

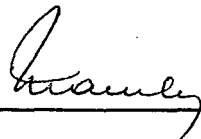
RESOLUÇÃO DE PROGRAMAS PSEUDO-BOOLEANOS
NÃO-LINEARES VIA DECOMPOSIÇÃO LAGRANGEANA

Ismael Regis de Farias Júnior

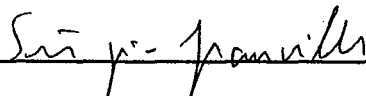
3

Tese submetida ao corpo docente da Coordenação dos Programas de Pós-Graduação em Engenharia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências em Engenharia de Sistemas e Computação.

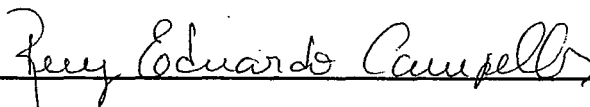
Aprovada por



Prof. Nelson Maculan Filho, D.Sc., D. Habil.
(Presidente)



Prof. Sérgio Granville, Ph.D.



Prof. Ruy Eduardo Campello, D. Sc.

Rio de Janeiro, RJ - Brasil

Agosto de 1990

de FARIAS, ISMAEL REGIS JÚNIOR

Resolução de Programas Pseudo-booleanos via decomposição lagrangeana. [Rio de Janeiro] 1990.

VIII, 94 p. 29.7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 1990).

Tese - Universidade Federal do Rio de Janeiro, COPPE.

1. Programação Matemática; Programação não-linear; Programação Quadrática; Programação Inteira; Programação Booleana; Programação Pseudo-Booleana; Relaxação Lagrangeana; Problema da Mochila Quadrática.

I. COPPE/UFRJ. II. Título (Série)

A Salomão de Lima Macedo,
meu avô.

AGRADECIMENTOS

Ao Professor Nelson Maculan Filho, pela proposição do tema e orientação deste trabalho.

Aos professores Sérgio Granville e Ruy Campello pela participação na banca examinadora desta tese.

A todos os colegas, professores e funcionários da Coppe, que com seu apoio muito contribuíram para este trabalho.

Aos meus amigos, dentre eles Janet Soares Moraes, Giovani dos Santos Lopes, Mônica Patrícia da Costa Viana e Edson Pinho da Silva, e de modo muito especial Ivan José do Couto Esteves e Samuel da Mata.

A meus pais, minha irmã e todos os meus parentes.

Sou acima de tudo grato a Deus, a quem devo tudo que tenho e tudo que sou.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências (M.Sc.).

RESOLUÇÃO DE PROGRAMAS PSEUDO-BOOLEANOS
VIA DECOMPOSIÇÃO LAGRANGEANA

Ismael Regis de Farias Júnior

Agosto de 1990

Orientador: Nelson Maculan Filho

Programa: Engenharia de Sistemas e Computação

É estudado o uso da estratégia de Decomposição Lagrangeana na obtenção de limites para o ótimo de problemas de programação pseudo-booleana com restrições lineares. Esta técnica é aplicada ao problema da Mochila Quadrática, para o qual formulamos um algoritmo enumerativo e obtemos resultados computacionais.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfilment of the requirements for the degree of Master of Science (M.Sc.).

RESOLUTION OF PSEUDO-BOOLEAN PROGRAMS
VIA LAGRANGEAN DECOMPOSITION

Ismael Regis de Farias Júnior

August, 1990

Thesis Supervisor: Nelson Maculan Filho

Department: Systems Engineering and Computing.

We study the strategy of Lagrangean Decomposition for the obtainence of bounds related to the optimum of Nonlinear Pseudo-Boolean Programs with linear constraints. This technique is applied to the Quadratic Knapsack Problem, to which we propose an enumerative algorithm and obtain computational results.

íNDICE

Capítulo I - Introdução.....	01
Capítulo II - Programação Pseudo-Booleana.....	03
II.1 - Álgebra de Boole.....	03
II.2 - Métodos de Programação Pseudo-Booleana não linear..	10
II.2.1 - Redução a outros problemas.....	10
II.2.2 - Métodos Algébricos.....	17
II.2.3 - Métodos Enumerativos.....	19
II.2.4 - Métodos de Corte.....	23
II.3 - Aplicações.....	24
Capítulo III - Dualidade Lagrangeana e seu uso em Programação Discreta.....	26
III.1 - Definições e resultados preliminares.....	26
III.2 - Propriedades da Função Dual.....	30
III.3 - Métodos de Subgradiente.....	42
III.3.1 - O Algoritmo da série divergente.....	44
III.3.2 - O Método do Passo Constante.....	47
III.3.3 - O Método de Relaxação.....	51
Capítulo IV - Utilização de Decomposição Lagrangeana na obtenção de Limites para o ótimo de Programas Pseudo-Booleanos não-lineares com vínculos lineares.....	55
IV.1 - Introdução.....	55
IV.2 - Resultados Teóricos.....	60
Capítulo V - O Problema da Mochila Quadrática.....	65

V.1 - Introdução.....	65
1.2 - Obtenção de limites via Decomposição Lagrangeana....	68
V.3 - Implementação Computacional.....	73
V.4 - Resultados Computacionais.....	75
V.5 - Conclusão.....	77
Referência Bibliográficas.....	81

CAPÍTULO I

INTRODUÇÃO

Foi no fim dos anos 30, quando Claude Shannon aplicou a álgebra de Boole ao estudo de circuitos de chaveamento com relés, que a teoria das álgebras booleanas encontrou sua primeira aplicação prática. Sendo circuitos de chaveamento, de um modo geral, compostos por dispositivos de dois estados eles naturalmente deveriam ser tratados por um ferramental matemático que manipulasse variáveis bivalentes. Tal ferramental encontrado foi a álgebra de dois elementos de Boole.

É de fato uma idéia natural usar variáveis bivalentes para estudar sistemas nos quais algumas componentes possuam apenas dois estados possíveis. Exemplos de tais sistemas abundam no mundo real e foi George B. Dantzig que em 1957 pela primeira vez chamou a atenção para a importância, variedade e profundidade dos problemas de Pesquisa Operacional envolvendo estes sistemas. Desde então um enorme número de trabalhos têm sido publicados a respeito deste assunto, tratando desde seus aspectos matemáticos mais abstratos até suas aplicações. Em particular, muita pesquisa passou a ser dedicada à obtenção de técnicas que permitam resolver estes problemas de maneira econômica.

Seguindo esta linha, neste trabalho estudamos o uso da estratégia de Decomposição Lagrangeana na resolução de Programas Pseudo-Booleanos não-lineares com restrições lineares, e o aplicamos ao problema da mochila quadrática.

A organização deste trabalho é a seguinte. No capítulo II é feita uma revisão das idéias mais importantes a respeito dos problemas de otimização Pseudo-Booleana não-linear, e são apresentadas algumas aplicações.

No capítulo III é feita uma digressão a respeito do uso de Relaxação Lagrangeana em otimização não-diferenciável.

No capítulo IV é estudada a estratégia de Decomposição Lagrangeana para a resolução de programas Pseudo-Booleanos não-lineares com restrições lineares.

No capítulo V é tratado o problema da mochila quadrática. Apresentamos um algoritmo enumerativo que utiliza a técnica descrita no capítulo IV para a obtenção de limites, e apresentamos resultados computacionais. No final do capítulo formulamos conclusões e sugestões para prosseguimento deste trabalho.

capítulo II

PROGRAMAÇÃO PSEUDO-BOOLEANA

Neste capítulo definimos Programa Pseudo-Booleano e obtemos alguns resultados básicos que serão utilizados no decorrer deste trabalho. Isto é feito na primeira seção. Na segunda seção fazemos uma breve revisão de alguns métodos existentes na literatura, e na terceira seção apresentamos algumas aplicações.

II.1) Álgebra de Boole

Nesta seção introduzimos os conceitos de Álgebra Booleana, função Booleana e função Pseudo-Booleana. Apresentamos alguns teoremas referentes à representação de funções Booleanas e Pseudo-Booleanas, e definimos programa Pseudo-Booleano.

Definição II.1 Definimos uma álgebra de Boole como um conjunto Σ com pelo menos dois elementos distintos σ_0 e σ_1 , que admite três operações, chamadas negação, disjunção e conjunção, denotadas por \bar{a} , $a \cup b$ e $a \cap b$ respectivamente, tal que $\forall a, b, c \in \Sigma$ temos

$$(i) \quad \bar{\bar{a}} \in \Sigma, \quad a \cup b \in \Sigma, \quad a \cap b \in \Sigma$$

$$(ii) \quad a \cup b = b \cup a, \quad a \cap b = b \cap a$$

$$(iii) \quad a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$$

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$$

$$(iv) \quad a \cup \sigma_0 = a, \quad a \cap \sigma_1 = a$$

$$(v) \quad a \cup \bar{a} = \sigma_1, \quad a \cap \bar{a} = \sigma_0$$

EXEMPLO II.1 O conjunto $\mathbb{B} = \{0,1\}$ com as operações

$$\bar{a} = 1 - a \quad (II.1)$$

$$a \cup b = a + b - a \cdot b \quad (II.2)$$

$$a \cap b = a \cdot b \quad (II.3)$$

$a, b \in \mathbb{B}$, e com $\sigma_0 = 0, \sigma_1 = 1$ é uma álgebra de Boole. ■

OBSERVAÇÃO II.1 Daqui para diante restringir-nos-emos à álgebra booleana de dois elementos definida no exemplo anterior. ■

DEFINIÇÃO II.2 Sejam \mathbb{B}^n o produto cartesiano de \mathbb{B} por si mesmo n vezes. Chamamos de função booleana a toda aplicação $f: \mathbb{B}^n \rightarrow \mathbb{B}$. ■

Em outras palavras, uma função booleana é uma função com valores e argumentos em \mathbb{B} . Este conceito é ilustrado no exemplo a seguir.

EXEMPLO II.2 A função f definida pela tabela abaixo

x	y	z	w	f
0	0	0	0	0
0	0	0	1	0
0	0	1	0	1
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	1
1	0	0	0	0
1	0	0	1	0
1	0	1	0	1
1	0	1	1	0
1	1	0	0	0
1	1	0	1	1
1	1	1	0	1
1	1	1	1	1

é uma função booleana. ■

Note que a função definida no exemplo acima pode também ser expressa por uma combinação de negações, disjunções e conjunções de seus argumentos. De fato, podemos escrever

$$\begin{aligned}
 f(x,y,z,w) = & (\bar{x} \cap \bar{y} \cap z \cap \bar{w}) \cup (\bar{x} \cap y \cap z \cap w) \cup \\
 & (x \cap \bar{y} \cap z \cap \bar{w}) \cup (x \cap y \cap \bar{z} \cap w) \cup \\
 & (x \cap y \cap z \cap \bar{w}) \cup (x \cap y \cap z \cap w) \quad \text{(II.4)}.
 \end{aligned}$$

Este resultado não é válido apenas para algumas funções mas trata-se de um resultado geral, conforme demonstraremos a seguir. Antes porém enunciaremos a definição abaixo, apresentada em HAMMER e RUDEANU (1968, p. 8), de uma expressão booleana.

DEFINIÇÃO II.3

- 1) 0 e 1 são expressões booleanas;
- 2) As variáveis x_1, \dots, x_n , $x_i \in \mathbb{B} \forall i \in \{1, \dots, n\}$, são expressões booleanas;
- 3) Se E_1 e E_2 são expressões booleanas, então \bar{E}_1 , $E_1 \cup E_2$ e $E_1 \cap E_2$ também o são. ■

Assim, por exemplo, quaisquer combinações de negações, disjunções e conjunções dos argumentos de uma função são também expressões booleanas.

OBSERVAÇÃO II.2 Adotaremos daqui para diante a notação

$$x^0 = \bar{x}, x^1 = x \quad (\text{II.5})$$

para a operação de negação, apresentada em HAMMER e RUDEANU (1968, p. 7). ■

TEOREMA II.1 Toda função booleana pode ser escrita como uma expressão booleana.

DEMONSTRAÇÃO Seja $\Lambda = \{(\alpha_1, \dots, \alpha_n) \in \mathbb{B}^n / f(\alpha_1, \dots, \alpha_n) = 1\}$. Sejam ainda x_1, \dots, x_n os argumentos de $f: \mathbb{B}^n \rightarrow \mathbb{B}$ e consideremos o conjunto de expressões $\varepsilon = \{x_1^{\alpha_1} \cap \dots \cap x_n^{\alpha_n} / (\alpha_1, \dots, \alpha_n) \in \Lambda\}$. Como

$$x_1^{\alpha_1} \cap \dots \cap x_n^{\alpha_n} = \begin{cases} 1, & \text{se } x_i = \alpha_i \quad \forall i \in \{1, \dots, n\} \\ 0, & \text{caso contrário} \end{cases} \quad (\text{II.6})$$

vem que as expressões em ε são todas iguais a 0 para qualquer ponto (x_1, \dots, x_n) que não pertença a Λ . Por outro lado, para todo ponto (x_1, \dots, x_n) pertencente a Λ existe uma expressão em ε cujo valor é 1. Uma vez que $1 \cup 0 = 1$ e $0 \cup \dots \cup 0 = 0$, f será igual à expressão formada pela disjunção de todos os elementos de ε . Denotando tal disjunção por $\alpha_1 \cup \dots \cup \alpha_n$ ($x_1^{\alpha_1} \cap \dots \cap x_n^{\alpha_n}$) podemos escrever

$$f(x_1, \dots, x_n) = \alpha_1 \cup \dots \cup \alpha_n \quad (x_1^{\alpha_1} \cap \dots \cap x_n^{\alpha_n}) \quad (\text{II.7})$$

o que demonstra o teorema acima. ■

OBSERVAÇÃO II.2 Podemos associar a uma mesma função diversas expressões diferentes. Por exemplo, além da expressão (II.4) podemos associar à função definida no exemplo (II.2) a expressão

$$\begin{aligned} f(x, y, z, w) = & (x \cup y \cup z \cup w) \cap (x \cup y \cup z \cup \bar{w}) \cap \\ & (x \cup y \cup \bar{z} \cup \bar{w}) \cap (x \cup \bar{y} \cup z \cup w) \cap \\ & (x \cup \bar{y} \cup z \cup \bar{w}) \cap (x \cup \bar{y} \cup \bar{z} \cup w) \cap \\ & (\bar{x} \cup y \cup z \cup w) \cap (\bar{x} \cup y \cup z \cup \bar{w}) \cap \\ & (\bar{x} \cup y \cup \bar{z} \cup \bar{w}) \cap (\bar{x} \cup \bar{y} \cup z \cup w) \quad (\text{II.8}) \quad \blacksquare \end{aligned}$$

OBSERVAÇÃO II.3 Com as operações de negação, disjunção e conjunção definidas no exemplo (II.1), o teorema acima implica que toda função booleana pode ser escrita como um polinômio de seus argumentos. ■

Estenderemos agora o conceito de função booleana para funções com argumentos em \mathbb{B} e valores no corpo dos reais. Visto que toda função booleana pode ser enquadrada nesta classe, nós as chamaremos de Funções Pseudo-Booleanas.

DEFINIÇÃO II.4 Chamamos de Função Pseudo-Booleana a toda aplicação $f: \mathbb{B}^n \rightarrow \mathbb{R}$. ■

como no caso de funções booleanas, também as funções pseudo-booleanas podem ser expressas (de maneira unívoca) como polinômio de seus argumentos. Este importante resultado é estabelecido no teorema a seguir, devido a Toma Gaspar.

TEOREMA II.2 (T. GASPAR) Toda função pseudo-booleana $f: \mathbb{B}^n \rightarrow \mathbb{R}$ pode ser escrita de maneira unívoca (após as devidas simplificações) como

$$f(x) = \sum_{k=1}^p c_k T_k, \quad c_k \in \mathbb{R}, \quad T_k = \prod_{j \in N_k} x_j, \quad N_k \subset \{1, \dots, n\},$$

$$k = 1, \dots, p \quad \text{(II.9)}$$

DEMONSTRAÇÃO HAMMER e RUDEANU (1968, pp. 21 - 22) ■

De posse dos conceitos acima podemos agora definir Programa Pseudo-Booleano.

DEFINIÇÃO II.5 Chamamos de Programa Pseudo-Booleano (PPB) a todo problema

$$(PPB) \left\{ \begin{array}{l} \text{minimizar } f(x) \\ \Omega = \{x \in B^n / g_i(x) \leq b_i, i = 1, \dots, m\} \end{array} \right.$$

onde $f: B^n \rightarrow \mathbb{R}$ e $g_i: B^n \rightarrow \mathbb{R}$ são funções booleanas e b_i é um número real, $\forall i \in \{1, \dots, m\}$.

f será chamada função-objetivo do problema, e as desigualdades $g_i(x) \leq b_i, i = 1, \dots, m$, serão chamadas vínculos ou restrições. Os pontos de Ω são chamados pontos viáveis. Aqueles que resolvem o (PPB) são chamados soluções ótimas do problema, e o valor de f numa solução ótima é chamado valor ótimo do (PPB). Se $\Omega = \emptyset$ dizemos que o problema é inconsistente. Dois programas pseudo-booleanos (PPB)¹ e (PPB)² com mesmo número de variáveis são ditos equivalentes se possuem as mesmas soluções ótimas e o mesmo valor ótimo.

OBSERVAÇÃO II.4 De acordo com o teorema (II.2) podemos expressar todo programa pseudo-booleano (de maneira unívoca) como

$$(PPB)_p \left\{ \begin{array}{l} \text{minimizar } \left\{ \sum_{k=1}^p c_k T_k \right\} \\ \Omega = \{x \in B^n / \sum_{l=1}^p a_{lk} T_l \geq b_l, l = 1, \dots, m\} \\ T_k = \prod_{j \in N^k} x_j, N^k \subset \{1, \dots, n\}, k = 1, \dots, p \\ T_{lk} = \prod_{j \in N^{lk}} x_j, N^{lk} \subset \{1, \dots, n\}, k = 1, \dots, p_l \end{array} \right.$$

e o chamamos Programa Pseudo-Booleano Polinomial.

