

PROGRAMAÇÃO LINEAR POR PARTES
MÉTODO PRIMAL-DUAL

Edson de Paula Ferreira

Orientador: Hermano de M.F. Tavares

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
DEZEMBRO DE 1978

UNICAMP
BIBLIOTECA CENTRAL

Agradeço

a todos os amigos que viveram o "outro lado"
da tese,

ao Rey, Akebo, Jurandir, Christiano, França,
Secundino, Raul, Anilton, Elieser, Dodô e
Niura pelo apoio,

ao Hermano, pela paciência.

SUMÁRIO

Desenvolvemos o método Primal-Dual para problemas cuja função objetivo a minimizar é convexa linear por partes, numa formulação baseada na dualidade segundo a teoria de Lagrange.

No capítulo 1 é feita uma nova caracterização do método Primal-Dual linear (DANTZIG, FORD, FULKERSON), e uma apresentação dos aspectos novos no tratamento de problemas lineares por partes.

No capítulo 2 apresentamos o problema linear por partes e fazemos o desenvolvimento do método em questão.

No capítulo 3 resolvemos um problema auxiliar pelo método primal linear por partes (YOUDINE) adequado às características particulares do problema.

No capítulo 4 reapresentamos os métodos Primal e Dual Simplex linear por partes (GARCIA) e fazemos uma análise conjunta nos aspectos qualitativos relevantes dos três métodos em questão.

INDICE

CAPÍTULO 1 - INTRODUÇÃO

1.1 - Primal-Dual Linear	1.1
1.2 - Primal-Dual Linear por Partes	1.5

CAPÍTULO 2 - MÉTODO PRIMAL-DUAL LINEAR POR PARTES (PROBLEMA PRINCIPAL, PP)

2.1 - Formulação do Problema	2.1
2.2 - O Problema Dual	2.4
2.3 - Obtenção da Derivada Direcional	2.9
2.4 - Condições de Otimalidade	2.14
2.5 - Solução Dual - Limites de Variação	2.16
2.6 - Mudança de Solução Dual - Busca Unidimensional	2.20
2.7 - A Nova Iteração	2.27
2.8 - Diagrama de Blocos - Problema Principal	2.29
2.9 - Diagrama de Blocos - Busca Unidimensional	2.32
2.10 - Diagrama de Blocos - Atualização de $\{ t_r \}$	2.34

CAPÍTULO 3 - OBTENÇÃO DA MÁXIMA DERIVADA DIRECIONAL (PROBLEMA AUXILIAR, PA)

3.1 - Introdução	3.1
3.2 - O Problema	3.3

3.3 - O Algoritmo	3.5
3.4 - Condições de Otimalidade	3.7
3.5 - Mudança de Solução Básica - Acréscimo na Variável Independente	3.10
3.6 - Mudança de Solução Básica - Decréscimo na Variável Independente	3.17
3.7 - Diagrama de Blocos	3.19

CAPÍTULO 4 - ANÁLISE QUALITATIVA DE MÉTODOS ALTERNATIVOS

4.1 - Introdução	4.1
4.2 - PLP - Primal Simplex	4.5
4.3 - PLP - Dual Simplex	4.8
4.4 - Primal, Dual, Primal-Dual, Quadro Comparativo	4.12

APÊNDICE - (PAD) É O DUAL DE (PA)

BIBLIOGRAFIA

NOTAÇÃO

$\{ : \}$: denota conjunto

Conjuntos de índices usuais: $Q, \bar{Q}, I, J$

$A(m,n)$: Matriz, m linhas, n colunas

A^j : j -ésima coluna de A

A_i : i -ésima linha de A

A_{ij} : Elemento de A

A^Q : Matriz, $\{A^j : j \in Q\}$

A^I : Matriz, $\{A^i : i \in I\}$, Base

$(A^I)^{-1}$: Inversa de A^I

Π : Matriz Identidade (m,m)

Vetores - Letras minúsculas grifadas

$\underline{b}(m,1)$: Vetor coluna, m elementos

$\underline{c}(1,n)$: Vetor linha, n elementos

$\underline{\lambda}^i, \underline{p}^i$: Vetores i -ésima iteração

b_j, c_j : Elementos de $\underline{b}$ e $\underline{c}$ respectivamente

Vetores especiais: $\underline{x}_Q, \underline{d}_k^Q, c_k^I$

Operações vetoriais: não especifica-se se o vetor é linha ou coluna

$\underline{A}\underline{x}$: Matriz A multiplicado pelo vetor $\underline{x}$

$\underline{p}A$: Vetor multiplicando matriz A

$\underline{e} \cdot \underline{y}$: Produto interno $\langle \underline{e}, \underline{y} \rangle$

Dimensões: Serão compatíveis com A, $\underline{b}$, $\underline{c}$ nas operações e comparações.

$k=i,j$: Índice k variando de i até j

$| \ , \ |$: Denota intervalo fechado

$\mathbb{R}^m$: Espaço vetorial m dimensional, real

CAPÍTULO 1: INTRODUÇÃO

O objetivo básico deste trabalho é o desenvolvimento de um método primal-dual para problemas, cuja função a minimizar é convexa e linear por partes. A formulação do método é voltada para uma melhor caracterização de seus aspectos qualitativos e a análise de métodos alternativos.

1.1 - Primal-Dual Linear

Dado o problema:

$$\begin{array}{ll} \text{MINIMIZAR} & F(\underline{x}) = \underline{c} \underline{x} \\ \text{S.a:} & A\underline{x} = \underline{b} \\ & \underline{x} \geq \underline{0} \end{array} \quad (P)$$

e seu dual

$$\begin{array}{ll} \text{MAXIMIZAR} & \phi(\underline{\lambda}) = \underline{\lambda} \underline{b} \\ \text{S.a:} & \underline{\lambda} A \leq \underline{c} \end{array} \quad (D)$$

Classicamente o método primal-dual é caracterizado como o que procede a partir de uma solução factível de (D), $\underline{\lambda}^0$, que, é "melhorada", a cada iteração, pela resolução de um problema primal restrito associado. Após um número finito de "melhoras", obtemos a solução ótima de (P) (SIMONNARD).

A idéia ao se desenvolver o método primal-dual, foi a de atingir, numa única fase, a factibilidade e a otimalidade; daí a construção do problema primal restrito ser feita sobre estes dois aspectos:

- Factibilidade:

É procurada como na fase inicial do algoritmo Simplex, acrescentando-se as variáveis artificiais $\underline{y}$, que são penalizadas numa função objetivo, obtendo-se:

$$\text{MINIMIZAR } h(\underline{y}) = \underline{e} \underline{y} \quad \underline{e} > 0$$

$$\text{S.a: } \underline{A}\underline{x} + \underline{y} = \underline{b}$$

$$\underline{y} > 0, \quad \underline{x} > 0$$

onde $\underline{e}$ pondera as componentes de $\underline{y}$.

- Otimalidade:

É assegurada de acordo com o teorema das folgas complementares:

Teorema das Folgas Complementares: Sendo $\underline{x}^*$ e $\underline{\lambda}^*$ soluções factíveis de (P) e (D), respectivamente, elas também serão ótimas se:

$$(\underline{c} - \underline{\lambda}^* \underline{A}) \underline{x}^* = 0 \quad (1.1)$$

Na i -ésima iteração, estando numa solução factível de (D), $\underline{\lambda}^i$, e definindo:

$$Q = \{j: \underline{\lambda}^i A^j = c_j\} \quad (1.2)$$

As variáveis x_j ficam divididas em dois conjuntos, $\underline{x}_Q$ e $\underline{x}_{\bar{Q}}$. Fazendo:

$$\underline{x}_{\bar{Q}} = \underline{0} \quad (1.3)$$

(1.1) é satisfeita para todo x_j .

Assim o problema primal restrito associado, que denominaremos Problema Auxiliar, fica definido como:

$$\begin{aligned} & \text{MINIMIZAR } h(\underline{y}) = \underline{e} \underline{y} \quad \underline{e} > 0 \\ & \text{S.a: } A^Q \underline{x}_Q + \underline{y} = \underline{b} \quad (\text{PA}) \\ & \underline{y} \geq 0, \quad \underline{x}_Q \geq 0 \end{aligned}$$

Sendo (PA) resolvido pelo algoritmo Simplex e $(\underline{x}_Q^*, \underline{y}^*)$ sua solução ótima, se $\underline{y}^* = 0$, $(\underline{x}_Q^*, \underline{x}_{\bar{Q}})$ é factível de (P) e satisfaz (1.1); pelo teorema das folgas complementares, $(\underline{x}_Q^*, \underline{x}_{\bar{Q}})$ é ótimo de (P).

Se $\underline{y}^* \neq 0$, sendo $\underline{u}^*$ o vetor multiplicador associado à solução ótima de (PA), modificamos a solução de (D), $\underline{\lambda}^i$, de acordo com:

$$\underline{\lambda}^{i+1} = \underline{\lambda}^i + t \underline{u}^*$$

onde t é tal que a factibilidade de (D) seja preservada.

Factibilidade de (P): se $\underline{u}^* A^j < 0$ para todo $j \in \bar{Q}$, t pode crescer arbitariamente e $\phi(\underline{\lambda}) \rightarrow \infty$, então (P) não tem solução factível.

Com a caracterização do método acima descrito, não podemos inferir uma qualidade a $\underline{u}^*$, além da simples verificação de que $\phi(\underline{\lambda})$ cresce segundo $\underline{u}^*$, o que ocorre para qualquer solução factível do problema dual de (PA), do qual $\underline{u}^*$ é solução ótima:

$$\text{MAXIMIZAR } \psi(\underline{u}) = \underline{u} \underline{b}$$

$$\text{S.a: } \underline{u} A^Q \leq 0 \quad (\text{PAD})$$

$$\underline{u} \leq \underline{e}$$

Para uma melhor caracterização do método, adotamos então o seguinte critério:

A partir de $\underline{\lambda}^0$, solução factível de (D), a cada iteração, mantendo a factibilidade de (D), mudamos de solução segundo a direção de máximo crescimento de $\phi(\underline{\lambda})$.

Desta maneira, o desenvolvimento do método transcorre "naturalmente", aparecendo (PAD) na procura da direção de máximo crescimento de $\phi(\underline{\lambda})$. Operacionalmente, resolvemos o seu dual (PA) pelo método primal linear Simplex e com o seu vetor multiplicador ótimo $\underline{u}^*$, fazemos a "melhor" implementação na solução do dual $(\underline{\lambda}^i)$.

1.2 - Primal-Dual Linear por Partes

Dado o problema:

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

(P)

$$\text{S.a: } \underline{Ax} = \underline{b}$$

$$\underline{\alpha} \leq \underline{x} \leq \underline{\beta}$$

- $A(m,n)$ de "rank" m , $f_j(x_j)$ convexa linear por partes.

O seu dual $\tilde{e}$:

$$\text{MAXIMIZAR } \phi(\underline{p})$$

(D)

$$\underline{p} \in \mathbb{R}^m$$

- $\phi(\underline{p})$ convexa, linear por partes, definida para todo $\underline{p}$.

Como alguns aspectos novos na abordagem do problema linear por partes, podemos citar:

- A dualidade $\tilde{e}$ estabelecida em termos da teoria de Lagrange (LAS-DON).

- Podemos inicializar a otimização com um $\underline{p}$ arbitrário.

- No problema auxiliar, a penalização das variáveis artificiais é feita segundo:

$$h(\underline{y}) = \sum_{j=1}^m e_j |y_j| \quad e_j > 0$$

do que resulta $\underline{y}$ irrestrito e o problema auxiliar linear por partes.

- Na análise de otimalidade o teorema das folgas complementares é generalizado para:

$\underline{x}^*$ e $\underline{p}^*$, soluções factíveis de (P) e (D) respectivamente, são ótimas se $\underline{x}^*$ minimiza a função Lagrangeano $L(\underline{x}, \underline{p}^*)$.

- Sendo separável a minimização do Lagrangeano, definimos "curva de otimalidade" associada à variável x_j , $(CO)_j$, como o lugar geométrico dos pontos $(x_j, \underline{p}^j)$ que minimizam $L(\underline{x}, \underline{p})$.

- Introduzimos a norma "retangular"

$$\| \underline{u} \| = \text{MÍNIMO} \left\{ q: -qe \leq \underline{u} \leq qe, \quad e_j > 0, \quad q > 0 \right\}$$

que permitiu formalizar a direção de máxima taxa de crescimento de $\phi(\underline{p})$, como de "derivada direcional" máxima.

- Na mudança de solução dual, $\underline{p}^{i+1} = \underline{p}^i + t\underline{u}^*$, a obtenção de t corresponde a uma busca unidimensional em $\phi(\underline{p})$ (linear por partes), cujo critério de parada envolve não mais a dual factibilidade, mas o crescimento contínuo de $\phi(\underline{p})$ com t .

- Estruturalmente, dividimos o método em duas partes denominadas Problema Principal (PP) e Problema Auxiliar (PA). Sendo:

(PP) - Busca Unidimensional, obtém nova solução $\underline{p}^{i+1}$.

(PA) - Algoritmo Primal, obtém direção de máximo crescimento $\underline{u}^*$.

(PA)

(PP)

$$\underline{p}^0 \longrightarrow \underline{u}^* \longrightarrow \underline{p}^{i+1} = \underline{p}^i + t\underline{u}^*$$

- Podemos então caracterizar o método como do tipo STEEPEST-
 -DESCENT ou "GRADIENTE ÓTIMO", para problemas lineares por partes.

- Se o problema (P) tem solução factível, o ótimo será uma solução tipo (α) e nunca tipo (β), porque as restrições definem um conjunto compacto (fechado e limitado); caso contrário a infactibilidade é detectada.

CAPÍTULO 2: MÉTODO PRIMAL-DUAL LINEAR POR PARTES (PROBLEMA PRINCIPAL, PP)

2.1 - Formulação do Problema

Uma função convexa linear por partes, pode ser formalmente definida por:

$$F(\underline{x}) = \max_{i \in I} \ell_i(\underline{x})$$

com $\ell_i(\underline{x}) = H_i \underline{x} - h_i$, i -ésima linha do sistema linear:

$$\underline{\ell}(\underline{x}) = H\underline{x} - \underline{h}$$

A programação linear por partes consiste no conjunto de métodos para o cálculo do mínimo duma função convexa linear por partes, definida num poliedro convexo.

Nosso problema pode então ser caracterizado genericamente por:

$$\text{MIN } F(\underline{x})$$

$$\text{S.a: } \underline{G}(\underline{x}) \leq \underline{0}$$

$$A\underline{x} = \underline{b}$$

$$\underline{\alpha} \leq \underline{x} \leq \underline{\beta}$$

Neste trabalho abordaremos a formulação particular:

Determine o valor de $\underline{x}$ que resolve,

$$\begin{aligned} \text{MIN} \quad F(\underline{x}) &= \sum_{i=1}^n f_j(x_j) \\ \text{S.a:} \quad A\underline{x} &= \underline{b} \\ \underline{\alpha} &\leq \underline{x} \leq \underline{\beta} \end{aligned} \quad (P)$$

sendo:

- $f_j(x_j)$ convexa linear por partes.
- A (m , n); $\underline{b}$ (m , 1); $\underline{x}$, $\underline{\alpha}$, $\underline{\beta}$ (n , 1).
- a matriz A (m , n) de "rank" m .
- (P) denominado Primal.

Representamos graficamente as $f_j(x_j)$:

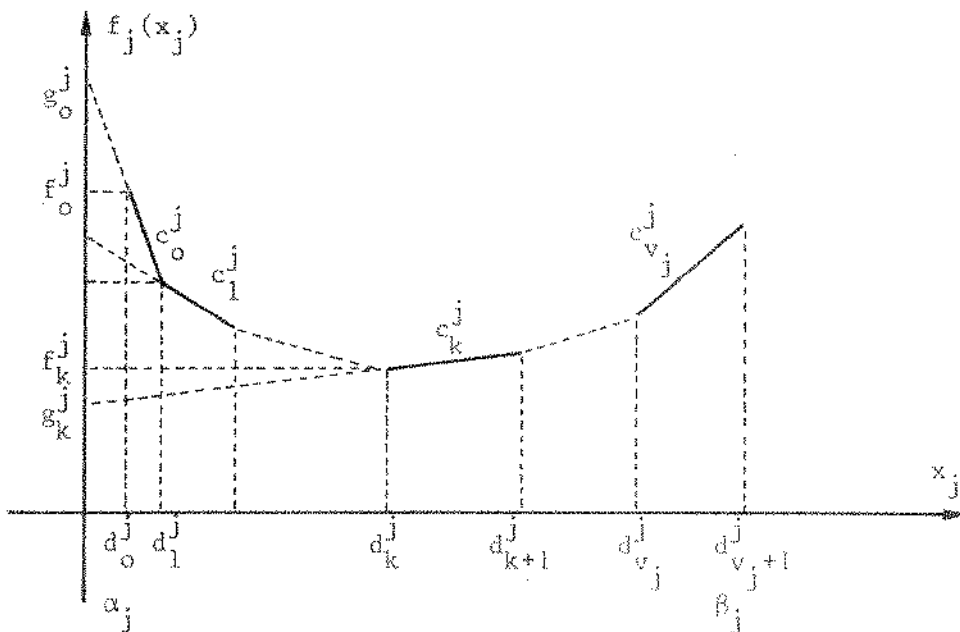


Figura 2.1

Na figura 2.1:

- v_j determina o número de intervalos lineares segundo x_j .
- c_k^j é a inclinação do intervalo de ordem k .
- d_k^j é o ponto crítico de $f_j(x_j)$, limite inferior do intervalo de ordem k .

Sendo $f_0^j = f_j(d_0^j)$, para $k = 1, v_j + 1$:

$$f_k^j = c_{k-1}^j (d_k^j - d_{k-1}^j) + f_{k-1}^j \quad (2.1)$$

Escrevemos para x_j no intervalo $|d_k^j, d_{k+1}^j|$, $k = 0, v_j$:

$$f_j(x_j) = c_k^j (x_j - d_k^j) + f_k^j \quad (2.2)$$

Sendo,

$$g_k^j = f_k^j - c_k^j d_k^j \quad (2.3)$$

obtemos para $x_j \in |d_k^j, d_{k+1}^j|$, $k = 0, v_j$:

$$f_j(x_j) = c_k^j x_j + g_k^j \quad (2.4)$$

Na figura 2.1:

- v_j determina o número de intervalos lineares segundo x_j .
- c_k^j é a inclinação do intervalo de ordem k .
- d_k^j é o ponto crítico de $f_j(x_j)$, limite inferior do intervalo de ordem k .

Sendo $f_0^j = f_j(d_0^j)$, para $k = 1, v_j + 1$:

$$f_k^j = c_{k-1}^j (d_k^j - d_{k-1}^j) + f_{k-1}^j \quad (2.1)$$

Escrevemos para x_j no intervalo $|d_k^j, d_{k+1}^j|$, $k = 0, v_j$:

$$f_j(x_j) = c_k^j (x_j - d_k^j) + f_k^j \quad (2.2)$$

Sendo,

$$g_k^j = f_k^j - c_k^j d_k^j \quad (2.3)$$

obtemos para $x_j \in |d_k^j, d_{k+1}^j|$, $k = 0, v_j$:

$$f_j(x_j) = c_k^j x_j + g_k^j \quad (2.4)$$

Definindo,

$$z_j = \underline{p} \Lambda^j \quad (2.6)$$

z_j denominado inclinação de referência, reescrevemos (2.5):

$$L(\underline{x}, \underline{p}) = \underline{p} \underline{b} + \sum_{j=1}^n (f_j(x_j) - z_j x_j) \quad (2.7)$$

A função dual $\tilde{\phi}$ é definida por:

$$\phi(\underline{p}) = \underset{\alpha \leq x \leq \beta}{\text{MINIMO}} L(\underline{x}, \underline{p}) \quad (2.8)$$

No decorrer do trabalho, mostramos que $\phi(\underline{p})$ é côncava, linear por partes, definida para todo $\underline{p}$ em R^m .

