1. Publicação nº	2. Versão	3. Data	5. Distribuição
INPE-3548-PRE/760		*Junho, 1985*	☐ Interna ☒ Externa
4. Origem *DIN/DEP*	Programa *POPES/INFOR*		☐ Restrita

6. Palavras chaves - selecionadas pelo(s) autor(es)
DUALIDADE EM PROGRAMAÇÃO MATEMÁTICA
FUNÇÕES PENALIDADE EXATAS
LAGRANGEANOS AUMENTADOS

7. C.D.U.: *512.852*

8. Título *INPE-3548-PRE/760*

UMA ABORDAGEM GERAL DE DUALIDADE/PENALIDADE EM PROGRAMAÇÃO MATEMÁTICA

10. Páginas: *13*

11. Última página: *12*

12. Revisada por

Horácio Hideki Yanasse

13. Autorizada por

Marco Antonio Raupp
Diretor Geral

9. Autoria *Luiz Antonio Nogueira Lorena*
** Paulo Roberto Oliveira*

Assinatura responsável

14. Resumo/Notas

Demonstra-se, usando conceitos generalizados de funções côncavas, conjugadas e superdiferenciais, os principais teoremas de dualidade entre um problema geral de maximização sujeito a restrições de desigualdades (problema primal) e seu "dual generalizado". Mostra-se também a equivalência entre o problema primal e um "problema com penalidades". O "dual generalizado" e o "problema com penalidades" são definidos usando classes gerais de funções que são reais finitas e não-decrescentes. Mostra-se ainda uma classe de funções que generaliza os Lagrandeanos Aumentados, e outra classe que generaliza as funções penalidades exatas para o problema primal.

15. Observações ** IM/UFRJ - COPPE/UFRJ*

Este trabalho será apresentado no XVIII SOBRAPO, que realizar-se-á em S.J. dos Campos, nos dias 6, 7 e 8 de novembro de 1985.

UMA ABORDAGEM GERAL DE DUALIDADE/PENALIDADE EM PROGRAMAÇÃO MATEMÁTICA

Luiz A. N. Lorena - INPE/CNPq

Paulo R. Oliveira - IM/UFRJ
COPPE/UFRJ

RESUMO

Demonstra-se, usando conceitos generalizados de funções côncavas, conjugadas e superdiferenciais, os principais teoremas de dualidade entre um problema geral de maximização sujeito a restrições de desigualdades (problema primal) e seu "dual generalizado". Mostra-se também a equivalência entre o problema primal e um "problema com penalidades". O "dual generalizado" e o "problema com penalidades" são definidos usando classes gerais de funções que são reais finitas e não-decrescentes. Mostra-se ainda uma classe de funções que generaliza os Lagrangeanos Aumentados, e outra classe que generaliza as funções penalidades exatas para o problema primal.

ABSTRACT

Using generalized concepts of concave functions, conjugate functions and superdifferentials, the main theorems of duality between a general problem of maximization subjected to inequality constraints (primal problem) and its "generalized dual" are shown. The equivalence between the primal problem and a "penalty problem" is also shown. The "generalized dual" and the "penalty problem" are defined using general classes of real finite and nondecreasing functions. Nrxt, a class of functions that generalizes the Augmented Lagrangeans, and a class that generalizes the exact penalty functions for the primal problem are described.

1. Introdução

O objetivo deste trabalho é apresentar uma abordagem geral de dualidade/penalidade em Programação Matemática.

Na Seção 2 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^n. Esses conceitos são usados para construir uma teoria geral de dualidade/penalidade para o seguinte problema geral de maximização, sujeito a restrições de desigualdade:

$$\text{(P)} \qquad v = \sup_{x \in X} f(x)$$

$$\text{suj. a } g(x) \leq b$$

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

Define-se então, o "problema dual" de (P):

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

e o "problema com penalidades"

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

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

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

A Seção 3 dedica-se à caracterização de 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.

2. Generalização de funções côncavas e conjugadas e de superdiferenciais.

Nesta seção são definidos os conceitos generalizados de funções côncavas e 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 $h \in H$ então $h' \in H$, onde $h'(d) = h(d) + r$, $\forall\, d \in R^m$, $r \in R$. De um modo abreviado, escreve-se $H + r \subseteq H$.

Em relação a uma dada função $f:R^n \to \overline{R}$ ($\overline{R} = R \cup \pm\infty$) pode-se considerar o seguinte subconjunto de H;

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

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 $d \in R^m$, então H^f é vazio.

