal que x é primal

viável, (u, s) é dual viável e $x, s > 0$. Se definimos a região viável primal-dual \mathcal{F} e seu interior \mathcal{F}^0 , como:

$$\mathcal{F} := \{(x, u, s) | Ax = b, A^T u - s = c, (x, s) \geq 0\}$$

e

$$\mathcal{F}^0 := \{(x, u, s) | Ax = b, A^T u - s = c, (x, s) > 0\},$$

podemos dizer que a cada iteração, os métodos primais-duais partem de um ponto $(x, u, s) \in \mathcal{F}^0$. A direção na qual se caminha a partir deste ponto é baseada na aplicação do método de Newton ao sistema de equações (9.15).

O Método de Newton

No método de Newton, dada uma função $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$, a busca de um zero da função, isto é, a busca de um ponto $\bar{\lambda} \in \mathbb{R}^N$ tal que $F(\bar{\lambda}) = 0$, é realizada através de um processo iterativo, no qual a cada iteração parte-se de um dado ponto $\lambda \in \mathbb{R}^N$ e caminha-se numa aproximação da direção $\Delta\lambda$ tal que $F(\lambda + \Delta\lambda) = 0$. Esta aproximação é calculada a partir de um modelo linear para a função F , dado pelo truncamento da sua expansão em Série de Taylor em torno de λ ,

$$F(\lambda + \Delta\lambda) \approx F(\lambda) + J(\lambda)\Delta\lambda,$$

onde J é o Jacobiano de F . Sendo assim, o cálculo da direção $\Delta\lambda$ a ser tomada a cada iteração é realizado através da solução do sistema linear $J(\lambda)\Delta\lambda = -F(\lambda)$.

Podemos agora reescrever o sistema (9.15) como $F(\lambda) = 0$, onde $\lambda := (x, u, s)$ e

$$F(\lambda) = F(x, u, s) := \begin{bmatrix} A^T u - s - c \\ Ax - b \\ XSe - \mu e \end{bmatrix}.$$

Neste caso o Jacobiano de F é dado por

$$J(\lambda) = J(x, u, s) = \begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix},$$

e a direção de Newton $\Delta\lambda := (\Delta x, \Delta u, \Delta s)$, quando calculada a partir de um ponto $(x, u, s) \in \mathcal{F}^0$, é dada pela solução do sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e - XSe \end{bmatrix} \quad (9.16)$$

Quando $\mu = 0$, a solução do sistema de equações (9.16) é denominada direção afim-escala. A justificativa deste nome está na expressão de Δx que soluciona o sistema. Para chegarmos a esta expressão, vamos inicialmente obter da terceira equação de (9.16), a seguinte expressão para Δs .

$$\Delta s = X^{-1}(\mu e - XSe - S\Delta x). \quad (9.17)$$

Substituindo (9.17) na primeira equação de (9.16), temos

$$A^T \Delta u + X^{-1} S \Delta x = \mu X^{-1} e - s. \quad (9.18)$$

Podemos agora resolver (9.18) para Δx , obtendo

$$\Delta x = XS^{-1}(c - A^T u + \mu X^{-1} e - A^T \Delta u), \quad (9.19)$$

onde consideramos $s = A^T u - c$.

Substituindo (9.19) na segunda equação de (9.16), temos

$$-AXS^{-1}A^T \Delta u = -AXS^{-1}(c - A^T u + \mu X^{-1} e). \quad (9.20)$$

Resolvendo (9.20) para Δu e substituindo em (9.18), temos

$$A^T \left((AXS^{-1}A^T)^{-1} AXS^{-1}(c - A^T u + \mu X^{-1} e) \right) + X^{-1} S \Delta x = \mu X^{-1} e - s,$$

ou ainda,

$$\begin{aligned} X^{-1} S \Delta x &= \mu X^{-1} e - s \\ &- A^T \left((AXS^{-1}A^T)^{-1} AXS^{-1}(c - A^T u + \mu X^{-1} e) \right). \end{aligned} \quad (9.21)$$

Definindo $D^2 = XS^{-1}$ e considerando novamente $s = A^T u - c$, podemos reescrever (9.21) como

$$\begin{aligned} \Delta x &= \left(D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2 \right) (c - A^T u + \mu X^{-1} e) \\ &= \left(D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2 \right) c \\ &+ \mu \left(D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2 \right) X^{-1} e \end{aligned} \quad (9.22)$$

Finalmente, podemos entender a escolha do nome afim-escala dado a direção (9.16), quando $\mu = 0$. Basta observarmos que, neste caso, (9.22) se reduz a

$$\Delta x = \left(D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2 \right) c,$$

que difere da direção do algoritmo afim-escala (9.11), apenas na escolha da matriz que define a mudança de escala.

Como já mencionado, se $\mu = 0$, o sistema de equações (9.15) em conjunto com as restrições de não negatividade para x e s representam as condições de otimalidade para os problemas (P) e (D). A direção de Newton (9.16) calculada para $\mu = 0$ deve portanto, apontar aproximadamente para o ótimo destes problemas. Como foi verificado, em geral, pouco se pode caminhar ao longo desta direção sem que uma das componentes de x ou s se torne negativa e conseqüentemente, o progresso obtido ao longo deste caminho é pequeno.

Por outro lado, a solução de (9.15) para um dado $\mu > 0$, é um ponto pertencente a trajetória central. Neste caso, a direção de Newton (9.16) deve apontar aproximadamente para trajetória central, ou seja, para o interior do ortante não negativo. Espera-se assim que um caminho maior possa ser percorrido ao longo desta direção antes que a fronteira do ortante seja encontrada. Pode-se considerar então, que a direção (9.16), calculada para algum $\mu > 0$, é obtida ao se desviar para o interior do ortante não negativo, a direção que aponta para o ótimo do problema. O objetivo do desvio é manter-se afastado da fronteira do ortante, de forma a permitir um maior passo ao longo da direção calculada, sem sair da região viável do problema.

Os algoritmos de pontos interiores primais-duais consideram em geral duas direções a caminhar a partir de um dado ponto (x, u, s) , ambas obtidas a partir de (9.16). A primeira é a já mencionada direção afim-escala que é obtida quando $\mu = 0$, e a segunda, denominada direção de centralização, é obtida quando μ assume o valor dado por

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i s_i = \frac{x^T s}{n}. \quad (9.23)$$

Devemos observar que o *gap* de dualidade associado às soluções x e (u, s) dos problemas (P) e (D), é dado por

$$b^T u - c^T x = x^T A^T u - c^T x = (uA - c)^T x = x^T s. \quad (9.24)$$

Verificamos então, de (9.23), que o *gap* associado a x e (u, s) pode ser escrito como $n\mu$, quando μ assume o valor que lhe é atribuído no cálculo da direção de centralização a partir do ponto (x, u, s) .

Notemos agora que a direção de centralização aponta aproximadamente para o ponto da trajetória central (x_μ, u_μ, s_μ) que satisfaz a relação $x_i s_i = \mu$ para todo $i = 1, \dots, n$. O *gap* de dualidade associado a este ponto central, dado por $x_\mu^T s_\mu$, é, portanto, igual a $n\mu$ e igual ao *gap* associado ao ponto de partida (x, u, s) .

Concluimos então que a direção de centralização aponta bem para o interior do ortante não negativo. Mais especificamente, ela aponta aproximadamente para o ponto central que tem o mesmo *gap* de dualidade que o ponto

de onde partimos. Esperamos assim que possamos caminhar razoavelmente na direção de centralização antes de encontrar a fronteira do ortante não negativo. Por outro lado, esperamos também que pouca ou nenhuma redução seja obtida no *gap* de dualidade ao longo desta direção.

Finalmente, considerando ao mesmo tempo os objetivos de diminuir o *gap* de dualidade e manter-se longe da fronteira do ortante não negativo, os algoritmos primais-duais partem a cada iteração de um dado ponto (x, u, s) e caminham numa direção obtida pela combinação das duas direções acima descritas. Esta combinação é dada pela solução do sistema abaixo,

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau \mu e - XSe \end{bmatrix} \quad (9.25)$$

onde $\mu = x^T s / n$ e $\tau \in [0, 1]$.

Quando τ assume um valor entre 0 e 1, a direção (9.25) se torna uma direção intermediária entre as direções afim-escala e de centralização. Diferentes escolhas de τ caracterizam diferentes algoritmos pertencentes à família dos algoritmos de pontos interiores primais-duais. Nas próximas seções, estudaremos alguns destes algoritmos.

9.5 Algoritmos de Trajetória Central

Na seção anterior definimos uma curva chamada trajetória central em função de um parâmetro μ . Comentamos que, se $\mathcal{F}^0 \neq \emptyset$, cada ponto desta curva corresponde à solução única, (x_μ, u_μ, s_μ) , do problema barreira (9.14) para um dado μ . É possível mostrar também que neste caso $\lim_{\mu \rightarrow 0} (x_\mu, u_\mu, s_\mu)$ existe, e que $\lim_{\mu \rightarrow 0} x_\mu$ é uma solução ótima de (P) e $\lim_{\mu \rightarrow 0} (u_\mu, s_\mu)$ é uma solução ótima de (D) [Mc 80].

Nesta seção estudaremos os algoritmos de trajetória central que se caracterizam por seguir esta trajetória na direção em que μ decresce. Os algoritmos seguem a trajetória, no sentido em que a cada uma de suas iterações são gerados pontos (x^k, u^k, s^k) que não necessariamente pertencem à trajetória, mas que se localizam próximos a ela. Estes pontos são estritamente positivos e satisfazem apenas as duas primeiras equações de (9.15), enquanto a terceira equação, ou a condição de μ -complementariedade, não necessariamente é exatamente satisfeita. Sendo assim, para medir a proximidade entre tais pontos e a trajetória central, consideramos justamente o quanto esta terceira equação é violada.

Definição: Dados $(x, u, s) \in \mathcal{F}^0$ e $\mu > 0$, a proximidade entre os pontos

(x, u, s) e (x_μ, u_μ, s_μ) é medida por:

$$\delta(x, s, \mu) := \frac{1}{\mu} \|XSe - \mu e\|. \quad (9.26)$$

Os métodos de trajetória central geram uma seqüência de pontos que se situam numa vizinhança da trajetória caracterizada por δ . Dado $\alpha \in (0, 1)$, definimos esta vizinhança como:

$$\mathcal{N}(\alpha) := \bigcup_{\mu \in (0, \infty)} \{(x, u, s) \in \mathcal{F}^0 \mid \delta(x, s, \mu) \leq \alpha\} \quad (9.27)$$

Notemos que dado o ponto (x, u, s) na trajetória central, a relação $XSe = \mu e$ é satisfeita para algum $\mu > 0$, e, conseqüentemente, $x^T s = n\mu$. Esta igualdade indica a relação entre pontos da trajetória e o *gap* de dualidade a eles associado. Uma relação análoga também pode ser obtida para pontos pertencentes à vizinhança $\mathcal{N}(\alpha)$. Para estes pontos, o *gap* de dualidade dado por $x^T s$ relaciona-se com δ pela expressão:

$$x^T s \leq (n + \delta(x, s, \mu)\sqrt{n})\mu \quad (9.28)$$

Para verificar esta expressão basta multiplicar por e^T , a relação $\frac{XSe}{\mu} = e + \beta$, onde $\|\beta\| = \delta(x, s, \mu)$, e aplicar a desigualdade de Cauchy-Schwartz. A importância desta relação está no fato de que ela nos dá um bom critério de parada para os algoritmos.

O Algoritmo de Trajetória Central de Passos Curtos

Estamos agora preparados para apresentar algoritmos que geram pontos sempre na vizinhança $\mathcal{N}(\alpha)$. Estes algoritmos, chamados de algoritmos de trajetória central de passos curtos, adotam valores próximos de 1 para τ . Desta forma, ao tomarmos o passo $(\Delta x, \Delta u, \Delta s)$ a partir de um ponto em $\mathcal{N}(\alpha)$, o novo ponto alcançado também pertencerá à vizinhança.

Para entendermos o comportamento do algoritmo de trajetória central de passos curtos, devemos observar que, se a cada iteração, o ponto de partida está em $\mathcal{N}(\alpha)$, então o ponto (x, u, s) está próximo do ponto da trajetória central (x_μ, u_μ, s_μ) . Ao tomarmos o valor de τ próximo de 1, teremos $\tau\mu \approx \mu$ e, portanto, (x, u, s) também estará próximo de $(x_{\tau\mu}, u_{\tau\mu}, s_{\tau\mu})$. Essa proximidade garante que a direção de Newton (9.25) seja uma boa aproximação para a direção que aponta para a trajetória central e, conseqüentemente, que o novo ponto obtido pertença também à proximidade da trajetória definida por $\mathcal{N}(\alpha)$.

Uma vez garantida que todas as iterações do algoritmo pertencem a $\mathcal{N}(\alpha)$, concluímos que a cada iteração, gera-se soluções viáveis para os problemas (P) e (D) , cujo *gap* de dualidade pode ser estimado por (9.28).

Apresentaremos agora o algoritmo primal-dual de trajetória central de passos curtos que foi introduzido por Kojima, Mizuno e Yoshise [KoMiYo 89] e Monteiro e Adler [MoAd 89] independentemente. A análise de complexidade apresentada a seguir é devida ao segundo grupo de autores. Nela verificamos que a complexidade do algoritmo é de $\sqrt{n} \log 1/\epsilon$, onde ϵ mede a precisão da solução obtida.

Algoritmo 9.2 (Trajetória Central de Passos Curtos) *Dados:* $\epsilon > 0$, $\alpha = 0.4$, $\tau = 1 - \frac{1}{\sqrt{n}}$, $(x^0, u^0, s^0) \in \mathcal{N}(\alpha)$;

$k := 0$;

Repita

Faça $\tau_k = \tau$, $\mu_k = x^{kT} s^k / n$;

Calcule a direção de Newton (9.25), ou seja, resolva o sistema

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta u^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau_k \mu_k e - X^k S^k e \end{bmatrix};$$

Faça

$$(x^{k+1}, u^{k+1}, s^{k+1}) := (x^k, u^k, s^k) + (\Delta x^k, \Delta u^k, \Delta s^k);$$

$$k := k + 1;$$

Até que $\mu_k < \epsilon$.

No lema 9.2, verificaremos que a medida do *gap* de dualidade, μ , é reduzida de forma linear a cada iteração do algoritmo de trajetória central de passos curtos. Este resultado é fundamental para que comprovemos a complexidade polinomial do algoritmo. O resultado do próximo lema será utilizado na sua demonstração.

Lema 9.1 *A direção $(\Delta x, \Delta u, \Delta s)$ definida por (9.25) satisfaz a*

$$\Delta x^T \Delta s = 0.$$

Demonstração

Das duas primeiras equações de (9.25), temos:

$$\begin{aligned} A^T \Delta u &= \Delta s, \\ A \Delta x &= 0. \end{aligned}$$

Portanto,

$$\Delta x^T \Delta s = \Delta x^T A^T \Delta u = (A \Delta x)^T \Delta u = 0. \quad \blacksquare$$

Lema 9.2 *Seja $(x, u, s) \in \mathcal{N}(\alpha)$, $(\Delta x, \Delta u, \Delta s)$, o passo de Newton dado por (9.25), $(\tilde{x}, \tilde{u}, \tilde{s}) := (x, u, s) + \sigma(\Delta x, \Delta u, \Delta s)$, para $\sigma \in [0, 1]$ e $\tilde{\mu} := \tilde{x}^T \tilde{s}/n$. Então*

$$\tilde{\mu} = (1 - \sigma(1 - \tau))\mu$$

Demonstração

$$\begin{aligned} n\tilde{\mu} &= \tilde{x}^T \tilde{s} \\ &= (x + \sigma \Delta x)^T (s + \sigma \Delta s) \\ &= x^T s + \sigma \Delta x^T s + \sigma x^T \Delta s + \sigma^2 \Delta x^T \Delta s \\ &= x^T s + \sigma(\Delta x^T s + x^T \Delta s), \end{aligned} \tag{9.29}$$

onde a última igualdade utiliza o resultado do lema anterior.

Da última equação de (9.25), temos:

$$S \Delta x + X \Delta s = \tau \mu e - X S e.$$

Multiplicando esta expressão por e^T e observando que $\mu e^T e = \mu n$, temos:

$$s^T \Delta x + x^T \Delta s = \tau n \mu - x^T s = (\tau - 1)n\mu. \tag{9.30}$$

Das relações (9.29) e (9.30), temos então

$$n\tilde{\mu} = (1 - \sigma(1 - \tau))n\mu.$$

Finalmente, dividindo-se esta última relação por n , obtem-se o resultado do lema. \blacksquare

Com o resultado obtido no lema anterior, comprovamos a redução linear no parâmetro μ , sempre que um passo é dado na direção de Newton. Esta redução é proporcional ao tamanho do passo, o qual é medido pelo parâmetro $\sigma \in [0, 1]$. No caso particular do algoritmo de trajetória central de passos

curtos, um passo unitário é dado na direção de Newton a cada iteração. Neste caso então, ao substituírmos o valor $\sigma = 1$ no resultado do lema, verificamos facilmente que a redução do parâmetro μ , a cada iteração do algoritmo, é dada por

$$\tilde{\mu} = \tau\mu. \quad (9.31)$$

Teorema 9.2 *Suponha que $(x^0, s^0, \mu_0) \in \mathcal{N}(\alpha)$ e*

$$\mu_0 \leq \frac{1}{\epsilon^\beta}, \quad (9.32)$$

para alguma constante positiva β .

Então, o algoritmo de trajetória central de passos curtos pára em $O(\sqrt{n} \log \frac{1}{\epsilon})$ iterações.

Demonstração

Utilizando-se a expressão (9.31) e o valor atribuído a τ no algoritmo, temos:

$$\begin{aligned} \mu_k &= \left(1 - \frac{1}{\sqrt{n}}\right) \mu_{k-1} \\ &= \left(1 - \frac{1}{\sqrt{n}}\right)^k \mu_0. \end{aligned}$$

Sendo assim, o algoritmo pára quando k é tal que

$$\left(1 - \frac{1}{\sqrt{n}}\right)^k \mu_0 < \epsilon.$$

Aplicando-se a função logaritmica a esta relação, temos

$$k \log \left(1 - \frac{1}{\sqrt{n}}\right) + \log \mu_0 < \log \epsilon.$$

Considerando (9.32), observamos que a relação acima será satisfeita se

$$k \log \left(1 - \frac{1}{\sqrt{n}}\right) + \beta \log \frac{1}{\epsilon} < \log \epsilon.$$

Como $\log(1 + \gamma) < \gamma$, para todo $\gamma > -1$, a relação acima continuará sendo verificada se

$$-k \frac{1}{\sqrt{n}} \leq \log \epsilon - \beta \log \frac{1}{\epsilon},$$

ou, equivalentemente,

$$-k \frac{1}{\sqrt{n}} \leq -(1 + \beta) \log \frac{1}{\epsilon},$$

O critério de convergência do algoritmo, $\mu_k < \epsilon$, é satisfeito então se

$$k \geq \sqrt{n}(1 + \beta) \log \frac{1}{\epsilon}. \quad \blacksquare$$

Uma vez demonstrada a complexidade polinomial do algoritmo, resta-nos apenas verificar que ao tomar-se um passo na direção de Newton (9.25), a partir de um dado ponto $(x, u, s) \in \mathcal{N}(\alpha)$, a nova iteração $(x, u, s) + (\Delta x, \Delta u, \Delta s)$ será também um ponto pertencente a $\mathcal{N}(\alpha)$.

No restante da análise do algoritmo, estaremos, portanto, interessados em fornecer um limite superior para a norma $\|\Delta X \Delta S e\|$. Este limite constitui um ingrediente importante na análise. Para entendermos tal importância, devemos lembrar que a cada iteração do algoritmo, partimos de um ponto (x, u, s) e calculamos uma aproximação para o passo que nos levaria ao ponto (x_μ, u_μ, s_μ) da trajetória central. Esta aproximação é estabelecida pela linearização da terceira equação de (9.16), a qual dá origem ao passo de Newton (9.25). Nesta linearização, o termo $\Delta X \Delta S e$ é justamente o termo desprezado. Conseqüentemente, a eficiência do passo de Newton estará associada ao tamanho de sua norma. Mantendo-a dentro de um certo limite, garantimos que a aproximação dada pelo passo de Newton, apesar de não nos levar exatamente para um ponto na trajetória central, nos leva para um ponto próximo a ela, onde esta proximidade é medida pela vizinhança $\mathcal{N}(\alpha)$. O lema a seguir formaliza o objeto deste parágrafo.

Lema 9.3 *Seja $(x, u, s) \in \mathcal{N}(\alpha)$; $(\tilde{x}, \tilde{u}, \tilde{s}) := (x, u, s) + \sigma(\Delta x, \Delta u, \Delta s)$ para $\sigma \in [0, 1]$, onde $(\Delta x, \Delta u, \Delta s)$ é o passo de Newton dado por (9.25); e $\tilde{\mu} := \frac{\tilde{x}^T \tilde{s}}{n}$. Então*

$$\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq (1 - \sigma)\alpha\mu + \sigma^2 \|\Delta X \Delta S e\|$$

Demonstração

$$\begin{aligned} \|\tilde{X} \tilde{S} e - \tilde{\mu} e\| &= \|X S e + \sigma S \Delta X e + \sigma X \Delta S e + \sigma^2 \Delta X \Delta S e - \tilde{\mu} e\| \\ &= \|X S e + \sigma(\tau \mu e - X S e) + \sigma^2 \Delta X \Delta S e - \tilde{\mu} e\| \quad (9.33) \\ &= \|X S e + \sigma(\tau \mu e - X S e) + \sigma^2 \Delta X \Delta S e\| \end{aligned}$$

$$- (1 - \sigma(1 - \tau))\mu e\| \quad (9.34)$$

$$\begin{aligned} &= \|(1 - \sigma)(XSe - \mu e) + \sigma^2 \Delta X \Delta Se\| \\ &\leq (1 - \sigma)\|XSe - \mu e\| + \sigma^2 \|\Delta X \Delta Se\| \\ &\leq (1 - \sigma)\alpha\mu + \sigma^2 \|\Delta X \Delta Se\| \end{aligned} \quad (9.35)$$

onde (9.33) utiliza a terceira equação de (9.25), (9.34) utiliza o resultado do lema 9.2 e (9.35) decorre de $(x, u, s) \in \mathcal{N}(\alpha)$. ■

Uma vez apresentado o lema anterior, resta-nos agora verificar que a medida $\|\Delta X \Delta Se\|$ é pequena o suficiente para garantir que o ponto $(\tilde{x}, \tilde{u}, \tilde{s})$ pertença à $\mathcal{N}(\alpha)$. Este resultado será demonstrado no teorema 9.3. Os lemas apresentados a seguir constituem ferramentas a serem utilizadas na sua demonstração.

Lema 9.4 *Se $u, v \in \mathbb{R}^p$ são tais que $u^T v = 0$, então*

$$\|UVe\| \leq \frac{1}{\sqrt{8}}\|u + v\|^2.$$

Demonstração

Considerando a igualdade

$$u_i v_i = \frac{1}{4} \left((u_i + v_i)^2 - (u_i - v_i)^2 \right),$$

para todo $i = 1, \dots, p$, temos:

$$\begin{aligned} \|UVe\|^2 &= \sum_{i=1}^p (u_i v_i)^2 \\ &= \frac{1}{16} \sum_{i=1}^p \left((u_i + v_i)^2 - (u_i - v_i)^2 \right)^2 \\ &\leq \frac{1}{16} \sum_{i=1}^p \left((u_i + v_i)^4 + (u_i - v_i)^4 \right) \end{aligned} \quad (9.36)$$

$$\leq \frac{1}{16} \left[\left(\sum_{i=1}^p (u_i + v_i)^2 \right)^2 + \left(\sum_{i=1}^p (u_i - v_i)^2 \right)^2 \right] \quad (9.37)$$

$$\begin{aligned} &= \frac{1}{16} \left(\|u + v\|^4 + \|u - v\|^4 \right) \\ &= \frac{1}{16} \left(2\|u + v\|^4 \right) \end{aligned} \quad (9.38)$$

$$= \frac{1}{8} \left(\|u + v\|^4 \right) \quad (9.39)$$

onde:

(9.36) se verifica porque se temos $a, b \geq 0$, então $(a - b)^2 \leq a^2 + b^2$;

(9.37) utiliza a relação

$$\sum_{i=1}^p a_i^2 \leq \left(\sum_{i=1}^n a_i \right)^2,$$

para $a_i \geq 0$, $i = 1, \dots, p$, a qual é um caso particular da desigualdade de Holder.

(9.38) se verifica porque u e v são ortogonais ($u^T v = 0$), e conseqüentemente $\|u + v\| = \|u - v\|$. ■

Lema 9.5 *Seja $(x, u, s) \in \mathcal{N}(\alpha)$ e $(\Delta x, \Delta u, \Delta s)$, o passo de Newton dado por (9.25). Então*

$$\|\Delta X \Delta S e\| \leq \frac{\alpha^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \alpha)} \mu$$

Demonstração

Iniciamos a demonstração aplicando uma mudança de escala sobre a terceira equação de (9.25), com o intuito de escrevê-la de forma mais conveniente. A mudança de escala é definida ao multiplicar-se a equação pela matriz $(XS)^{-1/2}$. Considerando D , a matriz diagonal cujos elementos da diagonal são dados por $\sqrt{x_i/s_i}$ para todo $i = 1, \dots, n$, ou seja, $D = S^{-1/2} X^{1/2}$, o resultado deste produto pode ser escrito como

$$D^{-1} \Delta x + D \Delta s = (XS)^{-1/2} (-XSe + \tau \mu e). \quad (9.40)$$

Observando que $\Delta X \Delta S = (D \Delta X)(D^{-1} \Delta S)$ e utilizando (9.40) e os resultados dos lemas 9.1 e 9.4, temos:

$$\begin{aligned} \|\Delta X \Delta S e\| &= \|(D \Delta X)(D^{-1} \Delta S) e\| \\ &\leq \frac{1}{\sqrt{8}} \|D \Delta x + D^{-1} \Delta s\|^2 \\ &= \frac{1}{\sqrt{8}} \|(XS)^{-1/2} (-XSe + \tau \mu e)\|^2 \\ &= \frac{1}{\sqrt{8}} \sum_{i=1}^n \frac{(-x_i s_i + \tau \mu)^2}{x_i s_i} \\ &\leq \frac{\sum_{i=1}^n (-x_i s_i + \tau \mu)^2}{\sqrt{8}(1 - \alpha) \mu} \\ &= \frac{\|XSe - \tau \mu e\|^2}{\sqrt{8}(1 - \alpha) \mu}, \end{aligned} \quad (9.41)$$

onde a última desigualdade se verifica porque $(x, u, s) \in \mathcal{N}(\alpha)$. Consequentemente, $|x_i s_i - \mu| \leq \alpha\mu$, e $x_i s_i \geq (1 - \alpha)\mu$, para todo $i = 1, \dots, n$.

Temos ainda que

$$\begin{aligned}
\|XSe - \tau\mu e\|^2 &= \|(XSe - \mu e) + (1 - \tau)\mu e\|^2 \\
&= \|XSe - \mu e\|^2 + 2(1 - \tau)\mu e^T(XSe - \mu e) \\
&\quad + (1 - \tau)^2 \mu^2 e^T e \\
&= \|XSe - \mu e\|^2 + 2(1 - \tau)\mu(x^T s - \mu n) \\
&\quad + (1 - \tau)^2 \mu^2 n \\
&= \|XSe - \mu e\|^2 + (1 - \tau)^2 \mu^2 n \\
&\leq \alpha^2 \mu^2 + (1 - \tau)^2 \mu^2 n,
\end{aligned} \tag{9.42}$$

onde a última igualdade se verifica porque $\mu = x^T s / n$ e a última desigualdade se verifica porque $(x, u, s) \in \mathcal{N}(\alpha)$.

Substituindo (9.42) em (9.41) temos o resultado do lema. \blacksquare

Lema 9.6 *Seja $(x, u, s) \in \mathcal{N}(\alpha)$; $(\tilde{x}, \tilde{u}, \tilde{s}) := (x, u, s) + \sigma(\Delta x, \Delta u, \Delta s)$ para $\sigma \in [0, 1]$, onde $(\Delta x, \Delta u, \Delta s)$ é o passo de Newton dado por (9.25); e $\tilde{\mu} := \frac{\tilde{x}^T \tilde{s}}{n}$. Então, considerando-se que α e τ assumem os valores adotados no algoritmo 9.2, temos:*

$$\|\tilde{X}\tilde{S}e - \tilde{\mu}e\| \leq \alpha\tilde{\mu},$$

Demonstração

$$\|\tilde{X}\tilde{S}e - \tilde{\mu}e\| \leq (1 - \sigma)\alpha\mu + \sigma^2 \|\Delta X \Delta S e\| \tag{9.43}$$

$$\leq (1 - \sigma)\alpha\mu + \sigma^2 \frac{\alpha^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \alpha)} \mu \tag{9.44}$$

$$\leq (1 - \sigma)\alpha\mu + \sigma^2 \tau \alpha \mu \tag{9.45}$$

$$\leq (1 - \sigma + \sigma\tau)\alpha\mu \tag{9.46}$$

$$= \alpha\tilde{\mu}, \tag{9.47}$$

onde:

(9.43) utiliza o resultado do lema 9.3,

(9.44) utiliza o resultado do lema 9.5,

(9.45) pode ser facilmente verificada quando α e τ assumem os valores adotados no algoritmo 9.2,

(9.46) se verifica porque $\sigma \in [0, 1]$, e

(9.47) utiliza o resultado do lema 9.2. ■

Teorema 9.3 *Seja $(x, u, s) \in \mathcal{N}(\alpha)$ e $(\Delta x, \Delta u, \Delta s)$, o passo de Newton dado por (9.25). Então*

$$(\tilde{x}, \tilde{u}, \tilde{s}) := (x, u, s) + \sigma(\Delta x, \Delta u, \Delta s) \in \mathcal{N}(\alpha)$$

para todo $\sigma \in [0, 1]$.

Demonstração

O resultado do lema 9.6 indica que $\delta(\tilde{x}, \tilde{s}, \tilde{\mu}) \leq \alpha$. Resta-nos apenas mostrar que $(\tilde{x}, \tilde{u}, \tilde{s}) \in \mathcal{F}^0$. As restrições $A\tilde{x} = b$ e $A^T\tilde{u} - \tilde{s} = c$ são facilmente verificadas já que, de (9.25), temos que $A\Delta x = 0$ e $A^T\Delta u - \Delta s = 0$. E, finalmente, a não negatividade de \tilde{x} e \tilde{s} é garantida ao considerar-se que se $\delta(\tilde{x}, \tilde{s}, \tilde{\mu}) \leq \alpha$, então $|\tilde{x}_i\tilde{s}_i - \tilde{\mu}| \leq \alpha\tilde{\mu}$, ou seja, $\tilde{x}_i\tilde{s}_i \geq (1 - \alpha)\tilde{\mu} = (1 - \alpha)(1 - \sigma(1 - \tau))\mu > 0$, já que $\alpha, \sigma, \tau \in [0, 1]$. Portanto, $(\tilde{x}, \tilde{s}) > 0$ para todo $\sigma \in [0, 1]$. ■

Exemplo 9.2 Resolver o problema de programação linear (9.13) utilizando o algoritmo de trajetória central de passos curtos.

