## 3. Funções Convexas e Teoremas de Separação

A convexidade desempenha um papel crucial no estudo dos problemas de minimização. Depois de definirmos as funções convexas e descrevermos as suas propriedades mais elementares, mostrando, nomeadamente, que as funções convexas são funções localmente lipschitzianas (ou seja, de Lipschitz numa vizinhança de cada ponto). Provaremos ainda a existência e unicidade da solução óptima do seguinte problema de minimização:

$$\frac{1}{2}\|\bar{x} - x_0\|^2 + f(\bar{x}) = \inf_{x \in X} \left(\frac{1}{2}\|x - x_0\|^2 + f(x)\right),$$

onde f é uma função estrita, semi-contínua inferior de X em $\mathbf{R} \cup \{+\infty\}$.

Como caso particular derivar-se-à um teorema para a obtenção da melhor aproximação de $x_0$ de entre os pontos de um conjunto convexo fechado, o qual tem consequências muito interessantes, entre as quais se destacam os teoremas de separação que têm fortes implicações na teoria da dualidade e em Análise Convexa.

Ao longo deste capítulo vamos considerar que X denota um espaço vectorial com produto interno.

### 3.1 - Definição

Dizemos que uma função f de X em $\mathbf{R} \cup \{+\infty\}$ é convexa se, para qualquer combinação convexa $x = \sum_{i=1}^{n} \lambda_i x_i$ de pontos $x_i \in X$ se obtém a desigualdade

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i).$$

Dizemos que f é côncava se -f é convexa, e que f é afim se é concava e convexa.

Vamos começar por caracterizar as funções convexas.

### 3.2 - Teorema

Sendo f uma função de X em $\mathbf{R} \cup \{+\infty\}$, as seguintes condições são equivalentes:

a) f é convexa;

b) o epigrafo de f é convexo.

Prova:

Vamos começar por provar que a) implica b). Sendo $(x,\alpha)$ e $(y,\beta)$ dois pontos do epigrafo de f, vem que $f(x)\leq\alpha$ e $f(y)\leq\beta$. Logo tem-se que $\forall\lambda\in[0,1]$ $\lambda f(x)\leq\lambda\alpha$ e ainda que $(1-\lambda)f(y)\leq(1-\lambda)\beta$. Assim, adicionando-se estas duas desigualdades obtém-se $\lambda f(x)+(1-\lambda)f(y)\leq\lambda\alpha+(1-\lambda)\beta$, e (pela convexidade de f) que $f(\lambda x+(1-\lambda)y)\leq\lambda\alpha+(1-\lambda)\beta$. Donde se conclui que $\lambda(x,\alpha)+(1-\lambda)(y,\beta)\in Ep(f)$.

Para completarmos esta prova, vamos finalmente provar que b) implica a). Considerando-se os pontos $x_1, x_2, \ldots, x_n\in X$, dado que $(x_i,f(x_i))\in Ep(f)$ $\forall i\in\{1,\ldots,n\}$, pela convexidade de Ep(f) conclui-se que $\forall\lambda_1,\lambda_2,\ldots,\lambda_n\geq 0$ tais que $\sum_{i=1}^{n}\lambda_i=1$, vem que $\sum_{i=1}^{n}\lambda_i(x_i,f(x_i))=(\sum_{i=1}^{n}\lambda_i x_i, \sum_{i=1}^{n}\lambda_i f(x_i))\in Ep(f)$, o que implica que $f(\sum_{i=1}^{n}\lambda_i x_i)\leq\sum_{i=1}^{n}\lambda_i f(x_i)$, ou seja, f é convexa.

◊

Uma consequência imediata deste teorema é o seguinte corolário.

**3.3 - Corolário**

Um subconjunto K de X é convexo sse a sua função indicadora é convexa.

Prova:

Com efeito, $Ep(\psi_K)=Kx\mathbf{R}_+$ é convexo sse K é convexo.

◊

**3.4 - Teorema**

Suponha-se que as funções f, g, $f_i$ (para $i\in I$) de X em $\mathbf{R}\cup\{+\infty\}$ são convexas. Então

a) f+g é convexa;

b) se $\alpha>0$ então $\alpha f$ é convexa;

c) se A é uma aplicação linear do espaço vectorial Y em X, então $f\circ A$ é convexa;

d) se $\phi$: $\mathbf{R}\rightarrow\mathbf{R}$ é convexa e não decrescente então $\phi\circ f$ é convexa;

e) $\sup_{i\in I} f_i$ é convexa.

Prova:

Para esta prova considere-se que $\lambda\in[0,1]$ e que $x,y\in X$.

a) Pela convexidade de f e g vem que

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

Como consequência, conclui-se que $(f+g)(\lambda x+(1-\lambda)y)\leq\lambda(f(x)+g(x))+(1-\lambda)(f(y)+g(y))$, ou seja, $(f+g)(\lambda x+(1-\lambda)y)\leq\lambda(f+g)(x)+(1-\lambda)(f+g)(y)$.

b) Pela convexidade de f sabe-se que $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$. Logo para $\alpha>0$ vem que

$$(\alpha f)(\lambda x+(1-\lambda)y)=\alpha f(\lambda x+(1-\lambda)y) \leq \alpha[\lambda f(x)+(1-\lambda)f(y)]=\lambda\alpha f(x)+(1-\lambda)\alpha f(y)$$

$$\Leftrightarrow$$

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

c) Sendo u e v dois pontos arbitrários de Y vem que

$$f(A(\lambda u+(1-\lambda)v))=f(\lambda(A(u))+(1-\lambda)(A(v))\leq\lambda f(A(u))+(1-\lambda)f(A(v)),$$

o que é equivalente a ter-se $foA(\lambda u+(1-\lambda)v)\leq\lambda foA(u)+(1-\lambda)foA(v)$.

d) Um vez que $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$ e a função $\phi$ é não decrescente, pode concluir-se que $\phi(f(\lambda x+(1-\lambda)y)\leq\phi(\lambda f(x)+(1-\lambda)f(y))$. Por outro lado, pela convexidade de $\phi$, vem que $\phi(\lambda f(x)+(1-\lambda)f(y))\leq\lambda\phi(f(x))+(1-\lambda)\phi(f(y))$. Logo pode concluir-se que

$$\phi of(\lambda x+(1-\lambda)y)=\phi(f(\lambda x+(1-\lambda)y) \leq \lambda\phi(f(x))+(1-\lambda)\phi(f(y))=\lambda\phi of(x)+(1-\lambda)\phi of(y).$$

e) Esta prova decorre de $Ep(\sup_{i\in I} f_i) = \cap_{i\in I} Ep(f_i)$, e de $Ep(f_i)$ ser um conjunto convexo $\forall i\in I$, o que permite concluir que $\cap_{i\in I} Ep(f_i)$ é convexo.

◊

**3.5 - Teorema**

Se f é uma função convexa de X em $\mathbf{R}\cup\{+\infty\}$, então as suas secções inferiores $S(f,\lambda)$ são convexas.

Prova:

Sejam x e y dois pontos da secção inferior $S(f,\lambda)$. Então tem-se que $f(x)\leq\lambda$ e $f(y)\leq\lambda$. Logo $\forall\alpha\in[0,1]$, $\alpha f(x)\leq\alpha\lambda$ e $(1-\alpha)f(y)\leq(1-\alpha)\lambda$. Adicionando estas duas desigualdades vem que $\alpha f(x)+(1-\alpha)f(y)\leq\alpha\lambda+(1-\alpha)\lambda=\lambda$. Pela convexidade de f sabe-se que $f(\alpha x+(1-\alpha)y)\leq\alpha f(x)+(1-\alpha)f(y)$ e como consequência vem que $f(\alpha x+(1-\alpha)y)\leq\lambda$, o que prova que $\alpha x+(1-\alpha)y\in S(f,\lambda)$.

◊

Note-se que o recíproco do teorema anterior não é verdadeiro. Com efeito existem funções não convexas cujas secções inferiores são convexas.

As funções cujas secções inferiores são convexas designam-se por **funções quase-convexas**.

