04/03

Convexidade

A referência para essa parte é "Convex Analysis" do T. Rockafellar.

Variedade afim

No que se segue, V, será sempre um espaço vetorial real de dimensão finita.

Definição 1 *Se x e y são vetores distintos de V, a reta que passa por x e y é $L = \{rx + (1-r)\,y : r \in \mathbb{R}\}$.*

Definição 2 *O conjunto $M \subset V$ é variedade afim se $(1-r)\,x + ry \in M$ sempre que $x, y \in M, r \in \mathbb{R}$.*

Em outras palavras M é variedade afim se contém toda reta que passa por cada par de pontos distintos de M.

Lema 1 *Se M é afim então se $x_i \in M, 1 \le i \le n$ e $\sum_{i=1}^{n} \lambda_i = 1$ então $\sum_{i=1}^{n} \lambda_i x_i \in M$.*

Demonstração: Para $n = 2$ a conclusão vale pela definição de variedade afim. Suponhamos agora que $\sum_{i=1}^{n+1} \lambda_i = 1$ e $x_i \in M, i \le n+1$. Sem perda de generalidade, $\lambda_{n+1} \ne 1$. Então $\sum_{i=1}^{n} \lambda_i = 1 - \lambda_{n+1} \ne 0$ e $\frac{1}{\sum_{i=1}^{n} \lambda_i} \sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} x_i \in M$ pela hipótese de indução. Então

$$\sum_{i=1}^{n+1} \lambda_i x_i = (1 - \lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} x_i + \lambda_{n+1} x_{n+1} \in M.$$

Comentário 1 *O conjunto vazio e o espaço vetorial V são variedades afins. Se $M = \{m\}$ então M é afim.*

Exemplo 1 *Todo subespaço vetorial de V é uma variedade afim de V.*

Teorema 1 *Toda variedade afim é a translação de um subespaço vetorial.*

Demonstração: Seja F subespaço vetorial de V. Então $a + F$ é variedade afim: se $x, y \in a + F$ e r é real, temos $(1-r)\,x + ry = (1-r)\,(a+f) + r\,(a+g) = a + (1-r)\,f + rg \in a + F$. Recíprocamente, suponhamos M variedade afim de V. Seja $a \in M$ e definamos $F = M - a$. Vamos verificar que F é um subespaço vetorial. Sejam $x, y \in F$.

Então $\lambda x + a = \lambda\,(x+a) + (1-\lambda)\,a \in M$ implica $\lambda x \in F$ para todo real λ. Agora $\frac{1}{2}\,(x+a) + \frac{1}{2}\,(y+a) = \frac{x+y}{2} + a \in M$ e então $\frac{x+y}{2} \in F$ e então $x + y = 2\frac{x+y}{2} \in F$. Portanto F é subespaço vetorial.

Comentário 2 *O espaço vetorial associado à variedade afim M é $M - M$.*

Para verificar isso note que $a+F = M$. Então $F = F-F = (a + F)-(a + F) = M - M$.

Comentário 3 *A intersercção de variedades afins é uma variedade afim. Assim para $S \subset V$ exist* $\operatorname{aff} S$ *a menor variedade afim que contém S.*

Lema 2 $\operatorname{aff} S = \{\sum_{i=1}^{m} \lambda_i x_i : m \geq 1, \sum_{i=1}^{m} \lambda_i = 1, x_i \in S\}$.

Demonstração: Seja $M = \{\sum_{i=1}^{m} \lambda_i x_i : m \geq 1, \sum_{i=1}^{m} \lambda_i = 1, x_i \in S\}$. Então $M \supset S$. Sejam $x, y \in M$, r real. Então

$$rx+(1-r)\,y = r\left(\sum_{i=1}^{m} \lambda_i x_i\right)+(1-r)\left(\sum_{k=1}^{p} \mu_k x'_k\right), \sum_{i=1}^{m} \lambda_i = 1 = \sum_{k=1}^{p} \mu_k, x_i, x'_i \in S.$$

Logo $rx+(1-r)\,y = \sum_i r\lambda_i x_i + \sum_k (1-r)\,\mu_k x'_k \in M$ pois a soma dos coeficientes é 1. Com isso obtemos $M \supset \operatorname{aff} S$. Por outro lado $\operatorname{aff} S \supset M$ terminando a demonstração.

Definição 3 *A dimensão de uma variedade afim M é a dimensão do subespaço vetorial associado, $L(M) = M - M$.*

Definição 4 *Os vetores $b_0, \ldots, b_m$ são afim independentes se $b_1 - b_0, \ldots, b_m - b_0$ são linearmente independentes.*

Nesse caso $\operatorname{con}\{b_0, \ldots, b_m\}$ *é, por definição, um simplexo de dimensão m. E cada b_i é vértice do simplexo. O ponto $\frac{1}{m+1}(b_0 + b_1 + \ldots + b_m)$ é o baricentro do simplexo.*

Definição 5 *Um hiperplano no $\mathbb{R}^n$ é uma variedade afim de dimensão $n-1$.*

Teorema 2 *Para β real e $b \in \mathbb{R}^n \setminus \{0\}$, $H = \{x \in \mathbb{R}^n : \langle x, b\rangle = \beta\}$ é um hiperplano. Além disso todo hiperplano do $\mathbb{R}^n$ possui uma tal representação com b, β únicos a menos de um múltiplo em comum.*

Demonstração: Seja H um hiperplano no $\mathbb{R}^n$. Seja $F = H - H$ o subespaço vetorial associado. Se $\{b_1, \ldots, b_{n-1}\}$ é uma base ortogonal de F e completando temos $\{b_1, \ldots, b_n\}$ base ortogonal do $\mathbb{R}^n$. É claro que $F = \{x : \langle x, b_n\rangle = 0\}$. Agora se $H = a + F$ e $\beta = \langle a, b_n\rangle$ temos se $b := b_n$

$$H = \{x \in \mathbb{R}^n : \langle x, b\rangle = \beta\}. \tag{*}$$

Recíprocamente se H está definido por (*), $b \neq 0$ seja $a \in \mathbb{R}^n$ tal que $\langle a, b\rangle = \beta$. Então se $\langle x, b\rangle = \beta = \langle a, b\rangle$ temos $x - a \perp b$. O subespaço $F = b^{\perp}$ tem dimensão $n - 1$. Suponhamos $H = \{x : \langle x, b'\rangle = \beta'\}$ outra representação do hiperplano H. Mas então com F é ortogonal à b' temos $b' \in [b]$ ou seja $b' = \mu b$, $\mu \neq 0$. E $\langle x, b'\rangle = \mu\langle x, b\rangle = \mu\beta = \beta'$.

