

ESTUDOS EM PROGRAMAÇÃO LINEAR

TESE DE DOUTORADO
apresentado à

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE CAMPINAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

por

RAUL VINHAS RIBEIRO

ESTUDOS EM PROGRAMAÇÃO LINEAR

Orientador: HERMANO DE MEDEIROS FERREIRA TAVARES

NOVEMBRO-80

UNICAMP
BIBLIOTECA CENTRAL

A G R A D E C I M E N T O S

Faço meu todo um parágrafo do Jura na página de agradecimentos de sua tese:

As idéias contidas neste trabalho não foram frutos exclusivos de uma única pessoa. Não surgiram de um estalo da imaginação. Ao contrário, germinaram num processo de muitas dúvidas e certezas compartilhadas com os amigos, dia após dia.

Apesar de haver muitos "em particulares", em particular nossos agradecimentos ao Authiê, o francês amigo, com ele fizemos a parte II desta tese e não saberíamos dizer o que nela é idéia nossa ou dele.

Ao Akebo, Dino, França, Hermano e Jura, amigos e companheiros de tantos anos, nas dificuldades, nas pequenas vitórias e derrotas de todo dia, nossos agradecimentos pela presença sempre constante, pelos estímulos nas horas necessárias.

À Matilde, a companheira, ao Velho Ribeiro e Dona Yone, nossos agradecimentos pela torcida, pelas rezas aos santos fortes por este filho incrédulo, mas que valeu.

À Taka, meus agradecimentos por esta ótima apresentação de nosso trabalho.

Este trabalho contou com o suporte financeiro do CNPq.

À Matilde , Marisa, Mariana e Thiago

À meus pais e irmãos

À dona Geralda, Ivanil e Heleno

S U M Á R I O

O presente trabalho, dividido em três partes, trata na primeira delas de mostrar como formalmente os métodos de programação linear, os métodos simplex e gradiente reduzido em PL, podem ser vistos como problemas de norma mínima ou projeção. É apresentado ainda o método simplex para restrições de desigualdade e pivoteamento por coluna. Na segunda parte apresentamos um método para problemas de PL com estrutura bloco angular, dual do GGUB (generalized generalized upper bound) e mostramos que o método de ROSEN é uma variante de nosso método que a cada iteração deve satisfazer uma propriedade adicional. Fazemos também uma interpretação geométrica do método. Na última parte fazemos uma apresentação do método primal para programação linear por partes, procurando explicitá-lo como generalização do método simplex e finalmente fazemos, baseado na parte II desta tese, uma especialização do método para problemas lineares por parte com estrutura bloco angular.

Í N D I C E

PARTE I: MÉTODOS DE DIREÇÕES FACTÍVEIS EM PL

Pág.

INTRODUÇÃO

CAPÍTULO I

MÉTODOS DE DIREÇÕES FACTÍVEIS

I. O Problema

II. Resolução de (P) por Métodos de Direções Factíveis

CAPÍTULO II

MÉTODOS DE DIREÇÕES FACTÍVEIS COM MUDANÇA DE BASE NO
ESPAÇO DAS VARIÁVEIS

I. Introdução

II. Transformação de Base: Problema (P)

III. Transformação de Base: Problema(D)

IV. O Problema da Geração de Direção Usável

V. O Método Simplex como um Método de Projeção

CAPÍTULO III

MÉTODOS SIMPLEX E GRADIENTE REDUZIDO COM RESTRIÇÕES DE
DESIGUALDADE

I. Introdução

II. Método Simplex

III. Quadro Simplex

IV. Método de Gradiente Reduzido

CAPÍTULO IV

CONCLUSÕES

PARTE II: PROBLEMAS DE PROGRAMAÇÃO LINEAR COM ESTRUTURA
BLOCO ANGULAR

Pág.

INTRODUÇÃO

CAPÍTULO V

O PROBLEMA DE PL COM ESTRUTURA BLOCO ANGULAR

I. O Problema

II. Notação Utilizada

III. Solução Básica Correspondente à Base ^B

CAPÍTULO VI

RESOLUÇÃO DO PROBLEMA BLOCO ANGULAR: MÉTODO PRIMAL

I. Introdução

II. Método Primal Simplex para (P)

III. Atualização da Base

CAPÍTULO VII

RESOLUÇÃO DO PROBLEMA BLOCO ANGULAR: MÉTODO DUAL

I. Dual do Problema Bloco Angular

II. Fase 1 - Procura de uma Base Dual Admissível

III. Fase 2 - Otimização do Problema (D)

IV. Estratégia de Aplicação do Método Dual

V. Variante do Método Dual Proposto

CAPÍTULO VIII

INTERPRETAÇÃO GEOMÉTRICA DO MÉTODO DUAL, CONCLUSÃO DA
PARTE II:

I. Interpretação Geométrica

II. Conclusão

P A R T E I

MÉTODOS DE DIREÇÕES FACTÍVEIS EM PL

INTRODUÇÃO

No Capítulo I, após apresentarmos o problema de Programação Linear (PPL), definimos os métodos de direções factíveis que podem ser, seguindo ZOUTENDIJK [5], enquadrados em três classes: Métodos de Otimização, Métodos Factíveis e Métodos Globais. Os métodos de otimização, para que tenham solução finita, implicam na definição de uma norma ou pseudo-norma de forma a limitar o conjunto das direções factíveis. Anotando a direção por s , a norma $||s||$, em geral, utilizada é l_1 , l_2 ou l_∞ .

Para as normas acima prova-se que procurar uma direção factível de crescimento da função objetivo é equivalente a resolver um problema de norma mínima (PNM) ou, no caso da norma l_2 , a resolver um problema de projeção (PP).

Vamos generalizar a norma acima para $||Ms||$, onde M é uma matriz quadrada regular. Nosso intento é mostrar que para uma matriz M particular, os métodos simplex e gradiente reduzido, podem ser vistos como problemas de norma mínima ou de projeção.

Neste sentido, no Capítulo II, apresentamos um método de direções factíveis que pode ser colocado em dois passos. Dado uma solução factível, fazemos uma transformação de base no espaço das variáveis de maneira que o conjunto das direções factíveis possa ser expresso de forma mais simples. No segundo passo é resolvido um problema de procura de direção factível. Na verdade este método é uma apresentação algo diferente do método simplex ou gradiente reduzido que nos permitirá no item V mostrar que são casos particulares do (PNM) ou (PP).

No Capítulo II é feita uma interpretação geométrica do método simplex e gradiente reduzido que nos permite chegar a uma formalização dos dois métodos para o caso de restrições de desigualdades, formalização esta efetuada no Capítulo III.

No Capítulo IV são apresentadas as conclusões gerais desta primeira parte.

MÉTODOS DE DIREÇÕES FACTÍVEISI. O PROBLEMA

O problema que nos interessa estudar, que chamaremos problema primal, é:

$$(P) \begin{cases} \text{MIN} & z = cx \\ \text{s.a.} & Ax \leq b \end{cases}$$

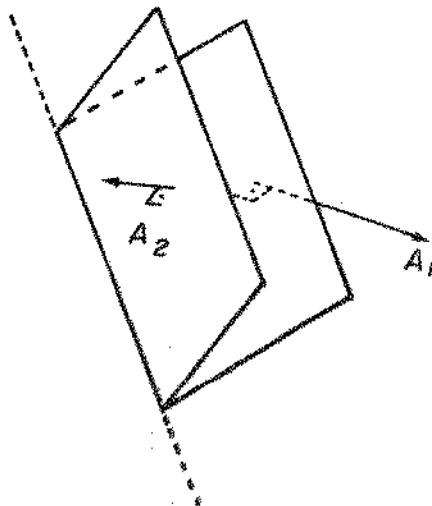
onde $c(1,n)$, $A(m,n)$ $x(n,1)$ e $m \geq n$. Suporemos que "rank" $(A) = n$. Definimos ainda:

$$S_1 = \{x/Ax \leq b\}$$

No caso em que $m < n$, S_1 é um conjunto convexo, fechado porém ilimitado. O conjunto $S_1' = \{x/Ax = b\}$ define uma variedade linear no R^n . A solução do problema, neste caso só é finita, por exemplo, se cx é paralelo a A_1x ou se c é ortogonal à S_1' (fig.1). De forma mais precisa, a solução é finita se podemos escrever:

$$-c = \lambda A, \lambda \geq 0$$

Este caso não nos interessa.



Um vértice de S_1 é constituído por n restrições saturadas. Os gradientes destas restrições definem uma base linha (GRAVES | 26 |) factível. Um valor de x que sature n_j restrições mas tal que não satisfaça as outras $n-m$ restrições não é um vértice de S_1 , mas define uma base linha infactível. O problema é degenerado se a um vértice de S_1 corresponder mais de uma base, ou seja, se para um valor de x mais do que n restrições são saturadas. Suporemos que (P) é não degenerado.

Podemos obter o problema dual de (P) construindo o lagrangeano.

$$L(x,p) = cx - p(Ax-b) \quad , \quad p \leq 0$$

A função dual é definida como

$$\begin{aligned} \phi(p) &= \text{MIN}_{x \in \mathbb{R}^n} L(x,p) \\ &= \begin{cases} pb & \text{se } pA=c \\ -\infty & \text{em caso contrário} \end{cases} \end{aligned}$$

e o problema dual pode ser escrito como

$$(D) \begin{cases} \text{MAX } \phi = pb \\ \text{s.a. } pA=c \\ p \leq 0 \end{cases}$$

Definimos ainda $S_2 = \{p/pA=c \text{ e } p \leq 0\}$

II. RESOLUÇÃO DE (P) POR MÉTODOS DE DIREÇÕES FACTÍVEIS

Dado um ponto x_0 factível e um subconjunto $I \subset \{1, 2, \dots, m\}$ de índices de restrições ativas em x_0 , um método de direcção factível para o problema (P) pode ser definido por

(1) Procurar s tal que

$$\begin{cases} cs \leq 0 & (1) \\ A_I s \leq 0 & (2) \end{cases}$$

Uma direção s que satisfaça (2) é chamada direção factível; se além disso satisfizer (1), é então denominada usável.

(2) Procurar o máximo valor do escalar t tal que ainda tenhamos

$$A(x + ts) \leq b, \quad t \geq 0$$

ZOUTENDJK | 5 | classifica os métodos de direções factíveis em três grupos:

(a) Métodos de Otimização com uma subclasse dada pelos métodos de projeção. Nesta classe a direção s a ser encontrada deve ser ótima segundo algum critério.

(b) Métodos Factíveis, neste caso, procuramos tão somente uma solução para (1) e (2).

(c) Métodos Globais, além das informações locais em x_0 dadas por (1) e (2), levamos em consideração as outras restrições. Esta classe de métodos não nos interessa aqui.

Apresentamos a seguir a descrição de alguns métodos.

Um método de otimização de procura de direção factível pode ser definido pelo problema:

$$(GD) \begin{cases} \text{MIN} & cs \\ \text{s.a.} & A_I s \leq 0 \\ & ||Ms|| \leq 1 \end{cases}$$

que chamaremos de problema de geração de direção usável (GD). A direção s deve satisfazer alguma norma ou pseudo-norma de modo que o problema tenha solução finita. A matriz $M(n,n)$ deve ser regular. Os métodos se diferenciam pela norma utilizada. Seja $\lambda = Ms$, temos então:

(a) Norma L_1 : $\sum |\lambda_j| \leq 1$

(b) Norma L_2 : $\lambda^T \lambda \leq 1$

(c) Norma L_∞ : $-1 \leq \lambda_j \leq 1$

(d) Pseudo-Normas:
$$\begin{cases} \lambda_j \leq 1 & \text{se } c_j \leq 0 \\ \lambda_j \geq -1 & \text{se } c_j \geq 0 \end{cases}$$

$$\begin{cases} \lambda_j \leq c_j & \text{se } c_j \leq 0 \\ \lambda_j \geq -c_j & \text{se } c_j \geq 0 \end{cases}$$

A direção s encontrada por estes métodos será a de máximo decréscimo da função para a norma dada.

Para M matriz identidade, ZOUTENDIJK | 5 |/prova que (GD) é equivalente ao problema de norma mínima, apresentado no teorema abaixo, para os casos (a), (b) e (c).

Teorema 1: Se uma direção usável existe; i.é, se (1)-(2) é consistente, então a solução ótima de (GD) é equivalente, a menos de um fator de proporcionalidade, ao problema:

$$(PNM) \begin{cases} \text{MIN } ||Ms|| \\ \text{s.a. } A_I s \leq 0 \\ cs \leq -1 \end{cases}$$

Temos ainda que para o caso (b), (GD) é equivalente ao problema de projeção:

Teorema 2: Se $||Ms|| = (s^T M^T M s)^{1/2}$ então (GD) é equivalente, a menos de um fator de proporcionalidade, ao problema:

$$(PP) \begin{cases} \text{MIN } |-M^{-1T} c^T - Ms| \\ \text{s.a. } A_I s \leq 0 \end{cases}$$

O problema (GD) para as normas (a), (c) e (d) é um problema linear e para a norma (b) é um problema quadrático. No primeiro caso um método de resolução seria o simplex. Para o segundo caso teríamos:

$$\begin{cases} \text{MIN } cs \\ \text{s.a. } A_I s \leq 0 \\ s^T Q s \leq 1 \end{cases}$$

onde $Q = M^T M$ é simétrica e positiva definida. As condições de otimalidade para o problema são:

$$\begin{cases} -c^T = A_I^T u + \beta Qs \\ u^T A_I s = 0 \\ u \geq 0 \\ \beta \geq 0 \end{cases}$$

fazendo $v = -BA_I s \geq 0$ e $N = A_I Q^{-1} A_I^T$, podemos reescrever as condições anteriores como:

$$\left\{ \begin{array}{l} -Nu + v = A_I Q^{-1} c^T \quad (3) \\ u^T v = 0 \quad (4) \\ u \geq 0 \quad (5) \\ v \geq 0 \quad (6) \end{array} \right.$$

chamadas de condições de complementariedade. Obter βs implica em encontrar uma solução para o problema de complementariedade linear. Tendo-se esta solução então:

$$\beta s = -Q^{-1} c^T - Q^{-1} A_I^T u \quad (7)$$

BAZARAA e SHETTY [30], ZOUTENDIJK [5] e LEMKE [31] apresentam um algoritmo de resolução deste problema, da autoria de LEMKE. O algoritmo converge em número finito de passos desde que N satisfaça certas condições e que não haja degenerescência, ou seja, que qualquer que seja a base para (3), o lado direito desta equação seja diferente de zero.

Entre os métodos factíveis nos interessa em particular o método de projeção de ROSEN [6], [14] que se aplica ao problema.

$$(PPR) \left\{ \begin{array}{l} \text{MIN } |-M^{-1T} c^T - Ms| \\ \text{s.a. } A_I s = 0 \end{array} \right.$$

A resolução deste problema nos dá a projeção de $-M^{-1T} c^T$ no sub-espaço linear formado pela intersecção das restrições ativas em x_0 , neste sentido é um método factível mas não de otimização pois não leva em conta todas as informações locais (na verdade, em x_0 , o conjunto das direções factíveis é dado por $\{s/A_I s \leq 0\}$) e não necessariamente obteremos a direção de máximo decréscimo da função.

As condições de otimalidade para o problema são:

$$-c^T - Qs = A_I^T u \quad (8)$$

$$A_I s = 0 \quad / \quad (9)$$

e $Q = M^T M$. Premultiplicando-se (8) por $A_I Q^{-1}$ obtemos

$$-A_I Q^{-1} c^T = A_I Q^{-1} A_I^T u$$

e

$$u = -(A_I Q^{-1} A_I^T)^{-1} A_I Q^{-1} c^T \quad (10)$$

que substituindo em (8) nos permite obter s:

$$s = -Q^{-1} \left[I - A_I^T (A_I Q^{-1} A_I^T)^{-1} A_I Q^{-1} \right] c^T \quad (11)$$

ou

$$s = -Q^{-1} P_I(Q) c^T \quad (12)$$

onde $P_I(Q)$ é uma matriz de projeção. É fácil comprovar que $P_I(Q)$ é idem potente, quadrada, simétrica, positiva semi-definida cujos únicos auto-valores nulos correspondem aos auto-vetores dados pelas colunas de A_I^T . A matriz Q é positiva definida e $Q^{-1} P_I(Q)$ têm também como auto-vetores correspondentes a auto-valores nulos às colunas de A_I^T . Podemos concluir que

$$cs = -c \left[Q^{-1} P_I(Q) \right] c^T \leq 0 \quad (13)$$

a igualdade valendo somente se $s = 0$, ou equivalentemente se $-c^T = A_I^T u$.

Se $s \neq 0$, então esta direção é de decréscimo para o problema (P) em x_0 . Se $s = 0$ e $u \geq 0$ então x_0 é um ótimo para (P). Se $s = 0$ e $u \not\geq 0$, podemos relaxando uma das restrições, seja A_i com multiplicador u_i negativo, obter uma direção usável pela projeção do gradiente sobre o subespaço determinado pelas restrições restantes. Temos de (8) e do fato que $s = 0$.

$$-c^T = A_I^T u \quad (15)$$

e seja A_I^T a matriz obtida relaxando-se a restrição A_i , então

$$-c^T - Q\bar{s} = A_I^T u \quad (16)$$

onde \bar{s} é a projeção de $-c^T$ usando-se A_I . Temos $\bar{s} \neq 0$ pois de outra forma (16) seria um caso especial de (15) com $u_i = 0$, o que é impossível desde que as linhas de A_I sejam linearmente independentes, o que assumimos como hipótese. De (13) podemos concluir que $c\bar{s} < 0$. Premultipliquemos (15) por \bar{s}^T

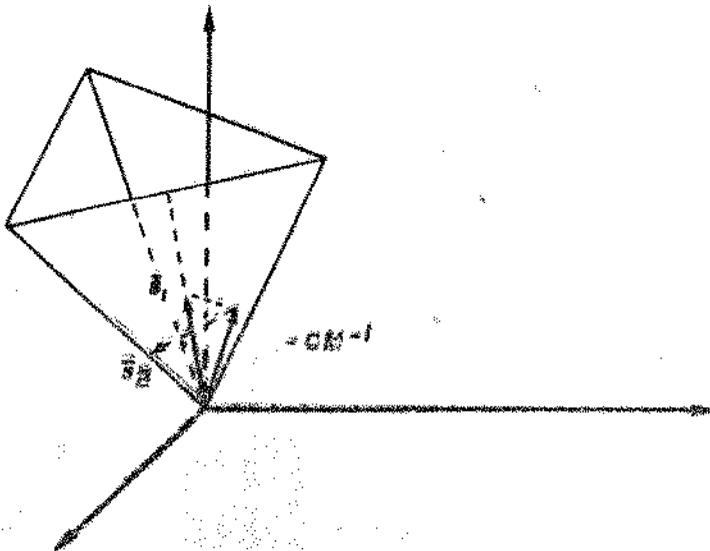
$$-\bar{s}^T c^T = \bar{s}^T A_I^T u > 0 \quad (17)$$

como $\bar{s}^T A_I^T = 0$, então

$$\bar{s}^T A_i^T u_i > 0 \quad (18)$$

mas $u_i < 0$, então $\bar{s}^T A_i^T < 0$ e a direção \bar{s} é de descida e factível pois $A_I \bar{s} = 0$ e $A_i \bar{s} < 0$.

O método de LEMKE, pelo teorema 2, é de projeção, comparando-o com o método de ROSEN, vemos que o primeiro projeta $-cM^{-1}$ no cone $\{s/A_I s < 0\}$ e o segundo no subespaço $\{s/A_I s = 0\}$. Na figura 2 \bar{s}_1 é a projeção no cone e \bar{s}_2 no subespaço de $-cM^{-1}$.



Nota: Se x_0 for um vértice do problema (P), o método de ROSEN para M matriz identidade tem interpretação geométrica se melhante ã do método simplex, mas formalmente o procedimento é diferente. Deveremos mostrar adiante que para uma matriz M particular, este método é exatamente o simplex.

MÉTODOS DE DIREÇÕES FACTÍVEIS COM MUDANÇAS DE BASE NO
ESPAÇO DAS VARIÁVEIS

I. INTRODUÇÃO

Apresentaremos alguns métodos de geração de direções factíveis que podem, para efeito de análise, serem apresentados em dois passos. No primeiro passo, fazemos uma transformação de base no espaço das variáveis de forma que o conjunto das direções factíveis possa ser expresso de uma forma mais fácil. No segundo passo é resolvido um problema de procura de uma direção usável.

Fazemos a apresentação do primeiro passo através de um exemplo, o que permite desde logo a interpretação geométrica do método. Apresentamos este primeiro passo para o problema (P) e para o problema (D).

