

UNIVERSIDADE NOVA DE LISBOA

Faculdade de Economia

DUALIDADE CONJUGADA
EM PROGRAMAÇÃO MATEMÁTICA CONVEXA

Maximiano Reis Pinheiro

Working Paper nº 40

UNIVERSIDADE NOVA DE LISBOA

Faculdade de Economia

Campo Grande, nº 185

1700 LISBOA

Junho,

O autor deseja manifestar o seu agradecimento
ao Professor Doutor Dias Coelho pelo seu estí
mulo e apoio, que tornaram possíveis a publi-
cação deste trabalho.

0. INTRODUÇÃO

1. CONJUNTOS CONVEXOS E FUNÇÕES CONVEXAS E CÔNCAVAS

1.1. CONJUNTOS CONVEXOS

1.1.1. Conjuntos Convexos

1.1.2. Conjuntos Fechados

1.1.3. Hiperplanos e Semiespaços de Suporte

1.2. FUNÇÕES CONVEXAS E CÔNCAVAS

1.2.1. Funções Convexas

1.2.2. Funções Côncavas

1.2.3. Funções Lineares

1.2.4. Semicontinuidade Inferior e Superior/Funções Fechadas

1.2.5. Subdiferenciabilidade e Diferenciabilidade de Funções
Convexas e Côncavas

2. FUNÇÕES CONJUGADAS

2.1.1. Funções Convexas Conjugadas e suas Principais Propriedades

2.1.2. Pontos Conjugados/Propriedades Marginais das Funções
Conjugadas/Transformação de Legendre

2.1.3. Conjugadas de Funções Côncavas

3. DUALIDADE CONJUGADA EM PROGRAMAÇÃO MATEMÁTICA CONVEXA

3.1. O PROBLEMA DE PROGRAMAÇÃO CONVEXA

3.1.1. O Problema Generalizado de Programação Convexa

3.1.2. Bifunção Associada e Função Perturbação

3.1.3. Algumas Propriedades da Bifunção Associada e da Função Perturbação

3.1.4. Relação entre Função Auxiliar Lagrangeana e Função Perturbação

3.2. CONJUGAÇÃO E DUALIDADE EM PROGRAMAÇÃO CONVEXA

3.2.1. Bifunções Adjuntas

3.2.2. Programas Duais

3.2.3. Alguns Resultados de Dualidade

3.2.4. Bifunções Adjuntas e Transformações Lineares Adjuntas/Dual dum Programa Linear

3.2.5. Uma Perspectiva Gráfica da Dualidade Conjugada

3.2.6. Relação entre Dualidade Conjugada e Teoria de Lagrange/Kuhn-Tucker

3.2.7. Dualidade Conjugada e Dualidade de Wolfe

3.2.8. Nota Histórica: o Teorema de Fenchel

3.2.9. Uma Generalização: a Inclusão doutro Tipo de Perturbações.

0. INTRODUÇÃO

— O presente texto pretende ser uma versão melhorada e mais completa dum estudo sobre dualidade conjugada em programação matemática convexa, tema já abordado no capítulo 3 do "working paper" nº 25 desta mesma colecção. Tal como nesse paper, pretende-se apenas apresentar um texto didáctico e integrado sobre o assunto, assente nas contribuições originais de Fenchel, Rockafellar e outros. Utilizando uma linguagem o mais simples possível, procurou-se expor um assunto complexo de modo a torná-lo acessível a leitores com um mínimo de preparação matemática.

— Integrado num plano de estudo mais vasto, a apresentação deste "Working paper" deverá ser seguida a breve trecho pela apresentação da dissertação de mestrado em Economia do autor.

Entitulada "Dualidade Conjugada e Quase-Conjugada nas Teorias do Produtor e Consumidor", a referida dissertação tem como objectivo mostrar que as relações de conjugação em análise convexa podem também, para além de constituírem um instrumento válido de análise de dualidade em programação convexa, ser utilizadas com o objectivo diferente de obter e explorar com assinalável êxito a dualidade em teoria microeconómica.

— No que se segue trabalhar-se-á, a menos que explicitado o contrário, no espaço R^n , cujos vectores representaremos, como é usual, por énplos ordenados $x = (x_1, \dots, x_n)^T$. Nesse sentido, e admitindo a base ortonormal $\{e_1, \dots, e_j, \dots, e_n\}$, em que e_j é a j -ésima coluna da matriz identidade de ordem n ,

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j$$

será, como sabemos, o produto escalar entre os vectores x e y . Recor-

demos, a propósito, que definindo $\cos(x,y) = \langle x,y \rangle / \|x\| \cdot \|y\|$, em que $\|x\| = \sqrt{\sum_j x_j^2} = \langle x,x \rangle^{1/2}$ é a norma euclidiana de x , dois vectores não nulos x e y não ditos ortogonais se $\cos(x,y) = 0$, o que acontece, portanto, quando e só quando $\langle x,y \rangle = 0$.

1. CONJUNTOS CONVEXOS E FUNÇÕES CONVEXAS E CONCAVAS

1.1. CONJUNTOS CONVEXOS

1.1.1. Conjuntos Convexos

— Um conjunto $C \subset \mathbb{R}^n$ é convexo se para qualquer $0 \leq \lambda \leq 1$ se verifica:

$$(1-\lambda)x + \lambda y \in C, \quad \forall x,y \in C$$

Significa o anterior que um conjunto é convexo se, contendo dois vectores x e y , contem igualmente todos os vectores do segmento de recta que os une, isto é, contem também todas as suas combinações lineares convexas.

— Entre os conjuntos convexos, e pela importância que assumirão no que se segue, podemos realçar os hiperplanos e os semiespaços por eles gerados. Um hiperplano é uma variedade afim de dimensão $n-1$, isto é, corresponde a uma translação de um subespaço de dimensão $n-1$, e pode ser expresso analiticamente em $\mathbb{R}^n$ como :

$$H = \{ x \mid \langle x,b \rangle = \beta \} \quad \text{com } b \in \mathbb{R}^n \text{ e } \beta \in \mathbb{R}.$$

Semiespaços são conjuntos gerados pelos hiperplanos substituindo a igualdade da equação de definição por uma desigualdade do tipo $\geq, >, \leq$ ou $<$.

1.1.2. Conjuntos Fechados

— Um conjunto $S \subset \mathbb{R}^n$ diz-se fechado se coincidir com o seu fecho $\text{Cl}S$

(do inglês "closure"): $S = \text{cl} S$. Fecho do conjunto S é o conjunto de pontos tais que, para cada $x \in S$, e por mais pequeno que seja $\varepsilon > 0$, tenhamos $S \cap E_\varepsilon(x) \neq \emptyset$, em que $E_\varepsilon(x)$ é a hiperesfera em $\mathbb{R}^n$ de centro x e raio ε :

$$E_\varepsilon(x) = \{y \mid \|y-x\| = \langle y-x, y-x \rangle^{1/2} \leq \varepsilon\}$$

— Interior do conjunto S ($\text{int } S$) é o conjunto $\{x \mid \exists \varepsilon > 0 : E_\varepsilon(x) \subset S\}$. Complementarmente, fronteira de S é o conjunto de pontos que pertencem a $\text{cl} S$ mas não pertencem a $\text{int } S$. Interior relativo de um conjunto $S \subset \mathbb{R}^n$ são o interior e a fronteira de S quando referidos não a $\mathbb{R}^n$ mas à menor variedade afim que o contém.

1.1.3. Hiperplanos e semiespaços de suporte

— Seja um conjunto não vazio $S \subset \mathbb{R}^n$ e seja $\bar{x}$ num ponto da sua fronteira. Um hiperplano $H = \{x \mid \langle b, x - \bar{x} \rangle = 0\}$ é designado por hiperplano de suporte ou apoio em $\bar{x}$ ao conjunto S desde que:

$$i) \text{ ou } S \subset H^{\geq} = \{x \mid \langle b, x - \bar{x} \rangle \geq 0\}$$

$$ii) \text{ ou } S \subset H^{\leq} = \{x \mid \langle b, x - \bar{x} \rangle \leq 0\}$$

isto é, desde que S esteja contido num dos semiespaços fechados gerados por H . O semiespaço gerado pelo hiperplano de suporte e que contém S é chamado de semiespaço de suporte ou apoio ao conjunto S no ponto $\bar{x}$.

1.2. FUNÇÕES CONVEXAS E CÔNCAVAS

1.2.1. Funções convexas

— Epigrafo de numa função $f(x)$ de $\mathbb{R}^n$ em $[-\infty, +\infty]$ (admitindo que as

funções reais podem também assumir os "valores" $-\infty$ e $+\infty$) é o subconjunto de $\mathbb{R}^{n+1}$ definido por:

$$\text{epi } f(x) = \{ (x, r) \mid x \in \mathbb{R}^n, r \in \mathbb{R}, r \geq f(x) \}$$



Uma função é convexa se o seu epigrafo for um conjunto convexo. (1)

— Domínio efectivo de uma função $f(x)$, ou mais simplesmente, dom f , nio de $f(x)$, é o conjunto de pontos de $\mathbb{R}^n$ tais que $f(x) < +\infty$. Dito de outra forma, $\text{dom } f(x)$ é a projecção de $\text{epi } f(x)$ sobre $\mathbb{R}^n$.

Uma função convexa diz-se própria se simultaneamente $\exists x \in \mathbb{R}^n$ $f(x) < +\infty$ e $\forall x \in \mathbb{R}^n$ $f(x) > -\infty$, isto é, se o epigrafo de $f(x)$ não é vazio nem contém linhas completas verticais.

— Se $f(x)$ é convexa, os seus conjuntos de nível inferior

$$f_{\alpha}^{\leq} = \{ x \mid f(x) \leq \alpha \} \quad \text{ou} \quad f_{\alpha}^{<} = \{ x \mid f(x) < \alpha \}$$

são conjuntos convexos. Esta relação fundamental entre função convexa e o seu epigrafo convexo não é recíproca, pois as funções quase-convexas, e por definição, geram também conjuntos de nível inferior convexos.