Comentário 4 *b é uma normal ao hiperplano H.*

Teorema 3 *Seja B matriz $m \times n$ e $b \in \mathbb{R}^m$. Então $M = \{x \in \mathbb{R}^n : Bx = b\}$ é variedade afim. Além disso toda variedade afim é dessa forma.*

Demonstração: Sejam $x, y \in M$, r real. $B\left(rx + (1-r)\,y\right) = rBx + (1-r)\,By = rb + (1-r)\,b = b$. Por outro lado se M é afim própria seja $L = M - M$. Seja $\{b_1, \ldots, b_m\}$ base de $L^{\perp}$. Então

$$L = \left(L^{\perp}\right)^{\perp} = \{x : x \perp b_i, i \leq m\} = \{x : \langle x, b_i \rangle = 0, i \leq m\} = \{x : Bx = 0\}.$$

Sendo B matriz $m \times n$ com linhas $b_1, \ldots, b_m$. Se $M = L + a = \{x : B\,(x - a) = 0\} = \{x : Bx = b\}$ sendo $b = Ba$.

Corolário 1 *Toda variedade afim de V é a intersecção de uma família finita de hiperplanos.*

Definição 6 *O conjunto $C \subset V$ é convexo se para todos c', c'' em C e $0 < r < 1$, $rc' + (1-r)\,c'' \in C$.*

Todo subespaço vetorial é convexo e toda variedade afim é convexa. Se $f \in V^* \setminus \{0\}$ então

$$\begin{array}{l}
\{x \in V : f(x) = \gamma\}, \text{ hiperplano} \\
\{x \in V : f(x) < \gamma\}, \text{ semi–espaço aberto} \\
\{x \in V : f(x) > \gamma\}, \text{ semi–espaço aberto} \\
\{x \in V : f(x) \leq \gamma\} \text{ semi–espaço fechado} \\
\{x \in V : f(x) \geq \gamma\}, \text{ semi–espaço fechado}
\end{array}$$

são convexos.

Comentário 5 *Note que as noções acima de aberto, fechado se referem ao aspecto geométrico e ao topológico pois f é um funcional linear contínuo.*

Lema 3 *A intersecção de convexos é um convexo.*

Demonstração: Seja $C = \cap_{i \in I} C_i$, cada C_i convexo de V. Se $x, y \in C$ e $r \in (0,1)$, de $x, y \in C_i$ vem $rx + (1-r)\,y \in C_i$ para todo i e logo $rx + (1-r)\,y \in C$.

Comentário 6 *O conjunto de soluções de um sistema de igualdades/desigualdades lineares é um conjunto convexo.*

Definição 7 *Um convexo poliédrico é a intersecção finita de semi–espaços fechados do espaço V.*

Definição 8 *Um politopo é a envoltória convexa de um conjunto finito de pontos.*

Definição 9 *A dimensão de um conjunto convexo C é a dimensão da variedade afim gerada por C.*

Se $x_1, \ldots, x_p$ são vetores e $r_i \geq 0, \sum_{i=1}^p r_i = 1$ então $\sum_{i=1}^p r_i x_i$ é uma combinação convexa de $x_1, \ldots, x_p$. Os conjuntos convexos, por definição, são fechados por combinações convexas de dois elementos ($p = 2$).

Lema 4 *Suponha C convexo e $x_1, \ldots, x_p \in C$. Então a combinação convexa $\sum_{i=1}^p r_i x_i \in C$.*

Demonstração: Seja C convexo. Para $p = 2$ o resultado acima vale por definição. Se valer para $p \geq 2$. Suponhamos $x_1, \ldots, x_p, x_{p+1} \in C$ e $r_i \geq 0, \sum_{i=1}^{p+1} r_i = 1$. Seja $x = \sum_{i=1}^{p+1} r_i x_i$ uma combinação convexa com $p+1$ termos. Note que pela hipótese de indução, $y = \frac{1}{\sum_{i=1}^p r_i} \sum_{i=1}^p r_i x_i \in C$. Mas então

$$x = \left(\sum_{i=1}^p r_i \right) y + \left(1 - \sum_{i=1}^p r_i \right) x_{p+1} = \sum_{i=1}^{p+1} r_i x_i \in C.$$

Portanto o resultado vale para $p+1$ terminando a indução e a demonstração.

Definição 10 *Se $S \subset V$, a envoltória convexa de S, $\operatorname{con} S$, é a intersecção dos subconjuntos convexos de V que contém S.*

Comentário 7 *Temos*

$$\operatorname{con} S = \left\{ \sum_{i=1}^p r_i s_i : p \geq 1, r_i \geq 0, s_i \in S, \sum_{i=1}^p r_i = 1 \right\}.$$

Exemplo 2 *Suponhamos $S := \{b_1, \ldots, b_p\} \subset V$. Então*

$$\operatorname{con} S = \left\{ \sum_{i=1}^p \lambda_i b_i : \lambda_i \geq 0, \sum_{i=1}^p \lambda_i = 1 \right\}.$$

Com efeito as combinações com mais de p elementos facilmente se reduz a uma combinação com p elementos: basta somar os coeficientes de cada b_i repetido.

