UMA NOVA CLASSE DE FUNÇÕES PENALIDADE E SUA APLICAÇÃO A PROBLEMAS DISCRETOS

Luiz Antonio Nogueira Lorena

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

Aprovada por:

Prof. Paulo Roberto Oliveira
(Presidente)

Prof. Nelson Maculan Filho

Prof. Carlos Humes Jr.

Prof. Demétrio Alonso Ribeiro

Prof. Horacio Hideki Yanasse

RIO DE JANEIRO, RJ - BRASIL
JUNHO DE 1985

LORENA, LUIZ ANTONIO NOGUEIRA

Uma Nova Classe de Funções Penalidade e sua Aplicação a Problemas Discretos. |Rio de Janeiro| 1985

VII, 56. 29,7cm (COPPE/UFRJ, D.Sc., Engenharia de Sistemas e Computação, 1985)

Tese - Univ. Fed. Rio de Janeiro. Fac. Engenharia

1. Programação Matemática I. COPPE/UFRJ II. Título (série).

A Josélia

Ana Carolina e

Luiz Henrique

Agradecimentos

Meu especial agradecimento ao meu amigo e orientador Paulo Roberto Oliveira que tanto contribuiu para a realização deste trabalho.

Ao Prof. Nelson Maculan Filho que me abriu os caminhos para o doutorado na COPPE.

Aos Doutores Horácio Hideki Yanasse e Carlos Hummes Júnior pelas valiosas sugestões para os resultados da Seção 1 do Capítulo 2.

Ao Instituto de Pesquisas Espaciais (INPE) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro e a Universidade Federal do Rio de Janeiro pela oportunidade para realização do doutorado.

Aos professores e colegas do Programa de Engenharia de Sistemas e Computação.

Resumo da Tese Apresentada à COPPE/UFRJ como parte dos requisitos necessários a obtenção do grau de Doutor em Ciências (D.Sc.)

UMA NOVA CLASSE DE FUNÇÕES PENALIDADE E SUA APLICAÇÃO A PROBLEMAS DISCRETOS

Luiz Antonio Nogueira Lorena

Junho de 1985

Orientador: Paulo Roberto de Oliveira
Programa: Engenharia de Sistemas e Computação

RESUMO

Inicialmente faz-se uma generalização dos conceitos de funções côncavas, conjugadas e de superdiferenciais e supergradientes de uma função de valor real definida em R^m. Esses conceitos são usados para mostrar uma teoria geral de Dualidade/Penalidade para um problema (P) de maximização sujeito a restrições de desigualdades. Define-se um problema dual (D) e um problema com penalidade (PP) e estabelecem-se condições para a equivalência entre os problemas.

Procura-se então uma classe de funções H que satisfaça as condições de equivalência dos problemas (P) e (PP). Sugere-se a classe $H_\alpha = \{h_\alpha(d) = \sum_{i=1}^{m} [\lambda_i(d_i - b_i) + \beta_i |d_i - b_i|^\alpha]$ para $\lambda_i, \beta_i \in R_+$; $\alpha > 0$; $d_i, b_i \in R\}$.

Estes resultados são aplicados a problemas discretos, especialmente ao problema da mochila 0-1, onde foi justificado uma série de heurísticas comumente usadas para essa classe de problemas.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

A NEW CLASS OF PENALTY FUNCTIONS AND THEIR APPLICATION TO DISCRETE PROBLEMS

Luiz Antonio Nogueira Lorena

Junho, 1985

Chairman: Paulo Roberto de Oliveira
Department: Sistems Engineering and Computer Science

ABSTRACT

A generalization of concave and conjugate functions, superdifferentials and supergradients of a real function defined on R^m is initially presented. These concepts are used to show a general Duality/Penalty theory for a maximization problem with inequality constraints (P). A dual problem (D) and a penalty problem (PP) are introduced and conditions for the equivalence between the problems are given.

Next, a search is made for a class of functions H which satisfies the equivalence conditions for the problems (P) and (PP). The following class of functions is suggested: $H_\alpha = \{h_\alpha(d) = \sum_{i=1}^{m} [\lambda_i \ (d_i - b_i) + \beta_i |d_i - b_i|^\alpha]$, for λ_i, $\beta_i \varepsilon R_+$; $\alpha > 0$; d_i, $b_i \varepsilon R\}$.

These results are applied to discrete problems, specially to 0-1 Knapsack problems, with the justification of some heuristics applied this class of problems.

ÍNDICE

CAPÍTULO I : INTRODUÇÃO

1. INTRODUÇÃO

O objetivo deste trabalho é apresentar uma nova classe de funções penalidade exatas, aplicadas a Problemas de Programação Matemática, em especial a Problemas de Programação Inteira zero-um.

No Capítulo II faz-se uma generalização dos conceitos de funções côncavas e de funções conjugadas, superdiferencial e supergradiente de uma função de valor real definida em R^m. Esses conceitos são usados para construir uma teoria geral de dualidade/penalidade para o seguinte problema geral de maximização, sujeito a restrições de desigualdades

$$(P) \qquad v = \sup_{x \in X} f(x) \quad \text{suj. a} \quad g(x) \leq b \qquad (1)$$

onde $X \subseteq R^n$, $f: R^n \dashrightarrow R$, $g: R^n \dashrightarrow R^m$ e $b \in R^m$.

Define-se o "problema dual" de (P)

$$(D) \qquad w = \inf_{h \in H \subseteq H_+^m} \sup_{x \in X} L(x,h) \qquad (2)$$

e o "problema com penalidades"

$$\text{(PP)} \qquad P_h = \sup_{x \in X} L(x,h), \; h \in H \subseteq H_+^m. \tag{3}$$

onde $L(x,h)=f(x)-h(g(x)+h(b))$ é uma generalização do conceito de função Lagrangeana, H_+^m é o conjunto de funções reais não-decrescentes definidas em R^m, com valor finito, e H é um subconjunto pré-determinado, não-vazio de H_+^m.

A dualidade forte (v=w) entre (P) e (D) e a equivalência ($v=p_h^*$ e mesmas soluções ótimas) entre (P) e (PP) para algum h*, estarão ligadas respectivamente aos conceitos generalizados de funções côncavas e superdiferenciais no "ponto" b (vetor das restrições de (P)).

No Capítulo III dedica-se a caracterizar um conjunto de funções $H \subseteq H_+^m$ usado na definição dos problemas (D) e (PP) que garanta a dualidade forte ou a equivalência entre os problemas acima mencionados. É dado maior ênfase a equivalência entre (P) e (PP), sugerindo-se a nova classe de funções penalidades

$$H_\alpha = \{h_\alpha : R^m \longrightarrow R : h_\alpha(d) = \sum_{i=1}^{m} [\lambda_i(d_i-b_i) + \beta_i |d_i - b_i|^\alpha, \tag{4}$$

para $\lambda_i,\ \beta_i \varepsilon R_+;\ \alpha>0;\ b_i \varepsilon R\}$, e derivam-se condições sobre os vetores λ e β para a equivalência entre os problemas.

Os resultados obtidos são então aplicados a Problemas Discretos, especialmente ao Problema de Programação Linear Inteira zero-um. Para esses problemas sugere-se usar as funções da classe $H_1(\alpha=1)$ de forma aproximada, tentando obter suas soluções ótimas através do "problema com penalidade", embora sem obter a igualdade entre seus valores ótimos.

Usando essa aproximação para problemas multidimensionais do tipo "knapsack" zero-um (veja Salkin, [38]), uma série de heurísticas comunente usadas para essa classe de problemas foram colocadas dentro de uma classe geral, sendo portanto justificadas teoricamente. Baseando-se ainda no mesmo princípio propõe-se uma heurística para o problema linear inteiro com variáveis bivalentes tipo zero-um.

No Capítulo IV conclui-se situando o trabalho entre outros já existentes e especificando a contribuição feita. Algumas perspectivas para trabalhos futuros são também apresentadas.

2. REVISÃO DA LITERATURA

Nesta seção procura-se situar o contexto do trabalho revisando, de modo breve, os principais resultados de trabalhos relacionados na literatura.

A teoria de dualidade para Problemas de Programação Não-Linear apresentou grande desenvolvimento nas décadas de 60 e 70. Geoffrion [13] unificou uma série de trabalhos anteriores (Dorn, [06]; Wolfe, [44]; Stoer, [41]; Mangasarian e Ponstein, [31]; Rockafellar, [36]; Fenchel,[09, 10], em seu excelente trabalho sobre dualidade em Programação Convexa, abordando conceitos como o de "estabilidade" para um problema convexo e "condições de otimalidade" para os problemas "primal" e "dual". O problema "dual" foi introduzido através de uma formulação "minimax" e uma interessante abordagem gráfica foi utilizada.

Anteriormente Everett [08] estudou a aplicação do Lagrangeano a problemas gerais de Programação Matemática, obtendo resultados importantes, de uso ainda atual (veja Magazine e Oguz, [30]). Para problemas não-convexos, Everett verificou a possível existência de "gaps" (saltos) de dualidade, isto é, regiões inacessíveis através da otimização do dual.

Gould [16] propõe uma formulação diferente do "dual" de forma que o "gap" seja evitado. Entretanto seu "dual" não se mostrou computacionalmente viável.

O estudo da teoria de dualidade para Programação Inteira teve grande impulso com o artigo de Geoffrion [14] que criou o termo "Relaxação Lagrangeana". Seu resultado principal foi destacar a possibilidade de uso do dual para obtenção de bons limitantes para Problemas de Programação Inteira.

Talvez os primeiros a motivar o uso do dual em Programação Inteira foram Held e Karp [18, 19] para uma abordagem (uso de árvore geradora mínima) do Problema do Caixeiro Viajante. Fisher [11] realizou uma resenha, mostrando os bons resultados de aplicações da Relaxação Lagrangeana em problemas de Combinatória na década de 70.

Baseando-se em um trabalho sobre "super-aditividade" (funções super-aditivas) de Gomory e Johnson [15] e Johnson [23] (veja também Araoz, [01]) foi proposta uma nova teoria de dualidade para Programação Inteira, originada com Johnson [22], e desenvolvida por Burdet e Johnson [03] e Jeroslow [20, 21]. O principal resultado dessa teoria foi a identificação de que para uma classe de funções super-aditivas, quando são usadas na definição do Lagrangeano, a solução ótima do primal é encontrada quando se otimiza o dual, isto é, evita-se o "gap" de dualidade. Infelizmente as funções super-aditivas convenientes não são óbvias de serem usadas, embora Wolsey [45] tenha mostrado que os principais algoritmos usados em Programação Inteira (uso da Relaxação Lagrangeana, da Teoria de Grupos, Planos de Corte e o "Branch-and-Bound"), produzem caracterizações diferentes da função super-aditiva ótima. Ou seja, quando é aplicado um dos algoritmos para Programação Inteira, indiretamente está-se resolvendo o dual.

No contexto de Programação Não-Linear, Rockafellar [37] propõe uma forma computacional de Lagrangeano Aumentado" que evita o "gap" de dualidade. Dolecki e Kurcyusz [05] usam conceitos generalizados de funções conve xas, conjugadas e subdiferenciais para construir uma teoria geral de dualidade/penalidade em espaços vetoriais. Tind e Wolsey [42] trazem alguns de seus resultados para Problemas de Programação Matemática (no espaço R^n). Em ambos os trabalhos nada computacional foi ao menos sugerido.

Neste trabalho desenvolvem-se os principais resultados dos trabalhos de Dolecki e Kurcyusz e Tind e Wolsey, objetivando identificar classes de funções que garantam respectivamente a dualidade forte e/ou a equivalência entre os problemas (P) e (D) e (P) e (PP) definidos anteriormente. Deseja-se ainda que os problemas (D) e (PP) sejam computacionalmente viáveis. É dado ênfase ao Problema de Programação Linear Inteira zero-um (em especial ao Problema tipo "knapsack" zero-um), que se mostrou de mais fácil aplicação.