(1) Existem outras formas de definir f convexa, menos gerais mas mais comuns:

a) No caso de $f(x)$ não assumir o "valor" $-\infty$, então $f(x)$ é convexa se, para qualquer $0 \leq \lambda \leq 1$

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

b) Se $f(x)$ bidiferenciável em $C \subset \mathbb{R}^n$, então $f(x)$ é convexa em C se a matriz hessiana de f for semidefinida positiva para todo o $x \in C$.

— Designamos por função supremo por pontos de um conjunto de funções $\{ z_i(x) \mid i \in I \}$ a função que a cada x faz corresponder o maior dos valores $z_i(x)$:

$$z(x) = \sup_x \{ z_i(x) \mid i \in I \}$$

Percebe-se sem dificuldade que o supremo por pontos de funções convexas é ainda numa função convexa (Rockafellar, 1970f, teor.5.5,p.35).

— Uma função convexa especialmente útil é a função indicadora de um conjunto convexo. Defina-se função indicadora dum qualquer conjunto C como a função

$$\delta(x|C) = \begin{cases} 0 & \text{se } x \in C \\ +\infty & \text{se } x \notin C \end{cases}$$

Como se percebe, a função indicadora é convexa se o conjunto C que a gera for um conjunto convexo. A função indicadora permite fazer a extensão extremamente cómoda do domínio duma função convexa a $\mathbb{R}^n$:

$$f(x) \text{ com } x \in S \Leftrightarrow f(x) + \delta(x|S) \text{ com } x \in \mathbb{R}^n.$$

1.2.2. Funções côncavas

— Uma função $g(x)$ de $\mathbb{R}^n$ em $[-\infty, +\infty]$ é côncava se $-g(x)$ for uma função convexa. As características de uma função côncava decorrem directamente da sua simétrica ser uma função convexa.

1.2.3. Funções lineares

— Atendendo à definição de funções convexas e côncavas é praticamente imediato que as funções lineares (também designadas por funções afins) $l(x) = \langle x, b \rangle - \beta$ ($b \in \mathbb{R}^n$, $\beta \in \mathbb{R}$) são as únicas funções simultaneamen

te convexas e côncavas.

Verifica-se ainda que o grafo de $l(x)$ corresponde a um hiperplano em $\mathbb{R}^{n+1}$:

$$\{ (x, l(x)) \mid l(x) = \langle x, b \rangle - \beta \}$$

$$\Leftrightarrow \{ (x, l(x)) \mid \langle (x, l(x)), (b, -1) \rangle = \beta \}$$

1.2.4. Semicontinuidade inferior e superior / funções fechadas

— Uma função $Z(x)$ diz-se semicontínua inferiormente (s.c.i.) num ponto x se $Z(x) \leq \lim_{x_i \rightarrow x} Z(x_i)$ qualquer que seja a sequência $x_1, \dots, x_i, \dots$ que convirja para x , existindo, mesmo que não finito, $\lim_{x_i \rightarrow x} Z(x_i)$. Uma função s.c.i em todos os pontos do seu domínio diz-se função semicontínua inferiormente. Duma forma equivalente, uma função diz-se s.c.i se os seus conjuntos de nível $\{ x \mid Z(x) \leq \alpha \}$ são conjuntos fechados, isto é, se o seu epigrafo é um conjunto fechado.

— Uma função $Z(x)$ diz-se semicontínua superiormente (s.c.s.) se $-Z(x)$ é uma função s.c.i, ou equivalentemente, se os conjuntos de nível $\{ x \mid Z(x) \geq \alpha \}$ são conjuntos fechados.

— Uma função convexa diz-se fechada se for uma função semicontínua inferiormente. Analogamente, uma função côncava diz-se fechada se for uma função semicontínua superiormente. Uma função contínua é uma função simultaneamente s.c.i e s.c.s.

1.2.5. Subdiferenciabilidade e diferenciabilidade de funções convexas

— Definem-se subgradientes duma função convexa $f(x)$ num ponto x como

os vectores $s \in \mathbb{R}^n$ tais que:

$$f(w) \geq f(x) + \langle w-x, s \rangle, \quad \forall w \in \mathbb{R}^n.$$

Tal significa que s será subgradiente de f no ponto x se o grafo da função $d(w) = f(x) + \langle w-x, s \rangle$ (com x e s fixos) fôr um hiperplano de suporte ao epigrafo de f no ponto x . (2)

— O conjunto de todos os subgradientes de f no ponto x (que pode ser vazio) designa-se por subdiferencial de f no ponto x e representar-se-á por $\partial f(x)$. No caso de $\partial f(x)$ ser constituído por um único elemento, a função f é diferenciável no ponto x e o gradiente da função nesse ponto $\nabla f(x)$ coincide com o único subgradiente.

— É condição necessária e suficiente para $\bar{x}$ ser mínimo da função convexa $f(x)$ (ou máximo da função côncava) que $0 \in \partial f(\bar{x})$, o que se compreende pela definição de subdiferencial duma função convexa. Sendo f diferenciável, a anterior condição reduz-se a $0 = \nabla f(\bar{x})$.

-
- (2) Se $g(x)$ fôr uma função côncava, e como $-g(x)$ é convexa, teremos s como subgradiente de $g(x)$ se $-s$ fôr subgradiente de $-g(x)$:

$$-g(w) \geq -g(x) + \langle w-x, -s \rangle$$

$\Leftrightarrow$

$$g(w) \leq g(x) + \langle w-x, s \rangle$$

2. FUNÇÕES CONJUGADAS

2.1.1. Funções convexas conjugadas e suas principais propriedades

— Como nota Rockafellar (1970 [f], p. 102), há 2 formas alternativas de caracterizar qualquer curva ou superfície clássica: ou como o lugar geométrico dos seus pontos, isto é, a curva por ela própria, ou então como o envelope das suas tangentes.

Esta dualidade de possíveis formas de encarar a mesma curva ou superfície vai, no caso da análise convexa, levar à noção da conjugação.

— Começemos por estabelecer o seguinte importante resultado:

TEOREMA [Rockafellar, 1970 [f], p.105]

Seja $f(x)$ uma função convexa fechada; então $f^*(x)$ é a função supremo de pontos de todas as funções lineares por ela majoradas (3).

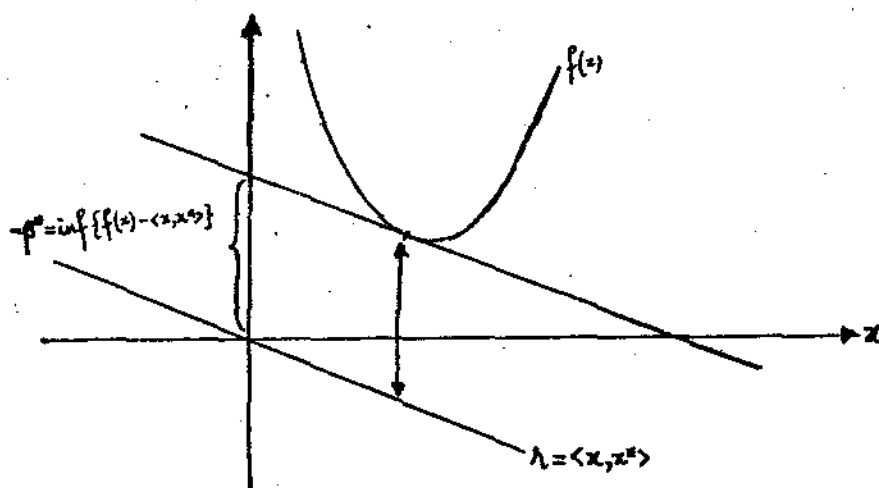
Ora, podemos representar genericamente uma função linear em $\mathbb{R}^n$ não vertical como:

$$r = \langle x^*, x \rangle - \beta$$

em que $(-\beta)$ é a ordenada na origem e x^* o vector de coeficientes.

Atendendo ao anterior teorema, o nosso problema é: dado x^* e dada a função convexa $f(x)$, como determinar a ordenada na origem $(-\beta)$ por forma a que a função linear $r = \langle x^*, x \rangle - \beta$ seja tangente a $f(x)$, isto é, por forma a que $H = \{(x, r) | r = \langle x^*, x \rangle - \beta\}$ defina a equação dum hiperplano de suporte ao epigrafo de $f(x)$?

(3) Se $f(x)$ não for convexa e/ou fechada, o supremo por pontos de todas as funções lineares por ela majoradas corresponde à "convexização fechada" da mesma, isto é, será uma função cujo epigrafo é a envolvente convexa fechada de $f(x)$.



A ilustração gráfica mostra-nos claramente a resposta, extensível a $\mathbb{R}^{n+1}$: basta atribuir à ordenada na origem o valor da ínfima distância entre a função $f(x)$ e a função linear que passa pela origem $r = \langle x^*, x \rangle$

$$-\beta^* = \inf_x \{f(x) - \langle x^*, x \rangle\}$$

$\Leftrightarrow$

$$\beta^* = \sup_x \{\langle x^*, x \rangle - f(x)\}$$

Como é obvio β^* vai depender do x^* fixado, pelo que podemos definir:

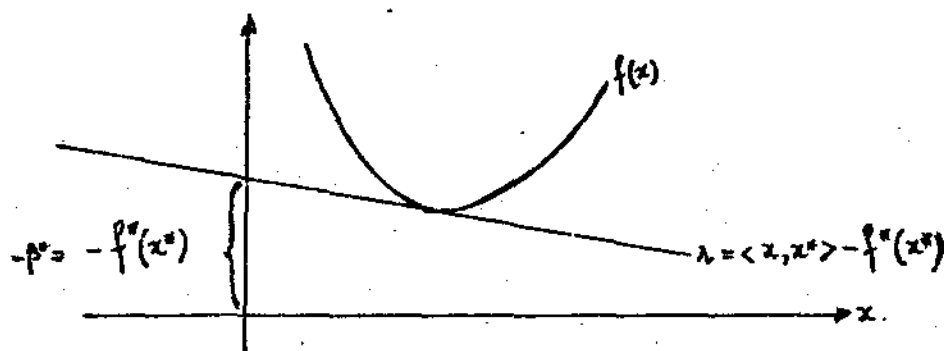
Definição: Seja $f(x)$ uma função convexa definida num conjunto convexo C . A função conjugada de $f(x)$ é a função

$$f^*(x^*) = \sup_{x \in C} \{\langle x^*, x \rangle - f(x)\}$$

cujo domínio efectivo é o conjunto de vectores tais que

$$f^*(x^*) < +\infty.$$

— $f^*(x^*)$ dá então, para cada x^* , o simétrico da maior ordenada na origem da função linear por forma a que esta seja ainda majorada por $f(x)$.



