

UNIVERSIDADE NOVA DE LISBOA

*Faculdade de Economia*

DUALIDADE CONJUGADA E QUASE CONJUGADA NAS

TEORIAS DO PRODUTOR E CONSUMIDOR

PARTE I : FUNDAMENTOS DA ANÁLISE CONVEXA

Maximiano Reis Pinheiro

Working Paper Nº 41

UNIVERSIDADE NOVA DE LISBOA

FACULDADE DE ECONOMIA

Campo Grande, 185

1700 LISBOA

Janeiro de 1986

O autor agradece ao Professor Doutor José Dias Coelho, seu orientador, todo o estímulo e apoio prestados ao longo dos dois anos que já dura o projecto de trabalho em que se insere este estudo . O autor agradece ainda aos Professores Doutores Fernando Brito Soares e Diogo Lucena, e ao Dr. Vítor Gaspar, as sugestões fornecidas e a disponibilidade manifestada.

Todos os erros e omissões que possam existir são de exclusiva responsabilidade do autor.

"There are many alternative approaches to duality theory as there are individuals working in the field of duality theory. These different approaches may be approximately classified, at the risk of gross simplification, into three groups".

"The first group of approaches is based on the conjugacy correspondence developed by FENCHEL (1940<sup>(1)</sup>, 1953<sup>(2)</sup>) and extended by ROCKAFELLAR (1970). [...] the precursor of this conjugacy correspondence is of course the classical Legendre transformation. This line of reasoning was taken up by SAMUELSON (1953) and LAU (1969<sup>(3)</sup>) for the differentiable case, and by JORGENSON and LAU (1973) for the general case [...]"

"The second group of approaches is based on the symmetric duality between gauge functions, or distance functions, or polar cones of convex sets. SHEPHARD (1953) was the pioneer of this group. He proved the duality theorem between production and cost functions under the assumption of differentiability employing the concept of distance functions. Other scholars who have used similar approaches include GORMAN (1968<sup>(4)</sup>), McFADDEN (1973<sup>(5)</sup>), HANOCH (1973<sup>(6)</sup>) and JACOBSEN (1970, 1972<sup>(7)</sup>), among others.

- 
- (1) - FENCHEL, W. (1949) "On conjugate Convex Functions", Canadian J. Math., 1, 73-77.
  - (2) - FENCHEL, W. (1953) Convex Cones, Sets and Functions, Dept. Math Princeton U., Princeton.
  - (3) - LAU, L. J. (1969) "Some applications of Profit Functions". Mem. 86, Center for Research in Economic Growth, Stanford Univ., Stanford.
  - (4) - GORMAN, W. M. (1968) "Measuring the Quantities of Fixed Factors", in J.N. WOLFE (ed.), Value, Capital and Growth: Papers in Honour of Sir John Hicks, Chicago: Aldine.
  - (5) - McFADDEN, D.L. (1973) "Cost Revenue, and Profit Functions" e "The General Linear Profit Function" in McFADDEN, D.L (ed) An Econometric Approach of Production Theory, North-Holland.
  - (6) - HANOCH, G. (1973) "Generation of New Production Functions Through Duality" in McFADDEN (ed) An Econometric Approach to Production Theory, North-Holland.
  - (7) - JACOBSEN, S.E. (1972) "On Shephard's Duality Theorem", Journal of Econ.Theory, 4, 458-464.

"The third group of approaches is based on the duality between the set of production possibilities and its support functions. The work of UZAWA (1964) and McFADDEN (1966<sup>(8)</sup>) and DIEWERT (1971<sup>(9)</sup>), 1973<sup>(10)</sup>), may be classified amongst this group".

"Now all three groups of approaches must yield identical results if the same assumptions are made on the set of production possibilities. The choice among them is then essentially a matter of elegance, expositional facility, and taste".

in LAU, L.J. "Comments" anexos a DIEWERT,  
W.E. (1974)

— Neste estudo a dualidade conjugada convexa de FENCHEL e ROCKAFELLAR, e a dualidade quase-conjugada quase-convexa de PASSY e PRISMAN, são utilizadas para a obtenção das principais relações de dualidade das teorias do produtor e consumidor. Esclareça-se desde já que o objectivo não é fazer um "survey" da literatura sobre dualidade na teoria microeconómica, mas explorar uma via possível para a dedução integrada de tais relações de dualidade. Pretender-se-á ainda mostrar que a abordagem adoptada tem importantes vantagens de generalidade que permitem, nomeadamente, absorver como desenvolvimentos seus as outras duas vias referidas na citação inicial de LAU.

---

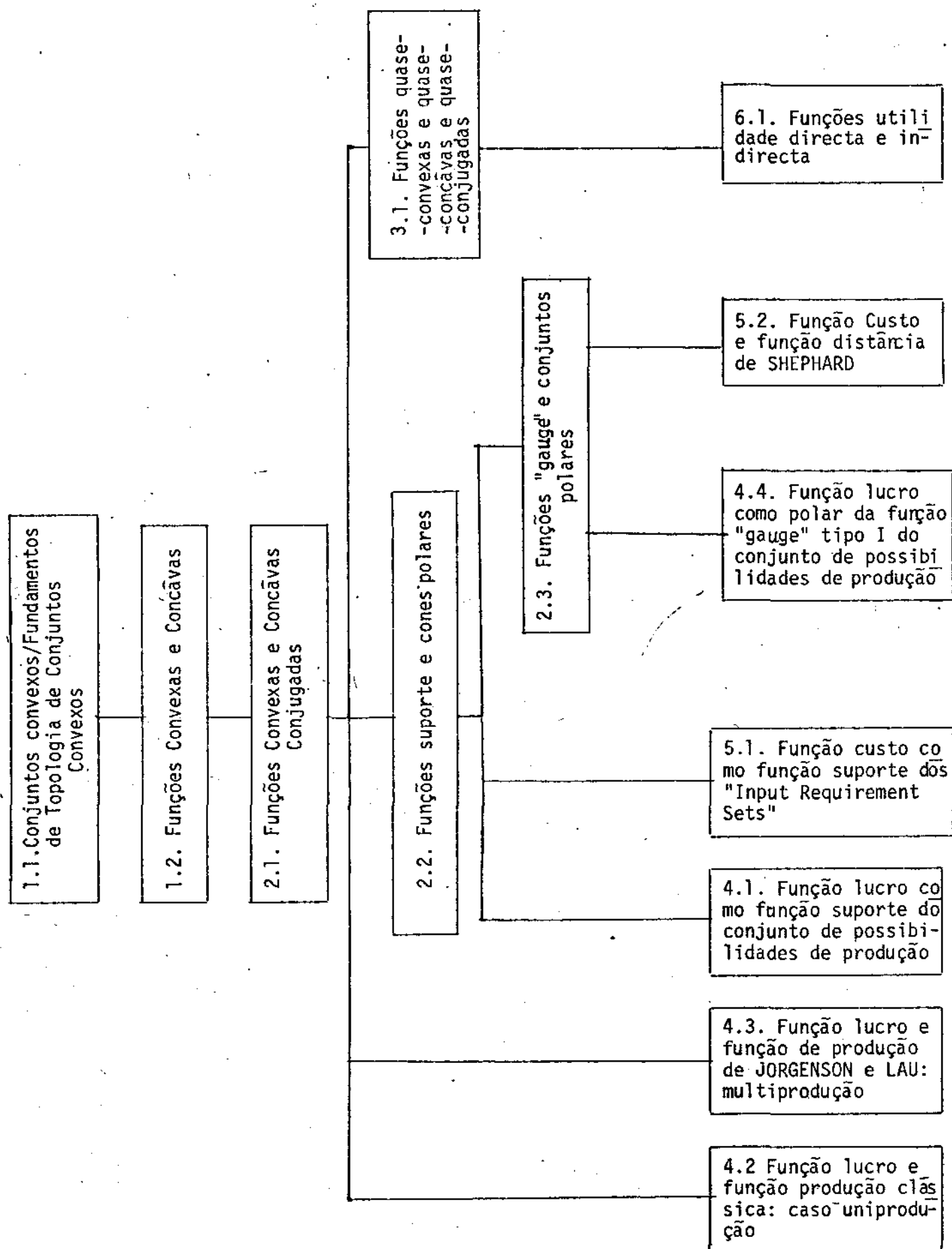
(8) - McFADDEN, D.L. (1966) "Cost, Revenue, and Profit Functions: A Cursory Review" Working Paper nº 86, Institute of Business and Economic Research, University of California at Berkeley.

(9) - DIEWERT, W.E. (1971) "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function", The Journal of Political Economy, 79, 481-507.

(10) - DIEWERT, W. E. (1973) "Functional Forms for Profit and Transformation Functions", Journal of Economic Theory.

—— O esquema de encadeamento lógico das várias secções dos seis capítulos deste estudo é apresentado na página seguinte. O primeiro capítulo (secções 1.1 e 1.2) estabelece os fundamentos de análise convexa necessários à compreensão de tudo o resto. A secção 2.1 sintetiza os resultados fundamentais da dualidade conjugada convexa, constituindo por isso o núcleo da dissertação. As secções 2.2 e 2.3 são desenvolvimentos particularizados da secção 2.1, que vêm confirmar que as 2ª e 3ª vias apontadas por LAU podem ser integradas na 1ª. Os capítulos 4 e 5 são aplicações directas do exposto nas secções 2.1, 2.2 e 2.3 às teorias do produtor e consumidor. Note-se, a propósito, que as duas últimas alíneas da secção 2.3 constituem uma extensão da nossa autoria, para conjuntos não contendo a origem, das funções "gauge"<sup>estudadas</sup> por ROCKAFELLAR (1970). Esta extensão tornava-se necessária para abarcar na nossa análise as relações de dualidade entre função custo e função distância de SHEPHARD (secção 5.2). Quanto ao capítulo 3, ele segue de perto a contribuição de PASSY e PRISMAN (1984), que vem permitir uma extensão interessantíssima das relações de conjugação, orientada para a análise quase-convexa. A sua apresentação neste estudo decorre do facto de ela permitir traduzir a dualidade entre funções utilidade directa e indirecta (capítulo 6).

—— Houve na elaboração deste estudo uma grande preocupação em produzir um texto auto-sustentado. Por outras palavras, foi feito um esforço no sentido de tornar mínimos os requisitos matemáticos para a leitura e compreensão do texto. Como seria de



esperar, contudo, esta nossa preocupação de começar pelo princípio, e de fundamentar adequadamente cada passo dado, não é isenta de custos. O maior deles será, provavelmente, a extensão do texto principal e dos anexos de demonstrações.

---

Ainda no âmbito desta introdução, resta referir que ao longo da dissertação, e a menos que explicitado o contrário, consideraremos o espaço  $\mathbb{R}^n$  cujos vectores representaremos, como é usual, por ênuplos ordenados  $x = (x_1, \dots, x_n)^T$ . Vamos admitir neste espaço a base natural ortonormada  $\{e_1, \dots, e_j, \dots, e_n\}$ , em que  $e_j$  é a  $j$ -ésima coluna da matriz identidade de ordem  $n$ . O produto escalar entre 2 vectores  $x^1$  e  $x^2$  representar-se-á por  $\langle x^1, x^2 \rangle = \sum_{j=1}^n x_j^1 x_j^2$ . Recordemos, a propósito, que definindo  $\cos(x^1, x^2) = \langle x^1, x^2 \rangle / \|x^1\| \cdot \|x^2\|$ , em que  $\|x\| = \langle x, x \rangle^{1/2}$  é a norma euclídiana de  $x$ , dois vectores não nulos  $x^1$  e  $x^2$  são ditos ortogonais se  $\cos(x^1, x^2) = 0$ , o que acontece, portanto, se  $\langle x^1, x^2 \rangle = 0$ .

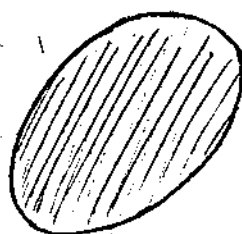
## 1.1. CONJUNTOS CONVEXOS/FUNDAMENTOS DE TOPOLOGIA DE CONJUNTOS CONVEXOS

### 1.1.1. Conjuntos convexos e envolvente convexa

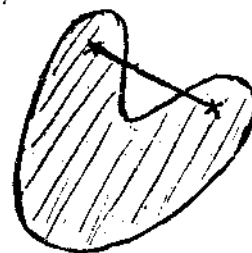
—— Definição: Um conjunto  $C \subset \mathbb{R}^n$  é convexo se

$$(1-\lambda)x^1 + \lambda x^2 \in C, \quad \forall x^1, x^2 \in C, \quad 0 \leq \lambda \leq 1 \quad ||$$

Por palavras, um conjunto é convexo se, contendo quaisquer dois vetores  $x^1$  e  $x^2$ , contem também todas as suas combinações lineares convexas, isto é, contem igualmente todos os pontos do segmento de recta que une  $x^1$  e  $x^2$ . (11)



conjunto convexo



conjunto não convexo

Figura 1.1.1.A: Ilustração de conjuntos convexos e não convexos

—— Entre as propriedades elementares dos conjuntos convexos destacamos:

Proposição 1.1.1.I: Intersecção de conjuntos convexos é ainda um conjunto convexo. ||

Proposição 1.1.1.II: Sejam 2 conjuntos  $C_1, C_2 \subset \mathbb{R}^n$  e seja  $C_1 + C_2 = \{x^1 + x^2 \mid x^1 \in C_1, x^2 \in C_2\}$ . Se  $C_1$  e  $C_2$  são conjuntos convexos então  $C_1 + C_2$  é um conjunto convexo. ||

Proposição 1.1.1.III: Sejam  $C \subset \mathbb{R}^n$  e  $\mu C = \{\mu x \mid x \in C\}$  para dado  $\mu \in \mathbb{R}$ . Se  $C$  é um conjunto convexo, então  $\mu C$  também é  $\forall \mu \in \mathbb{R}$ . ||

(11) - Por convenção o conjunto vazio é considerado um conjunto convexo, pois a sua consideração como tal não contraria nenhuma das propriedades dos conjuntos convexos.

Proposição 1.1.1. IV: Seja  $A$  uma transformação linear de  $\mathbb{R}^n$  em  $\mathbb{R}^m$ , e seja  $C \subset \mathbb{R}^n$  um conjunto convexo. Então o conjunto  $AC = \{Ax \mid x \in C\}$  é um conjunto convexo em  $\mathbb{R}^m$ . ||

———— Definição: Envoltente convexa de um conjunto  $D \subset \mathbb{R}^n$ ,  $\text{conv } D$ , é o mais pequeno conjunto convexo que contém  $D$ . ||

Proposição 1.1.1.V: A envoltente convexa de um conjunto  $D \subset \mathbb{R}^n$  é a intersecção de todos os conjuntos convexos que contém  $D$ . ||

### 1.1.2. Conjuntos afins e subespaços

———— Definição: Um conjunto  $V \subset \mathbb{R}^n$  é dito conjunto afim, ou variedade linear, se  $(1-\lambda)x^1 + \lambda x^2 \in V$ ,  $\forall x^1, x^2 \in V$ ,  $\lambda \in \mathbb{R}$ . ||

Tendo em conta que  $(1-\lambda)x^1 + \lambda x^2 = x^1 + \lambda(x^2 - x^1)$ , verifica-se que se um conjunto afim contém dois pontos, contém igualmente toda a linha recta que passa por eles. Note-se que a definição de conjunto afim corresponde à de conjunto convexo, excepto no facto de ao escalar  $\lambda$  ser permitido assumir qualquer valor real e não apenas no intervalo  $[0,1]$ .

Proposição 1.1.2.I: Um conjunto afim é um conjunto convexo. ||

Definição: Envoltente afim de um conjunto convexo  $C \subset \mathbb{R}^n$ ,  $V(C)$ , é o menor conjunto afim que contém  $C$ . ||

———— Definição: Um conjunto  $S \subset \mathbb{R}^n$  é um subespaço de  $\mathbb{R}^n$  se é fechado na adição e na multiplicação por escalares reais, isto é, se

$$(i) \quad x^1 + x^2 \in S, \quad \forall x^1, x^2 \in S;$$

$$(ii) \quad \mu x \in S, \quad \forall x \in S, \mu \in \mathbb{R}. \parallel$$

Note-se que, por ser fechado na multiplicação por escalar, um subespaço contém necessariamente a origem.

Definição: Dimensão de um subespaço  $S \subset \mathbb{R}^n$  é o número máximo de vectores linearmente independentes que nele podemos encontrar, em que os vectores  $x^1, \dots, x^m \in S$  são linearmente independentes se  $\mu_1 x^1 + \dots + \mu_m x^m = 0$  apenas com  $\mu_1, \dots, \mu_m$  escalares reais todos nulos.  $\parallel$

O próximo resultado estabelece a relação entre os conceitos de conjunto afim e de subespaço.

Proposição 1.1.2.II: Os subespaços são conjuntos afins que contêm a origem, e os conjuntos afins que contêm a origem são subespaços.  $\parallel$

Definição: Dois conjuntos afins  $V_1, V_2 \subset \mathbb{R}^n$  são ditos paralelos se cada um deles pode ser expresso como o resultado de uma translação do outro, isto é, se existe  $a \in \mathbb{R}^n$  tal que:

$$V_1 = V_2 + \{a\} = \{x+a \mid x \in V_2\} \quad \text{e} \quad V_2 = V_1 + \{-a\} = \{x-a \mid x \in V_1\} . \parallel$$

Proposição 1.1.2. III: Cada conjunto afim não vazio é paralelo a um único subespaço.  $\parallel$

Estamos agora em condições de definir dimensão de um conjunto afim por relação com a dimensão do único subespaço que lhe é paralelo.

Definição: Dimensão de um conjunto afim não vazio é a dimensão do subespaço que lhe é paralelo.  $\parallel$

Por convenção, a dimensão do conjunto afim vazio é -1. Os conjuntos afins de dimensão 0, 1 e 2 são respectivamente pontos, rectas e planos em  $\mathbb{R}^n$ .

### 1.1.3. Hiperplanos e semiespaços

————— Definição: Um hiperplano em  $\mathbb{R}^n$  é um conjunto afim de dimensão  $n-1$ . ||

Proposição 1.1.3.I: O conjunto  $H = \{x \mid \langle x, b \rangle = \beta\}$ , com  $b$  vector de  $\mathbb{R}^n$  não nulo e  $\beta$  escalar real, é um hiperplano de  $\mathbb{R}^n$ , e qualquer hiperplano em  $\mathbb{R}^n$  pode ser representado desta forma. ||

É conveniente realçar o seguinte:

(i) Quando  $\beta=0$ , o hiperplano passa pela origem e coincide com o subespaço que lhe é paralelo.

(ii) Devido ao facto de qualquer hiperplano ter dimensão  $n-1$ , existe um grau de liberdade na respectiva equação de definição, pelo que os conjuntos  $\{x \mid \langle x, b \rangle = \beta\}$  e  $\{x \mid \langle x, \frac{1}{\alpha} b \rangle = \frac{\beta}{\alpha}\}$ , com  $\alpha$  escalar não nulo, definem o mesmo hiperplano.

(iii) O vector  $b$  (e, consequentemente, qualquer vector  $\frac{1}{\alpha} \cdot b$  com  $\alpha \neq 0$ ), da equação de definição do hiperplano  $H = \{x \mid \langle x, b \rangle = \beta\}$ , é um vector ortogonal a todos os vectores do subespaço  $S = \{x \mid \langle x, b \rangle = 0\}$ . Por isso, o vector  $b$  é designado por vector normal ao hiperplano  $H$ .

(iv) Hiperplanos com vectores  $b$  colineares são paralelos ao mesmo subespaço.

————— O conceito de hiperplano é completado pelo de semiespaços por ele gerados.

Definição: Os conjuntos  $H^> = \{x \mid \langle x, b \rangle > \beta\}$ ,  $H^< = \{x \mid \langle x, b \rangle < \beta\}$ ,  $H^> = \{x \mid \langle x, b \rangle > \beta\}$  e  $H^< = \{x \mid \langle x, b \rangle < \beta\}$  são designados por semiespaços gerados pelo hiperplano  $H = \{x \mid \langle x, b \rangle = \beta\}$ .

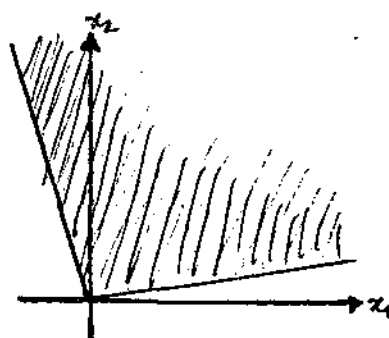
É especialmente importante a seguinte propriedade elementar dos semiespaços.

Proposição 1.1.3.II: Os semiespaços são conjuntos convexos. ||

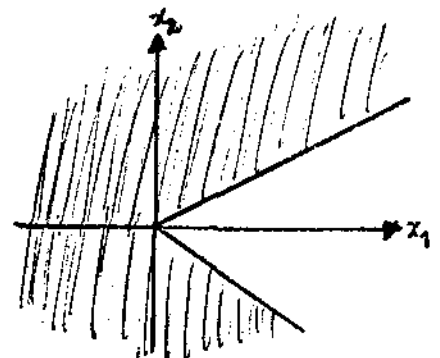
#### 1.1.4. Cones Convexos e respectivos cones polares

Definição: Um conjunto  $K \subset \mathbb{R}^n$  é um cone se for fechado na multiplicação por escalar real positivo, isto é, se  $\lambda x \in K, \forall x \in K, \lambda > 0$ . ||

Um cone  $K$  pode ser encarado como a união de semilinhas rectas emanadas da origem, a qual pode estar ou não contida em  $K$ . Um cone convexo será um cone que seja um conjunto convexo. (12)



cone convexo



cone não convexo

Figura 1.1.4. A: Ilustração de cones convexos e não convexos

(12)— Conforme nota ROCKAFELLAR (1970, p.13), muitos autores só consideram um cone como convexo se, para além de ser conjunto convexo, contiver também a origem, isto é, for fechado na multiplicação por escalares não negativos. Não será, contudo, a nossa posição.

Proposição 1.1.4.I: A intersecção de cones convexos é ainda um cone convexo. ||

Definição: Seja  $C \subset \mathbb{R}^n$  um conjunto convexo. O cone  $K = \{\lambda x \mid \lambda > 0, x \in C\} = \{x \mid \lambda > 0, x \in \lambda C\}$  é designado por cone gerado por  $C$ . ||

Proposição 1.1.4.II: O cone gerado por um conjunto convexo  $C \subset \mathbb{R}^n$  é o mais pequeno cone convexo que contém  $C$ . ||

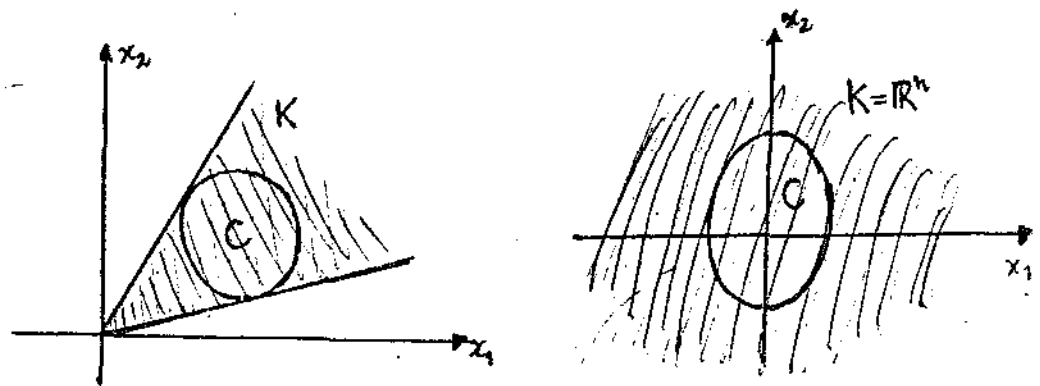


Figura 1.1.4.B: Ilustrações de cones gerados por conjuntos convexos.

Pela definição de cone convexo e pela proposição 1.1.4.I são cones convexos, entre outros, os seguintes conjuntos:

- (i) os subespaços de  $\mathbb{R}^n$  (pois, por definição, são conjuntos fechados na multiplicação por escalares reais);
- (ii) os hiperplanos que passam pela origem (que são, por definição, subespaços de dimensão  $n-1$ );
- (iii) os semiespaços gerados por hiperplanos que passam pela origem<sup>(13)</sup>;
- (iv) as intersecções de semiespaços gerados por hiperplanos que passam pela origem (devido à proposição 1.1.4.I).

Um exemplo importante de (iv) é o ortante não negativo de  $\mathbb{R}^n$ , que

(13)— Tomemos, por exemplo, o semiespaço  $H^+_{\langle x, b \rangle} = \{x \mid \langle x, b \rangle \geq 0\}$ . É imediato que este conjunto é fechado na multiplicação por escalar positivo. O conjunto  $H^+$  é portanto um cone. Sendo conjunto convexo (Proposição 1.1.3.II),  $H^+$  é um cone convexo.

representaremos por  $\mathbb{R}_+^n$ .