É dada uma solução inicial interior viável para o problema em $\mathcal{N}(0.4)$: $x_1 = 1.6559$, $x_2 = 4.2932$, $u_1 = 2.2632$, $u_2 = 3.3758$, $u_3 = 1.3144$. Em seguida, reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, \dots, 5$ e s_j , $j = 1, \dots, 5$. A solução inicial interior viável correspondente é

$$\begin{aligned} x^0 &= (1.6559 \quad 4.2932 \quad 2.3441 \quad 1.7068 \quad 4.4458)^T, \\ u^0 &= (2.2632 \quad 3.3758 \quad 1.3144)^T, \\ s^0 &= (3.2065 \quad 1.005 \quad 2.26320 \quad 3.3758 \quad 1.3144)^T. \end{aligned}$$

O valor da função objetivo em x^0 é 26.4337 e o *gap* de dualidade inicial é dado por

$$\mu^0 = (x^{0T} s^0)/5 = 5.3067.$$

Para definir a direção de busca, resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} \Delta x^0 \\ \Delta u^0 \\ \Delta s^0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau\mu_0 e - X^0 S^0 e \end{bmatrix};$$

onde

$$\begin{aligned}
A &= \begin{pmatrix} 1 & & 1 & & \\ & 1 & & 1 & \\ 3 & 2 & & & 1 \end{pmatrix}, \\
X^0 &= \begin{pmatrix} 1.6559 & & & & \\ & 4.2932 & & & \\ & & 2.3441 & & \\ & & & 1.7068 & \\ & & & & 4.4458 \end{pmatrix}, \\
S^0 &= \begin{pmatrix} 3.2065 & & & & \\ & 1.005 & & & \\ & & 2.26320 & & \\ & & & 3.3758 & \\ & & & & 1.3144 \end{pmatrix}, \\
\tau &= 1 - \frac{1}{\sqrt{5}} = 0.5528 \quad e \\
\tau\mu_0 e - X^0 S^0 e &= (-2.3762 \quad -1.3798 \quad -2.3716 \quad -2.8283 \quad -2.9103)^T.
\end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}
\Delta x^0 &= (0.0342 \quad 0.7613 \quad -0.0342 \quad -0.7613 \quad -1.6251)^T, \\
\Delta u^0 &= (-0.9788 \quad -0.1513 \quad -0.1741)^T, \\
\Delta s^0 &= (-1.5012 \quad -0.4996 \quad -0.9788 \quad -0.1513 \quad -0.1741)^T.
\end{aligned}$$

O novo ponto é dado por

$$\begin{aligned}
x^1 = x^0 + \Delta x^0 &= (1.6901 \quad 5.0545 \quad 2.3099 \quad 0.9455 \quad 2.8207)^T, \\
u^1 = u^0 + \Delta u^0 &= (1.2844 \quad 3.2245 \quad 1.1403)^T, \\
s^1 = s^0 + \Delta s^0 &= (1.7054 \quad 0.5051 \quad 1.2844 \quad 3.2245 \quad 1.1403)^T.
\end{aligned}$$

O valor da função objetivo em x^1 é 30.3428 e o *gap* de dualidade é dado por

$$\mu^1 = (x^{1T} s^1)/5 = 2.9335.$$

A primeira iteração do algoritmo está completa. Como o *gap* de dualidade ainda é bastante grande, seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo de trajetória central de passos curtos, no qual consideramos a precisão $\epsilon = 10^{-4}$.

```

% Algoritmo de Trajetoria Central de Passos Curtos

n=5;
m=3;
k=0;
epsilon=10^(-4);
alfa = 0.4;
tal = 1 - (1/sqrt(n));
A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];
%ler solucao inicial sol=[x,u,s]:
sol=input('Entre com a solucao inicial sol=[x,u,s]:');
x=sol(1:5);
s=sol(9:13);
mi=(x'*s)/n;
while (mi>= epsilon)
    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ...
        diag(s) zeros(n,m) diag(x)];
    ld=[zeros(n,1); zeros(m,1) ; (tal*mi*ones(n,1))-(diag(x)*s)];
    dir=M\ld;
    sol=sol+dir;
    x=sol(1:5);
    s=sol(9:13);
    mi=(x'*s)/n;
    k=k+1;
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo, estão representados na tabela 9.2. A figura 9.9 mostra a trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema. ■

Comparando as figuras 9.9 e 9.8, observamos o quanto as iterações realizadas pelo algoritmo de trajetória central de passos curtos se aproximam da trajetória central definida para o problema (9.13). Verificamos que tanto a trajetória central quanto a curva definida pelas iterações do algoritmo convergem para o vértice ($x_1 = 2, x_2 = 6$), do poliedro que representa a região viável do problema. Este ponto é de fato a única solução ótima do problema considerado.

Vamos ilustrar com o exemplo a seguir o comportamento do algoritmo de trajetória central de passos curtos e da própria trajetória central no caso em que o problema de programação linear tem infinitas soluções ótimas.

k	0	1	2	3	4	5	6
x_1	1.6559	1.6901	1.7815	1.8866	1.9430	1.9693	1.9831
x_2	4.2932	5.0545	5.4902	5.7109	5.8367	5.9092	5.9497
k	7	8	9	10	11	12	13
x_1	1.9907	1.9949	1.9972	1.9984	1.9991	1.9995	1.9997
x_2	5.9721	5.9846	5.9915	5.9953	5.9974	5.9986	5.9992
k	14	15	16	17	18	19	
x_1	1.9999	1.9999	2.0000	2.0000	2.0000	2.0000	
x_2	5.9996	5.9998	5.9999	5.9999	6.0000	6.0000	

Tabela 9.2: Iterações geradas pelo Algoritmo de Trajetória Central de Passos Curtos

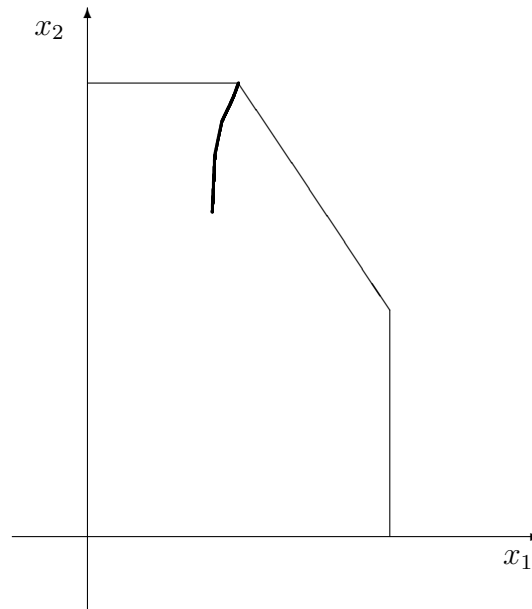


Figura 9.9: Algoritmo de Trajetória Central de Passos Curtos

Exemplo 9.3 Resolver o problema de programação linear abaixo utilizando o algoritmo de trajetória central de passos curtos.

$$\begin{aligned}
 (P) \quad & \text{maximizar} \quad 2x_1 - 2x_2 \\
 & \text{sujeito a:} \quad \begin{aligned} x_1 + x_2 &\leq 2 \\ x_1 - x_2 &\leq 1 \\ x_1, x_2 &\geq 0 \end{aligned}
 \end{aligned} \tag{9.48}$$

O dual de (P) é dado por

$$\begin{aligned}
 (D) \quad & \text{minimizar} \quad 2u_1 + u_2 \\
 & \text{sujeito a:} \quad \begin{aligned} u_1 + u_2 &\geq 2 \\ u_1 - u_2 &\geq -2 \\ u_1, u_2 &\geq 0 \end{aligned}
 \end{aligned} \tag{9.49}$$

É dada uma solução inicial interior viável para o problema em $\mathcal{N}(0.4)$: $x_1 = 0.5725$, $x_2 = 0.7818$, $u_1 = 5.4027$, $u_2 = 2.7760$. Em seguida, reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, 4$ e s_j , $j = 1, \dots, 4$, obtendo

$$\begin{aligned}
 (P) \quad & \text{maximizar} \quad 2x_1 - 2x_2 \\
 & \text{sujeito a:} \quad \begin{aligned} x_1 + x_2 + x_3 &= 2 \\ x_1 - x_2 + x_4 &= 1 \\ x_1, x_2, x_3, x_4 &\geq 0 \end{aligned}
 \end{aligned}$$

e

$$\begin{aligned}
 (D) \quad & \text{minimizar} \quad 2u_1 + u_2 \\
 & \text{sujeito a:} \quad \begin{aligned} u_1 + u_2 - s_1 &= 2 \\ u_1 - u_2 - s_2 &= -2 \\ u_1 - s_3 &= 0 \\ u_2 - s_4 &= 0 \\ s_1, s_2, s_3, s_4 &\geq 0 \end{aligned}
 \end{aligned}$$

A solução inicial interior viável correspondente é

$$\begin{aligned}
 x^0 &= (0.5725 \quad 0.7818 \quad 0.6458 \quad 1.2093)^T, \\
 u^0 &= (5.4027 \quad 2.7760)^T, \\
 s^0 &= (6.1787 \quad 4.6267 \quad 5.4027 \quad 2.7760)^T.
 \end{aligned}$$

O valor da função objetivo em x^0 é -0.4186 e o *gap* de dualidade inicial é dado por

$$\mu^0 = (x^{0T} s^0)/4 \approx 3.50.$$

k	0	1	2	3	4	5	6	7
x_1	0.5725	0.6766	1.1407	1.2046	1.2405	1.2574	1.2658	1.2701
x_2	0.7818	0.6524	0.3592	0.3172	0.2951	0.2847	0.2795	0.2769
k	8	9	10	11	12	13	14	15
x_1	1.2722	1.2732	1.2738	1.2740	1.2742	1.2742	1.2743	1.2743
x_2	0.2756	0.2749	0.2746	0.2745	0.2744	0.2743	0.2743	0.2743

Tabela 9.3: Exemplo 9.3 - Iterações geradas pelo Algoritmo de Trajetória Central de Passos Curtos - Variáveis Primais

k	0	1	2	3	4	5	6	7
u_1	5.4027	1.1444	0.9435	0.4788	0.2426	0.1211	0.0606	0.0303
u_2	2.7760	1.2596	1.4259	1.6923	1.8431	1.9218	1.9609	1.9804
k	8	9	10	11	12	13	14	15
u_1	0.0151	0.0076	0.0038	0.0019	0.0009	0.0005	0.0002	0.0001
u_2	1.9902	1.9951	1.9976	1.9988	1.9994	1.9997	1.9998	1.9999

Tabela 9.4: Exemplo 9.3 - Iterações geradas pelo Algoritmo de Trajetória Central de Passos Curtos - Variáveis Duais

Utilizamos para resolver o problema, o mesmo programa em MATLAB apresentado no exemplo anterior, substituindo apenas os dados de entrada. Os valores de x_1 e x_2 obtidos pelo algoritmo estão representados na tabela 9.3. A figura 9.10 mostra a trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema.

Observamos na figura, que a curva definida pelas iterações do algoritmo não converge para nenhum dos vértices do politopo que representa a região viável do problema, isto é, ela não converge para nenhuma solução básica viável do problema. Na verdade esta curva está convergindo para o centro analítico da aresta do politopo, que corresponde ao segmento de reta que une os vértices $(x_1 = 1, x_2 = 0)$ e $(x_1 = 1.5, x_2 = 0.5)$. Todos os pontos neste segmento de reta são soluções ótimas do problema. Este exemplo tem uma segunda característica interessante. O problema dual definido em (9.49), também tem apenas duas variáveis, u_1 e u_2 e portanto pudemos plotar na figura 9.11 as iterações realizadas pelo algoritmo no espaço das variáveis duais. Observamos que o problema dual tem apenas uma solução ótima, $u_1 = 0, u_2 = 2$. A curva representada na figura 9.11 está portanto convergindo para a única solução ótima do problema dual. O valor das variáveis nestas iterações está representado na tabela 9.4. ■

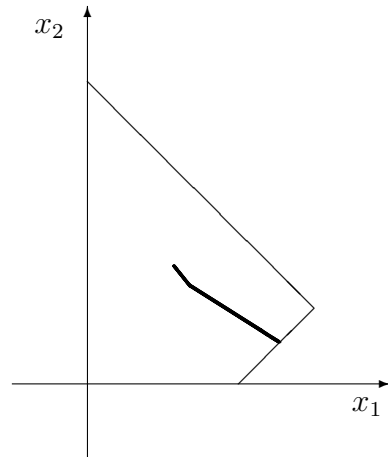


Figura 9.10: Exemplo 9.3 - Algoritmo de Trajetória Central de Passos Curtos
- Variáveis Primais

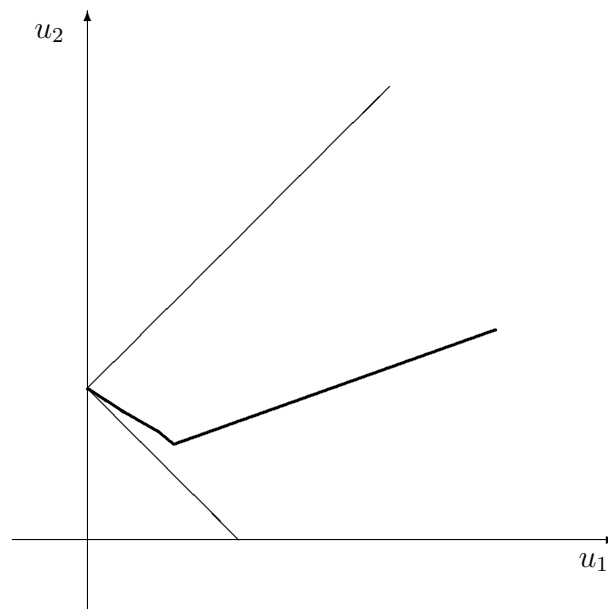


Figura 9.11: Exemplo 9.3 - Algoritmo de Trajetória Central de Passos Curtos
- Variáveis Duais

O Algoritmo Preditor-Corretor

Em (9.25) definimos a direção de Newton para $\tau \in [0, 1]$. Observamos que quando τ assume os valores extremos 0 e 1, a direção de Newton é conhecida respectivamente como direção afim-escala e direção de centralização. A primeira aponta aproximadamente para o ótimo dos problemas P e D , no entanto, ao caminharmos nela, podemos nos aproximar muito rapidamente da fronteira do ortante não negativo. Na verdade, se tomarmos um passo unitário na direção afim-escala a partir de um ponto em $\mathcal{N}(\alpha)$, o novo ponto pode até mesmo não satisfazer as restrições de não negatividade das variáveis x e s . Por outro lado, ao tomarmos um passo unitário na direção de centralização a partir de um ponto em $\mathcal{N}(\alpha)$, temos a garantia de que o novo ponto também pertencerá a $\mathcal{N}(\alpha)$, ou seja, ele será viável e estará longe da fronteira do ortante. No entanto, este passo mantém aproximadamente constante o *gap* de dualidade. No algoritmo de trajetória central de passos curtos atribuímos para τ um valor menor do que 1, sem no entanto, perder a garantia fornecida pela direção de centralização. Ou seja, continuamos exigindo que ao tomarmos um passo unitário na direção de Newton a partir de um ponto em $\mathcal{N}(\alpha)$, o novo ponto também pertencerá à vizinhança. Esta exigência faz com que o valor de τ continue muito próximo de 1 e, conseqüentemente, a convergência do algoritmo em direção ao ótimo do problema é, em geral, lenta, como pode ser visto nos exemplos da seção anterior.

Resultados melhores foram obtidos na prática com a aplicação do chamado algoritmo preditor-corretor. Neste algoritmo, trabalhamos com duas vizinhanças da trajetória central definidas por (9.27) para dois diferentes valores de α . Tipicamente utiliza-se $\alpha_1 = 0.25$ e $\alpha_2 = 0.5$, de forma que a primeira vizinhança é um subconjunto da segunda. Dois tipos diferentes de iteração se intercalam então, da seguinte forma:

Iterações Ímpares: Constituem o chamado passo corretor no qual partimos de um ponto em $\mathcal{N}(\alpha_2)$ e tomamos um passo unitário na direção de centralização. É possível mostrar que o novo ponto obtido pertence a $\mathcal{N}(\alpha_1)$. Este passo tem então como objetivo, centralizar.

Iterações Pares: Constituem o chamado passo preditor no qual partimos de um ponto em $\mathcal{N}(\alpha_1)$ e caminhamos na direção afim-escala o máximo possível sem, no entanto, sair da vizinhança $\mathcal{N}(\alpha_2)$. Mais especificamente, partimos do ponto $(x, u, s) \in \mathcal{N}(\alpha_1)$ e tomamos o novo ponto $(\tilde{x}, \tilde{u}, \tilde{s}) := (x, u, s) + \sigma(\Delta x, \Delta u, \Delta s)$, onde $(\Delta x, \Delta u, \Delta s)$ é o passo de Newton (9.25) calculado para $\tau = 0$ e $\sigma \in [0, 1]$ assume o maior valor para o qual $(\tilde{x}, \tilde{u}, \tilde{s}) \in \mathcal{N}(\alpha_2)$. Este passo tem como objetivo, diminuir o *gap* de dualidade.

O algoritmo preditor-corretor mantém a complexidade de $O(\sqrt{n} \log \frac{1}{\epsilon})$ do algoritmo de trajetória central de passos curtos. Ele foi apresentado pela primeira vez na forma mencionada por Mizuno, Todd e Ye em [MiToYe 93].

Exemplo 9.4 Resolver o problema de programação linear (9.13) utilizando o algoritmo preditor-corretor.

É dada uma solução inicial interior viável para o problema em $\mathcal{N}(0.5)$: $x_1 = 1.6569$, $x_2 = 3.6706$, $u_1 = 3.8104$, $u_2 = 4.1066$, $u_3 = 1.7468$. Em seguida, reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, \dots, 5$ e s_j , $j = 1, \dots, 5$. A solução inicial interior viável correspondente é

$$\begin{aligned} x^0 &= (1.6569 \quad 3.6706 \quad 2.3431 \quad 2.3294 \quad 5.6879)^T, \\ u^0 &= (3.8104 \quad 4.1066 \quad 1.7468)^T, \\ s^0 &= (6.0508 \quad 2.6003 \quad 3.8104 \quad 4.1066 \quad 1.7468)^T. \end{aligned}$$

O valor da função objetivo em x^0 é 23.3237 e o *gap* de dualidade inicial é dado por

$$\mu^0 = (x^{0T} s^0)/5 = 9.6000.$$

Iniciamos o algoritmo com o passo corretor. Para definir a direção de busca deste passo, resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} \Delta x^0 \\ \Delta u^0 \\ \Delta s^0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu_0 e - X^0 S^0 e \end{bmatrix};$$

onde

$$\begin{aligned} A &= \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ 3 & 2 & & & \\ & & & & \\ & & & & 1 \end{pmatrix}, \\ X^0 &= \begin{pmatrix} 1.6569 & & & & \\ & 3.6706 & & & \\ & & 2.3431 & & \\ & & & 2.3294 & \\ & & & & 5.6879 \end{pmatrix}, \\ S^0 &= \begin{pmatrix} 6.0508 & & & & \\ & 2.6003 & & & \\ & & 3.8104 & & \\ & & & 4.1066 & \\ & & & & 1.7468 \end{pmatrix}, \\ \mu_0 e - X^0 S^0 e &= (-0.4258 \quad 0.0555 \quad 0.6721 \quad 0.0340 \quad -0.3357)^T. \end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}\Delta x^0 &= (-0.0598 \quad 0.0618 \quad 0.0598 \quad -0.0618 \quad 0.0557)^T, \\ \Delta u^0 &= (0.1896 \quad 0.1236 \quad -0.0761)^T, \\ \Delta s^0 &= (-0.0387 \quad -0.0287 \quad 0.1896 \quad 0.1236 \quad -0.0761)^T.\end{aligned}$$

O novo ponto, que pertence a $\mathcal{N}(0.25)$ é dado por

$$\begin{aligned}x^1 = x^0 + \Delta x^0 &= (1.5972 \quad 3.7324 \quad 2.4028 \quad 2.2676 \quad 5.7436)^T, \\ u^1 = u^0 + \Delta u^0 &= (4.0000 \quad 4.2302 \quad 1.6707)^T, \\ s^1 = s^0 + \Delta s^0 &= (6.0121 \quad 2.5716 \quad 4.0000 \quad 4.2302 \quad 1.6707)^T.\end{aligned}$$

O valor da função objetivo em x^1 é 23.4536 e o *gap* de dualidade é dado por

$$\mu^1 = (x^{1T} s^1)/5 = 9.6000.$$

Como já era esperado pela análise teórica, não houve diminuição no *gap* de dualidade no passo corretor.

Partimos agora para o passo preditor. Para definir a direção de busca deste passo, resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^1 & 0 & X^1 \end{bmatrix} \begin{bmatrix} \Delta x^1 \\ \Delta u^1 \\ \Delta s^1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -X^1 S^1 e \end{bmatrix};$$

onde

$$\begin{aligned}X^1 &= \begin{pmatrix} 1.5972 & & & & \\ & 3.7324 & & & \\ & & 2.4028 & & \\ & & & 2.2676 & \\ & & & & 5.7436 \end{pmatrix} e \\ S^1 &= \begin{pmatrix} 6.0121 & & & & \\ & 2.5716 & & & \\ & & 4.0000 & & \\ & & & 4.2302 & \\ & & & & 1.6707 \end{pmatrix}.\end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}\Delta x^1 &= (0.0903 \quad 1.3024 \quad -0.0903 \quad -1.3024 \quad -2.8758)^T, \\ \Delta u^1 &= (-3.8496 \quad -1.8006 \quad -0.8342)^T, \\ \Delta s^1 &= (-6.3521 \quad -3.4689 \quad -3.8496 \quad -1.8006 \quad -0.8342)^T.\end{aligned}$$

Uma vez calculada a direção a seguir, devemos agora calcular o tamanho do passo que pode ser dado nesta direção sem que se saia da vizinhança $\mathcal{N}(0.5)$. Iniciamos este cálculo pelo teste da razão, uma vez que o novo ponto deve também ser interior viável, ou seja, as componentes de x e s devem se manter positivas.

$$\begin{aligned}\lambda_1 &= \min_{j=1,\dots,n} \{-x_j^1/\Delta x_j^1 | \Delta x_j^1 < 0\} = 1.7411, \\ \lambda_2 &= \min_{j=1,\dots,n} \{-s_j^1/\Delta s_j^1 | \Delta s_j^1 < 0\} = 0.7413, \\ \lambda &= 0.995 * \min\{\lambda_1, \lambda_2\} = 0.7376.\end{aligned}$$

Como $(x_1 + \lambda\Delta x_1, u_1 + \lambda\Delta u_1, s_1 + \lambda\Delta s_1) \notin \mathcal{N}(0.5)$, calculamos o menor q inteiro e positivo, tal que

$$(x_1 + (0.95)^q \lambda \Delta x_1, u_1 + (0.95)^q \lambda \Delta u_1, s_1 + (0.95)^q \lambda \Delta s_1) \in \mathcal{N}(0.5).$$

Encontramos $q = 5$ e consequentemente o novo ponto é

$$\begin{aligned}x^2 &= x^1 + (0.95)^5 \lambda \Delta x^1 = (1.6487 \quad 4.4758 \quad 2.3513 \quad 1.5242 \quad 4.1022)^T, \\ u^2 &= u^1 + (0.95)^5 \lambda \Delta u^1 = (1.8028 \quad 3.2025 \quad 1.1946)^T, \\ s^2 &= s^1 + (0.95)^5 \lambda \Delta s^1 = (2.3866 \quad 0.5917 \quad 1.8028 \quad 3.2025 \quad 1.1946)^T.\end{aligned}$$

O valor da função objetivo em x^2 é 27.3251 e o *gap* de dualidade é dado por

$$\mu^1 = (x^{2T} s^2)/5 = 4.1208.$$

Observamos que houve uma diminuição no *gap* de dualidade no passo preditor.

A primeira iteração do algoritmo está completa. Como o *gap* de dualidade ainda é bastante grande, seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo preditor-corretor, no qual consideramos a precisão $\epsilon = 10^{-4}$.

```
%Algoritmo Preditor-Corretor
```

```
n=5;
m=3;
k=0;
epsilon=10^(-4);
alfa1=0.25;
alfa2=0.5;
```

```

A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];

%ler solucao inicial sol=[x,u,s]:
sol=input('Entre com a solucao inicial sol=[x,u,s]:');
x=sol(1:5);
s=sol(9:13);
mi=(x'*s)/n;

while (mi>= epsilon)
    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ...
        diag(s) zeros(n,m) diag(x)];
    ld=[zeros(n,1); zeros(m,1) ; (mi*ones(n,1))-(diag(x)*s)];
    dir=M\ld;
    sol=sol+dir;
    x=sol(1:5);
    s=sol(9:13);
    mi=(x'*s)/n;

    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ...
        diag(s) zeros(n,m) diag(x)];
    ld=[zeros(n,1); zeros(m,1) ; -(diag(x)*s)];
    dir=M\ld;

    aux=[10];
    for i=1:5
        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end
    for i=9:13
        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end

    perc = min(1,0.995*min(aux));
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    s1=sol1(9:13);
    mi1=(x1'*s1)/n;

```

	k	1	2	3	4	5	6
passo	x_1	1.5972	1.6772	1.8484	1.9682	1.9978	2.0000
corretor	x_2	3.7324	4.8167	5.5683	5.9070	5.9934	6.0000
passo	x_1	1.6487	1.7621	1.9557	1.9983	2.0000	2.0000
preditor	x_2	4.4758	5.5089	5.8969	5.9923	6.0000	6.0000

Tabela 9.5: Iterações geradas pelo Algoritmo Preditor-Corrtetor

```

while (norm(x1.*s1-mi1*ones(n,1))> alfa2*mi1)
    perc=0.95*perc;
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    s1=sol1(9:13);
    mi1=(x1'*s1)/n;
end
sol = sol1;
x=x1;
s=s1;
mi=mi1;
k=k+1;
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo a cada iteração k estão representados na tabela 9.5. A trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema, está representada na figura 9.12. ■

O Algoritmo de Trajetória Central de Passos Longos

A utilização da norma Euclideana na definição de δ em (9.26) restringe bastante o conjunto de pontos que pertencem à vizinhança $\mathcal{N}(\alpha)$. Mesmo para valores de α próximos de 1, o conjunto de pontos em \mathcal{F}^0 é bem maior que o conjunto de pontos em $\mathcal{N}(\alpha)$. Como resultado, os algoritmos que permanecem sempre nesta vizinhança, podem caminhar pouco a cada iteração em direção ao ótimo.

Algoritmos práticos utilizam outras normas para definir vizinhanças da trajetória central mais abrangentes. As duas vizinhanças mais interessantes são:

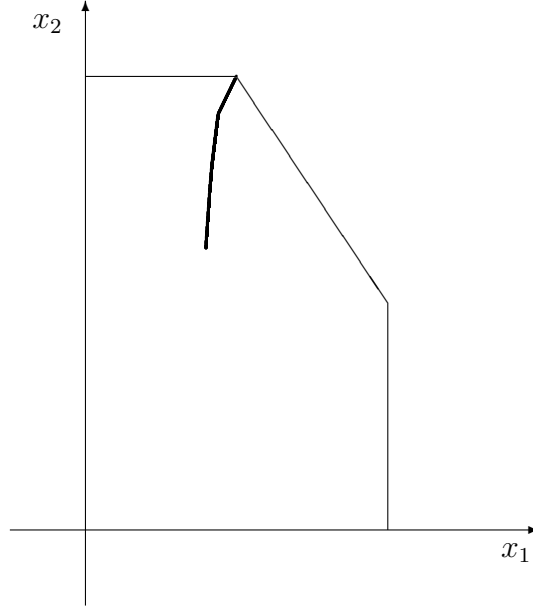


Figura 9.12: Algoritmo Preditor-Corretor

$$\mathcal{N}_\infty(\alpha) := \bigcup_{\mu \in (0, \infty)} \{(x, u, s) \in \mathcal{F}^0 \mid \delta_\infty(x, s, \mu) \leq \alpha\},$$

onde

$$\delta_\infty(x, s, \mu) := \frac{1}{\mu} \|XSe - \mu e\|_\infty,$$

e

$$\mathcal{N}_{-\infty}(\alpha) := \bigcup_{\mu \in (0, \infty)} \{(x, u, s) \in \mathcal{F}^0 \mid \delta_{-\infty}(x, s, \mu) \leq \alpha\},$$

onde

$$\delta_{-\infty}(x, s, \mu) := \frac{1}{\mu} \|XSe - \mu e\|_{-\infty}, \quad (9.50)$$

e dado $v \in \mathbb{R}^N$, temos por definição que $\|v\|_{-\infty} \leq \beta$ se e somente se $v_i \geq -\beta$ para todo $i = 1, \dots, N$.

Podemos verificar facilmente que para um dado α estas duas vizinhanças contêm um número de pontos viáveis dos problemas (P) e (D) consideravelmente maior que $\mathcal{N}(\alpha)$. A vizinhança $\mathcal{N}_{-\infty}(\alpha)$, em particular, torna-se bem próxima de \mathcal{F}^0 à medida que α tende para 1.

Algoritmos de trajetória central de passos longos utilizam estas vizinhanças mais abrangentes. Neles, o parâmetro τ recebe um valor menor do que o recebido no algoritmo de trajetória central de passos curtos e, portanto, a tentativa de decrescer o *gap* de dualidade é mais agressiva a cada iteração. Como eles têm mais espaço para trabalhar, o progresso em direção ao ótimo

é mais rápido. Ao tomarmos um valor pequeno para τ , no entanto, não garantimos mais que o ponto resultante de um passo unitário na direção de Newton (9.25) continue pertencendo à vizinhança da trajetória. Neste caso, utilizamos um procedimento análogo ao descrito no passo preditor do algoritmo preditor-corretor, no qual realizamos uma busca linear que estabelece o maior passo permitido na direção de Newton, para que o novo ponto alcançado permaneça na vizinhança. Os algoritmos de trajetória central de passos longos têm uma melhor performance na prática do que os algoritmos de passos curtos. Sua complexidade, no entanto, aumenta para $O(n \log \frac{1}{\epsilon})$ iterações.

Exemplo 9.5 Resolver o problema de programação linear (9.13) utilizando o algoritmo de trajetória central de passos longos.

