

UNIVERSIDADE NOVA DE LISBOA

Faculdade de Economia



CONJUGAÇÃO E DUALIDADE
EM PROGRAMAÇÃO CONVEXA

Maximiano Reis Pinheiro

Working Paper nº 25

UNIVERSIDADE NOVA DE LISBOA

Faculdade de Economia

Campo Grande, 185

1 700 LISBOA

Junho, 1984

CONSIDERAÇÕES PRÉVIAS

— Este texto resulta dum trabalho que apresentei no âmbito da cadeira de Optimização e Análise Económica I do curso de pós-graduação em economia da Faculdade de Economia da Universidade Nova de Lisboa. Pretende constituir um texto pedagógico sobre a questão das relações de conjugação e de dualidade em análise convexa, e fornecer pistas de utilização desta abordagem na obtenção de relações de dualidade nas teorias do consumidor e do produtor. Poderá deste modo ser encarado como uma 1ª etapa dum projecto mais ambicioso de publicação dum paper em que estas relações de dualidade microeconómicas sejam analisadas exaustivamente segundo esta abordagem, assim como seja estendida a questão da conjugação e dualidade à optimização dinâmica.

— Desejo demonstrar publicamente o meu agradecimento ao Prof. Doutor J. Dias Coelho, responsável pelas cadeiras de Optimização da pós-graduação, pelo estímulo prestado e pelas correcções e sugestões fornecidas. Obviamente todos os erros que possam persistir são da minha exclusiva responsabilidade. Uma palavra de reconhecimento deve ser ainda dirigida a Teresa Loyd Braga, colega no curso de pós-graduação, pelo muito que ajudou a leitura do seu trabalho, referenciado na bibliografia.

Faculdade de Economia,

Maximiano Reis Phry

INDICE

PARTE I - ALGUMAS NOÇÕES PRELIMINARES DE ANÁLISE CONVEXA

1.1. CONJUNTOS CONVEXOS

1.1.1 Conjuntos convexos e envolvente convexa	1
1.1.2 Conjuntos afins e subespaços	2
1.1.3 Hiperplanos	3
1.1.4 Semiespaços	3
1.1.5 Fecho, interior e fronteira dum conjunto convexo	4
1.1.6 Interior e fronteira relativos	4
1.1.7 Hiperplanos e semiespaços de suporte	5

1.2. FUNÇÕES CONVEXAS E CONCAVAS

1.2.1 Funções convexas	7
1.2.2 Algumas relações entre funções e conjuntos convexos/função indicadora	9
1.2.3 Funções concâvas	10
1.2.4 Funções afins	11
1.2.5 Semicontinuidade e continuidade	11
1.2.6 Fechos de funções convexas e concavas	12

PARTE II - CONJUGAÇÃO

2.1. FUNÇÕES CONJUGADAS

2.1.1 Funções convexas como supremo por pontos de funções lineares afins	13
2.1.2 Conjugada duma função convexa	14
2.1.3 Propriedades da conjugada duma função convexa	16
2.1.4 Desigualdade de Fenchel/pontos conjugados	18
2.1.5 Conjugadas de funções concâvas	20
2.1.6 Exemplo de aplicação: função produção e função lucro ..	22

2.2. FUNÇÕES SUPORTE DE CONJUNTOS CONVEXOS

2.2.1 Função suporte como conjugada de função indicadora	24
2.2.2 Exemplo de aplicação: função de produção e função custo	25

PARTE III - DUALIDADE EM PROGRAMAÇÃO CONVEXA

3.1. O PROBLEMA DE PROGRAMAÇÃO CONVEXA

3.1.1 O problema generalizado de programação convexa	27
3.1.2 Bifunção associada e função perturbação	27
3.1.3 Algumas propriedades da bifunção associada e da função perturbação	29
3.1.4 Relação entre função auxiliar Lagrangeana e função per- turbação/Multiplicadores de Lagrange e teoria de Kuhn- -Tucker	30

3.2. CONJUGAÇÃO E DUALIDADE EM PROGRAMAÇÃO CONVEXA

3.2.1 Bifunções adjuntas	33
3.2.2 Programas duais	34
3.2.3 Bifunções adjuntas e transformações lineares adjuntas/ /dual dum programa linear	35
3.2.4 Uma perspectiva gráfica da dualidade conjugada	37
3.2.5 Relação entre dualidade conjugada e teoria de Lagrange/ /Kuhn-Tucker	39

ANEXO À PARTE III

A. TEOREMA DE DUALIDADE DE FENCHEL

A.1 Teorema de dualidade de Fenchel/sua interpretação gráfi- ca	41
A.2. Teorema de Fenchel e teoria geral de dualidade convexa	43

BIBLIOGRAFIA

- [1], ROCKAFELLAR, R.T., (1970), Convex Analysis, Princeton University Press, Princeton.
- [2], LUENBERGER, D.G., (1969), Optimization by Vector Space Methods, John Willey and Sons, New-York.
- [3], BERGE, C. e GHOUILLA-HOURI, A. (1962), Programmes, Jeux et Réseaux de Transport, Dunod, Paris
- [4], BRAGA, T.L., "Notas sobre Dualidade no âmbito da Análise convexa/algumas aplicações à economia" - Relatório final de Optimização e Análise Económica I, 1981/82, (manuscrito), FE-UNL.
- [5], BAZARAA, M.S. e SHETTY, C.M., (1979), Nonlinear Programming-theory and Algorithms, Wiley and Sons.
- [6], VARIAN, H.R., (1978), Microeconomic Analysis, W.W. Norton e Ca, New-York.
- [7], BLACKORBY, C.; PRIMONT, D.; RUSSEL, R.R., (1978), Duality Separability and Functional Structure..., North-Holland, N. York.
- [8], DIEWERT, W.E. (1974), "Applications of Duality Theory" in Frontiers of Quantitative Economics, Vol. II, ed. INTRILIGATOR, M.D. e KENDRICK, D.A., North-Holland.

PARTE I

**ALGUMAS NOÇÕES PRELIMINARES
DE ANÁLISE CONVEXA**

1.1. CONJUNTOS CONVEXOS

— Consideremos o espaço $\mathbb{R}^n$ cujos vectores representaremos, como é usual, por énplos ordenados $x = (x_1, \dots, x_n)^T$. Admitindo a base ortonormada $\{e_1, \dots, e_j, \dots, e_n\}$, em que e_j é a j -ésima coluna da matriz identidade de ordem n ,

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

é, como sabemos, o produto escalar entre os vectores x e y . Recordemos, 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 são ditos ortogonais se $\cos(x, y) = 0$, o que acontece, portanto, quando e só quando $\langle x, y \rangle = 0$.

1.1.1. Conjuntos convexos e envolvente convexa

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

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

Por palavras, um conjunto é convexo se, contendo quaisquer dois vectores x e y , também contém todos os pontos do segmento de recta que os une, ou seja, contém também todas as suas combinações lineares convexas.

— Envolvente convexa dum conjunto $S \subset \mathbb{R}^n$, $\text{conv } S$, é o mais pequeno conjunto convexo que contém S , ou, identicamente, é a intersecção de todos os conjuntos convexos que contém S . Sublinhe-se que

a intersecção de conjuntos convexos é ainda um conjunto convexo.

1.1.2. Conjuntos afins e subespaços

— Um conjunto convexo $M \subset \mathbb{R}^n$ é dito conjunto afim, ou variedade de linear, se para qualquer $\lambda \in \mathbb{R}$ temos

$$(1-\lambda)x + \lambda y \in M \quad ; \quad \forall x, y \in M.$$

Tendo em conta que

$$(1-\lambda)x + \lambda y = x + \lambda (y-x)$$

significa isto que um conjunto afim, se contém 2 pontos x e y , contém igualmente toda a linha recta que passa por eles.

— Subespaços são conjuntos afins que contêm a origem. Definindo como paralelos dois conjuntos afins M e L desde que L resulte duma translação de M , ou seja, desde que $L = \{x+a \mid x \in M\}$ com $a \in \mathbb{R}^n$, verifica-se facilmente que cada conjunto afim não vazio é paralelo a um único subespaço (Rockafellar, [1], p.4). Dimensão dum conjunto afim pode ser então definida como a dimensão do subespaço que lhe é paralelo, sendo dimensão dum subespaço P o número máximo de vectores linearmente independentes que nele podemos encontrar em que os vectores $x^1, \dots, x^n \in P$ são linearmente independentes sse $\mu_1 x^1 + \dots + \mu_m x^m = 0$ apenas com $\mu_1, \dots, \mu_m$ escalares todos nulos.

Consoante os semiespaços contêm ou não contêm os hiperplanos que os geram, recebem a designação respectivamente de semiespaços fechados ou abertos. Como é óbvio, em qualquer dos casos, estamos perante conjuntos convexos.

1.1.5. Fecho, interior e fronteira de um conjunto convexo

— Seja a "esfera" de centro a $\mathbb{R}^n$ e raio $\epsilon \in \mathbb{R}_+$ dada por:

$$E_\epsilon(a) = \{x \in \mathbb{R}^n / d(x,a) \leq \epsilon\}$$

em que $d(x,a) = \langle x-a, x-a \rangle^{1/2} = \sqrt{\sum_j (x_j - a_j)^2}$ é a distância euclídiana entre x e a . Fecho dum conjunto $S \subset \mathbb{R}^n$, $\text{cls } S$ (do inglês "closure"), é o conjunto de pontos tais que, para cada $a \in S$, por mais pequeno que seja $\epsilon > 0$, temos $S \cap E_\epsilon(a) \neq \emptyset$:

$$\text{cls } S = \{x / \forall \epsilon > 0 \quad S \cap E_\epsilon(x) \neq \emptyset\}$$

Se $S = \text{cls } S$, então S diz-se um conjunto fechado. É de notar que a intersecção de conjuntos convexos fechados é ainda um conjunto convexo fechado.