——— Definição: Cone polar de um cone convexo  $K \subset \mathbb{R}^n$  é o conjunto

$$K^0 = \{x^* \mid \langle x, x^* \rangle \leq 0, \forall x \in K\} \quad . ||$$

Proposição 1.1.4.III: O cone polar de um cone convexo  $K \subset \mathbb{R}^n$  é um cone convexo. ||

Repare-se que  $K^0$  é formado pelo vector nulo e por todos os vectores não nulos  $x^*$  tais que  $\cos(x, x^*) = \langle x, x^* \rangle / \|x\| \cdot \|x^*\| \leq 0, \forall x (\neq 0) \in K$ .

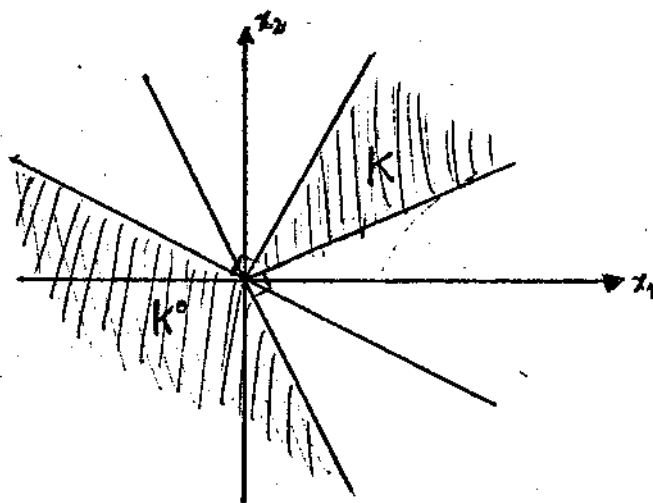


Figura 1.1.4.C: Cone convexo  $K$  e seu cone polar  $K^0$

### 1.1.5. Fundamentos de topologia de conjuntos convexos em $\mathbb{R}^n$

——— Representemos por  $E_\epsilon(x)$  a hipersfera em  $\mathbb{R}^n$  de centro  $x$  e raio  $\epsilon > 0$ :  $E_\epsilon(x) = \{y \mid d(x, y) \leq \epsilon\}$

em que  $d(x, y) = \|x - y\| = \langle x - y, x - y \rangle^{1/2}$  é a distância euclidiana entre  $x$  e  $y$  (dada pela norma euclidiana de  $x - y$ ). É fácil de comprovar que  $E_\epsilon(x)$  é um conjunto convexo.

Definição: Fecho de um conjunto  $D \subset \mathbb{R}^n$ , cl  $D$  (do inglês "closure") é o conjunto  $cl D = \{x \mid D \cap E_\epsilon(x) \neq \emptyset, \forall \epsilon \in D, \epsilon > 0\}$ . ||.

Proposição 1.1.5.I: Seja  $C \subset \mathbb{R}^n$  um conjunto convexo. Então  $\text{cl } D$  é um conjunto convexo.||

Definição: Um conjunto  $D \subset \mathbb{R}^n$  é fechado se coincide com o seu fecho, isto é, se  $D = \text{cl } D$ .||

Por convenção um conjunto vazio  $\emptyset$  é um conjunto fechado, já que a sua consideração como tal não contraria as propriedades dos conjuntos fechados não vazios. Note-se também que o próprio espaço  $\mathbb{R}^n$  é um conjunto fechado, como se confirma facilmente pela definição.

Proposição 1.1.5.II: Um conjunto  $D \subset \mathbb{R}^n$  é fechado se o limite de toda a sucessão convergente de seus elementos é ainda um elemento seu.||

Proposição 1.1.5.III: A intersecção de conjuntos fechados é um conjunto fechado.||

Proposição 1.1.5.IV: Seja um conjunto  $D \subset \mathbb{R}^n$ . Se  $D$  é fechado, então  $\mu D$  também é fechado,  $\forall \mu \in \mathbb{R}$ .||

Proposição 1.1.5.V: Seja  $C \subset \mathbb{R}^n$  um conjunto fechado. Então  $C + \{a\}$ , com  $a \in \mathbb{R}^n$ , é também um conjunto fechado.||

Proposição 1.1.5.VI: Seja um conjunto fechado  $T \subset \mathbb{R}^{n+m}$  e seja um dado  $\bar{y} \in \mathbb{R}^m$ . O conjunto  $X(\bar{y}) = \{x \mid x \in \mathbb{R}^n, y = \bar{y}, (x, y) \in T\}$  é um conjunto fechado.||

Definição: Um conjunto  $D \subset \mathbb{R}^n$  diz-se aberto se o seu conjunto complementar  $\bar{D} = \{x \mid x \notin D\}$  é um conjunto fechado.||

Definição: Interior de um conjunto  $D \subset \mathbb{R}^n$ ,  $\text{int } D$ , é o maior conjunto aberto contido em  $D$ .||

Equivalentemente,  $\text{int } D$  é o complementar do fecho do complementar de  $D$ :  $\text{int } D = \overline{\text{cl } D^c}$ .

Proposição 1.1.5.VII: Seja  $D \subset \mathbb{R}^n$ . Então  $\text{int } D = \{x \mid \exists \varepsilon > 0: E_\varepsilon(x) \subset D\}$  e  $D$  é aberto se  $D = \text{int } D$ . ||

Proposição 1.1.5.VIII: Seja  $C \subset \mathbb{R}^n$  um conjunto convexo. Então  $\text{int } C$  é também um conjunto convexo. ||

Definição: Fronteira de um conjunto  $D \subset \mathbb{R}^n$ ,  $\text{fr } D$ , é o conjunto de todos os pontos que pertencem a  $\text{cl } D$  e não pertencem a  $\text{int } D$ :

$$\text{fr } D = \text{cl } D \cap \overline{\text{int } D} = \{x \mid x \in \text{cl } D, x \notin \text{int } D\} . \quad ||$$

——— Definição: Interior e fronteira relativos de um conjunto  $D \subset \mathbb{R}^n$ , respectivamente  $\text{rint } D$  e  $\text{rfr } D$ , são o interior e a fronteira de  $D$  quando referidos não a  $\mathbb{R}^n$  mas sim à envolvente afim de  $D$ , isto é, ao menor conjunto afim que contém  $D$ . ||

Proposição 1.1.5.IX: Seja  $D \subset \mathbb{R}^n$  e seja  $V(D)$  a envolvente afim de  $D$ . Então:  $\text{rint } D = \{x \mid \exists \varepsilon > 0: (E_\varepsilon(x) \cap V(D)) \subset D\}$  e  $\text{rfr } D = \{x \mid x \in \text{cl } D, x \notin \text{rint } D\}$ . ||

Corolário 1.1.5.IX.1.: Seja  $D \subset \mathbb{R}^n$ . Então  $\text{int } D \subset \text{rint } D$ . ||

Proposição 1.1.5.X: Seja  $D \subset \mathbb{R}^n$  um conjunto convexo. Então  $\text{rint } D$  é um conjunto convexo. ||

Definição: Um conjunto  $D \subset \mathbb{R}^n$  diz-se relativamente aberto se  $D = \text{rint } D$ . ||

Proposição 1.1.5.XI: Seja  $D \subset \mathbb{R}^n$  um conjunto aberto. Então  $D$  é também relativamente aberto. ||

A noção de interior relativo é especialmente útil em casos em que o conjunto  $D$  está contido num conjunto afim de dimensão inferior a  $n$ . Com efeito, é óbvio pela proposição 1.1.5.VII que nesta situa-

ção temos  $\text{int } D = \emptyset$  (já que a hipersfera  $E_\epsilon(x)$  nunca poderá estar contida num conjunto com envolvente afim com dimensão menor que  $n$ ). Compreende-se assim a necessidade de um conceito menos exigente que o de interior, mas que mantenha as suas características mais importantes.

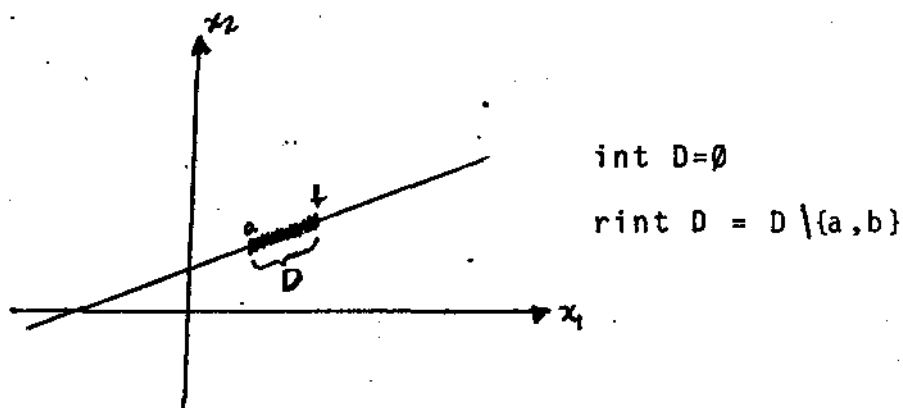


Figura 1.1.5.A: Ilustração de interior e interior relativo de um conjunto

#### 1.1.6. Alguns resultados de separação de conjuntos convexos

— A noção de separação de conjuntos é uma das mais férteis e úteis da análise convexa, como teremos oportunidade de constatar.

Definição: Sejam  $D_1$  e  $D_2$  dois conjuntos não vazios em  $\mathbb{R}^n$ . Um hiperplano  $H = \{x \mid \langle x, b \rangle = \beta\}$  separa  $D_1$  e  $D_2$  se  $D_1$  está contido num dos semiespaços gerados por  $H$  e  $D_2$  está contido no semiespaço fechado oposto, isto é, se

$$(i) \quad \text{ou } \langle x, b \rangle \geq \beta, \forall x \in D_1 \wedge \langle x, b \rangle \leq \beta, \forall x \in D_2;$$

$$(ii) \quad \text{ou } \langle x, b \rangle < \beta, \forall x \in D_1 \wedge \langle x, b \rangle \geq \beta, \forall x \in D_2.$$

O mesmo hiperplano  $H$  separa fortemente  $D_1$  e  $D_2$  se:

- (i) ou  $\langle x, b \rangle > \beta, \forall x \in D_1 \wedge \langle x, b \rangle \leq \beta, \forall x \in D_2;$
- (ii) ou  $\langle x, b \rangle \geq \beta, \forall x \in D_1 \wedge \langle x, b \rangle < \beta, \forall x \in D_2;$
- (iii) ou  $\langle x, b \rangle < \beta, \forall x \in D_1 \wedge \langle x, b \rangle \geq \beta, \forall x \in D_2;$
- (iv) ou  $\langle x, b \rangle \leq \beta, \forall x \in D_1 \wedge \langle x, b \rangle > \beta, \forall x \in D_2. ||$

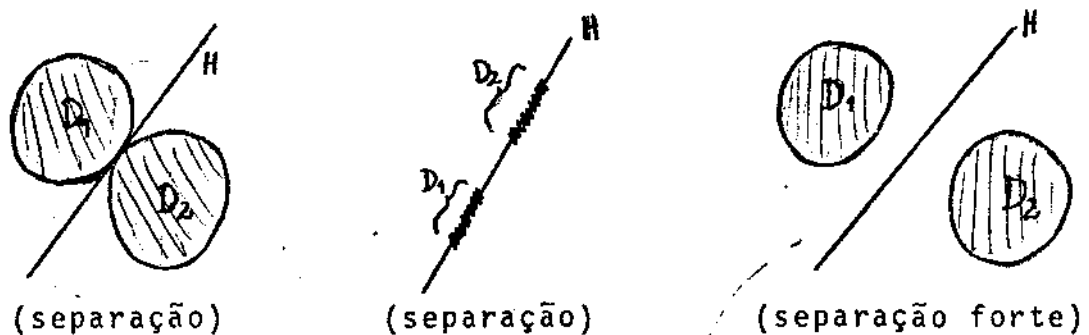


Figura 1.1.6.A; Ilustração de separação e separação forte.

Outros tipos de separação poderiam ser definidos, mas não o fazemos porque para os resultados que nos interessam eles não seriam relevantes.<sup>(14)</sup>

Proposição 1.1.6.I: Seja  $C$  um conjunto convexo fechado e seja  $y \notin C$ .

Então existe um único ponto  $\bar{x} \in C$  com distância mínima a  $y$ . Mais ainda, o ponto  $\bar{x}$  minimiza a distância entre  $C$  e  $y$  sse  $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0, \forall x \in C. ||$

Proposição 1.1.6.II: Seja um conjunto convexo não vazio e fechado  $C \subset \mathbb{R}^n$ , e seja  $y \notin C$ . Então existe um hiperplano que separa fortemente  $y$  e  $C$ , isto é, existirá um vector  $b \in \mathbb{R}^n$  não nulo e um  $\beta \in \mathbb{R}$  tal que:  
 $\langle y, b \rangle > \beta$  e  $\langle x, b \rangle \leq \beta, \forall x \in C. ||$

(14)-Para definição de outros tipos de separação (estrita e própria) ver, por exemplo, BAZARAA e SHETTY (1979, p.44)

Corolário 1.1.6.II.1.: Seja  $C$  um conjunto convexo fechado em  $\mathbb{R}^n$ . Então  $C$  é a intersecção de todos os semiespaços fechados que o contêm. ||

No caso de  $C$  não ser um conjunto fechado, a intersecção dos semiespaços fechados que o contêm, necessariamente um conjunto fechado (pela proposição 1.1.5.III), será  $C \cap C$ .

1.1.7. Conjuntos suporte, conjuntos polares e hiperplanos e semiespaços de suporte

——— Definição: Seja um conjunto convexo não vazio  $C \subset \mathbb{R}^n$ . O conjunto suporte de  $C$  é o conjunto de todos os semiespaços fechados que contêm  $C$ . ||

O corolário 1.1.6.II.1 estabelece que se o conjunto  $C$  é fechado, então ele pode ser obtido pela intersecção de todos os elementos do seu conjunto suporte.

——— Dois subconjuntos do conjunto suporte de um conjunto convexo assumem especial interesse na análise a efectuar na secção 2.3. São eles os conjuntos polares inferior e superior.

Definição: Os conjuntos polares inferior e superior de um conjunto convexo  $C \subset \mathbb{R}^n$  são respectivamente os conjuntos  $C_-^0 = \{x^* | \langle x, x^* \rangle \leq 1, \forall x \in C\}$  e  $C_+^0 = \{x^* | \langle x, x^* \rangle \geq 1, \forall x \in C\}$ . ||

Como vimos na alínea 1.1.3., cada hiperplano de  $\mathbb{R}^n$  que não passa pela origem pode ser expresso através de uma equação normalizada

$$\langle x, x^* \rangle = 1 \quad \text{com} \quad x^* \neq 0 \in \mathbb{R}^n.$$

Assim sendo, o conjunto  $C_-^0$  (resp.  $C_+^0$ ) pode ser encarado como o

conjunto de semiespaços inferiores (superiores) fechados que contêm  $C$  e que são gerados por hiperplanos que não passam pela origem.<sup>(15)</sup>

As propriedades de  $C_-^0$  e  $C_+^0$  serão estudadas na secção 2.3.. Nessa secção explicitar-se-á igualmente a relação entre as noções de cone polar e conjunto polar inferior.

— Do conjunto suporte de um conjunto convexo são também de destacar, pela sua importância, alguns elementos que têm a particularidade de os respectivos hiperplanos geradores serem tangentes ao conjunto convexo em questão. São os semiespaços (e hiperplanos) de suporte que a seguir se definem e que assumirão uma extrema relevância nos desenvolvimentos posteriores.

Definição: Seja um conjunto convexo não vazio  $C \subset \mathbb{R}^n$ , e seja  $\bar{x}$  um ponto da sua fronteira.. Um hiperplano  $H = \{x \mid \langle x - \bar{x}, b \rangle = 0\} = \{x \mid \langle x, b \rangle = \langle \bar{x}, b \rangle\}$ , com  $b \in \mathbb{R}^n$ , é designado por hiperplano de suporte (ou apoio) do conjunto  $C$  no ponto  $\bar{x}$  desde que:

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

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

O semiespaço fechado gerado pelo hiperplano de suporte e que contém  $C$  é designado por semiespaço de suporte (ou de apoio) a  $C$  no ponto  $\bar{x}$ . ||

Um semiespaço de suporte a um conjunto convexo  $C$  é então um semiespaço que contém  $C$  e que tem na sua fronteira pelo menos um ponto

---

(15) - Como curiosidade, note-se que quando  $0 \in C$ , então  $C_+^0 = \emptyset$ . Com efeito, para  $x=0$  não existirá nenhum vector  $x^*$  tal que  $\langle x, x^* \rangle \geq 1, \forall x$ , já que teremos necessariamente  $\langle 0, x^* \rangle = 0 < 1$

da fronteira de  $C$  (repare-se que para  $x=\bar{x}$  vem  $\langle x-\bar{x}, b \rangle = 0$ , pelo que o hiperplano de suporte passa pelo ponto fronteiro  $\bar{x}$ ).

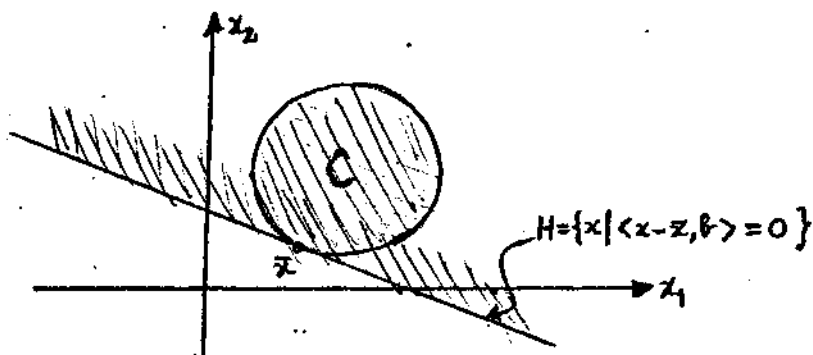


Figura 1.1.7.A: Ilustração de um hiperplano de suporte e respectivo semiespaço de suporte.

— A seguinte proposição deriva directamente do resultado de separação 1.1.6.II.

Proposição 1.1.7.I: 'Seja  $C \subset \mathbb{R}^n$  um conjunto convexo não vazio e se ja  $\bar{x}$  um ponto da sua fronteira. Então existe pelo menos um hiperplano de suporte a  $C$  no ponto  $\bar{x}$ . ||

Como veremos, as noções de hiperplano e semiespaço de suporte vão ser de crucial importância em todo o estudo de dualidade a desenvolver nos capítulos 2 e 3, bem como no estabelecimento de uma análise subdiferencial convexa (e quase-convexa) que vai permitir generalizar, para funções não diferenciáveis, a análise diferencial clássica.

## 1.2. FUNÇÕES CONVEXAS E CÔNCAVAS

### 1.2.1. Introdução

—— Por comodidade de tratamento, vamos permitir que as funções com que trabalharemos assumam não apenas valores reais, mas igualmente os "valores"  $+\infty$  e  $-\infty$ . Dito de outra forma, vamos trabalhar com funções que tomam valores em  $\mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$ . Para as operações envolvendo os símbolos  $\pm\infty$  adoptaremos as seguintes regras:

$$(i) \quad k + \infty = +\infty + k = +\infty \quad \text{para } k > -\infty;$$

$$(ii) \quad k - \infty = -\infty + k = -\infty \quad \text{para } k < +\infty;$$

$$(iii) \quad k \cdot (+\infty) = (+\infty) \cdot k = \begin{cases} +\infty & \text{para } k > 0 \\ -\infty & \text{para } k < 0; \end{cases}$$

$$(iv) \quad k \cdot (-\infty) = (-\infty) \cdot k = \begin{cases} -\infty & \text{para } k > 0 \\ +\infty & \text{para } k < 0; \end{cases}$$

$$(v) \quad 0 \cdot (+\infty) = 0 \cdot (-\infty) = (+\infty) \cdot 0 = (-\infty) \cdot 0 = 0 \quad (\text{com } 0 \text{ não infinito, mas constante});$$

$$(vi) \quad \inf \emptyset = +\infty$$

$$(vii) \quad \sup \emptyset = -\infty$$

Adicionalmente, onde necessárias, serão impostas condições que evitem a verificação de ambiguidades do tipo  $\infty - \infty$  ou  $-\infty + \infty$ .

### 1.2.2. Funções convexas: definição e algumas propriedades

—— Definição: Epigrafo de uma função  $f(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é o

subconjunto de  $\mathbb{R}^{n+1}$ :  $\text{epi} f(x) = \{(x, r) \mid x \in \mathbb{R}^n, r \in \mathbb{R}, r \geq f(x)\}$ . ||

Definição: Uma função  $f(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é convexa se o seu epi-grafo é um conjunto convexo. ||

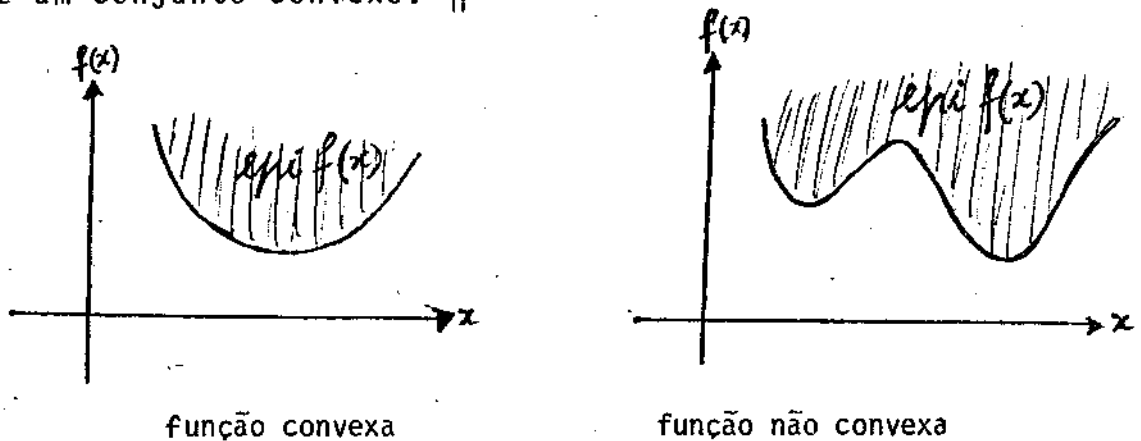


Figura 1.2.2.A: Ilustração de função convexa e não convexa

— A seguinte condição necessária e suficiente é apresentada por muitos autores como definição de função convexa.

Proposição 1.2.2.1: Seja  $f(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty]$ .  $f(x)$  é uma função convexa sse  $f[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f(x^1) + \lambda f(x^2)$ ,  $\forall x^1, x^2 \in \mathbb{R}^n, 0 \leq \lambda \leq 1$ . ||

Corolário 1.2.2.1.1. (Desigualdade de Jensen): Seja  $f(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty]$ .  $f(x)$  é uma função convexa sse para qualquer conjunto de escalares  $0 \leq \lambda_1, \dots, \lambda_m \leq 1$  tais que  $\lambda_1 + \dots + \lambda_m = 1$  temos:

$$f[\lambda_1 x^1 + \dots + \lambda_m x^m] \leq \lambda_1 f(x^1) + \dots + \lambda_m f(x^m), \forall x^1, \dots, x^m \in \mathbb{R}^n.$$

— Do conjunto de todas as funções convexas é conveniente destacar aquelas que correspondem a situações que podemos encarar como de maior interesse.

Definição: Uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diz-se própria se

o seu epigrafo é não vazio e não contém linhas verticais, isto é, se

$$\exists x \in \mathbb{R}^n: f(x) < +\infty \text{ e se } f(x) > -\infty, \forall x \in \mathbb{R}^n. \parallel$$

Proposição 1.2.2.II.: Sejam  $f_1(x)$  e  $f_2(x)$  funções convexas próprias de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então  $f_1(x) + f_2(x)$  é também uma função convexa própria.  $\parallel$

——— Outras propriedades importantes das funções convexas são estabelecidas nas próximas proposições.

Proposição 1.2.2.III.: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Os seus conjuntos de nível, ou secções inferiores,  $S_f^{\leq}(\alpha) = \{x \mid f(x) \leq \alpha\}$  e  $S_f^{<}(\alpha) = \{x \mid f(x) < \alpha\}$  são conjuntos convexos para qualquer  $\alpha \in [-\infty, +\infty]$ .  $\parallel$

Note-se que o inverso do anterior resultado não é sempre verdade, já que existem funções não convexas com todos os conjuntos de nível inferior convexos (funções quase-convexas, a estudar no capítulo 3).