CAPÍTULO II : DUALIDADE/PENALIDADE EM PROGRAMAÇÃO MATEMÁTICA

1. GENERALIZAÇÃO DE FUNÇÕES CÔNCAVAS, CONJUGADAS E DE SUPERDIFERENCIAIS

Nesta seção são definidos os conceitos generalizados de funções côncavas, conjugadas e de superdiferenciais, que serão usados posteriormente para a construção da teoria de dualidade/penalidade.

Considera-se um conjunto H de funções reais finitas definidas em R^m, e que possui a propriedade de translação, isto é, se $h \varepsilon H$ então $h' \varepsilon H$, onde $h'(d) = h(d) + r$, $\forall\, d \varepsilon R^m$, $r \varepsilon R$. De um modo abreviado, escreve-se $H + r \subseteq H$.

Em relação a uma dada função $f: R^m \dashrightarrow \bar{R}$ ($\bar{R} = R \cup +\infty$), pode-se considerar o seguinte subconjunto de H,

$$H^f = \{h \varepsilon H:\ h(d) >= f(d),\ \forall\, d \varepsilon R^m\}.$$

Definição 1: H^f é o conjunto das funções de H que são majorantes da função f.

É imediato que se $f(d) = +\infty$, para algum $d \varepsilon R^m$, então H^f é vazio.

O conjunto H^f pode ser usado para generalizar o conceito de concavidade.

Definição 2: Uma função $f: R^m \dashrightarrow \overline{R}$ é chamada H-côncava em d_0 se

$$f(d_0) = \inf_{h \in H^f} h(d_0).$$

Ou, de forma equivalente, f é H-côncava em d_0 se para cada par (d_0, r), $r > f(d_0)$, existir $h \in H^f$ tal que $h(d_0) < r$.

Definição 3: Se $f: R^m \dashrightarrow \overline{R}$ é H-côncava em d, $\forall\, d \in R^m$, diz-se que f é H- côncava.

Considerando-se por exemplo que H é o conjunto das funções afins de R^m, uma função f será H-côncava se for côncava própria ou a função $-\infty$, ou a função $+\infty$.

Pode-se ainda relacionar $h \in H$ e f através do conceito generalizado de função conjugada de f.

Definição 4: Para uma função $f: R^m \dashrightarrow \overline{R}$, define-se a função H-conjugada de f, $f^*: H \dashrightarrow \overline{R}$, por

$$f^*(h) = \inf_{d \in R^m} \{h(d) - f(d)\}.$$

Em uma interpretação geométrica $f^*(h)$ será a "distância vertical" entre h e f (observe que $f^*(h)$ poderá ser negativo).

O lema seguinte apresenta algumas propriedades importantes de f^*.

Lema 1: (i) $f^*(h') = f^*(h) + r$,

onde $h'(d) = h(d) + r$, $\forall\, d \in R^m$ e algum $r \in R$;

(ii) $f^*(h) \geqq 0$ se e somente se $h \in H^f$;

(iii) $f(d) + f^*(h) \leqq h(d)$.

Prova: imediata a partir das definições de f^* e H^f.

Pode-se ainda construir a função H-conjugada segunda de f.

Definição 5: Para uma função $f: R^m \dashrightarrow \bar{R}$, define-se sua função H-conjugada segunda, $f^{**}: R^m \dashrightarrow \bar{R}$, por

$$f^{**}(d) = \inf_{h \in H} \{h(d) - f^*(h)\}.$$

Investigar a relação existente entre a função f e sua H-conjugada segunda f^{**} será muito útil ao desenvolvimento da teoria de dualidade/penalidade. Da propriedade (iii) do Lema 1, conclui-se que $f(d) \leqq f^{**}(d)$, $\forall\, d \in R^m$. No Lema e proposição que seguem, caracteriza-se uma função H-côncava em um ponto d, através do conceito de função H-conjugada segunda.

Lema 2: Para uma função $f:R^m \dashrightarrow \overline{R}$,

$$f^{**}(d) = \inf_{h \varepsilon H^f} h(d), \forall\, d \varepsilon R^m.$$

Prova: Tem-se da definição 5 que

$$f^{**}(d) = \inf_{h \varepsilon H} \{h(d) - f^*(h)\}.$$

Ou, de forma equivalente

$$f^{**}(d) = \inf_{h \varepsilon H} \{h(d) + r - f^*(h) - r\},$$

$\forall\, r \varepsilon R$. Na operação do infimo, escolhe-se para cada h, $r_h = -f^*(h)$. Dada a propriedade de translação tem-se $0 = f^*(h) + r_h = f^*(h')$, onde $h'(d) = h(d) + r$. Logo, pelo Lema 1, (ii),

$$f^{**}(d) = \inf_{h' \varepsilon H^f} h'(d)$$

Proposição 1: Uma função $f:R^m \dashrightarrow \overline{R}$ é H-côncava em d_0, se e somente se, $f(d_0) = f^{**}(d_0)$.

Prova: imediata a partir do Lema 2 e da definição 5.

Para o conjunto H selecionou-se o subconjunto H^f das funções $h \varepsilon H$ majorantes de f. Estabelece-se agora um outro subconjunto de H, que generaliza a noção de supergradiente de f em um ponto de R^m.

Definição 6: Para uma função $f: R^m \dashrightarrow \overline{R}$ o conjunto de funções $h \varepsilon H$ que satisfaz a desigualdade $h(d) - h(d_0) \geqq f(d) - f(d_0)$, $\forall\ d \varepsilon R^m$ é chamado H-superdiferencial de f em d_0.

Para o conjunto H-superdiferencial de f em d_0 será usada a notação $\partial_H^{d_0}(f)$. Diz-se que f é H-superdiferenciável em d_0 se $\partial_H^{d_0}(f) \neq \emptyset$.

Existe uma relação clara entre os conceitos de H-superdiferenciável e H-conjugada.

Proposição 2: As seguintes afirmações são equivalentes:

(i) $h \varepsilon \partial_H^{d_0}(f)$;

(ii) $h(d_0) - f(d_0) = \inf\limits_{d \varepsilon R^m} \{h(d) - f(d)\}$;

(iii) $h(d_0) = f(d_0) + f^*(h)$.

Prova: imediata a partir das definições 4 e 6.

Definição 7: Se $h \varepsilon \partial_H^{d_0}(f)$ e $h(d_0) = f(d_0)$, será uma H-supergradiente de f em d_0.

Como $H+r \subseteq H$, temos da definição 6 que $\partial_H^{d_0}(f)+r \subseteq \partial_H^{d_0}(f)$ e assim se $\partial_H^{d_0}(f) \neq \emptyset$ este admitirá ao menos uma H-supergradiente de f em d_0.

O Lema seguinte mostra que uma H-supergradiente de f em d_0 é majorante de f.

Lema 3: Se h é uma H-supergradiente de f em d_0 então $h \in H^f$.

Prova: imediata a partir das definições 1, 6 e 7.

A seguinte proposição estabelece a relação existente entre os conceitos de função H-superdiferenciável e função H-côncava em um ponto de R^m.

Proposição 3: Se f é H-superdiferenciável em d_0, então é H -côncava em d_0.

Prova: Como f é H-superdiferenciável em d_0, existirá ao menos uma função H-supergradiente de f em d_0. Seja $\bar{h}$ essa função, isto é, $\bar{h} \in H^f$ (pelo Lema 3) e $\bar{h}(d_0)=f(d_0)$, ou $f(d_0)=\inf\limits_{h \in H^f} h(d_0)$.

2. TÓPICOS EM TEORIA DE DUALIDADE/PENALIDADE

Nessa seção considera-se o problema geral (P) e deri am-se os problemas (D) e (PP) (apresentados no Capítulo 1) usando os conceitos generalizados de função côncava, fun ção conjugada e de superdiferencial. O objetivo principal da seção será mostrar resultados de dualidade forte para (P) e (D) e também um resultado de equivalência entre (P) e (PP).

Seja $H_+^m(\overline{H}_+^m)$ o conjunto de funções $h:R^m \dashrightarrow R(\overline{R})$ que são não-decrescentes, isto é, $h(d^1) \geqq h(d^2)$, $\forall\, d^1, d^2 \in R^m$ tais que $d_i^1 \geqq d_i^2$, $i= 1,\ldots, m$.

Uma função importante de $\overline{H}_+^m$ é a conhecida função perturbação de (P), definida por

$$\emptyset(d)= \sup_{x \in X} \{f(x):g(x) \leqq d\}.$$

É importante que $\emptyset(b)= v$ (valor ótimo do problema (P)).

Observa-se que H_+^m possui a propriedade de trans isto é, $H_+^m + r \subseteq H_+^m$. Considerando que o conjunto H, definido na seção 1 desse Capítulo, poderá ser um subconjunto de H_+^m, e que no desenvolvimento que segue deseja-se efetivamente que isso ocorra, usa-se a mesma notação para um subconjunto de H_+^m, isto é, $H \subseteq H_+^m$. O lema a seguir caracteriza subconjuntos de H_+^m.

Lema: $H \subseteq H_{+}^{m}$ se e só se

$$\forall\, h \in H,\ \forall\, d_0 \in R^{m},\ h(d_0) = \inf_{\substack{d \in R^m \\ d_0 \leqq d}} h(d)$$

Prova: imediata

Relacionando o resultado do Lema 4 com o problema (P) pode-se considerar os $x \in X$ tais que $g(x) = d_0$, caracterizando um subconjunto H de H_{+}^{m} em relação a (P).

Na proposição seguinte mostra-se que a conjugada da função perturbação está fortemente relacionada com as funções f e g usadas na definição do problema (P), considerando-se as funções $h \in H \subseteq H_{+}^{m}$.

Proposição 4: $\forall\, h \in H \subseteq H_{+}^{m}$,

$$-\emptyset^*(h) = \sup_{x \in X} \{f(x) - h(g(x))\}.$$

Prova: $$\sup_{x \in X} \{f(x) - h(g(x))\} = \sup_{x \in X} \{f(x) - \inf_{\substack{d \\ g(x) \leqq d}} h(d)\} =$$

$$\sup_{d} \{\sup_{\substack{x \\ g(x) \leq d}} f(x) - h(d)\} = \sup_{d} \{\emptyset(d) - h(d)\} = -\emptyset^*(h).$$

Corolário 1: $H^{\emptyset}$ é não-vazio se e só se

$$f(x) \leqq h(g(x)),\ \forall\ x \in X \text{ e algum } h\ H \subseteq H_+^m.$$

Pode-se definir agora a seguinte função:

$$L(x,h) = f(x) - h(g(x)) + h(b),\ \forall\ x \in X,\ h \in H \subseteq H_+^m,$$

$l(x,h)$ é a função Lagrangeana para $x \in X$ e $h \in H \subseteq H_+^m$.

Por exemplo, considere-se que H é o conjunto das funções lineares não-decrescentes, isto é $h(d) = \sum_{u=1}^{m} u_i d_i$, $\forall\ d \in R^m$ e algum $u \in R_+^m$, então $L(x,h) = f(x) - \sum_{i=1}^{m} u_i (g_i(x) - b_i)$, que é a formulação tradicional, onde os u'is são os conhecidos multiplicadores de Kuhn-Tucker. Em geral para $g(x) \leqq b$, $x \in X$ (isto é, x viável para (P)), $L(x,h) \geqq f(x)$. Assim $\sup_{x \in X} L(x,h) \geqq f(x)$ e $\sup_{x \in X} L(x,h) \geqq v$, isto é, $\sup_{x \in X} L(x,h)$, $h \in H$ é um limitante superior para o valor ótimo do problema (P). Pode-se procurar o menor desses limitantes através do problema.