— Interior dum conjunto S , $\text{int } S$, é o conjunto de pontos tais que existe algum $\epsilon > 0$ que leva a que a esfera $E_\epsilon(x)$ esteja contida em S :

$$\text{int } S = \{x / \exists \epsilon > 0, E_\epsilon(x) \subset S\}$$

Complementarmente, designa-se por fronteira de S o conjunto de pontos que pertencem a $\text{cls } S$ mas não pertencem a $\text{int } S$.

1.1.6. Interior e fronteira relativos

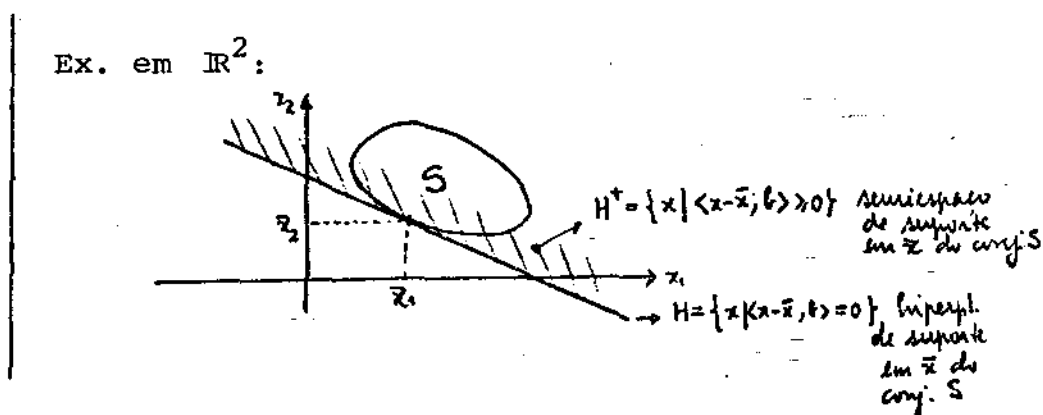
— Qualquer conjunto S contido num conjunto afim de dimensão inferior a n tem, como é óbvio, interior vazio ($\text{int } S = \emptyset$), tal co

mo o caracterizamos atrás. Daí que seja muitas vezes útil em análise convexa definir a noção de interior relativo.

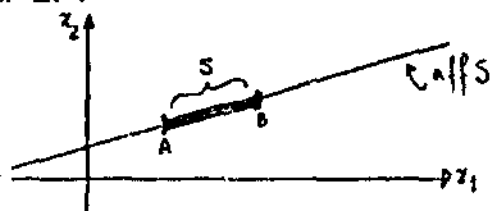
— Interior relativo dum conjunto S , $\text{rint } S$, é o interior que resulta quando encaramos S , não como um subconjunto de $\mathbb{R}^n$, mas apenas como um subconjunto do mais pequeno conjunto afim que o contém (e que designaremos por envolvente afim de S , $\text{aff } S$) (2). Dentro da mesma lógica, a fronteira relativa dum conjunto S é constituída por todos os pontos que pertencem a $\text{cl } S$ mas não pertencem a $\text{rint } S$.

1.1.7. Hiperplanos e semiespaços de suporte

— Seja um conjunto não vazio $S \subset \mathbb{R}^n$ e seja $\bar{x}$ um ponto da sua fronteira. Um hiperplano $H = \{x / \langle x - \bar{x}, b \rangle = 0\}$, com $b \in \mathbb{R}^n$, é chamado um hiperplano de suporte em $\bar{x}$ do conjunto S desde que: 1) ou $S \subset H^+ = \{x / \langle x - \bar{x}, b \rangle \geq 0\}$; 2) ou $S \subset H^- = \{x / \langle x - \bar{x}, b \rangle \leq 0\}$. O semiespaço, definido pelo hiperplano de suporte, que contém S é designado por semiespaço de suporte. Por palavras, um semiespaço de suporte a um conjunto S é um semiespaço que contém S e que tem na sua fronteira pelo menos um ponto da fronteira de S (repare-se que para $x = \bar{x}$ vem necessariamente $\langle x - \bar{x}, b \rangle = 0$, pelo que o hiperplano de suporte passa pelo ponto fronteiro $\bar{x}$).



(2) — Ex. em $\mathbb{R}^2$:



Neste caso
 $\text{int } S = \emptyset$
 $\text{e } \text{rint } S = S \setminus \{A, B\}$

— TEOREMA 1.A (Rockafellar, [1], teorema 11.5, pag. 99)

. Um conjunto convexo fechado é a intersecção de todos os semi-espacos fechados que o contêm.

Ou, equivalentemente, restringindo-nos aos semi-espacos de suporte:

— TEOREMA 1.B (Berge e Houri, [3], corolário 22 teor. separação, p. 55)

. Um conjunto convexo fechado é a intersecção dos seus semi-espacos de suporte.

Como veremos mais adiante (parte II), é a partir desta relação dual entre um dado conjunto convexo fechado e a respectiva caracterização alternativa, dada pela intersecção dos seus semi-espacos de suporte, que assenta a ideia de conjugação.

1.2. FUNÇÕES CONVEXAS E CONCAVAS

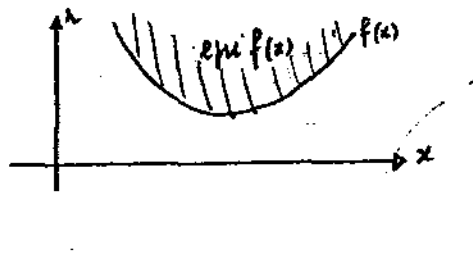
1.2.1. Funções convexas

— Epigrafo duma função $f(x)$ de $\mathbb{R}^n$ em $[-\infty, +\infty]$ (3) é o subconjunto de $\mathbb{R}^{n+1}$ definido por:

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

Função convexa será então qualquer função $f(x)$ cujo epigrafo seja um conjunto convexo.

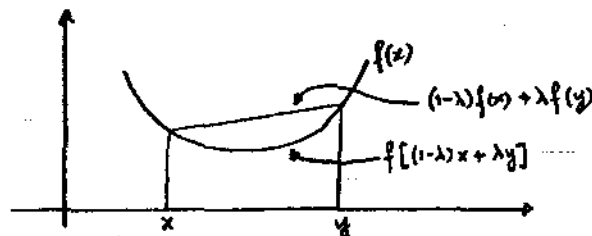
Ex. em $\mathbb{R}^2$:



— Como sabemos, existem outras maneiras de definir função convexa, embora menos gerais que a anterior:

a) no caso de $f(x)$ não tomar o "valor" $-\infty$, então $f(x)$ é convexa sse, para qualquer $0 \leq \lambda \leq 1$, temos

$$f[(1-\lambda)x + \lambda y] \leq (1-\lambda)f(x) + \lambda f(y) \quad (4)$$



Sublinhe-se, mais uma vez, que esta definição só é válida quando f nunca for igual a $-\infty$.

(3) — No que se segue, vamos admitir que as funções reais podem tomar, não só valores finitos, mas também $-\infty$ e $+\infty$.

(4) — Daqui a relação entre o nome de conjuntos e funções convexas: em ambas as definições se utiliza a noção de combinação linear convexa.

b) TEOREMA 1.C (Rockafellar, [1], p. 27)

. Seja f uma função com derivadas parciais contínuas até à 2ª ordem definidas num subconjunto aberto $C \subset \mathbb{R}^n$. Então $f(x)$ é convexa sse a sua matriz hessiana

$$H = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right] \text{ for semidefinida positiva,} \\ (n \times n) \quad \forall x \in C$$

Repare-se também que a caracterização de função convexa dada pelo anterior teorema é menos geral que a definição pela convexidade do epigrafo, pois exige a existência de derivadas parciais contínuas até à 2ª ordem.

— Poderemos sempre, por comodidade, "estender" a $\mathbb{R}^n$ uma função convexa, definida originalmente num subconjunto $S \subset \mathbb{R}^n$, fazendo $f(x) = +\infty$, $\forall x \notin S$. De qualquer forma, designaremos por domínio efectivo duma função convexa, $\text{dom } f$, apenas a projecção em $\mathbb{R}^n$ do seu epigrafo, ou seja, o conjunto S em que ela estava originalmente definida:

$$\text{dom } f = \{x / \exists r, (x, r) \in \text{epi } f(x)\} = \{x / f(x) < +\infty\}$$

Sendo f uma função convexa, o seu epigrafo é convexo, pelo que a sua projecção sobre $\mathbb{R}^n$, ou seja, o seu domínio efectivo, será necessariamente convexo.

— Uma função convexa diz-se própria se o seu epigrafo não é vazio, nem contém linhas verticais, isto é, se

$$\begin{cases} \exists x \in \mathbb{R}^n: f(x) < +\infty \\ \forall x \in \mathbb{R}^n, f(x) > -\infty \end{cases}$$

É fácil de verificar que a soma de funções convexas próprias é ainda uma função convexa própria (Rockafellar, [1], teorema 5.2, p.33).