II.2) MÉTODOS DE PROGRAMAÇÃO PSEUDO-BOOLEANA NÃO-LINEAR

Nesta seção fazemos uma breve revisão das principais técnicas existentes na literatura para resolução de programas pseudo-booleanos não-lineares.

Segundo HANSEN (1979) e HANSEN, JAUMARD e MATHON (1989) estas técnicas podem ser colocadas em quatro grupos:

- (i) Redução a outros problemas
- (ii) Métodos algébricos
- (iii) Métodos enumerativos
- (iv) Métodos de corte.

No primeiro caso tenta-se reformular o problema, tornando-o em outro mais simples, ou para o qual já existam métodos de resolução conhecidos e eficientes. Nos demais casos, tenta-se obter as soluções a partir da formulação original do problema.

II.2.1) REDUÇÃO A OUTROS PROBLEMAS

Nem sempre é conveniente resolver um programa pseudo-booleano na forma em que ele se apresenta. Por isso às vezes, antes de resolver um certo (PPB) é interessante modificá-lo em outra forma e resolver o novo problema (talvez usando uma das três outras técnicas enumeradas no início desta seção).

Conforme visto na observação (II.4) da seção

anterior, todo programa pseudo-booleano não-linear pode ser escrito como um programa polinomial nas mesmas variáveis bivalentes e com o mesmo número de restrições.

BAR (1972) e HAMMER e ROSENBERG (1972) observaram que todo $(PPB)_p$ é equivalente a um problema da mochila polinomial nas mesmas variáveis bivalentes. De fato, seja Φ uma função booleana dada por

$$\Phi(x) = \begin{cases} 0, & \text{se } x \text{ é um ponto viável} \\ 1, & \text{caso contrário} \end{cases} \quad (\text{II.10}).$$

Conforme visto na observação (II.3) da seção anterior, Φ pode ser escrita como um polinômio em seus argumentos, e o $(PPB)_p$ em questão poderá ser escrito como

$$(PPB)_M \quad \begin{cases} \text{minimizar } f(x) \\ x \in \Omega_M \\ \Omega_M = \{ x \in \mathbb{B}^n / \Phi(x) = 0 \}. \end{cases}$$

É óbvio entretanto, que o trabalho necessário para se obter Φ poderá não ser compensador.

Também todo programa pseudo-booleano polinomial com coeficientes racionais pode ser escrito como um programa pseudo-booleano polinomial sem vínculos. HAMMER ROSENBERG e RUDEANU (1963b) e HANSEN (1979) mostraram que todo vínculo de igualdade $h_i(x) = 0$, $i = 1, \dots, m_1$ pode ser removido à moda lagrangeana. De fato, relaxando o problema em tais vínculos e adicionando à função-objetivo o termo de

penalidade

$$C \left(\sum_{k=1}^p \max(0, c_k) - \sum_{k=1}^p \min(0, c_k) + 1 \right) \cdot \sum_{i=1}^m (h_i(x))^2$$

obtemos um problema equivalente ao primeiro. Para os vínculos em que $h_i(x) \geq 0 \forall x \in \mathbb{B}^n$ não há necessidade de se elevar $h_i(x)$ ao quadrado.

Por outro lado, vínculos de desigualdade envolvendo polinômios com coeficientes inteiros podem também ser transformados em vínculos de igualdade às custas da introdução de novas variáveis. Para isso, seja o vínculo

$$g_i(x) = \sum_{k=1}^{p_i} a_{ik} \cdot T_{ik} \geq b_i \quad (\text{II.11}).$$

Tomando

$$l = \lceil \log_2 \left(\sum_{k=1}^{p_i} \max(a_{ik}, 0) - b_i \right) \rceil \quad (\text{II.12}),$$

(II.11) pode ser escrito como

$$g_i(x) - \sum_{k=0}^l 2^k y_{k+1} = b_i, \quad y_k \in \mathbb{B} \quad \forall k \in \{1, \dots, l+1\} \quad (\text{II.13}).$$

Assim, é sempre possível eliminar os vínculos de um programa pseudo-booleano polinomial com coeficientes racionais. O novo problema entretanto pode ter um número muito maior de termos e variáveis que o primeiro.

ROSENBERG (1975) mostrou que todo programa pseudo-booleano polinomial sem vínculos pode ser escrito como um programa pseudo-booleano quadrático sem vínculos. Para isto, basta repetir tantas vezes quantas forem

necessárias a regra seguinte. Em cada termo com um produto de mais de duas variáveis, escolher duas variáveis x_i e x_j e substituir seu produto pela variável x_{n+l+1} (onde l é o número de vezes que esta regra foi anteriormente aplicada), introduzindo em seguida o termo corretivo

$$(\tilde{f} - \sum_{i=1}^p \min(0, c_i) + 1) \cdot (x_i x_j + (3 - 2x_j - 2x_i) \cdot x_{n+l+1}),$$

onde \tilde{f} é o valor da função-objetivo em um ponto viável qualquer, o qual garante que não haverá solução ótima do novo problema com $x_{n+l+1} \neq x_i x_j$.

Além disso, toda função pseudo-booleana quadrática pode ser feita convexa. HAMMER e RUBIN (1970) observaram que

$$\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j - \alpha \sum_{i=1}^n x_i (1 - x_i) \quad (\text{II.14}).$$

Tomando α grande o suficiente para que os autovalores da matriz $[(c_{ij}) + \alpha I]$ sejam positivos obtemos uma função convexa.

ROSENBERG (1972) mostrou que toda função multilinear definida no hiper-cubo unitário possui um ponto de mínimo em pelo menos um de seus vértices.

Assim sendo, todo program pseudo-booleano não-linear pode ser expresso como um programa quadrático convexo definido no hiper-cubo unitário. Esta simplificação pode entretanto ser proibitivamente cara.

Alguns autores trataram programas

pseudo-boleanos numa forma diferente de $(PPB)_p$. FLORIAN e ROBILLARD (1971), MICHELON e MACULAN (1988), SAIPE (1975) estudaram Programas Pseudo-Boleanos Hiperbólicos, ou seja problemas na forma

$$(PPB)_H \quad \begin{cases} \text{minimizar} & \frac{c_0 + c \cdot x}{\alpha_0 + \alpha \cdot x} \\ x \in \Omega_H & \\ \Omega_H = \{ x \in \mathbb{B}^n / Ax \geq b \} & \end{cases}$$

Para uma revisão destes problemas ver SOUZA (1989) e referências ali citadas.

Em certos casos particulares podemos resolver um (PPB) através de algoritmos para outros problemas. PICARD e RATLIFF (1975) mostraram que toda função pseudo-boleana quadrática de $n-2$ variáveis na forma

$$c_0 + \sum_{j=2}^{n-1} c_{jj} x_j - \sum_{i=2}^{n-1} \sum_{\substack{j=2 \\ i \neq j}}^{n-1} c_{ij} x_i x_j, \quad c_{ij} \geq 0, \quad i, j=2, \dots, n-1$$

pode ser minimizada por um algoritmo de fluxo em redes aplicado a uma rede com n vértices.

Dentre todos os métodos de redução a outros problemas os mais difundidos são aqueles chamados de linearização. Estes métodos foram introduzidos por FORTET (1959, 1960), DANTZIG (1958), WATTERS (1967), ZANGWILL (1965).

FORTET (1959, 1960) sugere trocar cada produto de variáveis T_k (ou T_{ik}) por uma nova

variável x_{n+k} , e adicionar os vínculos

$$\sum_{j \in N_k} x_j - x_{n+k} \leq |N_k| - 1 \quad (\text{II.15})$$

e

$$-\sum_{j \in N_k} x_j + |N_k| x_{n+k} \leq 0 \quad (\text{II.16}).$$

O primeiro vínculo implica que $x_{n+k} = 1$ quando $T_k = 1$ e o segundo faz com que $x_{n+k} = 0$ quando $T_k = 0$. O número de novas variáveis e novos vínculos entretanto pode ser muito grande, mesmo para problemas pequenos. GLOVER e WOOLSEY (1973) propuseram as regras seguintes com o intuito de diminuir o número de vínculos.

Sejam $N = \bigcup_k N_k$, $Q \subset N$, e $T = \{j \in N - Q / N_k = \{j\} \cup N_k \text{ para algum } k\}$. Seja R o conjunto dos subconjuntos de T com $|T| - 1$ elementos tal que $\forall W \in R \exists N_k = W \cup Q$.

(i) Seja $S = \{N_k / N_k = Q \cup \{j\}, j \in T\}$. Substitua $\forall N_k \in S$ os vínculos (II.15) por

$$|S| \sum_{j \in Q} x_j + \sum_{j \in T} x_j - \sum_{k/N_k \in S} x_{n+k} \leq |S| |Q| \quad (\text{II.17}),$$

mantendo inalterados os vínculos (II.16).

(ii) Seja $S = \{N_k / N_k = W \cup Q, W \in R\}$. Troque os os vínculos (II.15) $\forall N_k \in S$ por

$$|S| \sum_{j \in Q} x_j + (|S| - 1) \sum_{j \in T} x_j - \sum_{k/N_k \in S} x_{n+k} \leq |S| (|S| + |Q| - 2) \quad (\text{II.18}),$$

mantendo os vínculos (II.16) inalterados.

(iii) Seja S como em (i). Troque os vínculos (II.16)

$\forall N_k \in S$ por

$$- \sum_{j \in T} x_j + \sum_{k/N_k \in S} x_{n+k} \leq 0 \quad (\text{II.19})$$

$$- |S| x_q + \sum_{k/N_k \in S} x_{n+k} \leq 0, \quad x_q = \prod_{j \in Q} x_j \quad (\text{II.20}),$$

mantendo os vínculos (II.16) inalterados.

(iv) Seja S como em (ii). Troque os vínculos de

(II.16) $\forall N_k \in S$ por

$$- |S| x_q + \sum_{k/N_k \in S} x_{n+k} \leq 0, \quad x_q = \prod_{j \in Q} x_j \quad (\text{II.21})$$

$$- \sum_{j \in T} x_j - |S|^2 x_t + (|S| - 1) \sum_{k/N_k \in S} x_{n+k} \leq 0, \quad x_t = \prod_{j \in T} x_j \quad (\text{II.22}),$$

mantendo os vínculos (II.19) inalterados.

II.2) MÉTODOS ALGÉBRICOS

Os primeiros algoritmos algébricos para minimização de uma função pseudo-booleana parecem ter sido os de FORTET (1959, 1960) e CAMION (1960). Dentre todos, o mais difundido é o "Algoritmo Básico" de HAMMER, ROSENBERG E RUDEANU (1963).

O Algoritmo Básico procura obter uma solução ótima para um programa pseudo-booleano sem restrições como segue. Conforme vimos no teorema (II.2) podemos escrever $f(x)$ como

$$f_1(x_1, \dots, x_n) = f(x_1, \dots, x_n) = x_1 \Delta_1(x_2, \dots, x_n) + \eta_1(x_2, \dots, x_n) \quad (\text{II.23}).$$

Dai, teremos uma solução ótima (x_1^*, \dots, x_n^*) com $x_1^* = 1$ se $\Delta_1(x_2^*, \dots, x_n^*) < 0$. Definimos então uma função $\Psi_1(x_2, \dots, x_n)$ dada por

$$\Psi_1(x_2, \dots, x_n) = \begin{cases} \Delta_1(x_2, \dots, x_n), & \text{se } \Delta_1(x_2, \dots, x_n) < 0, \\ 0, & \text{caso contrário} \end{cases} \quad (\text{II.24}).$$

Expressando Ψ_1 em forma polinomial, definimos

$$f_2(x_2, \dots, x_n) = \Psi_1(x_2, \dots, x_n) + \eta_1(x_2, \dots, x_n) \quad (\text{II.25}).$$

e reduzimos o problema original de n variáveis para outro de $n - 1$ variáveis. Após havermos aplicado este processo $n - 1$

vezes obtemos uma função $f_n(x_n)$. Obtendo um ponto de mínimo x_n^* , podemos a partir daí obter uma solução ótima através da fórmula recursiva

$$x_i^* = \begin{cases} 1, & \text{se } \Psi_i(x_{i+1}^*, \dots, x_n^*) < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (\text{II.26}).$$

Na versão original do algoritmo básico as funções Ψ_i são obtidas da seguinte maneira. Primeiro lineariza-se $\Delta_i(x_{i+1}, \dots, x_n)$ trocando termos de produtos de variáveis por novas variáveis y_j . Em seguida computa-se a função característica Φ_i (HAMMER E RUDEANU (1968, p.83)) da desigualdade $\Delta_i(y_1, \dots, y_p) < 0$ (por definição Φ_i é uma função booleana igual a 1 em todo ponto (y_1, \dots, y_p) em que $\Delta_i(y_1, \dots, y_p) < 0$ e 0, caso contrário). Eliminam-se então as variáveis y_i e simplifica-se Φ_i através de operações booleanas. Ψ_i então será dada por

$$\Psi_i(x_{i+1}, \dots, x_n) = \Phi_i(x_{i+1}, \dots, x_n) \Delta_i(x_{i+1}, \dots, x_n) \quad (\text{II.27}).$$

CRAMA, HANSEN e JAUMARD (1989) apresentaram um modo alternativo para a obtenção das funções Ψ_i . Também INAGAKI e FUKUMURA (1967), e ROSENBERG (1973) apresentaram variantes e extensões do algoritmo básico para problemas com vínculos (outras referências podem ser encontradas em HANSEN, JAUMARD e MATHON (1989)).

II.2.3) MÉTODOS ENUMERATIVOS

Um dos primeiros algoritmos enumerativos a aparecer na literatura foi o algoritmo lexicográfico de LAWLER e BELL (1966). Este algoritmo resolve problemas na forma

$$(PPB)_{LB} \begin{cases} \text{minimizar } f(x) \\ x \in \Omega \\ \Omega = \{x \in \mathbb{B}^n / g_{i1}(x) - g_{i2}(x) \geq 0, i=1, \dots, m\} \end{cases}$$

onde f, g_{11}, \dots, g_{m2} são funções monotonamente não-decrescentes (uma função é monótona não-decrescente se $x^1 \geq x^2 \Rightarrow h(x^1) \geq h(x^2)$). Note que todo programa pseudo-booleano pode ser expresso na forma $(PPB)_{LB}$. De fato, se $g_i(x) \geq 0$ é um vínculo, basta escrever g_i como um polinômio de seus argumentos, agrupar os termos com coeficientes positivos em g_{i1} e os termos com coeficientes negativos em $-g_{i2}$. Quanto a f , podemos trocar a função-objetivo por outra igual a uma nova variável x_0 , e incluir ao problema a restrição $x_0 \geq f(x)$. Para uma discussão do algoritmo lexicográfico, ver GARFINKEL e NEMHAUSER (1972, pp. 344-348). O algoritmo lexicográfico de Lawler e Bell foi estendido por MAO e WALLINGFORD (1968, 1969) e mais recentemente por WANG (1988).

No fim dos anos 60 e durante os anos 70 diversos algoritmos de "Branch-&-Bound" para programas pseudo-boleanos não-lineares foram propostos: BERMAN

(1969), GINSBURGH e VAN PEETERSEN (1969), HAMMER (1971, 1973, 1974), HAMMER e RUBIN (1969), HAMMER e PELED (1972), HANSEN (1969, 1969b, 1970, 1972, 1974), HILLIER (1969), LAUGHUNN (1970), WALUCKIEWCZ, SLOMINSKI E FANER (1973), VIZVARI (1975), ZAK (1978), SCHOH e LYSKA (1978), BRESNEV (1979). Como de costume estes algoritmos fazem ramificações (Branching) para quebrar o problema em subproblemas menores, e testes para podar (Fathom) a árvore de enumeração em nós a partir dos quais o problema é inviável ou não pode fornecer uma solução melhor que a melhor até então encontrada. Alguns algoritmos usam também outros testes para mostrar que algumas variáveis têm necessariamente valor zero ou um em todas as soluções ótimas ou em todas as soluções viáveis.

Diversas regras de ramificação têm sido propostas. Comumente alguma variável livre é escolhida e seu valor é fixado em zero ou um de acordo com algum critério segundo o qual se espere reduzir a enumeração.

O ponto mais delicado nestes algoritmos (como em geral em algoritmos de Branch-&-Bound) é a obtenção de limites para o valor ótimo da função-objetivo. Estes valores precisam ser próximos do valor ótimo o suficiente para que possam eliminar o maior número possível de soluções viáveis, mas ao mesmo tempo o método de obtenção destes precisa ser suficientemente barato para não tornar o algoritmo proibitivamente caro. Limites podem ser reforçados por penalidades (uma penalidade superior p_j^1 (inferior p_j^c) é um incremento ao valor do limite quando uma variável x_j livre é fixada em um (zero)).

EXEMPLO II.3 (HANSEN, JAUMARD e MATHON (1989))

$$\underline{f} = \sum_{k=1}^p \min(0, c_k) \quad (\text{II.28})$$

é um limite inferior para $f(x)$. Além disso, as penalidades

$$p_j^c = \sum_{(k=1, \dots, p / j \in N_k)} \min(0, c_k) \quad (\text{II.29})$$

$$p_j^1 = \max(0, c_r) + \sum_{(k, k')} \min(|c_k|, |c_{k'}|) \quad (\text{II.30}),$$

onde $N_r = \{j\}$, (k, k') é tal que $N_k = \{1\}$ e $N_{k'} = \{1, j\}$ e $c_k c_{k'} < 0$, podem ser associadas com \underline{f} . ■

Dizemos que uma penalidade é aditiva quando a soma das penalidades obtidas da fixação dos valores de diversas variáveis pode ser adicionada ao limite corrente, fornecendo um novo limite inferior. As penalidades do exemplo abaixo, devido a HANSEN, JAUMARD e MATHON (1989), são aditivas, enquanto as penalidades do exemplo anterior não o são.