Com esta caracterização de funções quase-convexas, com facilidade se verifica que uma função estrita $f:X\rightarrow\mathbf{R}\cup\{+\infty\}$ é quase convexa sse

$$\forall x,y\in X\ \forall\lambda\in[0,1]\ f(\lambda x+(1-\lambda)y)\leq\max\{f(x),f(y)\}.$$

Com efeito suponha-se que todas as secções de f são convexas. Então $\forall x,y\in X$ fazendo $\alpha:=\max\{f(x),f(y)\}$, vem que $x,y\in S(f,\alpha)$, e consequentemente, pela convexidade de $S(f,\alpha)$, $\forall\lambda\in[0,1]$ vem que

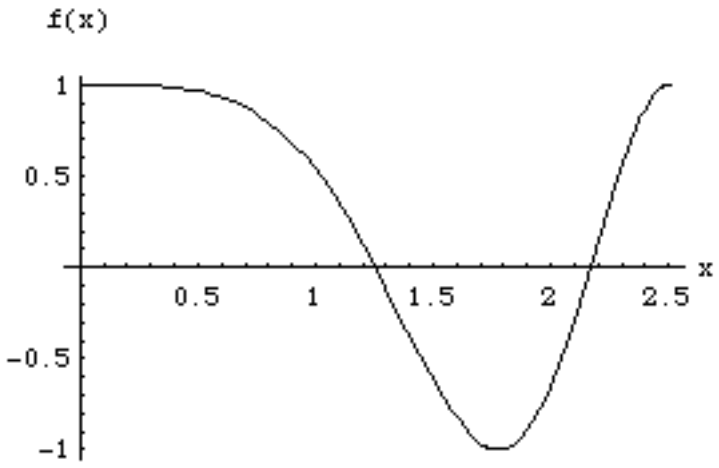
$$f(\lambda x+(1-\lambda)y)\leq\alpha=\max\{f(x),f(y)\}.$$

Reciprocamente, suponha-se que $\forall x,y\in X\ \forall\lambda\in[0,1]$ se verifica a desigualdade $f(\lambda x+(1-\lambda)y)\leq\beta:=\max\{f(x),f(y)\}$. Seja $S(f,\alpha)$ uma secção arbitrária de f e sejam x,y dois pontos arbitrários de $S(f,\alpha)$. Então dado que $f(x)\leq\alpha$ e $f(y)\leq\alpha$, pode concluir-se que $\beta:=\max\{f(x),f(y)\}\leq\alpha$, o que implica que $x,y\in S(f,\beta)$. Como consequência, tendo em conta a hipótese, vem que $f(\lambda x+(1-\lambda)y)\leq\beta\leq\alpha$. Logo $\lambda x+(1-\lambda)y\in S(f,\alpha)$.

Por exemplo, sendo $I=[0,\sqrt{2\pi}]$, a função:

$$f(x)=\begin{cases}\cos(x^2) & \text{se } x\in I\\ +\infty \text{ no caso contrário}\end{cases}$$

é uma função quase-convexa que não é convexa. Na verdade, embora as secções inferiores de f(x) sejam conjuntos convexos (neste caso intervalos) não se trata de uma função convexa. O gráfico de f(x) (relativo aos pontos do seu domínio) que se representa a seguir ilustra bem estas afirmações.




**3.6 - Definição**

Uma função estrita $f: X\rightarrow \mathbf{R}\cup\{+\infty\}$ é **estritamente convexa** se para $\lambda\in\,]0,1[$ e quaisquer dois pontos distintos x e y do Dom(f) se verifica que

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

Esta condição permite-nos estabelecer critérios de unicidade para a existência da solução óptima de um problema de optimização.

**3.7 - Teorema**

Seja f uma função convexa estrita de X em $\mathbf{R}\cup\{+\infty\}$. Então o conjunto M de soluções óptimas $x^*\in X$ do problema $f(x^*)=\inf_{x\in X} f(x)$ é convexo. Se f é estritamente convexa então M contém no máximo um ponto.

Prova:

Seja $\alpha:=\inf_{x\in X} f(x)$. A primeira afirmação é consequência de se ter $M=\bigcap_{\lambda>\alpha} S(f,\lambda)$ o que implica que M seja a intersecção de conjuntos convexos. Se f é estritamente convexa e se $x^*_1$ e $x^*_2$ são duas soluções óptimas do problema da determinação de $\alpha=\inf_{x\in X} f(x)$, então temos $\alpha \le f(\frac{x^*_1+x^*_2}{2}) < \frac{f(x^*_1)+f(x^*_2)}{2} = \alpha$, o que é impossível.

◊

**3.8 - Teorema**

Seja Y um conjunto convexo não vazio e $g: X\times Y\to\mathbf{R}\cup\{+\infty\}$ uma função convexa. Então a função f de X em $\mathbf{R}\cup\{+\infty\}$, definida por

$$f(x):=\inf_{y\in Y} g(x,y),$$

é convexa [9].

Prova:

Sejam $\varepsilon>0$, $\lambda\in\,]0,1[$ e $x_1,x_2\in X$. Suponha-se que $x_1$ e $x_2$ pertencem ao domínio de f (i. e, $f(x_1)<+\infty$ e $f(x_2)<+\infty$). Então existem $y_1$ e $y_2$ tais que

$$g(x_i,y_i) \le f(x_i)+\varepsilon \qquad (i=1,2).$$

Uma vez que g é uma função convexa, deduz-se que

$$g(\alpha x_1+(1-\alpha)x_2,\alpha y_1+(1-\alpha)y_2) \le \alpha g(x_1,y_1)+(1-\alpha)g(x_2,y_2) \le \alpha f(x_1)+(1-\alpha)f(x_2)+\varepsilon.$$

Porém, uma vez que $f(\alpha x_1+(1-\alpha)x_2)$ é menor ou igual a $g(\alpha x_1+(1-\alpha)x_2,\alpha y_1+(1-\alpha)y_2)$, vem que

$$f(\alpha x_1+(1-\alpha)x_2) \le \alpha f(x_1)+(1-\alpha)f(x_2)+\varepsilon \qquad (13),$$

pelo que basta fazer tender $\varepsilon$ para zero, para que se obtenha a desigualdade desejada.

Finalmente deve notar-se que se $x_1$ ou $x_2$ (pelo menos um destes pontos) não pertence ao domínio de f, então a desigualdade (13) é trivialmente satisfeita.

◊

**3.9 - Teorema**

Considerando-se n funções convexas $f_i$ de X em $\mathbf{R}\cup\{+\infty\}$ e a aplicação F de $K:=\bigcap_{1\le i\le n} \mathrm{Dom}(f_i)$ em $\mathbf{R}\cup\{+\infty\}$ definida por

---

[9] Observe-se que, de acordo com a definição de f ($\forall x\in X\ f(x)\in\mathbf{R}\cup\{+\infty\}$), $\forall x\in X\ \inf_{y\in Y} g(x,y)\in\mathbf{R}\cup\{+\infty\}$.

$$\forall x\in K \qquad F(x):=(f_1(x), \ldots ,f_n(x)),$$

verifica-se que os conjuntos $F(K)+\mathbf{R}^n_+$ e $F(K)+Int(\mathbf{R}^n_+)$ são convexos.

Prova:

Vamos fazer a prova apenas para o segundo dos conjuntos. O cone $Int(\mathbf{R}^n_+)$ é constituído pelos vectores de $\mathbf{R}^n$, u, de componentes estritamente positivas. Escolhidos dois pontos de $F(K)+Int(\mathbf{R}^n_+)$, $y_i=F(x_i)+u_i$ (para i=1,2). Se $\alpha\in ]0,1[$ podemos escrever:

$$y = \alpha y_1 + (1-\alpha)y_2 = \alpha F(x_1)+(1-\alpha)F(x_2)+\alpha u_1+(1-\alpha)u_2= F(\alpha x_1+(1-\alpha)x_2)+u,$$

para o que basta tomar $u = \alpha u_1+(1-\alpha)u_2+ \alpha F(x_1)+(1-\alpha)F(x_2)-F(\alpha x_1+(1-\alpha)x_2)$. Uma vez que a convexidade das funções $f_i$ implica que se tenha:

$$\begin{bmatrix} \alpha f_1(x_1)+(1-\alpha)f_1(x_2) \\ \alpha f_2(x_1)+(1-\alpha)f_2(x_2) \\ \cdots \\ \alpha f_n(x_1)+(1-\alpha)f_n(x_2) \end{bmatrix} - \begin{bmatrix} f_1(\alpha x_1+(1-\alpha)x_2) \\ f_2(\alpha x_1+(1-\alpha)x_2) \\ \cdots \\ f_n(\alpha x_1+(1-\alpha x_2)) \end{bmatrix} \geq 0$$

e uma vez que as componentes de $u_1$ e $u_2$ são estritamente positivas, podemos concluir que as componentes de u também são estritamente positivas. Logo $y\in F(K)+Int(\mathbf{R}^n_+)$.

◊

As normas [10] e semi-normas [11] de espaços vectoriais são exemplos de funções convexas. Mais geralmente, qualquer função f de X em $\mathbf{R}\cup\{+\infty\}$ subaditiva (ou seja, tal que $\forall x,y\in X\ f(x+y)\leq f(x)+f(y)$) e homogénea positiva de grau 1 (i. e, $\forall x\in X\ \forall\lambda>0$, $f(\lambda x)=\lambda f(x)$ [12]) é uma função convexa homogénea positiva e reciprocamente.