— O seguinte teorema sintetiza as principais propriedades (não marginais) da conjugada duma função convexa:

TEOREMA (Rockafellar, 1970, [f], p.104)

Seja $f(x)$ uma função convexa. A conjugada $f^*(x^*)$ é então uma função convexa fechada própria se $f(x)$ é própria. Além disso:

$$(cl f)^* = f^* \text{ e } f^{**} = cl f \quad (4)$$

2.1.2. Pontos conjugados/Propriedades marginais das funções conjugadas/transformação de Legendre

— Da definição:

$$f^*(x^*) = \sup_x \{ \langle x^*, x \rangle - f(x) \}$$

$\Leftrightarrow$

$$f^*(x^*) \geq \langle x^*, x \rangle - f(x), \quad \forall x$$

$\Leftrightarrow$

$$f(x) + f^*(x^*) \geq \langle x^*, x \rangle, \quad \forall x$$

Esta última desigualdade recebe o nome de desigualdade de Fenchel, e permite dar uma interpretação alternativa à relação de conjugação (5). Dois pontos $\bar{x}$ e $\bar{x}^*$ são referidos como pontos conjugados se

$$f(\bar{x}) + f^*(\bar{x}^*) = \langle \bar{x}^*, \bar{x} \rangle$$

isto é, $\bar{x}$ é conjugado de $\bar{x}^*$ se simultaneamente se atingem os supremos nas expressões seguintes:

$$f^*(\bar{x}^*) = \sup_x \{ \langle \bar{x}^*, x \rangle - f(x) \} = \langle \bar{x}^*, \bar{x} \rangle - f(\bar{x})$$

e

$$f(\bar{x}) = \sup_{x^*} \{ \langle x^*, \bar{x} \rangle - f^*(x^*) \} = \langle \bar{x}^*, \bar{x} \rangle - f^*(\bar{x}^*)$$

(4) Sendo f convexa, $cl f$ representa a maior função convexa fechada majorada por $f(x)$ (similarmente, para g côncava, $cl g$ representa a menor função côncava fechada minorada por g).

(5) Ver, por exemplo, Rockafellar, 1970 [f], p.104, ou Pinheiro, 1984, secção 2.1.4., ou ainda Pinheiro, 1985.

— Atendendo a estas últimas igualdades e às propriedades dos subgradiantes referidas na alínea 1.2.5., x e x^* serão pontos conjugados se

$$\begin{cases} 0 \in \partial_x [\langle x^*, x \rangle - f(x)] \\ 0 \in \partial_{x^*} [\langle x^*, x \rangle - f^*(x^*)] \end{cases} \Leftrightarrow (6) \quad \begin{cases} x^* \in \partial f(x) \\ x \in \partial f^*(x^*) \end{cases}$$

Sendo $f(x)$ e $f^*(x^*)$ diferenciáveis teremos para pontos conjugados:

$$\begin{cases} x^* = \nabla f(x) \\ x = \nabla f^*(x^*) \end{cases}$$

— Refira-se que a função conjugada de uma função convexa e diferenciável em $\mathbb{R}^n$ está fortemente relacionada com a sua transformação de Legendre. Isto porque, no caso de $x^* = \nabla f(x)$ definir uma função injectiva em x , poderemos obter a função inversa $x = (\nabla f)^{-1}(x^*)$ e logo, substituindo em $f^*(x^*) = \langle x^*, \bar{x} \rangle - f(\bar{x})$,

$$f^*(x^*) = \langle x^*, (\nabla f)^{-1}(x^*) \rangle - f((\nabla f)^{-1}(x^*)) \quad \dots (2.I)$$

O membro direito da anterior igualdade é exactamente a fórmula da transformação de Legendre da função f . No caso de x^* não definir uma correspondência biunívoca entre x^* e x , poderemos mesmo assim estabele

(6) Tomemos, por exemplo, o caso

$$0 \in \partial_x [\langle x^*, x \rangle - f(x)]$$

Sendo $f(x)$ por hipótese uma função convexa, a função (de x) $g(x) = \langle x^*, x \rangle - f(x)$ é uma função concáva, pelo que 0 será seu subgradiente no ponto x se:

$$g(w) \leq g(x) + \langle w - x, 0 \rangle$$

$$\Leftrightarrow g(w) \leq g(x)$$

$$\Leftrightarrow \langle x^*, w \rangle - f(w) \leq \langle x^*, x \rangle - f(x)$$

$$\Leftrightarrow f(w) \geq f(x) + \langle w - x, x^* \rangle$$

$$\Leftrightarrow x^* \in \partial f(x) \quad \text{c.q.d.}$$

cer uma formulação "fraca" da transformação de Legendre, também por substituição directa em $f^*(\bar{x}^*) = \langle \bar{x}^*, \bar{x} \rangle - f(x)$:

$$f^*(\nabla f(x)) = \langle \nabla f(x), x \rangle - f(x) \quad \dots(2.II)$$

Esta expressão, que é válida para todo o x onde a função f é diferenciável, ser-nos-á útil mais adiante.

2.1.3. Conjugadas de funções côncavas

— Um desenvolvimento similar ao efectuado nas alíneas anteriores po de ser feito para funções côncavas. Tenha-se, contudo, um especial cuidado no encara da analogia entre os dois tipos de funções, já que os resultados para as funções concâvas não resultam directa e exclusivamente da aplicação da simetria que faz com que, se g é côncava, $-g$ seja convexa. Existe, com efeito, também uma mudança de sinal na seguinte definição:

Definição: Seja $g(x)$ uma função côncava definida num conjunto convexo D . A função conjugada de $g(x)$ é

$$g^*(x^*) = \inf_{x \in D} \{ \langle x^*, x \rangle - g(x) \}$$

e tem por domínio o conjunto de pontos x^* tais que $g^*(x^*) > -\infty$.

— Restará referir que as propriedades das funções conjugadas de funções convexas mantêm-se para as conjugadas de funções côncavas (com a excepção da desigualdade de Fenchel que aparece invertida).

3. DUALIDADE CONJUGADA EM PROGRAMAÇÃO MATEMÁTICA CONVEXA

3.1 O PROBLEMA DE PROGRAMAÇÃO CONVEXA

3.1.1 O problema generalizado de programação convexa

— Seja o problema básico de programação convexa:

$$\begin{array}{ll} \min f(x) \\ \text{s.a. } \begin{cases} G(x) \leq 0 \\ x \geq 0 \end{cases} \end{array} \quad \dots(3.I)$$

com $x \in \mathbb{R}^n$, $f(x)$ uma função real convexa, e $G(x)$ um vector de m funções também reais e convexas (7). Associada com (3.I) é conveniente considerar a família de problemas

$$\begin{array}{ll} \min f(x) \\ \text{s.a. } \begin{cases} G(x) \leq u \\ x \geq 0 \end{cases} \end{array} \quad \dots(3.II)$$

Com u vector arbitrário de $\mathbb{R}^m$. Como é óbvio, a solução de (3.II), a que Rockafellar chama de programa generalizado de programação convexa, depende da perturbação u introduzida.

(7) O problema (3.I) não é restritivo na sua forma. Com efeito, basta ver que:

1) $y(x) \leq b \Leftrightarrow y(x) - b \leq 0$

2) $y(x) \geq b \Leftrightarrow -y(x) + b \leq 0$

3) $y(x) = b \Leftrightarrow \begin{cases} y(x) - b \leq 0 \\ -y(x) + b \leq 0 \end{cases}$

4) $x_j \leq 0 \Leftrightarrow v_j \geq 0$ com $v_j = -x_j$

5) x_j livre $\Leftrightarrow x_j = v_{j1} - v_{j2}$ com $v_{j1}, v_{j2} \geq 0$

3.1.2 Bifunção associada e função perturbação

— Definamos bifunção de $\mathbb{R}^m$ em $\mathbb{R}^n$ como a correspondência unívoca entre cada vector $u \in \mathbb{R}^m$ e uma função Fu de $\mathbb{R}^n$ em $[-\infty, +\infty]$. O valor de Fu para um determinado vector $x \in \mathbb{R}^n$ será designado por $(Fu)(x)$, assim como a função de $\mathbb{R}^{m+n}$ em $[-\infty, +\infty]$

$$f(u, x) = (Fu)(x)$$

será designada por função grafo de F . Muito simplesmente, uma bifunção não é mais que uma "função" que, em vez de gerar valores, como qualquer função ordinária, gera funções de $\mathbb{R}^n$ cada vez que lhe é fornecido um determinado vector $u \in \mathbb{R}^m$. O termo "bifunção" resulta portanto do facto de ela corresponder afinal a uma 1ª fase duma relação mais ampla de $\mathbb{R}^{m+n}$ em $[-\infty, +\infty]$ dada pela função grafo.

— Para os nossos objectivos de estudo interessa-nos uma bifunção especial, que consiga traduzir a família de problemas (3.II). Façamos para isso

$$(Fu)(x) = f(x) + \delta(x | x \geq 0; G(x) \leq u) = \begin{cases} f(x) & \text{se } x \geq 0 \text{ e } G(x) \leq u \\ +\infty & \text{caso contrário} \end{cases}$$

A esta bifunção daremos o nome de bifunção associada ao programa conexo generalizado (3.II). Repare-se que:

1º) $(Fu)(x)$ assim explicitada é uma bifunção pois cada concretização de u transforma $(Fu)(x)$ numa função de $\mathbb{R}^n$ em $[-\infty, +\infty]$.

