

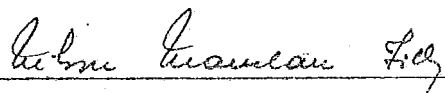
SOBRE A SOLUÇÃO DOS PROBLEMAS DE PROGRAMAÇÃO LINEAR

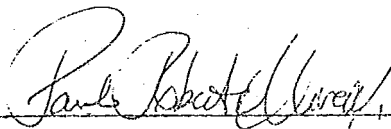
4

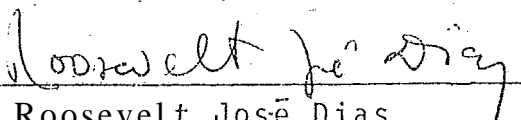
Renato Craveiro de Souza

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS (D. Sc.)

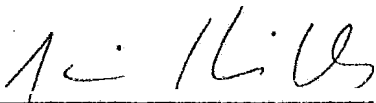
Aprovada por:


Prof. Reiçon Maculan Filho


Prof. Paulo Roberto de Oliveira


Prof. Roosevelt José Dias


Prof. Ruderico Ferraz Pimentel


Prof. Jair Koiller

RIO DE JANEIRO, RJ - BRASIL

JANEIRO de 1981

i

À minha esposa Elaine e à meus filhos Kildare, Kilvia e Kilmer que com amor compreensão, me deram as condições necessárias à elaboração deste trabalho.

AGRADECIMENTOS

Desejo agradecer ao Professor Nelson Maculan Filho, **não** somente pelo incentivo que me deu durante todo o período que estive sob sua orientação, mas principalmente pela amizade com que acompanhou meu trabalho.

A todos aqueles que direta ou indiretamente colaboraram para a conclusão deste Tese.

Aos colegas da Universidade Federal do Ceará, pela oportunidade concedida.

Aos meus pais os ensinamentos de disciplina e perseverança no trabalho.

Ainda à COPPE, CAPES e ao CNPq pelo apoio financeiro.

BIOGRAFIA DO AUTOR

Renato Craveiro de Souza, nascido em 3 de março de 1938, na cidade de Limoeiro do Norte, Estado do Ceará. Ingressou em março de 1961 na Faculdade Católica de Filosofia do Ceará, diplomando-se Bacharel em Matemática em dezembro de 1963, licenciando em Ciências pela mesma Faculdade em dezembro de 1964. Desde março de 1964 trabalha como professor na Universidade Federal do Ceará, e atualmente é professor adjunto do Departamento de Estatística e Matemática Aplicada. Em março de 1974 ingressou na COPPE/UFRJ obtendo em dezembro de 1975 o título de Mestre em Ciências em Engenharia de Sistemas e Computação.

RESUMO

Tentamos neste trabalho apresentar um Algoritmo para resolver Problemas de Programação Linear, de uma maneira não convencional. Inicialmente no capítulo I, fizemos uma explanação da Programação Linear, com intuito comparativo relativamente aos capítulos subseqüentes, bem como, de ferramenta para o desenvolvimento dos mesmos. Baseados em teoremas essencialmente simples, desenvolvemos, no capítulo II, um algoritmo o qual tem por objetivo principal, detectar em seu primeiro passo, um vetor o qual deverá ter sua presença assegurada na base final, e assim o fará em seus passos subseqüente, até termos no m-ésimo passo uma base Ótima. Aproveitamos o ensejo e inserimos de uma maneira um tanto didática o Algoritmo de Khachiyan para a resolução de Problemas de Programação Linear, haja visto ser um algoritmo não convencional, que nos Últimos anos tem recebido especial atenção dos que lidam com a pesquisa operacional e a ciência da computação. Salientamos outrossim que sua inserção visa exclusivamente dar maior divulgação e quiçá propicie trabalhos de aprimoramentos que lhe confira utilizações práticas de uma maneira mais ampla.

ABSTRACT

In this thesis we try to present an algorithm to solve Linear Programming problems, by a non conventional way. Initially in chapter I, we explain Linear Programming, with comparative intuition relative to the subsequent chapters, as well as, tools for development of the above.

Based on essentially simple theorems, we develop, in chapter II, an algorithm which has as its principal objective, to detect in its first step, a vector which should have its presence assured in the final basis vectors, and thus it will make in its subsequent steps, until we have in the m -th step an optimal basic vector group.

Taking advantage of the development, we insert in a didactic way the Algorithm of Khachiyan for the solution of Linear Programming problems, seeing that this is a non conventional Algorithm, that in recent years has received special attention of those who work in operational research and computer science. We also point out that its insertion is exclusively to give greater information and maybe give incentives for other research, with intent to perfect which opens other ways of utilization.

ÍNDICE

	<u>Págs.</u>
<u>CAPÍTULO I - TÓPICOS DE PROGRAMAÇÃO LINEAR.....</u>	1
I.1 - Introdução.....	1
I.2 - Formulação do Problema de Programação Linear.....	2
I.3 - Inequações e Equações.....	4
I.4 - Não Negatividade das Variáveis.....	4
I.5 - Problema de Minimizar e Maximizar.....	4
I.6 - Método do Simplex.....	6
I.7 - Base Artificial.....	23
I.8 - Variáveis Escalares.....	24
I.9 - Dualidade.....	25
<u>CAPÍTULO II - UMA NOVA ALTERNATIVA SOBRE A SOLUÇÃO DOS</u> <u>PROBLEMAS DE PROGRAMAÇÃO LINEAR.....</u>	27
II.1 - Definição do Problema.....	27
II.2 - Pseudoinversa, Operadores Projeção e seu Complemento Ortogonal.....	28
II.3 - Configuração de A_k^+ em Termos de A_{k-1}^+ e μ_{k-1}	31
II.4 - Nova Representação das Variáveis x_j	36
- Cone.....	40
II.6 - Exemplo Numérico.....	64

	<u>Págs.</u>
<u>CAPÍTULO III - ALGORITMO DE KHACHIYAN</u>	75
III.1 - Introdução.....	75
III.2 - Sistema de Inequações Lineares Estritas.....	77
III.3 - Afinidade entre Algoritmo de Khachiyan e o Problema de Programação Linear.....	105
<u>CAPÍTULO IV - CONCLUSÕES</u>	117
<u>BIBLIOGRAFIA</u>	121

CAPÍTULO I

TÓPICOS DE PROGRAMAÇÃO LINEAR

I.1 - INTRODUÇÃO

O desenvolvimento das forças produtivas e a planificação das indústrias, e de outros fatores de natureza econômica, adquirem no decorrer dos anos, cada vez mais importância, tornando suas soluções progressivamente mais difíceis. Foi nos países hoje mais desenvolvidos que, para solucionar estes problemas, surgiram os métodos matemáticos, e em particular, a Programação Linear. A primeira grande contribuição nesta área, foi o método do Simplex, desenvolvido em 1947 por GEORGE DANTZIG [10 p. 15] e sua equipe junto ao Departamento da Força Aérea dos Estados Unidos.

Por volta de 1952, com o advento dos computadores de alta velocidade, a programação linear, fazendo uso desse recurso como suporte, permitiu sua utilização prática no campo da organização e planificação da indústria, dando-lhes regimes ótimos de produção, gerando conseqüentemente maiores lucros, etc. Portanto aplicações de programação linear à economia, ao setor militar, e a outros domínios, não cessam de se estender. Foi sob o auspício do método do Simplex, que a Programação Linear aprimorou-se, até que, nos anos de 1960 surgiu o método primal-dual, desenvolvido por GOMORY e BALINSK [17], e similarniente um outro método foi desenvolvido independentemente

por GRAVES | 18 |.

A programação linear lida com métodos para detectar soluções ótimas de problemas, onde suas variáveis estão linearmente relacionadas entre si, e sujeitas a determinadas condições de restrição.

I.2 - FORMULAÇÃO DO PROBLEMA DE PROGRAMAÇÃO LINEAR (PPL)

Podemos formular o problema de programação linear (PPL) como segue:

$$\text{Minimizar } f = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (\text{I.2.1})$$

$$\text{Sujeito a } \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \end{aligned} \quad (\text{I.2.2})$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

De uma maneira mais compacta podemos ainda ter

$$\text{Minimizar } f = \sum_{j=1}^n c_j x_j$$

$$\text{Sujeito a } \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, 2, \dots, m \quad (\text{I.2.3})$$

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

ou

$$\text{Minimizar } f = c^T x$$

$$\text{Sujeito a } Ax = b$$

$$x \geq 0$$

Onde $c^T = (c_1, c_2, \dots, c_n)$, $x^T = (x_1, x_2, \dots, x_n)$
 $b^T = (b_1, b_2, \dots, b_m)$ e A uma matriz com m -linhas e n -colunas,
tendo seus elementos denotados por a_{ij} , e posto m . O conjunto
das variáveis x_1, x_2, \dots, x_n satisfazendo todas as restrições,
é chamado ponto viável ou vetor viável. Chamaremos de região
viável ao conjunto constituído de todos os pontos viáveis.

Uma solução básica para nosso problema será uma
solução obtida quando, fazendo $n-m$ variáveis iguais a zero, re-
solvemos em relação às variáveis restantes, sempre que o deter-
minante dos coeficientes a eles correspondentes seja diferente
de zero. As m variáveis se chamam variáveis básicas.

O problema de Programa Linear quando apresentado
como em 1.2.3, é dito estar na forma standard. Muitas vezes dese-
jamos maximizar ou minimizar uma função linear em presença de
restrição de igualdade e/ou inequações. Daí a necessidade de pas-
sar de um problema para outro equivalentemente.

1.3 - INEQUAÇÕES E EQUAÇÕES

Suponhamos que uma restrição seja dada por

$$\sum_{j=1}^n a_{ij}x_j \leq b_i \quad \left(\sum_{j=1}^n a_{ij}x_j \geq b_i \right).$$

Esta restrição pode ser transformada numa equação, bastando para tal somar (subtrair) uma variável escalar não negativa x_{n+i} ($x_{n+i} \geq 0$) dando-nos

$$\sum_{j=1}^n a_{ij}x_j + x_{n+i} = b_i \quad \left(\sum_{j=1}^n a_{ij}x_j - x_{n+i} = b_i \right).$$

Inversamente, uma equação pode ser transformada em duas inequações

$$\sum_{j=1}^n a_{ij}x_j \leq b_i \quad \text{e} \quad \sum_{j=1}^n a_{ij}x_j \geq b_i.$$

1.4 - NÃO NEGATIVIDADE DAS VARIÁVEIS

O método do Simplex é desenvolvido para solucionar PPL onde as variáveis são não negativas. Para tê-las nas condições desejadas, basta usar os artifícios que seguem:

1º) Se x_j não tem restrição de sinal, basta substituí-lo por $x_j' - x_j''$ onde $x_j' \geq 0$ e $x_j'' \geq 0$.

2º) Se $x_j \geq k_j$ fazer $x_j' = x_j - k_j$.

3º) Se $x_j \leq h_j$; $h_j \leq 0$, fazer $x_j' = h_j - x_j$

1.5 - PROBLEMAS DE MINIMIZAR E MAXIMIZAR

Utilizando a relação

$$\text{Mínimo } \sum_{j=1}^n c_j x_j = - \text{Máximo } \sum_{j=1}^n (-c_j) x_j$$

em que $f(x) = \sum_{j=1}^n c_j x_j$ representa a função objetivo a otimizar,

podemos sempre expressar o problema na forma de minimização (ou de maximização). Depois que a otimização do novo problema \bar{e} concluído, o valor da função objetivo do problema \bar{e} (-1) vezes o valor ótimo do problema novo.

A Tabela (1.1) que segue dá em resumo as diversas maneiras como as formas canônica e standard podem se apresentar.

		FORMAS	
		STANDARD	CANÔNICA
PROBLEMAS	MINIMIZAÇÃO	Minimize $\sum_{j=1}^n c_j x_j$ Sujeito a $\sum_{j=1}^n a_{ij} x_j = b_i$ $x_j \geq 0 \quad i=1 \dots m$ $j=1 \dots n$	Minimize $\sum c_j x_j$ Sujeito a $\sum_{j=1}^n a_{ij} x_j \geq b_i$ $x_j \geq 0 \quad i = 1 \dots m$ $j=1 \dots n$
	MAXIMIZAÇÃO	Maximize $\sum_{j=1}^n c_j x_j$ Sujeito a $\sum_{j=1}^n a_{ij} x_j = b_i$ $i=1 \dots m$ $x_j \geq 0, j=1 \dots n$	Maximize $\sum_{j=1}^n c_j x_j$ Sujeito a $\sum_{j=1}^n a_{ij} x_j \leq b_i$ $i=1 \dots m$ $x_j \geq 0, j=1 \dots n$

TABELA 1.1

1.6 - MÉTODO DO SIMPLEX

A =

Vejamos agora através da solução do problema que segue a idéia do método do Simplex.

b =

Minimizar $f = x_2 - 3x_3 + 2x_5$

c =

Sujeito a

$$\begin{aligned} x_1 + 3x_2 - x_3 + 2x_5 &= 21 \\ -2x_2 + 4x_3 + x_4 &= 36 \\ -4x_2 + 8x_5 + x_6 &= 30 \end{aligned} \quad (1.6.1)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, 6$$

Podemos ainda dar ao problema proposto, a seguinte disposição

$$\begin{aligned} 21 &= x_1 + 3x_2 - x_3 + 2x_5 \\ 36 &= -2x_2 + 4x_3 + x_4 \\ 30 &= -4x_2 + 8x_5 + x_6 \\ 0 &= -x_2 + 3x_3 - 2x_5 + f \\ x_j &\geq 0 \quad j = 1, 2, \dots, 6 \end{aligned} \quad (1.6.2)$$

onde desejamos encontrar o ótimo de f .

Note que o sistema (1.6.2) está na forma standard onde

$$A = \begin{bmatrix} 1 & 3 & -1 & 0 & 2 & 0 \\ 0 & -2 & 4 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & 8 & 1 \end{bmatrix} = [P_1 \ P_2 \ P_3 \ P_4 \ P_5 \ P_6]$$

$$b = [21 \ 36 \ 30]^T = P_0 \quad c^T = [0 \ 1 \ -3 \ 0 \ 2 \ 0]$$

e a solução básica \bar{e} é dada por:

$$\begin{aligned} f = 0 \quad x_1 = 21 \quad x_4 = 36 \quad x_6 = 30 \\ x_2 = x_3 = x_5 = 0 \end{aligned}$$

Vejamos agora o que ocorre quando x_3 varia. Com este propósito (I.6.2) é re-escrito como segue:

$$\begin{aligned} 21 + x_3 &= x_1 + 3x_2 + 2x_5 \\ 36 - 4x_3 &= -2x_2 + x_4 \\ 30 - 0x_3 &= -4x_2 - 8x_5 + x_6 \\ -3x_3 &= -x_2 - 2x_5 + f \end{aligned} \tag{I.6.3}$$

Variando unicamente x_3 e considerando $x_2 = x_5 = 0$, podemos por um momento suprimi-las, e (I.6.3) se transforma:

$$\begin{aligned} 21 + x_3 &= x_1 \\ 36 - 4x_3 &= x_4 \\ 30 - 0x_3 &= x_6 \\ -3x_3 &= f \end{aligned} \tag{I.6.4}$$

se analisando o sistema (I.6.4) temos que se x_3 cresce de zero para 1, na última equação de (I.6.3) o valor da função objetivo decresce de zero para -3. Portanto se estamos pesquisando o mínimo de f , uma boa idéia é crescer x_3 . Conseqüentemente uma generalização nos salta à vista, isto é, crescendo uma variável não básica a qual tem um coeficiente positivo na última equação de (I.6.2), o valor da função objetivo decresce. Além do mais, constatamos que aqueles termos positivos da função objetivo em (I.6.2) que têm maior valor proporcionam o maior decréscimo na função objetivo, por unidade de crescimento de uma variável não básica.

Este será o critério de entrada de um vetor na base, no método do Simplex. Assim uma escolha razoável para uma variável não básica, deverá ser entre aquelas que tem coeficientes positivos na linha da função objetivo, precisamente aquela variável que apresentar o maior valor deverá ser escolhida.

A 3.^a equação de (I.6.4) nos diz que se x_3 toma qualquer valor, x_6 será sempre positivo. O mesmo ocorre com x_7 para qualquer x_3 não negativo. Neste momento não podemos deixar passar despercebido que a análise feita sobre a 1.^a e 3.^a equações de (I.6.3) = (I.6.4) é sempre válida qualquer que seja x_3 .

Em situações semelhantes tal análise deverá ser omitida, isto é, quando em $P_3 = [x_{13}, x_{23}, x_{33}]^T \cdot [-1, 4, 0]^T$ x_{13} for não positiva. Conseqüentemente limitaremos nossa análise

se sobre aqueles $x_{i3} > 0$. A 2.^a equação de (I.6.4) restringe o valor de x_3 , isto \bar{e} , para termos $x_4 \geq 0$, x_3 não deverá crescer além de $(36/4) = 9$. Esta \bar{e} a idéia que norteia o critério de saída de um vetor da base, no método do Simplex.

Concluimos que o x_3 poderá crescer até 9, pois este valor de x_3 nos garante que as variáveis básicas continuarão não negativas.

Assim para $x_3 = 9$ temos:

$$\begin{aligned} 30 &= x_1 \\ 0 &= x_4 \\ 30 &= x_6 \\ -27 &= f \end{aligned} \tag{I.6.5}$$

Nesta solução $x_4 = 0$, pois foi a 1.^a variável básica que se anulou quando x_3 cresceu. Portanto os valores de (I.6.5), juntamente com $x_3 = 9$ e $x_2 = x_5 = 0$ dão outra solução para o sistema (I.6.2) preservando a não negatividade das variáveis. Observa-se que nesta solução a variável básica x_4 em (I.6.2) transforma-se em zero enquanto a não básica x_3 toma o valor 9. Esta solução pode evidentemente ser obtida a partir de (I.6.2) quando dividimos a 2.^a equação por 4 (coeficiente de x_3) e eliminamos x_3 das demais equações, Tal operação \bar{e} dita pivoteamento (4 \bar{e} o pivô). Concluindo o pivoteamento temos:

$$\begin{aligned}
 30 &= x_1 + \frac{10}{4} x_2 + \frac{1}{4} x_4 + 2x_5 \\
 9 &= -\frac{2}{4} x_2 + x_3 + \frac{1}{4} x_4 \\
 30 &= -4x_2 + 8x_5 + x_6 \\
 -27 &= \frac{2}{4} x_2 - \frac{3}{4} x_4 - 2x_5 + f
 \end{aligned} \tag{I.6.6}$$

Seguindo o raciocínio já descrito, uma vez que em (I.6.6) s̄o existe um termo positivo na linha da função objetivo, em $x_4 = x_5 = 0$ (I.6.6) \bar{e} equivalente a:

$$30 - \frac{10}{4} x_2 = x_1$$

$$9 + \frac{2}{4} x_2 = x_3$$

$$30 + 4x_2 = x_6$$

$$-27 - \frac{2}{4} x_2 = f$$

Como \bar{e} fácil de ver, para manter as variáveis não negativas, x_2 não pode exceder $30/(10/4) = 12$. Para $x_2 = 12$ te mos os seguintes valores para as variáveis

$$0 = x_1$$

$$15 = x_3$$

$$78 = x_6$$

$$-33 = f$$

$$x_4 = x_5 = 0$$

Se agora voltarmos a (1.6.6) e efetuarmos o pivoteamento (o coeficiente de x_2 na 1ª equação é o pivô) obtaremos:

$$12 = -\frac{4}{10} x_1 + x_2 + \frac{1}{10} x_4 + \frac{8}{10} x_5$$

$$15 = \frac{2}{10} x_1 + x_3 + \frac{3}{10} x_4 + \frac{4}{10} x_5$$

(I.6.7)

$$78 = \frac{16}{10} x_1 + \frac{4}{10} x_4 + \frac{112}{10} x_5 + x_6$$

$$-33 = \frac{-2}{10} x_1 - \frac{8}{10} x_4 - \frac{24}{10} x_5 + f$$

Como (I.6.7) não tem nenhum elemento positivo na última linha, concluimos que não é mais possível decrescer f . Assim o Ótimo será dado por

$$f = -33$$

$$x_2 = 12, x_3 = 15, x_6 = 78$$

$$x_1 = x_4 = x_5 = 0$$

Voltemos a (I.6.2) e observemos que se os coeficientes de x_3 nas 3 primeiras equações fossem todos não positivos (1.6.4) tomaria o seguinte aspecto:

$$\begin{aligned}
 21 + x_3 &= x_1 \\
 36 + 4x_3 &= x_4 \\
 30 + 0x_3 &= x_6 \\
 -3x_3 &= f
 \end{aligned}
 \tag{I.6.8}$$

Segue-se de (I.6.8) que, quando $x_3 \rightarrow \infty$, $f \rightarrow -\infty$ já que x_1 , x_4 e x_6 serão sempre positivas, qualquer que seja x_3 . Assim, teríamos um problema com solução ilimitada.

Como já percebemos, o algoritmo do Simplex tem por meta encontrar em cada passo uma nova solução viável cujo valor correspondente da função objetivo seja menor que o valor da função objetivo na solução anterior. Desta maneira prosseguimos até encontrar uma solução mínima, após um número finito de passos. Cada passo, pode ser dividido em três partes, e seguir:

- 1ª) Seleção de uma variável não básica a qual transforma-se em variável básica (critério de entrada de um vetor na base).
- 2ª) Seleção de uma variável básica a qual transforma-se em variável não básica (critério de saída de um vetor da base).
- 3ª) Transformação do sistema (pivoteamento).

Suponhamos agora que nosso problema (I.2.1)

(1.2.2) seja posto na forma:

$$\text{Minimizar } f = c^T x \quad (1.6.9)$$

$$\text{Sujeito a } [P_1, \dots, P_n] x = P_0 \quad (1.6.10)$$

$$x \geq 0 \quad (1.6.11)$$

onde $[P_1, \dots, P_n] = A$, $P_0 = b$ e

$$P_j = [x_{1j}, \dots, x_{mj}]^T \quad j = 0, 1, \dots, n$$

com $x_{i0} = b_i$ e $x_{ij} = a_{ij}$

Como estamos admitindo que A tem posto m , $\rho(A) = m$, então denotemos por B a matriz quadrada formada por m colunas linearmente independentes de A . Chamaremos de N a matriz formada pelas $n-m$ colunas de A . Então podemos escrever a matriz A como segue!

$$A = [B, N]$$

onde, sem perda de generalidade, após uma renumeração dos vetores colunas, se necessário, podemos dizer que

$$B = [P_1, \dots, P_m] \text{ e } N = [P_{m+1}, \dots, P_n]$$

Assim sendo, chamaremos de x_B o vetor coluna cujas componentes são as variáveis x_j associados as colunas de B . Isto nos leva à seguinte partição no conjunto de índices das variáveis: $I_B \cup I_N = \{1, 2, \dots, n\}$, $I_B \cap I_N = \emptyset$, dando-nos

$$x^T = [x_B^T, x_N^T] \text{ e } c^T = [c_B^T, c_N^T].$$

Agora o problema (169) - (I.6.11) pode ser re-escrito:

$$\text{Minimizar } f = c_B^T x_B + c_N^T x_N \quad (\text{I.6.12})$$

$$\text{Sujeito a } Bx_B + Nx_N = P_0 \quad (\text{I.6.13})$$

$$\text{ou } [P_1, \dots, P_m]x_B + [P_{m+1}, \dots, P_n]x_N = P_0 = b$$

$$x_B \geq 0 \text{ e } x_N \geq 0 \quad (\text{I.6.14})$$

Para não nos alongarmos muito na teoria, vamos admitir que $B = I$, o que nos leva ao seguinte quadro.