1.2.2. Algumas relações entre funções e conjuntos convexos / função indicadora

— A partir da relação fundamental entre uma dada função convexa e o respectivo epígrafo convexo é possível estabelecer algumas relações importantes para a análise que se segue:

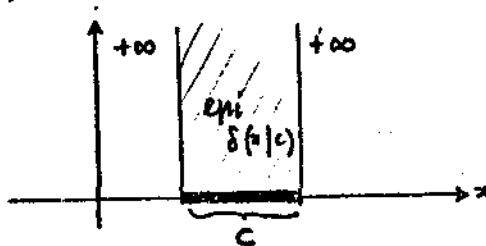
- a) Por um lado, se f é convexa, então os conjuntos de nível abertos $\{x/f(x) < \alpha\}$ ou fechados $\{x/f(x) \leq \alpha\}$, com α escalar, são conjuntos convexos (5).
- b) Por outro lado, seja F um conjunto convexo em $\mathbb{R}^{n+1}$; então a função $Z(x) = \inf\{r/(x,r) \in F\}$ é uma função convexa $Z(x)$. Este resultado compreende-se facilmente se encarmos a fronteira inferior de F como o fronteira do epígrafo duma função convexa $Z(x)$ não explicitada à priori.

— Ainda a propósito de relação entre conjuntos e funções convexas, defina-se função indicadora dum qualquer conjunto C como a função:

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

Logo, e como se percebe, a função indicadora é convexa apenas se o conjunto C for convexo.

Ex. em $\mathbb{R}^1$:



Como veremos, a função indicadora vai ser especialmente útil ao permitir fazer, duma forma extremamente cómoda, a extensão a $\mathbb{R}^n$ de qualquer função convexa $f(x)$ definida num subconjunto $S \subset \mathbb{R}^n$. Com efeito,

$$f(x) \text{ com } x \in S \iff f(x) + \delta(x/C) \text{ com } x \in \mathbb{R}^n$$

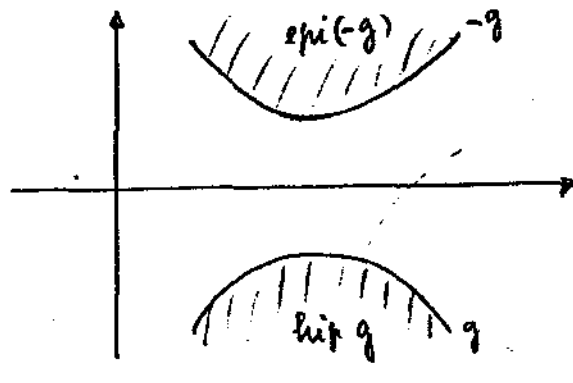
(5) — O recíproco não é verdadeiro. As funções quase-convexas também geram conjuntos de nível $\{x/(x) \leq \alpha\}$ convexos.
(<)

1.2.3. Funções concâvas

— Uma função $g(x)$ de $\mathbb{R}^n$ em $[-\infty, +\infty]$ é concâva se $-g(x)$ fôr uma função convexa. Sendo assim, poderemos retirar facilmente as características e propriedades das funções concâvas através da função convexa que é a sua simétrica. Nomeadamente, o epígrafo de $-g$ é um conjunto convexo. Refira-se a propósito que se dá o nome de hipografo de g ao simétrico do epigrafo de $-g$, isto é:

$$\text{hip } g = - \text{epi } (-g) = \{ (x, r) / x \in \mathbb{R}^n, r \in \mathbb{R}, r \leq g(x) \}$$

Ex. em $\mathbb{R}^2$:



— Enquanto que para funções convexas eram convexos os conjuntos de nível definidos por $\{x / f(x) \leq \alpha\}$, agora serão convexos os conjuntos de nível de funções concâvas definidos como se segue:

$$\{x / g(x) > \beta\} \text{ ou } \{x / g(x) \geq \beta\} \quad (6)$$

— Podemos também estender a $\mathbb{R}^n$ uma função concâva, definida em $S \subset \mathbb{R}^n$, atribuindo à função o valor $-\infty$ para todo o $x \notin S$. Tal significa fazer, e na lógica do que vimos no fim da alínea anterior, utilizando a função indicadora:

$$g(x), x \in S \iff g(x) - \delta(x/S), \quad x \in \mathbb{R}^n$$

(6) — Com efeito, $\{x / \underbrace{g(x)}_{f. \text{ convexa}} \leq \alpha\}$ é convexo, pelo que $\{x / g(x) \geq \underbrace{-\alpha}_{\beta}\}$ também o será.

1.2.4. Funções afins

— Funções afins são funções lineares do tipo $h(x) = \langle x, a \rangle - \beta$ com $a \in \mathbb{R}^n$ e $\beta \in \mathbb{R}$. Repare-se que, ao serem lineares, verifica-se facilmente que as funções afins são simultaneamente convexas e concavas.

— Tendo em conta o que referimos na alínea 1.1.3. (hiperplanos), conclui-se que o conjunto de pontos $(x, h(x))$, isto é, que o grafo de $h(x)$, corresponde precisamente a um hiperplano definido em $\mathbb{R}^{n+1}$. Com efeito,

$$\begin{aligned} & \{(x, h(x)) \mid h(x) = \langle x, a \rangle - \beta\} = \\ & = \{(x, h(x)) \mid \beta = (-1) \cdot h(x) + \langle x, a \rangle\} = \\ & = \{(x, h(x)) \mid \langle (x, h(x)), (a, -1) \rangle = \beta\} \end{aligned}$$

que é claramente a equação de definição dum hiperplano em $\mathbb{R}^{n+1}$.

1.2.5. Semicontinuidade e continuidade

— Uma função $Z(x)$ diz-se semicontínua inferiormente num ponto x sse

$$Z(x) \leq \lim_{x_i \rightarrow x} Z(x_i)$$

qualquer que seja a sequência $x_1, \dots, x_i, \dots, x_n$ que convirja para x , existindo, mesmo que não finito, $\lim_{x_i \rightarrow x} Z(x_i)$. Analogamente uma função $Z(x)$ diz-se semicontínua superiormente num ponto x sse $Z(x) \geq \lim_{x_i \rightarrow x} Z(x_i)$ para todas as sequências $x_1, \dots, x_n$ convergentes para x tais que exista finito $\lim_{x_i \rightarrow x} Z(x_i)$. Uma função semicontínua inferiormente (superiormente) em todos os pontos do seu domínio diz-se semicontínua inferiormente (superiormente).

— Quando uma função é simultaneamente semicontínua inferior e superiormente diz-se que é uma função contínua.

1.2.6. Fechos de funções convexas e concâvas

—— Designaremos por "envolvente semicontínua inferior" de uma função $Z(x)$ (lower semicontinuous hull) a maior função semicontínua inferior majorizada por $Z(x)$. Tal equivale a definir "envolvente se micontínua inferior" como a função cujo epígrafo é o fecho do epígrafo de $Z(x)$.

—— Fecho de uma função convexa $f(x)$, cl f , pode então ser de finido como a envolvente semicontínua inferior de $f(x)$ (que é também uma função convexa) desde que $f(x)$ não assuma nunca o valor $-\infty$. Se $f(x) = -\infty$ para algum x , então cl f é definida como a função cons tante $-\infty$, $\forall x \in \mathbb{R}^n$.

—— Similarmente, podemos definir "envolvente semicontínua su perior", assim como fecho de funções concâvas, pelo que nos dispensamos de o fazer.

—— Finalmente, restará referir que uma função que coincide com o seu fecho diz-se função fechada, isto é, $Z(x)$ é fechada sse

$$Z(x) = \text{cl } Z(x).$$

PARTE II

CONJUGAÇÃO

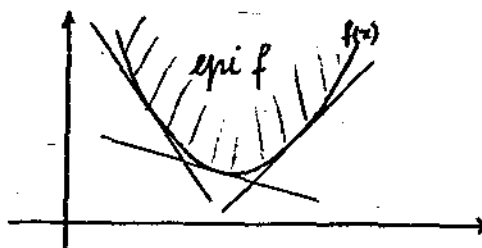
2.1. FUNÇÕES CONJUGADAS

— Como nota Rockafellar ([1], pag. 102), há duas 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 de conjugação.

2.1.1) Funções convexas como supremo por pontos de funções lineares afins

— Sublinhámos no fim da secção 1.1 que um conjunto convexo fechado pode ser encarado como a intersecção dos seus semi-espacos de suporte. Sendo assim, podemos encarar o epigrafo duma função convexa fechada, que é um conjunto convexo fechado, como a intersecção dos seus semi-espacos de suporte.

Ex. em $\mathbb{R}^2$:



— Atendendo a que cada hiperplano definido em $\mathbb{R}^{n+1}$ corresponde, como vimos na alínea afim de $\mathbb{R}^n$ em $\mathbb{R}$, poderemos transpor a anterior linguagem de conjuntos para uma linguagem de funções. Primeiro, contudo, definamos supremo por pontos de um conjunto de funções $\{Z_i(x), i \in I\}$ como a função que cada x faz corresponder o maior dos valores $Z_i(x)$:

$$Z(x) = \sup \{Z_i(x) / i \in I\}$$

Percebe-se sem dificuldade que o supremo por pontos de funções convexas é ainda uma função convexa.

— TEOREMA 2.A (Rockafellar, [1], cor. 12.1.1., pag. 103)

. Seja $f(x)$ uma função convexa fechada; então f pode ser encarada como o supremo por pontos de todas as funções lineares afins majoradas por $f(x)$. (7)

Será no desenvolvimento desta perspectiva que chegaremos à noção de conjugada duma função convexa.