2º) Dado u , resolver o problema (3.II) é equivalente a determinar $\inf_x (Fu)(x) = \inf_x [f(x) + \delta(x | x \geq 0; G(x) \leq u)]$

— Definamos então função perturbação como a função de $\mathbb{R}^n$ em $[-\infty, +\infty]$ que a cada vector u faz corresponder a solução óptima de (3.II) para esse vector de perturbações:

$$W(u) = \inf_x \{f(x) \mid x \geq 0; G(x) \leq u\} = \inf_x [f(x) + \delta(x \mid x \geq 0; G(x) \leq u)]$$

$$\Leftrightarrow W(u) = \inf_x (F_u)(x).$$

Sendo assim, o problema original (3.I) pode ser encarado como o problema de determinação de $W(0)$, ou equivalentemente, de determinação de $\inf_x (F_0)(x)$.

x

3.1.3 Algumas propriedades da bifunção associada e da função perturbação

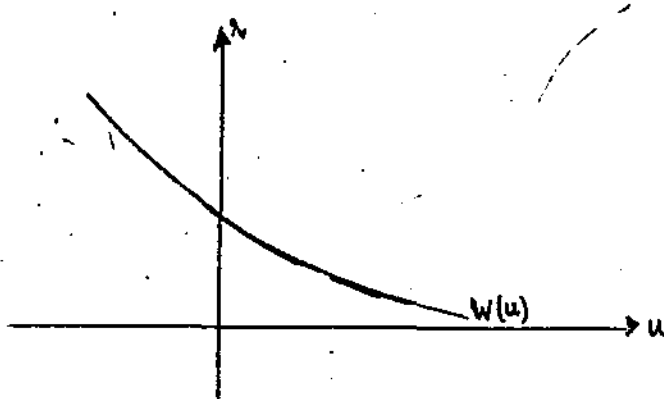
— Propriedade 1.1: A bifunção associada ao programa convexo generalizado (D) é convexa, em que se considera uma bifunção convexa se a respectiva função grafo o é. Ora, $f(u, x) = f(x) + \delta(x \mid x \geq 0; G(x) \leq u)$ é convexa dado serem convexas $f(x)$ por hipótese e $\delta(x \mid \cdot)$ por definição (a soma de funções convexas é ainda uma função convexa).

Propriedade 1.2: O programa convexo generalizado (3.II) é unicamente determinado pela respectiva bifunção associada (F_u) , isto é, podemos sempre recuperar o programa (3.II) a partir do conhecimento de F_u . Para tal tenha-se em atenção que, para um dado u , $S = \{x \mid (F_u)(x) < +\infty\}$ define o conjunto das soluções admissíveis, e que, para cada $x \in S$, temos $f(x) = (F_u)(x)$.

— Propriedade 2.1: A função perturbação $W(u)$ dum programa convexo é uma função convexa (demonstração, por exemplo, em Luenberger, 1969, pág. 216).

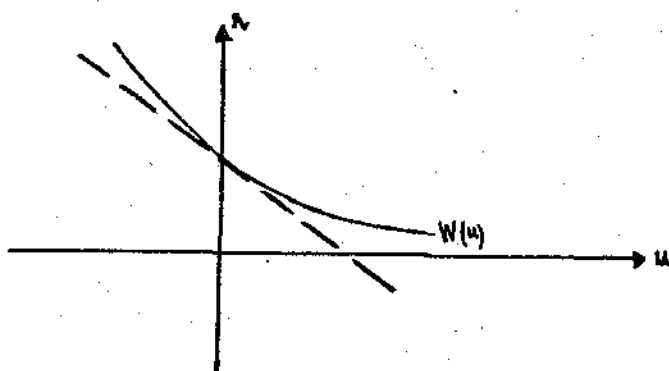
Propriedade 2.2: A função $W(u)$ é não crescente, isto é, se $u_1 \leq u_2$ então $W(u_1) \geq W(u_2)$. Esta propriedade é imediata, atendendo a que restringindo o conjunto das soluções admissíveis (isto é, diminuindo as perturbações u), nunca se poderá melhorar a solução óptima do problema.

— Atendendo às propriedades 2.1 e 2.2, uma representação gráfica típica de $W(u)$, para $u \in \mathbb{R}^1$, será:



3.1.4 Relação entre função auxiliar Lagrangeana e função perturbação

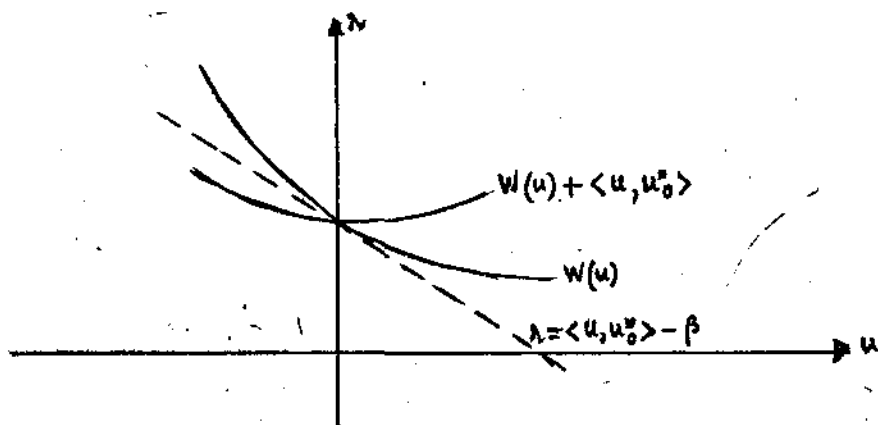
— Sendo $W(u)$ convexa (propriedade 2.1) haverá necessariamente um hiperplano tangente a $w(u)$ para $u=0$, hiperplano este que será majorado pela função perturbação em todo o seu domínio de definição.



Se rodássemos a cabeça por forma a que o hiperplano tangente se tornasse no eixo horizontal, teríamos então $W(u)$ minimizada para $u=0$. Dito de outra forma, se somarmos a $W(u)$ uma função linear apropriada $\langle u, u_0^* \rangle$, a combinação resultante $W(u) + \langle u, u_0^* \rangle$ é minimizada para $u=0$, e teremos

$$\inf\{W(u) + \langle u, u_0^* \rangle\} = W(0) + \langle 0, u_0^* \rangle = W(0) \rightarrow \text{sol. óptima pro blema (3.1).}$$

Ilustração em $\mathbb{R}^2$:



Mas:

$$\begin{aligned} \inf_u [W(u) + \langle u, u^* \rangle] &= \inf_u \left[\overbrace{\inf_x \{f(x) \mid G(x) \leq u, x \geq 0\}}^{W(u)} + \langle u, u^* \rangle \right] = \\ &= \inf_u \left[\inf_x \{f(x) + \langle u, u^* \rangle \mid G(x) \leq u, x \geq 0\} \right] = \\ &= \inf_{x \geq 0} [f(x) + \langle G(x), u^* \rangle] = \inf_{x \geq 0} L(x, u^*) \end{aligned}$$

isto é, ao minimizarmos, conforme pretendido, $W(u) + \langle u, u^* \rangle$, estamos afinal a minimizar a função auxiliar Lagrangeana $L(x, u^*) = f(x) + \langle G(x), u^* \rangle$, u^* não ortante não negativo. Daqui resulta que o vector u_0^* que faz com que $W(u) + \langle u, u^* \rangle$ atinja o mínimo para $u=0$, seja afinal o vector de multiplicadores de Lagrange associados à solução óptima do problema (3.1). Nas alíneas 3.2.5. e 3.2.6. reobteremos este importante resultado, mas já integrado na teoria geral da dualidade convexa, que passamos a apresentar.

3.2 CONJUGAÇÃO E DUALIDADE EM PROGRAMAÇÃO CONVEXA

" Um facto fundamental sobre programas convexos generalizados é que estes "problemas de minimização com perturbações" têm um dual, que é um certo "programa de maximização com perturbações", um programa côncavo generalizado. Na maioria dos casos, os dois programas duais têm o mesmo valor óptimo da função objectivo e as soluções óptimas para um são os multiplicadores de Lagrange para o outro" (Rockafellar, [f], 1970, pág.307).

3.2.1 Bifunções adjuntas

— A teoria da dualidade em programação convexa pode ser desenvolvida a partir do conceito de "adjunta" de uma bifunção. Como veremos na alínea 3.2.4, existe uma relação estreita, donde resulta a designação, entre adjuntas de bifunções e transformações lineares adjuntas.

— Definimos adjunta duma bifunção convexa $(Fu)(x)$ de $\mathbb{R}^m$ em $\mathbb{R}^n$ como a bifunção $(F^*x^*)(u^*)$ de $\mathbb{R}^n$ em $\mathbb{R}^m$ dada por

$$(F^*x^*)(u^*) = \inf_{x, u} \{ (Fu)(x) - \langle x, x^* \rangle + \langle u, u^* \rangle \}$$

A adjunta de uma bifunção côncava é definida da mesma maneira, excepto que o ínfimo é substituído por supremo. Sublinhe-se que a correspondência entre uma bifunção e a respectiva bifunção adjunta não é mais que uma pequena modificação da correspondência entre funções conjugadas.

Com efeito, seja $f(u, x)$ a função grafo da bifunção convexa $(Fu)(x)$.

Encarando

$$\langle x, x^* \rangle - \langle u, u^* \rangle = \langle (u, x); (-u^*, x^*) \rangle$$

como o produto escalar dos vectores (u, x) e $(-u^*, x^*)$ de $\mathbb{R}^{m+n}$, temos

$$\begin{aligned} (F^*x^*)(u^*) &= \inf_{x, u} \{ f(u, x) - \langle (u, x); (-u^*, x^*) \rangle \} = \\ &= - \sup_{x, u} \{ \langle (u, x); (-u^*, x^*) \rangle - f(u, x) \} = \\ &= - f^*(-u^*, x^*) \end{aligned}$$

em que f^* é a conjugada de f ⁽⁸⁾. Como f é convexa por hipótese,

$f^*(-u^*, x^*)$ é, pelo que sabemos, uma função convexa fechada, e lo-

go a bifunção adjunta de F

$$(F^*x^*)(u^*) = - f^*(-u^*, x^*)$$

será côncava e também fechada.

— Recorrendo às funções grafo, consegue-se mostrar sem dificuldade que $F^{**} = \text{cl}F$:

$$\begin{aligned} F^{**} &= \sup_{x^*, u^*} \{ (F^*x^*)(u^*) - \langle u^*, u \rangle + \langle x^*, x \rangle \} = \\ &= \sup_{x^*, u^*} \{ \langle (-u^*, x^*), (u, x) \rangle - f^*(-u^*, x^*) \} \\ &= f^{**} = (\text{cl}f)(u, x) = \text{cl}F \end{aligned}$$

— Os anteriores resultados podem ser sintetizados sob a forma de teorema:

(8) Repare-se que a correspondência entre uma bifunção e a sua adjunta não é exactamente a relação entre 2 funções conjugadas apenas porque há uma troca de sinal.

