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. Nelson 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

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 Federa? do Ceará, pela oportunidade concedida.

Aos meus pais os ensinamentos de disciplina e perserverança 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 Algorítmo 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 subsequente, 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 tke above.

Based on essentially simple theorems, we develope, in chapterII, an algorithm which has as it's principal objective, to detect in it's first step, a vector which should have its presence asured 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 didatic 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 niaybe give incentives for other research, with intent to perfect which opens other ways of utilization.

ÍNDICE

Págs.

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 à prática no campo da organização e planificação da indústria, dando-lhes regimes ótimos de produção, gerando consequentemente maiores lucros, etc. Portanto aplicações de programação linear à economia, ao setor militar, e a outros domínios, não cessam de se extender. 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 similarmente 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:

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

$$\text{Sujeito a} \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \\ x_1 \geq 0,\ x_2 \geq 0,\ \dots,\ x_n &\geq 0 \end{aligned} \qquad (I.2.2)$$

De uma maneira mais compacta podemos ainda ter

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

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

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

ou

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

$$\text{Sujeito a } Ax = b$$
$$x \geq 0$$

Onde $c^T = (c_1, c_2, \ldots, c_n)$, $x^T = (x_1, x_2, \ldots, x_n)$ $b^T = (b_1, b_2, \ldots, b_m)$ e A uma matriz com m-linhas e n-colunas, tendo seus elementos denotados por a_{ij}, e posto m. O conjunto das variáveis $x_1, x_2, \ldots, x_n$ 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, fazenda n-m variáveis iguais a zero, resolvemos em relação às variáveis restantes, sempre que o determinante 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 I.2.3, é dito estar na forma standard. Muitas vezes desejamos maximizar ou minimizar uma função linear em presença de restrição de igualdade e/ou inequações. Daí a necessidade de passar de um problema para outro equivalentemente.

1.3 - INEQUAÇÕES E EQUAÇÕES

Suponhamos que uma restrição seja dada por $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ $(\sum_{j=1}^{n} a_{ij}x_j \geq b_i)$. 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 $\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i$ $(\sum_{j=1}^{n} a_{ij}x_j - x_{n+i} = b_i)$. Inversamente, uma equação pode ser transformada em duas inequações $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ e $\sum_{j=1}^{n} a_{ij}x_j \geq b_i$.

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 é concluido, o valor da função objetivo do problema é (-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 $\sum_{j=1}^{n} a_{ij} x_j = b_i$ $x_j \geq 0 \quad i=1\ldots m$ $j=1\ldots n$	Minimize $\sum c_j x_j$ Sujeito a $\sum_{j=1}^{n} a_{ij} x_j \geq b_i$ $x_j \geq 0 \quad i = 1\ldots m$ $j=1\ldots n$
	MAXIMIZAÇÃO	Maximize $\sum_{j=1}^{n} c_j x_j$ Sujeito a $\sum_{j=1}^{n} a_{ij} x_j = b_i$ $i=1\ldots m$ $x_j \geq 0, j=1\ldots n$	Maximize $\sum_{j=1}^{n} c_j x_j$ Sujeito a $\sum_{j=1}^{n} a_{ij} x_j \leq b_i$ $i=1\ldots m$ $x_j \geq 0, j=1\ldots n$

TABELA 1.1

I.6 - MÉTODO DO SIMPLEX

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

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

Sujeito a

$$\begin{aligned} x_1 + 3x_2 - x_3 \quad\quad + 2x_5 \quad &= 21 \\ - 2x_2 + 4x_3 + x_4 \quad\quad &= 36 \\ - 4x_2 \quad\quad + 8x_5 + x_6 &= 30 \end{aligned} \qquad (I.6.1)$$

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

Podemos ainda dar ao problema proposto, a seguinte disposição

$$\begin{aligned} 21 &= x_1 + 3x_2 - x_3 \quad\quad + 2x_5 \\ 36 &= \quad - 2x_2 + 4x_3 + x_4 \\ 30 &= \quad - 4x_2 \quad\quad + 8x_5 + x_6 \\ 0 &= \quad - x_2 + 3x_3 \quad - 2x_5 \quad + f \end{aligned} \qquad (I.6.2)$$

$$x_j \geq 0 \qquad j = 1, 2, \ldots, 6$$

onde desejamos encontrar o ótimo de **f**.

Note que o sistema (I.6.2) está na forma standard onde

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

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

e a solução básica é dada por:

$$f = 0 \quad x_1 = 21 \quad x_4 = 36 \quad x_6 = 30$$
$$x_2 = x_3 = x_5 = 0$$

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

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

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

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

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. Consequentemente 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 valar proporcionam omaior decrescimento 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 3ª equação de (I.6.4) nos diz que se x_3 toma qualquer valor, x_6 será sempre positivo. O mesmo ocorre com x_1 para qualquer x_3 não negativo. Neste momento não podemos deixar passar despercebido que a análise feita sobre a 1ª e 3ª 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 $P_3 = [x_{13}, x_{23}, x_{33}]^T = [-1, 4, 0]^T$ x_{i3} for não positiva. Consequentemente limitaremos nossa análise sobre aqueles $x_{i3} > 0$. A 2ª equação de (I.6.4) restringe o valor de x_3, isto é, para termos $x_4 \geq 0$, x_3 não deverá crescer além de $(36/4) = 9$.. Esta é a idéia que norteia o critério de saída de um vetor da base, no método do Simplex.