É dada uma solução inicial interior viável para o problema em $\mathcal{N}_{-\infty}(0.8)$: $x_1 = 1.1307$, $x_2 = 3.3837$, $u_1 = 2.5332$, $u_2 = 3.9880$, $u_3 = 1.9028$. Em seguida, reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, \dots, 5$ e s_j , $j = 1, \dots, 5$. A solução inicial interior viável correspondente é

$$\begin{aligned} x^0 &= (1.1307 \quad 3.3837 \quad 2.8693 \quad 2.6163 \quad 7.8404)^T, \\ u^0 &= (2.5332 \quad 3.9880 \quad 1.9028)^T, \\ s^0 &= (5.2415 \quad 2.7936 \quad 2.5332 \quad 3.9880 \quad 1.9028)^T. \end{aligned}$$

O valor da função objetivo em x^0 é 20.3108 e o *gap* de dualidade inicial é dado por

$$\mu^0 = (x^{0T} s^0)/5 = 9.6000.$$

Para definir a direção de busca resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} \Delta x^0 \\ \Delta u^0 \\ \Delta s^0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau \mu_0 e - X^0 S^0 e \end{bmatrix};$$

onde

$$A = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ 3 & 2 & & & \\ & & & 1 & \end{pmatrix},$$

$$\begin{aligned}
X^0 &= \begin{pmatrix} 1.1307 & & & & \\ & 3.3837 & & & \\ & & 2.8693 & & \\ & & & 2.6163 & \\ & & & & 7.8404 \end{pmatrix}, \\
S^0 &= \begin{pmatrix} 5.2415 & & & & \\ & 2.7936 & & & \\ & & 2.5332 & & \\ & & & 3.9880 & \\ & & & & 1.9028 \end{pmatrix}, \\
\tau &= (1 - (1\sqrt{5}))/10 = 0.0553, \\
\tau\mu_0 e - X^0 S^0 e &= (-5.3961 \ -8.9219 \ -6.7377 \ -9.9031 \ -14.3878)^T.
\end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}
\Delta x^0 &= (0.1371 \ 1.3910 \ -0.1371 \ -1.3910 \ -3.1932)^T, \\
\Delta u^0 &= (-2.2272 \ -1.6649 \ -1.0601)^T, \\
\Delta s^0 &= (-5.4076 \ -3.7851 \ -2.2272 \ -1.6649 \ -1.0601)^T.
\end{aligned}$$

Uma vez calculada a direção a seguir, devemos agora calcular o tamanho do passo que pode ser dado nesta direção sem que se saia da vizinhança $\mathcal{N}_{-\infty}(0.8)$. Como verificamos que $(x_0 + \Delta x_0, u_0 + \Delta u_0, s_0 + \Delta s_0) \notin \mathcal{N}_{\infty}(0.8)$, calculamos o menor q inteiro e positivo, tal que $(x_0 + (0.95)^q \Delta x_0, u_0 + (0.95)^q \Delta u_0, s_0 + (0.95)^q \Delta s_0) \in \mathcal{N}_{\infty}(0.4)$. Encontramos $q = 7$ e consequentemente o novo ponto é

$$\begin{aligned}
x^1 &= x^0 + (0.95)^7 \Delta x^0 = (1.2265 \ 4.3551 \ 2.7735 \ 1.6449 \ 5.6104)^T, \\
u^1 &= u^0 + (0.95)^7 \Delta u^0 = (0.9778 \ 2.8254 \ 1.1625)^T, \\
s^1 &= s^0 + (0.95)^7 \Delta s^0 = (1.4652 \ 0.1503 \ 0.9778 \ 2.8254 \ 1.1625)^T.
\end{aligned}$$

O valor da função objetivo em x^1 é 25.4549 e o *gap* de dualidade é dado por

$$\mu^1 = (x^1{}^T s^1)/5 = 3.2666.$$

A primeira iteração do algoritmo está completa. Como o *gap* de dualidade ainda é bastante grande, seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo de trajetória central de passos longos, no qual consideramos a precisão $\epsilon = 10^{-4}$.

```

% Algoritmo de Trajetoria Central de Passos Longos

n=5;
m=3;
k=0;
epsilon=10^(-4);
alfa = 0.8;
tal = (1 - (1/sqrt(n)))/10;
A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];
%ler solucao inicial sol=[x,u,s]:
sol=input('Entre com a solucao inicial sol=[x,u,s]:');
x=sol(1:5)
s=sol(9:13);
mi=(x'*s)/n;
while (mi>= epsilon)
    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ...
        diag(s) zeros(n,m) diag(x)];
    ld=[zeros(n,1); zeros(m,1) ; (tal*mi*ones(n,1))-(diag(x)*s)];
    dir=M\ld;
    perc=1;
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    s1=sol1(9:13);
    mi1=(x1'*s1)/n;
    while (min(x1.*s1-mi1*ones(n,1))< -alfa*mi1)
        perc=0.95*perc;
        sol1=sol+perc*dir;
        x1=sol1(1:5);
        s1=sol1(9:13);
        mi1=(x1'*s1)/n;
    end
    sol = sol1;
    x=x1;
    s=s1;
    mi=mi1;
    k=k+1;
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo a cada iteração k estão representados na tabela 9.6. A trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema, está representada

k	0	1	2	3	4	5	6	7
x_1	1.1307	1.2265	1.3320	1.6593	2.0050	2.0001	2.0000	2.0000
x_2	3.3837	4.3551	5.1780	5.8038	5.9742	5.9986	5.9999	6.0000

Tabela 9.6: Iterações geradas pelo Algoritmo de Trajetória Central de Passos Longos

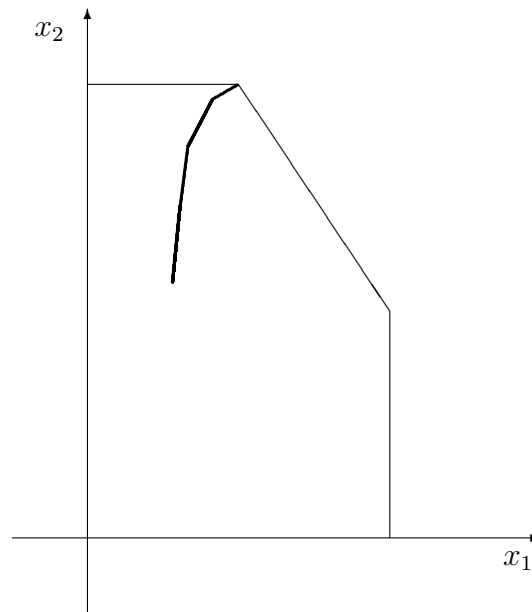


Figura 9.13: Algoritmo de Trajetória Central de Passos Longos

na figura 9.13. ■

Algoritmos de Trajetória Central Inviáveis

Os algoritmos de trajetória central apresentados até agora requerem uma solução inicial (x^0, u^0, s^0) , tal que $Ax^0 = b$, $x^0 > 0$ e $A^T u^0 - s^0 = c$, $s^0 > 0$. Ou seja, para inicializarmos os algoritmos, precisamos de uma solução viável e estritamente positiva tanto para o problema primal, quanto para o problema dual. É possível, no entanto, utilizar o método de Newton para desenvolver algoritmos que partem a cada iteração de pontos não viáveis. A única imposição feita sobre o ponto de partida destes algoritmos é a positividade de x e s .

Consideremos os resíduos primal e dual associados a um ponto (x, u, s) , definidos respectivamente por:

$$\rho_b := b - Ax \quad \text{e} \quad \rho_c := c - A^T u + s.$$

A direção de Newton, que aproxima a direção que nos leva ao ponto da trajetória central (x_μ, u_μ, s_μ) , a partir de um ponto (x, u, s) , é dada pela solução do seguinte sistema linear:

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta s \end{bmatrix} = \begin{bmatrix} \rho_c \\ \rho_b \\ \mu e - X S e \end{bmatrix} \quad (9.51)$$

onde as duas primeiras componentes do lado direito da equação (9.16) deixam de ser nulas para assumirem o valor dos resíduos associados ao ponto de partida.

Consideremos então algoritmos que partem, a cada iteração, de um ponto (x, u, s) e caminham na direção de Newton, dada pela solução de (9.51). Devemos observar que ao caminharmos nesta direção, estamos ao mesmo tempo interessados na centralização e na diminuição dos resíduos primal e dual. Facilmente verificamos que, se um passo unitário é dado na direção de Newton, o novo ponto encontrado será primal e dual viável, já que, pela equação (9.51) e pelas definições de ρ_c e ρ_b , temos:

$$A^T(u + \Delta u) - (s + \Delta s) = A^T u - s + A^T \Delta u - \Delta s = A^T u - s + \rho_c = c,$$

e,

$$A(x + \Delta x) = Ax + A\Delta x = Ax + \rho_b = b.$$

A partir de um ponto primal e dual viável, passos sucessivos na direção de Newton, nos levarão sempre a pontos também viáveis, já que, neste caso teremos $\rho_c = \rho_b = 0$ em (9.51).

Algoritmos de trajetória central inviáveis partem a cada iteração de um ponto (x, u, s) tal que $x > 0$ e $s > 0$, e caminham na direção (9.51) se mantendo sempre numa vizinhança da trajetória central. Esta vizinhança é uma extensão da vizinhança $\mathcal{N}_\infty(\alpha)$, definida para o algoritmo de trajetória central de passos longos, que contem pontos que violam as restrições $Ax = b$ e $A^T y - s = c$. Na vizinhança estendida, as normas dos resíduos primal e dual são limitadas por uma constante vezes μ . Desta forma, conforme $\mu \rightarrow 0$, os resíduos ficam cada vez menores e nos aproximamos portanto, da região viável dos problemas (P) e (D) . A definição da vizinhança estendida é dada por:

$$\mathcal{N}_{-\infty}(\alpha, \beta) := \bigcup_{\mu \in (0, \infty)} \left\{ \delta_{-\infty}(x, s, \mu) \leq \alpha, \frac{\|(\rho_b, \rho_c)\|}{\|(\rho_b^0, \rho_c^0)\|} \leq \frac{\mu}{\mu^0} \beta, (x, s) > 0 \right\}$$

onde $\alpha \in (0, 1)$, $\beta \geq 1$, $\delta_{-\infty}(x, s, \mu)$ está definida em (9.50) e ρ_b^0 e ρ_c^0 são os resíduos relativos ao ponto de partida do algoritmo, (x^0, u^0, s^0) .

Os algoritmos de trajetória central inviáveis geram uma solução para os problemas (P) e (D) , (x^k, u^k, s^k) , para a qual $\mu^k \leq \epsilon$, em não mais que $O(n^2 \log \frac{1}{\epsilon})$ iterações.

Exemplo 9.6 Resolver o problema de programação linear (9.13) utilizando o algoritmo de trajetória central inviável.

Reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, \dots, 5$ e s_j , $j = 1, \dots, 5$. É dada uma solução inicial não viável para o problema em $\mathcal{N}_{-\infty}(0.8, 1.4)$:

$$\begin{aligned} x^0 &= (1 \ 1 \ 1 \ 1 \ 1)^T, \\ u^0 &= (2 \ 2 \ 2)^T, \\ s^0 &= (2 \ 2 \ 2 \ 2 \ 2)^T. \end{aligned}$$

Neste ponto, temos $\mu^0 = (x^0{}^T s^0)/5 = 2$ e os resíduos primal e dual são dados por:

$$\begin{aligned} \rho_b &:= b - Ax^0 = \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} - \begin{pmatrix} 1 & & 1 & & \\ & 1 & & 1 & \\ 3 & 2 & & & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 12 \end{pmatrix}, \\ \rho_c &:= c - A^T u^0 + s^0 = \begin{pmatrix} 3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & & 3 \\ & 1 & 2 \\ 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Para definir a direção de busca resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} \Delta x^0 \\ \Delta u^0 \\ \Delta s^0 \end{bmatrix} = \begin{bmatrix} \rho_c \\ \rho_b \\ \tau \mu_0 e - X^0 S^0 e \end{bmatrix};$$

onde

$$\begin{aligned}
X^0 &= \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}, \\
S^0 &= \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}, \\
\tau &= 0.1, \\
\tau\mu_0 e - X^0 S^0 e &= (-1.8 \quad -1.8 \quad -1.8 \quad -1.8 \quad -1.8)^T.
\end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}
\Delta x^0 &= (1.7800 \quad 3.2700 \quad 0.2200 \quad 0.7300 \quad 0.1200)^T, \\
\Delta u^0 &= (-2.2400 \quad -3.2600 \quad -2.0400)^T, \\
\Delta s^0 &= (-5.3600 \quad -8.3400 \quad -2.2400 \quad -3.2600 \quad -2.0400)^T.
\end{aligned}$$

Uma vez calculada a direção a seguir, devemos agora calcular o tamanho do passo que pode ser dado nesta direção sem que se saia da vizinhança $\mathcal{N}_{-\infty}(0.8, 1.4)$. Iniciamos este cálculo pelo teste da razão, uma vez que no novo ponto as componentes de x e s devem se manter positivas. Como todas as componentes de Δx^0 são positivas, só realizamos o teste da razão considerando o passo dado em s .

$$\lambda = 0.995 * \min_{j=1, \dots, n} \{-s_j^0 / \Delta s_j^0 | \Delta s_j^0 < 0\} = 0.2386.$$

Como $(x_0 + \lambda\Delta x_0, u_0 + \lambda\Delta u_0, s_0 + \lambda\Delta s_0) \notin \mathcal{N}_{-\infty}(0.8, 1.4)$, calculamos o menor q inteiro e positivo, tal que $(x_0 + (0.95)^q \lambda \Delta x_0, u_0 + (0.95)^q \lambda \Delta u_0, s_0 + (0.95)^q \lambda \Delta s_0) \in \mathcal{N}_{-\infty}(0.4, 1.4)$. Encontramos $q = 2$ e consequentemente o novo ponto é

$$\begin{aligned}
x^1 &= x^0 + (0.95)^2 \lambda \Delta x^0 = (1.3833 \quad 1.7042 \quad 1.0474 \quad 1.1572 \quad 1.0258)^T, \\
u^1 &= u^0 + (0.95)^2 \lambda \Delta u^0 = (1.5176 \quad 1.2980 \quad 1.5607)^T, \\
s^1 &= s^0 + (0.95)^2 \lambda \Delta s^0 = (0.8458 \quad 0.2040 \quad 1.5176 \quad 1.2980 \quad 1.5607)^T.
\end{aligned}$$

Neste ponto temos

$$\mu^1 = (x^1{}^T s^1) / 5 = 1.2420.$$

A primeira iteração do algoritmo está completa. Como μ ainda é bastante grande, seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo de trajetória central inviável, no qual consideramos a precisão $\epsilon = 10^{-4}$.

```
%Algoritmo de Trajetoria Central Inviavel

n=5;
m=3;
k=0;
epsilon=10^(-4);
alfa = 0.8;
beta = 1.4;
tal = 0.4;
A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];
b=[4 6 18]';
c=[3 5 0 0 0]';
%ler solucao inicial sol=[x,u,s]:
sol=input('Entre com a solucao inicial sol=[x,u,s]:');
x=sol(1:5)
u=sol(6:8);
s=sol(9:13);
mi=(x'*s)/n;
mi0=mi;
rob=b-A*x;
roc=c-(A'*u)+s;
rob0=rob;
roc0=roc;
normrr0=norm([rob0;roc0],inf)/mi0;

while (mi>= epsilon)
    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ...
        diag(s) zeros(n,m) diag(x)];
    ld=[roc; rob; (tal*mi*ones(n,1))-(diag(x)*s)];
    dir=M\ld;
    %teste da razao
    aux=[10];
    for i=1:5
        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end
    for i=9:13
```

```

        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end
end

perc = min(1,0.995*min(aux));
sol1=sol+perc*dir;
x1=sol1(1:5);
u1=sol1(6:8);
s1=sol1(9:13);
mi1=(x1'*s1)/n;
rob1=b-A*x1;
roc1=c-(A'*u1)+s1;

while((min(x1.*s1)<(1-alfa)*mi1)|(min([rob1;roc1]) ..
    <-beta*normrr0*mi1)|(mi1>(1-0.01*perc)*mi))

    perc=0.95*perc;
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    u1=sol1(6:8);
    s1=sol1(9:13);
    mi1=(x1'*s1)/n;
    rob1=b-A*x1;
    roc1=c-(A'*u1)+s1;
end
sol = sol1;
x=x1
u=u1;
s=s1;
mi=mi1;
rob=rob1;
roc=roc1;
k=k+1;
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo estão representados na tabela 9.7. A figura 9.14 mostra a trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema.

Observamos que apesar do ponto (x_1, x_2) pertencer à região viável primal, desde a primeira iteração do algoritmo, como pode ser visto na figura 9.7;

k	0	1	2	3	4
x_1	1.0000	1.3833	1.4617	1.4795	1.5599
x_2	1.0000	1.7042	2.5206	3.3535	4.9700
k	5	6	7	8	
x_1	1.8569	1.9954	1.9999	2.0000	
x_2	5.7649	5.9976	5.9997	6.0000	

Tabela 9.7: Iterações geradas pelo Algoritmo de Trajetória Central Inviável

o ponto inicial utilizado, $(x^0, u^0, s^0) = (1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2)$, não é primal nem dual viável. ■

9.6 Algoritmos de Redução Potencial

Na seção anterior apresentamos os algoritmos de trajetória central. Verificamos que estes algoritmos geram pontos sempre numa vizinhança da trajetória central, como forma de evitar a fronteira do ortante não negativo. Nesta seção estudaremos os algoritmos de redução potencial. Estes algoritmos também objetivam aumentar o valor da função objetivo de (P) a cada iteração, sem se aproximar da fronteira da região viável do problema. Seguindo este intuito, o progresso do algoritmo em direção ao ótimo é medido pela seguinte função potencial:

$$f(x, s) := \rho \log x^T s - \sum_{i=1}^n \log x_i s_i \quad (9.52)$$

onde ρ é uma constante maior que n .

Lembrando de (9.24), que $s^T x = b^T u - c^T x$, verificamos que o primeiro termo da função potencial mede o *gap* de dualidade. E considerando a equivalência

$$\sum_{i=1}^n \log x_i s_i = \sum_{i=1}^n \log x_i + \sum_{i=1}^n \log s_i,$$

verificamos que o segundo termo da função penaliza a proximidade à fronteira da região viável tanto do problema primal quanto do dual.

As funções potenciais sempre foram consideradas nos métodos de pontos interiores. O algoritmo de Karmarkar já empregava a redução de uma função potencial que constitui uma pequena variação de (9.52). Como sugerido pelo nome, os algoritmos de redução potencial visam a cada iteração, reduzir o

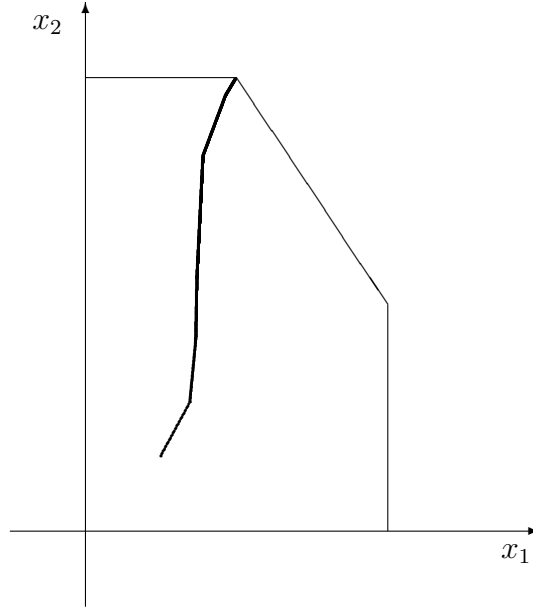


Figura 9.14: Algoritmo de Trajetória Central Inviável

valor da função potencial de um determinado valor constante. A motivação para tal procedimento está no próximo teorema, que nos fornece um resultado global quanto a complexidade desses algoritmos.

Teorema 9.4 *Considere um ponto inicial primal viável $x^0 \in \mathcal{X}^0$, um ponto inicial dual viável $(u^0, s^0) \in \mathcal{S}^0$ e uma tolerância $\epsilon > 0$. Suponha que um algoritmo gera uma seqüência de pontos $x^k \in \mathcal{X}^0$ e $(u^k, s^k) \in \mathcal{S}^0$ e reduz a cada iteração o valor da função potencial (9.52) de uma grandeza maior ou igual a $\delta > 0$. Então este algoritmo gera uma solução com gap de dualidade*

$$x^{kT} s^k \leq \epsilon$$

para todo $k \geq K$, onde

$$K = \frac{1}{\delta} \left((\rho - n) \log \frac{1}{\epsilon} + f(x^0, s^0) \right) \quad (9.53)$$

Demonstração

Da definição da função potencial (9.52), temos

$$\begin{aligned} f(x, s) &= \rho \log x^T s - \sum_{i=1}^n \log x_i s_i \\ &= (\rho - n) \log x^T s - \sum_{i=1}^n \log \frac{x_i s_i}{x^T s} \\ &\geq (\rho - n) \log x^T s. \end{aligned}$$

Logo,

$$(\rho - n) \log x^{kT} s^k \leq f(x^k, s^k) \leq f(x^0, s^0) - k\delta,$$

ou,

$$\log x^{kT} s^k \leq \frac{f(x^0, s^0) - k\delta}{\rho - n}.$$

Para que o critério $x^{kT} s^k \leq \epsilon$ seja alcançado, basta então que k satisfaça a relação

$$\frac{f(x^0, s^0) - k\delta}{\rho - n} \leq \log \epsilon,$$

isto é,

$$-k\delta \leq (\rho - n) \log \epsilon - f(x^0, s^0)$$

ou,

$$k \geq \frac{1}{\delta} \left((\rho - n) \log \frac{1}{\epsilon} + f(x^0, s^0) \right). \quad \blacksquare$$

O algoritmo de redução potencial primal-dual de Kojima, Mizuno e Yoshise [KoMiYo 91], apresentado e analisado a seguir, utiliza a direção de Newton (9.16) como direção de busca. A sua análise de complexidade consiste basicamente em verificar que determinada escolha do parâmetro τ garante a redução da função potencial a cada iteração, de um valor constante.

Dada uma solução inicial $(x^0, u^0, s^0) \in \mathcal{F}^0$ e $\epsilon > 0$, este algoritmo requer $O(\sqrt{n} \log \frac{1}{\epsilon})$ iterações para reduzir o *gap* de dualidade $x^T s$ a um valor menor que ϵ .

Algoritmo 9.3 (Redução Potencial) *Dados:* $\rho > n$, $\tau = \frac{n}{\rho}$, $(x^0, u^0, s^0) \in \mathcal{F}^0$;

$k := 0$;

Repita enquanto $x^{kT} s^k > \epsilon$

Faça $\tau_k = \tau$, $\mu_k = x^{kT} s^k / n$;

Calcule a direção de Newton (9.25), ou seja, resolva o sistema

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta u^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau_k \mu_k e - X^k S^k e \end{bmatrix};$$

Faça

$$\begin{aligned} (x^{k+1}, u^{k+1}, s^{k+1}) &:= (x^k, u^k, s^k) + \alpha_k(\Delta x^k, \Delta u^k, \Delta s^k); \\ k &:= k + 1; \end{aligned} \quad (9.54)$$

Fim (Repita)

No passo (9.54), α_k é obtido por uma busca linear, ou seja,

$$\alpha_k = \operatorname{argmin}_{\alpha \in [0,1]} \{f(x^k + \alpha \Delta x^k, s^k + \alpha \Delta s^k) | x^k + \alpha \Delta x^k > 0, s^k + \alpha \Delta s^k > 0\}.$$

No teorema 9.5, verificaremos a redução da função potencial $f(x, s)$, a cada iteração do algoritmo (9.3). Os resultados dos lemas colocados a seguir serão utilizados na demonstração do teorema.

Lema 9.7 *Seja $u \in \mathbb{R}^n$, tal que $\|u\|_\infty < 1$. Então*

$$\sum_{i=1}^n \log(1 + u_i) \geq e^T u - \frac{\|u\|_2}{2(1 - \|u\|_\infty)}.$$

Demonstração

Desenvolvendo-se a função $\log(1 + u_i)$ em série de Taylor, obtem-se

$$\begin{aligned} \log(1 + u_i) &= \sum_{j=1}^{\infty} \frac{(-1)^{j+1} u_i^j}{j} \\ &\geq u_i - \frac{1}{2} \sum_{j=2}^{\infty} |u_i|^j \\ &= u_i - \frac{u_i^2}{2(1 - |u_i|)}. \end{aligned}$$

Considerando agora que $|u_i| \leq \|u\|_\infty$ para todo i , temos

$$\sum_{i=1}^n \log(1 + u_i) \geq \sum_{i=1}^n u_i - \frac{1}{2(1 - \|u\|_\infty)} \sum_{i=1}^n u_i^2. \quad \blacksquare$$

Lema 9.8 *Seja*

$$\tilde{x} = x + \alpha \Delta x \quad e \quad \tilde{s} = s + \alpha \Delta s.$$

Defina

$$V = X^{\frac{1}{2}} S^{\frac{1}{2}} \quad e \quad D = X^{\frac{1}{2}} S^{-\frac{1}{2}},$$

e considere a constante β , tal que

$$\beta \geq \alpha \max \left(\|X^{-1} \Delta x\|_\infty, \|S^{-1} \Delta s\|_\infty \right). \quad (9.55)$$

Então,

$$\begin{aligned} f(\tilde{x}, \tilde{s}) - f(x, s) &\leq -\alpha \frac{\rho}{n\mu} \|D^{-1} \Delta x + D \Delta s\|^2 \\ &\quad + \frac{\alpha^2}{2(1-\beta)} \frac{1}{v_{\min}^2} \|D^{-1} \Delta x + D \Delta s\|^2. \end{aligned}$$

Demonstração

Utilizando a definição da função potencial (9.52), temos

$$\begin{aligned} f(\tilde{x}, \tilde{s}) - f(x, s) &= \rho(\log \tilde{x}^T \tilde{s} - \log x^T s) \\ &= \sum_{i=1}^n \log \tilde{x}_i + \sum_{i=1}^n \log x_i \\ &\quad - \sum_{i=1}^n \log \tilde{s}_i - \sum_{i=1}^n \log s_i. \end{aligned}$$

Logo,

$$f(\tilde{x}, \tilde{s}) - f(x, s) = \rho \log \frac{\tilde{x}^T \tilde{s}}{x^T s} - \sum_{i=1}^n \log \frac{\tilde{x}_i}{x_i} - \sum_{i=1}^n \log \frac{\tilde{s}_i}{s_i}.$$

Consideremos agora as seguintes relações:

$$\begin{aligned} \frac{\tilde{x}^T \tilde{s}}{x^T s} &= \frac{(x + \alpha \Delta x)^T (s + \alpha \Delta s)}{x^T s} \\ &= \frac{1}{x^T s} \left(x^T s + \alpha (x^T \Delta s + s^T \Delta x) + \alpha^2 \Delta x^T \Delta s \right) \\ &= 1 + \alpha \frac{(x^T \Delta s + s^T \Delta x)}{x^T s}, \end{aligned}$$

onde a última igualdade utiliza o resultado do lema 9.1,

$$\frac{\tilde{x}_i}{x_i} = \frac{x_i + \alpha \Delta x_i}{x_i} = 1 + \alpha \frac{\Delta x_i}{x_i},$$

e

$$\frac{\tilde{s}_i}{s_i} = \frac{s_i + \alpha \Delta s_i}{s_i} = 1 + \alpha \frac{\Delta s_i}{s_i}.$$

Temos então

$$\begin{aligned} f(\tilde{x}, \tilde{s}) - f(x, s) &= \rho \log \left(1 + \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} \right) \\ &\quad - \sum_{i=1}^n \log \left(1 + \alpha \frac{\Delta x_i}{x_i} \right) \\ &\quad - \sum_{i=1}^n \log \left(1 + \alpha \frac{\Delta s_i}{s_i} \right). \end{aligned}$$

Considerando que x , \tilde{x} , s e \tilde{s} são vetores estritamente positivos, temos que

$$\frac{\tilde{x}^T \tilde{s}}{x^T s} > 0 \Rightarrow 1 + \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} > 0 \Rightarrow -\alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} < 1$$

Aplicando então a conhecida relação $\log(1 - t) \leq -t$, para $t < 1$, temos

$$\log \left(1 + \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} \right) \leq \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s}.$$

Utilizando agora o resultado do lema 9.7, temos que, se α é pequeno o suficiente para satisfazer as relações

$$\alpha \|X^{-1} \Delta x\|_{\infty} \leq 1 \quad \text{e} \quad \alpha \|S^{-1} \Delta s\|_{\infty} \leq 1, \quad (9.56)$$

as seguintes desigualdades são verificadas:

$$-\sum_{i=1}^n \log \left(1 + \alpha \frac{\Delta x_i}{x_i} \right) \leq -\alpha e^T X^{-1} \Delta x + \frac{\alpha^2 \|X^{-1} \Delta x\|^2}{2(1 - \alpha \|X^{-1} \Delta x\|_\infty)},$$

e

$$-\sum_{i=1}^n \log \left(1 + \alpha \frac{\Delta s_i}{s_i} \right) \leq -\alpha e^T S^{-1} \Delta s + \frac{\alpha^2 \|S^{-1} \Delta s\|^2}{2(1 - \alpha \|S^{-1} \Delta s\|_\infty)}.$$

Conseqüentemente,

$$\begin{aligned} f(\tilde{x}, \tilde{s}) - f(x, s) &\leq \rho \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} \\ &\quad - \alpha e^T X^{-1} \Delta x + \frac{\alpha^2 \|X^{-1} \Delta x\|^2}{2(1 - \alpha \|X^{-1} \Delta x\|_\infty)} \\ &\quad - \alpha e^T S^{-1} \Delta s + \frac{\alpha^2 \|S^{-1} \Delta s\|^2}{2(1 - \alpha \|S^{-1} \Delta s\|_\infty)}. \end{aligned}$$

Seja a constante β , tal que

$$\beta \geq \alpha \max \left(\|X^{-1} \Delta x\|_\infty, \|S^{-1} \Delta s\|_\infty \right),$$

temos então

$$\begin{aligned} f(\tilde{x}, \tilde{s}) - f(x, s) &\leq \rho \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} \\ &\quad - \alpha e^T (X^{-1} \Delta x + S^{-1} \Delta s) \\ &\quad + \frac{\alpha^2}{2(1-\beta)} (\|X^{-1} \Delta x\|^2 + \|S^{-1} \Delta s\|^2). \end{aligned} \tag{9.57}$$

Utilizando as definições de V e D , temos

$$\begin{aligned} &\rho \frac{x^T \Delta s + s^T \Delta x}{x^T s} - e^T (X^{-1} \Delta x + S^{-1} \Delta s) \\ = &\frac{\rho}{x^T s} e^T (X \Delta s + S \Delta x) - e^T (X^{-1} \Delta x + S^{-1} \Delta s) \\ = &\frac{\rho}{x^T s} e^T (V D \Delta s + V D^{-1} \Delta x) - e^T (V^{-1} D^{-1} \Delta x + V^{-1} D \Delta s) \\ = &-\frac{\rho}{n\mu} \left[e^T \left(\frac{n\mu}{\rho} (V^{-1} D^{-1} \Delta x + V^{-1} D \Delta s) - V D^{-1} \Delta x - V D \Delta s \right) \right] \\ = &-\frac{\rho}{n\mu} \left[e^T \left(\frac{n\mu}{\rho} V^{-1} (D^{-1} \Delta x + D \Delta s) - V (D^{-1} \Delta x + D \Delta s) \right) \right] \\ = &-\frac{\rho}{n\mu} \left[e^T \left(\frac{n\mu}{\rho} V^{-1} - V \right) (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \left[\left(\frac{n\mu}{\rho} V^{-1} e - V e \right)^T (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \left[\left(V^{-1} \left(\frac{n\mu}{\rho} e - V^2 e \right) \right)^T (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \left[\left(V^{-1} \left(\frac{n\mu}{\rho} e - X S e \right) \right)^T (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \left[(V^{-1} (S \Delta x + X \Delta s))^T (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \left[(D^{-1} \Delta x + D \Delta s)^T (D^{-1} \Delta x + D \Delta s) \right] \\ = &-\frac{\rho}{n\mu} \|D^{-1} \Delta x + D \Delta s\|^2, \end{aligned} \tag{9.58}$$

onde a nona igualdade é consequência da terceira equação do sistema de Newton (9.25). Temos também

$$\begin{aligned}\|X^{-1}\Delta x\|^2 + \|S^{-1}\Delta s\|^2 &= \|V^{-1}D^{-1}\Delta x\|^2 + \|V^{-1}D\Delta s\|^2 \\ &\leq \|V^{-1}\|^2 (\|D^{-1}\Delta x\|^2 + \|D\Delta s\|^2) \\ &\leq \frac{1}{v_{\min}^2} (\|D^{-1}\Delta x\|^2 + \|D\Delta s\|^2).\end{aligned}\quad (9.59)$$

onde a última desigualdade decorre da relação $\|A\| \leq \sqrt{\|A\|_1\|A\|_\infty}$, válida para toda matriz A , e demonstrada em ([GoVo 96, corolário 2.3.2, pp.58]). Finalmente, utilizando novamente o resultado do lema 9.1, temos que

$$\begin{aligned}\|D^{-1}\Delta x + D\Delta s\|^2 &= (D^{-1}\Delta x + D\Delta s)^T(D^{-1}\Delta x + D\Delta s) \\ &= \|D^{-1}\Delta x\|^2 + 2\Delta x^T\Delta s + \|D\Delta s\|^2 \\ &= \|D^{-1}\Delta x\|^2 + \|D\Delta s\|^2.\end{aligned}\quad (9.60)$$

Portanto, utilizando as relações (9.57), (9.58), (9.59), (9.60), temos

$$\begin{aligned}f(\tilde{x}, \tilde{s}) - f(x, s) &\leq -\alpha \frac{\rho}{n\mu} \|D^{-1}\Delta x + D\Delta s\|^2 \\ &\quad + \frac{\alpha^2}{2(1-\beta)} \frac{1}{v_{\min}^2} \|D^{-1}\Delta x + D\Delta s\|^2\end{aligned}\quad \blacksquare$$

Lema 9.9 *Seja $x, s > 0$ e $\rho \geq n + \sqrt{n}$. Então*

$$\|D^{-1}\Delta x + D\Delta s\| \geq \frac{n\mu}{\rho} \frac{1}{v_{\min}} \frac{\sqrt{3}}{2}.$$

Demonstração

$$\begin{aligned}\frac{\rho^2}{n^2\mu^2} \|D^{-1}\Delta x + D\Delta s\|^2 &= \|V^{-1}e - \frac{\rho}{n\mu}v\|^2 \\ &= \|V^{-1}e\|^2 - 2\frac{\rho}{n\mu}e^TV^{-1}v + \frac{\rho^2}{n^2\mu^2}\|v\|^2.\end{aligned}$$

Considerando que

$$V = (XS)^{\frac{1}{2}} \Rightarrow \|v\|^2 = \sum_{i=1}^n x_i s_i = x^T s = n\mu$$

e

$$e^TV^{-1}v = e^Te = n,$$

temos

$$\begin{aligned}
\frac{\rho^2}{n^2\mu^2} \|D^{-1}\Delta x + D\Delta s\|^2 &= \|V^{-1}e\|^2 - 2\frac{\rho}{\mu} + \frac{\rho^2}{n\mu} \\
&= \|V^{-1}e\|^2 + \frac{\rho^2 - 2n\rho}{n\mu} \\
&= \|V^{-1}e\|^2 + \frac{(\rho-n)^2 - n^2}{n\mu} \\
&\geq \|V^{-1}e\|^2 + \frac{(n+\sqrt{n-n})^2 - n^2}{n\mu} \\
&= \|V^{-1}e\|^2 + \frac{n-n^2}{n\mu} \\
&= \|V^{-1}e\|^2 - \frac{n^2}{n\mu} + \frac{1}{\mu} \\
&= \|V^{-1}e\|^2 - \frac{2n^2}{n\mu} + \frac{n^2}{n\mu} + \frac{1}{\mu} \\
&= \|V^{-1}e\|^2 - \frac{2n^2}{n\mu} + \frac{n\mu}{\mu^2} + \frac{1}{\mu} \\
&= \|V^{-1}e\|^2 - \frac{2n}{\mu} + \frac{1}{\mu^2} \|v\|^2 + \frac{1}{\mu} \\
&= \|V^{-1}e\|^2 - \frac{2}{\mu} e^T V^{-1} v + \frac{1}{\mu^2} \|v\|^2 + \frac{1}{\mu} \\
&= \|V^{-1}e - \frac{1}{\mu} v\|^2 + \frac{1}{\mu} \\
&\geq \left(\frac{1}{v_{\min}} - \frac{1}{\mu} v_{\min} \right)^2 + \frac{1}{\mu} \\
&= \frac{1}{\mu^2} \left\{ \left(\frac{\mu}{v_{\min}} - v_{\min} \right)^2 + \mu \right\} \\
&= \frac{1}{\mu^2} \left\{ \left(\frac{\mu}{2v_{\min}} - v_{\min} \right)^2 + \frac{3}{4} \frac{\mu^2}{v_{\min}^2} \right\} \\
&\geq \frac{3}{4} \frac{1}{v_{\min}^2}
\end{aligned}$$

onde a segunda desigualdade mais uma vez decorre da relaça~o

$$\|A\| \leq \sqrt{\|A\|_1 \|A\|_\infty}. \quad \blacksquare$$

Teorema 9.5 *Seja $\rho \geq n + \sqrt{n}$. Ent~ao existe $\tilde{\alpha}$, tal que*

$$f(x + \tilde{\alpha}\Delta x, s + \tilde{\alpha}\Delta s) - f(x, s) \leq -0.18.$$

Demonstra~ao

Seja

$$\tilde{\alpha} = \frac{v_{\min}}{2\|D^{-1}\Delta x + D\Delta s\|} \quad \text{e} \quad \tilde{\beta} = \frac{1}{2}.$$

Utilizando (9.59) e (9.60), temos

$$\|X^{-1}\Delta x\|^2 + \|S^{-1}\Delta s\|^2 \leq \frac{1}{v_{\min}^2} \|D^{-1}\Delta x + D\Delta s\|^2.$$

Portanto,

$$\begin{aligned}
\tilde{\alpha} \|X^{-1}\Delta x\|_\infty \leq \tilde{\alpha} \|X^{-1}\Delta x\| &\leq \tilde{\alpha} \frac{1}{v_{\min}} \|D^{-1}\Delta x + D\Delta s\| \\
&= \frac{\frac{v_{\min}}{2\|D^{-1}\Delta x + D\Delta s\|}}{v_{\min}} \|D^{-1}\Delta x + D\Delta s\| \\
&= \frac{1}{2} \\
&= \tilde{\beta}
\end{aligned}$$

e

$$\begin{aligned}
\tilde{\alpha} \|S^{-1} \Delta s\|_{\infty} \leq \tilde{\alpha} \|S^{-1} \Delta s\| &\leq \tilde{\alpha} \frac{1}{v_{\min}} \|D^{-1} \Delta x + D \Delta s\| \\
&= \frac{\tilde{\alpha} \frac{1}{v_{\min}}}{2 \|D^{-1} \Delta x + D \Delta s\|} \frac{1}{v_{\min}} \|D^{-1} \Delta x + D \Delta s\| \\
&= \frac{1}{2} \\
&= \frac{1}{\beta}.
\end{aligned}$$

Conseqüentemente, $\tilde{\alpha}$ e $\tilde{\beta}$ satisfazem as especificações (9.56) e (9.55) respectivamente, e podemos então, utilizar os resultados dos lemas e teoremas anteriores.