(D) $$w = \inf_{h \in H} \sup_{x \in X} L(x,h).$$

(D) é o problema dual de (P). (P) é chamado problema primal. Diz-se que (D) é viável se existir $h \in H$ tal que $w \neq +\infty$.

A viabilidade de (D) pode ser caracterizada ainda através do seguinte lema.

Lema 5: As seguintes afirmações são equivalentes

(i) (D) é viável

(ii) $h(g(x)) \geqq f(x)$, $\forall$ $x \varepsilon X$ e algum $h \varepsilon H \subseteq H_+^m$,

(iii) $H^{\emptyset}$ é não-vazio.

Prova: imediata a partir da definição de (D) e do corolário 1.

Os comentários feitos para a construção do problema (D) e o Lema 5 mostram a seguinte proposição.

Proposição 5: (Dualidade fraca)

Se x é viável para (P) e h é viável para (D) então $v \leq w$.

A procura de condições para a dualidade forte ($v = w$) nos levará a consideração de que $\emptyset$ deve ser H-côncava em b, como segue.

Lema 6: $\emptyset^{**}(b) = w$.

Prova: $\emptyset^{**}(b) = \inf_{h \varepsilon H}\{h(b) - \emptyset^{*}(h)\}$, mas, pela proposição 4,

$$-\emptyset^*(h) = \sup_{x \in X} \{f(x) - h(g(x))\},$$

assim,

$$\emptyset^{**}(b) = \inf_{h \in H} \{h(b) + \sup_{x \in X} [f(x) - h(g(x))]\} =$$

$$= \inf_{h \in H} \sup_{x \in X} \{f(x) - h(g(x)) + h(b)\} =$$

$$= \inf_{h \in H} \sup_{x \in X} L(x,h) = w.$$

Proposição 6: (dualidade forte)

$v = w$ se e só se $\emptyset$ é H-côncava em b.

Prova: Supondo $v = w = \emptyset(b)$, pelo Lema 6, $w = \emptyset^{**}(b)$. Assim $\emptyset^{**}(b)$ e $\emptyset$ é H-côncava em b.

Supondo agora que $\emptyset$ é H-côncava em b, isto é, $\emptyset^{**}(b) = \emptyset(b) = v$.

Mas, pelo Lema 6, $w = \emptyset^{**}(b)$. Assim $v = w$.

Usando a relação existente entre os conceitos de H-superdiferencial e H-côncava (proposição 3), pode-se mostrar o seguinte corolário.

Corolário 2: Se na definição de $L(x,h)$, $h \in \partial_H^b(\emptyset)$ então $v = w$.

Prova: imediata a partir da proposição 3.

Mas um resultado mais forte pode ser obtido quando $\emptyset$ é H-superdiferenciável.

Proposição 7: (equivalência) $\sup\limits_{x \in X} L(x,h) = \emptyset(b)$ se e só se $h \in \partial_H^b(\emptyset)$.

Prova: $\sup\limits_{x \in X} L(x,h) = \sup\limits_{x \in X} \{f(x) - h(g(x))\} + h(b) =$

$= -\emptyset(h) + h(b)$. Mas $\emptyset(b) = -\emptyset^*(h) + h(b)$ equivale, conforme a proposição 2, a $h \in \partial_H^b(\emptyset)$.

Pode-se definir então o problema

$$\text{(PP)} \qquad p_h = \sup L(x,h), \; h \in H \subseteq H_+^m$$

Ficou evidenciada a importância de ser encontrada uma classe de funções $H \subseteq H_+^m$, para o qual $\emptyset$ é H-côncava ou

H-superdiferenciável em b, assegurando respectivamente a dualidade forte entre (P) e (D) ou a equivalência entre (P) e (PP). Na próxima seção procura-se então estabelecer essa classe de funções H.

3. DUALIDADE FORTE ENTRE (P) E (D) E/OU EQUIVALÊNCIA ENTRE (P) E (PP)

3.1 - A FUNÇÃO PERTURBAÇÃO É H-CÔNCAVA

Os seguintes resultados são particularizações de resultados que aparecem em Dolecki e Kurcyusz [05].

Seja o conjunto de funções.

$$H_n = \{h : h(d) = r + \rho . \eta(\| d-y \|),\ r \varepsilon R,\ \rho > 0,$$

$$d \varepsilon R^m,\ e\ y \varepsilon Y\}.$$

onde $Y \subseteq R^m$ é um conjunto denso,

$\| \cdot \|$ é uma norma de R^m, e

$\eta : R_+ \dashrightarrow R_+$ é uma função não-decrescente, tal que

(i) $\forall \delta > 0,\ \forall\ t_0 \geqq 0, \exists\ M(t_0, \delta)$ (limitante), $\forall\ t \leqq \delta$,

$$\eta(t + t_0)/\eta(t) \leqq M(t_0, \delta),$$

(ii) $\eta(t) \longrightarrow 0 \iff t \longrightarrow 0$.

A condição (i) indica que funções que crescem ao infinito em uma vizinhança do ponto t_0 não são consideradas.

São exemplos de funções do tipo η: t^α, $\alpha > 0$; αt, $\alpha > 0$.

As funções de H_η podem ser usadas para generalizar os Langrangeanos Aumentados, como será visto a seguir. O seguinte teorema é um caso particular do teorema 4.2 de Dolecki e Kurcyusz [05], e estabelece condições para que a função perturbação de (P) seja H_η-côncava em b (dualidade forte entre (P) e (D) para $h \varepsilon H_\eta$).

Teorema 1: Se $\emptyset$ é semi-continua superiormente em b e $H_\eta^\emptyset$ não é vazio, então $\emptyset$ é H_η-côncava em b.

Tomando, por exemplo, $\eta(t) = t^2$ e $\|\cdot\| = \|\cdot\|_2$, se $\emptyset$ é semi-continua superiormente em b e satisfaz

$$\emptyset(d) \leqq r + \rho_0 \, \| d - y_0 \|_2^2, \quad \forall \, d \varepsilon R^m \ldots \qquad \ldots (1)$$

para algum $r \varepsilon R$, $\rho_0 > 0$, $y_0 \varepsilon Y$, então $\emptyset$ é H_η-côncava em b. A condição (1), para $\rho_0 < +\infty$, é a conhecida "condição de crescimento quadrático" para (P) (Rockafellar, [37]). Pode-se notar a liberdade de escolha para a norma e a função η.

Define-se então o conjunto H_ℓ por

$$H_\ell = \{h \varepsilon H_n : h(g(x)) = \inf_{\substack{d \varepsilon R^m \\ g(x) \leqslant d}} h(d), \; x \varepsilon X\} \qquad \ldots(2)$$

Assim $H_\ell \subseteq H_+^m$ pelo Lema 4. A condição (1) para H_ℓ mostra que (D) é viável, isto é, existe uma função $h \not\equiv +\infty$ que satisfaz (D).

A função Lagrangeana será

$$L_n(x,h) = f(x) - \inf_{\substack{d \varepsilon R^m \\ g(x) \leqq d}} h(d) + h(b)$$

Tomando novamente $\eta(t) = t^2$ e $\|\cdot\| = \|\cdot\|_2$, para $b = 0$ tem-se

$$L_n(x,b) = f(x) - \sum_{i=1}^{m} \{\rho \max{}^2[g_i(x), \; y_i/\rho] \, .$$

$$2y_i \max[g_i(x), \; y_i/\rho]\},$$

que é o Langrangeano Aumentado de Rockafellar (Rockafellar [37].

Portanto, se $\emptyset$ é semi-contínua superiormente em b e H_n-limitada existirá a dualidade forte entre (P) e (D) para $L(x,h) = L_n(x,h)$, conforme a proposição 6.

3.2 - $\emptyset$ É H-SUPERDIFERENCIÁVEL

Inicialmente supõe-se existir uma classe de funções $H \subseteq H_+^m$ tal que $H+r \subseteq H$, $r \varepsilon R$ e funções $\tilde{h} \varepsilon H$ tais que $\tilde{h} \varepsilon \partial_H^b(\emptyset)$. Mais adiante definem-se condições para existência de tal classe H, e faz-se uma sugestão para os elementos dessa classe.

Teorema 2: $\tilde{x} \varepsilon X$ é uma solução ótima de (P) e $\tilde{h} \varepsilon \partial_H^b(\emptyset)$ se e só se as seguintes afirmações são satisfeitas:

(i) $g(\tilde{x}) \leq b$,

(ii) $\sup_{x \varepsilon X} \{f(x) - \tilde{h}(g(x))\} = f(\tilde{x}) - \tilde{h}(g(\tilde{x}))$,

(iii) $\tilde{h} \varepsilon H$,

(iv) $\tilde{h}(g(\tilde{x})) = \tilde{h}(b)$.

Prova: (i) e (iii) são imediatas, prova-se a seguir (ii) e (iv):

$$\tilde{h}\varepsilon\partial_H^b(\emptyset),\ \tilde{h}(d)-\tilde{h}(d) \geqq \emptyset(d)-\emptyset(b),\ \forall\ d\varepsilon R^m$$

$$\emptyset(b)-\tilde{h}(b) \geqq \emptyset(d)-\tilde{h}(d),\ \forall\ d\varepsilon R^m$$

$$\emptyset(b)-\tilde{h}(b) = \sup_{d\varepsilon R^m} \{\emptyset(d)-\tilde{h}(d)\} =$$

$$= -\emptyset^*(\tilde{h}) = \sup_{x\varepsilon X} \{f(x)-\tilde{h}(g(x))\},\ \text{ou}$$

$$\emptyset(b)-\tilde{h}(b) \geqq f(x)-\tilde{h}(g(x)),\ \forall\ x\varepsilon X.$$

Mas $\tilde{x}$ é ótimo em (P), $f(\tilde{x}) = \emptyset(b)$,

$f(\tilde{x})-\tilde{h}(b) \geqq f(x)-\tilde{h}(g(x))$, $\forall\ x\varepsilon X$

$\tilde{h}(g(x))-\tilde{h}(b) \geqq f(x)-f(\tilde{x})$, $\forall\ x\varepsilon X$.

Então, para $x = \tilde{x}$,

$\tilde{h}(g(\tilde{x}))-\tilde{h}(b) \geqq f(\tilde{x})-f(\tilde{x}) = 0$, ou $\tilde{h}(g(\tilde{x})) \geqq \tilde{h}(b)$.

Mas $\tilde{h}$ é não-decrescente e $g(\tilde{x}) \leqq b$,

assim $\tilde{h}(g(\tilde{x})) \leqq \tilde{h}(b)$.

Portanto

$$\tilde{h}(g(\tilde{x})) = \tilde{h}(b) \text{ e } \sup_{x \in X} \{f(x) - \tilde{h}(g(x))\} = f(\tilde{x}) - \tilde{h}(g(\tilde{x})).$$

Prova de que $\tilde{x}$ é ótimo para (P):

(da condição (ii)) $f(\tilde{x}) - \tilde{h}(g(\tilde{x})) \geqq f(x) - \tilde{h}(g(x)), \ \forall \ x \in X$

$$f(\tilde{x}) \geqq f(x) + \tilde{h}(g(\tilde{x})) - \tilde{h}(g(x))$$

(da condição iv)) $f(\tilde{x}) \geqq f(x) + \tilde{h}(b) - \tilde{h}(g(x))$, ou

$$f(\tilde{x}) \geqq f(x), \ \forall \ x \in X, \ g(x) \leqq b \text{ e } \tilde{h} \in H.$$

Prova de que $\tilde{h} \in \partial_H^b(\emptyset)$:

(da condição (ii)) $-\emptyset^*(h) = \sup_{d \in R^m} \{\emptyset(d) - h(d)\} =$

$$f(\tilde{x}) - \tilde{h}(g(\tilde{x}))$$

ou $\emptyset(d) - \tilde{h}(d) \leqq f(\tilde{x}) - \tilde{h}(g(\tilde{x})), \ \forall \ d \in R^m$

$$\emptyset(d) - \tilde{h}(d) \leqq \emptyset(b) - \tilde{h}(b), \ \forall \ d \in R^m$$

Assim $\tilde{h}(d) - \tilde{h}(b) \geqq \emptyset(d) - \emptyset(b), \ \forall \ d \in R^m$ e $\tilde{h} \in \partial_H^b(\emptyset)$.