Definição: Seja  $\{f_i(x) \mid i \in I\}$  um conjunto de funções convexas de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  (com um número finito ou infinito de elementos). A função supremo por pontos deste conjunto de funções convexas é a função

$$f(x) = \text{Sup} \{f_i(x) \mid i \in I\}. \parallel$$

Proposição 1.2.2.IV.: Seja  $\{f_i(x) \mid i \in I\}$  um conjunto de funções convexas de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então a respectiva função supremo por pontos é uma função convexa.  $\parallel$

Proposição 1.2.2.V.: Seja  $F \subset \mathbb{R}^{n+1}$  um conjunto convexo. A função  $f(x) = \inf\{r \mid (x, r) \in F\}$  é uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . ||  
Este último resultado intui-se facilmente se encararmos a fronteira inferior de  $F$  como a fronteira do epigrafo de uma função convexa  $f(x)$  não explicitada a priori.

—— Dever-se-á sublinhar que, devido à convenção que fizemos de trabalharmos com os símbolos  $\pm\infty$ , poderemos sempre fazer a extensão a  $\mathbb{R}^n$  de uma função convexa definida originalmente num subconjunto  $C \subset \mathbb{R}^n$ . Para tal bastará considerar  $f(x) = +\infty, \forall x \notin D$ . Ao subconjunto de  $\mathbb{R}^n$  onde a função  $f(x)$  estava originalmente definida dar-se-á a designação de domínio efectivo.

Definição: Domínio efectivo da função convexa  $f(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$   $\text{dom } f(x)$ , é a projecção sobre  $\mathbb{R}^n$  do seu epigrafo, ou seja,  $\text{dom } f(x) = \{x \mid \exists r < +\infty : (x, r) \in \text{epi } f(x)\} = \{x \mid f(x) < +\infty\}$ . ||

Proposição 1.2.2.VI.: O domínio efectivo de uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é um conjunto convexo. ||

### 1.2.3. Função indicadora de um conjunto convexo

—— Definição: Função indicadora de um conjunto  $D \subset \mathbb{R}^n$  é a função:

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

Proposição 1.2.3.I.: Seja  $D \subset \mathbb{R}^n$ . A função indicadora do conjunto  $D$  é convexa sse o conjunto  $D$  é convexo, sendo convexa própria-sse  $D$  é convexo e não vazio. ||

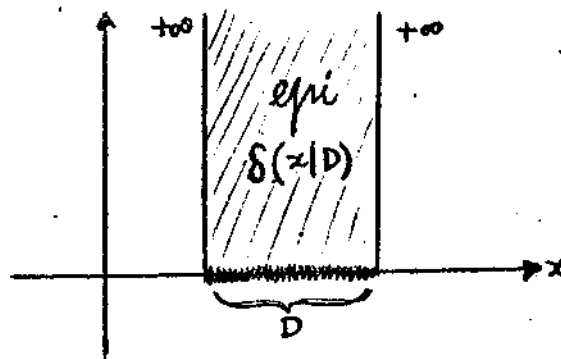


Figura 1.2.2.A: Ilustração da função indicadora de um conjunto convexo

O conceito de função indicadora vai ser especialmente útil ao permitir fazer, duma forma extremamente cômoda, a extensão a  $\mathbb{R}^n$  de qualquer função convexa definida num conjunto  $C \subset \mathbb{R}^n$ . Com efeito,

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

Sendo  $C$  não vazio e  $f$  convexa própria, as proposições 1.2.3.I e 1.2.2.II permitem concluir que  $f(x) + \delta(x|C)$  é também uma função convexa própria.

#### 1.2.4. Funções convexas linearmente homogêneas

—— Definição: Uma função  $f(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é homogênea de grau  $\alpha$  se, para cada  $x$ ,  $f(\lambda x) = \lambda^\alpha \cdot f(x)$ ,  $\forall \lambda > 0$ . No caso de  $\alpha = 1$  (função homogênea de grau 1) a função é também designada por linearmente homogênea. ||

Proposição 1.2.4.I.: O epígrafo de uma função linearmente homogênea de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é um cone em  $\mathbb{R}^{n+1}$ . ||

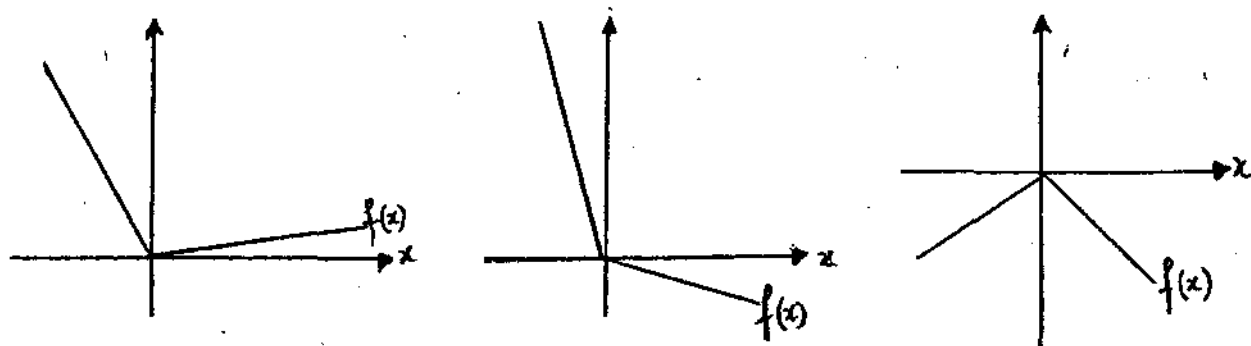


Figura 1.2.4.A.: Ilustração de funções linearmente homogêneas convexas e não convexas

A relação entre homogeneidade de grau 1 e convexidade é estabelecida no seguinte resultado:

Proposição 1.2.4.II.: Seja  $f(x)$  uma função linearmente homogênea de  $\mathbb{R}^n$  em  $]-\infty, +\infty]$ .  $f(x)$  é convexa sse  $f(x^1 + x^2) \leq f(x^1) + f(x^2)$ ,  $\forall x^1, x^2 \in \mathbb{R}^n$ . ||

Como teremos oportunidade de verificar, algumas das mais importantes relações de dualidade a desenvolver no capítulo 2 assentarão em funções convexas linearmente homogêneas. Um conceito que nesse contexto nos será necessário é o de função convexa linearmente homogênea gerada por uma dada função convexa  $f(x)$ . Porém, antes de o apresentarmos, teremos que introduzir duas operações especiais com funções convexas. Utilizando a terminologia de ROCKAFELLAR (1970, p. 33-35), designá-las-emos por multiplicações escalares esquerda e direita.

Definição: A operação de multiplicação escalar esquerda da função  $f(x)$  por  $\lambda \in \mathbb{R}$ , é a operação que faz corresponder à função  $f(x)$

a função:

$$(\lambda f)(x) = \lambda f(x) \quad . \quad ||$$

Proposição 1.2.4.III.: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e  $\lambda > 0$ . Então  $(\lambda f)(x)$  é também uma função convexa, isto é, a operação de multiplicação escalar esquerda preserva a convexidade para  $\lambda > 0$ . ||

Definição: A operação de multiplicação escalar direita da função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  por  $\lambda \in \mathbb{R}$ , é a operação que faz corresponder a  $f(x)$  a função  $(f\lambda)(x) = \lambda \cdot f(x/\lambda)$ . ||

Note-se que se  $f(x)$  for uma função linearmente homogênea  $(f\lambda)(x) = \lambda f(x/\lambda) = f(x)$ , ou seja, a operação de multiplicação escalar direita mantém inalteradas funções linearmente homogêneas.

Proposição 1.2.4.IV: Seja  $f(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $\lambda > 0$ . Então  $\text{epi } (f\lambda)(x) = \lambda (\text{epi } f(x))$ , ou seja, a multiplicação escalar direita de uma função corresponde à multiplicação do respectivo epígrafo desde que  $\lambda > 0$ . ||

—— Definição: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . A função convexa linearmente homogênea gerada por  $f(x)$  é a maior função convexa linearmente homogênea majorada por  $f(x)$ . ||

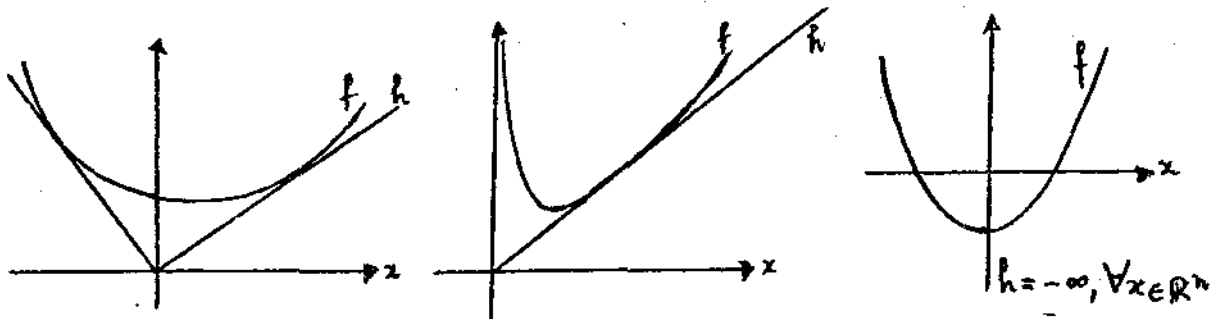


Figura 1.2.4.B.: Ilustração de funções convexas linearmente homogêneas  $h(x)$  geradas por funções convexas  $f(x)$ .

Atendendo à proposição 1.2.4.I, o epigrafo da função convexa linearmente gerada por uma função convexa  $f(x)$  será o cone gerado pelo epigrafo de  $f(x)$ .

Proposição 1.2.4.V. Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $h(x)$  é a função convexa linearmente homogênea gerada por  $f(x)$  sse  $h(x) = \inf\{(f\lambda)(x) \mid \lambda \geq 0\}$ . ||

Proposição 1.2.4.VI.: A função indicadora de um conjunto convexo é linearmente homogênea sse o conjunto é um cone convexo. ||

### 1.2.5. Funções côncavas

— Definição: Hipografo de uma função  $g(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é o subconjunto de  $\mathbb{R}^{n+1}$ :  $\text{hip } g(x) = \{(x, r) \mid x \in \mathbb{R}^n, r \in \mathbb{R}, r \leq g(x)\}$ . ||

Definição: Uma função  $g(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é côncava se o seu hipografo é um conjunto convexo. ||

A relação entre funções côncavas e convexas é estabelecida pelo seguinte resultado.

Proposição 1.2.5.I.: Sejam  $f(x)$  e  $g(x)$  funções de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é côncava sse  $-g(x)$  é convexa, e  $f(x)$  é convexa sse  $-f(x)$  é côncava. ||

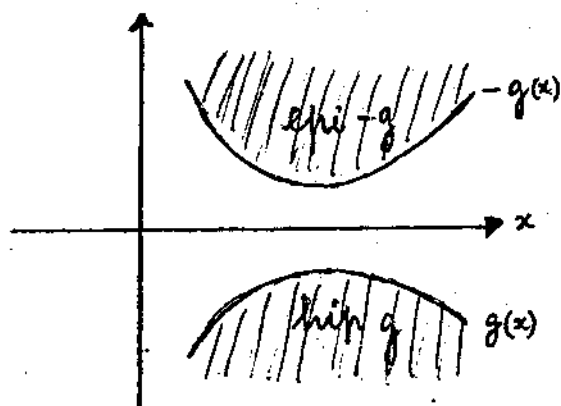


Figura 1.2.5.A: Ilustração da simetria entre hipografo de uma função côncava e epigrafo da sua simétrica que é convexa.

— As propriedades gerais das funções côncavas decorrem diretamente e imediatamente das propriedades das suas simétricas que são funções convexas.

Proposição 1.2.5.II: Seja  $g(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é côncava sse  $g[(1-\lambda)x^1 + \lambda x^2] \geq (1-\lambda)g(x^1) + \lambda g(x^2) \quad \forall x^1, x^2 \in \mathbb{R}^n, 0 \leq \lambda \leq 1$ . ||

Corolário 1.2.5.II.1. (desigualdade de Jensen): Seja  $g(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é uma função côncava sse para qualquer conjunto de escalares  $0 \leq \lambda_1, \dots, \lambda_m \leq 1$  tais que  $\lambda_1 + \dots + \lambda_m = 1$  temos:

$$g[\lambda_1 x^1 + \dots + \lambda_m x^m] \geq \lambda_1 g(x^1) + \dots + \lambda_m g(x^m), \quad \forall x^1, \dots, x^m \in \mathbb{R}^n. ||$$

Definição: Uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diz-se própria se o seu hipografo é não vazio e não contém linhas verticais, isto é, se  $\exists x \in \mathbb{R}^n: g(x) > -\infty$  e se  $g(x) < +\infty, \forall x \in \mathbb{R}^n$ . ||

Proposição 1.2.5.III: Seja  $g(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é côncava própria sse  $-g(x)$  é convexa própria. ||

Proposição 1.2.5.IV: Sejam  $g_1(x)$  e  $g_2(x)$  funções côncavas próprias de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então  $g_1(x) + g_2(x)$  é também uma função côncava própria.

Proposição 1.2.5.V: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Os seus conjuntos de nível superior, ou secções superiores,  $S_g^{\geq}(\alpha) = \{x | g(x) \geq \alpha\}$  e  $S_g^{>}(\alpha) = \{x | g(x) > \alpha\}$  são conjuntos convexas para cada  $\alpha \in [-\infty, +\infty]$ . ||

Definição: Seja  $\{g_i(x) | i \in I\}$  um conjunto de funções côncavas de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  (com um número de elementos finito ou infinito). A função ínfimo por pontos deste conjunto de funções côncavas é a

função

$$g(x) = \inf \{g_i(x) \mid i \in I\}.$$

Proposição 1.2.5.VI: Seja  $\{g_i(x) \mid i \in I\}$  um conjunto de funções côncavas de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então a respectiva função ínfimo por pontos é uma função côncava. ||

Proposição 1.2.5.VII: Seja  $G \subset \mathbb{R}^{n+1}$  um conjunto convexo. A função  $g(x) = \sup \{r \mid (x, r) \in G\}$  é uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . ||

Definição: Domínio efectivo de uma função côncava  $g(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ ,  $\text{dom } g(x)$ , é a projecção sobre  $\mathbb{R}^n$  do seu hipografo, ou seja  $\text{dom } g(x) = \{x \mid \exists r > -\infty : (x, r) \in \text{hip } g(x)\} = \{x \mid g(x) > -\infty\}$ . ||

Proposição 1.2.5.VIII: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então  $\text{dom } g(x) = \text{dom } (-g(x))$ . ||

Proposição 1.2.5.IX: O domínio efectivo de uma função côncava é um conjunto convexo. ||

— As anteriores propriedades juntar-se-ão ainda as seguintes, válidas para funções côncavas que sejam também linearmente homogêneas. Tal como para as acima enunciadas, elas decorrem directamente da simetria entre funções convexas e côncavas.

Proposição 1.2.5.X: O hipografo de uma função linearmente homogênea de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é um cone em  $\mathbb{R}^{n+1}$ . ||

Proposição 1.2.5.XI: Seja  $g(x)$  uma função linearmente homogênea de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é côncava sse  $g(x^1 + x^2) \geq g(x^1) + g(x^2), \forall x^1, x^2 \in \mathbb{R}^n$ . ||

Proposição 1.2.5.XII: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e  $\lambda > 0$ . Então  $(\lambda g)(x)$  é também uma função côncava, isto é, a operação

de multiplicação escalar esquerda preserva a concavidade para  $\lambda > 0$ . ||

Proposição 1.2.5.XIII: Seja  $g(x)$  uma função de  $R^n$  em  $[-\infty, +\infty]$  e seja  $\lambda > 0$ . Então a hip  $(g\lambda)(x) = \lambda(\text{hip } g(x))$ , ou seja, a multiplicação escalar direita de uma função corresponde à multiplicação do respectivo hipografo desde que  $\lambda > 0$ . ||

Definição: Seja  $g(x)$  uma função côncava de  $R^n$  em  $[-\infty, +\infty]$ . A função côncava linearmente homogênea gerada por  $g(x)$  é a menor função côncava linearmente homogênea minorada por  $g(x)$ . ||

Proposição 1.2.5. XIV: Seja  $g(x)$  uma função côncava de  $R^n$  em  $[-\infty, +\infty]$ .  $h(x)$  é a função côncava linearmente gerada pela função  $g(x)$  sse  $h(x) = \sup\{(g\lambda)(x) \mid \lambda > 0\}$  ||

#### 1.2.6. Funções lineares não verticais.

— Definição: Funções lineares não verticais, ou funções afins não verticais, são funções  $\ell(x) = \langle x, a \rangle - \beta$ , com  $a \in R^n$  e  $\beta \in R$ . ||

As funções lineares não verticais gozam das seguintes propriedades fundamentais.

Proposição 1.2.6.1: Os hiperplanos de  $R^{n+1}$  definidos por  $\{(x, r) \mid \langle (x, r), (b_1, b_0) \rangle = \alpha\}$  com  $x \in R^n$ ,  $r \in R$ ,  $b \in R^n$ ,  $b_0 \neq 0 \in R$  e  $\alpha \in R$ , correspondem a funções lineares não verticais de  $R^n$  em  $[-\infty, +\infty]$ :  $r = \langle x, \frac{-b}{b_0} \rangle + \frac{\alpha}{b_0}$ . ||

Atendendo à anterior proposição, definamos hiperplanos não verticais em  $R^{n+1}$  como os hiperplanos que correspondem a alguma função linear não vertical de  $R^n$ .

Definição: Um hiperplano de  $R^{n+1}$   $\{(x, r) \mid \langle (x, r), (b, b_0) \rangle = \alpha\}$ , com  $x \in R^n$ ,  $r \in R$ ,  $b \in R^n$ ,  $b_0 \neq 0 \in R$  e  $\alpha \in R$ , diz-se hiperplano não vertical em  $R^{n+1}$ . ||

Proposição 1.2.6.II: As funções lineares não verticais de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  são simultaneamente convexas e côncavas. Se elas não forem as funções constantes  $+\infty$  ou  $-\infty$ , então os seus grafos são hiperplanos não verticais em  $\mathbb{R}^{n+1}$ .||

### 1.2.7. Funções semicontínuas/funções convexas e côncavas fechadas

——— Definição: Uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diz-se semicontínua inferiormente (s.c.i.) num ponto  $x$  e relativamente a um de elementos  $D$  conjunto  $D \subset \mathbb{R}^n$  se  $z(x) \leq \lim_{x^i \rightarrow x} z(x^i)$  qualquer que seja a sucessão  $\{x^i\}$  convergente para  $x$ , existindo, finito ou infinito,  $\lim_{x^i \rightarrow x} z(x^i)$ . Se a função  $z(x)$  for s.c.i. em cada  $x \in D$  relativamente ao mesmo conjunto  $D$ , diz-se que  $z(x)$  é s.c.i. relativamente a  $D$ . No caso particular de  $D = \mathbb{R}^n$  omite-se normalmente a expressão "relativamente a  $\mathbb{R}^n$ ".||

Repare-se que na anterior definição somente se exige que  $z(x) \leq \lim_{x^i \rightarrow x} z(x^i)$  quando a sucessão  $\{x^i\}$  seja formada exclusivamente por elementos de  $D$ , nada se obrigando quando tal não acontece.

——— Proposição 1.2.7.I: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então as seguintes condições são equivalentes:

- (i)  $z(x)$  é uma função s.c.i.;
- (ii)  $S_z^{\leq}(\alpha) = \{x \mid z(x) \leq \alpha\}$  é um conjunto fechado para cada  $\alpha \in [-\infty, +\infty]$ ;
- (iii) O epigrafo de  $z(x)$  é um conjunto fechado.||

Corolário 1.2.7.I.1: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  tal que os respectivos conjuntos de nível inferior  $S_z^{\leq}(\alpha) = \{x \mid z(x) \leq \alpha\}$  são fechados para cada  $\alpha \leq \bar{\alpha}$ . Então a função  $z(x)$  é s.c.i. relativamente

a  $S_z^{\leq}(\bar{\alpha})$ . ||

— Entre as propriedades das funções s.c.i. destacamos:

Proposição 1.2.7.II: Sejam  $z_1(x)$  e  $z_2(x)$  funções de  $R^n$  em  $[-\infty, +\infty]$ , ambas s.c.i. relativamente a um mesmo conjunto  $D \subset R^n$ . Então  $z_1(x) + z_2(x)$  é também s.c.i. relativamente a  $D$ . ||

Proposição 1.2.7.III. Seja  $\mu$  um escalar real positivo e seja  $z(x)$  uma função de  $R^n$  em  $[-\infty, +\infty]$  s.c.i. relativamente a  $D \subset R^n$ . Então  $\mu \cdot z(x)$  é também uma função s.c.i. relativamente a  $D$ . ||

Proposição 1.2.7.IV: Seja  $\{z_i(x) | i \in I\}$  um conjunto de funções s.c.i. de  $R^n$  em  $[-\infty, +\infty]$ . Então a função supremo por pontos  $z(x) = \sup \{z_i(x) | i \in I\}$  é s.c.i.. ||

— Definição: Envolvente semicontínua inferior de uma função  $z(x)$  de  $R^n$  em  $[-\infty, +\infty]$  é a maior função s.c.i. (não necessariamente finita) majorada por  $z(x)$ . ||

Atendendo à proposição 1.2.7.I(iii), a envolvente s.c.i. de uma função  $z(x)$  é a função cujo epigrafo é o fecho do epigrafo de  $z(x)$ .

Definição: Por fecho de uma função convexa  $f(x)$  de  $R^n$  em  $[-\infty, +\infty]$ , cl  $f(x)$ , entende-se: (i) ou a envolvente s.c.i. de  $f(x)$  se  $f(x) > -\infty$ ,  $\forall x \in R^n$ ; (ii) <sup>ou</sup> a função constante  $-\infty$  se  $f(x) = -\infty$  para algum  $x \in R^n$ . ||

A situação (ii) da anterior definição destina-se a dar maior generalidade à proposição 2.1.2.I do capítulo 2, permitindo que esta abarque funções convexas impróprias.

Definição: Uma função convexa  $f(x)$  de  $R^n$  em  $[-\infty, +\infty]$  é fechada se coincide com o seu fecho:  $f(x) = \text{cl } f(x)$ . ||

Note-se, nomeadamente, que:

- (i) As funções constantes  $+\infty$  e  $-\infty$  são funções fechadas;
- (ii) Uma função convexa própria é fechada se e somente se coincide com a sua envolvente s.c.i., isto é, se é uma função s.c.i..

—— Tal como definimos semicontinuidade inferior, podemos agora definir simetricamente semicontinuidade superior.

Definição: Uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diz-se semicontínua superiormente (s.c.s.) num ponto  $x$  e relativamente a um conjunto  $D \subset \mathbb{R}^n$  se  $z(x) \geq \lim_{\substack{x^i \rightarrow x \\ x^i \in D}} z(x^i)$  qualquer que seja a sucessão  $\{x^i\}$  <sup>de elementos de</sup>  $D$  convergente para  $x$ , existindo, finito ou infinito,  $\lim_{x^i \rightarrow x} z(x^i)$ . Se a função  $z(x)$  for s.c.s. em cada  $x \in D$  relativamente ao mesmo conjunto  $D$ , diz-se que  $z(x)$  é s.c.s. relativamente a  $D$ . No caso particular de  $D = \mathbb{R}^n$  omite-se normalmente a expressão "relativamente a  $\mathbb{R}^n$ ". ||

Proposição 1.2.7.V: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $z(x)$  é s.c.s. no ponto  $x$  e relativamente a  $D \subset \mathbb{R}^n$  sse  $-z(x)$  é s.c.i. no mesmo ponto e relativamente ao mesmo conjunto. ||

Devido à simetria acabada de apontar entre semicontinuidade superior e inferior, podemos estabelecer muito facilmente as propriedades para funções s.c.s. correspondentes às estabelecidas para funções s.c.i. nas proposições 1.2.7.I a 1.2.7.IV..

Proposição 1.2.7.VI: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então as seguintes condições são equivalentes:

- (i)  $z(x)$  é uma função s.c.s.;
- (ii)  $S_z^{\geq}(\alpha) = \{x \mid z(x) \geq \alpha\}$  é um conjunto fechado para cada  $\alpha \in [-\infty, +\infty]$ ;
- (iii) O hipografo de  $z(x)$  é um conjunto fechado. ||

Corolário 1.2.7.VI.1: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  tal que os respectivos conjuntos de nível superior  $S_z^{\geq}(\alpha) = \{x \mid z(x) \geq \alpha\}$  são fechados para cada  $\alpha \in \mathbb{R}$ . Então a função  $z(x)$  é s.c.s. em todos os elementos de  $S_z^{\geq}(\bar{\alpha})$ . ||

Proposição 1.2.7.VII : Sejam  $z_1(x)$  e  $z_2(x)$  funções de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  ambas s.c.s. relativamente a um mesmo conjunto  $D \subset \mathbb{R}^n$ . Então  $z_1(x) + z_2(x)$  é também s.c.s. relativamente a  $D$ . ||

Proposição 1.2.7.VIII: Seja  $\mu$  um escalar real positivo e seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  s.c.s. relativamente a  $D \subset \mathbb{R}^n$ . Então  $\mu \cdot z(x)$  é também uma função s.c.s. relativamente a  $D$ . ||

Proposição 1.2.7.IX: Seja  $\{z_i(x) \mid i \in I\}$  um conjunto de funções s.c.s. de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então a função ínfimo por pontos  $z(x) = \inf \{z_i(x) \mid i \in I\}$  é s.c.s. ||

— Definição: Envolvente semicontínua superior de uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é a menor função s.c.s. (não necessariamente finita) minorada por  $z(x)$ . ||

Atendendo à proposição 1.2.7.VI(iii), a envolvente s.c.s. de uma função  $z(x)$  é a função cujo hipografo é o fecho do hipografo de  $z(x)$

Definição: Por fecho de uma função côncava  $g(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ ,  $clg(x)$ , entende-se: (i) ou a envolvente s.c.s. de  $g(x)$  se  $g(x) < +\infty$ ,  $\forall x \in \mathbb{R}^n$ ; (ii) ou a função constante  $+\infty$  se  $g(x) = +\infty$  para algum  $x \in \mathbb{R}^n$ . ||

Definição: Uma função côncava  $g(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  é fechada se coincide com o seu fecho:  $g(x) = clg(x)$ . ||

Atendendo à simetria entre por um lado funções convexas e côncavas (proposição 1.2.5.1), e por outro lado funções s.c.i. e s.c.s.,

(proposição 1.2.7.V), é fácil de concluir que uma função é côncava fechada sse a sua simétrica é convexa fechada.