Concluímos 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:

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

Nesta solução $x_4 = 0$, pois foi a 1ª 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 lá dividimos a 2ª equação por 4 (coeficiente de x_3) e eliminamos x_3 das demais equações, Tal operação é dita pivoteamento (4 é o pivô). Concluindo o pivoteamento temos:

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

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

$$\begin{aligned}
30 - \frac{10}{4}x_2 &= x_1 \\
9 + \frac{2}{4}x_2 &= x_3 \\
30 + 4x_2 &= x_6 \\
-27 - \frac{2}{4}x_2 &= f
\end{aligned}$$

Como é fácil de ver, para manter as variáveis não negativas, x_2 não pode exceder $30/(10/4) = 12$. Para $x_2 = 12$ temos os seguintes valores para as variáveis

$$\begin{aligned}
0 &= x_1 \\
15 &= x_3 \\
78 &= x_6 \\
-33 &= f \\
x_4 &= x_5 = 0
\end{aligned}$$

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

$$
\begin{aligned}
12 &= -\frac{4}{10} x_1 + x_2 + \frac{1}{10} x_4 + \frac{8}{10} x_5 \\
15 &= \frac{2}{10} x_1 + x_3 + \frac{3}{10} x_4 + \frac{4}{10} x_5 \\
78 &= \frac{16}{10} x_1 + \frac{4}{10} x_4 + \frac{112}{10} x_5 + x_6 \\
-33 &= \frac{-2}{10} x_1 - \frac{8}{10} x_4 - \frac{24}{10} x_5 + f
\end{aligned}
\qquad (I.6.7)
$$

Como (I.6.7) não tem nenhum elemento positivo na última linha, concluímos 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:

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

Segue-se de (I.6.8) que, quando $x3 \to \infty$, $f \to -\infty$ 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, c 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:

Minimizar $f = c^T x$ (I.6.9)

Sujeito a $[P_1, \ldots, P_n]x = P_0$ (I.6.10)

$x \geq 0$ (I.6.11)

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

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

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

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 ?-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

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

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: $I_B \cup I_N = (1, 2, \ldots, n\}$, $I_B \cap I_N = \phi$, dando-nos

$x^T = [x_B^T, x_N^T]$ e $c^T = [c_B^T, c_N^T]$.

Agora o pro'blema (169) - (I.6.11) pode ser re-escrito:

Minimizar $f = c_B^T x_B + \bar{c}_N^T x_N$ (I.6.12)

Sujeito a $Bx_B + Nx_N = P_0$ (I.6.13)

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

$x_B \geq 0$ e $x_N \geq 0$ (I.6.14)

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

'RO SIMPLEX: PASSO INICIAL

i	BASE	C	$\mathbf{a}_0$	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	$\mathbf{a}_1$	C_1	x_{10}	1	0		0	.	0	$x_{1,m+1}$	.	$x_{1,j}$		$x_{1,k}$		$x_{1,n}$
2	$\mathbf{a}_2$	C_2	x_{20}	0	1		0	.	0	$x_{2,m+1}$	.	$x_{2,j}$		$x_{2,k}$		$x_{2,n}$
⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮		⋮		⋮
ℓ	P_ℓ	C_ℓ	$x_{\ell 0}$	0	0		1	.	0	$x_{\ell,m+1}$	.	$x_{\ell,j}$		$x_{\ell,k}$		$x_{\ell,n}$
⋮	⋮	⋮	.	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮		⋮		⋮
m	$\mathbf{a}_m$	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 $z_0 = \sum_{i=1}^{m} c_i x_{i0}$,

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

Voltando a equação (I.6.6), observamos que os coeficientes x_i na última linha são os $z_j - c_j$, e lá nos escolhemos o vetor a entrar na base, justamente aquele que correspondia ao maior coeficiente de x_i 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_{i0}/x_{ik} para $x_{ik} > 0$.

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 $x_{ik} \leq 0$, encontramo-nos 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: 2º 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	.	P_m	P_{m+1}		P_j		P_k	.	n
1	P_1	C_1	x'_{10}	1	0	.	$x'_{1\ell}$	.	0	$x'_{1,m+1}$	.	x'_{ij}	.	0	.	x'_{1n}
2	P_2	C_2	x'_{20}	0	1	.	$x'_{2\ell}$	.	0	$x'_{2,m+1}$	.	x'_{2j}	.	0	.	x'_{2n}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ℓ	P_k	C_k	x'_{k0}	0	0		$x'_{\ell\ell}$	.	.	$x'_{\ell,m+1}$	.	$x_{\ell,j}$	.	1	.	$x'_{\ell,n}$
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
m	P_m	C_m	x'_{m0}	0	0		$x'_{m\ell}$	.	1	$x'_{m,m+1}$	.	.	.	0	.	$x'_{m,n}$
m+1			z'_0	0	0	.	$Z'_\ell - C_\ell$		0	$Z'_{m+1} - C_{m+1}$	.	$Z'_j - C_j$	.	0	.	$Z'_n - C_n$

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

$$\underset{j=1,\ldots,n}{\mathrm{Max}} (z_j - z_j) = z_k - c_k$$

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 coni

$$\min_{i} \left(\frac{x_{i0}}{x_{ik}},\ x_{ik} > 0\right)$$

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 é finito, pudemos 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 é infinito, podemos construir uma nova solução viável com m+1 variáveis positivas cujo valor da função objetivo pode fazer-se arbitrariamente negativo.