Teorema 4 (Carathéodory) *Seja $S \subset \mathbb{R}^n$. Então*

$$\operatorname{con} S = \left\{ \sum_{i=1}^{n+1} r_i s_i : s_i \in S, r_i \geq 0, \sum_i r_i = 1 \right\}.$$

Demonstração: Seja $x = \sum_{i=1}^{m} r_i s_i$ sendo $m > n+1$, uma combinação convexa, $s_i \in S$. Temos $s_2 - s_1, \ldots, s_m - s_1$ linearmente dependentes pois $m-1 > n$. Então

$$\lambda_1 (s_1 - s_m) + \ldots + \lambda_{m-1} (s_{m-1} - s_m) = 0, \exists i \leq m-1, \lambda_i \neq 0.$$

Seja $\lambda_m = -\sum_{j=1}^{m-1} \lambda_j$. Então $\lambda_1 s_1 + \ldots + \lambda_{m-1} s_{m-1} + \lambda_m s_m = 0$ e $\sum_{i=1}^{m} \lambda_m = 0$. Note que pelo menos algum $\lambda_i > 0$. Seja $t \geq 0$ e note que

$$x = \sum_{i=1}^{m} r_i s_i - t \sum_{i=1}^{m} \lambda_i s_i = \sum_{i=1}^{m} (r_i - t\lambda_i) s_i.$$

Agora t deve ser tal que $r_i - t\lambda_i \geq 0$ para todo i. Seja $t = \min \left\{ \frac{r_i}{\lambda_i} : \lambda_i > 0 \right\}$. Temos que $\sum_{i=1}^{m} (r_i - t\lambda_i) = \sum_i r_i - t \sum_i \lambda_i = 1$. Finalmente para i tal que $t = \frac{r_i}{\lambda_i}$ temos $r_i - t\lambda_i = 0$.

Logo escrevemos x como uma combinação linear de no máximo $m-1$ termos. Esse processo pode ser repetido até m alcançar $n+1$ e termina aí.

Corolário 2 *Se $K \subset V$ for compacto e $\dim V < \infty$ então $\operatorname{con} K$ é compacto.*

Demonstração: Seja $n = \dim V$. Seja $\Delta = \left\{ \lambda \in [0,1]^{n+1} : \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$. A função $f : \Delta \times K^{n+1} \to V$,

$$(\lambda_1, \ldots, \lambda_{n+1}, k_1, \ldots, k_{n+1}) \to \sum_{i=1}^{n+1} \lambda_i k_i$$

é contínua. Portanto tem imagem compacta. Pelo teorema de Carathéodory $f\left(\Delta \times K^{n+1}\right) = \operatorname{con} K$. Logo $\operatorname{con} K$ é compacto.

06/03

Teorema 5 *O fecho de um conjunto convexo é convexo.*

Demonstração: Seja $S \subset V$ convexo no espaço normado V. Sejam $x, y \in \overline{S}$ e $0 < r < 1$. Existem seqüências $x_n \in S$, $y_n \in S$ e $\lim_n x_n = x, \lim_n y_n = y$. Então $rx_n + (1-r) y_n \in S$ e converge para $rx + (1-r) y$. Portanto $rx + (1-r) y \in \overline{S}$.

Teorema 6 *A soma vetorial de convexos é convexa.*

Demonstração: Se C_1 e C_2 são convexos, $C_1 + C_2$ é convexo pois $x = c_1 + c_2$ e $y = c_1' + c_2'$ então

$$rx + (1-r)\,y = rc_1 + (1-r)\,c_1' + rc_2 + (1-r)\,c_2' \in C_1 + C_2.$$

Lema 5 *Se C for convexo e $r_1, r_2 \geq 0$ então $(r_1 + r_2)\,C = r_1 C + r_2 C$.*

Demonstração: É imediato que $(r_1 + r_2)\,C \subset r_1 C + r_2 C$. Seja agora $y = r_1 c_1 + r_2 c_2$, $c_1, c_2 \in C$, Então $\frac{y}{r_1+r_2} = \frac{r_1}{r_1+r_2} c_1 + \frac{r_2}{r_1+r_2} c_2 \in C$. Logo $y \in (r_1 + r_2)\,C$.

Definição 11 *Para C convexo,*

i) $L = \cup_{n=1}^{\infty} n\,(C - C) = \cup_{r \geq 0} r\,(C - C)$ *é um espaço vetorial..*

ii) $\operatorname{aff} C = x_0 + L$ *para* $x^0 \in C$.

Demonstração: (i) Seja $z = r\,(c - c')$, $r > 0, c, c'$ em C. Para $n > r$,

$$\frac{r}{n}\,(c - c') = \left(\frac{r}{n} c + \left(1 - \frac{r}{n}\right) c\right) - \left(\frac{r}{n} c' + \left(1 - \frac{r}{n}\right) c\right) \in C - C \implies z \in n\,(C - C)\,.$$

Portanto $\cup_{r \geq 0} r\,(C - C) = \cup_{n=1}^{\infty} n\,(C - C)$. Seja $x \in L$. Então $-x \in L$ pela simetria de $C - C$. Portanto para verificarmos que L é fechado por produto por escalar λ podemos supor $\lambda > 0$. Mas então $\lambda x \in \cup_{r \geq 0} \lambda r\,(C - C) = \cup_{r \geq 0} r\,(C - C) = L$. Sejam $x, y \in L$ e r real. Temos $x = n\,(c - c')$, $y = m\,(c'' - c''')$ sendo $c, c', c'', c''' \in C$. Então

$$x + y = nc + mc'' - (nc' + mc''') = (n + m)\,(c'''' - c''''') \in L.$$

(ii) Se $c \in C$ temos $c = x^0 + (c - x^0) \in x^0 + L$. Portanto $C \subset x^0 + L$ e então $\operatorname{aff} C \subset x^0 + L$. Seja $H + x^0 = \operatorname{aff} C$. Agora se $c, c' \in C$, $c - c' \in H$ e logo $L \subset H$. Portanto $L = H$.

Interior relativo de um convexo[1]

Um intervalo na reta tem interior não–vazio mas no plano o interior é vazio. Uma propriedade importante dos convexos de dimensão finita, é que sempre tem interior relativo não–vazio. Seja $B = \{x \in \mathbb{R}^n : |x| < 1\}$.