Proposição 1.2.7.X: Seja uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $z(x)$  é côncava fechada sse  $-z(x)$  é convexa fechada. ||

### 1.2.8. Funções contínuas/propriedades das funções convexas e côncavas contínuas

— A síntese dos conceitos de semicontinuidade inferior e superior origina o conceito clássico de continuidade de uma função.

Definição: Uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diz-se contínua num ponto  $x$  relativamente a um conjunto  $D \subset \mathbb{R}^n$  se é simultaneamente s.c.i. e s.c.s. em  $x$  relativamente a  $D$ .  $z(x)$  é contínua relativamente a  $D$  se é contínua em cada  $x \in D$  relativamente a  $D$ . No caso particular de  $D = \mathbb{R}^n$  omitir-se-á a expressão "relativamente a  $\mathbb{R}^n$ ". ||

Proposição 1.2.8.I: Seja uma função  $z(x)$  de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $z(x)$  é contínua no ponto  $x$  relativamente a  $D \subset \mathbb{R}^n$  sse  $z(x) = \lim_{x^i \rightarrow x} z(x^i)$  para toda a sucessão  $\{x^i\}$  de elementos de  $D$  que convirja para  $x$ , existindo, finito ou infinito,  $\lim_{x^i \rightarrow x} z(x^i)$ . ||

Corolário 1.2.8.I,1: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então  $z(x)$  é contínua num ponto  $x$  relativamente a  $D \subset \mathbb{R}^n$  sse  $-z(x)$  é contínua em  $x$  relativamente a  $D$ . ||

Corolário 1.2.8.I,2: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$ . Então  $z(x)$  é contínua num ponto  $x$  relativamente a  $D \subset \mathbb{R}^n$  sse  $\forall \epsilon > 0 \exists \delta > 0$ :  
 $(\|x^i - x\| \leq \delta \wedge x^i \in D) \Rightarrow |z(x^i) - z(x)| \leq \epsilon$ . ||

——— É possível garantir que uma função contínua tem máximo e mínimo finitos num determinado conjunto desde que este satisfaça certas condições de regularidade.

Definição: Seja um conjunto  $D \subset \mathbb{R}^n$ .  $D$  é limitado superiormente se existe algum vector  $x^m \in \mathbb{R}^n$  tal que  $x \leq x^m, \forall x \in D$ .  $D$  é limitado inferiormente se existe algum vector  $x^m \in \mathbb{R}^n$  tal que  $x \geq x^m, \forall x \in D$ .  $D$  é limitado se é simultaneamente limitado superior e inferiormente. ||

Proposição 1.2.8.II (Teorema de WEIERSTRASS): Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$  contínua relativamente a um conjunto  $D \subset \mathbb{R}^n$  não vazio, fechado e limitado. Então existe pelo menos um vector  $x^1 \in D$  para o qual a função  $z$  atinge o seu valor máximo (finito) em  $D$ , assim como existe também um vector  $x^2 \in D$  para o qual  $z$  atinge o seu valor mínimo (finito) em  $D$ . ||

——— Os próximos teorema e corolário mostram que uma função convexa ou concava é necessariamente contínua relativamente ao interior relativo do seu domínio efectivo. Pontos de descontinuidade, a existirem, estarão situados na fronteira relativa deste.

Proposição 1.2.8.III: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $f(x)$  é contínua relativamente a qualquer conjunto convexo  $C \subset \mathbb{R}^n$  relativamente aberto e contido em  $\text{dom } f$ . Em particular,  $f$  é contínua relativamente a  $\text{rint } \text{dom } f$ . ||

Corolário 1.2.8.III.1: Seja  $g(x)$  uma função concava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ .  $g(x)$  é contínua relativamente a qualquer conjunto convexo  $C \subset \mathbb{R}^n$  relativamente aberto e contido em  $\text{dom } g$ . Em particular  $g$  é contínua re

lativamente a  $\text{rint dom } g$ .||

Corolário 1.2.8.III.2: Seja  $g(x)$  uma função linear não vertical de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então  $g(x)$  é uma função contínua.||

— Antes de apresentarmos um outro resultado importante de continuidade de funções convexas e concavas definamos dois conceitos de que necessitaremos nesse resultado.

Definição: Seja um conjunto de  $m$  vectores linearmente independentes de  $\mathbb{R}^n$ :  $\{x^1, \dots, x^m\}$ . A envolvente convexa deste conjunto, isto é, o conjunto de todas as combinações lineares convexas de  $x^1, \dots, x^m$ , é designado por simplex  $m$ -dimensional de vértices  $x^1, \dots, x^m$ .||

Um simplex unidimensional, bidimensional e tridimensional é, respectivamente, um ponto, um segmento de recta e um triângulo em  $\mathbb{R}^n$ .

Definição: Um conjunto  $L \subset \mathbb{R}^n$  é um conjunto localmente simplicial ("locally simplicial") se para cada  $x \in L$  existe um número finito de simplex's  $S_1, \dots, S_r \subset L$  tais que para algum  $\epsilon > 0$  temos:

$$E_\epsilon(x) \cap (S_1 \cup \dots \cup S_r) = E_\epsilon(x) \cap L. ||$$

Proposição 1.2.8.IV: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $L$  um qualquer subconjunto localmente simplicial do domínio efectivo de  $f$ . Então  $f(x)$  é uma função semicontínua superiormente relativamente a  $L$ , pelo que  $f$  é contínua relativamente a  $L$  se fôr fechada.||

Corolário 1.2.8.IV.1: Seja  $g(x)$  uma função côncava e seja  $L$  um qualquer subconjunto localmente simplicial do domínio efectivo de  $g$ . Então  $g(x)$  é uma função semicontínua inferiormente relativamente

a  $L$ , pelo que  $g$  é contínua relativamente a  $L$  se for fechada.||

A importância que vai assumir na nossa análise a anterior proposição e respectivo corolário é a consequência directa de, entre outros conjuntos,<sup>(16)</sup> o ortante não negativo de  $\mathbb{R}^n$  ser um conjunto localmente simplicial.

Proposição 1.2.8.V: O conjunto  $\mathbb{R}_{0+}^n = \{x | x \in \mathbb{R}^n, x \geq 0\}$  é um conjunto localmente simplicial.||

Como algumas das funções com que depararemos nos capítulos 4,5 e 6 são fechadas e têm domínio <sup>efectivo</sup>  $\mathbb{R}_{0+}^n$ , a proposição 1.2.8.IV (ou corolário 1.2.8.IV.1.) e a proposição 1.2.8.V permitirão concluir que são funções contínuas relativamente a  $\mathbb{R}_{0+}^n$ .

### 1.2.9. Funções diferenciáveis/propriedades das funções convexas e côncavas diferenciáveis

—— Definição: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja um ponto  $x$  onde  $z$  é finita. A função  $z$  é diferenciável no ponto  $x$  se existe um único vector  $\nabla z(x)$  tal que

$$z(y) = z(x) + \langle \nabla z(x), y-x \rangle + o(\|y-x\|)$$

em que  $o(\cdot)$  é um infinitésimo de ordem superior a  $\|y-x\|$ . Equivalente mente,  $z$  é diferenciável no ponto  $x$  se

$$\lim_{\|y-x\| \rightarrow 0^+} \frac{z(y) - z(x) - \langle \nabla z(x), y-x \rangle}{\|y-x\|} = 0$$

(16)-ROCKAFELLAR (1970,p.184) demonstra que todos os conjuntos poliédricos convexos (isto é, conjuntos resultantes da intersecção dum número finito semiespaços) são conjuntos localmente simpliciais.

O vector  $\nabla z(x)$ , se existir, é designado por um vector gradiente de  $z$  no ponto  $x$ . ||

Definição: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $x$  um ponto onde  $z$  é finita. Designa-se por derivada direccional da função  $z$ , no ponto  $x$  e na direcção e sentido dados pelo vector  $d \in \mathbb{R}^n$ , o seguinte limite, se existir finito ou infinito:

$$z'(x, d) = \lim_{\lambda \rightarrow 0^+} \frac{z(x + \lambda d) - z(x)}{\lambda} . ||$$

Proposição 1.2.9.1: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $x$  um ponto onde  $z$  é diferenciável. Então  $z'(x, d)$ , qualquer que seja  $d \in \mathbb{R}^n$ , é dada por:  $z'(x, d) = \langle \nabla z(x), d \rangle$  . ||

Corolário 1.2.9.1.1: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $x$  um ponto onde  $z$  é diferenciável. O vector gradiente de  $z$  no ponto  $x$  tem como componentes os seguintes limites

$$\lim_{\lambda \rightarrow 0^+} \frac{z(x + \lambda e_j) - z(x)}{\lambda} \quad (j=1, \dots, n),$$

em que  $e_j$  é o vector dado pela  $j$ -ésima coluna da matriz identidade de ordem  $n$ . ||

Definição: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  diferenciável num ponto  $x$ . Por derivada parcial de  $z$  no ponto  $x$  e na variável  $x_j$  entende-se o limite

$$\frac{\partial z}{\partial x_j}(x) = \lim_{\lambda \rightarrow 0^+} \frac{z(x + \lambda e_j) - z(x)}{\lambda} . ||$$

O corolário 1.2.9.1.1, estabelece então que o vector gradiente duma função  $z$  num ponto  $x$ , se existe, é constituído pelas derivadas parciais de  $z$  nesse ponto  $x$ .

— Proposição 1.2.9.II: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $]-\infty, +\infty]$  diferenciável no ponto  $x$ . Então:

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

Corolário 1.2.9.II.1. Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty[$  diferenciável no ponto  $x$ . Então

$$g(y) \leq g(x) + \langle \nabla g(x), y-x \rangle, \quad \forall y \in \mathbb{R}^n. ||$$

Os próximos teorema e corolário estabelecem uma condição necessária e suficiente de convexidade ou concavidade para funções diferenciáveis num certo conjunto convexo  $C \subset \mathbb{R}^n$ .

Proposição 1.2.9.III: Seja  $f(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$  diferenciável num conjunto convexo  $C \subset \mathbb{R}^n$ . A função  $f(x)$  é convexa em  $C$  sse

$$f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle, \quad \forall x, y \in C. ||$$

Corolário 1.2.9.III.1. Seja  $g(x)$  uma função de  $\mathbb{R}^n$  em  $[-\infty, +\infty[$  diferenciável num conjunto convexo  $C \subset \mathbb{R}^n$ . A função  $g(x)$  é côncava em  $C$  sse

$$g(y) \leq g(x) + \langle \nabla g(x), y-x \rangle, \quad \forall x, y \in C. ||$$

Também como corolário da proposição 1.2.9.III podemos apresentar uma útil condição necessária e suficiente de ótimo para funções convexas e côncavas diferenciáveis.

Corolário 1.2.9.III.2: Seja  $z(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$  diferenciável e convexa (resp. côncava) num conjunto convexo aberto  $C \subset \mathbb{R}^n$ .  $\bar{x} \in C$  é mínimo (máximo) global finito de  $z$  em  $C$  sse  $\nabla z(\bar{x}) = 0. ||$

Definição: Seja  $z(x)$  uma função diferenciável num conjunto aberto  $D \subset \mathbb{R}^n$ , tal que todas as funções derivadas parciais de  $z$  são por sua vez diferenciáveis em  $D$ . A função  $z(x)$  diz-se neste caso bidiferenciável em  $D$  e a matriz das derivadas parciais de 2ª ordem num ponto  $x \in D$

$$H(x) = \left[ \frac{\delta^2 z}{\delta x_i \delta x_j}(x) \right] \quad (i, j=1, \dots, n)$$

é designada por matriz hesseana de  $z$  nesse ponto. ||

A proposição 1.2.9.III (resp. corolário 1.2.9.III.1) estabelece, como vimos, uma condição necessária e suficiente de convexidade (concavidade) para funções diferenciáveis. Os seguintes dois resultados são mais fortes nas suas hipóteses e estabelecem condições necessárias e suficientes respectivamente de convexidade e concavidade para uma função bidiferenciável.

Proposição 1.2.9.IV: Seja  $f(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$  bidiferenciável num conjunto convexo aberto  $C \subset \mathbb{R}^n$ . A função  $f(x)$  é convexa em  $C$  sse a sua matriz hesseana for semidefinida positiva para todo o  $x \in C$ .<sup>(17)</sup> ||

Corolário 1.2.9.IV.1. Seja  $g(x)$  uma função de  $\mathbb{R}^n$  em  $]-\infty, +\infty[$  diferenciável num conjunto convexo aberto  $C \subset \mathbb{R}^n$ . A função  $g(x)$  é côncava em  $C$  sse a sua matriz hesseana for semidefinida negativa para todo o  $x \in C$ . ||

#### 1.2.10. Subdiferenciabilidade de funções convexas e côncavas

No caso de funções convexas é possível construir uma análise que permite estudar as suas propriedades marginais sem exigir

(17)-Uma matriz quadrada  $Q$  de ordem  $n$  é semidefinida positiva se  $\langle v, Qv \rangle \geq 0, \forall v \in \mathbb{R}^n$ .

diferenciabilidade. Tal análise, assente nos conceitos de subgradiente e subdiferencial, tem a grande vantagem de se reduzir à análise diferencial clássica quando a função em questão é diferenciável.

Definição: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Subgradiente de  $f$  num ponto  $x \in \mathbb{R}^n$  é qualquer vector  $s \in \mathbb{R}^n$  tal que:

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

Definição: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Subgradiente de  $g$  num ponto  $x \in \mathbb{R}^n$  é qualquer vector  $s \in \mathbb{R}^n$  tal que:

$$g(y) \leq g(x) + \langle s, y-x \rangle, \quad \forall y \in \mathbb{R}^n. \quad ||$$

Proposição 1.2.10.I: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . O vector  $s$  é um subgradiente de  $g(x)$  no ponto  $x \in \mathbb{R}^n$  sse  $-s$  é um subgradiente de  $-g(x)$  no mesmo ponto  $x$ . ||

Proposição 1.2.10.II: Seja  $f(x)$  uma função convexa (resp.  $g(x)$  uma função côncava) de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . O vector  $s \in \mathbb{R}^n$  é um subgradiente de  $f$  (de  $g$ ) sse o hiperplano não vertical em  $\mathbb{R}^{n+1}$ ,  $\{(y, r) \mid r = f(x) + \langle s, y-x \rangle\}$  ( $\{(y, r) \mid r = g(x) + \langle s, y-x \rangle\}$ ) define um hiperplano de suporte a  $\text{epi } f$  (a  $\text{hip } g$ ) no ponto  $x$ . ||

Corolário 1.2.10.II.1: Seja  $f(x)$  uma função convexa (resp.  $g(x)$  uma função côncava) de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ . Então para cada  $x \in \text{int dom } f$  ( $\text{int dom } g$ ) existe pelo menos um subgradiente da função nesse ponto. ||

Proposição 1.2.10.III: Seja  $f(x)$  uma função convexa de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$

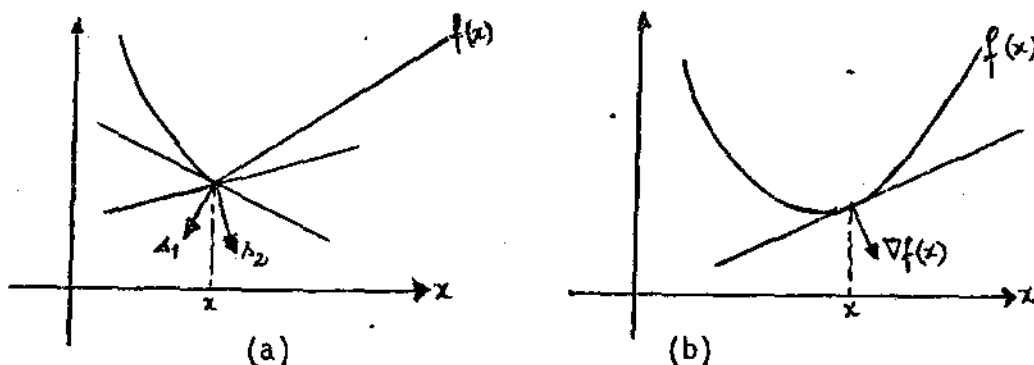
e seja  $x$  um ponto onde  $f$  é finita. Então  $s \in \mathbb{R}^n$  é um subgradiente de  $f$  em  $x$  sse  $f'(x, d) \geq \langle s, d \rangle$ ,  $\forall d \in \mathbb{R}^n$  ||

Corolário 1.2.10.III.1: Seja  $g(x)$  uma função côncava de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$  e seja  $x$  um ponto onde é finita. Então  $s \in \mathbb{R}^n$  é um subgradiente de  $g$  em  $x$  sse  $g'(x, d) \leq \langle s, d \rangle$ ,  $\forall d \in \mathbb{R}^n$ . ||

— Definição: O conjunto de todos os subgradientes da função convexa  $f(x)$  (resp. côncava de  $g(x)$ ) no ponto  $x$  designa-se por subdiferencial de  $f$  (de  $g$ ) no ponto  $x$  e representar-se-á por  $\partial f(x)$  (por  $\partial g(x)$ ). ||

A proposição que se segue revela uma relação fundamental entre subdiferenciabilidade e diferenciabilidade, mostrando que aquela se reduz a esta quando a função é diferenciável.

Proposição 1.2.10.IV: Seja  $f(x)$  uma função convexa (resp.  $g(x)$  uma função côncava) de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ , e diferenciável num ponto  $x$  onde é finita. Então  $\nabla f(x)$  ( $\nabla g(x)$ ) é o único subgradiente da função nesse ponto  $x$  <sup>(20)</sup> ||



(a) função convexa não diferenciável no ponto  $x$ : vários subgradientes.

(b) função " diferenciável no ponto  $x$ : único subgradiente  $= \nabla f(x)$ .

Figura 1.2.10.A: Ilustração da proposição 1.2.10.IV.

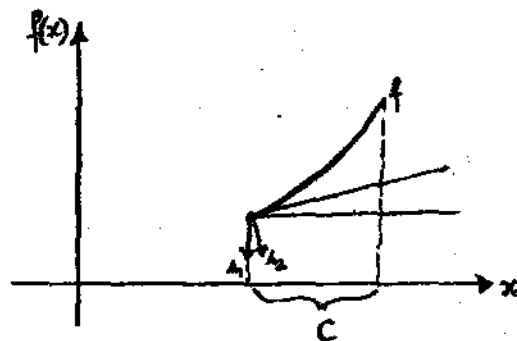
(20)- Poder-se-á demonstrar também o inverso, isto é, que se uma função convexa ou côncava tem um único subgradiente  $s$  num dado ponto  $x$ , então ela é diferenciável com vector gradiente igual a  $s$ . Não demonstraremos este resultado porque ele não será necessário (veja-se ROCKAFELLAR (1970, p. 242).

Proposição 1.2.10.V: Seja  $f(x)$  uma função convexa (resp.  $g(x)$  uma função côncava) de  $\mathbb{R}^n$  em  $[-\infty, +\infty]$ , e seja  $C \subset \text{dom } f$  ( $C \subset \text{dom } g$ ) um conjunto convexo. O vector  $\bar{x} \in C$  é mínimo (máximo) global de  $f$  (de  $g$ ) em  $C$  sse  $0 \in \partial f(\bar{x})$  ( $0 \in \partial g(\bar{x})$ ). ||

Note-se, para terminar este 1º capítulo, que a proposição 1.2.10.V engloba como caso particular o corolário 1.2.9.III.2. Com efeito, se a função em questão é diferenciável temos, pela proposição 1.2.10.IV, que o vector gradiente  $\bar{e}$ , em cada ponto, o único subgradiente da função<sup>(21)</sup>.

---

(21) - Note-se que no corolário 1.2.9.III se exige que  $C$  seja aberto, enquanto que na proposição 1.2.10.V isso não é exigido. Tal deriva do facto de para  $x \in \text{int } C$  o subdiferencial não ter um único elemento, não sendo por isso diferenciável a função,



SECÇÃO 1.1.

——— Dem. da proposição 1.1.1.I: Seja um conjunto de conjuntos convexos em  $\mathbb{R}^n$ :  $\{C_i | i \in I\}$ . Sejam também quaisquer 2 vectores  $x^1$  e  $x^2$  pertencentes a  $\bigcap_{i \in I} C_i$ . Se  $x^1$  e  $x^2$  pertencem a  $\bigcap_{i \in I} C_i$ , então também pertencem a cada  $C_i$ ,  $i \in I$ . Como são convexos por hipótese, então  $(1-\lambda)x^1 + \lambda x^2$ , qualquer que seja  $0 \leq \lambda \leq 1$ , também pertence a cada  $C_i$ , pelo que pertence a  $\bigcap_{i \in I} C_i$ .  $\bigcap_{i \in I} C_i$  é por isso um conjunto convexo. ||

——— Dem. da proposição 1.1.1.II: Elementar a partir da definição de conjunto convexo. ||

——— Dem. da proposição 1.1.1.III: Elementar a partir da definição de conjunto convexo. ||

——— Dem. da proposição 1.1.1.IV: Elementar a partir da definição de conjunto convexo. ||

——— Dem. da proposição 1.1.1.V: É imediato pela proposição 1.1.1.I que a intersecção de todos os conjuntos convexos que contêm  $D$  é ainda um conjunto convexo que contém  $D$ . É também óbvio que será o mais pequeno conjunto convexo que contém  $D$ , porque se não não seria a intersecção de todos os conjuntos convexos que contêm  $D$ . ||

——— Dem. da proposição 1.1.2.I: Está no texto que antecede a

proposição. ||

——— Dem. da proposição 1.1.2.II: Teremos que demonstrar que :

1) Subespaços são conjuntos afins que contêm a origem; 2) Conjuntos afins contendo a origem são subespaços.

1) Sejam  $x^1$  e  $x^2$  quaisquer 2 vectores dum subespaço  $S \subset \mathbb{R}^n$ . Por definição, os subespaços são fechados na multiplicação por escalares reais. Logo:

$$x^1 \in S \Rightarrow (1-\lambda)x^1 \in S, \forall \lambda \in \mathbb{R}$$

$$x^2 \in S \Rightarrow \lambda x^2 \in S, \forall \lambda \in \mathbb{R}$$

Como por definição  $S$  é também fechado na adição, vem que:

$$(1-\lambda)x^1 + \lambda x^2 \in S, \forall \lambda \in \mathbb{R}$$

Como o anterior é válido quaisquer que sejam  $x^1, x^2 \in S$ , um subespaço é portanto um conjunto afim (que contem a origem por definição de subespaço).

2) Consideremos agora que  $V$  é um conjunto afim contendo a origem. É imediato que é fechado na multiplicação por escalar já que

$$\lambda x^1 = (1-\lambda).0 + \lambda x \in V, \forall x \in V, \lambda \in \mathbb{R} \text{ (pois } 0 \in V).$$

Quanto a  $V$  ser fechado na adição repare-se que, porque é conjunto afim,

$$\frac{1}{2}(x^1 + x^2) = (1-\frac{1}{2})x^1 + \frac{1}{2}x^2 \in V, \forall x^1, x^2 \in V.$$

Logo, e porque, como vimos mais acima,  $V$  é fechado na multiplicação por qualquer escalar real:

$$x^1 + x^2 = 2 \left[ \frac{1}{2}(x^1 + x^2) \right] \in V, \forall x^1, x^2 \in V.$$

Ao ser fechado na adição e na multiplicação por escalares reais,  $V$  é um subespaço. ||

——— Dem. da proposição 1.1.2.III: Teremos que mostrar que : 1) Existe pelo menos um subespaço paralelo a cada conjunto afim; 2) Um conjunto afim não pode ser paralelo a dois subespaços distintos.