2.1.2. Conjugada de uma função convexa

— Seja um hiperplano em $\mathbb{R}^{n+1}$

$$H = \{(r, x) / r \cdot r^* + \langle x^*, x \rangle = \beta, r^* \in \mathbb{R}, x^* \in \mathbb{R}^n\}$$

em que, para efeitos de ilustração gráfica, se considera o eixo dos r 's como o eixo vertical. Devido à existência, por definição, de um grau de liberdade nas equações dos hiperplanos (conf. alínea 1.1.3), podemos, sem perda de generalidade, fazer $r^* = -1$, desde que não estejamos interessados em hiperplanos paralelos ao eixo vertical, e desde que escolhamos x^* e β apropriadamente (basta dividir toda a equação por $(-r^*)$ no caso de r^* ser diferente de (-1)). Fica então como equação genérica dum hiperplano não vertical:

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

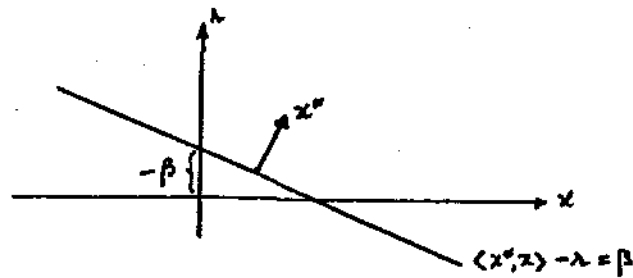
Repare-se que resolvendo a equação anterior em relação à coordenada do eixo vertical fica:

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

que não é mais que a expressão duma função linear afim com ordenada na origem $(-\beta)$.

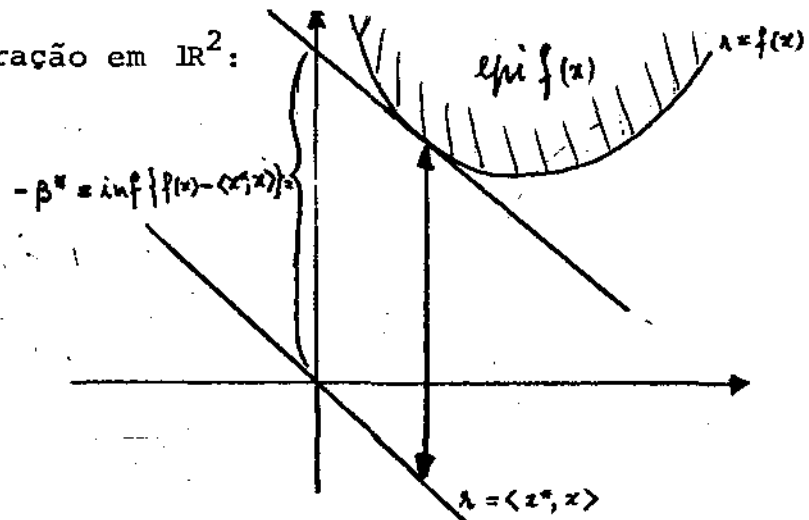
(7) — Se f não for convexa e fechada, o supremo por pontos de todas as funções lineares afins majorizadas por $f(x)$, será, não f , mas sim $cl(\text{conv } f)$, em que por $\text{conv } f$ se entende a função cujo epígrafo é a envolvente convexa do epígrafo de f .

Ex. em $\mathbb{R}^2$



— A nossa questão é: para um dado x^* e uma dada função convexa de $\mathbb{R}^n$ em $[-\infty, +\infty]$, como determinar a ordenada na origem ($-\beta$) por forma a que a função linear afim $r = \langle x^*, x \rangle - \beta$ seja tangente a $f(x)$, isto é, por forma a que $\langle x^*, x \rangle - r = \beta$ defina a equação dum hiperplano de suporte ao epígrafo de $f(x)$?

Ilustração em $\mathbb{R}^2$:



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

$$\begin{aligned} -\beta^* &= \inf_x \{f(x) - \langle x^*, x \rangle\} \\ \Leftrightarrow \beta^* &= \sup_x \{\langle x^*, x \rangle - f(x)\} \end{aligned}$$

Como é óbvio, β^* vai depender do x^* fixado.

— DEFINIÇÃO:

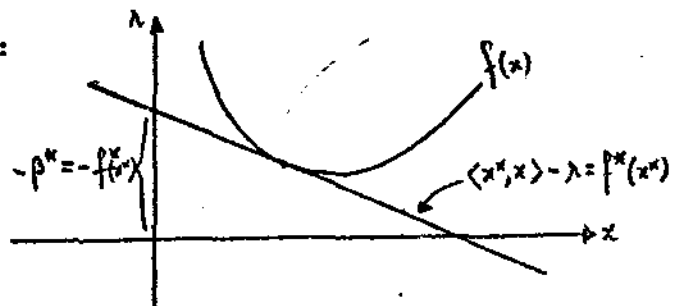
. Seja $f(x)$ uma função convexa definida num conjunto convexo C . Designaremos por função conjugada de $f(x)$ a função:

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

. O domínio efectivo de $f^*(x^*)$ será o conjunto formado pelos vectores x^* tais que $f^*(x^*) < +\infty$.

— A função conjugada de uma função convexa $f(x)$ dará então, para cada x^* , conforme pretendido, o simétrico da maior ordenada na origem da função afim $\langle x^*, x \rangle - \beta$ por forma a que esta seja ainda majorada por $f(x)$.

Ilustração em $\mathbb{R}^2$:



$\langle x^*, x \rangle - r = f^*(x^*)$ traduz portanto, para os vários x^* , as equações de definição dos hiperplanos de suporte ao epigrafo convexo de $f(x)$.

2.1.2. Propriedades da conjugada duma função convexa

— O seguinte teorema sintetiza as principais propriedades da conjugada duma função convexa:

— TEOREMA 2.B (Rockafellar, [1], teorema 12.2, pag. 104)

. Seja $f(x)$ uma função convexa. A conjugada f^* é então uma função convexa fechada, própria sse f é própria. Além disso $(cl f)^* = f^*$ assim como $f^{**} = cl f$

— Realçemos devidamente as várias características importantes da conjugada duma função convexa que são referidas no anterior teorema:

- 1º. A conjugada duma função convexa é uma função também convexa (para além de no Rockafellar, pode ser encontrada a demonstração em Luenberger, [2], pag 196, e em Berge e Houri, [3], pag. 95).
- 2º. A conjugada f^* é própria sse f é própria. Refira-se a propósito que as únicas duas funções impróprias conjugadas são precisamente as funções constantes $+\infty$ e $-\infty$ (o que se confirma facilmente pela definição).
- 3º. f^* é necessariamente uma função fechada, mesmo que f não o seja (além de no Rockafellar, demonstração em Luenberger, [2], pag. 196).
- 4º. A conjugada da conjugada é a própria função desde que a função de partida seja fechada. Com efeito, seja f uma função fechada. Como vimos no fim da alínea anterior, as funções afins que são tangentes a $f(x)$ têm como expressão analítica

$$\langle x^*, x \rangle - r = f^*(x^*) \iff r = \langle x^*, x \rangle - f^*(x^*)$$

Sendo assim, a função f pode ser encarada como o supremo por pontos de colecção destas funções afins (uma para cada valor diferente de x^*). Quer dizer:

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

o que corresponde precisamente à definição de conjugada de $f^*(x^*)$. Logo

$$f(x) = [f^*(x^*)]^* = f^{**}$$

— Mostrámos pois que existe uma relação dual entre uma dada função convexa e as funções afins que são por ela majoradas, sumariadas elegantemente pela função conjugada. A prova disto mesmo é que é possível recuperar a função original, no caso desta ser fechada, a partir da respectiva função conjugada. Mesmo no caso de a função original não ser fechada, recupera-se o essencial que é seu fecho clí

— As funções conjugadas gozam ainda de outras propriedades para além das que constam no teorema 2.B. São exemplos disso, resultando directamente da aplicação da definição, as seguintes (Rockafellar, [1], §12):

a) A conjugação inverte o sentido das desigualdades funcionais, isto é: $f_1 \leq f_2 \implies f_2^* > f_1^*$.

b) Seja $\psi^*(x^*)$ a conjugada duma função convexa $\psi(x)$. Então $f(x) = \psi[A(x-a)] + \langle x, a^* \rangle + \alpha$

(Com A matriz de $n \times n$, a^* e a vectores de $\mathbb{R}^n$, α escalar)

Admite como conjugada $f^*(x^*) = \psi^*[A^{*-1}(x^*-a^*)] + \langle x^*, a \rangle + \alpha^*$ com A^* matriz adjunta de A e $\alpha^* = -\alpha - \langle a, a^* \rangle$.

c) $\inf_{x^*} f(x) = -f^*(0)$, assim como, dualmente, $\inf_{x^*} f^*(x^*) =$

$-f(0)$. Logo, $\inf_x f(x) = 0 = f(0)$ verifica-se sse

$\inf_{x^*} f^*(x^*) = 0 = f^*(0)$. Por palavras, tal significa

que a conjugação preserva a classe das funções convexas fechadas não negativas que se anulam na origem.

d) Uma função convexa é simétrica, isto é,

$f(-x) = f(x)$, $\forall x$, sse a sua conjugada também é simétrica:

$f^*(-x^*) = f^*(x^*)$, $\forall x^*$

2.1.4. Desigualdade de Fenchel/pontos conjugados

— A teoria da conjugação que temos vindo a desenvolver nas alíneas anteriores pode ser vista como uma teoria para encontrar a "melhor" das desigualdades do tipo:

$$f(x) + \psi(y) \geq \langle x, y \rangle, \quad \forall x, y \quad \dots (A)$$

com f e ψ funções convexas de $\mathbb{R}^n$ em $]-\infty, +\infty]$.