<u>TEOREMA 2</u>: Se para qualquer solução básica viável $x=(x_{10}\cdots x_{m0})^T$ as condições $z_j - c_j \leq 0$ são satisfeitas para $j = 1, 2, \ldots, n$, 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,

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

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

$$x_B = \overline{x}_B = B^{-1}b, \qquad (I.6.16)$$

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

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

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

Por (I.6.14):

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

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

Fazendo

$$z_j = c^T B^{-1} P_J \text{ temos que}$$

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

o que nos dá

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

O sistema (I.6.12) - (1.6.14) toma o seguinte aspecto

$$f = c^T B^{-1}b - \sum_{j\varepsilon I_N} (z_j - c_j)\, x_j \qquad (I.6.18)$$

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

$$x_B \geq 0, \text{ e } x_J \geq 0, j \in I_N \qquad (I.6.20)$$

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

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

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

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

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

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

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

então criamos inicialmente uma linha no Quadro I.2, a $(m+2)$-ésima linha, na qual colocaremos correspondendo a cada coluna j, ($j = 1, \ldots, n$) o coeficiente de w, $\sum_{i}^{m} x_{ij}$ da equação (I.7.1). Na $(m+1)$-ésima linha colocaremos para cada coluna j ($j = 1, \ldots, n$) $- c_j$. 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

Maximizar $f = c^T x$ (I.8.1)

Sujeito a $Ax \leq b$ (I.8.2)

$x \geq 0$

Aqui, como já chamamos à obtenção em (I.3), acrescentanios as variáveis escalares e (I.8.1) - (I.8.2) transforma-se

Maximizar $f = c^{*T} x^*$ (I.8.3)

Sujeito a $A^* x^* = b$ (I.8.4)

onde

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

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.

I.9.1 - Definição Geral dos Problemas Duais

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

$$\text{Minimize } f = \sum_{j \varepsilon N} c_j \, x_j \qquad \text{(I.9.1)}$$

Sujeito a

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

$$\sum_{j \in N} a_{ij} x_j = b_i \qquad i \in \overline{M}_1 \tag{I.9.2}$$

$$x_j \geq 0 \qquad i \in N_1$$

$$x_j \text{ qualquer} \qquad j \in \overline{N}_1$$

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

O problema Dual: determinar um vetor linha $W = (w_1, w_2, \ldots, w_m)$ tal que

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

Sujeito a

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

$$w_i \text{ qualquer} \qquad i \in \overline{M}_1$$

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

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

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

Então o teorema da dualidade afirma

<u>TEOREMA</u>: O mínimo de (I.9.1) sujeito a (I.9.2) e igual ao máximo de (I.9.3) sujeito a (I.9.4).

CAPÍTULO II

UMA NOVA ALTERNATIVA SOBRE A SOLUÇÃO DOS PROBLEMAS DE PROGRAMAÇÃO LINEAR

II.1 - DEFINIÇÃO DO PROBLEMA

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

II.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 mxn,

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

$b \in R^m$

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

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 - PSEUDOINVERSA, 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^i_{j-1} queremos indicar a matriz obtida a partir de A_j quando dela omitimos a i-ésirna 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

$$b = V + {}^2x_i \, a^i + {}^2x_j \, a^j \tag{II.2.1}$$

Mostremos então que:

$$V = P_2 b$$

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

onde $A_2^+ = |A_2^T \; A_2|^{-1} A_2^T$ é a pseudoinversa de A_2 quando suas colunas são linearmente independentes.

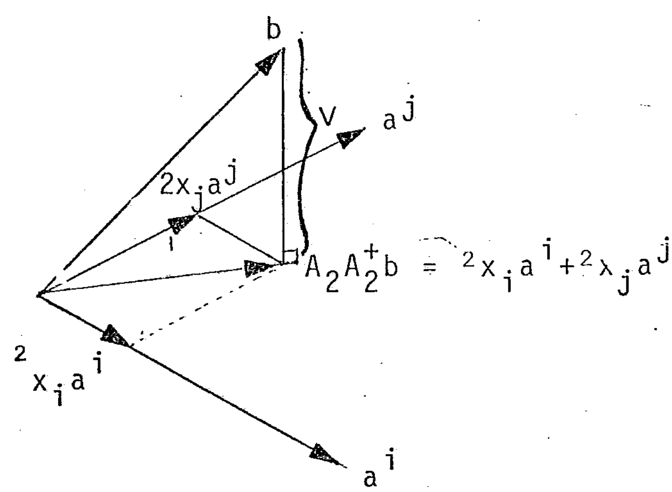

Fig. II.1

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

$$<b, a^i> = <a^i, a^i> \cdot {}^2x_i + <a^i, a^j> \cdot {}^2x_j$$

$$<b, a^j> = <a^j, a^i> \cdot {}^2x_i + <a^j, a^j> \cdot {}^2x_j$$

Pondo na forma matricial

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

Tomando a inversa

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

isto é

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

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 $2_2 = [I - A_2 A_2^+]$ temos:

$V = P_2 b$ 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.

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

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

II.3 - CONFIGURAÇÃO DE A_k^+ EM TERMOS DE A_{k-1}^+ e P_{k-1}

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

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

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:

$$\left[\begin{array}{c} <b, a^i> \\ <b, a^j> \\ \hline <b, a^k> \end{array}\right] = \left[\begin{array}{cc:c} <a^i, a^i> & <a^i, a^j> & <a^i, a^k> \\ <a^j, a^i> & <a^j, a^j> & <a^j, a^k> \\ \hline <a^k, a^i> & <a^k, a^j> & <a^k, a^k> \end{array}\right] \left[\begin{array}{c} {}^3x_i \\ {}^3x_j \\ \hline {}^3x_k \end{array}\right]$$

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

onde

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

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

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

De (II.3.2) tiramos

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

$$\begin{cases} [A_2^T A_2] \begin{bmatrix} {}^3x_i \\ {}^3x_j \end{bmatrix} + B_3\ {}^3x_k = \begin{bmatrix} \langle b, a^i\rangle \\ \langle b, a^j\rangle \end{bmatrix} & \text{(II.3.3)} \\ [D_3 - B_3^T (A_2^T A_2)^{-1} B_3]\ {}^3x_k = \langle b, a^k\rangle - B_3^T [A_2^T A_2]^{-1} \begin{bmatrix} \langle b, a^i\rangle \\ \langle b, a^j\rangle \end{bmatrix} & \text{(II.3.4)} \end{cases}$$