Do lema 9.8, temos

$$\begin{aligned}
f(x + \tilde{\alpha} \Delta x, s + \tilde{\alpha} \Delta s) &- f(x, s) \\
&\leq -\tilde{\alpha} \frac{\rho}{n\mu} \|D^{-1} \Delta x + D \Delta s\|^2 \\
&+ \frac{\tilde{\alpha}^2}{2(1-\tilde{\beta})} \frac{1}{v_{\min}^2} \|D^{-1} \Delta x + D \Delta s\|^2 \\
&= -\frac{\tilde{\alpha} \frac{\rho}{n\mu}}{2 \|D^{-1} \Delta x + D \Delta s\|} \|D^{-1} \Delta x + D \Delta s\|^2 \\
&+ \frac{\tilde{\alpha}^2}{4 \|D^{-1} \Delta x + D \Delta s\|^2} \frac{1}{2(1-\tilde{\beta})} \frac{1}{v_{\min}^2} \|D^{-1} \Delta x + D \Delta s\|^2 \\
&= -\frac{v_{\min}}{2} \frac{\rho}{n\mu} \|D^{-1} \Delta x + D \Delta s\| + \frac{1}{8(1-\tilde{\beta})}.
\end{aligned}$$

Utilizando agora o resultado do lema 9.9, temos

$$\begin{aligned}
f(x + \tilde{\alpha} \Delta x, s + \tilde{\alpha} \Delta s) - f(x, s) &\leq -\frac{v_{\min}}{2} \frac{\rho}{n\mu} \frac{\sqrt{3}}{2v_{\min}} \frac{n\mu}{\rho} + \frac{1}{8(1-\tilde{\beta})} \\
&= -\frac{\sqrt{3}}{4} + \frac{1}{8(1-\tilde{\beta})}.
\end{aligned}$$

Finalmente, considerando $\tilde{\beta} = \frac{1}{2}$, temos

$$f(x + \tilde{\alpha} \Delta x, s + \tilde{\alpha} \Delta s) - f(x, s) \leq -0.18. \quad \blacksquare$$

Considerando agora $\rho = n + \sqrt{n}$ e $\delta = 0.18$ em (9.53), chegamos a ordem do número de iterações realizadas pelo algoritmo 9.3, $O(\sqrt{n} \log \frac{1}{\epsilon})$.

Exemplo 9.7 Resolver o problema de programação linear (9.13) utilizando o algoritmo de redução potencial.

É dada uma solução inicial interior viável para o problema: $x_1 = 1$, $x_2 = 1$, $u_1 = 2$, $u_2 = 2$, $u_3 = 2$. Em seguida, reescrevemos os problemas primal e dual na forma padrão, com a adição de variáveis de folga x_i , $i = 3, \dots, 5$ e s_j , $j = 1, \dots, 5$. A solução inicial interior viável correspondente é

$$\begin{aligned}x^0 &= (1 \ 1 \ 3 \ 5 \ 13)^T, \\u^0 &= (2 \ 2 \ 2)^T, \\s^0 &= (5 \ 1 \ 2 \ 2)^T.\end{aligned}$$

O *gap* de dualidade em (x^0, s^0) é $x^{0T}s^0 = 48$ e

$$\mu_0 = x^{0T}s^0/n = 9.6.$$

O valor da função potencial neste ponto inicial é dado por:

$$f(x^0, s^0) = \rho \log x^{0T}s^0 - \sum_{i=1}^n \log x_i^0 s_i^0 = 281.3782,$$

onde $\rho = 2n = 75$. Para definir a direção de busca resolvemos o sistema linear

$$\begin{bmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ S^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} \Delta x^0 \\ \Delta u^0 \\ \Delta s^0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau\mu_0 e - X^0 S^0 e \end{bmatrix};$$

onde

$$\begin{aligned}A &= \begin{pmatrix} 1 & & 1 & & \\ & 1 & & 1 & \\ 3 & 2 & & & 1 \end{pmatrix}, \\X^0 &= \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 3 & & \\ & & & 5 & \\ & & & & 13 \end{pmatrix},\end{aligned}$$

$$S^0 = \begin{pmatrix} 5 & & & & \\ & 1 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix},$$

$$\begin{aligned}\tau &= n/\rho = 0.0667, \\ \tau\mu_0 e - X^0 S^0 e &= (-4.3600 \ -0.3600 \ -5.3600 \ -9.3600 \ -25.3600)^T.\end{aligned}$$

A solução do sistema linear é dada por

$$\begin{aligned}
\Delta x^0 &= (0.1206 \quad 2.6309 \quad -0.1206 \quad -2.6309 \quad -5.6236)^T, \\
\Delta u^0 &= (-1.7063 \quad -0.8197 \quad -1.0856)^T, \\
\Delta s^0 &= (-4.9631 \quad -2.9909 \quad -1.7063 \quad -0.8197 \quad -1.0856)^T.
\end{aligned}$$

Uma vez calculada a direção a seguir, devemos agora calcular o tamanho do passo que pode ser dado nesta direção. Iniciamos este cálculo pelo teste da razão, já que as componentes de x e s devem se manter não negativas:

$$\begin{aligned}
\lambda_1 &= \min_{j=1,\dots,n} \{-x_j^0 / \Delta x_j^0 | \Delta x_j^0 < 0\} = 1.9005, \\
\lambda_2 &= \min_{j=1,\dots,n} \{-s_j^0 / \Delta s_j^0 | \Delta s_j^0 < 0\} = 0.3344, \\
\lambda &= \min\{1, 0.995 * \min\{\lambda_1, \lambda_2\}\} = 0.3327
\end{aligned}$$

Calculamos o valor da função potencial no ponto $(x^0 + \lambda \Delta x^0, s^0 + \lambda \Delta s^0)$:

$$\begin{aligned}
f(x^0 + \lambda \Delta x^0, s^0 + \lambda \Delta s^0) &= \rho \log(x^0 + \lambda \Delta x^0)^T (s^0 + \lambda \Delta s^0) \\
&\quad - \sum_{i=1}^n \log(x_i^0 + \lambda \Delta x_i^0) * (s_i^0 + \lambda \Delta s_i^0) \\
&= 259.5658.
\end{aligned}$$

Como

$$f(x^0 + \lambda \Delta x^0, s^0 + \lambda \Delta s^0) < f(x^0, s^0), \quad (9.61)$$

tomamos o novo ponto como:

$$\begin{aligned}
x^1 = x^0 + \lambda \Delta x^0 &= (1.0401 \quad 1.8752 \quad 2.9599 \quad 4.1248 \quad 11.1292)^T, \\
u^1 = u^0 + \lambda \Delta u^0 &= (1.4324 \quad 1.7273 \quad 1.6388)^T, \\
s^1 = s^0 + \lambda \Delta s^0 &= (3.3489 \quad 0.0050 \quad 1.4324 \quad 1.7273 \quad 1.6388)^T.
\end{aligned}$$

O *gap* de dualidade em no novo ponto é $x^{1T} s^1 = 33.0959$ e

$$\mu_1 = x^{1T} s^1 / n = 6.6192.$$

A primeira iteração do algoritmo está completa. Como o *gap* de dualidade ainda é bastante grande, seguimos com uma nova iteração. Deixamos a próxima iteração do algoritmo como exercício para o leitor.

Lembramos que, caso a relação (9.61) não fosse satisfeita, deveríamos calcular o menor q inteiro e positivo, tal que

$$f(x^0 + (0.95)^q \lambda \Delta x^0, s^0 + (0.95)^q \lambda \Delta s^0) < f(x^0, s^0).$$

O novo ponto, neste caso, seria

$$(x^1, u^1, s^1) := (x^0 + (0.95)^q \lambda \Delta x^0, u^0 + (0.95)^q \lambda \Delta u^0, s^0 + (0.95)^q \lambda \Delta s^0).$$

Apresentamos abaixo o programa em MATLAB que implementa o algoritmo de redução potencial, no qual consideramos a precisão $\epsilon = 10^{-4}$.

```

% Algoritmo de Reducao Potencial

n=5;
m=3;
k=0;
epsilon=10^(-8);
ro=2*n;
tal=n/ro;

A=[1 0 1 0 0;0 1 0 1 0;3 2 0 0 1];
%ler solucao inicial sol=[x,u,s]:
sol=input('Entre com a solucao inicial sol=[x,u,s]:');
x=sol(1:5)
s=sol(9:13);
gap=x'*s;
f = ro*log(gap) - sum(log(x.*s));

while (gap>= epsilon)
    M=[zeros(n,n) A' -eye(n,n); A zeros(m,m) zeros(m,n); ..
        diag(s) zeros(n,m) diag(x)];
    ld=[zeros(n,1); zeros(m,1) ; (tal*(gap/n)*ones(n,1)) ..
        -(diag(x)*s)];
    dir=M\ld;
    %teste da razao
    aux=[10];
    for i=1:5
        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end
    for i=9:13
        if(dir(i)<0)
            aux = [aux; -sol(i)/dir(i)];
        end
    end
    perc = min(1,0.995*min(aux));
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    s1=sol1(9:13);
    gap1=x1'*s1;
    f1 = ro*log(gap1) - sum(log(x1.*s1));

```

k	0	1	2	3	4	5	6	7
x_1	1.0000	1.0401	0.8092	1.1349	2.0069	1.9997	2.0000	2.0000
x_2	1.0000	1.8752	5.6794	5.9668	5.9830	5.9983	5.9999	6.0000

Tabela 9.8: Iterações geradas pelo Algoritmo de Redução Potencial

```

while (f1>f)
    perc=0.95*perc;
    sol1=sol+perc*dir;
    x1=sol1(1:5);
    s1=sol1(9:13);
    gap1=x1'*s1;
    f1 = ro*log(gap1) - sum(log(x1.*s1));
end
sol = sol1;
x=x1;
s=s1;
gap=gap1;
f=f1;
k=k+1;
end

```

Os valores de x_1 e x_2 obtidos pelo algoritmo a cada iteração k estão representados na tabela 9.8. A trajetória definida pelos pontos gerados a cada iteração do algoritmo, na região viável do problema, está representada na figura 9.15. ■

9.7 Exercícios

1. Considere a transformação linear $A \in \mathbb{R}^{m \times n}$. Demonstre que um vetor $v \in \mathbb{R}^n$ é ortogonal a $\mathcal{I}(A^T)$ se e só se $v \in \mathcal{N}(A)$.
2. Fazer programas em MATLAB para desenvolver os passos da segunda iteração de cada um dos algoritmos abaixo na resolução do problema (9.13).
 - (a) algoritmo Afim-Escala,
 - (b) algoritmo de trajetória central de passos curtos,
 - (c) algoritmo preditor-corretor,

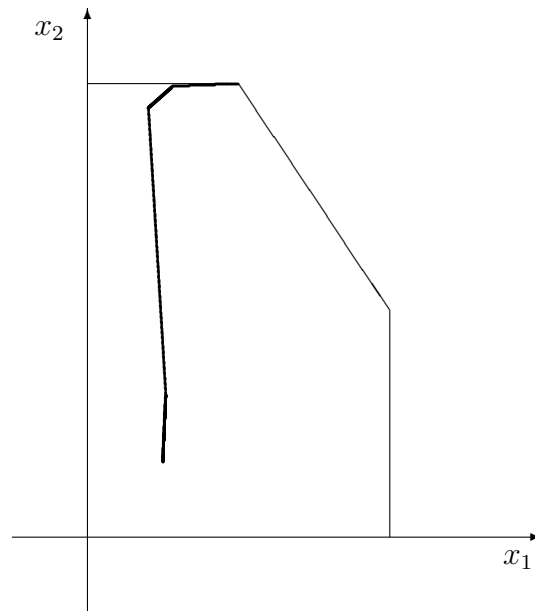


Figura 9.15: Algoritmo de Redução Potencial

- (d) algoritmo de trajetória central de passos longos,
- (e) algoritmo de trajetória central inviável,
- (f) algoritmo de redução potencial.

Capítulo 10

Introdução à Programação Inteira

Quando nos problemas de programação linear obrigarmos algumas ou todas as variáveis de decisão a só admitirem valores inteiros, estaremos diante de um problema de *programação linear inteira*. Como exemplo, seja

$$(P) : \text{maximizar } z = x_1 - 3x_2 - 4x_3$$

sujeito a:

$$\begin{array}{rcccccl} 2x_1 & + & x_2 & - & x_3 & \leq & 4 \\ 4x_1 & - & 3x_2 & & & \leq & 2 \\ 3x_1 & + & 2x_2 & + & x_3 & \leq & 3 \end{array}$$

$$\begin{array}{rcl} x_1 & \geq & 0 \\ & x_2 & \geq 0 \\ & & x_3 \geq 0 \end{array}$$

x_2 e x_3 inteiros.

O problema (P) restringe x_2 e x_3 a valores inteiros não negativos, enquanto que x_1 é um real qualquer não negativo. (P) pode também ser denominado de um *problema de programação linear mista*, pois nem todas as variáveis são restritas a valores inteiros.

Poderíamos imaginar a solução de (P) ignorando as restrições de integridade, para isto o problema seria considerado como sendo de programação linear, visto nos capítulos anteriores. Caso a solução obtida fornecesse valores inteiros para todas as variáveis de decisão, teríamos também resolvido o problema original (P) . No entanto, se algumas variáveis tomam valores fracionários na solução do problema linear, quando deveriam ser inteiras, a primeira idéia é tentar arredondar esses valores aos inteiros mais próximos

de maneira que as soluções inteiras obtidas sejam viáveis. Infelizmente este procedimento poderá fornecer soluções inteiras distantes do ótimo. A título de ilustração consideraremos o seguinte exemplo.

$$(P) : \text{maximizar } z = x_1 + 19x_2$$

sujeito a:

$$x_1 + 20x_2 \leq 50$$

$$x_1 + x_2 \leq 20$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

x_1 e x_2 inteiros.

Deixando de lado as restrições de integralidade o problema (P) se tornará:

$$(\bar{P}) : \text{maximizar } z = x_1 + 19x_2$$

sujeito a:

$$x_1 + 20x_2 \leq 50$$

$$x_1 + x_2 \leq 20$$

$$x_1 \geq 0$$

$$x_2 \geq 0.$$

(\bar{P}) é um problema de programação linear que pode ser solucionado utilizando o método do simplex. Para isto devemos acrescentar as variáveis de folga $x_3 \geq 0$ e $x_4 \geq 0$ da seguinte maneira.

$$(\bar{P}) : \text{maximizar } z = x_1 + 19x_2$$

sujeito a:

$$x_1 + 20x_2 + x_3 = 50$$

$$x_1 + x_2 + x_4 = 20$$

$$x_j \geq 0, \quad j = 1, 2, 3, 4.$$

Resolveremos (\bar{P}) na próxima seção e, ao mesmo tempo, apresentaremos o método do simplex por operações entre colunas.

10.1 Esquematizando o Método do Simplex por Operações entre Colunas

Para resolver o problema (\bar{P}) da seção anterior apresentaremos, para facilitar a visão das iterações do método do simplex, uma disposição dos dados em quadros proposta por Gomory [Go 58a]. Sem perda de generalidade (\bar{P}) poderá ser escrito sob a seguinte forma.

$$(\bar{P}) : \text{maximizar } z$$

sujeito a:

$$\begin{aligned} z &= 0 & -1(-x_1) & -19(-x_2) \\ x_1 &= 0 & -1(-x_1) & +0(-x_2) \\ x_2 &= 0 & +0(-x_1) & -1(-x_2) \\ x_3 &= 50 & +1(-x_1) & +20(-x_2) \\ x_4 &= 20 & +1(-x_1) & +1(-x_2) \\ x_j &\geq 0, \quad j = 1, 2, 3, 4. \end{aligned}$$

As equações acima poderão ser representadas esquematicamente pelo seguinte quadro:

$Q1$	1	$-x_1$	$-x_2$
z	0	-1	-19
x_1	0	-1	0
x_2	0	0	-1
x_3	50	1	20
x_4	20	1	1

O quadro acima denominado $Q1$, possui na primeira linha da esquerda para direita sua denominação $Q1$ associada à coluna das variáveis, 1 associado à coluna dos termos independentes (segundo membro), $-x_1$ associado à coluna da forma $-(1 \ 0 \ -B^{-1}a_1)$, e $-x_2$ associado à coluna da forma $-(0 \ 1 \ -B^{-1}a_2)$. As variáveis x_1 e x_2 são não básicas. Para $x_1 = x_2 = 0$ teremos $x_3 = 50$ e $x_4 = 20$. Assim sendo a matriz $(a_3 \ a_4) = (e_1 \ e_2)$ é básica primal viável de (\bar{P}) . Sobre o quadro $Q1$ poderemos aplicar os testes de otimalidade do método do simplex, assim como a escolha do pivô visando a melhorar o valor da função objetivo z .

Retomemos o quadro inicial $Q1$.

$Q1$	1	$-x_1$	$-x_2$
z	0	-1	-19
x_1	0	-1	0
x_2	0	0	-1
x_3	50	1	20*
x_4	20	1	1

onde * representa o elemento pivô, isto é, a coluna associada à variável x_2 entrará na base substituindo a coluna associada à variável x_3 .

Por operações de pivoteamento (método de eliminação de Gauss-Jordan), já descritas sob forma matricial no capítulo 3, obteremos os quadros seguintes.

$Q2$	1	$-x_1$	$-x_3$
z	$\frac{95}{2}$	$-\frac{1}{20}$	$\frac{19}{20}$
x_1	0	-1	0
x_2	$\frac{5}{2}$	$\frac{1}{20}$	$\frac{1}{20}$
x_3	0	0	-1
x_4	$\frac{35}{20}$	$\frac{19}{20}$ *	$-\frac{1}{20}$

$Q3$	1	$-x_4$	$-x_3$
z	$48\frac{8}{19}$	$\frac{1}{19}$	$\frac{18}{19}$
x_1	$18\frac{8}{19}$	$\frac{20}{19}$	$-\frac{1}{19}$
x_2	$1\frac{11}{19}$	$-\frac{1}{19}$	$\frac{1}{19}$
x_3	0	0	-1
x_4	0	-1	0

O quadro $Q3$ é ótimo (maximização), pois

$$z_4 - c_4 = \frac{1}{19} \geq 0 \quad \text{e} \quad z_3 - c_3 = \frac{18}{19} \geq 0.$$

Esse quadro nos fornece a seguinte solução ótima para (\bar{P}) ,

$$x_1 = 18\frac{8}{19}, \quad x_2 = 1\frac{11}{19}, \quad x_3 = x_4 = 0,$$

forneendo $z = 48\frac{8}{19}$.

Denominemos $val(\cdot)$ o valor da função objetivo no ótimo de (\cdot) e $\lfloor \alpha \rfloor$ o maior inteiro $\leq \alpha$. Então sabemos que, para o exemplo que estamos tratando, $\lfloor val(\bar{P}) \rfloor \geq val(P)$, pois (\bar{P}) é uma relaxação de (P) , o conjunto de soluções possíveis de (P) está contido no de (\bar{P}) e todos os coeficientes da função objetivo de (P) são inteiros. Sabemos que $\lfloor 48\frac{8}{19} \rfloor = 48$, logo se houver uma solução viável de (P) fornecendo para z um valor igual a 48, esta será ótima de (P) .

Tentaremos arredondar a solução ótima de (\bar{P}) visando à busca de uma solução ótima de (P) .

- $x_1 = 19$, $x_2 = 2$, solução inviável para (P) .
- $x_1 = 18$, $x_2 = 1$, $z = 37$.

- $x_1 = 18, x_2 = 2$, solução inviável para (P) .
- $x_1 = 19, x_2 = 1, z = 38$.

Outras soluções não provenientes de arredondamento da solução ótima de (\bar{P}) .

- $x_1 = 0, x_2 = 2, z = 38$.
- $x_1 = 10, x_2 = 2, z = 48$ (solução ótima de (P)).

Pudemos observar nesse exemplo que o puro arredondamento da solução ótima de (\bar{P}) não fornece o ótimo de (P) , nem mesmo uma boa aproximação!

A seguir apresentaremos os métodos de planos de corte ou, simplesmente, métodos de cortes.

10.2 Métodos de Planos de Corte

Consideremos novamente o seguinte problema:

$$(P) : \text{maximizar } x_0 = \sum_{j=1}^n c_j x_j$$

sujeito a:

$$\sum_{j=1}^n a_{ij} x_j = b_i; \quad i = 1, 2, \dots, m;$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n;$$

$$x_j \text{ inteiro, } j \in S \subseteq \{1, 2, \dots, n\}.$$

Onde c_j, a_{ij} e b_i são números reais dados e x_j variável de decisão.

Definamos $F(\cdot)$ como o conjunto de soluções viáveis de (\cdot) . Quando não considerarmos as restrições x_j inteiro, $j \in S \subseteq \{1, 2, \dots, n\}$ em (P) , teremos a relaxação linear de (P) , denominada (\bar{P}) .

Os métodos de planos de corte são motivados visando à determinação da envoltória convexa das soluções viáveis do problema de programação inteira (P) . Uma vez essa envoltória convexa obtida, aplicamos o método do simplex para otimizar a função objetivo de (P) sujeito às restrições que definem a envoltória convexa das soluções viáveis do problema de programação inteira (P) . Sabemos da grande dificuldade em obter a envoltória convexa dos pontos inteiros que satisfaçam desigualdades lineares. O poliedro que define essa envoltória convexa pode possuir um número enorme de desigualdades

lineares, às vezes, este número é de ordem exponencial em função do número de variáveis, por exemplo, 2^n , onde n é o número de variáveis de (P) .

Define-se uma desigualdade linear válida para (P) como uma desigualdade linear $\{x \in \mathcal{R}^n \mid d^T x \leq \alpha\}$, para $d \in \mathcal{R}^n$, e $\alpha \in \mathcal{R}$, tal que para todo $\bar{x} \in F(P)$, tem-se $d^T \bar{x} \leq \alpha$.

Os métodos de planos de corte introduzem a cada iteração uma ou mais desigualdades válidas à relaxação linear de (P) , já designada acima por (\bar{P}) , isto é, introduzem desigualdades válidas ao problema obtido quando eliminamos as restrições de integralidade de (P) . Essas desigualdades válidas eliminam uma solução fracionária de (\bar{P}) , sem eliminar as soluções inteiras viáveis de (P) .

Uma desigualdade válida interessante para $F(P)$ é aquela para qual

$$\{x \in \mathcal{R}^n \mid d^T x = \alpha\} \cap F(\bar{P}) \neq \emptyset.$$

Um método clássico de corte foi proposto em 1958 por Gomory [Go 58a]. Passamos a apresentá-lo para (P) , quando $S = \{1, 2, \dots, n\}$, isto é, todas as variáveis têm que ser inteiras. O método começa resolvendo (\bar{P}) pelo algoritmo do simplex, se a solução obtida for inteira, teremos também resolvido (P) . Caso contrário introduziremos uma desigualdade válida para (P) , que eliminará a solução (\bar{P}) —ótima e um novo problema linear será obtido, o processo se repete até a obtenção de uma solução inteira ou a inexistência dessa solução.

Utilizando as notações dos capítulos anteriores e que $\lfloor \alpha \rfloor$ representa o maior inteiro menor ou igual a $\alpha \in \mathcal{R}$, supomos que tivéssemos resolvido (\bar{P}) com o algoritmo do simplex e que B base ótima de (\bar{P}) : $\bar{x}_B = B^{-1}b \geq 0$, e $z_j - c_j \geq 0$, $j \in I_N$. Existe um $k \in \{1, 2, \dots, m\}$ para o qual $\bar{x}_{B(k)} \notin \mathcal{Z}$, desta maneira a solução ótima de (\bar{P}) não resolve (P) .

Consideremos

$$x_{B(k)} = \bar{x}_{B(k)} - \sum_{j \in I_N} y_{kj} x_j. \quad (10.1)$$

Podemos ainda escrever

$$x_{B(k)} + \sum_{j \in I_N} y_{kj} x_j = \bar{x}_{B(k)},$$

onde $x_{B(k)} \geq 0$, $\bar{x}_{B(k)} \geq 0$, $x_j \geq 0$, $j \in I_N$, então

$$x_{B(k)} + \sum_{j \in I_N} \lfloor y_{kj} \rfloor x_j \leq \bar{x}_{B(k)}.$$

Ainda mais temos que $x_{B(k)}$ e x_j , $j \in I_N$ têm que ser inteiros em (P) . Assim

$$x_{B(k)} + \sum_{j \in I_N} \lfloor y_{kj} \rfloor x_j \leq \lfloor \bar{x}_{B(k)} \rfloor. \quad (10.2)$$

É fácil observar que todas as soluções de $F(P)$ são também viáveis de (10.2). Para transformar (10.2) numa igualdade, introduzimos uma variável de folga $s \geq 0$ e obtemos:

$$x_{B(k)} + \sum_{j \in I_N} \lfloor y_{kj} \rfloor x_j + s = \lfloor \bar{x}_{B(k)} \rfloor. \quad (10.3)$$

Levando em consideração as equações (10.1) e (10.3) teremos

$$\bar{x}_{B(k)} - \sum_{j \in I_N} y_{kj} x_j + \sum_{j \in I_N} \lfloor y_{kj} \rfloor x_j + s = \lfloor \bar{x}_{B(k)} \rfloor$$

ou

$$s = \lfloor \bar{x}_{B(k)} \rfloor - \bar{x}_{B(k)} + \sum_{j \in I_N} (y_{kj} - \lfloor y_{kj} \rfloor) x_j.$$

Definamos $f_{kj} = y_{kj} - \lfloor y_{kj} \rfloor$, $j \in I_N$ e $f_{k0} = \bar{x}_{B(k)} - \lfloor \bar{x}_{B(k)} \rfloor$, é fácil verificar que $0 \leq f_{kj} < 1$, e $0 < f_{k0} < 1$. Assim

$$s = -f_{k0} + \sum_{j \in I_N} f_{kj} x_j,$$

e $s \geq 0$ fornecendo

$$\sum_{j \in I_N} f_{kj} x_j \geq f_{k0}. \quad (10.4)$$

A desigualdade (10.4) é uma desigualdade linear válida para (P) . Esta desigualdade é conhecida como sendo o corte de Gomory, ver [Go 58a].

Para ilustrar o método consideremos o seguinte exemplo.

Exemplo 10.1

$$(P) : \text{maximizar } z = 2x_1 + x_2 + 0x_3 + 0x_4 + 0x_5$$

sujeito a:

$$\begin{array}{rrrrrr} x_1 & + & x_2 & + & x_3 & & = & 5 \\ -x_1 & & x_2 & & & + & x_4 & = & 0 \\ 6x_1 & + & 2x_2 & & & & + & x_5 & = & 21 \\ x_j & \geq & 0, & j = 1, 2, 3, 4, 5 \\ x_j & \text{inteiro} & j = 1, 2, 3, 4, 5. \end{array} \quad (10.5)$$

Para resolver (\bar{P}) utilizaremos o algoritmo do simplex, onde

$$A = (a_1 \ a_2 \ a_3 \ a_4 \ a_5) = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 6 & 2 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 0 \\ 21 \end{pmatrix},$$

$$c^T = (2 \ 1 \ 0 \ 0 \ 0).$$

Consideraremos

$$B = (a_1 \ a_2 \ a_4) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 6 & 2 & 0 \end{pmatrix}, \text{ logo } B^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} \\ \frac{3}{2} & 0 & -\frac{1}{4} \\ -2 & 1 & \frac{1}{2} \end{pmatrix},$$

$$\bar{x}_B = B^{-1}b = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} \\ \frac{3}{2} & 0 & -\frac{1}{4} \\ -2 & 1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 5 \\ 0 \\ 21 \end{pmatrix} = \begin{pmatrix} \frac{11}{4} \\ \frac{9}{4} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \bar{x}_{B(1)} \\ \bar{x}_{B(2)} \\ \bar{x}_{B(3)} \end{pmatrix}.$$

Temos $\bar{x}_{B(1)} = \bar{x}_1$, $\bar{x}_{B(2)} = \bar{x}_2$, e $\bar{x}_{B(3)} = \bar{x}_4$. É fácil de observar que esta base B está associada com uma solução ótima de (\bar{P}) , isto é, $\bar{x}_1 = \frac{11}{4}$, $\bar{x}_2 = \frac{9}{4}$, $\bar{x}_4 = \frac{1}{2}$, and $\bar{x}_3 = \bar{x}_5 = 0$. Esta não é uma solução inteira. Tomemos a linha associada a $\bar{x}_{B(3)}$ para gerar um corte de Gomory. Assim

$$x_{B(3)} = x_4 = \frac{1}{2} + 2x_3 - \frac{1}{2}x_5.$$

Desta restrição obtemos

$$f_{30} = \frac{1}{2}, f_{33} = 0 \text{ e } f_{35} = \frac{1}{2},$$

o corte gerado de Gomory será

$$-\frac{1}{2}x_5 + s_1 = -\frac{1}{2}, \text{ e } s_1 \geq 0.$$

Acrescentaremos uma nova linha e uma nova coluna a matriz A . A nova base formada pelas colunas associadas a x_1 , x_2 , x_4 , e s_1 é dual viável, assim começaremos o problema de otimização utilizando o algoritmo dual do simplex. A regra para a obtenção do pivô para este novo problema diz que a coluna associada à variável s_1 deixa a base e a coluna associada à variável x_5 entra na base.