Com efeito, seja $x_1,x_2\in X$. Suponha-se que $f:X\rightarrow\mathbf{R}\cup\{+\infty\}$ é uma função subaditiva e homogénea positiva de grau 1, então $\forall\lambda\in ]0, 1[$ (pela subaditividade de f) $f(\lambda x+(1-\lambda)y)\leq f(\lambda x)+f((1-\lambda)y)$ e $f(\lambda x)+f((1-\lambda)y)=\lambda f(x)+(1-\lambda)f(y)$ (pela homogeneidade positiva de f). Reciprocamente, suponha-se que f é convexa e homogénea positiva, então

---

[10] Uma norma definida em X é uma função $\|.\|:X\rightarrow\mathbf{R}$ que $\forall x,y\in X\ \forall\lambda\in\mathbf{R}$ verifica as seguintes condições: a) $\|x\|>0$ se $x\neq 0$ e $\|0\|=0$, b) $\|\lambda x\|=|\lambda|.\|x\|$ e c) $\|x+y\|\leq\|x\|+\|y\|$).

[11] Uma seminorma definida em X é uma função de X em $\mathbf{R}$ que verifica todas as condições da norma à excepção da referida em a).

[12] Na verdada, para que uma função f de X em $\mathbf{R}\cup\{+\infty\}$ seja homogénea positiva , basta que se verifique (a) $f(\lambda x)\leq\lambda f(x)\ \forall x\in X$ e $\forall\lambda>0$. Com efeito, dado que $\lambda x\in X$ e $\lambda^{-1}>0$, vem que $f(x)=f(\lambda^{-1}\lambda x)\leq\lambda^{-1}f(\lambda x)$, donde se tira que (b) $\lambda f(x)\leq f(\lambda x)$. Finalmente de (a) e (b) vem que $\lambda f(x)\leq f(\lambda x)\leq\lambda f(x)\ \forall x\in X$ e $\forall\lambda>0$, o que é equivalente a ter-se $f(\lambda x)=\lambda f(x)\ \forall x\in X$ e $\forall\lambda>0$.

(pela convexidade de f) $f(x+y)=f(\frac{2}{2}x+\frac{2}{2}y)\leq\frac{1}{2}f(2x)+\frac{1}{2}f(2y)$ e (pela homogeneidade positiva de f) $\frac{1}{2}f(2x)+\frac{1}{2}f(2y)=f(x)+f(y)$.

Mostraremos ainda que as funções convexas contínuas num ponto são funções lipschitzianas numa vizinhança desse ponto.

**3.10 - Definição**

Uma função f de X em **R** diz-se lipschitziana se existe uma constante $c>0$ (chamada constante de Lipschitz) tal que $d(f(x),f(y))\leq c\cdot d(x,y)$ quaisquer que sejam $x,y\in X$. Uma função g de um subconjunto aberto $\Omega$ ($X\supseteq\Omega$) em **R** diz-se localmente lipschitziana se para cada ponto $x\in\Omega$ existe uma vizinhança de x na qual g é lipschitziana.

Note-se que toda a função lipschitziana é localmente lipschitziana em qualquer subconjunto aberto de X. Por outro lado uma função localmente lipschitziana num aberto A é contínua em todo o ponto $a\in A$. Com efeito, sendo c a constante de Lipschitz associada a f, para $\varepsilon>0$, tome-se $\delta\leq\frac{\varepsilon}{c}$ tal que $B(a,\delta)$ é uma bola aberta contida em A na qual f é lipschitziana com constante de Lipschitz igual a c. Nestas condições $d(x,a)<\delta \Rightarrow d(f(x),f(a))<c\cdot d(x,a)<c\cdot\delta\leq\varepsilon$. Facilmente se verifica ainda que toda a combinação linear de funções lipschitzianas ($f_1, f_2, \ldots, f_n$ de X em **R**), $f=\sum_{i=1}^{n}\alpha_i f_i$ (onde os $\alpha_i$ são escalares arbitrários) é uma função lipschitziana.

Por exemplo a função $f: \mathbf{R}\rightarrow\mathbf{R}$, definida por $f(x)=x^n$ (onde n é um inteiro positivo) é lipschitziana em cada parte limitada de **R**. Com efeito para $|x|\leq M$, vem que $|f'(x)|=n|x^{n-1}|\leq n\cdot M^{n-1}$. Assim, dados dois pontos arbitrários u e v de módulo não superior a M, uma vez que f é derivável, pelo teorema do valor médio existe z (entre u e v) tal que tal que $f(u)-f(v)=f'(z)(u-v)$. Como consequência vem que $|f(u)-f(v)|\leq c\cdot|u-v|$, onde $c=n\cdot M^{n-1}$. Alternativamente, poder-se-ia obter a mesma conclusão fazendo

$|f(u)-f(v)|=|u^n-v^n|=|u-v|\cdot|u^{n-1}+u^{n-2}v+\ldots+v^{n-1}|\leq|u-v|(|u|^{n-1}+|u|^{n-2}|v|+\ldots+|v|^{n-1})\leq c\cdot|u-v|$

com $c=n\cdot M^{n-1}$.

**3.11 - Teorema**

Sendo $f : X \to \mathbf{R}\cup\{+\infty\}$ uma função estrita e convexa, as seguintes condições são equivalentes

a) f admite um majorante num subconjunto aberto (necessariamente contido no domínio de f).

b) f é localmente lipschitziana no interior do seu domínio.

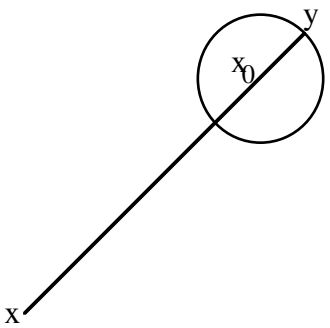
Prova:

Dado que a prova de que b) implica a) é óbvia, vamos apenas provar que a) implica b).

Seja $\eta>0$ e suponha-se que f é majorada na aderência da bola $x_0+\eta B(0,1)$ (necessàriamente contida no domínio de f), pela constante $M<+\infty$.

1. Vamos começar por provar que f é contínua em $x_0$.

A cada ponto $x\in X\setminus\{x_0\}$, vamos associar o ponto y pertencente à fronteira da bola $x_0+\eta B(0,1)$ (donde $f(y)\leq M$), tal que




$x_0=\theta y+(1-\theta)x$, o que implica $y=\frac{x_0-(1-\theta)x}{\theta}$, onde $\theta=\frac{\|x_0-x\|}{\eta+\|x_0-x\|}<1$ (tenha-se em conta que $y=x+(\|x_0-x\|+\eta)\frac{x_0-x}{\|x_0-x\|}$, pelo que $y=\frac{\|x_0-x\|+\eta}{\|x_0-x\|}x_0-\frac{\eta}{\|x_0-x\|}x \Leftrightarrow \frac{\|x_0-x\|}{\|x_0-x\|+\eta}y+\frac{\eta}{\|x_0-x\|+\eta}x=x_0$).

Nestas condições, pela convexidade de f, obtêm-se as inequações equivalentes

$$f(x_0)=f(\theta y+(1-\theta)x)\leq\theta M+(1-\theta)f(x) \Leftrightarrow f(x_0)-\theta f(x_0)\leq\theta M-\theta f(x_0)+(1-\theta)f(x)$$

as quais, por sua vez, são equivalentes às inequações

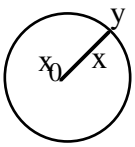
$$(1-\theta)f(x_0)\leq\theta(M-f(x_0))+(1-\theta)f(x) \Leftrightarrow f(x_0)\leq\frac{\theta}{1-\theta}(M-f(x_0))+f(x).$$

Substituindo, na última inequação obtida, $\theta$ pelo seu valor vem que

$$\forall x\in X,\quad f(x_0)-f(x)\leq\frac{M-f(x_0)}{\eta}\|x-x_0\| \qquad (14),$$

a qual também se verifica para $x=x_0$.

Tome-se agora $x\in\{x_0+\eta B(0,1)\}\setminus\{x_0\}$ e y pertencente à fronteira da bola $x_0+\eta B(0,1)$ (pelo que, novamente, se verifica $f(y)\leq M$), tal que $x=\theta y+(1-\theta)x_0$, donde vem que $y=\frac{x-(1-\theta)x_0}{\theta}$, com $\theta=\frac{\|x-x_0\|}{\eta}\leq 1$.