PRIMEIRO SIMPLEX: PASSO INICIAL

i	BASE	C	x_0	C_1	C_2	C_ℓ	C_m	C_{m+1}	C_j	C_k	C_n
				P_1	P_2	P_ℓ	P_m	P_{m+1}	P_j	P_k	P_n
1	ρ_1	C_1	x_{10}	1	0	0	0	$x_{1,m+1}$	$x_{1,j}$	$x_{1,k}$	$x_{1,n}$
2	ρ_2	C_2	x_{20}	0	1	0	0	$x_{2,m+1}$	$x_{2,j}$	$x_{2,k}$	$x_{2,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
ℓ	ρ_ℓ	C_ℓ	$x_{\ell 0}$	0	0	1	0	$x_{\ell,m+1}$	$x_{\ell,j}$	$x_{\ell,k}$	$x_{\ell,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
m	ρ_m	C_m	x_{m0}	0	0	0	1	$x_{m,m+1}$	$x_{m,j}$	$x_{m,k}$	$x_{m,n}$
m+1			Z_0	0	0	0	0	$Z_{m+1} - C_{m+1}$	$Z_j - C_j$	$Z_k - C_k$	$Z_n - C_n$

Quadro I.2

Observe que a $(m+1)$ -ésima linha do Quadro I.2 foi reservado para os elementos z_0 e $z_j - c_j$ onde $z_0 = \sum_{i=1}^m c_i x_{i0}$,

$$z_j = \sum_{i=1}^m c_i x_{ij}, \quad j = 1, 2, \dots, n.$$

Voltando a equação (I.6.6), observamos que os coeficientes x_j na última linha são os $z_j - c_j$, e lá nos escolhemos o vetor a entrar na base, justamente aquele que corresponde ao maior coeficiente de x_j na última linha. Aqui procedemos da mesma forma, escolheremos aquela variável que corresponde a $\max(z_j - c_j)$, para entrar na base. Este é precisamente o critério de entrada de um vetor na base. Se existe empate, devemos escolher o vetor com menor índice j .

Suponhamos, por exemplo que

$$\max(z_j - c_j) = z_k - c_k > 0$$

Então o vetor P_k será introduzido na base. Lembre-se que na solução do problema (I.6.1), para encontrar a variável (vetor) que deixa a base, nós procuramos o mínimo dos quocientes x_{j0}/x_{jk} para $x_{jk} > 0$.

Admitamos aqui que

$$\min_i \left(\frac{x_{i0}}{x_{ik}} \mid x_{ik} > 0 \right) = \frac{x_{\ell 0}}{x_{\ell k}}$$

Então P_ℓ deixará a base, entrando o vetor P_k em seu lugar.

Se todos os $x_{ik} \leq 0$, encontramos como em (I.6.8), onde o valor da função objetivo pode tornar-se arbitrariamente negativo. Se tal não ocorre fazemos o pivoteamento em torno do pivo $x_{\ell k}$ e obtemos o Quadro 1.3. que segue.

QUADRO DO SIMPLEX: 29 PASSO

i	BASES	C	P ₀	C ₁	C ₂	C _ℓ	C _m	C _{m+1}	C _j	C _k	C
				P ₁	P ₂			P _ℓ			
1	P ₁	C ₁	x ₁₀ ⁱ	1	0	x _{1ℓ} ⁱ	0	x _{1,m+1} ⁱ	x _{1j} ⁱ	0	x _{1n} ⁱ
2	P ₂	C ₂	x ₂₀ ⁱ	0	1	x _{2ℓ} ⁱ	0	x _{2,m+1} ⁱ	x _{2j} ⁱ	0	x _{2n} ⁱ
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ℓ	P _k	C _k	x _{k0} ⁱ	0	0	x _{ℓℓ} ⁱ	⋮	x _{ℓ,m+1} ⁱ	x _{ℓ,j}	1	x _{ℓ,n} ⁱ
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
m	P _m	C _m	x _{m0} ⁱ	0	0	x _{mℓ} ⁱ	1	x _{m,m+1} ⁱ	⋮	0	x _{m,n} ⁱ
m+1			z ₀ ⁱ	0	0	Z _ℓ ⁱ -C _ℓ	0	Z _{m+1} ⁱ -C _{m+1}	Z _j ⁱ -C _j	0	Z _n ⁱ -C _n

QUADRO I.3

Prosseguimos com este raciocínio até que $z_j - c_j \leq 0$ para todo j .

Podemos sintetizar o procedimento nas 3 regras que seguem:

REGRA 1: Selecionar, como variável não básica para entrar na base, uma com maior coeficiente positivo na linha da função objetivo. Se esta linha não contém nenhum coeficiente positivo, uma solução Ótima foi encontrada. Se o j -ésimo elemento da $(m+1)$ -ésima linha é indicada por $z_j - c_j$, a nova variável básica é determinada por

$$\text{Max } (z_j - c_j) = z_k - c_k \\ j=1, \dots, n$$

REGRA 2: Selecionar como variável para deixar a base, uma correspondente à menor razão entre o valor da variável básica e o correspondente coeficiente da nova variável básica nas linhas nos quais estes coeficientes são positivos. Se a nova variável básica não tem coeficientes positivos, o problema tem uma solução infinita.

Se x_{i0} representar o valor da variável básica na i -ésima linha e x_{ik} representar coeficiente correspondente da nova variável básica x_k , a variável básica a deixar a base é a associada com

$$\text{Min } \left(\frac{x_{i0}}{x_{ik}}, x_{ik} > 0 \right)$$

REGRA 3: Transforme a tabela tomando o coeficiente da nova variável básica na linha da variável que deixa a base, com o pivô.

Para complementar, citaremos dois teoremas, os quais dão a comprovação do que já foi dito.

TEOREMA 1: Se para qualquer j fixo, $z_j - c_j > 0$, então podemos construir um conjunto de soluções viáveis tal que $z < z_0$ para qualquer elemento do conjunto, onde o limite inferior z pode ser finito ou infinito.

Se o limite inferior \bar{z} é finito, podemos construir uma sequência de soluções viáveis com m variáveis, tendo a função objetivo menor valor que o precedente. Como o número máximo de soluções básicas de $Ax = b$ é menor ou igual a $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ fatalmente encontraremos a solução ótima.

Se o limite inferior \bar{z} é infinito, podemos construir uma nova solução viável com $m+1$ variáveis positivas cujo valor da função objetivo pode fazer-se arbitrariamente negativo.

TEOREMA 2: Se para qualquer solução básica viável $x = (x_1, \dots, x_m)^T$ as condições $z_j - c_j \leq 0$ são satisfeitas para $j = 1, 2, \dots, n$, então o problema admite um programa (Ótimo) mínimo.

Agora voltando as equações (I.6.12)-(I.6.14) temos que:

$$Bx_B = b - Nx_N$$

Como $A = [B, N]$ tem posto completo, B^{-1} existe. Assim,

$$x_B = B^{-1}b - B^{-1}Nx_N \quad (\text{I.6.15})$$

Caso façamos $x_N = 0$ teremos um valor para o vetor

$$x_B = \bar{x}_B = B^{-1}b, \quad (\text{I.6.16})$$

que nós já designamos ser uma solução básica para o sistema (I.6.13), e se tivermos $\bar{x}_B \geq 0$ então teremos uma solução básica viável. Podemos ainda reescrever (I.6.12) como segue:

$$f = c_B^T [B^{-1}b - B^{-1}Nx_N] + c_N^T x_N$$

$$f = c_B^T B^{-1}b - (c_B^T B^{-1}N - c_N^T)x_N$$

Por (I.6.14):

$$f = c_B^T B^{-1} b - [c_B^T B^{-1} (P_{m+1} \dots P_n) - (c_{m+1} \dots c_n)] x_N$$

$$f = c_B^T B^{-1} b - \sum_{j \in I_N} (c_B^T B^{-1} P_j - c_j) x_j$$

Fazendo

$$z_j = c_B^T B^{-1} P_j \text{ temos que}$$

$$z_j - c_j = c_B^T B^{-1} P_j - c_j \quad (I.6.17)$$

o que nos dá

$$f = c_B^T B^{-1} b - \sum_{j \in I_N} (z_j - c_j) x_j$$

O sistema (I.6.12) - (I.6.14) toma o seguinte aspecto

$$f = c_B^T B^{-1} b - \sum_{j \in I_N} (z_j - c_j) x_j \quad (I.6.18)$$

$$x_B = B^{-1} b - \sum_{j \in I_N} y_j x_j \quad (I.6.19)$$

$$x_B \geq 0, \quad e \cdot x_j = 0, \quad j \in I_N \quad (I.6.20)$$

onde

$$y_j = B^{-1} P_j$$

1.7 - BASE ARTIFICIAL

Como nem sempre dispomos de uma base canônica ($B = I$) para iniciar o método do Simplex, criou-se para solucionar tal impasse a técnica da base artificial. Então o problema proposto será aumentado para

$$\text{Minimizar } f = [c^T \quad w^T] \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\text{Sujeito a } [A \quad I] \begin{bmatrix} x \\ y \end{bmatrix} = b$$

$$x \geq 0, y \geq 0$$

onde $y^T = (x_{n+1}, \dots, x_{n+m})$, $w^T = w(1, 1, \dots, 1)$,

com w tão grande quanto se queira. O desenvolvimento do Simplex com base artificial é similar ao já apresentado, com uma inovação. Como

$$z_j - c_j = w \sum_{i=1}^m x_{ij} - c_j \quad (I.7.1)$$

então criamos inicialmente uma linha no Quadro I.2, a $(m+2)$ -ésima linha, na qual colocaremos correspondendo a cada coluna j , ($j = 1, \dots, n$) o coeficiente de w , $\sum_{i=1}^m x_{ij}$ da equação (I.7.1). Na $(m+1)$ -ésima linha colocaremos para cada coluna j ($j = 1, \dots, n$) $-c_j$. O método inicia-se eliminando um vetor artificial da base, o qual jamais deverá ser escolhido para retornar a base.

Continuamos selecionando um vetor para entrar na base, usando os elementos da $(m+2)$ -ésima linha até que:

- 1) Todos os vetores artificiais tenham sido eliminados da base. Prosseguimos com o método regular do Simplex.
- 2) Que não exista, na $(m-2)$ -ésima linha elementos positivos. Para maiores detalhes veja | 15 |.

1.8 - VARIÁVEIS ESCALARES

Suponhamos agora que desejamos

$$\text{Maximizar } f = c^T x \quad (1.8.1)$$

$$\text{Sujeito a } Ax \leq b \quad (1.8.2)$$

$$x \geq 0$$

Aqui, como já chamamos à obtenção em (1.3), acrescentamos as variáveis escalares e (1.8.1) - (1.8.2) transforma-se

$$\text{Maximizar } f = c^{*T} x^* \quad (1.8.3)$$

$$\text{Sujeito a } A^* x^* = b \quad (1.8.4)$$

onde

$$x^{*T} = [x^T, y^T], \quad c^{*T} = [c^T, 0] \quad A^* = [A, I]$$

Agora usamos (I.6) para solucionar o problema. Quando o problema original envolver a forma $Ax \geq b$, $b \geq 0$, usamos as variáveis escalares e pelo menos uma variável artificial. Veja | 15 |.

1.9 - DUALIDADE

Dado um problema de Programação Linear, o qual denotamos de Primal, a ele podemos fazer corresponder um outro problema de otimização, o problema Dual.

Eles são interligados, de tal forma que a solução ótima de qualquer um deles, nos proporciona informações a respeito da solução ótima do outro.

1.9.1 - Definição Geral dos Problemas Duais

O problema Primal: determinar um vetor coluna $x = (x_1, x_2, \dots, x_N)^T$ tal que

$$\text{Minimize } f = \sum_{j \in N} c_j x_j \quad (\text{I.9.1})$$

Sujeito a

$$\sum_{j \in N} a_{ij} x_j \geq b_i \quad i \in M_1$$

$$\sum_{j \in N} a_{ij} x_j = b_i \quad i \in \bar{M}_1 \quad (I.9.2)$$

$$x_j \geq 0 \quad j \in N_1$$

$$x_j \text{ qualquer } j \in \bar{N}_1$$

$$M = M_1 \cup \bar{M}_1, N = N_1 \cup \bar{N}_1$$

O problema Dual: determinar um vetor linha

$W = (w_1, w_2, \dots, w_m)$ tal que

$$\text{Maximize } g = \sum_{i \in M} (b_i) w_i \quad (I.9.3)$$

Sujeito a

$$w_i \geq 0 \quad i \in M_1$$

$$w_i \text{ qualquer } i \in \bar{M}_1$$

$$\sum_{i \in M} (a_{ij}) w_i \leq c_j \quad j \in N_1 \quad (I.9.4)$$

$$\sum_{i \in M} (a_{ij}) w_i = c_j \quad j \in \bar{N}_1$$

$$M = M_1 \cup \bar{M}_1, N = N_1 \cup \bar{N}_1$$

Então o teorema da dualidade afirma

TEOREMA: O mínimo de (I.9.1) sujeito a (I.9.2) é igual ao máximo de (I.9.3) sujeito a (I.9.4).

CAPÍTULO IUMA NOVA ALTERNATIVA SOBRE A SOLUÇÃO
DOS PROBLEMAS DE PROGRAMAÇÃO LINEAR11.1 - DEFINIÇÃO DO PROBLEMA

Tentamos neste trabalho estabelecer algoritmos para resolver problemas de programação linear de uma maneira não convencional.

11.1.1 - Dados

Seja o problema de programação linear:

Minimize $Z = c^T x$, sujeito a

$$Ax = b$$

$$x \geq 0 \quad \text{onde}$$

$c^T \in R^n$, $x \in R^n$, A matriz $m \times n$,

$A = (a_{ij}) = (a^1, a^2, \dots, a^n)$, $a^j \in R^m$, $j = 1, 2, \dots, n$

$b \in R^m$

$*c^T = (c_1, c_2, \dots, c_n)$, $x^T = (x_1, \dots, x_n)$ e $b^T = (b_1, \dots, b_m)$

O algoritmo a ser desenvolvido tem como objetivo primordial detectar em seu primeiro passo um vetor coluna de A , o qual deverá ter sua presença assegurada na base final. Prosseguindo, o algoritmo no 2º passo nos dará mais um vetor da base

final e assim o fará em seus passos subsequentes, até termos no m -ésimo passo uma base completa, a qual será a base ótima.

11.2 - PSEUBOINVERSA, OPERADORES PROJEÇÃO E SEU COMPLEMENTO ORTOGONAL

No decorrer deste trabalho denotaremos por A_j a matriz composta de j vetores colunas de A , e com A_{j-1}^i queremos indicar a matriz obtida a partir de A_j quando dela omitimos a i -ésima coluna. A pseudo-inversa de A_j será denotada por A_j^+ . $B_j = A_j A_j^+$ e $P_j = I - B_j$ indicarão respectivamente os operadores projeção e seu complemento ortogonal.

Sejam agora, $V \perp a^i$, $V \perp a^j$, $a^i \neq \lambda a^j$ tal que

$$b = V + \alpha x_i a^i + \alpha x_j a^j \quad (\text{II.2.1})$$

Mostremos então que:

$$V = P_2 b$$

$$B_2 b = A_2 A_2^+ b = \alpha x_i a^i + \alpha x_j a^j$$

onde $A_2^+ = (A_2^T A_2)^{-1} A_2^T$ é a pseudo-inversa de A_2 quando suas colunas são linearmente independentes.

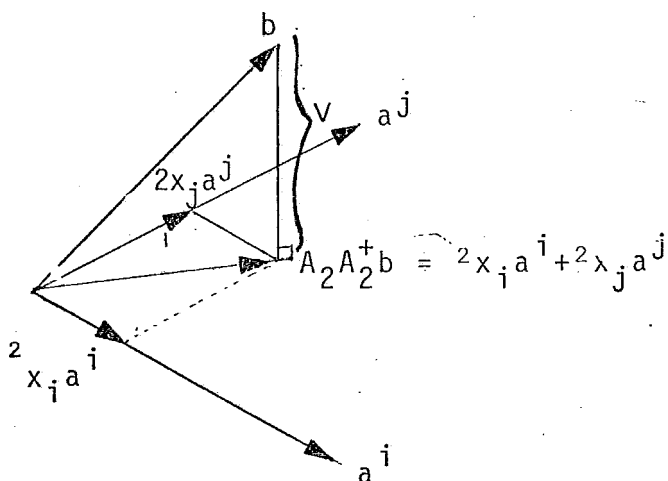


Fig. II.1

Multiplicando (II.2.1) pelos vetores a^i e a^j obtemos respectivamente:

$$\langle b, a^i \rangle = \langle a^i, a^i \rangle \cdot 2x_i + \langle a^i, a^j \rangle \cdot 2x_j$$

$$\langle b, a^j \rangle = \langle a^j, a^i \rangle \cdot 2x_i + \langle a^j, a^j \rangle \cdot 2x_j$$

Pondo na forma matricial

$$\begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix} = \begin{bmatrix} \langle a^i, a^i \rangle & \langle a^i, a^j \rangle \\ \langle a^j, a^i \rangle & \langle a^j, a^j \rangle \end{bmatrix} \begin{bmatrix} 2x_i \\ 2x_j \end{bmatrix}$$

Tomando a inversa

$$\begin{bmatrix} 2x_i \\ 2x_j \end{bmatrix} = \begin{bmatrix} \langle a^i, a^i \rangle & \langle a^i, a^j \rangle \\ \langle a^j, a^i \rangle & \langle a^j, a^j \rangle \end{bmatrix}^{-1} \begin{bmatrix} (a^i)^T \\ (a^j)^T \end{bmatrix} b$$

isto é

$$\begin{bmatrix} {}^2x_i \\ \vdots \\ {}^2x_j \end{bmatrix} = [A_2^T \quad A_2]^{-1} A_2^T b = A_2^+ b \quad (\text{II.2.2})$$

Então de(II.2.1) tiramos

$$V = b - A_2 \begin{bmatrix} {}^2x_i \\ \vdots \\ {}^2x_j \end{bmatrix} \text{ e por (II.2.2)}$$

$$V = b - A_2 A_2^+ b = [I - A_2 A_2^+] b$$

Fazendo $P_2 = [I - A_2 A_2^+]$ temos:

$$V = P_2 b \text{ e novamente por (II.2.1) e (II.2.2)}$$

$$B_2 b = A_2 A_2^+ b = {}^2x_i a^i + {}^2x_j a^j$$

Portanto a demonstração do teorema 1, que segue, torna-se Óbvia.

Teorema 1: Para algum vetor b , $B_k b$ é a projeção de b sobre o subespaço determinado por $\{a^1, \dots, a^k\}$ e $P_k b$ é a projeção de b sobre o complemento ortogonal do subespaço determinado por $\{a^1, \dots, a^k\}$.

(Para maiores esclarecimentos, veja corolário 3.5, página 20 de [1]).

11.3 - CONFIGURAÇÃO DE A_k^+ EM TERMOS DE $A_{k-1}^+ \frac{e_k P}{1}$

Mostraremos agora que se as colunas de A_k são LI, então

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - \frac{A_{k-1}^+ a^k (a^k)^T P_{k-1}}{(a^k)^T P_{k-1} a^k} \\ \frac{(a^k)^T P_{k-1}}{(a^k)^T P_{k-1} a^k} \end{bmatrix} \quad (\text{II.3.1})$$

Inicialmente desenvolveremos um raciocínio para $k = 3$, haja visto que sua generalização é imediata. Seguindo a mesma trilha que executamos para decompor b em sua projeção ortogonal e seu complemento ortogonal, obtemos:

$$b = v + {}^3x_i a^i + {}^3x_j a^j + {}^3x_k a^k$$

Portanto:

$$\begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \\ \langle b, a^k \rangle \end{bmatrix} = \begin{bmatrix} \langle a^i, a^i \rangle & \langle a^i, a^j \rangle & \langle a^i, a^k \rangle \\ \langle a^j, a^i \rangle & \langle a^j, a^j \rangle & \langle a^j, a^k \rangle \\ \langle a^k, a^i \rangle & \langle a^k, a^j \rangle & \langle a^k, a^k \rangle \end{bmatrix} \begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix}$$

$$\begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \\ \langle b, a^k \rangle \end{bmatrix} = \begin{bmatrix} A_2^T A_2 & B_3 \\ B_3^T & D_3 \end{bmatrix} \begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} \quad (\text{II.3.2})$$

onde

$$A_2^T A_2 = \begin{bmatrix} \langle a^i, a^i \rangle & \langle a^i, a^j \rangle \\ \langle a^j, a^i \rangle & \langle a^j, a^j \rangle \end{bmatrix}$$

$$B_3 = A_2^T a^k \rightarrow B_3^T = [a^k]^T A_2 \quad e$$

$$D_3 = \langle a^k, a^k \rangle$$

De (II.3.2) tiramos

$$\begin{bmatrix} A_2^T A_2 & \vdots & B_3 \\ 0 & \ddots & D_3 - B_3^T (A_2^T A_2)^{-1} B_3 \end{bmatrix} \begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \\ \langle b, a^k \rangle - B_3^T [A_2^T A_2]^{-1} \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix} \end{bmatrix} \rightarrow$$

$$\begin{cases} [A_2^T A_2] \begin{bmatrix} {}^3x_i \\ {}^3x_j \end{bmatrix} + B_3 {}^3x_k = \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix} \end{cases} \quad (\text{II.3.3})$$

$$\begin{cases} [D_3 - B_3^T (A_2^T A_2)^{-1} B_3] {}^3x_k = \langle b, a^k \rangle - B_3^T [A_2^T A_2]^{-1} \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix} \end{cases} \quad (\text{II.3.4})$$

De (II.3.3) obtemos:

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \end{bmatrix} = (A_2^T A_2)^{-1} \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix} - [A_2^T A_2]^{-1} B_3 {}^3x_k$$

Explicitando o valor de 3x_k em (II.3.4) obtemos:

$${}^3x_k = \frac{\begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix}}{D_3 - B_3^T [A_2^T A_2]^{-1} B_3} \quad (\text{II.3.5})$$

Decorre de (II.3.5) que

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \end{bmatrix} = \frac{\begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \end{bmatrix}}{D_3 - B_3^T [A_2^T A_2]^{-1} B_3} \begin{bmatrix} [A_2^T A_2]^{-1} z_3 [b, a^k] - \omega_3^T [A_2^T A_2]^{-1} \\ [A_2^T A_2]^{-1} z_3 [b, a^k] - \omega_3^T [A_2^T A_2]^{-1} \end{bmatrix} \quad (\text{II.3.6})$$

Agora por (II.3.5) e (II.3.6) temos:

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} [A_2^T A_2]^{-1} + \frac{[A_2^T A_2]^{-1} B_3 B_3^T [A_2^T A_2]^{-1}}{\Delta_3} \\ - \frac{B_3^T [A_2^T A_2]^{-1}}{\Delta_3} \\ \frac{1}{\Delta_3} \end{bmatrix} \begin{bmatrix} \langle b, a^i \rangle \\ \langle b, a^j \rangle \\ \langle b, a^k \rangle \end{bmatrix}$$

onde $\Delta_3 = D_3 - B_3^T [A_2^T A_2]^{-1} B_3$. Prosseguindo

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} [A_2^T A_2]^{-1} + \frac{[A_2^T A_2]^{-1} A_2^T a^k (a^k)^T A_2 [A_2^T A_2]^{-1}}{\Delta_3} \\ - \frac{(a^k)^T A_2 [A_2^T A_2]^{-1}}{\Delta_3} \\ \frac{1}{\Delta_3} \end{bmatrix} \begin{bmatrix} A_2^T \\ (a^k)^T \end{bmatrix} \begin{bmatrix} [b] \\ (a^k)^T \end{bmatrix}$$

$$\begin{bmatrix} 3 \times i \\ 3 \times j \\ 3 \times k \end{bmatrix} = \begin{bmatrix} (A_2^T A_2)^{-1} A_2^T + \frac{[A_2^T A_2]^{-1} A_2^T a^k (a^k)^T A_2 [A_2^T A_2]^{-1} A_2^T}{\Delta_3} - \frac{(A_2^T A_2)^{-1} A_2^T a^k (a^k)^T}{\Delta_3} \\ \frac{(a^k)^T - (a^k)^T A_2 [A_2^T A_2]^{-1} A_2^T}{\Delta_3} \end{bmatrix} [b]$$

$$\begin{bmatrix} 3 \times i \\ 3 \times j \\ 3 \times k \end{bmatrix} = \begin{bmatrix} A_2^+ + \frac{A_2^+ a^k (a^k)^T A_2 A_2^+ - A_2^+ a^k (a^k)^T}{\Delta_3} \\ \frac{(a^k)^T [I - A_2 A_2^+]}{\Delta_3} \end{bmatrix} [b]$$

$$\begin{bmatrix} 3 \times i \\ 3 \times j \\ 3 \times k \end{bmatrix} = \begin{bmatrix} A_2^+ - \frac{A_2^+ a^k (a^k)^T [I - A_2 A_2^+]}{\Delta_3} \\ \frac{(a^k)^T P_2}{\Delta_3} \end{bmatrix} [b] \quad (II.3.7)$$

Mas

$$\Delta_3 = D_3 - B_3^T [A_2^T A_2]^{-1} B_3 = (a^k)^T a^k - (a^k)^T A_2 [A_2^T A_2]^{-1} A_2^T a^k$$

$$\Delta_3 = (a^k)^T [I - A_2 A_2^+] a^k = \langle a^k, P_2 a^k \rangle$$

Podemos agora escrever (II.3.7) como segue: (para

$k = 3$)

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} A_{3-1}^+ - \frac{A_{3-1}^+ a^3 (a^3)^T P_{3-1}}{(a^3)^T P_{3-1} a^3} \\ \\ \frac{(a^3)^T P_{3-1}}{(a^3)^T P_{3-1} a^3} \end{bmatrix} \quad [b] \quad (II.3.8)$$

O que nos leva a generalizar:

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - \frac{A_{k-1}^+ a^k (a^k)^T P_{k-1}}{\langle a^k, P_{k-1}, a^k \rangle} \\ \\ \frac{(a^k)^T P_{k-1}}{\langle a^k, P_{k-1}, a^k \rangle} \end{bmatrix} \quad (II.3.9)$$

Podemos ainda concluir

$${}^k x_k = \frac{\langle a^k, P_{k-1} b \rangle}{\langle a^k, P_{k-1} a^k \rangle} \quad (II.3.10)$$

O índice superior e a esquerda indica o passo onde o nosso futuro algoritmo está trabalhando.