Definição 12 *O interior relativo do convexo C, denotado* $\operatorname{ri} C$, *é o interior de C como subespaço topológico de* $\operatorname{aff} C$: *$z \in \operatorname{ri} C$ se existir $r > 0$ tal que $(z + rB) \cap \operatorname{aff} C \subset C$.*

[1] Ou tudo que você sempre quis saber sobre o interior relativo e não teve coragem de perguntar.

Definição 13 *A fronteira relativa de* C *é* $\overline{C} \setminus \operatorname{ri} C$.

Exemplo 3 *Seja* $S = \{(x, y) \geq 0 : x + y \leq 1\}$. *Então se* $L = [0, 1] \times \{0\}$ *temos* $\operatorname{ri} L = \{(x, 0) : 0 < x < 1\}$. *E* $L(S) =$ *"o plano" e* $\operatorname{ri} S = \{(x, y) >> 0, x + y < 1\}$.

Lema 6 *Seja* $\{v_1, \ldots, v_n\}$ *base de* V. *Então* $T : V \to \mathbb{R}^n$ *tal que*

$$T(z) = (\lambda_1, \ldots, \lambda_n), z = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

é contínua e tem inversa contínua.

Demonstração: Seja $\phi(\lambda) = \sum_{i=1}^n \lambda_i v_i$. Note que $\phi(\lambda) = 0 \iff \lambda = 0$. É imediato que $\phi = T^{-1}$ e é contínua. Seja

$$\delta = \min \left\{ \left| \sum_{i=1}^n \lambda_i v_i \right| : |\lambda|_1 = 1 \right\}, |\lambda|_1 := \sum_{i=1}^n |\lambda_i|.$$

Pela compacidade de $\{\lambda \in \mathbb{R}^n : \sum_i |\lambda_i| = 1\}$ o mínimo existe e $\delta > 0$. Portanto se $\lambda \neq 0$, $\left|\frac{\lambda}{|\lambda|_1}\right|_1 = 1$ e

$$\left| \phi\left(\frac{\lambda}{|\lambda|_1}\right) \right| \geq \delta \implies |\phi(\lambda)| \geq \delta |\lambda|_1.$$

Seja $z \in V$ e $\lambda = T(z)$. Então $|z| = |\phi(\lambda)| \geq \delta |\lambda|_1 = \delta |T(z)|$. Demonstrando a continuidade de T.

Teorema 7 *Seja* $S = \operatorname{con}\{b_0, b_1, \ldots, b_n\}$ *um simplexo de dimensão* n. *Então* $\operatorname{ri} S \neq \emptyset$.

Demonstração: Seja $L = L(S) = [b_1 - b_0, \ldots, b_n - b_0]$. Definamos $v_i = b_i - b_0$ e $b^* = \frac{1}{n+1} \sum_{i=0}^n b_i$ o baricentro do simplexo. Para $z = \sum_{i=1}^n \lambda_i v_i$,

$$b^* + z = \frac{1}{n+1} \sum_{i=0}^n b_i + \sum_{i=1}^n \lambda_i (b_i - b_0) = \left(\frac{1}{n+1} - \sum_{i=1}^n \lambda_i \right) b_0 + \sum_{i=1}^n (\frac{1}{n+1} + \lambda_i) b_i.$$

se

$$\frac{1}{n+1} + \lambda_i \geq 0, 1 \leq i \leq n; \qquad (*)$$
$$\frac{1}{n+1} - \sum_{i=1}^n \lambda_i \geq 0,$$

temos $b^* + z \in S$ pois $\frac{n}{n+1} + \sum_{i=1}^n \lambda_i + \frac{1}{n+1} - \sum_{i=1}^n \lambda_i = 1$. O conjunto

$$\Delta = \left\{ \lambda \in \mathbb{R}^n_{++} : \frac{1}{n+1} - \sum_{i=1}^n \lambda_i > 0 \right\}$$

é aberto e então $b^* \in \operatorname{ri} S$ pois $T^{-1}(\Delta)$ é aberto tal que $(b^* + T^{-1}(\Delta)) \cap \operatorname{aff} S \subset S$.

Teorema 8 *Seja $C \subset \mathbb{R}^n$ convexo não–vazio. Então* $\operatorname{ri} C \neq \emptyset$.

Demonstração: Seja $S = \operatorname{con}\{b_0, b_1, \ldots, b_m\}$ um simplexo de C com a dimensão m maior possível. Pelo teorema anterior $\operatorname{ri} S \neq \emptyset$. Se $C \setminus \operatorname{aff} S$ for não vazio então existe $b_{m+1} \in C$ com $b_{m+1} - b_0 \notin L(S)$. Contradição com a escolha de m. Logo $\operatorname{aff} C \subset \operatorname{aff} S$ e então vale a igualdade. Mas então $\operatorname{ri} C \neq \emptyset$ pois $S \subset C$ e $\operatorname{aff} S = \operatorname{aff} C$.

Teorema 9 *Seja C convexo no $\mathbb{R}^n$. Se $x \in \operatorname{ri} C$ e $y \in \overline{C}$ então se $0 \leq r < 1$, $(1-r)\,x + ry \in \operatorname{ri} C \subset C$.*

Demonstração: Vou fazer somente o caso $L(C) = \mathbb{R}^n$. Nesse caso temos $x \in \operatorname{int} C$. Seja $\delta > 0$ tal que $x + \delta B \subset C$. Seja $0 < r < 1$. Para $\epsilon > 0$,

$$\begin{aligned} (1-r)\,x + ry + \epsilon B \subset (1-r)\,x + r\,(C + \epsilon B) = \\ (1-r)\left[x + \epsilon\,(1+r)\,(1-r)^{-1} B\right] + rC \subset (1-r)\,C + rC = C \end{aligned}$$

se $\epsilon > 0$ for tal que $\epsilon\,(1+r)\,(1-r)^{-1} < \delta$.

Corolário 3 $\operatorname{ri} C$ *é convexo.*

Corolário 4 1. *O fecho de* $\operatorname{ri} C$ *é* $\overline{C}$.