TEOREMA 3.B (Rockafellar [f], 1970, p.310)

Seja F uma qualquer bifunção convexa ou concava de $\mathbb{R}^m$ em $\mathbb{R}^n$.
então a sua bifunção adjunta F^* é uma bifunção fechada do tipo
oposto, definida de $\mathbb{R}^n$ em $\mathbb{R}^m$, própria se F é própria, e F^{**}
= $\text{cl}F$. Em particular, $F^{**}=F$ se F é fechada.

3.2.2 Programas duais

— Seja F uma bifunção convexa de $\mathbb{R}^m$ em $\mathbb{R}^n$, seja (P) o programa convexo generalizado a ela associado. O programa côncavo generalizado associado com a bifunção adjunta F^* (côncava) será designado por programa dual de (P) , e será representado por (P^*) . Representamos também por (P_0) e (P_0^*) os correspondentes problemas não generalizados.

— Note-se que, se F é fechada, o dual do dual do programa convexo generalizado, que é o programa associado à bifunção F^{**} , é precisamente o programa primal donde partiramos, já que $F^{**}=\text{cl}F$. Note-se ainda que, enquanto em (P) minimizamos $(Fu)(x)$ em $x \in \mathbb{R}^n$, em (P^*) maximizamos $(F^*x^*)(u^*)$ em $u^* \in \mathbb{R}^m$. Sendo assim, enquanto o valor ótimo do programa convexo generalizado (P_0) é, como sabemos, $W(0) = \inf_x (F0)(x)$, o valor ótimo do seu problema dual será $\sup_{x^*} (F^*0)(u^*)$.

3.2.3 Alguns resultados de dualidade

— Antes de esclarecermos as potencialidades e a generalidade da teoria da dualidade conjugada convexa atrás delineada, comecemos por estabelecer algumas comparações e relações entre os valores óptimos dos problemas duais.

— Pelo exposto nas secções 3.1.2 e 3.2.1, dados x, u, x^* e u^* :

$$(FU)(x) = f(u, x)$$

$$(F^*x^*)(u^*) = -f^*(-u^*, x^*)$$

em que $f(u, x)$ é a função grafo de $(FU)(x)$. Ora, pela desigualdade de Fenchel (secção 2.1.2) sabemos que

$$f(u, x) + f^*(-u^*, x^*) \geq \langle (u, x), (-u^*, x^*) \rangle, \quad \forall (u, x), (-u^*, x^*) \in \mathbb{R}^{m+n}$$

$$\text{donde: } (FU)(x) - (F^*x^*)(u^*) \geq \langle (u, x), (-u^*, x^*) \rangle, \quad \forall (u, x), (-u^*, x^*) \in \mathbb{R}^{m+n}$$

Em particular, fazendo $u=x^*=0$, vem :

$$(FO)(x) \geq (F^*O)(u^*), \quad \forall x \in \mathbb{R}^n, u^* \in \mathbb{R}^m$$

pelo que podemos escrever:

$$\inf_x (FO)(x) \geq \sup_{u^*} (F^*O)(u^*)$$

Os seguintes teorema e corolário derivam directamente do anterior resultado:

TEOREMA 3.C - TEOREMA DE DUALIDADE FRACA (AVRIEL, 1976, teor.5.5, p.114)

. O valor óptimo do problema primal convexo não generalizado nunca será inferior ao valor óptimo do respectivo problema dual côncavo não generalizado (P^*) , isto é,

$$\inf_x (FO)(x) \geq \sup_{u^*} (F^*O)(u^*)$$

COROLÁRIO 3.C.1. (AVRIEL, 1976, cor.5.6, p.114)

. Se $(FO)(\bar{x}) = (F^*O)(\bar{u}^*)$, então $\bar{x}$ e $\bar{u}^*$ são soluções óptimas respectivamente de (P_0) e (P^*) .

— O problema convexo não generalizado (P_0) é dito consistente se existe algum x tal que $(F_0)(x) < +\infty$. Análogamente, o problema concavo dual não generalizado (P_0^*) é consistente se existir algum u^* tal que $(F^*O)(u^*) > -\infty$. No caso de não serem consistentes, estes problemas dizem-se inconsistentes (9). Estabelecida esta terminologia, podemos obter o seguinte resultado a partir do teorema 3.C:

COROLÁRIO 3.0.2 (AVRIEL, 1976, cor. 5.7, p. 114)

. Se $\inf_x (F_0)(x) = -\infty$ então (P_0^*) é inconsistente. Se $\sup_{u^*} (F^*O)(u^*) = +\infty$ então (P_0) é inconsistente.

— O teorema 3.C (completado pelos seus corolários) permite-nos nomeadamente afirmar que o valor óptimo do problema (P_0) (respectivamente (P_0^*)) serve de limite superior (resp. inferior) ao valor óptimo do seu problema dual (P_0^*) (resp. (P_0)), e que quando as duas soluções dos dois problemas têm o mesmo valor, elas são as soluções óptimas. De seguida vamos ver a questão de certa forma inversa e mais "forte": em que condições é que podemos ter a garantia que os valores óptimos de (P_0) e (P_0^*) são coincidentes? Os resultados a seguir enunciados permitem-nos chegar à conclusão que tal acontecerá (em programação convexa) desde que respeitadas adequadas hipóteses de regularidade (10).

— Defina-se o problema convexo generalizado (P) como estável se: (i) ou a sua função perturbação é finita para $u = 0$ ($W(0)$ finito) e não existe nenhuma direcção y tal que $W'(0, y) = -\infty$; (ii) ou $W(0) = -\infty$.

(9) Repare-se que será inconsistente um problema impossível (conjunto soluções admissíveis vazio), bem como um problema para o qual as únicas soluções admissíveis não pertençam ao domínio efectivo da função objectivo.

(10) Em programação não convexa, ou em programação convexa que não respeite as condições de regularidade, poderemos ter uma discrepância entre os valores óptimos dos problemas duais, que recebe a designação de "duality gap".

Por outras palavras, (P) é estável se ou o valor óptimo do problema não generalizado (Po) é finito e a função perturbação $W(u)$ não é infinitamente inclinada na vizinhança de 0, ou então se o valor óptimo de (Po) é $-\infty$. Utilizando a noção de estabilidade é possível estabelecer condições necessárias e suficientes para a igualdade dos valores óptimos de (Po) e (P*o), consubstanciadas no teorema 3.D, que designaremos por teorema de dualidade forte:

TEOREMA 3.D - TEOREMA DE DUALIDADE FORTE (AVRIEL, 1976, parte do teor. 5.11, p.119)

Seja $f(u,x)$ (função grafo de $(Fu)(x)$) uma função convexa própria em $\mathbb{R}^{m+n}$, e seja $W(0)$ finito. Então o programa (P*) tem uma solução óptima e

$$W(0) = \inf_x (FO)(x) = \sup_{u^*} (F^*O)(u^*)$$

sse (P) é estável.

— Os seguintes teorema e corolário exemplificam outro tipo de resultados de dualidade que se podem obter na sequência do teorema de dualidade forte exigindo condições de regularidade diferentes (11).

(11) Diferentes na medida em que poderão haver problemas instáveis que as verificam [cf. AVRIEL, 1976, ex.5.2.2., p.121].

TEOREMA 3.E (AVRIEL, 1976, teor. 5.18, p. 125)

- . Seja $f(u, x)$ (função grafo de $(FU)(x)$) uma função convexa própria e fechada em $\mathbb{R}^{m+n}$.
- . Então são equivalentes as três seguintes proposições:
 - (i) $\inf_x (FO)(x) = \sup_{u^*} (F^*O)(u^*)$;
 - (ii) a função de perturbação do problema primal convexo (P) é fechada em $u = 0$;
 - (iii) a função de perturbação do problema dual concavo (P^*) é fechada em $x^* = 0$.

Atribuíamos a designação de normal aos problemas (P) e (P^*) se se verificam as condições (ii) e (iii) do teorema (3.E). Usando esta definição é possível estabelecer o seguinte corolário:

COROLÁRIO (AVRIEL, 1976, cor. 5.19, p. 125)

- . Sejam as condições do teorema 3.E. O problema (P) é normal e o seu valor ótimo é finito sse (P^*) é normal e o seu valor ótimo é finito. Inversamente, se os valores ótimos dos problemas não generalizados (Po) e (P^*o) são iguais, então tanto (P) como (P^*) são normais.

3.2.4. bifunções adjuntas e transformações lineares adjuntas/ dual dum programa linear

— Como apontámos no início da alínea 3.2.1., existe uma estreita relação entre as noções de adjunta duma bifunção e a de transformação linear adjunta. Clarifiquemos melhor esta relação antes de prosseguirmos o nosso estudo da dualidade conjugada.

— Consideremos F como a bifunção convexa

$$(FU)(x) = \delta(x | x = Au) = \begin{cases} 0 & \text{se } x = Au \\ +\infty & \text{se } x \neq Au \end{cases}$$

que designaremos por bifunção indicadora da transformação linear A de $\mathbb{R}^m$ em $\mathbb{R}^n$. Calculando F^* pela definição da bifunção adjunta obtemos directamente:

$$(F^*x^*)(u^*) = \inf_{x, u} \{ \delta(x | x = Au) - \langle x, x^* \rangle + \langle u, u^* \rangle \} =$$

(substituindo x por Au)

$$= \inf_u \{ -\langle Au, x^* \rangle + \langle u, u^* \rangle \} =$$

(pela definição de matriz adjunta $A^*(12)$):

$$= \inf_u \{ \langle u, u^* - A^*x^* \rangle \} =$$

$$= \begin{cases} 0 & \text{se } u^* = A^*x^* \\ -\infty & \text{se } u^* \neq A^*x^* \end{cases}$$

$$= -\delta(u^* | u^* = A^*x^*)$$

Quer isto dizer que a adjunta duma bifunção indicadora duma transformação linear A corresponde ao simétrico da bifunção indicadora da respectiva transformação adjunta A^* , confirmando-se, portanto, a estreita relação entre adjunta duma bifunção e de adjunta duma transformação linear.