Deve ter-se em atenção que

$$y=x_0+\eta\frac{x-x_0}{||x-x_0||},$$

pelo que

$$y=\frac{\eta}{||x-x_0||}x-\frac{\eta-||x-x_0||}{||x-x_0||}x_0 \Leftrightarrow \frac{||x-x_0||}{\eta}y+\frac{\eta-||x-x_0||}{\eta}x_0=x.$$

Mais uma vez, a convexidade de f implica que se tenha

$$f(x)=f(\theta y+(1-\theta)x_0)\leq\theta M+(1-\theta)f(x_0) \Leftrightarrow f(x)\leq\theta(M-f(x_0))+f(x_0),$$

e, consequentemente, substituindo θ pelo seu valor, que se tenha

$$\forall x\in x_0+\eta B(0,1), \qquad f(x)-f(x_0)\leq\frac{M-f(x_0)}{\eta}||x-x_0|| \qquad (15).$$

As desigualdades (14) e (15) implicam

$$\forall x\in x_0+\eta B(0,1), \qquad |f(x)-f(x_0)|\leq\frac{M-f(x_0)}{\eta}||x-x_0|| \qquad (16),$$

donde se pode concluir que f é contínua em $x_0$.

2. Vamos agora provar que f é lipschitziana na bola $x_0+\beta B(0,1)$, onde $\beta<\eta$. Considerem-se dois pontos $x_1$ e $x_2$ da bola $x_0+\beta B(0,1)$, um inteiro, n, maior do que $\frac{||x_1-x_2||}{\eta-\beta}$ e divida-se o segmento $x_1x_2$ em n partes, de modo a que se obtenham os pontos

$$y_j:=x_1+\frac{j}{n}(x_2-x_1) \text{ (para } j=0, 1, \dots, n).$$

Note-se que $y_0=x_1$ e $y_n=x_2$ e ainda que os pontos $y_j$ pertencem à bola $x_0+\beta B(0,1)$. Nestas condições verifica-se que as bolas $y_j+(\eta-\beta)B(0,1)$ estão contidas na bola $x_0+\eta B(0,1)$ [13]. Como consequência a função f é majorada por M nas bolas

$$y_j+(\eta-\beta)B(0,1) \text{ (para } j=0, 1, \dots, n).$$

Assim, da desigualdade (16), com $x_0$ substituído por $y_j$ e η substituído por η-β, obtém-se

$$|f(y_{j+1})-f(y_j)| \leq \frac{M-f(y_j)}{\eta-\beta}||y_{j+1}-y_j||$$

uma vez que $||y_{j+1}-y_j|| \leq \frac{||x_1-x_2||}{n} \leq \eta-\beta$, i. e, uma vez que $y_{j+1}\in y_j+(\eta-\beta)B(0,1)$.

Por outro lado, a desigualdade (14) implica que, para j=0, … ,n, se tenha

$$f(x_0)-f(y_j) \leq \frac{M-f(x_0)}{\eta}||y_j-x_0|| \leq M-f(x_0),$$

dado que $||y_j-x_0||\leq\eta$.

Então $\forall j\in\{0, 1, \dots, n\}$, vem que

---

[13] Com efeito, se $y\in y_j+(\eta-\beta)B(0,1)$ então $d(y,x_0)\leq d(y,y_j)+d(y_j,x_0)<(\eta-\beta)+\beta=\eta$, o que implica que $y\in x_0+\eta B(0,1))$.

$$|f(y_{j+1})-f(y_j)| \le \frac{M-f(x_0)+f(x_0)-f(y_j)}{\eta-\beta}\,\|y_{j+1}-y_j\| \le \frac{M-f(x_0)+M-f(x_0)}{\eta-\beta}\,\|y_{j+1}-y_j\|,$$

ou seja,

$$|f(y_{j+1})-f(y_j)| \le 2\frac{M-f(x_0)}{\eta-\beta}\,\|y_{j+1}-y_j\|.$$

Uma vez que $\|x_1-x_2\|=\sum_{j=0}^{\eta-1}\|y_{j+1}-y_j\|$, vem que

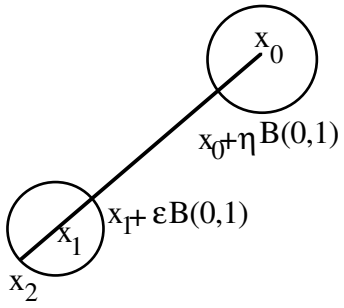
$$|f(x_1)-f(x_2)| \le \sum_{j=0}^{\eta-1}|f(y_{j+1})-f(y_j)| \le 2\frac{M-f(x_0)}{\eta-\beta}\|x_1-x_2\|,$$

o que nos permite concluir que f é lipschitziana na bola $x_0+\beta B(0,1)$.

3. Finalmente vamos mostrar que f é lipschitziana numa vizinhança de cada ponto pertencente ao interior do seu domínio, para o que, tendo em conta o que se demonstrou anteriormente, é suficiente provar que f é majorada numa vizinhança de cada ponto do interior do seu domínio.

De acordo com a hipótese, existe $x_0$ pertencente ao domínio de f e $\eta>0$ tal que f é majorada na aderência da bola $x_0+\eta B(0,1)$.

Seja $x_1$ um ponto do interior do domínio de f e seja $\varepsilon>0$ tal que a bola $x_1+\varepsilon B(0,1)$ está contida no domínio de f. Considerando-se o ponto $x_2=x_0+(\varepsilon+\|x_1-x_0\|)\frac{x_1-x_0}{\|x_1-x_0\|}$, que pertence à aderência da bola $x_1+\varepsilon B(0,1)$, verifica-se que $x_1=\lambda x_0+(1-\lambda)x_2$ (ou seja, $x_2=x_0+\frac{1}{1-\lambda}(x_1-x_0)$), para $\lambda=\frac{\varepsilon}{\varepsilon+\|x_1-x_0\|}$.



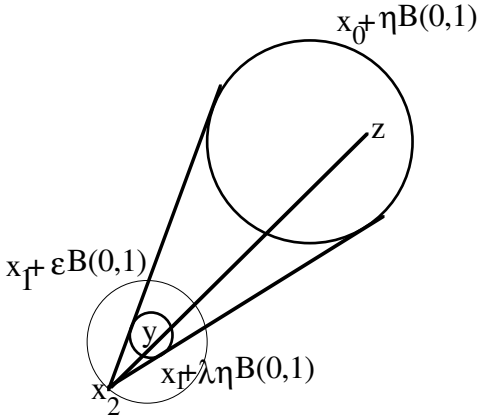

Com efeito, de $x_2=x_0+(\varepsilon+\|x_1-x_0\|)\frac{x_1-x_0}{\|x_1-x_0\|}$ tira-se que

$$x_2 = (1-\frac{\varepsilon+\|x_1-x_0\|}{\|x_1-x_0\|})x_0 + \frac{\varepsilon+\|x_1-x_0\|}{\|x_1-x_0\|}\,x_1 \Leftrightarrow x_2 + \frac{\varepsilon}{\|x_1-x_0\|}\,x_0 = \frac{\varepsilon+\|x_1-x_0\|}{\|x_1-x_0\|}\,x_1,$$

ou seja, $x_1 = \frac{\varepsilon}{\varepsilon+\|x_1-x_0\|}\,x_0 + \frac{\|x_1-x_0\|}{\varepsilon+\|x_1-x_0\|}\,x_2$.

Vamos finalmente provar que a função f, na bola $x_1+\lambda\eta B(0,1)$, é majorada por $\lambda M+(1-\lambda)f(x_2)$.

Primeiramente note-se que, se $y\in x_1+\lambda\eta B(0,1)$, então $z:=\frac{1}{\lambda}(y-(1-\lambda)x_2)$ pertence à bola $x_0+\eta B(0,1)$.




Com efeito

$$d(z,x_0)=\|z-x_0\|=\|\frac{1}{\lambda}y - \frac{1-\lambda}{\lambda}x_2 - x_0\|=\|\frac{1}{\lambda}(y-x_2)+(x_2-x_0)\|=\|\frac{1}{\lambda}(y-x_1+x_1-x_0+x_0-x_2)-(x_0-x_2)\|,$$

donde vem que

$$d(z,x_0)=\|\frac{1}{\lambda}(y-x_1) + \frac{1}{\lambda}(x_1-x_0) + \frac{1}{\lambda}(x_0-x_2) - \frac{\lambda}{\lambda}(x_0-x_2)\|=\|\frac{1}{\lambda}(y-x_1) - \frac{1}{\lambda}(x_0-x_1) + \frac{1-\lambda}{\lambda}(x_0-x_2)\|,$$

logo, tendo em conta a definição de $x_2$, tira-se que $x_0-x_2=\frac{1}{1-\lambda}(x_0-x_1)$, pelo que se obtém

$$d(z,x_0) = \|\frac{1}{\lambda}(y-x_1) - \frac{1}{\lambda}(x_0-x_1) + \frac{1}{\lambda}(x_0-x_1)\| =\frac{1}{\lambda}\|y-x_1\| <\frac{\lambda\eta}{\lambda} = \eta.$$

Como consequência vem que $f(z)\leq M$ e, pela convexidade de f e tendo em conta que $y=\lambda z+(1-\lambda)x_2$, tem-se que

$$f(y) = f(\lambda z+(1-\lambda)x_2) \leq \lambda f(z) + (1-\lambda)f(x_2) \leq \lambda M + (1-\lambda)f(x_2)$$

o que completa a prova.