2. *O interior relativo de* $\overline{C}$ *é* $\operatorname{ri} C$.

Demonstração: (1) É imediato do teorema 9. (2) Seja $z \in \operatorname{ri}\overline{C}$. E $x \in \operatorname{ri} C$. Para $\mu > 1$, suficientemente próximo de 1, $y = (1-\mu)\,x + \mu z = z - (\mu - 1)\,(x - z)$ ainda pertence a $\operatorname{ri}\overline{C} \subset \overline{C}$. Então se $\lambda = \mu^{-1}$,

$$z = \frac{1}{\mu} y - \left(\frac{1}{\mu} - 1\right) x = \lambda y + (1-\lambda)\,x \in \operatorname{ri} C,$$

pelo teorema 9.

Corolário 5 *Sejam C_1 e C_2 convexos do $\mathbb{R}^n$. Então*

$$\overline{C_1} = \overline{C_2} \iff \operatorname{ri} C_1 = \operatorname{ri} C_2.$$

Equivalentemente, $\operatorname{ri} C_1 \subset C_2 \subset \overline{C_1}$.

Demonstração: Suponhamos $\overline{C_1} = \overline{C_2}$. Pelo corolário anterior, $\operatorname{ri} C_1 = \operatorname{ri}\overline{C_1} = \operatorname{ri}\overline{C_2} = \operatorname{ri} C_2$.

Corolário 6 *Se C é convexo, todo aberto que intersecta $\overline{C}$ intersecta* $\operatorname{ri} C$.

Demonstração: Fixe $x^0 \in \mathrm{ri}\ C$. Seja U aberto tal que $U \cap \overline{C} \neq \emptyset$. Existe então $c \in U \cap C$. Então para $r \in (0,1)$ suficientemente próximo de 1, $(1-r)\,x^0 + rc \in U$. Mas $(1-r)\,x^0 + rc \in \mathrm{ri}\ C$ terminando a demonstração.

Corolário 7 *Suponhamos que C_1 não–vazio é um subconjunto convexo da fronteira relativa de C_2. Ou seja $C_1 \subset \overline{C_2} \setminus \mathrm{ri}\, C_2$. Então* $\dim C_1 < \dim C_2$.

Demonstração: Se $\dim C_1 = \dim C_2$ então de $\mathrm{aff}\, C_1 \subset \mathrm{aff}\, C_2$ vem $\mathrm{aff}\, C_1 = \mathrm{aff}\, C_2$. Então se $x \in C_1$ é tal que $(x+U) \cap \mathrm{aff}\, C_1 \subset C_1$ vem

$$(x+U) \cap \mathrm{aff}\, C_2 \subset \overline{C_2} \implies x \in \mathrm{ri}\, C_2$$

em contradição com a hipótese.

Comentário 8 *O próximo teorema simplifica a verificação de que um ponto está no interior relativo.*

Teorema 10 *Seja C convexo. Então $z \in \mathrm{ri}\ C$ se e somente se para todo $x \in C$ existe $\mu > 1$ tal que $(1-\mu)\,x + \mu z \in C$.*

Demonstração: Caso $z \in \mathrm{ri}\ C$. Seja $\epsilon > 0$ tal que $(z + \epsilon B) \cap \mathrm{aff}\, C \subset C$. Então $z - t\,(x-z) \in C$ se $0 < t < \frac{\epsilon}{|x-z|}$. Logo se $\mu = 1 + t$ temos $y := \mu z + (1-\mu)\,x = z - t\,(x-z) \in C$. Recíproca: Seja x no interior relativo de C e $\mu > 1$ tal que $y = (1-\mu)\,x + \mu z \in C$. Então definindo $\lambda = \frac{1}{\mu} \in (0,1)$, $z = (1-\lambda)\,x + \lambda y$. Logo pelo teorema 9, $z \in \mathrm{ri}\ C$.

Corolário 8 *Para C convexo. Então $z \in \mathrm{int}\, C$ se e somente se para todo vetor y existe $\epsilon > 0$, $z + \epsilon y \in C$.*

Demonstração: Pelo teorema anterior obtemos que $z \in \mathrm{ri}\ C$. Mas pela hipótese, $L\,(C) \supset \{e_1, e_2, \ldots, e_n\}$ e logo $\mathrm{aff}\, C = \mathbb{R}^n$.

08/03

Teorema 11 *Seja C_i convexo, $i \in I$. Suponhamos que $\cap_{i \in I}\, \mathrm{ri}\ C_i \neq \emptyset$. Então:*

1. *O fecho de $\cap_{i \in I} C_i$ é igual a $\cap_{i \in I} \overline{C_i}$;*
2. *Se I for finito, o interior relativo de $\cap_i C_{i \in I}$ é $\cap_{i \in i}\, \mathrm{ri}\ C_i$.*

Demonstração: Fixemos $a \in \cap_{i\in I} \text{ri } C_i$. Se $y \in \cap_{i\in I}\overline{C_i}$, $(1-r)\,a + ry \in \text{ri } C_i$ para todo i se $r \in (0,1)$. E se r tende a 1 o limite é y. Portanto

$$\cap_{i\in I}\overline{C_i} \subset \overline{\cap_{i\in i} \operatorname{ri} C_i} \subset \overline{\cap_{i\in i} C_i} \subset \cap_{i\in I}\overline{C_i}.$$

Então vale (1). E $\cap_{i\in I} \text{ri } C_i$ e $\cap_{i\in I} C_i$ tem o mesmo fecho. Eles tem então o mesmo interior relativo:

$$\operatorname{ri} \cap_{i\in i} C_i \subset \cap_{i\in I} \operatorname{ri} C_i.$$

Seja $z \in \cap_{i\in I} \operatorname{ri} C_i$. Cada segmento de reta em $\cap_{i\in I} C_i$ com um extremo z pode ser prolongado. A intersecção finita desses prolongamentos ainda é um prolongamento. Portanto $z \in \operatorname{ri} \cap_{i\in i} C_i$.