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:

$$ {}^3x_k = \frac{<b,\ a^k> - B_3^T[A_2^TA_2]^{-1} \begin{bmatrix} <b,\ a^i> \\ <b,\ a^j> \end{bmatrix}}{D_3 - B_3^T[A_2^TA_2]^{-1}\ B_3} \qquad \text{(II.3.5)} $$

Decorre de (II.3.5) que

$$ \begin{bmatrix} {}^3x_i \\ {}^3x_j \end{bmatrix} = [A_2^TA_2]^{-1} \begin{bmatrix} <b,\ a^i> \\ <b,\ a^j> \end{bmatrix} - \frac{[A_2^TA_2]^{-1} B_3 \left[<b,\ a^k> - B_3^T[A_2^TA_2]^{-1} \begin{bmatrix} <b,\ a^i> \\ <b,\ a^j> \end{bmatrix} \right]}{D_3 - B_3^T[A_2^TA_2]^{-1}\ B_3} \qquad \text{(II.3.6)} $$

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

$$ \begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} [A_2^TA_2]^{-1} + \dfrac{[A_2^TA_2]^{-1}B_3B_3^T[A_2^TA_2]^{-1}}{\Delta_3} & -\dfrac{[A_2^TA_2]^{-1}B_3}{\Delta_3} \\ -\dfrac{B_3^T[A_2^TA_2]^{-1}}{\Delta_3} & \dfrac{1}{\Delta_3} \end{bmatrix} \begin{bmatrix} <b,a^i> \\ <b,a^j> \\ <b,a^k> \end{bmatrix} $$

onde $\Delta_3 = D_3 - B_3^T[A_2^TA_2]^{-1}\ B_3$. Prosseguindo

$$ \begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} [A_2^TA_2]^{-1} + \dfrac{[A_2^TA_2]^{-1}A_2^T\ a^k(a^k)^TA_2[A_2^TA_2]^{-1}}{\Delta_3} & -\dfrac{[A_2^TA_2]^{-1}A_2^Ta^k}{\Delta_3} \\ -\dfrac{(a^k)^TA_2[A_2^TA_2]^{-1}}{\Delta_3} & \dfrac{1}{\Delta_3} \end{bmatrix} \begin{bmatrix} A_2^T \\ (a^k)^T \end{bmatrix} [b] $$

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} \begin{bmatrix} (A_2^TA_2)^{-1}A_2^T + \dfrac{[A_2^TA_2]^{-1}A_2^Ta^k(a^k)^TA_2[A_2^TA_2]^{-1}A_2^T}{\Delta_3} - \dfrac{(A_2^TA_2)^{-1}A_2^Ta^k(a^k)^T}{\Delta_3} \\ \dfrac{(a^k)^T - (a^k)^TA_2[A_2^TA_2]^{-1}A_2^T}{\Delta_3} \end{bmatrix} |b|$$

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} A_2^+ + \dfrac{A_2^+a^k(a^k)^TA_2A_2^+ - A_2^+a^k(a^k)^T}{\Delta_3} \\ \dfrac{(a^k)^T[I - A_2A_2^+]}{\Delta_3} \end{bmatrix} [b]$$

$$\begin{bmatrix} {}^3x_i \\ {}^3x_j \\ {}^3x_k \end{bmatrix} = \begin{bmatrix} A_2^+ - \dfrac{A_2^+a^k(a^k)^T[I - A_2A_2^+]}{\Delta_3} \\ \dfrac{(a^k)^T\ P_2}{\Delta_3} \end{bmatrix} [b] \qquad (II.3.7)$$

Mas

$$\Delta_3 = D_3 - B_3^T[A_2^TA_2]^{-1}B_3 = (a^k)^Ta^k - (a^k)^TA_2[A_2^TA_2]^{-1}A_2^T\ a^k$$

$$\Delta_3 = (a^k)^T[I - A_2A_2^+]a^k = \langle a^k, P_2\ a^k \rangle$$

Podemos agora escrever (II.3.7) como segue: (para $k = 3$)

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

O que nos leva a generalizar:

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

.Podemos ainda concluir

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

O índice superior e à 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, $i = 1, \ldots, k-1$, o que contraria (no nosso caso) a hipótese dos $a^i = 1, \ldots, k$ 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:

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

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

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

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

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

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

Portanto no passo m teremos:

$$\begin{bmatrix} {}^mx_1 \\ {}^mx_2 \\ \vdots \\ {}^mx_{m-1} \end{bmatrix} = I_{m-1} \begin{bmatrix} {}^1x_1 \\ {}^2x_2 \\ \vdots \\ {}^{m-1}x_{m-1} \end{bmatrix} - \begin{bmatrix} A_1^+a & A_2^+a & \dots & A_{m-2}^+a^{m-1} & A_{m-1}^+a^m \\ 0 & & & & \\ 0 & 0 & & & \\ \vdots & \vdots & & & \\ 0 & 0 & & 0 & \end{bmatrix} \begin{bmatrix} {}^2x_2 \\ {}^3x_3 \\ \vdots \\ {}^{m-1}x_{m-1} \\ {}^mx_m \end{bmatrix} \qquad (II.4.1)$$

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

Consideremos a seguir que:

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

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

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

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

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

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

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

Generalizando temos:

$$ {}^jX^1 = {}^{j-1}X + {}^jx_j \; A_{j-1}^+ \; a^j \qquad j = 2, \ldots, m \tag{II.4.2}$$

ou

$$ {}^jX^1 = A_{j-1}^+ [b - {}^jx_j \; a^j] \qquad j = 2, \ldots, m \tag{II.4.3}$$

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

$$ {}^jX^1 = [{}^jx_1 \;\; {}^jx_2 \; \ldots \; {}^jx_{j-2} \;\; {}^jx_{j-1}]^T $$

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

II.5 - CONE C

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

$C[A]$ 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

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

II.5.b - Hiperplano em R^n

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

Se Z = 0 o hiperplano passa pela origem.