◊

**3.12 - Corolário**

Se o interior do domínio de uma função convexa f de $\mathbf{R}^n$ em $\mathbf{R}\cup\{+\infty\}$ é não vazio, então f é localmente lipschitziana no interior do Dom(f).

Prova:

Considere-se a bola $x_0+\eta B(0,1)$ contida no domínio de f. Podemos então encontrar n pontos $x_i\in x_0+\eta B(0,1)$ tais que $x_i-x_0$ (i=1, … , n) são vectores linearmente independentes. Consequentemente o conjunto, S, das combinações lineares convexas $S=\{x=\sum_{i=0}^{n}\lambda_i x_i: \lambda_i>0, i=0,\ldots,n \text{ e } \sum_{i=0}^{n}\lambda_i=1\}$ é aberto e está contido no domínio de f.

Uma vez que f admite um majorante no conjunto aberto S que corresponde a

$$\max\{f(x_i): i\in\{0,\ldots,n\}\},$$

(tenha-se em conta que $f(\sum_{i=0}^{n}\lambda_i x_i) \leq \sum_{i=0}^{n}\lambda_i f(x_i) \leq \max_{0\leq i\leq n} f(x_i)$), da aplicação do teorema 3.11, decorre a prova pretendida.

◊

Vamos agora considerar o problema de minimização

$$f_\lambda(x) := \inf_{y\in X} [f(y) + \frac{1}{2\lambda}\|y-x\|^2] \qquad (17),$$

onde X é um espaço de Hilbert [14], f é uma função de X em $\mathbf{R}\cup\{+\infty\}$, $\|.\|$ denota a norma de Hilbert (induzida pelo produto interno definido em X) e $\lambda$ é um parâmetro positivo.

Nos resultados que se seguem, X* identifica o espaço das funcionais lineares reais contínuas sobre X e, para cada $y\in X^*$ e para cada $x\in X$, $\langle y,x\rangle$ denota y(x) (nestas condições, $X^*=\{y:X\to\mathbf{R}$ tais que $y(x)=\langle y,x\rangle\}$). A norma considerada em X* será identificada por $\|.\|^*$ e é definida por $\|y\|^*=\sup\{\langle y,x\rangle: \|x\|=1\}$.

Nestas condições, para o operador $\langle y,x\rangle$ definido em $X^*xX$, é válida a desigualdade de Cauchy-Schwarz. Com efeito, da definição de $\|.\|^*$ decorre a desigualdade $\langle y,\frac{x}{\|x\|}\rangle \le \|y\|^*$ a qual, por sua vez, implica $\langle y,x\rangle \le \|y\|^*\|x\|$.

Note-se que, pelo teorema da representação de Riesz [2], qualquer que seja a funcional linear contínua, g, sobre um espaço de Hilbert, H (no nosso caso X, onde se supõe definido o produto interno real, $u|v\in\mathbf{R}\ \forall u,v\in X$), existe um único $u\in H$ tal que $g(v)=u|v\ \forall v\in H$ (pelo que, podemos identificar g(v)=u(v) por $\langle u,v\rangle=u|v$). Nestas condições, qualquer desigualdade que envolva o operador $\langle,\rangle$ de $X^*xX$, pode ser estudada, tendo em conta que $\langle y,x\rangle=y|x$. Como consequência, a dedução de que $\|y\|^*=\|y\|$ pode fazer-se conforme a seguir se descreve.

$$\|y\| = y|\frac{y}{\|y\|} = \langle y,\frac{y}{\|y\|}\rangle \le \sup_{x\in X}\langle y,\frac{x}{\|x\|}\rangle = \|y\|^* = \sup_{x\in X} y|\frac{x}{\|x\|} \le \sup_{x\in X}\|y\|.\|\frac{x}{\|x\|}\| = \|y\|.$$

[14]Um espaço de Hilbert é um espaço com produto interno completo (i. e, no qual todas as sucessões de Cauchy convergem).

**3.13 - Teorema**

Se f é uma função estrita, semi-contínua inferior e convexa de X em $\mathbf{R}\cup\{+\infty\}$, então existe $p\in X^*$ e $\alpha\in\mathbf{R}$ tal que:

$$\forall y\in X, \qquad f(y) \geq \langle p,y\rangle+\alpha.$$

Prova:

Uma vez que f é estrita, basta considerar $x_0\in X$ tal que $f(x_0)$ é finito e separar o ponto $(x_0,f(x_0)-1)$ de Ep(f) por um hiperplano $H_0$ de $X\times\mathbf{R}$. Da convexidade de f decorre que Ep(f) é convexo e da sua semi-continuidade inferior decorre que Ep(f) é fechado. Logo Ep(f) é um fechado convexo de $X\times\mathbf{R}$ e $(x_0,f(x_0)-1)\notin Ep(f)$. Nestas condições, sabe-se que existe $c\in X^*$ e existem escalares $\delta_0$ e $\delta_1$ que determinam o hiperplano:

$$H_0=\{(y,\beta)\in X\times\mathbf{R}: \langle c,y\rangle+\delta_0\beta=\delta_1\}\ [15].$$

A existência deste hiperplano de separação é garantida pelo corolário 3.15 que se apresenta mais adiante [16].

Evidentemente que $\delta_0\neq0$, caso contrário $(x_0,f(x_0))$ e $(x_0,f(x_0)-1)$ não poderiam estar em meios espaços opostos, conforme se requer. Nestas condições, considere-se, sem perda de generalidade, que $\delta_0>0$ (se não for o caso, basta multiplicar a equação $\langle c,y\rangle+\delta_0\beta=\delta_1$ por -1). Consequentemente, $H_0$ é o gráfico da aplicação afim $g:X\to\mathbf{R}$ definida por $g(y) = \langle\frac{-1}{\delta_0}c,y\rangle + \frac{\delta_1}{\delta_0}$.

Suponha-se $\exists x\in X$ tal que $f(x)<g(x)=\beta$, então $\langle c,x\rangle+\delta_0 f(x)<\delta_1$. Consequentemente, dado que Ep(f) está contido no mesmo meio espaço e $(x,f(x)+\lambda)\in Ep(f)$ $\forall\lambda\geq0$, obtém-se a inequação $\langle c,x\rangle+\delta_0(f(x)+\lambda)\leq\delta_1$ $\forall\lambda\geq0$, a qual não se pode verificar para $\lambda$ suficientemente grande (basta tomar $\lambda>\frac{1}{\delta_0}(\delta_1-\langle c,x\rangle)-f(x)$). Assim podemos concluir que $\forall y\in X$ $f(y)\geq g(y)$.

Logo, fazendo $p = \frac{-1}{\delta_0}c$ e $\alpha = \frac{\delta_1}{\delta_0}$, vem que $\forall y\in X \qquad f(y) \geq \langle p,y\rangle+\alpha$.

◊

**3.14 - Teorema**

---

[15]Designa-se por hiperplano toda translação de um subespaço vectorial de codimensão 1, ou equivalentemente, todo o conjunto de nível de alguma funcional linear real contínua não identicamente nula.

[16]Note-se que, embora o teorema 3.14 (a seguir) que dá origem ao corolário 3.15 se apoie no teorema que estamos a demonstrar, como veremos, isso apenas se verifica, nos casos em que a função f do problema (17) toma valores negativos, para garantir que $f_\lambda$ é limitada inferiormente. Nos casos em que f é positiva ou zero, como é o caso da função indicadora de um conjunto convexo utilizada no corolário 3.15 , a prova do teorema 3.14 não necessita de se apoiar neste teorema (3.13).

Suponha-se que $f:X\to\mathbf{R}\cup\{+\infty\}$ é uma função estrita, convexa e semi-contínua inferior. Então existe uma única solução para o problema (17):

$$f_\lambda(x) = f(y_\lambda)+\frac{1}{2\lambda}\|y_\lambda-x\|^2.$$

Esta solução, $y_\lambda$, é caracterizada pela seguinte desigualdade variacional

$$\forall y\in X \qquad \frac{1}{\lambda}\langle y_\lambda-x,y_\lambda-y\rangle + f(y_\lambda) - f(y) \le 0.$$

Prova:

1. Primeiramente vamos provar que $f_\lambda(x)$ é uma função limitada inferiormente, ou seja, $\forall x\in X\ f_\lambda(x)>-\infty$.