Observe que $P_2 P_2 = P_2$ e $P_2^T = P_2$, então

$$\Delta_3 = \langle a^k, P_2 a^k \rangle = (a^k)^T P_2 P_2 a^k = (a^k)^T P_2^T P_2 a^k$$

$$\Delta_3 = [P_2 a^k]^T [P_2 a^k] = ||P_2 a^k||^2$$

o que nos leva a afirmar:

$$\Delta_k = \langle a^k, P_{k-1} a^k \rangle = \|P_{k-1} a^k\|^2$$

Notamos ainda que

$$P_{k-1} a^k = (I - A_{k-1} A_{k-1}^+) a^k = a^k - A_{k-1} A_{k-1}^+ a^k = 0 \leftrightarrow$$

$$\leftrightarrow A_{k-1} A_{k-1}^+ a^k = a^k$$

Sabemos que $A_{k-1} A_{k-1}^+ a^k$ é a projeção de a^k sobre $R(A_{k-1} A_{k-1}^+)$. Assim se $a^k \in R(A_{k-1} A_{k-1}^+)$ então $A_{k-1} A_{k-1}^+ a^k = a^k$ o que equivale a dizer que a^k é combinação linear dos a^i , $i = 1, \dots, k-1$, o que contraria (no nosso caso) a hipótese dos a^i , $i = 1, \dots, k$ serem LI. Concluimos assim que nosso $\Delta_k = \|P_{k-1} a^k\|^2$ nunca será nulo ($\Delta_k > 0$).

II.4 - NOVA REPRESENTAÇÃO DAS VARIÁVEIS j_{x_i}

A partir de (II.3.8) - (II.3.10), para $k = 2$ temos:

$$\begin{bmatrix} {}^2x_1 \\ {}^2x_2 \end{bmatrix} = \begin{bmatrix} A_1^+ - \frac{A_1^+ a^2 (a^2)^T P_1}{\langle a^2, P_1 a^2 \rangle} \\ \frac{(a^2)^T P_1}{\langle a^2, P_1 a^2 \rangle} \end{bmatrix} [b] = \begin{bmatrix} A_1^+ b - A_1^+ a^2 \frac{\langle a^2, P_1 b \rangle}{\langle a^2, P_1 a^2 \rangle} \\ \frac{\langle a^2, P_1 b \rangle}{\langle a^2, P_1 a^2 \rangle} \end{bmatrix}$$

$${}^2x_1 = {}^1x_1 - {}^2x_2 A_1^+ a^2$$

$${}^2x_2 = \frac{\langle (a^2), P_1 b \rangle}{\langle a^2, P_1 a^2 \rangle}$$

No passo $k = 4$, se $A_4 = [a^1, a^2, a^3, a^4]$ temos:

$$\begin{bmatrix} {}^4x_1 \\ {}^4x_2 \\ {}^4x_3 \end{bmatrix} = \begin{bmatrix} \frac{\langle a_1^2, b \rangle}{|a^1|^2} - {}^2x_2 A_1^+ a^2 \\ \frac{\langle a^2, P_1 b \rangle}{\langle a^2, P_1 a^2 \rangle} \\ {}^3x_3 \end{bmatrix} - \begin{bmatrix} {}^3x_3 A_2^+ a^3 \\ {}^4x_4 A_3^+ a^4 \end{bmatrix}$$

$$\begin{bmatrix} {}^4x_1 \\ {}^4x_2 \\ {}^4x_3 \end{bmatrix} = I_3 \begin{bmatrix} {}^1x_1 \\ {}^2x_2 \\ {}^3x_3 \end{bmatrix} - \begin{bmatrix} A_1^+ a^2 & A_2^+ a^3 & A_3^+ a^4 \\ 0 & & \\ 0 & 0 & \end{bmatrix} \begin{bmatrix} {}^2x_2 \\ {}^3x_3 \\ {}^4x_4 \end{bmatrix}$$

Portanto no passo m teremos:

$$\begin{bmatrix} {}^m x_1 \\ {}^m x_2 \\ \vdots \\ {}^m x_{m-1} \end{bmatrix} = I_{m-1} \begin{bmatrix} {}^1 x_1 \\ {}^2 x_2 \\ \vdots \\ {}^{m-1} x_{m-1} \end{bmatrix} - \begin{bmatrix} A_1^+ a & A_2^+ a & \dots & A_{m-2}^+ a^{m-1} & A_{m-1}^+ a^m \\ 0 & 0 & & & \\ 0 & 0 & & & \\ \vdots & \vdots & & & \\ 0 & 0 & & 0 & \end{bmatrix} \begin{bmatrix} {}^2 x_2 \\ {}^3 x_3 \\ \vdots \\ {}^{m-1} x_{m-1} \\ {}^m x_m \end{bmatrix}$$

(II.4.1)

$${}^m x_m = \frac{\langle a^m, P_{m-1} b \rangle}{\langle a^m, P_{m-1} a^m \rangle} \quad (\text{II.4.1}')$$

Consideremos a seguir que:

$$A_1 = [a^1], A_2 = [a^1, a^2], A_3 = [a^1, a^2, a^3], A_4 = [a^1, a^2, a^2, a^4]$$

$$A_1^+ = {}^1 C_1^T, A_2^+ = \begin{bmatrix} {}^{12} C_1^T \\ {}^{12} C_2^T \end{bmatrix}, A_3^+ = \begin{bmatrix} {}^{123} C_1^T \\ {}^{123} C_2^T \\ {}^{123} C_3^T \end{bmatrix}$$

Então (II.4.1) para $m = 4$, toma o aspecto que segue:

$$\begin{bmatrix} {}^4 x_1 \\ {}^4 x_2 \\ {}^4 x_3 \end{bmatrix} = \begin{bmatrix} {}^1 x_1 \\ {}^2 x_2 \\ {}^3 x_3 \end{bmatrix} - \begin{bmatrix} {}^1 C_1^T a^2 & {}^{12} C_1^T a^3 & {}^{123} C_1^T a^4 \\ 0 & {}^{12} C_2^T a^3 & {}^{123} C_2^T a^4 \\ 0 & 0 & {}^{123} C_3^T a^4 \end{bmatrix} \begin{bmatrix} {}^2 x_2 \\ {}^3 x_3 \\ {}^4 x_4 \end{bmatrix}$$

$$\begin{bmatrix} {}^4 x_1 \\ {}^4 x_2 \\ {}^4 x_3 \end{bmatrix} = \begin{bmatrix} [({}^1 x_1 - {}^2 x_2 {}^1 C_1^T a^2) - {}^3 x_3 {}^{12} C_1^T a^3] - {}^4 x_4 {}^{123} C_1^T a^4 \\ [{}^2 x_2 - {}^3 x_3 {}^{12} C_2^T a^3] - {}^4 x_4 {}^{123} C_2^T a^4 \\ {}^3 x_3 - {}^4 x_4 {}^{123} C_3^T a^4 \end{bmatrix} = \begin{bmatrix} {}^3 x_1 \\ {}^3 x_2 \\ {}^3 x_3 \end{bmatrix} - {}^4 x_4 \begin{bmatrix} {}^{123} C_1^T a^4 \\ {}^{123} C_2^T a^4 \\ {}^{123} C_3^T a^4 \end{bmatrix}$$

Como
$$\begin{bmatrix} {}^3x_1 \\ \vdots \\ {}^3x_2 \\ \vdots \\ {}^3x_3 \end{bmatrix} = A_3^+ b \quad \text{então}$$

$$\begin{bmatrix} {}^4x_1 \\ \vdots \\ {}^4x_2 \\ \vdots \\ {}^4x_3 \end{bmatrix} = A_3^+ b - {}^4x_4 A_3^+ a^4 = A_3^+ [b - {}^4x_4 a^4]$$

Generalizando temos:

$${}^jx^1 = {}^{j-1}x + {}^jx_j A_{j-1}^+ a^j \quad (II.4.2)$$

$$j = 2, \dots, m$$

ou

$${}^jx^1 = A_{j-1}^+ [b - {}^jx_j a^j] \quad (II.4.3)$$

$$j = 2, \dots, m$$

Onde com ${}^jx^1$ queremos indicar o vetor jx , no j -ésimo passo, quando dele omitimos a última componente; isto é:

$${}^jx^1 = [{}^jx_1 \quad {}^jx_2 \quad \dots \quad {}^jx_{j-2} \quad {}^jx_{j-1}]^T$$

Fixaremos agora alguns conceitos e notações, para que possamos enunciar os teoremas que se tornam necessários.

II.5 - CONE C

$C[a^1, \dots, a^n]$ é o cone gerado pelo conjunto de vetores colunas $\{a^1, \dots, a^n\}$

$C[A]$ é o cone gerado pelos vetores colunas da matriz A.

II.5.a - Raio Vetor Gerado por v

Dado v o raio vetor r é o conjunto

$$r = \{x \mid x = Av \quad v \neq 0, \text{ e } A \geq 0\}$$

II.5.b - Hiperplano em R^n

$$H = \{x \mid cx = Z, \quad c \neq 0\}$$

Se $Z = 0$ o hiperplano passa pela origem.

II.5.c - Um hiperplano H Divide R^n em Semi-Espaços

i) Aberto $S_1 = \{x \mid cx < Z\}$ $S_2 = \{x \mid cx > Z\}$

ii) Fechado $S_3 = \{x \mid cx \leq Z\}$ $S_4 = \{x \mid cx \geq Z\}$

II.5.d - Poliedro

Um subconjunto P de um espaço vetorial real de dimensão finita \bar{E} é chamado poliedro, se P é a interseção de um número finito de semi-espacos fechados.

II.5.e - Face

Uma face de um poliedro P é a interseção de P com um hiperplano suporte.

II.5.f - Cone Poliedr ico Convexo

Um cone poli drico convexo C ,   a envolt ria convexa de um conjunto finito de raios vetores.

Portanto se os (vetores) pontos $a^i \neq 0 \quad i=1, \dots, h$ geram raios vetores, ent o C   a cole o dos pontos

$$y = \sum_{i=1}^h \lambda_i a^i \quad \lambda_i \geq 0 \quad \forall_i$$

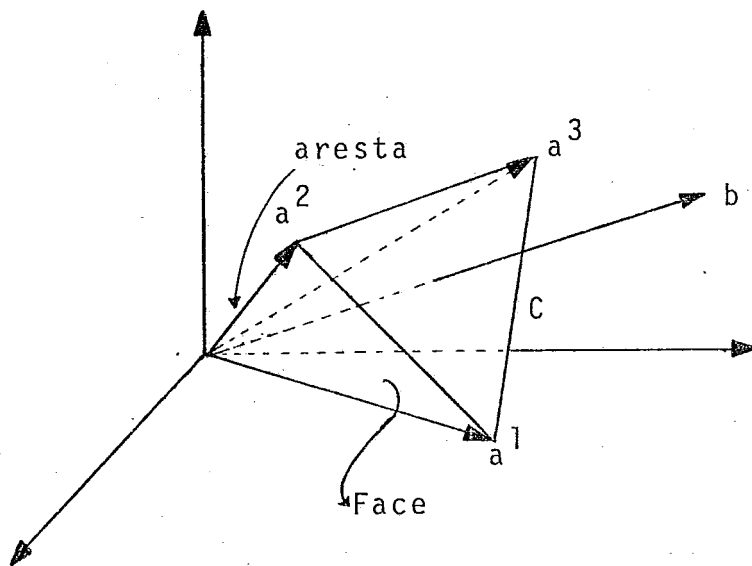


Fig. II.2

Assim se $A \bar{e}$ uma matriz $m \times k$, então o conjunto dos pontos

$$y = Ax = \sum_{j=1}^k x_j a^j, x_j \geq 0 \forall_j$$

\bar{e} um cone poliédrico convexo em R^m . . Observa-se que as colunas de A geram os raios vetores cuja combinação convexa produz o cone poliédrico. Como $C(a^1, \dots, a^k)$ \bar{e} um cone poliédrico segue-se que existe uma solução não negativa $x \geq 0$ para

$$Ax = b$$

se e somente se $b \bar{e}$ um elemento do cone gerado pelas colunas de A . Podemos agora dizer que $P_k b \in C[C_{k+1}, \dots, C_n]$ se

$\langle a^j, P_k b \rangle \geq 0 \quad j = k + 1, \dots, n$, onde

$$A = (a^1, \dots, a^n), \quad A^+ = \begin{array}{c} \left[\begin{array}{c} C_1^T \\ \vdots \\ C_k^T \\ C_{k+1}^T \\ \vdots \\ C_n^T \end{array} \right] \end{array}, \quad b \in C[a^1, \dots, a^n]$$

e $\dim(C) = n$.

De fato, como $C[C_{k+1}, \dots, C_n]$ é um subcone de $C[C_1, \dots, C_k, C_{k+1}, \dots, C_n]$ de dimensão $n-k$, onde $0 \leq k \leq n-1$, então $P_k b$ pertencerá ao cone gerado por $C[C_{k+1}, \dots, C_n]$ se $P_k b$ puder ser escrito como uma combinação linear não negativa de $\{C_{k+1}, \dots, C_n\}$; isto é

$$P_k b = \gamma_{k+1} C_{k+1} + \dots + \gamma_n C_n \quad \gamma_j \geq 0,$$

$$j = k + 1, \dots, n.$$

Portanto basta mostrar que:

$$\gamma_j = \langle a^j, P_k b \rangle \quad \forall j, \quad j = k + 1, \dots, n$$

Pelo teorema 1, $\langle a^j, P_k b \rangle = 0 \quad \forall j = 1, \dots, k$.

Por outro lado

$$A^+ A = \begin{bmatrix} C_1^T \\ \vdots \\ C_i^T \\ \vdots \\ C_n^T \end{bmatrix} [a^1, \dots, a^k, a^{k+1}, \dots, a^n] = I \rightarrow \begin{cases} C_i^T a^j = 0 & \forall i \neq j \\ C_i^T a^j = 1 & \forall i = j \end{cases} \quad (\text{II.5.1})$$

$$\langle a^j, P_k b \rangle = \langle a^j, (\gamma_{k+1} C_{k+1} + \dots + \gamma_j C_j + \dots + \gamma_n C_n) \rangle$$

$$\langle a^j, P_k b \rangle = \sum_{i=k+1}^n \gamma_i \langle C_i, a^j \rangle = \gamma_j \langle C_j, a^j \rangle = \gamma_j$$

$$\langle a^j, P_k b \rangle = \gamma_j \quad j = k+1, \dots, n.$$

Como elucidação dessas idéias, vejamos o seguinte

exemplo no R^3 ; sejam

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, A_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} C_1^T \\ C_2^T \\ C_3^T \end{bmatrix} \quad \rightarrow = 1a^1 + 2a^2 + 6a^3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

OBS: Quando estamos no passo um, o índice superior e à esquerda indicado em ${}^i P_j$ corresponde àquele vetor com o qual ${}^i P_j$

foi construído. Estaremos utilizando esta notação até o passo 3.

Sabemos que:

$${}^1P_1b = \langle a^2, {}^1P_1b \rangle C_2 + \langle a^3, {}^1P_1b \rangle C_3$$

$${}^2P_1b = \langle a^1, {}^2P_1b \rangle C_1 + \langle a^3, {}^2P_1b \rangle C_3$$

$${}^3P_1b = \langle a^1, {}^3P_1b \rangle C_1 + \langle a^2, {}^3P_1b \rangle C_2$$

$$\text{Se } {}^1P_1b \in C[C_2, C_3] \rightarrow \begin{cases} \langle a^2, {}^1P_1b \rangle \geq 0 \\ \langle a^3, {}^1P_1b \rangle \geq 0 \end{cases} \quad (\text{II.5.2})$$

$${}^2P_1b \in C[C_1, C_3] \rightarrow \begin{cases} \langle a^1, {}^2P_1b \rangle \geq 0 \\ \langle a^3, {}^2P_1b \rangle \geq 0 \end{cases} \quad (\text{II.5.3})$$

$${}^3P_1b \in C[C_1, C_2] \rightarrow \begin{cases} \langle a^1, {}^3P_1b \rangle \geq 0 \\ \langle a^2, {}^3P_1b \rangle \geq 0 \end{cases} \quad (\text{II.5.4})$$

$${}^1P_1 = I - a^1 [a^1]^+ = I - \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} [1/2 \quad 0 \quad -1/2] = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

$${}^1P_1b = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \rightarrow \begin{cases} \langle a^2, {}^1P_1b \rangle = 0 \\ \langle a^3, {}^1P_1b \rangle = 2 \end{cases} \quad (\text{II.5.2}) \text{ é satisfeita}$$

$${}^2P_1 = I - a^2 [a^2]^+ = I - \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} [0 \quad 1/2 \quad -1/2] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

$${}^2P_1 b = \begin{bmatrix} 1 \\ 2,5 \\ 2,5 \end{bmatrix} \rightarrow \begin{cases} \langle a^1 \quad {}^2P_1 b \rangle = -1,5 \\ \langle a^3 \quad {}^2P_1 b \rangle = 2,5 \end{cases}$$

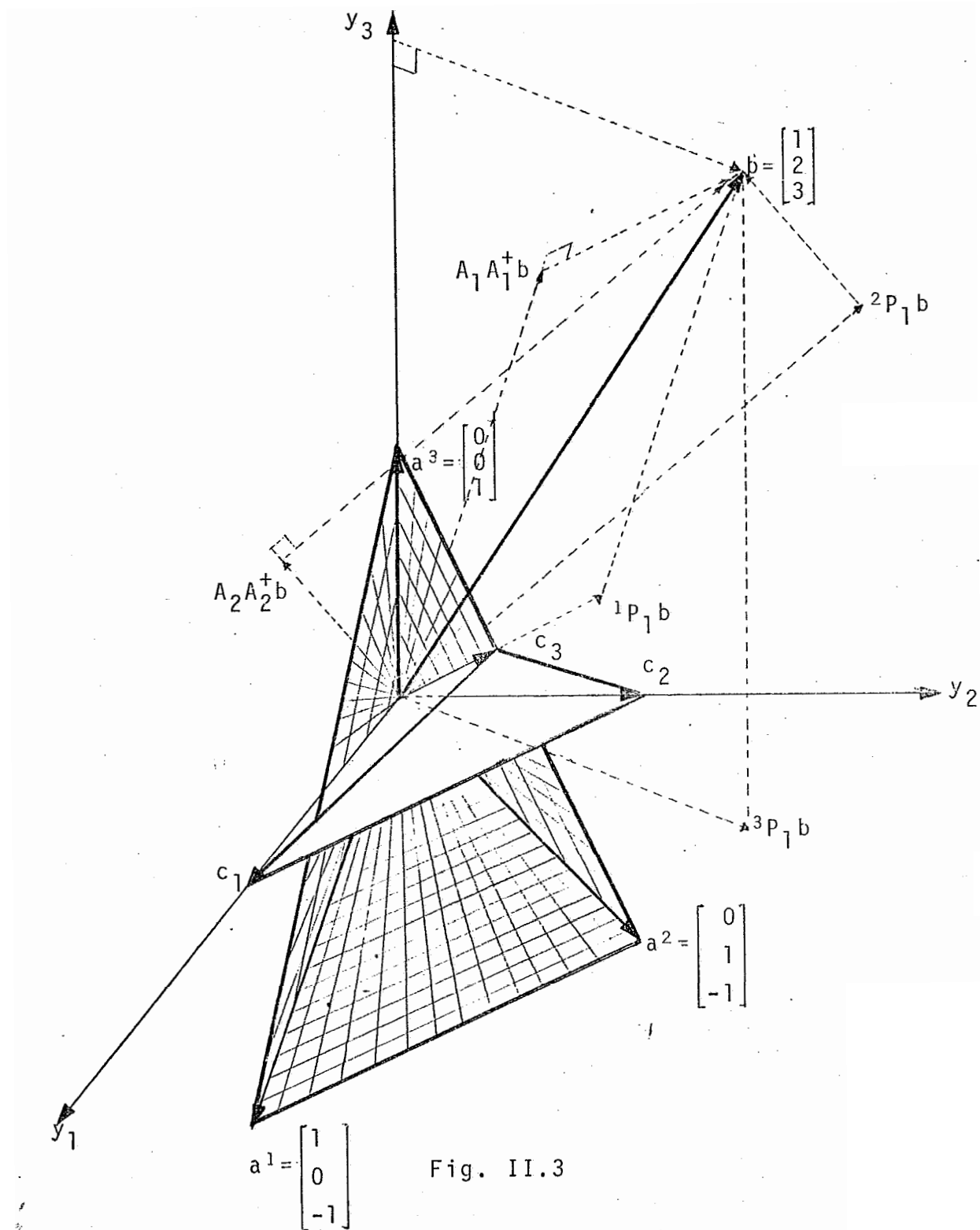
Assim (II.5.3) não é satisfeita.

$${}^3P_1 = I - a^3 [a^3]^+ = I - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [0 \quad 0 \quad 1] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$${}^3P_1 b = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \rightarrow \begin{cases} \langle a^1 \quad {}^3P_1 b \rangle = + 1 \\ \langle a^2 \quad {}^3P_1 b \rangle = + 2 \end{cases} \quad \text{satisfaz (II.5.4)}$$

Concluimos assim que:

$${}^1P_1 b \in C[C_2 \ C_3], \quad {}^2P_1 b \notin C[C_1 \ C_3] \quad \text{e} \quad {}^3P_1 b \in C[C_1 \ C_2]$$



Teorema 2: (Teorema da projeção de Perez)

Seja $\{a^1, \dots, a^k\}$ um conjunto de k vetores LI em \mathbb{R}^m ($1 < k < m$). A projeção de a^i sobre o subespaço ortogonal ao gerado por $\{a^1, \dots, a^{i-1}, a^{i+1}, \dots, a^k\}$ é

$$\frac{C_i}{\|C_i\|^2}$$

onde C_i^T é a i -ésima linha de A .

Antes de entrarmos no mérito da demonstração do teorema, veja-o no \mathbb{R}^3 através da Fig. 11.3.

Prova: Se as colunas de $A = [a^1, \dots, a^k]$ são LI então

$$A^+ = (A^T A)^{-1} A^T = D A^T \text{ onde } D = (A^T A)^{-1} = \begin{bmatrix} d_1 \\ \vdots \\ d_k \end{bmatrix}$$

é uma matriz $(k \times k)$.

$A^T = [a^1, a^2, \dots, a^m]$ é uma matriz $k \times m$ e

$$A^+ = D A^T = \begin{bmatrix} d_1 \\ \vdots \\ d_k \end{bmatrix} [a^1, a^2, \dots, a^m] = \begin{bmatrix} C_1^T \\ \vdots \\ C_i^T \\ \vdots \\ C_k^T \end{bmatrix}$$

Então $C_i^T = d_i A^T$

Já mostramos que $P_{k-1}^j b = \gamma_j C_j$, conseqüentemente

$p_{k-1}^j a^j = a_j c_j$ e daí, por (II.5.1) temos

$$\langle a^i, p_{k-1}^j a^j \rangle = \alpha_j \langle a^i, c_j \rangle = \begin{cases} 0 & \forall i \neq j \\ \alpha_j & \forall i = j \end{cases}$$

o que nos leva a afirmar

$$\langle a^i, c_j \rangle = \begin{cases} 0 & \forall i \neq j \\ 1 & \forall i = j \end{cases}$$

Assim

$$c_i^T c_j = (d_i A^T) c_j = d_i (A^T c_j) = d_i e_j = d_{ij}$$

onde e_j é um vetor onde todas componentes são zero, exceto a j -ésima que é 1.