— Seja W o conjunto de todos os pares de funções (f, ψ) para os quais a anterior desigualdade é válida. O "melhor" dos pares $(f, \psi) \in W$ é aquele que não pode ser mais "apertado", isto é, é o par tal que, para $(f', \psi') \in W$ que

$$\begin{cases} f' \leq f \\ \psi' \leq \psi \\ f' + \psi' \geq \langle x, y \rangle \end{cases}$$

então é porque $f' = f$ e $\psi' = \psi$. Ora, partindo de

$$f(x) + \psi(y) \geq \langle x, y \rangle$$

conclui-se claramente que:

$$1^\circ. f(x) \geq \langle x, y \rangle - \psi(y), \quad \forall_{x,y} \iff f(x) \geq \sup_y \{ \langle x, y \rangle - \psi(y) \} = \psi^*(x).$$

$$2^\circ. \psi(y) \geq \langle x, y \rangle - f(x), \quad \forall_{x,y} \iff \psi(y) \geq \sup_x \{ \langle x, y \rangle - f(x) \} = f^*(y).$$

Sendo assim, para que a desigualdade (A) seja respeitada, teremos que ter, necessariamente

$$f(x) \geq \psi^*(x)$$

$$\psi(y) \geq f^*(y)$$

Logo, o "melhor", o mais "apertado" par (f, ψ) corresponde precisamente ao par formulado pelas funções conjugadas (ψ^*, f^*) , o que traduz por tanto uma outra visão da questão da conjugação.

— Repare-se que, em particular e pelo que acabámos de ver, a desigualdade

$$f(x) + f^*(x^*) \geq \langle x, x^* \rangle, \quad \forall_{x, x^*} \quad \dots (B)$$

é respeitada para qualquer função convexa e respectiva conjugada. Quer isto dizer que, fixada uma qualquer função convexa, a função mais "apertada" que com ela pode formar o "melhor" par respeitador da desigualdade, é precisamente a sua função conjugada. A desigualdade (B) é dado o nome de desigualdade de Fenchel. Os pares (x_0, x_0^*) para os quais a desigualdade de Fenchel é respeitada como igualdade

recebem a designação de pontos conjugados, isto é, x_0 e x_0^* são pontos conjugados sse

$$f(x_0) + f^*(x_0^*) = \langle x_0, x_0^* \rangle$$

— Seja $f(x)$ uma função diferenciável, cujo gradiente no ponto x representaremos $\nabla f(x)$. A partir da definição de funções conjugadas retira-se directamente que

$$\nabla f(x_0) = x_0^*$$

$$\nabla f^*(x_0^*) = x_0$$

Completam esta linha de raciocínio os seguintes teorema e corolário (Berge e Houri, [3], pag. 96 e 97):

— TEOREMA 2.C

. Para que uma função convexa diferenciável $f(x)$ definida em $\mathbb{R}^n$ admita uma função conjugada, é necessário e suficiente que, para todo o ponto $x_0^* \in \mathbb{R}^n$, exista um ponto $x_0 \in \mathbb{R}^n$ tal que $\nabla f(x) = x_0^*$.

— COROLÁRIO

. Seja uma função convexa diferenciável $f(x)$ que admite uma função conjugada $f^*(x^*)$. Para que o ponto x_0^* seja conjugado do ponto x_0 é necessário e suficiente que tenhamos $\nabla f(x_0) = x_0^*$.

2.1.5. Conjugadas de funções concâvas

— Um desenvolvimento similar ao das anteriores alíneas pode ser feito para as funções concâvas. Tenha, contudo, um especial cuidado no encarar 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 concâva, então $-g$ seja simétrica (8).

(8) — Com efeito, há também uma mudança de sinal na definição de conjugada.

— DEFINIÇÃO:

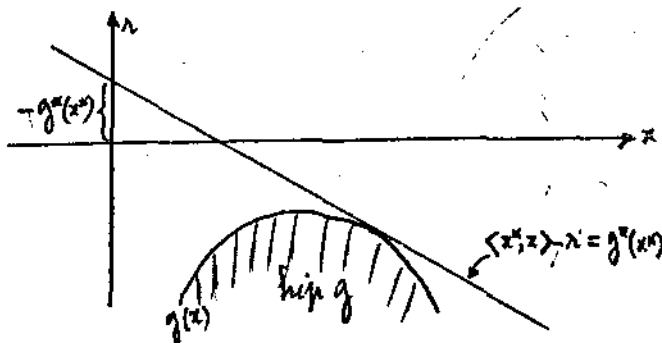
. Seja $g(x)$ uma função concâva definida num conjunto convexo D . A função conjugada de $g(x)$ é

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

. O domínio efectivo de $g^*(x^*)$ será o conjunto de pontos x^* tais que $g^*(x^*) > -\infty$.

— Graficamente, poderemos visualizar $g^*(x^*)$, tal como para as funções convexas, como o simétrico das ordenadas na origem das funções lineares afins tangentes a $g(x)$.

Ilustração em $\mathbb{R}^2$:



— Ao nível das propriedades das conjugadas de funções concâvas, e tal como para as funções convexas, temos que:

- a) g^* é uma função concâva fechada;
- b) $(cl\ g)^* = g^*$
- c) $g^{**} = cl\ g$

— Finalmente convirá referir que a desigualdade de Fenchel, no caso de funções concâvas se verifica com o sinal invertido:

$$g(x) + g^*(x^*) \leq \langle x, x^* \rangle, \quad \forall_{x, x^*}$$

mantendo-se, apesar disso, a definição de pontos conjugados, assim como a validade para funções concâvas do teorema e do corolário apresentados na alínea anterior.

2.1.6. Exemplo de aplicação: função produção e função lucro

— Seja $y = g(x)$ a função de produção de uma empresa, definindo a máxima quantidade de output y que se pode obter quando se utilizam as quantidades de inputs traduzidas pelo vector $x \in \mathbb{R}^n$. É portanto uma função que representa a tecnologia disponível. Admitamos que $g(x)$ é uma função concáva fechada e que a empresa é "price-taker" quer no mercado de factores, quer no mercado do produto final.

— Representando por $p \in \mathbb{R}$ e $W \in \mathbb{R}^n$ respectivamente o preço do output e o vector de preços dos inputs, a função lucro da empresa que dá o máximo lucro possível para cada combinação de preços (p, W) com que depara a empresa no mercado, será:

$$\Pi(p, W) = \sup_x \{p \cdot g(x) - \langle W, x \rangle\}.$$

Ora,

$$\Pi(p, W) = \sup_x \{p \cdot g(x) - \langle W, x \rangle\} \iff \Pi(p, W)/p = -\inf \{ \langle W/p, x \rangle - g(x) \} \iff$$

$g^*(W/p)$ conjugada de $g(x)$

$$= \boxed{g^*(W/p) = -\Pi(p, W)/p}$$

Como a função lucro é necessariamente uma função comogênea de grau 1 em (p, W) , podemos escrever:

$$g^*(W/p) = -\Pi(1, W/p)$$

$$g^*(W/r) = -\pi(W/r)$$

com π função lucro nos preços normalizados. Significa isto que a conjugada da função produção não é mais que o simétrico da função lucro nos preços normalizados dos factores. Como g é concáva vem que $g^* = -\pi(W/p)$ é concáva, e logo $\pi(W/p)$ é conxexa. Todavia, o mais importante da relação dual estabelecida resulta de ser perfeitamente possível recuperar a função de produção concáva fechada a partir da função lucro. Com efeito e como mostrámos nas alíneas anteriores $g^{**} = g$ e logo

$$g(x) = [g^*(W/r)]^* = \inf_{W/p} \{ \langle x, W/p \rangle - [-\pi(W/p)] \}.$$

— É também possível mostrar que pelo menos um dos resultados do lema de Hotelling pode ser encarado como uma relação entre pontos conjugados. O lema de Hotelling estabelece nomeadamente que (ver por exemplo Varian, [6]):

$$x_i(p, W) = -\delta \Pi(p, W) / \delta W_i \quad (i = 1, \dots, n)$$

Ora, desenvolvendo, fica:

$$x_i(p, W) = \delta(-\Pi/p) / \delta(W_i/p) \iff x(p, W) = \nabla_{W/p}(-\Pi/p)$$

$$\iff x(p, W) = \nabla g^*(W/p)$$

o que quer dizer, portanto, e se atendermos ao corolário do teorema 2.C, que as funções procura de factores por parte da firma mais não fazem que dar, para cada vector de preços normalizados W/p , o respectivo ponto conjugado no espaço dos inputs — $x(p, W)$.

2.2. FUNÇÕES SUPORTE DE CONJUNTOS CONVEXOS

2.2.1. Funções suporte como conjugada da função indicadora

— Seja a função indicadora dum conjunto convexo C , que definimos na alínea 1.2.2. como a função convexa:

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

a conjugada desta função indicadora poderá ser obtida fazendo:

$$\begin{aligned} \delta^*(x^*/C) &= \sup_x \{ \langle x^*, x \rangle - \delta(x/C) \} \\ &= \delta^*(x^*/C) = \sup_{x \in C} \{ \langle x^*, x \rangle \} \end{aligned}$$

a esta definição assim definida como a conjugada da função indicadora dá-se o nome de função suportado conjunto C . Ela traduz um tipo bastante comum de problemas, em que se pretende maximizar uma função linear $\langle x, x^* \rangle$ num domínio convexo C (16).