EXEMPLO II.4 (HANSEN, JAUMARD e MATHON(1989)) As
penalidades aditivas

$$p_j^c = \sum_{(k=1, \dots, p / j \in N_k)} \alpha_{jk} \min(0, c_k) \quad (\text{II.31}),$$

com $0 \leq \alpha_{jk} \leq 1$ e $\sum_{j \in N_k} \alpha_{jk} = 1, \forall k \in \{1, \dots, p\}$.

$$p_j^1 = - \text{máx}(0, c_r) \quad (\text{II.32}),$$

com $N_r = \{j\}$, podem ser associadas a f . ■

Uma outra maneira de obtermos limites é através do conceito de planos inferiores.

DEFINIÇÃO II.6 Dada uma função pseudo-booleana $f(x)$, definimos um plano inferior de f como uma função linear $l(x) = l_0 + \sum_{j=1}^n l_j x_j$, tal que $l(x) \leq f(x) \forall x \in \mathbb{B}^n$. ■

EXEMPLO II.5 Dada uma função $f(x) = \sum_{k=1}^p c_k T_k$, a função $l(x) = \sum_{k=1}^p \text{mín}(0, c_k) \sum_{j \in N} x_j$ é um plano inferior para f . ■

LU e WILLIAMS (1978), como extensão do conceito de Dualidade de Teto de HAMMER, HANSEN e SIMEONE (1984), obtiveram também um outro procedimento para obtenção de limites inferiores para problemas de minimização.

No capítulo IV estudaremos um outro procedimento desenvolvido por MICHELON e MACULAN (1988) como extensão de GUIGNARD e KIM (1987) para obtenção de limites para programas pseudo-booleanos, e no capítulo V apresentaremos um algoritmo enumerativo para a resolução do problema da mochila quadrática, utilizando os limites desta maneira obtidos.

MÉTODOS DE CORTE

Os métodos de planos de corte tiveram origem na linearização de programas pseudo-booleanos através de redução destes a problemas de cobertura generalizada proposta por GRANOT e HAMMER (1971).

GRANOT e HAMMER (1971), BALAS e JEROSLOW (1972), BALAS e MAZZOLA (1984, 1984b) mostraram que os vínculos de um programa pseudo-booleano são equivalentes a um conjunto de desigualdades de cobertura generalizada. Linearizando a função-objetivo (ver parágrafo II.2.1), reduzimos o (PPB) a um problema de cobertura generalizada. O número de vínculos obtidos, no entanto, pode ser absurdamente grande, tornando impraticável a solução do novo problema (ver GRANOT, GRANOT e KALLBERG (1979)).

GRANOT, GRANOT e KALLBERG (1979), GRANOT e GRANOT (1980), GRANOT, GRANOT e VAESSEN (1982) propuseram então um algoritmo que vai introduzindo, à medida que se torna necessário, as desigualdades de cobertura generalizada. Um esquema bem mais eficiente foi proposto por BALAS e MAZZOLA (1984, 1984b).

II.3) APLICAÇÕES

São inúmeras as aplicações da programação pseudo-booleana.

1) LAUGHUNN (1970), MAO e WALLINGFORD (1968), PETERSON e LAUGHUNN (1971), SEPPALA (1967), SCHECHTER e HAMMER (1971), AHRENS (1974), ORON (1979), RAGADE, HIPEL e UNNY (1976), e principalmente HILLIER (1969), estudaram problemas de escolha de investimentos através de programas pseudo-booleanos.

2) Teoria de localização (HAMMER (1968), RHYS (1970)).

3) Horário Escolar (v. da MATA (1989)).

4) STECKE (1983) estuda manufatura flexível via (PPB) e resolve problemas por linearização (ver também COCA-BALTA, MARIETTO e CLEIMAN (1990)).

5) KOLESAR (1980) estuda o teste de perda de visão pela suspeita de glaucoma com um modelo em variáveis binárias.

6) GILMORE e GOMORY (1963) modelam problemas de corte através de programas pseudo-booleanos hiperbólicos (Cutting-Stock Problems).

Diversas referências fornecem uma lista

exaustiva de aplicações. Dentre elas mencionamos CRAVEN (1988), SCHAIABLE(1981), HANSEN (1979), HANSEN, JAUMARD e MATHON (1989).

CAPÍTULO III

DUALIDADE LAGRANGEANA E SEU USO EM PROGRAMAÇÃO DISCRETA

Neste capítulo estudaremos o uso de Dualidade Lagrangeana em programação matemática, dando especial ênfase aos resultados mais importantes para a resolução de duais de programas discretos.

III.1) DEFINIÇÕES E RESULTADOS PRELIMINARES

Sejam $f: \mathbb{R}^n \rightarrow \mathbb{R}$ e $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções contínuas e X um subconjunto compacto do \mathbb{R}^n . Seja ainda o problema

$$\text{(CPP)} \quad \begin{cases} \text{minimizar } \langle f(x) \rangle \\ x \in \Omega \\ \Omega = \{ x \in X \mid g(x) = 0 \}, \quad \Omega \neq \emptyset \end{cases}$$

o qual será referido como problema primal, e para o qual sabemos, pelo teorema de Weierstrass (Lima, 1981, p. 44) haver uma solução ótima finita x^* com valor $f^* = f(x^*)$.

DEFINIÇÃO III.1 Definimos a Função de Lagrange

$L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ do problema primal (PP) como

$$L(x, u) = f(x) + u \cdot g(x), \quad x \in X, \quad u \in \mathbb{R}^m.$$

As componentes u_i de u chamamos Multiplicadores de Lagrange das componentes g_i de g . ■

OBSERVAÇÃO III.1 Se x é um ponto viável de (PP), então

$$L(x, u) = f(x), \quad \forall u \in \mathbb{R}^m. \quad \blacksquare$$

DEFINIÇÃO III.2 Definimos a função dual $w: \mathbb{R}^m \rightarrow \mathbb{R}$ de (PP) como

$$w(u) = \underset{x \in X}{\text{mínimo}} \{ L(x, u) \} \quad \blacksquare$$

TEOREMA III.1 $w(u) \leq f^* \quad \forall u \in \mathbb{R}^m$.

DEMONSTRAÇÃO Como x^* é um ponto viável de (PP) vem, pela observação (III.1),

$$w(u) = \underset{x \in X}{\text{mínimo}} \{ L(x, u) \} \leq L(x^*, u) = f(x^*) = f^*. \quad \blacksquare$$

DEFINIÇÃO III.3 Dizemos que um problema

$$(PR) \underset{x \in \Omega_r}{\text{minimizar}} \{ f_r(x) \}$$

é uma Relaxação de (PP) se $\Omega \subset \Omega_r$ e $\underset{x \in \Omega}{\text{mínimo}} \{ f(x) \} \leq \underset{x \in \Omega_r}{\text{mínimo}} \{ f_r(x) \} \quad \forall x \in \Omega$. ■

Muitas vezes a resolução da relaxação que

fornece os valores de w em $u \in \mathbb{R}^m$ é bem mais simples que a resolução do problema primal. Conforme vimos, estes valores fornecem um limite inferior para o valor ótimo do problema primal, podendo ser usados em algum algoritmo para resolução de (PP). Estes limites serão tanto melhores quanto maior for o valor de $w(u)$, podendo seu valor máximo em algumas circunstâncias ser até mesmo igual ao valor ótimo do problema primal.

DEFINIÇÃO III.4 Dizemos que o problema

$$\text{(PD) maximizar } \langle w(u) \rangle_{u \in \mathbb{R}^m}$$

é um problema dual de (PP). ■

Infelizmente no caso geral o valor ótimo w^* de (PD) não é igual a f^* . O que podemos afirmar seguramente é que vale o teorema abaixo.

TEOREMA III.2 (DUALIDADE FRACA) $w^* \leq f^*$. ■

Este teorema é consequência óbvia do teorema anterior.

Os resultados e conceitos acima podem ser ilustrados geometricamente como segue. Vamos supor (a fim de permitir um desenho no plano) $m = 1$, e consideremos a transformação $x \rightarrow (g(x), f(x))$ de domínio X e imagem em $Y \subset \mathbb{R}^2$.

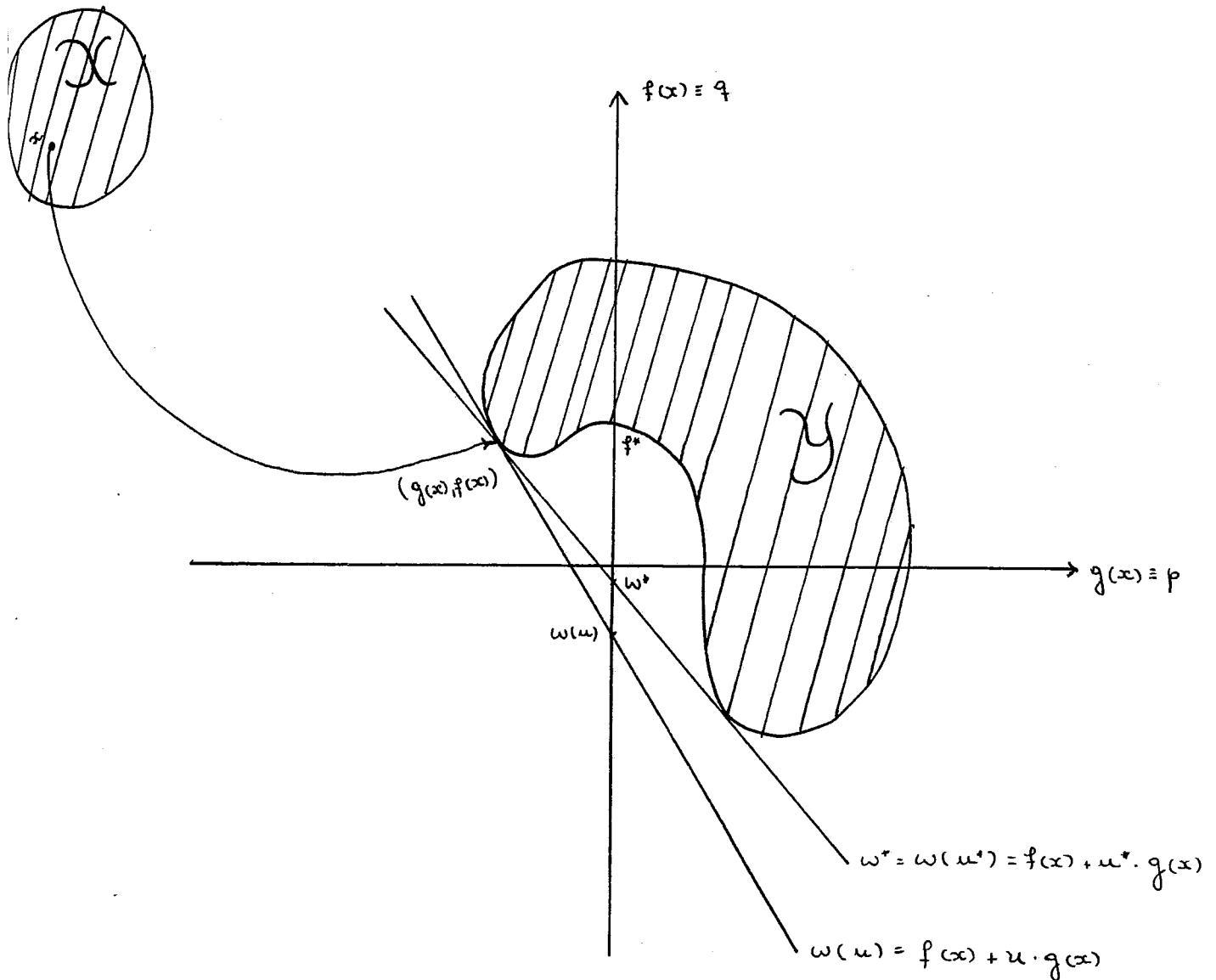


FIGURA III.1

O problema primal consiste em encontrar um ponto no eixo das ordenadas que minimize $f(x)$. Este ponto é f^* .

Tomemos agora um multiplicador de Lagrange $u \in \mathbb{R}$. Para obter $w(u)$ temos que minimizar $f(x) + u g(x)$ em todo $x \in X$. Fazendo $p = g(x)$ e $q = f(x)$, $x \in X$, isto equivale a minimizar $q + up$ em Y . Por outro lado, $q + up = b$ é a equação de uma reta de inclinação $-u$ passando por

$(0, b)$. Assim sendo, para minimizar $q + up$ temos que transportar esta reta paralelamente a ela tão para baixo quanto possível. Quando tal acontecer, a ordenada do ponto de interseção com o eixo dos p 's será igual a $w(u)$. O problema dual consiste em encontrar a inclinação u^* que maximize essa ordenada.

III.2) PROPRIEDADES DA FUNÇÃO DUAL

Neste parágrafo estudaremos algumas propriedades da função dual, em especial aquelas mais úteis na construção de algoritmos para sua maximização.

TEOREMA III.3 A função dual é côncava e finita.

DEMONSTRAÇÃO Sendo f e g contínuas no \mathbb{R}^n , e X um conjunto limitado, vem que w é finita.

Também, $\forall \lambda \in [0, 1]$, $u^1, u^2 \in \mathbb{R}^m$, obtemos

$$\begin{aligned} w(\lambda u^1 + (1 - \lambda)u^2) &= \underset{x \in X}{\text{mínimo}} \{f(x) + (\lambda u^1 + (1 - \lambda)u^2) \cdot \\ &g(x)\} \geq \lambda \underset{x \in X}{\text{mínimo}} \{f(x) + u^1 \cdot g(x)\} + (1 - \lambda) \underset{x \in X}{\text{mínimo}} \{f(x) \\ &+ u^2 \cdot g(x)\} = \lambda w(u^1) + (1 - \lambda) w(u^2). \quad \blacksquare \end{aligned}$$

Nem sempre a função dual é diferenciável. Consideremos o exemplo na figura da página seguinte onde $X = \{x^1, x^2, x^3\}$ e $m = 1$.

Neste exemplo a função dual é linear por partes e não-diferenciável nos pontos de interseção das

retas r_1 , r_2 e r_3 .

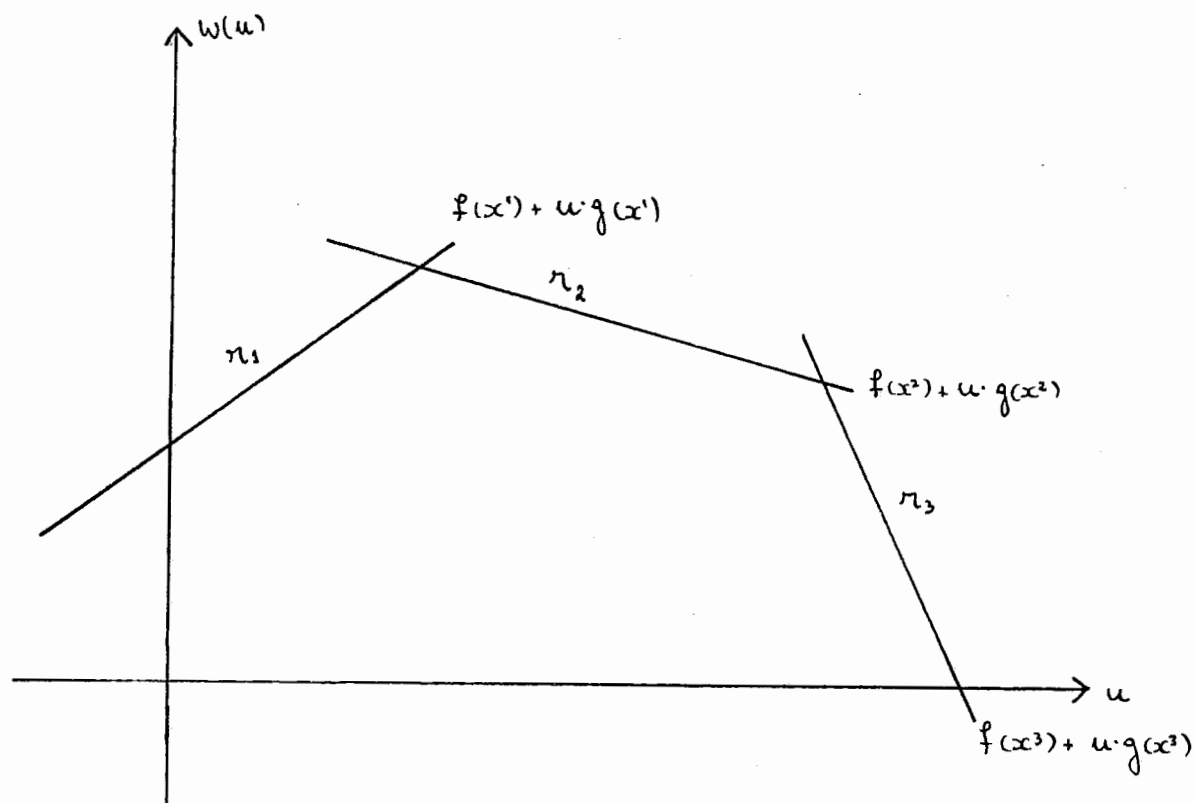


FIGURA III.2

TEOREMA III.4 Sejam $\tilde{u} \in \mathbb{R}^m$ e definamos $X(u) = \{ x \in \mathbb{R}^n / w(u) = f(x) + u \cdot g(x) \}$, $u \in \mathbb{R}^m$. Suponhamos que $X(\tilde{u})$ seja igual ao conjunto unitário \tilde{X} . Então w é diferenciável em \tilde{u} com gradiente $\nabla w(\tilde{u}) = g(\tilde{x})$.

DEMONSTRAÇÃO Bazaraa e Shetty (1979, p. 189). ■

A não diferenciabilidade de w é um incômodo

para a resolução do problema dual. Entretanto, sendo w côncava pode-se contornar o problema generalizando o conceito de gradiente. Além disso, pode-se ainda definir em cada ponto $u \in \mathbb{R}^m$ a derivada direcional de w .