— Uma aplicação desta mesma relação acabada de mostrar pode ser estabelecida para o caso particular e clássico da programação linear, o qual servirá também como exercício simples de construção da bifunção adjunta e do programa

(12) Chama-se adjunta duma transformação linear A , e representa-se por A^* , a transformação (também linear) definida por:

$$\langle Az, v \rangle = \langle z, A^*v \rangle, \quad \forall z, v.$$

dual. Tome-se o programa linear

$$\begin{aligned} & \inf_x \langle a^*, x \rangle \\ \text{s.a. } & \begin{cases} a - Ax \leq 0 \\ x \geq 0 \end{cases} \end{aligned} \quad \dots (3.III)$$

a generalização de (3.III) será:

$$\begin{aligned} & \inf_x \langle a^*, x \rangle \\ \text{s.a. } & \begin{cases} a - Ax \leq u \\ x \geq 0 \end{cases} \end{aligned} \quad \dots (3.IV)$$

Como bifunção associada a (3.IV) teremos:

$$(F^*U)(x) = \langle a^*, x \rangle + \delta(x | x \geq 0 ; a - Ax \leq u)$$

Calculando pela definição a adjunta F^* , vem:

$$\begin{aligned} (F^*x^*)(u^*) &= \inf_{u,x} \{ \langle a^*, x \rangle + \delta(x | x \geq 0 ; a - Ax \leq u) - \langle x, x^* \rangle + \langle u, u^* \rangle \} = \\ &= \inf_{v \geq 0, x \geq 0} \{ \langle x, a^* - x^* \rangle + \underbrace{\langle a - Ax + v, u^* \rangle}_u \} \text{ (com } v \text{ variável-desvio)} \\ &= \inf_{v \geq 0, x \geq 0} \{ \langle a, u^* \rangle + \langle v, u^* \rangle + \langle x, a^* - x^* - A^* u^* \rangle \} = \\ &= \langle a, u^* \rangle + \inf_{v \geq 0} \{ \langle v, u^* \rangle \} + \inf_{x \geq 0} \{ \langle x, a^* - x^* - A^* u^* \rangle \} = \\ &= \begin{cases} \langle a, u^* \rangle & \text{se } u^* \geq 0 \text{ e } a^* - x^* - A^* u^* \geq 0 \\ -\infty & \text{caso contrário} \end{cases} \\ &= \langle a, u^* \rangle - \delta(u^* | u^* \geq 0 ; a^* - A^* u^* \geq x^*) \end{aligned}$$

O problema dual de (3.III) será então o problema de maximização de $(F^*0)(u^*)$, isto é:

$$\begin{aligned} & \sup_{u^*} \langle a, u^* \rangle \\ \text{s.a. } & \begin{cases} a^* - A^* u^* \geq 0 \\ u^* \geq 0 \end{cases} \end{aligned}$$

com A^* matriz adjunta de A .

3.2.5. Uma perspectiva gráfica da dualidade conjugada

— Deixando o caso simples da programação linear, e voltando ao caso geral não linear da programação convexa, é particularmente importante, para uma perspectiva gráfica da dualidade, o seguinte teorema, que estabelece a dualidade em termos de funções perturbação dos problemas primal e dual generalizados, e que decorre facilmente da definição de bifunções adjuntas:

TEOREMA 3.F (Rockafellar, [f], 1970, pág.314)

- Seja F uma bifunção convexa de $\mathbb{R}^m$ em $\mathbb{R}^n$ e seja F^* a sua bifunção côncava adjunta de $\mathbb{R}^n$ em $\mathbb{R}^m$. Temos então, por um lado, que

$$[-W(u)]^* = [-\inf_x (Fu)(x)]^* = (F^*O)(u^*) ; \quad (F^*O)^* = -cl[W(u)]$$

ou seja, a conjugada do simétrico da função perturbação do problema primal é precisamente o programa dual não generalizado.

- Por outro lado, e representando por $\bar{W}(x^*) = \sup_{u^*} (F^*x^*)(u^*)$ a função perturbação do problema dual generalizado, temos também que

$$[-\bar{W}(x^*)]^* = [-\sup_{u^*} (F^*x^*)(u^*)]^* = (FO)(x) ; \quad (FO)^* = -cl[\bar{W}(x^*)]$$

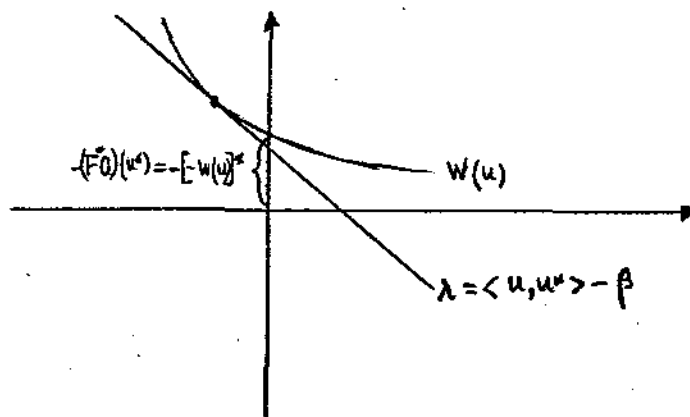
— Consideremos apenas o primeiro resultado estabelecido no teorema:

$$(F^*O)(u^*) = [-W(u)]^*$$

Conforme apontamos no ponto 2, a conjugada de $-W(u)$ dar-nos-á, para cada u^* previamente fixado, o simétrico da menor ordenada na origem possível das funções afins $r = \langle u, u^* \rangle - \beta$ por forma a que elas sejam minoradas por $-W(u)$. Como é fácil de perceber, tal equivale a dizer que nos dá, para cada u^* previamente fixado, a maior ordenada na origem possível das funções afins $r = -\langle u^*, u \rangle - \beta$ por forma a que elas

sejam majoradas por $W(u)$.

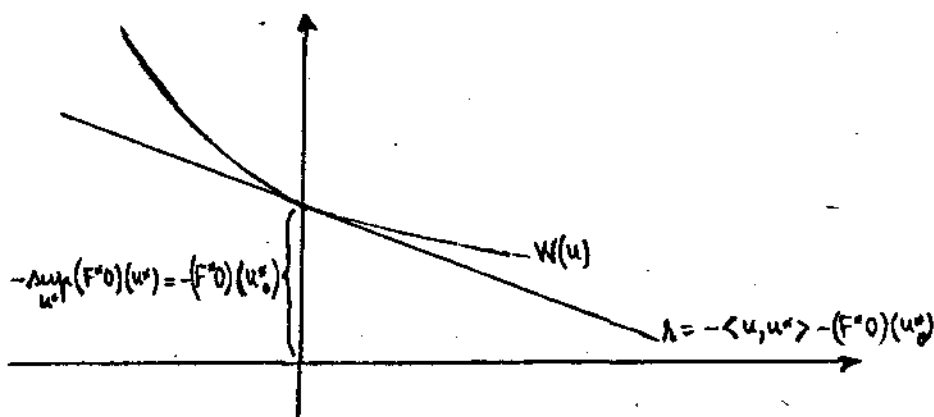
Ilustração em $\mathbb{R}^2$ para um dado u^* previamente fixado:



Logo, $(F^*O)(u^*)$ dá-nos, para cada u^* , o máximo da ordenada na origem das funções lineares afins $r = -\langle u, u^* \rangle - \beta$ majorados por $W(u)$.

— A maximização de $(F^*O)(u^*)$, que é o programa côncavo dual do programa convexo (Po) , traduz portanto o problema de maximização das ordenadas na origem das funções afins tangentes a $W(u)$ (ordenadas na origem estas que são, por sua vez, e como vimos atrás, as máximas possíveis por forma a que as famílias de funções afins paralelas sejam majoradas por $W(u)$). Sendo assim, é imediato que a solução ótima do problema dual vai ser precisamente o vector u_0^* tal que $r = -\langle u, u_0^* \rangle - (F^*O)(u_0^*)$ defina uma função afim tangente a $W(u)$ para $u=0$.

Ilustração em $\mathbb{R}^2$:



Tendo em conta o que referimos na alínea 3.1.4 a propósito dos multiplicadores de Lagrange, conclui-se do que acabámos de ver que a solução óptima do problema dual não é mais que o vector dos valores no óptimo dos multiplicadores de Lagrange do problema primal.

3.2.6 Relação entre dualidade conjugada e teoria de Lagrange / Kuhn-Tucker.

— Consideremos novamente o 1º resultado do teorema (3.F):

$$(F^*O)(u^*) = [-W(u)]^* \quad \dots (3.V)$$

Desenvolvendo pela definição de conjugada, e tomando em atenção que $-W(u)$ é uma função côncava, vem:

$$(F^*O)(u^*) = \inf_x \{ \langle u, u^* \rangle - [-W(u)] \} = \inf_x \{ W(u) + \langle u, u^* \rangle \} =$$

(conforme demonstrado na alínea 3.1.4)

$$= \inf_{x \geq 0} \{ f(x) + \langle G(x), u^* \rangle \}, \quad \forall u^* \geq 0$$

$$= \inf_{x \geq 0} L(x, u^*), \quad \forall u^* \geq 0 \quad (13)$$

Quer isto dizer que o programa dual corresponde afinal à maximização relativamente a u^* do mínimo da função de Lagrange relativamente a x , isto é:

(13) - Alternativamente, este mesmo resultado pode ser obtido recorrendo à determinação da bifunção adjunta do problema convexo generalizado primal. Com efeito, (Rockafellar, [f], 1970, p.320) mostra que:

$$(F^*x^*)(u^*) = \begin{cases} -\sup_{x \geq 0} \{ \langle x, x^* \rangle - [f(x) + \langle G(x), u^* \rangle] \} & \text{se } u^* \geq 0 \\ -\infty & \text{caso contrário} \end{cases}$$

$$(F^*O)(u^*) = \begin{cases} \inf_{x \geq 0} \{ f(x) + \langle G(x), u^* \rangle \} & \text{se } u^* \geq 0 \\ -\infty & \text{caso contrário} \end{cases} = \inf_{x \geq 0} L(x, u^*), \quad \forall u^* \geq 0 \quad \text{c.q.d.}$$

$$\sup_{u^* \geq 0} (F^*O)(u^*) = \sup_{u^* \geq 0} \inf_{x \geq 0} L(x, u^*) \quad \dots (3.VI)$$

Dito de outra forma, o óptimo do problema dual não generalizado corresponde ao ponto de sela no ortante positivo de $\mathbb{R}^{m+n}$ de Lagrangeana do problema primal. Estabelece-se assim uma relação estreita entre a dualidade conjugada e a dualidade Lagrangeana, justificando a conclusão gráfica do fim da alínea anterior respeitante à identidade entre a solução óptima do problema dual e o vector dos valores no óptimo dos multiplicadores de Lagrange do problema primal.