— TEOREMA 2.D (Rockafellar, [1], p. 112, teorema 13.1)

. Seja C um conjunto convexo. Então $x \in C$ sse

$$\langle x, x^* \rangle \leq \delta^*(x^*/C), \quad \forall_{x^*}$$

Significa o anterior teorema que um dado conjunto convexo fechado C pode ser expresso como o conjunto de soluções dum sistema de desigualdades gerada pela sua função suporte:

$$C = \{ x / \langle x, x^* \rangle \leq \delta^*(x^*/C), \quad \forall_{x^*} \}$$

Dito de outra forma, é possível recuperar completamente o conjunto C a partir da sua função suporte, o que não admira, já que a conjugada da função suporte é por definição a função indicadora do conjunto em questão.

(16) A minimização duma função linear num conjunto convexo também pode ser traduzida por $\inf \{ \langle x, x^* \rangle / x \in C \} = - \sup \{ \langle x, -x^* \rangle / x \in C \} = -\delta^*(-x^*/C)$

2.2.2. Exemplo de aplicação: função produção e função custo (17)

— Tornemos a considerar a empresa concorrencial com função de produção $y = g(x)$ concava e fechada da alínea 2.1.6.

— A função custo da empresa é definida como dando o mínimo custo para produzir o nível de output y :

$$C(w, y) = \inf_x \langle w, x \rangle$$

$$\text{s.a. } g(x) \geq y$$

$$\Leftrightarrow C(w, y) = \inf_x \{ \langle w, x \rangle / x \in C \}$$

em que $C = \{x / g(x) \geq y\}$, designado por "Input Requirement Set", é um conjunto convexo fechado já que $g(x)$ é por hipótese uma função concava fechada.

$$\text{— Ora, } C(w, y) = \inf_x \{ \langle w, x \rangle / x \in C \}$$

$$\Leftrightarrow -C(w, y) = \sup_x \{ \langle -w, x \rangle / x \in C \}$$

$$\Leftrightarrow \delta^*(-w/C) = -C(w, y)$$

Por palavras, $-C(w, y)$ é a função suporte do conjunto C que define as características da tecnologia da empresa. Resulta do anterior que será possível recuperar a função de produção da empresa a partir do conhecimento da sua função custos. Com efeito:

$$[-C(w, y)]^* = \delta(x/c) = \{x / g(x) \geq y\}$$

Identicamente, poderemos também obter C recorrendo ao teorema 2.D, fazendo:

$$C = \{x / \langle -w, x \rangle \leq \delta^*(-w/c), \quad \forall_w\}$$

$$\Leftrightarrow C = \{x / \langle -w, x \rangle \leq -C(w, y), \quad \forall_w\}$$

$$\Leftrightarrow C = \{x / \langle x, w \rangle \geq C(w, y), \quad \forall_w\} \quad (18)$$

(17) Devido à analogia formal entre a análise do produtor e do consumidor, o exemplo que se segue pode ser transposto para a análise do consumidor, para o que bastará encarar $g(x)$ como a função utilidade.

e logo, atendendo à definição de função de produção:

$$g(x) = \sup \{y / \langle w, x \rangle \geq C(w, y), \quad \forall w\}$$

— Tal como mostrámos que as funções procura de factores estabelecem correspondências entre pontos conjugados, é possível mostrar agora, a partir do lema de Shephard, que as funções procura condicionais ou derivadas (19) também podem ser encaradas da mesma maneira. O lema de Shephard estabelece que:

$$\partial C(w, y) / \partial w_i = x_i(w, y) \quad (i = 1, \dots, n)$$

Ora,

$$\nabla_w C(w, y) = x(w, y)$$

$$\iff \nabla_{-w} [-C(w, y)] = x(w, y)$$

$$\iff \nabla_{(-w)} \delta^*(-w, y) = x(w, y)$$

Quer isto dizer, e atendendo ao corolário do teorema 2.C, que as funções procura condicionais fazem corresponder, para cada nível de output, a cada vector w (ou $-w$), o seu vector conjugado no espaço dos outputs $x(w, y)$.

(18) — Pelo que vimos e atendendo às propriedades das funções conjugadas, é fácil de perceber que quando os "input requirement sets" não são originalmente convexos e fechados, o resultado da sua recuperação a partir da função custo corresponderá a uma convexização fechada dos mesmos. Como as partes não convexas da fronteira destes conjuntos não são relevantes do ponto de vista económico, a função custo sintetisa mesmo neste caso todos os aspectos relevantes de tecnologia.

(19) — Funções procura de factores para um dado nível de output. Na teoria do consumidor recebem a designação de funções procura hicksianas.

PARTE III

DUALIDADE EM PROGRAMAÇÃO

CONVEXA

3.1. O PROBLEMA DE PROGRAMAÇÃO CONVEXA

3.1.1. O problema generalizado de programação convexa

— Seja o problema básico de programação convexa:

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

com $x \in \mathbb{R}^n$, $f(x)$ uma função real convexa, e $G(x)$ um vector de m funções também reais e convexas (9). Relacionado com (C) é conveniente considerar a família de problemas

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

com u vector arbitrário de $\mathbb{R}^m$. Como é óbvio, a solução de (D), a que Rochafellar chama de programa generalizado de programação convexa, depende da "perturbação" u introduzida.

3.1.2. Bifunção associada e função perturbação

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

(9) — O problema (A) não é restritivo na sua forma. Com efeito, basta ver que:

- 1) $\psi(x) \leq b \iff \psi(x) - b \leq 0$
- 2) $\psi(x) \geq b \iff -\psi(x) + b \leq 0$
- 3) $\psi(x) = b \iff \begin{array}{l} \psi(x) - b \leq 0 \\ -\psi(x) + b \leq 0 \end{array}$
- 4) $x_j \leq 0 \iff u_j \geq 0 \text{ com } u = -x_j$
- 5) $x_j \text{ livre} \iff x_j = u_{j1} - u_{j2} \text{ com } u_{j2} \geq 0$

será designada por "função grafo" (graph function) de F. Muito simplesmente, uma função 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 formado um determinado vector u $\mathbb{R}^m$. O termo bi-função" resulta portanto do facto de ela corresponder afinal a uma primeira 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 função especial, que consiga traduzir a família de problemas (D). 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 convexo generalizado (D). 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 (D) 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 $\mathbb{R}^m$ em $[-\infty, +\infty]$, que a cada vector u faz corresponder a solução óptima de (D) para esse vector de perturbações:

$$\begin{aligned} w(u) &= \inf_x \{f(x)/x \geq 0; G(x) \leq u\} = \inf_x [f(x) + \delta(x/x \geq 0; G(x) \leq u)] \\ &= w(u) = \inf_x (Fu)(x) \end{aligned}$$

Sendo assim, o problema original (C) pode ser encarado como a determinação de $w(0)$, ou, equivalentemente, a determinação de $\inf_x (F0)(x)$.

3.1.3. Algumas propriedades da bifunção associada e da função per-turbação

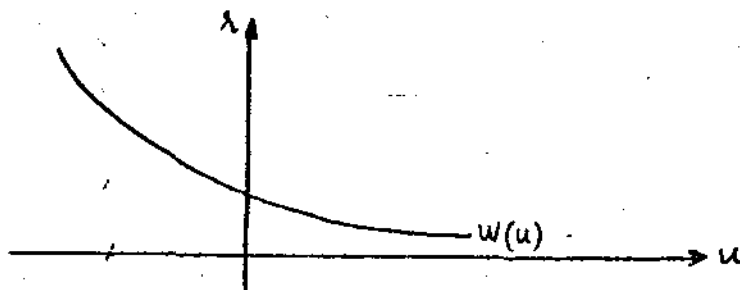
— Propriedade 1.1: A bifunção associada ao programa convexo generalizado (D) é convexa, em que se considera uma bifunção convexa sse a respectiva função grafo o é. Ora, $f(u, x) = f(x) + \delta(x/x \geq 0; G(x) \leq u)$ é convexa dado serem convexas por hipótese e $\delta(x/.)$ por de-finição (como sabemos,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

Propriedade 1.2: O programa convexo generalizado (D) é u-nicamente determinado pela respectiva bifunção associada (F_u), isto é, podemos sempre recuperar o programa (D) a partir do conhecimento de F_u . Para tal tenha-se em atenção que, para um dado u , $S = \{x / (F_u)(x) < +\infty\}$ 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 Luenber-ger, [2], pag. 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, dimi-nuindo as perturbações u), nunca poderei melhorar a solução ótima do problema.

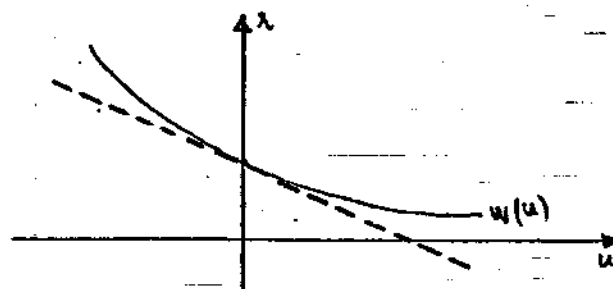
— Atendendo às propriedades 2.1 e 2.2, uma representação gráfica típica de $w(u)$, para $u \in \mathbb{R}^1$, será:



3.1.4. Relação entre função auxiliar Lagrangeana e função perturbação/multiplicadores de Lâgrange e teoria de Kuhn-Tucker

— 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.