As condições (i)-(iv) são conhecidas como "condições de otimalidade" para (P), e são uma generalização das condições de Kuhn-Tucker para Problemas de Programação Não-Linear.

Corolário 3: Se $\tilde{h}\varepsilon\partial_H^b(\emptyset)$ e existe ao menos uma $\tilde{x}\varepsilon X$, $g(\tilde{x}) \leqq b$ tal que $\tilde{h}(g(\tilde{x}))= \tilde{h}(b)$, então $\tilde{x}$ é ótimo para (P).

Assim, a condição de otimalidade (ii) sempre ocorrerá quando $\tilde{h}\varepsilon\partial_H^b(\emptyset)$. A folga complementar $(\tilde{h}(g(\tilde{x}))= \tilde{h}(b))$ decorre do fato de que $H \subseteq H_+^m$.

Corolário 4: Se $\tilde{h}$ é uma H-supergradiente de $\emptyset$ em b, e existe ao menos um $\tilde{x}\varepsilon X$, $g(\tilde{x}) \leqq b$ tal que $\tilde{h}(g(\tilde{x}))= \tilde{h}(b)$, então:

(i) $f(\tilde{x})= \tilde{h}(g(\tilde{x}))= \tilde{h}(b)= \emptyset(b)$,

(ii) $\tilde{x}$ é ótimo para (P),

(iii) $\sup\limits_{x\varepsilon X} \{f(x)-\tilde{h}(g(x))\}= 0$

A Figura 1 ilustra os resultados do teorema 2 e dos corolários 3 e 4.

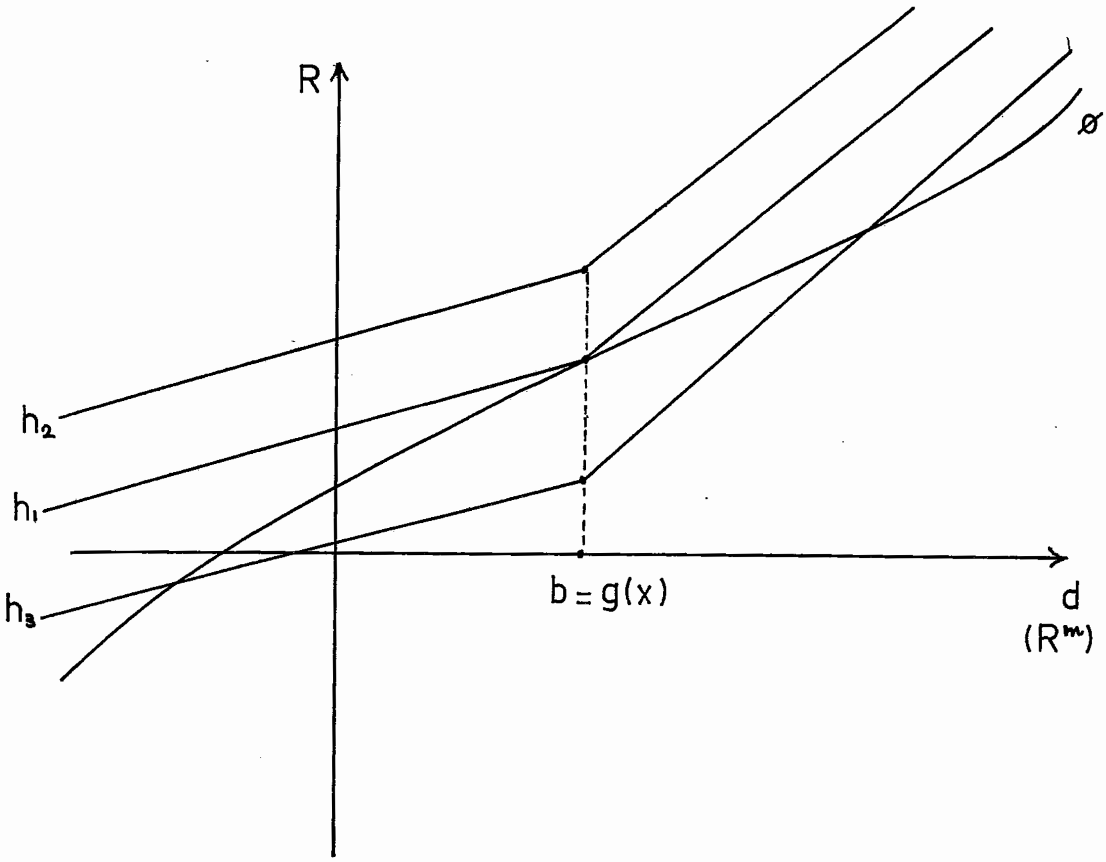

Figura 1: Relação entre $f(x)$, $h_i(g(x))$ e $h_i(b)$, $\forall$ $x \varepsilon X$ e $h_i \varepsilon \partial_H^b(\emptyset)$, i= 1,2,3.

Como pode-se notar não existe alusão ao fato de como encontrar tal $\tilde{x} \varepsilon X$, ótimo para (P). Pode-se usar o problema (PP) para esse propósito, pois, considerando-se os resultados da proposição 7 e do teorema 2 tem-se que:

se $h \varepsilon \partial_H^b(\emptyset)$ então $\sup\limits_{x \varepsilon X} L(x,h) = \emptyset(b)$, mas

$$\sup_{x\varepsilon X} L(x,h) = \sup_{x\varepsilon X} \{f(x)-h(g(x))\} + h(b)$$

$$= f(\tilde{x})-h(g(\tilde{x}))+h(b),$$

e $\tilde{x}$ é ótimo para (P) se a folga complementar é satisfeita.

Conhecida uma função $h\varepsilon\partial_H^b(\emptyset)$ pode-se obter uma função $\bar{h}\varepsilon H$ que é uma H-supergradiente de $\emptyset$ em b. Define-se

$$\bar{h}(g(x)) = \sup_{x\varepsilon X} L(x,h) + h(g(x)) - h(b),\ h\varepsilon H \qquad \dots(3)$$

Proposição 8: Se existe $h\varepsilon H$ tal que $h\varepsilon\partial_H^b(\emptyset)$ então existirá $\bar{h}\varepsilon H$ (definida em 3) que é uma H-supergradiente de $\emptyset$ em b.

Prova: Se $h\varepsilon\partial_H^b(\emptyset)$, de acordo com a proposição 7,

$$\sup_{x\varepsilon X} L(x,h) = \emptyset(b),$$

assim, de (3), $\bar{h}(b) = \sup L(x,h) = \emptyset(b)$.

Falta mostrar que $\bar{h}\varepsilon\partial_H^b(\emptyset)$, ou que

$$\bar{h}(d) - \bar{h}(b) \geqq \emptyset(d) - \emptyset(b).$$

Mas

$$\bar{h}(d) = \emptyset(b) + h(d) - h(b)$$

$$\bar{h}(d) - \emptyset(b) = h(d) - h(b)$$

$$\bar{h}(d) - \bar{h}(b) = h(d) - h(b)$$

e como $h \varepsilon \partial_H^b(\emptyset)$,

$$h(d) - h(b) \geqq \emptyset(d) - \emptyset(b),$$

portanto

$$\bar{h}(d) - \bar{h}(b) \geqq \emptyset(d) - \emptyset(b).$$

Na Figura 2 pode-se observar um caso típico de h e $\bar{h} \varepsilon \partial_H^b(\emptyset)$. Como era de se esperar $\bar{h}(b) = \sup\limits_{x \varepsilon X} L(x,h) = p_{\bar{h}}$ (valor ótimo do problema (PP)) = v (valor ótimo de (P)).

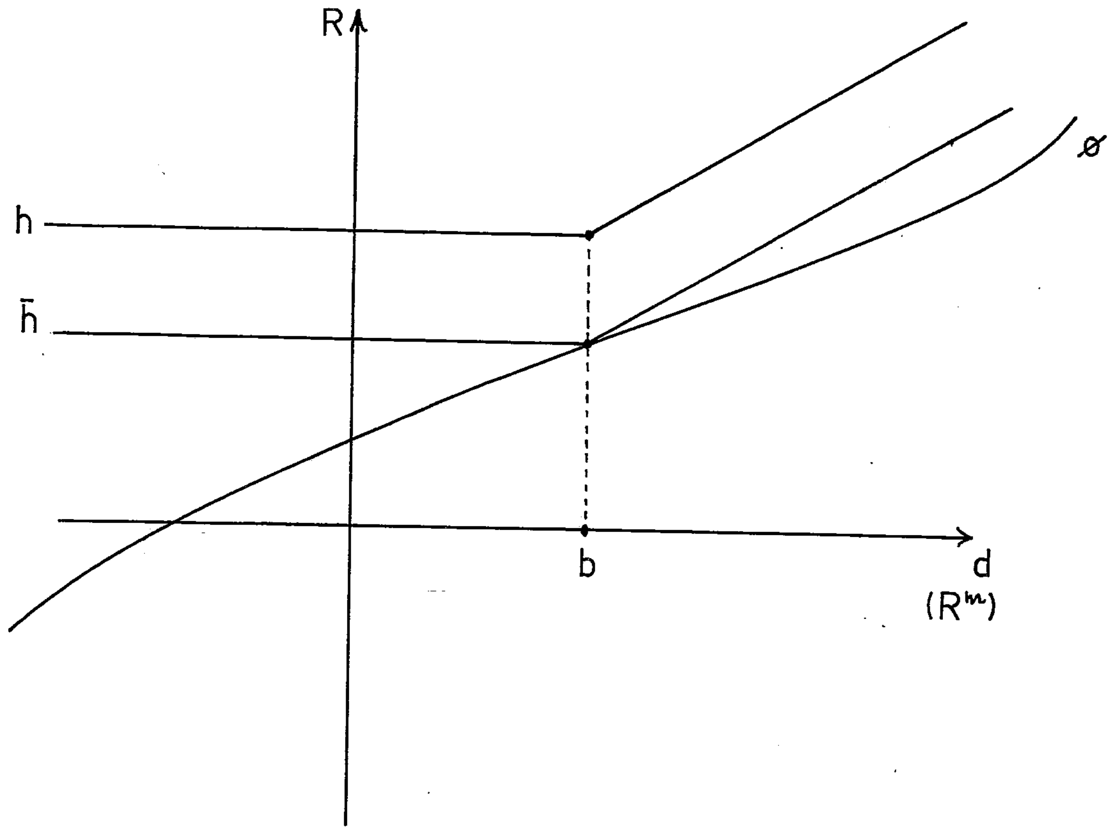

Figura 2: h e $\bar{h} \varepsilon \partial_H^b(\emptyset)$.

A seguir faz-se uma sugestão para a classe de funções H que satisfaz ao Teorema 2 e a Proposição 7. Para essa classe pode-se otimizar o problema (PP) em substituição a (P).