DEFINIÇÃO III.5 Dizemos que $\xi \in \mathbb{R}^m$ é um subgradiente de w em $u \in \mathbb{R}^m$ se

$$w(u) \geq w(v) + \xi \cdot (u - v), \quad \forall v \in \mathbb{R}^m. \quad \blacksquare$$

DEFINIÇÃO III.6 Sejam $u, d \in \mathbb{R}^n$ com $\|d\| \neq 0$. Definimos a derivada direcional de w ao longo de d como

$$w'(u; d) = \lim_{t \rightarrow 0^+} \frac{w(u + td) - w(u)}{t},$$

desde que o limite acima exista. \blacksquare

TEOREMA III.5 O limite acima existe $\forall u, d \in \mathbb{R}^n, \|d\| \neq 0$.

DEMONSTRAÇÃO Basta mostrar que o limite

$$\lim_{t \rightarrow 0^+} \frac{w(u + td) - w(u)}{t}$$

existe.

De fato, sejam $\lambda_2 > \lambda_1 > 0$. Vem que

$$w(u + \lambda_1 d) = w\left(\frac{\lambda_1}{\lambda_2}(u + \lambda_2 d) + \frac{\lambda_2 - \lambda_1}{\lambda_2}u\right)$$

Visto que w é côncava,

$$w(u + \lambda_1 d) \geq \frac{\lambda_1}{\lambda_2} w(u + \lambda_2 d) + \frac{\lambda_2 - \lambda_1}{\lambda_2} w(u) =$$

$$\frac{\lambda_1}{\lambda_2} w(u + \lambda_2 d) + w(u) - \frac{\lambda_1}{\lambda_2} w(u) \Rightarrow$$

$$\frac{w(u + \lambda_1 d) - w(u)}{\lambda_1} \geq \frac{w(u + \lambda_2 d) - w(u)}{\lambda_2}$$

e portanto $-\frac{w(u + \lambda_1 d) - w(u)}{\lambda_1}$ é uma função não-decrescente de $\lambda > 0$. Assim, o limite em questão existe e é igual a

$$\inf_{\lambda > 0} \left\{ -\frac{w(u + \lambda d) - w(u)}{\lambda} \right\} \quad \blacksquare$$

A resolução do problema relaxado (PR) fornece automaticamente um subgradiente em $u \in \mathbb{R}^m$, conforme mostrado a seguir

TEOREMA III.6 Dado $u \in \mathbb{R}^m$, $g(x)$ é um subgradiente para w em $u \forall x \in X(u)$.

DEMONSTRAÇÃO Sejam $x \in X(u)$ e $v \in \mathbb{R}^m$. Vem

$$w(u) = f(x) + u \cdot g(x) = f(x) + v \cdot g(x) + (u - v) \cdot g(x) \geq$$

$$w(v) + (u - v) \cdot g(x). \quad \blacksquare$$

O teorema acima garante que w possui um subgradiente em todo ponto em que é definida. Também dele e do teorema (III.4) deprendemos facilmente que em pontos de não-diferenciabilidade existem mais do que um subgradiente. De fato,

TEOREMA III.7 w é diferenciável em um ponto $u \in \mathbb{R}^m$ se e somente se houver um único subgradiente de w em u .

DEMONSTRAÇÃO Seja $x^* \in X(u)$. Se houver um único subgradiente de w em u , então necessariamente $X(u) = \{x^*\}$ e de acordo com o teorema (III.4) w é diferenciável em u com gradiente $\nabla w(u) = g(x^*)$.

Suponhamos agora que w seja diferenciável em u . Da definição de subgradiente de uma função côncava, e do gradiente, obtemos

$$w(u + \lambda d) \geq w(u) + \lambda \xi \cdot d$$

$$w(u + \lambda d) = w(u) + \lambda \nabla w(u) \cdot d + \lambda \|d\| \alpha(u; d),$$

onde $\lambda \in \mathbb{R}$, $d \in \mathbb{R}^m$ e $\alpha: \mathbb{R}^m \rightarrow \mathbb{R}$ é uma função tal que

$\lim_{\lambda \rightarrow 0} \alpha(u; \lambda d) = 0 \quad \forall u \in \mathbb{R}^m$, e ξ é um subgradiente de w em u .

Subtraindo a igualdade da desigualdade vem

$$0 \geq \lambda (\xi - \nabla w(u)) \cdot d - \lambda \|d\| \alpha(u; \lambda d).$$

Dividindo a desigualdade acima por $\lambda > 0$ e fazendo $\lambda \rightarrow 0$ obtemos

$$(\xi - \nabla w(u)) \cdot d \leq 0$$

Tomando agora $d = \xi - \nabla w(u)$, vem

$$\xi = \nabla w(u)$$

e assim, em pontos de diferenciabilidade, w possui um único subgradiente, a saber $\nabla w(u)$. ■

Dado que em pontos de não diferenciabilidade da função dual há mais de um subgradiente, introduziremos a definição abaixo.

DEFINIÇÃO III.7 Dado $u \in \mathbb{R}^m$, definimos o subdiferencial $\partial w(u)$ como o conjunto de todos os subgradientes de w em u . ■

Dado um ponto u , já conhecemos, pelo teorema (III.6), uma parte do subdiferencial de w em u . De fato, seja $Y(u) = \{y \in \mathbb{R}^m / y = g(x), \text{ para algum } x \in X(u)\}$. Vem que

TEOREMA III.8 $Y(u) \subset \partial w(u) \forall u \in \mathbb{R}^m$. ■

A demonstração é óbvia, a partir do teorema (III.6).

Na resolução de um problema de maximização de uma função, é sempre importante de alguma maneira tentar caracterizar pontos de máximo e direções de subida.

No caso em que esta função é côncava, diferenciável e definida em todo \mathbb{R}^m , o máximo local coincide com o máximo global e acontece em um ponto u^* dado por $\nabla w(u^*) = 0$ (BAZARAA e SHETTY (1979, pp.94 e 96)). Além disso, em um ponto u que não seja ótimo, sua direção de maior crescimento é $\nabla w(u)$ (LANG (1968, p. 75)).

Em pontos de não-diferenciabilidade esta caracterização é feita através de uma generalização dos resultados acima.

TEOREMA III.9 u^* é um ponto de máximo de w se e somente se

$0 \in \partial w(u)$.

DEMONSTRAÇÃO u^* é uma solução ótima para o problema dual

$$\Leftrightarrow w(u^*) \geq w(u) \quad \forall u \in \mathbb{R}^m \Leftrightarrow 0 \in \partial w(u). \quad \blacksquare$$

A direção de maior crescimento local de uma função é dada pela direção que maximiza sua derivada direcional. Conforme já comentado acima, esta direção no caso de funções diferenciáveis é dada pelo gradiente. Vamos agora estender este resultado para o caso da função dual. Veremos que em pontos de diferenciabilidade ela se reduz ao resultado descrito acima para funções diferenciáveis. Antes porém veremos dois resultados intermediários.

LEMA III.1 Sejam $u, d \in \mathbb{R}^m$. Então $\exists y \in Y(u)$ tal que $w'(u; d) \geq d \cdot y$.

DEMONSTRAÇÃO Seja a sucessão $\{t_k\}$, $t_k \rightarrow 0^+$, e tomemos os pontos $u + t_k d$. Como X é compacto, existe $x_k \in X(u + t_k d)$ e além disso a sucessão $\{x_k\}$ admite uma subsucessão $\{x_k\}_{k \in K}$ convergente para um ponto $\tilde{x} \in X$.

Por outro lado,

$$f(x) + (u + t_k d) g(x) \geq f(x_k) + (u + t_k d) g(x_k).$$

Passando ao limite na subsucessão $k \in K$,

$$f(x) + u g(x) \geq f(\tilde{x}) + u g(\tilde{x}), \quad \forall x \in X.$$

Logo, $\tilde{x} \in X(u)$.

Por outro lado,

$$w(u + t_k d) - w(u) = [f(x_k) + u g(x_k) - w(u)] + t_k d \cdot g(x_k) \geq$$

$$t_k d \cdot g(x_k)$$

ou,

$$\frac{w(u + t_k d) - w(u)}{t_k} \geq d \cdot g(x_k).$$

Passando ao limite ao longo da subsequência $k \in K$,

$$w'(u; d) \geq d \cdot g(\tilde{x}),$$

o que prova a assertiva acima. ■

TEOREMA III.10 $w'(u; d) = \underset{y \in \partial w(u)}{\text{mínimo}} \langle d, y \rangle$.

DEMONSTRAÇÃO Para todo $y \in \partial w(u)$, para todo $t > 0$, e para todo $d \in \mathbb{R}^m$,

$$w(u + td) \leq w(u) + tyd \Rightarrow \frac{w(u + td) - w(u)}{t} \leq y \cdot d \Rightarrow$$

$$\lim_{t \rightarrow 0^+} \frac{w(u + td) - w(u)}{t} \leq y \cdot d \Rightarrow w'(u; d) \leq y \cdot d \quad \forall y \in \partial w(u) \Rightarrow$$

$$w'(u; d) \leq \underset{y \in \partial w(u)}{\text{mínimo}} \langle y, d \rangle$$

Pelo lema (III.1),

$$w'(u; d) = \min_{y \in \partial v(u)} \langle y, d \rangle. \quad \blacksquare$$

Estamos agora em condições de encontrar a direção de maior crescimento local de u .

$$\text{TEOREMA III.11} \quad \max_{\|k\| \leq 1} \langle w'(u; d) \rangle = \min_{y \in \partial v(u)} \|y\|$$

$$\text{DEMONSTRAÇÃO} \quad \max_{\|k\| \leq 1} \langle w'(u; d) \rangle = \max_{\|k\| \leq 1} \langle \min_{y \in \partial v(u)} \langle y, d \rangle \rangle =$$

$$\min_{y \in \partial v(u)} \langle \max_{\|k\| \leq 1} \langle y, d \rangle \rangle$$

Pela desigualdade de Cauchy-Schwarz (LIMA (1981, p. 4)),

$$\max_{\|k\| \leq 1} \langle y, d \rangle = \|y\|$$

e portanto

$$\max_{\|k\| \leq 1} \langle w'(u; d) \rangle = \min_{y \in \partial v(u)} \|y\| \quad \blacksquare$$

Assim sendo, a direção de maior crescimento local de w a partir de um ponto u que não seja um ponto de máximo é a direção do subgradiente de menor norma. Em pontos de diferenciabilidade o diferencial é igual a $\langle \nabla w \rangle$, e portanto a direção de maior crescimento é a do gradiente no

ponto.

A determinação do subdiferencial de w em cada ponto é, desta forma, de suma importância para a resolução do problema dual. É apenas conhecendo $\partial w(u)$ em cada ponto $u \in \mathbb{R}^m$ que podemos dizer se um certo ponto é uma solução ótima ($0 \in \partial w(u)$), e se não qual a direção de maior crescimento (direção do subgradiente de menor norma em $\partial w(u)$).

TEOREMA III.12 $\partial w(u) = \text{CONV} \langle Y(u) \rangle, \forall u \in \mathbb{R}^m$.

DEMONSTRAÇÃO Seja \tilde{y} um ponto de $\partial w(u)$ que não pertença a $\text{CONV} \langle Y(u) \rangle$. Como $Y(u)$ é fechado e $\text{CONV} \langle Y(u) \rangle$ é convexa, vem (BAZARAA e SHETTY (1979, p.45)) que $\exists d \in \mathbb{R}^m, d \neq 0$, e $\exists \alpha \in \mathbb{R}$, tais que

$$d \tilde{y} < \alpha$$

e

$$d y \geq \alpha \quad \forall y \in \text{CONV} \langle Y(u) \rangle.$$

Portanto, pelo lema (III.1),

$$w'(u; d) \geq \alpha$$

e pelo teorema (III.10),

$$w'(u; d) \leq d \tilde{y} < \alpha,$$

o que é absurdo. Assim, *

$$\partial w(u) \subset \text{CONV}\{Y(u)\}.$$

Por outro lado, pelo teorema (III.8), e uma vez que $\partial w(u)$ é convexo e fechado (LASDON (1970, p. 510)),

$$\text{CONV}\{Y(u)\} \subset \partial w(u).$$

Portanto,

$$\partial w(u) = \text{CONV}\{Y(u)\} \quad \blacksquare$$

Note que nem sempre um subgradiente de w em um ponto u_0 é uma direção de subida. O exemplo na figura abaixo, onde estão desenhadas as curvas de nível de uma função dual w (agora fizemos $m = 2$) ilustra isto

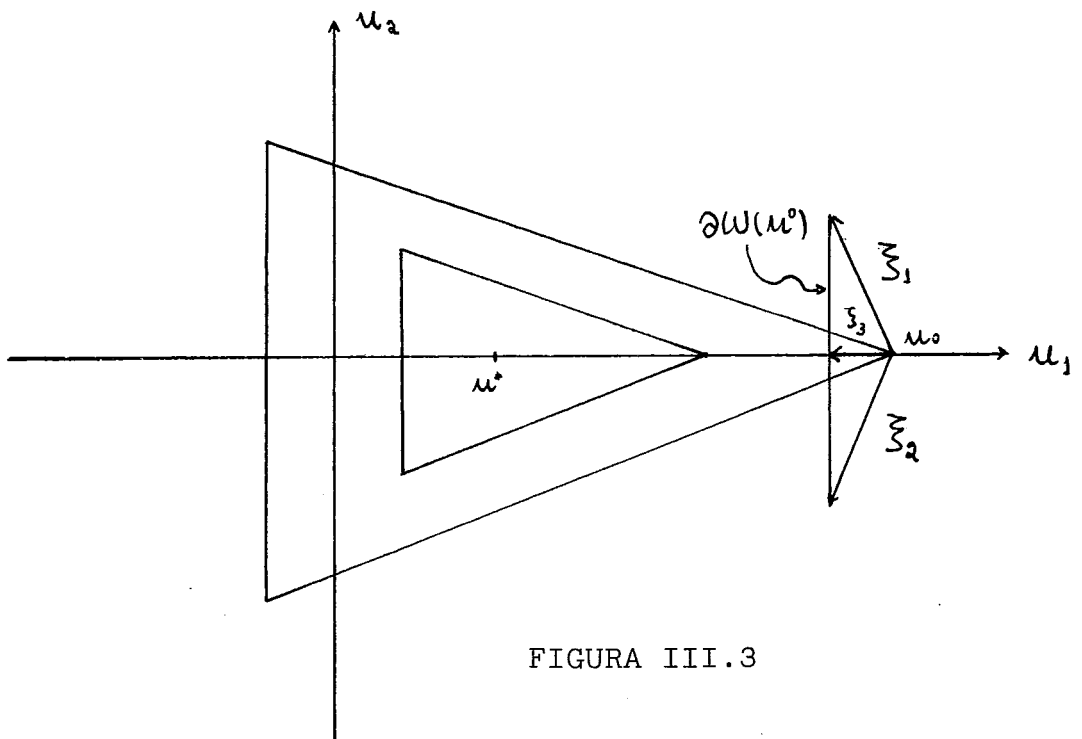


FIGURA III.3

No exemplo acima ξ_3 , o subgradiente de norma mínima de $\partial w(u_0) = \text{CONV}\{\xi_1, \xi_2\}$, é realmente a direção de maior crescimento. Entretanto nem ξ_1 , nem ξ_2 são direções de subida. O melhor que se pode dizer de ξ_1 e ξ_2 é que formam um ângulo menor que 90° com o segmento que une u^0 a u^* , e que portanto são direções de descida para a distância ao ótimo. Isto, na verdade, é tudo que se pode dizer de maneira genérica do subgradiente em um ponto $u \neq u^*$.

TEOREMA III.13 Seja $d(u) = \|u - u^*\|$, $u \in \mathbb{R}^m$. Qualquer subgradiente $\xi \in \partial w(u)$ é uma direção de descida para d .

DEMONSTRAÇÃO Sendo ξ o subgradiente de uma função côncava,

$$w(u^*) \leq w(u) + \xi \cdot (u^* - u) \Rightarrow \xi \cdot (u - u^*) \leq w(u) - w(u^*) \leq 0 \Rightarrow \xi \cdot (u - u^*) \leq 0$$

Mas, como $\nabla d(u) = u - u^*$, vem que

$$\xi \cdot \nabla d(u) \leq 0,$$

e portanto ξ é uma direção de descida para d . ■

Podemos definir o cosseno do ângulo entre $u - u^*$ e o "pior" subgradiente ξ em u como

$$l(u) = \text{mínimo} \left\{ \frac{\xi \cdot (u^* - u)}{\|\xi\| \|u^* - u\|} \right\} \quad (\text{GOFFIN (1977)}).$$

No pior caso teremos

$$l = \text{mínimo } \{l(u)\}.$$

Chamamos l de número condicionante, e conforme veremos na próxima seção ele é importante na determinação da convergência de algoritmos.

Tentar caracterizar o subdiferencial em cada ponto, a fim de encontrar o subgradiente de menor norma, para então saber se este é um ponto de mínimo ou então descobrir a direção de crescimento máximo da função dual está inteiramente fora de nosso alcance. É computacionalmente muito caro. Entretanto, ao resolver o problema relaxado (PR) obtemos, conforme mostrado no teorema (III.6), automaticamente um subgradiente. Caminhando na direção deste subgradiente, caso não estejamos em um ponto de máximo, estaremos nos aproximando de u^* . Este procedimento, que generaliza aquele adotado em algoritmos de gradiente, é a essência dos chamados métodos de subgradiente, os quais estudaremos na seção seguinte. Ao contrário de todo subdiferencial, a obtenção de um subgradiente através da resolução do problema relaxado é bastante barata, o que torna estes métodos computacionalmente factíveis.

II.3) MÉTODOS DE SUBGRADIENTE

Os métodos de subgradiente podem ser genericamente descritos como segue.