Definindo:

$$Q = \left\{ j : z_j = c_k^j \text{ para algum } k \in \{0, \dots, v_j\} \right\} \quad (2.9)$$

$$\bar{Q} = \{1, \dots, n\} - Q$$

Podemos escrever $\bar{Q}$ como a união dos conjuntos abaixo:

$$\begin{aligned}
 - \quad j: \quad c_{k-1}^j < z_j < c_k^j \quad \text{para algum } k \in \{1, \dots, v_j\} \\
 - \quad j: \quad z_j < c_0^j \\
 - \quad j: \quad z_j > c_{v_j}^j
 \end{aligned} \tag{2.10}$$

(*) Note que a partição $\{\bar{Q}, Q\}$ é função do p em questão.

No cálculo da função dual, a minimização em (2.8) é separável conforme:

$$- \quad j \in \bar{Q}$$

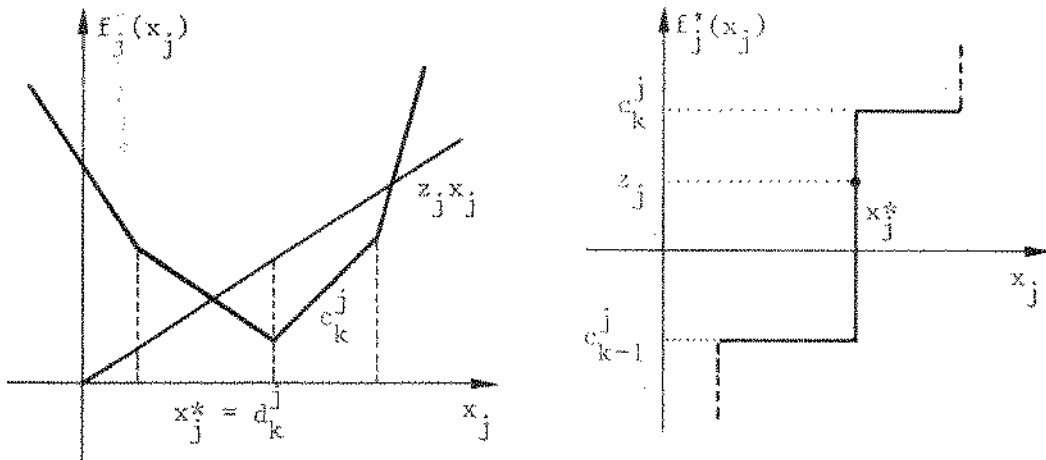


Figura 2.3

$(f_j(x_j) - z_j x_j)$ tem seu mínimo em $x_j^* = d_k^j$

Se $z_j < c_0^j$ ou $z_j > c_{v_j}^j$, o mínimo ocorrerá respectivamente em $x_j = \alpha_j$ ou $x_j = \beta_j$.

- $j \in Q$

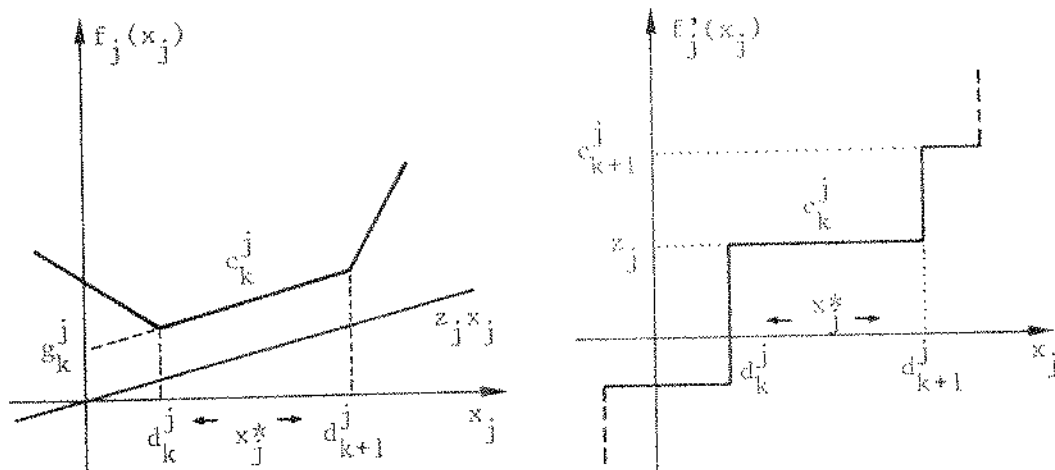


Figura 2.4

$(f_j(x_j) - z_j x_j)$ tem seu mínimo igual a g_k^j em qualquer ponto do intervalo $[d_k^j, d_{k+1}^j]$.

Do exposto acima concluímos que, calculado z_j para dado $\underline{p}$, x_j que otimiza o Lagrangeano $\tilde{e}$ o correspondente a z_j sobre a curva de otimalidade abaixo:

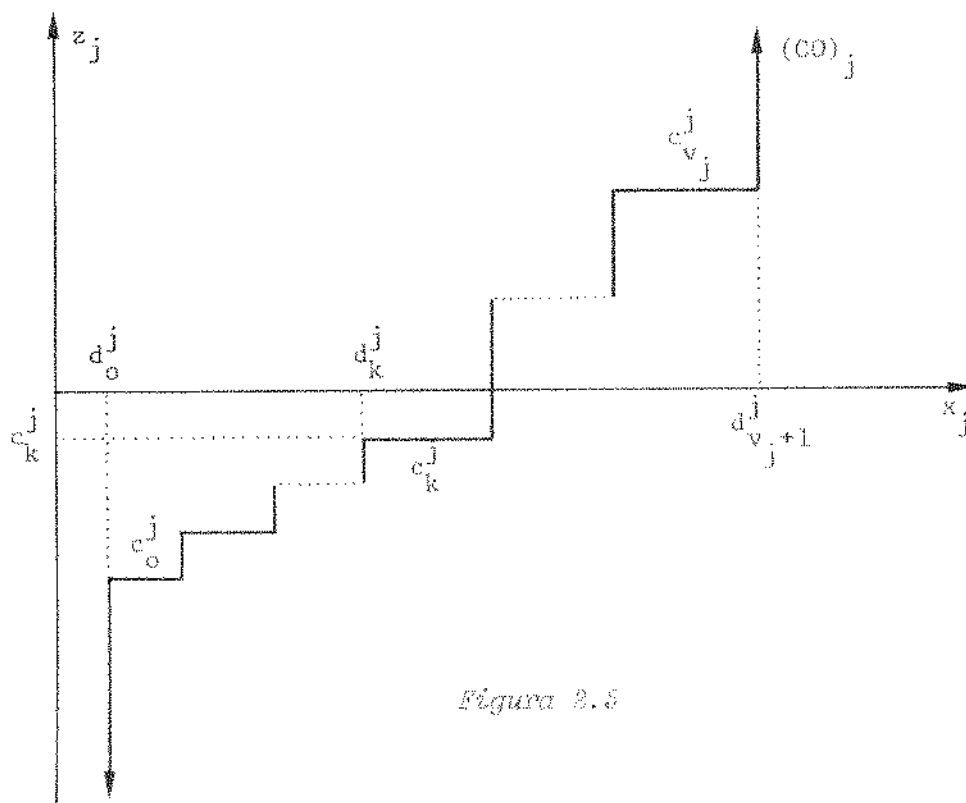


Figura 2.5

Obtemos o valor da função dual:

$$\phi(\underline{p}) = \underline{p} \underline{b} + \sum_{j \in \bar{Q}} (f_j(d_k^j) - z_j d_k^j) + \sum_{j \in Q} g_k^j \quad (2.11)$$

O problema dual é definido por:

$$\begin{array}{ll} \text{MAX} & \phi(\underline{p}) \\ \underline{p} & \end{array} \quad (2.12)$$

Como $\phi(\underline{p})$ é obtida numa minimização em $\underline{x}$ canalizada, $\phi(\underline{p})$ existe para qualquer $\underline{p}$ finito.

Resulta o problema dual:

$$\begin{array}{ll} \text{MAX} & \underline{p} \underline{b} + \sum_{j \in \bar{Q}} (f_j(d_k^j) - z_j d_k^j) + \sum_{j \in Q} g_k^j \\ \underline{p} & \end{array} \quad (D)$$

$\underline{p}$ qualquer

2.3 - Obtenção da Derivada Direcional

Considerando (P) e (D) definidos anteriormente.

Como (P) é convexo não há intervalo de dualidade entre o "par dual". Se $\underline{x}^*$ e $\underline{p}^*$ são as respectivas soluções ótimas, então $F(\underline{x}^*) = \phi(\underline{p}^*)$.

Começamos nossa abordagem pelo problema dual, com uma solução arbitrária $\underline{p}^0$. (D) irrestrito.

A partir de $\underline{p}^0$ temos que tentar aumentar $\phi(\underline{p})$. Se não for possível, $\underline{p}^0$ é ótimo, se for, vamos procurar fazê-lo na direção que fornece máxima derivada direcional.

A derivada direcional em questão refere-se à norma "retangular":

$$\|\underline{u}\| = \text{MIN} \left\{ q: -qe \leq \underline{u} \leq qe, \quad e > 0, \quad q > 0 \right\}$$

Seja $\underline{p} = \underline{p}^0 + t\underline{u}$, onde:

- t é suficientemente pequeno.
- $\|\underline{u}\| = 1$.
- z_j , $\bar{Q}$ e Q são relativos a $\underline{p}^0$.

Escrevemos o Lagrangeano de (P) em $\underline{p}^0 + t\underline{u}$:

$$\begin{aligned} L(\underline{x}, \underline{p}^0 + t\underline{u}) &= (\underline{p}^0 + t\underline{u})\underline{b} + \sum_{j \in \bar{Q}} \left\{ f_j(x_j) - (\underline{p}^0 + t\underline{u}) A^j x_j \right\} + \\ &+ \sum_{j \in Q} \left\{ f_j(x_j) - (\underline{p}^0 + t\underline{u}) A^j x_j \right\} \end{aligned} \quad (2.13)$$

Supondo que $\bar{Q}$ não seja alterado (primeira restrição a t):

$$\begin{aligned} \phi(\underline{p}^0 + t\underline{u}) &= (\underline{p}^0 + t\underline{u})\underline{b} + \sum_{j \in \bar{Q}} \left\{ f_j(d_k^j) - (\underline{p}^0 + t\underline{u})A^j d_k^j \right\} + \\ &+ \underset{\substack{\text{MINIMO} \\ \underline{\alpha}_Q \leq x_Q \leq \underline{\beta}_Q}}{\sum_{j \in Q}} \left\{ f_j(x_j) - (\underline{p}^0 + t\underline{u})A^j x_j \right\} \end{aligned} \quad (2.14)$$

A minimização em (2.14) é separável, definindo para $j \in Q$:

$$w_j = \underline{u} A^j \quad (2.15)$$

$$Q_1 = \{j: w_j = 0\} \quad (2.16)$$

$$Q_2 = \{j: w_j > 0\} \quad (2.17)$$

$$Q_3 = \{j: w_j < 0\} \quad (2.18)$$

Assim, considerando

$$\underset{\substack{\text{MINIMO} \\ \underline{\alpha}_Q \leq x_Q \leq \underline{\beta}_Q}}{\sum_{j \in Q}} f_j(x_j) - (z_j + tw_j)x_j \quad (2.19)$$

- Se $j \in Q_1$, z_j não se altera e $x_{Q_1}^*$ continua livre no intervalo

$$[d_k^{Q_1}, d_{k+1}^{Q_1}] \quad (\text{figura 2.4}).$$

- Se $j \in Q_2$, z_j aumenta e o valor $x_j^* = d_{k+1}^j$ minimiza em 2.19.

Graficamente:

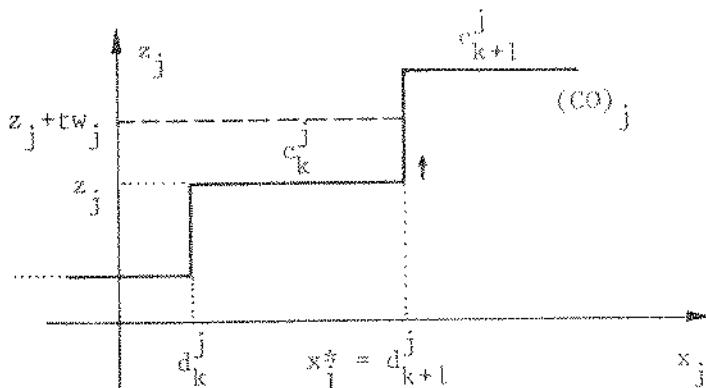


Figura 2.6

Supomos que z_j não ultrapasse c_{k+1}^j (segunda restrição a t).

- Se $j \in Q_3$, z_j diminui e o valor $x_j^* = d_k^j$ minimiza em 2.19. Gra_{ficamente}:

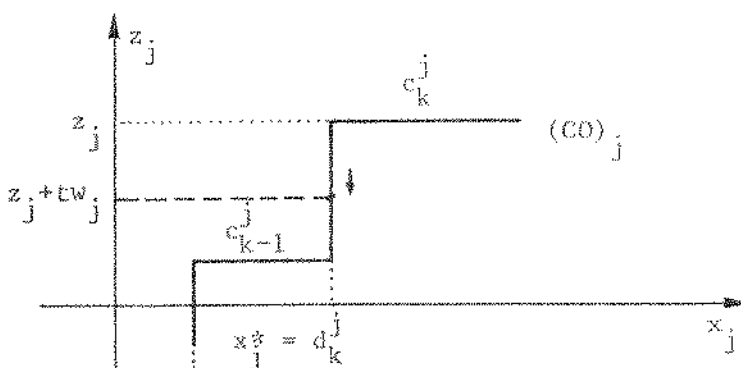


Figura 2.7

Supomos que z_j não ultrapasse c_{k-1}^j (terceira restrição a t).

Assim a equação (2.19) resulta em:

$$g_k^j + \sum_{j \in Q_2} \left\{ g_k^j - t_{\underline{u}} A^j d_{k+1}^j \right\} + \sum_{j \in Q_3} \left\{ g_k^j - t_{\underline{u}} A^j d_k^j \right\} \quad (2.20)$$

Substituindo (2.20) em (2.14):

$$\begin{aligned} \phi(\underline{p}^o + t_{\underline{u}}) &= \underline{p}^o \underline{b} + \sum_{j \in \bar{Q}} \left\{ f_j(d_k^j) - \underline{p}^o A^j d_k^j \right\} + \sum_{j \in Q} g_k^j + \\ &+ t_{\underline{u}} \underline{b} - \sum_{j \in \bar{Q}} t_{\underline{u}} A^j d_k^j - \sum_{j \in Q_2} t_{\underline{u}} A^j d_{k+1}^j + \\ &- \sum_{j \in Q_3} t_{\underline{u}} A^j d_k^j \end{aligned} \quad (2.21)$$

$$\phi(\underline{p}^o + t_{\underline{u}}) = \phi(\underline{p}^o) + t_{\underline{u}}(\underline{b} - A^{\bar{Q}} \bar{Q} - A^{Q_2} d_{k+1}^{Q_2} - A^{Q_3} d_k^{Q_3}) \quad (2.22)$$

$$\frac{\phi(\underline{p}^o + t_{\underline{u}}) - \phi(\underline{p}^o)}{t} = \underline{u}(\underline{b} - A^{\bar{Q}} \bar{Q} - A^{Q_2} d_{k+1}^{Q_2} - A^{Q_3} d_k^{Q_3}) \quad (2.23)$$

Considerando (2.23), definimos como a derivada direcional de $\phi(\underline{p})$ na direção $\underline{u}$:

$$\psi(\underline{u}) = \underline{u}(\underline{b} - A \frac{\bar{Q}}{d_k} \bar{Q} - A \frac{Q_2}{d_{k+1}} Q_2 - A \frac{Q_3}{d_k} Q_3) \quad (2.24)$$

Verifica-se que para $\underline{u}$, $\bar{Q}$, Q_1 , Q_2 e Q_3 definidos como anteriormente, $\psi(\underline{u})$ é solução do problema:

$$\psi(\underline{u}) = \underset{d_k^Q \leq x_Q \leq d_{k+1}^Q}{\text{MINIMO}} \left\{ \underline{u}(\underline{b} - A \frac{\bar{Q}}{d_k} \bar{Q} - A \frac{Q}{x_Q}) \right\} \quad (2.25)$$

Para acharmos a direção que fornece a "máxima" derivada direcional, resolvemos:

$$\begin{array}{ll} \text{MAXIMIZAR} & \psi(\underline{u}) \\ \text{s.t.} & \|\underline{u}\| = 1 \end{array}$$

Nosso problema resulta em:

$$\begin{array}{ll} \text{MAX}_{\underline{u}} & \left\{ \text{MIN}_{\frac{x_Q}{d_k}} \left\{ \underline{u}(\underline{b} - A \frac{\bar{Q}}{d_k} \bar{Q} - A \frac{Q}{x_Q}) \right\} \right\} \\ & \text{(PAD)} \end{array}$$

$$\text{s.t.} \quad \|\underline{u}\| = 1$$

$$d_k^Q \leq x_Q \leq d_{k+1}^Q$$

Prova-se no Apêndice 1 que o problema anterior (PAD) é o dual de:

$$\begin{aligned} \text{MIN} \quad h(\underline{y}) &= \sum_{j=1}^m e_j |y_j| \\ \text{S.a:} \quad A_{\underline{x}_Q}^Q + \underline{y} &= \underline{b} - A_{\underline{d}_k}^{\overline{Q}} \end{aligned} \quad (\text{PA})$$

$$d_k^Q \leq \underline{x}_Q \leq d_{k+1}^Q$$

denominado Problema Auxiliar de (P).

2.4 - Condições de Otimalidade

Seja $(\underline{d}_k^{\overline{Q}}, \underline{x}_Q^*, \underline{y}^*)$ solução ótima do (PA):

Se $\underline{y}^* = 0$, então $(\underline{d}_k^{\overline{Q}}, \underline{x}_Q^*)$ é solução ótima de (P)

(P) corresponde a minimizar $F(\underline{x})$ sujeito às restrições:

$$- A_{\underline{x}_Q}^Q = \underline{b} - A_{\underline{x}_Q}^{\overline{Q}}$$

$$- \underline{\alpha} \leq \underline{x} \leq \underline{\beta}$$

Se $\underline{y}^* = 0$ obviamente $(\underline{d}_k^{\overline{Q}}, \underline{x}_Q^*)$ é factível de (P).

Sendo $h(\underline{y}^*) = 0$ e se $\underline{u}^*$ é solução ótima de (PAD):

$$\phi(\underline{u}^*) = 0$$

ou

$$\frac{\phi(\underline{p}^* + t\underline{u}^*) - \phi(\underline{p}^*)}{t} = 0$$

onde $\underline{p}^*$ é a solução do dual (D) correspondente a $(\underline{d}_k^Q, \underline{x}_Q^*)$.

Concluimos que:

$$\phi(\underline{p}^* + t\underline{u}^*) = \phi(\underline{p}^*)$$

onde $\underline{u}^*$ é a direção de máximo crescimento. Podemos então, dizer que para $\underline{u}$ arbitrário:

$$\phi(\underline{p}^* + t\underline{u}) \leq \phi(\underline{p}^*)$$

donde $\underline{p}^*$ é um máximo local de $\phi(\underline{p})$. Como $\phi(\underline{p})$ é côncava, $\underline{p}^*$ é máximo global. A solução correspondente $\underline{x}^*$, é ótima do primal (P).

Concluindo:

- Se $\underline{y}^* = \underline{0}$, o problema em questão (P) está resolvido.
- Se $\underline{y}^* \neq \underline{0}$, calculamos $\underline{u}^*$, direção de máxima derivada direcional, e alteramos $\underline{p}$ na direção $\underline{u}^*$:

$$\underline{p}^{i+1} = \underline{p}^i + t\underline{u}^* \quad (2.26)$$

(*) i -ésima iteração.