Seja B_1 a nova base,

$$B_1 = (a_1 \ a_2 \ a_4 \ a_5) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 6 & 2 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix},$$

logo

$$B_1^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{3}{2} & 0 & -\frac{1}{4} & -\frac{1}{2} \\ -2 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix},$$

$$\bar{x}_{B_1} = B_1^{-1}b = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{3}{2} & 0 & -\frac{1}{4} & -\frac{1}{2} \\ -2 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 5 \\ 0 \\ 21 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_4 \\ \bar{x}_5 \end{pmatrix}.$$

B_1 está associada a uma solução ótima, mas esta solução ainda não é inteira. Tomemos a linha associada a \bar{x}_1 para gerar o novo corte de Gomory, obteremos

$$s_2 = -\frac{1}{2} + \frac{1}{2}x_3 + \frac{1}{2}s_1, \quad s_2 \geq 0.$$

A regra do simplex nos indica que a coluna associada à variável s_2 deixará a base e a coluna associada à variável x_3 entrará. Seja B_2 a nova base,

$$B_2 = (a_1 \ a_2 \ a_3 \ a_4 \ a_5) = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 6 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix},$$

logo

$$B_2^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{2} & -1 \\ \frac{3}{2} & 0 & -\frac{1}{4} & -\frac{1}{2} & 3 \\ 0 & 0 & 0 & 0 & -2 \\ -2 & 1 & \frac{1}{2} & 1 & -4 \\ 0 & 0 & 0 & -2 & 0 \end{pmatrix},$$

$$\bar{x}_{B_2} = B_2^{-1}b = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{2} & -1 \\ \frac{3}{2} & 0 & -\frac{1}{4} & -\frac{1}{2} & 3 \\ 0 & 0 & 0 & 0 & -2 \\ -2 & 1 & \frac{1}{2} & 1 & -4 \\ 0 & 0 & 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 0 \\ 21 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \\ \bar{x}_5 \end{pmatrix}.$$

Esta última solução é uma solução ótima inteira, resolvendo (P) . Para esse pequeno exemplo tivemos que gerar apenas dois cortes de Gomory para resolvê-lo. O método de Gomory possuirá convergência finita quando todos os dados de (P) forem inteiros, mas para isso tem-se que seguir uma ordem especial na geração dos cortes apresentada por Gomory em [Go 58a], ver [Sa 75, Ma 78].

A seguir apresentaremos um método de solução para (P) quando todas as variáveis forem restritas a tomarem valores zero ou um.

10.3 Método de Balas para Otimização Linear 0-1

Balas apresentou em [Ba 63] e [Ba 65] um método de enumeração implícita para a solução do problema de programação linear inteira 0 – 1 ou programação linear bivalente, no qual somente as operações de soma, de subtração e de comparação são utilizadas.

Consideremos o seguinte problema

$$(P) : \text{minimizar } x_0 = \sum_{j=1}^n c_j x_j$$

sujeito a:

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, m$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, \dots, n,$$

onde c_j, a_{ij} e b_i são números reais dados e x_j variável bivalente de decisão.

O problema (P) definido acima é de programação linear bivalente $(0, 1)$.

Poderemos sempre supor, sem perda de generalidade, que $c_j \geq 0$, $j = 1, 2, \dots, n$, pois caso exista um k para o qual $c_k < 0$, faremos uma mudança de variável: $x_k = 1 - t_k$, $t_k \in \{0, 1\}$. Seja $I = \{1, 2, \dots, n\}$, assim teremos $x_0 = \sum_{j \in I - \{k\}} c_j x_j - c_k t_k + c_k$.

E passaremos a minimizar

$$x_0 - c_k = \sum_{j \in I - \{k\}} c_j x_j - c_k t_k.$$

Durante toda esta seção suporemos que $c_j \geq 0$, $j = 1, 2, \dots, n$.

Definiremos também $c = (c_1 \ c_2 \ \dots \ c_n)$, $b^T = (b_1 \ b_2 \ \dots \ b_m)$ e $A = (a_{ij})_{m \times n}$, uma matriz com m linhas e n colunas; $x^T = (x_1 \ x_2 \ \dots \ x_n)$.

O problema (P) poderá ser também escrito sob a seguinte forma:

$$(P) : \text{minimizar } x_0 = cx \tag{10.6}$$

sujeito a:

$$Ax \leq b \tag{10.7}$$

$$x \in \{0, 1\}^n. \tag{10.8}$$

Uma solução de (10.8) será representada pelo vetor $x^p = (x_1^p \ x_2^p \ \dots \ x_n^p)$, ou também pelo conjunto $J_p = \{j \mid x_j = 1\}$, por exemplo

$$x^7 = (0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1)^T \in \{0, 1\}^8, \quad \text{ou} \quad J_7 = \{2, 3, 4, 7, 8\}.$$

Uma solução x^q de (10.8) é dita descendente de x^p se $J_p \subset J_q$, por exemplo

$$x^q = (1 \ 1 \ 1 \ 0 \ 1 \ 1)^T \text{ é descendente de } x^p = (0 \ 0 \ 1 \ 0 \ 1 \ 1)^T.$$

Em alguns problemas certas variáveis são fixadas *a priori*, vejamos nos exemplos a seguir:

$$3x_1 + 7x_2 + 2x_3 + x_4 \leq 5 \text{ implica que } x_2 = 0.$$

$$2x_1 + 8x_2 + x_3 + x_4 \geq 9 \text{ implica que } x_2 = 1.$$

No caso em que ambas restrições acima façam parte de (10.7), o problema (P) será vazio.

Proposição 10.1 *Se $x^p \in \{0, 1\}^n$, então $\sum_{j \in J_p} c_j x_j \leq \sum_{j \in J_k} c_j x_j$ para todas as soluções x^k descendentes de x^p .*

Demonstração:

Basta lembrarmos que $c_j \geq 0$, $j = 1, 2, \dots, n$ e que $J_p \subset J_k$. ■

Proposição 10.2 *Se $x^0 = (0 \ 0 \ \dots \ 0)^T$ satisfizer (10.7) então x^0 será também uma solução ótima de (P).*

Demonstração:

$x_0 = \sum_{j=1}^n c_j x_j \geq 0$, mas $cx^0 = 0$, assim sendo x_0 é uma solução ótima de (P). ■

Observação: o esquema de enumeração que será apresentado a seguir supõe que a solução inicial x^0 seja tal que $J_0 = \phi$, isto é, $x^0 = (0 \ 0 \ \dots \ 0)^T$; pois caso x^0 satisfaça também 10.7, neste caso $b_i \geq 0$, $i = 1, 2, \dots, m$, então x^0 será uma solução ótima de (P).

10.3.1 Esquema de Enumeração

Suponhamos que estejamos em uma solução x^p de (10.8) e que \bar{x}_0 seja o melhor valor da função objetivo de (P) associada a uma solução viável de (P), isto é, existe x^k , tal que $\bar{x}_0 = \sum_{j \in J_k} c_j$ e que x^k satisfaça (10.7) e (10.8). Caso não tenhamos ainda encontrado uma solução viável, colocamos $\bar{x}_0 = +\infty$.

A partir da solução x^p desejamos obter x^q descendente de x^p , tal que $|J_q| = |J_p| + 1$ ou melhor $J_q = J_p \cup \{l\}$.

Consideremos as seguintes hipóteses.

1. x^p é viável de (P) , neste caso não nos interessa buscar uma solução x^q descendente de x^p , pois $\sum_{j \in J_p} c_j \leq \sum_{j \in J_q} c_j$, ver propriedade 1.
2. Se

$$\sum_{j \in J_p} c_j + c_l \geq \bar{x}_0, \quad \forall l \notin J_p,$$

isto quer dizer que todas as soluções descendentes de x^p fornecerão valores à função objetivo (10.6) sempre superiores ou iguais a \bar{x}_0 , neste caso não nos interessará também enumerar as soluções descendentes de x^p .

3. Se existir um i tal que

$$b_i - \sum_{j \in J_p} a_{ij} - \sum_{j \notin J_p} \min\{0, a_{ij}\} < 0,$$

nunca haverá uma solução descendente de x^p viável de (P) , desta maneira não teremos interesse em enumerar as descendentes de x^p .

As três condições discutidas acima serão definidas como condições de parada em x^p .

Quando tivermos parado em x^p por uma das três condições, teremos enumerado implicitamente todas as soluções descendentes de x^p .

Caso em x^p não satisfaçamos nenhuma condição de parada, temos que procurar uma solução descendente de x^p , por exemplo, x^q , tal que $J_q = J_p \cup \{l\}$, onde, evidentemente, $l \notin J_p$.

Consideremos $s_i = b_i - \sum_{j=1}^n a_{ij}x_j$ e $s_i \geq 0, i = 1, 2, \dots, m$, s_i será a variável de folga associada à restrição i .

Seja $s_i^p = b_i - \sum_{j \in J_p} a_{ij}$, isto é, s_i^p representa o valor de s_i quando $x = x^p$.

Em x^p podemos definir os seguintes conjuntos:

$$A_p = \left\{ k \mid \sum_{j \in J_p} c_j + c_k \geq \bar{x}_0, \quad k \notin J_p \right\},$$

$$D_p = \{k \mid \text{se para todo } i \text{ com } s_i^p < 0, \quad a_{ik} \geq 0, \quad k \notin J_p\},$$

e

$$C_p = \{1, 2, \dots, n\} - (J_p \cup A_p \cup D_p).$$

O conjunto C_p fornecerá os índices das variáveis candidatas a tomarem valor igual a um, isto é, um índice $l \in C_p$ para formar $J_q = J_p \cup \{l\}$.

Seja $d_j^p = \sum_{i=1}^m \min\{0, s_i^p - a_{ij}\}$, $j \in C_p$ e

$$d_l^p = \max_{j \in C_p} \{d_j^p\}.$$

Dizemos que d_l^p é a “menor” soma das inviabilidades.

Caso $d_l^p = 0$, então a solução descendente associada ao conjunto $J_q = J_p \cup \{l\}$ será viável de (P) .

No caso em que haja mais de um índice para o qual $d_j^p = 0$, isto é, o conjunto

$$L_p = \{j \in C_p \mid d_j^p = 0\}$$

possui cardinalidade maior ou igual a dois. O índice $l \in L_p$ escolhido para formar a solução descendente de x^p será o l associado a

$$c_l = \min_{j \in L_p} \{c_j\}.$$

É fácil verificar que no caso de que nenhuma solução de parada seja verificada teremos $C_p \neq \phi$.

Supomos agora que em x^q descendente de x^p , o conjunto C_q seja vazio, ou ainda, uma das três condições de parada seja verificada. Teremos que voltar de x^q para x^p e para isso atualizaremos C_p de duas maneiras:

1. $C_p := C_p - \{l\}$, onde l é tal que $J_q = J_p \cup \{l\}$,
2. A_p poderá ser modificado caso \bar{x}_0 também o seja, acarretando outra modificação em C_p .

O retorno de x^q para x^p é denominado de *backtracking*.

A enumeração termina completamente quando $C_0 = \phi$. Lembremos que \bar{x}_0 é sempre atualizado durante a enumeração. Caso (P) seja vazio então $\bar{x}_0 = +\infty$ no final da enumeração.

Exemplo 10.2 Este exemplo foi tomado de [Ba 65].

$$(P) : \text{minimizar } x_0 = 5x_1 + 7x_2 + 10x_3 + 3x_4 + x_5$$

sujeito a:

$$\begin{array}{rrrrrrrrcl} -x_1 & + & 3x_2 & - & 5x_3 & - & x_4 & + & 4x_5 & \leq & -2 \\ 2x_1 & - & 6x_2 & + & 3x_3 & + & 2x_4 & - & 2x_5 & \leq & 0 \\ & & x_2 & - & 2x_3 & + & x_4 & + & x_5 & \leq & -1 \end{array}$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3, 4, 5.$$

Definiremos as variáveis de folga s_i :

$$\begin{array}{rcccccccc} s_1 & = & -2 & + & x_1 & - & 3x_2 & + & 5x_3 & + & x_4 & - & 4x_5 & \geq & 0 \\ s_2 & = & 0 & - & 2x_1 & + & 6x_2 & - & 3x_3 & - & 2x_4 & + & 2x_5 & \geq & 0 \\ s_3 & = & -1 & & & - & x_2 & + & 2x_3 & - & x_4 & - & x_5 & \geq & 0 \end{array}$$

Etapa inicial

$$x^0 = (0 \ 0 \ 0 \ 0 \ 0)^T, \ J_0 = \phi, \ s_1^0 = -2, \ s_2^0 = 0, \ s_3^0 = -1.$$

Como $s_i^p < 0$, $i = 1, 3$, então x^0 é inviável. Suporemos $\bar{x}_0 = +\infty$, isto é, não conhecemos nenhuma solução viável para (P) . As duas primeiras condições de parada não são verificadas, vejamos a terceira:

$$-2 + 1 + 0 + 5 + 1 + 0 = 5 \geq 0,$$

$$0 + 0 + 6 + 0 + 0 + 2 = 8 \geq 0,$$

$$-1 + 0 + 0 + 2 + 0 + 0 = 1 \geq 0,$$

também não é verificada.

Necessitamos encontrar uma solução descendente de x^0 . Para isso definiremos os seguintes conjuntos: $A_0 = \phi$, pois $\bar{x}_0 = +\infty$, $D_0 = \{2, 5\}$, logo

$$C_0 = \{1, 2, 3, 4, 5\} - (A_0 \cup D_0 \cup J_0) = \{1, 3, 4\},$$

isto é, as variáveis x_1 , x_3 e x_4 são as candidatas a tomarem o valor um (apenas uma entre elas o tomará). Calculemos d_j^0 , $j \in C_0$:

$$d_1^0 = -1 - 2 - 1 = -4; \ d_3^0 = 0 - 3 + 0 = -3; \ d_4^0 = -1 - 2 - 2 = -5$$

e

$$d_3^0 = \max\{d_1^0, d_3^0, d_4^0\} = -3;$$

ir à *etapa 1*.

Etapa 1

$J_1 = J_0 \cup \{3\} = \{3\}$, pois $J_0 = \phi$. Teremos então $s_1^1 = 3$, $s_2^1 = -3$, $s_3^1 = 1$. Assim sendo $x^1 = (0 \ 0 \ 1 \ 0 \ 0)^T$ não é viável, \bar{x}_0 continua igual a $+\infty$. As duas primeiras condições de parada não são satisfeitas e a terceira

$$3 + 1 + 0 + 1 + 0 = 5 \geq 0,$$

$$-3 + 0 + 6 + 0 + 2 = 5 \geq 0,$$

$$1 + 0 + 0 + 0 + 0 = 1 \geq 0,$$

também não é verificada. Passaremos então a buscar uma solução descendente de x^1 .

$A_1 = \phi$, pois $\bar{x}_0 = +\infty$. Lembremos que $D_1 = \{1, 4\}$, logo

$$C_1 = \{1, 2, 3, 4, 5\} - (\{1, 4\} \cup \{3\}) = \{2, 5\}.$$

Sabemos também que $d_2^1 = 0 + 0 + 0 = 0$, $d_5^1 = -1 - 1 + 0 = -2$, logo $d_2^1 = \max\{d_2^1, d_5^1\} = 0$ (a próxima solução descendente será viável), ir à *etapa 2*.

Etapa 2

$J_2 = J_1 \cup \{2\} = \{2, 3\}$, $x^2 = (0 \ 1 \ 1 \ 0 \ 0)^T$, $s_1^2 = 0$, $s_2^2 = 3$, $s_3^2 = 0$, x^2 é viável, $cx^2 = c_2 + c_3 = 7 + 10 = 17$, logo \bar{x}_0 será atualizado, $\bar{x}_0 = 17$. A primeira regra de parada é satisfeita. Ir à *etapa 3*.

Etapa 3 (backtracking)

Na realidade voltamos à *etapa 1*, onde $J_1 = \{3\}$ e x^1 não é viável, mas agora $\bar{x}_0 = 17$. Impediremos que a variável x_2 seja candidata a tomar valor igual a um, ela ficará zerada, para que não encontremos novamente a solução associada à *etapa 2*. Neste caso a terceira condição de parada é verificada:

$$3 + 1 + 1 + 0 = 5 \geq 0,$$

$$-3 + 0 + 0 + 2 = -1 < 0,$$

$$1 + 0 + 0 + 0 = 1 \geq 0.$$

Ir para a *etapa 4*.

Etapa 4 (backtracking)

Voltamos à *etapa inicial*, onde $J_0 = \phi$ e a variável x_3 será fixada a zero. Lembremos que x_2 é uma variável candidata novamente a tomar o valor um. Sabemos que $\bar{x}_0 = 17$. A terceira condição de parada é verificada:

$$-2 + 1 + 0 + 1 + 0 = 0 \geq 0$$

$$0 + 0 + 6 + 0 + 2 = 8 \geq 0$$

$$-1 + 0 + 0 + 0 + 0 = -1 < 0.$$

Como estamos na *etapa inicial* e não temos mais possibilidades de buscar soluções descendentes viáveis com $x_3 = 0$, assim sendo temos a parada final.

A melhor solução encontrada é $x^2 = (0 \ 1 \ 1 \ 0 \ 0)^T$, fornecendo $\bar{x}_0 = 17$. Logo x^2 é a solução ótima.

Observemos no exemplo que na *etapa 2* temos $J_2 = \{3, 2\}$ e quando voltamos à *etapa 1* (*etapa 4*), $J_1 = \{3\}$ fixando $x_2 = 0$, por convenção podemos designar esta etapa pelo conjunto $\{3, -2\}$, isto significa que $x_3 = 1$ e $x_2 = 0$ para toda busca dos descendentes. ■

10.3.2 Convergência do Método de Balas

A seguir forneceremos uma enumeração implícita finita baseada nos trabalhos de Glover [Gl 65] e Geoffrion [Ge 67]. Nunca enumeraremos explicitamente a mesma solução mais de uma vez, assim sendo a enumeração termina.

Enunciaremos o algoritmo de Balas, utilizando uma estrutura de pilha. Essa pilha representa o conjunto dos índices associados às variáveis fixadas.

Seja a pilha π , para a qual $p(j)$ será sua j -ésima componente, tal que:

$$p(j) > 0 \text{ se } x_{p(j)} = 1 \text{ e}$$

$$p(j) < 0 \text{ se } x_{p(j)} = 0.$$

Por exemplo,

$$\pi = [-3, 2, -7, -4],$$

isto é,

$$x_3 = 0, \ x_2 = 1, \ x_7 = 0, \ x_4 = 0$$

são os valores fixos, todas as soluções descendentes x^k da solução associada a π não terão os índices 3, 7 e 4 pertencendo a J_k .

Procedimento de Balas

Fase 0 (inicialização)

$$\pi = \phi; \quad \bar{x}_0 = +\infty;$$

Fase 1 Se uma das condições de parada for verificada (no caso de ser a primeira, isto é, π está associada a uma solução viável de (P) , x^p , neste caso se $cx^p < \bar{x}_0$ far-se-á $\bar{x}_0 = cx^p$ e a melhor solução até o momento é x^p), ir para a fase 2;

Caso contrário, ir para a fase 3;

- Fase 2 Enquanto o último elemento da pilha for negativo, removê-lo da pilha;
 Se a pilha for vazia, ir para a fase 4;
 Caso o último elemento da pilha for positivo, trocar seu sinal e ir para a fase 1;
- Fase 3 Escolher uma variável x_k pelo critério aconselhado e colocar na pilha o elemento k , ir para a fase 1;
- Fase 4 Pare, solução associada a \bar{x}_0 será ótima se $\bar{x}_0 \neq +\infty$; o problema será vazio se $\bar{x}_0 = +\infty$;

No exemplo tratado anteriormente teríamos:

$$\begin{aligned}\pi &= \phi \text{ (início);} \\ \pi &= [3]; \\ \pi &= [3, 2]; \\ \pi &= [3, -2]; \\ \pi &= [-3]; \\ \pi &= \phi \text{ (fim).}\end{aligned}$$

10.3.3 Otimização Não Linear 0-1

Nos problemas de Otimização Não Linear em variáveis 0-1, quando os termos não lineares são produtos de variáveis 0-1, podemos linearizá-los da seguinte maneira.

Seja o produtório de p variáveis $x_j \in \{0, 1\}$, $j = 1, 2, \dots, p$, isto é, $\prod_{j=1}^p x_j$. Utilizando uma transformação de Fortet [Fo 60], teremos:

$$\begin{aligned}y &= \prod_{j=1}^p x_j, \\ y &\geq 0, \\ y &\leq x_j, \quad j = 1, 2, \dots, p, \\ \sum_{j=1}^p x_j - p + 1 &\leq y.\end{aligned}$$

Verificamos facilmente se $x_j = 0$ teremos $y = 0$, e se $x_j = 1$, para $j = 1, 2, \dots, p$, teremos $y = 1$. Assim sendo $y \in \{0, 1\}$.

Exemplo 10.3

$$\begin{array}{ll}\text{minimizar} & x_0 = 3x_1 + 4x_2^5 - 5x_1x_2x_3 + 3x_1x_3, \\ \text{sujeito a :} & \\ & x_j \in \{0, 1\}, \quad j = 1, 2, 3.\end{array}$$

Como $x_2 \in \{0, 1\}$, logo $x_2^5 = x_2$.

Façamos $y = x_1x_2x_3$ e $t = x_1x_3$.

O problema poderá ser escrito da seguinte forma:

$$\begin{array}{ll}\text{minimizar} & x_0 = 3x_1 + 4x_2 - 5y + 3t, \\ \text{sujeito a:} & \\ & y \leq x_1, \quad y \leq x_2, \quad y \leq x_3, \quad x_1 + x_2 + x_3 - 2 \leq y, \\ & t \leq x_1, \quad t \leq x_3, \quad x_1 + x_3 - 1 \leq t, \\ & y \geq 0, \quad t \geq 0 \text{ e } x_j \in \{0, 1\}, \quad j = 1, 2, 3.\end{array}$$

Ou ainda:

$$\begin{array}{ll}\text{minimizar} & x_0 = 3x_1 + 4x_2 - 5y + 3t, \\ \text{sujeito a:} & \\ & y \leq x_1, \quad y \leq x_2, \quad y \leq x_3, \quad x_1 + x_2 + x_3 - 2 \leq y, \\ & t \leq x_1, \quad t \leq x_3, \quad x_1 + x_3 - 1 \leq t, \\ & y \in \{0, 1\}, \quad t \in \{0, 1\} \text{ e } x_j \in \{0, 1\}, \quad j = 1, 2, 3.\end{array}$$

Poderemos resolver esta última formulação do problema pelo método de Balas. ■

10.4 Métodos de “Branch-and-Bound”

Os métodos de “branch-and-bound” foram desenvolvidos a partir do trabalho pioneiro de Land e Doig [LaDo 60]. O termo “branch-and-bound” foi empregado pela primeira vez em [LiMuSwKa 63].

Apresentaremos a seguir a versão proposta por Dakin [Da 65] para os métodos de enumeração implícita do tipo “branch-and-bound”.

10.4.1 Ilustração dos Métodos de “Branch-and-Bound”

Seja

$$(P) : \text{maximizar } x_0 = c_1x + c_2y$$

sujeito a:

$$A_1x + A_2y \leq b$$

$$x \geq 0$$

$$y \geq 0$$

$$x \text{ inteiro},$$

onde $c_1^T \in \mathcal{R}^p$, $c_2^T \in \mathcal{R}^q$, $x \in \mathcal{R}^p$, $y \in \mathcal{R}^q$, $b \in \mathcal{R}^m$, $A_2 \in \mathcal{R}^{m \times p}$ e $A_1 \in \mathcal{R}^{m \times q}$.

Sabemos que (P) pode também ser escrito como se segue:

$$(P) : \text{maximizar } x_0 \tag{10.9}$$

$$c_1x + c_2y - x_0 \geq 0 \tag{10.10}$$

$$A_1x + A_2y \leq b \tag{10.11}$$

$$x \geq 0 \tag{10.12}$$

$$y \geq 0 \tag{10.13}$$

$$x \text{ inteiro.} \tag{10.14}$$

O conjunto formado pelas restrições (10.10) até (10.13), caso não seja vazio, é um conjunto poliédrico convexo em \mathcal{R}^{p+q+1} . Projetemos este conjunto poliédrico sobre o plano $x_l \times x_0$, onde x_l é uma componente de x . Esta projeção está ilustrada à figura 10.1.

Podemos observar que $\bar{x}_0 = \text{val}(\bar{P})$, onde (\bar{P}) é a relaxação linear de (P) , isto é, a restrição de integralidade (10.14) não é considerada. Suponhamos que $\bar{x}_l \notin \mathcal{Z}$ seja o valor de x_l na solução ótima de (\bar{P}) . Verificamos, na figura 10.1, que x_0^1 é o maior valor que x_0 poderá assumir quando somente a variável x_l deva ser inteira. Dependendo da ilustração x_0^2 poderia ser maior do que x_0^1 . Logo poderemos concluir que o maior valor de x_0 supondo apenas a variável x_l restrita a valores inteiros será observado para $x_l = \lfloor \bar{x}_l \rfloor$ ou para $x_l = \lfloor \bar{x}_l \rfloor + 1$.

Denominemos $F(\cdot)$ o conjunto dos pontos que satisfazem as restrições do problema (\cdot) .

O método proposto por Dakin [Da 65] começa verificando se a solução de (\bar{P}) satisfaz às restrições de integralidade (10.14). Caso afirmativo teremos também resolvido (P) . Caso contrário tomaremos uma variável x_l para a qual o seu valor x_l^0 não seja inteiro no ótimo de (\bar{P}) , construiremos dois problemas descendentes diretos:

- (P_1) , para o qual $F(P_1) = F(\bar{P}) \cap \{x \in \mathcal{R}^p \mid x_l \leq \lfloor x_l^0 \rfloor\}$;
- (P_2) , para o qual $F(P_2) = F(\bar{P}) \cap \{x \in \mathcal{R}^p \mid x_l \geq \lfloor x_l^0 \rfloor + 1\}$.

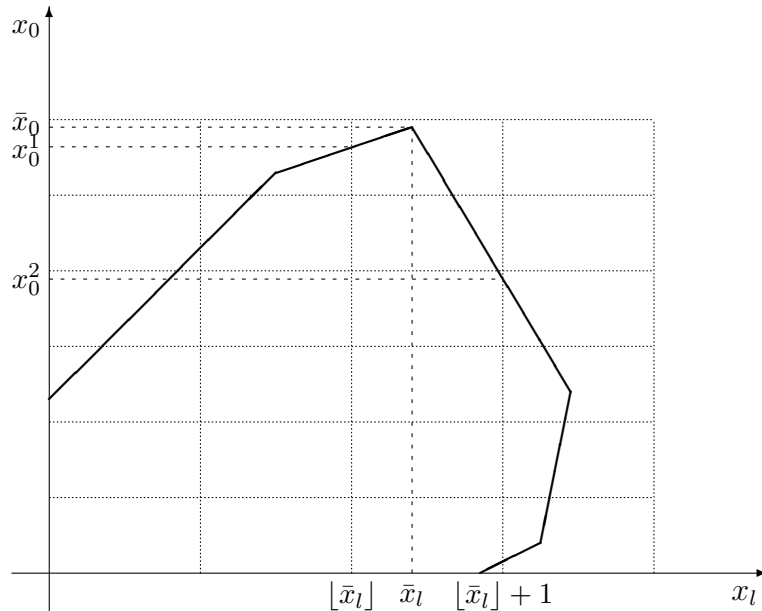


Figura 10.1: Projeção sobre $x_l \times x_0$

Introduziremos a noção de nós abertos e fechados associados aos problemas (P_i) . Cada problema (P_i) será associado a um nó (que será também designado de (P_i)) de uma árvore binária desenvolvida para enumerar (implicitamente) as soluções de (P) . Um problema (P_i) ou seu nó associado é dito fechado se:

- $F(P_i) = \phi$;
- (P_i) possui uma solução ótima onde todas as componentes de x são inteiras;
- $val(P_i) \leq \hat{x}_0$ (no caso de maximização de x_0), onde \hat{x}_0 é o melhor valor obtido para x_0 , tal que $(x, y) \in F(P)$, até o desenvolvimento atual da árvore de enumeração. Dizemos que \hat{x}_0 é a melhor cota inferior para $val(P)$ até o momento considerado da enumeração.

Outras condições poderão ser introduzidas visando a fechar um nó.

Um nó será aberto se tivermos que construir e resolver seus dois nós descendentes diretos.

Podemos iniciar com um \hat{x}_0 associado a uma solução heurística para (P) , isto é, $(x, y) \in F(P)$, ou começar com $\hat{x}_0 = -\infty$, para o caso de maximização. Há métodos heurísticos interessantes para várias classes de problemas combinatórios, ver [CaMa 94].

Teremos uma lista de nós abertos e uma outra de nós fechados. Quando a lista de nós abertos se tornar vazia, o método de *branch-and-bound* terminará. Algo que será melhor entendido através de um exemplo mais abaixo.

Sejam (P_q) e (P_{q+1}) os descendentes diretos de (P_i) :

- (P_q) , para o qual $F(P_q) = F(P_i) \cap \{x \in \mathcal{R}^p \mid x_l \leq \lfloor x_l^i \rfloor\}$;
- (P_{q+1}) , para o qual $F(P_{q+1}) = F(P_i) \cap \{x \in \mathcal{R}^p \mid x_l \geq \lfloor x_l^i \rfloor + 1\}$.

Onde x_l^i é o valor não inteiro da variável x_l na solução de (P_i) .

Como $F(P_q) \subseteq F(P_i)$ e $F(P_{q+1}) \subseteq F(P_i)$ então teremos que $val(P_q) \leq val(P_i)$ e $val(P_{q+1}) \leq val(P_i)$.

É fácil verificar que se $F(P_q) \neq \phi$, teremos que $x_l^q = \lfloor x_l^i \rfloor$ e que se $F(P_{q+1}) \neq \phi$, teremos que $x_l^{q+1} = \lfloor x_l^i \rfloor + 1$, onde x_l^q é o valor de x_l no ótimo de (P_q) e x_l^{q+1} é o valor de x_l no ótimo de (P_{q+1}) .

Exemplo 10.4 Seja

$$(P) : \text{maximizar } x_0 = -4x_1 - 5x_2$$

sujeito a:

$$\begin{array}{rcccccccl} x_1 & + & 4x_2 & - & x_3 & & & = & 5 \\ 3x_1 & + & 2x_2 & & & - & x_4 & = & 7 \end{array}$$

$$\begin{array}{rcl} x_1 & \geq & 0 \\ x_2 & \geq & 0 \\ x_3 & \geq & 0 \\ x_4 & \geq & 0 \end{array}$$

x_1 e x_2 inteiros.

Inicialmente teremos que resolver (\bar{P}) , isto é, consideraremos a solução da relaxação linear de (P) , as restrições de integralidade não serão levadas em consideração. Para isso utilizaremos o método do simplex. Para facilitar o desenvolvimento deste exemplo definiremos os seguintes vetores:

$$\begin{aligned} a_1 &= (1 \ 3)^T, \ a_2 = (4 \ 2)^T, \ a_3 = (-1 \ 0)^T, \ a_4 = (0 \ -1)^T, \\ b &= (5 \ 7)^T, \ c = (-4 \ -5 \ 0 \ 0) = (c_1 \ c_2 \ c_3 \ c_4). \end{aligned}$$

Verificamos que a matriz $B = (a_3 \ a_4) = -I$ é inversível e $B^{-1} = -I$.
Utilizando as notações anteriores temos que

$$\bar{x}_B = B^{-1} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} -5 \\ -7 \end{pmatrix} = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} \not\geq 0,$$

$$u = (c_3 \ c_4)B^{-1} = (0 \ 0)B^{-1} = (0 \ 0),$$

$$z_1 - c_1 = (0 \ 0) \begin{pmatrix} 1 \\ 3 \end{pmatrix} - (-4) = 4 \geq 0,$$

$$z_2 - c_2 = (0 \ 0) \begin{pmatrix} 4 \\ 2 \end{pmatrix} - (-5) = 5 \geq 0.$$

A base B é dual viável de (\bar{P}) . Assim sendo, utilizaremos o método dual do simplex para a solução de (\bar{P}) .

Faremos a coluna a_4 sair da base, pois a_4 está associada à variável x_4 , que está com valor negativo, $x_4 = -7$. Qual será a coluna que entrará na base no lugar de a_4 ? Calcularemos y_{21} e y_{22} .

$$\begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix} = y_1 = B^{-1}a_1, \quad \begin{pmatrix} y_{12} \\ y_{22} \end{pmatrix} = y_2 = B^{-1}a_2,$$

ou ainda,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} ** \\ -3 \end{pmatrix} = \begin{pmatrix} ** \\ y_{21} \end{pmatrix} \Rightarrow y_{21} = -3$$

e

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} ** \\ -2 \end{pmatrix} = \begin{pmatrix} ** \\ y_{22} \end{pmatrix} \Rightarrow y_{22} = -2.$$

Façamos o teste da razão:

$$\min \left\{ \left| \frac{z_1 - c_1}{y_{21}} \right|, \left| \frac{z_2 - c_2}{y_{22}} \right| \right\} = \min \left\{ \left| \frac{4}{-3} \right|, \left| \frac{5}{-2} \right| \right\} = \left| \frac{4}{-3} \right| = \left| \frac{z_1 - c_1}{y_{21}} \right|,$$

logo a_1 entrará na base no lugar de a_4 .