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

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

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

II.5.d - Poliedro

Um subconjunto P de um espaço vetorial real de dimensão finita é chamado poliedro, se P é a interseção de um número finito de semi-espaços fechados.

II.5.e - Face

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

II.5.f - Cone Poliedríco Convexo

Um cone poliédrico convexo C, é a envoltória convexa de um conjunto finito de raios vetores.

Portanto se os (vetores)pontos $a^i \neq 0$ $i=1,\ldots,h$ geram raios vetores, então C é a coleção dos pontos

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

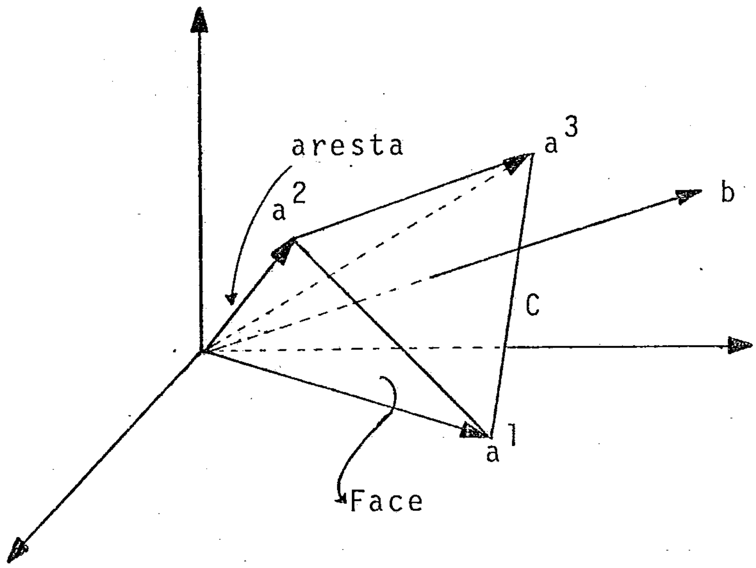

Fig. II.2

Assim se A é uma matriz $m \times k$, então o conjunto dos pontos

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

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

$$Ax = b$$

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

$<a^j, P_k b> \geq 0 \quad j = k + 1, \ .., n$, onde

$$A = (a^1, \ldots, a^n), \ A^+ = \begin{bmatrix} C_1^T \\ \vdots \\ C_k^T \\ C_{k+1}^T \\ \vdots \\ C_n^T \end{bmatrix}, \ b \ \varepsilon \ C[a^1, \ldots, a^n]$$

e $\dim(C) = n$.

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

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

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

Portanto basta mostrar que:

$$\gamma_j = <a^j, P_k b> \quad \forall_j, \ j = k + 1, \ \ldots, n$$

Pelo teorema 1, $<a^j \, P_k b> = 0 \; \forall_j = 1, \ldots, k.$

Por outro lado

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

$$<a^j, P_k b> = <a^j . (\gamma_{k+1} C_{k+1} + \ldots \gamma_j C_j + \ldots \gamma_n C_n)>$$

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

$$<a^j, P_k b> = \gamma_j \qquad j = k + 1, \ldots, n.$$

Como elucidação dessas idéias, vejamos o seguinte exemplo no R^3; sejam

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

<u>OBS</u>: Quando estamos no passo um, o índice superior e à esquerda indicado em ${}^iP_1$ corresponde àquele vetor com o qual ${}^iP_1$

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

Sabemos que:

$$^{1}P_1 b = \langle a^2, {}^{1}P_1 b\rangle\ C_2 + \langle a^3, {}^{1}P_1 b\rangle\ C_3$$

$$^{2}P_1 b = \langle a^1, {}^{2}P_1 b\rangle\ C_1 + \langle a^3, {}^{2}P_1 b\rangle\ C_3$$

$$^{3}P_1 b = \langle a^1, {}^{3}P_1 b\rangle\ C_1 + \langle a^2, {}^{3}P_1 b\rangle\ C_2$$

$$\text{Se } {}^{1}P_1 b \in C[C_2\ \ C_3] \rightarrow \begin{cases} \langle a^2, {}^{1}P_1 b\rangle \geq 0 \\ \langle a^3, {}^{1}P_1 b\rangle \geq 0 \end{cases} \qquad (II.5.2)$$

$$^{2}P_1 b \in C[C_1\ \ C_3] \rightarrow \begin{cases} \langle a^1, {}^{2}P_1 b\rangle \geq 0 \\ \langle a^3, {}^{2}P_1 b\rangle \geq 0 \end{cases} \qquad (II.5.3)$$

$$^{3}P_1 b \in C[C_1\ \ C_2] \rightarrow \begin{cases} \langle a^1, {}^{3}P_1 b\rangle \geq 0 \\ \langle a^2\ {}^{3}P_1 b\rangle \geq 0 \end{cases} \qquad (II.5.4)$$

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

$$^{1}P_1 b = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \rightarrow \begin{cases} \langle a^2, {}^{1}P_1 b\rangle = 0 \\ \langle a^3, {}^{1}P_1 b\rangle = 2 \end{cases} \quad (II.5.2) \text{ é satisfeita}$$

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

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

Assim(II.5.3) **não é satisfeita.**

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

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

Concluimos **assim** que:

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

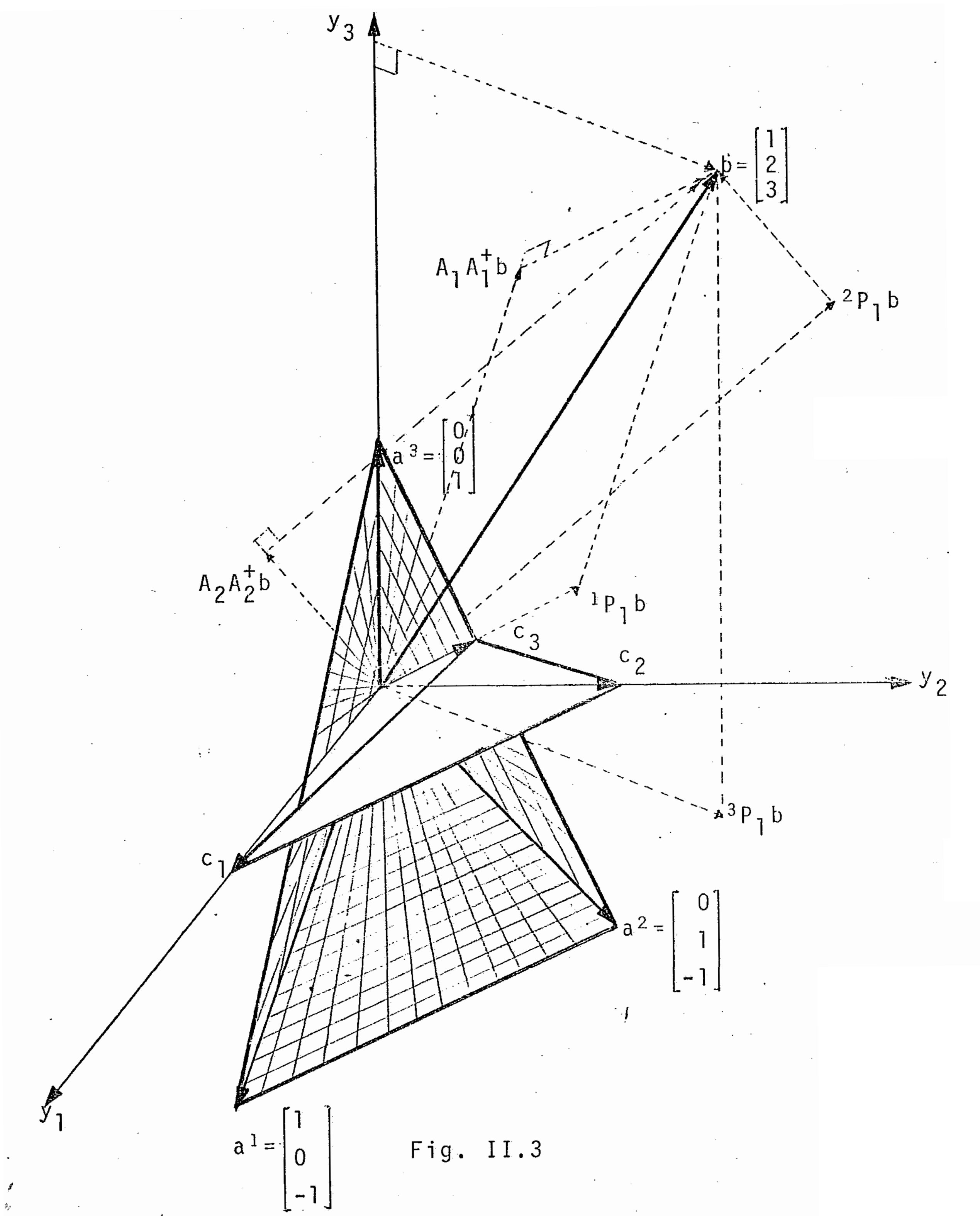

Fig. II.3

Teorema 2: (Teorema da projeção de Perez)

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

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

onde C_i^T é a i-ésima linha de A.

Antes de entrarmos no mérito da demo'nstração do teorema, veja.-o no R^3 através da Fig. 11.3.

<u>Prova</u>: Se as colunas de $A = [a^1, \ldots, a^k]$ são LI então

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

é uma matriz (kxk).

$A^T = [a^1, a^s, \ldots, a^m]$ é uma matriz kxm e

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

Então $C_i^T = d_i A^T$

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

$P^j_{k-1}\, a^j = \alpha_j\, C_j$ e daí, por(II.5.1) temos

$$<a^i\, P^j_{k-1}\, a^j> = \alpha_j < a^i C_j > = \begin{cases} 0 & \forall i \neq j \\ \alpha_j & \forall i = j \end{cases}$$

o que nos leva a afirmar

$$<a^i, C_j> = \begin{cases} 0 & \forall i \neq j \\ 1 & \forall i = j \end{cases}$$

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 $d_{ij} = C_j^T\, C.$ 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 $\phi_j = \dfrac{C_i^T\, C_j}{||C_i||^2}$ e observando que $\dfrac{C_i^T\, C_i}{||C_i||^2} = 1$, temos

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

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

Aplicando a projeção em a^i vem:

$$P_{1\ldots i-1,i+1\ldots k} \, a^i = P_{1\ldots i-1\ i+1\ldots k} \frac{\text{[redacted]}}{||C_i||^2} - \sum_{j \neq i} \phi_j \, P_{1\ldots i-1,i+1\ldots k}$$

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

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

$$P_{1\ldots i-1\ i+1\ldots k} \, a^i = [I - B_{1\ldots i-1\ i+1\ldots k}] \frac{C_i}{||C_i||^2}$$

$$= \frac{C_i}{||C_i||^2} - B_{1\ldots i-1,i+1\ldots k} \frac{C_i}{||C_i||^2}$$

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

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 $A_m = (a^1, \ldots, a^m) \rightarrow A_m^+ = \begin{bmatrix} C_1^T \\ \vdots \\ C_m^T \end{bmatrix}$

Pelo teorema 2 $P_{m-1}^i \; a^i = \frac{C_i}{||C_i||^2}$

Então $<b, P_{m-1}^i \; a^i> = <b, \frac{C_i}{||C_i||^2}>$

Assim

$$x_i = <C_i, b> \geq 0 \leftrightarrow <b, P_{m-1}^i \; a^i> \geq 0$$

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

Para mostrar que $x = [C_1^T b \; \ldots \; C_j^T b]^T = A^+ b$ é o único vetor de norma mínima entre aqueles que minimizam $||b-Ax||^2$, veja |1, pp. 19-20|.

