Pub. Mat. UAB
Nº 22 Nov. 1980
Actes VII JMHL

# SOBRE UN TEOREMA DE DUALIDADE

J.C. Namorado, Clímaco

Centro de Matemática
Universidade de Coimbra

**Abstract** - In this work we discuss some results presented in [1], [2], [3] and we compare the "conventional dual theory" - using "Lagrange Multipliers" - with a "new dual theory" stated in [1].

## 1. Introdução

Quando pretendemos utilizar métodos de optimização, em problemas de engenharia ou de planeamento, há, por vezes, vantagem em considerar várias funções objectivo conflituosas entre si. Um dos processos de abordar estes problemas é construir uma função utilidade que incorpore os diversos objectivos. Em geral, as funções utilidade pertencem a uma classe de funções, com propriedades específicas, a que chamamos "funções monotonicamente estruturadas".

A estrutura desta classe de funções pode ser explorada para desenvolver métodos que permitam optimizá-las com menos esforço computacional do que se utilizássemos algoritmos clássicos. A este respeito veja-se [1], [2], [4].

Em [1] foi estabelecido um novo tipo de "problema dual", com a particularidade de ser definido tanto para "problemas primais" com restrições, como para problemas sem restrições (dualidade-A).

Em [2] desenvolvemos esta teoria e comparámos as "limitações inferiores" (lower bounds) geradas a partir da teoria da dualidade convencional que utiliza "Multiplicadores de Lagrange" (dualidade-L) e a partir da "dualidade-A". Tivemos em vista o estudo da possibilidade da sua aplicação, para estabelecer condições de paragem, quando se utiliza um processo iterativo para optimizar o "problema primal".

No parágrafo 2 introduz-se a noção de "função monotonicamente estruturada".

Nos parágrafos 3 e 4 focamos alguns dos principais resultados que acabámos de referir. No parágrafo 3 sugerimos ainda um novo tipo de demonstração para o teorema de dualidade apresentado. Em 3 e 4 tornam-se claras as relações existentes entre os dois tipos de problemas duais.

No parágrafo cinco apresentamos a extensão, para condições mais gerais, dum lema para "funções monotonicamente estruturadas" apresentado em [1]. Saliente-se que a aplicação deste lema à "teoria do controlo" é potencialmente interessante [1,2].

Ao longo do texto vamos considerar que o problema de optimização em estudo consiste na minimização duma dada "função objectivo" num determinado domínio. Consideramos ainda, por razões de simplicidade, que existe mínimo não só da "função objectivo" referida, como de todos os problemas auxiliares de minimização considerados, nos respectivos domínios. Note-se que esta suposição é feita por razões de simplicidade.

## 2 - Função Monotonicamente Estruturada

**Definição** - Seja $F$ um sub-conjunto dum espaço linear de dimensão finita ($F \subseteq X$). A função $f: X \to R$ diz-se monotonicamente estruturada se e só se $f(x) = h \circ v(x)$ e:

a - $v(x): X \longrightarrow R^m: x \mapsto [v_1(x), \dots, v_m(x)]^T$ em que $v_i: X \to R$ $(i=1,\dots,m)$ são funções convexas.

b - $h(.): R^m \to R$ é monótona estritamente crescente.

## 3 - Teorema de Dualidade

Em [2] apresentámos uma extensão dum teorema de dualidade introduzido em [1]. Vamos aqui referir uma versão ligeiramente simplificada do teorema.

**Teorema** - Sejam:

1 - $f(x) = h \circ v(x)$ uma "função monotonicamente estruturada" definida em $X$, sendo $X$ um espaço linear de dimensão finita, $h(.)$ sub-diferenciável e convexa, $v(x) = [v_1(x), \dots, v_m(x)]^T$ em que $v_1, \dots, v_m$ são funções diferenciáveis $(\nabla v_i(x) = \partial v_i(x) / \partial x)$.

2 - $P: X \longrightarrow W$ em que $W$ é um espaço linear de dimensão finita, $P$ é linear (i.e. $P \varepsilon L(X,W)$) e $\dim W \geq \min\{m-1, \dim X\}$

3 - $\mathscr{f} = h \circ \nu : X^m \to R$, $\nu : X^m \to R^m$:

$$\mathscr{x} = (x_1, \dots, x_m) \mapsto [v_1, (x_1), \dots, v_m(x_m)]^T$$

4 - $\mathcal{F} = \mathcal{F}(P) \triangleq \{(x_1, \ldots, x_m) \in X^m : Px_1 = \ldots = Px_m\}$

Então,

$$\max_{P} \min_{\mathcal{X}\in\mathcal{F}(P)} f(\mathcal{X}) = \min_{x\in F} f(x) = f(x^*) \tag{1}$$

e um P que maximiza (1) é:

$$P^* = \left[\nabla v_1(x^*), \ldots, \nabla v_{m-1}(x^*)\right]^T, \quad \forall x^* \in \arg\min_{x\in F} f(x)$$

A demonstração que apresentâmos em [2] é uma extensão, com pequenas variantes, da apresentada em [1]. Uma nova demonstração, que sugere a possibilidade de relaxar este teorema para condições mais gerais, pode fazer-se utilizando resultados da teoria da dualidade convencional.