Já que $d_{ij} = c_j^T c_i$, então

$$d_i = [(c_i^T c_1) (c_i^T c_2) \dots (c_i^T c_i) \dots c_i^T c_k] e$$

$$\frac{c_i}{\|c_i\|^2} = \frac{(d_i A^T)^T}{\|c_i\|^2} = \frac{Ad_i^T}{\|c_i\|^2} = \sum_{j=1}^k \left(\frac{c_i^T c_j}{\|c_i\|^2} \right) a^j$$

Fazendo $\phi_j = \frac{c_i^T c_j}{\|c_i\|^2}$ e observando

que $\frac{c_i^T c_i}{\|c_i\|^2} = 1$, temos

$$\frac{c_i}{\|c_i\|^2} = \sum_{\substack{j=1 \\ j \neq i}}^k \phi_j a^j + a^i \quad (\text{II.5.5})$$

$$\text{ou } a^i = \frac{c_i}{\|c_i\|^2} - \sum_{\substack{j=1 \\ j \neq i}}^k \phi_j a^j \quad (\text{II.5.6})$$

Aplicando a projeção em a^i vem:

$$P_{1 \dots i-1, i+1 \dots k} a^i = P_{1 \dots i-1, i+1 \dots k} \frac{c_i}{\|c_i\|^2} - \sum_{j \neq i} \phi_j P_{1 \dots i-1, i+1 \dots k}$$

onde a 2ª parcela do 2º membro é nula, tendo-se em vista que a projeção de a^j sobre um subespaço ortogonal de a^j é evidentemente nula.

Como

$P_{1 \dots i-1, i+1 \dots k} = I - B_{1 \dots i-1, i+1 \dots k}$, segue-se que

$$\begin{aligned} P_{1 \dots i-1, i+1 \dots k} a^i &= [I - B_{1 \dots i-1, i+1 \dots k}] \frac{c_i}{\|c_i\|^2} \\ &= \frac{c_i}{\|c_i\|^2} - B_{1 \dots i-1, i+1 \dots k} \frac{c_i}{\|c_i\|^2} \end{aligned}$$

o que nos leva a concluir a prova do teorema, já que novamente a 2ª parcela do 2º membro é nula, haja visto ser $\frac{c_i}{\|c_j\|^2}$ ortogonal ao subespaço gerado por $\{a^1, \dots, a^{i-1}, \dots, a^k\}$.

Teorema 3: O sistema $Ax = b$ onde A é uma matriz de posto completo $m \times j$, com $j \geq m$, tem solução mínima quadrada $x \geq 0$ se e somente se existem pelo menos m vetores colunas de A LI, tal que a projeção de cada um desses vetores colunas sobre o subespaço ortogonal gerado pelos $m-1$ vetores restantes tem um produto escalar não negativo com b .

Prova: Sejam $A_m = (a^1, \dots, a^m) \rightarrow A_m^+ = \begin{bmatrix} c_1^T \\ \vdots \\ c_m^T \end{bmatrix}$

$$\text{Pelo teorema 2 } P_{m-1}^i a^i = \frac{c_i}{\|c_j\|^2}$$

$$\text{Então } \langle b, P_{m-1}^i a^i \rangle = \langle b, \frac{c_i}{\|c_j\|^2} \rangle$$

Assim

$$x_i = \langle c_i, b \rangle \geq 0 \leftrightarrow \langle b, P_{m-1}^i a^i \rangle \geq 0$$

$$\forall i = 1, \dots, m$$

Para mostrar que $x = [c_1^T b \dots c_j^T b]^T = A^+ b$ é o único vetor de norma mínima entre aqueles que minimizam $\|b - Ax\|^2$, veja [1, pp. 19-20].

Voltando à Figura 1.3 observa-se que

$$\begin{array}{l}
 \text{Se } C_3 \left\{ \begin{array}{l} \perp a^1 \\ \perp a^2 \end{array} \right. \rightarrow \left\{ \begin{array}{l} C_3 \text{ é perpendicular ao plano determinado por } a^1 \text{ e } a^2, \\ \text{além do mais } \langle b, {}^{21}P_2 a^3 \rangle = \langle b^T, \frac{C_3}{\|C_3\|^2} \rangle > 0. \text{ Isto significa dizer que } b \text{ pertence ao semi-espaço cujo hiperplano fronteira passando pela origem é o plano determinado por } a^1 \text{ e } a^2. \quad (\text{II.5.7.a}) \end{array} \right. \\
 \\
 \text{Se } C_2 \left\{ \begin{array}{l} \perp a^1 \\ \perp a^3 \end{array} \right. \rightarrow \left\{ \begin{array}{l} C_2(C_1) \text{ é perpendicular ao plano determinado por } a^1 \text{ e } a^3 \text{ (} a^2 \text{ e } a^3 \text{)}. \text{ Além do mais se } \langle b, {}^{13}P_2 a^2 \rangle = \frac{b^T C_2}{\|C_2\|^2} > 0 \text{ (} \langle b, {}^{23}P_2 a^1 \rangle = \frac{b^T C_1}{\|C_1\|^2} > 0 \text{)} \text{ isto significa dizer que } b \text{ pertence ao semi-espaço cujo hiperplano fronteira passando pela origem é o plano determinado por } a^1 \text{ e } a^3 \text{ (} a^2 \text{ e } a^3 \text{)}. \quad (\text{II.5.7.b}) \end{array} \right. \\
 \left(\text{Se } C_1 \left\{ \begin{array}{l} \perp a^2 \\ \perp a^3 \end{array} \right. \right)
 \end{array}$$

conclusão: Por (II.5.7.a) e (II.5.7.b) concluímos que b está no interior do cone determinado por a^1 , a^2 e a^3 . Portanto b pode ser escrita como combinação linear positiva de a^1 , a^2 e a^3 ; isto é

$$[a^1, a_2, a_3] x = b \text{ ou}$$

$$Ax = b$$

$$x > 0$$

Como nosso objetivo é resolver o PLP diretamente, em termos de projeções, o teorema que segue visa caracterizar as faces, através do operador projeção.

Teorema 4: Para $m \geq 3$, a^1, \dots, a^{m-1} geram uma face do cone $C[a^1, \dots, a^n] \leftrightarrow \langle a^j, P_{m-1} b \rangle \geq 0$

$\forall j = 1, \dots, n$ (se b pertence ao cone).

Prova: Suponhamos primeiramente que a^1, \dots, a^{m-1} geram uma face do cone $C[a^1, \dots, a^n]$.

Se com C denotamos o cone, e com F_γ a face, então existe um hiperplano sobre o qual a face F_γ repousa, digamos

$H_\gamma = \{x \text{ t.q. } \langle x, \mu_\gamma \rangle = 0\}$, de tal modo que $F_\gamma = H_\gamma \cap C$, onde C está contido no semi-espço $\{x \text{ t.q. } \langle x, \mu_\gamma \rangle \geq 0\}$ e μ_γ é um vetor unitário ortogonal ao hiperplano H_γ .

Como $b \in C$ e $a^j \in C \forall j$, temos respectivamente $\langle b, \mu_\gamma \rangle \geq 0$ e $\langle a^j, \mu_\gamma \rangle \geq 0 \forall j$. Observa-se que $\mu_\gamma \in R(P_{m-1})$, portanto

$$P_{m-1}b = \frac{\langle \mu_\gamma, b \rangle}{\|\mu_\gamma\|^2} \mu_\gamma \quad \text{Conseqüentemente}$$

$$\langle a^j, P_{m-1}b \rangle = \langle a^j, \frac{\langle \mu_\gamma, b \rangle}{\|\mu_\gamma\|^2} \mu_\gamma \rangle = \frac{\langle \mu_\gamma, b \rangle}{\|\mu_\gamma\|^2} \langle a^j, \mu_\gamma \rangle \geq 0 \quad \forall j$$

Inversamente, admitamos agora que $\langle a^j, P_{m-1}b \rangle \geq 0 \forall j$ e que a^1, \dots, a^{m-1} não geram uma face do cone. Então existe no mínimo um índice k tal que $\langle a^k, \mu_\gamma \rangle < 0$. Mas isto significa dizer que

$$\langle a^k, P_{m-1}b \rangle = \frac{\langle \mu_\gamma, b \rangle}{\|\mu_\gamma\|^2} \langle a^k, \mu_\gamma \rangle < 0$$

o que é uma contradição. Observa-se que esta prova esta fundamentada a a^1, \dots, a^{m-1} , $m \geq 3$ gerarem uma face do cone em um espaço m -dimensional, mas o teorema 5 que segue levantar; tal restrição. Observe ainda que se a^i gera uma aresta do cone $C[a^1, \dots, a^n]$, então não podemos garantir que $\langle b, \mu_\gamma \rangle \langle a^j, \mu_\gamma \rangle \geq 0 \forall j$, pois para cada i existe um feixe de hiperplanos suportes, conseqüentemente diversos μ_γ , um para cada hiperplano do feixe. Portanto pode existir a^i tal que $\langle a^j, P_{m-1}b \rangle < 0$ e entretanto es-

te a^i gera uma aresta do cone. Veja o exemplo numérico, página 65 onde a^2 e a^3 geram arestas do cone e $\langle a^1, {}^2P_1 b \rangle \langle a^4, {}^3P_1 b \rangle < 0$. então $\langle a^1, {}^2P_1 b \rangle < 0$, $\langle a^4, {}^3P_1 b \rangle < 0$.

Corolário: Se $\langle a^j, {}^iP_1 b \rangle \geq 0 \forall j$, então a^i gera uma aresta do cone.

Prova: Consequência imediata do teorema'.

Teorema 5: Se a^1, \dots, a^{m-k} , $k = 2, \dots, m-1$ geram uma face do subcone gerado por $a^1, \dots, a^{m-(k-1)}$ então $\langle a^j, {}^jP_{m-k} b \rangle \geq 0 \quad j = 1, \dots, n$.

Prova: Temos que

$$P_n = \begin{cases} P_{n-1} & \text{se } a^n \text{ é uma combinação linear de } a^1, \dots, a^{n-1} \\ P_{n-1} - \frac{[P_{n-1} a^n]}{\|P_{n-1} a^n\|} \cdot \frac{[P_{n-1} a^n]^T}{\|P_{n-1} a^n\|} & \text{caso contrário} \end{cases} \quad (\text{II.5.8})$$

Prosseguindo, para $k = 2$ temos por (II.5.8)

$$P_{m-1} = \begin{cases} P_{m-2} & \text{se } a^{m-1} \text{ é uma combinação linear de } a^1, \dots, a^{m-2} \\ P_{m-2} - \frac{[P_{m-2} a^{m-1}]}{\|P_{m-2} a^{m-1}\|} \cdot \frac{[P_{m-2} a^{m-1}]^T}{\|P_{m-2} a^{m-1}\|} & \text{caso contrário} \end{cases}$$

Agora pelo teorema 4 temos $\langle a^j, {}^jP_{m-1} b \rangle \geq 0 \quad \forall j = 1, \dots, n$, assim

$\langle a^j, P_{m-1}b \rangle = \langle a^j, P_{m-2}b \rangle \geq 0$, se a^{m-1} é uma' comblinear de a^1, \dots, a^{m-2} , caso contrário

$$\langle a^j, P_{m-1}b \rangle = \langle a^j, P_{m-2}b \rangle - \frac{\langle a^j, P_{m-2}a^{m-1} \rangle}{\|P_{m-2}a^{m-1}\|} \cdot \frac{\langle a^{m-1}, P_{m-2}b \rangle}{\|P_{m-2}a^{m-1}\|} \geq 0 \quad (II.5.9)$$

$\forall j = 1, \dots, n$

De (II.5.9) tiramos

$$\langle a^j, P_{m-2}b \rangle \geq \frac{\langle a^j, P_{m-2}a^{m-1} \rangle}{\|P_{m-2}a^{m-1}\|} \cdot \frac{\langle a^{m-1}, P_{m-2}b \rangle}{\|P_{m-2}a^{m-1}\|} \geq 0 \quad \forall j = 1, \dots, n$$

$\leftrightarrow \langle a^j, P_{m-2}a^{m-1} \rangle$ e $\langle a^{m-1}, P_{m-2}b \rangle$ tiverem o mesmo sinal.

Como $\langle a^j, P_{m-2}a^{m-1} \rangle \langle a^{m-1}, P_{m-2}b \rangle = \langle a^{m-1}, P_{m-2}a^j \rangle \langle a^{m-1}, P_{m-2}b \rangle$ e, tendo-se em vista que os vetores $P_{m-2}a^j$ e $P_{m-2}b$ são colineares, e estão num mesmo semi-espaço segue-se que $\langle a^{m-1}, P_{m-2}a^j \rangle$ e $\langle a^{m-1}, P_{m-2}b \rangle$ tem o mesmo sinal, o que nos leva a afirmar que o seu produto

$$\langle a^j, P_{m-2}a^{m-1} \rangle \langle a^{m-1}, P_{m-2}b \rangle \geq 0 \quad \forall j = 1, \dots, n$$

Assim para $k = 2$

$$\langle a^j, P_{m-2}b \rangle \geq 0 \quad \forall j = 1, \dots, n.$$

Continuando, admitamos agora que a proposição seja válida para $k = L$, e mostremos que ela continua verdadeira pa

ra $k = L + 1$.

De maneira completamente análoga, temos

$$\langle a^j, P_{m-(L+1)} b \rangle \geq \frac{\langle a^j, P_{m-(L+1)} a^{m-L} \rangle}{\|P_{m-(L+1)} a^{m-L}\|} \cdot \frac{\langle a^{m-L}, P_{m-(L+1)} b \rangle}{\|P_{m-(L+1)} a^{m-L}\|} \geq 0$$

$\forall j = 1, \dots, m$

Haja visto que o numerador do 2º membro \bar{e} sempre não negativo, pelos mesmos motivos já esclarecidos no caso anterior.

ALGORITMO

STEP 1

- a) Selecione um conjunto L_1 , constituído de todos aqueles vetores, tal que $\langle a^j, b \rangle \geq 0$. Se $L_1 = \emptyset$, então não existe sol. viável.
- b) Seja L_2 um subconjunto dos índices dos vetores de L_1 , tal que $\langle a^i, P_j b \rangle \geq 0 \quad \forall i = 1, \dots, n$.
- c) Para cada $j \in L_2$ substitua b por b_j ($b_j = A^j x_j$) e escolha um, tal que o mínimo do funcional seja.

$\min_j Z = c_j x_j$ onde

$$x_j = [A^j]^+ b = \frac{\langle a^j, b \rangle}{\|a^j\|^2}$$

STEP 2

Tendo selecionado t vetores, para $t = 1, \dots, m-2$, escolha o próximo conforme:

a) Selecione um conjunto de k vetores tais que

$$\bar{a}_{1\dots t,j}^+ b \geq 0 \text{ para } j = 1, \dots, k$$

b) Para cada um dos k vetores acima compute

$$\langle a^L, \bar{a}_{1\dots t,j}^+ b \rangle \text{ para } L \in \{1\dots N\} - \{1\dots t, j\}$$

c) Para aqueles j tais que $\langle a^L, \bar{a}_{1\dots t,j}^+ b \rangle > 0$ substitua b por b_j ($b_j = x_1 A^1 + \dots + x_t A^t + x_j A^j$), e se existem muitos, admita um tal que o mínimo do funcional, seja

$\min_j Z = x_1 c_1 + \dots + x_t c_t + x_j c_j$ onde

$$\begin{bmatrix} x_1 \\ \vdots \\ x_t \\ x_j \end{bmatrix} = [A_{1\dots t,j}]^+ b \geq 0 \quad (\text{d'ado por a})$$

STEP 3:

Se já foram selecionados $m-1$ vetores o m -ésimo será um, o qual minimize o funcional, isto é, se com $Z_{1\dots m-1,k}$ representamos o valor do funcional quando usamos os vetores $A^1 \dots A^{m-1} A^k$, então:

$$Z_{1\dots m-1,m} = \min Z_{1\dots m-1,k}$$

Vejamos agora como o algoritmo funciona. Suponhamos que estavamos com o seguinte problema de programação linear

$$\begin{aligned} \min \quad & c^T x \\ \text{sujeito a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

onde $x = [x_1, x_2, x_3, x_4]^T$ e $A_{3 \times 4} = [A^1, A^2, A^3, A^4]$

No 1º Step o algoritmo substitue o problema proposto por $k = 4$ problemas aproximados.

$$P_{1j} \quad \left\{ \begin{array}{l} \min \quad c^T x \\ \beta_j \\ \text{sujeito a} \quad [A^1, A^2, A^3, A^4] x = b_{1j} \quad (b_{1j} = A^j x_j) \\ x \geq 0 \\ \beta_j = \{j\} \end{array} \right.$$

Admitamos que o mínimo do funcional ocorreu para $j = 1$. No 2º Step o algoritmo substitue o problema original por

$k = 3$ problemas aproximados

$$P_{2j} \quad \begin{cases} \min c^T x \\ \beta_j \\ \text{suju. } [A^1, A^2, A^3, A^4]x = b_{2j} \quad (b_{2j} = A^1 x_1 + A^j x_j) \\ x > 0 \\ \beta_j = \{1, j\} \end{cases} \quad j=2,3,4$$

Admitamos que o m̃nimo do funcional ocorreu para $j = 2$.

Mo 39 Step o algoritmo calcula

$$P_{3j} \quad \begin{cases} \min c^T x \\ \beta_j \\ Ax = b \\ \beta_j = \{1, 2, j\} \end{cases} \quad j=3,4$$

Note que o problema original \bar{e}

$$\bar{P}_3 = \begin{cases} \min c^T x \\ \text{suju. } Ax = b \\ x > 0 \end{cases}$$

Suponhamos que a sol. de P_{3j} seja $x = (x_1 \ x_2 \ x_3 \ 0)$ resta mostrar que a solução de P_{3j} \bar{e} a mesma de \bar{P}_3 (problema proposto).

Para chegarmos a tal conclusão, devemos provar que $c_j - z_j = 0$ para $j = 1, 2, 3$ e $c_j - z_j < 0$ para $j = 4$.

Vejamos então que o critério de otimalidade \bar{e} válido.

No Step 3 tínhamos:

$$Z_{1\dots m-1,m} \leq Z_{1\dots m-1,k} \rightarrow \quad (\text{II.5.10})$$

$$\rightarrow (x_1 c_1 + \dots + x_{m-1} c_{m-1}) + x_m c_m \leq (x_1^i c_1 + \dots + x_{m-1}^i c_{m-1}) + x_m^i c_m$$

Fazendo

$$\underline{x}_{m-1} = (x_1 \dots x_{m-1})$$

$$\underline{c}_{m-1} = (c_1 \dots c_{m-1})$$

$$\underline{x}_{m-1}^i = (x_1^i \dots x_{m-1}^i)$$

Temos:

$$\langle \underline{x}_{m-1}, \underline{c}_{m-1} \rangle + x_m c_m \leq \langle \underline{x}_{m-1}^i, \underline{c}_{m-1} \rangle + x_m^i c_m \quad (\text{II.5.11})$$

Podemos ainda escrever

$$\begin{bmatrix} A_{m-1} \\ \vdots \\ A_m \end{bmatrix} \begin{bmatrix} \underline{x}_{m-1} \\ x_m \end{bmatrix} = b \rightarrow A_{m-1} \underline{x}_{m-1} + A_m x_m = b \rightarrow$$

$$\rightarrow A_{m-1} \underline{x}_{m-1} = b - A_m x_m \rightarrow A_{m-1}^+ A_{m-1} \underline{x}_{m-1} = A_{m-1}^+ (b - A_m x_m)$$

$$\rightarrow \underline{x}_{m-1} = A_{m-1}^+ b - x_m A_{m-1}^+ A_m \quad \text{isto é}$$

$$\underline{x}_{m-1} = A_{m-1}^+ b - x_m A_{m-1}^+ a^m \quad (\text{II.5.12})$$

de modo analógico

$$\underline{x}'_{m-1} = A_{m-1}^+ b - x_j A_{m-1}^+ a^j \quad (\text{II.5.13})$$

Considerando os valores de \underline{x}_{m-1} e \underline{x}'_{m-1} encontrados em (II.5.12) e (II.5.13), temos que (II.5.11) p'ode ser escrito como segue

$$\begin{aligned} & \langle (A_{m-1}^+ b - x_m A_{m-1}^+ a^m), \underline{c}_{m-1} \rangle + x_m c_m \leq \langle (A_{m-1}^+ b - x_j A_{m-1}^+ a^j), \underline{c}_{m-1} \rangle + x_j c_j \\ & \langle A_{m-1}^+ b, \underline{c}_{m-1} \rangle - x_m \langle A_{m-1}^+ a^m, \underline{c}_{m-1} \rangle + x_m c_m \leq \langle A_{m-1}^+ b, \underline{c}_{m-1} \rangle - x_j \langle A_{m-1}^+ a^j, \underline{c}_{m-1} \rangle + x_j c_j \\ & x_m c_m - x_m \langle A_{m-1}^+ a^m, \underline{c}_{m-1} \rangle \leq x_j c_j - x_j \langle A_{m-1}^+ a^j, \underline{c}_{m-1} \rangle \\ & x_m [\langle A_{m-1}^+ a^m, \underline{c}_{m-1} \rangle - c_m] + x_j [\langle A_{m-1}^+ a^j, \underline{c}_{m-1} \rangle - c_j] \geq 0 \end{aligned} \quad (\text{II.5.14})$$

Mas por (II.4.1)

$$x_m = \frac{\langle a^m, P_{m-1} b \rangle}{\langle a^m, P_{m-1} a^m \rangle}$$

e

$$x_j = \frac{\langle a^j, P_{m-1} b \rangle}{\langle a^j, P_{m-1} a^j \rangle}$$

Este é o valor de x_j se ele estiver na base, portanto para $x_j > 0$ temos:

$$\frac{x_m}{x_j} = \frac{\langle a^m, P_{m-1} b \rangle}{\langle a^m, P_{m-1} a^m \rangle} \cdot \frac{\langle a^j, P_{m-1} a^j \rangle}{\langle a^j, P_{m-1} b \rangle}$$

$$= \frac{\langle a^j, P_{m-1} b \rangle \langle a^m, P_{m-1} a^j \rangle}{\langle a^j, P_{m-1} b \rangle \langle a^m, P_{m-1} a^m \rangle}$$

o que nos dá:

$$\frac{x_m}{x_j} = \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle} \quad (\text{II.5.15})$$

Multiplicando (II.5.14) por (-1) e dividindo por x_j obtemos:

$$\frac{x_m}{x_j} [-\langle A_{m-1}^+ a^m, c_{m-1} \rangle + c_m] + \langle A_{m-1}^+ a^j, c_{m-1} \rangle - c_j \leq 0$$

Por (II.5.15):

$$\langle A_{m-1}^+ a^j, c_{m-1} \rangle = \langle A_{m-1}^+ a^m \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle}, c_{m-1} \rangle +$$

$$\frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle} c_m - c_j \leq 0$$

$$[c_1 \dots c_{m-1}] [A_{m-1}^+ a^j - A_{m-1}^+ (a^m) \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle}] + \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle} c_m - c_j \leq 0$$

$$[c_1 \dots c_{m-1} : c_m] \begin{bmatrix} A_{m-1}^+ a^j - A_{m-1}^+ a^m \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle} \\ \frac{\langle a^m, P_{m-1} a^j \rangle}{\langle a^m, P_{m-1} a^m \rangle} \end{bmatrix} - c_j \leq 0$$

isto é

$$[c_1 \dots c_m] \begin{bmatrix} A_{m-1}^+ - A_{m-1}^+ a^m \frac{(a^m)^T P_{m-1}}{\langle a^m, P_{m-1} a^m \rangle} \\ \frac{(a^m)^T P_{m-1}}{\langle a^m, P_{m-1} a^m \rangle} \end{bmatrix} a^j - c_j \leq 0 \quad (\text{II.5.16})$$

Mas por (II.3.1)

$$A_m^+ = \begin{bmatrix} A_{m-1}^+ - A_{m-1}^+ a^m \frac{(a^m)^T P_{m-1}}{\langle a^m, P_{m-1} a^m \rangle} \\ \frac{(a^m)^T P_{m-1}}{\langle a^m, P_{m-1} a^m \rangle} \end{bmatrix} = A_m^{-1}$$

Logo (II.5.16) nos dá:

$$[c_1 \dots c_m] A_m^{-1} a^j - c_j = z_j - c_j \leq 0$$

Para todos aqueles j que teriam condições de entrada na base e darem soluções viáveis.