3.2.1 - FUNÇÕES PENALIDADES EXATAS

Define-se a seguinte classe de funções:

$$H\alpha = \{h_\alpha : R^m \dashrightarrow R : h\alpha(d) = \sum_{i=1}^{m} [\lambda_i (d_i - b_i) + \beta_i |d_i - b_i|^\alpha,$$

para λ_i, β_i ε R_+, $\alpha > 0$, $b_i \varepsilon R\}$.

Inicialmente examina-se a função $h_1 \varepsilon H_\alpha$, isto é,

$h_1(d) = \sum_{i=1}^{m} [\lambda_i (d_i - b_i) + \beta_i |d_i - b_i|]$. (veja o trabalho de Lorena e Oliveira, [27]).

Para que h_1 satisfaça as condições da Proposição 7, é necessário primeiro que ela pertença ao conjunto H_+^m, isto é deve-se ter $0 \leqq \beta_i \leqq \lambda_i$, i= 1,..., m. Resta estabelecer a condição que $h_1 \varepsilon \partial^b_{H_+^m}(\emptyset)$, isto é:

$$h_1(d) - h_1\ (b) \geqq \emptyset(d) - \emptyset(b).$$

Como $h_1(b) = 0$,

$$h_1(d) \geqq \emptyset(d) - \emptyset(b), \text{ onde por definição}$$

$$h_1(d) = \begin{cases} \sum (\lambda_i - \beta_i)\ (d_i - b_i), \text{ se } d_i \leqq b_i, \\ \\ \sum (\lambda_i + \beta_i)\ (d_i - b_i), \text{ se } d_i \geqq b_i. \end{cases}$$

Pode-se então estabelecer os parâmetros k_1, k_2; $k_1 \leqq \min_i(\lambda_i+\beta_i)$ e $k_2 \geqq \max_i(\lambda_i-\beta_i)$, $i=1,\ldots, m$, tais que para $y_i \triangleq d_i-b_i$, é fácil verificar que:

$$k_1 \geqq \sup_{y\varepsilon R^m} \left\{\sup_{y_i>0} \left[\frac{\emptyset(d) - \emptyset(b)}{y_i}\right]\right\}, \qquad \ldots(4)$$

$$k_2 \leqq \inf_{y\varepsilon R^m} \left\{\inf_{y_i<0} \left[\frac{\emptyset(d) - \emptyset(b)}{y_i}\right]\right\}. \qquad \ldots(5)$$

Observa-se diretamente que como $k_2 > 0$, um limitante inferior para o ínfimo na expressão (5) é zero, portanto $k_2= 0$ satisfaz a expressão. Assim fazendo-se $\lambda_i= \beta_i$, $i=1,\ldots,$ m em h_1, garante-se a folga complementar para o problema (P) usando h_1 (ver Corolário 3).

A existência de um valor finito para k_1 na expressão (4) está ligado a "condições de qualificação" ou "condições de estabilidade" do problema (P) (veja Geoffrion, [13], Rockafellar, [37] e Clarke, [04]).

Estas observações mostram a seguinte proposição:

Proposição 9: Se para $h_1 \varepsilon H_1$, $k_2= 0$ existir um valor finito k_1 (da expressão (4)), (P) será equivalente a (PP), com $h= h_1$.

Seja definir uma função $Q: R_+ \dashrightarrow R_+$ com as seguintes propriedades

(i) $Q(0) = 0$,

(ii) $Q(r) > 0$ para $r > 0$,

(iii) $\infty > Q'(0+) \triangleq \lim\limits_{r \to 0+} \dfrac{Q(r) - Q(0)}{r} > 0.$

Considerando $Q(h_1(d))$, $d \geqq b$, vale destacar que

$$k_1 \leqq Q'(h_1(b+)) = \lim_{d_i \to b_i +} \frac{h_1(d) - h_1(b)}{\|d-b\|}.$$

Pode-se então generalizar a proposição 9 para uma função h que satisfaça as condições:

(i) $Q(h(d)) = 0$, para $d \leqq b$,

(ii) $Q(h(d)) > 0$, para $d_i > b_i$, e

(iii) $\infty > Q'(h(b+)) > 0$.

As seguintes funções satisfazem as condições (i)-(iii) e são exemplos de generalizações possíveis:

$$g_1(d) \triangleq \sum_{i=1}^{m} \gamma_i \max (0, y_i),$$

$$q_2(d) \triangleq \left[\sum_{i=1}^{m} \gamma_i \, (\max \, (0,y_i)^q\right]^{1/q}, \text{ e}$$

$$q_3(d) \triangleq \sum_{i=1}^{m} \gamma_i \, [\exp \, (\max \, (0,y_i)) - 1].$$

Mas, por exemplo:

$$q_4(d) \triangleq \sum_{i=1}^{m} \gamma_i \, [\max \, (0,y_i)]^2$$

não satisfaz a condição (iii) (veja Han e Mangasarian [17]).

As condições (i)-(iii) mostram que a função h, que satisfaz as condições da proposição 7 para equivalência entre (P) e (PP) é não-diferenciável em b. Existem algoritmos adaptados, principalmente à função g_1, em que a não-diferenciabilidade não oferece dificuldade (veja Garcia-Palomares, [12]). Lorena [26] mostrou uma forma aproximada de uso para h_1 em problemas discretos.

Para as outras componentes da classe H_α, isto é, para $0 < \alpha < 1$ e $\alpha > 1$, embora algumas sejam diferenciáveis em b, não é possível fazer com que $h_\alpha \varepsilon H_+^m$ controlando somente os parametros λ_i e β_i, i=1,..., m. Portanto, dificilmente, para um problema geral (P) $h_\alpha \varepsilon \partial_{H_+}^{b} m(\emptyset)$.

Os resultados até a proposição 9 levam em consideração o problema geral (P). Na próxima seção considera-se um problema mais específico, de programação discreta, para o qual também será possível usar as outras funções de H_α, para $\alpha \neq 1$, de forma adaptada.

4. APLICAÇÃO A PROBLEMAS DISCRETOS

Nesta seção completam-se os resultados sobre a classe H. Considera-se o problema:

$$(PD) \qquad \begin{aligned} vd = \max\ & f(x) \\ & x \in XD \\ & g(x) \leqq b, \end{aligned}$$

onde $XD \subset R^n$ é um conjunto discreto e finito, e o problema particular de programação linear inteira:

$$(PLI) \qquad \begin{aligned} vi = \max\ & cx \\ & x \in X \\ & Ax \leqq b, \end{aligned}$$

onde $X = \{x \geqq 0,\ x \text{ é inteiro}\}$.

Mostra-se a equivalência entre (PLI) e o seguinte problema

$$(PLI) \qquad pli = \max_{x \in X} cx - \lambda(Ax-b) - \beta\,|Ax-b|_1$$

desde que (PLI) possua solução ótima com valor ótimo finito e os vetores A^j e b sejam inteiros (A^j é a coluna j da matriz Amxn).

Sabe-se que a função perturbação $\emptyset_{pd}$ de (PD) é semicontínua superior e observa-se que em geral sua solução ótima é estritamente viável (Meyer, [32]).

Define-se a função $\bar{h}_\alpha$ por:

$$\bar{h}_\alpha(g(x)) = \max_{x \varepsilon XD} L(x,h_\alpha) + h_\alpha(g(x)) - h_\alpha(b),\ h_\alpha \varepsilon H_\alpha.$$

Verifica-se facilmente que se $h_\alpha \varepsilon \partial^b_{H_+} m(\emptyset_{pd})$, então $\bar{h} \varepsilon \partial^b_{H_+} m(\emptyset_{pd})$ e $\bar{h}_\alpha(b) = \max_{x \varepsilon XD} L(x,h_\alpha) = \emptyset_{pd}(b) = vd$ (conforme a proposições 7 e 8), isto é, $\bar{h}_\alpha$ será uma H^m_+-supergradiente de $\emptyset_{pd}$ em b.

A Figura 3 ilustra $\bar{h}_\alpha$ e h_α, para α= 1,2 e 1/2 para um problema do tipo (PD) com valor ótimo finito.

Nos gráficos (a), (c) e (e) da Figura 3, as condições da Proposição 7 estão satisfeitas, o que não ocorre nos gráficos (b), (d) e (f), mas mesmo nesses casos a solução viável ótima de (PD) poderá ser encontrada controlando adequadamente os valores de λ e β.

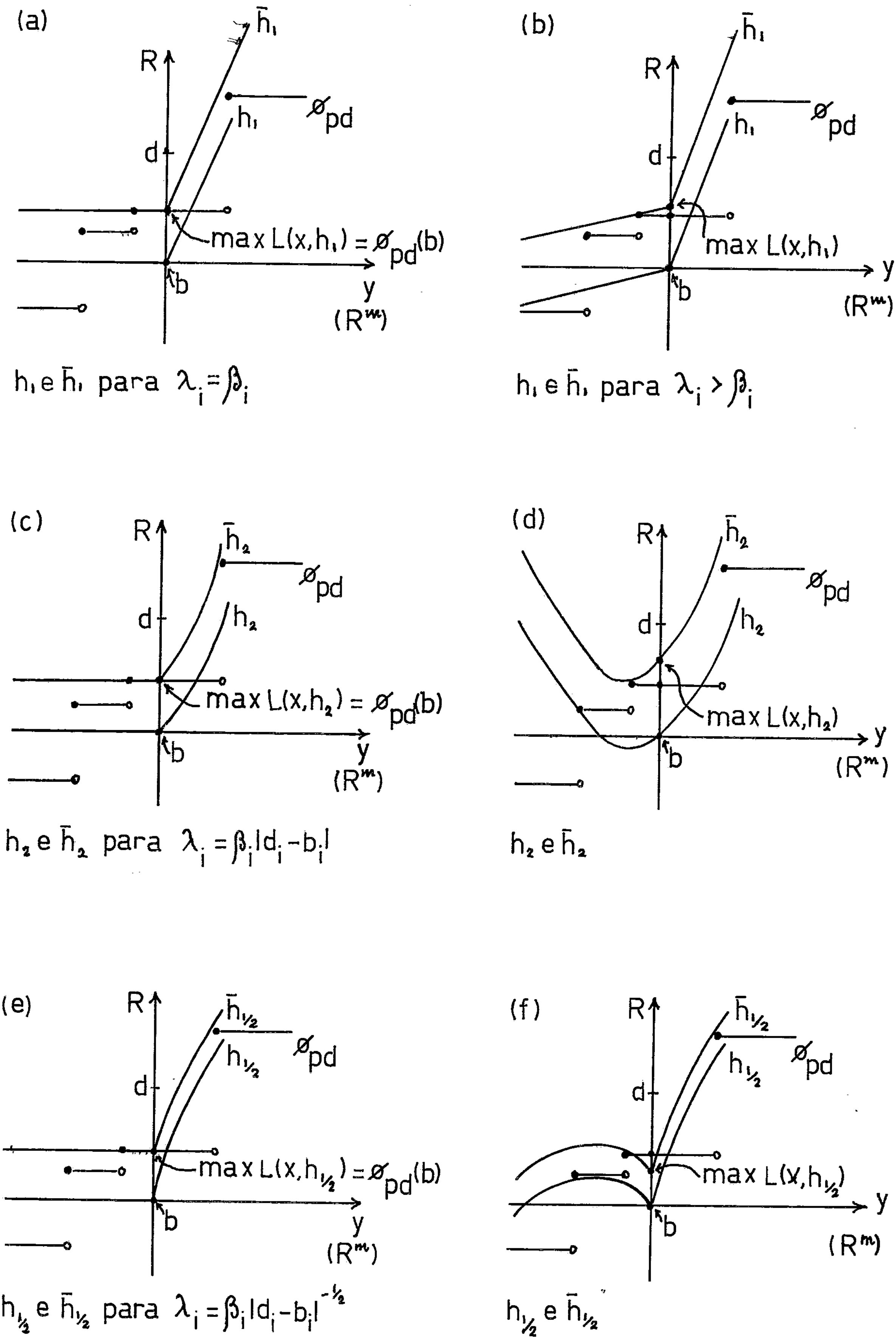

Figura 3: Representação gráfica de h_α e $\bar{h}_\alpha$, $\alpha=1, 2$ e $1/2$ para o problema (PD)

Na expressão 4 derivou-se um valor para k_1 que possivelmente fosse finito para o problema geral (P). Vê-se agora que para o problema de programação linear inteira (PLI), sob certas condições, pode-se garantir que k_1 é finito e mostrar a equivalência de (PLI) com (PPLI) usando a função $h_1 \varepsilon H_1$.