II. TRANSFORMAÇÃO DE BASE: PROBLEMA (P)

Utilizaremos como exemplo para (P) o problema abaixo, representado na figura 3:

$$(P) \begin{cases} \text{MIN} & [-1 \quad -2/3]x & (19) \\ \text{s.a.} & \begin{bmatrix} -1 & 1 \\ -1 & 0 \\ 1 & 1 \\ 2/3 & -1/3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 5/3 \\ 0 \\ 11/3 \\ 10/9 \\ 0 \end{bmatrix} & (20) \end{cases}$$

onde $A_1 = [-1 \ 1]$, $A_2 = [-1 \ 0]$, $A_3 = [1 \ 1]$, $A_4 = [2/3 \ -1/3]$, $A_5 = [0 \ -1]$. A é a matriz formada por estes vetores linhas, nesta ordem; $b^T = [5/3 \ 0 \ 11/3 \ 10/9 \ 0]$ e $c = [-1 \ -2/3]$. Se

ja $S_1 = \{x/Ax \leq b\}$

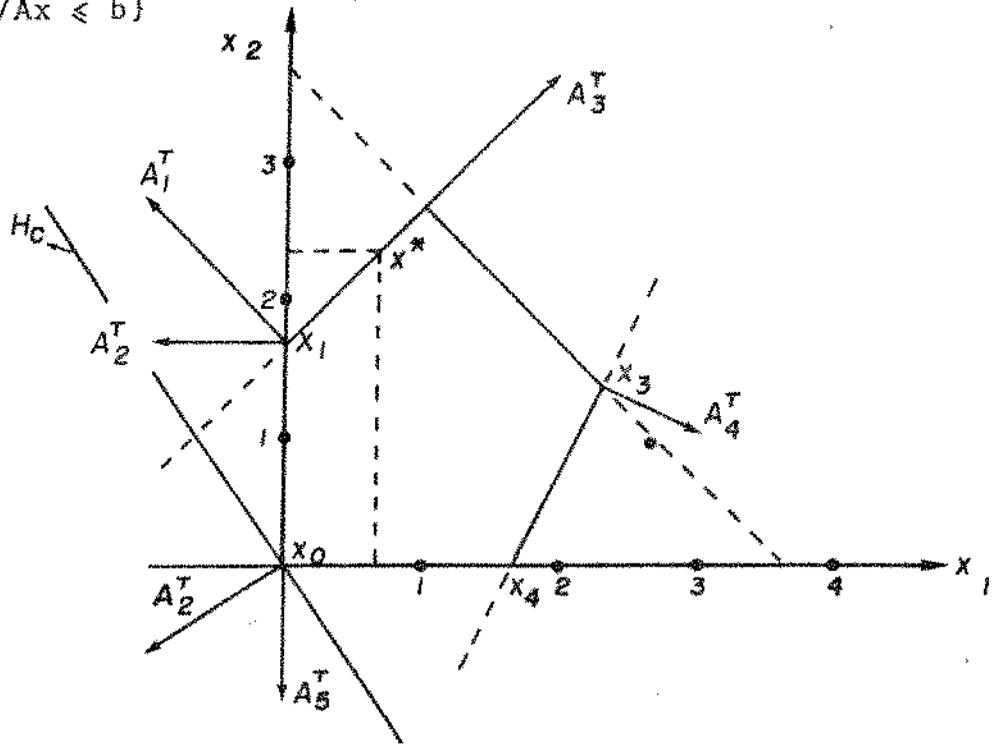
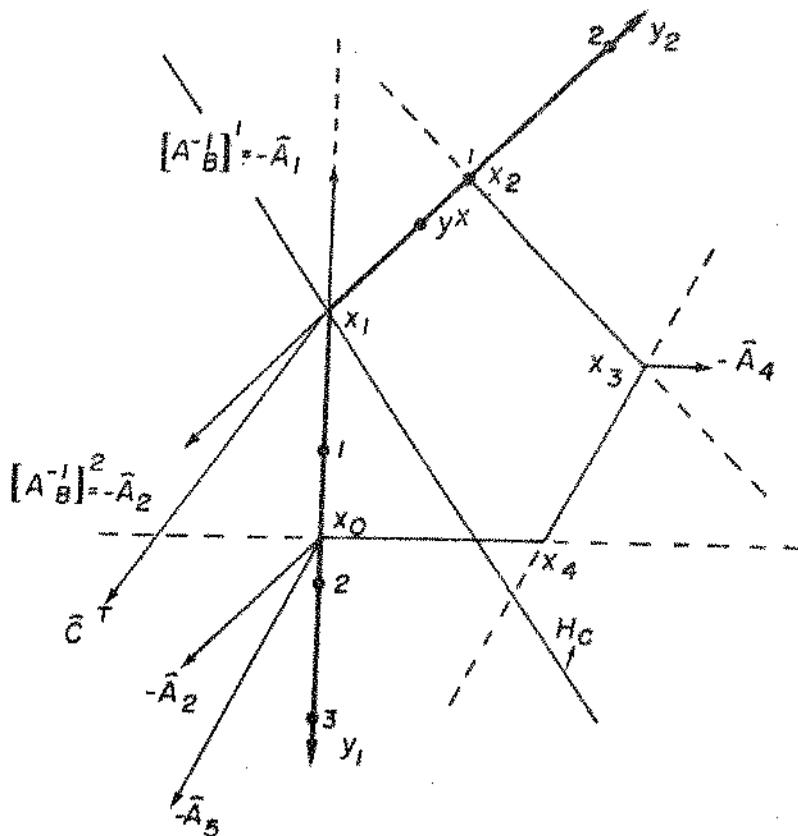


FIG 3

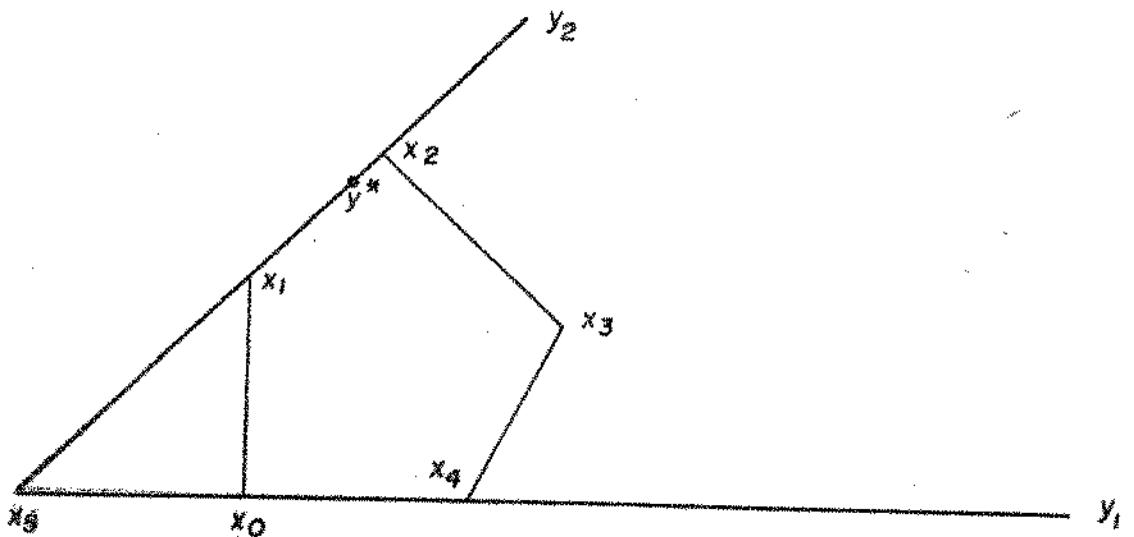
Vamos supor que conhecemos um ponto factível $x^{*T} = [2/3 \ 7/3]$ e queremos a partir deste ponto encontrar uma direção usável, ou seja, queremos $s^T = [s_1 \ s_2]$ tal que a primeira restrição não seja violada (fig. 3). Como as outras restrições estão folgadas, não precisamos nos preocupar com elas por enquanto. Queremos então s tal que

$$A_i(x + ts) \leq b_i$$

A nossa tarefa fica facilitada se considerarmos um novo sistema de coordenadas com origem no vértice x_1 e eixos nas direções das arestas que contêm este ponto. O sentido dos eixos é tal que S_1 fique contido no primeiro quadrante. Esta escolha é arbitrária, sendo feita desta forma por conveniência de exposição. Vamos anotar por $y_B^T = [y_1 \ y_2]$ a representação dos pontos do plano em termos da nova base, então o conjunto das direções factíveis é tal que $y_1 \geq 0$. B é o conjunto de índices das restrições cuja intersecção é o vértice x_1 , portanto $B = \{1, 2\}$. Esta transformação de base nos facilita a procura de uma direção usável, pois qualquer que seja o método utilizado, sua resolução é mais simples pois as únicas restrições sobre as direções são do tipo $y_i \geq 0$ (fig. 4).



Observamos que temos a possibilidade de escolher uma base tal que o ponto de origem seja infactível. Na (fig. 5) temos uma base com origem em $x_5 \notin S_1$, notar que x^* está em uma das arestas, eixo do sistema de coordenadas.



Temos então inicialmente uma base definida pelos vetores $(1\ 0)^T$ e $(0\ 1)^T$ e uma origem no ponto $x_0 = (0,0)$. Queremos passar para uma nova base definida por vetores na direção das arestas de S_1 em x_1 , com origem neste ponto. A matriz A_B^{-1} , tal que A_B é constituído pelos vetores linhas A_1 e A_2 , tem como colu

nas, vetores nas direções destas arestas. Vejamos porque: seja $[A_B^{-1}]^r$ a coluna r de A_B^{-1} , então este vetor é ortogonal a cada vetor definido pelas linhas de A_B a menos da linha r , portanto $[A_B^{-1}]^r$ é paralelo a cada hiperplano $H_i: A_i x = b_i, i \neq r$, logo é paralelo à intersecção destes hiperplanos, ou seja, tem a mesma direção de uma das arestas. As colunas de A_B^{-1} podem constituir os vetores da base que desejamos.

Denominaremos A_B de matriz básica e as restrições com índices em B de restrições básicas. As restrições com índices em \bar{B} , \bar{B} conjunto complementar de B em relação a $\{1, 2, \dots, m\}$, e $A_{\bar{B}}$ são chamadas respectivamente de restrições não básicas e matriz não básica. Para o nosso problema, temos então:

$$A_B = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} \quad (21)$$

e sua inversa

$$A_B^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \quad (22)$$

É fácil mostrar que o sentido dos eixos, de forma que o primeiro quadrante contenha S_1 , deve ser contrário ao sentido dos vetores definidos pelas colunas de A_B^{-1} .

O ponto x_1 é tal que $A_B x_1 = b_B$ e portanto $x_1 = (A_B)^{-1} b_B$. No exemplo $x_1 = (0 \quad 1^2/3)$.

Queremos então que nosso sistema de coordenadas seja definido pelos vetores colunas de $(-A_B)^{-1}$, com origem no ponto x_1 , intersecção de $H_i, i \in B$. Sempre que não houver motivo de dúvidas chamaremos a matriz $-A_B^{-1}$ de base. A transformação de representação de um vetor na base E para a base $-A_B^{-1}$ é dada por:

$$y_B = -A_B^{-1} x$$

Se, além disso, quisermos mudar a origem para o ponto x_1 , então

$$y_B = -A_B(x - x_1)$$

ou

$$y_B = -A_B x + b_B \quad (23)$$

A transformação em sentido inverso é dada por

$$x = (A_B)^{-1} (b_B - y_B) \quad (24)$$

O problema (P) pode ser reescrito como

$$\begin{cases} \text{MIN } c(A_B)^{-1} (b_B - y_B) \\ \text{s.a. } A(A_B)^{-1} (b_B - y_B) \leq b \end{cases}$$

Fazendo

$$\hat{c} = c(A_B)^{-1} \quad (25)$$

$$\hat{A}_{\bar{B}} = A_{\bar{B}} A_B^{-1} \quad (26)$$

$$\hat{b}_{\bar{B}} = b_{\bar{B}} - A_{\bar{B}} A_B^{-1} b_B \quad (27)$$

$$z_0 = \hat{c} b_B \quad (28)$$

podemos escrever o problema acima como

$$(P') \begin{cases} \text{MIN } -\hat{c} y_B + z_0 & (29) \\ \text{s.a. } -\hat{A}_{\bar{B}} y_B \leq \hat{b}_{\bar{B}} & (30) \\ y_B \geq 0 & (31) \end{cases}$$

Seja $H_c : cx$ o hiperplano minimizador, então $[-\hat{A}_{\bar{B}}]_i$ e \hat{c} são as representações dos vetores diretores dos hiperplanos H_i , $i \in \bar{B}$ e H_c na nova base. Note que, como os vetores da nova

base não formam um ângulo de 90° entre si, então os vetores diretores dos hiperplanos não são perpendiculares a estes.

Se conhecermos $x^* \in S_1$, então $y_B^* = -A_B x^* + b_B$ é factível para (P') . Se x^* é tal que as restrições $i \in B_1 \subset B$ são ativas e todas as demais folgadas, então y_B^* é tal que $y_{B_1}^* = 0$ e $y_{B_2}^* > 0$, $B_1 \oplus B_2 = B$.

Dado então y_B^* , o problema de encontrar uma direção usável é

$$\left\{ \begin{array}{l} \text{Encontrar} \quad \lambda / -\hat{c}\lambda < 0 \end{array} \right. \quad (32)$$

$$\left\{ \begin{array}{l} \text{s.a.} \quad \lambda B_1 \geq 0 \end{array} \right. \quad (33)$$

$$\left\{ \begin{array}{l} \lambda B_2 \quad \text{irrestrito} \end{array} \right. \quad (34)$$

Encontrando λ , o passo nesta direção é dado pelo maior t tal que ainda tenhamos

$$\left\{ \begin{array}{l} -\hat{A}_{\bar{B}}(y_B + t\lambda) \leq \hat{b}_{\bar{B}} \end{array} \right. \quad (35)$$

$$\left\{ \begin{array}{l} y_{B_2} + t\lambda \geq 0 \end{array} \right. \quad (36)$$

ou seja tal que alguma restrição $-\hat{A}_i y_B \leq \hat{b}_i$, $i \in \bar{B}$ ou $y_i \geq 0$, $i \in B_2$ se torne ativa (a primeira delas). No primeiro caso a restrição $i \in \bar{B}$ entra na base, saindo desta alguma restrição não ativa; esta operação não é necessária no segundo caso. Em seguida atualizamos B_1 e B_2 e resolvemos novamente o problema de procura de direção usável.

Para o nosso exemplo, dos dados do problema e (22), (23), (25), (26) e (27) podemos calcular:

$$\hat{A}_B = \begin{bmatrix} 1 & -2 \\ -1/3 & -1/3 \\ -1 & 1 \end{bmatrix}, \quad \hat{b}_B = \begin{bmatrix} 2 \\ 5/3 \\ 5/3 \end{bmatrix}, \quad y_B^* = \begin{bmatrix} 0 \\ 2/3 \end{bmatrix}, \quad \hat{c} = \begin{bmatrix} -2/3 & 5/3 \end{bmatrix}$$

$$\text{e } z_0 = -10/9$$

e de (29), (30) e (31) o problema pode ser escrito como

$$\left\{ \begin{array}{l} \text{MIN} \quad \left[\begin{array}{cc} 2/3 & -5/3 \end{array} \right] y_B - 10/9 \\ \text{s.a.} \quad \left[\begin{array}{cc} -1 & 2 \\ 1/3 & 1/3 \\ 1 & -1 \end{array} \right] y_B \leq \left[\begin{array}{c} 2 \\ 5/3 \\ 5/3 \end{array} \right] \\ y_B \geq 0 \end{array} \right. \quad (37)$$

do qual conhecemos uma solução factível $y_B^{*T} = (0 \ 2/3)$.

De (32), (33) e (34) podemos escrever o problema de procura de direcção usável a partir de y_B^*

$$\left\{ \begin{array}{l} \text{Encontrar } \lambda / \left[\begin{array}{cc} 2/3 & -5/3 \end{array} \right] \left[\begin{array}{c} \lambda_1 \\ \lambda_2 \end{array} \right] < 0 \\ \text{s.a.} \quad \lambda_1 \geq 0 \\ \lambda_2 \text{ irrestrito} \end{array} \right. \quad (38)$$

III. TRANSFORMAÇÃO DE BASE: PROBLEMA (D)

Reescrevamos, para maior comodidade, o problema (D):

$$(D) \quad \left\{ \begin{array}{l} \text{MAX} \quad \phi = pb \\ \text{s.a.} \quad pA = c \\ p \leq 0 \end{array} \right.$$

e seja $S_2 = \{p/pA = c \text{ e } p \leq 0\}$ e vamos supor que conhecemos um ponto p^* tal que $p^* \in S_2$.

Qualquer solução factível de $pA = c$ pode ser escrita como

$$p = p^* + w \quad (39)$$

onde $w(1,m) \in N(A^T)$, espaço nulo de A^T , de dimensão $m-n$. Seja $G(m-n,n)$ uma base de $N(A^T)$, podemos então escrever

$$w = \xi G \quad (40)$$

onde $\xi = (1,m-n)$. Para que p seja factível é preciso que

$$p = p^* + \xi G \leq 0 \quad (41)$$

ou seja

$$\xi G \leq -p^* \quad (42)$$

O valor da função objetivo para um valor de ξ que satisfaça a relação acima é

$$\phi = pb = (p^* + \xi G)b = \xi Gb + p^*b \quad (43)$$

De (42) e (43) o problema (D) pode ser escrito como:

$$\begin{cases} \text{MAX} & \xi Gb + p^*b \\ \text{s.a.} & \xi G \leq -p^* \end{cases} \quad (44)$$

O problema que temos agora é encontrar uma base G para $N(A^T)$. Partilhemos A em duas submatrizes A_B e $A_{\bar{B}}$, $A_B(n,n)$ e B tal que $p_B^* < 0$. Seja $\hat{A}_{\bar{B}} = A_{\bar{B}}A_B^{-1}$, os vetores linhas da matriz

$$G_B = \begin{bmatrix} -\hat{A}_{\bar{B}} & I \end{bmatrix} \quad (45)$$

constituem uma base para $N(A^T)$ pois

$$\begin{bmatrix} -\hat{A}_{\bar{B}} & I \end{bmatrix} \begin{bmatrix} A_B \\ A_{\bar{B}} \end{bmatrix} = -A_{\bar{B}}A_B^{-1}A_B + A_{\bar{B}} = 0$$

Substituindo (45) em (44) e notando que para a base em (45) $p_{\bar{B}} = p_{\bar{B}}^* + \xi_{\bar{B}}$, podemos escrever

$$(D') \begin{cases} \text{MAX} & p_{\bar{B}} \hat{b}_{\bar{B}} + \phi_0 & (46) \\ \text{s.a.} & -p_{\bar{B}} \hat{A}_{\bar{B}} \leq -\hat{c} & (47) \\ & p_{\bar{B}} \leq 0 & (48) \end{cases}$$

com $\hat{b}_{\bar{B}}$, \hat{c} como definidos no item anterior e $\phi_0 = \hat{c} b_{\bar{B}}$. Uma solução factível para (D') é $p_{\bar{B}} = p_{\bar{B}}^*$. Para esta solução verificamos que as restrições (47) são folgadas, em (48) são ativas as soluções com índices em $\bar{B}_1 \subset \bar{B}$ tal que $p_{\bar{B}_1}^* = 0$. O problema de se achar uma direção usável pode então ser escrito como:

$$\begin{cases} \text{Encontrar} & \gamma / \gamma \hat{b}_{\bar{B}} > 0 & (49) \\ \text{s.a.} & \gamma \bar{B}_1 \leq 0 & (50) \\ & \gamma \bar{B}_2 \text{ irrestrito} & (51) \end{cases}$$

$$e \bar{B}_1 \oplus \bar{B}_2 = \bar{B}$$

Encontrado γ , o passo nesta direção é dado pelo maior t tal que ainda tenhamos:

$$\begin{cases} -(p_{\bar{B}} + t \gamma) \hat{A}_{\bar{B}} \leq -\hat{c} & (52) \\ p_{\bar{B}_2} + t \gamma \leq 0 & (53) \end{cases}$$

ou seja tal que alguma restrição $-p_{\bar{B}} \hat{A}_{\bar{B}}^j \leq \hat{c}$ ou $\gamma_j \geq 0$, $i \in \bar{B}_2$ se torne ativa (a primeira delas). No primeiro caso a variável básica cujo índice ocupa a posição j em \bar{B} sai da base, entrando nesta qualquer variável não básica que seja não nula. Notar que o raciocínio acima é válido porque

$$p_{\bar{B}} = \hat{c} - p_{\bar{B}} \hat{A}_{\bar{B}} \quad (54)$$

No segundo caso atualizamos \bar{B}_1 e \bar{B}_2 e resolvemos novamente o problema de procura de direção usável.

Notamos que (P') e (D') são duais entre si, o que é natural desde que (P) e (D) o são, e que $y_B \in R^n$, $p_B \in R^{m-n}$, $p \in N(A^T) \subset R^m$ e $\dim[N(A^T)] = m-n$.

Vejamos a interpretação geométrica deste item. Para isso vamos precisar de (44) com G definido por (45), fazendo a substituição obtemos:

$$(D'') \begin{cases} \text{MAX} & \xi_B \hat{b}_B + p^* b \\ \text{s.a.} & -\xi_B \hat{A}_B \leq -p_B^* \\ & \xi_B \leq -p_B^* \end{cases}$$

A interpretação será feita por meio do exemplo abaixo representado na (fig. 6).

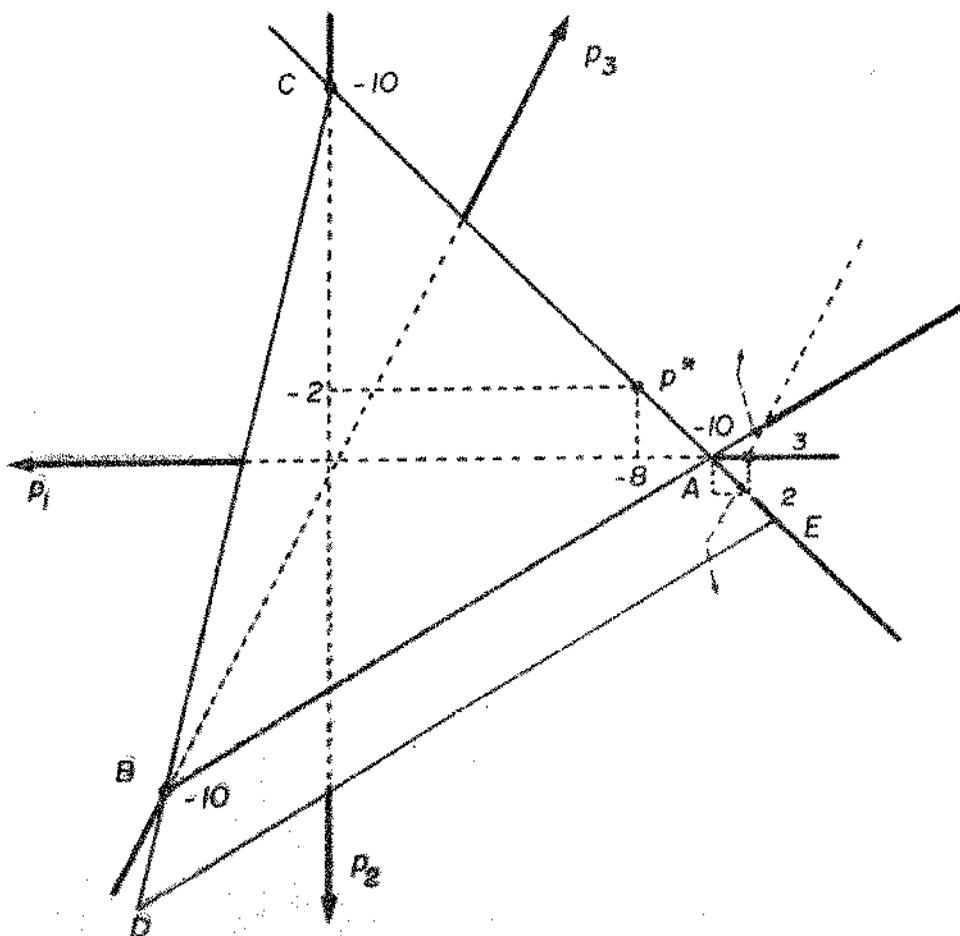


Fig. 6

$$(D) \left\{ \begin{array}{l} \text{MAX } \phi = [p_1 \ p_2 \ p_3] \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\ \text{s.a. } [p_1 \ p_2 \ p_3] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = -10 \\ p \leq 0 \end{array} \right. \quad (55)$$

onde $A^T = [1 \ 1 \ 1]$, $b^T = [1 \ -1 \ 0]$ e $c = [-10]$ e conhecemos um ponto factível $p^* = [-8 \ -2 \ 0]$

Como $p_1^* < 0$, escolhemos $B = \{1\}$ e portanto $\bar{B} = \{2,3\}$. Podemos então calcular

$$A_B = A_1 = 1 \quad \text{e} \quad A_B^{-1} = 1$$

$$\hat{A}_{\bar{B}} = A_{\bar{B}} A_B^{-1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot 1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$G = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\hat{b}_{\bar{B}} = b_{\bar{B}} - A_{\bar{B}} A_B^{-1} b_B = \begin{bmatrix} -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot 1 \cdot 1 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

$$\hat{c} = c A_B^{-1} = -10 \cdot 1 = -10$$

$$\phi_0 = \hat{c} b_B = -10 \cdot 1 = -10$$

$$E = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$p^* b = -6$$

e podemos escrever (D') e (D''):

$$(D') \left\{ \begin{array}{l} \text{MAX } [p_2 \ p_3] \begin{bmatrix} -2 \\ -1 \end{bmatrix} = -10 \\ \text{s.a. } [p_2 \ p_3] \begin{bmatrix} -1 \\ -1 \end{bmatrix} \leq 10 \\ p_{\bar{B}} \leq 0 \end{array} \right. \quad (58)$$

$$(59)$$

$$(60)$$

$$(D'') \begin{cases} \text{MAX} & [\xi_2 \quad \xi_3] \begin{bmatrix} -2 \\ -1 \end{bmatrix} - 6 & (61) \\ \text{s.a.} & [\xi_2 \quad \xi_3] \begin{bmatrix} -1 \\ -1 \end{bmatrix} \leq + 8 & (62) \\ & [\xi_2 \quad \xi_3] \leq [+2 \quad 0] & (63) \end{cases}$$

O problema (D'') é representado na figura 6 no plano que passa pelos pontos A, B e C, variedade linear do espaço nulo de A^T . A base neste plano é dada pelos vetores $[-1 \ 1 \ 0]$ e $[-1 \ 0 \ 1]$, nas direções respectivamente de \overline{CA} e \overline{BA} , a origem do sistema de referência é o ponto A. O conjunto factível S_1 é dado pelos pontos do triângulo CDE. A restrição (62) é representada pelos pontos à esquerda de \overline{BC} e a restrição (63) pelos pontos acima de \overline{DE} . Uma solução factível para este problema é $\xi_B = 0$, para esta solução a única restrição ativa é $\xi_3 = 0$. Neste caso as bases também são escolhidas à cada iteração de forma que as únicas restrições ativas sejam as do tipo $\xi_i \leq p_i^*$, $i \in B$. As bases tem seus vetores nas direções das arestas e a origem é sempre um ponto extremo, como no problema (P').