Exemplo Numérico

$$\text{Min. } z = 4x_1 + 1x_2 + 8x_3 + 5x_4$$

$$\text{Suj.a } x_1 + 3x_2 + 5x_3 + 6x_4 = 3$$

$$6x_1 + 4x_2 + 2x_3 + x_4 = 2$$

$$5x_1 + 3x_2 + 4x_3 + 6x_4 = 4$$

$$x_i \geq 0 \quad i = 1, 2, 3, 4$$

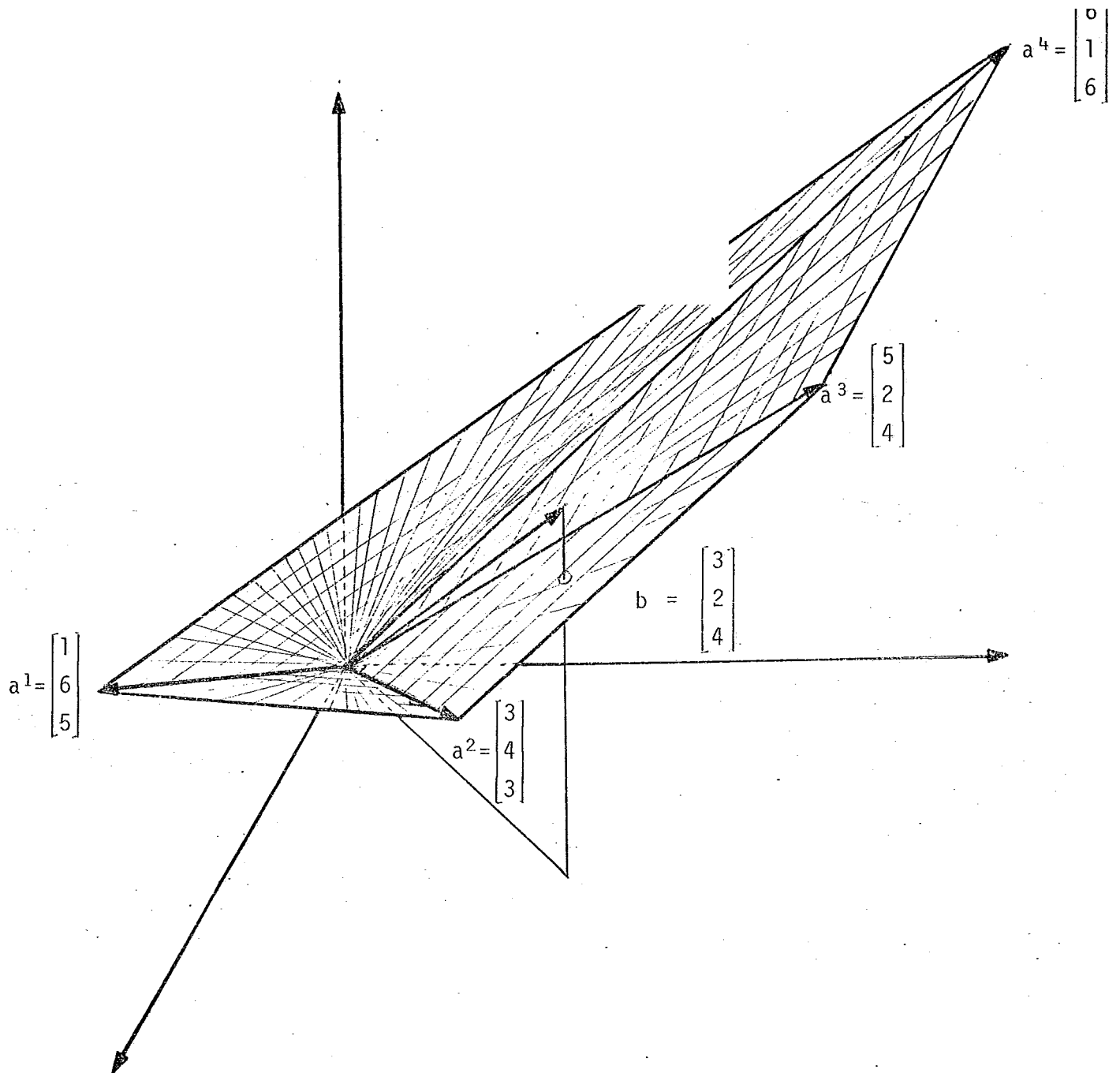


Fig. II.4

Solução:

Sejam:

$$a^1 = \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix} \quad a^2 = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix} \quad a^3 = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix} \quad a^4 = \begin{bmatrix} 6 \\ 1 \\ 6 \end{bmatrix} \quad b = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

$$x = [x_1, x_2, x_3, x_4]^T \quad c^T = [c_1, c_2, c_3, c_4] = [4, 1, 8, 5]$$

Step 1:

a) Como os a^j $j = 1, \dots, 4$ e b tem suas componentes positivas, concluímos que

i) $\langle a^j, b \rangle > 0 \quad \forall j = 1, \dots, 4$

ii) $L_1 = \{a^1, a^2, a^3, a^4\}$

b) Cálculo de L_2

$${}^1P_1 = I - a^1 [a^1]^+ = I - \frac{a^1 (a^1)^T}{\|a^1\|^2} \rightarrow {}^1P_1 b = \frac{1}{62} \begin{pmatrix} 156 \\ -56 \\ 36 \end{pmatrix} \rightarrow$$

$$\begin{cases} \langle a^1, {}^1P_1 b \rangle = 0 \\ \langle a^2, {}^1P_1 b \rangle = 5,677 \\ \langle a^3, {}^1P_1 b \rangle = 13,096 \\ \langle a^4, {}^1P_1 b \rangle = 17,677 \end{cases}$$

Assim o índice 1 pertence a L_2 .

$${}^2P_1 = I - a^2 [a^2]^+ = I - \frac{a^2 (a^2)^T}{\|a^2\|^2} \rightarrow {}^2P_1 b = \frac{1}{34} \begin{pmatrix} 24 \\ -36 \\ 24 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{cases} \langle a^1, {}^2P_1 b \rangle = -2,117 < 0 \\ \langle a^2, {}^2P_1 b \rangle = 0 \\ \langle a^3, {}^2P_1 b \rangle = 4,235 \\ \langle a^4, {}^2P_1 b \rangle = 7,411 \end{cases}$$

$${}^3P_1 = I - a^3 [a^3]^+ = I - \frac{a^3 (a^3)^T}{\|a^3\|^2} \rightarrow {}^3P_1 b = \frac{1}{45} \begin{pmatrix} -20 \\ 28 \\ 11 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{cases} \langle a^1, {}^3P_1 b \rangle = 4,511 \\ \langle a^2, {}^3P_1 b \rangle = 1,888 \\ \langle a^3, {}^3P_1 b \rangle = 0 \\ \langle a^4, {}^3P_1 b \rangle = -0,577 < 0 \end{cases}$$

$${}^4P_1 = I - a^4 [a^4]^+ = I - \frac{a^4 (a^4)^T}{\|a^4\|^2} \rightarrow {}^4P_1 b = \frac{1}{73} \begin{pmatrix} -9 \\ 108 \\ -9 \end{pmatrix} \rightarrow$$

$$\begin{cases} \langle a^1, {}^4P_1 b \rangle = 8,136 > 0 \\ \langle a^2, {}^4P_1 b \rangle = 5,178 > 0 \\ \langle a^3, {}^4P_1 b \rangle = 1,712 > 0 \\ \langle a^4, {}^4P_1 b \rangle = 0 \end{cases}$$

Temos assim que $L_2 = \{1, 4\}$

c) Como $L_2 = \{1, 4\}$, temos então duas aproximações do problema proposto

$$P_{11}: \begin{cases} \min c^T \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} \\ \text{su}j. A \begin{bmatrix} x_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{cases} = b_{11} \quad \text{onde } b_{11} = [a^1] x_1$$

$$P_{14}: \begin{cases} \min c^T \begin{bmatrix} 0 \\ 0 \\ 0 \\ x_4 \end{bmatrix} \\ \text{su}j. A \begin{bmatrix} 0 \\ 0 \\ 0 \\ x_4 \end{bmatrix} \end{cases} = b_{14} \quad \text{onde } b_{14} = [a^4] x_4$$

Portanto, através desses problemas selecionamos para ficar na base final, a^j , $j \in \{1, 4\}$ que nos der o menor valor para o funcional (a função objetiva). Desejamos assim encontrar

$\min \{c_1 x_1, c_4 x_4\}$ onde

$$x_1 = [a^1]^+ b = \frac{\langle a^1, b \rangle}{\|a^1\|^2}, \quad x_4 = [a^4]^+ b = \frac{\langle a^4, b \rangle}{\|a^4\|^2}$$

$$x_1 = \begin{matrix} 1 & 3 \\ 5 & 4 \end{matrix} \cdot \frac{1}{62} = \frac{35}{62}; \quad x_4 = \begin{matrix} 6 & 3 \\ 6 & 4 \end{matrix} \cdot \frac{1}{73} = \frac{44}{73}$$

$$\min\left\{\frac{4 \times 35}{62}, \frac{5 \times 44}{73}\right\} = \min\{2,258, 3,013\} = 2,258$$

Então o vetor a^1 vai para base final.

Step 2:

a) Cálculo de $[a^1, a^2]^+ b$, $[a^1, a^3]^+ b$, $[a^1, a^4]^+ b$

$$[a^1, a^2]^+ b = [(a^1, a^2)^T (a^1, a^2)]^{-1} [a^1, a^2]^T b = \frac{1}{344} \begin{bmatrix} -92 & 36 & 44 \\ 144 & -4 & -24 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -0,0813 \\ 0,9534 \end{bmatrix}$$

$$[a^1, a^3]^+ b = [(a^1, a^3)^T (a^1, a^3)]^{-1} [a^1, a^3]^T b = \frac{1}{1421} \begin{bmatrix} -140 & 196 & 77 \\ 273 & -98 & 63 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0,1970 \\ 0,6157 \end{bmatrix}$$

$$[a^1, a^4]^+ b = [(a^1, a^4)^T (a^1, a^4)]^{-1} (a^1, a^4)^T b = \frac{1}{2762} \begin{bmatrix} -179 & 396 & 113 \\ 330 & -190 & 162 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0,2559 \\ 0,4554 \end{bmatrix}$$

Temos assim que neste passo, k é igual a 2 ($k = 2$). Note que

$$[a^1, a^2][a^1, a^2]^+ b = -0,0813 a^1 + 0,9534 a^2, \text{ isto } \bar{e},$$

a projeção de b sobre o plano determinado por a^1 e a^2 , não é uma combinação linear não negativa de a^1 e a^2 . Tal ocorrência poderia ter sido constatado geometricamente através da Figura II.4.

b) Calculemos agora

$$\left\{ \begin{array}{l} \langle a^j, {}^{13}P_2 b \rangle \\ j = 2, 4 \end{array} \right. \text{ e } \left\{ \begin{array}{l} \langle a^j, {}^{14}P_2 b \rangle \\ j = 2, 3 \end{array} \right. , \text{ Vejamos:}$$

$${}^{13}P_2 b = [I - (a^1, a^3)(a^1, a^3)^+] b = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 5 & 4 \end{pmatrix}^{-1} \begin{bmatrix} -0,2758 \\ -0,4137 \\ 0,5517 \end{bmatrix} \rightarrow$$

$$\left\{ \begin{array}{l} \langle a^2, {}^{13}P_2 b \rangle = \langle (4), (-0,4137) \rangle = -0,8271 < 0 \\ \langle a^4, {}^{13}P_2 b \rangle = \langle (1), (-0,4137) \rangle = 1,2417 \end{array} \right.$$

$${}^{14}P_2 = [I - (a^1, a^4)(a^1, a^4)^+] = \frac{1}{2762} \begin{bmatrix} 961 & 744 & -1084 \\ 744 & 576 & -840 \\ -1085 & -840 & 1255 \end{bmatrix}$$

$${}^{14}P_2 b = \frac{1}{2762} \begin{bmatrix} 31 \\ 24 \\ 35 \end{bmatrix} = \begin{bmatrix} 0,0112 \\ 0,0086 \\ 0,0126 \end{bmatrix} \rightarrow$$

$$\rightarrow \left\{ \begin{array}{l} \langle a^2, {}^{14}P_2 b \rangle = 0,0302 \\ \langle a^3, {}^{14}P_2 b \rangle = 0,0228 \end{array} \right.$$

Como \bar{s}_0 existe um \bar{u} nico $L = 4$ tal que

$$\langle a^j, {}^{1\ell}P_2 b \rangle \geq 0 \quad \forall j \neq 1, \ell$$

Então o vetor a^4 vai para base final

Step 3

Devemos agora calcular o

$$\min\{c_{143}^T \cdot x_{143}, c_{142}^T \cdot x_{142}\}$$

$$\text{onde } x_{143} = (x_1, x_4, x_3)^T = [a^1, a^2, a^3]^+ b$$

$$x_{142} = (\bar{x}_1, \bar{x}_4, \bar{x}_2)^T = [a^1, a^4, a^2]^+ b$$

$$c_{143}^T = (c_1, c_4, c_3) = (4, 5, 8)$$

$$c_{142}^T = (c_1, c_4, c_2) = (4, 5, 1)$$

Mas por (II.4.3)

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} &= \begin{bmatrix} {}^3x_1 \\ {}^3x_4 \end{bmatrix} = {}^{14}A_2^+ [b - {}^3x_3 a^3] = \\ &= \frac{1}{2762} \begin{bmatrix} -179 & 396 & 113 \\ 330 & -190 & 162 \end{bmatrix} \begin{pmatrix} 2,9210 \\ 1,9684 \\ 3,9368 \end{pmatrix} = \begin{pmatrix} 0,2539 \\ 0,4444 \end{pmatrix} \end{aligned}$$

$$\text{onde } x_3 = {}^3x_3 = \frac{\langle a^3, {}^{14}P_2 b \rangle}{\langle a^3, {}^{14}P_2 a^3 \rangle} = \frac{0,0228}{1,4370} = 0,0158$$

Assim

$$x_{143} = \begin{bmatrix} x_1 \\ x_4 \\ x \end{bmatrix} = \begin{bmatrix} {}^3x_1 \\ {}^3x_4 \\ {}^3x_3 \end{bmatrix} = \begin{bmatrix} 0,2539 \\ 0,4444 \\ 0,0158 \end{bmatrix}$$

Proseguindo, por (II.4.3) temos

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_4 \end{bmatrix} = \begin{bmatrix} {}^3\bar{x}_1 \\ {}^3\bar{x}_4 \end{bmatrix} = {}^{14}A^+ [b - {}^3x_2 a^2]$$

$$= \frac{1}{2762} \begin{bmatrix} -179 & 396 & 113 \\ 330 & -190 & 162 \end{bmatrix} \begin{pmatrix} 2,9649 \\ 1,9532 \\ 3,9649 \end{pmatrix} = \begin{pmatrix} 0,2501 \\ 0,4524 \end{pmatrix}$$

$$\text{onde } x_2 = {}^3x_2 = \frac{\langle a^2, {}^{14}P_2 b \rangle}{\langle a^2, {}^{14}P_2 a^2 \rangle} = \frac{0,0302}{2,5611} = 0,0117$$

Portanto

$$x_{142} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_4 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} {}^3\bar{x}_1 \\ {}^3\bar{x}_4 \\ {}^3x_2 \end{bmatrix} = \begin{bmatrix} 0,2501 \\ 0,4524 \\ 0,0117 \end{bmatrix}$$

$$c_{143}^T x_{143} = \begin{matrix} 4 & 0,2539 \\ (5), & (0,4444) \\ 8 & 0,0158 \end{matrix} > = 3,3640$$

$$c_{142}^T x_{142} = \begin{matrix} 4 & 0,2501 \\ (5), & (0,4524) \\ 1 & 0,0117 \end{matrix} > = 3,2741$$

Observe também que

$$[a^1, a^4, a^3]^+ b = [a^1, a^4, a^3]^{-1} b = \frac{1}{63} \begin{bmatrix} -8 & 6 & 7 \\ -14 & -21 & 28 \\ 31 & 24 & -35 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0,2539 \\ 0,4444 \\ 0,0158 \end{bmatrix}$$

$$[a^1, a^4, a^2]^+ b = [a^1, a^4, a^2]^{-1} b = \frac{1}{84} \begin{bmatrix} -21 & 0 & 21 \\ 2 & -12 & 14 \\ 31 & 24 & -35 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0,2500 \\ 0,4523 \\ 0,0119 \end{bmatrix}$$

Portanto

$\min\{c_{143} x_{143}, c_{142} x_{142}\} = \min\{3,640, 3,2741\} = 3,2741$ o que nos leva a concluir que a base final \bar{e} é constituída dos vetores a^1, a^4, a^2

Vejamos agora que $z_3 - c_3 < 0$:

$$\frac{1}{84} \begin{bmatrix} -21 & 0 & 21 \\ 2 & -12 & 14 \\ 31 & 24 & -35 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} -21 \\ 42 \\ 63 \end{bmatrix} \rightarrow$$

$$[c_1, c_4, c_2]^T \begin{bmatrix} -21/84 \\ 42/84 \\ 63/84 \end{bmatrix} - c_3 = \frac{1}{84} \begin{pmatrix} -21 & 4 \\ 42 & 5 \\ 63 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix} - 8 = 2,25 - 8 < 0$$

Temos assim que o critério de otimalidade \bar{c} é satisfeito, e a solução ótima é:

$$x^* = (x_1^*, x_2^*, x_3^*, x_4^*)^T = (0,2500; 0,0119, 0, 0,4523)$$

$$z^* = 3,2741$$

No que diz respeito à teoria desenvolvida, vejamos os comentários no Capítulo IV, "Conclusões", página 117.

CAPÍTULO III

ALGORITMO DE KHACHIYAN

III.1 - INTRODUÇÃO

Nossa meta principal neste capítulo é apresentar de uma maneira didática, o recente algoritmo apresentado por Khachiyan [26] para resolver em tempo polinomial, um problema de programação linear. Visando tal propósito, seguiremos o roteiro apresentado por Aspvall, B. [3].

Como se perceber; no decorrer das seções subsequentes, a idéia inicial será a de estabelecer um critério para encontrar a solução de um sistema de inequações lineares estritas $Ax < b$, onde todos os elementos de A e b são inteiros, pros-

seguindo, estendê-lo para um sistema da forma $Ax \leq b$, e finalmente extrapolá-lo para a solução de um problema de programação linear. Seja

$$R = \{x | Ax < b\}$$

e $S \subset R$ um subconjunto de R .

Na Fig. III.1 como R é limitado, tomamos $S = R$.

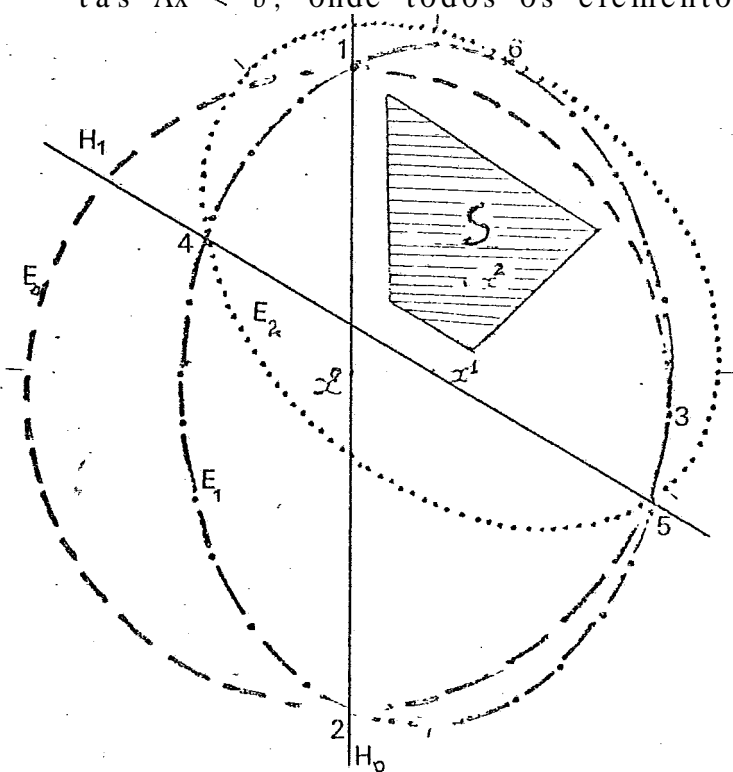


Fig. III.1

O algoritmo inicia com um elipsoide (esfera) E_0 , centrado na origem, com raio suficientemente grande de modo a conter S .

A linha geral do algoritmo é que em qualquer momento o elipsoide com o qual estivermos lidando, sempre conterá S .

A idéia geométrica do algoritmo, em qualquer iteração e a que descreveremos baseados na Fig. III.1. Se o centro do elipsoide não é satisfeito, como é o caso de x^0 , então um hiperplano H_0 paralelo a restrição não satisfeita em x^0 , e passando pelo centro de E_0 , é usado para cortar o elipsoide (esfera) E_0 pela metade. Como facilmente percebemos, uma metade de E_0 que chamaremos de semi-elipsoide, e totalmente inviável, e o outro semi-elipsoide nós circunscrevemos com o elipsoide E_1 passando pelos pontos 1, 2 e 3 e de volume mínimo. Note que nos pontos 3 e 6, os hiperplanos tangentes, são respectivamente paralelos aos hiperplanos H_0 e H_1 . Como x^1 , centro do elipsoide E_1 , não pertence ainda a S , então passamos novamente pelo centro de E_1 e paralelamente a uma restrição não satisfeita em x^1 , um hiperplano H_1 , dividindo-o em dois semi-elipsoides. Novamente envolvemos o semi-elipsoide contendo S , por um novo elipsoide de volume mínimo (e único) passando o pelos pontos 4, 5 e 6. Observamos na nossa Fig. III.1, que o centro x^2 do elipsoide E_2 já pertence a S . Caso contrário, repetiríamos o procedimento.

III.2 - SISTEMAS DE INEQUAÇÕES LINEARES ESTRITAS

No decorrer desta seção teremos como meta principal apresentar um algoritmo o qual visa detectar se o sistema de inequações lineares estritas é solúvel. Além do mais, o algoritmo é construtivo dando-nos o vetor viável quando ele existe.

Com

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &< b_i \\ i &= 1 \dots m, m \geq 2, n \geq 2 \end{aligned} \quad (\text{III.2.1})$$

Queremos denotar o sistema de inequações lineares estritas com coeficientes inteiros. Percebe-se que no pior caso tal algoritmo requer $O(n^3(m+n)L)$ operações (+, -, \times , \div , e \vee), onde

$$L = \sum_{i,j} [\log_2(|a_{ij}|+1)] + \sum_i [\log_2(|b_i|+1)] + [\log_2 mn] + 1 \quad (\text{III.2.2})$$

É o número de símbolos 0 e 1 que são necessários para escrever (III.2.1) em código binário, e $[a]$ significa o menor inteiro maior ou igual a a . O mesmo problema proposto na forma matricial é:

$$Ax < b \quad (\text{III.2.3})$$

onde $A \in Z^{m \times n}$ e $b \in Z^m$. Quando quisermos lidar, com uma particu

lar inequação de (III.2.1), usaremos a representação

$$a_i^T x < b_i \quad (\text{III.2.4})$$

onde $0 \neq a_i \in \mathbb{Z}^n$ e $b_i \in \mathbb{Z}$. (Percebe-se assim que todos nossos vetores são vetores coluna).

Matematicamente, o algoritmo de Khachiyan soluciona (III.2.1), construindo uma sequência de matrizes $B^{(k)}$, $k = 0, 1, 2, \dots$ cada uma de dimensão $n \times n$, as quais são usadas para criar a sequência de soluções x_k , $k = 0, 1, \dots$ a qual converge para a solução de (III.2.1) (se ela existe).

Algoritmo (Khachiyan Algoritmo)

Passo 1 (Início) Faça $x^{(0)} \leftarrow 0$, $B^{(0)} \leftarrow 2^{2L} I$, e $k \leftarrow 0$

Passo 2 (Terminou ?) Se $x^{(k)}$ é solução para (III.2.1) o algoritmo para dando-nos a solução viável $x^{(k)}$. Se a solução não foi ainda encontrada e $k < k^* = 4(n+1)^2 L$, vá para o passo 3. Caso contrário o algoritmo para afirmando que tal solução não existe.

Passo 3 (Próxima Iteração) Escolha uma inequação em (III.2.4) que não seja satisfeita em $x^{(k)}$ isto é, $a_i^T x^{(k)} \geq b_i$ e faça

$$x^{(k+1)} \leftarrow x^{(k)} - \alpha \frac{B^{(k)} a_i}{\sqrt{a_i^T B^{(k)} a_i}} \quad (\text{III.2.5})$$

e

$$B^{(k+1)} \leftarrow \frac{1}{\beta} (B^{(k)} - \gamma \frac{(B^{(k)} a_i)(B^{(k)} a_i)^T}{a_i^T B^{(k)} a_i}) \quad (\text{III.2.6})$$

Faça $k \leftarrow k+1$ e vá para passo 2.