A função $\phi(\underline{p})$ realmente cresce na direção $\underline{u}^*$. Sendo $\underline{y}^* \neq 0$:

$$\psi(\underline{u}^*) = \sum_{j=1}^m e_j |y_j^*| > 0$$

então

$$\phi(\underline{p}^i + t\underline{u}^*) - \phi(\underline{p}^i) = t\psi(\underline{u}^*) > 0$$

e

$$\phi(\underline{p}^{i+1}) > \phi(\underline{p}^i)$$

(2.27)

(*) Para t conveniente.

2.5 - Solução Dual - Limites de Variação

Fazemos um apanhado das restrições impostas a t , a fim de determinar a amplitude do intervalo de variação de $\underline{p}$ na direção $\underline{u}^*$, no qual o desenvolvimento anterior é válido.

- Primeira restrição

$\bar{Q}$ não se altera. Temos que garantir que as restrições (2.10) não sejam violadas. $\bar{Q}$ é alterado com a ocorrência de uma das situações abaixo:

$$- w_j > 0 \quad e \quad z_j + tw_j = c_k^j$$

$$- w_j < 0 \quad e \quad z_j + tw_j = c_{k-1}^j$$

com $k \in \{0, \dots, v_j+1\}$.

Graficamente:

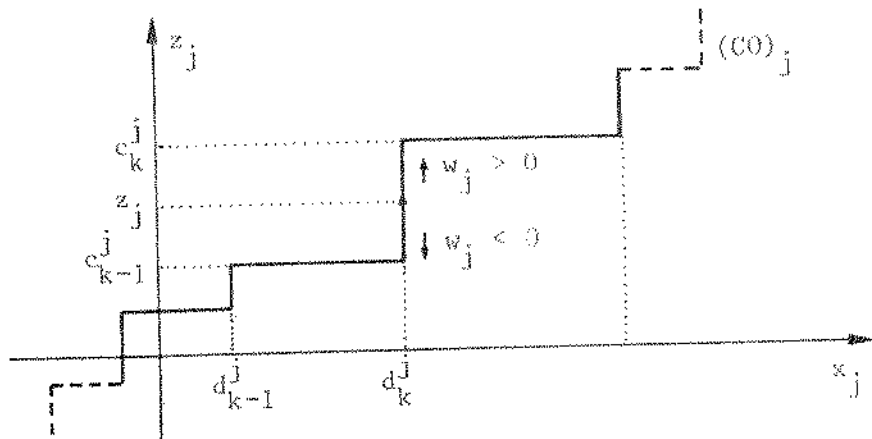


Figura 2.8

Para este caso definimos então os limites de t :

$$t(j) = \frac{c_k^j - z_j}{w_j}, \quad w_j > 0, \quad k \in \{0, \dots, v_j\} \quad (2.28)$$

$$t(j) = \frac{c_{k-1}^j - z_j}{w_j}, \quad w_j < 0, \quad k \in \{1, \dots, v_j+1\} \quad (2.29)$$

~ Segunda restrição

Para $j \in Q$ e $w_j > 0$, temos que manter $z_j < c_{k+1}^j$.

$$t(j) = \frac{c_{k+1}^j - c_k^j}{w_j}, \quad k \in \{0, \dots, v_j-1\} \quad (2.30)$$

- Terceira restrição

Para $j \in Q$, $w_j < 0$ temos que manter $z_j > c_{k-1}^j$.

$$t(j) = \frac{c_{k-1}^j - c_k^j}{w_j}, \quad k \in \{1, \dots, v_j\} \quad (2.31)$$

Graficamente:

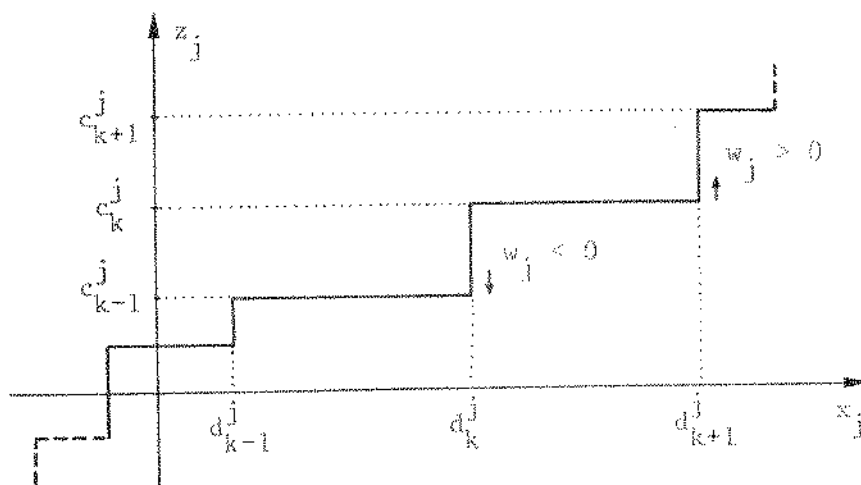


Figura 2.9

Seja t^* a amplitude do intervalo em questão:

$$t^* = \min_j \{ t(j) \}$$

Note que na definição de $t(j)$ para cada caso, se estamos em alguma extremidade do intervalo de variação de x_j , conforme o sinal de w_j , t não é limitado, isto é, podemos crescê-lo arbitrariamente. São os casos:

- $j \in \bar{Q}$:

$$w_j > 0 \quad \text{e} \quad k = v_j + 1 \quad (x_j = \beta_j)$$

$$w_j < 0 \quad \text{e} \quad k = 0 \quad (x_j = \alpha_j)$$

- $j \in Q$:

$$w_j > 0 \quad \text{e} \quad k = v_j$$

$$w_j < 0 \quad \text{e} \quad k = 0$$

(*) $j \in \bar{Q}$, k refere-se a ponto crítico.

$j \in Q$, k refere-se a inclinação.

Definimos:

$$t(j) = \infty$$

(2.32)

Graficamente:

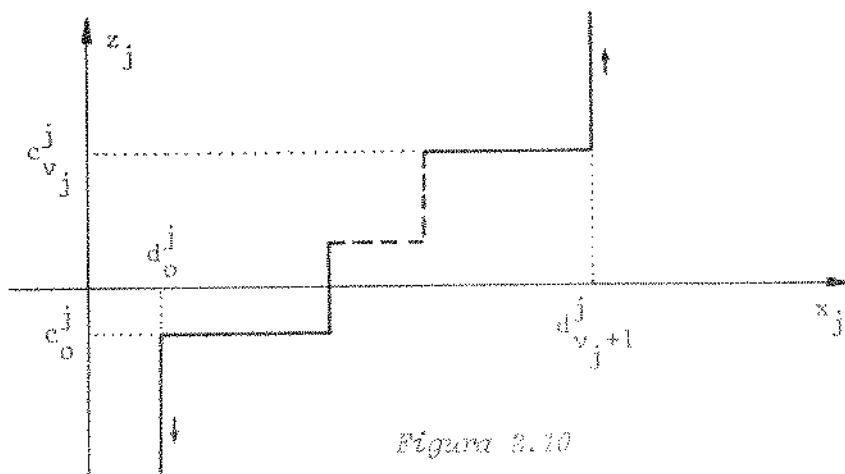


Figura 2.10

Assim, se $\{j: t(j) \text{ é definido}\}$ for vazio, $t^* = \infty$ e o valor da função dual cresce arbitrariamente pois:

$$\phi(\underline{p}^i + t\underline{u}^*) - \phi(\underline{p}^i) = t\psi(\underline{u}^*)$$

onde t pode ser feito arbitrariamente grande. Concluímos que o problema primal NÃO TEM SOLUÇÃO FACTÍVEL (LASDON).

2.6 - Mudança de Solução Dual - Busca unidimensional

Até agora caminhamos na direção $\underline{u}^*$ até atingirmos o limite t^* , garantindo o crescimento de $\phi(\underline{p})$. Com a nova solução do dual obtida, voltaríamos ao (PP), atualizaríamos Q e $\bar{Q}$..., e calcularíamos a derivada direcional máxima nesta nova "face", resolvendo outro (PA). Antes disso, vamos considerar a possibilidade de, ultrapassando t^* , $\phi(\underline{p})$ continuar crescendo segundo $\underline{u}^*$. Assim procedemos a uma busca unidimensional de maneira que nesta iteração, a nova solução do dual seja

$$\underline{p}^{i+1} = \underline{p}^i + t_s \underline{u}^*$$

com t_s obtido de:

$$\begin{array}{ll} \text{MAX} & \phi(\underline{p}^i + t\underline{u}^*) \\ t > 0 \end{array}$$

Através da ordenação crescente dos elementos $t(j)$, construiremos a sequência:

$$\{ t_r \}$$

dos consecutivos limites de t .

Denominamos $\underline{p}_r$ a solução do dual obtida quando r atinge o r -ésimo limite. Assim:

$$\underline{p}_1 = \underline{p}^i + t_1 \underline{u}^* \quad (2.33)$$

$$\underline{p}_{r+1} = \underline{p}_r + (t_{r+1} - t_r) \underline{u}^*$$

Logicamente, nesta busca unidimensional a referência $\underline{z}$ é modificada segundo $\underline{p}^i + t \underline{u}^*$ e temos que, sucessivamente, ir mudando o valor de $\underline{x}$ de maneira que os pontos (x_j, z_j) estejam sobre as respectivas curvas de otimalidade.

Temos que:

$$x_j = d_k^j \quad \text{se } j \in \bar{Q}$$

$$x_j \in [d_k^j, d_{k+1}^j] \quad \text{se } j \in Q$$

Resolvido (PA), calculado $\underline{u}^*$, se dermos um ligeiro acréscimo em t , Q se divide em Q_1 , Q_2 e Q_3 de modo que:

$$x_j \in [d_k^j, d_{k+1}^j] \quad \text{se } j \in Q_1$$

$$x_j = d_{k+1}^j \quad \text{se } j \in Q_2$$

$$x_j = d_k^j \quad \text{se } j \in Q_3$$

Fazemos para $j \in Q_2$ uma atualização em k tal que:

$$\text{NOVO } k = k+1 \quad (2.34)$$

Assim para $j \in \bar{Q}$, Q_2 e Q_3 , $x_j = d_k^j$.

Fazemos uma primeira atualização nos conjuntos Q e $\bar{Q}$ definindo:

$$\begin{aligned} \text{NOVO } Q &= Q_1 \\ \text{NOVO } \bar{Q} &= \bar{Q} + Q_2 + Q_3 \end{aligned} \quad (2.35)$$

Crescendo o valor de t , atingindo t_r , para algum j , z_j iguala algum c_k^j , donde $x_j \in [d_\tau^j, d_{\tau+1}^j]$ com:

$$\begin{aligned} \tau &= k & \text{se } j \in \bar{Q} & \text{ e } w_j > 0 \\ \tau &= k+1 & \text{se } j \in \bar{Q} & \text{ e } w_j < 0 \end{aligned} \quad (2.36)$$

(*) No intervalo de ordem k , k refere-se ao limite inferior do intervalo ou à sua inclinação.

Se t_r é ultrapassado ligeiramente, x_j muda para:

$$\begin{aligned} x_j &= d_{k+1}^j & j \in \bar{Q} & , \quad w_j > 0 \\ x_j &= d_{k-1}^j & j \in \bar{Q} & , \quad w_j < 0 \end{aligned} \quad (2.37)$$

Assim podemos escrever que a cada limite t_r "ultrapassado", para o correspondente $j \in \bar{Q}$, k é atualizado de modo que:

$$\text{NOVO } k = \gamma = k + \text{SIGN}(w_j) \quad (2.38)$$

Atualizado k , para a variável x_j , calculamos novo limitante (se ele existir), que é convenientemente inserido na sequência $\{t_r\}$.

Atingido t_s , ótimo na direção $\underline{u}^*$, a atualização de k é feita conforme (2.36).

Procedemos agora ao cálculo de t_s , na busca unidimensional propriamente dita, o que corresponde a achar o elemento da sequência $\{t_r\}$ além do qual não se verifica crescimento em $\phi(p)$. Definimos:

$$\underline{\bar{b}} = \underline{b} - \Lambda \frac{\bar{Q} \bar{Q}}{\underline{k}} \quad (2.39)$$

Inicialmente, quando $t < t_1$, temos a taxa de crescimento (derivada direcional máxima):

$$\psi(\underline{u}^*) = \underline{u}^* \underline{\bar{b}}$$

Se ultrapassarmos t_1 , $\underline{\bar{b}}$ é alterado de acordo com a modificação no x_j correspondente a t_1 , e para cada $\underline{\bar{b}}$ atualizado, este permanece constante entre dois limites consecutivos de t .

Supondo atingido p_r , sendo x_{jr} a variável correspondente, dando um pequeno acréscimo (ξ) a t , obtemos:

$$p = p_r + \xi \underline{u}^*$$

Definindo:

$$\bar{R} = \bar{Q} - \{jr\}$$

escrevemos o valor da função dual:

$$\phi(\underline{p}) = \underline{p} \underline{b} + \sum_{j \in \bar{R}} \left\{ f_j(d_k^j) - \underline{p} A_{d_k^j}^j \right\} + f_{jr}(d_Y^{jr}) - \underline{p} A_{d_Y^{jr}}^{jr} + \sum_{j \in Q} g_k^{jr}$$

$$\phi(\underline{p}_r) = \underline{p}_r \underline{b} + \sum_{j \in \bar{R}} \left\{ f_j(d_k^j) - \underline{p}_r A_{d_k^j}^j \right\} + f_{jr}(d_k^{jr}) - \underline{p}_r A_{d_k^{jr}}^{jr} + \sum_{j \in Q} g_k^{jr}$$

No limite t_r , jr deixa o conjunto $\bar{Q}$, daí $\underline{p}_r A_{d_k^{jr}}^{jr} = c_k^{jr}$ e g_k^{jr} é constante no intervalo $|d_t^{jr}, d_{t+1}^{jr}|$. Donde:

$$f_{jr}(d_Y^{jr}) - \underline{p}_r A_{d_Y^{jr}}^{jr} = f_{jr}(d_k^{jr}) - \underline{p}_r A_{d_k^{jr}}^{jr}$$

$$\phi(\underline{p}) - \phi(\underline{p}_r) = (\underline{p} - \underline{p}_r) (\underline{b} - A_{d_k}^{\bar{R}, \bar{R}} - A_{d_Y}^{jr, jr}) \quad (2.40)$$

Definindo a "taxa de crescimento" e o "gradiente":

$$TC = \underline{u}^* \underline{\bar{b}} \quad (2.41)$$

$$GRAD = \underline{\bar{b}}$$

Vemos que $\underline{b} - A_{d_k}^{\bar{R}, \bar{R}} - A_{d_Y}^{jr, jr}$, correspondente ao $\underline{\bar{b}}$ atualizado, é o "novo gradiente".

Assim:

$$\text{NGRAD} = \underline{b} - \Lambda \frac{\bar{R}}{d_k} - \Lambda^{jr} d_Y^{jr} - \Lambda^{jr} d_k^{jr} + \Lambda^{jr} d_k^{jr}$$

$$\text{NGRAD} = \underline{b} - \Lambda \frac{\bar{Q}}{d_k} - \Lambda^{jr} (d_Y^{jr} - d_k^{jr})$$

$$\text{NGRAD} = \text{GRAD} - \Lambda^{jr} (d_Y^{jr} - d_k^{jr})$$

(2.42)

$$\text{NTC} = \text{TC} - \underline{u} * \Lambda^{jr} (d_Y^{jr} - d_k^{jr})$$

Verifica-se facilmente que o produto $\underline{u} * \Lambda^{jr} (d_Y^{jr} - d_k^{jr})$ é sempre positivo, donde:

$$\text{NTC} < \text{TC}$$

(2.43)

Vemos daí que, se na função dual caminhamos segundo alguma direção de crescimento desta, a cada mudança de face a taxa de crescimento diminui.

Do acima exposto depreende-se que $\phi(\underline{p})$ é realmente côncava linear por partes.

Como critério de "parada" do procedimento acima descrito, aparece naturalmente:

$$\text{NTC} < 0$$

(2.44)

Podemos ilustrar geometricamente a relação $\tilde{NTC} < TC$. Escrevemos:

$$\left\langle \frac{\underline{u}^*}{\|\underline{u}^*\|}, \text{NGRAD} \right\rangle < \left\langle \frac{\underline{u}^*}{\|\underline{u}^*\|}, \text{GRAD} \right\rangle$$

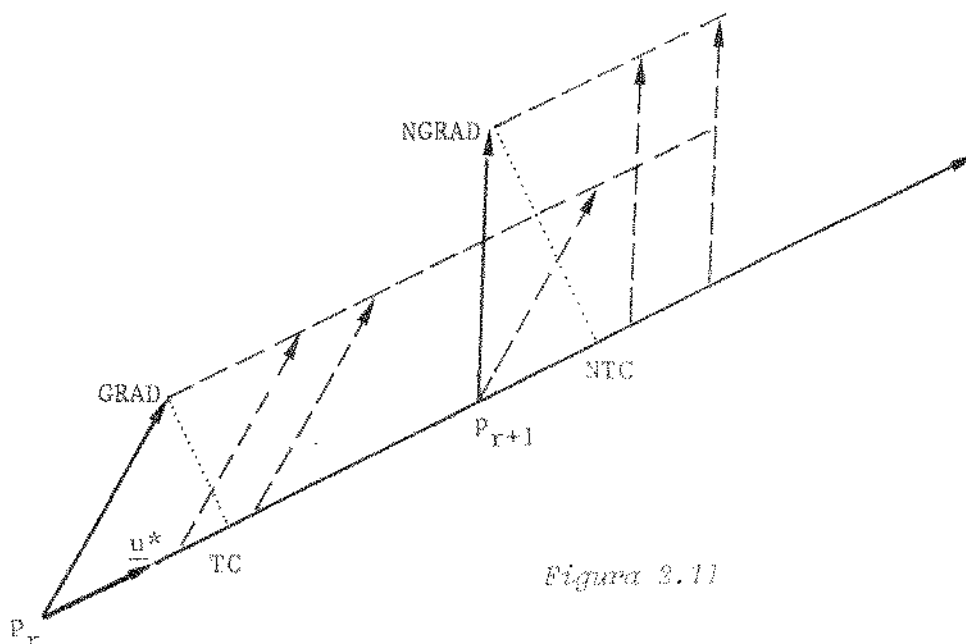


Figura 3.11

A taxa de crescimento é a projeção do gradiente na direção $\underline{u}^*$. O gradiente é constante entre dois limites consecutivos de t e a cada limite atingido ele muda, não se sabe a priori como, mas a sua projeção na direção $\underline{u}^*$ diminui.

Durante o processo anterior, se após o limite t_r a taxa de crescimento for ainda positiva e não houver nenhum outro limite definido, t pode crescer arbitrariamente e o problema primal não tem solução factível.

2.7 - A Nova Iteração

Na busca unidimensional

$$\begin{array}{l} \text{MAX} \quad \phi(\underline{p}^i + t\underline{u}^*) \\ t > 0 \end{array}$$

descrita no item anterior, podemos obter um dos resultados:

- Fazer t arbitrariamente grande segundo uma "taxa de crescimento" positiva, o que acarreta:

$$\text{SUP} \left\{ \phi(\underline{p}^i + t\underline{u}^*) \quad , \quad t \rightarrow \infty \right\} = \infty$$

donde concluímos que o problema (P) é infactível (LSDOM).

- Obtermos o limite de crescimento de t na direcção $\underline{u}^*$ (t_s) finito.
- Neste caso fazemos:

$$\underline{p}^{i+1} = \underline{p}^i + t_s \underline{u}^*$$

Sendo x_{js} a variável correspondente a t_s , temos que $z_{js} = c_{js}$. Assim o índice js sai do conjunto $\bar{Q}$ e vai completar o Q da próxima iteração.

Durante a busca unidimensional, os valores de k e $\bar{b}$ são atualizados a cada t_r , de modo que no final desta temos k e $\bar{b}$ correspondendo ao novo Q .