Corolário 9 *Seja C convexo e M variedade afim tal que $M \cap \text{ri } C \neq \emptyset$. Então*

$$\operatorname{ri}(M \cap C) = M \cap \operatorname{ri} C,$$
$$\overline{M \cap C} = M \cap \overline{C}.$$

Demonstração: $\operatorname{ri}(M) = M$ e M é fechada.

Corolário 10 *Seja C_1 convexo. E $C_2 \subset \overline{C_1}$ convexo que não está contido na fronteira relativa de C_1. Então $\operatorname{ri}(C_2) \subset \operatorname{ri}(C_1)$.*

Demonstração: Se $\operatorname{ri}(C_2) \cap \operatorname{ri}(C_1) = \emptyset$, $\operatorname{ri}(C_2) \subset \overline{C_1} \setminus \operatorname{ri}(C_1)$. E C_2 estaria contido na fronteira relativa. Então $\operatorname{ri}(C_2) \cap \operatorname{ri}(C_1) \neq \emptyset$,

$$\operatorname{ri}(C_2) \cap \operatorname{ri}(C_1) = \operatorname{ri}(C_2) \cap \operatorname{ri}(\overline{C_1}) = \operatorname{ri}(C_2 \cap \overline{C_1})) = \operatorname{ri}(C_2)$$
$$\implies \operatorname{ri}(C_2) \subset \operatorname{ri}(C_1).$$

Proposição 1 $\operatorname{ri}(C_1 + C_2) = \operatorname{ri}(C_1) + \operatorname{ri}(C_2)$.

Demonstração: Seja $a_i \in \operatorname{ri}(C_i)$, $i = 1, 2$. Para $z = c_1 + c_2 \in C_1 + C_2$, seja $\mu_i > 1$ tal que $(1-\mu_i)\,c_i + \mu_i a_i \in C_i, i = 1, 2$. Seja $\mu = \min\{\mu_1, \mu_2\}$. Então

$$(1-\mu)\,c_i + \mu a_i \in C_i, i = 1, 2 \implies (1-\mu)\,z + \mu\,(a_1 + a_2) \in C_1 + C_2$$
$$\implies a_1 + a_2 \in \operatorname{ri}(C_1 + C_2).$$

Ou seja vale $\supset$. Para a inclusão reversa,

$$\overline{\operatorname{ri}(C_1) + \operatorname{ri}(C_2)} \supset \overline{\operatorname{ri}(C_1)} + \overline{\operatorname{ri}(C_2)} = \overline{C_1} + \overline{C_2} \supset C_1 + C_2 \supset \operatorname{ri}(C_1) + \operatorname{ri}(C_2)$$

e então $C_1 + C_2$ e $\operatorname{ri}(C_1) + \operatorname{ri}(C_2)$ tem o mesmo fecho e portanto o mesmo interior relativo. Logo $\operatorname{ri}(C_1 + C_2) = \operatorname{ri}(\operatorname{ri}(C_1) + \operatorname{ri}(C_2)) \subset \operatorname{ri}(C_1) + \operatorname{ri}(C_2)$.

Separação de convexos

Definição 14 *Sejam C_1 e C_2 convexos não–vazios.*

a) *O hiperplano H separa C_1 e C_2 se C_1 está num semi–espaço fechado de H e C_2 no outro.*

b) *H separa propriamente se $(C_1 \cup C_2) \setminus H \neq \emptyset$.*

c) *Separa C_1 e C_2 fortemente se existe $\epsilon > 0$ tal que $C_1 + \epsilon B$ está num semi–espaço aberto de H e $C_2 + \epsilon B$ está contido no outro.*

d) *A separação é estrita se C_1 e C_2 estão em semi–espaços abertos distintos.*

Em termos analíticos, existe um funcional linear não–nulo $f(z) = \langle z, b \rangle$ e um escalar β tais que

$$C_1 \subset \{x : f(x) \leq \beta\},$$
$$C_2 \subset \{x : f(x) \geq \beta\}.$$

Note que trocando f por $-f$ trocamos o lado de separação.

Teorema 12 *Sejam C_1 e C_2 não–vazios do $\mathbb{R}^n$.*

Separação própria *Existe um hiperplano separando–os propriamente se existir vetor b tal que*

$$\inf\{\langle x, b \rangle : x \in C_1\} \geq \sup\{\langle x, b \rangle : x \in C_2\};$$
$$\sup\{\langle x, b \rangle : x \in C_1\} > \inf\{\langle x, b \rangle : x \in C_2\}.$$

Separação forte

$$\inf\{\langle x, b \rangle : x \in C_1\} > \sup\{\langle x, b \rangle : x \in C_2\}.$$

Demonstração: Para a separação própria: Seja β entre $\sup\{\langle x, b \rangle : x \in C_2\}$ e $\inf\{\langle x, b \rangle : x \in C_1\}$. Então β é finito, $b \neq 0$ e $H = \{x : \langle x, b \rangle = \beta\}$ é o hiperplano que separa propriamente. Para a separação forte sejam β e $\delta > 0$ tais que

$$\inf\{\langle x, b \rangle : x \in C_1\} - \delta > \beta > \delta + \sup\{\langle x, b \rangle : x \in C_2\}.$$

Seja $0 < \epsilon < \frac{\delta}{|b|}$. Para $x \in C_1 + \epsilon B$ e $y \in B$ temos

$$\langle x, b \rangle = \langle c_1 + \epsilon y, b \rangle \geq \inf\{\langle x, b \rangle : x \in C_1\} - \epsilon |b| > \inf\{\langle x, b \rangle : x \in C_1\} - \delta > \beta.$$

Analogamente,

$$\beta > \delta + \sup\{\langle x, b \rangle : x \in C_2\} \geq \langle x, C_2 + \epsilon B \rangle.$$

Lema 7 *Seja $H = \{x : \langle x, b\rangle = \beta\}$ um hiperplano. E C convexo tal que*

$$C \cap \{x : \langle x, b\rangle < \beta\} \neq \emptyset \text{ e}$$
$$C \cap \{x : \langle x, b\rangle > \beta\} \neq \emptyset.$$

Então $C \cap H \neq \emptyset$.