ALGORITMO III.1

Início

 $k \leftarrow 0$ inicializar p^0 e u^0 , e obter $g^0 \in \partial w(u^0)$ Se $g^0 \neq 0$

Então

Terminou \leftarrow falso

Senão

Terminou \leftarrow verdadeiro

Fim-se

Corpo

Enquanto \neg terminou Faça

$$u^{k+1} \leftarrow u^k + p^k \frac{g^k}{\|g^k\|}$$

 $k \leftarrow k + 1$

Efetuar teste de parada

Devolver o resultado do teste de parada na
variável terminou

Fim-Enquanto

Fim

p^k é um passo de deslocamento na direção do subgradiente. Como normalmente o procedimento acima não conseguirá determinar se $0 \in \partial w(u^k)$, para algum k , este não poderá ser o único critério de parada, e por isso outros testes deverão compor o "Teste de parada", como por exemplo o número de iterações, o tamanho do passo, caso este seja não-crescente a cada iteração, e a diferença entre o melhor valor $w(u^k)$ obtido e um limite \bar{w} superior previamente conhecido. ■

Dentro deste esquema geral descreveremos três alternativas correspondentes a três maneiras diferentes de fixarmos o passo p^k :

- (i) Série divergente (POLYAK (1969))
- (ii) Passo constante (MINOUX (1983))
- (iii) Relaxação (HELD, WOLFE e CROWDER (1974))

O método da série divergente foi o primeiro algoritmo de subgradiente a ser proposto, mas seu interesse é principalmente histórico. O método do passo constante possui interesse sobretudo teórico. Sua convergência é muito lenta. O método de relaxação é um método semi-empírico. Não é possível estabelecer uma prova geral de convergência para ele. Entretanto é o mais bem sucedido, e será aquele que adotaremos na implementação do algoritmo proposto no capítulo V para a resolução do problema da mochila quadrática.

III.3.1) O ALGORITMO DA SÉRIE DIVERGENTE

Este método fundamenta-se no seguinte teorema.

TEOREMA III.14 (POLYAK (1969)) Se $p_k \geq 0$ for tal que

$$p_k \rightarrow 0 \text{ e } \sum_{k=1}^{\infty} p_k = \infty$$

então,

$$\liminf_{k \rightarrow \infty} (w(u^k)) = w^*$$

Em outras palavras, w^* é um ponto de acumulação da sequência $\{w(u^k)\}$.

DEMONSTRAÇÃO Basta provar que $\forall \varepsilon > 0 \exists k \in \mathbb{N} / u^k \in V(\varepsilon)$
 $= \{u \in \mathbb{R}^m / w(u) \geq w^* - \varepsilon\}$. Mostraremos isto por absurdo.
 Para tal, suponhamos haver um $\varepsilon > 0$ tal que $\forall k \in \mathbb{N}$ tem-se

$$w(u^k) < w^* - \varepsilon$$

Sendo w contínua, vem que

$$\exists \delta > 0 / \|u - u^*\| \leq \delta \Rightarrow w(u) \geq w^* - \varepsilon.$$

Fazendo

$$\bar{u}^k = u^* - \delta \frac{g^k}{\|g^k\|}$$

teremos sempre

$$w(\bar{u}^k) \geq w^* - \varepsilon,$$

visto que $\|u^k - u^*\| = \delta$. Mas como w é côncava, tem-se $\forall k \in \mathbb{N}$

$$g^k(\bar{u}^k - u^k) \geq w(\bar{u}^k) - w(u^k) > 0.$$

Por outro lado,

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 + p_k^2 + 2p_k g^k \frac{(u^k - u^*)}{\|g\|} - 2\delta p_k$$

ou ainda,

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 + p_k(p_k - 2\delta),$$

visto que $g^k (\bar{u}^k - u^k) > 0$.

Como $p_k \rightarrow 0$, vem que existe $K \in \mathbb{N} / k > K \Rightarrow p_k < \delta$. Para tais valores de k ,

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \delta p_k.$$

Dai, para um certo $r \in \mathbb{N}_+$,

$$\delta \sum_{k=K}^{K+r} p_k \leq \|u^K - u^*\|^2 - \|u^{K+r+1} - u^*\|^2 \leq \|u^K - u^*\|^2.$$

Assim, tomando r suficientemente grande, $\sum_{k=1}^{\infty} p_k$ não é divergente, o que é um absurdo. Assim, o teorema acima fica provado. ■

O grande problema do método da série divergente está em sua velocidade de convergência. Ela é necessariamente baixa, pois é sempre pior que a velocidade de decréscimo do termo geral de uma série divergente.

III.3.2) O MÉTODO DO PASSO CONSTANTE

TEOREMA II.15 (MINOUX (1983)) Sejam $\delta_k = \|u^k - u^*\|$ e l como na seção (III.2). Tomemos $p_k = p = l\varepsilon$, $\varepsilon > 0$. Nestas condições, $\forall \varepsilon > 0$ o algoritmo do passo constante obtém um ponto u^k tal que $\delta_k \leq \varepsilon$ num número finito de iterações majorado por

$$\bar{k} = \frac{d^0 - \varepsilon}{\varepsilon(1 - (1 - l^2)^{1/2})}$$

DEMONSTRAÇÃO Suponhamos que $\delta^0 > \varepsilon$ (do contrário a demonstração seria trivial). Vem que

$$\delta_0 > \frac{p}{l}$$

Assim, é possível tomar um ponto ξ

$$\xi = \frac{p}{|l| \delta_0} u^* + \frac{|l| \delta_0 - p}{|l| \delta_0} u_0$$

no "segmento" que une u_0 a u^* , distanciado de u^0 por $p/|l|$ (ver figura III.4 que ilustra isso em duas dimensões).

Temos que

$$\begin{aligned} (\xi - u_1) \cdot (u_1 - u_0) &= (\xi - u_0 + u_0 - u_1) \cdot (u_1 - u_0) = (\xi - \\ &u_0) \cdot (u_1 - u_0) + \|u_0 - u_1\|^2 = \frac{p}{|l| \delta_0} (u^* - u_0) \cdot (u_1 - u_0) + \end{aligned}$$

$$\|u_0 - u_1\|^2.$$

Como

$$u_1 = u_0 + p \frac{g^0}{\|g^0\|}$$

Vem que

$$\|u_0 - u_1\| = p$$

E portanto,

$$(x - u_1) \cdot (u_1 - u_0) = \frac{p}{|1| \delta_0} (u^* - u^0) \cdot p \frac{g^0}{\|g^0\|} + p^2$$

Definindo

$$\theta = \arccos \frac{(u^* - u^0) \cdot g^0}{\delta_0 \|g^0\|}$$

Temos

$$(x - u_1) \cdot (u_1 - u_0) = \frac{p^2}{|1|} \cos\theta + p^2$$

Da definição de θ sabemos que $\cos\theta \geq -1$. Portanto,

$$(x - u_1) \cdot (u_1 - u_0) \geq \frac{p^2}{|1|} (-1) + p^2 \geq 0$$

Mas,

$$\|\xi - u_1\|^2 = \|\xi - u_1\|^2 + 2(\xi - u_1) \cdot (u_1 - u_0) + \|u_1 - u_0\|^2$$

ou

$$\|\xi - u_0\|^2 \geq \|\xi - u_1\|^2 + \|u_1 - u_0\|^2$$

Logo,

$$\frac{p^2}{l^2} \geq p^2 + \|u_1 - \xi\|^2$$

ou ainda,

$$\|u_1 - \xi\|^2 \leq \left(\frac{p^2}{l^2} - p^2 \right)^{1/2}$$

donde

$$\delta_1 \leq \delta_0 - \frac{p}{l} (1 - (1 - l^2)^{1/2}) < \delta_0$$

Podemos repetir este procedimento enquanto δ_k permaneça maior que ε . Como $(p/l)(1 - (1 - l^2)^{1/2}) > 0$, o número máximo de passos que precisamos para ter $\delta_k \leq \varepsilon$ é

$$\bar{k} = \frac{\delta_0 - \varepsilon}{(p/l)(1 - (1 - l^2)^{1/2})} = \frac{\delta_0 - \varepsilon}{\varepsilon (1 - (1 - l^2)^{1/2})}$$

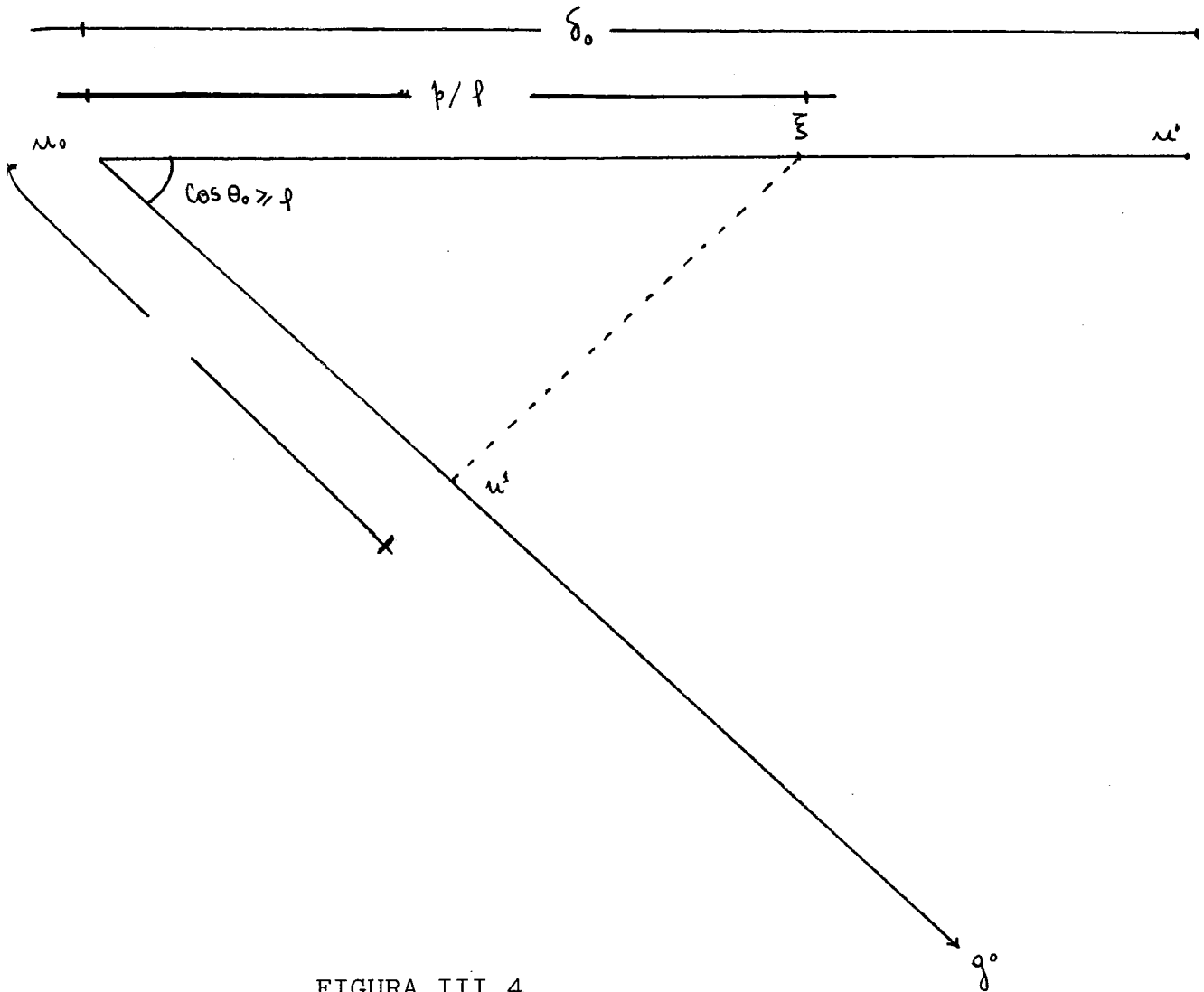


FIGURA III.4

o que demonstra o teorema acima. ■

O interesse prático no método do passo constante é muito limitado, pois sua convergência é muito lenta. Para $l = 0,25$, por exemplo, seriam necessárias 31.178 iterações para se alcançar uma precisão de $\delta_0/100$!

III.3.3) O MÉTODO DE RELAXAÇÃO

Conforme observado por A. J. HOFFMAN, este método pode ser visto como uma aplicação do método de AGMON, MOTZIKIN e SCHOENBERG (1954). Neste método define-se um passo variável dado por

$$p_k = p \frac{w^* - w(u^k)}{|g^k|}$$

onde p é um coeficiente de relaxação entre 0 e 2.

TEOREMA III.16 Se $0 < p < 2$, então $\delta_{k+1} < \delta_k$.

DEMONSTRAÇÃO Sendo w côncava,

$$w^* - w(u^k) \leq g^k (u^* - u^k).$$

Logo,

$$p_k \leq p \delta_k \cos \theta_k$$

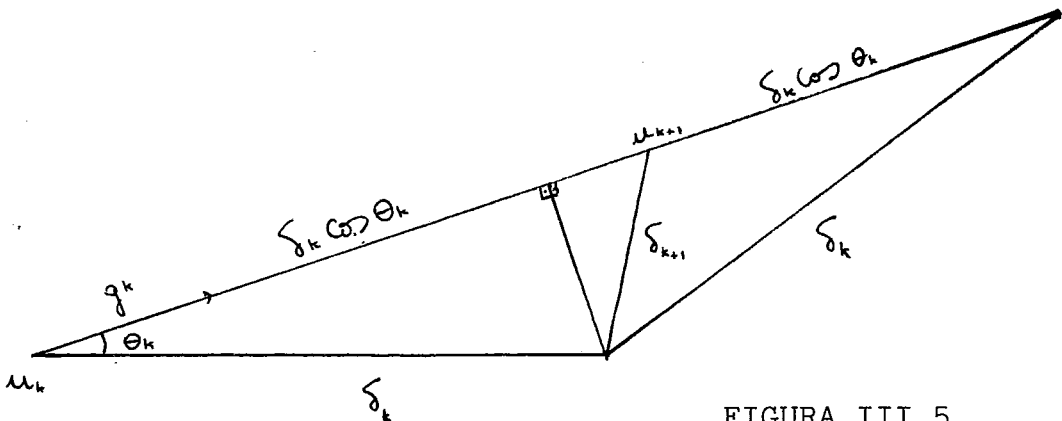


FIGURA III.5

$$0 < p < 2 \Rightarrow \delta_{k+1} < \delta_k \quad \blacksquare$$

Com isto, fica estabelecido que a distância ao ótimo diminui estritamente. Podemos facilmente fazer a prova efetiva da convergência com $p = 1$ e introduzindo a hipótese suplementar da existência de uma constante $0 < c < 1$ tal que para todo u e para todo subgradiente $g \in \partial w(u)$ se tenha

$$w^* - w(u) \geq c \langle g, u^* - u \rangle.$$

Visto que

$$1 \leq \frac{\langle g, u^* - u \rangle}{\|g\| \delta_k}$$

obtemos

$$p_k \geq c \delta_k$$

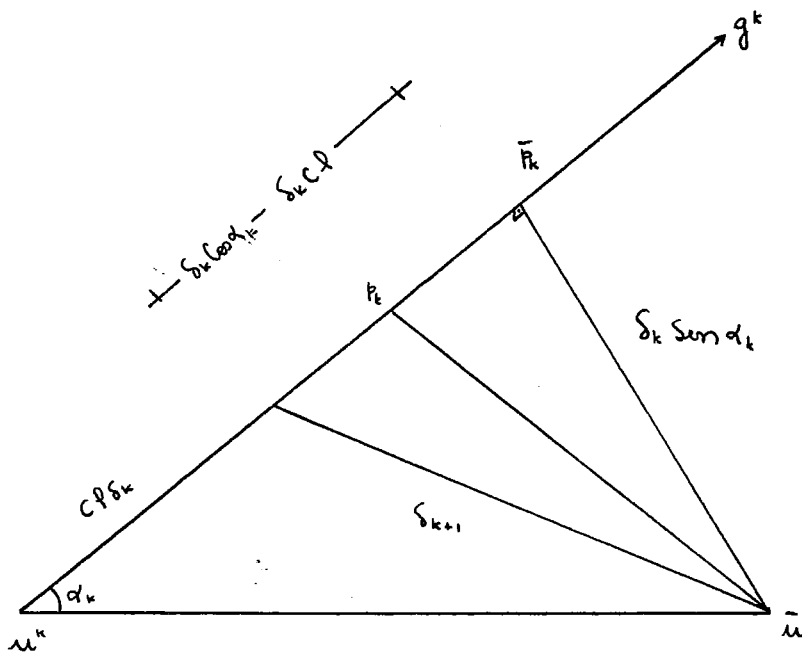


FIGURA III.6

Também,

$$\delta_{k+1}^2 \leq \delta_k^2 \sin^2 \alpha_k + \delta_k^2 (\cos \alpha_k - c)^2$$

e portanto,

$$\delta_{k+1}^2 \leq \delta_k^2 (1 - 2c \cos \alpha_k + c^2 \cos^2 \alpha_k) \leq \delta_k^2 (1 - c^2)$$

Donde se conclui que este tipo de escolha conduz a uma velocidade de convergência majorada por $(1 - c^2)^{1/2}$.

Normalmente, o valor w^* não é conhecido a priori e temos que substituí-lo na definição do passo no método de relaxação por um valor estimado $\bar{w} \geq w^*$ (\bar{w} pode ser obtido através do cálculo da função-objetivo do problema primal em um ponto primal-viável).

HELD, WOLFE e CROWDER (1974) usaram o procedimento acima, definindo um passo dado por

$$p_k = p \frac{\bar{w} - w(u^k)}{\|g^k\|}$$

onde $p = 2$ nas $2n$ primeiras iterações. Em seguida p passa a ser reduzido à metade após n iterações. É novamente reduzido à metade após $n/2$ iterações. De novo após $n/4$ iterações, e assim sucessivamente até o número de iterações em que p permanece constante se tornar muito pequeno. A partir daí, p é reduzido à metade a cada iteração.