Visto que (ver [1]), $\min_{\mathcal{X}\in\mathcal{F}(P)} f(\mathcal{X}) \leqq \min_{x\in X} f(x)$, $\forall P \in L(X,W)$, resta provar que $P^*$ é tal que

$$\min_{\mathcal{X}\in\mathcal{F}(P^*)} f(\mathcal{X}) = f(x^*)$$

Para isso, consideramos o seguinte problema (idêntico ao "problema primal"):

$$\min_{\mathcal{X}\in X^m} \left\{ h\begin{bmatrix} v_1(x_1) \\ \vdots \\ v_m(x_m) \end{bmatrix} : x_1 - x_2 = 0, \ldots, x_{m-1} - x_m = 0 \right\} \tag{2}$$

em que $x_i = \left[x_i^1, \ldots, x_i^n\right]^T$ para $i=1,\ldots,m$.

A partir de (2) construimos:

$$\min_{\mathcal{X}\in X^m} \left\{ h\begin{bmatrix} v_1(x_1) \\ \vdots \\ v_m(x_m) \end{bmatrix} : P^*(x_1 - x_2) = 0, \ldots, P^*(x_{m-1} - x_m) = 0 \right\} \tag{3}$$

Como o problema (3) é convexo, a função $\Pi(y)$ também é convexa (ver [2]):

$$\Pi(y) = \min_{\mathcal{X}\in X^m} \left\{ h\begin{bmatrix} v_1(x_1) \\ \cdot \\ \cdot \\ \cdot \\ v_m(x_m) \end{bmatrix} : \begin{bmatrix} P^* & 0 & \ldots & 0 \\ 0 & P^* & & \vdots \\ \vdots & & & 0 \\ 0 & \ldots & 0 & P^* \end{bmatrix} \begin{bmatrix} x_1 - x_2 \\ \\ \\ x_{m-1} - x_m \end{bmatrix} = y \right\}$$

em que $y = \left[y_1, \ldots, y_{m-1}\right]^T \in Y_1$

$$Y_1 = \left\{ y \in R^{(m-1)\times(n-1)} \exists (x_1,\dots,x_m) \in X^m : \begin{bmatrix} P^* & 0 & \dots & 0 \\ 0 & P^* & & \\ \vdots & & & 0 \\ 0 & \dots & 0 & P^* \end{bmatrix} \begin{bmatrix} x_1 - x_2 \\ \vdots \\ x_{m-1} - x_m \end{bmatrix} = y \right\}$$

Uma vez que o problema (3) é estável existe o "dual-L" de (3).

A partir deste ponto, com alguma manipulação algébrica, pode demonstrar-se o pretendido.

## 4 - Comparação das "limitações inferiores" geradas utilizando a "dualidade-L" e a "dualidade-A".

Por vezes o problema dual é útil para estabelecer condições de paragem quando se usa um algoritmo iterativo para optimizar o problema primal. Fixando $\varepsilon$(erro permitido) o processo iterativo deve interromper-se quando $f(x_k)-\min_{x\varepsilon F} f(x) \leq \varepsilon$. Como, em geral, $\min_{x\varepsilon F} f(x)$ não é conhecido à priori, é importante conhecer uma "limitação inferior" que convirja o mais rapidamente possível para $f(x^*)$ quando $x$ tende para $x^*$. "Limitações inferiores"com esta propriedade podem obter-se utilizando a "teoria da dualidade convencional" e a "dualidade -A".

Em [2] fazemos a comparação das "limitações inferiores" geradas por estes dois processos.

Vamos referir brevemente um caso particular. Consideremos o "problema primal":

$$\min(P_a) = \min_{x\varepsilon F \underline{\underline{\Delta}} X \underline{\underline{\Delta}} R^2} \left\{ h \begin{bmatrix} v_1(x) \\ v_2(x) \end{bmatrix} \right\}$$

em que as funções obedecem às condições impostas no teorema do parágrafo 2.