Lema 7: $\emptyset_{pli}(d) < \infty$, $\forall\ d \varepsilon R^m$, se e só se $\{\lambda \geqq 0,\ \lambda A \geqq c\} \neq \emptyset$ (isto é, o dual da relaxação de programação linear de (PLI) é viável e (PLI) possui solução ótima.

Prova: Supondo $\emptyset_{pli}(d) < \infty$, $\forall\ d \varepsilon R^m$, existirá uma função $h_1 \varepsilon H_1$ não-decrescente, com $\lambda \geqq 0$ e $\beta = 0$, tal que:

$$h_1(d) \geqq \emptyset_{pli}(d),\ \forall\ d \varepsilon R^m;\ \text{e}$$

$$h_1(d) \geqq cx,\ [\forall (d,x) : d \geqq Ax,\ x \varepsilon X],\ \text{ou}$$

$$h_1(Ax) \geqq cx,\ \forall\ x \varepsilon X;$$

$$\lambda Ax \geqq cx,\ \forall\ x \varepsilon X \text{ ou } \lambda A \geqq c.$$

Fazendo o caminho inverso, chega-se a conclusão que

$$\emptyset_{pli}(d) < \infty \text{ se } \{\lambda \geqq 0,\ \lambda A \geqq c\} \neq \emptyset.$$

Proposição 10: Se $\{\lambda \geqq 0,\ \lambda A \geqq c\} \neq \emptyset$, A^j e b forem vetores inteiros e (PLI) possui solução ótima, então (PLI) será equivalente a (PPLI) usando uma função $h_1 \varepsilon H_1$.

Prova: Se $\{\lambda \geqq 0, \lambda A \geqq c\} \neq \emptyset$, então $\emptyset_{pli}(d) < \infty$, $\forall$ d R^m. Como A e b são vetores inteiros $\emptyset_{pli}(b) =$ vi é finito (Meyer, [32]); além disso:

$$k_1 = \max_{d-b} \left\{ \max_{d_i > b_i} \left[\frac{\emptyset_{pli}(d) - \emptyset_{pli}(b)}{d_i - b_i}\right] \right\}$$

é finito, pois para $d_i = Ax_i$, d_i e b_i são inteiros e, portanto, $\min_i \{d_i - b_i\} = 1$.

Seja λ^* a solução ótima do dual da relaxação de programação linear de (PLI), com valor ótimo d, e $\emptyset_{pl}$ a função perturbação da relaxação de programação linear.

Proposição 11: Para (PLI), $k_1 \geqq \max_i \{\lambda_i^*\}$.

Prova: Basta fazer no desenvolvimento da expressão (4) $\beta_i = 0$, $\forall$ i, donde tem-se $k_1 \geqq \max_i \{\lambda_i^*\}$.

A proposição 11 mostra que λ_i^* poderá ser usado para estabelecer um limite inferior de k em problemas do tipo (PLI), embora em aplicações práticas seja recomendável uma estimativa inicial de k_1 de cálculo rápido que λ^*.

Observe ainda que pode-se usar $\lambda = \max_i \{\lambda_i\}$ e $\beta = \max_i \{\beta_i\}$ em H_α sem alterar os resultados obtidos até aqui. Portanto, considera-se no próximo capítulo $\lambda = [\lambda, \lambda, \ldots, \lambda]$ e $\beta = [\beta, \beta, \ldots, \beta]$.

CAPÍTULO III: IMPLEMENTAÇÃO

1. INTRODUÇÃO

Torna-se evidente que para a implementação de um algoritmo que use a função h_1, um passo essencial é a obtenção de uma estimativa para k_1 (da expressão (4)), pois esta é a garantia de que na otimização de (PP) são evitadas soluções inviáveis de (P). Nesta expressão aparece a função perturbação de (P), que é desconhecida, e, em geral, mais difícil de obter que a própria resolução de (P). Assim, dependendo de (P) pode-se usar a seguinte estimativa de k_1:

$$k_{est} \geqq \max_i \left\{ \frac{f(x^*) - \mu}{g_i(x^*) - b_i}, \ (g_i(x^*) - b_i) \neq 0 \right\}$$

(Lorena, [26]), onde x^* é qualquer solução inviável de (P) e μ uma subestimativa de $L(x,h_1)$.

A partir dessa constatação e dos resultados do Capítulo II, o seguinte algoritmo poderia ser usado para a resolução do problema (P).

Algoritmo 1:

Passo 1: estimar k_1 por k_{est},

Passo 2: fazer $\lambda_i = k_{est}, \ \forall \ i,$

$\beta_i = \lambda_i, \ \forall \ i,$

Passo 3: otimizar $L(x, h_1)$ obtendo x^*,

se x^* é viável pare,

senão faça $k_{est} = \delta . k_{est}$ $(1 < \delta < \infty)$ e volte ao Passo 2.

A aplicação do Algoritmo 1 estará comprometida pelo fato de que a função objetivo de (PP) será, em geral, não diferenciável, como foi visto no Capítulo II.

No Passo 2 do Algoritmo 1 foi usada a condição $\lambda_i = \beta_i = k_{est}$, $\forall$ i. Tomando agora novamente o problema discreto (PD), observa-se que na Figura 3 - caso b, foi levantada a hipótese de obter-se a solução ótima de (PD) mesmo com $\emptyset_{pd}(b) < L(x,h_1)$. Consequentemente o Passo 2 do Algoritmo 1 poderia ser reescrito como

Passo 2: fazer $\lambda_i = k_{est}$, $\forall$ i;

$\beta_i = k_{est} - E$, $\forall$ i, $(E>0)$.

Deve-se ter uma idéia do tamanho de $E = \max_i (\lambda_i - \beta_i)$. Obviamente $E < k_{est}$, pois $\beta_i > 0$, $\forall$ i. Observa-se ainda na Figura 3 - caso b, que h_1 deverá ser uma H-supergradiente de $\emptyset_{pd}$ em $g(\tilde{x})$, onde $\tilde{x}$ é uma solução ótima de (PD). Formaliza-se a seguir essa observação.

Define-se

$$L_V(x,h_1) = f(x) - h_1(g(x)) + h_1(g(\tilde{x})), \ \forall\, x \in X,\ h_1 \in H_1 \subset H_+^m,$$

onde $\tilde{x}$ é uma solução ótima de (PD).

Proposição 12: $\sup\limits_{x \in X} L_V(x,h_1) = \emptyset_{pd}\,(g(\tilde{x}))$

se e só se $h_1 \in \partial_{H_1}^{g(\tilde{x})}(\emptyset_{pd})$.

Prova: Similar a da Proposição 7.

A seguir verifica-se a condição que deverá existir sobre os parâmetros λ e β ($\lambda_i = \lambda$, $\beta_i = \beta$, $\forall$ i) para que $h \in \partial_{H_1}^{g(\tilde{x})}(\emptyset_{pd})$.

Lema 8: Seja $\bar{x}$ uma solução estritamente viável de (PD) que não é ótima, então $|h_1(g(\tilde{x})) - h_1(g(\bar{x}))|$ será diretamente proporcional a $|\sum\limits_i (g_i(\tilde{x}) - g_i(\bar{x}))|$ com constante de proporcionalidade $(\lambda - \beta)$.

Prova: $|h_1(g(\tilde{x})) - h_1 g(\bar{x}))| =$

$$|\sum_i [\lambda(g_i(\tilde{x}) - b_i) + \beta|g_i(\tilde{x}) - b_i|] -$$

$$\sum_i [\lambda(g_i(\bar{x}) - b_i) + \beta|g_i(\bar{x}) - b_i|]| =$$

$$|(\lambda - \beta) \sum_i (g_i(\tilde{x}) - b_i) - (\lambda-\beta) \sum_i (g_i(\bar{x}) - b_i)| =$$

$$(\lambda - \beta)|\sum_i (g_i(\tilde{x}) - g_i(\bar{x}))|.$$

Como $(\lambda-\beta) > 0$, o resultado segue.

Proposição 13: Se $\tilde{x}$ é uma solução ótima de (PD), $(\lambda-\beta)$ é "suficientemente pequeno" e $k_{est} \geqq k_1$ (finito), então $\sup_{x \in X} L_v(x,h\ ) = $
$= \emptyset_{pd}(g(\tilde{x}))$.

Prova: Deve-se mostrar que para $(\lambda-\beta)$ "suficientemente pequeno"

$$\emptyset_{pd}(g\tilde{x})) - \emptyset_{pd}(d) \geqq h_1(g(\tilde{x})) - h\ (d), \qquad \ldots(6)$$

para todo $d \leqq b$.

Se d=b, a relação (6) é satisfeita, pois, $h_1(b) = 0$, $\emptyset_{pd}(g(\tilde{x})) = \emptyset_{pd}(b)$ e $h_1(g(\tilde{x})) < 0$.

Supondo $d= g(\bar{x}) < b$:

Se $\bar{x}= \tilde{x}$, a relação (6) é trivialmente satisfeita,

se $\bar{x} \neq \tilde{x}$ é ótima, $\emptyset_{pd}(g(\tilde{x}))= \emptyset_{pd}(g(\tilde{x})$, e $h_1(g(\bar{x})) < 0$, portanto (6) é satisfeita;

se $\bar{x} \neq \tilde{x}$ não é ótima, $\emptyset_{pd}(g(x)) > \emptyset_{pd}(g(\bar{x}))$ pela definição de $\emptyset_{pd}$. Assim, para $|h_1(g(\tilde{x}))-h_1(g(\bar{x}))|$ "suficientemente pequeno" (ou de acordo com o Lema anterior $(\lambda-\beta)$ "suficientemente pequeno")

$$\emptyset_{pd}(g(\tilde{x}))-\emptyset_{pd}(g(\bar{x})) \geqq |h_1(g(\tilde{x})) - h_1(g(\bar{x}))|$$
$$\geqq h_1(g(\tilde{x})) - h_1(g(\bar{x})).$$

Assim, pela Proposição 13, deve-se procurar fazer quando da otimização de $L(x,h_1)$, $(\lambda-\beta)$ "pequeno", e procurar garantir a viabilidade com $k_{est} > k_1$.

Com referência ao Passo 2 do Algoritmo 1, mesmo com esse resultado não há, meios de fugir da não-diferenciabilidade de h_1, embora para (PD), esse não seja o maior problema, mas sim os problemas de enumeração que aparecem ao tentar sua otimização. Na próxima seção considera-se um problema particular de programação discreta, o problema multidimensional da mochila tipo zero-um. O objetivo será aplicar as idéias contidas na proposição 13 na geração de heurísticas para aproximar a solução do problema. Tal procedimento é justificado, pois tal problema pertence a classe NP-hard (Magazine e Oguz, [30]).