Apesar de não se conhecer provas gerais de convergência deste algoritmo, os resultados computacionais

obtidos com ele têm sido tão bons que ele é seguramente o método mais utilizado para resolução do dual de problemas combinatórios de grande porte. Sua eficiência pode também ser bastante melhorada através de uma escolha judiciosa de seus parâmetros.

No capítulo V usaremos o método da relaxação na resolução do dual do problema da mochila quadrática.

CAPÍTULO IV

UTILIZAÇÃO DE DECOMPOSIÇÃO LAGRANGEANA NA OBTENÇÃO DE LIMITES PARA O ÓTIMO DE PROGRAMAS PSEUDO-BOOLEANOS NÃO-LINEARES COM VÍNCULOS LINEARES

Neste capítulo discutiremos o uso da técnica de decomposição lagrangeana na obtenção de limites para o ótimo de programas pseudo-booleanos não-lineares com restrições lineares, tendo em vista sua utilização em algoritmos enumerativos. Comparamos também os valores desses limites com aqueles obtidos via relaxação contínua usual.

Os resultados apresentados neste capítulo foram pela primeira vez obtidos por MICHELON e MACULAN (1988b), que aplicaram de forma pioneira a técnica de decomposição lagrangeana ao estudo de programas discretos não-lineares.

Na primeira seção deste capítulo introduziremos a técnica de decomposição lagrangeana, e na segunda seção apresentaremos resultados teóricos importantes para a utilização desta técnica. Estes resultados serão aplicados no capítulo V quando discutiremos a resolução do problema da mochila quadrática via decomposição lagrangeana.

IV.1) INTRODUÇÃO

Seja o programa pseudo-booleano não-linear com restrições lineares (PP)

$$(PP) \quad \begin{cases} \text{minimizar} & \langle f(x) \rangle \\ x \in \Omega_1 \cap \Omega_2 \cap \mathbb{B}^n \\ \Omega_1 = \{x \in \mathbb{R}^n / Ax - b = 0\} \\ \Omega_2 = \{x \in \mathbb{R}^n / Cx - d = 0\} \end{cases}$$

onde A é uma matriz com m_1 linhas e n colunas, b é um vetor-coluna com m_1 linhas, C é uma matriz com m_2 linhas e n colunas, e d é um vetor-coluna com m_2 linhas.

Este programa pode ser reescrito como

$$(PP)' \quad \begin{cases} \text{minimizar} & \langle f(y) \rangle \\ (y, x) \in \Omega' \\ \Omega' = \{y \in \Lambda, x \in \mathbb{B}^n / Ay - b = 0, y = x, Cx - d = 0\} \end{cases}$$

onde Λ é um subconjunto do \mathbb{R}^n que contém \mathbb{B}^n .

Relaxando as restrições $y = x$ de Ω' à moda lagrangeana, obtemos a relaxação

$$(PR) \quad \begin{cases} \text{minimizar} & \langle f(y) + u \cdot (y - x) \rangle \\ y \in \Omega_1 \cap \Lambda, x \in \Omega_2 \cap \mathbb{B}^n \end{cases}$$

e assim, podemos obter a função dual w relativa à relaxação (PR) acima em um dado ponto $u \in \mathbb{R}^n$ através da resolução de dois problemas: (PRuy) e (PRux).

$$(PRuy) \quad \begin{cases} \text{minimizar} & \langle f(y) + u \cdot y \rangle \\ y \in \Omega_1 \cap \Lambda \end{cases}$$

$$(PR_{ux}) \quad \begin{cases} \text{minimizar } \langle -u, x \rangle \\ x \in \Omega_2 \cap B^n \end{cases}$$

ou seja,

$$w(u) = \underset{y \in \Omega_1 \cap A}{\text{mínimo}} \langle f(y) + u \cdot y \rangle + \underset{x \in \Omega_2 \cap B^n}{\text{mínimo}} \langle -u, x \rangle$$

Vimos no capítulo III que o valor de uma função dual em qualquer ponto é sempre menor ou igual ao ótimo do programa primal correspondente. Isto obviamente também é verdade para a função dual w acima. De fato, seja $w \in \mathbb{R}^n$. Vem que

$$w(u) \leq \langle f(x^*) + u \cdot x^* \rangle + \langle -u, x^* \rangle = f^* + u \cdot x^* - u \cdot x^* = f^*,$$

onde x^* é uma solução ótima de (PP) e f^* é o valor ótimo da função-objetivo.

Desta forma, teremos também que o máximo da função dual será menor ou igual a f^* , ou seja, o problema dual

$$(PD) \quad \begin{cases} \text{maximizar } \{w(u)\} \\ u \in \mathbb{R}^n \end{cases}$$

fornece um limite inferior para o ótimo do programa primal. Além disso, a decomposição

$$(IV.1) \quad w(u) = \underset{y \in \Omega' \cap A}{\text{mínimo}} \{f(y) + u \cdot y\} + \underset{x \in \Omega \cap B^n}{\text{mínimo}} \{-u \cdot x\}$$

onde $\Omega = \Omega_1 \cap \Omega_2$ e $\Omega \supset \Omega$.

fornece uma solução viável para (PP). Neste caso então, (PD) fornece um limite superior para f^* . Estes limites podem ser usados, por exemplo, em um algoritmo de enumeração implícita para resolução de (PP).

OBSERVAÇÃO IV.1 Restringir-nos-emos daqui para diante à consideração da função dual (IV.1). Assim, quando nos referirmos ao problema dual (PD) estaremos implicitamente considerando o máximo da função (IV.1) no \mathbb{R}^n .

TEOREMA IV.1 Quando $\Omega = \Omega$, O problema dual (PD) fornece limites inferiores para f^* melhores, ou no pior caso iguais, aos obtidos através da relaxação contínua usual.

DEMONSTRAÇÃO Tomemos $A = [0,1]^n$ ou $A = \mathbb{R}^n$. Seja (\overline{PP}) a relaxação contínua de (PP)

$$(\overline{PP}) \quad \begin{cases} \text{minimizar } \{f(x)\} \\ x \in \Omega \cap A \end{cases}$$

Sejam \bar{x} uma solução ótima de (\overline{PP}) e \bar{f} o valor ótimo correspondente, ou seja

$$\bar{f} = f(\bar{x}) = \underset{x \in \Omega \cap A}{\text{mínimo}} \{f(x)\}.$$

Vem que

$$w^* \geq w(0) = \underset{y \in \Omega_A}{\text{mínimo}} \{f(y) + 0 \cdot y\} + \underset{x \in \Omega_B^n}{\text{mínimo}} \{-0 \cdot x\} =$$

$$\underset{y \in \Omega_A}{\text{mínimo}} \{f(y)\} = \bar{f} \quad \blacksquare$$

Este resultado tão promissor dá-nos uma idéia da potencialidade do uso de decomposição lagrangeana em programação pseudo-booleana. Mas ele é apenas "a ponta do iceberg".

A decomposição de (PR) em dois problemas independentes permite-nos uma obtenção computacional mais rápida de w em um ponto u . De fato, podemos processar (PR_{uy}) e (PR_{ux}) paralelamente, obtendo assim $w(u)$ em menos tempo.

Além disso, uma escolha judiciosa de Λ , Ω_1 , e Ω_2 pode transformar um problema difícil (no caso (PP)) em dois problemas bem mais simples ((PR_{ux}) e (PR_{uy})).

A idéia de criar cópias de algumas variáveis em programação matemática é anterior ao seu uso em relaxação lagrangeana (GLOVER e MULVEY (1975, 1980), por exemplo já usavam esta idéia para resolver problemas de fluxos em redes). Ela foi usada pela primeira vez em relaxação lagrangeana por SHEPARDSON e MARSTEN (1980), e por RIBEIRO (1983) conforme sugestão de Minoux. Mais tarde, GUIGNARD e KIM (1987) exploraram diversos aspectos teóricos desta nova relaxação e sugeriram para ela o nome de decomposição lagrangeana. Até então só se havia considerado o caso

linear, quando MICHELON e MACULAN (1988b) estudaram o caso não-linear. Seguiram-se a este trabalho BENHAMAMOUCH e PLATEAU (1989) que aplicaram a técnica de decomposição lagrangeana ao problema de alocação de recursos, e GUIGNARD (1989) que estendeu alguns resultados de MICHELON e MACULAN (1988).

Muita coisa ainda há por explorar no que se refere à aplicação desta técnica em problemas discretos não-lineares. Até agora, por exemplo, não encontramos na literatura qualquer trabalho que reportasse resultados computacionais avaliando o uso de decomposição lagrangeana na resolução de programas pseudo-booleanos. Faremos isto no capítulo V, quando discutiremos o problema da mochila quadrática. Antes porém apresentaremos, na próxima seção, alguns outros resultados teóricos importantes para o uso de decomposição lagrangeana em programas pseudo-booleanos não-lineares.

OBSERVAÇÃO IV.2 Daqui para diante tomaremos $\Lambda = \mathbb{R}^n$ ou $\Lambda = [0,1]^n$. ■

IV.2) RESULTADOS TEÓRICOS

Vimos na seção anterior que o ótimo do problema dual, quando $\Omega' = \Omega$, é sempre melhor ou igual ao limite inferior fornecido pela relaxação contínua usual. Vamos avaliar um pouco mais esse limite. Para isso, definiremos inicialmente o problema

$$(PP)^* \quad \left\{ \begin{array}{l} \text{minimizar } \{f(x)\} \\ x \in \text{CONV}(\Omega \cap B^n) \end{array} \right.$$

Note que o valor ótimo da relaxação $(PP)^*$ é sempre menor ou igual ao valor ótimo do problema primal, visto que $\Omega \cap B^n \subset \text{CONV}(\Omega \cap B^n)$.

Conforme observado por GEOFFRION (1974) é impraticável tentar obter uma expressão do poliedro definido por $\text{CONV}(\Omega \cap B^n)$ através de um sistema de desigualdades. Nosso interesse no problema $(PP)^*$, no entanto, não é prático mas teórico. Usaremos a definição de $(PP)^*$ para demonstrar alguns resultados.

Sejam \tilde{y} e \tilde{x} soluções ótimas para (PR_{uy}) e (PR_{ux}) , respectivamente, num dado ponto $u \in \mathbb{R}^n$ do domínio de w . Ou seja,

$$w(u) = f(\tilde{y}) + u \cdot \tilde{y} - u \cdot \tilde{x}.$$

De acordo com o teorema (III.6) $\tilde{y} - \tilde{x}$ é um subgradiente de w em u . Sabemos, pelo teorema (III.9), que $u \in \mathbb{R}^n$ será uma solução ótima para o problema dual se e somente se $0 \in \partial w(u)$. Assim, se para algum $u^* \in \mathbb{R}^n$ obtivermos $\tilde{y} = \tilde{x}$ este u^* será uma solução ótima para o problema dual. Mas não apenas isto.

TEOREMA IV.2 Se tivermos $\tilde{y} = \tilde{x}$ para algum $u^* \in \mathbb{R}^n$, então \tilde{x} é uma solução ótima para o problema primal e $w^* = f^*$.

DEMONSTRAÇÃO Conforme já vimos na seção anterior,

\tilde{x} será uma solução viável para (PP). Além disso, $f^* \geq w^*$.

Agora,

$$w^* = f(\tilde{y}) + u \cdot \tilde{y} - u \cdot \tilde{x} = f(\tilde{x}) + u \cdot \tilde{x} - u \cdot \tilde{x} = f(\tilde{x}).$$

Assim, \tilde{x} é uma solução ótima para (PP), e $w^* = f^*$. ■

Nem sempre teremos a felicidade de obter um limite tão bom quanto o limite acima. Apresentaremos agora um caso mais geral.

TEOREMA IV.3 Se f é convexa, então o valor ótimo do problema dual é igual ao valor ótimo de (PP)*.

DEMONSTRAÇÃO Podemos reescrever (PP)* como

$$(PP)^* \quad \begin{cases} \text{minimizar } \{f(y)\} \\ y \in \Omega \cap A \\ y = x \\ x \in \text{CONV}(\Omega \cap B^n) \end{cases}$$

Seja χ o valor ótimo de (PP)*. Visto que $\text{CONV}(\Omega \cap B^n) \subset \Omega \cap A$ e f é convexa,

$$\chi = \max_{u \in \mathbb{R}^n} \left\{ \min_{y \in \Omega \cap A, x \in \text{CONV}(\Omega \cap B^n)} \{f(y) + u \cdot (y - x)\} \right\}$$

(BAZARAA e SHETTY (1979, pp. 183-184))

$$= \max_{u \in \mathbb{R}^n} \left\{ \min_{y \in \Omega \cap A, x \in \Omega \cap B^n} \{f(y) + u \cdot (y - x)\} \right\} = w^*$$

E portanto o valor ótimo de (PP)^{*} é igual a w^* . ■

OBSERVAÇÃO IV.3 Note que neste caso também o limite inferior obtido é melhor ou igual àquele fornecido pela relaxação contínua usual. ■

Conforme já vimos, na melhor das hipóteses $w^* = f^*$, e no pior caso $w^* = \bar{f}$. Nesta segunda situação o uso de decomposição lagrangeana não oferece nenhuma vantagem no que diz respeito ao limite que se pretende conseguir.

DEFINIÇÃO IV.1 Se $\Omega' = \Omega$ Diremos que a decomposição lagrangeana (PR) possui a propriedade da integralidade se

$$\min_{x \in \Omega \cap B^n} \langle -u, x \rangle = \min_{x \in \Omega \cap \Lambda} \langle -u, x \rangle$$

para todo $u \in \mathbb{R}^n$. ■

TEOREMA IV.4 Se $\Omega_1 = \Omega_2$, f for convexa, e a decomposição lagrangeana possuir a propriedade da integralidade,

$$\bar{f} = w^*.$$

DEMONSTRAÇÃO

$$w^* = \max_{u \in \mathbb{R}^n} \left(\min_{y \in \Omega \cap \Lambda} \langle f(y) + u, y \rangle + \min_{x \in \Omega \cap B^n} \langle -u, x \rangle \right) =$$

$$\max_{u \in \mathbb{R}^n} \left(\min_{y \in \Omega \cap \Lambda} \langle f(y) + u, y \rangle + \min_{x \in \Omega \cap \Lambda} \langle -u, x \rangle \right) =$$

$$\min_{y \in \Omega_A} \langle f(y) \rangle = \bar{f}.$$

■

CAPÍTULO V

O PROBLEMA DA MOCHILA QUADRÁTICA

Neste capítulo aplicaremos os resultados dos capítulos anteriores, e em especial a técnica estudada no capítulo IV, na resolução do problema da mochila quadrática.

Na primeira seção colocaremos o problema e apresentaremos algumas aplicações. Na segunda seção discutiremos o uso de decomposição lagrangeana na obtenção de limites do ótimo e apresentaremos um algoritmo de enumeração implícita que usa estes limites para resolver o problema. Na terceira seção descreveremos a implementação deste algoritmo, e na quarta seção apresentaremos resultados computacionais, além de uma análise deles.

V.1 INTRODUÇÃO

DEFINIÇÃO V.1 Definimos o problema da mochila quadrática (MQ) como

$$(MQ) \quad \begin{cases} \text{minimizar } \left(\frac{1}{2} x^t \cdot H \cdot x + c^t \cdot x \right) \\ x \in \Omega \\ \Omega = \{ x \in \mathbb{B}^n / a^t \cdot x \geq b \} \end{cases}$$

onde x , c e a são vetores-coluna com n linhas, H é uma matriz quadrada de ordem n , e b um escalar. ■

OBSERVAÇÃO V.1 Sem perda de generalidade podemos considerar H uma matriz simétrica, a um vetor com componentes não-negativas, e $a_i \leq b \forall i \in \{1, \dots, n\}$. Suporemos ainda que $\Omega \neq \emptyset$. Conforme vimos no capítulo II podemos também, sem perda de generalidade, considerar H positiva-definida; e na prática podemos supor todas as entradas de H , as componentes de a e c , e b como números inteiros. ■

O problema da mochila quadrática foi tratado pela primeira vez por GALLO, HAMMER e SIMEONE (1980), que o apresentaram na forma

$$(MQ)_{GHS} \quad \begin{cases} \text{maximizar } x^t \cdot H \cdot x \\ x \in \Omega \\ \Omega = \{x \in \mathbb{B}^n / a^t \cdot x \leq b\} \end{cases}$$

com $H_{ij} \geq 0 \forall i, j \in \{1, \dots, n\}$. Note que os termo lineares da função-objetivo foram absorvidos na diagonal principal de H .

GALLO, HAMMER e SIMEONE (1980) resolveram o problema da mochila quadrática $(MQ)_{GHS}$ através de um algoritmo de enumeração implícita que usava como limites superiores o máximo de planos-superiores de $x^t \cdot H \cdot x$ em $\Lambda \supset \Omega$, e como limites inferiores valores de $x^t \cdot H \cdot x$ em um ponto de Ω .

Conforme vimos no capítulo II um plano-superior de $x^t \cdot H \cdot x$ em $\Lambda \supset \Omega$ é um hiper-plano $\xi^t \cdot x \geq x^t \cdot H \cdot x, \forall x \in \Lambda$, onde ξ é um vetor-coluna com n linhas.