A nova

$$B = \begin{pmatrix} -1 & 1 \\ 0 & 3 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}.$$

Teremos que

$$\bar{x}_B = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} -\frac{8}{3} \\ \frac{7}{3} \end{pmatrix} \not\geq 0,$$

a_3 sairá da base. Façamos novamente as operações de atualização com a nova base.

$$\begin{aligned} u &= (c_3 \ c_1)B^{-1} = (0 \ -4) \begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} = (0 \ -\frac{4}{3}), \\ z_2 - c_2 &= ua_2 - c_2 = \left(0 \ -\frac{4}{3}\right) \begin{pmatrix} 4 \\ 2 \end{pmatrix} - (-5) = -\frac{8}{3} + 5 = \frac{7}{3}, \\ z_4 - c_4 &= ua_4 - c_4 = \left(0 \ -\frac{4}{3}\right) \begin{pmatrix} 0 \\ -1 \end{pmatrix} - 0 = \frac{4}{3}. \end{aligned}$$

Temos que calcular os novos valores de y_{12} e y_{14} :

$$\begin{pmatrix} y_{12} \\ y_{22} \end{pmatrix} = y_2 = B^{-1}a_2, \quad \begin{pmatrix} y_{14} \\ y_{24} \end{pmatrix} = y_4 = B^{-1}a_4,$$

ou ainda,

$$\begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -\frac{10}{3} \\ ** \end{pmatrix} = \begin{pmatrix} y_{12} \\ ** \end{pmatrix} \Rightarrow y_{12} = -\frac{10}{3}$$

e

$$\begin{pmatrix} -1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ ** \end{pmatrix} = \begin{pmatrix} y_{14} \\ ** \end{pmatrix} \Rightarrow y_{14} = -\frac{1}{3}.$$

Façamos o teste da razão:

$$\begin{aligned} \min \left\{ \left| \frac{z_2 - c_2}{y_{12}} \right|, \left| \frac{z_4 - c_4}{y_{14}} \right| \right\} &= \\ \min \left\{ \left| \frac{7/3}{-10/3} \right|, \left| \frac{4/3}{-1/3} \right| \right\} &= \left| \frac{7/3}{-10/3} \right| = \frac{7}{10} = \left| \frac{z_2 - c_2}{y_{12}} \right|, \end{aligned}$$

logo a_2 entrará na base no lugar de a_3 .

Obtivemos a nova

$$B = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix}.$$

Teremos que

$$\bar{x}_B = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{8}{10} \\ \frac{18}{10} \end{pmatrix} \geq 0,$$

Esta última solução é ótima de (\bar{P}) , ela é primal e dual viável de (\bar{P}) . No entanto, $\bar{x}_1 \notin \mathcal{Z}$ e $\bar{x}_2 \notin \mathcal{Z}$, isto é, essas duas variáveis restritas a serem inteiras não o são, assim sendo a solução de (\bar{P}) não resolve (P) .

Para iniciar o método de enumeração escolheremos a variável x_1 para realizarmos as duas ramificações a partir de (\bar{P}) . Lembremos que na solução de (\bar{P}) obtivemos $x_1 = \bar{x}_1 = \frac{18}{10} = 1,8 \notin \mathcal{Z}$. Formaremos os dois problemas (P_1) e (P_2) da seguinte maneira:

- $F(P_1) = F(\bar{P}) \cap \{x \in \mathcal{R}^4 \mid x_1 \leq 1\}$ e
- $F(P_2) = F(\bar{P}) \cap \{x \in \mathcal{R}^4 \mid x_1 \geq 2\}$.

O próximo passo será a solução de (P_1) e (P_2) . Para isso será utilizado o método dual do simplex especializado para tratar restrições canalizadas. É importante notar que a solução básica ótima de (\bar{P}) é também uma solução básica dual viável de (P_1) e (P_2) .

Inicia-se o procedimento de enumeração fixando $\hat{x}_0 = -\infty$ (melhor valor da função objetivo obtida até este momento). No caso de não se ter uma solução viável para (P) , faz-se $\hat{x}_0 = -\infty$.

Na resolução dos (P_i) , $i = 1, 2, \dots$ serão utilizadas as notações do capítulo 8. No caso de (P_1) tem-se que $I_\alpha = \{3, 4\}$, $I_\beta = \phi$, $\alpha(1) = \alpha(2) = \alpha(3) = \alpha(4) = 0$, $\beta(1) = 1$, $\beta(2) = \beta(3) = \beta(4) = +\infty$. Observa-se que a variável x_1 está associada à segunda linha da matriz $B^{-1}N$, cujos elementos são calculados a seguir.

$$\begin{pmatrix} y_{13} \\ y_{23} \end{pmatrix} = y_3 = B^{-1}a_3, \quad \begin{pmatrix} y_{14} \\ y_{24} \end{pmatrix} = y_4 = B^{-1}a_4,$$

ou ainda,

$$\begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{3}{10} \\ \frac{2}{10} \end{pmatrix} = \begin{pmatrix} y_{13} \\ y_{23} \end{pmatrix} \Rightarrow y_{13} = -\frac{3}{10}$$

e

$$\begin{pmatrix} \frac{3}{10} & -\frac{1}{10} \\ -\frac{2}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{10} \\ -\frac{4}{10} \end{pmatrix} = \begin{pmatrix} y_{14} \\ y_{24} \end{pmatrix} \Rightarrow y_{24} = -\frac{4}{10}.$$

$$L_\alpha = \{j \in I_\alpha \mid y_{2j} > 0\} = \{3\} \quad \text{e} \quad L_\beta = \phi.$$

Como 3 é o único elemento de L_α e L_β é vazio, a_3 entrará na base substituindo a_1 .

Assim sendo $I_\alpha = \{4\}$, $I_\beta = \{1\}$, Lembrar que $\alpha(1) = \alpha(2) = \alpha(3) = \alpha(4) = 0$, $\beta(1) = 1$, $\beta(2) = \beta(3) = \beta(4) = +\infty$.

A nova base será

$$B = \begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix}.$$

Sabemos que

$$x_B = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = B^{-1}b - y_1x_1 - y_4x_4.$$

Neste caso $x_1 = \beta(1) = 1$ e $x_4 = \alpha(4) = 0$, logo

$$\hat{x}_B = \begin{pmatrix} \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = B^{-1}b - y_1 = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} - \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix};$$

$$\hat{x}_B = \begin{pmatrix} \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ 9 \end{pmatrix} - \begin{pmatrix} \frac{3}{2} \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

A solução ótima obtida para (P_1) é $x_1 = 1$, $x_2 = 2$, $x_3 = 4$, $x_4 = 0$, fornecendo $val(P_1) = -14$. O ótimo de (P_1) é uma solução viável de (P) , logo o problema (P_1) não terá descendentes, isto é, o nó associado a este problema será fechado. Por outro lado o valor atual da função objetivo de (P) é $\hat{x}_0 = -\infty$, mas $val(P_1) > -\infty$, atualiza-se $\hat{x}_0 = val(P_1) = -14$.

A resolução de (P_2) será também feita a partir da solução ótima de (\bar{P}) , como já visto acima. Sabe-se que $I_\alpha = \{3, 4\}$, $I_\beta = \phi$, $L_\alpha = \{j \in I_\alpha \mid y_{2j} < 0\} = \{4\}$ e $L_\beta = \phi$. Só existindo um elemento em $L_\alpha = \{4\}$, então a_1 sairá da base e a_4 entrará. Para representar a nova solução básica tem-se: $I_\alpha = \{1, 3\}$, $\alpha(1) = 2$, $\alpha(2) = \alpha(3) = \alpha(4) = 0$, $\beta(1) = \beta(2) = \beta(3) = \beta(4) = +\infty$ e $I_\beta = \phi$.

A nova base será

$$B = (a_2 \ a_4) = \begin{pmatrix} 4 & -0 \\ 2 & -1 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix}.$$

Sabemos que

$$x_B = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = B^{-1}b - y_1x_1 - y_3x_3.$$

Neste caso $x_1 = \alpha(1) = 2$ e $x_3 = \alpha(3) = 0$, logo

$$\hat{x}_B = \begin{pmatrix} \hat{x}_2 \\ \hat{x}_4 \end{pmatrix} = B^{-1}b - 2y_1 = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} - 2 \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{2}{4} & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix};$$

$$\hat{x}_B = \begin{pmatrix} \hat{x}_2 \\ \hat{x}_4 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} \\ -\frac{18}{4} \end{pmatrix} - 2 \begin{pmatrix} \frac{1}{4} \\ -\frac{10}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{2}{4} \end{pmatrix} \geq 0.$$

A solução ótima obtida para (P_2) é $x_1 = 2$, $x_2 = \frac{3}{4} = 0,75$, $x_3 = 0$, $x_4 = \frac{1}{2} = 0,5$, fornecendo $val(P_2) = -\frac{47}{4} = -11,75$. Como $val(P_2) = -11,75 > \hat{x}_0 = -14$, o nó associado ao problema (P_2) não poderá ser fechado. Consideraremos seus descendentes (P_3) e (P_4) que podem ser:

- $F(P_3) = F(P_2) \cap \{x \in \mathcal{R}^4 \mid x_2 \leq 0\}$ e
- $F(P_4) = F(P_2) \cap \{x \in \mathcal{R}^4 \mid x_2 \geq 1\}$.

Para solucionar (P_3) utilizaremos o método dual do simplex com variáveis canalizadas, algo já feito acima. A solução ótima de (P_2) é dual viável de (P_3) . A coluna a_2 sairá da base. Sabe-se que $I_\alpha = \{1, 3\}$, $I_\beta = \phi$, $\alpha(1) = 2$, $\alpha(2) = \alpha(3) = \alpha(4) = 0$, $\beta(2) = 0$, $\beta(1) = \beta(3) = \beta(4) = +\infty$. É preciso lembrar que a base associada á solução ótima de (P_2) é

$$B = (a_2 \ a_4) = \begin{pmatrix} 4 & -0 \\ 2 & -1 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} \frac{1}{4} & 0 \\ -\frac{2}{4} & -1 \end{pmatrix}.$$

Determina-se:

$$y_1 = \begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix} = B^{-1}a_1 = \begin{pmatrix} \frac{1}{4} \\ -\frac{10}{4} \end{pmatrix}, \quad y_3 = \begin{pmatrix} y_{13} \\ y_{23} \end{pmatrix} = B^{-1}a_3 = \begin{pmatrix} -\frac{1}{4} \\ -\frac{2}{4} \end{pmatrix}.$$

Ter-se-á $L_\alpha = \{j \in I_\alpha \mid y_{1j} > 0\} = \{1\}$ e $L_\beta = \phi$. Assim sendo a_1 entrará na base no lugar de a_2 .

A nova base será

$$B = (a_1 \ a_4) = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix},$$

cuja inversa

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix}.$$

Sabemos que

$$x_B = \begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = B^{-1}b - y_2x_2 - y_3x_3.$$

Neste caso $x_2 = \beta(2) = 0$ e $x_3 = \alpha(3) = 0$, logo

$$\begin{aligned} \hat{x}_B &= \begin{pmatrix} \hat{x}_1 \\ \hat{x}_4 \end{pmatrix} = B^{-1}b = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix}; \\ \hat{x}_B &= \begin{pmatrix} \hat{x}_1 \\ \hat{x}_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}. \end{aligned}$$

Para (P_3) obteve-se a solução $x_1 = 5$, $x_2 = 0$, $x_3 = 0$, $x_4 = 8$, fornecendo $val(P_3) = -20$. Essa é uma solução viável de (P) , mas pior que a solução viável já obtida no nó associado a (P_1) , tal que $val(P_1) = -14$. Assim sendo o nó associado a (P_3) será fechado.

Como foi feito para (P_3) , a solução de (P_4) será encontrada a partir de (P_2) . Verifica-se que $L_\alpha = \{j \in I_\alpha \mid y_{1j} < 0\} = \{3\}$ e $L_\beta = \phi$. Logo a_3 entrará na base no lugar de a_2 . Sabe-se que $\alpha(1) = 2$, $\alpha(2) = 1$, $\alpha(3) = \alpha(4) = 0$, $\beta(1) = \beta(2) = \beta(3) = \beta(4) = +\infty$. A nova base será

$$B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = B^{-1}.$$

Sabemos que

$$x_B = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = B^{-1}b - y_1x_1 - y_2x_2.$$

Neste caso $x_1 = \alpha(1) = 2$ e $x_2 = \alpha(2) = 1$, logo

$$\begin{aligned}
\hat{x}_B &= \begin{pmatrix} \hat{x}_3 \\ \hat{x}_4 \end{pmatrix} \\
&= B^{-1}b - 2y_1 - y_4 \\
&= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} - 2 \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\
&\quad - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix}; \\
\hat{x}_B &= \begin{pmatrix} \hat{x}_3 \\ \hat{x}_4 \end{pmatrix} = \begin{pmatrix} -5 \\ -7 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ -3 \end{pmatrix} - \begin{pmatrix} -4 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.
\end{aligned}$$

Obtém-se assim uma solução ótima de (P_4) , $x_1 = 2$, $x_2 = 1$, $x_3 = x_4 = 0$, fornecendo $val(P_4) = -13$. Esta solução é também viável de (P) , logo o nó associado a (P_4) será fechado, mas $val(P_4) = -13 > \hat{x}_0 = -14$, assim sendo far-se-á $\hat{x}_0 = val(P_4) = -13$. Como não há mais nós abertos a solução de (P_4) fornece o ótimo de (P) . ■

10.5 Exercícios

1. Suponhamos que estejamos em uma etapa do método de enumeração de Balas associada à solução x^p de (10.8). Temos que para uma restrição i de (10.7):

$$\sum_{j \in J_p} a_{ij} + \sum_{j \notin J_p} a_{ij}x_j \leq b_i$$

ou ainda

$$\sum_{j \notin J_p} a_{ij}x_j \leq b_i - \sum_{j \in J_p} a_{ij} = s_i^p. \quad (10.15)$$

Demonstrar que $\forall x_j \in \{0, 1\}$, $j \notin J_p$, solução de (10.15), para as quais

$$\sum_{k \notin J_p} \min\{0, a_{ik}\} + |a_{ij}| > s_i^p,$$

então $x_j = 0$ se $a_{ij} > 0$ e $x_j = 1$ se $a_{ij} < 0$. Dar um exemplo. Referência [Ge 69].

2. Demonstrar que o algoritmo de Balas aqui apresentado, utilizando a estrutura de pilha converge. Referências: artigos [Gl 65], [Ge 67] e o livro [Ta 75], às páginas 85-138.

3. Resolver pelo algoritmo de Balas o seguinte problema.

$$(P) : \text{minimizar } x_0 = 4x_1 + 6x_2 + 2x_3 + 6x_4 + 7x_5 - 7x_6$$

sujeito a:

$$\begin{array}{rrrrrrrrrr} 6x_1 & - & 4x_2 & & & - & 6x_4 & & & + & 2x_6 & \geq & 0 \\ 3x_1 & + & x_2 & + & 7x_3 & + & 7x_4 & + & 8x_5 & - & 8x_6 & \geq & 17 \\ -9x_1 & + & 4x_2 & + & 5x_3 & + & 2x_4 & + & 8x_5 & - & 8x_6 & \geq & 15 \end{array}$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3, 4, 5, 6.$$

4. Escrever e executar um programa para computador do método de Balas.
5. Introduzindo variáveis bivalentes $(0-1)$, fornecer uma formulação para resolver o seguinte problema de programação matemática:

$$(P) : \text{minimizar } \sum_{j=1}^3 f_j(x_j)$$

sujeito a:

$$\begin{array}{rrrrrr} 3x_1 & + & 2x_2 & + & 2x_3 & \leq & 18 \\ x_1 & - & x_2 & + & x_3 & \leq & 8 \\ x_1 & + & x_2 & + & x_3 & \geq & 0,5 \\ x_j & \geq & 0, & j = 1, 2, 3, \end{array}$$

onde

$$f_j(x_j)_{j=1,2,3} = \begin{cases} x_j & \text{se } 0 \leq x_j \leq 1 \\ 0,5x_j + 0,5 & \text{se } 1 \leq x_j \leq 2 \\ 1,5 & \text{se } x_j \geq 2. \end{cases}$$

(Obs.: deseja-se o mínimo de uma função côncava restrita a um conjunto convexo)

6. O custo de fabricação de x_i unidades de um determinado produto em uma localidade i é $d_i + c_i x_i$ ($d_i > 0$ e $c_i > 0$) se $x_i > 0$ e zero se $x_i = 0$. O custo para transportar uma unidade deste produto da localidade i para a localidade j é $a_{ij} > 0$. A demanda do produto em j é $b_j \geq 0$ e a produção máxima em i é de g_i unidades. Supondo $i \in I$ e $j \in J$ (I e J conjuntos finitos conhecidos), formular um problema de programação linear mista (com variáveis contínuas e inteiras) que determine a distribuição do produto, satisfazendo as demandas e minimizando o custo total de fabricação e de transporte (supor que $\sum_{i \in I} g_i > \sum_{j \in J} b_j$ e que haja sempre uma ligação de $i \in I$ para $j \in J$).

7. Resolver pelo método de Balas:
 minimizar $x_0 = 3x_1 - 2x_2 + x_3 + 2x_4$,
 sujeito a, $x_1 + x_2 + 2x_3 + 2x_4 \geq 5$, $x_j \in \{0, 1\}$, $j = 1, 2, 3, 4$.
8. Transformar o problema abaixo de tal maneira que as variáveis x_1 e x_2 sejam substituídas por variáveis 0 – 1. A transformação é única? Justifique. Fornecer aquela com o menor número de variáveis e menor número de restrições.
 Maximizar $x_0 = 4x_1 + 5x_2 + 9x_3 + 5x_4$, sujeito a:
 $x_1 + 3x_2 + 9x_3 + 6x_4 \leq 16$, $6x_1 + 6x_2 + 2x_3 + 7x_4 \leq 19$,
 $x_j \geq 0$, $j = 1, 2, 3, 4$, x_1 e x_2 inteiros.
9. Transformar o problema de programação não-linear 0 – 1 em um problema de programação linear mista:
 Minimizar $x_0 = 3x_1 + 2x_2 + x_3 + 5x_1x_2 - x_2x_3 + 7x_1x_2x_3$, sujeito a:
 $x_j \in \{0, 1\}$, $j = 1, 2, 3$.
 Aplicar um método de *branch-and-bound* para resolver este problema de programação linear mista.
10. Resolver utilizando o método de cortes de Gomory o seguinte problema:
 minimizar $x_0 = 4x_1 + 5x_2$, sujeito a:
 $3x_1 + x_2 \geq 2$, $x_1 + 4x_2 \geq 5$, $3x_1 + 2x_2 \geq 7$,
 $x_1 \geq 0, x_2 \geq 0, x_1$ inteiro e x_2 inteiro.
 Ilustrar no plano $x_1 \times x_2$ os cortes gerados em cada iteração.
11. Utilizar um método de *branch-and-bound* para resolver também o problema da 7ª questão. Esquematizar a arborescência gerada para esta resolução.
12. Se $3x_1 + 6x_2 + 18x_3 \geq 7$ representar um corte de Gomory para um problema de programação linear em que x_1, x_2 e x_3 só podem tomar valores inteiros, poderemos melhorá-lo? Justificar.
13. Sem resolver o problema
 maximizar $x_0 = 4x_1 + 5x_2 - 3x_3$, sujeito a:
 $3x_1 - x_2 + 7x_3 \leq 10$, $3x_1 + 5x_2 + 5x_3 = 15$, $4x_1 + 6x_2 + 10x_3 = 11$,
 $x_j \geq 0$ e inteiro, para $j = 1, 2, 3$;
 dizer se há uma solução viável. Justificar.
14. Procurar algumas formulações matemáticas para o *problema do caixeiro viajante* simétrico (sobre um grafo não orientado) e assimétrico (sobre um grafo orientado).

Capítulo 11

Dualidade em Programação Inteira

Acreditamos que o trabalho de Lorie e Savage [LoSa 55] seja a primeira publicação da noção de dualidade em programação inteira, onde os métodos de relaxação lagrangeana são utilizados visando à solução dos problemas de otimização combinatória. Um outro trabalho pioneiro foi publicado por Everett [Eve 63]. No entanto, os trabalhos de Held e Karp ([HeKa 70], [HeKa 71]) que, realmente, desenvolveram essa área do conhecimento científico. Geoffrion [Ge 74] cunhou o termo "relaxação lagrangeana". Outros trabalhos importantes dos anos 70 não podem ser esquecidos: [Fi73], [Fi 81], [FiSh 74], [FiNoSh 75], [Sh 71], [Sh 79].

11.1 Relaxação Lagrangeana

Nossa exposição será desenvolvida para os problemas de programação linear com todas as variáveis 0-1, mas podemos estender, sem muitas dificuldades, os resultados apresentados para os problemas de programação linear inteira mais gerais. Ver, por exemplo, [Pl 2000].

Seja o problema de programação inteira:

$$(P) : \text{ minimizar } x_0 = cx \quad (11.1)$$

sujeito a:

$$Ax \leq b \quad (11.2)$$

$$Dx \leq d \quad (11.3)$$

$$x \in \{0, 1\}^n, \quad (11.4)$$

onde $c^T \in \mathcal{R}^n$, $x \in \mathcal{R}^n$, A uma matriz $m \times n$, $b \in \mathcal{R}^m$, D uma matriz $p \times n$, $d \in \mathcal{R}^p$.

Associemos às restrições (11.2) o vetor linha $u = (u_1 \ u_2 \ \dots \ u_m)$, tal que $u \geq 0$, conhecido como vetor dos multiplicadores de Lagrange associados às restrições em questão. Então para $u \geq 0$ escrevemos o seguinte problema:

$$\mathcal{L}(u) = \min \{ cx + u(Ax - b) \} \quad (11.5)$$

sujeito a:

$$Dx \leq d \quad (11.6)$$

$$x \in \{0, 1\}^n. \quad (11.7)$$

O problema (11.5)-(11.7) será denominado uma relaxação lagrangeana de (P) .

Suporemos que as restrições (11.2), (11.3) e (11.4) formam um conjunto não vazio.

Proposição 11.1 $\mathcal{L}(u) \leq \text{val}(P)$.

Demonstração

Suponhamos que x^* seja uma solução ótima de (P) , isto é, $\text{val}(P) = cx^*$. Sabemos também que $\mathcal{L}(u) \leq cx^* + u(Ax^* - b)$, como $u \geq 0$ e $Ax^* - b \leq 0$, então poderemos escrever

$$\mathcal{L}(u) \leq cx^* + u(Ax^* - b) \leq cx^* = \text{val}(P). \quad \blacksquare$$

Esta propriedade nos diz que $\mathcal{L}(u)$ é uma cota inferior de $\text{val}(P)$. Visando à busca da maior cota possível, consideramos o seguinte problema.

$$(D) : \begin{array}{l} \text{maximizar } \mathcal{L}(u) \\ \text{sujeito a : } u \geq 0. \end{array}$$

(D) é considerado um problema dual de (P) . Se u^* for uma solução ótima de (D) , teremos $\text{val}(D) = \mathcal{L}(u^*)$.

Geralmente $\text{val}(D) < \text{val}(P)$ para os problemas de programação inteira, dizemos que existe um salto primal-dual. Em lugar de considerarmos o vetor dos multiplicadores de Lagrange u , poderemos utilizar um espaço de funções de Lagrange generalizadas como apresentado em [TiWo 78], [Wo 81] e [NeWo 88]; esse enfoque elimina o salto primal-dual, no entanto, a solução do novo problema dual não é prática em termos computacionais.

11.2 Solução do Problema Dual

Construiremos a seguir dois métodos para resolver (D) : uma técnica de geração de colunas do tipo Dantzig e Wolfe [DanWo 60] e um método utilizando sub-gradientes sugerido por Held, Wolfe e Crowder [HeWoCr 74].

11.2.1 Técnica de geração de colunas

O problema (D) pode ser escrito sob a forma de um problema de programação linear:

$$(D) : \text{maximizar } \mathcal{L}(u) = w$$

sujeito a:

$$\begin{aligned} w &\leq cx^k + u(Ax^k - b), \quad k = 1, 2, \dots, K, \\ u &\geq 0, \end{aligned}$$

onde $\{x^1, x^2, \dots, x^K\} = \{x \in \{0, 1\}^n \mid Dx \leq d\}$. Devemos notar que w e u são as incógnitas de (D) ; os vetores x^k de componentes 0-1, soluções viáveis de (11.3) e (11.4), são considerados, inicialmente, conhecidos. Poderíamos pensar em resolver o problema (D) acima por técnicas de geração de linhas, ver [Las 70], mas optamos por trabalhar com seu problema dual de programação linear. O problema (D) pode ser escrito sob a forma de um problema de programação linear:

$$(D) : \text{maximizar } \mathcal{L}(u) = w$$

sujeito a:

$$\begin{aligned} w - u(Ax^k - b) &\leq cx^k, \quad k = 1, 2, \dots, K, \\ u &\geq 0. \end{aligned}$$

Tomando o problema dual de (D) teremos:

$$(DD) : \text{minimizar } \sum_{k=1}^K (cx^k) \lambda_k$$

sujeito a:

$$\begin{aligned} \sum_{k=1}^K \lambda_k &= 1 \\ - \sum_{k=1}^K (Ax^k - b) \lambda_k &\geq 0 \\ \lambda_k &\geq 0, \quad k = 1, 2, \dots, K, \end{aligned}$$

onde λ_k , $k = 1, 2, \dots, K$ representam as variáveis duais. Podemos ainda escrever

$$(DD) : \text{minimizar } \sum_{k=1}^K (cx^k) \lambda_k \quad (11.8)$$

sujeito a:

$$\sum_{k=1}^K \lambda_k = 1 \quad (11.9)$$

$$\sum_{k=1}^K (Ax^k) \lambda_k \leq b \quad (11.10)$$

$$\lambda_k \geq 0, \quad k = 1, 2, \dots, K. \quad (11.11)$$

Para resolvermos (11.8)-(11.11), introduziremos um vetor

$$s = (s_1 \ s_2 \ \dots \ s_m) \geq 0,$$

cujas componentes são as variáveis de folga das restrições (11.10), assim sendo:

$$(DD) : \text{minimizar } \sum_{k=1}^K (cx^k) \lambda_k \quad (11.12)$$

sujeito a:

$$\sum_{k=1}^K \lambda_k = 1 \quad (11.13)$$

$$\sum_{k=1}^K (Ax^k) \lambda_k + s = b \quad (11.14)$$

$$\lambda_k \geq 0, \quad k = 1, 2, \dots, K \quad (11.15)$$

$$s \geq 0. \quad (11.16)$$

A matriz que define as restrições de igualdades (11.13) e (11.14) será esquematizada abaixo:

$$\hat{A} = \begin{pmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ Ax^1 & Ax^2 & \dots & Ax^K & e_1 & e_2 & \dots & e_m \end{pmatrix}$$

Os vetores e_i , $i = 1, 2, \dots, m$ representam os vetores unitários, isto é, a única componente diferente de zero é a i -ésima, sendo esta igual a um.

Seja B uma matriz $(m+1) \times (m+1)$ formada por $m+1$ colunas de \hat{A} , tal que B seja inversível. Suponhamos também que a solução básica associada a

B seja primal viável de (11.12)-(11.16). A otimalidade desta solução básica será verificada, como já foi estudado anteriormente, através de sua viabilidade dual.

Os coeficientes associados às variáveis λ_k na função objetivo são iguais a cx^k , ao passo que os associados às variáveis de folga s_i são iguais a zero. Denominaremos de q_B o vetor linha, cujas componentes são os coeficientes das variáveis básicas na função objetivo. O vetor linha v representando os multiplicadores do simplex será então escrito como $v = q_B B^{-1}$. Podemos particionar v da seguinte forma: $v = (v_0 \ v^1)$, onde $v_0 \in \mathcal{R}$ e $(v^1)^T \in \mathcal{R}^m$, isto é, v_0 está associado à restrição (11.13) e as componentes de v^1 às restrições (11.14).

Caso uma das componentes de v^1 , por exemplo, v_i^1 seja positiva então a coluna $\begin{pmatrix} 0 \\ e_i \end{pmatrix}$ deve entrar na base, pois $v \begin{pmatrix} 0 \\ e_i \end{pmatrix} = v_i^1$. Como supusemos que $v_i^1 > 0$, logo se a variável s_i entrar na base o valor da função objetivo (11.12) diminuirá se s_i tomar um valor positivo na nova solução básica.

Suporemos agora que $v^1 \leq 0$, e passaremos a determinar:

$$\max_{k=1,2,\dots,K} \left\{ (v_0 \ v^1) \begin{pmatrix} 0 \\ Ax^k \end{pmatrix} - cx^k \right\}.$$

A expressão acima pode ainda ser escrita:

$$t(x^l) = \max_{k=1,2,\dots,K} \{ (v^1 A - c)x^k + v_0 \}, \quad (11.17)$$

x^l é uma solução deste problema. Caso $t(x^l) > 0$ a coluna $\begin{pmatrix} 0 \\ Ax^l \end{pmatrix}$ entrará na base. Caso contrário, isto é, $t(x^l) \leq 0$ a base corrente B é uma solução ótima de (DD).

Como não conhecemos todos os pontos x^k , $k = 1, 2, \dots, K$, teremos que resolver o problema de programação matemática:

$$(PA) : \text{maximizar } (v^1 A - c)x + v_0 \quad (11.18)$$

sujeito a:

$$Dx \leq d \quad (11.19)$$

$$x \in \{0, 1\}^n. \quad (11.20)$$

Seja $val(PA) = (v^1 A - c)x^l + v_0$. Como já foi mencionado acima, se $val(PA) > 0$ a variável λ_l entrará na nova base. Caso contrário a atual base B é uma solução ótima.

As regras de pivoteamento para o algoritmo primal do simplex já foram vistas nos primeiros capítulos. Mais adiante apresentaremos um exemplo completo para a solução de (DD) . Lembremos que $val(D) = val(DD)$. Além disso não devemos esquecer que v^1 associado ao ótimo de (DD) é uma solução ótima de (D) , isto é, $u^* = v^1$. verificar que v_0 associado ao ótimo de (DD) é igual a $val(D)$.

Se acrescentarmos ao problema (11.8)-(11.11) as restrições de integralidade para as variáveis λ_k , $k = 1, 2, \dots, K$, isto é, $\lambda_k \in \mathcal{Z}$, $k = 1, 2, \dots, K$, teremos um problema equivalente a (P) . Por que?

Seja a relaxação linear de (PA) :

$$(\overline{PA}) : \text{maximizar } (v^1 A - c)x + v_0 \quad (11.21)$$

sujeito a:

$$Dx \leq d \quad (11.22)$$

$$x \in [0, 1]^n. \quad (11.23)$$

Por outro lado sabemos também que a relaxação linear de (P) será

$$(\bar{P}) : \text{minimizar } x_0 = cx \quad (11.24)$$

sujeito a:

$$Ax \leq b \quad (11.25)$$

$$Dx \leq d \quad (11.26)$$

$$x \in [0, 1]^n. \quad (11.27)$$

Proposição 11.2 $val(\bar{P}) \leq val(D) \leq val(P)$.

Demonstração

Da propriedade 11.1 temos que $val(D) \leq val(P)$. Para verificar que $val(\bar{P}) \leq val(D)$ basta notar que

$$\{x \in \mathcal{R}^n \mid Dx \leq d, x \in \{0, 1\}^n\} \subseteq \{x \in \mathcal{R}^n \mid Dx \leq d, x \in [0, 1]^n\}. \quad \blacksquare$$

Observação: Se todos os vértices do politopo formado pelas restrições de (\overline{PA}) pertencerem a $\{0, 1\}^n$, então $val(\overline{PA}) = val(PA)$. Neste caso sabermos, a priori, que $val(\bar{P}) = val(D)$. Esta particularidade é conhecida como a **propriedade de integralidade** de Geoffrion [Ge 74].

Quando a relaxação lagrangeana não possui a propriedade de integralidade, a solução computacional de (PA) pode ser muito difícil, pois temos

que resolver um problema de programação inteira 0-1, com menos restrições que o problema inicial (P) .

É interessante a utilização da relaxação lagrangiana com a propriedade de integralidade, apesar de a cota inferior obtida valer $val(\bar{P})$. Na realidade, durante a solução de (DD) geramos pontos x^k que podem ser também viáveis de (P) . Dessa maneira temos também uma cota superior para $val(P)$. Isto é, $val(P) \leq cx^k$.

11.2.2 Método utilizando sub-gradientes

Definiremos sub-gradiente e forneceremos um método específico para maximizar $\mathcal{L}(u)$, sujeito a $u \geq 0$.

Proposição 11.3 $\mathcal{L}(u)$ é uma função côncava e afim por partes.