1) Seja  $V$  um conjunto afim e  $v$  um qualquer dos seus elementos. Comecemos por mostrar que  $V + \{-v\}$  é um conjunto afim que contem a o rigem. Sejam quaisquer  $x^1, x^2 \in V$ . Então  $(x^1 - v), (x^2 - v) \in V + \{-v\}$ . Ora, para qualquer  $\lambda \in \mathbb{R}$  :

$$(1-\lambda)(x^1 - v) + \lambda(x^2 - v) = [(1-\lambda)x^1 + \lambda x^2] - v \in V + \{-v\}$$

pois por definição de conjunto afim  $(1-\lambda)x^1 + \lambda x^2 \in V$ . Logo  $V + \{-v\}$  é um conjunto afim e é paralelo a  $V$ . Mas como  $v \in V$  então  $0 = v - v \in V + \{-v\}$ , pelo que  $V + \{-v\}$  é um conjunto afim que contem a origem e que é paralelo a  $V$ . Pela proposição 1.1.2.II,  $V + \{-v\}$  é um subespaço paralelo a  $V$ .

2) Sejam 2 subespaços  $S_1$  e  $S_2$  paralelos a  $V$ . Para  $S_1$  e  $S_2$  serem pa ralelos a  $V$  é óbvio que terão que ser paralelos entre si. Logo  $S_1 = S_2 + \{a\}$  para algum  $a \in \mathbb{R}^n$ . Como por definição  $0 \in S_2$  vem que  $a \in S_1$ . Como também por definição  $S_1$  é fechado na adição de elementos seus (e  $a \in S_1$ ) temos:  $S_1 \supset S_1 + \{a\} = S_2$ . Um raciocínio similar leva a concluir que igualmente  $S_2 \supset S_1$ . Logo  $S_1 = S_2$ , isto é, não poderemos ter dois subespaços distintos paralelos a  $V$ . ||

——— Dem. da proposição 1.1.3.I: Seja um subespaço  $S \subset \mathbb{R}^n$  de dimensão  $n-1$ . O seu complemento ortogonal  $S^\perp$ , definido como o conjun

to dos vectores ortogonais a todos os vectores de  $S$ , tem dimensão 1 (pois da teoria da ortogonalidade em  $\mathbb{R}^n$  sabemos que  $\dim S + \dim S^\perp = n$ ). Seja  $b$  um vector não nulo e pertencente a  $S^\perp$  (como  $S^\perp$  tem dimensão 1 todos os outros vectores de  $S^\perp$  são colineares a  $b$ ). O subespaço  $S = (S^\perp)^\perp$  é, portanto, formado por todos os vectores  $x$  ortogonais ao vector  $b$ , isto é,  $S = \{x | \langle x, b \rangle = 0\}$ .

Atendendo a que cada conjunto afim é paralelo a um único subespaço (proposição 1.1.2.III), e que um hiperplano é um conjunto afim de dimensão  $n-1$ , os hiperplanos paralelos a  $S$  podem ser obtidos através de translações de  $S$  para algum  $a \in \mathbb{R}^n$ :

$$\begin{aligned} H &= S + \{a\} = \{x+a | x \in S\} = \{x+a | \langle x, b \rangle = 0\} = \{x+a | \langle x+a-a, b \rangle = 0\} = \\ &= \{x+a | \langle x+a, b \rangle = \langle a, b \rangle\} = \{x | \langle x, b \rangle = \langle a, b \rangle\}. \end{aligned}$$

Fazendo  $\beta = \langle a, b \rangle$  vem que o conjunto  $H = \{x | \langle x, b \rangle = \beta\}$  é um hiperplano em  $\mathbb{R}^n$ . Decorre também do anterior que todos os hiperplanos paralelos ao subespaço  $S$  podem ser traduzidos desta forma, para algum valor de  $\beta$ . Como todos os hiperplanos são paralelos a um subespaço (proposição 1.1.2.III), qualquer hiperplano pode ser assim traduzido. ||

——— Dem. da proposição 1.1.3.II: Elementar a partir das definições de conjunto convexo e semiespaço. ||

——— Dem. da proposição 1.1.4.I: Teremos <sup>que</sup> mostrar apenas que a intersecção de cones é ainda um cone, já que pela proposição 1.1.1.I a intersecção de conjuntos convexos é um conjunto convexo. Seja então um conjunto de cones  $\{K_i | i \in I\}$  e seja um qualquer  $x \in \bigcap_{i \in I} K_i$ .

Como  $x$  pertence a cada  $K_i$ , e porque estes são cones, então  $\lambda x \in K_i$ ,  $\forall i \in I$ ,  $\lambda > 0$ . Então também  $\lambda x \in \bigcap_{i \in I} K_i$ ,  $\forall \lambda > 0$ , isto é,  $\bigcap_{i \in I} K_i$  é um cone. ||

——— Dem. da proposição 1.1.4.II: Pela definição de cone e de conjunto convexo é imediato que  $K$  é um cone convexo. Fazendo  $\lambda=1$  conclui-se também que  $K \supset C$ . Mostremos por absurdo que  $K$  é o mais pequeno cone que contém  $C$ . Seja  $x \in K$  e tal que  $\{y | y = \mu x, \mu > 0\} \cap C = \emptyset \Leftrightarrow \{y | y = \mu x, y \in C, \mu > 0\} = \emptyset \Leftrightarrow \mu \{x | \mu > 0, x \in \frac{1}{\mu} \cdot C\} = \emptyset \Leftrightarrow \frac{1}{\lambda} \{x | \lambda > 0, x \in \lambda C\} = \emptyset$ . Mas o anterior contraria a hipótese  $x \in K$ . Logo,  $K$  é o mais pequeno cone convexo que contém  $C$ . ||

——— Dem. da proposição 1.1.4.III: Elementar pelas definições de cone e de conjunto convexo. ||

——— Dem. da proposição 1.1.5.I: Se  $\text{cl}C$  é vazio, então é trivialmente convexo. Suponha-se então que  $\text{cl}C$  é não vazio e sejam  $x^1, x^2$  dois quaisquer elementos de  $\text{cl}C$ . Temos que mostrar que para qualquer  $0 \leq \lambda \leq 1$

$$z = (1-\lambda)x^1 + \lambda x^2 \in \text{cl}C.$$

Dado um qualquer  $\epsilon > 0$  tão pequeno quanto queiramos, escolham-se  $y^1, y^2 \in C$  tais que  $\|x^1 - y^1\| \leq \epsilon$  e  $\|x^2 - y^2\| \leq \epsilon$ . Então também

$$(1-\lambda) \cdot \|x^1 - y^1\| + \lambda \|x^2 - y^2\| \leq \epsilon \Leftrightarrow \|(1-\lambda)x^1 - (1-\lambda)y^1\| + \|\lambda x^2 - \lambda y^2\| \leq \epsilon$$

Pela desigualdade de SCHWARTZ vem ainda que:

$$\|[(1-\lambda)x^1 + \lambda x^2] - [(1-\lambda)y^1 + \lambda y^2]\| \leq \epsilon$$

Fazendo  $z = (1-\lambda)x^1 + \lambda x^2$  e  $w = (1-\lambda)y^1 + \lambda y^2$  fica:  $\|z-w\| \leq \varepsilon \Leftrightarrow d(z,w) \leq \varepsilon$ .

Como  $y_1, y_2 \in C$  vem que  $w \in C$ . Logo, e como  $\varepsilon > 0$  é tão pequeno quanto queiramos,  $C \cap E_\varepsilon(z) \neq \emptyset$ , isto é,  $z \in cl C$ . ||

——— Dem. da proposição 1.1.5.II: Pela definição de fecho de um conjunto, o limite de qualquer sucessão convergente de elementos de  $D$  pertence a  $cl D$ . Suponhamos que  $D$  é não fechado. Seja então  $\bar{x}$  tal que  $\bar{x} \notin D$  e  $\bar{x} \in cl D$ . Como há elementos de  $D$  tão próximos quanto queiramos de  $\bar{x}$  (porque senão  $\bar{x}$  não pertenceria a  $cl D$ ), podemos considerar  $\bar{x}$  como ponto de acumulação de  $D$ , isto é, como limite de pelo menos uma sucessão de elementos de  $D$ . Conclusão: se  $D$  é fechado, qualquer sucessão de elementos de  $D$  tem como limite um ponto de  $cl D = D$ ; se  $D$  é não fechado, então existem sucessões de elementos de  $D$  que convergem para vectores não pertencentes a  $D$ . ||

——— Dem. da proposição 1.1.5.III: Seja um conjunto de conjuntos fechados  $\{D_i | i \in I\}$ . Se  $\bigcap_{i \in I} D_i = \emptyset$ , então é fechado. Se  $\bigcap_{i \in I} D_i \neq \emptyset$  seja  $x \in cl \bigcap_{i \in I} D_i$ . Então  $[\bigcap_{i \in I} D_i] \cap E_\varepsilon(x) \neq \emptyset$ . Consequentemente:  $D_i \cap E_\varepsilon(x) \neq \emptyset, \forall i \in I \Leftrightarrow x \in cl D_i, \forall i \in I$ . Como cada  $D_i$  é fechado por hipótese, então  $x \in D_i, \forall i \in I \Rightarrow x \in \bigcap_{i \in I} D_i$ . Já que o anterior é válido qualquer que seja  $x \in cl \bigcap_{i \in I} D_i$ , então  $cl \bigcap_{i \in I} D_i = \bigcap_{i \in I} D_i$ . ||

——— Dem. da proposição 1.1.5.IV: Elementar pela definição de conjunto fechado e de fecho de um conjunto. ||

——— Dem. da proposição 1.1.5.V: Se  $C$  é fechado, vem pela definição, que  $C = \{x | C \cap E_\varepsilon(x) \neq \emptyset, \forall x \in C, \varepsilon > 0\}$ . Logo  $C + \{a\} = \{x+a | C \cap E_\varepsilon(x) \neq \emptyset, \forall x \in C, \varepsilon > 0\} = \{x | (C + \{a\}) \cap E_\varepsilon(x) \neq \emptyset, \forall x \in (C + \{a\}), \varepsilon > 0\}$ , pelo

que  $C + \{a\}$  é um conjunto fechado. ||

——— Dem. da proposição 1.1.5.VI:  $X(\bar{y}) = \{x | (x, y) \in T(\bar{y})\}$ , com  $T(\bar{y}) = T \cap \{(x, y) | y \geq \bar{y}\} \cap \{(x, y) | y \leq \bar{y}\}$ . É óbvio pela definição que  $\{(x, y) | y \geq \bar{y}\}$  e  $\{(x, y) | y \leq \bar{y}\}$  são conjuntos fechados. Pela proposição 1.1.5.III,  $T(\bar{y})$  é um conjunto fechado. Ora,  $T(\bar{y}) + \{(0, -\bar{y})\} = \{(x, 0) | x \in X(\bar{y})\}$  é, pela proposição 1.1.5.V, um conjunto fechado. Consequentemente, é imediato, pela definição de conjunto fechado, que  $X(\bar{y})$  é fechado. ||

——— Dem. da proposição 1.1.5.VII: Pela definição de fecho, o fecho do complementar de  $D$  é:  $cl\bar{D} = \{x | \bar{D} \cap E_\varepsilon(x) \neq \emptyset, \forall \varepsilon > 0\}$ . Logo:  $\text{int } D = (\overline{cl\bar{D}}) = \{x | x \notin cl\bar{D}\} = \{x | \exists \varepsilon > 0: \bar{D} \cap E_\varepsilon(x) = \emptyset\} = \{x | \exists \varepsilon > 0: E_\varepsilon(x) \subset D\}$ . Demonstremos agora que  $D$  é aberto sse  $D = \text{int } D$ . Como por definição  $\text{int } D$  é o maior conjunto aberto contido em  $D$ , se  $D$  é aberto então  $D = \text{int } D$ . Por outro lado, se  $D = \text{int } D$ , e pela definição de  $\text{int } D$ ,  $D$  é um conjunto aberto. ||

——— Dem. da proposição 1.1.5.VIII: Se  $\text{int } D = \emptyset$  então é trivialmente convexo. Suponha-se então que  $\text{int } C$  é não vazio e sejam quaisquer  $x^1, x^2 \in \text{int } C$ . Temos que mostrar que, para qualquer  $0 \leq \lambda \leq 1$ ,  $z = (1-\lambda)x^1 + \lambda x^2 \in \text{int } C$ . Como  $x^1, x^2 \in C$ , pela proposição 1.1.5.VII existe um  $\varepsilon > 0$  tal que  $E_\varepsilon(x^1) \subset C$  e  $E_\varepsilon(x^2) \subset C$ . Seja um qualquer vector  $w$  tal que  $\|w\| \leq \varepsilon$ . Então, e porque  $E_\varepsilon(x^1)$  e  $E_\varepsilon(x^2)$  estão contidas em  $C$ , vem que  $(x^1 + w), (x^2 + w) \in C$ . Como  $C$  é um conjunto convexo,  $z + w = (1-\lambda)x^1 + \lambda x^2 + w = (1-\lambda)(x^1 + w) + \lambda(x^2 + w) \in C$ ,  $\forall w \in \mathbb{R}^n$  com  $\|w\| \leq \varepsilon$ . Isto significa que  $E_\varepsilon(z) \subset C$ , pelo que  $z \in \text{int } C$  (proposição 1.1.5.VII). ||

— Dem. da proposição 1.1.5.IX: Imediata pelas definições e pela proposição 1.1.5.VII. ||

— Dem. do corolário 1.1.5.IX.1: Imediata pelas proposições 1.1.5.VII e 1.1.5.IX. ||

— Dem. da proposição 1.1.5.X: Se  $\text{rint } D = \emptyset$ , então é trivialmente convexo. Seja então  $\text{rint } D \neq \emptyset$ , e sejam  $x^1, x^2 \in \text{rint } D$ . Precisamos de mostrar que qualquer que seja  $0 < \lambda < 1$ , também  $(1-\lambda)x^1 + \lambda x^2 \in \text{rint } D$ . Pela proposição 1.1.5.IX tal será equivalente a mostrar que existe um  $\varepsilon > 0$  tal que, qualquer que seja  $0 < \lambda < 1$ , temos:

$$\begin{aligned} & [E_\varepsilon[(1-\lambda)x^1 + \lambda x^2] \cap V(D)] \subset D \Leftrightarrow \\ & \Leftrightarrow [\{(1-\lambda)x^1 + \lambda x^2\} + (E_\varepsilon(0) \cap S(D))] \subset D \quad \dots (A) \end{aligned}$$

em que  $S(D) = V(D) - \{(1-\lambda)x^1 + \lambda x^2\}$ . Repare-se que  $S(D)$  é um conjunto afim paralelo a  $V(D)$ . Porque  $x^1, x^2 \in D$  e porque  $D$  é convexo, então  $(1-\lambda)x^1 + \lambda x^2 \in D$ , o que implica que  $(1-\lambda)x^1 + \lambda x^2 \in V(D)$ . Consequentemente,  $S(D)$  contém a origem, pelo que é um subespaço (Proposição 1.1.2.II). Note-se também que  $x^2 \in \text{rint } D \Rightarrow x^2 \in D \Rightarrow x^2 \in [D + (E_\varepsilon(0) \cap S(D))]$ ,  $\forall \varepsilon > 0$ . Então,  $\forall \varepsilon > 0$

$$\begin{aligned} & [\{(1-\lambda)x^1 + \lambda x^2\} + (E_\varepsilon(0) \cap S(D))] \subset [\{(1-\lambda)x^1\} + \lambda [D + (E_\varepsilon(0) \cap S(D))] + \\ & + (E_\varepsilon(0) \cap S(D))] = [\{(1-\lambda)x^1\} + (1+\lambda) [E_\varepsilon(0) \cap S(D)] + \lambda D = \\ & = (1-\lambda) \cdot [x^1 + \frac{1+\lambda}{1-\lambda} [E_\varepsilon(0) \cap S(D)]] + \lambda D = (1-\lambda) [x^1 + (E_\delta(0) \cap S(D))] + \lambda D \\ & \dots (B) \end{aligned}$$

em que  $\delta = \varepsilon(1+\lambda) / (1-\lambda)$  (note-se que  $\frac{1+\lambda}{1-\lambda} \cdot S(D) = S(D)$  porque  $S(D)$  é um subespaço e, por isso, fechado na multiplicação por escalar). Mas como  $x^1 \in \text{rint } D$  temos que existe algum  $\delta > 0$  tal que  $[x^1 + (E_\delta(0) \cap S(D))] \subset D$ . Logo, por (B), teremos para algum  $\delta > 0$ :

$$\begin{aligned} & [\{(1-\lambda)x^1 + \lambda x^2\} + (E_\varepsilon(0) \cap S(D))] \subset [(1-\lambda) [x^1 + (E_\delta(0) \cap S(D))] + \lambda D] \subset \\ & \subset (1-\lambda) D + \lambda D = D \end{aligned}$$

Está demonstrado que  $(1-\lambda)x^1 + \lambda x^2 \in \text{rint } D$ ,  $\forall x^1, x^2 \in \text{rint } D$ ,  $0 \leq \lambda \leq 1$ . ||

——— Dem. da proposição 1.1.5.XI: Se  $D$  é um conjunto aberto, então, pela proposição 1.1.5.VII,  $D = \text{int } D$ . Pelo corolário 1.1.5.IX.1 resulta que  $D = \text{int } D \subset \text{rint } D$ . Mas pela proposição 1.1.5.IX é óbvio que  $D \supset \text{rint } D$ . Logo,  $D = \text{rint } D$ , pelo que  $D$  é um conjunto relativamente aberto. ||

——— Dem. da proposição 1.1.6.I: Façamos  $\inf \{ \|y - x\| \mid x \in C \} = \gamma$ . Ao anterior ínfimo está necessariamente associada uma sucessão  $\{x^k\}$  tal que

$$\|y - x^k\| \rightarrow \gamma \quad \dots (A)$$

$$\begin{aligned} \text{Ora, } \|x^k - x^m\|^2 &= \langle x^k - x^m, x^k - x^m \rangle = \langle x^k, x^k \rangle + \langle x^m, x^m \rangle - 2\langle x^k, x^m \rangle = \\ &= 2\langle x^k - y, x^k - y \rangle + 2\langle x^m - y, x^m - y \rangle - \langle x^k + x^m - 2y, x^k + x^m - 2y \rangle \\ \Leftrightarrow \|x^k - x^m\|^2 &= 2\|x^k - y\|^2 + 2\|x^m - y\|^2 - 4\left\|\frac{x^k + x^m}{2} - y\right\|^2 \quad \dots (B) \end{aligned}$$

Como  $x^k, x^m \in C$  e  $C$  é um conjunto convexo, então também  $\frac{x^k + x^m}{2} \in C$ .

Logo, porque definimos  $\gamma = \inf \{ \|y - x\| \mid x \in C \}$ , teremos que

$$\left\| \frac{x^k + x^m}{2} - y \right\|^2 \geq \gamma^2$$

$\Leftrightarrow$  (atendendo a (B)):

$$\|x^k - x^m\| \leq 2\|x^k - y\|^2 + 2\|x^m - y\|^2 - 4\gamma^2 \quad \dots (C)$$

Fazendo  $k$  e  $m$  suficientemente grandes teremos (por (A)) que  $\|x^k - y\|$  e  $\|x^m - y\|$  serão suficientemente próximos de  $\gamma$  (por valores superiores), ou, o que é o mesmo (atendendo a (C))  $\|x^k - x^m\|$  será suficientemente

próximo de zero. Pela condição necessária e suficiente de convergência de sucessões de CAUCHY-BOLZANO, podemos então concluir que a sucessão  $\{x^k\}$  é convergente para algum vector  $\bar{x}$ . Como por hipótese  $C$  é fechado, então, pela proposição 1.1.5.II,  $\bar{x} \in C$ , ou seja, existe um vector  $\bar{x} \in C$  que minimiza a distância ao vector  $y \notin C$ .

Para provar a unicidade suponhamos que existe um outro vector  $\bar{x}' \in C$  tal que  $\|y - \bar{x}\| = \|y - \bar{x}'\| = \gamma$ . ... (D)

Como  $\bar{x}, \bar{x}' \in C$  e como  $C$  é convexo,  $(\bar{x} + \bar{x}')/2 \in C$  pelo que

$$\left\| \frac{\bar{x} + \bar{x}'}{2} - y \right\|^2 \geq \gamma^2 \Leftrightarrow \left\| \frac{\bar{x} + \bar{x}'}{2} - y \right\| \geq \gamma \quad \dots (E)$$

Mas por outro lado, pela desigualdade de SCHWARTZ,

$$\begin{aligned} \left\| \frac{\bar{x} + \bar{x}'}{2} - y \right\| &\leq \frac{1}{2} \|y - \bar{x}\| + \frac{1}{2} \|y - \bar{x}'\| \\ \Leftrightarrow (\text{por (D)}) \quad \left\| \frac{\bar{x} + \bar{x}'}{2} - y \right\| &\leq \gamma \quad \dots (F) \end{aligned}$$

Reunindo (E) e (F) fica:  $\left\| \frac{\bar{x} + \bar{x}'}{2} - y \right\| = \gamma$

$$\Leftrightarrow \left\| \frac{(y - \bar{x}) + (y - \bar{x}')}{2} \right\| = \gamma \quad \dots (G)$$

Elevando ao quadrado obtemos:

$$\begin{aligned} \left\langle \frac{(y - \bar{x}) + (y - \bar{x}')}{2}, \frac{(y - \bar{x}) + (y - \bar{x}')}{2} \right\rangle &= \gamma^2 \\ \Leftrightarrow \langle y - \bar{x}, y - \bar{x} \rangle + \langle y - \bar{x}, y - \bar{x}' \rangle + 2 \langle y - \bar{x}, y - \bar{x}' \rangle &= 4\gamma^2 \end{aligned}$$

Como por (D),  $\langle y - \bar{x}, y - \bar{x} \rangle = \langle y - \bar{x}', y - \bar{x}' \rangle = \gamma^2$  fica:

$$\langle y - \bar{x}, y - \bar{x}' \rangle = \gamma^2$$

pelo que, continuando a ter em conta (D):

$$\cos (y - \bar{x}, y - \bar{x}') = \frac{\langle y - \bar{x}, y - \bar{x}' \rangle}{\|y - \bar{x}\| \cdot \|y - \bar{x}'\|} = \frac{\gamma^2}{\gamma \cdot \gamma} = 1$$

donde se conclui que os vectores  $(y-\bar{x})$  e  $(y-\bar{x}')$  são colineares, isto é,

$$y-\bar{x} = \lambda(y-\bar{x}') \quad \text{para algum } \lambda \in \mathbb{R}.$$

Como por (D)  $\|y-\bar{x}\| = \|y-\bar{x}'\|$ , teremos necessariamente que  $|\lambda| = 1$ .

Claramente  $\lambda \neq -1$ , pois se assim não fosse teríamos

$$y-\bar{x} = (-1) \cdot (y-\bar{x}') \Leftrightarrow y = \frac{\bar{x}+\bar{x}'}{2} \in C$$

o que não acontece por hipótese. Então  $\lambda=1$ , isto é,

$$y-\bar{x} = y-\bar{x}' \Leftrightarrow \bar{x}=\bar{x}'$$

Está estabelecida a unicidade.

Resta demonstrar que  $\langle x-\bar{x}, y-\bar{x} \rangle \leq 0 \quad \forall x \in C$ , é condição necessária e suficiente para  $\bar{x} \in C$  minimizar a distância entre  $y$  e  $C$ .

Para provar a suficiência admita-se que  $\langle x-\bar{x}, y-\bar{x} \rangle \leq 0$  para algum  $\bar{x}$  e para todo o  $x \in C$ . Teremos,  $\forall x \in C$ :

$$\begin{aligned} \|y-x\|^2 &= \|y-\bar{x} + \bar{x} - x\|^2 = \langle (y-\bar{x}) - (x-\bar{x}), (y-\bar{x}) - (x-\bar{x}) \rangle = \\ &= \|y-\bar{x}\|^2 + \|x-\bar{x}\|^2 - 2\langle x-\bar{x}, y-\bar{x} \rangle \end{aligned}$$

$$\Leftrightarrow (\text{pois por hipótese } \langle x-\bar{x}, y-\bar{x} \rangle \leq 0, \forall x \in C)$$

$$\|y-x\|^2 \geq \|y-\bar{x}\|^2 + \|x-\bar{x}\|^2, \quad \forall x \in C$$

Como  $\|x-\bar{x}\|^2 \geq 0$  fica:

$$\|y-x\| \geq \|y-\bar{x}\|, \quad \forall x \in C$$

isto é,  $\bar{x}$  minimiza a distância ao ponto  $y \notin C$ .

Para provar a necessidade, notemos que por  $C$  ser convexo, e dado que

$$\bar{x} \in C: \quad (1-\lambda)\bar{x} + \lambda x = \bar{x} + \lambda(x-\bar{x}) \in C, \quad \forall x \in C, \quad 0 \leq \lambda \leq 1.$$

Então, atendendo a <sup>que</sup>  $\bar{x}$  minimiza a distância a  $y$ ,

$$\|y - [\bar{x} + \lambda(x-\bar{x})]\|^2 \geq \|y-\bar{x}\|^2, \quad \forall x \in C, \quad 0 \leq \lambda \leq 1 \quad \dots(H)$$

Ora,  $\|y - [\bar{x} + \lambda(x - \bar{x})]\|^2 = \|y - \bar{x}\|^2 + \lambda^2 \|x - \bar{x}\|^2 - 2\lambda \langle x - \bar{x}, y - \bar{x} \rangle$ .