Podemos representar as modificações ocorridas nos conjuntos Q e $\bar{Q}$, entre duas iterações consecutivas, da seguinte maneira:

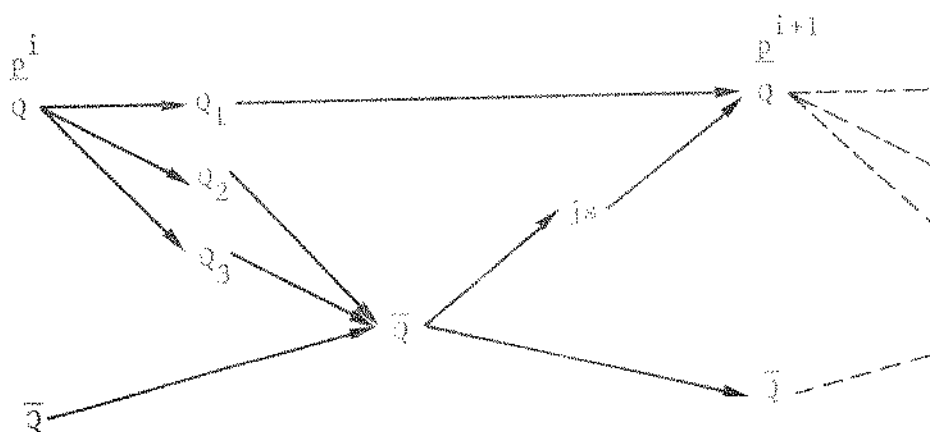


Figura 2.18

Sendo também calculado a variação no valor da função dual, $\Delta\phi$, escrevemos:

$$\phi(\underline{p}^{i+1}) = \phi(\underline{p}^i) + \Delta\phi$$

Para o cálculo da nova inclinação de referência:

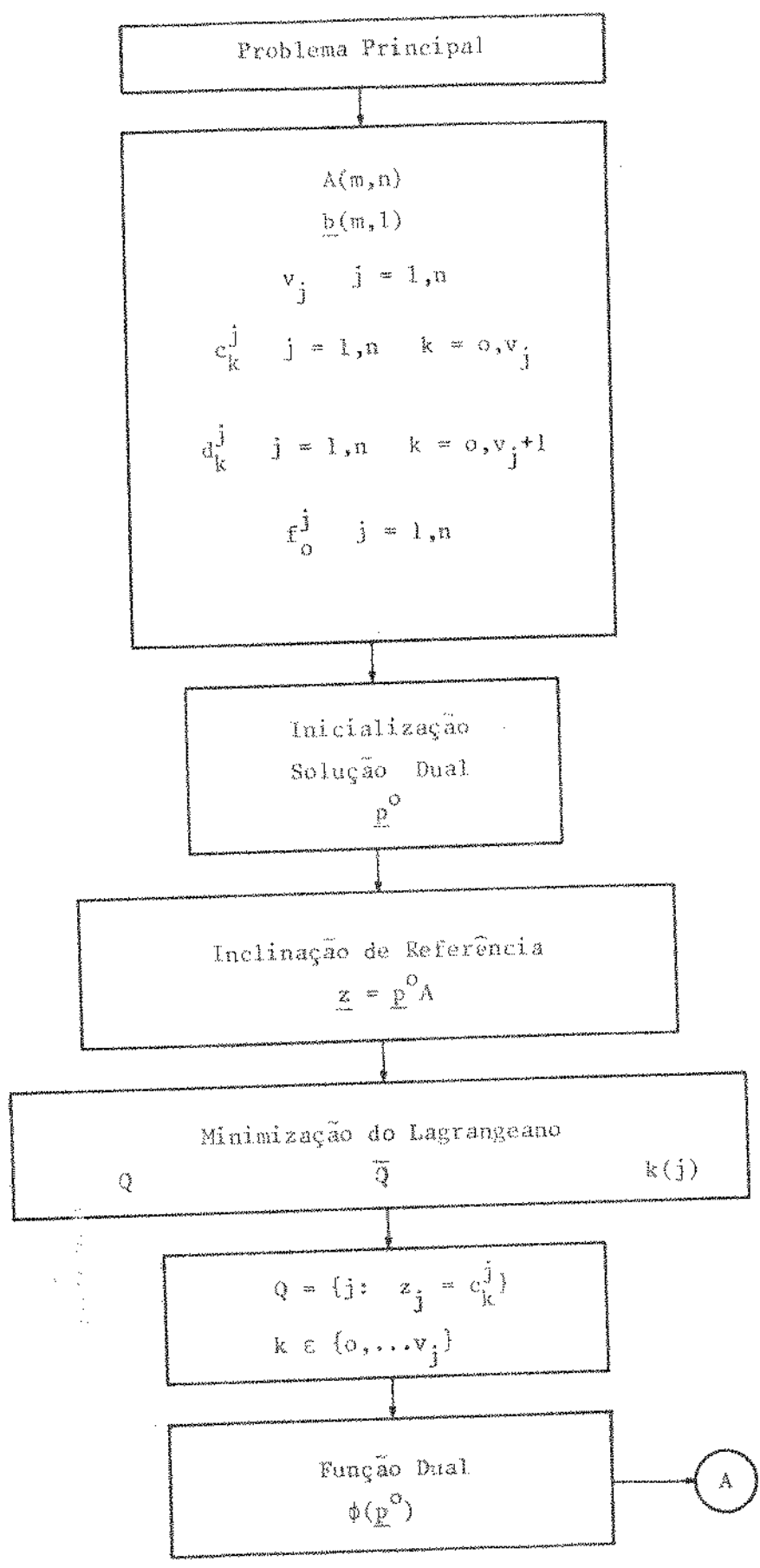
$$\underline{z} = \underline{p}^{i+1} A$$

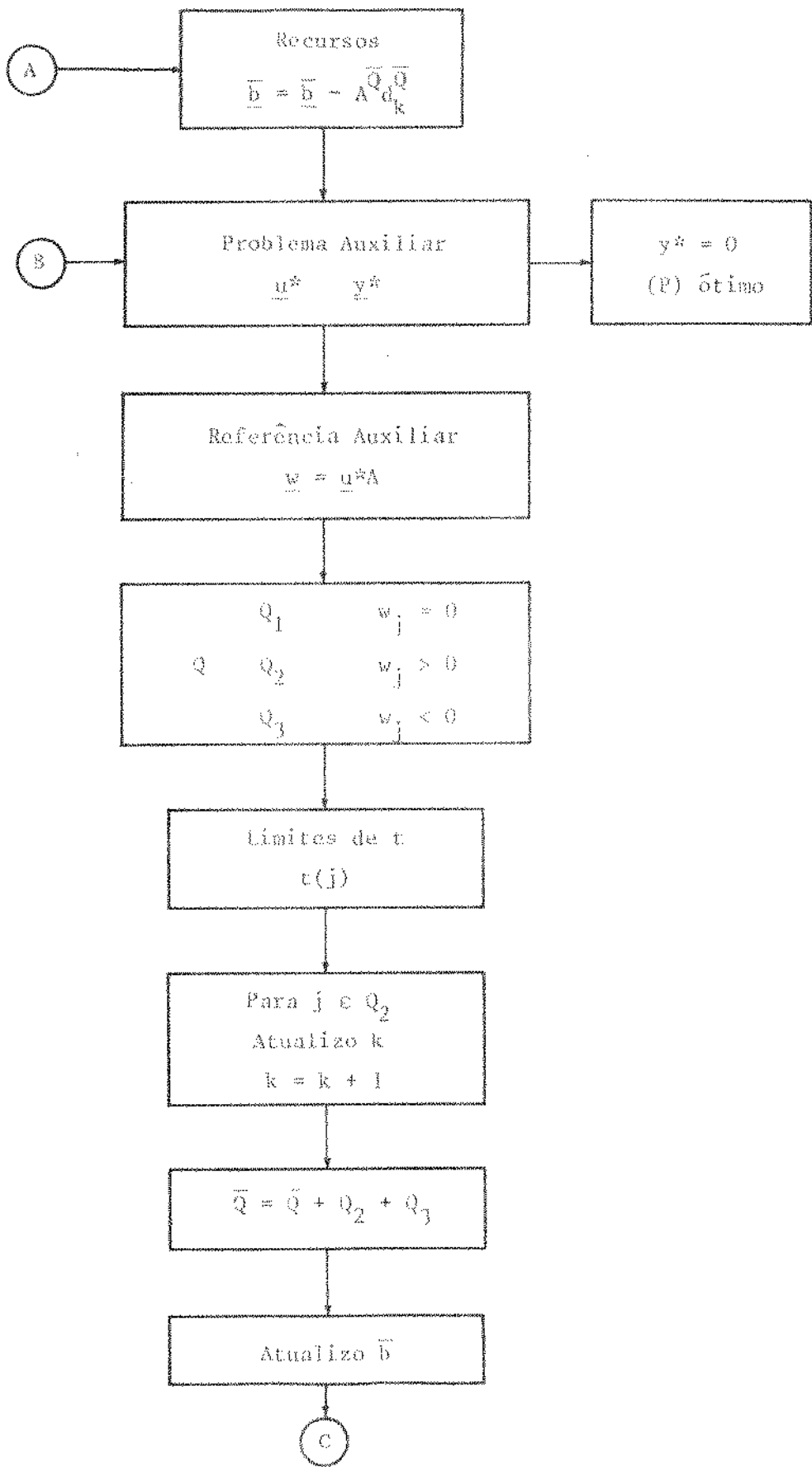
$$\underline{z} = (\underline{p}^i + t_s \underline{u}^*) A$$

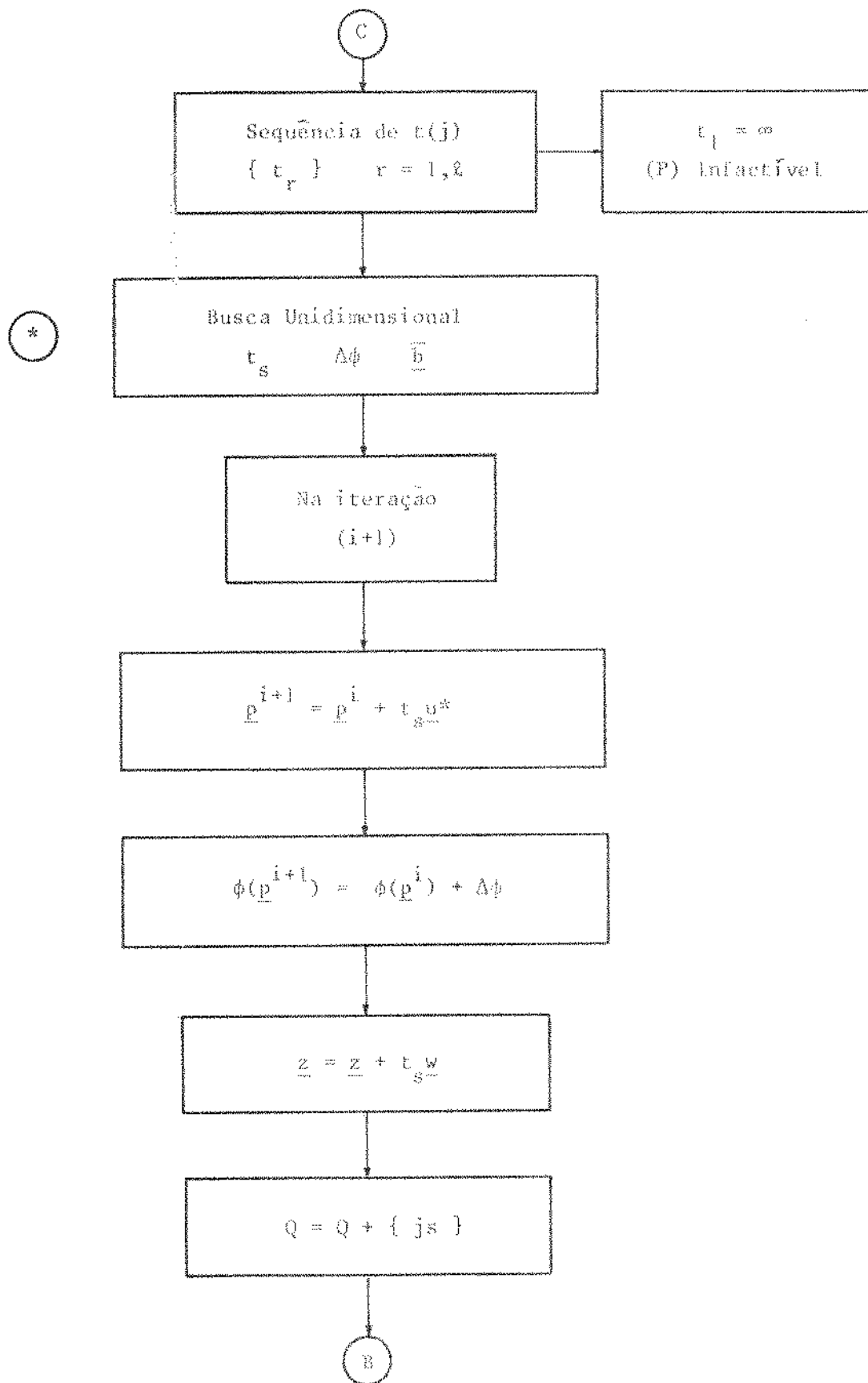
$$\underline{z} = \underline{z} + t_s \underline{w}$$

Assim com Q , k , $\underline{b}$ e $\underline{z}$, resolvemos o novo problema auxiliar, donde se $\underline{y}^* = 0$ (P) está resolvido, caso contrário repetimos o processo.

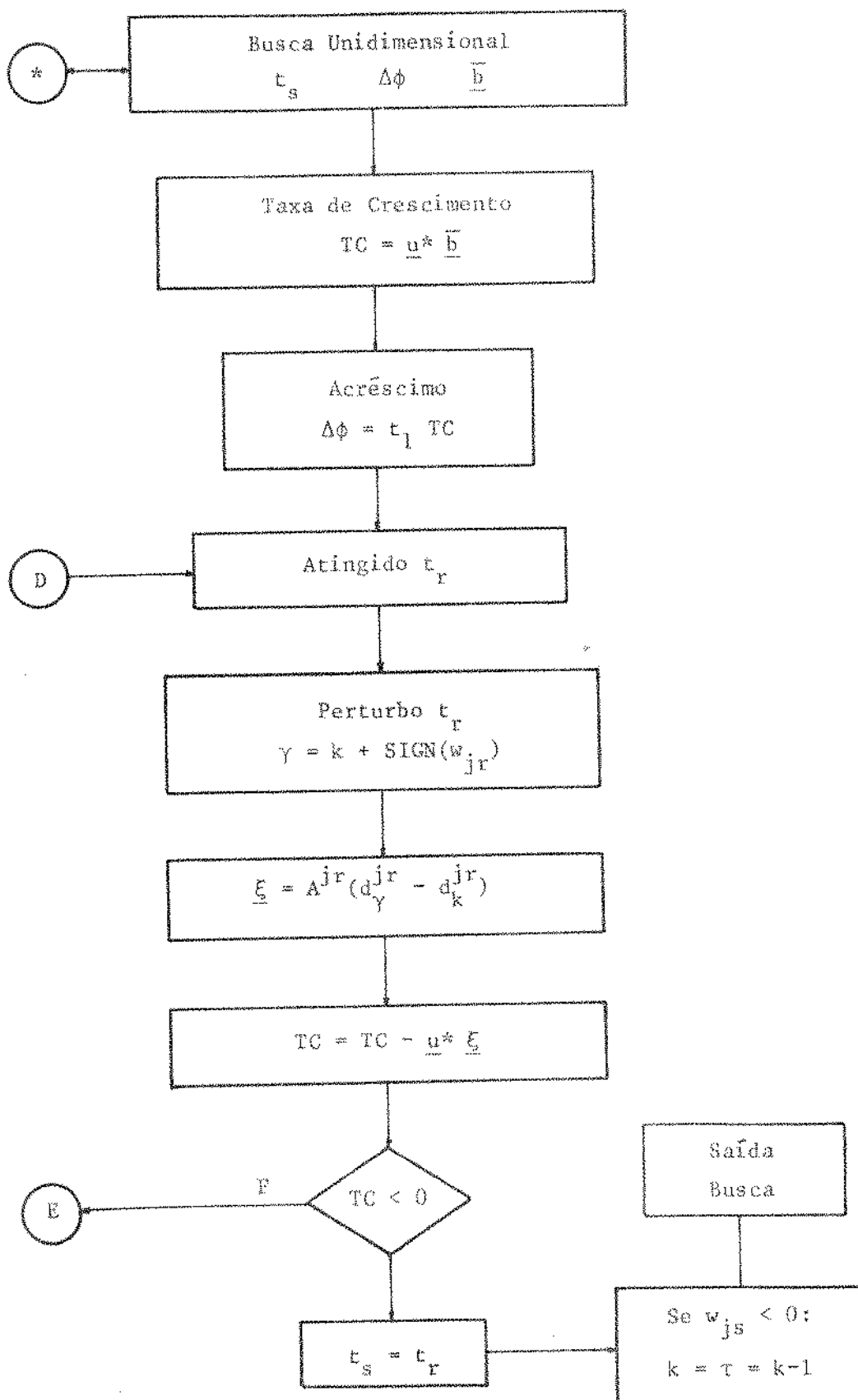
2.8 - Diagrama de Blocos - Problema Principal