2. PROBLEMA MULTIDIMENSIONAL DA MOCHILA ZERO-UM

Seja o problema

$$\text{(PMO-1)} \quad \begin{aligned} &\max \sum_{j=1}^{n} c_j\, x_j \\ &\sum_{j=1}^{n} a_{ij} x_j <= b_i, \ i = 1, \ldots, m \\ &x_j \in \{0,1\} \end{aligned}$$

onde a_{ij}, b_i, c_j > = 0, $\forall$ i, j,

O problema (PMO-1) é o conhecido problema multidimensional da mochila, com variáveis bivalentes do tipo zero-um. O "problema com penalidades" usando $h_1 \in H_1 \subset H_t^m$, que estará relacionado a (PMO-1) será

(PPMO-1)

$$\max \sum_j c_j x_j - \sum_i \left[\lambda \left(\sum_j a_{ij}\, x_j - b_i\right) + \beta \left|\sum_i a_{ij} x_j - b_i\right|\right]$$

$$x_j \in \{0,1\}$$

Pode-se aproximar (PPMO-1) por

$$\max \sum_j (c_j - (\lambda \pm \beta) \sum_i a_{ij})\, x_j - (\lambda \pm \beta) \sum_i b_i \qquad (7)$$

$$x_j \in \{0,1\}$$

onde usa-se $\begin{cases} (\lambda - \beta), & \text{se } \sum_j a_{ij}\; x_j \leq b_i, \\ (\lambda + \beta), & \text{se } \sum_j a_{ij}\; x_j > b_i. \end{cases}$

A otimização de (7) poderá ser feita atribuindo

$$\begin{cases} x_j = 1, \text{ para } (\lambda \pm \beta) \leq c_j/\sum_i a_{ij} \\ x_j = 0, \text{ caso contrário.} \end{cases} \qquad (8)$$

Os resultados da proposição 13, mostram que para obtenção de uma solução ótima para (PMO-1) deve-se fazer na otimização de (7), $(\lambda - \beta)$ "suficientemente pequeno" e $(\lambda + \beta)$ "suficientemente grande" $(\geq k_1)$. A expressão (8) apresenta uma nova condição, específica de (PMO-1) que deverá ser considerada em conjunto com as anteriores. Como se observa, usando a expressão (8), existirá um compromisso para a seleção de variáveis a serem atribuidas o valor 1, porque para $(\lambda - \beta)$

"pequeno" a expressão será satisfeita para um grande nº de índices j, enquanto que $(\lambda + \beta)$ "grande" terá o efeito oposto.

A seguir apresenta-se uma heurística gulosa, que seleciona as variáveis cujo índice corresponde aos maiores $c_j/\sum_i a_{ij}$, e que poderá ser usada para aproximar a solução de (PMO-1).

Algoritmo 2:

Inicialização: $x_j = 0$, $j = 1, \ldots, n$

Passo 1: Ordenar os $c_j/\sum_i a_{ij}$ do maior para o menor,

Passo 2: Obedecendo a ordenação fazer $x_j = 1$ enquanto a solução for viável.

Observa-se que a viabilidade ($(\lambda + \beta)$ "suficientemente grande") estará comprometida no algoritmo 2, para alguns índices j selecionados no Passo 2, isto é, a razão (8) poderá não ser grande o suficiente para certos índices j. Certamente o critério $(\lambda - \beta)$ "pequeno" estará quase plenamente satisfeito.

É interessante observar que, para m=1, isto é, quando no problema (PMO-1) tem-se somente uma restrição, a solução será a mesma obtida com a Relaxação de Programação Linear do problema (solução do dual de (PMO-1) - veja Maculan [29]).

Assim, como se deseja obter a melhor solução viável, pode-se modoficar o Algoritmo 2 de forma que a cada passo sejam consideradas as variáveis com maior possibilidade de constituírem uma melhor solução viável.

Pode-se ter por exemplo o algoritmo:

Algoritmo 3:

Inicialização: $x_j=0$, $j=1, \ldots, n$

Passo 1: Ordenar os $c_j * \sum_i (b_i - a_{ij}) / \sum_i a_{ij}$ do maior para o menor,

Passo 2: Obedecendo a ordenação fazer $x_i=1$ enquanto a solução for viável.

O termo $\sum_i (b_i - a_{ij})$ proporciona uma medida da "contribuição individual" da variável j para a viabilidade. Com a modificação na expressão (8), a ordenação do algoritmo 3 passa a evidenciar as seguintes variáveis:

(i) A variável j com custo c_j "grande", o que é conveniente a um problema de maximização. A menos que a_{ij} seja "muito próximo" a b_i, $\forall_i$;

(ii) A variável j com custo c_j "pequeno", mas com a_{ij} também "pequeno" em relação a b_i. Tal variável, embora com grande possibilidade de vir a compor uma solução viável, dificilmente seria selecionada no algoritmo 2.

A seguir apresenta-se algumas de heuristicas para o problema multidimensional da mochila zero-um, todas polinomiais e que produzem bons limitantes a baixo tempo computacional. A característica comum a esses algoritmos, destacada nesse estudo, é que todos estão baseados no princípio que foi descrito até aqui, isto é, na aplicação aproximada da função penalidade h_1 ao problema da mochila.

Basicamente todos procuram considerar os dois princípios: $(\lambda - \beta)$ "pequeno" e a viabilidade.

Algoritmo 4: (Kochenberger, McCarl e Wyman [25]).

Inicialização: $x_j=0$, $j=1, \ldots, n$

Passo 1: $a_{ij} \longleftarrow a_{ij}/(b_i - \sum_{j=1}^{n} a_{ij})$

Calcular o maior $c_j/\sum_i a_{ij}$,

fazer $x_j=1$,

Passo 2: Se a solução é viável voltar ao Passo 1, senão pare.

Algoritmo 5: (Senju e Toyoda [39]).

Inicialização: $x_j=1$, $j=1, \ldots, n$

Passo 1: $a_{ij} \longleftarrow a_{ij}/b_i$

$$\text{Calcular o menor } c_j / \sum_i s_i\, a_{ij}$$

onde $s_i = \max\ (0, \sum_k a_{ik} - 1)$, $i=1, \ldots, m$

k = índice das variáveis iguais a 1

fazer $x_j=0$.

Passo 2: Se a solução é inviável volte ao Passo 1,

Passo 3: Checar se existe variável $x_j=0$, que pode ser colocada igual a 1, sem violar a viabilidade.

Zanakis [46] realizou um estudo estatístico e computacional dos Algoritmos 4 e 5, concluindo que ambos são muito rápidos computacionalmente (como era de se esperar, pois ambos são algoritmos polinomiais) e estatisticamente produzem uma margem pequena de erros (entre 0,2 a 2,6% relativo aos melhores resultados heurísticos).

Recentemente, Magazine e Oguz [30], apresentaram o seguinte algoritmo, que, segundo os autores, combina o algoritmo de Senju e Toyoda com as idéias de "Lagrangeano generalizado" de Everett. Mas, como se verifica a seguir, pode ser considerado um algoritmo que usa os conceitos da aplicação de função penalidade ao problema da mochila.

Algoritmo 6: (Magazine e Oguz [30]).

Inicialização: $x_j=1$; $j=1, \ldots, n$

$$a_{ij} \longleftarrow a_{ij}/b_i; \quad j=1, \ldots, n \text{ e } i=1, \ldots, m$$

$$y_i = \sum_{j=1}^{n} a_{ij}, \quad i=1, \ldots, m$$

Passo 1: Achar i tal que $y_{i*} = \max\limits_i y_i$

Passo 2: Calcular o menor (para $j : x_i=1$)

$$\frac{c_j - a_{i*j}}{\sum\limits_i a_{ij}}$$

e fazer $x_j=0$, $y_i=y_i - a_{ij}$,

Passo 3: Se $y_i > 1$ para algum i ($i = 1, \ldots, m$) volte ao Passo 1,

Passo 4: Checar se existe alguma variável x_j,

$i \in \left\{j:x_j=0\right\}$ que pode ser colocada igual a 1 sem violar a condição $y_i <= 1$.

Os autores fazem ainda uma comparação de seu algoritmo com os algoritmos 4 e 5, concluindo por sua "vantagem" para alguns problemas, e, principalmente, por dispor também de um limitante superior ao valor ótimo do Problema (PMO-1) e dos "multiplicadores de Lagrange" correspondentes a cada restrição do problema.

Obando [33] em sua tese propõe o seguinte algoritmo em que soluções ε-ótimas também são obtidas.

Algoritmo 7: (Obando [33]).

Passo 1: Resolver o problema dual da relaxação de programação linear de (PLIO-1), usando a teoria de Programação Linear para Problemas de grande porte. Aproximar $x_j = 1$ ou $x_j = 0$;

Passo 2: Checar se existe variável x, $j \in \left\{j : x_j = 0\right\}$ que pode ser colocada igual a 1 sem violar a condição de viabilidade.

Obando usa o critério: $\max \left(\dfrac{c_j}{\sum_i (a_{ij}/b_i)} \right)$

Obando testou seu algoritmo para problemas da mochila zero-um com uma restrição e para alguns problemas multidimensionais encontrados na literatura. Em geral os limitantes obtidos foram bons, embora o tempo computacional seja invaria velmente grande (devido ao Passo 1 de seu algoritmo), comparado ao tempo computacional dos algoritmos 2 a 6.

Os problemas testados incluem principalmente os de Petersen [35], Senju e Toyoda [39] e Weingartner e Ness [43]. Seus tempos computacionais (Burroughs 6700) situaram-se na faixa de 6 segundos, enquanto que para os algoritmos 2 a 6 os tempos estão na faixa de 0,2 a 1,8 segundos.

Testes computacionais (Burroughs 6800) realizados com os problemas citados usando os algoritmos 2 e 3, mostraram limitantes melhores para o problema unidimensional e invariavelmente limitantes piores (que os de Obando) para os problemas multidimensionais. Isso é justificado pelo fato de que os algoritmos ainda não considerarem de maneira eficiente o critério da viabilidade (como nos algoritmos 4 e 6) usada no desenvolvimento da expressão (8).

Em geral, pode-se dizer que para o caso unidimensional, os Algoritmos 2 a 7 se equivalem na obtenção de limitantes para (PMO-1), embora não em termos computacionais. Para problemas multidimensionais, os Algoritmos 4 e 6 são superiores, em limitantes e tempos computacionais (usando o Burroughs 6800).

3. PROBLEMA LINEAR INTEIRO ZERO-UM

O problema linear inteiro zero-um será

$$\max \sum_{j=1}^{n} c_j \, x_j$$

(PLIO-1) $$\sum_{j=1}^{n} a_{ij} \, x_j \leq b_i \; , \; i=1, \; \ldots \; , \; m$$

$$x_i \in \left\{0,1\right\}.$$

O problema (PLIO-1) geral será aquele em que a_{ij}, b_i e c_j não sofrem restrições de não negatividade. Sem perda de generalidade, pode-se fazer $c_i \geq 0$, mediante uma mudança de variáveis ($x_j' = 1 - x_j$).

O problema (PPLIO-1) fica aproximado por

$$\max \sum_i \left[c_j - (\lambda + \beta) \sum_i a_{ij}\right] x_j + (\lambda \pm \beta) \sum_i b_i$$

$$x_j \in \left\{0,1\right\}$$

donde, otimizando

$$\begin{cases} x_j = 1, \text{ para } (\lambda \pm \beta) \leq \dfrac{c_j}{\sum_i a_{ij}} \; , \; \sum a_{ij} > 0 \; \ldots \\ \\ x_j = 0, \text{ caso contrário.} \end{cases} \qquad (9)$$

Observe-se que como $c_j(\lambda \pm \beta) >= 0$, para $\sum_i a_{ij} \leq 0$ a expressão $c_j - (\lambda \pm \beta) \sum_i a_{ij} \geq 0$ é satisfeita. Combinando

-se os resultados da proposição 13 com a expressão (9) pode-se formular o seguinte Algoritmo.