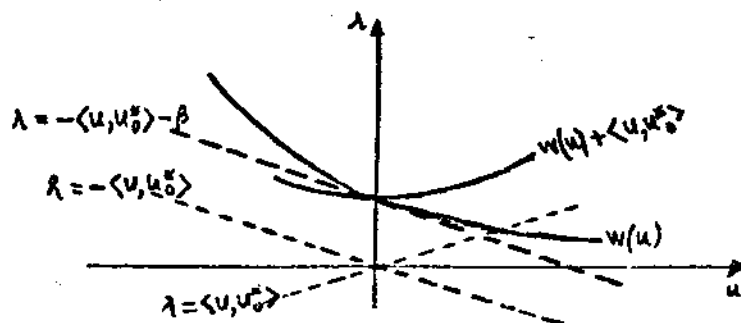
Ilustração em $\mathbb{R}^2$:



Se rodássemos a cabeça por forma a que o hiperplano tangente se tornasse a nova 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

$$\inf \{w(u) + \langle u, u_0^* \rangle\} = w(0) + \langle 0, u_0^* \rangle = w(0) \rightarrow \text{sol. óptima problema (C).}$$

Ilustração em $\mathbb{R}^2$:



Ora, demonstra-se com facilidade que, para quaisquer $u^* \geq 0$ e $u \in \mathbb{R}^m$,

$$\inf_{x \geq 0} [w(u) + \langle u, u^* \rangle] = \inf_{x \geq 0} [f(x) + \langle G(x), u^* \rangle] \quad (10)$$

isto é, ao minimizarmos, conforme pretendido, $w(u) + \langle u, u^* \rangle$, estamos afinal a minimizar a função auxiliar Lagrangeana $L = f(x) + \langle G(x), u^* \rangle$. 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 dos multiplicadores de Lagrange associados à solução óptima do problema (C) (11).

— Sintetisemos e precisemos o que foi dito através do seguinte teorema:

— TEOREMA 3.A (Luenberger, [2], pag. 217/218, versão ligeiramente modificada)

. Seja o problema (C). Assume-se a existência dum ponto $x^1 > 0$ tal que $G(x^1) < 0$ (isto é, assume-se que o conjunto das soluções admissíveis tem interior não vazio). Considere-se η_0 como o valor óptimo do problema (C), existente e finito por hipótese.

. Então, existe também um vector $u_0^* \geq 0$ de $\mathbb{R}^m$, tal que

$$\eta_0 = \inf_{x \geq 0} \{f(x) + \langle G(x), u_0^* \rangle\} = \inf_{x \geq 0} L(x, u_0^*)$$

mais ainda, se fôr x_0 a solução óptima do problema (C), então

$$\langle G(x_0), u_0^* \rangle = 0 \text{ e logo } f(x_0) = L(x_0, u_0^*).$$

$$(10) \quad \text{— Por um lado temos } \inf_{x \geq 0} \{f(x) + \langle G(x), u^* \rangle\} \leq \inf_{x \geq 0} \{f(x) + \langle u, x^* \rangle / G(x) \leq u, x \geq 0\} = \\ = \inf_{x \geq 0} \{f(x) / G(x) \leq u, x \geq 0\} + \langle u, u^* \rangle =$$

Por outro lado, para cada $x^1 \geq 0$, temos, com $u^1 = G(x^1)$,

$$f(x^1) + \langle G(x^1), u^* \rangle \geq \inf_{x \geq 0} \{f(x) + \langle u^1, u^* \rangle / G(x) \leq u^1, x \geq 0\} = \\ = \inf_{x \geq 0} \{f(x) / G(x) \leq u^1, x \geq 0\} + \langle u^1, u^* \rangle = \\ = w(u^1) + \langle u^1, u^* \rangle$$

$$\Leftrightarrow \inf_{x^1 \geq 0} \{f(x^1) + \langle G(x^1), u^* \rangle\} \geq \inf_{x^1 \geq 0} \{w(u^1) + \langle u^1, u^* \rangle\}$$

Logo, teremos que ter a igualdade $\inf_{x \geq 0} \{f(x) + \langle G(x), u^* \rangle\} = \inf_{x \geq 0} \{w(u) + \langle u, u^* \rangle\}$ (Luenb., [2], p. 224).

Uma versão equivalente dos resultados do teorema 3.A pode ser estabelecida em termos de ponto de sela da Lagrangeana:

— COROLÁRIO (por exemplo em Luenberger, [2], pag. 219)

. Continuem a admitir-se as hipóteses do teorema 3.A, e seja também x_0 a solução ótima. Então existe um vector $u_0^* \geq 0$ tal que a função Lagrangeana

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

tem um ponto de sela em (x_0, u_0^*) , isto é:

$$L(x_0, u^*) \leq L(x_0, u_0^*) \leq L(x, u_0^*), \quad \forall x \geq 0, u^* \geq 0$$

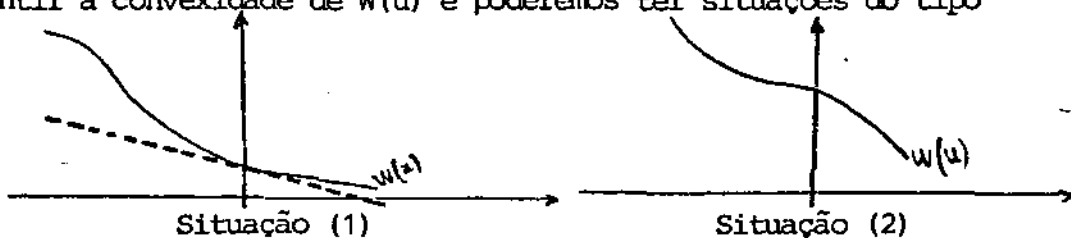
— Repare-se, a propósito que as condições de Kuhn-Tucher, sendo condições de 1ª ordem do problema

$$\sup_{u^* \geq 0} \inf_{x \geq 0} L(x, u^*)$$

são por isso também condições de 1ª ordem do problema (C). A hipótese que fizemos de o conjunto das soluções admissíveis ter interior não vazio, e já que estamos a trabalhar em programação convexa, garante-nos que as condições de K.-T. são neste caso necessárias de ótimo (condição de qualificação de Slater); A convexidade do problema garante-nos igualmente que elas são também condições suficientes (teorema de suficiência de K.-T.). (12)

(11) — Como curiosidade, repare-se que os multiplicadores de Lagrange, ao definirem o hiperplano tangente a $w(u)$ para $u=0$, $-\langle u, u_0^* \rangle - r = \beta$, constituem por isso o simétrico do vector gradiente de $w(u)$ neste mesmo ponto: $\nabla w(0) = -u_0^*$. Sendo assim, podemos interpretá-los tal como foram definidos, como dando o simétrico da sensibilidade do valor ótimo do problema (C) à introdução de perturbações positivas infinitesimais.

(12) — No caso de omitirmos as hipóteses de convexidade já não poderemos garantir a convexidade de $W(u)$ e poderemos ter situações do tipo



No situação 1, e apesar da não convexidade, a técnica indicada ainda se aplica devido ao hiperplano tangente a $w(u)$ para $u=0$ ser majorizado por aquela função em todo o seu domínio. Todavia, na situação 2 a técnica de Lagrange/K.-T. já não nos conduziria ao ótimo do problema (C).

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 concavo generalizado. Na maioria dos casos, os dois programas duais têm o mesmo valor ótimo da função objectivo e as soluções óptimas para um são os multiplicadores de Lagrange para outro" (Rokafellar, [1], pag. 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 3.2.3, existe uma relação estreita, donde resulta a designação, entre adjuntas de bifunção e transformações lineares adjuntas.

— Definimos adjunta duma bifunção convexa $(Fu)(x)$ de $\mathbb{R}^n$ 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 duma bifunção concava 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 (13). Como f é convexa por hipótese, $f^*(-u^*,x^*)$ é, pelo que sabemos, uma função convexa fechada, e logo a bifunção adjunta de F

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

será concava 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:

— TEOREMA 3.B (Rockafellar, [1], pag. 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 sse 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 $\mathbb{R}^n$, e seja (P) o programa convexo gsnralizado a ela associado. O programa concavo generalizado associado com a bifunção adjunta F^* (concava) será designado por programa dual de (P) , e será representado por (P^*) .

— 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 partimos, já que $F^{**} = cl F$. 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 não generalizado (C) é, como sabemos, $W(0) = \inf_x (F0)(x)$, o valor ótimo do seu problema dual será $\sup_{u^*} (F^*0)(u^*)$.