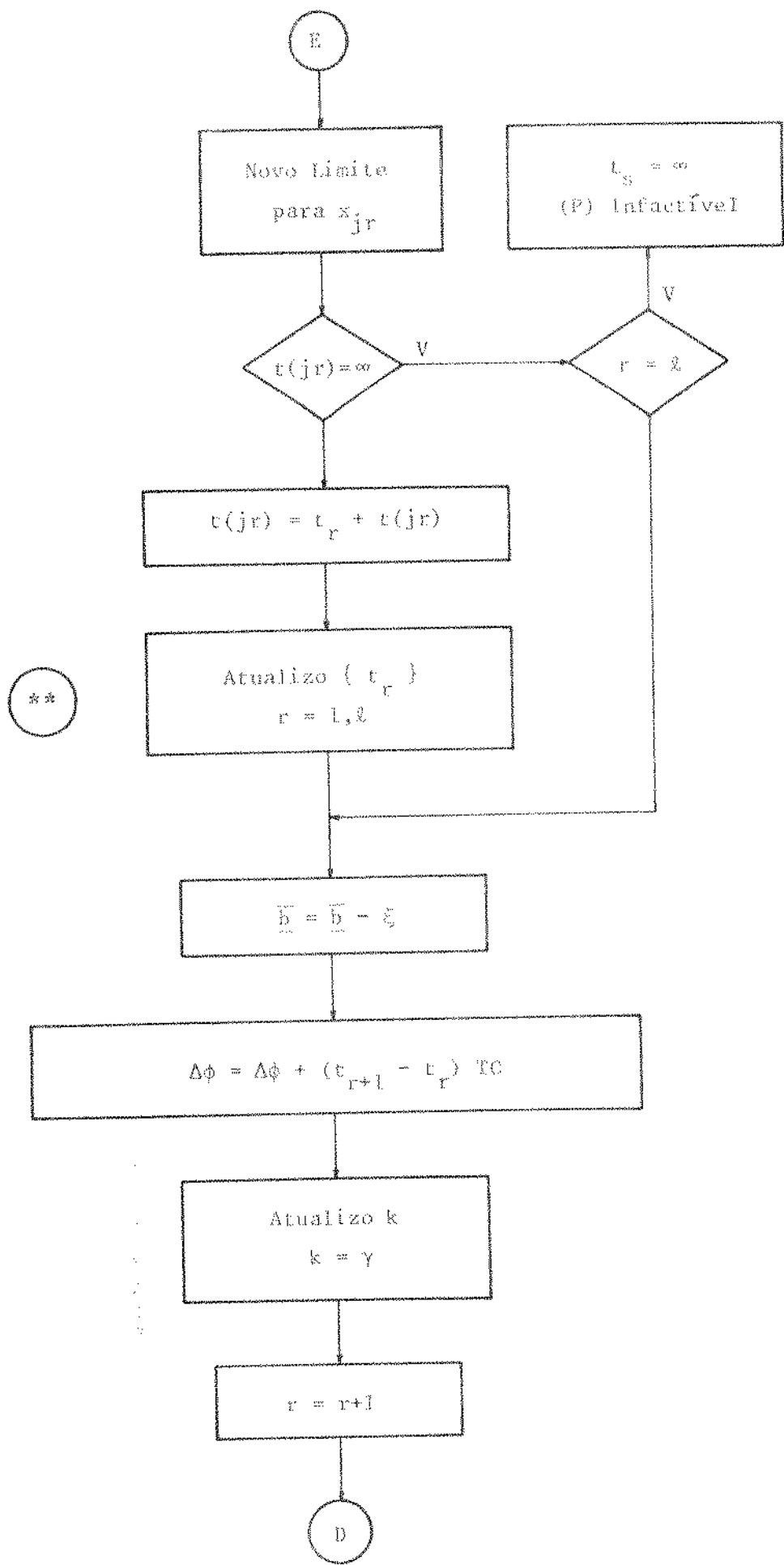


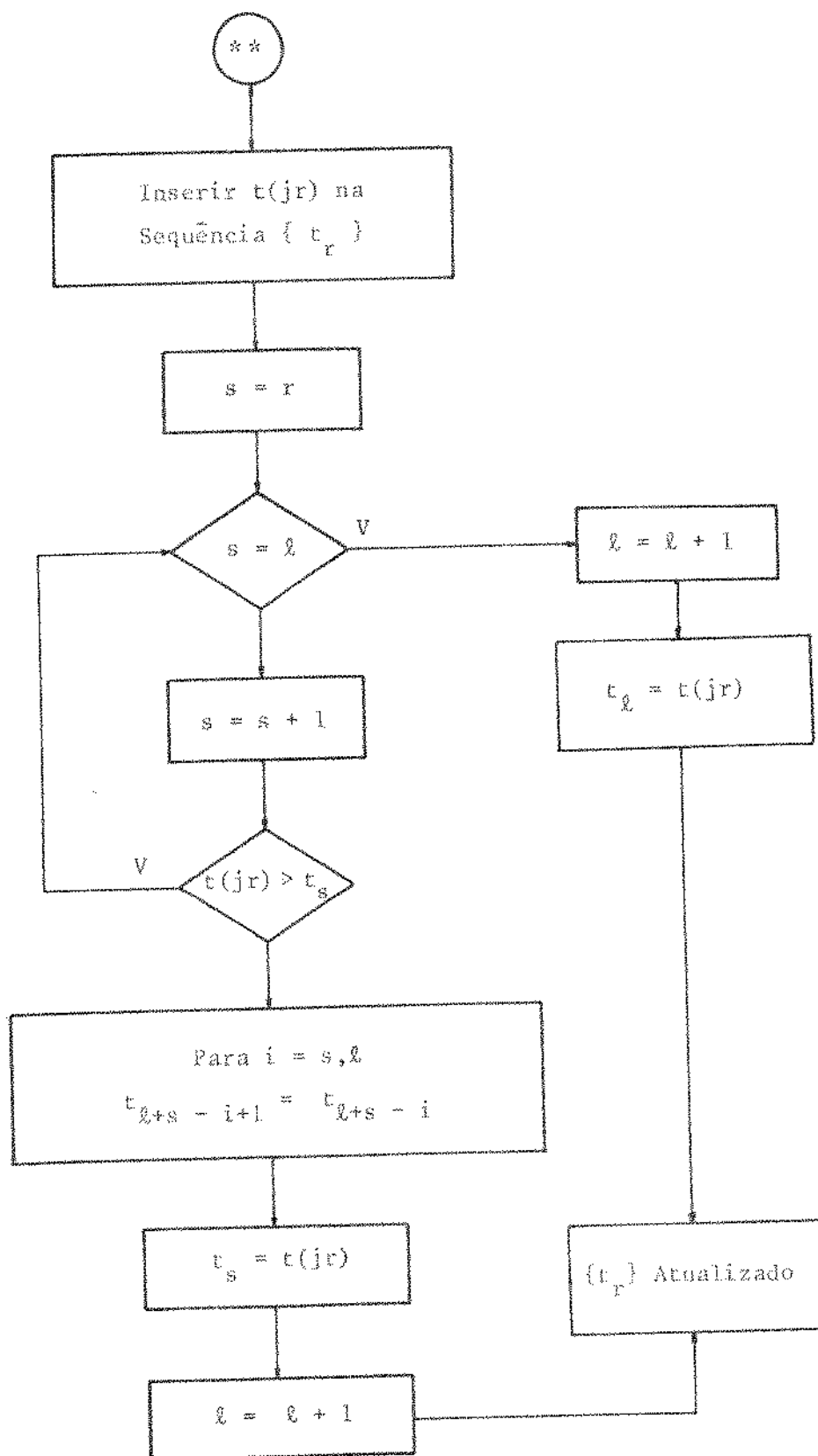




2.9 - Diagrama de Blocos - Busca Unidimensional





2.10 - Diagrama de Blocos - Atualização de $\{t_r\}$ 

CAPÍTULO 3: OBTENÇÃO DA MÁXIMA DERIVADA DIRECIONAL
(PROBLEMA AUXILIAR, PA)

3.1 - Introdução

No desenvolvimento anterior, estando num ponto arbitrário do dual, quando da procura duma direção que corresponde a maior taxa de crescimento para $\phi(\underline{p})$, se o Q é não vazio, deparamos com o problema (PAD), cuja solução ótima fornece a tal direção:

$$\begin{aligned} \text{MAX}_{\underline{u}} \quad & \left\{ \text{MIN}_{\underline{x}_Q} \left\{ \underline{u}(\underline{b} - A_{\underline{d}_k}^Q - A_{\underline{x}_Q}^Q) \right\} \right\} \\ \text{S.a:} \quad & \|\underline{u}\| = 1 \end{aligned} \quad (\text{PAD})$$

$$\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q$$

Este procedimento é equivalente ao de se obter o "vetor multiplicador" referente à solução básica ótima do problema (PA):

$$\begin{aligned} \text{MIN} \quad & h(\underline{y}) = \sum_{j=1}^m e_j |y_j| \\ \text{S.a:} \quad & A_{\underline{x}_Q}^Q + \underline{y} = \underline{b} - A_{\underline{d}_k}^Q \end{aligned} \quad (\text{PA})$$

$$\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q$$

Norma retangular adotada:

$$\|u\| = \text{MIN} \{t: -te \leq u \leq te, \quad e > 0, \quad t > 0\}$$

No apêndice 1, prova-se que (PAD) é o dual de (PA).

Do exposto, mostramos a não existência de intervalo de dualidade entre (PA) e (PAD):

Resolvido (PA), tem-se a sua solução ótima, $(\underline{y}^*, \underline{x}_Q^*)$, satisfazendo:

$$A^Q \underline{x}_Q^* + \underline{y}^* = \underline{b} - A^{\overline{Q}}_{\underline{d}} \overline{Q}$$

(PAD) resulta em:

$$\text{MAX}_{\underline{u}} \left\{ \underline{u} (\underline{b} - A^{\overline{Q}}_{\underline{d}} \overline{Q} - A^Q \underline{x}_Q^*) \right\} = \text{MAX}_{\underline{u}} \left\{ \underline{u} \underline{y}^* \right\}$$

cuja solução, separável, é:

$$\begin{aligned} y_j^* &= 0, & u_j^* &\in [-e_j, e_j] \\ y_j^* &> 0, & u_j^* &= e_j \\ y_j^* &< 0, & u_j^* &= -e_j \end{aligned} \tag{3.0}$$

Assim o valor ótimo de (PAD) é:

$$\begin{aligned} \psi(\underline{u}^*) &= \underline{u}^* (\underline{b} - A^{\overline{Q}}_{\underline{d}} \overline{Q} - A^Q \underline{x}_Q^*) = \\ &= \sum_{y_j^* > 0} e_j y_j^* + \sum_{y_j^* < 0} (-e_j) y_j^* \end{aligned}$$

Sem perda de generalidade, considerando j tal que $y_j^* = 0$, escrevemos:

$$\psi(\underline{u}^*) = \sum_{j=1}^m e_j |y_j^*| = h(\underline{y}^*)$$

Na resolução de (PA), utilizamos basicamente o algoritmo primal linear por partes, numa formulação adequada às suas características particulares.

Quanto à solução ótima de (PA), ela nunca será tipo β . Qualquer um dos fatores abaixo estabelece isto:

- Função objetivo penaliza as m variáveis irrestritas.
- A cada iteração, $\underline{y}$ tem que satisfazer a $A^Q \underline{x}_Q + \underline{y} = \underline{b} - A^{\bar{Q}}_{d_k} \bar{Q}$, nunca crescendo arbitrariamente.

3.2 - O Problema

O problema (PA), pode ser caracterizado como um primal restrito:

- Linear e canalizado nas variáveis $\underline{x}_Q$.
- Linear por partes e irrestrito, com somente um ponto crítico, nas variáveis $\underline{y}$.

No problema principal, a cada $j \in Q$ obtido, definimos $j(q) = j$. Se Q é vazio, ($q = 0$), a solução de (PA) é trivial, ($\underline{y}^* = \underline{b} - A^{\bar{Q}}_{d_k} \bar{Q}$).

$$\begin{aligned} \text{Definindo: } x_{j+n} &= y_j & j &= 1, m \\ e_{j+n} &= e_j & j &= 1, m \\ \bar{b} &= \underline{b} - A^{\bar{Q}}_{d_k} \bar{Q} \end{aligned} \tag{3.1}$$

$A = [A^Q, \Pi]$, inicialmente

$$h_j(x_j) = 0 \quad j = 1, q$$

$$h_j(x_j) = e_j |x_j| \quad j = n+1, n+m$$

Reescrevemos (PA):

$$\begin{aligned} \text{MIN} \quad h(\underline{x}) &= \sum_{j=n+1}^{n+m} h_j(x_j) \\ \text{S.a:} \quad A\underline{x} &= \underline{b} \\ d_k^j &\leq x_j \leq d_{k+1}^j \quad j = 1, q \end{aligned} \quad (\text{PA})$$

Observações:

- Para simplificar, sendo Q constituído de q elementos, se j varia em Q , escrevemos $j = 1, q$ e não $j(i)$ com $i = 1, q$.

- A variável $\underline{x}$ é um vetor coluna de $m+q$ componentes.

- I é a matriz identidade (m, m) .

Representemos graficamente as $h_j(x_j)$ e as respectivas curvas de otimalidade, $(CO)_j$:

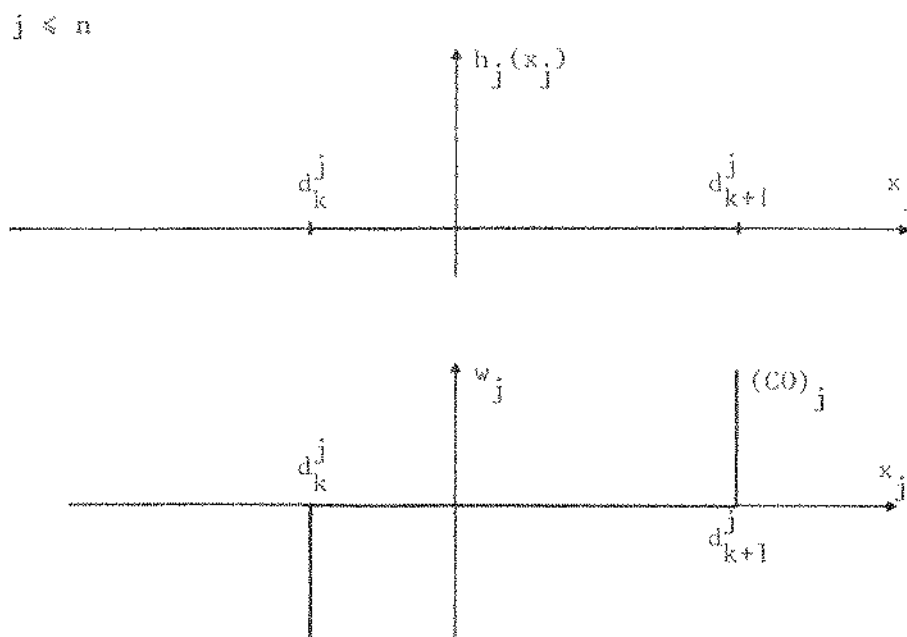


Figura 3.1

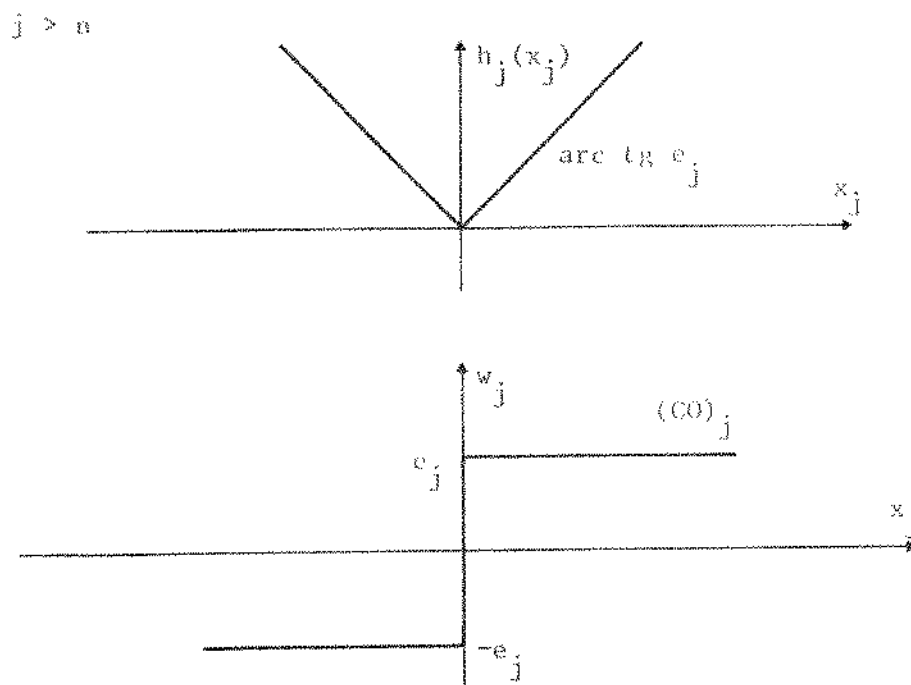


Figura 3.2

3.3 - O Algoritmo

A partir de uma solução básica factível, o algoritmo percorre soluções básicas factíveis, num número finito de passos, até a solução ótima. A cada passo fazemos um teste de otimalidade, que, se satisfeito, finaliza o algoritmo, se não, procuramos uma nova solução que diminua $h(\underline{x})$.

No problema auxiliar, por construção, A tem "rank" m .

Sendo $\{I, J\}$ a base, I o conjunto de índices das variáveis básicas, de $A\underline{x} = \underline{\bar{b}}$ obtemos:

$$A^J \underline{x}_J + A^I \underline{x}_I = \underline{\bar{b}}$$

$$\underline{x}_I = (A^I)^{-1} \underline{\bar{b}} - (A^I)^{-1} A^J \underline{x}_J$$

$$\underline{x}_I = \underline{\bar{b}} - \tilde{A}^J \underline{x}_J \quad (3.2)$$

Colocamos as variáveis $\underline{x}_j$ em pontos críticos $(\underline{d}_k^J)$. Assim, para $j \in J$,

$$x_j = 0 \quad \text{se } j > n \quad (3.3)$$

$$x_j = \underline{d}_k^j \quad \text{se } j \leq n \quad (3.4)$$

Temos a solução básica:

$$\underline{x} = \begin{bmatrix} \underline{x}_J \\ \underline{x}_I \end{bmatrix} = \begin{bmatrix} \underline{d}_k^J \\ \underline{b} - \hat{A}^J \underline{d}_k^J \end{bmatrix} \quad (3.5)$$

Note que a uma base $\{I, J\}$, correspondem várias soluções básicas.

A solução (3.5) é factível se $x_j \in |d_k^j, d_{k+1}^j|$, $j \in I$, $j < n$.

No problema auxiliar, sendo a base inicial:

$$J = Q$$

$$I = \{n+1, \dots, n+m\}$$

como x_j , $j \in I$ é irrestrita, uma solução básica factível de partida é de obtenção trivial.

Considerações:

- A solução (3.5) é degenerada se x_j , $j \in I$ coincide com algum ponto crítico $\underline{d}_k^j$.

- (PA) é degenerado se alguma solução básica factível for degenerada.

- Supomos (PA) não degenerado.

3.4 - Condições de Otimalidade

Definimos o vetor de inclinações $\underline{c}_k^I(1, m)$:

$$\underline{c}_k^j = 0 \quad , \quad j \leq n \quad (3.6)$$

$$\underline{c}_k^j = e_j \quad , \quad j > n \quad , \quad x_j > 0 \quad (3.7)$$

$$\underline{c}_k^j = -e_j \quad , \quad j > n \quad , \quad x_j < 0 \quad (3.8)$$

Sendo:

$$L(\underline{x} \quad , \quad \underline{u}) = h(\underline{x}) + \underline{u}(\bar{b} - A\underline{x})$$

se para dado $\underline{u}$, $\underline{x}^*$ minimiza $L(\underline{x} \quad , \quad \underline{u})$, sujeito a $\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q$, então os pontos $(\underline{x}_j^* \quad , \quad w_j)$, com $w_j = \underline{u} A^j$, estão sobre as respectivas curvas de otimalidade. Se, além disso, $\underline{x}^*$ é factível de (PA), então $\underline{x}^*$ é ótimo de (PA). Como o algoritmo primal só trabalha com soluções factíveis, resulta a condição de otimalidade a cada iteração:

$\underline{x}^*$ É ÓTIMO DO (PA), SE OS PONTOS $(\underline{x}_j^* \quad , \quad w_j)$, $j = 1, n+m$, ESTÃO SOBRE AS RESPECTIVAS CURVAS DE OTIMALIDADE.

Para $j \in I$, supondo não degenerescência, x_j está em algum intervalo linear de $h_j(x_j)$. Para obtermos os pontos $(x_j \quad , \quad w_j)$ sobre $(CO)_j$, impomos:

$$w_j = \underline{c}_k^j \quad j \in I \quad (3.9)$$

daí, $\underline{u} A^I = \underline{c}_k^I$ ou

$$\underline{u} = \underline{c}_k^I (A^I)^{-1} \quad (3.10)$$

Desta maneira x_j básico satisfaz a condição de otimalidade.

Sendo $w_j = \underline{u} A^j$, de (3.10) obtemos $w_j = \underline{c}_k^I (A^I)^{-1} A^j$, ou

$$w_j = \underline{c}_k^I \tilde{A}^j \quad (3.11)$$

Com $\underline{u}$ calculado em (3.10), obtemos w_j , $j \in J$, e testamos os pontos (x_j, w_j) .