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

DEFINIÇÃO 2 : uma função $f:R^n \to \overline{R}$ é chamada H - côncava em d_0 se

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

DEFINIÇÃO 3 : Se $f:R^n \to \overline{R}$ 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 $f:R^n \to \overline{R}$, define-se a função H-conjugada de f,

$f^*:H \to \overline{R}$, por

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

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

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

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

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 $f: R^m \to \overline{R}$, define-se sua função H-conjugada segunda,

$f^{**}: R^m \to \overline{R}$, 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 $f(d) \leq f^{**}(d)$, $\forall\, d \in R^m$. No Lema e proposição que se seguem, caracteriza-se uma função H-côncava em um ponto d, através do conceito de função H-conjugada segunda.

LEMA 2: Para uma função $f: R^m \to \overline{R}$,

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

Prova: Tem-se da Definição 5 que

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

Ou, de forma equivalente,

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

$\forall\, r \in R$. Na operação do ínfimo, escolhe-se para cada h, $r_h = -f^*(h)$.

Dada a propriedade de translação tem-se $0 = f^*(h) + r_h = f^*(h')$, onde $h'(d) = h(d) + r_h$. Logo, pelo Lema 1, (ii),

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

PROPOSIÇÃO 1: Uma função $f: R^m \to \bar{R}$ é H-côncava em d_o se e somente se.

$$f(d_o) = f^{**}(d_o).$$

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

Para o conjunto H selecionou-se o subconjunto H^f das funções $h \in H$ 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 $f: R^m \to \bar{R}$ o conjunto de funções $h \in H$ que satisfaz a desigualdade

$$h(d) - h(d_o) \geq f(d) - f(d_o),\ \forall\ d \in R^m,$$

é chamado H-superdiferencial de f em d_o.

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

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

PROPOSIÇÃO 2: As seguintes afirmações são equivalentes:

(i) $h \in \partial_H^{d_o}(f)$;

(ii) $h(d_o) - f(d_o) = \inf_{d \in R^m} \{h(d) - f(d)\}$;

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

Prova: imediata a partir das Definições 4 e 6.

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

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

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

LEMA 3: Se h é uma H-supergradiente de f em d, 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_o, então é H-côncava em d_o.

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

$$f(d_o) = \inf_{h \in H^f} h(d_o),$$

3 - Tópicos em Teoria de Dualidade/Penalidade

Nessa seção considera-se o problema geral (P) e derivam-se os problemas (D) e (PP) (apresentados na seção 1) usando os conceitos generalizados de função côncava, função 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 $H_+^m(\bar{H}_+^m)$ o conjunto de funções $h:R^n \to R$ $(\bar{R})$ que são não-decrescentes, isto é, $h(d^1) \geq h(d^2)$, $\forall\, d^1, d^2 \in R^m$ tais que $d_i^1 \geq d_i^2$, $i = 1,\ldots,m$.

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

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

É imediato que $\phi(b) = v$ (valor ótimo do problema (P)).

Observa-se que H_+^m possui a propriedade de translação, isto é, $H_+^m + r \subseteq H_+^m$. Considerando que o conjunto H, definido na Seção 1, 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 4: $H \subseteq H_+^m$ se e só se

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

Prova: imediata.

Relacionando o resultado do Lema 4 com o problema (P) pode-se considerar os $x \in X$ tais que $g(x) = d_o$.

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: $\forall\, h \in H \subseteq H_+^m$,

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

Prova: $\sup\limits_{x \in X} \{f(x) - h(g(x))\} = \sup\limits_{x \in X} \{f(x) - \inf\limits_{\substack{d \\ g(x) \leq d}} h(d)\} =$

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

COROLÁRIO 1: H^ϕ é não-vazio se e só se

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

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, $h(d) = \sum_{i=1}^{m} u_i\, d_i$, $\forall\, d \in R^m$ 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_i's são os conhecidos multiplicadores de Kuhn-Tucker. Em geral para $g(x) \leq b$, $x \in X$ (isto é, x viável para (P)), $L(x,h) \geq f(x)$. Assim $\sup_{x \in X} L(x,h) \geq f(x)$ e $\sup_{x \in X} L(x,h) \geq v$, 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.

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

(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) $h(g(x)) \geq f(x)$, $\forall\, x \in X$ e algum $h \in H \subseteq H_+^m$,

(iii) H^{ϕ} 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 (P) é viável e (D) é viável então $v \leqq w$.

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

LEMA 6: $\phi^{**}(b) = w$.

Prova: $\phi^{**}(b) = \inf_{h \in H} \{h(b) - \phi^*(h)\}$,

mas, pela Proposição 4,

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

Assim,

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

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

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

PROPOSIÇÃO 6: (dualidade forte)

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

Prova: Supondo $v = w = \phi(b)$,
pelo Lema 6, $w = \phi^{**}(b)$.
Assim, $\phi^{**}(b) = \phi(b)$ e ϕ é H-côncava em b.
Supondo agora que ϕ é H-côncava em b, isto é, $\phi^{**}(b) = \phi(b) = v$.
Mas, pelo Lema 6, $w = \phi^{**}(b)$. 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 $L(x,h)$, $h \in \partial_H^b(\phi)$ então $v = w$.