GALLO, HAMMER e SIMEONE (1980) sugeriram 4 planos-superiores com os quais obtiveram limites para o valor ótimo de (MQ)_{GHs} usados em um algoritmo de enumeração implícita (embre que eles supuzeram $H_{ij} \geq 0 \forall i, j \in \{1, \dots, n\}$):

$$(i) \quad \zeta_j^1 = \sum_{i=1}^n H_{ij}$$

(ii) $\zeta_j^2 = \sum_i^{(h)} H_{ij}$, onde h é o número máximo de 1's em uma solução viável, e $\sum_i^{(h)} H_{ij}$ é a soma dos h maiores elementos da coluna j de H

$$(iii) \quad \zeta_j^3 = \max_{x \in \{0,1\}^n} \left(\sum_{i=1}^n H_{ij} x_i \right)$$

$$(iv) \quad \zeta_j^4 = \max_{x \in B^n} \left(\sum_{i=1}^n H_{ij} x_i \right)$$

O problema da mochila quadrática possui diversas aplicações. WITZGALL (1975) formula o problema de localização de estações de transmissão de mensagens eletrônicas como (MQ). De maneira semelhante este modelo foi usado para estudar problemas de localização de estações de trem (LAND (1975)), problemas de localização de terminais de carga e de aeroportos (RHYS (1970)).

Outras aplicações possíveis são: seleção de portfólio (LAUGHUNN (1970)), redução do número de pluviômetros em uma estação pluviométrica, e a determinação de se um dado grafo possui um clique de ordem k (GALLO, HAMMER e SIMEONE (1980)).

V.2) OBTENÇÃO DE LIMITES VIA DECOMPOSIÇÃO LAGRANGEANA

Dado o problema da mochila quadrática (MQ)

$$\left\{ \begin{array}{l} \text{minimizar } \langle \frac{1}{2} x^t \cdot H \cdot x + c^t \cdot x \rangle \\ x \in \Omega_{MQ} \\ \Omega_{MQ} = \{x \in \mathbb{B}^n / a^t \cdot x \geq b\} \end{array} \right.$$

conforme vimos no capítulo IV podemos reescrevê-lo como (MQ)₁

$$\left\{ \begin{array}{l} \text{minimizar } \langle \frac{1}{2} y^t \cdot H \cdot y + c^t \cdot y \rangle \\ (y, x) \in \Omega_{MQ1} \\ \Omega_{MQ1} = \{y \in [0,1]^n, x \in \mathbb{B}^n / a^t \cdot y \geq b, y = x, a^t \cdot x \geq b\} \end{array} \right.$$

ou (MQ)₂

$$\left\{ \begin{array}{l} \text{minimizar } \langle \frac{1}{2} y^t \cdot H \cdot y + c^t \cdot y \rangle \\ (y, x) \in \Omega_{MQ2} \\ \Omega_{MQ2} = \{y \in \mathbb{R}^n, x \in \mathbb{B}^n / a^t \cdot y \geq b, y = x, a^t \cdot x \geq b\} \end{array} \right.$$

ou ainda (MQ)₃

$$\left\{ \begin{array}{l} \text{minimizar } \langle \frac{1}{2} y^t \cdot H \cdot y + c^t \cdot y \rangle \\ (y, x) \in \Omega_{MQ3} \\ \Omega_{MQ3} = \{y \in \mathbb{R}^n, x \in \mathbb{B}^n / y = x, a^t \cdot x \geq b\} \end{array} \right.$$

Relaxando (MQ)₁, (MQ)₂ e (MQ)₃ à moda

lagrangeana nas restrições $y = x$ e decompondo os problemas relaxados, obtemos as funções duais correspondentes

$$w_1(u) = \underset{y \in \Omega_{R_1 y}}{\text{mínimo}} \left\{ \frac{1}{2} y^t \cdot H \cdot y + (c^t + u^t) \cdot y \right\} + \underset{x \in \Omega_{R_x}}{\text{mínimo}} \{ -u^t \cdot x \}$$

$$w_2(u) = \underset{y \in \Omega_{R_2 y}}{\text{mínimo}} \left\{ \frac{1}{2} y^t \cdot H \cdot y + (c^t + u^t) \cdot y \right\} + \underset{x \in \Omega_{R_x}}{\text{mínimo}} \{ -u^t \cdot x \}$$

$$w_3(u) = \underset{y \in \Omega_{R_3 y}}{\text{mínimo}} \left\{ \frac{1}{2} y^t \cdot H \cdot y + (c^t + u^t) \cdot y \right\} + \underset{x \in \Omega_{R_x}}{\text{mínimo}} \{ -u^t \cdot x \}$$

onde

$$\Omega_{R_1 y} = \{ y \in [0, 1]^n / a^t \cdot y \geq b \}$$

$$\Omega_{R_2 y} = \{ y \in \mathbb{R}^n / a^t \cdot y \geq b \}$$

$$\Omega_{R_3 y} = \mathbb{R}^n$$

$$\Omega_{R_x} = \{ x \in \mathbb{B}^n / a^t \cdot x \geq b \}$$

Com relação a w_3 , visto que H é positiva definida, podemos escrever que

$$w_3(u) = -\frac{1}{2} (c^t + u^t) \cdot H^{-1} \cdot (c + u) + \underset{x \in \Omega_{R_x}}{\text{mínimo}} \{ -u^t \cdot x \}$$

Assim, para obter o valor de w_1 ou w_2 em um dado ponto $u \in \mathbb{R}^n$, basta resolver um problema quadrático no contínuo e um problema de mochila linear. Mais simples ainda, quanto a w_3 , basta obter a inversa de H e uma solução ótima de um problema de mochila linear.

De acordo com o que vimos nos capítulos III e IV, para um dado $\tilde{u} \in \mathbb{R}^n$, se \tilde{y} é uma solução ótima do problema quadrático e contínuo e \tilde{x} uma solução ótima do problema da mochila linear, então $\tilde{y} - \tilde{x}$ é um subgradiente da função dual em \tilde{u} .

Quanto aos problemas duais (MQD)₁, (MQD)₂ e (MQD)₃, correspondentes a w_1 , w_2 e w_3

$$(MQD)_1 \quad \left\{ \begin{array}{l} \text{maximizar } w_1(u) \\ u \in \mathbb{R}^n \end{array} \right.$$

$$(MQD)_2 \quad \left\{ \begin{array}{l} \text{maximizar } w_2(u) \\ u \in \mathbb{R}^n \end{array} \right.$$

$$(MQD)_3 \quad \left\{ \begin{array}{l} \text{maximizar } w_3(u) \\ u \in \mathbb{R}^n \end{array} \right.$$

de acordo com o que vimos no capítulo IV ambos possuem o mesmo valor ótimo, ou seja

$$\max_{u \in \mathbb{R}^n} w_1(u) = \max_{u \in \mathbb{R}^n} w_2(u) = \max_{u \in \mathbb{R}^n} w_3(u)$$

Portanto, teoricamente a escolha de w_1 , w_2 ou w_3 para a obtenção de um limite inferior maior é

indiferente.

A fim de avaliar o desempenho computacional das decomposições acima na resolução do problema da mochila quadrática implementamos o algoritmo abaixo, que usa os valores ótimos de $(MQD)_1$, $(MQD)_2$ ou $(MQD)_3$ como limites inferiores.

ALGORITMO V.1

INÍCIO

LS $\leftarrow \infty$ (LS é o limite superior corrente)

CORPO

Enquanto há vértices vivos ("live vertices") Faça

Se o vértice atual for inviável

Então

Podar o vértice atual (Fathom)

Se houver vértices vivos

Então

Backtracking para o nó vivo mais recentemente gerado

Fim-se

Senão

Obter, através da resolução de um dual, um limite inferior (que será denotado por LI) e um limite superior para o subproblema correspondente ao vértice atual. Sejam w^* e $\frac{1}{2} \tilde{x}^t \cdot H \cdot \tilde{x} + c^t \cdot \tilde{x}$, respectivamente, estes valores.

LI $\leftarrow w^*$

Se LS $> \frac{1}{2} \tilde{x}^t \cdot H \cdot \tilde{x} + c^t \cdot \tilde{x}$

Então

$$LS \leftarrow \frac{1}{2} \tilde{x}^t \cdot H \cdot \tilde{x} + c^t \cdot \tilde{x}$$

Fim-se

Se $LI \geq LS$

Então

O vértice corrente é podado

Se houver vértices vivos

Então

Backtracking para o vértice vivo mais recentemente gerado.

Fim-se

Senão

Escolher uma variável de ramificação ("Branching variable").

Ramificar o vértice corrente, gerando dois vértices vivos, um dos quais tem a variável de ramificação fixada com valor zero e o outro com valor um. Escolher um dos vértices para se tornar o vértice atual.

Fim-se

Fim-se

Fim-Enquanto

FIM

Tomar a solução viável que forneceu LS como uma solução ótima e tomar LS como valor ótimo do problema primal. ■

V.3) IMPLEMENTAÇÃO COMPUTACIONAL

Descreveremos sucintamente nesta seção a implementação computacional do algoritmo (V.1). Deixaremos para a próxima seção a apresentação dos resultados obtidos, e para a seção (V.5) sua análise.

AMBIENTE OPERACIONAL

O algoritmo (V.1) foi implementado em Fortran IV em ambiente VM/CMS e equipamento IBM/4381.

O ALGORITMO DE ENUMERAÇÃO IMPLÍCITA

Foram implementadas três versões V_1 , V_2 , V_3 , conforme a função dual usada: w_1 , w_2 , w_3 , respectivamente.

Para a exploração da árvore de enumeração adotamos LIFO como regra de prioridade na escolha do nó corrente. Para variável de ramificação escolhemos sempre a de menor índice que possuísse um valor diferente de zero ou um na solução da parte quadrática e contínua do problema dual, e na ramificação escolhemos sempre o vértice com esta variável fixada em zero para se tornar o vértice corrente.

A RESOLUÇÃO DO PROBLEMA DUAL

Para resolução do problema dual implementamos o algoritmo (III.1) variando o passo como sugerido em HELD, WOLFE e CROWDER (1974) (v. seção III.3.3)

RESOLUÇÃO DA PARTE DISCRETA E LINEAR DA RELAXAÇÃO LAGRANGEANA

A parte linear e discreta da relaxação lagrangeana, que fornece o valor da função dual em um dado ponto, é um problema de mochila pseudo-booleana linear. Para sua resolução utilizamos uma implementação do algoritmo de MARTELLO e TOTH (1987) cedida pelos autores.

RESOLUÇÃO DA PARTE CONTÍNUA E QUADRÁTICA DA RELAXAÇÃO LAGRANGEANA

Podemos transformar um problema de programação quadrática no contínuo com restrições lineares em um problema de complementaridade linear, o qual pode ser resolvido pelo algoritmo de LEMKE (1968) (BAZARAA e SHETTY (1979, capítulo 11)). Para resolução da parte contínua e quadrática da relaxação lagrangeana nas versões V_1 e V_2 do algoritmo $V.1$ implementamos o algoritmo de Lemke. Segundo GALLO, HAMMER e SIMEONE (1980), na maioria das aplicações a matriz H é bastante esparsa. Assim sendo, para inversa da base utilizamos a implementação da decomposição LU descrita em REID (1976), e para atualização da inversa da base utilizamos o "método de redução de Bump" de REID (1975) (ver também CHVATÁL (1983, capítulo 24)). Quanto a V_2 , a inversão de H foi feita segundo o mesmo algoritmo de decomposição LU mencionado acima.

V.4) RESULTADOS COMPUTACIONAIS

Nesta seção apresentamos resultados obtidos com o processamento das três versões implementadas no algoritmo (V.1). Foram processadas 10 instâncias cada, para 5, 10, 15, 20 e 25 variáveis.

Na tabela V.1 apresentamos o tempo máximo (TM) e o tempo médio (tm) de processamento em segundos para V_1 , V_2 e V_3 , e para $n = 5, 10, 15, 20$ e 25 . Na tabela V.2 apresentamos o número máximo (NMD) e o número médio (nm) de vértices processados para V_1 , V_2 e V_3 , e para $n = 5, 10, 15, 20, 25$. Na tabela (V.3) apresentamos o tempo médio de processamento em segundos dividido pelo número de nós processados para V_1 , V_2 e V_3 , e para $n = 5, 10, 15, 20, 25$. Finalmente na tabela V.4 apresentamos uma "classificação" das três versões mostrando para cada n quantos problemas foram resolvidos no vértice inicial.

TABELA V.1

n	VERSAO			
		V_1	V_2	V_3
5	tm	0,74	1,14	0,12
	TM	1,94	1,74	0,25
10	tm	8,31	8,24	0,63
	TM	33,12	13,59	1,94
15	tm	21,18	19,52	3,39
	TM	63,72	100,88	7,04
20	tm	5,32	64,75	9,05
	TM	28,35	346,33	20,44
25	tm	34,51	99,31	23,19
	TM	157,07	723,40	60,77

Tempo médio (tm) e tempo máximo (TM) de processamento em segundos de CPU. ■

TABELA V. 2

n	VERSÃO			
		V ₁	V ₂	V ₃
5	nm	3	5	6
	NM	6	8	12
10	nm	5	14	13
	NM	10	20	34
15	nm	8	34	38
	NM	22	136	78
20	nm	3	77	75
	NM	12	404	172
25	nm	7	96	137
	NM	30	674	366

Número médio (nm) e número máximo de vértices (NM) processados. ■

TABELA V. 3

n	VERSÃO			
		V ₁	V ₂	V ₃
5		0,24	0,22	0,02
10		1,66	0,58	0,04
15		2,64	0,57	0,08
20		1,77	0,84	0,12
25		4,93	1,03	0,16

Tempo médio de processamento em segundos por número médio de nós processados. ■

TABELA V. 4

n	VERSÃO		
	V ₁	V ₂	V ₃
5	4	3	1
10	2	1	0
15	2	0	0
20	8	1	0
25	4	0	0

TOTAL : V₁ - 40%V₂ - 8%V₃ - 2%

Número de problemas resolvidos no nó inicial. ■

V.5) CONCLUSÃO

Vimos na seção (V.2) que ambas as funções duais usadas nas versões V₁ e V₃ possuem um mesmo máximo. Na prática entretanto, o método de subgradiente que usamos (relaxação) pode não nos conduzir a este valor. No entanto, caso tomemos uma região de viabilidade para o termo quadrático e contínuo da decomposição não muito diferente da região de viabilidade do termo linear e discreto, teremos uma chance maior de alcançarmos este ponto u^* de máximo, uma vez que, neste caso, em geral teremos $\tilde{y} - \tilde{x}$ mais próximo da origem em cada problema dual. Além disso, como no caso V₁ são boas chances de que exista u^* tal que $\tilde{y} = \tilde{x}$,

podemos esperar que, na prática os limites obtidos na otimização da função dual usada na versão V_1 do algoritmo (V.1) sejam bem melhores que aqueles obtidos na versão V_3 . Assim sendo, podemos entender por que na versão V_1 tivemos um número bem menor de vértices pesquisados por problema que na versão V_3 , e por que tivemos um percentual tão alto de problemas resolvidos no vértice inicial no caso V_1 (ver tabelas (V.2) e (V.4)).

É de se esperar entretanto que a obtenção do ótimo da função dual em um dado vértice seja mais barata na versão V_3 que na versão V_1 , devido ao custo computacional menor da parte quadrática e contínua da versão V_3 . De fato, como mostra a tabela (V.3) é exatamente isto que acontece. Vemos aí que não só o custo de cada problema dual em V_3 na implementação feita é bem menor, mas é tão menor que apesar do número de vértices pesquisados ser muito maior, o tempo total de processamento de V_3 (ver tabela (V.1)) é em geral bem menor que o de V_1 .

Quanto a V_2 apesar de (como se era de esperar) o número de nós pesquisados em geral ser menor que o número de nós pesquisados em V_3 (ver tabela (V.2)) e do percentual de problemas resolvidos no vértice inicial ser bem maior, o desempenho da implementação da versão V_3 se mostrou muito melhor que o da versão V_2 (ver tabela (V.1)). Também o desempenho de V_1 foi bem melhor que o de V_2 . A razão disto é que a melhora no número total de nós pesquisados na versão V_2 não é significativa em relação a V_3 , e portanto o número de nós pesquisados em V_2 é pouca coisa melhor que o número de nós pesquisados em V_3 e muito maior que o número de nós

pesquisados em V_1 . (ver tabela (V.2)). Como o tempo de processamento de cada problema dual da versão V_2 é bem maior que o tempo de processamento de cada dual na versão V_3 (ver tabela (V.3)), V_2 acaba se tornando desvantajosamente caro.

Como prosseguimento deste trabalho podemos sugerir várias alternativas que talvez venham a melhorar ainda mais os resultados obtidos. Para começar, podemos pesquisar outros algoritmos para programas quadráticos que se mostrem mais eficientes que os usados nas três implementações. Além disso, conforme vimos no capítulo II, todo programa quadrático no hiper-cubo unitário possui uma solução ótima em um de seus vértices. Assim sendo, a decomposição

$$w_4(u) = \underset{y \in \{0,1\}^n}{\text{mínimo}} \left\{ \frac{1}{2} y^t \cdot H \cdot y + c^t \cdot y \right\} + \underset{x \in \Omega_{Rx}}{\text{mínimo}} \left\{ - u^t \cdot x \right\}$$

cujo máximo é igual ao das outra três analisadas, pode na prática fornecer limites inferiores ainda melhores que os de V_1 , além de ser mais barata.

Podemos também tentar melhorar o desempenho do algoritmo de subgradiente usado, testando valores diferentes para seus parâmetros.

Quanto ao método de subgradiente usado, podemos trocá-lo por outro que garanta a subida do valor da função dual a cada iteração (MARTINS (1985)).

Podemos ainda tentar melhorar o desempenho do algoritmo enumerativo. Uma alternativa para isso seria, por exemplo, o uso de penalidades.

Michelon e Maculan (1988) propuseram dois algoritmos de corte (mais especificamente dois algoritmos de reforço do dual (ver BÁRCIA (1985)) que resolvem problemas pseudo-boleanos não-lineares com restrições lineares. A implementação e avaliação de seu desempenho computacional é promissoramente interessante.

No caso específico do problema da mochila quadrática, seria interessante fazer uma comparação do desempenho computacional dos algoritmos implementados neste trabalho com o de GALLO, HAMMER e SIMEONE (1980).