Sendo a variável x_j não básica, ela assume um valor crítico. A partir das figuras 3.1 e 3.2, escrevemos $x_j \in (CO)_j$ se:

$$\begin{aligned} \text{para } j \leq n : \quad x_j &= d_k^j \quad \text{e} \quad w_j \leq 0, \quad \text{ou} \\ &\vdots \\ x_j &= d_{k+1}^j \quad \text{e} \quad w_j \geq 0 \end{aligned} \quad (3.12)$$

$$\text{para } j > n : \quad x_j = 0 \quad \text{e} \quad -e_j \leq w_j \leq e_j$$

Supondo que (x_ℓ, w_ℓ) , $\ell \in J$, não pertence a $(CO)_\ell$, ocorre um dos casos:

$$\begin{aligned} \text{para } \ell \leq n \quad & \left\{ \begin{array}{ll} \text{(i)} & x_\ell = d_k^\ell \quad \text{e} \quad w_\ell > 0 \\ \text{(ii)} & x_\ell = d_{k+1}^\ell \quad \text{e} \quad w_\ell < 0 \end{array} \right. \\ \text{para } \ell > n \quad & \left\{ \begin{array}{ll} \text{(iii)} & x_\ell = 0 \quad \text{e} \quad w_\ell > e_\ell \\ \text{(iv)} & x_\ell = 0 \quad \text{e} \quad w_\ell < -e_\ell \end{array} \right. \end{aligned} \quad (3.13)$$

Definimos:

$$\begin{aligned} c_k^\ell &= 0, & \text{casos (i) e (ii)} \\ c_k^\ell &= e_\ell, & \text{caso (iii)} \\ c_k^\ell &= -e_\ell, & \text{caso (iv)} \end{aligned} \quad (3.14)$$

Graficamente um ponto (x_ℓ, w_ℓ) pode estar localizado relativamente à curva de otimalidade da seguinte maneira:

$\ell \leq n$

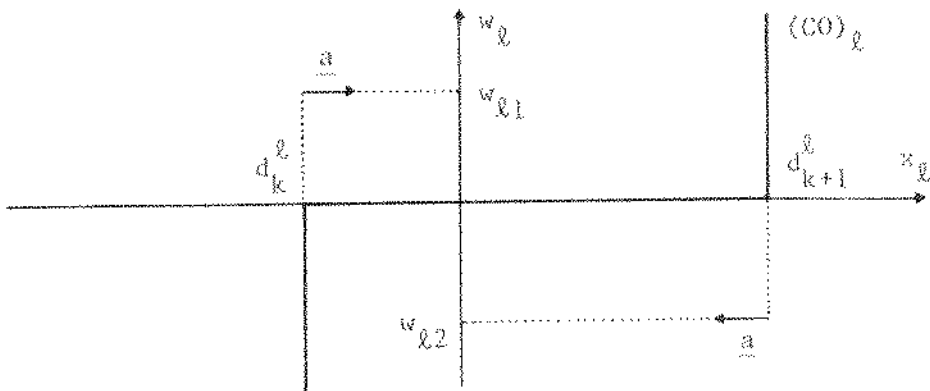


Figura 3.3

$\ell > n$

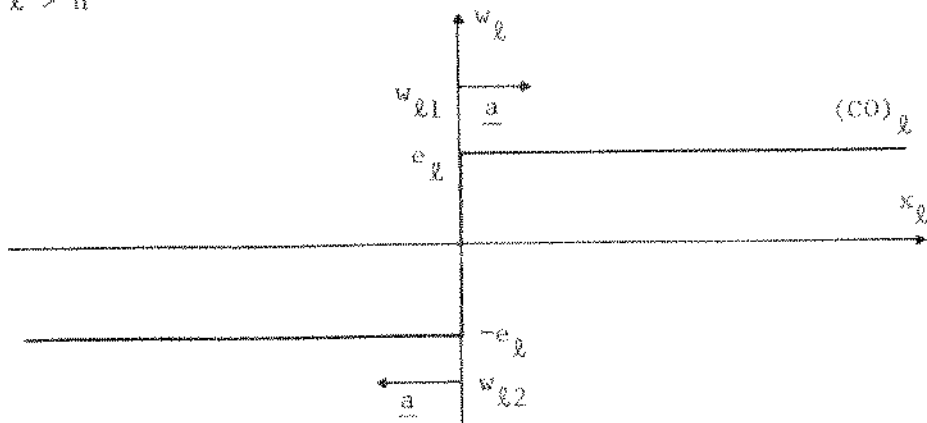


Figura 3.4

Observações:

- ℓ_1 , caso (i) ou (iii)
- ℓ_2 , caso (ii) ou (iv)
- $h(\underline{x})$ diminui se x_j é alterado segundo $\underline{a}$.

3.5 - Mudança de Solução Básica - Acréscimo na Variável Independente

Corresponde aos casos (i) e (iii).

Queremos aumentar x_ℓ , construindo uma nova solução básica factível para (PA), associada a um valor menor de $h(\underline{x})$.

Definimos inicialmente o vetor:

$$\underline{x}(\theta) = \begin{cases} x_j - \theta \hat{A}_t^\ell & j \in I \\ x_\ell + \theta & \\ x_j & j \in J - \{ \ell \} \end{cases}$$

onde t indica a posição de j em I .

Considerando que:

- as variáveis não básicas permanecem inalteradas (exceto x_ℓ),
- as variáveis x_j , $j > n$ são irrestritas,
- a restrição $A\underline{x}(\theta) = \bar{b}$ é satisfeita para θ arbitrário,

concluimos que a factibilidade de $\underline{x}(\theta)$ é restrita para:

$$\left\{ j : j \leq n \text{ e } j \in I + \{ \ell \} \right\} \quad (3.15)$$

$$d_k^j \leq x_j(\theta) \leq d_{k+1}^j$$

Analisando o crescimento de θ , restrito a (3.15), vemos que as variáveis x_j , $j > n$, percorrem intervalos lineares e se o ponto crítico é atingido, a respectiva inclinação se altera de $2e_j$.

Graficamente, para $j \in I$, os pontos (x_j, w_j) se movem relativamente a curva de otimalidade da seguinte maneira:

$j \leq n$

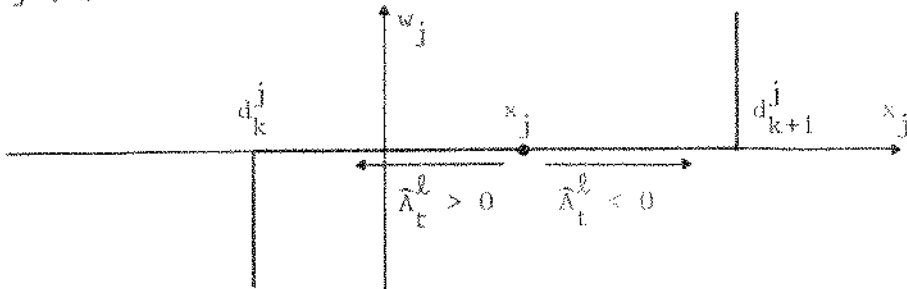


Figura 3.6

$j > n$

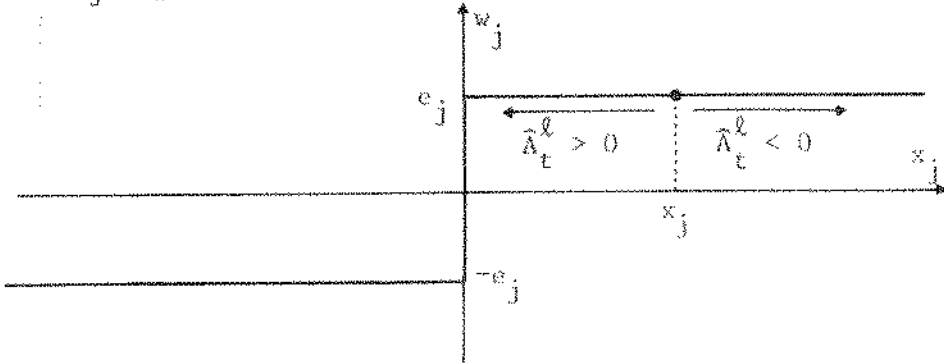


Figura 3.6

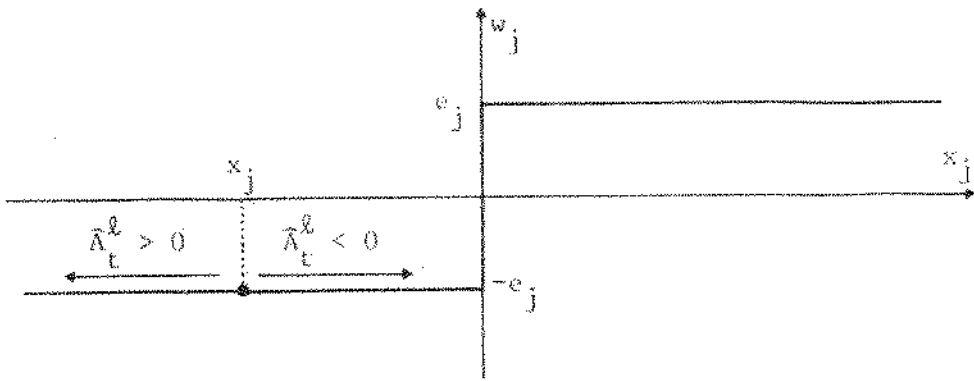


Figura 3.7

O cálculo da variação de $h(\underline{x})$ é feito a cada trecho linear desta, isto é, por intervalo de θ tal que, entre dois θ consecutivos, nenhum x_j ultrapassa o ponto crítico correspondente.

Observando as figuras 3.5, 3.6 e 3.7, escrevemos os valores de θ que dão margem a esta situação:

$$j \leq n, \quad \hat{A}_t^\ell > 0 \quad \theta(j) = \frac{x_j - d_k^j}{\hat{A}_t^\ell} \quad (3.16)$$

$$j \leq n, \quad \hat{A}_t^\ell < 0 \quad \theta(j) = \frac{x_j - d_{k+1}^j}{\hat{A}_t^\ell} \quad (3.17)$$

$$j > n, \quad \begin{cases} \hat{A}_t^\ell > 0, & x_j > 0 \\ \hat{A}_t^\ell < 0, & x_j < 0 \end{cases} \quad \theta(j) = \frac{x_j}{\hat{A}_t^\ell} \quad (3.18)$$

Com relação a x_ℓ , das figuras 3.3 e 3.4, concluímos que se $\ell \leq n$:

$$\theta(\ell) = d_{k+1}^\ell - d_k^\ell \quad (3.19)$$

Ordenando os $\theta(j)$ acima calculados, construímos a sequência $\{\theta_r\}$.

Não considerando degenerescência, supomos distintos todos os elementos de $\{\theta_r\}$.

Crescendo θ sucessivamente, a cada θ_r atingido, para a variável x_j correspondente, não calculamos nenhum outro limite, porque só ocorre uma das duas situações:

$$- j \leq n, \quad x_j = d_k^j \quad \text{ou} \quad d_{k+1}^j$$

Este θ_{rj} é o limite de crescimento de θ (equação 3.15).

$$- j > n, \quad x_j = 0$$

A partir daí, x_j cresce indefinidamente com θ .

Até agora nos preocupamos em ter a equação (3.15) satisfeita. Tendo em vista que o objetivo é obter nova solução que, além de factível, seja básica e melhore o valor de $h(\underline{x})$, consideramos se, com θ percorrendo $\{\theta_r\}$, $h(\underline{x})$ está diminuindo a cada passo.

Definimos:

$$\Delta = c_k^\ell - w_\ell \quad (3.20)$$

Observações:

- c_k^ℓ como em (3.14).
- Em (3.12) vemos que $\Delta < 0$.
- $|\Delta|$ fornece uma medida de quão longe x_ℓ está da otimalidade.

Supondo $\theta \in] 0, \theta_1]$:

$$h(\underline{x}(\theta)) - h(\underline{x}) = \theta (c_k^\ell - c_k^I \hat{\Lambda}^\ell) \quad (3.21)$$

da equação (3.11), $w_\ell = c_k^I \hat{\Lambda}^\ell$.

$$h(\underline{x}(\theta)) - h(\underline{x}) = \theta (c_k^\ell - w_\ell) = \theta \Delta < 0 \quad (3.22)$$

Vemos que Δ é o crescimento por unidade de θ no primeiro intervalo. $h(\underline{x})$ diminuiu.

Se θ ultrapassa θ_1 , em (3.21) c_k^{j1} se altera, conforme figuras 3.6 e 3.7, de c_k^j para:

$$c_k^j = c_k^j - 2e_j \text{Sign}(\hat{\Lambda}_t^\ell) \quad (3.23)$$

sendo $x_{j1} = 0$ referente a $\theta = \theta_1$.

E assim para todo θ_r escrevemos:

$$c_k^{jr} - \overline{c}_k^{jr} = 2e_{jr} \text{Sign} (\hat{A}_{tr}^\ell) \tag{3.24}$$

Definimos σ_r :

$$\sigma_r = (c_k^{jr} - \overline{c}_k^{jr}) \hat{A}_{tr}^\ell \tag{3.25}$$

$$\sigma_r = 2e_{jr} \text{Sign} (\hat{A}_{tr}^\ell) \hat{A}_{tr}^\ell \tag{3.26}$$

Vemos que σ_r é sempre positivo.

Resulta que no intervalo $|\theta_1, \theta_2|$:

$$h(\underline{x}(\theta)) - h(\underline{x}) = \Delta\theta_1 + \Delta_1 (\theta - \theta_1)$$

onde $\Delta_1 = \Delta + \sigma_1$, logo:

$$h(\underline{x}(\theta)) - h(\underline{x}) = \Delta\theta + \sigma_1 (\theta - \theta_1)$$

Genericamente num intervalo $|\theta_r, \theta_{r+1}|$:

$$h(\underline{x}(\theta)) - h(\underline{x}) = \Delta\theta + \sum_{i=1}^r \sigma_i (\theta - \theta_i) \tag{3.27}$$

$$\Delta_r = \Delta + \sum_{i=1}^r \sigma_i \tag{3.28}$$

Vê-se que Δ , inicialmente negativo, sofre acréscimos positivos, σ_r , a cada θ_r atingido. Quando Δ_r se tornar não negativo, $h(\underline{x})$ não diminui mais. Se θ_s o valor de θ tal que isto aconteça.

Desta maneira $x(\theta_s)$ é a melhor solução factível para $h(\underline{x})$ nesta iteração. Observe também que dado a forma como foram obtidos os valores de x_j , $j \in I + \{l\}$, $x(\theta_s)$ é básica.

Assim:

- x_l entra na base e x_{js} sai.
- Se $js = l$, temos a mesma base mas soluções básicas diferentes.

O procedimento anterior pode ser ilustrado graficamente da seguinte maneira:

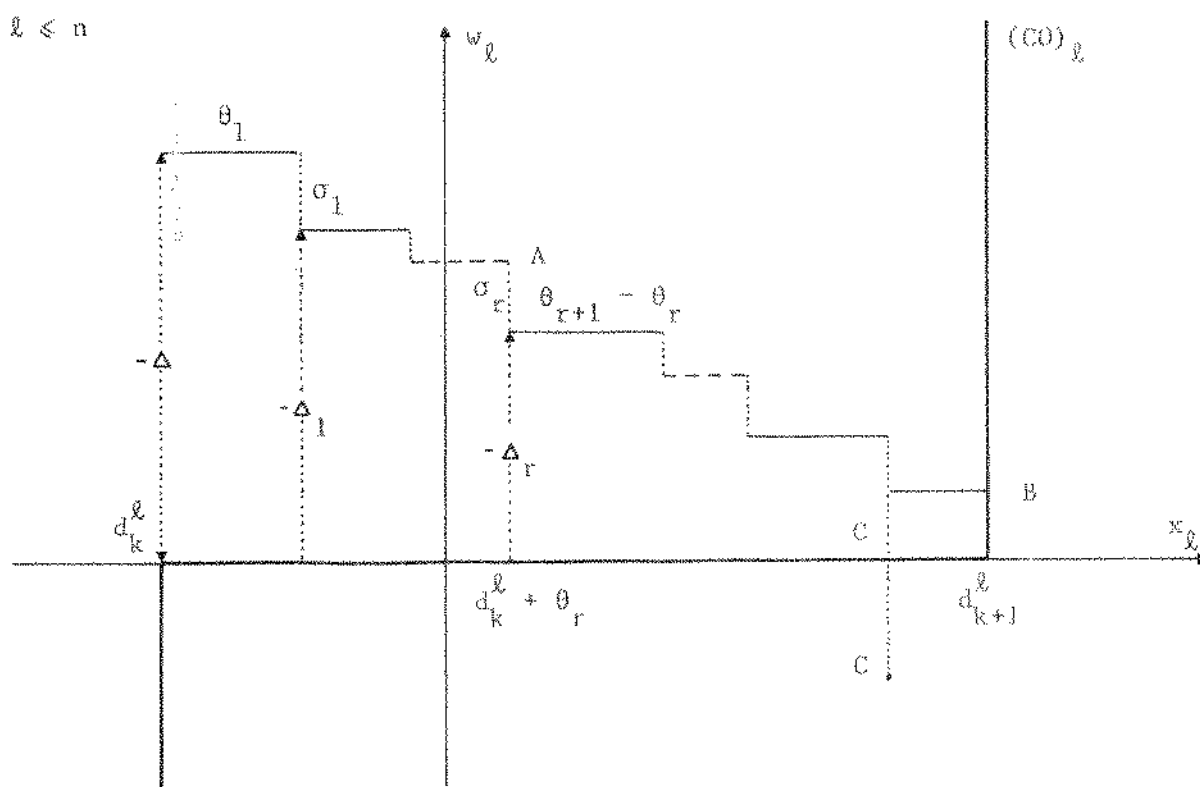


Figura 3.8

3.6 - Mudança de Solução Básica - Decréscimo na Variável Independente

Corresponde aos casos (ii) e (iv).

Queremos diminuir x_ℓ , constituindo uma nova solução básica factível para (PA), associada a um valor menor de $h(\underline{x})$.

A alteração é feita segundo:

$$\underline{x}(\theta) = \begin{cases} x_j + \theta \hat{\Lambda}_t^\ell & j \in I \\ x_\ell - \theta & \\ x_j & j \in J - \{ \ell \} \end{cases}$$

onde t indica a posição de j em I .

O estudo é análogo ao anterior com as seguintes modificações:

- Valores críticos de θ :

$$j \leq n, \quad \hat{\Lambda}_t^\ell > 0 \quad \theta(j) = \frac{d_{k+1}^j - x_j}{\hat{\Lambda}_t^\ell} \quad (3.29)$$

$$j \leq n, \quad \hat{\Lambda}_t^\ell < 0 \quad \theta(j) = \frac{d_k^j - x_j}{\hat{\Lambda}_t^\ell} \quad (3.30)$$

$$j > n, \quad \begin{cases} \hat{\Lambda}_t^\ell > 0, & x_j < 0 \\ \hat{\Lambda}_t^\ell < 0, & x_j > 0 \end{cases} \quad \theta(j) = \frac{-x_j}{\hat{\Lambda}_t^\ell} \quad (3.31)$$

(*) Se $\ell \leq n$: $\theta(\ell) = d_{k+1}^\ell - d_k^\ell$.