Prova: imediata a partir da Proposição 3.

Mas um resultado mais forte pode ser obtido quando ϕ é H-superdiferenciável.

PROPOSIÇÃO 7 (equivalência): $\sup_{x\in X} L(x,h) = \phi(b)$ se e só se $h \in \partial_H^b(\phi)$.

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

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

Pode-se definir então o problema

(PP) $$p_h = \sup_{x\in X} L(x,h),\ h \in H \subseteq H_+^m.$$

Ficou evidenciada a importância de ser encontrada uma classe de funções $H \subseteq H_+^m$, para a qual ϕ é H-côncava ou H-superdiferenciável em b, assegurando respectivamente a dualidade forte entre (P) e (D) ou a equivalência entre (P) e (PP).

No trabalho de Lorena (1985) apresenta-se as seguintes classes, para as quais ϕ será H-côncava e H-superdiferenciável, respectivamente.

$$H_n = \{h: h(d) = r + \rho.\psi(||d-y||),\ r \in R,\ \rho > 0,$$
$$d \in R^m,\ e\ y \in Y\},$$

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

$||.||$ é um norma de R^m, e

$\psi: R_+ \to R_+$ é uma função não-decrescente, tal que

(i) $\forall\ \delta > 0,\ \forall\ t_o \geq 0,\ \exists M(t_o,\delta)$ (limitante), $\forall t \geq \delta$,

$$\psi(t + t_o)/\psi(t) \leq M(t_o,\delta),$$

(ii) $\psi(t) \to 0 \longleftrightarrow t \to 0$.

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

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

$$H_\alpha = \{h_\alpha: R^m \to R : h_\alpha(d) = \sum_{i=1}^{m} [\lambda_i(d_i-b_i)+\beta_i|d_i-b_i|^\alpha],$$
$$\text{para } \lambda_i,\beta_i \in R_+,\ \alpha > 0,\ b_i \in R\}.$$

Examinou-se a função $h_1 \in H_\alpha$, isto é;

$h_1(d) = \sum_{i=1}^{m} [\lambda_i(d_i-b_i)+ \beta_i|d_i-b_i|]$, no trabalho de Lorena e Oliveira (1984), concluindo-se que h_1 generaliza algumas funções penalidade exatas. A classe H_n generaliza os Lagrangeanos aumentados (Lorena, 1985).

O seguinte teorema generaliza as condições de Kuhn-Tucker para problemas de Programação não-linear. [Lorena, 1985].

TEOREMA 1: $x^* \in X$ é uma solução ótima de (P) e $\bar{h} \in \partial^b_{H_+^m}(\emptyset)$ se e só se as seguintes afirmações são satisfeitas:

(i) $g(x^*) \leq b$,

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

(iii) $\bar{h} \in H_+^m$,

(iv) $\bar{h}(g(x^*)) = \bar{h}(b)$.

4. Conclusões

Os conceitos estendidos de função côncava, conjugada e superdiferencial, que aparecem na Seção 2, são particularizações para R^n das definições de Dolecki e Kurcyusz (1978), bem como a classe de funções H_n da Seção 3. A seção 3 apresenta um conjunto de extensões que aparecem, em parte, no trabalho de Tind e Wolsey (1981), embora os autores não avancem muito nessa direção.

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 à função perturbação do problema primal. Tind e Wolsey (1981) apresentam uma classe tão geral que a função perturbação está contida nessa classe.

5. Agradecimentos

Aos Doutores Horácio Hideki Yanasse (INPE - Divisão de Estatística e Pesquisa Operacional) e Carlos Humes Júnior (I.M.E./ U.S.P.) pelas valiosas sugestões para os resultados da Seção 2.

6. Bibliografia

DOLECKI, S.; KURCYUSZ, S. On Ø-convexity in extremal problems SIAM J. Control and Optimization, 16: 277-300, 1978.

LORENA, L.A.N. Uma nova classe de funções penalidade de sua aplicação a problemas discretos. Tese de Doutoramento no Programa de Engenharia de Sistemas e Computação, Rio de Janeiro, COPPE/UFRJ, 1985.

LORENA, L.A.N.; OLIVEIRA, P.R.. Equivalência entre problemas de maximização sujeito a restrições de desigualdades. Anais do XVII Simpósio Brasileiro de Pesquisa Operacional, 254-270, 1984.

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