Voltando à Figura I.3 observa-se que

Se $C_3 \begin{cases} \perp a^1 \\ \perp a^2 \end{cases} \rightarrow$ C_3 é perpendicular ao plano determinado a^1 e a^2. além do mais $\langle b^{21}P_2 a^3\rangle = \langle b^T, \dfrac{C_3}{||C_3||^2}\rangle > 0$. Isto significa dizer que b pertence ao semi-espaço cujo hiperplano fronteira passando pela origem é o plano determinado por a^1 e a^2. (II.5.7.a)

Se $C_2 \begin{cases} \perp a^1 \\ \perp a^3 \end{cases} \rightarrow$ $\left(\text{Se } C_1 \begin{cases} \perp a^2 \\ \perp a^3 \end{cases} \rightarrow\right)$ $C_2(C_1)$ é perpendicular ao plano determinado por a^1 e a^3 (a^2 e a^3). Além do mais se $\langle b, {}^{13}P_2 a^2\rangle = \dfrac{b^T C_2}{||C_2||^2} > 0$ $(\langle b, {}^{23}P_2 a^1\rangle = \dfrac{b^T C_1}{||C_1||^2} > 0)$ isto significa dizer que b pertence ao semi espaço cujo hiperplano fronteira passando pela origem é o plano determinado por a^1 e a3 (a^2 e a^3). (II.5.7.b)

conclusão: Por (II.5.7.a) e (II.5.7.b) concluinos que b esta 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

$[a^1, a_2, a_3]\ x = b$ 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 $m \geq 3$, $a^1, \ldots, a^{m-1}$ geram uma face do cone $C[a^1, \ldots, a^n] \leftrightarrow \langle a^j, P_{m-1}b\rangle \geq 0$

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

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

Se com C denotamos o cone, e com F_1 a face, então existe um hiperplano sobre o qual a face F_1 repousa, digamos $H_1 = \{x \text{ t.q. } <x,\mu_1> = 0\}$, de tal modo que $F_1 = H_1 \cap C$, onde C está contido no semi-espaço $\{x \text{ t.q. } <x,\mu_1> \geq 0\}$ e μ_1 é um vetor unitário ortogonal ao hiperplano H_1.

Como $b \in C$ e $a^j \in C$ $\forall j$, temos respectivamente $<b,\mu_1> \geq 0$ e $<a^j, \mu_1> > 0$ $\forall j$. Observa-se que $\mu \in R(P_{m-1})$, portanto $P_{m-1}b = \frac{}{||\mu_1||^2} \mu_1$ Consequentemente

$$<a^j, P_{m-1}b> = <a^j, \frac{<\mu_1,b>}{||\mu_1||^2} \mu_1> = \frac{<\mu_1,b>}{||\mu_1||^2} <a^j, \mu_1> \geq 0 \quad \forall j$$

Inversamente, admitamos agora que $<a^j,P_{m-1}b> \geq 0$ $\forall j$ e que $a^1, \ldots, a^{m-1}$ não geram uma face do cone. Então existe no mínimo um índice k tal que $<a^k \mu_1> < 0$. Mas isto significa dizer que

$$<a^k,P_{m-1}\ b> = \frac{<\mu_1, b>}{||\mu_1||} \frac{<a_k,\mu_1>}{||\mu_1||} < 0$$

o que é uma contradição. Observa-se que esta prova esta fundamentada a $a^1, \ldots, a^{m-1}$, $m \geq 3$ 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 $C[a^1, \ldots, a^n]$, então não podemos garantir que $<b,\mu_1><a^j,\mu_1> \geq 0$ $\forall j$, pois para cada i existe um feixe de hiperplanos suportes, consequentemente diversos μ_1, um para cada hiperplano do feixe. Portanto pode existir a^i tal que $<a^j, {}^iP_1b> < 0$ e entretanto este a^i gera uma aresta do cone. Veja o exemplo numérico, página 65 onde a^2 e a^3 geram arestas do cone e $<a^1, {}^2P_1 b><a^4\ {}^3P_1 b> < 0$. então $<a^1, {}^2P_1 b> < 0$, $<a^4, {}^3P_1 b> < 0$.

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

Prova: Consequência imediata do teorema'.