Demonstração

Ver [GoMi 79], página 497. ■

Antes de continuarmos, apresentaremos um exemplo. Seja

$$(P) : \text{minimizar } x_0 = -6x_1 - 8x_2 + 5x_3$$

sujeito a:

$$-4x_1 - 4x_2 - x_3 \leq -2$$

$$2x_2 - 2x_3 \leq 1$$

$$-2x_2 + 2x_3 \leq 1$$

$$2x_1 + 2x_2 + 2x_3 \leq 5$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Associamos a variável $u \geq 0$ à restrição $-4x_1 - 4x_2 - x_3 \leq -2$ e obteremos a seguinte expressão:

$$l(x^k, u) = -6x_1^k - 8x_2^k + 5x_3^k + (-4x_1^k - 4x_2^k - x_3^k + 2)u,$$

onde $x^k = (x_1^k \ x_2^k \ x_3^k)$, $k = 1, 2, \dots, K$ são as soluções viáveis de

$$2x_2 - 2x_3 \leq 1$$

$$-2x_2 + 2x_3 \leq 1$$

$$2x_1 + 2x_2 + 2x_3 \leq 5$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

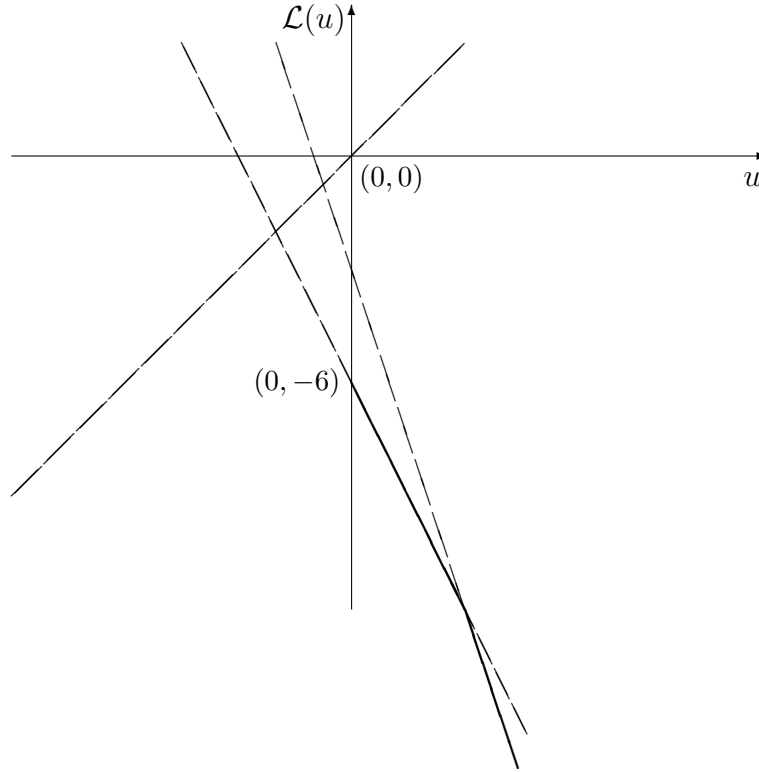


Figura 11.1: A função dual

Neste caso muito simples verificamos que apenas três pontos, $x^1 = (0 \ 0 \ 0)$, $x^2 = (1 \ 0 \ 0)$ e $x^3 = (0 \ 1 \ 1)$, satisfazem

$$2x_2 - 2x_3 \leq 1$$

$$-2x_2 + 2x_3 \leq 1$$

$$2x_1 + 2x_2 + 2x_3 \leq 5$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

A função objetivo do problema dual dessa relaxação lagrangeana poderá ser escrita:

$$\mathcal{L}(u) = \min_{k=1,2,3} \{ l(x^k, u) \},$$

onde $l(x^1, u) = 2u$, $l(x^2, u) = -6 - 2u$ e $l(x^3, u) = -3 - 3u$.

Na figura 11.1 representamos $\mathcal{L}(u)$, para $u \in \mathcal{R}_+$.

Sabemos que o dual associado à relaxação em questão será

$$(D) : \text{maximizar } \mathcal{L}(u)$$

sujeito a:

$$u \geq 0.$$

Podendo ser escrito sob a forma de programação linear:

$$(D) : \text{maximizar } \mathcal{L}(u) = w$$

sujeito a:

$$\begin{aligned} w &\leq 2u \\ w &\leq -6 - 2u \\ w &\leq -3 - 3u \\ w &\geq 0. \end{aligned}$$

Resolvendo (D) acima pelo método do simplex obteremos $u = 0$ e $w = -6$, isto é, $\text{val}(D) = -6$. Algo que pode ser verificado na figura 13.1. Como o ponto $x^2 = (1 \ 0 \ 0)$ é viável de (P) e fornece a x_0 um valor -6 , concluímos pela propriedade 13.1 que $x^2 = (1 \ 0 \ 0)$ é um ótimo de (P) .

É interessante verificar que $\text{val}(\bar{P}) = -11, 5$, cota bem pior do que $\text{val}(D)$ para este exemplo. A propriedade da integralidade não é válida.

Definição 11.1 Um vetor $\gamma \in \mathcal{R}^n$ é um sub-gradiente de uma função côncava $f : \mathcal{R}^n \rightarrow \mathcal{R}$, no ponto x^0 se $f(x) \leq f(x^0) + \gamma^T(x - x^0)$ para todo $x \in \mathcal{R}^n$.

Se a função f for diferenciável em x^0 então teremos apenas um sub-gradiente que será igual ao gradiente de f em x^0 .

Na figura 11.2, consideramos o gráfico de uma função f côncava não diferenciável em x^1 , mas diferenciável em x^2 . Os hiperplanos $h_j(x^i) = f(x^i) + \gamma_j^T(x - x^i)$, para $j = 1, 2, 3$, são também ilustrados na figura 11.2

Definição 11.2 O subdiferencial $\partial f(x^0)$ de f em x^0 é o conjunto de todos os sub-gradientes de f em x^0 .

Se f é diferenciável em x^0 então $\partial f(x^0) = \{\nabla f(x^0)\}$, onde $\nabla f(x^0)$ representa o gradiente de f em x^0 .

Podemos notar também que $\partial \mathcal{L}(u) \neq \emptyset$.

Proposição 11.4 $\partial f(x^0) \neq \emptyset$ é convexo e fechado.

Demonstração

Para demonstrar a convexidade de $\partial f(x^0)$, tomaremos γ_1 e $\gamma_2 \in \partial f(x^0)$, então

$$f(x) \leq f(x^0) + \gamma_1^T(x - x^0) \quad (11.28)$$

e

$$f(x) \leq f(x^0) + \gamma_2^T(x - x^0). \quad (11.29)$$

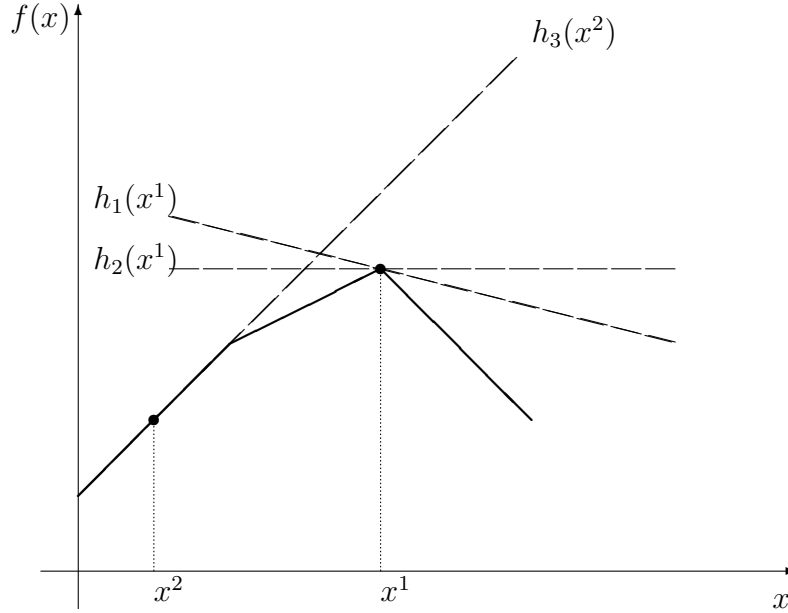


Figura 11.2: Idéia de sub-gradientes

Consideremos $0 \leq \alpha \leq 1$, multiplicamos (11.28) por α e (11.29) por $1 - \alpha$, após essas operações somemos membro a membro as duas desigualdades:

$$\alpha f(x) + (1 - \alpha)f(x) \leq \alpha f(x^0) + (1 - \alpha)f(x^0) + [\alpha\gamma_1^T + (1 - \alpha)\gamma_2^T](x - x^0)$$

ou ainda

$$f(x) \leq f(x^0) + [\alpha\gamma_1^T + (1 - \alpha)\gamma_2^T](x - x^0),$$

logo $\gamma = \alpha\gamma_1 + (1 - \alpha)\gamma_2$, $0 \leq \alpha \leq 1$, é também um sub-gradiente de f em x^0 .

Para mostrar que $\partial f(x^0)$ é um conjunto fechado, consideraremos uma sequência $\{\gamma_j\}$ de elementos de $\partial f(x^0)$ convergindo para $\bar{\gamma} \notin \partial f(x^0)$.

Logo existe $\bar{x} \in \mathcal{R}^n$, tal que: $f(\bar{x}) > f(x^0) + \bar{\gamma}^T(\bar{x} - x^0)$, ou ainda

$$f(\bar{x}) - f(x^0) - \bar{\gamma}^T(\bar{x} - x^0) \geq \epsilon > 0. \quad (11.30)$$

Por outro lado sabemos também que

$$-f(\bar{x}) + f(x^0) + \gamma_j^T(\bar{x} - x^0) \geq 0, \quad \forall \gamma_j \in \partial f(x^0). \quad (11.31)$$

Somando membro a membro (11.30) e (11.31) temos que

$$(\gamma_j - \bar{\gamma})^T(\bar{x} - x^0) \geq \epsilon.$$

Utilizando a desigualdade de Schwarz teremos

$$\begin{aligned}\epsilon &\leq (\gamma_j - \bar{\gamma})^T (\bar{x} - x^0) \leq \|\gamma_j - \bar{\gamma}\| \|\bar{x} - x^0\|, \\ \|\gamma_j - \bar{\gamma}\| &\geq \frac{\epsilon}{\|\bar{x} - x^0\|}, \forall \gamma_j \in \partial f(x^0);\end{aligned}$$

uma contradição com a convergência da sequência $\{\gamma_j\}$ de elementos de $\partial f(x^0)$ para $\bar{\gamma} \notin \partial f(x^0)$. ■

Proposição 11.5 *Se $f : \mathcal{R}^n \rightarrow \mathcal{R}$ é uma função côncava, então x^0 maximiza f se e somente se $0 \in \partial f(x^0)$.*

A demonstração é deixada para o(a) leitor(a).

Retornemos à função \mathcal{L} definida em (11.5), (11.6) e (11.7), suporemos que \bar{x} seja uma solução viável de (11.6) e (11.7), tal que $\mathcal{L}(u^0) = c\bar{x} + u^0(A\bar{x} - b)$. Assim sendo consideraremos a seguinte propriedade.

Proposição 11.6 *$A\bar{x} - b$ é um sub-gradiente de \mathcal{L} em u^0 .*

Demonstração

Podemos escrever

$$\begin{aligned}\mathcal{L}(u^0) + (u - u^0)(A\bar{x} - b) &= c\bar{x} + u^0(A\bar{x} - b) + (u - u^0)(A\bar{x} - b) \\ &= c\bar{x} + u(A\bar{x} - b) \geq \mathcal{L}(u), \quad u \geq 0.\end{aligned}$$

■

A propriedade 11.6 nos fornece um maneira de obter um sub-gradiente de \mathcal{L} em um dado ponto u^0 .

Um método utilizando sub-gradientes descrito em [HeWoCr 74] para resolver o problema dual

$$(D) : \text{maximizar } \mathcal{L}(u) \tag{11.32}$$

sujeito a:

$$u \geq 0, \tag{11.33}$$

será apresentado a seguir.

Partiremos de um ponto $u^0 \geq 0$ arbitrário, definiremos uma sequência de pontos:

$$u^{k+1} = u^k + \theta_k (\gamma^k)^T, \quad \text{onde } \gamma^k \in \partial \mathcal{L}(u^k).$$

É importante notar que γ^k pode não ser uma direção de subida para \mathcal{L} em u^k , ver [Mino 86].

Uma escolha do passo θ_k , para $\gamma^k \neq 0$ foi estudada em [HeWoCr 74].

$$\theta_k = \rho_k \frac{\bar{\mathcal{L}} - \underline{\mathcal{L}}}{\|\gamma^k\|^2}, \text{ para } 0 < \rho_k \leq 2, \quad (11.34)$$

onde $\bar{\mathcal{L}}$ é uma cota superior de $val(D)$ e $\underline{\mathcal{L}}$ uma cota inferior de $val(D)$. Essas cotas serão atualizadas durante as iterações do método, visando a diminuir o salto $\bar{\mathcal{L}} - \underline{\mathcal{L}}$. A escolha de ρ_k , segundo Beasley [Bea 92], deve ser igual a 2 no início do método iterativo e caso nas últimas p iterações o valor de \mathcal{L} não aumentar, dividiremos ρ_k por 2. Beasley [Bea 92] diz que em suas extensas experiências computacionais um bom valor para p é 30.

Sabemos que u^{k+1} pode ser negativo quando fazemos $u^{k+1} = u^k + \theta_k(\gamma^k)^T$. Assim sendo, o algoritmo será escrito com segue:

$$u_i^{k+1} = \max\{ 0, u_i^k + \theta_k \gamma_i^k \},$$

isto é, quando uma componente $u_i^k + \theta_k \gamma_i^k$ for negativa, faremos $u_i^{k+1} = 0$.

Exemplo 11.1 Resolução de (D) por geração de colunas.

Seja

$$(P) : \text{minimizar } x_0 = -5x_1 - 6x_2 - 3x_3$$

sujeito a

$$3x_1 + 4x_2 + 2x_3 \leq 4 \quad (11.35)$$

$$x_1 + x_2 + x_3 \leq 2 \quad (11.36)$$

$$x_1 + x_3 \leq 1 \quad (11.37)$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3. \quad (11.38)$$

Denominemos de X o conjunto formado pelas restrições (11.36), (11.37) e (11.38). É fácil verificar que

$$X = \{ \begin{array}{l} x^1 = (0 \ 0 \ 0)^T, \ x^2 = (1 \ 0 \ 0)^T, \ x^3 = (0 \ 1 \ 0)^T, \\ x^4 = (0 \ 0 \ 1)^T, \ x^5 = (1 \ 1 \ 0)^T, \ x^6 = (0 \ 1 \ 1)^T \end{array} \}.$$

Dualizaremos a restrição (11.35) à qual associaremos a variável dual u . Podemos então considerar

$$l(x, u) = -5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 2x_3 - 4)u.$$

Lembremos que

$$\begin{array}{ll} (D) : & \text{maximizar } \mathcal{L}(u) \\ \text{sujeito a:} & u \geq 0. \end{array}$$

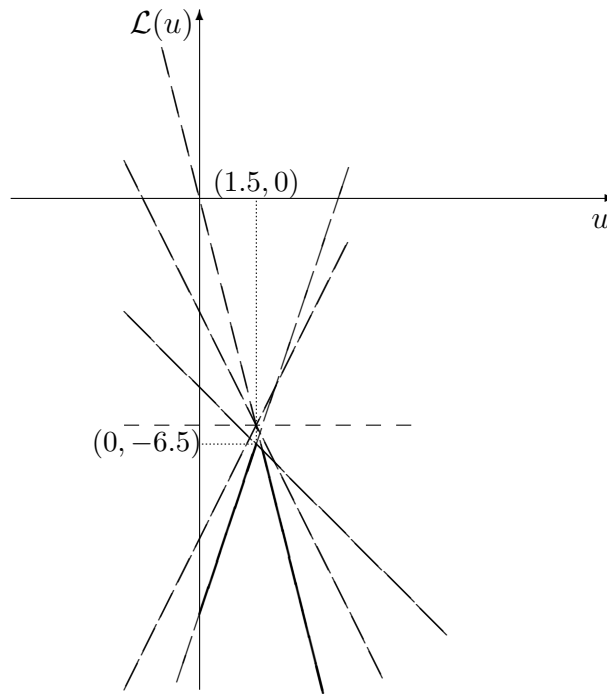


Figura 11.3: Função dual para o exemplo

Sabemos que :

$$\begin{aligned} l(x^1, u) &= -4u, \quad l(x^2, u) = -5 - u, \quad l(x^3, u) = -6, \\ l(x^4, u) &= -3 - 2u, \quad l(x^5, u) = -11 + 3u, \quad l(x^6, u) = -9 + 2u. \end{aligned}$$

$\mathcal{L}(u) = \min\{ l(x, u) \mid x \in X \}$, cujo gráfico pode ser visto na figura 11.3.

Verificamos, graficamente, na figura 11.3 que u^* , tal que $\mathcal{L}(u^*) = \text{val}(D)$ é determinado pela interseção de $l(x^2, u)$ com $l(x^5, u)$, isto é,
 $-5 - u^* = -11 + 3u^*$, logo $u^* = \frac{6}{4} = 1,5$.

Nosso intuito aqui é o de resolver (D) por geração de colunas, assim sendo escreveremos (D) sob uma outra forma:

$$(D) : \text{maximizar } \mathcal{L}(u) = t$$

sujeito a:

$$\begin{aligned} t &\leq l(x^i, u), \quad i = 1, 2, 3, 4, 5, 6 \\ u &\geq 0. \end{aligned}$$

Ou ainda:

$$(D) : \text{maximizar } \mathcal{L}(u) = t$$

sujeito a:

$$t \leq -5x_1^i - 6x_2^i - 3x_3^i + (3x_1^i + 4x_2^i + 2x_3^i - 4)u, \quad i = 1, 2, 3, 4, 5, 6$$

$$u \geq 0.$$

Escreveremos ainda:

$$(D) : \text{maximizar } \mathcal{L}(u) = t$$

sujeito a:

$$t + (-3x_1^i - 4x_2^i - 2x_3^i + 4)u \leq -5x_1^i - 6x_2^i - 3x_3^i, \quad i = 1, 2, 3, 4, 5, 6$$

$$u \geq 0.$$

Cujo dual:

$$(DD) : \text{minimizar } \sum_{i=1}^6 (-5x_1^i - 6x_2^i - 3x_3^i) \lambda_i$$

sujeito a:

$$\sum_{i=1}^6 \lambda_i = 1$$

$$\sum_{i=1}^6 (-3x_1^i - 4x_2^i - 2x_3^i + 4) \lambda_i \geq 0$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6.$$

Após algumas operações teremos:

$$(DD) : \text{minimizar } \sum_{i=1}^6 (-5x_1^i - 6x_2^i - 3x_3^i) \lambda_i$$

sujeito a:

$$\sum_{i=1}^6 \lambda_i = 1$$

$$\sum_{i=1}^6 (3x_1^i + 4x_2^i + 2x_3^i) \lambda_i \leq 4$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6.$$

Como observação importante definiremos o seguinte problema:

$$(DDI) : \text{minimizar } \sum_{i=1}^6 (-5x_1^i - 6x_2^i - 3x_3^i) \lambda_i$$

sujeito a:

$$\sum_{i=1}^6 \lambda_i = 1$$

$$\sum_{i=1}^6 (3x_1^i + 4x_2^i + 2x_3^i) \lambda_i \leq 4$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6.$$

$$\lambda_i \in \mathcal{Z}, \quad i = 1, 2, 3, 4, 5, 6.$$

É fácil verificar que (DDI) e (P) são equivalentes. Por outro lado

$$val(DD) \geq val(\bar{P}).$$

Por que?

Voltemos para a solução de (DD) . Introduziremos a variável de folga $s \geq 0$:

$$(DD) : \text{minimizar } \sum_{i=1}^6 (-5x_1^i - 6x_2^i - 3x_3^i) \lambda_i$$

sujeito a:

$$\sum_{i=1}^6 \lambda_i = 1$$

$$\sum_{i=1}^6 (3x_1^i + 4x_2^i + 2x_3^i) \lambda_i + s = 4$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6; \quad s \geq 0.$$

Uma solução básica inicial associada à solução $x^1 = (0 \ 0 \ 0) \in X$ e à variável de folga s (sendo λ_1 a variável associada a x^1) é viável. Isto é, a base B inicial seria

$$B = \begin{pmatrix} 1 & 0 \\ 3x_1^1 + 4x_2^1 + 2x_3^1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

No entanto, a título de ilustração utilizaremos o método das duas fases para encontrar uma solução inicial básica viável de (DD) .

Seja o problema auxiliar da primeira fase:

(*FASE1*) : minimizar $w = \lambda_a$

sujeito a:

$$\sum_{i=1}^6 \lambda_i + \lambda_a = 1$$

$$\sum_{i=1}^6 (3x_1^i + 4x_2^i + 2x_3^i) \lambda_i + s = 4$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6; \quad s \geq 0 \text{ e } \lambda_a \geq 0.$$

$\lambda_a = 1$ e $s = 4$ estão associadas a uma base primal viável $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ de (*FASE1*). O vetor custo associado a esta base é $(1 \ 0)$ e o vetor dos multiplicadores do simplex $v = (1 \ 0) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$.

Logo $v = (1 \ 0)$, os z_i associados às variáveis λ_i , serão escritos:

$$z_i = v \begin{pmatrix} 1 \\ 3x_1^1 + 4x_2^1 + 2x_3^1 \end{pmatrix} = (1 \ 0) \begin{pmatrix} 1 \\ 3x_1^1 + 4x_2^1 + 2x_3^1 \end{pmatrix} = 1.$$

Como o custo associado à variável λ_i é nulo, teremos, utilizando a notação deste livro, que $z_i - c_i = 1 - 0 = 1$. Utilizando as técnicas de geração de colunas apresentadas no capítulo 7 para a solução de (*FASE1*), teremos que resolver:

(*PA*) : maximizar $\sum_{i=1,2,3,4,5,6} \lambda_i$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Qualquer x^i , $i = 1, 2, 3, 4, 5, 6$, será uma solução ótima deste (*PA*). Tomemos $x^1 = (x_1^1 \ x_2^1 \ x_3^1) = (0 \ 0 \ 0)$. Sabemos que $val(PA) = 1 > 0$, logo faremos $\begin{pmatrix} 1 \\ 3x_1^1 + 4x_2^1 + 2x_3^1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \times 0 + 4 \times 0 + 2 \times 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ entrar na nova base. Para saber qual coluna sairá temos que pré-multiplicar esta coluna pela inversa da base atual. Ou seja:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Para esquematizar as operações de troca de base, por pivoteamento (eliminação de Gauss-Jordan) utilizaremos o seguinte quadro Q :

Q			1	λ_j
w	1	$c_B B^{-1}$	\bar{w}	$z_j - c_j$
λ_B	0	B^{-1}	λ_B	y_j

Para o nosso caso, teremos o quadro inicial $Q1$:

$Q1$				1	λ_1
w	1	1	0	1	1
λ_a	0	1	0	1	1^*
s	0	0	1	4	0

 \Rightarrow

$Q2$				1	λ_1
w	1	0	0	0	0
λ_1	0	1	0	1	1
s	0	0	1	4	0

onde $*$ indica o pivô. O quadro $Q2$ fornece uma base primal viável para o problema (DD) , pois $val(FASE1) = \bar{w} = 0$.

Teremos que formar um novo quadro associado ao problema (DD) , para darmos início à segunda fase do método do simplex. A base agora está associada à coluna de λ_1 e à coluna de s . Os custos associados às variáveis básicas são as componentes do vetor

$$(-5x_1^1 - 6x_2^1 - 3x_3^1 \ 0) = (-5 \times 0 - 6 \times 0 - 3 \times 0 \ 0) = (0 \ 0).$$

Sabemos ainda que

$$v = (v_0 \ v^1) = (0 \ 0) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = (0 \ 0).$$

Denominando de $f.o.$ o valor da função objetivo de (DD) a cada iteração, formaremos o quadro inicial para a segunda fase.

$Q3$				1	λ_j
$f.o.$	1	0	0	0	
λ_1	0	1	0	1	
s	0	0	1	4	

Para preencheremos a última coluna do quadro $Q3$ teremos que resolver:

$$(PA) : \text{maximizar } t = (v_0 \ v^1) \begin{pmatrix} 1 \\ 3x_1 + 4x_2 + 2x_3 \end{pmatrix} - (-5x_1 - 6x_2 - 3x_3)$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Ou ainda:

$$(PA) : \text{maximizar } t = (0 \ 0) \begin{pmatrix} 1 \\ 3x_1 + 4x_2 + 2x_3 \end{pmatrix} - (-5x_1 - 6x_2 - 3x_3)$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Finalmente

$$(PA) : \text{maximizar } t = 5x_1 + 6x_2 + 3x_3$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Cuja solução é $x_1 = 1, x_2 = 1, x_3 = 0$, fornecendo $t = 5 \times 1 + 6 \times 1 + 3 \times 0 = 11 > 0$. Esta solução é o ponto x^5 . A coluna gerada é

$$\begin{pmatrix} 1 \\ 3 \times 1 + 4 \times 1 + 2 \times 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

Pré multiplicando esta coluna pela inversa da base:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

O quadro $Q3$ ficará então

$Q3$				1	λ_5		$Q4$				1	λ_5
$f.o.$	1	0	0	0	11	\Rightarrow	$f.o.$	1	0	$-11/7$	$-44/7$	0
λ_1	0	1	0	1	1		λ_1	0	1	$-1/7$	$3/7$	0
s	0	0	1	4	7^*		λ_5	0	0	$1/7$	$4/7$	1

onde (*) indica o pivô.

Novo problema gerador de coluna:

$$(PA) : \text{maximizar } t = (0 \quad -11/7) \begin{pmatrix} 1 \\ 3x_1 + 4x_2 + 2x_3 \end{pmatrix} - (-5x_1 - 6x_2 - 3x_3)$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Ou ainda:

$$(PA) : \text{maximizar } t = \frac{2}{7}x_1 - \frac{2}{7}x_2 - \frac{1}{7}x_3$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Cuja solução é $x_1 = 1, x_2 = x_3 = 0$, fornecendo $t = \frac{2}{7} > 0$, representando o ponto x^2 . A coluna que entrará na nova base será

$$\begin{pmatrix} 1 \\ 3 \times 1 + 4 \times 0 + 2 \times 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Pré-multiplicando esta coluna pela inversa da base atual:

$$\begin{pmatrix} 1 & -1/7 \\ 0 & 1/7 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 4/7 \\ 3/7 \end{pmatrix}.$$

Assim sendo

$Q4$				1	λ_2
$f.o.$	1	0	$-11/7$	$-44/7$	$2/7$
λ_1	0	1	$-1/7$	$3/7$	$(4/7)^*$
λ_5	0	0	$1/7$	$4/7$	$3/7$

↓

$Q5$				1	λ_2
$f.o.$	1	$-1/2$	$-3/2$	$-13/2$	0
λ_2	0	$7/4$	$-1/4$	$3/4$	1
λ_5	0	$-3/4$	$1/4$	$1/4$	0

Novamente geraremos uma coluna:

$(PA) : \text{maximizar } t =$

$$(-1/2 \quad -3/2) \begin{pmatrix} 1 \\ 3x_1 + 4x_2 + 2x_3 \end{pmatrix} - (-5x_1 - 6x_2 - 3x_3)$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Isto é,

$$(PA) : \text{maximizar } t = -1/2 + \frac{1}{2}x_1 + 0x_2 + 0x_3$$

sujeito a:

$$x_1 + x_2 + x_3 \leq 2$$

$$x_1 + x_3 \leq 1$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3.$$

Podemos verificar que as soluções $x_1 = 1, x_2 = x_3 = 0$, e $x_1, x_2 = 1, x_3 = 0$ são soluções ótimas, fornecendo $t = -1/2 + 1/2 = 0$, logo $Q5$ é um quadro ótimo. Deste quadro se tem que $\lambda_2 = 3/4, \lambda_5 = 1/4$ fornecendo $val(DD) = -13/2 = -6,5 = val(D)$. Logo $val(P) \geq -6,5$, mas como os coeficientes da função objetivo de (P) são inteiros, podemos escrever que $val(P) \geq -6$. verificamos que para $x^3 = (0 \ 1 \ 0)$, $x_0 = -6$, assim sendo x^3 é uma solução ótima de (P) . ■

Exemplo 11.2 Resolução de (D) por um método utilizando sub-gradientes. Retomaremos o mesmo problema (P) do exemplo acima. Seja

$$(P) : \text{minimizar } x_0 = -5x_1 - 6x_2 - 3x_3$$

sujeito a

$$3x_1 + 4x_2 + 2x_3 \leq 4 \tag{11.39}$$

$$x_1 + x_2 + x_3 \leq 2 \tag{11.40}$$

$$x_1 + x_3 \leq 1 \tag{11.41}$$

$$x_j \in \{0, 1\}, \quad j = 1, 2, 3. \tag{11.42}$$

Associaremos à restrição (11.39) a variável $u_1 \geq 0$ e à restrição (11.40) a variável $u_2 \geq 0$. Denominemos de X o conjunto formado pelas restrições (11.41) e (11.42).

Seja

$$l(x, u) = -5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 3x_3 - 4)u_1 + (x_1 + x_2 + x_3 - 2)u_2.$$

Assim sendo, a função dual será: $\mathcal{L}(u) = \min \{l(x, u) \mid x \in X\}$.

O problema dual será mais uma vez escrito como

$$(D) : \text{maximizar } \mathcal{L}(u)$$

sujeito a:

$$u \geq 0.$$

Para resolvermos (D) acima, iniciaremos o método de otimização com a solução $u^0 = (u_1^0 \ u_2^0) = (0 \ 0)$ e utilizaremos $u_i^{k+1} = \max\{0, u_i^k + \theta_k \gamma_i^k\}$, $i = 1, 2$. Lembremos que $\gamma_k \in \partial \mathcal{L}(u^k)$ e $\theta_k = \rho_k \frac{\bar{\mathcal{L}} - \underline{\mathcal{L}}}{\|\gamma^k\|^2}$, para $0 < \rho_k \leq 2$, como definido em (11.34).

Se $\mathcal{L}(u^k) = -5x_1^k - 6x_2^k - 3x_3^k + (3x_1^k + 4x_2^k + 3x_3^k - 4)u_1^k + (x_1^k + x_2^k + x_3^k - 2)u_2^k$, então

$$\gamma_k = \begin{pmatrix} \gamma_1^k \\ \gamma_2^k \end{pmatrix} = \begin{pmatrix} 3x_1^k + 4x_2^k + 3x_3^k - 4 \\ x_1^k + x_2^k + x_3^k - 2 \end{pmatrix} \in \partial \mathcal{L}(u^k).$$

$\bar{\mathcal{L}} \geq \text{val}(P) \geq \underline{\mathcal{L}}$, para obtermos um valor inicial para $\bar{\mathcal{L}}$, basta conseguirmos uma solução viável de (P); por exemplo, $x_1 = 1, x_2 = x_3 = 0$, fornecendo $x_0 = -5$. Logo tomaremos $\bar{\mathcal{L}} = -5$. O valor inicial de $\underline{\mathcal{L}}$ poderá ser

$$\underline{\mathcal{L}} = \mathcal{L}(u^0) = \min\{-5x_1 - 6x_2 - 3x_3 \mid x_1 + x_3 \leq 1, x_j \in \{0, 1\}, j = 1, 2, 3\}.$$

É fácil observar que $\mathcal{L}(u^0) = -5 \times 1 - 6 \times 1 - 3 \times 0 = -11$, isto é, $x_1^0 = x_2^0 = 1, x_3^0 = 0$; faremos então $\underline{\mathcal{L}} = -11$. Tomaremos

$$\begin{aligned} \gamma^0 = \begin{pmatrix} \gamma_1^0 \\ \gamma_2^0 \end{pmatrix} &= \begin{pmatrix} 3x_1^0 + 4x_2^0 + 3x_3^0 - 4 \\ x_1^0 + x_2^0 + x_3^0 - 2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \times 1 + 4 \times 1 + 3 \times 0 - 4 \\ 1 \times 1 + 1 \times 1 + 1 \times 0 - 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}. \end{aligned}$$

Tomaremos $\rho = 2$, $\|\gamma^0\|^2 = 3^2 + 0^2 = 9$. Logo $\theta_0 = 2 \frac{-5 - (-11)}{9} = \frac{4}{3}$. calculemos então

$$u_1^1 = \max\{0, \frac{4}{3} \times 3\} = \frac{4}{3} \times 3 = 4, \quad u_2^1 = \max\{0, \frac{4}{3} \times 0\} = 0.$$

$\mathcal{L}(u^1)$ será dado por

$$\min \{-5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 3x_3 - 4)u_1^1 + (x_1 + x_2 + x_3 - 2)u_2^1 \mid x \in X\},$$

ou ainda

$$\mathcal{L}(u^1) = \min \{-5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 3x_3 - 4)4 \mid x \in X\},$$

$$\mathcal{L}(u^1) = \min\{7x_1 + 10x_2 + 5x_3 - 16 \mid x_1 + x_3 \leq 1, x_j \in \{0, 1\}\}.$$

O valor de $\mathcal{L}(u^1)$ será obtido quando $x_1 = x_2 = x_3 = 0$, $\mathcal{L}(u^1) = -16$.

Como $-16 < -11$ não modificaremos o valor de $\underline{\mathcal{L}}$, isto é, $\underline{\mathcal{L}}$ continuará valendo -11 . Verificamos também que $x_1 = x_2 = x_3 = 0$ é uma solução viável de (P) , fornecendo $x_0 = 0$. Continuaremos com $\bar{\mathcal{L}} = -5$.