3.2.3. 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 dum 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

$$(Fu)(x) = \delta(x/x = Au) = \begin{cases} 0 & \text{se } x = Au \\ + & \text{se } x \neq Au \end{cases}$$

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:

$$\begin{aligned} (F^*x^*)(u^*) &= \inf_{x,u} \{ \delta(x/x=Au) - \langle x, x^* \rangle + \langle u, u^* \rangle \} = \\ &\quad (\text{substituindo } x \text{ por } Au) \\ &= \inf_u \{ -\langle Au, x^* \rangle + \langle u, u^* \rangle \} = \\ &\quad (\text{pela definição de matriz adjunta } A^*(14)) \\ &= \inf_u \{ \langle u, u^* - A^*x^* \rangle \} = \\ &= \begin{cases} 0 & \text{se } u^* = A^*x^* \\ -\infty & \text{se } u^* \neq A^*x^* \end{cases} \\ &= -\delta(u^*/u^* = A^*x^*) \end{aligned}$$

(14) — Chama-se adjunta dum transformação linear A , e representa-se por A^* , a transformação (também linear) definida por: $\langle Az, u^* \rangle = \langle z, A^*u^* \rangle$.

Quer isto dizer que a adjunta de uma 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 dual. Tome-se o programa linear

$$\begin{array}{ll} \text{Inf} & \langle a^*, x \rangle \\ x & \\ \text{s.a.} & \left\{ \begin{array}{l} a - Ax \leq 0 \\ x \geq 0 \end{array} \right. \end{array} \quad \dots (E)$$

A generalização de (E) será:

$$\begin{array}{ll} \text{Inf} & \langle a^*, x \rangle \\ x & \\ \text{s.a.} & \left\{ \begin{array}{l} a - Ax \leq u \\ x \geq 0 \end{array} \right. \end{array} \quad \dots (F)$$

Como bifunção associada a (F) teremos:

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

Calculando pela definição adjunta F^* , vem:

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

O problema dual de (E) será então o problema de maximização de (F^*0) 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 .

3.2.4. 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.C (versão modificada de Rockafellar, [1], pag. 314)

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

$$[-W(u)]^* = [-\inf_x (Fu)(x)]^* = (F^*0)(u^*); \quad (F^*0)^* = -cl[W(u)]$$

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

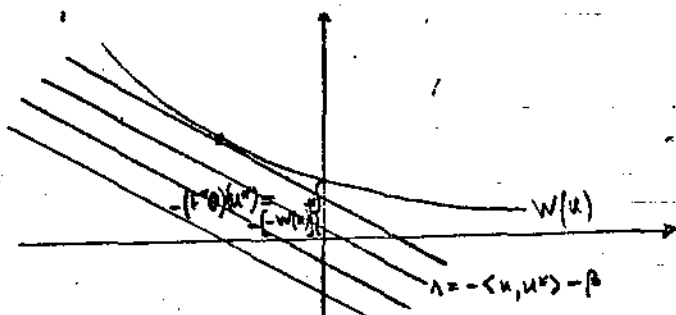
$$[-\bar{W}(x^*)]^* = [-\sup_{u^*} (F^*x^*)(u^*)]^* = (F0)^* = -cl[\bar{W}(x^*)]$$

— Consideremos apenas o primeiro resultado estabelecido no teorema:

$$(F^*0)(u^*) = [-W(u)]^*$$

Conforme apontamos na parte II (alínea 2.1.2 e 2.1.5), a conjugada de $-W(u)$ dar-nos-á, para cada u^* previamente fixado, o simétrico da menor ordenada na origem possível das funções afins $r = \langle u, u^* \rangle$ - 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 $W(u)$.

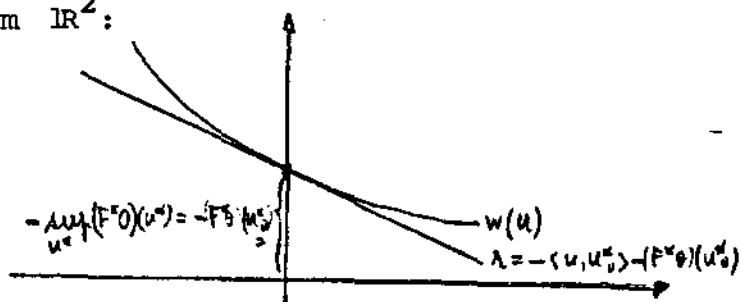
Ilustração em $\mathbb{R}^2$ para um dado u^* previamente fixado:



Logo, $(F^*0)(u^*)$ dá-nos, para cada u^* , o máximo da ordenada na origem das funções lineares afins $r = -\langle u, u^* \rangle - \beta$ majoradas por $W(u)$.

— A maximização de $(F^*0)(u^*)$, que é o programa concavo dual do programa convexo (C), 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 óptima do problema dual vai ser precisamente o vector u_0^* tal que $r = -\langle u, u_0^* \rangle - (F^*0)(u_0^*)$ 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 é mais que o vector dos valores no óptimo dos multiplicadores de Lagrange do problema primal.

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

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

$$(F^*0)(u^*) = [-W(u)]^*$$

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

$$\begin{aligned} (F^*0)(u^*) &= \inf_u \{ \langle u, u^* \rangle - [-W(u)] \} = \inf_u \{ W(u) + \langle u, u^* \rangle \} = \\ &\quad (\text{conforme demonstrado na alínea 3.1.4}) \\ &= \inf_{x \geq 0} \{ f(x) + \langle G(x); u^* \rangle \}, \quad \forall u^* \geq 0 \\ &= \inf_{x \geq 0} L(x, u^*), \quad \forall u^* \geq 0 \quad (15) \end{aligned}$$

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

$$\sup_{x^* \geq 0} (F^*0)(u^*) = \sup_{x^* \geq 0} \inf_{x \geq 0} L(x, u^*)$$

Dito de outra forma, o óptimo do problema dual não generalizado corresponde ao ponto de sela no ortante positivo de $\mathbb{R}^{m+n}$ da Lagrangeana do problema primal. Justifica-se assim 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.

(15) — Alternativamente, este mesmo resultado pode ser obtido recorrendo à determinação da bifunção adjunta do problema convexo generalizado primal. Com efeito, Rockafellar [1], p. 320) mostra que:

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

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

— Acabámos portanto de mostrar que existe uma estreita relação entre a chamada dualidade Lagrangeana e a teoria geral da dualidade conjugada. Tal constitui mais um importante factor a favor do carácter fortemente globalizante e unificador desta última, a qual consegue também englobar outras abordagens da dualidade não referidas como por exemplo a denominada dualidade de Wolfe. Estes desenvolvimentos, bem como alguns outros, entre os quais se poderá destacar uma análise da dualidade conjugada em optimização dinâmica (e suas interpretações económicas), serão objecto de publicação numa próxima oportunidade.

MEMO & PARTS III

A. TEOREMA DE DUALIDADE DO FENCHEL

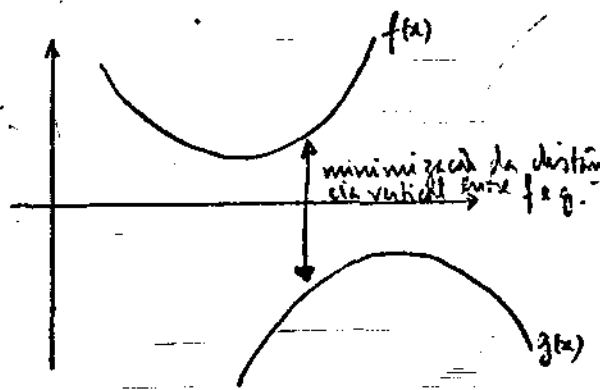
A. 1) Teorema de dualidade de Fenchel/sua interpretação gráfica

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

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

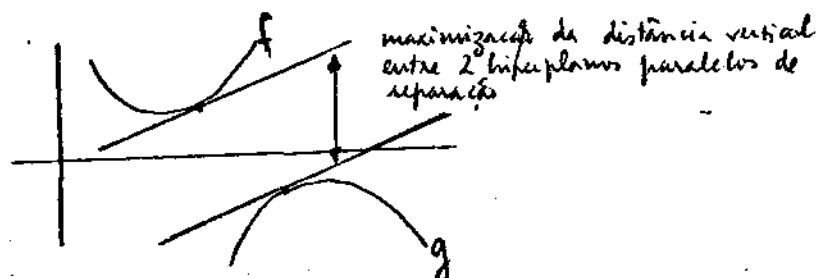
O problema (G) pode ser perspectivado graficamente como o problema de minimização da distância vertical entre 2 funções f e g .

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 (G) pode ser visto como o problema de maximização da distância vertical entre quaisquer dois hiperplanos de suporte de f e g que sejam paralelos.

Ilustração gráfica em $\mathbb{R}^2$:



— Ora, como vimos na parte 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 paralelos pode ser vista como a maximização da diferença entre as respectivas ordenadas na origem, temos então definido o problema dual do problema (G):

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

— Estamos então em condições de perceber a lógica do seguinte teorema, devido originalmente a Fenchel:

— TEOREMA DE DUALIDADE DE FENCHEL (TEOREMA A.1) (Rockafellar, [1], p. 327/Luenberger, [2], p. 201)

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

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

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 ínfi mo e o supremo são finitos.

A. 2) Teorema de Fenchel e teoria geral da dualidade convexa

— Como sublinha Rockafellar ([1], p. 327), o teorema de Fenchel pode ser deduzido independentemente da teoria geral da dualidade convexa, e recorrendo apenas à noção de função conjugada (nomeadamente na perspectiva da desigualdade de Fenchel) 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 da secção 3.2). Todavia, e alternativamente, como também demonstra Rockafellar ([1], pag. 329-331), os resultados de dualidade estabelecidos pioneiramente por Fenchel podem ser derivados da teoria geral da dualidade convexa, tal como ela foi desenvolvida na secção 3.2, que é indiscutivelmente mais operacional, mais rica analiticamente e mais conclusiva.