Devido aos resultados bem-sucedidos no desempenho da versão V_9 , cremos que este algoritmo será particularmente eficiente na resolução do problema da mochila quadrática inteira (problema este que, segundo nos consta, não foi ainda testado computacionalmente, ou sequer resolvido).

O uso de decomposição lagrangeana em programação discreta não-linear, além de se mostrar uma poderosa alternativa para resolução de tais problemas, oferece também um vasto campo para investigações, e cremos que o leitor poderá encontrar muitas outras além das acima sugeridas.

REFERÊNCIA BIBLIOGRÁFICAS

AGMON, S., MOTZIKIN, T. e SCHONBERG, I. (1954), "The Relaxation Method for Linear Inequalities," Canadian Journal of Mathematics 6, pp. 382-404.

AHRENS, W. (1974), "Die Loesung Eines Nichtlinearen Investitionsproblems Mit Hilfe Binaeres Optimierungsalgorithmen-Gezeigt Am Biespiel der Planung Regionaler Abwasserbehandlunssysteme," Zeitschrift fur Operations Research 18 (1974) B131-B147.

BALAS, E. e JEROSLOW, R. (1972), "Canonical Cuts on the Unity Hypercube," SIAM Journal on Applied Mathematics 23, 61-69.

BALAS, E. e MAZZOLA, J. (1984), "Nonlinear 0-1 Programming I: Linearization Techniques," Mathematical Programming 30, 1-21.

BALAS, E. e MAZZOLA, J. (1984b), "Nonlinear 0-1 Programming II: Dominance Relations and Algorithms," Mathematical Programming 30, 22-45.

BAR, G. (1972), "Zur Linearen Darstelbarkeit von Ausdrücken des Aussgenskalkuls," EIK 8, 353-378.

BÁRCIA, P. (1985), "Métodos de Reforço do Dual em Programação Inteira," Tese de Doutorado, Instituto

Superior Técnico da Escola de Engenharia de Sistemas da
Universidade Nova de Lisboa.

BAZARAA, M., e SHETTY, C.M. (1979), "Nonlinear Programming,
Theory and Algorithms," (J. Wiley & Sons, New York).

BENHAMAMOUCH, D. et PLATEAU, G. (1989), "Les Problemmes
d'allocation de ressources dans les systems distribues
(synthese)," Prepublication 89, Serie Informatique,
Universite Paris-Nord.

BERMAN, G. (1969), "A Branch and Bound Method for
Maximization of Pseudo-boolean functions," Faculty of
Mathematics, University of Waterloo, Canada.

BRESNEV, V. (1979), "Algorithms for the minimization of
polynomials with Boolean Variables," (Russian)
Problemy Kibernetiki (Moskow) 36, 225-246.

CAMION P. (1960), "Une Méthode de résolution par l'algebre de
Boole des Problèmes Combinatoires ou Interviennent des
Entiers," Cahiers du Centre d'Études de Recherche
Opérationnelle 2(3), 234-289.

CHVATAL, V. (1983), "Linear Programming," W.H. Freeman and
Company (N. York).

COCA-BALTA, A.G., MARIETTO, M.G.B. e CLEIMAN, D.F. (1990),
"Otimização Não-Linear 0-1 em Planejamento e Controle da

Produção Automatizada," Manuscrito, COPPE/UFRJ - Programa de Engenharia de Sistemas e Computação.

CRAMA, Y., HANSEN, P. e JAUMARD B. (1990), "The Basic Algorithm for Pseudo-boolean Programming Revisited," to appear in Discrete Applied Mathematics.

CRAVEN, B.D. (1988), "Fractional Programming," (Heidermann Verlag, Berlin).

DANTZIG, G.B. (1958), "On the Significance of Solving Linear Programming Problems with some Integer Variables," The Rand Corporation, document P1486, também publicado em *Econometrica* 28 (1960) 30-44.

FLORIAN, M. e ROBILLARD, P. (1971), "Programmation Hyperbolique en Variables Bivalentes," *Revue Française d'Informatique et de Recherche Opérationnelle* 5, 3-9.

FORTET, R. (1959) , "L'algèbre de Boole et ses applications en Recherche Opérationnelle," *Cahiers du Centre d'Études de Recherche Opérationnelle*, 1(4), 5-36.

FORTET, R. (1960), "Applications de L'algèbre de Boole en Recherche Opérationnelle," *Revue Française d'Informatique et de Recherche Opérationnelle* 4(14) 17-25.

GALLO, G., HAMMER, P. e SIMEONE, B., "Quadratic Knapsack Problems," *Math. Prog.* 12, 132-149.

GARFINKEL, R. e NEMHAUSER, G. (1972), "Integer Programming," (J. Willey & Sons, N. York).

GEOFFRION, A. (1974), "Lagrangean Relaxation for Integer Programming," Math. Prog. 2, 82-114.

GILLMORE, P.C. e GOMORY, R.E. (1963), "A Linear Programming Approach to the Cutting-Stock Problem - Part II," Op. Res. 11 863-888.

GINSBURGH, V. e VAN PEETERSEN (1969), "Un Algorithme de Programmation Quadratique en Variables Binaires," Rev. Française Informat. Recherche Op. 3(2) 57-74.

GLOVER, F. e MULVIEY, J. (1975), "The Equivalence of 0-1 Integer Programming Problem to Discrete Generalized and Pure Network Models," Report HBS 75-46, Harvard University, também Op. Res. 28 (1980) 829-933.

GLOVER, F. e WOOLSEY, E. (1973), "Further Reduction of Zero-One Polynomial Programs to Zero-One Linear Programming Problems," Op. Res. 21(1), 156-161.

GRANOT, D. e GRANOT, F. (1980), "Generalized Covering Relaxation in 0-1 Programming," Op. Res. 28(6), 1442-1450.

GRANOT, D., GRANOT, F. e KALLBERG, J., "Covering Relaxation for Positive 0-1 Polynomial Programming," Management

Science 25(3), 264-273.

GRANOT, D., GRANOT, F. e VAESSEN, W. (1982), "An accelerated Covering Relaxation Algorithm for Solving Positive 0-1 Polynomial Programs," Mathematical Programming 22, 350-357.

GRANOT, F. e HAMMER, P. (1971), "On the use of Boolean Functions in 0-1 Programming," Methods of Operations Research 12, 154-184.

GUIGNARD, M. (1989), "Lagrangean Decomposition for Integer problems with multiple subsets of linear constraints and nonlinear objective functions," Working Paper, University of Pennsylvania.

GUIGNARD, M. e KIM, S. (1987), "Lagrangean Decomposition: A model yeilding stronger lagrangean bounds," Mathematical Programming, vol. 39, pp. 215-228.

HAMMER, P.L. (1968), "Plant Location, A Pseudo-Boolean Approach," Israel Journal of Technology 350-362.

HAMMER, P.L. (1971), "A B.B.B. Method for Linear and Nonlinear Bivalent Programming," in B. Avi-Itzhak, ed., Developments in Operations Research (Gordon and Breach, N. York), 45-82

HAMMER, P.L. (1973), "BABO - A Boolean Approach for Bivalent

Optimization," in M. Avriel, M.J. Rijckaert and D.J. Wiece, eds., Optimization and Design, (Prentice Hall, Englewood Cliffs, NJ) 473-489.

HAMMER, P.L. (1974), "Boolean Procedures for Bivalent Programming," in P.L. Hammer and G. Zoutendijk, eds., Mathematical Programming in Theory and Practice (North Holland, Amsterdam) 311-363.

HAMMER, P.L., HANSEN, P. e SIMEONE, B. (1984), "Rooft Duality, Complementation and Persistency in Quadratic 0-1 Optimization," Mathematical Programming 28, 121-155.

HAMMER, P.L., e PELED, U. (1972), "On the Maximization of a Pseudo-boolean Function," Journal of the Association for the Computing Machinery (19), 265-282.

HAMMER, P.L., e ROSENBERG, I. (1972), "On Equivalent Forms of Pseudo-Boolean Programs," in S. Zaremba, ed., Applications of Number Theory to Numerical Analysis (Academic Press, New York) pp. 453-463.

HAMMER, P.L., ROSENBERG, I. e RUDEANU, S. (1963), "On the Determination of the Minima of Pseudo-boolean functions," (in Romanian) Studii si Cercetari Matematice 14, 359-364.

HAMMER, P.L., ROSENBERG, I. e RUDEANU, S. (1963b), "Application of Discrete Linear Programming to the Minimization of Boolean Functions," (in Russian) Rev.

Math. Pures Appl. 8, 459-475.

HAMMER, P.L., e RUBIN, A. A.(1970), "Some Remarks on Quadratic Programming with 0-1 Variables," Rev. Française Informat. Recherche Opérationnelle 4, 67-79.

HAMMER, P.L., e RUDEANU, S. (1968), ;"Boolean Methods in Operations Research and Related Areas", (Springer, Berlin).

HANSEN, P. (1969), "Un Algorithme S.E.P. pour les Programmes Pseudo-Booléens non Linéaires," Cahiers du Centre d'Études de Recherche Opérationnelle (11) 26-44.

HANSEN, P. (1969b), "Note sur L'extension de la méthode d'énumération implicite aux programmes non linéaires en variables zéro-un," Cahiers Centre Études Rech. Oper. 11, 162-166.

HANSEN, P. (1970), "Un Algorithme pour les Programmes Non Linéaires en Variables Zéro-un," Comptes Rendus Acad.Sci. Paris 270, 1700-1702.

HANSEN, P. (1972), : "Quadratic Zero-One Programming by Implicit Enumeration," in F.A. Lootsma, ed., Numerical Methods in Nonlinear Optimization (Academic Press, New York), 265-278.

HANSEN, P. (1974), "Programmes Mathématiques en Variables

0-1," Thèse d'agrégation, Université Libre de Bruxelles.

HANSEN, P. (1979), "Methods of Nonlinear 0-1 Programming,"
Annals of Discrete Mathematics 5, 53-70.

HANSEN, P., JAUMARD, A., e MATHON, V. (1989), "Constrained
Nonlinear 0-1 Programming," Working Paper, GERAD -
Université McGill, Montreal, Canada.

HELD, M., WOLFE, P., CROWDER, H. (1979), "Validations of
Subgradient Optimization," Mathematical Programming 6,
pp. 62-68.

HILLIER, F. (1969), "The Evaluation of Risky Interrelated
Investments," (North-Holland, Amsterdam)

INAGAKI, T. e FUKUMURA, T. (1967), "Pseudo-boolean
Programming with Constraints," Journal of Electronics and
Communications Engineering of Japan 50, 26-34.

KOLESAR, P. (1980), "Testing for Vision Loss in Glaucoma
Suspects," Management Sci. 26(5) 439-450.

LAND, A. (1975), "Personal communication in Gallo, Hammer e
Simeone (1980).

LANG, S., "Cálculo," vol. 2, Livros Técnicos e Científicos
S/A.

- LAUGHUNN, D. J. (1970), "Quadratic Binary Programming with Applications to Capital-Budgeting Problems," *Op. Res.* 14, 454-461.
- LAWLWER, E.L., e BELL, M.D., "A Method for Solving Discrete Optimization Problems," *Op. Res.* 14, 1098-1112.
- LEMKE, C. (1968), "On Complementary Pivot Theory," in *Mathematics of the Decision Science*, G.B. Dantzig e A.F. Veinott, eds.
- LIMA, E.L. (1981), "Curso de Análise, vol. 2," Projeto Euclides - IMPA, Livros Técnicos e Científicos S/A.
- LU, S-H e WILLIAMS, A. (1987), "Roof Duality for Polynomial Zero-one optimization," *Mathematical Programming* 37, 357-360.
- MAO, J. e WALLINGFORD, B. (1968), "An Extension of Lawlwr and Bell's Method of Discrete Optimization with Examples from Capital Budgeting," *Management Science* 15, 51-60.
- MAO, J. e WALLINGFORD, B. (1969), "Corrections and Comments on An Extension of Lawler and Bell's Method of Discrete Optimization with Examples from Capital Budgeting," *Management Science* 15, 481.
- MARTELLO, S. e TOTH, P. (1987), "A New Algorithm for the 0-1 Knapsack Problem," Working Paper, DEIS - University of

Bologna, Italy.

MARTINS, R.F. (1987), "Um Estudo de Alguns Métodos de Subgradiente para a Resolução do Dual de Problemas de Otimização Combinatória," Tese de Mestrado, Departamento de Engenharia Elétrica, PUC/RJ.

MATTA, S.S. da (1989), "Problema do Horário na Escola de Segundo Grau: Modelagem e Implementação," Tese de Mestrado - COPPE/UFRJ, Programa de Engenharia de Sistemas e Computação.

MICHELON, P. e MACULAN, N. (1988), "An Algorithm for the Mixed or Integer Fractional Programming," Paper presented in MP in Kioto, Kioto, Japan, September.

MICHELON, P. e MACULAN, N. (1988b), "Lagrangian Decomposition for Integer Nonlinear Programming with Linear Constraints," Technical Report, COPPE - Systems Engineering, Federal University of Rio de Janeiro.

MINOUX, M. (1983), "Programmation Mathématique: Théorie et Algorithmes," Dunod, Paris.

ORON, G. (1979), "An Algorithm for Optimizing Nonlinear Constrained Zero-One Problems to Improve Wastewater Treatment," Engineering Optimization (4) 109-115.

PETERSON, D.E. e LAUGHUNN, D. J. (1971), "Capital

- Expenditure Programming and Some Alternative Approaches to Risk," *Management Science* 17 320-336.
- PICARD, J.C. e RATLIFF, H.D. (1975), "Minimum Cuts and Related Problems," *Networks* 5, 357-370.
- POLYAK, B. (1969), "Minimization of nonsmooth Functionals," *USSR Comp. Math. and Math. Phys.* 9, pp. 14-29.
- RAGADE, R.K., HIPEL, K.W., UNY, T.E. (1976), "Metarationality in Benefit-Cost Analysis," *Water Resources Research* 12(15) 1069-1076.
- REID, J. (1975), "A sparsity Exploiting Variant of the Bartels-Golub Decomposition for Linear Programming Bases," Harwell Report.
- REID, J. (1976), "Fortran Subroutines for Handling Sparse Linear Programming Bases," Harwell Report, HL76/111 (C.13).
- RHYS, J.M. W. (1970), "A selection Problem of Shared Fixed Costs and Network Flows," *Management Science* 17(3), 200-207.
- RIBEIRO, C. (1983), "Algorithmes de Recherche de Plus Courts Chemins Contraintes: Etude Théorique, Implementation et Parallelisation," Doctoral Dissertation, Paris (1983).

ROSENBERG, I. (1972), "0-1 Optimization and Nonlinear Programming," *Revue Française d'Automatique, Informatique et Recherche Opérationnelle* 6, 95-97.

ROSENBERG, I. (1973), "A Variant of Pseudo-boolean Programming," *Revue Roumaine de Mathématiques Pures et Appliquées XVIII* (15), 721-725.

ROSENBERG, I. (1975), "Reduction of Bivalent Maximization to the Quadratic Case," *Cahiers du Centre d'Études de Recherche Opérationnelle* 17, 71-79.

SAIPE, A.L. (1975), "Solving a 0-1 Hyperbolic Program by Branch and Bound," *Naval Research Logistic Quarterly*, 22, 497-516.

SCHAIBLE, S. (1981), "A Survey of Fractional Programming," in S. Schaible and W. T. Ziemba, eds., *Generalized Concavity in Optimization and Economics*, Academic Press, New York, 417-440.

SCHECHTER, M. e HAMMER, P.L. (1970), "A Note on The Dynamic Planning of Investment Projects," *European Economic Review* 2(1) 111-121.

SCHOH, M. e LYSKA, W. (1978), "Kombinatorische Algorithmen zur Lösung Spezieller Nichtlinearen 0-1 Optimierungsaufgaben," *Math. Operationsforsch. Statist., Ser. Optimization* 9(1), 9-20.

- SEPPALA, Y. (1967), "Choosing Among Investment Possibilities With Stochastic Payoff Minus Expenditures," *Op. Res.* 15 978-979.
- SHEPARDSON, F. e MARSTEN, R. (1980), "A Lagrangean Relaxation Algorithm for the Two-duty Period Scheduling Problem," *Management Science* 26, 274-281.
- SOUZA, M. (1989), "Programação Fracionária Linear em Variáveis Bivalentes: Uma Contribuição ao Estudo de Algoritmos," Tese de Mestrado, Programa de Engenharia de Sistemas e Computação, COPPE/UFRJ, Rio de Janeiro.
- STECKE, K.E. (1983), "Formulation and Solution of Nonlinear Integer Production Planning Problems for Flexible Manufacturing Systems," *Management Sci.* 29(3) 273-288.
- VIZVARI, B. (1975), "Enumeration Algorithms in polynomial 0-1 Programming," *Alkamazott Matematikai Lapok* 1, 373-384.
- WALIKIEWICZ, S., SLOMINSKI, L. e FANER, M., "An improved Algorithm for Pseudo-Boolean Programming," in R. Conti and A. Ruberti, eds., fifth IFIP. Conference on Optimization Techniques, Part I (Springer, Heidelberg) pp. 493-504.
- WANG, K. (1988), "An Algorithm for Non-Linear 0-1 Programming and its Application in Structural Optimization," *J. Num. Method & Comp. Appl.* 9(1), 22-31

(Chinese J. Num. Math. & Appl. 10(2), (1989) 29-39).

WATTERS, L.J., (1967), "Reduction of Integer Polynomial Programming to Zero-one Linear Programming Problems," Op. Res. 15, 1171-1174.

WITZGALL, C. (1975), "Mathematical Methods of Site Selection for Electronic Messages (EMS)," NBS Internal Report.

ZAK, Y. A. (1978), "Algorithms for Nonlinear Pseudo-Boolean Programming," Engineering Cybernetics 16, 29-40.

ZANGWILL, W. (1965), "Media Selection by Decision Programming", Journal of Advertising Research 5, 30-36.