Se f é positiva ou zero, então $f_\lambda$ é também positiva ou zero e, por conseguinte, é limitada inferiormente. Suponhamos que f toma valores negativos.

Pelo teorema 3.13 existe $p\in X^*$ e $\alpha\in\mathbf{R}$ tais que

$$\forall y\in X \qquad f(y) \ge \langle p,y\rangle+\alpha \qquad (18).$$

Por outro lado, da desigualdade de Cauchy-Schwarz, $\langle p,x-y\rangle\le\|p\|^*\cdot\|x-y\|$, e tendo em conta que, para $\lambda>0$, $0\le(\|\lambda p\|^*-\|x-y\|)^2=\lambda^2\|p\|^{*2}+\|x-y\|^2-2\lambda\|p\|^*.\|x-y\|$, ou seja, $\|p\|^*.\|x-y\|\le\frac{\lambda}{2}\|p\|^{*2}+\frac{1}{2\lambda}\|x-y\|^2$, vem que

$$\langle p,x-y\rangle \le \frac{\lambda}{2}\|p\|^{*2}+\frac{1}{2\lambda}\|x-y\|^2$$

(19).

De (18) vem que $f(y)\ge\langle p,y-x\rangle + \alpha + \langle p,x\rangle$, donde se tira que

$$f(y) + \frac{1}{2\lambda}\|y-x\|^2 \ge \langle p,y-x\rangle + \alpha + \langle p,x\rangle + \frac{1}{2\lambda}\|y-x\|^2,$$

Combinando esta última desigualdade com (19), obtém-se

$$f(y) + \frac{1}{2\lambda}\|y-x\|^2 \ge \langle p,y-x\rangle + \alpha + \langle p,x\rangle + \langle p,x-y\rangle - \frac{\lambda}{2}\|p\|^{*2}$$

$$\Leftrightarrow$$

$$f(y) + \frac{1}{2\lambda}\|y-x\|^2 \ge \alpha + \langle p,x\rangle - \frac{\lambda}{2}\|p\|^{*2},$$

desigualdade que, para um ponto arbitrário $x\in X$, é válida para todo o y em X.

Consequentemente $\forall x\in X$ vem que

$$f_\lambda(x) \ge \alpha + \langle p,x\rangle - \frac{\lambda}{2}\|p\|^{*2} > -\infty.$$

2. Seguidamente, vamos mostrar que qualquer solução, $y_\lambda$, do problema (17)

$$f_\lambda(x) = f(y_\lambda)+\frac{1}{2\lambda}\|y_\lambda-x\|^2,$$

satisfaz a desigualdade variacional $\forall y\in X$, $\frac{1}{\lambda}\langle y_\lambda-x,y_\lambda-y\rangle + f(y_\lambda) - f(y) \le 0$.

Seja $z=y_\lambda+\theta(y-y_\lambda)=\theta y+(1-\theta)y_\lambda$, com $\theta\in\,]0,1[$. Então obtém-se a desigualdade

$$f_\lambda(x)=f(y_\lambda)+\frac{1}{2\lambda}\|y_\lambda-x\|^2\le f(y_\lambda+\theta(y-y_\lambda))+\frac{1}{2\lambda}\|y_\lambda+\theta(y-y_\lambda)-x\|^2$$

$$\leq (1-\theta) f(y_\lambda) + \theta f(y) + \frac{1}{2\lambda}\|y_\lambda - x\|^2 + \frac{\theta^2}{2\lambda}\|y - y_\lambda\|^2 + \frac{\theta}{\lambda}\langle y_\lambda - x, y - y_\lambda\rangle,$$

a qual implica,

$$\theta(f(y_\lambda) - f(y)) + \frac{\theta}{\lambda}\langle y_\lambda - x, y_\lambda - y\rangle \leq \frac{\theta^2}{2\lambda}\|y - y_\lambda\|^2 \Leftrightarrow f(y_\lambda) - f(y) + \frac{1}{\lambda}\langle y_\lambda - x, y_\lambda - y\rangle \leq \frac{\theta}{2\lambda}\|y - y_\lambda\|^2.$$

Como consequência, fazendo tender $\theta$ para zero, obtém-se a desigualdade pretendida.

3. Reciprocamente, vamos provar que, se $y_\lambda$ satisfaz a desigualdade

$$\frac{1}{\lambda}\langle y_\lambda - x, y_\lambda - y\rangle + f(y_\lambda) - f(y) \leq 0,$$

então $y_\lambda$ é tal que $f_\lambda(x) = f(y_\lambda) + \frac{1}{2\lambda}\|y_\lambda - x\|^2$.

Com efeito, tendo em conta que $\frac{1}{2}\|y_\lambda - x\|^2 - \frac{1}{2}\|y - x\|^2 \leq \langle y_\lambda - x, y_\lambda - y\rangle$ [17], obtém-se

$$f(y_\lambda) + \frac{1}{2\lambda}\|y_\lambda - x\|^2 - f(y) - \frac{1}{2\lambda}\|y - x\|^2 \leq f(y_\lambda) - f(y) + \frac{1}{\lambda}\langle y_\lambda - x, y_\lambda - y\rangle.$$

Dado que por hipótese $f(y_\lambda) - f(y) + \frac{1}{\lambda}\langle y_\lambda - x, y_\lambda - y\rangle \leq 0$, vem que

$$f(y_\lambda) + \frac{1}{2\lambda}\|y_\lambda - x\|^2 \leq f(y) + \frac{1}{2\lambda}\|y - x\|^2 \qquad \forall\, y \in X.$$

4. Vamos agora provar que existe uma solução, $y_\lambda$, para o problema (17).

Para esta prova, vamos considerar a sucessão de pontos $y_n \in X$ que satisfaz

$$f(y_n) + \frac{1}{2\lambda}\|y_n - x\|^2 \leq f_\lambda(x) + \frac{1}{n} \qquad (20),$$

a qual provaremos tratar-se de uma sucessão de Cauchy.

Pela fórmula da mediana [18] obtém-se

$$\left\|\frac{y_n - x}{2} + \frac{x - y_m}{2}\right\|^2 + \left\|\frac{y_n - x}{2} - \frac{x - y_m}{2}\right\|^2 = 2\left\|\frac{y_n - x}{2}\right\|^2 + 2\left\|\frac{x - y_m}{2}\right\|^2$$

$$\Leftrightarrow$$

$$\frac{1}{4}\|y_n - y_m\|^2 + \left\|\frac{y_n + y_m}{2} - x\right\|^2 = \frac{1}{2}\|y_n - x\|^2 + \frac{1}{2}\|y_m - x\|^2$$

e esta última equação é, por sua vez, equivalente à equação

$$\|y_n - y_m\|^2 = 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\left\|\frac{y_n + y_m}{2} - x\right\|^2$$

(21).

Consequentemente, tendo em conta que por (20) $\|y_p - x\|^2 \leq 2\lambda(f_\lambda(x) - f(y_p) + \frac{1}{p})$, aplicando (21) obtém-se

---

[17] Esta desigualdade decorre imediatamente de um desenvolvimento sugerido pelo Prof. Andrey Sarychev

$$\frac{1}{2}\|y - x\|^2 = \frac{1}{2}\|(y_\lambda - x) - (y_\lambda - y)\|^2 = \frac{1}{2}\|y_\lambda - x\|^2 - \langle y_\lambda - x, y_\lambda - y\rangle + \frac{1}{2}\|y_\lambda - y)\|^2 \geq \frac{1}{2}\|y_\lambda - x\|^2 - \langle y_\lambda - x, y_\lambda - y\rangle.$$

[18] Fazendo $u = \frac{y_n - x}{2}$ e $v = \frac{x - y_m}{2}$ obtém-se

$$\|u+v\|^2 + \|u-v\|^2 = \|u\|^2 + u|v + v|u + \|v\|^2 + \|u\|^2 - u|v - v|u + \|v\|^2 = 2\|u\|^2 + 2\|v\|^2 \Leftrightarrow \|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

$$\|y_n-y_m\|^2 \le 4\lambda(f_\lambda(x)-f(y_n)+\frac{1}{n}) + 4\lambda(f_\lambda(x)-f(y_m)+\frac{1}{m}) - 4\|\frac{y_n+y_m}{2} - x\|^2$$
$$\le 4\lambda(\frac{1}{n} + \frac{1}{m} +2f_\lambda(x) - f(y_n) - f(y_m)) - 4\|\frac{y_n+y_m}{2} - x\|^2 \qquad (22).$$