(D') pode ser interpretado a partir da projeção de S_1 e das curvas de nível da função objetivo em S_1 no espaço das variáveis não básicas. Como a projeção implica na correspondência biunívoca dos pontos de S_1 e de sua projeção, o problema (D'') pode ser redefinido no espaço das variáveis não básicas como (D').

IV. O PROBLEMA DA GERAÇÃO DE DIREÇÃO USÁVEL

Vimos que o problema de geração de direção usável, para o problema (P'), pode ser escrito como:

$$\begin{cases} \text{Encontrar} & \lambda / -\bar{c}\lambda < 0 \\ \text{s.a.} & \lambda_{B_1} \geq 0 \\ & \lambda_{B_2} \text{ irrestrito} \end{cases}$$

se o sistema acima for inconsistente então a solução corrente de

(P') é ótima.

Existem vários métodos de solução para o problema. Apresentamos a seguir alguns.

(a) Métodos de otimização: o problema neste caso pode ser escrito como

$$\begin{cases} \text{MAX} & \bar{c}\lambda & (64) \\ \text{s.a.} & \lambda_{B_1} \geq 0 & (65) \\ & \|\lambda\| \leq 1 & (66) \end{cases}$$

Observamos que se no ótimo deste problema obtivermos $\bar{c}\lambda = 0$ então a corrente solução de (P') é ótima, se $\lambda \neq 0$ então os problemas (P) ou (P') admitem múltiplas soluções.

Vejamos a solução deste problema para as várias normas:

Norma L₁: $\sum |\lambda_j| = 1$, o problema pode ser escrito:

$$\begin{cases} \text{MAX} & \bar{c}\lambda \\ \text{s.a.} & \lambda_{B_1} \geq 0 \\ & \sum |\lambda_j| = 1 \end{cases}$$

Façamos a seguinte transformação de variáveis:

$$\begin{cases} \lambda_j = \lambda_j^+ - \lambda_j^- \\ \lambda_j^+ \geq 0, \lambda_j^- \geq 0, \lambda_j^+ \lambda_j^- = 0 \end{cases}$$

e a restrição de norma pode ser escrita como

$$\sum |\lambda_j| = \sum \lambda_j^+ + \sum \lambda_j^- = 1$$

e o problema pode ser reescrito como

$$\left\{ \begin{array}{l} \text{MAX} \quad \bar{c}_{B_1} \lambda_{B_1}^+ + \bar{c}_{B_2} \lambda_{B_2}^+ - \bar{c}_{B_2} \lambda_{B_2}^- \\ \text{s.a.} \quad | \quad \sigma \quad \quad \quad \sigma \quad \quad \quad \sigma \quad \left[\begin{array}{c} \lambda_{B_1}^+ \\ \lambda_{B_2}^+ \\ \lambda_{B_2}^- \end{array} \right] = 1 \\ \\ \lambda_{B_1}^+ \geq 0 \quad \lambda_{B_2}^+ \geq 0 \quad \lambda_{B_2}^- \geq 0 \end{array} \right.$$

Como só há uma variável básica no problema, a restrição $\lambda_{B_2}^+ \cdot \lambda_{B_2}^- = 0$ é sempre satisfeita e podemos dispensá-la.

Para a variável r^+ na base, o custo relativo, no ótimo, das variáveis não básicas é

$$\begin{aligned} \bar{c}_i - \bar{c}_{r^+} &\leq 0 \quad ; \quad i \in B_1^+, B_2^+ \\ -\bar{c}_i - \bar{c}_{r^+} &\leq 0 \quad ; \quad i \in B_2^- \end{aligned}$$

Para a variável r^- na base, o custo relativo é

$$\begin{aligned} \bar{c}_i + \bar{c}_{r^-} &\leq 0 \quad ; \quad i \in B_1^+, B_2^+ \\ -\bar{c}_i + \bar{c}_{r^-} &\leq 0 \quad ; \quad i \in B_2^- \end{aligned}$$

Portanto, a variável básica ótima é aquela cujo custo é dado por:

$$\bar{c}_r = \text{MAX} \{ [\bar{c}_{B_1}]_i, | [\bar{c}_{B_2}]_i | \}$$

Sejam C_1 e C_2 os conjuntos das variáveis λ_j^+ e λ_j^- que no ótimo têm custo relativo nulo, neste caso o problema admite in finitas soluções. O conjunto solução é:

$$\left\{ \lambda / \sum_{j \in C_1} \lambda_j^+ + \sum_{j \in C_2} \lambda_j^- = 1, \quad \lambda_j = 0 \text{ para } j^+ \notin C_1 \text{ e } j^- \notin C_2 \right\}$$

A solução básica ótima é então

$$\begin{aligned}
 \lambda_r &= 1 && \text{se } \lambda_{r+} \text{ for básica} \\
 \lambda_r &= -1 && \lambda_{r-} \text{ for básica} \\
 \lambda_j &= 0 && \text{para todo } j \neq r
 \end{aligned}
 \tag{67}$$

Reparar que se $\lambda_r = \lambda_{r+}$ então $\tilde{c}_r > 0$ e se $\lambda_r = \lambda_{r-}$ então $\tilde{c}_r < 0$

Norma L_∞ : $-1 \leq \lambda_j \leq 1$, o problema é então

$$\begin{cases}
 \text{MAX } \tilde{c}\lambda \\
 \text{s.a. } 0 \leq \lambda_{B_1} \leq 1 \\
 \quad \quad -1 \leq \lambda_{B_2} \leq 1
 \end{cases}$$

A solução deste problema é:

$$i \in B_1 \begin{cases}
 \lambda_i = 0 & \text{se } \tilde{c}_i < 0 \\
 \lambda_i \in [0, 1] & \text{se } \tilde{c}_i = 0 \\
 \lambda_i = 1 & \text{se } \tilde{c}_i > 0
 \end{cases}
 \tag{68}$$

$$i \in B_2 \begin{cases}
 \lambda_i = -1 & \text{se } \tilde{c}_i < 0 \\
 \lambda_i \in [-1, 1] & \text{se } \tilde{c}_i = 0 \\
 \lambda_i = 1 & \text{se } \tilde{c}_i > 0
 \end{cases}$$

Norma L_2 : $\lambda^T \lambda = 1$, o problema é

$$\begin{cases}
 \text{MAX } \tilde{c}\lambda \\
 \text{s.a. } F\lambda \leq 0 \\
 \quad \quad \lambda^T \lambda \leq 1
 \end{cases}$$

onde $F = [-E \text{ (cardinal } B_1, \text{ cardinal } B_1), 0 \text{ (cardinal } B_1, \text{ cardinal } B_2)]$

A solução deste problema é obtida resolvendo-se o sistema (3), (4), (5) e (6) fazendo-se $N = FF^T$:

$$\left\{ \begin{array}{l} -FF^T u + v = Fc^T \\ u^T v = 0 \\ v \geq 0 \\ u \geq 0 \end{array} \right.$$

como $FF^T = E$, então

$$\left\{ \begin{array}{l} u - v = -\bar{c}_{B_1} \\ u^T v \geq 0 \\ v \geq 0 \\ u \geq 0 \end{array} \right.$$

A solução deste sistema é dada por

$$\left\{ \begin{array}{ll} u_i = -\bar{c}_i & \text{se } \bar{c}_i < 0 \\ u_i = 0 & \text{se } \bar{c}_i \geq 0 \end{array} \quad \begin{array}{l} i \in B_1 \\ i \in B_1 \end{array} \right. \quad (69)$$

a solução λ é obtida via equação (7):

$$B\lambda = c^T - F^T u$$

logo

$$\left\{ \begin{array}{ll} B\lambda_i = \bar{c}_i - \bar{c}_i = 0 & \text{se } \bar{c}_i < 0, i \in B_1 \\ B\lambda_i = \bar{c}_i & \text{em caso contrário} \end{array} \right. \quad (70)$$

(b) Métodos factíveis, neste caso não procuramos a melhor direção segundo algum critério, mas procuramos alguma solução para o sistema:

$$\left\{ \begin{array}{l} \bar{c}_\lambda > 0 \\ \lambda_{B_1} \geq 0 \end{array} \right.$$

A direção obtida depende do método de resolução.

Método Simplex: escolho $r \in B$ tal que $\hat{c}_r \neq 0$

$$\text{Se } r \in B_1 \text{ faço } \lambda_r = \begin{cases} 1 & \text{se } \hat{c}_r > 0 \\ 0 & \text{se } \hat{c}_r < 0, \text{ escolho novo } r \end{cases}$$

$$\text{Se } r \in B_2 \text{ faço } \lambda_r = \begin{cases} -1 & \text{se } \hat{c}_r < 0 \\ 1 & \text{se } \hat{c}_r > 0 \end{cases}$$

$$\lambda_j = 0, \text{ todo } j \neq r$$

Projeção de ROSEN: projeto \hat{c} em $\lambda_{B_1} = 0$, ou seja em $F\lambda = 0$. Temos então

$$\lambda = (I - F^T(F F^T)^{-1} F) \hat{c}^T$$

ou

$$\lambda = \begin{bmatrix} \hat{c}_{B_1}^T \\ \hat{c}_{B_2}^T \end{bmatrix} - \begin{bmatrix} -I \\ 0 \end{bmatrix} (-\hat{c}_{B_1}^T) = \begin{bmatrix} 0 \\ \hat{c}_{B_2}^T \end{bmatrix} \quad (71)$$

Se $B_1 = B$ ou se $\hat{c}_{B_2}^T = 0$ então $\lambda = 0$. O multiplicador de Kuhn-Tucker para este problema é dado por $u = (E E^T)^{-1} E \hat{c} = -\hat{c}_{B_1}^T$. Se $\lambda = 0$ e $-\hat{c}_{B_1}^T \geq 0$, não temos direção usável e a corrente solução de (P') é ótima. Se existe $-\hat{c}_i < 0$, $i \in B$, então retiramos a linha i de F , recalculamos u e λ , a direção encontrada é não nula e usável, dada por:

$$\lambda = \begin{bmatrix} 0 \\ \vdots \\ \hat{c}_i \\ \vdots \\ 0 \end{bmatrix} \quad (72)$$

Direção usável qualquer: escolhemos qualquer direção λ tal que

$$\begin{cases} \hat{c}_1 \lambda_j \geq 0 \\ \hat{c}_\lambda > 0 \\ \lambda_{B_1} \geq 0 \end{cases} \quad (73)$$

Para finalizar este item, observamos que as projeções obtidas pelos métodos de ROSEN ou de otimização com norma L_2 na base $(-A_B^{-1})$ são oblíquas em relação à base E.

V. O MÉTODO SIMPLEX COMO UM MÉTODO DE PROJEÇÃO

Observamos inicialmente que o primeiro passo dos métodos apresentados neste capítulo, a transformação de base, não é necessária, tendo sido apresentada apenas como recurso de exposição. Dado um ponto factível $x^*(n,1)$ precisamos tão somente saber quais restrições são ativas neste ponto.

Suponhamos que n' restrições são ativas, $n' \leq n$ (hipótese de não degenerescência). Estas restrições contribuem com seus gradientes na formação da matriz básica $A_B(n,n)$. Se $n' < n$, podemos arbitrariamente escolher os gradientes de $n-n'$ restrições não ativas para completar A_B . Os índices das restrições ativas definem B_1 e das não ativas na base definem B_2 . Estes dados já são suficientes para que possamos exprimir o problema de busca de direção factível em termos da base $(-A_B^{-1})$:

$$\left\{ \begin{array}{l} \text{Encontrar } \lambda / -cA_B^{-1} < 0 \\ \text{s.a.} \quad \lambda_{B_1} \geq 0 \\ \lambda_{B_2} \text{ irrestrito} \end{array} \right.$$

Encontrada uma direção $\bar{\lambda}$, a sua expressão em termos da base inicial E é

$$\bar{s} = -(A_B)^{-1} \bar{\lambda}$$

Podemos agora situar os métodos de otimização e de projeção de ROSEN deste capítulo no contexto dos métodos do capítulo I. Fazendo $M = A_B$, os métodos de otimização do capítulo I podem ser escritos como

$$\begin{cases} \text{MAX} & cs \\ \text{s.a.} & AB_1s \leq 0 \\ & ||A_Bs|| \leq 1 \end{cases}$$

Temos que $\lambda = -A_Bs$ e $s = -A_B^{-1}\lambda$ que substituídos no problema acima nos permite escrever:

$$\begin{cases} \text{MAX} & \tilde{c}\lambda \\ \text{s.a.} & \lambda_{B_1} \geq 0 \\ & ||\lambda|| \leq 1 \end{cases}$$

Para a projeção de ROSEN temos:

$$\begin{cases} \text{MIN} & |-A_B^{-1}c^T - A_Bs| \\ \text{s.a.} & AB_1s = 0 \end{cases}$$

ou seja

$$\begin{cases} \text{MIN} & |-\tilde{c}^T - \lambda| \\ \text{s.a.} & \lambda_{B_1} = 0 \end{cases}$$

Vejamos agora a relação dos métodos do capítulo I com o método simplex. Vamos supor que o ponto factível inicial x^* é um ponto extremo de S_1 , ou seja, $B_1 = B$. Para o método de otimização com norma L_1 , a direção usável encontrada é

$$\begin{aligned} \lambda_r &= 1 & \text{e} & \tilde{c}_r = \text{MAX}_{i \in B} \tilde{c}_i \\ \lambda_j &= 0 & & j \neq r \end{aligned}$$

que é a direção de um dos eixos de ordenadas, aresta de S_1 , esta direção é aquela que apresenta maior custo relativo. Dado que a solução inicial é um vértice, todos os pontos seguintes serão pontos extremos. Supondo não degenerescência, a cada iteração uma restrição básica passa a ser não ativa e outra não básica

se ativa. A matriz básica A_B pode ser atualizada pelos métodos utilizados em PL e finalmente o valor da variável a cada iteração é dado por $x = A_B^{-1}b_B$. Este é o método simplex aplicado em problemas com restrições de desigualdade, onde a escolha da restrição a sair da base é baseado no critério do maior custo relativo.

Para o método de projeção de ROSEN, supondo também $B_1 = B$, vale a mesma análise acima, com a diferença que a aresta pela qual evoluímos para a próxima solução não é necessariamente a de maior custo relativo.

Observamos ainda que no método de otimização com norma L_2 a direção encontrada é a mesma dada pelo método de gradiente reduzido.

Os métodos simplex e gradiente reduzido podem então ser vistos tanto do ponto de vista da interpretação geométrica como formal como métodos de projeção de gradiente como definidos no capítulo I para $M = A_B$. O método simplex é equivalente à projeção de ROSEN supondo uma solução inicial que seja ponto extremo. E o método de gradiente reduzido é equivalente à projeção de LEMKE. O método simplex, onde a escolha da restrição a sair da base é feita em termos do critério do maior custo relativo, e o método de gradiente reduzido são equivalentes ao problema de norma mínima (Teorema 1 - Capítulo I).

Os resultados dos itens IV e V podem ser estendidos sem dificuldades ao problema (D) (com restrições de igualdade).

MÉTODO SIMPLEX E GRADIENTE REDUZIDO
COM RESTRIÇÕES DE DESIGUALDADE

I. INTRODUÇÃO

Os métodos simplex e gradiente reduzido para PL com restrições de desigualdade surgem como decorrência direta do exposto no capítulo II. Apresentamos então no item II o método simplex para PL com as restrições mencionadas acima. No item III expomos o quadro simplex onde os pivoteamentos, diferentemente da apresentação clássica, são feitos por colunas e a seguir uma breve comparação entre o método simplex exposto no item anterior e o clássico. Apresentamos no item IV o método do gradiente reduzido para PL que, como vimos no capítulo, é uma generalização da programação linear: o ponto inicial factível não precisa ser um vértice do politopo definido pelo conjunto de restrições ou, de outra forma, uma solução básica, e as direções usáveis escolhidas não são necessariamente arestas do politopo.

Acrescentamos ainda que a forma de apresentação do método simplex deste capítulo pode ser de utilidade didática na medida em que é mais próxima das interpretações geométricas geralmente apresentadas num curso de programação linear.

II. MÉTODO SIMPLEX

Seja o problema (P) e x^* um vértice de S_1 (definido em 1.I), podemos partir A de forma que:

$$A_B x^* = b_B$$

$$A_{\bar{B}} x^* \leq b_{\bar{B}}$$

onde $A_B(n,n)$ e $A_{\bar{B}}(m-n,n)$ e, sem perda de generalidade, suporemos

$$B = \{1, 2, \dots, n\}$$

A solução básica $x^* = (A_B)^{-1}b_B = \hat{b}_B$, corresponde a uma solução ótima de (P) se, usando os resultados de Kuhn-Tucker, tivermos $p = \begin{bmatrix} p_B & | & p_{\bar{B}} \end{bmatrix}$ tal que:

$$p_B A_B = c \quad (74)$$

$$p_B \leq 0 \quad (75)$$

$$p_{\bar{B}} = 0 \quad (76)$$

ou seja $p_B = cA_B^{-1} \leq 0$. Em caso contrário devemos procurar outro vértice e portanto outra base factível de (P).

A solução factível que corresponde à base B é tal que

$$A_{\bar{B}}(A_B^{-1}b_B) < b_{\bar{B}}$$

e a folga nestas restrições, denotadas por $\delta_{\bar{B}}$ é dada por

$$\delta_{\bar{B}} = b_{\bar{B}} - A_{\bar{B}}A_B^{-1}b_B > 0 \quad (77)$$

Devemos alterar x^* em uma direção v tal que, para $t > 0$ e arbitrariamente pequeno, tenhamos

$$A_B(x^* + tv) \leq b_B$$

$$A_{\bar{B}}(x^* + tv) \leq b_{\bar{B}}$$

Do ponto de vista das restrições não básicas, supondo não degenerescência, qualquer direção é factível. Vamos fazer $v = -A_B^{-1}\lambda$. Do capítulo II sabemos que esta direção é factível para $t \geq 0$ se $\lambda \geq 0$. Temos que

$$A_B(x^* - tA_B^{-1}\lambda) = A_B x^* - t\lambda \leq b_B \quad (78)$$

Como queremos um procedimento simplex, façamos:

$$\lambda^T = (0 \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0 \ 0) \quad (79)$$

↑
— posição r

Para este valor de λ , em (78) só uma restrição deixa de ser ativa, pois

$$A_{B-\{r\}} x^* = b_{B-\{r\}}$$

e

$$A_r x^* - t \leq b_r$$

Devemos passar de uma solução factível a outra de forma a melhorar z . Temos que

$$\Delta z = c(x^* + tv) - cx^* = -t c A_B^{-1} \lambda$$

ou seja

$$\frac{\Delta z}{t} = -\hat{c}_r \quad (80)$$

Para a direção especificada em (79) teremos

$$\frac{\Delta z}{t} = -\hat{c}^r \quad (81)$$

e o índice r deve ser tal que

$$-\hat{c}^r \leq 0$$

se este índice não for encontrado, ou seja, se

$$-\hat{c}^j \geq 0 \quad \forall j \in \{1, 2, \dots, n\}$$

então x^* é ótimo.

Devemos dar um passo t de forma que nenhuma restrição com índice em \bar{B} seja violada. Devemos calcular o maior passo t tal que ainda tenhamos:

$$A_{\bar{B}}(x^* - t[A_{\bar{B}}^{-1}]^T) \leq b_{\bar{B}} \quad (82)$$

ou ainda

$$-t\hat{A}_{\bar{B}}^T \leq \hat{b}_{\bar{B}} \quad (83)$$

Desta relação podemos calcular o passo por

$$t_{\text{MAX}} = \min_{i / [\hat{A}_{\bar{B}}]_i^T < 0} \frac{\hat{b}_i}{[-\hat{A}_{\bar{B}}]_i^T} = \frac{\hat{b}_s}{[-\hat{A}_{\bar{B}}]_s^T} \quad (84)$$

Para este valor do passo a restrição s se torna ativa e entra na base, saindo desta a restrição r .

O método primal para (P) pode ser descrito, dado uma base B factível, pelos seguintes passos.

(1) Calcule: $p_B/p_B = c(A_B)^{-1} = \hat{c}$, $p_{\bar{B}} = 0$. Se $\hat{c} \leq 0$, a corrente solução é ótima, pare. Em caso contrário escolha $\hat{c}^r > 0$. A restrição r deve sair da base.

(2) Calcule: $\hat{A}_{\bar{B}} = A_{\bar{B}}A_B^{-1}$

$$\hat{b}_{\bar{B}} = b_{\bar{B}} - \hat{A}_{\bar{B}}b_B$$

(3) Calcule

$$t_{\text{MAX}} = \min_{i / [\hat{A}_{\bar{B}}]_i^T < 0} \frac{\hat{b}_i}{[-\hat{A}_{\bar{B}}]_i^T} = \frac{\hat{b}_s}{[-\hat{A}_{\bar{B}}]_s^T}$$

A linha s entra na base

(4) Atualize B , \bar{B} e A_B^{-1} e vá a (1)

Com a finalidade de tornar este item mais completo apresentamos os passos de resolução do problema (D). Seja A_B uma base factível para (D), então:

(1) Calcule $p_B/p_B = c(A_B)^{-1} = \hat{c}$, $p_{\bar{B}} = 0$

(2) Calcule: $\hat{A}_{\bar{B}} = A_{\bar{B}}A_B^{-1}$

$$\hat{b}_{\bar{B}} = b_{\bar{B}} - \hat{A}_{\bar{B}}b_B$$

Se $\hat{b}_{\bar{B}} \geq 0$ então a corrente solução é ótima, pare.
 Em caso contrário escolha $\hat{b}_s < 0, \in s \in B$. Entra na base a base a variável s

(3) Calcule:

$$t_{MAX} = \text{MIN}_{i/[A_{\bar{B}}]_s^i < 0} \frac{\hat{c}^i}{[A_{\bar{B}}]_s^i} = \frac{\hat{c}^r}{A_s^r}$$

A variável r sai da base

(5) Atualize B, \bar{B} e A_B^{-1} e vá a (1).