Substituindo em (H) fica:

$$\lambda^2 \|x - \bar{x}\|^2 - 2\lambda \langle x - \bar{x}, y - \bar{x} \rangle \geq 0, \quad \forall x \in C, 0 \leq \lambda \leq 1$$

Para  $\lambda = 0$  o anterior é trivialmente verdadeiro. Para  $0 < \lambda \leq 1$  divide-se por  $\lambda$  e passe-se ao limite quando  $\lambda \rightarrow 0$ . O resultado será

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0, \quad \forall x \in C.$$

Está completada a demonstração. ||

——— Dem. da proposição 1.1.6.II: Seja  $C$  um conjunto convexo não vazio e fechado e seja  $y \in C$ . Pela proposição 1.1.6.I existe um único vector  $\bar{x} \in C$  (minimizador da distância a  $y$ ) que respeita

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0, \quad \forall x \in C$$

$$\Leftrightarrow \langle x, y - \bar{x} \rangle \leq \langle \bar{x}, y - \bar{x} \rangle, \quad \forall x \in C$$

mas tendo em conta que  $\|y - \bar{x}\|^2 = \langle y - \bar{x}, y - \bar{x} \rangle = \langle y, y - \bar{x} \rangle - \langle \bar{x}, y - \bar{x} \rangle$ ,

fica:  $\|y - \bar{x}\|^2 \leq \langle y, y - \bar{x} \rangle - \langle x, y - \bar{x} \rangle, \quad \forall x \in C$

Fazendo  $b = y - \bar{x} \neq 0$  (pois  $y \notin C$ , e logo  $y \neq \bar{x}$ ), vem:

$$\langle y - x, b \rangle \geq \|y - \bar{x}\|^2, \quad \forall x \in C \Leftrightarrow \langle x, b \rangle \leq -\|y - \bar{x}\|^2 + \langle y, b \rangle, \quad \forall x \in C$$

Como  $\|y - \bar{x}\|^2 > 0$  (pois, como vimos acima,  $b = y - \bar{x} \neq 0$ ) obtemos:

$$\langle x, b \rangle < \langle y, b \rangle, \quad \forall x \in C$$

Fazendo  $\beta = \sup \{ \langle x, b \rangle \mid x \in C \}$  teremos, como queríamos mostrar:

$$\langle x, b \rangle \leq \beta, \quad \forall x \in C \quad \text{e} \quad \langle y, b \rangle > \beta,$$

pois o supremo será atingido necessariamente para um elemento de  $C$  (porque este é fechado). ||

— Dem. do corolário 1.1.6.II.1: É obvio que  $C$  está contido na intersecção dos semiespaços fechados que o contêm. Para provar inversamente que  $C$  contém a intersecção dos semiespaços fechados que o contêm, vamos proceder por absurdo. Suponhamos que existe um ponto  $y$  que pertence à intersecção dos semiespaços fechados que contêm  $C$ , mas que não pertence a  $C$ . Então, pela proposição 1.1.6.II, existirá um hiperplano que separa fortemente  $y$  e o conjunto  $C$ , isto é, será possível definir um novo semiespaço que contém  $C$  e que não contém  $y$ . Mas isto contradiz o facto de  $y$  pertencer a todos os semiespaços fechados que contêm  $C$ . Consequentemente,  $C$  contém a intersecção dos semiespaços fechados que contêm  $C$ . Está demonstrado que  $C$  é intersecção dos semiespaços fechados que o contêm. ||

— Dem. da proposição 1.1.7.I: Seja  $\bar{x}$  um ponto da fronteira de  $C$ . Existe pelo menos uma sucessão  $\{y^k\}$ , com cada  $y^k \notin C$ , tal que  $y^k \rightarrow \bar{x}$  (pois  $\bar{x}$  está na fronteira de  $C$ ). Pela proposição 1.1.6.II, existe para cada  $y^k$  um hiperplano  $\langle x, b^k \rangle = \beta$  tal que  $\langle b^k, y^k \rangle > \langle x, b^k \rangle$ ,  $\forall x \in C$ . Divida-se a anterior desigualdade por  $\|b^k\| > 0$ , donde fica:

$$\left\langle \frac{b^k}{\|b^k\|}, y^k - x \right\rangle > 0, \quad \forall x \in C \quad \dots (A)$$

Sendo  $\{b^k / \|b^k\|\}$  uma sucessão limitada por construção (pois qualquer  $b^k / \|b^k\|$  tem norma 1), existirá pelo menos uma subsucessão convergente da sucessão  $\{b^k / \|b^k\|\}$ . Restrinjamo-nos aos termos desta subsucessão convergente, e seja  $\bar{b} / \|\bar{b}\|$  o seu limite. Passando (A) ao limite quando  $y^k \rightarrow \bar{x}$  fica então

$$\left\langle \frac{\bar{b}}{\|\bar{b}\|}, \bar{x} - x \right\rangle \geq 0, \quad \forall x \in C$$

isto é, existe um hiperplano de suporte a  $C$  no ponto  $\bar{x}$  definido pela equação

$$\left\langle \frac{\bar{b}}{\|\bar{b}\|}, \bar{x} - x \right\rangle = 0. \quad \parallel$$

SECÇÃO 1.2.

——— Dem. da proposição 1.2.2.1: A condição  $f(x) > -\infty, \forall x \in \mathbb{R}^n$ , é imposta de modo a impedir ambiguidades de tipo  $\infty - \infty$ . Sejam então quaisquer dois vectores  $x^1, x^2$  de  $\mathbb{R}^n$ . Os vectores de  $\mathbb{R}^{n+1}$  dados por  $(x^1, f(x^1))$  e  $(x^2, f(x^2))$  pertencem por definição a  $\text{epi} f(x)$ . Se  $f(x)$  é convexa,  $\text{epi} f(x)$  é convexo, pelo que

$$(1-\lambda)(x^1, f(x^1)) + \lambda(x^2, f(x^2)) \in \text{epi } f(x)$$

$$\Leftrightarrow ([ (1-\lambda)x^1 + \lambda x^2 ], [(1-\lambda)f(x^1) + \lambda f(x^2)]) \in \text{epi } f(x)$$

$$\Leftrightarrow f[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f(x^1) + \lambda f(x^2)$$

Mostramos que  $\text{epi } f$  convexo implica a anterior desigualdade.

Admitamos agora esta última desigualdade:  $f[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f(x^1) + \lambda f(x^2), \forall x^1, x^2 \in \mathbb{R}^n$ .

Mostremos inversamente que ela implica que  $\text{epi } f(x)$  seja convexo.

Sejam quaisquer  $r_1 \geq f(x^1)$  e  $r_2 \geq f(x^2)$ . Temos então:

$$f[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)r_1 + \lambda r_2, \forall (x^1, r_1), (x^2, r_2) \in \text{epi } f(x)$$

$$\Leftrightarrow ([ (1-\lambda)x^1 + \lambda x^2 ], [(1-\lambda)r_1 + \lambda r_2]) \in \text{epi } f(x), \forall (x^1, r_1), (x^2, r_2) \in \text{epi } f.$$

$$\Leftrightarrow \text{epi } f \text{ é um conjunto convexo. } \parallel$$

——— Dem. do corolário 1.2.2.1.1: Seja  $f[\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_m x^m]$ . Sem perda de generalidade podemos considerar todos os  $\lambda_i \neq 0$  ( $i=1, \dots, m$ ), pois em caso contrário basta eliminá-los da expressão, di-

minuindo correspondentemente o valor do  $m$ . Se cada  $\lambda_i$  é não nulo, então também nenhum é igual a 1, a menos que  $m=1$ . Ora se  $m=1$ , o resultado é imediato pela proposição 1.2.2.I. Se  $m>1$ , é válido, para quaisquer  $x^1, x^2, \dots, x^m$ , que:

$$f[\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_m x^m] = f\left[\lambda_1 x^1 + (1-\lambda_1) \frac{\lambda_2 x^2 + \dots + \lambda_m x^m}{1-\lambda_1}\right]$$

Assim sendo, e pela proposição 1.2.2.I, se  $f$  é convexa temos:

$$\begin{aligned} f[\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_m x^m] &\leq \lambda_1 f(x^1) + (1-\lambda_1) f\left(\frac{\lambda_2 x^2 + \dots + \lambda_m x^m}{1-\lambda_1}\right) = \\ &= \lambda_1 f(x^1) + (1-\lambda_1) f\left[\frac{\lambda_2}{1-\lambda_1} x^2 + \left(1 - \frac{\lambda_2}{1-\lambda_1}\right) \frac{(1-\lambda_1)(\lambda_3 x^3 + \dots + \lambda_m x^m)}{1-\lambda_1 - \lambda_2}\right] \end{aligned}$$

Aplicando novamente o mesmo argumento fica:

$$\begin{aligned} f[\lambda_1 x^1 + \lambda_2 x^2 + \lambda_3 x^3 + \dots + \lambda_m x^m] &\leq \lambda_1 f(x^1) + \lambda_2 f(x^2) + \\ &+ \left(1 - \frac{\lambda_2}{1-\lambda_1}\right) f\left[\frac{(1-\lambda_1)(\lambda_3 x^3 + \dots + \lambda_m x^m)}{1-\lambda_1 - \lambda_2}\right] \end{aligned}$$

Prosseguindo, sucessivamente, obteremos (tendo em conta que  $\sum \lambda_i = 1$ )

$$f[\lambda_1 x^1 + \dots + \lambda_m x^m] \leq \lambda_1 f(x^1) + \dots + \lambda_m f(x^m).$$

Consideremos agora inversamente que se verifica a anterior desigualdade. Mostremos que ela implica que a função é convexa, isto é, que o epigrafo da função é um conjunto convexo. Para isso consideramos  $r_1 \geq f(x^1)$ ,  $r_2 \geq f(x^2)$ , ...,  $r_m \geq f(x^m)$ , isto é,  $(x^1, r_1)$ ,

$(x^2, r_2), \dots, (x^m, r_m)$  e epi  $f$ . Pela desigualdade vem:

$$f[\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_m x^m] \leq \lambda_1 r_1 + \lambda_2 r_2 + \dots + \lambda_m r_m, \quad \forall r_1 \geq f(x^1), \dots$$

$$\Leftrightarrow (\lambda_1 x^1 + \dots + \lambda_m x^m; \lambda_1 r_1 + \dots + \lambda_m r_m) \in \text{epi } f(x), \quad \forall (x^1, r_1) \in \text{epi } f, \dots$$

ou seja, epi  $f$  é um conjunto convexo, conforme se confirma facilmente a partir da definição (para o que basta desenvolver um argumento recursivo similar ao efectuado acima). ||

— Dem. da proposição 1.2.2.II: Sendo  $f_1$  e  $f_2$  funções convexas próprias fica, nomeadamente, afastada a hipótese de verificação de qualquer ambiguidade do tipo  $\infty - \infty$ . Para além disso, é aplicável a proposição 1.2.2.II. Desta proposição vem que,  $\forall x^1, x^2 \in \mathbb{R}^n$ ,  $0 \leq \lambda \leq 1$  (e porque  $f_1$  e  $f_2$  são convexas):

$$f_1[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f_1(x^1) + \lambda f_1(x^2)$$

$$f_2[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f_2(x^1) + \lambda f_2(x^2)$$

$$\begin{aligned} f_1[(1-\lambda)x^1 + \lambda x^2] + f_2[(1-\lambda)x^1 + \lambda x^2] &\leq (1-\lambda)(f_1(x^1) + f_2(x^1)) + \\ &+ \lambda(f_1(x^2) + f_2(x^2)) \end{aligned}$$

Consequentemente  $f_1 + f_2$  é uma função convexa. Como  $f_1$  e  $f_2$  são próprias é imediato pela definição que  $f_1 + f_2$  é própria. ||

— Dem. da proposição 1.2.2.III: Seja por exemplo o conjunto  $S_f^{\leq}(\alpha)$  para algum  $\alpha$  (para  $S_f^<(\alpha)$  a demonstração é perfeitamente idêntica).  $S_f^{\leq}(\alpha)$  pode ser encarado como a projecção sobre  $\mathbb{R}^n$  de um conjunto convexo dado pela intersecção de epi  $f$  e do semiespaço  $\{(x, r) | r \leq \alpha\}$  (ambos conjuntos convexas). Ora, a pro

jeccção sobre  $\mathbb{R}^n$  de um conjunto de  $\mathbb{R}^{n+1}$  consegue-se através de uma transformação linear com matriz  $A = [I_n; 0]$ , em que  $I_n$  é a matriz identidade de ordem  $n$ . Pela proposição 1.1.1.IV,  $S_f^\leq(\alpha)$  é então um conjunto convexo para cada  $\alpha \in [-\infty, +\infty]$ . ||

——— Dem. da proposição 1.2.2.IV: Como cada função  $f_i$  é convexa, os seus epígrafos são conjuntos convexos. Como o epígrafo de  $f(x)$  é dado por

$$\text{epi } f = \bigcap_{i \in I} \text{epi } f_i$$

é um conjunto convexo (prop. 1.1.1.I). A função  $f$  é por isso convexa. ||

——— Dem. da proposição 1.2.2.V: O epígrafo de  $f(x)$  é dado por

$$\text{epi } f(x) = F + \{(0, r) | r \geq 0\}$$

isto é, pela soma do conjunto  $F$  com o semirecta vertical ascendente emanada da origem (que é um conjunto convexo, como se confirma facilmente pela definição). A proposição 1.1.1.II permite concluir que  $\text{epi } f$  é um conjunto convexo, pelo que a função é convexa. ||

——— Dem. da proposição 1.2.2.VI: O domínio efectivo de uma função convexa é a projecção sobre  $\mathbb{R}^n$  do seu epígrafo convexo. Pelo que vimos na demonstração da proposição 1.2.2.III, a projecção sobre  $\mathbb{R}^n$  de um conjunto de  $\mathbb{R}^{n+1}$  é o resultado de uma transformação linear da matriz  $A = [I_n; 0]$ . Consequentemente, pela proposição 1.1.1.IV, o domínio efectivo será um conjunto convexo. ||

——— Dem. da proposição 1.2.3.I: Elementar a partir das definições de conjunto convexo, função convexa, função própria e função indicadora. ||

——— Dem. da proposição 1.2.4.I: Mostrar que o epigrafo de uma função linearmente homogênea é um cone significa mostrar que é fechado na multiplicação por escalar positivo. Seja  $(x, r) \in \text{epi } f(x)$ . Como  $r \geq f(x)$  e porque  $f(x)$  é linearmente homogênea vem que  $\lambda r \geq \lambda f(x) = f(\lambda x)$ . Logo também  $(\lambda x, \lambda r) = \lambda(x, r) \in \text{epi } f(x)$ , isto é, o epi  $f(x)$  é fechado na multiplicação por escalar positivo. ||

——— Dem. da proposição 1.2.4.II: Seja  $f(x)$  uma função linearmente homogênea de  $\mathbb{R}^n$  em  $]-\infty, +\infty]$ . Pela proposição 1.2.2.I ela é convexa sse, para qualquer  $0 \leq \lambda \leq 1$ :

$$f[(1-\lambda)z^1 + \lambda z^2] \leq (1-\lambda)f(z^1) + \lambda f(z^2), \quad \forall z^1, z^2 \in \mathbb{R}^n$$

$$\Leftrightarrow f[(1-\lambda)z^1 + \lambda z^2] \leq f[(1-\lambda)z^1] + f[\lambda z^2], \quad \forall z^1, z^2 \in \mathbb{R}^n$$

Fazendo  $x^1 = (1-\lambda)z^1$  e  $x^2 = \lambda z^2$  vem equivalentemente:

$$f(x^1 + x^2) \leq f(x^1) + f(x^2), \quad \forall x^1, x^2 \in \mathbb{R}^n. \quad ||$$

——— Dem. da proposição 1.2.4.III: Evidente a partir da definição de função convexa e da proposição 1.1.1.III. ||

——— Dem. da proposição 1.2.4.IV: Para dado  $\lambda > 0$  temos  $\lambda(\text{epi } f) = \{ \lambda(x, r) | r \geq f(x) \} = \{ (\lambda x, \lambda r) | r \geq f(x) \}$ . Substituindo  $\lambda x$  por  $x$  ( $x$  por  $x/\lambda$ ) e  $\lambda r$  por  $r$  ( $r$  por  $r/\lambda$ ) fica, equivalentemente:

$$\lambda(\text{epi } f) = \{(x, r) \mid r/\lambda \geq f(x/\lambda)\} = \{(x, r) \mid r \geq \lambda f(x/\lambda)\} = \text{epi}[\lambda f(x/\lambda)] = \\ = \text{epi}((f\lambda)(x)). \quad ||$$

——— Dem. da proposição 1.2.4.V: Pela proposição 1.2.4.I o epigrafo de uma função linearmente homogênea é um cone em  $\mathbb{R}^{n+1}$ . Logo, a maior função linearmente homogênea majorada por  $f(x)$  terá como epigrafo o cone convexo gerado por  $\text{epi } f(x)$ , isto é:

$$\text{epi } h(x) = \{(x, r) \mid \lambda > 0, r \geq f(x)\} = \{(x, r) \mid \lambda > 0, \frac{r}{\lambda} \geq f(\frac{x}{\lambda})\} = \\ = \{(x, r) \mid \lambda > 0, r \geq \lambda f(x/\lambda)\} = \{(x, r) \mid \lambda > 0, r \geq (f\lambda)(x)\}$$

Pela definição de epigrafo vem então:

$$h(x) = \inf \{ r \mid \lambda > 0, r \geq (f\lambda)(x) \} = \inf \{ (f\lambda)(x) \mid \lambda > 0 \}. \quad ||$$

——— Dem. da proposição 1.2.4.VI: Pela definição, uma função convexa é linearmente homogênea sse coincide com a função convexa linearmente homogênea por ela gerada. Por outras palavras, uma função convexa é linearmente homogênea sse o seu epigrafo é um cone convexo (pois é convexa sse o seu epigrafo é convexo, e é linearmente homogênea sse o seu epigrafo é um cone - proposição 1.2.4.I). Logo, a função indicadora de um conjunto convexo é linearmente homogênea sse o seu epigrafo for um cone (já que é convexo - proposição 1.2.3.I), isto é sse :

$$\{(x, r) \mid \delta(x|C) \leq r\} = \lambda \{(x, r) \mid \delta(x|C) \leq r\}, \quad \forall \lambda > 0 \\ \Leftrightarrow \{(x, r) \mid \delta(x|C) \leq r\} = \{(\lambda x, \lambda r) \mid \delta(x|C) \leq r\} = \\ = \{(x, r) \mid \delta(\frac{x}{\lambda}|C) \leq \frac{r}{\lambda}\} = \{(x, r) \mid \lambda \delta(x|\lambda C) \leq r\} = \\ = \{(x, r) \mid \delta(x|\lambda C) \leq r\}, \quad \forall \lambda > 0 \\ \Leftrightarrow \delta(x|C) = \delta(x|\lambda C), \quad \forall \lambda > 0$$

$$\Leftrightarrow C = \lambda C, \quad \forall \lambda > 0$$

ou seja,  $\delta(x|C)$  é linearmente homogênea sse  $C$  é um cone.  $\parallel$

——— Dem. da proposição 1.2.5.I: Vamos demonstrar apenas que  $g(x)$  é côncava sse  $-g(x)$  é convexa. O restante demonstra-se com o mesmo argumento. Seja então  $g(x)$  uma função côncava. Logo  $\text{hip } g$  é um conjunto convexo. Mas  $\text{hip } g(x) = \{(x, r) | r \leq g(x)\} = \{(x, r) | -r \geq -g(x)\} = -\{(-x, -r) | -r \geq -g(x)\} = -\{(x, r) | r \geq -g(-x)\} = -\text{epi}(-g(-x))$ , pelo que  $-\text{epi}(-g(-x))$  é um conjunto convexo. Pela proposição 1.1.1.III,  $\text{epi}(-g(-x))$  é também um conjunto convexo. Atendendo a que  $\text{epi}(-g(x)) = A.\text{epi}(-g(-x))$ , com  $A = [-I_n; 1]$  ( $I_n$  = matriz identidade de ordem  $n$ ), a proposição 1.1.1.IV permite concluir que  $\text{epi}(-g(x))$  é convexo, isto é, que  $-g(x)$  é uma função convexa. Um raciocínio inverso permite concluir que se  $-g(x)$  é convexa,  $g(x)$  é côncava.  $\parallel$

——— Dem. da proposição 1.2.5.II: Decorre directamente das proposições 1.2.2.I e 1.2.5.I.  $\parallel$

——— Dem. do corolário 1.2.5.II.1: Decorre directamente do corolário 1.2.2.I.1 e da proposição 1.2.5.I.  $\parallel$

——— Dem. da proposição 1.2.5.III: Imediata pelas definições de funções convexa e côncava própria e pela proposição 1.2.5.I.  $\parallel$

——— Dem. da proposição 1.2.5.IV: Decorre directamente das proposições 1.2.2.II e 1.2.5.III.  $\parallel$

——— Dem. da proposição 1.2.5.V: Decorre directamente das proposições 1.2.2.III e 1.2.5.I. ||

——— Dem. da proposição 1.2.5.VI: Decorre directamente das proposições 1.2.2.IV e 1.2.5.I. ||

——— Dem. da proposição 1.2.5.VII: Decorre directamente das proposições 1.2.2.V e 1.2.5.I . ||

——— Dem. da proposição 1.2.5.VIII: Imediata pelas definições de domínio efectivo de funções convexas e côncavas e pela proposição 1.2.5.I. . ||

——— Dem. da proposição 1.2.5.IX: Decorre directamente das proposições 1.2.2.VI, 1.2.5.I e 1.2.5.VIII. . ||

——— Dem. da proposição 1.2.5.X: Similar à demonstração da proposição 1.2.4.I, bastando substituir  $\text{epi } f(x) = \{(x,r) | r \geq f(x)\}$  por  $\text{hip } g(x) = \{(x,r) | r \leq g(x)\}$  . ||

——— Dem. da proposição 1.2.5.XI: Decorre directamente das proposições 1.2.4.II e 1.2.5.I. ||

——— Dem. da proposição 1.2.5.XII: Evidente a partir da definição de função côncava e da proposição 1.1.1.III. ||

——— Dem. da proposição 1.2.5.XIII: Similar à demonstração da proposição 1.2.4.IV, bastando substituir  $\text{epi } f(x) = \{(x,r) | r \geq f(x)\}$  por  $\text{hip } g(x) = \{(x,r) | r \leq g(x)\}$ . ||

——— Dem. da proposição 1.2.5.XIV: Idem relativamente à proposição 1.2.4.V. ||

——— Dem. da proposição 1.2.6.I: Seja a equação de definição do hiperplano em  $\mathbb{R}^{n+1}$ :

$$\langle (x,r), (b,b_0) \rangle = \alpha, \text{ com } b_0 \neq 0$$

$$\Leftrightarrow rb_0 = -\langle x,b \rangle + \alpha$$

Como por hipótese  $b_0 \neq 0$  vem equivalentemente:

$$r = \langle x, \frac{-b}{b_0} \rangle + \frac{\alpha}{b_0}$$

que é uma função linear não vertical em  $\mathbb{R}^n$ . ||

——— Dem. da proposição 1.2.6.II: Comece-se por notar que o subconjunto de  $\mathbb{R}^{n+1}$  dado por  $\{(x,l(x)) | l(x) = \langle x,a \rangle - \beta\}$  é um hiperplano não vertical em  $\mathbb{R}^{n+1}$ . Com efeito,  $\{(x,l(x)) | l(x) = \langle x,a \rangle - \beta\} = \{(x,r) | \langle (x,r), (a,-1) \rangle = \beta\}$ . Assim sendo, o epigrafo e o hipografo de  $l(x)$  são respectivamente os semiespaços fechados superior e inferior gerados por este hiperplano. Como pela proposição 1.1.3.II os semiespaços são conjuntos convexos,  $l(x)$  é simultaneamente uma função convexa e côncava. Repare-se que se  $l(x) = \pm\infty$  esta função não definirá nenhum hiperplano em  $\mathbb{R}^{n+1}$ , pois não existe nenhum par  $(a,\beta) \in \mathbb{R}^{n+1}$  tal que  $\langle x,a \rangle - \beta = \pm\infty$ . Apesar de tudo, mesmo nesta situação,  $l(x) = +\infty$  ( $l(x) = -\infty$ ) é simultaneamente convexa e côncava (resp. o seu hipografo) pois o seu epigrafo/é um conjunto vazio (é  $\mathbb{R}^{n+1}$ ) e o seu hipografo

é  $\mathbb{R}^{n+1}$  (é o conjunto vazio). ||

— Dem. da proposição 1.2.7.I: Vamos mostrar que : 1) (i)  $\Leftrightarrow$  (iii);  
2) (iii)  $\Rightarrow$  (ii); 3) (ii)  $\Rightarrow$  (i).