- Taxa inicial de deacr scimo:

$$\Delta = w_{\ell} - c_k^{\ell} \quad (3.32)$$

A cada θ_r atingido temos:

- Nova inclina  o:

$$\bar{c}_k^j = c_k^j + 2e_j \text{Sign}(\hat{A}_t^{\ell}) \quad (3.33)$$

- Nova taxa de deacr scimo:

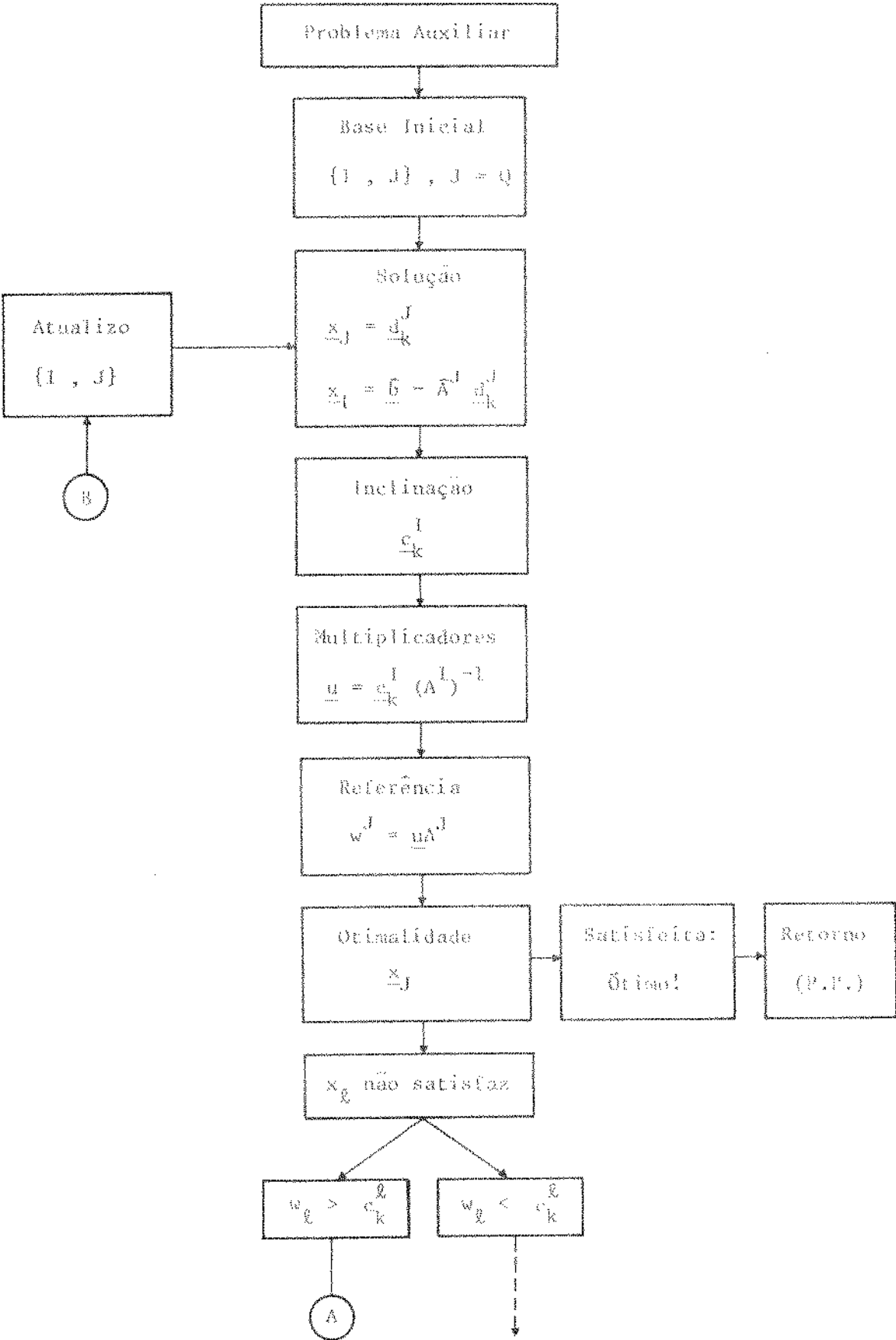
$$\sigma_r = (\bar{c}_k^j - c_k^j) \hat{A}_t^{\ell} \quad (3.34)$$

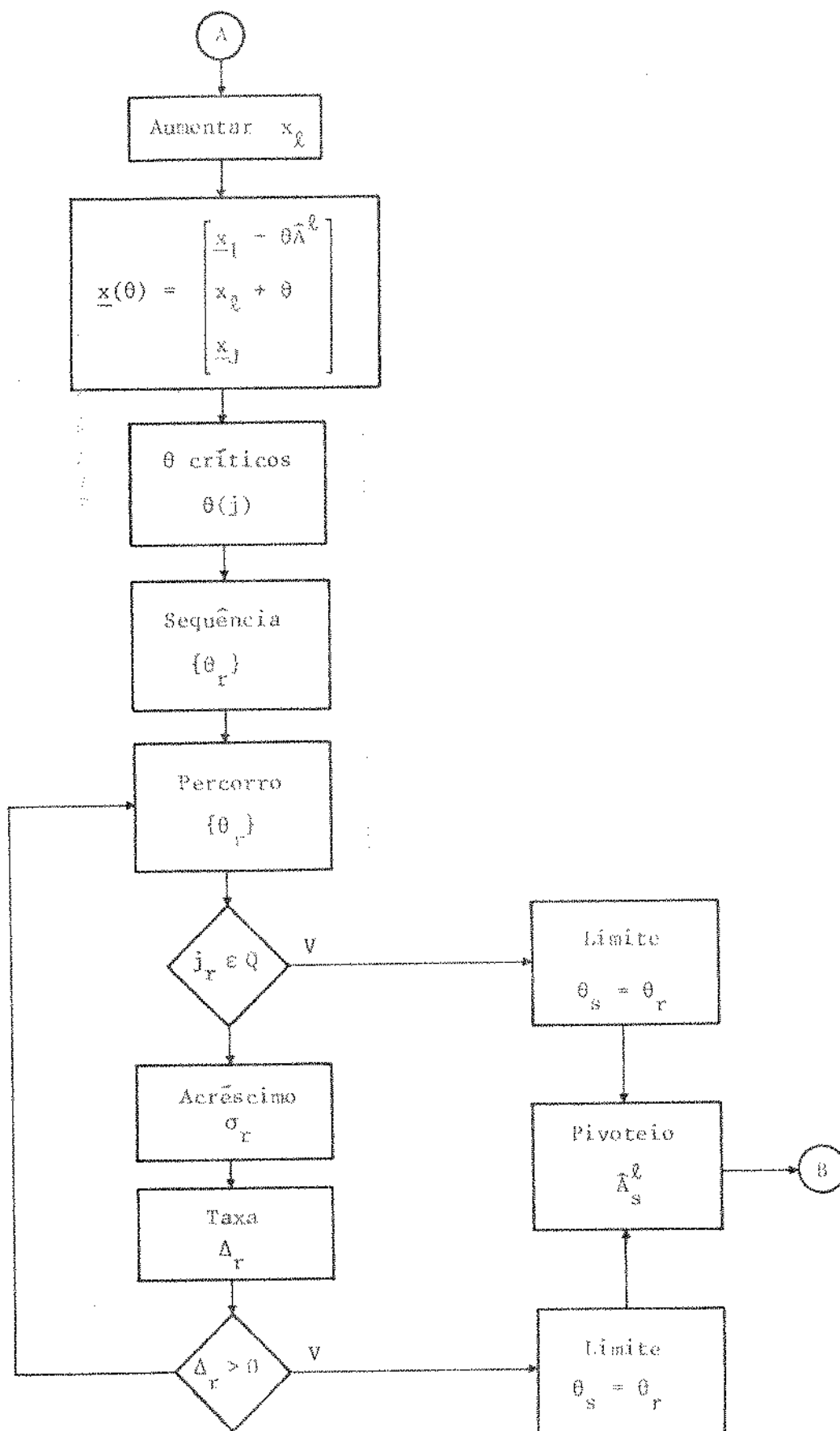
$$(*) \quad \sigma_r = 2e_j \text{Sign}(\hat{A}_t^{\ell}) \hat{A}_t^{\ell}$$

$$(*) \quad \Delta_r = \Delta + \sum_{i=1}^r \sigma_i$$

(*) Express o inalterada.

3.7 - Diagrama de Blocos





CAPÍTULO 4: ANÁLISE QUALITATIVA DE MÉTODOS ALTERNATIVOS

Fazemos uma abordagem dos métodos primal e dual-Simplex, relevando seus aspectos qualitativos, numa formulação voltada para a curva de otimalidade.

4.1 - Introdução

Reapresentamos a curva de otimalidade e estendemos da programação linear os conceitos de base, multiplicador associado a uma base e solução dual factível (LUENBERGER).

Na resolução do problema:

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

(P)

$$\text{S.a: } A\underline{x} = \underline{b}$$

$$\underline{\alpha} \leq \underline{x} \leq \underline{\beta}$$

- $A(m,n)$ de "rank" m , $f_j(x_j)$ convexa linear por partes.

Graficamente:

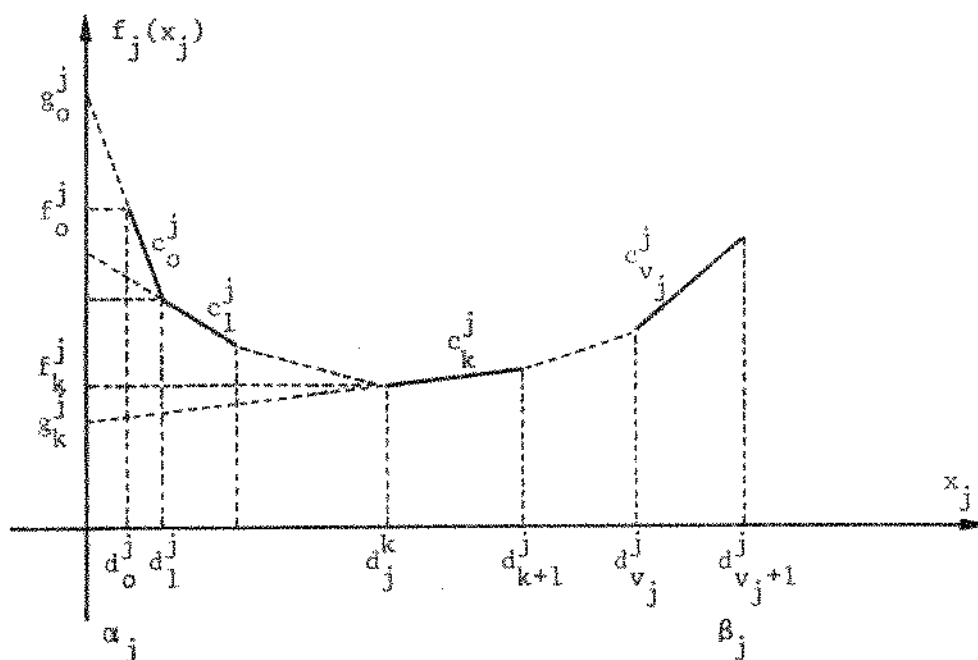


Figura 4.1

A otimalidade de uma solução é estabelecida pelo teorema (YOUFINE):

Se $\underline{x}^*$ solução factível de (P), $\underline{x}^*$ é ótima de (P), se para alguma solução $\underline{p}^*$ de (D), $\underline{x}^*$ minimiza $L(\underline{x}, \underline{p}^*)$.

A minimização de $L(\underline{x}, \underline{p}^*)$ é separável por x_j . Assim a curva de otimalidade pode ser caracterizada como o lugar geométrico dos pontos (x_j^*, z_j) , onde para dado $\underline{p}$, sendo $z_j = \underline{p}A_j^j$, x_j^* minimiza $(f_j(x_j) - z_j x_j)$ (Figura 4.2).

Os dois métodos em questão trabalham com soluções básicas. A idéia na utilização de soluções básicas é a de que se (P) tem uma solução factível ótima, ele tem uma solução básica factível ótima. Assim nos restringimos a análise das soluções básicas.

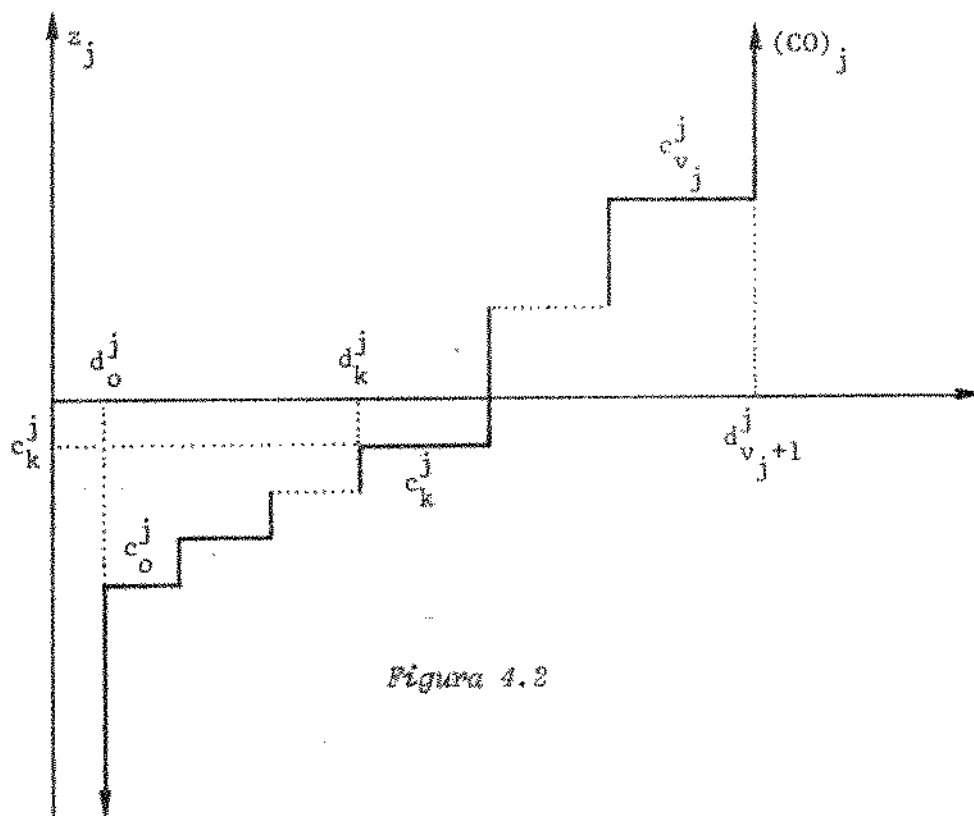


Figura 4.2

Na resolução do sistema linear $Ax = b$, com $A(m,n)$, $n > m$, dividimos as variáveis em básicas (I) e não básicas (J), fixando o valor de $(n - m)$ delas (x_J^*), em pontos críticos (d_k^J), de modo que as colunas da matriz A, correspondentes às m variáveis restantes (A^I), sejam linearmente independentes, obtendo:

$$x_J^* = d_k^J$$

$$x_I^* = \underline{b} - \hat{A}^J x_J^* \quad (4.1)$$

Sendo $\underline{b} = (A^I)^{-1} b$, $\hat{A}^J = (A^I)^{-1} A^J$.

Note que a uma mesma base correspondem várias soluções básicas. Se para algum $j \in I$, $x_j^* = d_k^j$, $k \in \{0, \dots, v_{j+1}\}$, a solução básica em questão é degenerada. Nesta análise só consideraremos soluções não degeneradas; idem $j \in J$, $z_j = c_k^j$, $k \in \{0, \dots, v_j\}$, degenerescência no dual.

Dizemos então, que as variáveis $\underline{x}_I$ são dependentes das variáveis $\underline{x}_J$. Na minimização do Lagrangeano $L(\underline{x}, \underline{p})$, se queremos torná-lo "insensível" às variáveis $\underline{x}_I$ nos respectivos intervalos, fazemos, para $j \in I$: $z_j = c_k^j$ ou

$$\underline{p}_k^I = \underline{c}_k^I \quad (4.2)$$

daí
$$\sum_{j \in I} (f_j(x_j) - z_j x_j) = g_k^j, \text{ constante}$$

Em programação linear, isto equivale a achar o vetor multiplicador que anule as componentes do vetor de custo relativo, correspondentes às variáveis básicas ($\underline{\hat{c}}^I$).

Estendendo a definição de cada $f_j(x_j)$ para pontos fora do intervalo $|\alpha_j, \beta_j|$, de maneira que para $x_j < \alpha_j$ definimos $c_{-1}^j = c_0^j$ e para $x_j > \beta_j$ definimos $c_{v_j+1}^j = c_{v_j}^j$, o conceito de vetor multiplicador fica agora associado, não somente a uma base $\{I, J\}$, mas também a uma estrutura de intervalos $\underline{k}$, sendo os k fixados para cada $j \in I$, $k \in \{-1, 0, \dots, v_j+1\}$. Assim definimos:

$$\underline{p}_k^I = \underline{c}_k^I (\underline{A}^I)^{-1} \quad (4.3)$$

vetor multiplicador da base $\{I, J\}$ associado a uma estrutura $\underline{k}$.

Com relação à otimalidade de uma variável x_j , para dado $\underline{p}$, ela é assegurada se $(x_j, z_j) \in (CO)_j$. Considerando uma solução básica $(\underline{x}_I^*, \underline{x}_J^*)$, o vetor multiplicador $\underline{p}_k^I$ (estrutura $\underline{k}$) fica bem definido (4.3). Sendo:

$$z_j^* = \underline{p}_k^I \underline{A}^j$$

para cada $j \in I$ tal que $x_j \in |\alpha_j, \beta_j|$, temos $(x_j^*, z_j^*) \in (CO)_j$, e considerando uma estrutura $\underline{\gamma}$ qualquer, de intervalos em I , com multiplicador $\underline{p}_\gamma^I$ (4.3), tal que $\gamma \in \{0, \dots, v_j\}$, suponha que ocorre um dos casos:

$$- \underline{x}_I^* \in [\underline{\alpha}_I, \underline{\beta}_I] \quad \text{e} \quad (\underline{x}_J^*, \underline{z}_J^*) \in (CO)_J \quad (4.5)$$

$$- \underline{x}_I^* \in [\underline{\alpha}_I, \underline{\beta}_I] \quad \text{e} \quad \text{existe } j \in J \text{ tq. } (\underline{x}_J^*, \underline{z}_J^*) \notin (CO)_j \quad (4.6)$$

$$- \text{Existe } \underline{\gamma} \text{ tq. } (\underline{x}_J^*, \underline{p}_{\underline{\gamma}}^{I,J}) \in (CO)_J, \text{ e existe } j \in I \text{ tq.} \\ (\underline{x}_J^*, \underline{p}_{\underline{\gamma}}^{j,J}) \notin (CO)_j \quad (4.7)$$

$$- \text{N\~ao existe } \underline{\gamma} \text{ tq. } (\underline{x}_J^*, \underline{p}_{\underline{\gamma}}^{I,J}) \in (CO)_J.$$

Desta maneira, relativamente a (P), $(\underline{x}_I^*, \underline{x}_J^*)$ tem respectivamente um dos aspectos:

- É b\'asica, fact\'ivel e \^otima.
- É b\'asica, fact\'ivel e n\~ao \^otima.
- N\~ao necessariamente fact\'ivel, corresponde \~a solu\~c\~ao dual fact\'ivel da programac\~ao linear, ou solu\~c\~ao b\'asica otimista (GARCIA).
- Nem fact\'ivel, nem otimista.

4.2 - PLP - Primal Simplex

O m\'etodo trabalha no problema primal com solu\~c\~oes b\'asicas fact\'iveis. Requer uma primeira fase de resolu\~c\~ao que consiste na procura de uma solu\~c\~ao b\'asica fact\'ivel de partida. Acrescentamos as vari\'aveis artificiais ($\underline{y}$) e trabalhamos com o sistema aumentado:

$$\underline{Ax} + \underline{Iy} = \underline{b}$$

levando primeiramente as vari\'aveis artificiais a zero.

Obtida uma solu\~c\~ao b\'asica fact\'ivel, a partir da\'i, a cada itera\~c\~ao, verificamos, para as vari\'aveis n\~ao b\'asicas, se os $(\underline{x}_j, \underline{z}_j)$ est\~ao sobre as respectivas curvas de otimalidade. Se isto ocorre para todo $j \in J$, esta solu\~c\~ao \u00e9 \^otima (4.5), caso contr\'ario (4.6), escolhemos alguma n\~ao b\'asica ($\underline{x}_q$) tal que isto

não ocorra e a alteramos mantendo fixas as demais não básicas (procedimento Simplex). Em termos da base $\{I, J\}$, se x_ℓ ao parar o faz noutro ponto crítico, não houve mudança de base; se x_ℓ pára entre pontos críticos (d_k^ℓ, d_{k+1}^ℓ) , x_ℓ entra na base.

Suponha x_ℓ tal que $z_\ell > c_k^\ell$; temos que aumentar x_ℓ construindo uma nova solução básica factível associada a um valor menor de $F(\underline{x})$. Definimos:

$$x_j(\theta) = \begin{cases} x_j - \theta \hat{A}_t^\ell & j \in I \\ x_\ell + \theta & \\ x_j & j \in J - \{ \ell \} \end{cases}$$

onde t indica a posição de x_j na base I .