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

Prova: Temos que

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

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

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

Agora pelo teorema 4 temos $<a^j, P_{m-1} b> \geq 0 \quad \forall j = 1, \ldots, n$, assim

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

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

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

De (II.5.9) tiramos

$$<a^j,P_{m-2}b> \geq \frac{<a^j,P_{m-2}a^{m-1}>}{||P_{m-2}a^{m-1}||} \cdot \frac{<a^{m-1},P_{m-2}b>}{||P_{m-2}a^{m-1}||} \geq 0 \quad \forall j = 1, \ldots, n$$

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

Como $<a^j,P_{m-2}a^{m-1}><a^{m-1},P_{m-2}b> = <a^{m-1},P_{m-2}a^j><a^{m-7},P_{m-2}b>$ e, tendo-se em vista que os vetores $P_{m-2}a^j$ e $P_{m-2}b$ são colineares, e estão num niesmo semi-espaço segue-se que $<a^{m-1},P_{m-2}a^j>$ e $<a^{m-1},P_{m-2}b>$ tem o mesmo sinal, o que nos leva a afirmar que o seu produto

$$<a^j,P_{m-2}a^{m-1}><a^{m-1},P_{m-2}b> \geq 0 \quad \forall j = 1, \ldots, n$$

Assim para k = 2

$$<a^j,P_{m-2}b> \geq 0 \quad \forall j = 1, \ldots, n.$$

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

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

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

Haja visto que o numerador do 2º membro é 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 $<a^j, b> \geq 0$. Se $L_1 = \phi$, então não existe sol. viável.

b) Seja L_2 um subconjunto dos índices dos vetores de L_1 tal que $<a^i, P_j b> \geq 0 \quad \forall i = 1, \ldots, n$.

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.

$$\min_j Z = c_j x_j \text{ onde}$$

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

STEP 2

Tendo selecionado **t** vetores, para **t** = 1, ..., m-2, escolha o próximo conforme:

a) Selecione um conjunto de k vetores tais que

$$A^+_{1...t,j}\, b \geq 0 \text{ para } j = 1, \ldots, k$$

b) Para cada um dos k vetores acima compute

$$\langle a^L P_{1...t,j} b \rangle \text{ para } L \in \{1...N\}-\{1...t,j\}$$

c) Para aqueles j tais que $\langle a^L, P_{1...t,j}\, b \rangle \geq 0$ substitua b por b_j ($b_j = x_1A^1+...+x_tA^t+x_jA^j$), e se existem muitos, admita um tal que o mínimo do funcional, seja

$$\min_j Z = x_1c_1 + \ldots x_tc_t + x_jx_j \text{ onde}$$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_t \\ x_j \end{bmatrix} = [A_{1...t,j}]^+ b \geq 0 \quad \text{(dado por a)}$$

STEP 3:

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

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

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

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

onde $x = [x_1, x_2, x_3, x_4]^T$ e $A_{3x4} = [A^1, A^2, A^3, A_4]$

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

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

Admitamos que o mínimo do funcional ocorreu para j = 1. No 2º Step o algoritmo substitue o probleiiia original por k = 3 problemas aproximados

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

Admitdamos que o mínimo do funcional ocorreu para j = 2.

No 3º Step o algoritmo calcula

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

Note que o problema original é

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

Suponhamos que a sol. de P_{3_j} seja $x = (x_1\ x_2\ x_3\ 0)$ resta mostrar que a solução de P_{3_j} é a mesma de $\overline{P}_3$ (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 é válido.

No Step 3 tinhamos:

$$Z_{1\ldots m-1,m} \leq Z_{1\ldots m-1,k} \rightarrow \tag{II.5.10}$$

$$\rightarrow (x_1c_1 + \ldots + x_{m-1}c_{m-1}) + x_m c_m \leq (x_1'c + \ldots + x_{m-1}'c_{m-1}) + x_m'c_m$$

Fazendo

$$\underset{\sim}{x}_{m-1} = (x_1 \; \ldots \; x_{m-1})$$

$$\underset{\sim}{c}_{m-1} = (c_1 \; \ldots \; c_{m-1})$$

$$\underset{\sim}{x}_{m-1}' = (x_1' \; \ldots \; x_{m-1}')$$

Temos:

$$<\underset{\sim}{x}_{m-1}, \underset{\sim}{c}_{m-1}> + x_m c_m \leq <\underset{\sim}{x}_{m-1}', c_{m-1}> + x_j c_j \tag{II.5.11}$$

Podemos ainda escrever

$$[A_{m-1} \; \vdots \; A_m] \begin{bmatrix} \underset{\sim}{x}_{m-1} \\ x_m \end{bmatrix} = b \rightarrow A_{m-1} \underset{\sim}{x}_{m-1} + A_m x_m = b \rightarrow$$

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

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

$$\underset{\sim}{x}_{m-1} = A_{m-1}^{+} b - x_m A_{m-1}^{+} a^m \tag{II.5.12}$$

. de modo analógo

$$\underset{\sim}{x}'_{m-1} = A^{+}_{m-1}b - x_j\, A^{+}_{m-1}a^j \tag{II.5.13}$$

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

$$<(A^{+}_{m-1}b-x_mA^{+}_{m-1}a^m),\underset{\sim}{c}_{m-1}> + x_mc_m \leq <(A^{+}_{m-1}b-x_jA^{+}_{m-1}a^j),\underset{\sim}{c}_{m-1}>+ x_jc_j$$

$$<A^{+}_{m-1}\, b,\underset{\sim}{c}_{m-1}>-x_m<A^{+}_{m-1}a^m,\underset{\sim}{c}_{m-1}>+x_mc_m \leq <A^{+}_{m-1}b,c_{m-1}>-x_j<A^{+}_{m-1}\underset{\sim}{c}_{M-1}>+x_jc_j$$

$$x_mc_m-x_m<A^{+}_{m-1}a^m,\underset{\sim}{c}_{m-1}> \leq x_jc_j-x_j<A^{+}_{m-1}a^j,c_{m-1}>$$

$$x_m\left[<A^{+}_{m-1}a^m,\underset{\sim}{c}_{m-1}>-c_m\right]+x_j\left[c_j-<A^{+}_{m-1}a^j,\underset{\sim}{c}_{m-1}>\right] \geq 0 \tag{II.5.14}$$

Mas por (II.4.1)

$$x_m = \frac{<a^m,P_{m-1}b>}{<a^m,P_{m-1}a^m>}$$

e

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

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