1) Sendo  $Z(x)$  s.c.i., e pela definição, teremos  $Z(x) \leq \lim_{x^i \rightarrow x} Z(x^i)$  qualquer que seja a sucessão  $\{x^i\}$ . Admita-se uma qualquer sucessão  $\{r^i\}$  tal que  $\lim r^i = r$  e  $r^i \geq Z(x^i)$  para todo o  $i$ . A condição  $Z(x) \leq \lim_{x^i \rightarrow x} Z(x^i)$  pode ser reexpressa equivalentemente como

$Z(x) \leq r$ ,  $\forall$  as sucessões  $\{x^i\} \rightarrow x$  e  $\{r^i\} \rightarrow r$  tais que  $r^i \geq Z(x^i)$  para todo o  $i$ .

Isto é: qualquer sucessão  $(x^i, r^i)$  de elementos do epigrafo de  $Z(x)$  tem como limite um elemento do epigrafo de  $Z$ :  $(x, r)$  (pois  $Z(x) \leq r$ ). Pela proposição 1.1.5.II, o epigrafo de  $Z$  é um conjunto fechado. Mostramos que (i)  $\Leftrightarrow$  (iii).

2) Seja (iii). Se o epigrafo de  $Z$  é fechado, então pela proposição 1.1.5.II, qualquer sucessão  $\{(x^i, r^i)\}$  de seus elementos ainda é um elemento seu. Faça-se  $r^i = \alpha$  para todo o  $i$ . Fica então que:

$Z(x) \leq \alpha$ ,  $\forall$  a sucessão  $\{x^i\} \rightarrow x$  tal que  $Z(x^i) \leq \alpha$  para todo o  $i$ .

Isto é: qualquer sucessão de vectores  $x^i \in S_Z^{\leq}(\alpha)$  tem como limite um elemento de  $S_Z^{\leq}(\alpha)$ . Pela proposição 1.1.5.II,  $S_Z^{\leq}(\alpha)$  é um conjunto fechado qualquer que tenha sido o  $\alpha$  considerado. Provamos que (iii)  $\Rightarrow$  (ii)

3) Admita-se (ii), ou seja, admita-se que  $S_Z^{\leq}(\alpha)$  é um conjunto fechado qualquer que seja  $\alpha \in [-\infty, +\infty]$ . Considerem-se quaisquer sucessões  $\{x^i\} \rightarrow x$  e  $\{r^i\} \rightarrow r = \lim_{x^i \rightarrow x} Z(x^i)$ . Sendo  $r = \lim_{x^i \rightarrow x} Z(x^i)$  haverá um  $i$  a partir do qual (pela definição de limite):

$$r - \varepsilon < Z(x^i) < r + \varepsilon, \forall \varepsilon > 0.$$

Tal implica que  $\forall \varepsilon > 0 \exists i_0 \in \mathbb{N} : i > i_0 \Rightarrow x^i \in \{y | Z(y) < r + \varepsilon\}$

Passando ao limite quando  $\epsilon \rightarrow 0^+$  teremos que:

$$x = \lim x^i \in \text{cl} \{y | Z(y) \leq r\} \Leftrightarrow x \in \text{cl} S_Z^{\leq}(r)$$

Como  $S_Z^{\leq}(r)$  é fechado (porque (ii) se verifica por hipótese) teremos:

$$x \in S_Z^{\leq}(r) \Leftrightarrow Z(x) \leq r = \lim Z(x^i),$$

ou seja,  $Z$  é s.c.i. Está também demonstrado que (ii)  $\Rightarrow$  (i). ||

——— Dem. do corolário 1.2.7.I.1: Faça-se  $W(x) = Z(x) + \delta(x | S_Z^{\leq}(\bar{\alpha}))$ .

$$\begin{aligned} \text{Temos que } S_W^{\leq}(\alpha) &= \{x | W(x) \leq \alpha\} = \{x | Z(x) + \delta(x | S_Z^{\leq}(\bar{\alpha})) \leq \alpha\} = \\ &= \begin{cases} S_Z^{\leq}(\alpha) & \text{se } \alpha < \bar{\alpha} \\ S_Z^{\leq}(\bar{\alpha}) & \text{se } \alpha \geq \bar{\alpha} \end{cases} \end{aligned}$$

Como  $S_Z^{\leq}(\alpha)$  para  $\alpha \leq \bar{\alpha}$  são conjuntos fechados por hipótese, a função  $W(x)$  é s.c.i. Consequentemente, é também s.c.i. relativamente a  $S_Z^{\leq}(\bar{\alpha})$ . Mas  $W(x) = Z(x)$  para todo o  $x \in S_Z^{\leq}(\bar{\alpha})$ , pelo que  $Z(x)$  é igualmente s.c.i. relativamente a  $S_Z^{\leq}(\bar{\alpha})$ . ||

——— Dem. da proposição 1.2.7.II: Elementar a partir da definição. ||

——— Dem. da proposição 1.2.7.III: Elementar a partir da definição. ||

——— Dem. da proposição 1.2.7.IV: Se  $Z(x) = \text{Sup} \{Z_i(x) | i \in I\}$  temos que

$$\text{epi } Z(x) = \bigcap_{i \in I} \text{epi } Z_i(x)$$

Ora, se as funções  $Z_i(x)$  são s.c.i., os respectivos epígrafos são conjuntos fechados (proposição 1.2.7.I). Como a intersecção de conjunto fechados é um conjunto fechado (prop. 1.1.5.III),  $\text{epi } Z(x)$  é também um conjunto fechado. Pela proposição 1.2.7.I, a função  $Z(x)$  é s.c.i. . ||

———— Dem. da proposição 1.2.7.V: Elementar pelas definições. ||

———— Dem. da proposição 1.2.7.VI: Decorre directamente das proposições 1.2.7.I e 1.2.7.V. ||

———— Dem. do corolário 1.2.7.VI.1: Decorre directamente do corolário 1.2.7.I.1 e da proposição 1.2.7.V. ||

———— Dem. da proposição 1.2.7.VII: Elementar a partir da definição. ||

———— Dem. da proposição 1.2.7.VIII: Elementar a partir da definição . ||

———— Dem da proposição 1.2.7.IX: Similar à demonstração da proposição 1.2.7.IV, bastando substituir  $\text{epi } Z(x)$  por  $\text{hip } Z(x)$ . ||

———— Dem. da proposição 1.2.7.X: Imediata pelas definições e pelas proposições 1.2.5.I e 1.2.7.V. ||

———— Dem. da proposição 1.2.8.I: Decorre directamente das definições. ||

————— Dem. do corolário 1.2.8.1.1: Imediato pela proposição 1.2.8.I. ||

————— Dem. do corolário 1.2.8.1.2: Imediato pela proposição 1.2.8.I e pela definição de limite de uma função. ||

————— Dem. da proposição 1.2.8.II: Consideremos todas as possíveis sucessões  $\{x^i\}$  convergentes e de elementos de  $D$ , tais que  $Z(x^{i+1}) \geq Z(x^i)$ ,  $\forall i \in \mathbb{N}$ . Como  $D$  é não vazio, existe pelo menos uma destas sucessões (no caso de  $D$  possuir um único elemento, ela será a sucessão constante). Como  $D$  é fechado, então todas estas sucessões convergentes terão como limites elementos de  $D$  (proposição 1.1.5.II). Representemo-los por  $\bar{x}$ . Porque a função  $Z(x)$  é contínua relativamente a  $D$ , então  $Z(\bar{x}) = \lim_{x^i \rightarrow \bar{x}} Z(x^i)$  (proposição 1.2.8.I). Teremos então, e por construção, que o máximo de  $Z$  em  $D$  será dado pelo  $\bar{x}$  associado ao maior  $Z(\bar{x})$  conseguido. Com efeito, e porque  $D$  é limitado, não há a possibilidade de existirem sucessões divergentes de elementos de  $D$  tais que  $Z(x^{i+1}) \geq Z(x^i)$ . Através de um argumento similar, e partindo de todas as sucessões de elementos de  $D$  tais que  $Z(x^{i+1}) \leq Z(x^i)$  mostra-se que existe mínimo finito de  $Z$  em  $D$ . ||

Dem. da proposição 1.2.8.III: Teremos que mostrar: 1) para funções convexas impróprias e 2) para funções convexas próprias.

1) Como sabemos, uma função convexa é imprópria se é a função constante  $+\infty$  ou se para algum  $x \in \mathbb{R}^n$  temos  $f(x) = -\infty$ . Se  $f(x)$  é a função constante  $+\infty$  o seu domínio efectivo é vazio, pelo que a proposição é trivial neste caso. Se  $f(x)$  é imprópria porque  $f(x) = -\infty$  para pelo menos algum  $x \in \mathbb{R}^n$ , consideremos que o interior relativo do seu domínio relativo é não vazio, pois caso contrário a proposição é trivial. Seja então  $x^1 \in \text{rint dom } f$ , tal que  $f(x^1) = -\infty$ . Mostremos que  $f(x) = -\infty$  para todo o  $x \in \text{rint dom } f$ . Suponhamos que existe um  $x^2 \in \text{dom } f$  tal que  $f(x^2) > -\infty$ . Por definição, dado que  $f$  é convexa,  $((1-\lambda)x^1 + \lambda x^2; (1-\lambda)f(x^1) + \lambda f(x^2)) \in \text{epi } f, \forall 0 \leq \lambda \leq 1$   
 $\Leftrightarrow f[(1-\lambda)x^1 + \lambda x^2] \leq (1-\lambda)f(x^1) + \lambda f(x^2), \forall 0 \leq \lambda \leq 1$ . Porque  $f(x^1) = -\infty$  vem que  $f[(1-\lambda)x^1 + \lambda x^2] = -\infty, \forall 0 \leq \lambda < 1$ . Como o anterior é válido qualquer que seja  $x^2 \in \text{dom } f$  tal que  $f(x^2) > -\infty$ , conclui-se que se  $f(x^1) = -\infty$  para algum  $x^1 \in \text{rint dom } f$ , então  $f(x) = -\infty, \forall x \in \text{rint dom } f$ . Ora, pela proposição 1.2.8.I, é imediato que uma função constante num determinado conjunto é contínua relativamente a ele (e relativamente a qualquer seu subconjunto).

2) Mostremos agora que a proposição é válida se  $f(x)$  é própria e se tem  $\text{rint dom } f \neq \emptyset$ . Seja um qualquer  $\bar{x} \in C$  com  $C$  relativamente aberto ( $C = \text{rint } C$ ) e contido no  $\text{rint dom } f$ . Mostrar que  $f$  é contínua em  $\bar{x}$  relativamente a  $C$  é mostrar (pelo corolário 1.2.8.I.2) que  $\forall \varepsilon > 0 \exists \delta > 0$ :  
 $(\|x - \bar{x}\| \leq \delta \wedge x \in C) \Rightarrow |f(x) - f(\bar{x})| \leq \varepsilon$ .

Seja  $V(C)$  a envolvente afim de  $C$ . Como  $C$  é relativamente aberto e  $\bar{x} \in C$ , então existe um  $\delta' > 0$  tal que  $(\|x - \bar{x}\| \leq \delta' \wedge x \in V(C)) \Rightarrow x \in C$ .

Construa-se

$$\theta = \max_{1 \leq i \leq n} \{ \max[f(\bar{x} + \delta' v^i) - f(\bar{x}); f(\bar{x} - \delta' v^i) - f(\bar{x})] \} \quad \dots (A)$$

em que  $\{v^1, \dots, v^m\}$  é uma base ortonormada do subespaço  $S(C)$  paralelo à variedade linear  $V(C)$ . Seja:

$$\delta = \min \left\{ \frac{\delta'}{m} ; \frac{\epsilon \delta'}{m\theta} \right\} \quad \dots (B)$$

Note-se que  $\delta \leq \delta'$  (pois  $m \geq 1$ ). Escolha-se um  $x \in V(C)$  com  $\|x - \bar{x}\| \leq \delta$ .

Como  $(x - \bar{x}) \in S(C)$ , podemos escrever  $(x - \bar{x})$  como combinação linear dos vectores  $v^1, \dots, v^m$ :  $(x - \bar{x}) = \sum_{i=1}^m \alpha_i \cdot v^i$  (pois  $\{v^1, \dots, v^m\}$  é uma base de  $S(C)$ ).

Faça-se  $\beta_i = \alpha_i / \delta'$ , donde vem:

$$x - \bar{x} = \delta' \sum_{i=1}^m \beta_i \cdot v^i \quad \dots (C)$$

$$\text{Logo: } \|x - \bar{x}\| = \langle x - \bar{x}, x - \bar{x} \rangle^{1/2} = \delta' \left\langle \sum_{i=1}^m \beta_i \cdot v^i ; \sum_{i=1}^m \beta_i \cdot v^i \right\rangle^{1/2} = \delta' \left( \sum_{i=1}^m \beta_i^2 \right)^{1/2}$$

... (D)

(já que  $\|v^i\| = 1$  e  $\langle v^i, v^j \rangle = 0$  para  $i \neq j$ ). Como escolhemos  $x$  tal que  $\|x - \bar{x}\| \leq \delta$  vem, por (B), que:

$$\|x - \bar{x}\| \leq \frac{\delta'}{m} \Leftrightarrow (\text{por (D)}) \left( \sum_{i=1}^m \beta_i^2 \right)^{1/2} \leq \frac{1}{m} \Rightarrow \beta_i \leq \frac{1}{m}$$

Note-se também que porque  $\|x - \bar{x}\| \leq \delta$  e por (D) obtemos:

$$\delta' \left( \sum_{i=1}^m \beta_i^2 \right)^{1/2} \leq \delta \Rightarrow (\text{por (B)}) \delta' \left( \sum_{i=1}^m \beta_i^2 \right)^{1/2} \leq \frac{\epsilon \delta'}{m\theta}$$

$$\Leftrightarrow \left( \sum_{i=1}^m \beta_i^2 \right)^{1/2} \leq \frac{\epsilon}{m\theta}$$

$$\Rightarrow \beta_i \leq \frac{\epsilon}{m\theta}$$

... (E)

Dada a convexidade de  $f$  temos que:

$$\begin{aligned}
 f(x) &= f(\bar{x} + (x - \bar{x})) = (\text{por (C)}) f(\bar{x} + \delta' \cdot \sum_{i=1}^m \beta_i \cdot v^i) = \\
 &= f\left[\frac{1}{m} \sum_{i=1}^m (\bar{x} + m \delta' \beta_i \cdot v^i)\right] \\
 &\leq \frac{1}{m} \sum_{i=1}^m f(\bar{x} + m \delta' \beta_i \cdot v^i) \quad (\text{pelo corol. 1.2.2.I.1}) \\
 &= \frac{1}{m} \sum_{i=1}^m f[(1 - m \beta_i) \bar{x} + m \beta_i (\bar{x} + \delta' v^i)] \\
 &\leq \frac{1}{m} \sum_{i=1}^m [(1 - m \beta_i) f(\bar{x}) + m \beta_i f(\bar{x} + \delta' v^i)] \quad (\text{pelo corol. 1.2.2.I.1}) \\
 &= f(\bar{x}) + \sum_{i=1}^m \beta_i \cdot [f(\bar{x} + \delta' v^i) - f(\bar{x})]
 \end{aligned}$$

Atendendo a (A), é obvio que  $f(\bar{x} + \delta' v^i) - f(\bar{x}) \leq \theta$  pelo que

$$f(x) - f(\bar{x}) \leq \theta \cdot \sum_{i=1}^m \beta_i$$

De (E) obtemos:

$$f(x) - f(\bar{x}) \leq \theta \cdot m \cdot \frac{\varepsilon}{m \theta} \Leftrightarrow f(x) - f(\bar{x}) \leq \varepsilon$$

Em síntese, mostramos que  $(\|x - \bar{x}\| \leq \delta \wedge x \in V(C)) \Rightarrow f(x) - f(\bar{x}) \leq \varepsilon$ .

... (F)

Mostremos agora que também  $-(f(x) - f(\bar{x})) = f(\bar{x}) - f(x) \leq \varepsilon$ . Seja  $y = 2\bar{x} - x$  ( $y \in V(C)$  porque  $x, \bar{x} \in C$ ), e note-se que:

$$\|y - \bar{x}\| = \|2\bar{x} - x - \bar{x}\| = \|x - \bar{x}\|$$

pelo que (já que por hipótese  $\|x - \bar{x}\| \leq \delta$ ):  $\|y - \bar{x}\| \leq \delta$ .

Consequentemente, por (F) temos:

$$f(y) - f(\bar{x}) \leq \varepsilon \quad \dots (G)$$

Mas como  $\bar{x} = \frac{1}{2}x + \frac{1}{2}y$  e porque  $f$  é convexa:

$$f(\bar{x}) \leq \frac{1}{2}f(y) + \frac{1}{2}f(x) \Leftrightarrow f(y) \geq 2f(\bar{x}) - f(x)$$

Substituindo em (G) fica:

$$(\|x - \bar{x}\| \leq \delta \wedge x \in V(C)) \Rightarrow f(\bar{x}) - f(x) \leq \varepsilon \quad \dots (H)$$

Reunindo (F) e (H) teremos:

$$(\|x - \bar{x}\| \leq \delta \wedge x \in V(C)) \Rightarrow |f(x) - f(\bar{x})| \leq \varepsilon$$

Ora o anterior implica, como queríamos mostrar:

$$(\|x - \bar{x}\| \leq \delta \wedge x \in C) \Rightarrow |f(x) - f(\bar{x})| \leq \varepsilon \quad . \quad \parallel$$

——— Dem. do corolário 1.2.8.III.1: Decorre directamente da proposição 1.2.8.III, corolário 1.2.8.I.1 e proposição 1.2.5.I.  $\parallel$

——— Dem. do corolário 1.2.8.III.2: Pela proposição 1.2.6.II, uma função linear não vertical é convexa (e côncava). É obvio que tem domínio efectivo  $\mathbb{R}^n$ . Logo, como  $\mathbb{R}^n$  é relativamente aberto, uma função linear não vertical é contínua em  $\mathbb{R}^n$  (pela proposição 1.2.8.III).  $\parallel$

——— Dem. da proposição 1.2.8.IV: Seja  $x \in L$ . Porque  $L$  é local

mente simplicial existem simplex's  $S_1, \dots, S_m \subset L$  tais que para algum  $\varepsilon > 0$  teremos  $E_\varepsilon(x) \cap (S_1 \cup \dots \cup S_m) = E_\varepsilon(x) \cap L \dots (A)$

Cada um destes simplex's  $S_i$  que contêm  $x$  pode ser "triangularizado" num número finito de outros simplex's  $T_{i,1}, \dots, T_{i,k}$  tendo  $x$  por vértice e tais que  $\bigcup_{j=1}^k T_{i,j} \supset S_i$ . Para tal basta considerar, por exemplo, cada  $T_{i,j}$  como o simplex que tem como vértices o vector  $x$  e  $(m-1)$  vértices de  $S_i$ .

Então, por construção e atendendo a (A), temos:

$$E_\varepsilon(x) \cap L = E_\varepsilon(x) \cap \left( \bigcup_{i,j} T_{i,j} \right)$$

Se pudermos demonstrar que  $f$  é s.c.s. relativamente a cada  $T_{i,j}$ , é imediato que também é s.c.s. relativamente a  $\bigcup_{i,j} T_{i,j}$  e, portanto, a  $S \subset \bigcup_{i,j} T_{i,j}$ . O argumento está assim reduzido a mostrar que se  $T_{i,j}$  é um simplex contido no dom  $f$  e tem  $x$  por um dos vértices, então é s.c.s. em  $x$  relativamente a  $T_{i,j}$ .

Sejam  $\{v^1, \dots, v^p\}$  os vértices de  $T_{i,j}$  para além de  $x$ . Por definição de simplex, cada vector  $y$  de  $T_{i,j}$  pode ser expresso como combinação linear dos vértices de  $T_{i,j}$ , isto é, para dados  $0 \leq \lambda_1, \dots, \lambda_p \leq 1$ :

$$y = (1 - \lambda_1 - \lambda_2 - \dots - \lambda_p)x + \lambda_1 v^1 + \dots + \lambda_p v^p$$

Como  $f$  é uma função convexa vem pelo corolário 1.2.2.I.1 que (1):

$$f(y) \leq (1 - \lambda_1 - \lambda_2 - \dots - \lambda_p)f(x) + \lambda_1 f(v^1) + \dots + \lambda_p f(v^p)$$

---

(1) - Note-se que neste caso o corolário 1.2.2.I.1 é válido mesmo quando  $f = -\infty$  para algum vector de  $T_{i,j}$ . Com efeito, e já que por construção  $T_{i,j} \subset \text{dom } f$ , então  $f(y) < +\infty, \forall y \in T_{i,j}$ . Está por isso afastada qualquer hipótese de ambiguidades  $\infty - \infty$  que constituíam a única razão da restrição  $f > -\infty$  na proposição 1.2.2.I e seu corolário.

Façamos  $y \rightarrow x$ , o que significa fazer  $\lambda_1, \dots, \lambda_m \rightarrow 0$ . Fica:

$$\lim_{y \rightarrow x} f(y) \leq f(x)$$

ou seja,  $f$  é s.c.s. em  $x$  relativamente a  $T_{ij}$ . ||

————— Dem. do corolário 1.2.8.IV.1: Decorre directamente da proposição 1.2.8.IV, corolário 1.2.8.I.1 e proposição 1.2.5.I. ||

————— Dem. da proposição 1.2.8.V: Seja  $\{e_1, \dots, e_n\}$  o conjunto dos vectores-coluna da matriz identidade de ordem  $n$ . Para cada  $x \in \mathbb{R}_{0+}^n$  e para cada  $\varepsilon > 0$  real, formemos o simplex

$$S = \text{conv} \{0; x + \varepsilon e_1; \dots; x + \varepsilon e_n\}$$

Teremos necessariamente

$$E_\varepsilon(x) \cap \mathbb{R}_{0+}^n = E_\varepsilon(x) \cap S, \quad \forall x \in \mathbb{R}_{0+}^n, \varepsilon > 0$$

$\mathbb{R}_{0+}^n$  é, portanto, um conjunto localmente simplicial. ||

————— Dem. da proposição 1.2.9.I: Seja  $Z$  diferenciável no ponto  $x$ . Por definição teremos, fazendo  $y = x + \lambda d$  e com  $d \neq 0$ :

$$\lim_{\|\lambda d\| \rightarrow 0^+} \frac{Z(x + \lambda d) - Z(x) - \langle \nabla Z(x), \lambda d \rangle}{\|\lambda d\|} = 0$$

$$\frac{1}{\|d\|} \lim_{\lambda \rightarrow 0^+} \frac{Z(x + \lambda d) - Z(x) - \lambda \langle \nabla Z(x), d \rangle}{\lambda} = 0$$

$$\frac{1}{\|d\|} \cdot [Z'(x,d) - \langle \nabla Z(x), d \rangle] = 0$$

$$Z'(x,d) = \langle \nabla Z(x), d \rangle. \quad \parallel$$

——— Dem. do corolário 1.2.9.I.1: Pela proposição 1.2.9.I temos que:

$$Z'(x,d) = \langle \nabla Z(x), d \rangle$$

Pela definição de derivada direccional vem por outro lado que :

$$Z'(x,d) = \lim_{\lambda \rightarrow 0^+} \frac{Z(x+\lambda d) - Z(x)}{\lambda}$$

Logo, fazendo  $d = e_j$  ( $j=1, \dots, n$ ) obtemos:

$$\nabla_j Z(x) = \langle \nabla Z(x), e_j \rangle = \lim_{\lambda \rightarrow 0^+} \frac{Z(x+\lambda e_j) - Z(x)}{\lambda} \quad \parallel$$

——— Dem. da proposição 1.2.9.II: Para  $y=x$  a igualdade é trivialmente verificada. Para  $y \neq x$ , e porque  $f$  é convexa, vem pela proposição 1.2.2.I:

$$f[(1-\lambda)x + \lambda y] \leq (1-\lambda)f(x) + \lambda f(y), \quad \forall y \in \mathbb{R}^n, 0 \leq \lambda \leq 1$$

$$\frac{f[x + \lambda(y-x)] - f(x)}{\lambda} \leq f(y) - f(x), \quad \forall y \in \mathbb{R}^n, 0 \leq \lambda \leq 1$$

Passando ao limite quando  $\lambda \rightarrow 0^+$  fica:  $f'(x, y-x) \leq f(y) - f(x), \forall y \in \mathbb{R}^n$

Ora, pela proposição 1.2.9.I, temos  $f'(x, y-x) = \langle \nabla f(x), y-x \rangle$ , pelo que  $\langle \nabla f(x), y-x \rangle \leq f(y) - f(x), \forall y \in \mathbb{R}^n \Leftrightarrow f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle, \forall y \in \mathbb{R}^n. \parallel$

——— Dem. do corolário 1.2.9.II.1: Decorre directamente das proposições 1.2.9.II e 1.2.5.I. ||

——— Dem. da proposição 1.2.9.III: Atendendo à proposição 1.2.9.II, temos que :  $f$  convexa em  $C \Rightarrow f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle$ ,  $\forall x, y \in C$ .