Demonstração: Sejam $x, y \in C$ tais que $\langle x, b\rangle < \beta < \langle y, b\rangle$. Então para $t = \frac{\beta - \langle x,b\rangle}{\langle y-x,b\rangle} \in (0,1)$, $\langle (1-t)\,x + ty, b\rangle = \beta$.

Teorema 13 *Seja $C \subset \mathbb{R}^n$ convexo não–vazio e relativamente aberto: $C = \operatorname{ri} C$. Seja $M \subset \mathbb{R}^n$ variedade afim disjunta de C. Existe então um hiperplano $H \supset M$ e tal que C está contido num dos semi–espaços abertos de H.*

Demonstração: Se M for um hiperplano então pelo lema anterior, C está contido num dos semi–espaços abertos de M. Se M não for hiperplano. Vamos obter uma variedade afim M' com dimensão maior do que a de M e ainda disjunta de C. Por meio de uma translação (de M e C) podemos supor $0 \in M$. Então $C - M \supset C$ e $0 \notin C - M$. Temos $\dim M^{\perp} \geq 2$ pois M não é hiperplano. Seja P subespaço vetorial de $M^{\perp}$, $\dim P = 2$. Seja $C' = P \cap (C - M)$. Se $C' \neq \emptyset$, C' é aberto em P pois de $\operatorname{ri}(C - M) = \operatorname{ri} C - \operatorname{ri} M = C - M$ vem $P \cap \operatorname{ri}(C - M) \neq \emptyset$ e pelo cor. 9,

$$\operatorname{ri} C' = P \cap \operatorname{ri}(C - M) = P \cap (C - M) = C'.$$

E temos $0 \notin C'$. Queremos encontrar um subespaço uni–dimensional $L \subset P$, $L \cap C' = \emptyset$. Nesse caso $M' = M + L$ é um subespaço com dimensão maior do que a de M e que não intersecta C. Se C' for vazio ou um ponto existe reta L que não intersecta C'. Se $\operatorname{aff} C'$ for uma reta que não passa pela origem escolhemos L paralela a $\operatorname{aff} C'$ passando pela origem. Se $\operatorname{aff} C'$ for uma reta passando pela origem, tomamos L perpendicular a ela e passando pela origem. Se $\dim \operatorname{aff} C' = 2$. Então C' é aberto topológico. Seja $K = \cup_{r>0} rC'$ o cone gerado por C'. Temos K aberto convexo. E $0 \notin K$. A intersecção de K com $S^1 = \{x \in P : |x| = 1\}$ é um "intervalo" de comprimento menor do que π (identificando P com $\mathbb{R}^2$) pois caso contrário conteria um par de antípodas e então a origem. Agora é só passar um reta pela origem que não passe por $S^1 \cap K$.

Teorema 14 *Sejam C_1 e C_2 convexos. Existe um hiperplano que separa C_1 e C_2 própriamente se e somente se os interiores relativos de C_1 e C_2 são disjuntos.*

Demonstração: $C = C_1 - C_2$ é convexo, Pela proposição 1, $\operatorname{ri} C = \operatorname{ri} C_1 - \operatorname{ri} C_2$ e então $0 \notin \operatorname{ri} C$. Existe um hiperplano contendo $M = \{0\}$ tal que $\operatorname{ri} C$ está contido num dos seus semi–espaços abertos. Logo C está contido num semi–espaço fechado. Resumindo: existe $b \in \mathbb{R}^n \setminus \{0\}$, tal que

$$0 \leq \inf_{x \in C} \langle x, b \rangle = \inf_{x_1 \in C_1} \langle x_1, b \rangle - \sup_{x_2 \in C_2} \langle x_2, b \rangle$$
$$0 < \sup_{x \in C} \langle x, b \rangle = \sup_{x_1 \in C_1} \langle x_1, b \rangle - \inf_{x_2 \in C_2} \langle x_2, b \rangle .$$

Então o teorema 12 implica a separação própria. Por outro lado essas condições implicam $0 \notin \operatorname{ri} C$ pois $D = \{x : \langle x, b \rangle \geq 0\} \supset C$ e $\operatorname{ri} D = \{x : \langle x, b \rangle > 0\}$ intersecta C e portanto (cor. 10) $\operatorname{ri} C \subset \operatorname{ri} D$.

Teorema 15 *Sejam C_1 e C_2 convexos não–vazios. Para existir um hiperplano separando–os fortemente, é necessário e suficiente que*

$$d\,(C_1, C_2) := \inf \{|x_1 - x_2| : x_1 \in C_1, x_2 \in C_2\} > 0.$$

Em outras palavras, $0 \notin \overline{C_1 - C_2}$.

Demonstração: Se um hiperplano separa fortemente C_1 e C_2 existe $\epsilon > 0$ tal que $(C_1 + \epsilon B) \cap (C_2 + \epsilon B) = \emptyset$. Mas então $d\,(C_1, C_2) \geq \epsilon$. Em particular $0 \notin \overline{C_1 - C_2}$. Suponhamos agora $0 \notin \overline{C_1 - C_2}$. Então $\epsilon B \cap (C_1 - C_2) = \emptyset$ para algum $\epsilon > 0$. Nesse caso $C_1 + \frac{\epsilon}{2} B$ e $C_2 + \frac{\epsilon}{2} B$ podem ser propriamente separados. E C_1 e C_2 são fortemente separados.

Lema 8 *Seja X fechado e K compacto em $\mathbb{R}^n$. Então $X - K$ é fechado.*

Demonstração: Seja $z_n = x_n - k_n \to z$. Passando para uma subseqüência se necessário podemos supor $k_n \to k \in K$. Mas então $x_n = z_n + k_n \to z + k \in X$. Logo $z \in X - K$.

Corolário 11 *Sejam C_1, C_2 convexos não–vazios, disjuntos e fechados, um deles limitado. Então existe um hiperplano separando–os fortemente.*

Demonstração: Com efeito, pelo lema anterior, $0 \notin \overline{C_1 - C_2} = C_1 - C_2$.