Evidentemente, se o algoritmo encontra um $x^{(k)}$, então (III.2.1) é solúvel. Posteriormente determinaremos que os valores α , β , γ e k^* são os que seguem.

$$\alpha = \frac{1}{n+1}, \quad n \geq 2 \quad (\text{III.2.7})$$

$$\beta = \frac{n^2-1}{n^2}, \quad n \geq 2 \quad (\text{III.2.8})$$

$$\gamma = \frac{2}{n+1}, \quad n \geq 2 \quad (\text{III.2.9})$$

$$k^* = 4(n+1)^2, \quad n \geq 2 \quad (\text{III.2.10})$$

Prosseguiremos com uma série de lemas com o objetivo de demonstrar, que uma parada do algoritmo, significa que $x^{(k)}$ é uma solução de (III.2.1), e que se o mesmo não para quando $k > k^*$, então (III.2.1) não é solúvel.

Sejam x' um ponto do R^n e B uma matriz simétrica definida positiva, e com

$$E = \{x \mid (x-x')^T B^{-1} (x-x') \leq 1\} \quad (\text{III.2.11})$$

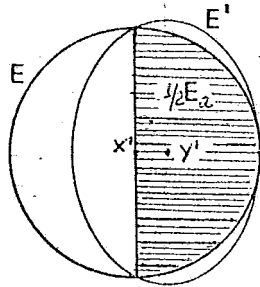


Fig. III.2

queremos definir o elipsoide com centro x' . Denotaremos com $1/2 E_a$ o semi elipsoide

$$\frac{1}{2} E_a = E \cap \{x \mid a^T (x-x') \leq 0\} \quad (\text{III.2.12})$$

e com E' o conjunto $\{y \mid (y-y')^T B'^{-1} (y-y') \leq 1\}$, onde

$$y' = x' - \alpha \frac{Ba}{\sqrt{a^T Ba}} \quad (\text{III.2.13})$$

$$B' = \frac{1}{\beta} \left(B - \gamma \frac{(Ba)(Ba)^T}{a^T B a} \right) \quad (\text{III.2.14})$$

$$a^T = (-1, 0 \dots 0)$$

Chamamos atenção dos leitores que, embora nossas demonstrações sejam no R^n , as ilustrações sempre serão no R^2 . Na ilustração Fig. III.2 estamos admitindo $B = I$ e $x' = (0, 0)$, $y' = (a, 0)$.

Mostraremos que E' é um elipsoide com centro y' (isto é, B' é uma matriz simétrica positiva definida), que $\frac{1}{2} E_a \subseteq E'$, e que $\lambda(E')$ o volume de E' é menor que $\lambda(E)$.

O lema 1 que segue nos assegurará que, se (III.2.1) tem soluções, parte destas soluções estará contida numa esfera E_0 de raio 2^L , centrada na origem.

Lema 1: Todo vértice V do poliedro

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned} \tag{III.2.15}$$

satisfaz $\|V\|_\infty < 2^L/m.n.$ Além do mais suas coordenadas são números racionais com denominadores menor que $2^L/mn$ em valor absoluto.

Prova: Consideremos o vértice genérico $V = (v_1, \dots, v_k, \dots, v_n)$ do poliedro (III.2.15). A ele está associada uma matriz básica cujo determinante denotaremos por D . Através da regra de Cramer, sabemos que cada componente v_k ou é zero ou pode ser expresso como D_k/D onde $D, D_k \neq 0$ são determinantes de matrizes cujos elementos pertencem ao conjunto $\{0, 1, a_{ij}, b_i\}$. Como a_{ij} e $b_i \in \mathbb{Z}$, concluímos que D e D_k são inteiros ($|D| \geq 1$). Note que quando acrescentamos as variáveis de folga para transformar o sistema de inequações (III.2.15) em um sistema de equações, automaticamente criamos a possibilidade de termos entre as colunas das matrizes associadas com D e D_k vetores unitários.

rios.

Portanto, denotaremos com d_j para $1 \leq j \leq m$ os vetores colunas das matrizes associadas com D e D_k , isto é, d_j assume uma das três possibilidades.

$d_j = (a_{1j}, \dots, a_{mj})$, $d_j = b$, $j \in J \subseteq \{1, \dots, m\}$ ou d_j é um vetor unitário para $j \notin J$.

A demonstração estará baseada nas afirmações que seguem:

$$I) \quad |D| \leq \prod_{1 \leq j \leq n} \|d_j\|_2$$

$$II) \quad \|d_j\|_2 \leq \prod_{1 \leq i \leq m} (|a_{ij}| + 1)$$

A afirmação (I) é uma consequência direta da desigualdade de Hadamard [14, pp. 252]. Vejamos agora (II):

Sabemos que se $x = (x_1, x_2)$ então

$$\sqrt{x_1^2 + x_2^2} \leq \sqrt{x_1^2} + \sqrt{x_2^2} = |x_1| + |x_2| \leq |x_1| + |x_2| + |x_1||x_2| + 1$$

isto é

$$\sqrt{x_1^2 + x_2^2} \leq (|x_1| + 1)(|x_2| + 1)$$

Suponhamos agora que se $x = (x_1, x_2, \dots, x_k)$ então

$$\sqrt{x_1^2 + \dots + x_k^2} \leq \prod_{i=1}^k (|x_i| + 1)$$

Mostraremos então que a propriedade acima é preservada se $x = (x_1, \dots, x_k, x_{k+1})$

$$\|x\|_2 = \sqrt{x_1^2 + \dots + x_k^2 + x_{k+1}^2} \leq \sqrt{x_1^2 + \dots + x_k^2} + \sqrt{x_{k+1}^2} \leq \prod_{i=1}^k (|x_i| + 1) + |x_{k+1}|$$

$$\|x\|_2 \leq \prod_{i=1}^k (|x_i| + 1) + |x_{k+1}| \prod_{i=1}^k (|x_i| + 1)$$

pois $\prod_{i=1}^k (|x_i| + 1) \geq 1$. Então

$$\|x\|_2 \leq \prod_{i=1}^k (|x_i| + 1)(|x_{k+1}| + 1) = \prod_{i=1}^{k+1} (|x_i| + 1)$$

Portanto

$$|D| \leq \prod_{1 \leq j} \|d_j\|_2 \leq \prod_{j \in J} \prod_{1 \leq i \leq m} (|a_{ij}| + 1) = \frac{2^L}{2m \cdot n \prod_{1 \leq i \leq m} (|b_i| + 1)} \leq \frac{2^L}{mn}$$

tendo-se em vista que por (III.2.2)

$$2^L = 2 \prod_{ij} (|a_{ij}| + 1) \prod_{1 \leq i \leq m} (|b_i| + 1) mn$$

o que nos dá

$$\prod_{ij} (|a_{ij}| + 1) = \frac{2^L}{2mn \prod_{1 \leq i \leq m} (|b_i| + 1)}$$

Prosseguindo, concluímos que qualquer que seja k ,

$$|v_k| = \frac{|D_k|}{|D|} \leq |D_k| \leq 2^L/m.n,$$

Assim $\|V\|_\infty \leq 2^L/m.n$

Corolário 1: Todo vértice V do politopo

$$Ax \leq b$$

$$x \geq 0$$

$$x \leq \lceil 2^L/n \rceil e$$

onde $e = (1, \dots, 1)^T$, tem coordenadas que são números racionais com denominadores em valor absoluto menor que $2^L/mn$.

Prova: Como já chamamos atenção, cada coordenada v_k é zero ou pode ser expressa como D_k/D , onde $|D| \leq \prod_{j \in J} \|d_j\|_2$. Uma observação adicional que se faz necessário é que agora d_j é uma coluna da matriz $(A^T, I)^T$ em lugar de A . Denotemos por d_i^j para $1 \leq i \leq m+n$, a linha da matriz cujas colunas são d_j , $j \in J$. Consequentemente

$$|D| \leq \prod_{1 \leq i \leq m+n} \|d_i^j\|_2 = \prod_{1 \leq i \leq m} \|d_i^j\|_2 < 2^L/mn$$

Lema 2: Se (III.2.1) tem solução, então o volume do espaço de solução interior a esfera $\|x\|_2 \leq 2^L$ é no mínimo $2^{-(n+1)L}$

Prova: Admitiremos para maior clareza, sem perda de generalidade, que (III.2.1) tem uma solução no octante positivo de modo que o poliedro

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned} \tag{III.2.16}$$

tem um ponto interior.

Como o poliedro não contém reta estendida infinitamente em ambas direções (simultaneamente), ele tem também um vértice V . Sabemos pelo Lema 1, que $\|V\|_{\infty} < \lceil 2^L/n \rceil$. Assim (III.2.16) tem um ponto interior x^* com $\|x^*\|_{\infty} < \lceil 2^L/n \rceil$, e conseqüentemente também o tem o politopo

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \\ x &\leq \lceil 2^L/n \rceil e \end{aligned} \tag{III.2.17}$$

onde $e = (1, 1, \dots, 1)^T$

Exemplo:

$$\begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} -1 \\ +2 \end{bmatrix}$$

$$x \geq 0$$

$$x_1 \leq \lceil 2^L/n \rceil, \quad x_2 \leq \lceil 2^L/n \rceil$$

Então por (III.2.2)

$$2^L = 2 \prod_{\substack{1 \leq i \leq 2 \\ 1 \leq j \leq 2}} (|a_{ij}|+1) \prod_{1 \leq i \leq 2} (|b_i|+1)_{m,n}$$

$$2^L = 2(2.2)(2.3)4 = 192 \rightarrow \lceil 2^L/n \rceil = 96$$

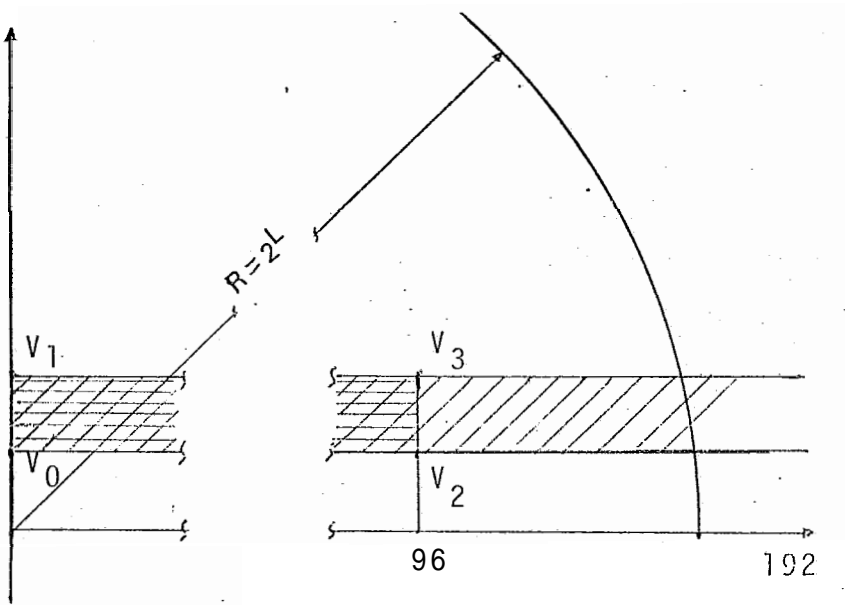


Fig. 111.3

O politopo (III.2.17) deve por essa razão ter pelo menos $n+1$ vértices V_0, \dots, V_n (Ex. V_0, V_1, V_2) os quais não estão simultaneamente sobre um hiperplano e seu volume é no mínimo [49 - pp. 124]

$$\frac{1}{n!} \left| \det \begin{pmatrix} 1 & \dots & 1 \\ V_0 & \dots & V_n \end{pmatrix} \right| > 0 \quad (\text{III.2.18})$$

o qual representa o volume do simplex com vértices V_0, \dots, V_n . Admitamos por um momento que $V_i \in E^3$. Então $V_i = (v_1, v_2, v_3)$ e a regra de Cramer nos dá

$$\left. \begin{array}{l} v_1 = D_1/D_i \\ v_2 = D_2/D_i \\ v_3 = D_3/D_i \end{array} \right\} \rightarrow V_i = \frac{1}{D_i} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \frac{\mu_i}{D_i} \text{ para } \mu_i = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}$$

Voltando ao corolário I temos que μ_i é um vetor de coordenadas inteiras, e D_i é um inteiro $< 2^L/n$. Portanto, fazendo

$$V_i = \frac{1}{D_i} \mu_i \quad . \text{ Temos}$$

$$\begin{aligned} \left| \det \begin{pmatrix} 1 & \dots & 1 \\ v_0 & & v_n \end{pmatrix} \right| &= \left| \det \begin{pmatrix} 1 & \dots & 1 \\ \mu_0/D_0 & & \mu_n/D_n \end{pmatrix} \right| = \frac{1}{|D_0| \dots |D_n|} \left| \det \begin{pmatrix} D_0 & \dots & D_n \\ \mu_0 & & \mu_n \end{pmatrix} \right| \\ &\geq \frac{1}{|D_0| \dots |D_n|} \geq \left(\frac{2^L}{n} \right)^{-(n+1)} \end{aligned}$$

$$\text{Já que } |D_i| > 0 \text{ e } \left| \det \begin{pmatrix} D_0 & \dots & D_n \\ \mu_0 & & \mu_n \end{pmatrix} \right| > 1$$

Assim o volume do politopo (III.2.17) é no mínimo

$$\left(\frac{2^L}{n} \right)^{-(n+1)} / n! > 2^{-(n+1)L}$$

Pois $\frac{n^{n+1}}{n!} > 1$.

Tendo-se em vista que o conteúdo dos três lemas subsequentes serem invariantes sob transformação afim, por esta razão admitiremos $x' = 0$, $B = I$ e $a = (-1, 0, \dots, 0)^T$ na prova dos mesmos; isto é o elipsoide E é a esfera unitária centrada na origem e interseptando o semiespaço $x_1 \geq 0$.

Em seguida daremos uma proposição a qual nos habilitará vislumbrar a razão porque os lemas 3, 4 e 5 são invariantes sobre transformações afim.

Proposição 1: Sejam B uma matriz simétrica definida positiva, $a \neq 0$ e x' vetor n -dimensional, e sejam α , β e $\gamma < 1$ constantes positivas.

Defina:

$$E = \{x \mid x^T x \leq 1\}$$

$$\bar{E} = \{x \mid (x-x')^T B^{-1} (x-x') \leq 1\}$$

Faça: $V = (-1, 0, \dots, 0)^T$ e defina

$$E' = \{x \mid (x+\alpha V)^T \beta (I - \gamma (VV^T))^{-1} (x + \alpha V) \leq 1\}$$

$$\bar{E}' = \{x \mid (x-x' + \alpha \frac{Ba}{\sqrt{a^T Ba}})^T \beta (B - \gamma \frac{(Ba)(Ba)^T}{a^T Ba})^{-1} (x-x' + \alpha \frac{Ba}{\sqrt{a^T Ba}}) \leq 1\}$$

Então a transformação afim que leva E para \hat{E} também remete E' para \bar{E}' .

Prova: Como B é uma matriz simétrica definida positiva, existe uma fatoração de Cholesky para $B = LL^T$, onde L é uma matriz não singular. Denotemos por Q uma matriz ortogonal (rotação) de tal modo que $v = Q^T L^T \cdot a / \|Q^T L^T a\|_2$ (observe que $\|Q^T L^T a\|_2^2 = a^T L Q Q^T L^T a = a^T L L^T a = a^T B a$). Definamos agora a aplicação afim

$$F(x) = x' + L Q x \quad \rightarrow$$

$$x = F[F^{-1}(x)] = x' + L Q F^{-1}(x) \quad \rightarrow$$

$$x - x' = L Q F^{-1}(x) \quad \rightarrow$$

$$[LQ]^{-1} [x - x'] = F^{-1}(x)$$

Portanto $F(x)$ define uma transformação afim inversível, com $F^{-1}(x) = Q^T L^{-1}(x - x')$. Segue-se assim que,

$$\begin{aligned} F(E) &= \{x \mid (F^{-1}(x))^T F^{-1}(x) \leq 1\} \\ &= \{x \mid [(LQ)^{-1}(x-x')]^T [(LQ)^{-1}(x-x')] \leq 1\} \\ &= \{x \mid (x-x')^T (L^T)^{-1} Q Q^T L^{-1}(x-x') \leq 1\} \\ &= \{x \mid (x-x')^T (L^T)^{-1} L^{-1}(x-x') \leq 1\} \\ &= \{x \mid (x-x')^T B^{-1}(x-x') \leq 1\} \quad \underline{\underline{E}} \end{aligned}$$

Mostremos agora que $F(x)$ também leva E' em \bar{E}' isto é $F(E') = \bar{E}'$. Vejamos:

$$\begin{aligned}
 F(E') &= \{x | (F^{-1}(x) + \alpha V)^T \beta [I - \gamma(VV^T)]^{-1} (F^{-1}(x) + \alpha V) \leq 1\} \\
 &= \{x | (Q^T L^{-1}(x - x') + \alpha \frac{Q^T L^T a}{\|Q^T L^T a\|_2})^T \beta (I - \gamma \frac{(Q^T L^T a)(Q^T L^T a)^T}{\|Q^T L^T a\|_2^2})^{-1} \\
 &\quad \cdot (Q^T L^{-1}(x - x') + \frac{\alpha Q^T L^T a}{\|Q^T L^T a\|_2}) \leq 1\}
 \end{aligned}$$

$$F(E') = \{x | [Q^T L^{-1}(x - x' + \frac{\alpha L Q Q^T L^T a}{\sqrt{a^T B a}})]^T \beta [I - \frac{\gamma (Q^T L^T a)(a^T L Q)}{a^T B a}]^{-1}$$

$$\cdot [Q^T L^{-1}(x - x' + \frac{\alpha L Q Q^T L^T a}{\sqrt{a^T B a}})] \leq 1\}$$

$$= \{x | (x - x' + \frac{\alpha B a}{\sqrt{a^T B a}})^T (Q^T L^{-1})^T \beta [I - \gamma \frac{(Q^T L^T a)(a^T L Q)}{a^T B a}]^{-1} (Q^T L^{-1})$$

$$\cdot [x - x' + \frac{\alpha B a}{\sqrt{a^T B a}}] \leq 1\}$$

$$= \{x | (x - x' + \alpha \frac{B a}{\sqrt{a^T B a}})^T \beta [(Q^T L^{-1})^{-1} (I - \gamma \frac{(Q^T L^T a)(a^T L Q)}{a^T B a}) (L^{T-1} Q)^{-1}]^{-1}$$

$$\cdot (x - x' + \alpha \frac{B a}{\sqrt{a^T B a}}) \leq 1\}$$

$$= \{x | (x - x' + \alpha \frac{B a}{\sqrt{a^T B a}})^T \beta [L Q Q^T L^T - \gamma \frac{(L Q Q^T a)(a^T L Q Q^T L^T)}{a^T B a}]^{-1} x$$

$$\cdot (x - x' + \alpha \frac{B a}{\sqrt{a^T B a}}) \leq 1\}$$

$$F(E') = \left\{ x \mid \left(x - x' + \alpha \frac{Ba}{\sqrt{a^T Ba}} \right)^T \left[B - \gamma \cdot \frac{(Ba)(Ba)^T}{a^T Ba} \right]^{-1} \left(x - x' + \frac{\alpha Ba}{\sqrt{a^T Ba}} \right) \leq 1 \right\} \triangleq \bar{E}'$$

Logo $F(E') = \bar{E}'$ (III.2.19)

Mostremos agora que $F_1(E) = E'$. cnde:

$$\left. \begin{array}{l} F_1(x) = x' + L_1 x \\ B' = L_1^T L_1 \\ x' = (\alpha, 0, \dots, 0) \\ V = (-1, 0, \dots, 0) \\ B' = \text{diag} \left[\frac{1-\gamma}{\beta}, \frac{1}{\beta}, \dots, \frac{1}{\beta} \right] \end{array} \right\} \rightarrow \left. \begin{array}{l} F_1^{-1}(x) = L_1^{-1}(x - x') \\ [B']^{-1} = [L_1^T]^{-1} L_1^{-1} = \beta [I - \gamma V V^T]^{-1} \\ x' = \alpha(1, 0, \dots, 0) = -\alpha V \\ B' = \frac{1}{\beta} \text{diag} [(1-\gamma), 1, \dots, 1] = \frac{1}{\beta} [\text{diag}(1, \dots, 1) - \gamma \text{diag}(1, 0, \dots, 0)] \\ B' = \frac{1}{\beta} [I - \gamma V V^T] \end{array} \right.$$

Vejamos:

$$F_1(E) = \{ x \mid [F_1^{-1}(x)]^T [F_1^{-1}(x)] \leq 1 \}$$

$$F_1(E) = \{ x \mid [L_1^{-1}(x + \alpha V)]^T [L_1^{-1}(x + \alpha V)] \leq 1 \}$$

$$F_1(E) = \{ x \mid (x + \alpha V)^T [L_1^T]^{-1} L_1^{-1} (x + \alpha V) \leq 1 \}$$

$$F_1(E) = \{ x \mid (x + \alpha V)^T [B']^{-1} (x + \alpha V) \leq 1 \}$$

$$F_1(E) = \{ x \mid (x + \alpha V)^T \beta [I - \gamma V V^T]^{-1} (x + \alpha V) \leq 1 \} \triangleq E'$$

Isto é $F_1(E) = E'$, mas por (III.2.19) $F(E') = \bar{E}'$, portanto

$F[F_1(E)] = F[E'] = \bar{E}'$, então

$$F[F_1(E)] = \bar{E}' = \{x \mid (x - x' + \frac{Ba}{\sqrt{a^T Ba}})^T \beta (B - \gamma \frac{(Ba)(Ba)^T}{a^T Ba})^{-1} (x - x' + \frac{a - Ba}{\sqrt{a^T Ba}}) \leq 1\}$$

$$\bar{E}' = \{x \mid [x - x^*]^T B^{*-1} [x - x^*] \leq 1\} \quad \text{onde}$$

$$x^* = x' - \frac{\alpha Ba}{\sqrt{a^T Ba}} \quad (\text{III.2.20})$$

$$B^* = \frac{1}{\beta} \left[B - \frac{\gamma (Ba)(Ba)^T}{a^T Ba} \right] \quad (\text{III.2.21})$$

São justamente as fórmulas de recorrência do algoritmo de Khachiyan. Nós observamos assim que as formulas anteriormente citadas definem num sistema geral afim, a esfera E , e o elipsoide derivado E' .

Isto é completamente geral quando um elipsoide com volume positivo puder ser expresso como a transformação afim inversível da esfera unitária. Visto que a relação " B' é "definida positiva" e " $\frac{1}{2} E_a \subseteq E'$ " são invariantes sobre transformações afim, como o é a razão entre os volumes $\lambda(E')/\lambda(E)$, não perderemos nenhuma generalidade ao trabalhar com $x' = 0$, $B = I$ e $a = (-1, 0, \dots, 0)^T$.

Lema 3: Se a matriz B é simétrica definida positiva, então B' também o é. (Isto é o conjunto E' define um elipsoide).