III. QUADRO SIMPLEX

Definimos o Quadro Simplex, para o problema linear com restrições de desigualdade (problema (P)), como a tabela constituída pelos elementos dos vetores c e b , da matriz A , e um escalar z^0 inicialmente nulo, dispostos da seguinte forma:

I					II
c_1	c_2	c_n		0
A_1^1	A_1^2	A_1^n		b_1
A_2^1	A_2^2	A_2^n		b_2
\vdots	\vdots		\vdots		
A_m^1	A_m^2	A_m^n		b_m

(85)

Vamos supor que conhecemos uma matriz básica A_B e que ela corresponde a uma solução básica factível inicial dada por $A_B^{-1}b_B = x^*$. Suponhamos que, sem perda de generalidade, que $B = \{1, 2, \dots, n\}$. Vamos pós-multiplicar a parte I do quadro por A_B^{-1}

e obtemos:

I				II	
c_1	c_2	c_n		0	A
1	1	1		b_1	B
.	.	.		b_2	
.	.	.		b_n	
\hat{A}_{n+1}^1	\hat{A}_{n+1}^2 \hat{A}_{n+1}^n		b_{n+1}	C
:	:	:		:	
\hat{A}_m^1	\hat{A}_m^2 \hat{A}_m^n		b_m	

(86)

Vamos subtrair da parte II do quadro a parte I pós-multiplicada por b_B , obtemos:

I				II	
\hat{c}_1	\hat{c}_2	\hat{c}_n		$z^0 = -\hat{c}b_B$	A
1	1	1			B
.	.	.			
.	.	.			
\hat{A}_{n+1}^1	\hat{A}_{n+1}^2 \hat{A}_{n+1}^n		\hat{b}_{n+1}	C
:	:	:		:	
\hat{A}_m^1	\hat{A}_m^2 \hat{A}_m^n		\hat{b}_m	

(87)

As informações que temos para proceder a primeira iteração do método simplex estão dispostas nas partes A e C do quadro (87). O problema que temos é saber se é possível atualizar o quadro para a próxima iteração por um procedimento de pivoteamento, como no caso do simplex com restrições de igualdade.

Seja uma base A_B e a matriz não básica correspondente

$A_{\bar{B}}$:

$$A(x^* - t[A_B^{-1}]^r) \leq b$$

e para $t = t_{MAX}$ as novas folgas nas restrições serão dadas por

$$\hat{b} = b - Ax^* + t_{MAX}\hat{A}^r$$

mas

$$t_{MAX} = \frac{\hat{b}_s}{[-\hat{A}_B]^r_s} = \frac{\hat{b}_{n+s}}{\hat{A}_{n+s}^r}$$

portanto

$$\hat{\hat{b}} = \hat{b} - \hat{A}^r \frac{\hat{b}_{n+s}}{\hat{A}_{n+s}^r} = \hat{b} - \hat{A}^r \hat{b}_{n+s} \quad (93)$$

De (81) vemos que a variação na função objetiva ao caminhar na direção especificada acima é dada por

$$\Delta z = -t\hat{c}^r = \hat{c}^r \frac{b_s}{\hat{A}_{n+s}^r}$$

portanto

$$-z' = z^0 - \hat{c}^r \frac{b_s}{\hat{A}_{n+s}^r} \quad (94)$$

As relações (93) e (94) são equivalentes a estendermos as operações de pivoteamento da parte I a parte II do quadro simplex.

Podemos então concluir que podemos atualizar o quadro simplex por meio de pivoteamentos a semelhança do método simplex para problemas com restrições de igualdade, mas estas operações serão feitas por colunas e não linhas como no caso clássico. É fácil ver que outras formas de atualização dos dados, necessárias a cada iteração, como a forma produto da inversa ou a decomposição LU podem ser usados para o problema (P) sem dificuldades.

Este quadro simplex é o mesmo usado para o problema (D). Reparar que os dois métodos necessitam os mesmos dados a cada iteração. Reparar também que o método primal é idêntico, em

termos das operações a efetuar ao método dual simplex clássico e o método dual é semelhante ao método primal clássico.

Cabe aqui uma discussão se haveria algum ganho em trabalharmos com os problemas de programação linear na forma de desigualdades. Vamos supor que para (P) um subconjunto de restrições seja $x \geq 0$. As restrições podem ser escritas como

$$\begin{aligned} A_1 x &\leq b_1 \\ -E_1 x &\leq 0 \end{aligned} \tag{95}$$

e $A_1(m_1, n_1)$, $E_1(n, n)$ e $m = m_1 + n$. Acrescentando as variáveis de folga teremos:

$$\begin{aligned} \begin{bmatrix} A_1 & E_2 \end{bmatrix} \begin{bmatrix} x \\ x_f \end{bmatrix} &= b_1 \\ x, x_f &\geq 0 \end{aligned} \tag{96}$$

com $E_2(m_1, m_1)$. Uma matriz básica do problema com restrições na forma (95) seria de ordem n e na forma (96) de ordem m_1 . Para $n = m_1$, em termos de memória e tempo de computação os algoritmos devem se comportar de modo semelhante. Se $m_1 < n$, a apresentação clássica do método simplex seria a melhor, para $m_1 > n$ seria melhor trabalhar com a forma apresentada neste capítulo.

IV. MÉTODO DO GRADIENTE REDUZIDO

Para o problema (P) seja x^* factível e A_B uma base qualquer factível ou infactível tal que todas restrições que são ativas em x^* têm índices em B . Sem perda de generalidade, supomos $B = (1, 2, \dots, n)$. Podemos partir A_B de forma que:

$$A_{B_1} x^* = b_{B_1} \tag{97}$$

$$A_{B_2} x^* < b_{B_2} \tag{98}$$

$$A_{\bar{B}} x^* < b_{\bar{B}} \tag{99}$$

onde $A_{B_1}(n_1, n)$, $A_{B_2}(n_2, n)$, $A_{\bar{B}}(m-n, n)$, $B_1 \oplus B_2 = B$ e $n = n_1 + n_2$.

Vamos anotar as folgas nas restrições de base pelo vetor $y_B^* = [y_{B_1}^* \ y_{B_2}^*]$, $y_{B_1}^* = 0$ e $y_{B_2}^* > 0$. Nas restrições não básicas as folgas, denotadas por $y_{\bar{B}}^* > 0$, podem ser obtidas como função de y_B^* . Podemos escrever

$$x^* = (A_B)^{-1} (b_B - y_B^*) \quad (100)$$

e

$$y_{\bar{B}}^* = b_{\bar{B}} - A_{\bar{B}} x^* \quad (101)$$

portanto

$$y_{\bar{B}}^* = \bar{b}_{\bar{B}} + \hat{A}_{\bar{B}} y_B^* \quad (102)$$

Então qualquer que seja x factível e B definido como em (97), (98) e (99), deveremos ter

$$y_{B_1} = 0 \quad (103)$$

$$y_{B_2} > 0 \quad (104)$$

$$\hat{A}_{\bar{B}} y_{\bar{B}} + \bar{b}_{\bar{B}} > 0 \quad (105)$$

Devemos alterar x^* em uma direção $v = -A_B^{-1} \lambda_B$. Para esta direção deveremos ter:

$$A_B (x^* - t A_B^{-1} \lambda_B) = A_B x^* - t \lambda_B \leq b_B \quad (106)$$

Com $t > 0$, a direção v satisfaz a relação acima se

$$\lambda_{B_1} \geq 0 \quad (107)$$

$$\lambda_{B_2} \text{ irrestrito} \quad (108)$$

De (80) temos:

$$\frac{\Delta z}{t} = -\hat{c}\lambda_B$$

As direções em que a função decresce são tais que

$$\hat{c}\lambda_B > 0 \quad (109)$$

Se não existir λ_B tal que (109) seja satisfeita então x^* é um ótimo de (P).

No capítulo II são apresentados vários métodos de se calcular λ_B . Supondo a existência de λ_B tal que $\hat{c}\lambda_B > 0$, a etapa a seguir é calcular o maior passo t tal que uma das folgas y_{B_2} ou $y_{\bar{B}}$ se anule em primeiro lugar. De (106) deduzimos que:

$$y_{B_2}^* + t\lambda_{B_2} \geq 0 \quad (110)$$

e calculamos

$$t_{r_1} = \min_{\lambda_i < 0, i \in B_2} \frac{y_i^*}{-\lambda_i} = \frac{y_{r_1}^*}{-\lambda_{r_1}} \quad (111)$$

Deveremos ter

$$b_{\bar{B}} - A_{\bar{B}}(x^* - tA_B^{-1}\lambda_B) = y_{\bar{B}}^* + t\tilde{A}_{\bar{B}}\lambda_B \geq 0$$

e calculamos

$$t_{r_2} = \min_{i \in \bar{B} / \tilde{A}_i \lambda_B < 0} \frac{y_i^*}{-\tilde{A}_i \lambda_B} = \frac{y_{r_2}^*}{-\tilde{A}_{r_2} \lambda_B} \quad (112)$$

O maior passo é então

$$t_{MAX} = \min \{t_{r_1}, t_{r_2}\}$$

e para este valor de t as novas folgas serão

$$y_{B_1} = t_{MAX} \lambda_B$$

$$y_{B_2} = y_{B_2}^* + t_{MAX} \lambda_{B_2}$$

$$y_B = y_B^* + t_{MAX} \bar{A}_B \lambda_B$$

Se $t_{MAX} = t_{r_1}$, conservo o conjunto de base B sem mudanças mas atualizo B_1 e B_2 e recalculo λ_B .

Se $t_{MAX} = t_{r_2}$, a restrição r_2 entra na base e retiro desta uma cuja folga não seja nula, um critério poderia ser, escolher a maior. Seja $s \in B$ tal que

$$y_s = \text{MAX}_{i \in B} y_i$$

sai então da base a restrição s.

O método pode ser descrito pelos seguintes passos:

- (1) Dado x^* factível, escolha uma base B tal que as restrições ativas tenham seus índices em B. Construa B_1 e B_2 . Calcule A_B^{-1} , $y_{B_2}^*$, y_B^* e faça $y_{B_1}^* = 0$.
- (2) Calcule $\bar{c} = c A_B^{-1}$
- (3) Determine $\lambda_B / \bar{c} \lambda_B > 0$ e $\lambda_{B_1} \geq 0$. Se tal λ_B não existe, a corrente solução de (P) é ótima, pare.
- (4) Calcule $t_{MAX} = \text{MIN}\{t_{r_1}, t_{r_2}\}$ onde

$$t_{r_1} = \text{MIN}_{\lambda_i < 0, i \in B_2} \frac{y_i^*}{-\lambda_i}$$

$$t_{r_2} = \text{MIN}_{\bar{A}_i \lambda_B < 0, i \in B} \frac{y_i^*}{-\bar{A}_i \lambda_B}$$

e calcule as novas folgas

$$y_{B_1}^* = t_{MAX} \lambda_B$$

$$y_{B_2}^* = y_{B_2}^* + t_{MAX} \lambda_{B_2}$$

$$y_B^* = y_B^* + t_{MAX} \lambda_B$$

(5) Se $t_{MAX} = tr_1$, atualize B_1 e B_2 e vá a (3)

Se $t_{MAX} = tr_2$, calcule

$$y_s^* = \max_{i \in B} y_i^*$$

(6) A restrição r_2 entra na base e sai a restrição s .
Atualize B , B_1 , B_2 e A_B^{-1} . Vá a (2).

Podemos utilizar o seguinte tableau para o método:

$$\left[\begin{array}{c|c} -A_B^{-1} & b_B - A_B^{-1} b_B + A_B^{-1} \lambda_B^* \\ \hline c_A^{-1} & z = c_A^{-1} b_B - c_A^{-1} y_B^* \end{array} \right] = \left[\begin{array}{c|c} -\hat{A}_B & \hat{b}_B + \hat{A}_B y_B^* = y_B^* \\ \hline -\hat{c} & z = \hat{c}(b - y_B^*) \end{array} \right]$$

P A R T E I I

PROBLEMAS DE PROGRAMAÇÃO LINEAR COM ESTRUTURA
BLOCO ANGULAR

INTRODUÇÃO

Nesta segunda parte deste trabalho, nos interessamos pela resolução de um problema de programação linear com estrutura bloco-angular. Vários métodos já foram propostos, entre os quais podemos citar o princípio de decomposição de DANTZIG-WOLFE [6], o método das variáveis limitadas generalizado (GUB) [6] [33] e o método de partição de ROSEN [6] [14].

O método GUB, estendido ao caso dos programas lineares com estrutura bloco angular (GGUB) [33], é um método primal baseado no método simplex. Este é especializado para se tomar em conta a estrutura particular da matriz de restrição. Nos capítulos VI e VII apresentamos este método de uma forma mais condensada, e nos parece, mais possante que aquela encontrada em [6], pp. 324-356. Isto se deve, sobretudo à notação adotada e às noções de "pivoteamento interno" e "pivoteamento externo" desenvolvidas no capítulo V.

Propomos a seguir, no capítulo VII, um método dual do GGUB, baseado no método dual simplex. Especializamos este algoritmo para ter em conta a estrutura bloco angular do problema e propomos um meio de obter facilmente uma base dual admissível inicial. Este método pode ser vantajoso no caso onde não conhecemos uma solução de base admissível do programa linear a resolver. Com efeito, se utilizamos o método primal, neste caso é necessário, a priori, utilizar a fase 1, e sabemos que esta pode eventualmente ser mais longa que a fase 2. No método dual que propomos, a obtenção de uma base dual admissível de partida é uma operação relativamente simples.

Mostramos que o método de partição de ROSEN pode ser considerada como uma variante do método dual que propomos. Nesta variante se exige a verificação de uma propriedade suplementar à cada etapa da maximização da função dual. É a primeira vez, pelo que sabemos, que uma ligação é estabelecida entre o método GGUB e o método de partição de ROSEN.

Finalmente, no capítulo VIII, fazemos uma interpretação geométrica do método dual proposto, onde utilizamos em particular a noção de projeção. Ela permite ilustrar o caso onde, por projeção em um espaço restrito, não é mais possível otimizar

a função dual. É necessário então um passo de otimização em um sub-sistema para sair do impasse. O método de ROSEN permite justamente eliminar a possibilidade de ocorrência deste caso mas complicando cada iteração do método de forma a satisfazer uma propriedade suplementar.

O PROBLEMA DE PL COM ESTRUTURA BLOCO-ANGULAR

I. O PROBLEMA

Um problema de programação linear pode ser apresentado sob a forma geral abaixo:

$$\begin{aligned} \text{MIN} \quad z &= cx \\ \text{s.a.} \quad Ax &= b \\ x &\geq 0 \end{aligned}$$

No caso particular em que a matriz A possui uma estrutura bloco-angular, o problema acima pode ser escrito como:

$$(P) \left\{ \begin{array}{l} \text{MIN } f(x) = z = [c_1 \quad c_2 \quad \dots \quad c_\ell] x \\ \text{s.a.} \quad \begin{bmatrix} A_{01} & A_{02} & \dots & A_{0\ell} \\ & A_1 & & \\ & & \ddots & \\ & & & A_\ell \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\ell \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_\ell \end{bmatrix} \end{array} \right. \quad (1)$$

$$x \geq 0 \quad (2)$$

que nós denominaremos de problema primal e tal que $A_0(m_0, n) = [A_{01} \ A_{02} \ \dots \ A_{0\ell}]$, $A_{0i}(m_0, n_i)$, $A_i(m_i, n_i)$, $x(n, 1)$, $c_i(1, n_i)$, $b_0(m_0, 1)$, $b_i(m_i, 1)$ com

$$n = \sum_{i=1}^{\ell} n_i \quad \text{e} \quad m = \sum_{i=1}^{\ell} m_i$$

Em consequência, a matriz de restrições A possui $m_0 + m$ linhas.

Hipótese: No estudo deste problema supomos ainda não redundância de restrições, isto é, a matriz A é de "rank" completo.

Como consequência desta hipótese podemos afirmar que cada submatriz A_i é de "rank" completo.

Base do Sistema (1)

Devemos agora caracterizar uma base A^B , onde B é um conjunto ordenado de $m + m_0$ índices que correspondem às colunas linearmente independentes de A . Seja

$$A_i = \begin{bmatrix} A_{0i} \\ \textcircled{0} \\ A_i \\ \textcircled{0} \end{bmatrix} \quad \text{e} \quad A_i^{B_i} = \begin{bmatrix} A_{0i}^{B_i} \\ \textcircled{0} \\ A_i^{B_i} \\ \textcircled{0} \end{bmatrix}$$

onde $A_i^{B_i}$ é uma submatriz constituída por todas as colunas de A_i que pertencem à base A^B . Nós chamaremos de "partição i " a submatriz de A cujas linhas tenham os mesmos índices que as linhas de A_i . Utilizaremos a notação \bar{B}_i para representar o conjunto de índices complementares de B_i em relação ao conjunto de índices de todas as colunas de A_i .

Teorema 1: Toda matriz básica A^B do problema (P) é tal que o "rank" de $A_i^{B_i}$ seja m_i .

Prova: Dada uma submatriz A_i , todas as demais ($A_j, j \neq i$) terão componentes nulas na partição i . Suponhamos que o "rank" de $A_i^{B_i}$ seja inferior à m_i , isto implica que as linhas de A^B na partição i são linearmente dependentes, o que é impossível dado que A^B é uma base.

II. NOTAÇÃO UTILIZADA

O teorema 1 nos permite partir A^B da seguinte forma

$$\left[\begin{array}{c|ccc} A_0^L & A_0^{I_1} & A_0^{I_2} & \dots & A_0^{I_\ell} \\ \hline & A_1^{I_1} & & & \\ & & A_2^{I_2} & & \\ & & & \dots & \\ & & & & A_\ell^{I_\ell} \end{array} \right] = \begin{bmatrix} A_0^L & A_0^I \\ A^L & A^I \end{bmatrix} = \begin{bmatrix} A_0^B \\ A^B \end{bmatrix} \quad (3)$$

com $I = I_1 \oplus I_2 \oplus \dots \oplus I_\ell$ e L subconjuntos ordenados de índices de B e tal que $A_i^{I_i}(m_i, m_i)$ é de "rank" m_i e $B = I \oplus L$. Anotaremos por L_i , conjunto de índices de colunas, à intersecção de L e B_i . Notamos que A^L não tem necessariamente a forma bloco diagonal.

Colunas Chaves e Não Chaves. Chamaremos de colunas chaves àquelas de índices em I e não chaves àquelas de índices em L . Seja B o conjunto de índices de colunas que não pertencem a base, a submatriz de A formada por estas colunas será notada por A^B . Temos ainda $A_0^{I_i}(m_0, m_i)$, $A_0^L(m_0, m_0)$, $A^L(m, m_0)$.

Dada uma base A^B , supondo A partida segundo colunas não chaves, chaves e não básicas, podemos dividir (1) em dois sub-sistemas de restrições: o sistema de restrições globais ao qual podemos acrescentar a função objetivo e o sistema de restrições locais. Cada bloco de restrições locais será chamado sub-sistema. Temos então

$$\left[\begin{array}{c|c|c} C^L & C^I & C^B \\ \hline A_0^L & A_0^I & A_0^B \end{array} \right] \begin{bmatrix} x_L \\ x_I \\ x_B \end{bmatrix} = \begin{bmatrix} z \\ b_0 \end{bmatrix} \quad (4)$$

para as restrições globais e

$$\left[\begin{array}{c|c|c} A^L & A^I & A^B \end{array} \right] \begin{bmatrix} x_L \\ x_I \\ x_B \end{bmatrix} = b \quad (5)$$

para as restrições locais. Pré-multiplicando-se (5) por $(A^I)^{-1}$ obtemos

$$\begin{bmatrix} \hat{A}^L & E & \hat{A}^B \end{bmatrix} \begin{bmatrix} x_L \\ x_I \\ x_{\bar{B}} \end{bmatrix} = b \quad (\text{RLC})$$

onde

$$\begin{aligned} \hat{A}^L &= (A^I)^{-1} A^L \\ E &= (A^I)^{-1} A^I \\ \hat{A}^B &= (A^I)^{-1} A^B \\ \hat{b} &= (A^I)^{-1} b \end{aligned} \quad (6)$$

O sistema (RLC) é a forma canônica do sistema de restrições locais em relação a base A^I . A operação acima de pré-multiplicação por $(A^I)^{-1}$ é equivalente a efetuar m operações de pivoteamento com elementos pivots pertencentes a A^I . A partir de (RLC) podemos obter x_I em função de x_L e $x_{\bar{B}}$:

$$x_I = \hat{b} - \hat{A}_{x_L}^L - \hat{A}_{x_{\bar{B}}}^B \quad (7)$$

Substituindo (7) em (4) obtemos:

$$\left[\begin{array}{c|c|c} c^L - c^L \hat{A}^L & 0 & c^{\bar{B}} - c^I \hat{A}^{\bar{B}} \\ \hline A_0^L - A_0^I \hat{A}^L & \textcircled{\oplus} & A_0^{\bar{B}} - A_0^I \hat{A}^{\bar{B}} \end{array} \right] \begin{bmatrix} x_L \\ x_I \\ x_{\bar{B}} \end{bmatrix} = \begin{bmatrix} z - c^I \hat{b} \\ b_0 - A_0^I \hat{b} \end{bmatrix}$$

Esta última operação é equivalente a estender as operações de pivoteamento de (5) a (4). Quando as operações de pivoteamento que permitem colocar um sistema sob forma canônica são estendidas a outro sistema, chamaremos esta extensão de "Pivoteamento Externo" e utilizamos o símbolo - (til) para representar os elementos deste sistema. Assim, podemos reescrever o sistema de restrições globais sob a forma:

$$\left[\begin{array}{c|c|c} \tilde{c}^L & 0 & \tilde{c}^{\bar{B}} \\ \hline \tilde{A}_0^L & \textcircled{\oplus} & \tilde{A}_0^{\bar{B}} \end{array} \right] \begin{bmatrix} x_L \\ x_I \\ x_{\bar{B}} \end{bmatrix} = \begin{bmatrix} \tilde{z} \\ \tilde{b}_0 \end{bmatrix} \quad (\text{PR})$$

COM

$$\begin{aligned}
 \bar{c}^L &= c^L - c^I (A^I)^{-1} A^L \\
 \bar{c}^B &= c^B - c^I (A^I)^{-1} A^B \\
 \bar{z} &= z - c^I (A^I)^{-1} b \\
 \bar{A}_0^L &= A_0^L - A_0^I (A^I)^{-1} A^L \\
 \bar{A}_0^B &= A_0^B - A_0^I (A^I)^{-1} A^B \\
 \bar{b}_0 &= b_0 - A_0^I (A^I)^{-1} b
 \end{aligned} \tag{8}$$

Teorema 2: \bar{A}_0^L é não singular

Prova: A matriz básica A^B de (1) sofreu transformações de pivoteamento que não modificam o "rank" da matriz. A nova matriz que tem "rank" $m_0 + m$ pode ser escrita

$$\left[\begin{array}{c|c} \bar{A}_0^L & \textcircled{0} \\ \hline \bar{A}^L & E \end{array} \right]$$

Como a matriz E é de "rank" m, \bar{A}_0^L deve ser de "rank" m_0 , sendo portanto não singular.