— Refira-se, a propósito, que a dimensão de preços-sombra atribuída aos multiplicadores de Lagrange, e logo à solução óptima do problema dual não generalizado (P^*O), se pode demonstrar facilmente a tendendo às propriedades marginais dos pontos conjugados:

$$u^* \in \delta[-W(u)] \text{ sse } u \in \delta[-W(u)]^*$$

Fazendo $u = 0$ e tendo em conta que (expressão 3.V) $[-W(u)]^* = (F^*O)(u^*)$, fica:

$$u^* \in \delta[-W(0)] \text{ sse } 0 \in \delta[(F^*O)(u^*)]$$

Se $W(u)$ e (F^*O) forem diferenciáveis respectivamente em 0 e u^* , teremos:

$$u^* = - \nabla_u W(0) \iff \nabla_{u^*} (F^*O)(u^*) = 0$$

ou seja:

$$u^* = - \nabla_u [\inf_x (FO)(x)] \iff \nabla_{u^*} (F^*O)(u^*) = 0$$

... (3.VII)

Como $(F^*O)(u^*)$ é uma função côncava, $V_{u^*}(F^*O)(u^*) = 0$ é condição necessária e suficiente para u^* ser ótimo do problema dual não generalizado. Portanto, pela expressão (3.VII), u^* ser solução ótima do problema dual não generalizado equivale a u^* ser igual a $-V_u[\inf_x (FO)(x)]$, isto é, equivale a traduzir, como queríamos demonstrar, o simétrico da sensibilidade ⁽¹⁴⁾ do valor ótimo do problema primal não generalizado (P_o) a variações infinitesimais do vector de perturbações ⁽¹⁵⁾.

— Mas é possível ir ainda mais longe na relação entre teoria de Kuhn-Tucker e a dualidade conjugada: os principais resultados daquela podem ser estabelecidos sem dificuldade recorrendo a esta última, para problemas de programação convexa. Com efeito, obtivemos mais atrás (expressão 3.VI) que:

$$\left[\begin{array}{c} \text{valor ótimo prob.dual} \\ \text{não generalizado} \\ (P^*o) \end{array} \right] = \sup_{u^* \geq 0} \inf_{x \geq 0} L(x, u^*) \quad \dots (3.VIII)$$

Também sabemos que, verificada determinada condição de regularidade, então (cf. alínea 3.2.3):

$$[\text{valor ótimo } (P_o)] = [\text{valor ótimo } (P^*o)] \quad \dots (3.IX)$$

Reunindo (3.VIII) e (3.IX) obtemos:

$$[\text{valor ótimo } (P_o)] = \sup_{u^* \geq 0} \inf_{x \geq 0} L(x, u^*)$$

Logo, as condições necessárias de 1ª ordem para a verificação do ponto

(14) Simétrico porque definimos a função auxiliar Lagrangeana como

$L = f(x) + \langle u^*, G(x) \rangle$ (cf. alínea 3.1.4).

(15) Para uma visualização geométrica desta interpretação dos multiplicadores de Lagrange, atente-se na relação entre função auxiliar Lagrangeana e função de perturbação que apresentamos na alínea 3.1.4. Com efeito, pelo que então concluímos, os multiplicadores de Lagrange definem o hiperplano tangente a $W(u)$ para $u = 0$.

de sela no ortante positivo da função Lagrangeana,

$$\left\{ \begin{array}{l} \nabla_x L(x, u^*) \geq 0 ; x \geq 0 ; \langle x, \nabla_x L(x, u^*) \rangle = 0 \\ \nabla_{u^*} L(x, u^*) \leq 0 ; u^* \geq 0 ; \langle u^*, \nabla_{u^*} L(x, u^*) \rangle = 0 \end{array} \right. \dots (3.X)$$

serão condições necessárias de óptimo do problema primal não generalizado (Po) (desde que cumpridas a tal condições de regularidade). Mais ainda: como $(F_0)(x)$ é uma função convexa, as condições necessárias são igualmente suficientes, simultaneamente para ponto de sela da Lagrangeana e para óptimo do problema convexo não generalizado (Po). Ora, as condições (3.X) são as bem conhecidas condições de Kuhn-Tucker, pelo que se mostrou que elas são condições necessárias e suficientes de óptimo em programação convexa se respeitada determinada condição de regularidade que garanta a igualdade (3.IX), por exemplo a estabilidade de (P) (16).

3.2.7. Dualidade conjugada e dualidade de Wolfe (17)

— O carácter geral e englobante da teoria da dualidade conjugada não se circunscreve à dualidade linear e à dualidade lagrangeana, pois também a usualmente designada dualidade de Wolfe, formulada por este autor em 1961 (18) sem recorrer à noção de conjugação, pode ser obtida a partir da abordagem que temos vindo a desenvolver na hipótese de diferenciabilidade da função objectivo e das restrições.

(16) Pelo que todas as diferentes condições de qualificação da teoria de Kuhn-Tucker, implicam em programação convexa a verificação da igualdade (3.IX).

(17) Baseado em Rockafellar, [d], 1968, p.413. O assunto está igualmente tratado em Avriel, 1976, p.137-139.

(18) Wolfe, P. (1961), "A duality theorem for Nonlinear Programming", Quart. Appl. Math., 19, p.239-244.

— Sejam os problemas convexos generalizados (P) e não generalizado (Po). Para maior simplicidade de exposição, deixemos de considerar em qualquer deles as restrições de sinal $x \geq 0$ (19). A função de perturbação associada ao problema convexo generalizado fica então a seguinte:

$$(Fu)(x) = f(x) + \delta(x | G(x) \leq u)$$

Começemos por calcular a sua bifunção adjunta:

$$\begin{aligned} (F^*x^*)(u^*) &= \inf_{u, x} \{ (Fu)(x) - \langle x, x^* \rangle + \langle u, u^* \rangle \} = \\ &= \inf_{u, x} \{ f(x) + \delta(x | G(x) \leq u) - \langle x, x^* \rangle + \langle u, u^* \rangle \} = \\ &= \inf_{\substack{u \geq 0 \\ u \leq 0}} \{ f(x) + \delta(x | G(x) + u = u) - \langle x, x^* \rangle + \langle u, u^* \rangle \} = \\ &= \inf_{\substack{x \\ v \geq 0}} \{ f(x) - \langle x, x^* \rangle + \langle G(x) + v, u^* \rangle \} = \\ &= \inf_x \{ f(x) - \langle x, x^* \rangle + \langle G(x), u^* \rangle \} + \underbrace{\inf_{v \geq 0} \langle v, u^* \rangle}_{\begin{cases} 0 & \text{se } u^* \geq 0 \\ -\infty & \text{caso cont.} \end{cases}} \\ &= \begin{cases} -\sup_x \{ \langle x, x^* \rangle - [f(x) + \langle G(x), u^* \rangle] \} & \text{se } u^* \geq 0 \\ -\infty & \text{caso contrário} \end{cases} \\ &= \begin{cases} -[f(x^*) + \langle G(x^*), u^* \rangle] * (x^*) & \text{se } u^* \geq 0 \\ -\infty & \text{caso contrário} \end{cases} \end{aligned}$$

Quer isto dizer que o problema dual côncavo generalizado consiste na maximização em u^* de $-[f(x^*) + \langle G(x^*), u^* \rangle] * (x^*)$ s.a $u^* \geq 0$. O problema dual não generalizado obter-se-á do problema generalizado fazendo $x^* = 0$:

$$\sup_x - (f(x) + \langle u^*, G(x) \rangle) * (0) \quad \dots (3.XI)$$

s.a $u^* \geq 0$

(19) O que, de qualquer forma, não é restritivo. Para que elas continuem a ser consideradas bastará $x \geq 0 \Leftrightarrow -x \leq 0$ nas restantes restrições $G(x) \leq 0$, associando também a cada uma delas uma perturbação u_i .

Pela expressão (2.II) da alínea 2.1.2. temos, para qualquer função convexa diferenciável f , que $f^*(\nabla f(x)) = \langle \nabla f(x), x \rangle - f(x)$ o que implica que

$$-f^*(0) = f(x) \text{ com } \nabla f(x) = 0$$

Aplicando este resultado à função, convexa em x , $f(x) + \langle u^*, G(x) \rangle$, o problema (3.XI) fica:

$$\left\{ \begin{array}{l} \sup_{u^*} f(x) + \langle u^*, G(x) \rangle \\ \text{s.a. } \begin{cases} u^* \geq 0 \\ \nabla f(x) + \langle u^*, \nabla G(x) \rangle = 0 \end{cases} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \sup_{u^*} f(x) + u_1^* g_1(x) + \dots + u_m^* g_m(x) \\ \text{s.a. } \begin{cases} u_i^* \geq 0 \quad (i=1, \dots, m) \\ \nabla f(x) + u_1^* \nabla g_1(x) + \dots + u_m^* \nabla g_m(x) = 0 \end{cases} \end{array} \right.$$

Este problema é precisamente o conhecido problema dual de Wolfe do problema convexo não generalizado $\min f(x)$ s.a. $G(x) \leq 0$. Utilizando a expressão de Rockaféllar ([d], 1968, p.413), o problema dual de Wolfe é portanto "quase" equivalente ao problema dual não generalizado (3.XI), não sendo "completamente" equivalente na medida em que exige a diferenciabilidade das funções objectivo e restrições (por isso o problema dual de Wolfe só é definido para o subconjunto de $(\text{dom } f \cap \text{dom } g_1 \cap \dots \cap \text{dom } g_m)$ onde existam $\nabla f(x)$, $\nabla g_1, \dots, \nabla g_m$).