Prova: Por (III.2.21)

$B' = \text{diag}(\frac{1-\gamma}{\beta}, \frac{1}{\beta}, \dots, \frac{1}{\beta})$ que é evidentemente si-

metrica. Resta-nos mostrar que $\mathbf{x}^T B' \mathbf{x} > 0$ para todo $\mathbf{x} \neq 0$. Pros-
seguindo

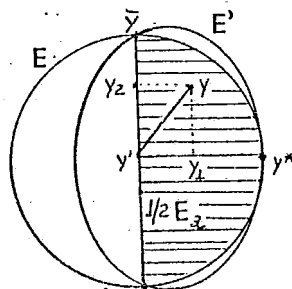
$$\mathbf{x}^T B' \mathbf{x} = \left(\frac{1-\gamma}{\beta}\right) x_1^2 + \frac{1}{\beta} (\|\mathbf{x}\|_2^2 - x_1^2) = \frac{1}{\beta} (\|\mathbf{x}\|_2^2 - \gamma x_1^2) > 0$$

pois $\gamma x_1^2 < x_1^2 \leq \|\mathbf{x}\|_2^2$, para $\gamma < 1$

Desejamos agora que os pontos $\mathbf{y}^* = (1, 0, \dots, 0)^T$
e $\bar{\mathbf{y}} = (0, \dots, 0, 1)^T$ estejam em E e E' . Para que tal ocorra
devemos exigir

$$(\mathbf{y}^* - \mathbf{y}')^T (B')^{-1} (\mathbf{y}^* - \mathbf{y}') = 1 \quad (\text{III.2.22})$$

$$(\bar{\mathbf{y}} - \mathbf{y}')^T (B')^{-1} (\bar{\mathbf{y}} - \mathbf{y}') = 1 \quad (\text{III.2.23})$$



Seja y um ponto genérico de E' tal que

$$[(y_1^{-\alpha}), y_2, \dots, y_n] \begin{bmatrix} \frac{\beta}{1-\gamma} & & 0 \\ & \beta & \\ 0 & & \beta \end{bmatrix} \begin{bmatrix} y_1^{-\alpha} \\ y_2 \\ y_n \end{bmatrix} = 1$$

$$(y_1^2 - 2\alpha y_1 + \alpha^2) \frac{\beta}{1-\gamma} + \beta \|y\|_2^2 - \beta y_1^2 = 1 \quad (\text{III.2.23})$$

$$\frac{(y_1^2 - 2\alpha y_1 + \alpha^2)}{1-\gamma} + \|y\|_2^2 - y_1^2 = \frac{1}{\beta} \quad (\text{III.2.24})$$

Para que (III.2.22) seja satisfeito basta fazer em (III.2.24) $y = y^* = (1, 0, \dots, 0)^T$. Consequentemente temos

$$\frac{1-2\alpha + \alpha^2}{1-\gamma} = \frac{1}{\beta} \quad (\text{III.2.25})$$

De maneira análoga, para que (III.2.23) seja satisfeita devemos ter

$$\frac{\alpha^2}{1-\gamma} + 1 = \frac{1}{\beta} \quad (\text{III.2.26})$$

Por (III.2.25) e (III.2.26) temos

$$\frac{1-2\alpha + \alpha^2}{1-\gamma} = \frac{\alpha^2 - \gamma + 1}{1-\gamma}$$

Assim devemos ter

$$\gamma = 2\alpha \quad (\text{III.2.27})$$

Lema 4: O semi-elipsoide $1/2 E_a$ esta completamente contido no elipsoide E' .

Próva: Devemos provar que $\forall y \in \frac{1}{2} E_a \rightarrow y \in E'$. Se $y \in \frac{1}{2} E_a$, então $\|y\|_2 \leq 1$ e $0 \leq -a^T y = y_1 \leq 1$ onde $a = (-1, 0, \dots, 0)$. Por (III.2.24).

$$\begin{aligned} (y')^T (B')^{-1} (y - y') &= \beta \frac{(y_1^2 - 2\alpha y_1 + \alpha^2)}{1 - \gamma} + \beta (\|y\|_2^2 - y_1^2) \beta = \\ &= (y_1^2 \frac{\beta}{1 - \gamma} - \beta y_1^2) - \frac{2\alpha\beta}{1 - \gamma} y_1 + \beta (\|y\|_2^2 - y_1^2) \beta = \\ &= y_1^2 \beta \left(\frac{\gamma}{1 - \gamma} \right) - \frac{2\alpha\beta}{1 - \gamma} y_1 + \beta (\|y\|_2^2 - 1) + \beta \left(\frac{1 - \gamma - \alpha^2}{1 - \gamma} \right) = \\ &= \frac{y_1 \beta}{1 - \gamma} (y_1 \gamma - 2\alpha) + \beta (\|y\|_2^2 - 1) + \beta \cdot \frac{1}{\beta} = \\ &= \frac{\beta \gamma}{1 - \gamma} y_1 (y_1 - 1) + \beta (\|y\|_2^2 - 1) + 1 \end{aligned} \quad (\text{III.2.28})$$

Pois por (III.2.25) e (III.2.27) temos respectivamente

$$\frac{1}{\beta} = \frac{1 - \gamma - \alpha^2}{1 - \gamma}$$

$$\gamma = 2\alpha$$

Como em (III.2.28) $\frac{\beta\gamma}{1-\gamma} > 0$, $(y_1 - 1) \leq 0$, $\beta > 0$ e

$(\|y\|_2^2 - 1) \leq 0$ concluímos que $(y-y')^T (B')^{-1} (y-y') \leq 1$, isto é $y' \in E'$.

Mostremos agora que o ponto onde E tangência o hiperplano $a^T x = d$ é dado por

$$x'' = x' - Ba / \sqrt{a^T Ba} \quad (\text{III.2.29})$$

Seja E dado por

$$\{y \mid (y-x')^T B^{-1} (y-x') \leq 1\}$$

e chamemos $g(y) = (y-x')^T B^{-1} (y-x') - 1 = 0$ (III.2.30)

É fácil ver que o gradiente de $g(y)$ no ponto x'' é dado por

$$\text{vg}(x'') = B^{-1}(x''-x') + [(x''-x')^T B^{-1}]^T$$

Este vetor deve ser igual a $-xa$, isto é

$$B^{-1}(x''-x') + [(x''-x')^T B^{-1}]^T = -\lambda a$$

Como B é simétrica então

$$2B^{-1}(x''-x') = -\lambda a$$

$$x'' - x' = -\lambda \frac{Ba}{2}$$

Voltando a (III.2.30) temos

$$(x'' - x')^T B^{-1} (x'' - x') = \frac{1}{2} (Ba)^T B^{-1} \frac{\lambda}{2} (Ba) = 1$$

$$\frac{\lambda^2 a^T B B^{-1} B}{4} a = 1 \rightarrow \lambda = \frac{2}{\sqrt{a^T B a}}$$

Assim

$$x'' - x' = \left(\frac{2}{\sqrt{a^T B a}} \right) \left(-\frac{Ba}{2} \right) = \frac{-Ba}{\sqrt{a^T B a}}$$

$$x'' = x' - \frac{Ba}{\sqrt{a^T B a}} \quad (\text{III.2.31})$$

Prosseguindo, vamos mostrar que se existirem dois elipsóides E' e E'' de mesmo volume, interseptando o hiperplano $a^T(x - x') = 0$ nos mesmos pontos como E o faz, e tangenciando o hiperplano $a^T x = d$ em x'' , então $E' = E''$.

Suponhamos que $E' \neq E''$ e que as suas elipses sejam dadas respectivamente por

$$(y - x')^T B'^{-1} (y - x') = 1 \rightarrow \begin{cases} g(y) = (y - x')^T B'^{-1} (y - x') - 1 = 0 \\ \nabla g(y) = 2B'^{-1} (y - x') \end{cases}$$

$$(y - x'')^T B''^{-1} (y - x'') = 1 \rightarrow \begin{cases} G(y) = (y - x'')^T B''^{-1} (y - x'') - 1 = 0 \\ \nabla G(y) = 2B''^{-1} (y - x'') \end{cases}$$

No ponto $y = x''$ temos por analogia com a equação (III.2.31) que

$$x'' - x' = - \frac{B'a}{\sqrt{a^T B a}} \rightarrow x'' = x' - (B'a) / \sqrt{a^T B a}$$

$$x'' - x^* = - \frac{B^*a}{\sqrt{a^T B^* a}} \rightarrow x'' = x^* - (B^*a) / \sqrt{a^T B^* a}$$

Como por (III.2.20)

$$x' = x^0 - a \frac{B'a}{\sqrt{a^T B'a}} \quad (\text{III.2.32})$$

$$x^* = x^0 - \alpha \frac{B^*a}{\sqrt{a^T B^*a}} \quad (\text{III.2.33})$$

Segue-se que

$$x'' = \left(x^0 - \frac{\alpha B^*a}{\sqrt{a^T B^*a}} \right) - \frac{B'a}{\sqrt{a^T B'a}} = \left(x^0 - \frac{\alpha B^*a}{\sqrt{a^T B^*a}} \right) - \frac{B^*a}{\sqrt{a^T B^*a}}$$

Portanto

$$(\alpha+1) \frac{B'(-a)}{\sqrt{a^T B'a}} = (\alpha+1) \frac{B^*a}{\sqrt{a^T B^*a}}$$

Agora por (III.2.37) e (III.2.33) concluímos que

$$x' = x^*$$

Observamos ainda que os semi-eixos b e b^* são iguais, pois

$$\left. \begin{aligned} b &= \|x'' - x'\| = \left\| \frac{B'a}{\sqrt{a^T B'a}} \right\| \\ b^* &= \|x'' - x^*\| = \left\| \frac{B^*a}{\sqrt{a^T B^*a}} \right\| \end{aligned} \right\} \rightarrow b = b^*$$

Conseqüentemente os semi-eixos a e a^* são diferentes visto $E' \neq E''$.

Verificaremos no caso de R^2 que, como o volume do elipsóide é a área da elipse ($s = \pi ab$), temos por hipótese

$$s = s^*$$

$$\pi ab = \pi a^* b$$

Como $b = b^*$, então $a^* = a$ o que contraria a hipótese.

Mostraremos logo mais, através do Lema 5, que em dois passos consecutivos do algoritmo os volumes dos elipsóides E' e E estão relacionados por $\lambda(E') = c(\alpha) \lambda(E)$, onde

$$c(\alpha) = (1-\alpha) \left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}}$$

Determinaremos agora o valor de α para que $\lambda(E')$ seja m̄nimo. Segue-se assim, que $\lambda(E')$ ser̄ m̄nimo quando $c(\alpha)$ o for. Vejamos

$$c'(\alpha) = \frac{dc(\alpha)}{d\alpha} = (-1) \left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}} + (1-\alpha) \frac{n-1}{2} \left[\frac{2(1-\alpha)(-1)(1-2\alpha) - (1-\alpha)^2(-2)}{(1-2\alpha)^2} \right] \left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}} - 1$$

$$c'(\alpha) = \left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}} \left[-1 + \frac{n-1}{2} (1-\alpha)^2 \left(\frac{-2(1-2\alpha) + 2(1-\alpha)}{(1-2\alpha)^2} \right) \left(\frac{(1-\alpha)^2}{1-2\alpha} \right)^{-1} \right]$$

$$c'(\alpha) = \left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}} \left[-1 + \frac{(n-1)\alpha}{1-2\alpha} \right]$$

Como $\left[\frac{(1-\alpha)^2}{1-2\alpha} \right]^{\frac{n-1}{2}} \neq 0$ ent̄o para que $c'(\alpha) = 0$ devemos ter

$$-1 + \frac{(n-1)\alpha}{1-2\alpha} = 0$$

$$(-1 + 2\alpha - n\alpha - \alpha) = 0$$

$$\alpha + n\alpha = 1$$

$$\alpha = \frac{1}{n+1}$$

Ent̄o $\gamma = \frac{2}{n+1}$, e $\beta = \frac{n^2-1}{n^2}$

Temos assim que $c(\alpha)$ ser̄ dado em funç̄o de n por

$$c(n) = \frac{n}{n+1} \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}}$$

Lema 5: Os volumes dos elipsoides E' e E satisfaz $\lambda(E') = c(n)\lambda(E)$

onde $c(n) = \frac{n}{n+1} \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}} < 2^{-1/2(n+1)}$

Prova: Novamente vamos admitir que $B = I$, de modo que B' seja dado por

$$B' = \frac{n^2}{n^2-1} \left(B - \frac{2}{n+1} \frac{(Ba)(Ba)^T}{a^T Ba} \right) \quad (\text{III.2.34})$$

Isto é

$$B' = \text{diag} \left(\frac{n^2}{(n+1)^2}, \frac{n^2}{n^2-1}, \dots, \frac{n^2}{n^2-1} \right) \quad (\text{III.2.35})$$

Através dos conceitos apresentados na proposição 1, podemos concluir que $E' = F_1(E)$ onde $F_1(x) = x' + Lx$ e $B' = LL^T$.

Sabemos que o volume de E' é igual ao Jacobiano da transformação vezes o volume de E , isto é

$$\lambda(E') = \det L \times \lambda(E).$$

Como $\det B' = \det L \cdot \det L^T = (\det L)^2$, então $\det L = \sqrt{\det B'}$.

e assim

$$\lambda(E') = \sqrt{\det B^T} \lambda(E).$$

or (III.2.35) temos

$$\sqrt{\det B^T} = \left(\frac{n}{n+1}\right) \left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} = c(n)$$

isto é

$$\lambda(E') = c(n) \lambda(E)$$

A expansão de Taylor da exponencial ($e^x > 1 + x$) nos dá

$$i) \quad \frac{n}{n+1} = 1 - \frac{1}{n+1} < e^{-\frac{1}{n+1}}$$

$$ii) \quad \frac{n^2}{n^2-1} = 1 + \frac{1}{n^2-1} < e^{\frac{1}{n^2-1}}$$

Agora (i) e (ii) nos dá uma majoração para $c(n)$

$$c(n) < \frac{1}{n+1} \frac{n-1}{e^{2(n^2-1)}} = e^{-\frac{1}{n+1} - \frac{n-1}{2(n^2-1)}} = e^{-\frac{1}{2(n+1)}}$$

Mas $e^{\frac{1}{2(n+1)}} > 2^{\frac{1}{2(n+1)}}$, então $e^{-\frac{1}{2(n+1)}} < 2^{-\frac{1}{2(n+1)}}$ o que nos le
va a garantir que

$$c(n) < 2^{-\frac{1}{2(n+1)}} < 1$$

Portanto

$$\left(\frac{n}{n+1}\right) \left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} = c(n) < 2^{-\frac{1}{2(n+1)}} < 1$$

Como $n > 2$ então

$$0,769 = \left(\frac{2}{3}\right) \left(\frac{4}{3}\right)^{1/2} = c(2) \leq c(n) < 1, \text{ isto é}$$

$$0,769 \leq c(n) < 1$$

Teorema 1: O algoritmo 1 determina um vetor $x^{(k)}$ viável se e somente se (III.2.1) \bar{e} satisfeito

Prova: Mostraremos o teorema somente num sentido, pois o oposto \bar{e} imediato. Admitamos então que o algoritmo I para sem'determinar a solução muito embora saibamos que o problema tem solução. O Lema 2 nos diz que o conjunto de soluções S , interior a $E^{(0)}$ tem volume $\lambda(S) \geq 2^{-(n+1)L}$

Observamos ainda que o conjunto S , \bar{e} um subconjunto de $E^{(\ell)}$, onde $\ell = 4(n+1)^2L$, visto que a inação $a_i^T x < b_i$ escolhida na k -ésima iteração do algoritmo, \bar{e} violada no ponto $x^{(k)}$, de modo que $a_i^T x < b_i \leq a_i^T x^{(k)}$. Portanto S está contido no conjunto $\{x | a_i^T (x - x^{(k)}) \leq 0\}$, o que nos leva a afirmar

$$S \subseteq E^{(k)} \cap \{x \mid a_1^T(x - x^{(k)}) \leq 0\} \subseteq E^{(k+1)}$$

$$0 \leq k \leq \ell = 4(n+1)^2L$$

Aplicando 1-vezes o Lema 5 temos

$$\lambda(S) \leq \lambda(E^{(\ell)}) = [c(n)]^\ell \lambda(E^0) = [c(n)]^{4(n+1)^2L} \lambda(E^0)$$

Como $\lambda(E^0) < (2 \times 2^L)^n < 2^{(n+1)L}$ pois $2^n < 2^L$, segue-se agora que

$$\lambda(S) < \frac{4(n+1)^2L}{2^{(n+1)}} \cdot 2^{(n+1)L} = 2^{-(n+1)L}$$

Levando-nos a conclusão que S é vazio, dando-nos uma contradição, e conseqüentemente a demonstração do teorema.

Observe que o valor de $k < R = 4(n+1)^2L$ foi um imperativo para chegarmos a contradição que prova o teorema. De fato, já vimos na demonstração do Teorema 1 que

$$\lambda(S) < 2^{-\frac{\ell}{2^{(n+1)}}} \cdot 2^{(1+n)L}$$

Então para calcular o valor de L que nos habilite a contradição desejada, basta exigirmos que

$$2^{-\frac{\ell}{2^{(n+1)}}} \cdot 2^{(1+n)L} = 2^{-(1+n)L}$$

$$2^{-\frac{\ell}{2^{(n+1)}}} = 2^{-2(1+n)L}$$

$$- \frac{\ell}{2(n+1)} = - 2(1+n)L \quad (\text{III.2.36})$$

De (III.2.36) concluímos que o valor de ℓ que nos leva a uma contradição na demonstração do Teorema é

$$\ell = 4(1+n)^2L$$

Assim o algoritmo de Khachiyan é repetido até que uma das duas condições ocorra: (i) nenhuma violação foi detectada; neste caso a solução viável foi encontrada, ou (ii) o número de iterações atingiu $4(1+n)^2L$; neste caso podemos concluir que o sistema de inequações não tem solução.

III.3 - AFINIDADE ALGORITMO DE KHACHIYAN E O PROBLEMA DE PROGRAMAÇÃO LINEAR (LP)

Ignorando a exclusividade de "inequações estritas" para tudo que foi apresentado até aqui, podemos estabelecer, através do Teorema da dualidade, meios que nos permitam resolver o problema de programação linear, usando o algoritmo de Khachiyan. O problema (LP) pode ser escrito como:

$$\left. \begin{array}{l} \text{Maximizar } C_{LP}^T x_{LP} \\ \text{sujeito a } A_{LP} x_{LP} \leq b_{LP} \end{array} \right\} \quad (\text{III.3.1})$$

Onde A_{LP} é uma matriz $m_{LP} \times n_{LP}$, x_{LP} é um vetor coluna n_{LP} -dimensional, b_{LP} é um vetor coluna m_{LP} -dimensional e C_{LP} é um vetor coluna n_{LP} -dimensional. Todos os elementos de A_{LP} , b_{LP} e C_{LP} são inteiros e conhecidos. Chamemos então o problema (III.3.1) de problema primal (consequentemente seu dual será:

$$\left. \begin{array}{l} \text{Minimizar } b_{LP}^T y_{LP} \\ \text{sujeito a } A_{LP}^T y_{LP} \geq C_{LP} \\ y_{LP} \geq 0 \end{array} \right\} \quad (\text{III.3.2})$$

Pelo Teorema da dualidade temos que $\max_x C_{LP}^T x_{LP} = \min b_{LP}^T y_{LP}$ se e somente se (III.3.1) tem ótimo finito.

O problema primal e o problema dual podem ser escritos como o:

$$\left. \begin{array}{l} C_{LP}^T x_{LP} - b_{LP}^T y_{LP} = 0 \\ A_{LP} x_{LP} \leq b_{LP} \\ -x_{LP} \leq 0 \\ -A_{LP}^T y_{LP} \leq -C_{LP} \\ -y_{LP} \leq 0 \end{array} \right\} \quad (\text{III.3.3})$$

Segue-se que o sistema de inequações (III.3.3) tem solução se e somente se (III.3.1) tem \bar{otimo} finito ver ^{|10|} e ^{|32|}.

Com:

$$A = \begin{bmatrix} C_{LP}^T & -b_{LP} \\ A_{LP} & 0 \\ -I_{x_{LP}} & 0 \\ 0 & -A_{LP} \\ 0 & -I_{y_{LP}} \end{bmatrix}, \quad x = \begin{bmatrix} x_{LP} \\ y_{LP} \end{bmatrix}, \quad e \quad b = \begin{bmatrix} 0 \\ b_{LP} \\ 0 \\ -c_{LP} \\ 0 \end{bmatrix} \quad (III.3.4)$$

(III.3.3) toma o seguinte formato

$$Ax \leq b \quad (III.3.5)$$

que \bar{e} essencialmente como (III.2.1) quando \bar{a} levantamos a exclusividade das inequações serem "estritas".

Resta-nos agora mostrar que o algoritmo de Khachiyan pode ser usada para determinar se o sistema $Ax \leq b$ tem solu \tilde{c} o, e se o tem, como encontrar uma.

Visando tal objetivo vejamos os lemas 6 e 7, os quais nos darão de uma maneira construtiva, uma solução.

Lema 6: Deixe x_0 ser um ponto de \mathbb{R}^n , e $\theta_i(x) = a_i^T x - b_i$. Então existe um ponto x_1 em \mathbb{R}^n tal que

a) $\theta_i(x_1) \leq \max(0, \theta_i(x_0))$

b) Os vetores $\{a_i \mid \theta_i(x_1) \geq 0\}$ geram todos os outros vetores a_j .

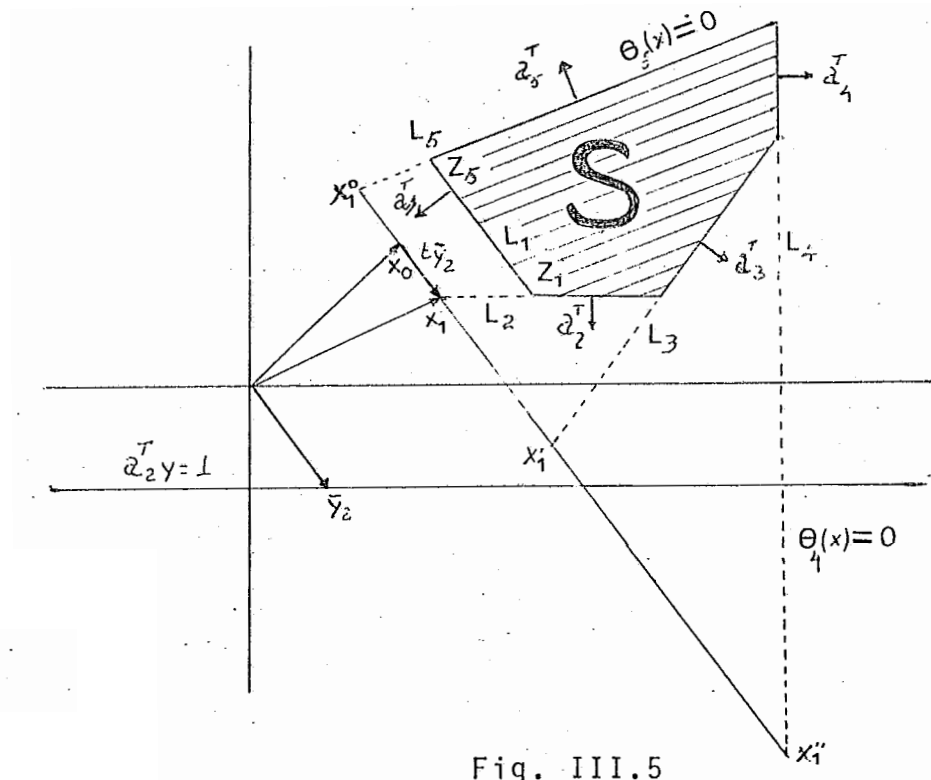


Fig. III.5

Antes de entrarmos no mērito da demonstração, vejamos no \mathbb{R}^2 um esboço geométrico. Seja

$$S = \{x \in \mathbb{R}^2 \mid a_i^T x \leq b_i \quad i = 1, 2, 3, 4, 5\}$$

O conjunto de soluções do problema. Deixe x_0 ser um ponto no R^2 tal que $\{a_i | \theta_i(x_0) \geq 0\} = \{a_1\}$ não gere todos os outros vetores a_j .

Nosso objetivo é movimentarmos paralelamente a reta (no R^n hiperplano) L_1 até encontrarmos outra reta L_i $i = 2, 3, 4, 5$ (hiperplanos) de tal modo que ao fim de cada movimento, deslocaremos de um ponto a outro, dando-nos progressivamente a oportunidade de chegarmos a um ponto x_1 onde o conjunto $\{a_i | \theta_i(x_1) \geq 0\}$ gere todos os outros vetores a_j .

Da Fig. 111.5, observamos que partindo de x_0 e paralelamente a L_1 , poderemos encontrar as retas L_2, L_3 e L_4 , respectivamente nos pontos x_1^0, x_1^I e x_1^{II} . Nota-se também que ao contrário de x_1^I e x_1^{II} , os pontos x_1^0 e x_1 são importantes, pois eles nos dão as informações adicionais que Z_1 interseção de L_1 com L_2 e Z_2 interseção de L_1 com L_5 são vértices de S .

Constatamos facilmente através da Fig. III.5 que $a_2^T(x_0 + t_{12} \bar{y}_2) = b_2$.