Algoritmo 8:

Inicialização: $x_j=1$, $j=1, \ldots, n$

Passo 1: Ordenar os $c_i/\sum_i a_{ij}$, $\sum_i a_{ij} > 0$ do menor para o maior.

Passo 2: Obedecendo a ordenação fazer $x_j=0$ enquanto $\sum_i a_{ij} > 0$ e a solução for inviável.

Passo 3: Se a solução for viável pare, senão, ordenar os $\sum_i a_{ij} \leq 0$ do maior para o menor e obedecendo a ordenação fazer $x_j=0$ enquanto a solução for inviável.

O Algoritmo 8 foi aplicado a alguns problemas de enumeração implícita encontrados em Salkin [38] (pag. 236) e um problema de cobertura de conjunto ("set covering"- pag.466). Os limites e os tempos computacionais foram bons, mas pode-se antecipar dificuldades no tratamento de problemas maiores, principalmente os de cobertura de conjunto, que proporcionam muitos empates na ordenação dos $c_j/\sum_i a_{ij}$.

Existe ainda a possibilidade de usar os critérios (8) ou (9) para fixar variáveis. Por exemplo, para Algoritmo 1, após o Passo 1 (ordenação) pode-se fixar $x_j=1$ para o índice j correspondente aos "maiores" $c_j/\sum_i a_{ij}$ e x_j para os "menores". A seguir aplica-se um Algoritmo de enumeração ao problema resultante.

A quantidade de variáveis a serem fixadas depende de experimentação. Certamente o tempo computacional será reduzido, em comparação a aplicação direta de um algoritmo de enumeração, embora não exista garantia de otimalidade.

CAPÍTULO IV : CONCLUSÕES

Este trabalho apresentou uma forma inédita de aplicação de uma nova classe de funções penalidade a Problemas Discretos, em especial ao Problema de Programação Linear Inteira zero-um. Muito pouco se tem notícia da aplicação de funções penalidade em Problemas Discretos. Blair e Jeroslow [02] apresentam uma conceituação teórica da aplicação do Lagrangeano Aumentado de Rockafellar ao Problema de Programação Linear Inteira Mista com restrições de igualdade. Os autores apenas citam resultados computacionais obtidos em O'neill [34].

Os conceitos estendidos de função côncava, conjugada e superdiferencial, que aparecem no Capítulo II, Seção 1, são particularizações para R^n das definições de Dolecki e Kurcyusz [05], bem como a classe de funções H_n do Capítulo II, Seção 3. A Seção 2 do Capítulo II, apresenta um conjunto de extensões que aparecem, em parte, no trabalho de Tind e Wolsey [42], embora os autores não avancem muito nessa direção.

Os resultados da Seção 3, Capítulo II, generalizam as condições de Kuhn-Tucker e se assemelham as condições de qualificação e/ou estabilidade para o problema geral considerado.

Da classe de funções H_α, a função h_1 faz parte da classe de funções superaditivas, e um resultado relevante é que h_1 é uma função que proporciona a dualidade forte entre os problemas apresentados e não é igual a função perturbação do problema primal. Tind e Wolsey apresentam uma classe tão geral que a função perturbação está contida nessa classe.

Com referência a aplicação de h_1 ao problema da mochila zero-um, pode-se dizer o seguinte. Shapiro [40] destaca em seu trabalho que a otimização de $L(x,h_1)$ onde $\beta_i = 0, \forall\ i$ (caso particular de h_1), foi inicialmente proposta a 30 anos atrás por Lorie e Savage [28], embora com essa abordagem nada poderá ser afirmado quanto ao resultado obtido (nem mesmo se a solução obtida é viável). Com a abordagem utilizada para h_1, com $\beta_i = \beta > 0, \forall\ i$, através da proposição 13 e do auxílio da abordagem de Lorie e Savage chegou-se a justificar as heurísticas apresentadas no Capítulo III, apresentando de uma forma conveniente o compromisso entre os parâmetros de penalidade $(\lambda-\beta)$ e $(\lambda+\beta)$ com o objetivo de aproximar a solução ótima do problema.

Embora mostrada a equivalência entre os problemas (P) e (PP), a conclusão da não-diferenciabilidade limitou as aplicações das funções de H_α. A aplicação aproximada de h_1 teve o mérito de justificar uma série de heurísticas, apresentando-as como aplicações de um conceito globalizante. A aplicação direta de h_1 e aproximada (ou direta) das outras funções de H_α, sugerem pesquisas futuras (observe que h_2 será diferenciável quando usada de forma adaptada, como na Figura 3-d), bem como a aplicação de H_α a outros problemas de Combinatória.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] ARAOZ J. *Poliedral Neopolarities.* Tese Ph.D. University of Notre Dame, 1972.

[2] BLAIR, C.E.; JEROSLOW, R.G. *An exact penalty method for mixed-integer programs.* Mathematics of Operations Research 9:14-18, 1981.

[3] BURDET, C.A.; JOHNSON, E.L. *A subadditive approach to solve linear integer programs.* Annals of Discrete Mathematics 1:117-144, 1977.

[4] CLARKE, F.H. *A new approach to Lagrange multipliers.* Mathematics of Operations Research 1:165-174, 1976.

[5] DOLECKI, S.; KURCYUSZ, S. *On ∅-convexity in extremal problems.* SIAM J. Control and Optimization 16:277-300, 1978.

[6] DORN, W.S. *Duality in Quadratic Programming.* Quart. Appl. Math. 21:237-243, 1960.

[7] ELSTER, K.H.; NEHSE, R. *Zur theorie der polarfunktionale.* Math. Operationsforsch Statist 5:3-21, 1974.

[8] EVERETT, H. *Generalized Lagrange multiplier method for solving problems of optimum allocation of resources.* Operations Research 11:399-417, 1963.

[9] FENCHEL, W. *Convex cones, sets and functions,* from notes by D.W. Blackett of lectures at Princeton University. Princeton University, Department of Mathematics. September, 1953.

[10] FENCHEL, W. *On conjugate convex functions.* Can. J. Math. 1:73-77, 1949.

[11] FISCHER, M.L. *The Lagrangian Relaxation method for solving Integer Programming Problems.* Management Science 27:1-18, 1981.

[12] GARCIA-PALOMARES, U.M. *Connections among Nonlinear Programming, minimax and exact penalty functions.* Relatório Técnico do Departamento de Processos y Siste-mas, Universidad Simon Bolivar, 1984.

[13] GEOFFRION, A.M. *Duality in Nonlinear Programming: a simplified applications oriented development.* SIAM Review 13:1-37, 1971.

[14] GEOFFRION, A.M. *Lagrangean Relaxation to Integer Programming.* Math. Progr. Study 2:82-114, 1974.

[15] GOMORY, R.E.; JOHNSON, E.L. *Some continuous functions related to corner polyedron.* Math. Programming 3:23-85, 1972.

[16] GOULD, F.J. *Extensions of Lagrange multipliers in Nonlinear Programming.* SIAM J. on Applied Math. 17:1280-1297, 1969.

[17] HAN, S.P.; MANGASARIAN, O.L. *Exact penalty functions in nonlinear programming.* Math. Progr. 17:251-269, 1979.

[18] HELD, M.; KARP, R. *The travelling salesman problem and minimum spanning trees.* Operations Research 18:1138, 1970.

[19] HELD, M.; KARP, R. *The travelling salesman problem and minimum spanning trees II.* Math. Progr. 1:6, 1971.

[20] JEROSLOW, R.G. *Cutting plane theory: algebraic methods.* Discrete Mathematics 23:121-150, 1978.

[21] JEROSLOW, R.G. *Minimal Inequalities.* Math. Progr. 17:1-15, 1979.

[22] JOHNSON, E.L. *Cyclic groups, cutting planes and shortest paths.* in: T.C. HU and S. ROBISON eds., Mathematical Programming - Academic Press, New York, 1973.

[23] JOHNSON, E.L. *On the group problem for mixed integer programming.* Math. Progr. Study 2:137-179, 1974.

[24] JOHNSON, E.L. *On the group problem and a subadditive approach to integer programming.* Annals of Discrete Mathematics 5:97-112, 1979.

[25] KOCHENBERGER, G.A.; McCARL, B.A.; WYMAN, F.P. *A heuristic for general integer programming.* Decision Science 5:1, 1974.

[26] LORENA, L.A.N. *Uso de uma função penalidade exata para otimização em problemas discretos.* XVI Simpósio Brasileiro de Pesquisa Operacional: 360-376, 1983.

[27] LORENA, L.A.N.; OLIVEIRA, P.R. *Equivalência entre problemas de maximização sujeitos a restrições de desigualdades.* XVII Simpósio Brasileiro de Pesquisa Operacional: 254-270, 1984.

[28] LORIE, J.; SAVAGE, L.I. *Three problems in capital reationing.* Journal of Business: 229-239, 1955.

[29] MACULAN, N. *Relaxation Lagrangienne: le probleme du knapsack 0-1.* Canadian Journal of Operational Research and Information Processing 21:315-327, 1982.

[30] MAGAZINE, M.J.; OGUZ, O. *A heuristic algorithm for the multidimensional zero-one knapsack problem.* European Journal of Operations Research 16:319-326, 1984.

[31] MANGASARIAN, O.L.; PONSTEIN, J. *Minimax and Duality in Nonlinear Programming.* J. Math. Anal. Appl. 11:504-518, 1965.

[32] MEYER, R.R. *On the existence of optimal solutions to integer and mixed-integer programming problems.* Math. Programming 7:223-235, 1974.

[33] OBANDO, F.E.V. *Um algoritmo eficiente para encontrar soluções aproximadas de problemas de programação bivalente (0-1), usando dualidade em programação intei ra.* Tese do Programa de Engenharia de Sistemas da COPPE: agosto de 1978.

[34] O'NEILL, P. *A penalty function approach to Integer and mixed integer programming.* Talk given at tenth International Symposium on Math. Programming. Montreal, August 27-31, 1979.

[35] PETERSEN, C. *Computational experience with variants of the Balas Algorithm applied to the selection of R and D projects.* Management Science 13:736-750, 1967.

[36] ROCKAFELLAR, R.T. *Duality in Nonlinear Programming.* Mathematics of the Decision Sciences - Part 1, G.B. DANTZIG and A.F. VEINOTT eds., American Mathematical Society - Providence: 401-422, 1968.

[37] ROCKAFELLAR, R.T. *Augmented Lagrange multiplier functions to duality in nonconvex programming.* SIAM J. on Control 12:268-283, 1974.

[38] SALKIN, H.M. *Integer Programming.* Addison-Wesley - -Reading, 1975.

[39] SENJU, S.; TOYODA, Y. *An approach to linear programming with zero-one variables.* Management Science 15:196-207, 1968.

[40] SHAPIRO, J.F. A survey of Lagrangean techniques for Discrete Optimization. *Annals of Discrete Math.* 5:113-138, 1979.

[41] STOER, J. *Duality in Nonlinear Programming and the minimax theorem.* Num. Math. 5:371-379, 1963.

[42] TIND, J.; WOLSEY, L.A. *An elementary survey of general duality theory in Mathematical Programming.* Math. Progr. 21:241-261, 1981.

[43] WEINGARTNER, H.; NESS, D. *Methods for solution of 0-1 knapsack problems.* Operations Research 15:83-103, 1967.

[44] WOLFE, P. *A duality theorem for Nonlinear Programming.* Quart. Appl. Math. 19:239-244, 1961.

[45] WOLSEY, L.A. *Integer programming duality: price functions and sensitivity analysis.* Math. Progr. 20:173-195, 1981.

[46] ZANAKIS, S.H. *Heuristic 0-1 Linear Programming: Comparisons of tree methods.* Management Science 24:1, 1977.