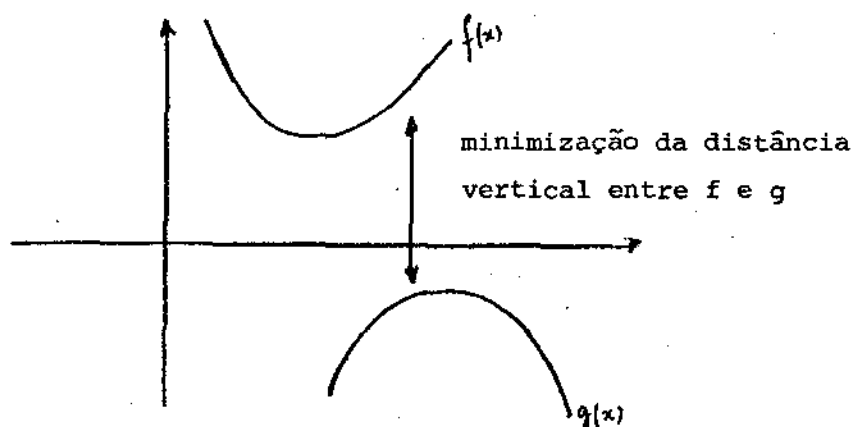
3.2.8 Nota histórica: o teorema de Fenchel

— Sejam $f(x)$ e $g(x)$ respectivamente uma função convexa e uma função côncava, definidas em determinados domínios convexos. Suponhamos agora que pretendíamos resolver o problema

$$\inf_{x \in \{\text{dom } f \cap \text{dom } g\}} [f(x) - g(x)] \quad \dots (3.XII)$$

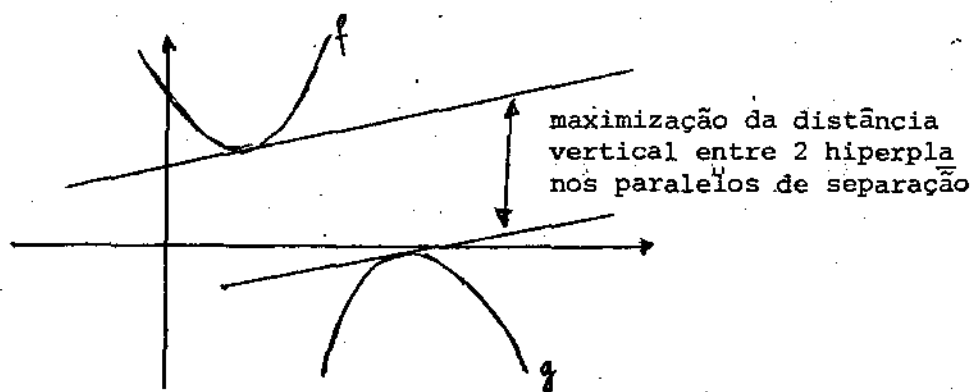
O problema (3.XII) pode ser perspectivado graficamente como o problema de minimização da distância vertical entre as 2 funções f e g .

Ilustração gráfica em $\mathbb{R}^2$:



Existe todavia, e como se perceberá facilmente, uma forma alternativa e dual de encarar a mesma questão: o problema (3.XII) pode ser visto como o problema de maximização da distância vertical entre quaisquer dois hiperplanos de suporte de $\text{epi } f$ e $\text{hip } g = \text{epi}[-g(x)]$ que sejam paralelos.

Ilustração gráfica em $\mathbb{R}^2$:



— Ora, como vimos no ponto II, para um dado x^* , $-f^*(x^*)$ e $-g^*(x^*)$ podem ser encaradas como as ordenadas na origem de funções afins paralelas (devido ao mesmo x^*) tangentes respectivamente a $f(x)$ e a $g(x)$. Como a maximização da distância vertical entre hiperplanos pode ser vista como a maximização da diferença entre as respectivas ordenadas na origem, temos então definido o problema dual do problema (3.XII).

$$\begin{aligned} & \sup_{x^*} \{ [-f^*(x^*)] - [-g^*(x^*)] \} \\ & = \sup_{x^*} \{ g^*(x^*) - f^*(x^*) \} \end{aligned}$$

— Estamos então em condições de perceber a lógica do seguinte teorema, devido originalmente a FENCHEL (20) :

TEOREMA 3.6 - TEOREMA DE DUALIDADE DE FENCHEL (Rockafellar, [f], 1970, p.327/Luenberger, 1969, p.201)

. Seja $f(x)$ uma função convexa própria definida em $\text{dom } f \subset \mathbb{R}^n$, e seja $g(x)$ uma função côncava própria definida em $\text{dom } g \subset \mathbb{R}^n$. Então:

$$\inf_{x \in \{\text{dom } f \cap \text{dom } g\}} \{ f(x) - g(x) \} = \sup_{x^* \in \{\text{dom } f^* \cap \text{dom } g^*\}} \{ g^*(x^*) - f^*(x^*) \}$$

desde que pelo menos uma das seguintes condições seja satisfeita:

- (a) $\text{rint } (\text{dom } f) \cap \text{rint } (\text{dom } g) \neq \emptyset$
- (b) f e g fechadas e $\text{rint } (\text{dom } f^*) \cap \text{rint } (\text{dom } g^*) \neq \emptyset$

. No caso de ambas serem satisfeitas garante-se que o ínfimo e o supremo são finitos.

— Como sublinha Rockafellar ([f], 1970, p. 327), o teorema de Fenchel pode ser deduzido independentemente da teoria geral da dualidade convexa, recorrendo apenas à noção de função conjugada e às noções elementares de hiperplanos de separação e de suporte (foi nesse sentido que desenvolvemos o nosso raciocínio gráfico anterior, onde não utilizámos qualquer resultado das secções 3.2.1. a 3.2.7). Todavia, e alternativamente, como também demonstra Rockafellar ([f], 1970, pág 329 - 331), os resultados de dualidade conjugada estabelecidos pionei

(20) FENCHEL, W., Convex Cones, sets and functions. Lecture Notes, Princeton Univ., 1951.

ramente por Fenchel podem ser derivados da teoria geral da dualidade conjugada convexa, tal como ela tem vindo a ser desenvolvida nas alíneas precedentes.

3.2.9. Uma generalização: a inclusão doutro tipo de perturbações

— Vamos terminar este estudo apontando uma generalização possível da análise que temos vindo a fazer da dualidade conjugada em programação convexa. O problema básico de programação convexa (3.1)

$$\begin{array}{ll} \min f(x) \\ \text{s.a. } \left\{ \begin{array}{l} G(x) \leq 0 \\ x \geq 0 \end{array} \right. \end{array}$$

foi na alínea 3.1.1. relacionado com uma família de problemas convexos,

$$\begin{array}{ll} \min f(x) \\ \text{s.a. } \left\{ \begin{array}{l} G(x) \leq u \\ x \geq 0 \end{array} \right. \end{array}$$

que diferiam do problema básico pela consideração duma perturbação u no membro direito das restrições que não de sinal. Esta família de problemas, que designámos por problema convexo generalizado, acabou por se mostrar extremamente útil aos nossos objectivos de estudo, tendo sido desenvolvida com base nela toda a teoria geral da dualidade conjugada convexa. Fique todavia a certeza que todos os resultados obtidos poderiam ter sido estabelecidos, sem qualquer dificuldade adicional, trabalhando com um problema com perturbações mais gerais, por exemplo,

$$\begin{array}{ll} \min f(x-u\hat{0}) \\ \text{s.a. } g_i(x-u\hat{1}) \leq w_i & (i=1, \dots, m) \\ x \geq 0 \end{array}$$

em que $u^0, u^1 \in \mathbb{R}^n$ e W_i ($i=1, \dots, m$) $\in \mathbb{R}$ são variáveis-perturbação. É neste sentido que aponta ROCKAFELLAR nas páginas 322-326 do seu livro ([f], 1970) ou, melhor ainda, é esta a linha seguida por AVRIEL (1976, secções 5.2 e seguintes). AVRIEL trabalha sempre com um problema "perturbado" o mais geral possível, embora menos intuitivo, confirmando-se que os resultados a que chega são perfeitamente análogos aos que se obtêm incluindo perturbações unicamente nos membros direi-tos das restrições.

BIBLIOGRAFIA

- . AVRIEL, M. (1976) Nonlinear Programming - Analysis and Methods, Prentice-Hall, New-Jersey.
- . BAZARAA, M.S. e SHETTY, C.M. (1979) Nonlinear Programming - Theory and Algorithms. J.Willey and Sons, New-York.
- . BERGE, C.; e GOUILLA-HOURI, A. (1962) Programmes, Jeux et Réseaux de Trnsport, Dunod, Paris.
- . LUENBERGER, D.G. (1969) Optimization by Vector Space Methods, J. Willey and Sons, New York.
- . PINHEIRO, M.R. [a] (1984) Conjugação e Dualidade em Programação Convexa, Working Paper nº 25 da Faculdade de Economia da Universidade Nova de Lisboa, Lisboa.
- . _____ [b] (1985) Dualidade Conjugada e Quase Conjugada nas teorias do Produtor e Consumidor, Dissertação de Mestrado em Economia, Faculdade de Economia da Universidade Nova de Lisboa, Lisboa.
- . ROCKAFELLAR, R.T. [a] (1964) "Duality Theorems for Convex Functions", Bull. Amer. Math. Soc., 70, pp. 189 - 192.
- . _____ [b] (1967) "Duality and Stability in Extremum Problems Involving Convex Functions", Pacific J.Math., 21, pp. 167 - 187.
- . _____ [c] (1967) "Convex Programming and Systems of Elementary Monotonic Relations", J. Math. Anal. Applic., 19, pp. 543 - 564.

[d] (1968) "Duality in Nonlinearity Programming" in Mathematics of the Decision Sciences, Part II, Lectures in Applied Mathematics, vol. II, Amer. Math. Society, pp. 401 - 422.

[e] (1969) "Convex Functions and Duality in Optimization - Problems and Dynamics" in KUHN, H.W. and SZEGÖ, G.P. Eds, Mathematical Systems Theory and Economics I, Springer-Verlag, Berlin, Heidelberg, New York.

[f] (1970) Convex Analysis, Princeton University Press, Princeton.