Portanto para livrarmos daqueles pontos indesejáveis, x_1^I e x_1^{II} basta escolhermos convenientemente t e \bar{y}_i . Vejamos:

Sejam a_i vetores não colineares com a_1 , e \bar{y}_i $i = 2, 3, 4, 5$ vetores que satisfazem respectivamente os sistemas

$$\bar{y}_i: \begin{cases} a_1^T y = 0 & (\text{cond. 'para } \bar{y}_1 \text{ ser paralelo as rest. que não} \\ & \text{são satisfeitas em } x_0, \text{ qualquer que seja } i) \\ a_i^T y = 1. \end{cases} \quad (\text{III.3.8})$$

$$i = 2, 3, 4, 5$$

Chamemos de

$$t_{1i} = \min D_{1i} \quad (\text{III.3.9})$$

onde

$$D_{1i} = \left\{ -\frac{\theta_j(x_0)}{a_j^T \bar{y}_i}, a_j^T \bar{y}_i > 0, j \in \{2, 3, 4, 5\} \right\}.$$

Então

$$D_{12} = \left\{ -\frac{\theta_2(x_0)}{a_2^T \bar{y}_2}, -\frac{\theta_3(x_0)}{a_3^T \bar{y}_2}, -\frac{\theta_4(x_0)}{a_4^T \bar{y}_2} \right\}$$

$$D_{13} = \left\{ -\frac{\theta_2(x_0)}{a_2^T \bar{y}_3}, -\frac{\theta_3(x_0)}{a_3^T \bar{y}_3}, -\frac{\theta_4(x_0)}{a_4^T \bar{y}_3} \right\}$$

$$D_{14} = \left\{ -\frac{\theta_2(x_0)}{a_2^T \bar{y}_4}, -\frac{\theta_3(x_0)}{a_3^T \bar{y}_4}, -\frac{\theta_4(x_0)}{a_4^T \bar{y}_4} \right\}$$

$$D_{15} = \left\{ -\frac{\theta_5(x_0)}{a_5^T \bar{y}_5} \right\}$$

É fácil ver que

$$D_{13} = (a_3^T \bar{y}_2) D_{12}, \quad D_{14} = (a_4^T \bar{y}_2) D_{12}$$

Portanto

$$t_{15} = - \frac{\theta_5(x_0)}{a_5^T \bar{y}_5}$$

$$t_{12} = - \frac{\theta_2(x_0)}{a_2^T \bar{y}_2}$$

$$t_{13} = \min D_{13} = (a_3^T \bar{y}_2) \min D_{12} = (a_3^T \bar{y}_2) t_{12} \quad (\text{III.3.10})$$

$$t_{14} = \min D_{14} = (a_4^T \bar{y}_2) \min D_{12} = (a_4^T \bar{y}_2) t_{12} \quad (\text{III.3.11})$$

Segue-se de (III.3.10) que

$$t_{13} = t_{12} (a_3^T \bar{y}_2) = - \frac{\theta_2(x_0)}{a_2^T \bar{y}_2} (a_3^T \bar{y}_2)$$

$$\frac{a_2^T \bar{y}_2 t_{13}}{a_3^T \bar{y}_2} = - \theta_2(x_0)$$

$$a_2^T \left(\frac{\bar{y}_2 t_{13}}{a_3^T \bar{y}_2} \right) + a_2^T x_0 - b_2 = 0$$

$$a_2^T \left[x_0 + \left(\frac{t_{13}}{a_3^T \bar{y}_2} \right) \bar{y}_2 \right] = b_2$$

Então

$$x_1 = x_0 + \left(\frac{t_{13}}{a_3^T \bar{y}_2} \right) \bar{y}_2$$

De (III.3.11), temos por analogia que

$$a_2^T \left[x_0 + \left(\frac{t_{14}}{a_4^T \bar{y}_2} \right) \bar{y}_2 \right] = b_2$$

Logo

$$x_1 = x_0 + \left(\frac{t_{14}}{a_4 \bar{y}_2} \right) \bar{y}_2$$

Concluimos, que qualquer que seja ($i = 2, 3, 4$) o sistema (III.3.8) que consideremos, após um deslocamento paralelo a restrição não satisfeita, (III.3.8) - (II.3.9) nos assegurará que sempre estaremos em x_1 .

Note que \bar{y}_i , $i = 2, 3, 4$ tem o mesmo sentido, e \bar{y}_5 o oposto. Portanto quando fazemos em (III.3.8) - (III.3.9) $i=5$, caminharemos no sentido oposto a \bar{y}_i $i = 2, 3, 4$, determinando o ponto x_1^0 .

Voltando a demonstração do Lema 6, seja x_0 um ponto qualquer do R^n . Trivialmente x_0 satisfaz (a). Se x_0 não satisfaz (b), é suficiente mostrar que podemos determinar \bar{x} satisfazendo (a) e tal que

$$\{i | \theta_i(\bar{x}) \geq 0\} \supset \{i | \theta_i(x_1) \geq 0\}$$

Repetindo este procedimento m vezes, obtemos x_1 satisfazendo (a) e (b). Vejamos.

Suponha (depois de uma reenumeração) que $\theta_1(x_0) \geq 0 \dots \theta_k(x_0) \geq 0$ (as k -primeiras restrições não sejam satisfeitas em x_0) e $\theta_{k+1}(x_0) < 0 \dots \theta_n(x_0) < 0$. Admitamos agora que a_{k^*} $k < k^* \leq m$ não seja combinação linear de $a_1 \dots a_k$. Deixe \bar{y} ser a solução do sistema

$$y : \begin{cases} a_i^T y = 0 & i = 1 \dots k \\ a_{k^*}^T y = 1 \end{cases}$$

e faça $\bar{x} = x_0 + t \bar{y}$ onde

$$t = \min \left\{ \frac{-\theta_j(x_0)}{a_j^T \bar{y}} \mid a_j^T \bar{y} > 0, k < j \leq m \right\}$$

Assim $0 < t \leq \frac{\theta_{k^*}(x_0)}{a_{k^*}^T \bar{y}} = \theta_{k^*}(x_0)$ e

$$\theta_i(\bar{x}) = \theta_i(x_0 + t \bar{y}) = a_i^T(x_0 + t \bar{y}) - b_i = t a_i^T \bar{y} + a_i^T x_0 - b_i$$

Isto é

$$\theta_i(\bar{x}) = t a_i^T \bar{y} + \theta_i(x_0) \begin{cases} = t \cdot 0 + \theta_i(x_0) & \text{se } 1 \leq i \leq k \\ \leq 0 & \text{se } k < i \leq m \end{cases}$$

Onde temos a igualdade para algum

$$k < i_0 \leq m.$$

Conseqüentemente

$$\{a_i \mid \theta_i(\bar{x}) \geq 0\} \supseteq \{a_i \mid \theta_i(x_0) \geq 0\} \cup \{a_{i_0}\}$$

Repetido sucessivamente este processo, encontraremos x_1 tal que (a) e (b) verificam-se simultaneamente.

Lema 7: O sistema de inequações lineares

$$a_i^T x \leq b_i, \quad i = 1, \dots, m \quad (\text{III.3.12})$$

tem solução se e somente se o sistema de inequações lineares estritas

$$a_i^T x < b_i + 2^{-L} \quad i = 1, \dots, m \quad (\text{III.3.13})$$

o tem.

Demonstração: (III.3.12) \rightarrow (III.3.13)) trivialmente se (III.3.12) tem solução, então (III.3.13) também o tem.

(III.3.13) \rightarrow (III.3.12)) deixe x_1 ser tal que (a) e (b) sejam simultaneamente satisfeitos, e ainda mais, seja solução do sistema de inequações lineares estritas (III.3.13), (assim $\theta_i(x_1) < 2^{-L}$ e $a_i^T x_1 \geq b_i$ para $1 \leq i \leq k$). Através de uma rotulação deixe a_1, \dots, a_r serem linearmente independente e gerarem a_{r+1}, \dots, a_k . Por (a) e (b) eles também geram a_{k+1}, \dots, a_m . Agora denotemos por Z a solução do sistema linear $a_j^T Z = b_j$ para $j = 1, \dots, r$. Mostremos então que Z satisfaz (III.3.12). Sabemos por Cramer que $a_i = \sum_{1 \leq j < r} (\bar{D}_j / \bar{D}) a_j$ para $1 \leq i \leq m$, onde $\bar{D} > 0$ e \bar{D}_j são inteiros com valor absoluto menor que $2^L / mn$.

Assim

$$a_i^T Z - b_i = \sum_{1 \leq j \leq r} \frac{D_j}{D} a_j^T Z - b_i$$

onde podemos admitir $D > 0$ pois

$$\text{se } D < 0 \text{ faça } \begin{cases} \bar{D} = -D \\ \bar{D}_j = -D_j \end{cases}$$

Então

$$\bar{D}(a_i^T Z - b_i) = \sum_{1 \leq j \leq r} \bar{D}_j a_j^T Z - \bar{D} b_i$$

Mas $a_j^T Z = b_j$, portanto

$$\bar{D}(a_i^T Z - b_i) = \sum_{1 \leq j \leq r} \bar{D}_j b_j - \bar{D} b_i$$

Então $\bar{D}(a_i^T x_1 - b_i)$ é um invariante. Prosseguindo, como

$$b_j = a_j^T x_1 - \theta_j(x_1)$$

$$b_i = a_i^T x_1 - \theta_i(x_1)$$

Temos

$$\bar{D}(a_i^T Z - b_i) = \sum_{1 \leq j \leq r} \bar{D}_j (a_j^T x_1 - \theta_j(x_1)) - \bar{D} (a_i^T x_1 - \theta_i(x_1))$$

$$= \left(\sum_{1 \leq j \leq r} \bar{D}_j a_j^T x_1 - \bar{D} a_i^T x_1 \right) + \bar{D} \theta_i(x_1) - \sum_{1 \leq j \leq r} \bar{D}_j \theta_j(x_1)$$

Como $\sum_{1 \leq j \leq r} \bar{D} a_j^T x_1 = \bar{D} a_i^T x_1$ e por hipótese (III.3.13)

$\theta_i(x_1) < 2^{-L}$ $i = 1, \dots, m$ segue-se

$$\bar{D}(a_i^T Z - b_i) < \bar{D} 2^{-L} + \sum_{1 \leq j \leq r} |\bar{D}_j| 2^{-L}$$

$$< \frac{1}{mn} + \frac{m}{mn} < 1$$

Como $\bar{D}(a_i^T Z - b_i)$ é inteiro, e $\bar{D} > 0$ concluímos que

$$a_i^T Z - b_i \leq 0$$

Z satisfaz (III.3.12).

CAPÍTULO IVCONCLUSÕES

Baseados em teoremas simples, o método desenvolvido no Capítulo II, tem como principal vantagem, fixar em cada passo do algoritmo um vetor da base-final. No método do simplex, temos que remover um vetor e adicionar um novo, em cada interação.

Fundamentalmente nossa teoria é válida para problemas de Programação Linear na forma standard. Uma outra suposição que se torna necessária é que a matriz A tenha posto completo, o que nos leva a garantir que o vetor b sempre estará na imagem de A ($b \in R(A)$). Além do mais exigimos que nenhum vetor coluna da matriz A pode estar no interior do cone poliédrico convexo, formado pela envoltoria convexa de raios vetores, constituídos pelos vetores colunas de A . Vejamos agora o número de operações necessárias.

Tomemos como termometro o número de multiplicações que o algoritmo requer. Visando encontrar uma cota superior para o número de multiplicações, denotemos por:

$(k-1)$ - O número de vetores na base ($k = 1, \dots, m$)

n_k - O número de vetores que tem possibilidade de entrar na base e dar solução viável, quando na base já exist

tem $(k-1)$ vetores.

$(n-kt)$ - 0 número de vetores que não estão na base.

n_k^* - 0 número de produtos escalares, no passo k , do tipo $\langle a^{k+1}, p_k b \rangle$

PE_k - 0 número de multiplicações efetuadas nos n_k^* produtos escalares.

É fácil concluir que se a base tem $(k-1)$ vetores, então existe somente n_k possibilidades de incorporar mais um vetor na base, o que nos leva a afirmar que para cada um dos n_k vetores temos que calcular o produto escalar do restante dos vetores de A com a projeção de b sobre o subespaço ortogonal. Assim $PE_k \leq m[n_k(n-k)]$, e uma cota é dada por

$$\text{Número de multiplicações} \leq m \left[4 \left(\sum_{k=1}^m (n-k+1) + n_k \right) - n_m \right] + \sum_{k=1}^m PE_k$$

Este método poderá vir a ser utilizado como ponto de partida para outros algoritmos mais eficientes e abrangentes ficando esta colaboração, como uma idéia para trabalhos que poderão ser desenvolvidos posteriormente.

Quanto ao algoritmo de Khachiyan (Capítulo III) podemos notar que o enfoque dado pelo mesmo está baseado nos trabalhos de N. Z. Shor [46] e [47], o qual nos dá uma nova

idēia de como lidar com problemas de programação linear. Num rápido contacto com o método do elipsóide muitos fatores nos parecem impraticáveis, por exemplo:

- 1) Que grau de dificuldade existe ao passar dos argumentos geométricos à sua formulação em termos computacionais ?
- 2) Como poderemos executar o processo de arredondamento o qual eventualmente nos leva a solução ótima ?
- 3) Em cada iteração temos que estocar o elipsóide, o que equivale a prever o armazenamento de n^2 números, cada um tendo muitos dígitos; será que isto não é muito ?

É paupável que a resposta para muitas questões que se possam colocar estarão ainda a mercê de muitas experiências e pesquisa, mas mesmo assim, podemos dizer algo a respeito delas:

- 1) Percebemos no decorrer do Capítulo III, que o enlace "fundamentos matemáticos - construção geométrica" é bastante simples. Recordamos que, para descrever o elipsóide E_k , nós usamos seu centro x_k e a matriz simétrica $n \times n$ B_k , definida positiva. Veja as fórmulas de recorrência (III.2.5 - III.2.6), onde vale ressaltar a ausência de inversão de matrizes. Portanto para chegarmos a x_{k+1} e B_{k+1} a partir de x_k e B_k , necessitamos de $O(n^2)$ operações aritméticas.

- 2) Como na prática são poucos os casos onde necessitamos exatamente do ótimo, isto não constituirá obstáculos na maioria dos problemas práticos.
- 3) Em vez de armazenar o elipsóide corrente em cada interação, podemos armazenar outros dados equivalentes. Muitos melhoramentos poderão ainda ser pesquisado nesta direção.

Para responder muitas outras perguntas que se possam colocar, temos ainda, como já foi citado, que depender de muitas experiências. Um fator que nos parece bastante negativo no método do elipsóide, é que o número de operações aritméticas depende do número de dígitos dos coeficientes (c_j , b_i , a_{ij}). Nos problemas práticos L é normalmente grande, daí os computadores nem sempre dispõem de um número de dígitos significativos necessários para representar 2^L , para que possamos garantir a convergência do método. É bem provável que existam problemas onde o tal método possa ser aplicado com bastante sucesso. Certamente problemas vinculados as áreas de programação matemática, análise numérica e computação, que aprimorados, constituirão a ferramenta indispensável para aplicação do método do elipsóide de uma maneira ampla, haja visto por exemplo, o algoritmo do simplex teve que esperar pelas técnicas de inversão de grandes e médias matrizes para se tornar realmente, de grande utilidade prática.

BIBLIOGRAFIA

- [1 | ALBERT, A. - Regression and the Moore-Penrose Pseudoinverse - Academic Press, New York and London (1972).
- [2 | ALBERT, A. - Conditions for Positive and Nonnegative Definiteness in Terms of Pseudoinverses - SIAM J. Appl. Math. 17 (1969), 434-440.
- [3 | ASPVALL, B., e STONE, R. E. - Khachiyan's Linear Programming Algorithm, - Report STAN-CS-79-776, Departments of Computer Science and Operations Research, Stanford University (1979).
- [4 | BALINSKI, M. L. - An Algorithm for Finding all vertices of Convex Polyhedral Sets - J. Soc. Indust. Appl. Math. Vol. 9, No. 1, March 1961.
- [5 | BAZARAA, M. S., JOHN, J. JARVIS - Linear Programming and Network Flows", John Wiley & Sons., 1977.
- [6 | BEN ISRAEL, A. e CHARMES, A. - An Explicit Solution of a Special Class of Linear Programming Problems. - Operations RES 16 (1968), 1166-1175.
- [7 | BISSHOPP, F. - Khachiyan's Algorithm for Linear Inequalities - Div. Applied Math., Brown University, Providence R. I. 02912, (1979).

- |⁸| BISSHOPP, F. - Shor's Algorithm and Linear Inequalities -
Division of Applied Mathematics, Brown University, Providence Rhode Island 02912 (1979).
- |⁹| BURDET, C. A. - Generation all the Faces of a Polyhedron -
SIAM J. Appl Math., Vol. 26, No. 3, May 1974.
- |¹⁰| DANTZIG, G. B. - Linear Programming and Extensions - Princeton University Press, Princeton New Jersey, (1963).
- |¹¹| DESOER, C. A., e B. H. WHALEN - A note on Pseudoinverses -
J. Soc. Indust. Appl. Math., Vol. 11, No. 2, June 1963.
- |¹²| DYER, M. E e L. G. PROLL - An Algorithm for Determining all Extreme Points of a Convex Polytope - Math. Programming 12 (1977), pp. 81-96.
- |¹³| GÁCS, P., e LOVASZ, L. - Khachiyan's Algorithm for Linear Programming - Technical Report Stan-CS-79-750, Computer Science Department, Stanford University.
- |¹⁴| GANTMACHER, F. R. - The Theory of Matrices - Vol. 1 - Chelsea publishing Company - New York, N. Y. (1974).
- |¹⁵| GASS, S. I. - Programacion Linear - C.E.C.S.A., 1972.

- [¹⁶] GOLDFARB, D. - On the Bartels-Golub Decomposition for Linear Programming Bases - North-Holland Publishing Company (1957), pp. 273-279.
- [¹⁷] GOMORY, R. e M. BALINSKI - A Mutual Primal-Dual Simplex Method - In Recent Advance in Mathematical Programming, McGraw-Hill, New York (1963).
- [¹⁸] GRAVES, G. - A Complet Constructive Algorithm for the General Mixed Linear Programming Problem - Naval Res. Logistics Quart., Vol. 12, No. 1, 1-34 (1965).
- [¹⁹] GREGORY, J. W. - An Analysis of Khachiyan's Algorithm for Linear Programming, in a Commercial Environment, Control Data Corporation, 1801 West Country Road B, Roseville, Minnesota 55113,
- [²⁰] GREVILLE, T. N. E. - Some Applications of Pseudoinverse of a Matrix - SIAM Review, Vol. 2, No. 1, January 1960.
- [²¹] GROETSCHEL, M., L. LOVASZ e A. SCHRIJVER - The Ellipsoid Method and its Consequences in Combinatorial Optimization - 1980, pp. 35. Report No. 80151 - or Institut Fuer Oekonometrie und Operations Research, Universitaet Bonn', Bonn Federal Republic of Germany.
- [²²] HOHENBALKEN, B. V, - Least Distance Methods for the Scheme of Polytopes - Math. Programming 15 (1978), pp. 1-11.

- [²³] JONES, P. C., e MARWILL, E. S. - A Variant of Khachiyan's Algorithm for Linear Programming, EG & G. Idaho, Inc. - Idaho Falls, Idaho (1979).
- [²⁴] JONES, P. C., e MARWILL, E. S. - Solving Linear Complementarity Problems with Khachiyan's Algorithm - January 1980, pp. 9.
- [²⁵] JOSHI, V. N. - A Note on the Solution of Rectangular Linear Systems by Iteration -.SIAM Review, Vol. 12, No. 3, July 1970.
- [²⁶] KHACHIYAN, L. G. - A Polynomial Algorithm in Linear Programming - Doklady Akademiia Nauk URSS, Vol. 244, No. 5 (1979) (Tradução feita para o inglês por Valch e Lawler, Berkeley University of California, 1979).
- [²⁷] LAU, H. T. - A Note on Polynomial Algorithm for Linear Programming - Department of Computer Science McGill University - Montreal.
- [²⁸] LAWSON, C. L., RICHARD, J. HANSON - Solving Least Squares Problems - Prentice Hall Inc. Englewood Cliffs', N. J. (1974).
- [²⁹] LEMKE, C. E. - The Dual Method of Solving the Linear Programming Problem - Carnegie Institute of Technology, Dept. of Math. Technical Report, No. 29 (1953).

- [³⁰] LOUD, W. S. - Generalized Inverses and Generalized Greem's Functions' - J. SIAM Appl. Math., Vol. 14, No. 1, March (1966).
- [³¹] LOVASZ, L. - A New Linear Programming Algorithm Better or Worse than the Simplex Method ? Bolyai Institute, University of Szeged. H-G720 Szeged, Hungary (1980).
- [³²] MACULAN, N. e PEREIRA, M. V, F. - Programação Linear - Editora Atlas S. A., São Paulo, SP (1980).
- [³³] MACULAN, N. - Algorithmo de Khachiyan para a Resolução de Problemas de Programação Linear - 29 Simpósio de Combinatória - IBILCE - UNESP, pp. 44-51 (1980).
- [³⁴] MCCALL, E. H. - A Study of the Khachiyan's Algorithm for Real-World Linear Programming - Computer Science Department University of Minnesota 136, Lind Hall, Minneapolis, Minnesota (1980).
- [³⁵] MIFFLIN, R. - A Stable Method for Solving Certain Constrained Least Squares Problems - Math. Programming, 16 (1979) pp, 141-158.
- [³⁶] MITCHELL, B. F., V. E. DEMYANOU e V. N. ALOZEMOV - Finding the Point of a Polyhedron Closest to the Origin - SIAM J. Control, Vol. 12, No, 1, February, 1974.

- [³⁷] NOLTEMEIER, H. - An Algorithm for the Determination of Longest Distances in a Graph - Math. Programming 9 (1975) pp. 350-357.
- [³⁸] PANNE, C. V. - Methods for Linear and Quadratic Programming - North-Holland/American Elsevier (1975).
- [³⁹] PICKEL, P. F. - Some Improvements to Khachiyan's Algorithm in Linear Programming - Math. Dept. Polytechnic Inst. of N. Y. Route 110, Farnsingdale N. Y. 11735, U. S. A. (1979).
- [⁴⁰] PYLE, L. D. - The Generalized Inverse in Linear Programming Basic Structure - SIAM J. Appl. Math., Vol. 22, No. 3, May (1972).
- [⁴¹] RANDOL BURTON - The Ellipsoid Method in Linear Programming Dep. of Mathematics Cuny Graduate Center, 33, West 42 St. N. Y., NY 10036 (1979).
- [⁴²] RAO, C. R. e MITRA, S. K. - Generalized Inverse of Matrices and its Applications - John Wiley & Sons Inc., New York, (1971).
- [⁴³] RUBIN, D. S. - Vertex Generation and Cardinality Constrained Linear Programs - Technical Notes, pp. 555-565 (1974).

- [44] SCOLNIK, H. - A New Approach to Linear Programming - A Paper Presented at the VIII Int, Symp. on Mathematical Programning, Stanford Univ. (1973).
- [45] SCROGGS, J. E. e PATRICK L. ODELL - An Alternate Definition of a Pseudoinverse of a Matrix - J. SIAM. Appl. Math. Vol. 14, No. 4, Jul. (1966).
- [46] SHOR, N. Z. - Convergence Rate of the Gradient Descent Method with Dilatation of the Space - Kibernetika, No. 2, (Traduzido para o inglês em Cybernetics 6).
- [47] SHOR, N. Z. - Cut-Off Method with Spale Extension in Convex Programming Problems, Kibernetika, No. 1 (traduzido para o inglês em Cybernetics 13).
- [48] SIMONNARD, M. - Programmacion Lineal Paraninfo - Madrid (1977).
- [49] SOMMERVILLE, D. M. Y. - An Introduction to the Geometry of N Dimension.
- [50] SPOSITO, V. - Solutions of a Special Class of Linear Programming Problems" - Op. Res. 21, 386-388 (1973).
- [51] THOMPSON, G., TONGE, F., e' ZIONTS, S. - Techniques for Removing Nonbinding Constraints and Exraneous Variables from Linear Programming Problems - Management Sc., Vol.

12, 588-608, (1966).

- |⁵²| TUCKER, A. W. - Dual Systems of Homogeneous Linear Relations - Linear Inequalities and Related Systems. Annals of Mathematics Studies, Princeton University Press.
- |⁵³| WOLFE, P. - A Bibliography for the Ellipsoid Algorithm IBM Research Center, Yorktown Heights, N. Y.
- |⁵⁴| WOLFE, P. - Finding the Nearest Point in a Polytope - Math. Programming 11 (1976) 128-149,
- |⁵⁵| ZARANTONELLO, E. H. - L'Algebre des Projecteurs Coniques - Proc. Am. Math. Soc. 37, pp. 232-243 (1972).
- |⁵⁶| ZLOBEC, S. e BEN-ISRAEL - Explicit Solutions of Interval Linear Programs - OPNS. Res. 21, pp. 390-392 (1973).