Dado que, pela definição de $f_\lambda(x)$, se verifica a inequação

$$f_\lambda(x) \le f(\frac{y_n+y_m}{2})+\frac{1}{2\lambda}\|\frac{y_n+y_m}{2} - x\|^2,$$

vem

$$2\lambda(f_\lambda(x) - f(\frac{y_n+y_m}{2})) \le \|\frac{y_n+y_m}{2} - x\|^2 \Leftrightarrow -4\|\frac{y_n+y_m}{2} - x\|^2 \le 8\lambda(f(\frac{y_n+y_m}{2}) - f_\lambda(x)).$$

Combinando esta última inequação com a inequação (22), obtém-se

$$\|y_n-y_m\|^2 \le 4\lambda(\frac{1}{n} + \frac{1}{m} +2f_\lambda(x) - f(y_n) - f(y_m)) + 8\lambda(f(\frac{y_n+y_m}{2}) - f_\lambda(x))$$
$$= 4\lambda(\frac{1}{n} + \frac{1}{m} + 2f(\frac{y_n+y_m}{2}) - f(y_n) - f(y_m)).$$

Finalmente, pela convexidade de f,

$$f(\frac{y_n+y_m}{2}) \le \frac{1}{2}(f(y_n) + f(y_m)) \Leftrightarrow 2f(\frac{y_n+y_m}{2}) - f(y_n) - f(y_m) \le 0,$$

obtém-se $\|y_n-y_m\|^2 \le 4\lambda(\frac{1}{n} + \frac{1}{m})$.

Como consequência, dado que X é completo, vem que $y_n$ converge para um ponto $y_\lambda$. Logo, da semi-continuidade inferior de f, decorre que

$$f(y_\lambda) + \frac{1}{2\lambda}\|y_\lambda-x\|^2 \le \liminf_{y_n\to y_\lambda} [f(y_n) + \frac{1}{2\lambda}\|y_n-x\|^2] \le \liminf_{n\to\infty} [f_\lambda(x) + \frac{1}{n}] = f_\lambda(x)$$

e, consequentemente, que $f_\lambda(x)=f(y_\lambda) + \frac{1}{2\lambda}\|y_\lambda-x\|^2$.

5. Resta, finalmente, provar que a solução do problema (17) é única. Suponha-se que existem duas soluções, y' e y", para o problema (17). Da desigualdade variacional

$$\frac{1}{\lambda} \langle y_\lambda-x, y_\lambda-y\rangle + f(y_\lambda) - f(y) \le 0,$$

válida para todo o y em X, em particular, vem que

$$\frac{1}{\lambda} \langle y'-x, y'-y"\rangle + f(y') - f(y") \le 0 \text{ e } \frac{1}{\lambda} \langle y"-x, y"-y'\rangle + f(y") - f(y') \le 0.$$

Da adição das desigualdades anteriormente obtidas decorre que

$$\frac{1}{\lambda} (\langle y'-x, y'-y"\rangle + \langle y"-x, y"-y'\rangle) \le 0 \Leftrightarrow \frac{1}{\lambda} \langle y'-x-y"+x, y'-y"\rangle = \frac{1}{\lambda} \|y'-y"\|^2 \le 0,$$

donde se conclui que y'=y".

◊

Aplicando este teorema ao caso em que a função f corresponde à função indicadora de um conjunto convexo arbitrário K ($\psi_K$=f), somos conduzidos à determinação do ponto de K mais próximo de x, i. e, à determinação da projecção de x sobre K. Com efeito, neste caso, $f_\lambda(x) = \frac{1}{2\lambda}d(x,K)^2$, onde $d(x,K):=\inf_{y\in K}\|y-x\|$ é a distância de x a K.

**3.15 - Corolário**

Seja K um subconjunto convexo fechado não vazio do espaço de Hilbert X. O problema de minimização

$$g(x) = \inf_{y\in K} \|y-x\|, \qquad (23)$$

tem uma única solução $y_x \in K$ que é caracterizada pelas desigualdades variacionais $\forall y\in K, \langle y_x-x, y_x-y\rangle \leq 0$.

Prova:

Com efeito, considerando a função $f_\lambda$ do teorema 3.14 para $\lambda=\frac{1}{2}$ e $f(y)=\psi_K(y)$, a determinação de uma solução para (23) é equivalente à determinação de uma solução para $f_{\frac{1}{2}}(x)= \inf_{y\in X} [\psi_K(y)+\|y-x\|^2]$, onde $\psi_K(y)$ corresponde à função indicadora de K

$$\psi_K(y) = \begin{cases} 0 \text{ se } y\in K \\ +\infty \text{ se } y\notin K \end{cases}.$$

Dado que $\psi_K(y)$ satisfaz as hipóteses de teorema anterior, ou seja, é estrita (dado que K é não vazio), convexa (dado que K é convexo) e semi-contínua inferior (dado que K é fechado), aplicando-se o teorema 3.14 conclui-se que existe uma solução $y_x$ (que necessàriamente pertence a K) a qual é caracterizada pela desigualdade variacional $\forall y\in X$ $\langle y_x-x, y_x-y\rangle - \psi_K(y) \leq 0$ [19] $\Rightarrow$ $\forall y\in K, \langle y_x-x, y_x-y\rangle \leq 0$.

◊

**3.16 - Definição**

A aplicação que a cada $x\in X$ faz corresponder o ponto $y_x$, que é solução do problema (23) ($proj_K:X\rightarrow K$ tal que $proj_K(x)=y_x$), designa-se por projector para obtenção da melhor aproximação de x em K, ou simplemesnte, projector sobre K. Por sua vez $y_x$ designa-se por projecção de x sobre K.

**3.17 - Teorema** (de separação)

Considere-se um subconjunto não vazio, convexo e fechado, K, do espaço de Hilbert X. Se x não pertence a K, então existe uma forma linear real contínua $p\in X^*$ e $\varepsilon>0$ tal que

$$\sup_{y\in K} \langle p,y\rangle \leq \langle p,x\rangle - \varepsilon.$$

Prova:

Pelo corolário 3.15 existe um único ponto $y_x=proj_K(x)$ (a projecção de x sobre K), para o qual se verifica a desigualdade variacional

[19] Observe-se que $\psi_K(y_x)=0$.

$$\langle x-y_x, y-y_x\rangle \le 0 \qquad \forall y\in K.$$

Logo, fazendo $u=x-y_x$, vem que

$$\langle u, y-y_x\rangle = \langle u, y+u-x\rangle = \langle u, y-x\rangle + \langle u,u\rangle \le 0 \ \forall y\in K \Leftrightarrow \|u\|^2 \le \langle u, x-y\rangle \ \forall y\in K.$$

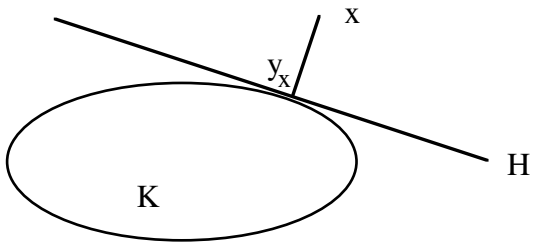
Uma vez que $x\notin K$ e $K$ é fechado, $\|u\|>0$. Logo $\exists \varepsilon=\|u\|^2>0$ tal que $\langle u,x-y\rangle - \varepsilon \ge 0$ $\forall y\in K$, ou seja, $\langle u,y\rangle \le \langle u,x\rangle - \varepsilon$ $\forall y\in K$. Donde, fazendo $p=u$, tal que $p(y)=\langle p,y\rangle$, vem que $\sup_{y\in K} \langle p,y\rangle \le \langle p,x\rangle - \varepsilon$.

◊

Note-se que fazendo $\alpha = \frac{\alpha}{2} + \frac{\alpha}{2} := \langle p,x\rangle - \sup_{y\in K} \langle p,y\rangle$, obtém-se

$$\beta := \langle p,x\rangle - \frac{\alpha}{2} = \sup_{y\in K} \langle p,y\rangle + \frac{\alpha}{2},$$

donde se pode concluir que o hiperplano $H=\{x\in X: \langle p,x\rangle=\beta\}$ separa $x$ de $K$, uma vez que $\langle p,x\rangle > \beta$ e $\langle p,y\rangle \le \beta$ $\forall y\in K$.




Se X é um espaço de dimensão finita, então podemos obter um teorema de separação mais abrangente, sem assumir que K é um conjunto fechado, o que é muito útil, tendo em conta que, na prática, muitas vezes, é difícil verificar se um dado conjunto é ou não fechado. Antes porém convirá que se façam algumas considerações acerca da propriedade da intersecção finita de famílias de conjuntos em espaços topológicos.