Base de Trabalho: O sistema (PR) e a matriz \bar{A}_0^L são chamados respectivamente de "Problema Reduzido" e "Base de Trabalho". Por pivoteamento podemos colocar (PR) sob forma canônica em relação a base \bar{A}_0^L , e estendendo estas operações de pivoteamento à função objetivo e ao sistema (RLC) obtemos:

$$\begin{bmatrix} 0 & 0 & \bar{c}^B \\ \textcircled{E}_0 & \textcircled{0} & \bar{A}_0^B \end{bmatrix} \begin{bmatrix} x_L \\ x_I \\ x_B \end{bmatrix} = \begin{bmatrix} \bar{z} \\ \bar{b}_0 \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} \textcircled{0} & E & \bar{A}_0^B \end{bmatrix} \begin{bmatrix} x_L \\ x_I \\ x_B \end{bmatrix} = \bar{b} \tag{10}$$

onde

$$\begin{aligned}
 \hat{c}^{\bar{B}} &= \bar{c}^{\bar{B}} - \bar{c}^L (\bar{A}_0^L)^{-1} \bar{A}_0^{\bar{B}} \\
 \hat{z} &= \bar{z} - \bar{c}^L (\bar{A}_0^L)^{-1} \bar{b}_0 \\
 \hat{\bar{A}}_0^{\bar{B}} &= (\bar{A}_0^L)^{-1} \bar{A}_0^{\bar{B}} \\
 \hat{b}_0 &= (\bar{A}_0^L)^{-1} \bar{b}_0 \\
 \hat{\bar{A}}^{\bar{B}} &= \bar{A}^{\bar{B}} - \bar{A}^L (\bar{A}_0^L)^{-1} \bar{A}_0^{\bar{B}} \\
 \hat{b} &= \bar{b} - \bar{A}^L (\bar{A}_0^L)^{-1} \bar{b}_0
 \end{aligned} \tag{11}$$

III. SOLUÇÃO BÁSICA CORRESPONDENTE À BASE A^B

Façamos $x_B = 0$ em (9) e (10). Podemos escrever:

$$\begin{aligned}
 x_L &= \hat{b}_0 = (\bar{A}_0^L)^{-1} [b_0 - A^I (A^I)^{-1} b] \\
 x_I &= \hat{b} = (A^I)^{-1} b - (A^I)^{-1} A^L (\bar{A}_0^L)^{-1} (b_0 - (A_0^I)^{-1} b)
 \end{aligned}$$

ou na forma matricial, sabendo que

$$x_B = \begin{bmatrix} x_L \\ x_I \end{bmatrix} = (A^B)^{-1} \begin{bmatrix} b_0 \\ b \end{bmatrix}$$

podemos deduzir

$$(A^B)^{-1} = \left[\begin{array}{c|c} (\bar{A}_0^L)^{-1} & -(\bar{A}_0^L)^{-1} A_0^I (A^I)^{-1} \\ \hline -(\bar{A}_0^L)^{-1} A^L (\bar{A}_0^L)^{-1} & (A^I)^{-1} + (\bar{A}_0^L)^{-1} A^L (\bar{A}_0^L)^{-1} A_0^I (A^I)^{-1} \end{array} \right] \tag{12}$$

Os resultados (9) e (10) são equivalentes a colocar o sistema global em forma canônica em relação à base A^B e a calcular os custos relativos. Colocar (1) sob forma canônica em relação à A^B é fazer

$$\hat{A} = (A^B)^{-1} A$$

De (9) e (10) temos

$$\hat{A} = \left[\begin{array}{c|c|c} E_0 & \textcircled{0} & \bar{A}^B \\ \hline \textcircled{0} & E & \bar{A}^B \end{array} \right]$$

Os custos relativos podem ser obtidos a partir do vetor multiplicador p tal que:

$$p A^B = c^B$$

que podemos repartir na forma

$$p_0 A_0^L + p A^L = c^L$$

$$p_0 A_0^I + p A^I = c^I$$

A solução deste sistema é

$$p_0 = \bar{c}^L (A_0^L)^{-1} \quad (13)$$

$$p = (c^I - p_0 A_0^I) (A^I)^{-1} \quad (14)$$

O custo relativo para as variáveis não básicas é dada por

$$\bar{c}^B = c^B - p A^B \quad (15)$$

Substituindo (13) em (14), obtemos

$$\bar{c}^B = c^B - p_0 \bar{A}_0^B$$

é obviamente

$$\frac{a^B}{c} = \frac{\tilde{a}^B}{\tilde{c}}$$

Conclusão: Observamos que o estudo do sistema $A x = b$ que acabamos de fazer é geral, não importando a forma da matriz A . No caso particular do nosso problema (P), o resultado é interessante porque A^I é uma matriz bloco diagonal. A inversa de A^B pode então ser obtida a partir das inversas de A_0^L e de A_i^{Ii} , $i = 1, 2, \dots, \ell$, de dimensões (em geral) bem menores que A^B .

RESOLUÇÃO DO PROBLEMA BLOCO ANGULAR: MÉTODO PRIMALI. INTRODUÇÃO

O método de resolução do problema (P) é muito semelhante ao primal simplex clássico. Como (P) tem estrutura bloco angular, devemos especializar o algoritmo do simplex de forma a explorar esta propriedade. Para isto, a cada iteração, devemos trabalhar com bases ordenadas de maneira a manter a estrutura apresentada em (3).

Vamos então supor que, por intermédio de uma fase 1, tenhamos obtido uma base ordenada A^B para o problema (P). Para obter a inversa desta base é suficiente inverter as matrizes \bar{A}_0^L e A_i^I ($i = 1, 2, \dots, \ell$). O cálculo de $(A^B)^{-1}$ é então feito a partir de (12).

II. MÉTODO PRIMAL SIMPLEX PARA (P)

Os passos do algoritmo simplex, na forma revisada, a cada iteração são:

Passo 1: Cálculo do vetor multiplicador p associado à base A^B .

Precisamos calcular somente a componente p_0 de p , o que é justificado pelo fato de que no passo 2 não é necessário p para calcularmos os custos relativos. Temos então:

$$p_0 = \bar{c}^L (\bar{A}_0^L)^{-1} \quad (13)$$

Passo 2: Escolha da coluna a entrar na base

Devemos calcular o custo relativo das variáveis não básicas:

$$\theta_r = \frac{\hat{b}_{m_0+r'}}{\hat{a}_{m_0+r'}^s} = \frac{[\hat{b}]_{r'}}{[\hat{A}]_{r'}^s}$$

Notemos que se $\hat{a}^s \leq 0$ então (P) não tem solução finita. Em caso contrário, a coluna a sair da base, A^p , é aquela cujo índice ocupa a posição r do conjunto ordenado B . O elemento pivot é $\hat{a}_{r'}^s$, em torno do qual construímos a matriz de pivoteamento, necessária a atualização da base.

Passo 4: Atualização da base

É neste passo que devemos introduzir as propriedades particulares do problema que tratamos. Os vetores e matrizes utilizados até aqui são elementos de (PR) e de (RLC) que podem ser obtidos facilmente a partir de $(\hat{A}_0^L)^{-1}$ e de $(A^I)^{-1}$. Os passos anteriores são os passos clássicos do algoritmo simplex onde nós utilizamos o fato que não é necessário trabalhar com $(A^B)^{-1}$ sob forma explícita, pois tudo pode ser feito a partir de $(\hat{A}_0^L)^{-1}$ e $(A^I)^{-1}$. No item a seguir introduzimos a especialização do método.

III. ATUALIZAÇÃO DA BASE

Devemos retirar A^p da base e substituí-lo por A^s mas de forma que a nova base $A^{B'}$ tenha a estrutura definida em (3). Temos então dois casos a analisar que correspondem às duas situações apresentadas no cálculo de θ_r .

Caso 1. $r < m_0$. A^p é tal que $p \in L$, i.é., a coluna a sair da base é não chave. A inversa da nova base é dada por:

$$(A^{B'})^{-1} = MP_1(r,s)(A^B)^{-1} \quad (20)$$

com $MP_1(r,s)$, matriz de pivoteamento, dada por:

$$\hat{c}^{\bar{B}} = \tilde{c}^{\bar{B}} = \bar{c}^{\bar{B}} - p_0 \bar{A}_0^{\bar{B}} \quad (16)$$

Se $\hat{c}^{\bar{B}} \geq 0$ então $A^{\bar{B}}$ é uma base ótima de (P). Em caso contrário, seja $\hat{c}^{\bar{S}} < 0$, a coluna $A^{\bar{S}}$, $s \in \bar{B}$, é candidata entrar na base.

Passo 3: Escolha da coluna a sair da base

Devemos expressar $A^{\bar{S}}$ e b em termos da base corrente $A^{\bar{B}}$, seja

$$\begin{aligned} \hat{A}^{\bar{S}} &= (A^{\bar{B}})^{-1} A^{\bar{S}} \\ &= \begin{bmatrix} (\bar{A}^L)^{-1} \bar{A}_0^{\bar{S}} \\ \hline \bar{A}^{\bar{S}} - \bar{A}^L (\bar{A}_0^L)^{-1} \bar{A}_0^{\bar{S}} \end{bmatrix} = \begin{bmatrix} \bar{A}_0^{\bar{S}} \\ \hline \bar{A}^{\bar{S}} \end{bmatrix} \end{aligned} \quad (17)$$

e

$$\begin{aligned} \hat{b} &= (A^{\bar{B}})^{-1} b \\ &= \begin{bmatrix} (\bar{A}_0^L)^{-1} \bar{b}_0 \\ \hline \bar{b} - \bar{A}^L (\bar{A}_0^L)^{-1} \bar{b}_0 \end{bmatrix} = \begin{bmatrix} \bar{b}_0 \\ \hline \bar{b} \end{bmatrix} \end{aligned} \quad (18)$$

Devemos a seguir calcular

$$\theta_r = \frac{\hat{b}_r}{\hat{A}_r^{\bar{S}}} = \min_{j/\hat{A}_j^{\bar{S}} > 0} \frac{\hat{b}_j}{\hat{A}_j^{\bar{S}}}$$

Podemos encontrar duas situações diferentes no cálculo de θ_r :

a) $r \leq m_0$, neste caso temos

$$\theta_r = \frac{\hat{b}_r}{\hat{A}_r^{\bar{S}}} = \frac{[\bar{b}_0]_r}{[\bar{A}_0^{\bar{S}}]_r} \quad (19)$$

b) $r > m_0$ ($r = m_0 + r'$), neste caso

Esta operação é equivalente a retirarmos \bar{A}_0^p da base de (PR) e a substituímos por \bar{A}_0^s , ou seja, é uma operação de pivoteamento no problema reduzido. I não se altera e em consequência $(A^I)^{-1}$ também não se altera. Devemos atualizar L e B .

Caso 2. $r = m_0 + r'$. A^p é então tal que $\rho \in I_n$, a coluna a sair da base é chave. Podemos distinguir dois subcasos:

Caso 2-a. $\rho \in I_h$ mas $s \notin B_h$. Isto implica que as colunas a permutar na base não pertencem ao mesmo sub-sistema. Após a permuta de A^p e de A^s teremos

$$B'_h = B_h - \{\rho\}$$

pois $s \notin B_h$. Pelo teorema 1, para que $A^{B'}$, $B' = B - \{\rho\} + \{s\}$, seja uma base, é preciso que o "rank" de $[A_h]^{B'_h}$ seja m_h . A matriz I'_h é de rank $m_h - 1$ pois $I'_h = I_h - \{\rho\}$. Em consequência existe no conjunto L_h um vetor $[A_h]^T$ linearmente independente dos vetores de $[A_h]^{I'_h}$. Para que isto seja possível é necessário que $[A_h]^T$ tenha uma componente não nula na direção $[A_h]^p$, i.é, que o elemento $[\hat{A}]^T_{\tau}$, seja necessariamente não nulo. Suporemos que τ é o t -ésimo índice de L .

O intercâmbio direto de A^s e A^p destrói a estrutura bloco-diagonal de A^I . Para evitar isto, começamos por trocar A^T e A^p , o que altera a ordenação de B sem modificar a estrutura de A^B . Após esta mudança A^p é tal que $\rho \in L'$, o novo conjunto das variáveis não chaves, aplica-se então o caso anterior.

O intercâmbio de posição entre as colunas τ e ρ é equivalente a trocar de posição as linhas t e r de $(A^B)^{-1}$.

$$(A^B)^{-1} = \left[\begin{array}{c|c} (\bar{A}_0^{-1}) & \cdot \\ \hline -\bar{A}^L (\bar{A}_0^L)^{-1} & \cdot \end{array} \right] \begin{array}{l} \leftarrow \text{linha } t \text{ (} t \leq m_0 \text{)} \\ \leftarrow \text{linha } r \text{ (} r = m_0 + r' \text{)} \end{array}$$

Após trocarmos as linhas t e r a nova base de trabalho $(\bar{A}_0^{-L'})^{-1}$ será dada por:

onde $j = \sum_{i=1}^{h-1} m_i$, $r'' = r' - j$ e os elementos da matriz de pivoteamento dados por

$$\alpha_{j+k} = \begin{cases} -[\tilde{A}]_{j+k}^s / (\tilde{A})_{j+r''}^s & k = 1, 2, \dots, m_h \quad k=r'' \\ 1 / [\tilde{A}]_{j+r''}^s & k = r'' \end{cases}$$

Finalmente, devemos atualizar I_h e \bar{B} .

RESOLUÇÃO DO PROBLEMA BLOCO ANGULAR: MÉTODO DUALI. DUAL DO PROBLEMA BLOCO ANGULAR

Para definir o dual do problema bloco-angular, devemos inicialmente construir o Lagrangiano de (P) dualizando em relação às restrições (1). Representaremos o vetor das variáveis duais (parâmetros de Lagrange) por $p = [p_0, p_1, \dots, p_\ell]$. Temos então:

$$\begin{aligned} L(x, p) &= cx + p (b - Ax) \\ &= (c - p_0 A_0 - pA)x + p_0 b_0 + pb \end{aligned} \quad (32)$$

com $p = [p_0, p_1, \dots, p_\ell]$ e $p_i = (1, m_i)$ para $i = 0, 1, \dots, \ell$

A função dual é definida por:

$$\begin{aligned} \phi(p) &= \min_{x \geq 0} L(x, p) \quad \text{para } p \in W \\ &= L(x^*, p) \end{aligned} \quad (33)$$

onde para p dado, x^* é a variável primal que minimiza o Lagrangiano e W é o conjunto das variáveis duais para as quais este mínimo existe. W é então o politopo definido por

$$c - p_0 A_0 - pA \geq 0$$

ou ainda

$$pA \leq c \quad (34)$$

O problema dual é então definido por

$$(D) \begin{cases} \text{MAX} & \phi(p) = p b \\ \text{s.a.} & p A \leq c \end{cases}$$

Seja A^B uma base de (1). A esta base podemos associar um vetor p tal que tenhamos

$$\bar{c}^B = c^B - p A^B = 0 \quad (35)$$

Se além disso, supondo não degenerescência do problema dual, tivermos

$$\bar{c}^B = c^B - p A^B > 0 \quad (36)$$

poderemos afirmar que p , solução admissível de (D); corresponde a um vértice do politopo W pois é solução de $m+m_0$ restrições linearmente independente saturadas de (34).

Base dual admissível de (P). No caso acima diremos que A^B é uma base dual admissível de (P).

Uma aresta de W corresponde a todo vetor p tal que m_0+m-1 restrições linearmente independentes de (34) sejam saturadas e tal que as outras desigualdades sejam não saturadas ($\bar{c}^j > 0$).

Apresentamos a seguir alguns resultados clássicos que são particularmente importantes na sequência deste trabalho.

Teorema 3. $\phi(p) \leq f(x)$ para todo $p \in W$ e para todo x tal que $A x = b, x \geq 0$.

Corolário 1. (a) Se $\text{MAX} \{\phi(p) / p \in W\} = +\infty$ então (P) não tem solução admissível.

(b) Se $\text{MIN} \{f(x) / A x = b, x \geq 0\} = -\infty$ então (D) não tem solução admissível.

Corolário 2. Se $\phi(p^*) = f(x^*)$ com x^* solução admissível do primal e $p^* \in W$ então x^* é solução ótima de (P) e p^* é solução ótima de (D).

O problema dual de (PL) pode ser resolvido trabalhando-se com as variáveis duais ou com as variáveis primais. Deveremos neste trabalho usar a segunda forma.

O método a seguir de resolução de (D) consiste em se deslocar de vértice em vértice de W de tal sorte que a cada passo $\phi(D)$ cresce. Devemos inicializar o método em um vértice de W que corresponde, como vimos acima, a uma base dual admissível de (P). Apresentamos então o método em duas fases. A primeira que chamaremos fase 1, consiste em procurar um vértice inicial de W , i.é., uma base dual admissível de (P). A fase 2 é aquela em que otimizamos $\phi(D)$.

II. FASE 1 - PROCURA DE UMA BASE DUAL ADMISSÍVEL

Como no método primal exposto anteriormente, trabalhamos com bases ordenadas de forma que elas possuam uma estrutura bloco-angular. A fase 1 é realizada em dois passos. No primeiro passo procuramos um conjunto I de colunas chaves, obtemos então um vetor D , solução admissível de (D), que não corresponde a um vértice de W . Nós não dispomos então ainda de uma base dual admissível de (P). Na segunda etapa procuramos um conjunto L de colunas não chaves para completar a construção da base.

Passo 1: Procura de um conjunto de colunas chaves

Seja o Lagrangiano definido em (32). Escolhemos um vetor p_0 qualquer e façamos $c' = c - p_0 A_0$ e $K = p_0 b_0$. Temos então

$$L(x, D) = (c' - pA)x + pb + K \quad (37)$$

Esta forma sugere a construção de ℓ pares de problemas primais e duais, que chamaremos auxiliares, do tipo:

Problema Primal	Problema dual
$(PAP)_i \left\{ \begin{array}{l} \text{MIN} \quad c'_i x_i \\ \text{s.a.} \quad A_i x_i = b_i \\ \quad \quad x_i \geq 0 \end{array} \right.$	$(PAD)_i \left\{ \begin{array}{l} \text{MAX} \quad p_i b_i \\ \text{s.a.} \quad p_i A_i \leq c'_i \end{array} \right.$

Vamos definir os conjuntos seguintes:

$$S_0 = \{x/A_0x = b_0, x > 0\}$$

$$S_i = \{x_i/A_i x_i = b_i, x > 0\}$$

$$S = \bigcap_{i=0}^{\ell} S_i$$

Se certos problemas primais $(PAP)_i$

a) Não têm soluções admissíveis, então os conjuntos correspondentes S_i são vazios, S é então vazio e (P) não tem solução admissível.

b) Têm solução $-\infty$, então, com a escolha de p_0 que fizemos, é impossível obter um vetor admissível. Neste caso podemos acrescentar a estes problemas uma restrição do tipo $\eta x_i + \alpha_i = M_i$ com $\eta = (1, 1, \dots, 1)$, α_i a variável de folga e M_i um escalar muito grande. Os conjuntos S_i correspondentes, aos quais acrescentamos as novas restrições, são agora limitados, o que implica a existência de uma solução ótima finita. Se a solução ótima do problema (P) , no qual acrescentamos as restrições acima é tal que uma variável de folga α_i não pertence a base, então o problema original não tem solução finita.

A resolução dos problemas $(PAP)_i$ geram um conjunto de bases A_i^I . O vetor multiplicador p é tal que:

$$\begin{aligned} c^I - pA^I &= (c^I - p_0 A_0^I) - pA^I \\ &= c^I - pA^I \\ \Delta &= c^I \\ &= 0 \end{aligned} \tag{38}$$

Anotando por J o conjunto das variáveis não básicas, a resolução dos $(PAP)_i$ nos permite igualmente afirmar que:

$$\begin{aligned}
c^J - pA^J &= (c^J - p_0 A_0^J) - pA^J \\
&= \bar{c}^J - p \bar{A}^J \\
\bar{c}^J &\geq 0
\end{aligned} \tag{39}$$

Por conseguinte $p = [p_0 \ p]$ é uma solução admissível de (D). Se existe uma solução admissível, existe então um ponto extremo, vértice de W , que corresponde à uma base dual admissível.

Notamos que não é necessário otimizar $(PAP)_i$, as expressões (38) e (39) são válidas qualquer que seja o conjunto de bases duais-admissíveis dos problemas auxiliares. Na realidade, interessa otimizar os problemas auxiliares porque isto permite saber se (P) não tem solução admissível como vimos em a). Esta otimização é igualmente necessária no método de ROSEN como veremos mais adiante.

Passo 2. Procura de um conjunto de colunas não chaves

Procuramos um vértice de W fazendo variar p_0 . Se p segue variações de p_0 de maneira a satisfazer (13) então $\bar{c}^I = 0$ permanecerá válido. Substituindo (13) em (32) obtemos:

$$L(x, p(p_0)) = (\bar{c}^J - p_0 \bar{A}^J) x_j + p_0 \bar{b}_0 + c^I \bar{b} \tag{40}$$

com

$$\bar{c}^J = c^J - c^I (A^I)^{-1} A^J$$

$$\bar{A}_0^J = A_0^J - A_0^I (A^I)^{-1} A^J$$

$$\bar{b}_0 = b_0 - A_0^I (A^I)^{-1} b$$

Este Lagrangiano nos sugere um problema equivalente que chamaremos problema reduzido:

$$(PR) \begin{cases} \text{MIN } \bar{c}^J x_J \\ \text{s.a. } \bar{A}_0^J x_J = \bar{b}_0 \\ x_J \geq 0 \end{cases}$$

A busca de uma base dual admissível de (PR) gera um vetor p_0 tal que:

$$\hat{c}^L = \bar{c}^L - p_0 \bar{A}_0^L = 0$$

$$\hat{c}^B = \bar{c}^B - p_0 \bar{A}_0^B \geq 0$$

O vetor $p = [p_0 \ p]$, com $p = (c^I - p_0 A_0^I) (A^I)^{-1}$, corresponde a um vértice de W , i.é., a uma base dual admissível de (P) com $B = L \oplus I$. Com efeito, A^B é um conjunto de $m_0 + m$ colunas linearmente independentes onde L é o conjunto das variáveis não-chaves tais que

$$\hat{c} = c^L - p_0 A_0^L - p A^L = c^L - p A^L$$

$$\hat{c} = \hat{c}^L = 0$$

e ainda

$$\hat{c} = c^B - p_0 A_0^B - p A^B = c^B - p A^B$$

$$= \hat{c}^B \geq 0$$

e

$$\hat{c}^I = 0$$

O problema (PR) também não precisa ser resolvido, é suficiente de fato procurarmos uma base dual admissível deste problema para obtermos um vértice de W .