O problema $(P_a)$ é equivalente a $(P_b)$:

$$\min(P_b) = \min_{x_1,x_2 \varepsilon R^2} \left\{ h \begin{bmatrix} v_1(x_1) \\ v_2(x_2) \end{bmatrix} : x_1 - x_2 = 0 \right\}$$

Usando a "dualidade-L" obtêm-se "limitações inferiores" da forma:

$$\min(P_c) = \min_{x_1,x_2 \varepsilon R^2} \left\{ h \begin{bmatrix} v_1(x_1) \\ v_2(x_2) \end{bmatrix} + \lambda(x)^T (x_1 - x_2) \right\}$$

$$\leq \min (P_b), \quad \forall \lambda(x) = \begin{bmatrix} \lambda^1(x) \\ \lambda^2(x) \end{bmatrix} \varepsilon R^2$$

Como $(P_b)$ é estável, para $\lambda(x) = \lambda(x^*)$ (com $x^*=x_1^* =x_2^*$) $\min (P_c) = \min(P_b)$.

Assumindo que as funções em causa são contínuas, é possível obter uma expressão para um $\lambda(x)$, com $\lambda(x) \to \lambda(x^*)$ quando $x \to x^*$. Para isso, calcula-se $\theta(x)$ tal que:

$$\min_{\theta \in R^2} \|A - B\theta\|^2$$

Obtém-se $\theta(x) = \left[B^T B\right]^{-1} B^T A$

em que $A = \left[\frac{\partial f(x)}{\partial x_1^1} , \frac{\partial f(x)}{\partial x_2^1} , \frac{\partial f(x)}{\partial x_1^2} , \frac{\partial f(x)}{\partial x_2^2}\right]^T$

com $f(x) = h \begin{bmatrix} v_1(x_1) \\ v_2(x_2) \end{bmatrix}$

$$B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}^T$$

$$\theta(x) = \left[\theta^1(x),\theta^2(x)\right]^T$$

Faz-se $\lambda(x) = - \theta(x)$.

Por outro lado, como:

$$\min_{x_1,x_2 \in R^2} \left\{ h \begin{bmatrix} v_1(x_1) \\ v_2(x_2) \end{bmatrix} + \lambda(x)^T (x_1-x_2) \right\} \leq$$

$$\leq \min_{x_1,x_2 \in R^2} \left\{ h \begin{bmatrix} v_1(x_1) \\ v_2(x_2) \end{bmatrix} : \lambda(x)^T(x_1-x_2)=0 \right\}, \quad \forall \lambda(x) \in R^2$$

é possível dizer que as "limitações inferiores" obtidas usando a "dualidade-A" são iguais ou melhores que as obtidas utilizando a "dualidade-L".

**5 - Extensão do Lema 2.1 (i) de [1].**

Lema - Seja $f(x) = h o v(x)$ pertencente à classe das "funções monotonicamente estruturadas" definidas em $F \subseteq X$, sendo $X$ um espaço linear de dimensão finita, F convexo e fechado, $m=2$, $v_1(.)$ e $v_2(.)$ diferenciáveis, $h(.)$ sub-diferenciável e $\forall \partial_1 h(.) \in \partial h(.)$, $\partial_1 h(.) > 0$.

Seja $S \underline{\Delta} \{x: x \in \arg\min_{x \in F} f(x)\}$

$S_i \underline{\Delta} \{x: x \in \arg\min_{x \in F} v_i(x)\}$ $i=1,2.$

Então, se $S_1 \cap S_2 = \phi$, $S \cap S_i = \phi$, $i=1,2.$

A demonstração que propomos para este lema (ver [3]) é completamente distinta da apresentada em [1] para a versão original do lema e utiliza uma função perturbação do tipo da introduzida em [4].

**Referências**


[1] - Allwright, J.C. "Monotonically-Structured Cost Functions: Minimization, Duality and Applications to Optimal Control", Proc.6$^{th}$ World IFAC Congress, Boston, Massachussetts, U.S.A., August, 1975.

[2] - Clímaco, J.C.N. "Minimization, Duality and Applications to Optimal Control of Monotonically Structured Cost Functions" M.Sc. dissertation, Imperial College, London 1977.

[3] - Clímaco, J.C.N. "Note on Some Applications of a Certain Perturbation Function", Revue Belge de Statistique, d'Informatique et de Recherche Operationnelle - aceite para publicação.

[4] - Geoffrion, A.M. -, "Duality in Nonlinear Programming: A Simplified Applications-Oriented Development", Siam Review, 1971, 13,1.

[5] - Geoffrion, A.M. -, "Solving Bicriterion Mathematical Programs", Operations Research 1967, 15, 39.



O autor não quer deixar de agradecer ao Instituto Nacional de Investigação Científica o subsídio que lhe foi concedido.