**3.18 - Definição**

Uma família (ou colecção) de conjuntos diz-se que tem a propriedade da intersecção finita, quando qualquer subfamília com um número finito de conjuntos tem intersecção não vazia.

Esta propriedade vai permitir a caracterização de espaços topológicos compactos.

**3.19 - Lema**

Um espaço topológico $<S,\tau>$ é compacto sse para toda a família de subconjuntos fechados de S, $F=\{F_\lambda: \lambda\in\Lambda\}$, com a propriedade da intersecção finita, se verifica que $\bigcap_{\lambda\in\Lambda} F_\lambda\neq\emptyset$.

Prova:

Seja $<S,\tau>$ um e.t. compacto e $F=\{F_\lambda: \lambda\in\Lambda\}$ uma família de subconjuntos fechados de S que tem a propriedade da intersecção finita. Suponha-se que $\bigcap_{\lambda\in\Lambda} F_\lambda=\emptyset$. Então $S=S\setminus\bigcap_{\lambda\in\Lambda} F_\lambda=\bigcup_{\lambda\in\Lambda}(S\setminus F_\lambda)$. Consequentemente $\bigcup_{\lambda\in\Lambda}(S\setminus F_\lambda)$ é uma cobertura aberta de S, a qual deve admitir uma subcobertura finita, seja ela, $\{S\setminus F_i: i\in I, |I|<\infty \text{ e } \Lambda\supseteq I\}$. Logo $S=\bigcup_{i\in I}(S\setminus F_i)=S\setminus(\bigcap_{i\in I}F_i)$, o que implica que $\bigcap_{i\in I}F_i=\emptyset$. Facto que contraria a hipótese de *F* gozar da propriedade da intersecção finita.

Reciprocamente, suponha-se que $\bigcap_{\lambda\in\Lambda} F_\lambda\neq\emptyset$, sempre que $F=\{F_\lambda:\lambda\in\Lambda\}$ é uma família de subconjuntos fechados de S que tem a propriedade da intersecção finita. Assuma-se ainda que $<S,\tau>$ não é compacto. Então existe uma cobertura aberta de $<S,\tau>$, $\{A_\lambda: \lambda\in\Lambda\}$, que não admite subcobertura finita. Consequentemente, para qualquer subconjunto finito, I, de $\Lambda$, vem que $S\setminus\bigcup_{i\in I}A_i=\bigcap_{i\in I}(S\setminus A_i)\neq\emptyset$. Logo a família de subconjuntos fechados de S, $\{S\setminus A_\lambda: \lambda\in\Lambda\}$, tem a propriedade da intersecção finita, pelo que $\bigcap_{\lambda\in\Lambda}(S\setminus A_\lambda)=S\setminus(\bigcup_{\lambda\in\Lambda}A_\lambda)\neq\emptyset$, o que contradiz a hipótese de ser uma cobertura.

◊

**3.20 - Teorema** (de separação em sentido lato)

Seja K um subconjunto convexo não vazio de um espaço com produto interno X de dimensão finita. Se $x_0$ não pertence a K, então existe uma forma linear [20] $p\in X^*\setminus\{0\}$ tal que

$$\sup_{y\in K}<p,y> \leq <p,x_0> \qquad (24).$$

Prova:

[20]Dado que X* tem dimensão finita, qualquer forma linear de X* é contínua.

Apesar de K não ser fechado, todos os invólucros convexos de famílias finitas de pontos são subconjuntos de K convexos e fechados, os quais podemos separar de $x_0$, por aplicação do teorema (de separação) 3.17.

A cada $y\in K$ associe-se o subconjunto $F_y$ de X*, $F_y=\{p\in X^*: \|p\|^*=1$ e $\langle p,y\rangle\leq\langle p,x_0\rangle\}$, contido na esfera de raio unitário $S=\{p\in X^*: \|p\|^*\leq 1\}$. Note-se que o conjunto de soluções (de norma unitária) de (24) é a intersecção dos conjuntos $F_y$, $\bigcap_{y\in K} F_y$. Consequentemente, o que temos de mostrar é que esta intersecção é não vazia. Para isso vamos utiliazr o facto de S ser compacto (o que decorre de X ter dimensão finita). Uma vez que os subconjuntos $F_y$ são fechados, basta provar que esta família de conjuntos satisfaz a propriedade da intersecção finita.

Seja a família finita de pontos de X, $y_1, \dots, y_n$, e considere-se o seu invólucro convexo, $C:=\{\sum_{i=1}^{n}\lambda_i y_i: \lambda_i\geq 0, i=1,\dots,n,$ e $\sum_{i=1}^{n}\lambda_i=1\}$. Dado que K é convexo, vem que C está contido em K e, consequentemente, $x_0\notin C$. Por outro lado, C é convexo e fechado (e até compacto). Logo o teorema (de separação) 3.17, implica que exista uma forma linear (a qual pode consider-se como tendo norma unitária) de X* tal que $\sup_{y\in C}\langle p,y\rangle < \langle p,x_0\rangle$.

Dado que os pontos $y_i$, para i=1, … , n, pertencem a C, então $\langle p,y_i\rangle \leq \langle p,x_0\rangle$, donde $p\in F_{y_i}$, para i=1, … , n, ou seja, $\bigcap_{i=1}^{n} F_{y_i}\neq\emptyset$. Como consequência a propriedade da intersecção finita é satisfeita, pelo que $\bigcap_{y\in K} F_y\neq\emptyset$. Logo existe uma forma linear $p\in\bigcap_{y\in K} F_y$ que satisfaz (24).

◊

Embora existam inúmeros corolários dos teoremas de separação, vamos apenas abordar um que se refere à separação de dois subconjuntos disjuntos M e N de um espaço vectorial, começando por notar que M é disjunto de N sse $0\notin$ M-N [21] (i. e, $M\cap N=\emptyset \Leftrightarrow 0\notin$ M-N). Note-se ainda que

$$\sup_{z\in M-N}\langle p,z\rangle = \sup_{x\in M;y\in N}\langle p,x-y\rangle = \sup_{x\in M}\langle p,x\rangle + \sup_{y\in N}\langle p,-y\rangle = \sup_{x\in M}\langle p,x\rangle - \inf_{y\in N}\langle p,y\rangle.$$

**3.21 - Corolário**

Considerem-se dois subconjuntos disjuntos não vazios, M e N, do espaço de Hilbert X.

a) Se M-N é convexo e fechado, então existe uma forma linear contínua $p\in X^*$ e $\varepsilon>0$ tal que $\sup_{x\in M}\langle p,x\rangle \leq \inf_{y\in N}\langle p,y\rangle - \varepsilon$.

b) Se X tem dimensão finita e M-N é convexo, então existe uma forma linear $p\in X^*\backslash\{0\}$ tal que $\sup_{x\in M}\langle p,x\rangle \leq \inf_{y\in N}\langle p,y\rangle$.

---

[21] M-N=$\{z=x-y: x\in M$ e $y\in N\}$.

Prova:

A prova decorre directamente da aplicação do teorema de separação 3.17 e o teorema de separação em sentido lato 3.20, para a obtenção da separação entre 0 e o conjunto M-N, respectivamente, na situação a) e b).

◊

Note-se no entanto que, embora no caso de M e N serem conjuntos convexos se possa concluir que M-N é convexo e, no caso de M ser compacto e N fechado, se possa concluir que M-N é fechado, existem exemplos onde M-N é convexo e M não é convexo, e também situações onde onde M-N é fechado, apesar de nem M nem N serem compactos.

**3.22 Exercício**

Sendo $f:X\rightarrow \mathbf{R}\cup\{+\infty\}$ uma função quase-convexa e sendo K o conjunto de soluções óptimas do problema: $\inf_{x\in X} f(x)$,

a) provar que K é convexo;

b) provar que f é semi-contínua em K.

**3.23 Exercício**

Sendo $f_1, f_2, \ldots, f_p$ funções lipschitzianas de X em **R**, provar que a função f definida pela combinação linear destas funções ($\exists \alpha_1,\ldots,\alpha_p \in R$ tais que $f(x)=\sum_{i=1}^{p}\alpha_i f(x)$) é também uma função lipschitziana.

**3.24 Exercício**

Dados dois subconjuntos M e N de um espaço de Hilbert X, provar cada uma das seguintes afirmações:

a) Se M e N são compactos, então M+N é compacto.

b) Se M é compacto e N é fechado então M+N é fechado.

c) O facto de M-N ser convexo não implica que M seja convexo.

d) O facto de M-N ser fechado não implica que M e N sejam compactos.