Ao final do passo 2, o conjunto de colunas cujos índices são $B = L \oplus I$ constitui uma base dual admissível de (P). Além disso esta base possui uma propriedade particular: por construção a ordem dos elementos de B é tal que A^B possui uma estrutura bloco angular.

III. FASE 2 - OTIMIZAÇÃO DO PROBLEMA (D)

Teste de Otimalidade: Deslocamo-nos de vértice em vértice de W e, para sabermos se atingimos o vértice ótimo, devemos estabelecer um teste de otimalidade. Como sabemos que um vértice de W corresponde a uma solução de base dual admissível de (P), é esta solução que será analisada para estabelecer o teste de otimalidade.

Chamemos x^* a solução básica dual admissível de (P) que corresponde ao vetor p^* corrente (vértice de W).

$$\text{Teorema 4. } cx^* = \phi(p^*) \quad (41)$$

$$\begin{aligned} \text{Prova: } cx^* &= c^B x_B^* + c^{\bar{B}} x_{\bar{B}}^* = (c^B - p^* A^B) x_B^* + p^* A^B x_B^* \\ &= \bar{c}^B x_B^* + p^* b = \phi(p^*) \end{aligned}$$

Sabemos que $A x^* = b$ e $cx^* = \phi(p^*)$, então pelo teorema 3 se $x^* \geq 0$ então x^* e p^* são respectivamente soluções de (P) e (D). Além disso sabemos que $x_B^* = 0$, devemos então a cada iteração testar se x_L^* e x_I^* são não negativos. Se tal é o caso, então a base corrente é ótima e em caso contrário é necessário procurar uma nova base dual admissível, i.é., um novo vértice de W .

Dado um vetor p^* que corresponde a uma base dual admissível A^B de (P), no mínimo do Lagrangiano $L(x, p^*)$ a variável primal x^* pode tomar qualquer valor no conjunto X_B definido por

$$X_B = \{x^*/x_B^* \geq 0 \quad e \quad x_{\bar{B}}^* = 0\}$$

A solução básica dual admissível que corresponde ao vértice p^* e que anotamos por x^* , é tal que, em geral, as componentes de x_B^* podem ser positivas, nulas ou negativas e $x_{\bar{B}}^* = 0$. O algoritmo dual simplex, que trabalha com as variáveis primais, não é outra coisa que a procura de uma base dual admissível tal que $x^* \in X_B$.

Mudança de Base. Toda coluna j da base tal que a variável correspondente x_j^* seja negativa é candidata a sair da base.

Devemos considerar dois casos:

Caso 1. $x_p < 0$ e a coluna A^p é tal que $\rho \in A$ coluna ρ ocupa a posição $r \leq m_0$ em B .

Para sabermos qual é a coluna que deve entrar na base devemos calcular a linha r de A^B expressa em termos da base de vetores da base A^B , ou seja:

$$\begin{aligned} (A^B)^{-1} A^B &= [A^B]_r \\ &= [\tilde{A}_0^B]_r \\ &= [(\tilde{A}_0^B)^{-1}]_r \tilde{A}_0^B \end{aligned} \quad (42)$$

Esta expressão mostra que é necessário obter a linha r de (PR) na forma canônica, com (PR) construído a partir da base corrente. A coluna que vai entrar na base é obtida calculando-se:

$$\epsilon = \frac{-\tilde{c}^s}{[\tilde{A}_0^B]_r^s} = \min_{j \in \bar{B} / [\tilde{A}_0^B]_r^j < 0} \frac{-\tilde{c}^j}{[\tilde{A}_0^B]_r^j} \quad (43)$$

Se $[\tilde{A}_0^B]_r \geq 0$ então $\phi(P)$ pode crescer indefinidamente e (P) não tem assim solução admissível. Em caso contrário, A^s é a coluna que entra na base.

É necessário então atualizar a base A^B . Seja B' a nova base, temos então:

$$(A^{B'})^{-1} = MP_1(r,s)(A^B)^{-1}$$

com $MP_1(r,s)$ uma matriz de pivoteamento cuja forma é dada por (21), donde deduzimos:

$$(\tilde{A}_0^{B'})^{-1} = MP_2(r,s)(\tilde{A}_0^B)^{-1}$$

e a matriz de pivoteamento é obtida (como no método primal) \tilde{a}

partir da coluna \tilde{A}_0^S de (PR) expressa em termos dos vetores da base \tilde{A}_0^L . Como o conjunto I não é modificado, A^I e $(A^I)^{-1}$ não mudam, é suficiente atualizar L, B e $(\tilde{A}_0^L)^{-1}$.

Em resumo, este primeiro caso corresponde a aplicação de uma iteração do algoritmo dual simplex em (PR) obtido a partir dos conjuntos L e I correntes.

Caso 2. $x_p < 0$ e A^p é tal que $p \in I_h$. A coluna A^p ocupa a posição $r = m_0 + r'$ em B. Aqui novamente temos dois casos a considerar.

Caso 2-a. $[\tilde{A}^L]_{r'} \neq 0$. Neste caso é possível, como nós vimos no método primal, efetuar uma permuta das colunas A^p e A^t , $t \in L_h$ e ocupando a posição t em B, com $[\tilde{A}^L]_{r'}^t \neq 0$.

Devemos atualizar $(\tilde{A}_0^L)^{-1}$ e $(A^I)^{-1}$. O procedimento de atualização da base é idêntico ao caso 2-a do método primal. Após esta permuta de colunas, a situação é a mesma do caso 1 e procedemos então da mesma maneira.

Caso 2-b. $[\tilde{A}^L]_{r'} = 0$. Aqui não podemos permutar a coluna A^p com uma outra coluna do conjunto L. Procuramos então diretamente a coluna que deve entrar na base. Para isto é necessário expressar a linha r de $A^{\bar{B}}$ (r' de $A^{\bar{B}}$) em função dos vetores da base corrente, ou seja:

$$\begin{aligned}
 [\hat{A}^{\bar{B}}]_r &= [(A^B)^{-1}]_r A^{\bar{B}} \\
 &= [\hat{A}^{\bar{B}}]_{r'} - [\hat{A}^L]_{r'} (\tilde{A}_0^L)^{-1} \tilde{A}_0^{\bar{B}} \\
 &= [\hat{A}^{\bar{B}}]_{r'} \\
 &= [[\hat{A}_h^{\bar{B}}]_{r'} \quad 0 \quad \dots \quad 0] \quad (44)
 \end{aligned}$$

Em consequência a coluna que deve entrar na base pertence necessariamente a $A^{\bar{B}h}$. Da mesma forma que no método primal $(\tilde{A}_0^L)^{-1}$ e L não sofrem mudanças. Devemos atualizar $(A_h^{Ih})^{-1}$, I_h e B_h . Note mos que a atualização de $(A_h^{Ih})^{-1}$ corresponde a um pivoteamento no sub-sistema h. Devemos então calcular:

$$\varepsilon = \frac{-\hat{c}^s}{[\hat{A}]_{r'}^s} = - \frac{\hat{c}^s}{[\hat{A}_h]_{r'}^s} = \text{MIN}_{j \in \bar{B}_h / [\hat{A}_h]_{r'}^j < 0} \frac{-\hat{c}^j}{[\hat{A}_h]_{r'}^j} \quad (45)$$

Se $[\hat{A}_h^{\bar{B}_h}]_{r'} \geq 0$, então $\phi(P)$ pode crescer indefinidamente e (P) não tem então solução admissível. Em caso contrário A^s entra na base.

A atualização de $(A_h^{I_h})^{-1}$ será feita por

$$(A_h^{I_h})^{-1} = MP_G(r', s) (A_h^{I_h})^{-1}$$

com $MP_G(r', s)$, uma matriz de pivoteamento como em (31).

IV. ESTRATÉGIA DE APLICAÇÃO DO MÉTODO DUAL

Notemos que se fazemos uma escolha aleatória da coluna que sai da base, podemos ser levados a permutar duas colunas A^ρ e A^τ (caso 2-a) com $x_\rho < 0$ e $x_\tau < 0$. Para evitar uma tal situação propomos um critério para escolher a coluna que deve deixar a base. A fase dois do método dual se divide então em três etapas.

Etapa 1. Aplicamos o teste de otimalidade somente às variáveis x_j com $j \in L$ até esgotarmos todas as possibilidades de aplicação do caso 1. Ao final desta etapa teremos então $x_L \geq 0$. Se $x_L \geq 0$ obtemos a solução ótima, senão:

Etapa 2. Aplicamos o teste de otimalidade às variáveis x_j com $j \in I$ e efetuamos todas as permutações de colunas possíveis entre I e L (caso 2-a), garantindo que uma coluna que saiu de I não volte a I no curso desta etapa.

Se conseguirmos efetuar ao menos uma permuta nesta etapa, voltamos à etapa 1. Caso contrário, passamos à etapa 3.

Etapa 3. Fazemos uma permuta direta em todos os subsistemas onde seja possível aplicar o caso 2-b e voltamos à etapa 2.

V. VARIANTE DO MÉTODO DUAL PROPOSTO

Suponhamos que no passo 1 da fase 1 tenhamos obtido um conjunto de colunas chaves a partir das bases ótimas de $(PAP)_i$. Isto implica que as colunas chaves são tais que as bases correspondentes dos problemas auxiliares são admissíveis para estes problemas.

Notemos que a solução ótima x^0 de (P) é tal que para o sub-sistema i temos:

$$\begin{aligned} A_i^{B_i^0} x_{B_i^0}^0 &= b_i \\ x_{B_i^0}^0 &\geq 0 \end{aligned}$$

com B_i^0 o sub conjunto de colunas de B^0 que correspondem ao bloco i (B^0 , é bem entendido, a base ótima de (P)). Posto que $x_{B_i^0}^0$ é uma solução admissível do sub-sistema i , do teorema fundamental da programação linear, existe ao menos uma solução básica admissível (e então uma partição I_i^0, L_i^0 de B_i^0) de tal sorte que

$$(A_i^{I_i^0})^{-1} b_i \geq 0$$

Propomos então, a cada iteração, passar de B a B' , com A^B e $A^{B'}$ bases duais admissíveis de (P), mas de forma que os conjuntos I e I' correspondentes sejam tais que A^I e $A^{I'}$ sejam bases admissíveis dos sub-sistemas. Em outros termos, $(A^I)^{-1} b_i \geq 0$ e $(A^{I'})^{-1} b_i \geq 0$. As modificações em relação à estratégia precedente são então as seguintes:

a) Na etapa 2 a permuta de colunas entre I_i e L_i é feita de forma a manter a admissibilidade da base dos sub-problemas i . As colunas que devemos permutar são obtidas pela resolução dos sub-problemas:

$$\left\{ \begin{array}{l} \text{MIN } \sum_{i \in k_i} x_i \\ \text{s.a. } A_i^{B_i} x_{B_i} = b_i \\ x_{B_i} \geq 0 \end{array} \right.$$

com $K_i = \{i \in I_i / x_i < 0\}$. A resolução deste problema gera uma partição L'_i, I'_i tal que

$$\begin{cases} I'_i \oplus L'_i = B_i \\ (A_i^{I'_i})^{-1} b_i \geq 0 \end{cases}$$

b) A etapa 3 é eliminada porque

$$x'_r = [\hat{b}_i]_{r'} - [\hat{A}_i^{L_i}]_{r'} x_{L_i} \geq 0$$

sempre que $[\hat{A}_i^{L_i}]_{r'} = 0$

Os problemas auxiliares (PAP)_i ou (PAD)_i que podem ser construídos, a cada iteração, se encontram sempre no ótimo, visto que as bases correspondentes são admissíveis com o vetor de custo relativo não negativo também a cada iteração.

Esta variante é o método de ROSEN para problemas lineares com estrutura bloco angular.

INTERPRETAÇÃO GEOMÉTRICA DO MÉTODO DUAL
CONCLUSÃO DA PARTE II

I. INTERPRETAÇÃO GEOMÉTRICA

Supondo que nós já tenhamos obtido uma solução de base dual admissível (fase 1 do método dual) fazemos uma interpretação geométrica da fase 2, do método dual, que corresponde à estratégia da resolução apresentada no Capítulo VII, item IV.

Princípio do Método. Vimos que uma solução de base dual admissível de (P) corresponde a um vértice do conjunto W . O método apresentado, baseado no algoritmo dual simplex consiste em passar de vértice em vértice de W (definido no E^{m+m_0}) de tal sorte que a função dual $\phi(p)$ seja aumentada a cada iteração, e isto de forma a manter uma estrutura bloco angular para as bases duais admissíveis de (P).

Podemos explorar vantajosamente esta estrutura projetando, a cada iteração, um subconjunto $W_R \subset W$ no espaço E^{m_0} das variáveis duais p_0 . Se o máximo de $\phi(p)$ não é atingido, uma nova iteração do método redefine um novo conjunto W_R no qual se encontra um vértice que dá a $\phi(p)$ um valor superior àquele obtido até o momento.

Interpretação da Etapa 1. Seja um vértice de W que anotaremos por Σ , obtido fazendo-se $p = p^*$ que corresponde à base dual admissível $B = L \oplus I$. Para explorar a estrutura bloco angular, decidimos não alterar o conjunto I das colunas chaves (i.é, manter $\hat{c}^I = 0$). Esta limitação impõe uma limitação na pesquisa de um vértice vizinho de Σ . Entre todas arestas partindo de Σ , somente consideramos aquelas que são geradas fazendo-se $\hat{c}^r \neq 0$ para um elemento $r \in L$.

Para manter $\hat{c}^I = 0$ e fazer $\hat{c}^{L-r} = 0$, $\hat{c}^r = \epsilon$, devemos escolher p tal que

$$p = p^* - \varepsilon (A^B)_r^{-1}, \quad \varepsilon > 0 \quad (46)$$

Para ε pequeno, a variação correspondente da função dual é dado por:

$$\phi(p^* - \varepsilon (A^B)_r^{-1}) = \phi(p^*) - \varepsilon x_r^+ \quad (47)$$

Se escolhermos $r \in L$ tal que $x_r^+ < 0$ é evidente que um deslocamento na direção da aresta correspondente W provoca um aumento de $\phi(p)$.

Utilizando (12) e sabendo-se que $r \leq m_0$, a direção de deslocamento no espaço dual é dado por:

$$(A^B)_r^{-1} = \left[\begin{array}{c|c} [(\bar{A}_0^L)^{-1}]_r & -[(\bar{A}_0^L)^{-1}]_r A_0^I (A^I)^{-1} \end{array} \right] \quad (48)$$

Substituindo (48) em (46) obtemos:

$$p_0 = p_0^* - \varepsilon [(\bar{A}_0^L)^{-1}]_r \quad (49)$$

$$\begin{aligned} p &= p^* + \varepsilon [(\bar{A}_0^L)^{-1}]_r A_0^I (A^I)^{-1} \\ &= (c^I - p_0^* A_0^I) (A^I)^{-1} + \varepsilon [(\bar{A}_0^L)^{-1}]_r A_0^I (A^I)^{-1} \\ &= (c^I - p_0 A_0^I) (A^I)^{-1} \end{aligned} \quad (50)$$

(50) mostra portanto que A^I permanece sem mudanças e (49) que, de fato, efetuamos um passo do algoritmo dual simplex em (PR).

Com efeito, o fato de impormos $\hat{c}^I = 0$ implica que nos limitamos ao estudo de um conjunto reduzido de vértice de W que inclui E . O poliedro constituído pela combinação convexa destes vértices é chamado W_R e sua projeção no espaço E^{m_0} das variáveis duais p_0 é um poliedro anetado por W_R , constituído pelas restrições do dual de (PR).

Existe uma correspondência biunívoca entre os vértices

de W_R e aqueles de $W_{R'}$. Com efeito, dois vértices adjacentes de W_R não podem se projetar em um mesmo ponto de E^m pois para passar de um outro devemos escolher a direção de deslocamento $[(A^B)^{-1}]_r$ cuja componente no espaço dos p_0 é $[(\tilde{A}^L)^{-1}]_{r \neq 0}$. Consequentemente, uma aresta de W_R se projeta em uma aresta de $W_{R'}$.

Cada iteração da etapa 1 consiste em um passo do algoritmo simplex aplicado a (PR) e corresponde então a um deslocamento sobre uma aresta de $W_{R'}$, projeção de uma aresta de W_R , até atingir o vértice adjacente.

A cada iteração aumentamos $\phi(P) = \phi(p_0, p(p_0))$ e o procedimento se repete até que obtemos o melhor vértice de $W_{R'}$ (e portanto o melhor vértice de W_R posto que a correspondência entre os vértices de W_R e $W_{R'}$ é biunívoca).

Não é necessário otimizar o problema (PR), pois a partir de um vértice vizinho de Σ , poderíamos definir um novo conjunto de colunas chaves (um novo conjunto $W_{R'}$), mas é vantajoso só redefinir W_R quando não for mais possível melhorar $\phi(P)$ com o conjunto W_R corrente, o que justifica o procedimento proposto na etapa 1.

Interpretação da etapa 2. Se os vértices ótimos de W_R não maximizam $\phi(P)$, i.é, se existe $x_i^+ < 0$ para $i \in I$, devemos definir um novo conjunto W_R cuja exploração permite aumentar $\phi(P)$. Para isto é necessário modificar o conjunto I de colunas chaves.

Óra, nós sabemos que na etapa 1, as arestas exploradas são aquelas que correspondem a um elemento $\tilde{c}^r \neq 0$ e que um deslocamento sobre estas arestas permitem aumentar $\phi(P)$ se $x_r^+ < 0$. Construímos então um novo conjunto de colunas chaves permutando elementos $i \in I$ (tal que $x_i^+ < 0$) com elementos $r \in L$.

Notamos que estas permutas não modificam a base A^B , mas somente a classificação das colunas chaves e não chaves (arestas chaves e arestas não chaves), consequentemente permanecemos no mesmo vértice Σ de W e a solução de base dual admissível permanece a mesma.

Se for possível efetuar ao menos uma tal permuta, definimos um novo conjunto W_R (portanto um novo problema reduzido (PR)), podemos retornar a etapa 1 para procurar o melhor vértice entre aqueles do novo conjunto W_R . O fato de termos podido

efetuar ao menos uma permuta \tilde{e} suficiente para garantir que retornando a etapa 1 podemos encontrar um melhor v \acute{e} rtice de W_R para aumentar $\phi(P)$.

Nesta etapa \tilde{e} vantajoso fazer o m \acute{a} ximo de permuta \tilde{c} o \tilde{e} s poss \acute{i} veis para um n \acute{u} mero t \tilde{a} o grande quanto poss \acute{i} vel de vari \tilde{a} veis no novo conjunto L seja tal que $x_r^+ < 0$, $r \in L$. O novo conjunto W_R assim definido ter \tilde{a} todas as chances de possuir um maior n \acute{u} mero de v \acute{e} rtices para aumentar $\phi(P)$.

Interpreta \tilde{c} o \tilde{e} da Etapa 3. Suponhamos que na etapa dois n \tilde{a} o foi poss \acute{i} vel permutar colunas chaves e n \tilde{a} o chaves e suponhamos que o \tilde{o} timo de $\phi(P)$ n \tilde{a} o foi ainda encontrado (existe ao menos um \tilde{i} ndice $i \in I$ tal que $x_i^+ < 0$).

Como n \tilde{a} o \tilde{e} poss \acute{i} vel encontrar um novo conjunto W_R , do qual um dos v \acute{e} rtices permitiria aumentar $\phi(P)$, isto implica que no conjunto W_R corrente todas as arestas que permitem sair de Σ para aumentar $\phi(P)$ s \tilde{a} o ortogonais ao espa \tilde{c} o de p_0 .

Com efeito para $r = m_0 + r'$, utilizando (12), a dire \tilde{c} o \tilde{e} de deslocamento no espa \tilde{c} o dual \tilde{e} dado por:

$$\left[(A^B)^{-1} \right]_r = \left[-[\tilde{A}^L]_{r'}, \left[(A^I)^{-1} \right]_{r'} + [\tilde{A}^L]_{r'} (\tilde{A}_0^L)^{-1} A_0^I (A^I)^{-1} \right] \quad (51)$$

e se $[\tilde{A}^L]_{r'} = 0$ quando $x_{m_0+r'}^+ < 0$, isto implica que as arestas que saem de Σ que permitem aumentar a fun \tilde{c} o \tilde{e} $\phi(P)$ s \tilde{a} o ortogonais ao espa \tilde{c} o de p_0 .

Devemos modificar as vari \tilde{a} veis duais de acordo com (46), e tendo em conta (51), obtemos:

$$p_0 = p^* + \epsilon [\tilde{A}^L]_{r'} = p_0^* \quad (52)$$

$$\begin{aligned} p &= p^* - \epsilon [(A^I)^{-1}]_{r'} - \epsilon [\tilde{A}^L]_{r'} (\tilde{A}_0^L)^{-1} A_0^I (A^I)^{-1} = \\ &= p^* - \epsilon [(A^I)^{-1}]_{r'} \quad (53) \end{aligned}$$

A propriedade de ortogonalidade implica um desacoplamento entre (PR) e (RLC). (52) mostra que n \tilde{a} o h \tilde{a} mudan \tilde{c} a de base em (PR) e (53) que efetuamos simplesmente uma itera \tilde{c} o \tilde{e} do algoritmo dual simplex em (RLC):

O fato de trabalharmos no espaço de p é particularmente interessante porque os ℓ sub-sistemas são inteiramente desacoplados. Suponhamos com efeito que $x_{m_0+r'}^+ < 0$ é uma variável do sub-sistema h . Sabendo que $(A^I)^{-1}$ tem uma estrutura bloco angular, (53) pode ser reescrito como:

$$P_k = P_k^*$$

$$P_h = P_h^* - \epsilon \left[A_h^I \right]^{-1} r'' \quad \text{com} \quad r'' = r' - \sum_{i=1}^{h-1} m_i$$

Esta expressão mostra bem que p é simplesmente atualizado por uma iteração do algoritmo dual simplex em um único sub-sistema h .

Para explorar esta propriedade de desacoplamento, devido à ortogonalidade, é interessante que efetuemos uma iteração do algoritmo dual simplex em todos os sub-sistemas onde isto é possível.

II. CONCLUSÃO

Nesta segunda parte deste trabalho nos interessamos pela resolução de problemas de programação linear com estrutura bloco angular. Estabelecemos uma ligação entre o método GGUB e o método de partição de ROSEN. Esta ligação foi estabelecida propondo-se um método dual baseado no algoritmo dual simplex, que pode ser interpretado como o método dual do GGUB e cujo procedimento se dá em três etapas:

. Na primeira etapa resolvemos um programa linear restrito (PR) de dimensão reduzida pelo algoritmo dual simplex. Se o ótimo do problema original (P) não foi atingido:

. Na segunda etapa reordenamos os elementos da base de (P) de forma a construir um novo (PR) cuja resolução permite aumentar o valor da função dual e retornamos à etapa 1. Contudo, se não é possível reordenar os elementos da base:

. Na terceira etapa um desacoplamento permite efetuar uma iteração do algoritmo simplex em um ou mais sub-sistemas, me

lhorando a função objetivo, e retornamos à etapa 2.