Tomaremos

$$\begin{aligned} \gamma^1 &= \begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \end{pmatrix} = \begin{pmatrix} 3x_1^1 + 4x_2^1 + 3x_3^1 - 4 \\ x_1^1 + x_2^1 + x_3^1 - 2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \times 0 + 4 \times 0 + 3 \times 0 - 4 \\ 1 \times 0 + 1 \times 0 + 1 \times 0 - 2 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}. \end{aligned}$$

$$\|\gamma^1\|^2 = (-4)^2 + (-2)^2 = 16 + 4 = 20 \quad \text{e} \quad \theta_1 = 2 \frac{-5 - (-11)}{20} = \frac{3}{5}.$$

$$u_1^2 = \max\{0, 4 + \frac{3}{5} \times (-4)\} = \frac{8}{5}, \quad u_2^2 = \max\{0, 0 + \frac{3}{5} \times (-2)\} = 0.$$

$\mathcal{L}(u^2)$ será dado por

$$\min \{-5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 3x_3 - 4)u_1^2 + (x_1 + x_2 + x_3 - 2)u_2^2 \mid x \in X\},$$

ou ainda

$$\mathcal{L}(u^2) = \min \{-5x_1 - 6x_2 - 3x_3 + (3x_1 + 4x_2 + 3x_3 - 4)\frac{8}{5} \mid x \in X\},$$

$$\mathcal{L}(u^2) = \min\{-\frac{1}{5}x_1 + \frac{2}{5}x_2 + \frac{1}{5}x_3 - \frac{32}{5} \mid x_1 + x_3 \leq 1, x_j \in \{0, 1\}\}.$$

$\mathcal{L}(u^2)$ é encontrado para $x_1^2 = 1$, $x_2^2 = x_3^2 = 0$, fornecendo $\mathcal{L}(u^2) = -\frac{33}{5} = -6,6$. Para a próxima iteração faríamos $\underline{\mathcal{L}} = -6,6$. Como a função objetivo de (P) possui todos os coeficientes inteiros, então $-6 \leq \text{val}(P) \leq -5$. Logo a solução viável de (P) , $x_1 = 0$, $x_2 = 1$, $x_3 = 0$, fornecendo $x_0 = -6$ é ótima de (P) .

É interessante notar que a função objetivo de (D) começou com o valor -11 decresceu para -16 e, finalmente, cresceu até $-6,6$. ■

11.3 Exercícios

1. Seja

$$(P) : \text{ maximizar } x_0 = 2x_1 + 3x_2 + 4x_3 \quad (11.43)$$

sujeito a:

$$12x_1 + 19x_2 + 30x_3 \leq 46 \quad (11.44)$$

$$49x_1 + 40x_2 + 31x_3 \leq 76 \quad (11.45)$$

$$x_j \in \{0, 1\}. \quad (11.46)$$

Determinar o valor ótimo do dual de (P) quando relaxarmos (11.44).
Determinar o valor ótimo do dual de (P) quando relaxarmos (11.45).
Calcular $val(\bar{P})$, onde (\bar{P}) é formado de (P) trocando-se (11.46) por $0 \leq x_j \leq 1, j = 1, 2, 3$. Encontrar um ótimo de (P) . Seja (Q) o problema dual-lagrangeano de (P) quando relaxamos ao mesmo tempo (11.44) e (11.45), mostrar que $val(Q) = val(\bar{P})$, sem resolver (Q) .

Referências Bibliográficas

- [Ar 93] A. ARBEL. Exploring Interior-Point Linear Programming - Algorithms and Software. MIT Press, Cambridge (MA), 1993.
- [Ba 63] E. BALAS. Programare liniara cu variabili bivalente. *Proceedings of the Third Scientific Session on Statistics*, Bucureste, Rumânia, 1963.
- [Ba 65] E. BALAS. An additive algorithm for solving linear programs with zero-one variables. *Operations Research*, vol. 13, 1965, pp. 517-526.
- [Ba 79] E. BALAS. Disjunctive programming. *Annals of Discrete Mathematics*, vol. 5, 1979, 5-51.
- [Ba 82] E. BALAS. Integer programming. Design Research Center, DRC-70-13-82, Carnegie-Mellon University, Pittsburgh, PA, USA.
- [Ba 98] E. BALAS. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, vol. 89, 1998, 3-44.
- [BaCeCo 93] E. BALAS, Sebastián CERIA and Gérard CORNUÉJOLS. A lift-and-project cutting plane algorithm for mixed 0 – 1 programs. *Mathematical Programming*, vol. 58, 1993, 295-324.
- [BaJoNeSaVa 98] Cynthia BARNHART, Ellis L. JOHNSON, George L. NEMHAUSER, Martin W. P. SAVELSBERGH and Pamela H. VANCE. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, vol.46, 1998, 316-329.

- [BaJa 77] M. S. BAZARAA e J. J. JARVIS. *Linear Programming and Network Flows*. John Willey and Sons, 1977.
- [Be 79] E. M. L. BEALE. Branch-and-bound methods for mathematical programming. *Annals of Discrete Mathematics*, vol. 5, 201-220, 1979.
- [Bea 92] J. E. BEASLEY. Lagrangean relaxation. Report, The Management School, Imperial College, Londres, UK, 1992.
- [Ben 62] J. F. BENDERS. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, vol. 4, 238-253, 1962.
- [Ber 70] Claude BERGE. *Graphes et hypergraphes*. Dunod, Paris, 1970.
- [BeTs 97] D. BERTSIMAS e J. N. TSITSIKLIS. *Introduction to Linear Optimization*. Athena Scientific, Belmont, Massachusetts, 1997.
- [Bl 77] R. G. BLAND. New finite pivoting rules for the simplex method. *Mathematics of Operations Research*, vol. 2, 1977, pp. 103-107.
- [Bo 79] Paulo Oswaldo BOAVENTURA NETTO. *Teoria e Modelos de Grafos*. Editora Edgard Blücher Ltda, São Paulo, 1979.
- [BrHaMa 77] S. P. BRADLEY, A. C. HAX e T. L. MAGNANTI. *Applied Mathematical Programming*. Addison Wesley Publishing Company, 1977.
- [BrOlBo 81] P. F. BREGALDA, A. A. F. de OLIVEIRA e C. T. BORNSTEIN. *Introdução à Programação Linear*. Editora Campus, Rio de Janeiro, 1981.
- [CaMa 94] Ruy E. CAMPELLO e Nelson MACULAN. *Algoritmos e Heurísticas – Desenvolvimento e Avaliação de Performance*. Editora da Universidade Federal Fluminense (EDUFF), Niterói, RJ, 1994.
- [ChCoHe 53] A. CHARNES, W. W. COOPER e A. HENDERSON. *An Introduction to Linear Programming*. John Wiley and Sons, 1953.

- [Ch 83] Vasec CHVÁTAL. *Linear Programming*. W. H. Freeman and Company, 1983, New York / San Francisco, USA.
- [Da 65] R. DAKIN. A tree search algorithm for mixed integer programming problems. *Computer Journal*, vol. 8, 1965, pp. 250-255.
- [Dan 51] George B. DANTZIG. Maximization of a linear function of variables subject to linear inequalities. *Activity Analysis of Production and Allocation*, editado por C. Koopmans, Wiley, New York, 1951, pp. 359-373.
- [Dan 91] George B. DANTZIG. Linear Programming. *History of Mathematical Programming – A collection of personal reminiscences*, editado por J.K. Lenstra, A.H.G. Rinnooy Kan and A. Schrijver, CWI, North-Holland, Amsterdam, 1991, pp. 19-31.
- [Dan 63] George B. DANTZIG. *Linear Programming and Extensions*. Princeton University Press, 1963, Princeton, NJ, USA.
- [Dan 66] George B. DANTZIG. All shortest routes in a graph. *Proc. I.C.C. Conference on Theory of Graphs*, Roma, Gordon and Breach, NY, 1966, pp. 91-92.
- [DanWo 60] George B. DANTZIG e P. WOLFE. The decomposition algorithm for linear programming. *Operations Research*, vol. 8, 1960, pp. 101-111.
- [De 74] Jacques DÉLORME. Contributions à la résolution du problème de recouvrement: méthodes de troncature. Docteur-ingénieur dissertation, Université Paris VI, 1974.
- [DeSo 89] Jacques DESROCHERS and François SOUMIS. A column-generation approach to the urban transit crew scheduling problem. *Transportation Science*, vol. 23, 1989, 1-13.
- [Di 67] I. I. DIKIN. Iterative solution of problems of linear and quadratic programming. *Soviet Mathematics Doklady*, vol. 8, 1967, pp. 674-675.
- [Di 74] I. I. DIKIN. On the speed of an iterative process. *Upravlyaemye Sistemi*, vol 12, 1974, pp. 54-60.

- [Dij 59] E. W. DIJKSTRA. A note on two problems in connection with graphs. *Numerische Mathematik*, vol. 1, 1959, pp. 269-271.
- [DoSaSo 58] R. DORFMAN, P. A. SAMUELSON e R. M. SOLOW. *Linear Programming and Economic Analysis*. McGraw-Hill, 1958.
- [Ev 79] Shimon EVEN. *Graph Algorithms*. Computer Science Press, 1979.
- [Eve 63] H. EVERETT. Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. *Operations Research*, vol. 11, 1963, pp.399-417.
- [Fa 02] J. FARKAS. Über die Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik*, vol. 124, 1902, pp. 1-27.
- [FiMc 55] A. FIACCO e G. MCCORMICK. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, New York, 1955.
- [Fi73] M. L. FISHER. Optimal solution of scheduling problem using Lagrange multipliers: Part I. *Operations Research*, vol. 21, 1973, pp. 1114-1127.
- [Fi 81] M. L. FISHER. The Lagrangian relaxation method for solving integer programming problems. *Management Science*, vol. 27, 1981, pp. 01-18.
- [FiSh 74] M. L. FISHER e J. F. SHAPIRO. Constructive duality in integer programming. *SIAM Journal of Applied Mathematics*, vol. 27, 1974, pp. 31-52.
- [FiNoSh 75] M. L. FISHER W.D. NORTHUP e J. F. SHAPIRO. Using duality to solve discrete optimization problems: Theory and computational experience. *Mathematical Programming Study*, vol. 3, 1975, pp. 56-94.
- [FoFu 62] L. R. FORD, Jr. e D. R. FULKERSON. *Flows in Networks*. Princeton University Press, 1962.
- [FoHiTo 74] J.J.H. FORREST, J.P.H. HIRST and J.A. TOMLIN. Practical solution of large mixed integer programming problems

- with UMPIRE. *Management Science*, vol. 20, 1974, 739-773.
- [Fo 60] R. FORTET. Applications de l'algèbre de Boole en recherche opérationnelle. *Revue Française de Recherche Opérationnelle*, vol. 4, 1960, 17-25.
- [Fou 890] J. B. J. FOURIER. *Oeuvres*. "Second extrait", G. DARBOUTX, Gauthiers-Villars, 1890, pp. 325-328.
- [Fr 55] K. R. FRISCH. The Logarithmic Potential Method of Convex Programming. Memorandum, University Institute of Economics, Oslo, Norway, 1955.
- [Ga 60] D. GALE. *The Theory of Linear Economic Models*. McGraw-Hill, 1960.
- [GaNe 72] Robert S. GARFINKEL and George L. NEMHAUSER. *Integer programming*. John Wiley & Sons, 1972.
- [Gar 60] W. W. GARVIN. *Introduction to Linear Programming*. McGraw-Hill, 1960.
- [Gas 58] S. I. GASS. *Linear Programming: Methods and Applications*. McGraw-Hill, 1958.
- [Ge 67] A. M. GEOFFRION. Integer programming by implicit enumeration and Balas's Method. *SIAM Review*, vol. 7, pp. 178-190, 1967.
- [Ge 69] A. M. GEOFFRION. An improved implicit enumeration approach for integer programming. *Operations Research*, vol. 17, pp. 437-454, 1969.
- [Ge 74] A. M. GEOFFRION. Lagrangian relaxation and its use in integer programming. *Mathematical Programming Study*, vol. 2, pp. 82-114, 1974.
- [GiMuWr 91] P. E. GILL, W. MURRAY e M. H. WRIGHT. *Numerical Linear Algebra and Optimization*. vol. 1, Addison-Wesley Publishing Company, 1991.
- [GiEl 71] S. E. GIRÃO BARROSO e A. R. von ELLENRIEDER. *Programação Linear*. Almeida Neves Editores, Rio de Janeiro, 1971.

- [Gl 65] F. GLOVER. A multiphase-dual algorithm for the zero-one integer programming problem. *Operations Research*, vol. 13, pp. 879-919, 1965.
- [GoYo 73] E. GOLDSTEIN e D. YODINE. *Problèmes Particuliers de la Programmation Linéaire*. MIR, Moscou, 1973.
- [GoVo 96] G. H. GOLUB e C. F. VAN LOAN. *Matrix Computations*. The Johns Hopkins University Press, 1996 (third edition).
- [Go 58a] R. GOMORY. An algorithm for integer solutions to linear programming. Princeton IBM Math. Research Report, 1958, publicado também em *Recent Advances In Mathematical Programming* (Graves and Wolfe editors), McGraw Hill, 1963.
- [Go 58b] R. GOMORY. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, vol. 64, 275-278, 1958.
- [GoMi 79] M. GONDRAN e M. MINOUX. *Graphes et algorithmes*. EYROLLES, Paris, França, 1979.
- [Gon 89] Clovis C. GONZAGA. *Algoritmos de Pontos Interiores em Programação Linear*. Primeira Escola Brasileira de Otimização, UFRJ, Rio de Janeiro, 1989.
- [Gon 92] Clovis C. GONZAGA. Path-following methods for linear programming. *SIAM Review*, vol. 34, n. 2, 167-224, 1992.
- [Gr 71] H. GREENBERG. *Integer programming*. Academic Press, 1971.
- [GuPrSe 2000] Christelle GUÉRET, Christian PRINS e Marc SEVAUX. *Programmation Linéaire – 65 problèmes d’optimisation modélisés et résolus avec Visual Xpress*. Eyrolles, Paris, 2000.
- [GuKi 87a] Monique GUIGNARD and S. KIM. Lagrangean decomposition for integer programming: Theory and applications. *RAIRO-Operations Research*, vol. 21, 307-324, 1987.
- [GuKi 87b] Monique GUIGNARD and S. KIM. Lagrangean decomposition: a model yielding stronger Lagrangian bounds. *Mathematical Programming*, vol. 32, 215-228, 1987.

- [GuRoTeVi 95] O. GÜLER, C. ROOS, T. TERLAKY, and J.-P. VIAL. A survey of the implications of the behavior of the central path for the duality theory of linear programming. *Management Science*, vol. 41, 1922-1934, 1995.
- [Ha 65] G. HADLEY. *Linear Programming*. Addison-Wesley Publishing Company, 1965.
- [HaMiLa 87] Pierre HANSEN, Michel MINOUX and Martine LABBÉ. Extension de la programmation linéaire généralisée au cas des programmes mixtes. *C. R. Acad. Sci. Paris*, t. 305, série I, 1987, 569-572.
- [HaJaPo 91] Pierre HANSEN, Brigitte JAUMARD and Marcus POGGI DE ARAGÃO. Un algorithme primal de programmation linéaire généralisée pour les programmes mixtes. *C. R. Acad. Sci. Paris*, t. 313, série I, 1991, 557-560.
- [Har 72] Frank HARARY. *Graph Theory*. Addison-Wesley Publishing Company, 1972.
- [HeKa 70] M. HELD e R. M. KARP. The traveling salesman problem and minimum spanning trees. *Operations Research*, vol. 18, pp. 1138-1162, 1970.
- [HeKa 71] M. HELD e R. M. KARP. The traveling salesman problem and minimum spanning trees: Part II. *Mathematical Programming*, vol. 1, pp. 06-25, 1971.
- [HeWoCr 74] M. HELD, P. WOLFE e H. D. CROWDER. Validation of subgradient optimization. *Mathematical Programming*, vol. 6, pp. 62-88, 1974.
- [He 94] D. den HERTOG. *Interior Point Approach to Linear Programming*. Kluwer Academic Publishers, 1994.
- [Hu 69] T. C. HU. *Integer Programming and Network Flows*. Addison Wesley, 1969.
- [HuCa 87] C. HUMES e A. F. P. CASTRO HUMES. *Uma Visão Conceitual da Programação Linear*. Sociedade Brasileira de Pesquisa Operacional (SOBRAPO), Rio de Janeiro, 1987.

- [Ib 77] T. IBARAKI. The power of dominance relations in branch-and-bound algorithms. *Journal of ACM*, vol. 24, 1977, 264-279.
- [Ib 82] T. IBARAKI. The power of upper and lower bounding functions in branch-and-bound algorithms. *Journal of the Operations Research Society of Japan*, vol. 25, 1982, 292-320.
- [IlTe 2002] T. ILLÉS and T. TERLAKY. Pivot versus interior point methods: Pros and cons. *European Journal of Operational Research*, vol. 140, no. 2, pp. 170-190, 2002.
- [JaLa 98] Éric JACQUET-LAGRÈZE. *Programmation Linéaire – Modélisation et mise en oeuvre informatique*. Economica, Paris, 1998.
- [Ja 97] Benjamin JANSEN. *Interior Point Techniques in Optimization*. Kluwer Academic Publishers, 1997.
- [JaCa 87] Maria Helena JARDIM CAMPOS. Solução por métodos de subgradientes de primeira ordem dos problemas de relaxação lagrangeana em programação combinatória. Ph.D. dissertation, COPPE, Federal University of Rio de Janeiro, 1987.
- [JaCaMa 88] Maria Helena JARDIM CAMPOS and Nelson MACULAN. Problemas de otimização relacionados ao corte de bobinas de papel. Atas do XVI Colóquio Brasileiro de Matemática, IMPA-CNPq, Rio de Janeiro, 1988, 159-182.
- [JaCaMa 95] Maria Helena JARDIM CAMPOS and Nelson MACULAN. Optimization problems related to the cut of paper reels: a dual approach. *Investigación Operativa*, vol. 5, 1995, 45-53.
- [Je 79] R. G. JEROSLOW. An introduction to the theory of cutting planes. *Annals of Discrete Mathematics*, vol. 5, pp. 71-95, 1979.
- [Ka 39] L. V. KANTOROVICH. *Métodos Matemáticos na Organização e no Planejamento da Produção* (em russo). Editora da Universidade Estatal de Leningrado, 1939,

- (tradução inglesa: *Management Science*, vol. 6, pp. 366-422, 1960).
- [Kar 84] N. KARMARKAR. A new polynomial algorithm for linear programming. *Combinatorica*, vol. 4, 1984, pp. 373-395.
- [Kha 79] L. KHACHIAN. A polynomial algorithm for linear programming. *Doklady Akademii Nauk SSSR*, vol. 244, pp. 191-194. Em russo. Tradução em inglês: *Soviet Mathematics Doklady*, vol. 20, 1979, 191-194.
- [KlMi 72] V. KLEE e G. MINTY. How good is the simplex algorithm? *Inequalities-III*, editor O. Shisha, Accademic Press, New York, pp.159-175, 1972.
- [KoMiYo 89] M. KOJIMA, S. MIZUNO e A. YOSHISE. A polynomial-time algorithm for a class of linear complementarity problems. *Mathematical Programming*, vol. 44, 1989, pp. 1-26.
- [KoMiYo 91] M. KOJIMA, S. MIZUNO e A. YOSHISE. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Mathematical Programming*, vol. 50, 331-342, 1991.
- [KoFi 69] A. A. KORBUT and Y. FINKELSTEIN. *Diskretnoje Programirovanie*. Nauka, Moscow, Russia, 1969.
- [Ko 80] L. B. KOVÁCS. *Combinatorial methods of Discrete Programming*. Akadémia Kiadó, Budapest, Hungary, 1980.
- [Ku 56] H. W. KUHN. Solvability and consistency for linear equations and inequalities. *American Mathematical Monthly*, vol. 63, 1956, pp. 217-232.
- [Kun 72] J. KUNTZMANN. *Théorie des Réseaux*. Dunod, Paris, 1972.
- [LaDo 60] A. H. LAND e A. G. DOIG. An automatic method for solving discrete programming problems. *Econometrica*, vol. 28, n. 3, 1960, pp. 497-520.
- [LaPo 79] A. H. LAND e S. POWELL. Computer codes for problems of integer programming. *Annals of Discrete Mathematics*, vol. 5, pp. 221-269, 1979.

- [Las 70] L. S. LASDON. *Optimization Theory for Large Systems*. Macmillan, 1970.
- [Law 76] E. L. LAWLER. *Combinatorial Optimization: Networks and Matroids*. Holt, Rinehart and Winston, 1976.
- [LeRiSc 91] J. K. LENSTRA, A. H. G. RINNOOY KAN e A. SCHRIJVER. *History of Mathematical Programming - S Collections of Personal Reminiscences*. CWI - North Holland, 1991.
- [LiMuSwKa 63] J. LITTLE, K. MURTY, D. SWEENEY e C. KAREL. An algorithm for the Traveling Salesman Problem. *Operations Research*, vol. 11, n. 5, 1963, pp. 972-989.
- [LoSa 55] J. LORIE e L. J. SAVAGE. Three problems in capital rationing. *Journal of Business*, vol.28, n. 6, 1955, pp. 229-239.
- [LoSc 91] L. LOVÁSZ and A. SCHRIJVER. Cones of matrices and set-functions and 0 – 1 optimization. *SIAM Journal on Optimization*, vol. 1, n. 2, 1991, 166-190.
- [LuBe 96] Abilio LUCENA and John E. BEASLEY. Branch and cut algorithms. *Advances in linear and integer programming*. Beasley, editor, Clarendon Press, Oxford (UK), 1996, 187-221.
- [Lu 89] D. G. LUENBERG. *Introduction to Linear and Nonlinear Programming*. Addison-Wesley Publishing Company, 1989 (segunda edição).
- [Mac 75] H. MACHADO. *Programação Linear*. 10^o Colóquio Brasileiro de Matemática, IMPA - CNPq, Rio de Janeiro, 1975.
- [Mc 80] L. MCLINDEN. An analogue of Moreau's proximation theorem, with applications to the nonlinear complementarity problem. *Pacific Journal of Mathematics*, vol.88, 101-161, 1980.
- [Ma 78] Nelson MACULAN. *Programação linear inteira*. Editora COPPE-UFRJ, Rio de Janeiro, Brazil, 1978.

- [Ma 83] Nelson MACULAN. Programmation linéaire en nombres entiers. Publication n. 486, Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Canad , 1983.
- [MaPe 80] N. MACULAN e M. V. F. PEREIRA. *Programa  o Linear*. Editora ATLAS, S o Paulo, 1980.
- [Ma 87] Nelson MACULAN. The Steiner problem in graphs. *Annals of Discrete Mathematics*, vol. 31, 1987, 185-212.
- [MaMiPl 92] Nelson MACULAN, Philippe MICHELON and G rard PLATEAU. Column-generation in linear programming with bounding variable constraints and its applications in integer programming. *Pesquisa Operacional*, vol. 12, n. 2, 1992, 45-57.
- [Maf 2000] Francesco MAFFIOLI. *Elementi di Programmazione Matematica*. Casa Editrice Ambrosiana, seconda edizione, Mil o, 2000.
- [MaTo 90] Silvano MARTELLO and Paolo TOTH. *Knapsak problems: algorithms and computer implementations*. John Wiley & Sons, 1990.
- [Mi 91] Philippe MICHELON. M thodes lagrangiennes pour programmation lin aire avec variables enti res. *Investigaci n Operativa*, vol.2, n. 2, 127-146, 1991.
- [MiMa 91] Philippe MICHELON and Nelson MACULAN. Lagrangean decomposition for integer nonlinear programming. *Mathematical Programming*, vol.52, 303-313, 1991.
- [Min 11] H. MINKOWSKI. *Gesammelte Abhandlungen*. Teubener, Leipzig, 1911.
- [Mino 86] M. MINOUX. *Mathematical Programming: Theory and Algorithms*. John Wiley and Sons, 1986.
- [MiToYe 93] S. MIZUNO, M. TODD e Y. YE. On adaptive step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, vol. 18, 1993, pp. 964-981.

- [MoAd 89] R. MONTEIRO e I. ADLER. Interior path-following primal-dual algorithms. Part I: Linear Programming. *Mathematical Programming*, vol. 44, 1989, pp. 27-41.
- [Mu 76] K. G. MURTY. *Linear and combinatorial programming*. John Wiley & Sons, 1976.
- [NeWo 88] G L. NEMHAUSER e L. A. WOLSEY. *Integer and combinatorial optimization*. Wiley, 1988.
- [NeRiT0 89] G L. NEMHAUSER, A. H. G. RINNOOY KAN e M. J. TODD (editores). *Optimizations. Handbooks in Operations Research and Management Science*, vol. 1, North-Holland, 1989.
- [Ni 80] N.J. NILSSON. *Principles of artificial intelligence*. Tioga Publishing Company, Palo Alto, USA, 1980.
- [Or 68] W. ORCHARD-HAYS. *Advanced Linar Programming Computing Techniques*. McGraw-Hill, 1968.
- [PaSt 82] C. H. PAPADIMITRIOU e K. STEIGLITZ. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, 1982.
- [PaRa 88] R. Gary PARKER and Ronald L. RARDIN. *Discrete Optimization*. Academic Press, 1988.
- [Pl 2000] Gérard PLATEAU. Optimisation Combinatoire. Cours de D.E.A. (Intelligence Artificielle et Optimisation Combinatoire), 5 parties, LIPN, Université Paris Nord, França, 2000.
- [Po 66] B. T. POLYAK. A general method for solving extremum problems. *Soviet Mathematics*, vol. 8, pp. 593-597, 1966.
- [Pu 75] A. L. PUCCINI. *Introdução à Programação Linear*. Livros Técnicos e Científicos (LTC), Rio de Janeiro, 1975.
- [ReMa 92] Hernaldo REINOSO e Nelson MACULAN. Lagrangean decomposition for integer programming: a new scheme. *INFOR*, vol. 52, n. 2, pp. 1-5, 1992.

- [RiMiPe 89] Celso C. RIBEIRO, Michel MINOUX and M. C. PENNA. An optimal column-generation-with-ranking algorithm for very large scale partitioning problems in traffic assignment. *European Journal of Operational Research*, vol. 41, 1989, 232-239.
- [Ro 70] R. T. ROCKAFELLER. *Convex Analysis*. Princeton University Press, 1970.
- [Sa 75] H. M. SALKIN. *Integer Programming*. Addison Wesley, 1975.
- [Sak 83] M. SAKAROVITCH. *Linear Programming*. Springer-Verlag, 1983.
- [Sc 86] A. SCHIJVER. *Theory of Linear and Integer Programming*. John Wiley and Sons, NY, 1986.
- [Sh 71] J. F. SHAPIRO. Generalized Lagrange multipliers in integer programming. *Operations Research*, vol. 19, pp. 68-79, 1971.
- [Sh 79] J. F. SHAPIRO. A survey of Lagrangian techniques for discrete optimization. *Annals of Discrete Mathematics*, vol. 5, pp. 133-138, 1979.
- [Si 72] Michel SIMONNARD. *Programmation Linéaire - Technique du Calcul Economique*. Dunod, 1973, Paris, França, 2 volumes, com a participação de Xavier CHOUTET.
- [Sim 58] M. H. SIMONSEN *Introdução à Programação Linear*. Notas de Matemática, n. 8, IMPA - CNPq, 1958.
- [Sp 79] K. SPIELBERG. Enumerative methods in integer programming. *Annals of Discrete Mathematics*, vol. 5, pp. 139-184, 1979.
- [St 72] J. E. STRUM. *Introduction to Linear Programming*. Holden-Day, 1972.
- [SyDeKo 83] Maciej M. SYSLO, Narsingh DEO e Janusz S. KOWALIK. *Discrete Optimization Algorithms with PASCAL programs*. Prentice-Hall, Inc., 1983, Englewood Cliffs, NJ, USA.

- [Ta 75] H. A. TAHA. *Integer Programming - Theory, Applications, and Computations*. Academic Press, 1975.
- [Th 82] Hervé THIRIEZ. *Comprendre et utiliser les modèles en gestion*. Les éditions d'organisation, 1982.
- [TiWo 78] J. TIND e L. A. WOLSEY. An unifying framework for duality theory in mathematical programming. Discussion paper 7834, C.O.R.E., Université Catholique de Louvain, Belgique, 1978.
- [Va 95] Pamela H. VANCE. Branch-and-price algorithms for the one-dimensional cutting stock problem. *Computational Optimization and Applications*, vol.9, 1998, 211-228.
- [VaBaJoNe 94] Pamela H. VANCE, Cynthia BARNHART, Ellis L. JOHNSON and George L. NEMHAUSER. Solving binary cutting stock problems by column generation and branch-and-bound. *Computational Optimization and Applications*, vol.3, 1994, 111-130.
- [Van 94] François VANDERBECK. Decomposition and column generation for integer programming. Thèse de doctorat en sciences appliquées, Université Catholique de Louvain, Belgium, 1994.
- [Vand 98] Robert J. VANDERBEL. *Linear Programming – Foundations and Extensions*. Kluwer Academic Publishers, 1998.
- [Var 72] P. P. VARAYA. *Notes on Optimization*. Van Nostrand Reinhold, 1972.
- [We 90] D. de WERRA. *Eléments de Programmation Linéaire avec Application aux Graphes*. Presses Polytechniques Roman-des, Lausanne, Suïça, 1990.
- [Wi 87] P. WINTER. Steiner problems in networks: A survey. *Networks*, vol. 17, 1987, 129-167.
- [Wo 81] L. A. WOLSEY. Integer programming duality: Price functions and sensitivity analysis. *Mathematical Programming*, vol. 20, pp. 173-195, 1981.
- [Wo 98] Laurence A. WOLSEY. *Integer Programming*. John Wiley & Sons, 1998.

- [Wr 97] Stephen J. WRIGHT. *Primal-Dual Interior-Point Methods*. SIAM, 1997.
- [Zi 74] S. ZIONTS. *Linear and Integer Programming*. Prentice-hall, 1974.

Índice Remissivo

- Algoritmo
 - Afim-Escala, 123, 128, 133
 - de Pontos Interiores, 72, 122, 145
 - de Redução Potencial, 178, 180
 - de Trajetória Central, 146
 - de Trajetória Central de Passos Curtos, 147, 148
 - de Trajetória Central de Passos Longos, 167
 - de Trajetória Central Inviável, 172
 - do Simplex, 20, 107
 - Dual do Simplex, 45, 47, 116
 - Preditor-Corretor, 162
- Análise de Sensibilidade, 51
- Ciclagem, 35
- Complexidade Polinomial, 122
- Condições
 - de Complementaridade, 142
 - de Karush-Kuhn-Tucker, 141
- Cone Polédrico, 80
- Conjunto Convexo, 27, 76
- Conjunto Polédrico, 27, 78, 87
- Convergência, 33, 207
- Desigualdade Válida, 197
- Dikin, 123
- Direção, 81
- Direção de Cauchy, 131
- Dualidade, 40
- Dualidade em Programação Inteira, 222
- Enumeração, 202
- Gap de Dualidade, 145, 148, 178
- Geração de Colunas, 91, 224
- Gomory, 197
- Hiperplano, 76
- Interpretação
 - das Variáveis Duais, 45
 - Geométrica, 27
- Jacobiano, 143
- Kachian, 122, 124
- Karmarkar, 122
- Matriz
 - de Projeção, 130
 - Inversa, 25
- Método
 - de Balas, 201, 207
 - de “Branch-and-Bound”, 209
 - de Eliminação de Fourier, 64
 - de Elipsóides, 124
 - de Newton, 143
 - de Planos de Corte, 196
 - de Sub-Gradientes, 228
 - do Simplex, 19
- Obtenção de Solução Viável, 64, 111
- Politopo, 78
- Ponto Central, 139
- Pós-Otimização, 53
- Problema

- Auxiliar, 96
- Barreira, 139
- Mestre, 93
- Programação
 - Inteira, 192
 - Linear, 7, 8, 24, 122
 - Não Linear 0-1, 208
- Propriedade de Integralidade, 227
- Raio, 78, 80
- Região Viável, 8
- Regra de Bland, 36
- Relaxação Lagrangeana, 222
- Restrições Canalizadas, 107
- Sistema Inconsistente, 73
- Solução
 - Básica, 9, 28, 108
 - Básica Dual, 42
 - Degenerada, 9, 24, 35
 - Ótima, 8
 - Viável, 8
- Sub-diferencial, 230
- Sub-gradiente, 230, 232
- Teorema
 - da Dualidade Forte, 43
 - da Dualidade Fraca, 41
 - da Existência, 44
 - das Folgas Complementares, 44
- Trajetória Central, 140
- Vértice, 77
- Vizinhança da Trajetória Central, 147, 167