Resta demonstrar a implicação inversa. Façamos  $x = (1-\mu)x^1 + \mu x^2$  com  $0 \leq \mu \leq 1$  e  $x^1, x^2 \in C$ . Consideremos alternadamente  $y=x^1$  e  $y=x^2$ , donde fica:

$$(i) \text{ (para } y=x^1) \quad f(x^1) \geq f(x) + \langle \nabla f(x), x^1-x \rangle$$

$$(ii) \text{ (para } y=x^2) \quad f(x^2) \geq f(x) + \langle \nabla f(x), x^2-x \rangle$$

Multiplicando (i) por  $(1-\mu)$  e (ii) por  $\mu$  fica, depois de somarmos ordenadamente:  $(1-\mu)f(x^1) + \mu f(x^2) \geq f(x) + \langle \nabla f(x), (1-\mu)x^1 + \mu x^2 - x \rangle$ .

Como  $x = (1-\mu)x^1 + \mu x^2$  obtemos:

$$(1-\mu)f(x^1) + \mu f(x^2) \geq f[(1-\mu)x^1 + \mu x^2]$$

Pela proposição 1.2.2.I podemos concluir que  $f$  é convexa. ||

——— Dem. do corolário 1.2.9.III.1: Decorre directamente das proposições 1.2.9.III e 1.2.5.I. ||

——— Dem. do corolário 1.2.9.III.2: Vamos demonstrar para o caso de  $Z(x)$  ser uma função convexa. Para funções côncavas a demonstração é similar, atendendo à simetria estabelecida pela proposição 1.2.5.I. Seja então uma função  $Z(x)$  diferenciável e convexa num conjunto  $C$ . Pela proposição 1.2.9.III vem que  $Z(y) \geq Z(x) + \langle \nabla Z(x), y-x \rangle$ ,  $\forall x, y \in C$ .

Ora,  $\bar{x}$  é mínimo global no conjunto  $C$  sse

$$Z(y) \geq Z(\bar{x}), \forall y \in C \Leftrightarrow \langle \nabla Z(\bar{x}), y-\bar{x} \rangle \geq 0, \forall y \in C.$$

Na medida em que  $C$  é um conjunto aberto,  $y-\bar{x}$  pode assumir qualquer sinal, pelo que o anterior é equivalente a  $\nabla Z(\bar{x}) = 0$ . ||

——— Dem. da proposição 1.2.9.IV: Pela fórmula de desenvolvimento em série de Taylor de uma função temos, dados  $x, y \in C$ :

$$f(y) = f(x) + \langle \nabla f(x); y-x \rangle + \frac{1}{2} \langle y-x; H[x + \theta(y-x)] (y-x) \rangle$$

para algum  $0 \leq \theta \leq 1$ .

1) Demonstremos em primeiro lugar que se  $H$  é semidefinida positiva para todo o  $x \in C$ , então  $f$  é convexa em  $C$ .

Comece-se por reparar que se  $x, y \in C$ , então também  $x + \theta(y-x) = (1-\theta)x + \theta y \in C$ , pois  $C$  é convexo. Se  $H$  é semidefinida positiva para todo o  $x \in C$  temos que:

$$\frac{1}{2} \langle y-x, H[x + \theta(y-x)] (y-x) \rangle \geq 0, \forall x, y \in C$$

Pelo desenvolvimento em série de Taylor tal equivale a

$$f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle, \forall x, y \in C$$

Pela proposição 1.2.9.III a função é convexa em  $C$ .

2) Demonstremos agora que se  $f$  for convexa e bidiferenciável em  $C$ , então  $H$  é semidefinida positiva em  $C$ . Vamos demonstrar por absurdo. Suponhamos que existe algum  $x + \theta(y-x) \in C$  tal que

$$\langle y-x, H[x + \theta(y-x)] (y-x) \rangle < 0$$

ou seja, que  $H$  não é semidefinida positiva para algum ponto de  $C$ . Para esse ponto viria, a partir do desenvolvimento em série de Taylor de 2ª ordem:

$$f(y) < f(x) + \langle \nabla f(x), y-x \rangle$$

o que contraria a proposição 1.2.9.III, ou seja, a função não poderia ser convexa se tal acontecesse. ||

Na medida em que  $C$  é um conjunto aberto,  $y-\bar{x}$  pode assumir qualquer sinal, pelo que o anterior é equivalente a  $\nabla Z(\bar{x}) = 0$ . ||

——— Dem. da proposição 1.2.9.IV: Pela fórmula de desenvolvimento em série de Taylor de uma função temos, dados  $x, y \in C$ :

$$f(y) = f(x) + \langle \nabla f(x); y-x \rangle + \frac{1}{2} \langle y-x; H[x + \theta(y-x)] (y-x) \rangle$$

para algum  $0 \leq \theta \leq 1$ .

1) Demonstremos em primeiro lugar que se  $H$  é semidefinida positiva para todo o  $x \in C$ , então  $f$  é convexa em  $C$ .

Comece-se por reparar que se  $x, y \in C$ , então também  $x + \theta(y-x) = (1-\theta)x + \theta y \in C$ , pois  $C$  é convexa. Se  $H$  é semidefinida positiva para todo o  $x \in C$  temos que:

$$\frac{1}{2} \langle y-x, H[x + \theta(y-x)] (y-x) \rangle \geq 0, \forall x, y \in C$$

Pelo desenvolvimento em série de Taylor tal equivale a

$$f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle, \forall x, y \in C$$

Pela proposição 1.2.9.III a função é convexa em  $C$ .

2) Demonstremos agora que se  $f$  for convexa e bidiferenciável em  $C$ , então  $H$  é semidefinida positiva em  $C$ . Vamos demonstrar por absurdo. Suponhamos que existe algum  $x + \theta(y-x) \in C$  tal que

$$\langle y-x, H[x + \theta(y-x)] (y-x) \rangle < 0$$

ou seja, que  $H$  não é semidefinida positiva para algum ponto de  $C$ . Para esse ponto viria, a partir do desenvolvimento em série de Taylor de 2ª ordem:

$$f(y) < f(x) + \langle \nabla f(x), y-x \rangle$$

o que contraria a proposição 1.2.9.III, ou seja, a função não poderia ser convexa se tal acontecesse. ||

————— Dem. do corolário 1.2.9.IV.1: Decorre directamente das proposições 1.2.9.IV e 1.2.5.I. ||

————— Dem. da proposição 1.2.10.I: Imediato pelas definições e pela proposição 1.2.5.I. ||

————— Dem. da proposição 1.2.10.II: Vamos demonstrar apenas para funções convexas. Para funções côncavas é análogo, atendendo à simetria estabelecida pela proposição 1.2.5.I. Comece-se por reparar que:

$$\{(y,r) | r=f(x) + \langle s, y-x \rangle\} = \{(y,r) | \langle (y,r), (s,1) \rangle = f(x) - \langle s, x \rangle\}$$

pelo que o anterior conjunto define um hiperplano não vertical em  $\mathbb{R}^{n+1}$ . Para  $y=x$  vem  $r=f(x)$ , pelo que o hiperplano passa no ponto  $(x, f(x))$ . Para além disso, pela definição de subgradiente,  $f(y) \geq f(x) + \langle s, y-x \rangle$ ,  $\forall y \in \mathbb{R}^n$ , donde resulta que  $(y,r) \in \text{epi } f \Leftrightarrow r \geq f(y) \Rightarrow r \geq f(x) + \langle s, y-x \rangle$ .

Consequentemente,  $\text{epi } f \subset \{(y,r) | r \geq f(x) + \langle s, y-x \rangle\}$ , o que significa que  $\text{epi } f$  está contido no semiespaço gerado pelo hiperplano  $r = f(x) + \langle s, y-x \rangle$  que passa no ponto  $(x, f(x))$ . Este hiperplano é, portanto, um hiperplano de suporte a  $\text{epi } f$  no ponto  $(x, f(x))$ .

Resta demonstrar que qualquer hiperplano de suporte ao  $\text{epi } f$  define um subgradiente de  $f$  no ponto em questão. Seja o hiperplano não vertical de  $\mathbb{R}^{n+1}$  dado por :  $H = \{(y,r) | \langle (y,r), (b, b_0) \rangle = \alpha\}$  com  $b_0 \neq 0$ . Dividindo tudo por  $b_0$  fazendo  $s = -b/b_0$  temos

$$H = \{(y,r) | r = \langle s, y \rangle + \frac{\alpha}{b_0}\}$$

H será hiperplano de suporte a epi f no ponto  $(x, f(x))$  sse:

$$(i) \quad (y, r) = (x, f(x)) \Rightarrow r = \langle s, y \rangle + \frac{\alpha}{b_0}$$

$$(ii) \quad H \supset \text{epi } f \Leftrightarrow r \geq \langle s, y \rangle + \frac{\alpha}{b_0}, \quad \forall (y, r) \in \text{epi } f$$

De (i) vem  $\alpha/b_0 = f(x) - \langle s, x \rangle$ . De (ii), e já incluindo este último resultado, fica  $r \geq \langle s, y \rangle + f(x) - \langle s, x \rangle$ ,  $\forall (y, r) \in \text{epi } f$

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

ou seja,  $s$  é um subgradiente de  $f$  no ponto  $x$ .  $\parallel$

——— Dem. do corolário 1.2.10.II.1: Demonstraremos para funções convexas.

Para funções concâvas é similar atendendo à proposição 1.2.5.I. Seja  $f(x)$  uma função convexa, pelo que o seu epigrafo é um conjunto convexo. Pela proposição 1.1.7.I existe pelo menos um hiperplano de suporte a cada ponto da fronteira de  $\text{epi } f(x)$ . Para  $x \in \text{int dom } f$  tais hiperplanos de suporte serão necessariamente não verticais, pelo que a proposição 1.2.10.II permite retirar a conclusão desejada. Os pontos da fronteira de  $\text{dom } f$  são excluídos pois poderiam originar apenas hiperplanos de suporte ao epi  $f$  que fossem verticais (e, conseqüentemente, não correspondendo a subgradientes nesse ponto).  $\parallel$

——— Dem. da proposição 1.2.10.III: Se  $s$  é um subgradiente de  $f$  no ponto  $x$  vem por definição:  $f(y) \geq f(x) + \langle s, y-x \rangle$ ,  $\forall y \in \mathbb{R}^n$ .

Faça-se  $y = x + \lambda d$  ( $\lambda > 0$ ), donde  $f(x + \lambda d) - f(x) \geq \langle s, \lambda d \rangle, \forall d \in \mathbb{R}^n$

$$\Leftrightarrow \frac{f(x + \lambda d) - f(x)}{\lambda} \geq \langle s, d \rangle, \forall d \in \mathbb{R}^n$$

Passando ao limite quando  $\lambda \rightarrow 0^+$ , fica:

$$f'(x, d) \geq \langle s, d \rangle, \forall d \in \mathbb{R}^n. \quad \parallel$$

——— Dem. do corolário 1.2.10.III.1: Decorre directamente das proposições 1.2.10.III e 1.2.5.I.  $\parallel$

——— Dem. da proposição 1.2.10.IV: Demonstraremos unicamente para o caso de  $f(x)$  função convexa. Para funções côncavas é análogo atendendo à proposição 1.2.5.I. Pela proposição 1.2.10.III,  $s$  é um subgradiente de  $f$  no ponto  $x$  sse:  $f'(x, d) \geq \langle s, d \rangle, \forall d \in \mathbb{R}^n$ . Mas pela proposição 1.2.9.I, se a função é diferenciável temos:

$$f'(x, d) = \langle \nabla f(x), d \rangle, \forall d \in \mathbb{R}^n$$

Fica então  $\langle \nabla f(x), d \rangle \geq \langle s, d \rangle, \forall d \in \mathbb{R}^n$

$$\Leftrightarrow \langle \nabla f(x) - s, d \rangle \geq 0, \forall d \in \mathbb{R}^n$$

$$\Leftrightarrow \nabla f(x) = s, \quad \forall d \in \mathbb{R}^n$$

Desta forma, não poderemos ter subgradientes diferentes de  $f$  no ponto  $x$  porque terão que igualar  $\nabla f(x)$ .  $\parallel$

——— Dem. da proposição 1.2.10.V: Vamos demonstrar para o caso de

$f(x)$  função convexa. Para funções côncavas é perfeitamente similar.

$\bar{x}$  é mínimo global de  $f$  em  $C$  sse:

$$f(y) \geq f(\bar{x}), \quad \forall y \in C$$

$$\Leftrightarrow f(y) \geq f(\bar{x}) + \langle 0, y - \bar{x} \rangle, \quad \forall y \in C$$

$$\Leftrightarrow 0 \in \partial f(\bar{x}). \quad \parallel$$

ÍNDICE

Pág.

**PARTE I : FUNDAMENTOS DE ANÁLISE CONVEXA**

INTRODUÇÃO

CAPÍTULO 1

- 1.1. Conjuntos convexos/Fundamentos de topologia de conjuntos convexos ..... 5
  - 1.1.1. Conjuntos convexos e envolvente convexa
  - 1.1.2. Conjuntos afins e subespaços
  - 1.1.3. Hiperplanos e semiespaços
  - 1.1.4. Cones convexos e respectivos cones polares
  - 1.1.5. Fundamentos de topologia de conjuntos convexos em  $\mathbb{R}^n$
  - 1.1.6. Alguns resultados de separação de conjuntos convexos
  - 1.1.7. Conjuntos suporte, conjuntos polares e hiperplanos e semiespaços de suporte
- 1.2. Funções convexas e côncavas ..... 19
  - 1.2.1. Introdução
  - 1.2.2. Funções convexas: definição e algumas propriedades
  - 1.2.3. Função indicadora de um conjunto convexo
  - 1.2.4. Funções convexas linearmente homogêneas
  - 1.2.5. Funções côncavas
  - 1.2.6. Funções lineares não verticais
  - 1.2.7. Funções semicontínuas/funções convexas e côncavas fechadas
  - 1.2.8. Funções contínuas/propriedades das funções convexas e côncavas contínuas
  - 1.2.9. Funções diferenciáveis/propriedades das funções convexas e côncavas diferenciáveis
  - 1.2.10. Subdiferenciabilidade de funções convexas e côncavas

**PARTE II : CONJUGAÇÃO E QUASE - CONJUGAÇÃO**

CAPÍTULO 2

- 2.1. Funções convexas e côncavas conjugadas ..... 45
  - 2.1.1. Funções convexas conjugadas: definição e interpretação geométrica

.../...

|        |                                                                              |    |
|--------|------------------------------------------------------------------------------|----|
| 2.1.2. | Principais propriedades não marginais da conjugada de uma função convexa     |    |
| 2.1.3. | Conjugação e desigualdade de Fenchel                                         |    |
| 2.1.4. | Vectores conjugados e propriedades marginais das funções convexas conjugadas |    |
| 2.1.5. | Conjugadas de funções côncavas e suas propriedades                           |    |
| 2.1.6. | Funções parcialmente conjugadas                                              |    |
| 2.2.   | <u>Funções suporte de conjuntos convexos</u>                                 | 55 |
| 2.2.1. | Função suporte de um conjunto convexo                                        |    |
| 2.2.2. | Função suporte como conjugada da função indicadora                           |    |
| 2.2.3. | Cones polares e funções suporte de cones convexos                            |    |
| 2.3.   | <u>Funções "gauge" e conjuntos polares</u>                                   | 59 |
| 2.3.1. | Funções "gauge" de tipo I de um conjunto convexo e suas propriedades         |    |
| 2.3.2. | Funções polares de funções gauge-I                                           |    |
| 2.3.3. | Função "gauge" de tipo II de um conjunto convexo e suas propriedades         |    |
| 2.3.4. | Funções polares de funções gauge-II                                          |    |

### CAPÍTULO 3

|        |                                                                                                                                     |    |
|--------|-------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.1.   | <u>Funções quase-convexas e quase-côncavas quase-conjugadas</u>                                                                     | 69 |
| 3.1.1. | Funções quase-convexas e quase-côncavas                                                                                             |    |
| 3.1.2. | Extensão a $\mathbb{R}^n$ de uma função quase-convexa ou quase-côncava/funções quase-convexas e quase-côncavas homogêneas de grau 0 |    |
| 3.1.3. | Conjuntos "evenly" convexos                                                                                                         |    |
| 3.1.4. | Cones "evenly" convexos e respectivos cones polares.                                                                                |    |
| 3.1.5. | Funções "evenly" quase-convexas ou "evenly" quase-côncavas                                                                          |    |
| 3.1.6. | Funções quase-conjugadas de funções quase-convexas homogêneas de grau 0                                                             |    |

- 3.1.7. Funções quase-conjugadas de funções quase-côncavas homogêneas de grau 0
- 3.1.8. Subgradiantes de funções quase-convexas e quase-côncavas homogêneas de grau 0
- 3.1.9. Propriedades marginais das funções quase-conjugadas

### PARTE III : APLICAÇÃO ÀS TEORIAS DO PRODUTOR E CONSUMIDOR

#### CAPÍTULO 4

- 4.1. Função lucro como função suporte do conjunto de possibilidades de produção ..... 87
  - 4.1.1. Hipóteses sobre a tecnologia da empresa
  - 4.1.2. Função lucro como função suporte do conjunto de possibilidades de produção/principais propriedades não marginais da função lucro
  - 4.1.3. Recuperação do conjunto de possibilidades de produção a partir da função lucro
  - 4.1.4. Propriedades marginais da função lucro/lema de Hotteling
- 4.2. Função lucro e função de produção clássica: caso uniprodução ..... 93
  - 4.2.1. Caso uniprodução: o conceito clássico de função de produção e suas propriedades
  - 4.2.2. Função lucro normalizada como conjugada da função de produção
  - 4.2.3. Recuperação da função de produção a partir da função lucro normalizada
  - 4.2.4. Principais propriedades marginais da função lucro normalizada/lema de Hotteling
- 4.3. Função lucro e função de produção generalizada de Jorgenson e Lau: multiprodução ..... 98
  - 4.3.1. Generalização do conceito clássico de função de produção
  - 4.3.2. Função lucro normalizada como conjugada da função de produção generalizada

- 4.3.3. Recuperação da função de produção generalizada a partir da função lucro normalizada
- 4.3.4. Propriedades marginais da função lucro normalizada/lema de Hotteling
- 4.4. Função lucro e função gauge-I do conjunto de possibilidades de produção ..... 102
  - 4.4.1. Função gauge-I do conjunto de possibilidades de produção
  - 4.4.2. A função lucro e a função gauge-I do conjunto de possibilidades de produção como funções gauge-I polares/obtenção de uma a partir da outra
  - 4.4.3. Propriedades marginais da função lucro/lema de Hotteling

## CAPITULO 5

- 5.1. Função custo como função suporte dos "Input Requirement Sets" ..... 106
  - 5.1.1. "Input Requirement Sets"
  - 5.1.2. Função custo como função suporte dos IRS
  - 5.1.3. Recuperação dos IRS a partir da função custo
  - 5.1.4. Propriedades marginais da função custo/lema de Shephard
  - 5.1.5. Uma hipótese adicional de "free disposal" de inputs/monotonia não decrescente em y da função custo
  - 5.1.6. Uniprodução: dualidade função de produção clássica/função custo
- 5.2. Função custo e função distância de Shephard ..... 112
  - 5.2.1. Função distância de Shephard
  - 5.2.2. Função distância e função custo como gauge-II polares

## CAPITULO 6

- 6.1. Dualidade entre funções de utilidade directa e indirecta ..... 118

- 6.1.1. Introdução: analogia formal entre as teorias do produtor e consumidor
- 6.1.2. Função de utilidade indirecta como quase-conjugada da função homogênea de grau 0 associada à função utilidade directa
- 6.1.3. Recuperação da função de utilidade directa a partir da função de utilidade indirecta
- 6.1.4. Propriedades marginais da função de utilidade indirecta/identidade de Roy
- 6.1.5. Dualidade função utilidade indirecta / função custo

## CONCLUSÃO

## BIBLIOGRAFIA

- AUBIN, J.P. (1984) L'analyse Non Linéaire et Ses Motivations Economiques, Masson, Paris.
- 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. Wiley and Sons, New York.
- BERGE, C. e GOUILLA-HOURI, A. (1962) Programmes, Jeux et Réseaux de Transport, Dunod, Paris.
- BLACKORBY, C., PRIMONT, D. e RUSSEL, R.R. (1978) Duality, Separability and Functional Structure: Theory and Economic Applications, North-Holland, Amsterdam, New York-Oxford.
- CORNWALL, R.R. (1984) Introduction to the Use of General Equilibrium Analysis, North-Holland, Amsterdam-New York-Oxford.
- DIEWERT, W.E. (1974) "Applications of Duality Theory" in INTRILIGATOR, M.D. e KENDRICK, D.A. (eds) Frontiers of Quantitative Economics, vol.II, North-Holland, Amsterdam - New York - Oxford.
- DIEWERT, W.E., AVRIEL, M. e ZANG, I (1981) "Nine Kinds of Quasi concavity and concavity", Journal of Economic Theory, 25, pp. 397-420.
- DIEWERT, W. E. (1982) "Duality Approaches to Microeconomic Theory", in ARROW, K.J. e INTRILIGATOR, M.D. (eds) Handbook of Mathematical Economics, vol. II, North-Holland, Amsterdam-New York-Oxford.
- EICHHORN, W., HENN, R., OPITZ, O. e SHEPHARD, R.W. (Eds) (1973) Production Theory, Proceedings of an International Seminar Held at The University of Karlsruhe, May-July, Springer-Verlag.

- HANOCH, G. (1978) "Symmetric Duality and Polar Functions", in FUSS M. e McFADDEN, D. (Eds) Production Economics: A Dual Approach to Theory and Applications, Vol I, North-Holland, Amsterdam-New York-Oxford.
- JACOBSEN, S.E. (1970) "Production Correspondences", Econometrica, vol 38, nº 5, Sept.
- JORGENSEN, D.W. e LAU, L.J. (1974) "The Duality of Technology and Economic Behaviour", The Review of Economic Studies vol 41, nº 126. April.
- LAU, L. J. (1978) "Applications of Profit Functions" in FUSS, M. e McFADDEN, D. (Eds) op. cit..
- LUENBERGER, D. G. (1969) Optimization by Space Vector Methods, J. Wiley and Sons, New-York.
- McFADDEN, D. (1978a) "Duality of Production, Cost and Profit Functions" in FUSS, M. e McFADDEN, D. (Eds), op. cit..
- McFADDEN, D. (1978b) "Convex Analysis", in FUSS, M. e McFADDEN, D. op. cit..
- PASSY, U. e PRISMAN, E.Z. (1984) "Conjugacy in Quasi-Convex Programming", Math. Programming, 30, pp. 121-146.
- PINHEIRO, M.R. (1984) Conjugação e Dualidade em Programação Convexa. Coleção de Working Papers da Faculdade de Economia da Universidade Nova de Lisboa, Lisboa.
- PINHEIRO, M.R. (1985) Dualidade Conjugada em Programação Matemática Convexa, Coleção de Working Papers da Faculdade de Economia da Universidade Nova de Lisboa, Lisboa.
- ROCKAFELLAR, R.T. (1964) "Duality Theorems for Convex Functions", Bull Amer. Math. Soc., 70, pp. 189-192.
- ROCKAFELLAR, R.T. (1967) "Conjugates and Legendre Transforms of Convex Functions", Canad. J. Math., 19, pp. 200-205.

- ROCKAFELLAR, R.T. (1968) "Duality in Nonlinear Programming" in Mathematics of Decision Sciences, Part II, Lectures in Applied Mathematics, vol II, Amer. Math. Soc. pp. 401-422.
- ROCKAFELLAR, R.T. (1969) "Convex Functions and Duality in Optimization-Problems and Dynamics" in KUHN, H. W. e SZEGÖ, G.P. (Eds) Mathematical Systems Theory and Economics I, Springer-Verlag, Berlin-Heidelberg- New York.
- ROCKAFELLAR, R.T. (1970) Convex Analysis, Princeton Univ. Press, Princeton, N.J.
- ROCKAFELLAR, R.T. (1981) The Theory of Subgradients and Its Applications to Problems of Optimization. Convex and Nonconvex Functions, Heldermann-Verlag, Berlin.
- SAMUELSON, P. (1953). "Prices of Factores and Goods in General Equilibrium", The Review of Economic Studies, 21, 1-20.
- SHEPHARD, R.W. (1953) Cost and Production Functions, Princeton Univ. Press, Princeton, N.J..
- SHEPHARD, R.W. (1970) Theory of Cost and Production Functions, Princeton Univ. Press, Princeton, N.J..
- UZAWA, H. (1964) "Duality Principles in the Theory of Cost and Production ", International Economic Review, vol 5, n. 2 May.
- VARIAN, H.R. (1978) Microeconomic Analysis, W.W. Norton & C., New York.
- WEYMARK, J. A. (1980) "Duality Results in Demand Theory", European Economic Review, 14, pp. 377-395.