Propuzemos uma interpretação geométrica deste método mostrando que a etapa 1 consiste em projetar no espaço da variável p_0 um subconjunto de vértices W_R do politopo das restrições W e a procurar o ótimo neste subconjunto. A etapa 2 consiste em construir um novo subconjunto W_R cuja projeção no espaço de p_0 é $W_{R'}$, a procura do melhor vértice em W_R , permite aumentar a função dual. Enfim, a etapa 3 corresponde ao caso em que as arestas de W , sobre quais poderíamos deslocar para aumentar $\phi(p)$, são ortogonais ao espaço de p_0 .

Mostramos que o método de partição de ROSEN pode ser encarado como uma variante do método dual proposto onde um procedimento particular na etapa 3 (que permite trabalhar com bases admissíveis dos sub-sistemas) suprime a eventualidade de ocorrer a etapa 3.

P A R T E I I I

PROGRAMAÇÃO LINEAR POR PARTES:

PROBLEMAS COM ESTRUTURA

BLOCO ANGULAR

INTRODUÇÃO

Com o desenvolvimento de métodos de otimização para problemas não diferenciáveis, os problemas com função objetivo ou restrições lineares por parte se tornam pesquisas de interesse que se traduzem na publicação de vários artigos a respeito | 1 | | 20 | | 23 | | 34 |.

Problemas lineares por parte podem surgir em várias situações distintas. Uma delas, por exemplo, seria o problema da variação de parâmetros da função objetivo ou do vetor de recursos em (PL) (análise paramétrica). Outra situação seria o problema dual, linear por partes, em problemas de (PL) quando na definição do Lagrangiano somente um subconjunto das restrições é dualizado. Em | 1 | e | 7 | encontramos um exemplo onde restrições lineares por parte aparecem em problemas de planejamento da produção.

Nos interessa uma classe particular de problemas lineares por parte. Chamaremos então de Programação Linear por Partes (PLP) a este ramo da programação matemática cujo objeto é esta classe particular de problemas.

A (PLP) se torna assunto de interesse para o grupo de otimização da FEC-UNICAMP por ser uma generalização natural da (PL) e oferecer aplicações interessantes, por exemplo, alguns problemas de (PL) de grande número de variáveis e restrições podem ter suas formulações adaptadas para (PLP), provocando substanciais reduções de cálculo e de memória computacional | 1 | | 7 |. Assim, algumas pesquisas, que se traduziram em teses de mestrado e uma de doutorado, se direcionaram no sentido de ampliar os métodos de resolução de problemas de (PLP). GOLSTEIN e YODINE | 1 |, pesquisadores soviéticos, apresentam um método primal para (PLP), generalização do primal simplex. SALES | 8 | e FERREIRA | 35 | propõem os métodos dual e primal dual respectivamente em teses de mestrado e FERNANDES | 7 |, em tese de doutorado, expõe um método dual para problemas de (PLP) com estrutura bloco angular. Uma tese anterior a estas, de SOUZA | 9 |, aplicava o método primal de GOLSTEIN e YODINE na resolução de um problema de planejamento de sistemas de potência. Nesta linha, de aplicação dos métodos, outra tese de mestrado se desenvolve no momento na área de planejamento de sistemas hidrotérmicos. Discutimos

no momento a possibilidade de aplicação destes estudos em (PLP) para problemas de otimização de redes com critério linear por parte.

A (PLP) é uma generalização de (PL), como dissemos acima, portanto estruturas particulares de restrições em (PLP) podem ser resolvidos estendendo-se os métodos de (PL) para estas estruturas particulares. Neste sentido os estudos da parte II deste trabalho serão úteis aqui para se resolver problemas de (PLP) com estrutura bloco angular.

No capítulo IX apresentamos a classe de problemas considerados pela (PLP) e o método primal de resolução de um problema desta classe. Procuramos usar na exposição uma notação e terminologia que tornassem o mais claro possível o parentesco deste método com o simplex clássico. No capítulo X expomos o método de resolução deste problema quando apresenta estrutura de restrições bloco angular e a seguir as conclusões desta parte.

PROGRAMAÇÃO LINEAR POR PARTESI. INTRODUÇÃO

Apresentamos no item II deste capítulo a classe de problemas de interesse da (PLP). Desta classe de problemas nos interessa um em particular cuja formulação é apresentado no item III. Alguns conceitos clássicos de (PL) tais como solução básica, forma canônica, degenerescência podem ser estendidos à (PLP). Esta extensão é feita no item IV onde apresentamos também o critério de otimalidade para o problema em estudo. No item V descrevemos uma iteração do método.

II. A CLASSE DE PROBLEMAS DE INTERESSE DE (PLP)

De modo geral podemos considerar como sendo pertinentes a (PLP) problemas com a seguinte fórmulação.

$$\left\{ \begin{array}{l} \text{OTIMIZAR} \quad F(x) = \sum_{j=1}^n f_j(x_j) \\ \text{s.a.} \quad \sum_{j=1}^n g_{ij}(x_j) \{ \leq, =, \geq \} g_i, \quad i=1,2,\dots,s \\ \\ A_x \{ \leq, =, \geq \} b \\ \\ \alpha \leq x \leq \beta \end{array} \right.$$

onde $f_j(x_j)$ e $g_{ij}(x_j)$ são funções lineares por parte, veja figura 1.

Da formulação exposta acima, podemos dizer que a programação linear por partes vai tratar de problemas pertencentes a uma faixa intermediária entre a programação linear e a programação convexa, de função objetivo separável, sujeita a restrições lineares ou lineares por partes, neste último caso as funções lineares por partes que definem as restrições são também

separáveis.

Em [1] e [7] encontramos uma apresentação mais extensa do problema formulado acima. Em particular nos ateremos aos problemas cujas restrições sejam tão somente lineares. Em [1] são feitas indicações no sentido de se resolver o caso em que ocorrem restrições não lineares.

III. O PROBLEMA

O problema de (PLP) de que nos ocuparemos pode ser descrito por

$$\begin{aligned} \text{MIN} \quad & F(x) \\ \text{s.a.} \quad & Ax = b \\ & x \in S \end{aligned}$$

onde S é um poliedro convexo, no nosso caso restrições de "capacidade" do tipo $\alpha \leq x \leq \beta$, $A(m,n)$ é "rank" m , $x(n,1)$, $b(m,1)$ e $F(x)$ é uma função convexa, separável e linear por partes, do tipo

$$F(x) = \sum_{j=1}^n f_j(x_j) \quad (1)$$

tal que

$$f_j(x_j) = \begin{cases} c_0^j x_j + g_{0,j} & \alpha_j \leq x_j \leq d_{1,j} \\ \vdots \\ c_k^j x_j + g_{k,j} & d_{k,j} \leq x_j \leq d_{k+1,j} \\ \vdots \\ c_{\ell_j,j}^j x_j + g_{\ell_j,j} & d_{\ell_j,j} \leq x_j \leq \beta \end{cases} \quad (2)$$

e o coeficiente angular de $f_j(x_j)$ no intervalo $[d_{k,j}; d_{k+1,j}]$ é dado por

$$c_k^j = c_0^j + \sum_{i=1}^k \sigma_i^j, \quad \sigma_i^j > 0 \quad (3)$$

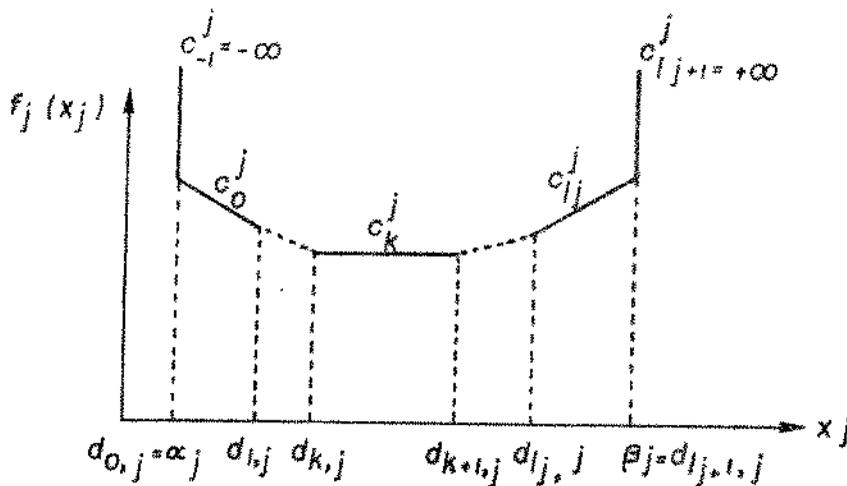
e

$$g_k^j = f_j(d_{k,j}) - c_k^j d_{k,j} \quad (4)$$

Denominaremos ao ponto $d_{k,j}$ de Ponto Crítico de $f_j(x_j)$. Podemos então reescrever o problema acima como:

$$(P) \begin{cases} \text{MIN } F(x) = \sum_{j=1}^n f_j(x_j) & (5) \\ \text{s.a.} & Ax = b & (6) \\ & \alpha \leq x \leq \beta & (7) \end{cases}$$

Nota: Convencionamos que $\alpha_j = d_{0,j}$ e faremos $c_{-1}^j = -\infty$, e $\beta_j = d_{l_{j+1},j}$ e faremos $c_{l_{j+1}}^j = +\infty$



IV. ALGUNS CONCEITOS: SOLUÇÃO BÁSICA E FORMA CANÔNICA, CRITÉRIO DE OTIMALIDADE

Dissemos que (PLP) é uma extensão de (PL). Vamos então estender alguns conceitos usados em (PL) para (PLP).

Solução básica. O sistema (6) pode ser escrito como

$$A^B x_B = b - A^{\bar{B}} x_{\bar{B}}$$

onde $A^B(m, m)$ é uma matriz regular. Prê-multiplicando a relação acima por $(A^B)^{-1}$ obtemos

$$x_B = \hat{b} - \hat{A}^{\bar{B}} x_{\bar{B}}$$

diremos que $x^T = x^{+T} = \begin{bmatrix} x_B^{+T} & x_{\bar{B}}^{+T} \end{bmatrix}$ é uma Solução Básica do sistema (6) se

$$\begin{bmatrix} x_B^+ \\ x_{\bar{B}}^+ \end{bmatrix} = \begin{bmatrix} \hat{b} - \hat{A}^{\bar{B}} d_{K,J} \\ d_{K,J} \end{bmatrix} \quad (8)$$

e x_B^+ , $x_{\bar{B}}^+$ são chamados respectivamente Variáveis Básicas e Não Básicas. Esta solução será factível para o problema (P) se também satisfizer (7).

A matriz A^B , como em (PL), será chamada Matriz Básica e sempre que não houver motivo de confusão chamaremos ao conjunto de índices B de base. Reparar que a uma base podem corresponder várias soluções básicas e que o número destas é finito, dado que o número de conjuntos B que podemos construir e o número de pontos críticos são finitos.

A solução básica x^+ será não degenerada se x_j^+ , $j \in B$, não coincide com nenhum dos pontos críticos $d_{k,j}$, ou seja se $d_{k,j} < x_j^+ < d_{k+1,j}$, $j \in B$.

O problema (P) é não degenerado se cada uma de suas soluções básicas factíveis for não degenerada. Neste trabalho suporemos (P) não degenerado, observando que o estudo de degenerescência em (PLP) pode ser feito como extensão da degenerescência em (PL), sem dificuldades.

Apresentamos a seguir, sem demonstração, a extensão do teorema fundamental da (PL) para (PLP):

Teorema 1. Se o problema (P) tiver uma solução factível ótima então existe uma solução básica factível ótima.

Forma canônica. O problema (P) é equivalente, para x factível e p(1,m), ao problema abaixo

$$\text{MIN } F(x) = \sum_{j=1}^n f_j(x_j) - pAx + pb \quad (9)$$

$$\text{s.a.} \quad Ax = b \quad (10)$$

$$x \geq 0 \quad (11)$$

Seja x^+ uma solução básica factível e não degenerada de (P). Sem perda de generalidade seja $B = \{1, 2, \dots, m\}$. Então x^+ é tal que

$$d_{K,B} < x_B^+ < d_{K+\nu, B} \quad (12)$$

$$x_{\bar{B}}^+ = d_{K, \bar{B}} \quad (13)$$

com $\nu = \{1, 1, \dots, 1\}$

Para $x = x_B^+$ no intervalo definido em (12) corresponde uma inclinação da função objetivo dada por c_K^B . Façamos

$$z^{+B} = p^{+B} A^B = c_K^B \quad (14)$$

donde

$$p^{+B} = c_K^B (A^B)^{-1} \quad (15)$$

e (9) pode ser reescrito como

$$F(x) = \left(\sum_{j \in B} f_j(x_j) \right) - c_K^B x_B + \left(\sum_{j \in \bar{B}} f_j(x_j) \right) - z^{+\bar{B}} x_{\bar{B}} \quad (16)$$

com $z^{+\bar{B}} \triangleq c_K^B A^{\bar{B}}$, denominado Inclinação de Referência. A razão deste nome ficará clara mais adiante ao apresentarmos as figuras 2, 3, 4 e 5.

Para pequenas variações de x em torno de x^+ , o primei-

ro termo de (16), a partir de (2), pode ser aproximado para:

$$\left(\sum_{j \in B} f_j(x_j) \right) - c_K^B x_B = c_K^B x_B + g_{K,B} - c_K^B x_B$$

$$= g_{K,B} \quad (17)$$

e o segundo termo, para x_j à esquerda e à direita de x_j^+ , $|x_j - x_j^+|$ pequeno, pode ser escrito respectivamente como:

$$\text{à esquerda: } \left(\sum_{j \in \bar{B}} f_j(x_j) \right) - z^{+\bar{B}} x_j = (c_{K-\eta}^{\bar{B}} - z^{+\bar{B}}) x_{\bar{B}} + cte_1$$

$$= \hat{c}_{K-\eta}^{\bar{B}} x_{\bar{B}} + cte_1 \quad (18)$$

$$\text{à direita: } \left(\sum_{j \in \bar{B}} f_j(x_j) \right) - z^{+\bar{B}} x_j = (c_K^{\bar{B}} - z^{+\bar{B}}) x_{\bar{B}} + cte_2$$

$$= \hat{c}_K^{\bar{B}} x_{\bar{B}} + cte \quad (19)$$

e chamaremos a $\hat{c}_{K-\eta}^{\bar{B}}$ e $\hat{c}_K^{\bar{B}}$ de Custo Relativo à Esquerda e Custo Relativo à Direita respectivamente.

Podemos então aproximar (P) por:

$$(P') \begin{cases} \text{MIN } F(x_{\bar{B}}) = z_0 + \hat{c}_Y^{\bar{B}}(x^+) x_{\bar{B}} \\ x_B + \bar{A}^{\bar{B}} x_{\bar{B}} = \bar{b} \\ \alpha \leq x \leq \beta \end{cases}$$

onde $z_0 = c_K^B (A^B)^{-1} b + g_{K,B} + cte_1 + cte_2$ e $\hat{c}_Y^{\bar{B}}(x^+) \in \{\hat{c}_{k-1}^{\bar{B}}, \hat{c}_k^{\bar{B}}\}$.

Se a variável x_j aumenta, vale o custo relativo à direita, e se ela diminui vale o custo relativo à esquerda.

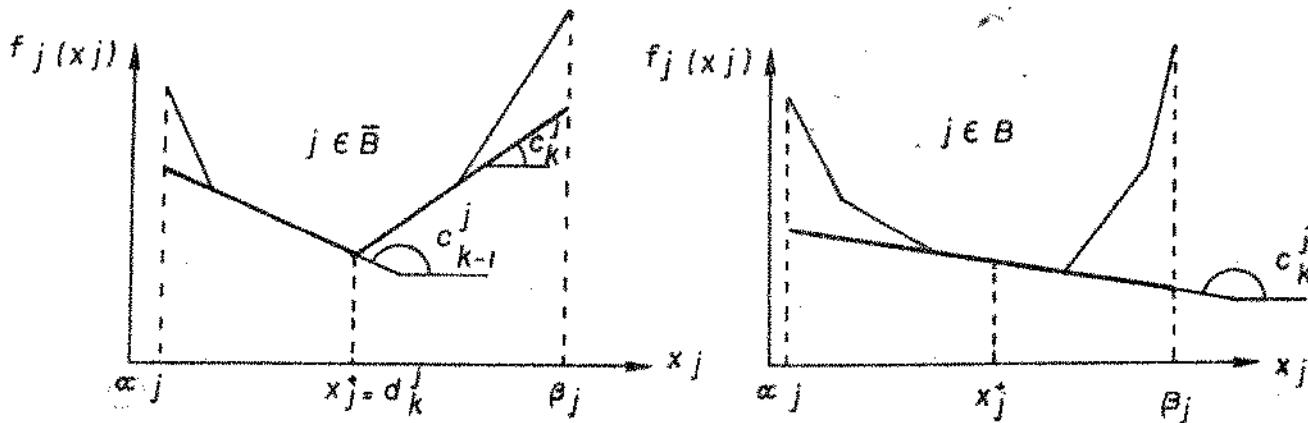
Ao problema (P') chamaremos de Forma Canônica, do problema (P). Reparar que esta forma canônica não está somente relacionada a uma base B mas também a uma solução básica pois os custos relativos dependem das localizações das variáveis básicas

e não básicas no intervalo $\alpha \leq x \leq \beta$. A forma canônica (P') é uma aproximação do problema (P), mas que é exata para x_B e $x_{\bar{B}}$ nos intervalos (Fig. 2 e Fig. 3):

$$d_{K,B} \leq x_B \leq d_{K+\nu,B} \quad (20)$$

$$d_{K-\nu,\bar{B}} \leq x_{\bar{B}} \leq d_{K+\nu,\bar{B}}$$

Esta forma do problema tem a vantagem, para pequenas variações em torno de x^* , da função só depender das variáveis não básicas, facilitando uma análise local da variação de $F(x)$.



Critério de Otimalidade: Deveremos explicitar um critério de otimalidade em termos de soluções básicas, pois, se o problema tem solução factível ótima está assegurada a existência de uma solução básica ótima.

Seja $x = x^*$ uma solução básica factível do problema (P'), vamos fazer ($\lambda(n-m, 1) \geq 0$):

$$x_{\bar{B}} = x_{\bar{B}}^* + t\lambda \quad (21)$$

Para $|t|$ pequeno a forma canônica (P') não se altera. As variáveis básicas serão afetadas de acordo com:

$$\begin{aligned}
 x_B &= \hat{b} - \hat{A}^B (x_B + t\lambda) \\
 &= x_B^+ - t\hat{A}^B \lambda
 \end{aligned}
 \tag{22}$$

ou de forma condensada:

$$\begin{bmatrix} x_B \\ x_{\bar{B}} \end{bmatrix} = \begin{bmatrix} x_B^+ \\ x_{\bar{B}}^+ \end{bmatrix} + t \begin{bmatrix} \hat{A}^B \\ \vartheta \\ \textcircled{0} \end{bmatrix} \lambda
 \tag{23}$$

A função objetivo, para t não nulo, pode ser escrita como:

$$F(x_B^+ + t\lambda) = F(x_B^+) + t c_{Y(x^+)}^B \lambda
 \tag{24}$$

onde

$$\gamma_j(x_j^+) \begin{cases} k & \text{se } t > 0 \\ k-1 & \text{se } t < 0 \end{cases}
 \tag{25}$$

Como $F(x_B)$ é separável podemos estudar suas variações em termos de variações em cada variável x_j , $j \in B$. Fazamos então:

$$\lambda = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \text{linha } r
 \tag{26}$$

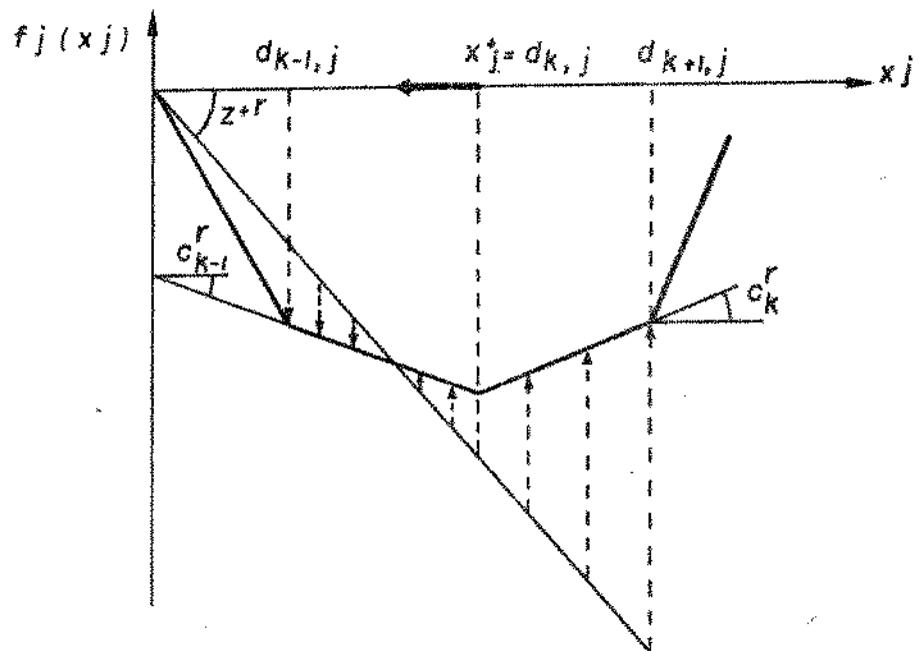
e podemos escrever

$$F(x_B^+ + t\lambda) = F(x_B^+) + t \hat{c}_{Y(x^+)}^r
 \tag{27}$$

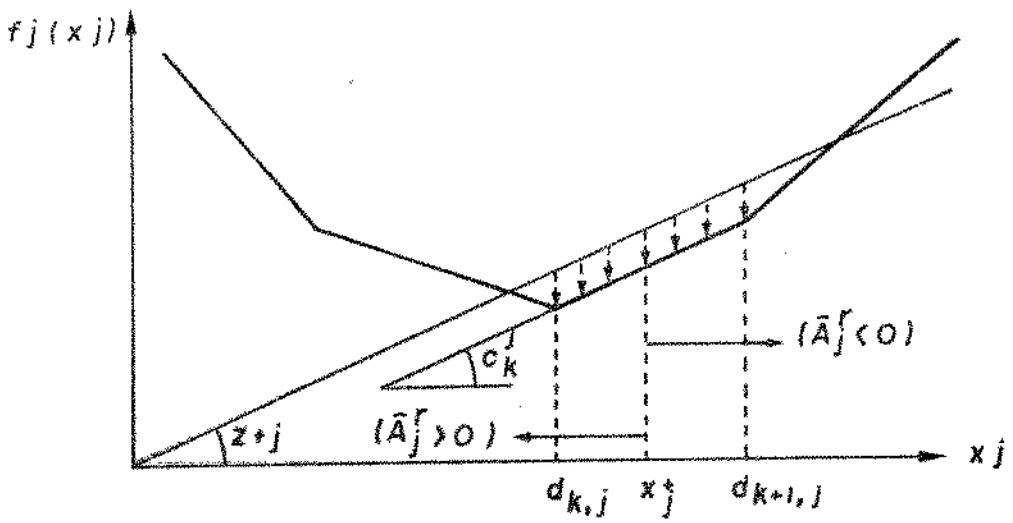
Caso 3. $z^{+r} < c_{k-1}^r$ ou

$$\hat{c}_k^r > \hat{c}_{k-1}^r > 0 \quad (30)$$

Aqui (Fig. 6) o termo $(f_r(x_r) - z^{+r}x_r)$ diminui ao diminuirmos t ($t \leq 0$).