A restrição $A\underline{x} = \underline{b}$ é satisfeita para θ arbitrário; desta maneira o limite para o crescimento de θ é determinado por um dos dois fatores:

- Alguma variável $x_j(\theta)$, $j \in I + \{ \ell \}$, atingir um limite, $x_j(\theta) = \alpha_j$ ou β_j .
- A função objetivo parar de diminuir com o crescimento de θ .

Durante o crescimento de θ , as variáveis x_j , $j \in I + \{ \ell \}$ passam por pontos críticos, mudando de intervalo; definimos então uma sequência $\{ \theta_r \}$ de valores críticos de θ , de maneira que entre θ_r consecutivos, as variáveis estão percorrendo intervalos lineares, isto é, $F(\underline{x}(\theta))$ varia linearmente com θ . Assim:

$$F(\underline{x}(\theta)) - F(\underline{x}) = \theta(c_k^\ell - z_\ell) \quad (4.8)$$

$$F(\underline{x}(\theta)) - F(\underline{x}) = \theta(c_k^\ell - \frac{I}{c_k} \hat{A}^\ell) \quad (4.9)$$

Sendo que a taxa de crescimento $(\Delta = c_k^\ell - z_\ell)$, é negativa, e que a cada θ_r atingido ela sofre um acréscimo positivo (σ_r) , quando esta se tornar não negativa, $F(\underline{x}(\theta))$ pára de diminuir.

Da figura anterior ressaltamos os seguintes aspectos:

- (x_ℓ, z_ℓ) se movimenta de maneira contínua na horizontal em (θ_{r-1}, θ_r) , e a cada θ_r , z_ℓ (vertical) varia bruscamente de $|\sigma_r|$.

- Considerando a trajetória descrita por (x_ℓ, z_ℓ) , vemos que a área entre esta curva e a curva de otimalidade é o decréscimo em $F(\underline{x})$ nesta iteração.

- Com relação às situações A, B, C e D apresentadas:

A: Alguma variável $x_{jr}(\theta)$ atinge um limite e bloqueia o crescimento de θ . x_{jr} sai da base e x_ℓ entra $(x_{jr}(\theta) = \alpha_j \text{ ou } \beta_j)$.

B: (x_ℓ, z_ℓ) atinge a curva de otimalidade $(CO)_\ell$. A base não se altera mas temos outra solução básica.

C: Para alguma variável x_{js} , Δ_s se torna não negativo. x_{js} sai da base e x_ℓ entra.

D: O vetor multiplicador p_k^I não se altera se o θ_r atingido corresponde a x_ℓ .

4.3 - PLP - Dual Simplex

O método dual Simplex pode ser caracterizado como um procedimento no problema primal (P), equivalente a aplicar o método primal Simplex no problema dual (D) (SIMONNARD). Trabalha com soluções básicas otimizadas (4.7), e uma solução básica otimista de partida pode ser obtida da seguinte maneira:

- Escolho uma base I e uma estrutura $\underline{y} \in \{0, \dots, v_j\}$.

- Obtenho $p_Y^I = \underline{c}_Y^I (A^I)^{-1}$.

- Calculo $z_j^* = p_Y^I A^j$.

- Para cada $j \in J$, com z_j^* , obtenho em $(CO)_j$ o x_j^* correspondente.

- Tenho então a solução básica otimista:

$$\underline{x}_J^* = \underline{d}_k^J$$

$$\underline{x}_I^* = \underline{b} - \underline{A}^J \underline{x}_J^*$$

A partir daí, a cada iteração, verificamos para as variáveis básicas se os $(\underline{x}_j^*, \underline{z}_j^*)$ estão sobre as respectivas curvas de otimalidade $(CO)_j$. Se isto ocorre para todo $j \in I$, esta solução é ótima (4.5), caso contrário (4.7) escolhemos alguma variável básica (x_ℓ) tal que isto não ocorra e alteramos $\underline{z}_\ell$, mantendo fixas as demais componentes de $\underline{z}_I$ (procedimento Simplex). Terminada a alteração em $\underline{z}_\ell$, em termos de base $\{I, J\}$, se x_ℓ está agora num ponto crítico, x_ℓ sai da base, caso contrário não há mudança de base.

Suponha $x_\ell < d_Y^\ell$, sendo $\underline{z}_\ell^* = \underline{p}_Y^I A_\ell^\ell$, temos que alterar $\underline{p}_Y^I$ construindo uma nova solução básica otimista, associada a um valor maior de $\phi(\underline{p})$. Definimos:

$$\underline{z}_\ell(\theta) = \underline{z}_\ell^* - \theta$$

$$\underline{z}_j(\theta) = \underline{z}_j^* \quad j \in I - \{ \ell \}$$

Resulta a variação em $\underline{p}_Y^I$:

$$\underline{p}(\theta) = \underline{p}_Y^I - \theta (\underline{A}^I)_\ell^{-1} \quad (4.10)$$

onde crescemos θ até que $\phi(\underline{p})$ pára de crescer com θ .

A referência $\underline{z}_j(\theta)$ das variáveis não básicas resulta em:

$$\underline{z}_j(\theta) = \underline{p}(\theta) \underline{A}^j \quad j \in J \quad (4.11)$$

Durante o crescimento de θ , os z_j , $j \in J + \{ \ell \}$, passam por pontos críticos determinados na curva de otimalidade pela mudança do correspondente x_j^* . Definimos então uma sequência $\{\theta_r\}$ de valores críticos de θ , de maneira que entre θ_r consecutivos $\phi(p)$ varia linearmente com θ . Assim:

$$\phi(\underline{p}(\theta)) - \phi(\underline{p}_Y^I) = \theta(d_Y^\ell - x_\ell^*)$$

Sendo que a taxa de crescimento ($\Delta = d_Y^\ell - x_\ell^*$), é positiva, e a cada θ_r atingido ela sofre um acréscimo negativo (σ_r); quando esta (Δ_r) se tornar não positiva, $\phi(p)$ pára de crescer.

Desta maneira Δ_r pode ser considerada como uma medida de factibilidade nesta iteração.

Cada vez que o θ_r atingido corresponde a uma variável não básica, o ponto crítico x_j^* é alterado de acordo com a curva de otimalidade de maneira que as variáveis não básicas permanecem ótimas durante todo o processo. Com θ percorrendo $\{\theta_r\}$, representamos graficamente a "evolução" de (x_ℓ, z_ℓ) relativamente à curva de otimalidade (figura 4.4).

Na figura 4.4 ressaltamos os seguintes aspectos:

- (x_ℓ, z_ℓ) se movimenta de maneira contínua na vertical em (θ_{r-1}, θ_r) e a cada θ_r , x_ℓ (horizontal) varia bruscamente de $|\sigma_r|$.

- Considerando a trajetória descrita por (x_ℓ, z_ℓ) , vemos que a área entre esta curva e a curva de otimalidade é o acréscimo em $\phi(p)$ nesta iteração.

- Com relação às situações A, B e C apresentadas:

A: (x_ℓ, z_ℓ) atinge a curva de otimalidade $(CO)_\ell$, a base não se altera mas temos outra solução básica.

B: Para alguma variável z_{js} , Δ_s se torna não positiva. x_{js} entra na base e x_ℓ sai.

C: x_j^* não se altera se o θ_r atingido corresponde a z_ℓ .

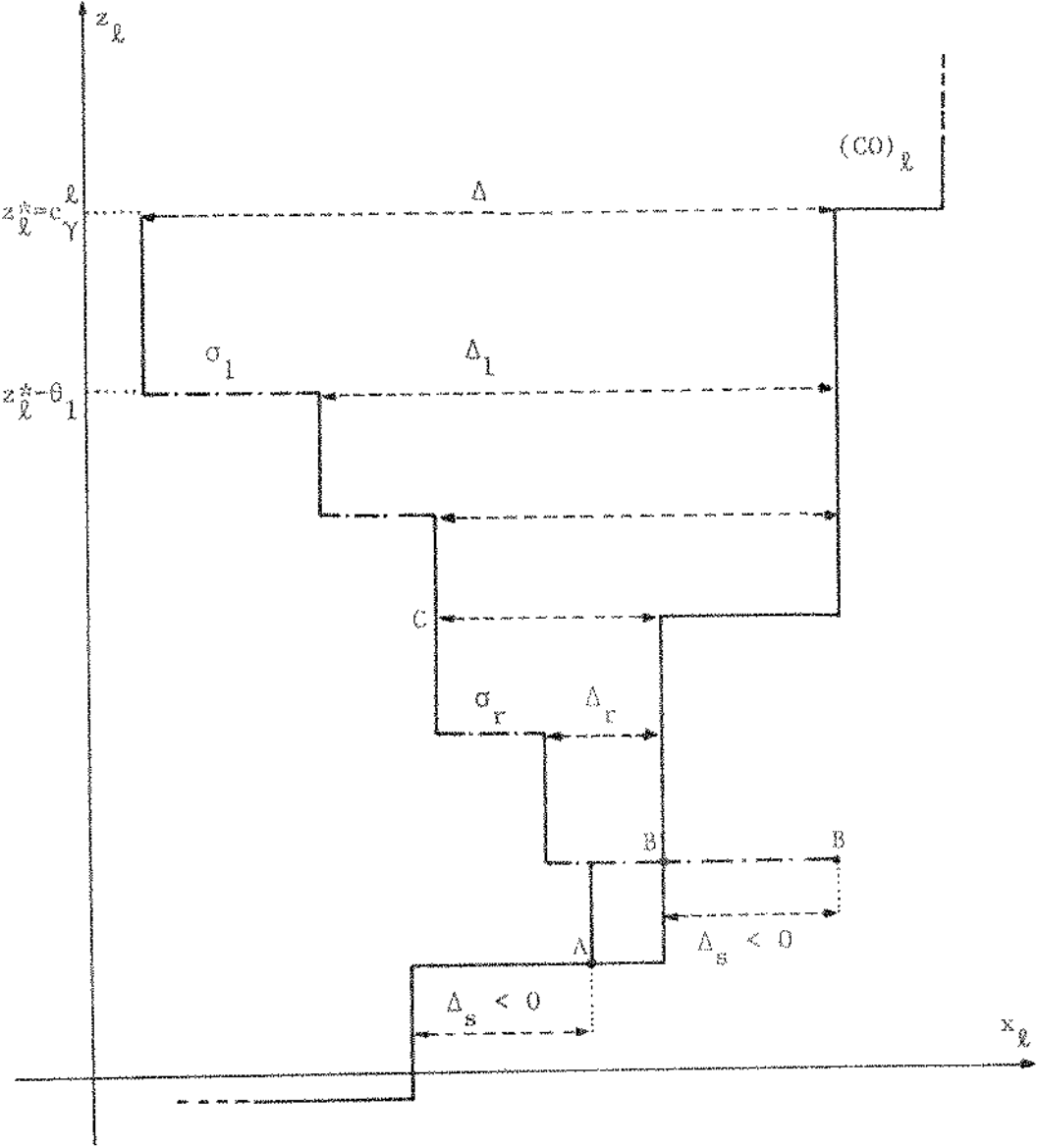


Figura 4.4

4.4 - Primal, Dual, Primal-Dual, Quadro Comparativo

Para os métodos primal-Simplex, dual-Simplex e Primal-Dual, usados na resolução do problema (P), fazemos um quadro comparativo com alguns aspectos qualitativos relevantes. Foram considerados os seguintes aspectos:

INICIALIZAÇÃO:

O método Primal-Simplex, no caso geral, exige uma primeira fase de resolução, que consiste na procura de uma solução básica factível; isto requer, em média, tanto trabalho quanto a otimização propriamente dita. Esta fase inicial corresponde a um problema de minimização que trabalha no sistema aumentado $Ax + y = b$, penalizando as variáveis artificiais (y). Temos então um método que antes de ser aplicado na resolução do problema (P), exige que o sistema $Ax = b$ não seja redundante, isto é, $A(m,n)$ de "rank" m , e que encontremos uma solução básica que satisfaça $Ax = b$ e $\alpha < x < \beta$ ou detectemos que (P) é infactível.

O método dual Simplex exige uma fase inicial mais simples. Corresponde ao procedimento descrito no item 4.3 onde após m pivoteamentos obtemos a solução básica otimista de partida ou detectamos redundância em $Ax = b$. Note que a restrição $\alpha < x < \beta$ é agora relaxada.

A inicialização do primal dual é trivial, pois trabalha no problema dual (irrestrito). A restrição $Ax = b$ é relaxada. Assim o método não requer soluções factíveis nem básicas; resolve problemas com $Ax = b$ redundante.

CAMINHO: Considerando que as soluções básicas são vértices no gráfico da função linear por partes e que temos arestas ligando vértices adjacentes, dizemos que os métodos primal Simplex e dual Simplex "caminham" nas arestas de $F(x)$ e $\phi(p)$ respectivamente. O método primal-dual caminha em $\phi(p)$ segundo a direção de máximo crescimento (u^*).

FACTIBILIDADE: Consideramos as restrições que são satisfeitas no recurso da otimização.

OTIMALIDADE: Consideramos os pontos (x_j, z_j) relativamente à curva de otimalidade $(CO)_j$. No caso do primal e dual Simplex seja $\{I, J\}$ a base.

ÓTIMO: Consideramos as condições que determinam o final da otimização, isto é, a solução ótima.

	Primal Simplex	Dual Simplex	Primal-Dual
Inicialização	Fase Inicial: Básica Factível	Fase Inicial: Básica Otimista	Trivial
Caminho	$F(\underline{x})$ Arestas	$\phi(\underline{p})$ Arestas	$\phi(\underline{p})$ Máximo Crescimento
Factibilidade	$\underline{Ax} = \underline{b}$ $\underline{\alpha} < \underline{x} < \underline{\beta}$	$\underline{Ax} = \underline{b}$	$\underline{\alpha} < \underline{x} < \underline{\beta}$
Otimalidade	$(\underline{x}_I, \underline{z}_I) \in (CO)_I$	$(\underline{x}_J, \underline{z}_J) \in (CO)_J$	$(\underline{x}_j, \underline{z}_j) \in (CO)_j$ $j = 1, n$
Ótimo	$(\underline{x}_J, \underline{z}_J) \in (CO)_J$	$(\underline{x}_I, \underline{z}_I) \in (CO)_I$	$\underline{Ax} = \underline{b}$

APÊNDICE 1: (PAD) É O DUAL DE (PA)

Dado o problema (P):

$$\text{MINIM } F(\underline{x})$$

$$\text{S.a: } A\underline{x} = \underline{b}$$

$$\underline{x} \in R$$

Definimos o seu problema dual como:

$$\text{MAXIM } \phi(\underline{u})$$

$$\text{S.a: } \underline{u} \in D$$

onde:

- $\underline{u}$, vetor de variáveis duais.
- $\phi(\underline{u}) = \text{MINIM}_{\underline{x} \in R} L(\underline{x}, \underline{u})$, função dual
- $D = \left\{ \underline{u}: \text{MINIM}_{\underline{x} \in R} L(\underline{x}, \underline{u}) \text{ existe} \right\}$
- $L(\underline{x}, \underline{u}) = F(\underline{x}) + \underline{u}(\underline{b} - A\underline{x})$, função Lagrangeano associada a (P).

Assim, dado o problema (PA):

$$\text{MINIM } h(\underline{y}) = \sum_{j=1}^m e_j |y_j| \quad , \quad \underline{e} > \underline{0}$$

$$\text{S.a: } A^Q_{\underline{x}_Q} + \underline{y} = \underline{b} - A^{\overline{Q}}_{\underline{d}_k} \overline{Q}$$

$$\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q$$

$$L(\underline{x}_Q, \underline{y}, \underline{u}) = \sum_{j=1}^m e_j |y_j| + \underline{u}(\underline{b} - A^{\overline{Q}}_{\underline{d}_k} \overline{Q} - A^Q_{\underline{x}_Q} - \underline{y}) \quad (\text{A.1})$$

$$L(\underline{x}_Q, \underline{y}, \underline{u}) = \sum_{j=1}^m (e_j |y_j| - u_j y_j) + \underline{u}(\underline{b} - A^{\overline{Q}}_{\underline{d}_k} \overline{Q} - A^Q_{\underline{x}_Q}) \quad (\text{A.2})$$

$$\phi(\underline{u}) = \text{MINIM}_{\underline{y}_j} \sum_{j=1}^m (e_j |y_j| - u_j y_j) + \quad (\text{A.3})$$

$$+ \text{MINIM}_{\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q} \left\{ \underline{u}(\underline{b} - A^{\overline{Q}}_{\underline{d}_k} \overline{Q} - A^Q_{\underline{x}_Q}) \right\} \quad (\text{A.4})$$

No cálculo de (A.3), resolvemos os problemas restritos:

$$\text{MINIM}_{y_j \geq 0} \sum_{j=1}^m (e_j - u_j) y_j \quad (\text{A.5})$$

$$\text{MINIM}_{y_j \leq 0} \sum_{j=1}^m -(e_j + u_j) y_j \quad (\text{A.6})$$

(A.5) e (A.6) assumem valor finito (zero) nos casos:

$$(e_j - u_j) \geq 0 \quad \rightarrow \quad u_j \leq e_j \quad (\text{A.7})$$

$$(e_j + u_j) \geq 0 \quad \rightarrow \quad u_j \geq -e_j \quad (\text{A.8})$$

(A.7) e (A.8) podem ser reunidos em:

$$|u_j| \leq e_j \quad (\text{A.9})$$

Como a minimização (A.4) assume valor finito para qualquer $\underline{u}$, obtemos:

$$D = \{ \underline{u}: |u_j| \leq e_j, \quad j = 1, m \} \quad (\text{A.10})$$

e sendo:

$$\| \underline{u} \| = \text{MINIM} \{ t: -te \leq \underline{u} \leq te, \quad e > 0, \quad t > 0 \} \quad (\text{A.11})$$

$$D = \{ \underline{u}: \| \underline{u} \| \leq 1 \} \quad (\text{A.12})$$

Resulta:

$$\phi(\underline{u}) = \text{MINIM}_{\substack{\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q}} \left\{ \underline{u}(\underline{b} - A_{\underline{d}_k^Q}^Q - A_{\underline{x}_Q}^Q) \right\} \quad (\text{A.13})$$

e o problema dual (PAD):

$$\text{MAXIM}_{\underline{u}} \left\{ \text{MINIM}_{\underline{x}_Q} \left\{ \underline{u}(\underline{b} - A_{\underline{d}_k^Q}^Q - A_{\underline{x}_Q}^Q) \right\} \right\}$$

$$\text{S.a:} \quad \| \underline{u} \| = 1$$

$$\underline{d}_k^Q \leq \underline{x}_Q \leq \underline{d}_{k+1}^Q$$

Bibliografia

- BEALE, E.M.L. "An alternative method of linear programming" *Proceedings of the Cambridge Philosophical Society*, Vol.50, Part 4, 1954.
- DANTZIG, G.B.; FORD, L.R. and FULKERSON, D.R. "A primal-dual algorithm for linear programs" Editors: KUHN, H.W. and TUCKER, A.W. *Linear Inequalities and Related Systems - Annals of Mathematics Study*, No.58, Article 7, Princeton University Press, 1956.
- DANTZIG, G.B. *Applications et Prolongements de la Programmation Linéaire* Dunod, Paris, 1966.
- GARCIA, A.S. "Método dual-Simplex linear por partes" Tese de Mestrado, UNICAMP, IMEC, 1978.
- HADLEY, G. *Linear Programming* Addison Wesley, 1962.
- LASDON, L.S. *Optimization Theory for Large Systems* MacMillan, 1970.
- LUENBERGER, D.G. *Introduction to Linear and Nonlinear Programming* Addison Wesley, 1973.
- SAKAROVITCH, M. *Notes on Linear Programming* Van Nostrand, 1971.
- SIMONNARD, M. *Programmation Linéaire* Dunod, Paris, 1972.
- YODINE, D. et GOLSTEIN, E. *Problemes Particuliers de la Programmation Linéaire* MIR, 1973.