Para as variáveis básicas temos, por construção, $z^{+B} = c_K^B$ e o termo $(f_j(x_j) - z^{+j}x_j)$, $j \in B$, é insensível à variações de x_j (Fig. 7).



Temos então as condições para estabelecer o critério de otimalidade. Se todas as variáveis não básicas estiverem na situação apresentada no caso 2, então qualquer variação de x_B^* em torno de x_B^* faz a função crescer. Podemos concluir que a

solução básica x^+ é um ponto de ótimo local. Como a função $F(x)$ é convexa, então x^+ é também um ótimo global.

Teorema 2. A solução básica x^+ é ponto de ótimo global se:

$$\hat{c}_{k-1}^B < 0 < \hat{c}_K^B \quad (31)$$

V. DESCRIÇÃO DE UMA ITERAÇÃO DO MÉTODO

Seja $x = x^+$ uma solução básica factível e suponhamos o problema (P) escrito em forma canônica em relação a esta solução. Testamos x^+ quanto à otimalidade por meio da relação (31), se não for ótima devemos alterar esta solução.

Vamos alterar x^+ em uma direção dada por (26) tal que somente uma componente do vetor x_B^+ , x_r^+ , seja alterada, e caminhemos nesta direção até que uma nova solução básica factível seja encontrada. Como queremos uma direção em que F decresça, a variável x_r^+ deve ser tal que seus custos relativos à esquerda e à direita sejam ambos negativos ou ambos positivos (caso 1 ou caso 3 do item anterior). De (23) temos então que:

$$\begin{bmatrix} x_B \\ x_B \end{bmatrix} = \begin{bmatrix} x_B^+ \\ x_B^+ \end{bmatrix} + t \begin{bmatrix} A^r \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

ou rearranjando as variáveis:

$$\begin{bmatrix} x_B \\ x_r \\ x_{B-(r)} \end{bmatrix} = \begin{bmatrix} x_B^+ \\ x_r^+ \\ x_{B-(r)}^+ \end{bmatrix} + t \begin{bmatrix} -A^r \\ 1 \\ 0 \end{bmatrix} \quad (32)$$

Façamos uma descrição detalhada para x_r no caso 1, i.é, $\hat{c}_{k-1}^r < \hat{c}_k^r < 0$. Neste caso para F diminuir é preciso aumentarmos t até um valor t_{MAX} tal que uma das variáveis básicas ou x_r^+ alcance um ponto crítico. Para além deste ponto crítico deixaria de valer a forma canônica referente à solução x^+ . O cálculo de t_{MAX} é feito através de:

. Para a variável não básica (veja Fig. 4).

$$x_r^+ + t_r = d_{k,r} + t_r = d_{k+1,r}$$

$$t_r = d_{k+1,r} - d_{k,r} \quad (33)$$

. Para as variáveis básicas, temos dois casos a considerar (veja Fig. 7).

. $\hat{A}_j^r > 0$, neste a variável básica decresce e temos:

$$x_j^+ - t_j \hat{A}_j^r = d_{k,j}$$

e portanto

$$t_j = \frac{x_j^+ - d_{k,j}}{\hat{A}_j^r} \quad (34)$$

. $\hat{A}_j^r < 0$, a variável básica cresce e temos

$$x_j^+ - t \hat{A}_j^r = d_{k+1,j}$$

e então

$$t_j = \frac{x_j^+ - d_{k+1,j}}{\hat{A}_j^r} \quad (35)$$

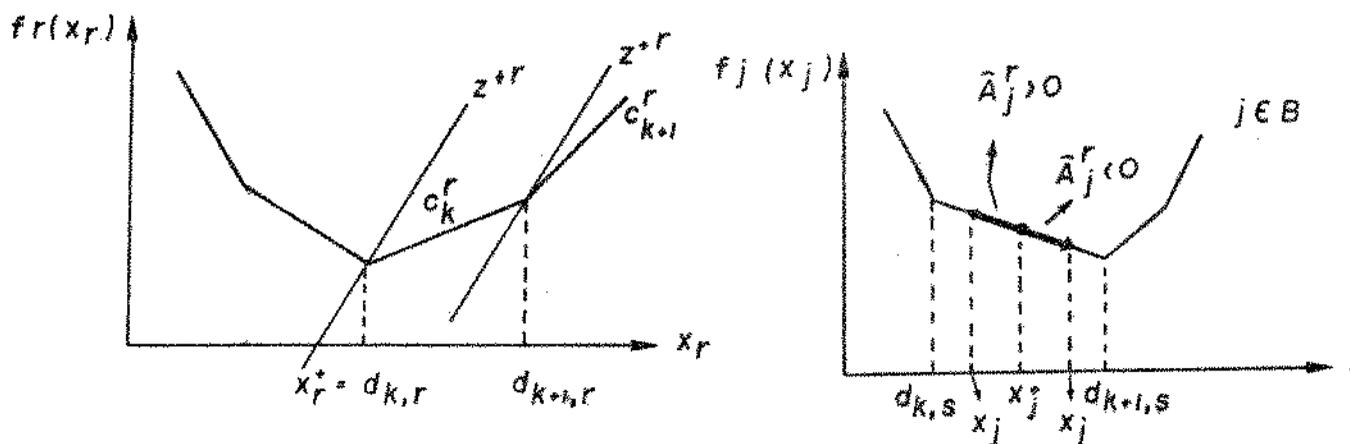
O valor máximo do passo t é dado então por

$$t_{MAX} = \text{MIN} \{t_r, t_j \text{ para } j \in B\} \quad (36)$$

Supondo não degenerescência, ou seja, somente uma das variáveis x_j , $j \in BU(r)$, chega a um ponto crítico, analisemos as duas possibilidades abaixo:

a) $t_{MAX} = t_r$. Podemos construir outra forma canônica cuja diferença em relação a anterior é que a variável não básica x_r assume o valor do ponto crítico imediatamente à direita, temos então:

$$\begin{cases} x_r = d_{k+1,r} \\ x_B = x_B^+ - t_r \tilde{A}^r \in (d_{k,B}; d_{k+1,B}) \end{cases} \quad (37)$$



A inclinação de referência não se altera e mudam apenas os custos relativos da variável x_r (Fig. 8). O novo custo relativo a esquerda é o antigo custo relativo à direita e é portanto negativo. O novo custo relativo à direita é dado por (σ_k^j definido em (3)):

$$\begin{aligned} \hat{c}_{k+1}^r &= c_{k+1}^r - z^{+r} \\ &= c_k^r + \sigma_{k+1}^r - z^{+r} \end{aligned}$$

e então

$$\hat{c}_{k+1}^r = \hat{c}_k^r + \sigma_{k+1}^r \quad (38)$$

Temos então quase que a mesma forma canônica: a única variação se verifica nos custos relativos da variável x_r se tivermos ainda $\hat{c}_k^r < \hat{c}_{k+1}^r < 0$ podemos continuar a aumentar x_r que

F ainda decresce. Se $\hat{c}_{k+1}^r > 0$, qualquer alteração em x_r aumenta F. Testamos as outras variáveis não básicas quanto à otimalidade e, caso não estejamos no ótimo, repetimos o processo.

b) $t_{MAX} = t_s, s \in B$. Os valores das variáveis são dados por

$$\left\{ \begin{array}{l} x_{B-\{s\}} = x_{B-\{s\}}^+ - t_s \hat{A}^r \quad (d_{k,B-\{s\}}; d_{k+1,B-\{s\}}) \\ x_s = \begin{cases} d_{k,s} & \text{se } \hat{A}_s^r > 0 \\ d_{k+1,s} & \text{se } \hat{A}_s^r < 0 \end{cases} \\ x_r = x_r^+ + t_s \end{array} \right. \quad (39)$$

Aqui podemos ter, eventualmente, duas opções:

. A variável x_r entra na base, sai a variável s . Atualizamos B, \bar{B} e $(A^B)^{-1}$. Calculamos a nova inclinação de referência $z^{+\bar{B}}$ e os novos custos relativos à esquerda e à direita das variáveis não básicas. Ou seja, construímos uma nova forma canônica para a nova solução básica factível e repetimos o processo.

. Tentamos continuar a busca unidirecional a condição de que um aumento de t ainda proporcione uma diminuição da função objetivo. Neste caso a variável x_s ainda permanece na base. Ao fazermos isto, x_s passa a um novo intervalo de definição de F com inclinação c_{k-1}^s se $\hat{A}_s^r > 0$ ou c_{k+1}^s se $\hat{A}_s^r < 0$.

Como alteramos a inclinação de uma variável básica temos então novas inclinações de referência para as variáveis não básicas (veja Fig. 10), mas nos interessa somente aquela referente a x_r que para $\hat{A}_s^r > 0$, pode ser calculada por:

$$\begin{aligned} z_{\text{nov}}^{+r} &= \left[c_k^1 \dots c_k^{s-1} c_k^s c_{k-1}^{s+1} c_k^{s+1} \dots c_k^n \right] \hat{A}^r \\ &= \left[c_k^B - (0 \dots 0 \sigma_k^s 0 \dots 0) \right] \hat{A}^r \end{aligned}$$

e portanto

$$z_{\text{nov}}^{+r} = z^{+r} - \sigma_k^s \hat{A}_s^r \quad (40)$$

Para $\tilde{A}_S^r < 0$ o cálculo é semelhante e teremos:

$$z_{\text{nov}}^{+r} = z^{+r} + \sigma_{k+1}^s \tilde{A}_S^s \quad (41)$$

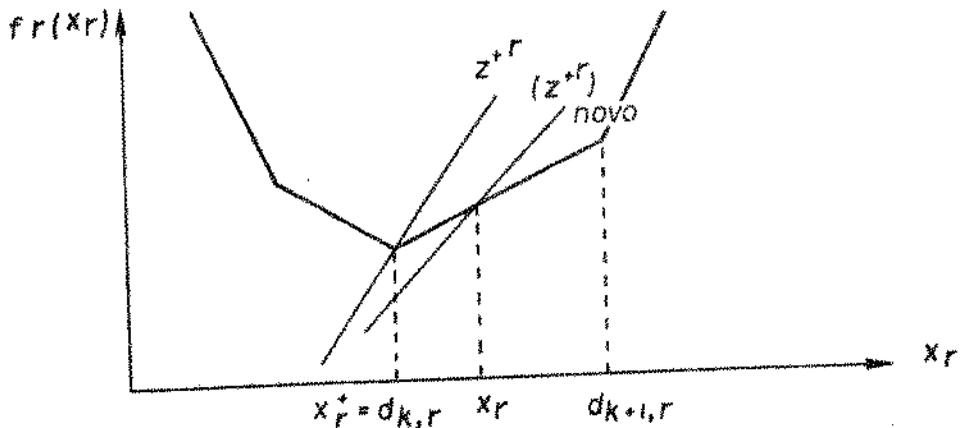
Neste caso, a variável x_r só tem um custo relativo pois pertence ao intervalo $(d_{k,r}; d_{k+1,r})$, dado por

$$(\tilde{c}_k^r)_{\text{nov}} = c_k^r - z_{\text{nov}}^{+r}$$

ou seja

$$(\tilde{c}_k^r)_{\text{nov}} = \begin{cases} \tilde{c}_k^r + \sigma_k^s \tilde{A}_S^r & \text{se } \tilde{A}_S^r > 0 \\ \tilde{c}_k^r - \sigma_{k+1}^s \tilde{A}_S^r & \text{se } \tilde{A}_S^r < 0 \end{cases} \quad (42)$$

Se $(\tilde{c}_k^r)_{\text{nov}} < 0$, aumentamos t até que uma das variáveis x_B ou x_r alcance novo ponto crítico, em caso contrário cessamos a busca unidirecional, x_r entra e x_s sai da base.



Para o caso em que $\tilde{c}_k > \tilde{c}_{k+1} > 0$ o desenvolvimento é análogo.

O PROBLEMA DE PLP COM ESTRUTURA
BLOCO ANGULAR

I. INTRODUÇÃO

O método primal de resolução de um problema de (PL) com estrutura bloco angular é estendido aqui a problemas de PLP com mesma estrutura. No item II apresentamos o problema de (PLP) com estrutura bloco angular e no item III o método de resolução. No item IV apresentamos as conclusões da 3a. parte deste trabalho.

II. O PROBLEMA

O problema de (PL) com estrutura bloco angular pode ser escrito como

$$\begin{aligned} \text{MIN } F(x) &= \sum_{i=1}^{\ell} F_i(x_i) \\ \text{s.a.} \quad &Ax = b \\ &x \geq 0 \end{aligned}$$

ou

$$\left\{ \begin{array}{l} \text{MIN } F(x) = F_1(x_1) + F_2(x_2) + \dots + F_{\ell}(x_{\ell}) \\ \text{s.a.} \quad \begin{bmatrix} A_{01} & A_{02} & \dots & A_{0\ell} \\ A_1 & & & \\ & A_2 & & \\ & & \dots & \\ & & & A_{\ell} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\ell} \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{\ell} \end{bmatrix} \\ x \geq 0 \end{array} \right.$$

e $A_0(m_0, n) = [A_{01} \ A_{02} \ \dots \ A_{0\ell}]$, $A_{0i}(m_0, n_i)$, $x(n, 1)$, $x_i(n_i, 1)$, $A_i(m_i, n_i)$, $b_i(m_i, 1)$ e $n = \sum_{i=1}^{\ell} n_i$. O número de linhas da matriz é $m + m_0$ onde

$$m = \sum_{i=1}^{\ell} m_i$$

Façamos

$$\begin{cases} q_0 = 0 \\ q_k = \sum_{i=1}^k n_i \end{cases} \quad (43)$$

então a função $F_k(x_k)$ pode ser descrita por

$$F_k(x_k) = \sum_{i=q_{k-1}+1}^{q_k} f_j(x_j) \quad (44)$$

onde $f_j(x_j)$ é uma função linear por parte como definida em (2).

Valem aqui todas as considerações feitas para a matriz A no capítulo V e procuraremos generalizar o método primal para problemas de (PL) bloco angular do capítulo VI de forma a resolvermos o problema (P) acima. Usaremos ao longo deste capítulo a notação e terminologia proposta na parte II deste trabalho.

III. RESOLUÇÃO DO PROBLEMA

Vamos supor que, via uma fase I, tenhamos obtido uma solução básica factível $x^{+T} = [x_B^{+T} \ x_{\bar{B}}^{+T}]$ tal que A^B é uma matriz básica de (P). O conjunto B pode ser partido em dois subconjuntos L e I dos índices das colunas chaves e não chaves respectivamente, ou seja, A^B tem uma estrutura bloco angular. Façamos:

$$z^{+B} = P^{+B} A^B = c_K^B \quad (45)$$

donde

$$p^{+B} = c_K^B (A^B)^{-1} \quad (46)$$

e que podemos decompor em:

$$\begin{cases} p_0^{+B} = \bar{c}_K^L (\bar{A}_0^L)^{-1} \\ p^{+B} = (c_K^I - p_0^{+B} A_0^I) (A^I)^{-1} \end{cases} \quad (47)$$

e seja

$$z^{+B} = p^{+B} A^B \quad (48)$$

levando (47) em (48) obtemos:

$$z^{+B} = p_0 \bar{A}_0^B + c_K^I (A^I)^{-1} A^B$$

Podemos então calcular os custos relativos à esquerda e à direita:

$$\begin{aligned} \hat{c}_{K-1}^B &= c_{K-1}^B - z^{+B} \\ &= \bar{c}_{K-1}^B - p_0 \bar{A}^B \end{aligned} \quad (49)$$

$$\begin{aligned} \hat{c}_K^B &= c_K^B - z^{+B} \\ &= \bar{c}_K^B - p_0 \bar{A}_0^B \end{aligned} \quad (50)$$

Se (49) e (50) satisfizerem (31) x^+ é uma solução básica factível ótima, em caso contrário, deveremos procurar outra solução básica,

A variável não básica candidata a entrar na base deve ter ambos os custos relativos positivos ou negativos. Seja x_r a variável escolhida e para efeito do desenvolvimento a seguir suponhamos que seus custos relativos sejam negativos ($\hat{c}_{k-1}^r < \hat{c}_k^r < 0$).

Temos, repetindo (32), que:

$$\begin{bmatrix} x_B \\ x_r \\ x_{B-\{r\}} \end{bmatrix} = \begin{bmatrix} x_B^+ \\ x_r^+ \\ x_{B-\{r\}}^+ \end{bmatrix} + t \begin{bmatrix} \hat{A}_j^r \\ 1 \\ 0 \end{bmatrix} \quad (51)$$

e devemos aumentar t até que uma das variáveis básicas ou x_r chegue a um ponto crítico. O tamanho do passo é calculado através de

. Para a variável não básica, de (33) obtemos:

$$t_r = d_{k+1,r} - d_{k,r} \quad (52)$$

. Para as variáveis básicas, temos dois casos

$\hat{A}_j^r > 0$, então de (34)

$$t_j = \frac{x_j^+ - d_{k,j}}{\hat{A}_j^r} \quad (53)$$

$\hat{A}_j^r < 0$, então de (35)

$$t_j = \frac{x_j^+ - d_{k+1,j}}{\hat{A}_j^r} \quad (54)$$

O valor máximo do passo t é dado então por

$$t_{MAX} = \text{MIN} \{t_r, t_j \text{ para } j \in B\}$$

Temos dois casos a considerar:

a) $t_{MAX} = t_r$, de (37) podemos escrever

$$\begin{cases} x_r = d_{k+1,r} \\ x_B = x_B^+ - t_r \hat{A}_j^r \in (d_{k,B}; d_{k+1,B}) \end{cases} \quad (55)$$

Para a variável x_r o novo custo relativo à esquerda é o antigo custo relativo à direita e é negativo. O novo custo relativo à direita, de (38), é dado por:

$$\hat{c}_{k+1}^r = \hat{c}_k^r + \sigma_{k+1}^r \quad (56)$$

Temos então uma nova solução básica factível. Devemos então testá-la quanto à otimalidade. Se não for ótima, repetimos o processo.

b) $t_{MAX} = t_s$, $s \in B$. Os novos valores das variáveis básicas e x_r , de (39), são dados por:

$$\left\{ \begin{array}{l} x_{B-\{s\}} = x_{B-\{s\}}^+ - t_s \hat{A}_s^r \in (d_{K,B-\{s\}}; d_{K+1,B-\{s\}}) \\ x_s = \begin{cases} d_{k,s} & \text{se } \hat{A}_s^r > 0 \\ d_{k+1,s} & \text{se } \hat{A}_s^r < 0 \end{cases} \\ x_r = x_r^+ + t_s \end{array} \right. \quad (57)$$

Se optarmos por continuar a busca unidirecional, temos necessidade de calcular o novo custo relativo de variável x_r , que como vimos agora é único. De (42) temos

$$(\hat{c}_k^r)_{\text{novo}} = \begin{cases} \hat{c}_k^r + \sigma_k^r \hat{A}_s^r & \text{se } \hat{A}_s^r > 0 \\ \hat{c}_k^r - \sigma_{k+1}^r \hat{A}_s^r & \text{se } \hat{A}_s^r < 0 \end{cases} \quad (58)$$

Se $(\hat{c}_k^r)_{\text{novo}} < 0$, aumentamos t até que uma nova variável alcance novo ponto crítico, em caso contrário, devemos atualizar a matriz básica.

É na atualização da base que especializamos o método para garantir uma estrutura bloco angular para A^B . Como no capítulo VI devemos considerar duas possibilidades:

1. $s \leq m_0$. A^s é tal que $s \in L$, ou seja, é uma coluna não

chave. Devemos atualizar $(\tilde{A}_0^L)^{-1}$, as operações efetuadas são as descritas no caso 1 da página ...?

. $s > m_0$. A^S é tal que $s \in I_h$, ou seja, é uma coluna chave. Temos dois casos a considerar:

. Se A^r é tal que $r \in I_h$, podemos permutar A^S com A^t , $t \in L_h$ e recaímos no caso anterior. Devemos ao permutar A^S e A^t atualizar $(\tilde{A}_0^L)^{-1}$, $(A^{I_h})^{-1}$, L e I como descrito no caso 2-a (pag. ?).

. Se A^r é tal que $r \in I_h$ e $|\hat{A}^{L_h}|_{r \neq 0}$ procedemos como no caso anterior (ver pag. , caso 2-b-1), mas se $|\hat{A}^{L_h}|_{r=0}$ então permutamos diretamente A^r e A^S . Devemos neste caso atualizar somente $(A^{I_h})^{-1}$, I e \bar{B} (ver pag. , caso 2-b-2).

Atualizado $(A^B)^{-1}$, B e \bar{B} , calculamos as novas inclinações de referência e repetimos o processo,

BIBLIOGRAFIA

- [1] GOLSTEIN, E. e YODINE, D. - "Problèmes Particuliers de la Lineaire", Edition Mir, Moscou, 1973.
- [2] LUENGERGER, D.G. - "Optimization by Vector Space Methods", Wiley, N. York, 1969.
- [3] AOKI, M. - "Introduction to Optimization Techniques", Macmillan, N. York, 1971.
- [4] LUENBERGER, D.G. - "Introduction to Linear and Nonlinear Programming", Addison-Wesley, California, 1973.
- [5] ZOULENDIJK, G. - "Mathematical Programming Methods", North-Holland, Amsterdam, 1976.
- [6] LASDON, L.S. - "Optimization Theory for Large Systems", Macmillan, N. York, 1970.
- [7] FERNANDES, J.F.R. - "Programação Linear por Partes: Resolução por Decomposição de um Problema de Grande Porte", Tese de Doutorado, FEC, UNICAMP, 1979.
- [8] GARCIA, A.S. - "Método Dual Simplex para Problemas com Critério Linear por Partes", Tese de Mestrado, IMECC, UNICAMP, 1978.
- [9] SOUZA, C.R. - "Aplicação de Programa Linear por Partes a Sistemas de Potência", Tese de Mestrado, FEC, UNICAMP, 1977.
- [10] AUTHIÉ, G., FERNANDES, J.F.R., FRANÇA, P.M., SOARES, S., TAVARES, H.M.F., - "Commande Optmale Decentralisee pour Systemes Lineaires Discrets à Retards Distribués" - IFAC - Workshop on Large Systems, Industrial Applications, Toulouse, 1977.
- [11] SOARES, S. - "Programação Matemática de Grande Porte: Uma Aplicação a Sistemas Hidrotérmicos de Potência", Tese de Doutorado, FEC, UNICAMP, 1978.

- [24] KNOWLES, T.W. - "An Artificial-Variable Elimination Method for Solving Block Diagonal Programming Problems", Oper. Res., 21, n° 3, 1973.
- [25] GRAVES, G.W., McBRIDE, R.D. - "The Factorization Approach to Large Scale Linear Programming", Math. Programming, Vol. 10, 91-110, 1976.
- [26] GRAVES, G.W. - "A Complete Constructive Algorithm for the General Mixed Linear Programming Problem", Naval Res. Log. Quartely, 1, n° 12, (1965).
- [27] ORCHARD-HAYS, W. - "Factoring LP Block-Angular Bases", Math. Programming, 4, 75-92, 1975.