Corolário 12 *Se C_1, C_2 convexos, não–vazios com fechos disjuntos. Se um deles for limitado, existe um hiperplano que os separa fortemente.*

11/03

Teorema 16 *Um conjunto convexo fechado é a intersercção dos semi–espaços fechados que o contém.*

Demonstração: Seja C convexo fechado. Sem perda de generalidade, C é não–vazio e $\neq \mathbb{R}^n$. Para $x \in \mathbb{R}^n \setminus C$ temos $C \cap \{x\} = \emptyset$. Portanto existe um hiperplano H_x que contém C no semi–espaço fechado à esquerda, H_x^- e $x \notin H_x$. A intersecção desses semi–espaços, $\cap_{x \in C^c} H_x^- = C$.

Definição 15 *Um hiperplano H é um hiperplano suporte do convexo C se $H \cap C \neq \emptyset$ e C está contido num dos semi–espaços fechados de H. O hiperplano é não–trivial se $C \setminus H$ for não–vazio.*

Teorema 17 *Seja C convexo e $D \neq \emptyset$ convexo contido em C. Existe um hiperplano não trivial suportando C e que contém D se e somente se $D \cap \operatorname{ri} C = \emptyset$.*

Demonstração: Um hiperplano não–trivial que contém $D \subset C$ e suporta C é um hiperplano que separa D e C própriamente. Pelo teorema 14 existe o hiperplano se e somente se $\operatorname{ri} C$ disjunto de $\operatorname{ri} D$. Mas isso é equivalente a D ser disjunto de $\operatorname{ri} C$ (cor. 10)

Notação 1 *Para $x = (x_1, \ldots, x_n)$ escrevemos $x < 0$ se $x_i \leq 0$ para todo i e $x \neq 0$. E $x << 0$ se $x_i < 0$ para todo $i \leq n$.*

Comentário 9 *Na linguagem matricial (usada nos exemplos abaixo), $x \in \mathbb{R}^m$ é um vetor coluna, $m \times 1$. O produto interno de x, b vetores $m \times 1$ é $x^t b$.*

Exemplo 4 (Gordan) *Seja A matriz $m \times n$. Então somente uma das alternativas a seguir é verdadeira:*

i) *Existe $x \in \mathbb{R}^n$ tal que $Ax << 0$;*

ii) *Existe $b \in \mathbb{R}^m_+$ e $b \neq 0$ tal que $A^t b = 0$.*

Suponhamos (i) e (ii) válidas. Então $x^t A^t b = 0 \implies (Ax)^t b = 0$ uma impossibilidade pois $Ax << 0$. Sejam $C_1 = \{Ax : x \in \mathbb{R}^n\} \subset \mathbb{R}^m$ e $C_2 = \{y \in \mathbb{R}^m : y << 0\}$. Se (i) for falso, $C_1 \cap C_2 = \emptyset$. Então existe um hiperplano que separa propriamente: $H = \{x \in \mathbb{R}^m : x^t b = \beta\}$ e $C_1 \subset \{x : x^t b \geq \beta\}$ e $C_2 \subset \{x : x^t b \leq \beta\}$. Então $0 \in C_1$ implica $\beta \leq 0$. E $\left(-\frac{1}{n}, \ldots, -\frac{1}{n}\right) \in C_2$ implica $0 \leq \beta$ e então $\beta = 0$. E necessariamente $b \geq 0$. Finalmente sendo C_1 um subespaço vetorial, $y^t b = 0$ se $y \in C_1$ e logo vale $(Ax)^t b = 0$ para todo $x \implies A^t b = 0$.

Exemplo 5 (lema de Farkas) *Seja A matriz $m \times n$ e b vetor $m \times 1$. Somente uma das alternativas a seguir é verdadeira:*

1. *Existe x $n \times 1$, $x \geq 0$, tal que $Ax = b$;*
2. *Existe μ vetor $m \times 1$, $\mu^t A \geq 0$ e $\mu^t b < 0$.*

Comentário 10 *Trocando o sinal de μ (2) equivale a (2)' $\mu^t A \leq 0$ e $\mu^t b > 0$.*

Demonstração: Primeiramente notemos que (1) e (2) não são válidas simultaneamente. Se

$$Ax = b, x \geqq 0,$$
$$\mu^t A \geqq 0, \mu^t b < 0.$$

Então $\mu^t Ax \geq 0$ pois $x \geq 0$. Logo $\mu^t b \geq 0$ em contradição com $\mu^t b < 0$. Seja $v_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^t$ a j–ésima coluna da matriz A. Então $Ax = b$ se e somente se

$$b \in \text{cone}\{v_1, v_2, \ldots, v_n\} = \left\{ \sum_{i=1}^{n} \lambda_i v_i : \lambda_i \geq 0, 1 \leq i \leq n \right\}.$$

Suponhamos então que (1) não vale. Então $b \notin K := \text{cone}\{v_1, v_2, \ldots, v_n\}$. Existe então μ, $m \times 1$ tal que

$$\mu^t b < 0 \leq \mu^t v_i, 1 \leq i \leq n.$$

E portanto vale (2).

Comentário 11 *A separação estrita acima depende de K ser fechado. Os lemas a seguir visam demonstrar isso.*

Lema 9 *Se $v_1, v_2, \ldots, v_n$ for l.i. então K é fechado.*

Demonstração: Seja $x^t = \sum_{i=1}^n \lambda_i^t v_i, \lambda_i^t \geq 0$, com limite x. Se $\lambda^t = (\lambda_i^t)_{i=1}^n$ possui subseqüência convergente, $\lambda^{k(t)}$, então $x = \sum_{i=1}^n \lim_t \lambda_i^{k(t)} v_i = \sum_{i=1}^n \lambda_i v_i \in K$ pois $\lambda_i = \lim_k \lambda_i^{k(t)} \geq 0$. Se (λ^t) não possui subseqüência convergente[2], $|\lambda^t| \to \infty$ e então, passando a uma subseqüência se necessário, $\frac{\lambda^t}{|\lambda^t|_S} \to \mu \geq 0, |\mu| = 1$. Mas então

$$\sum_i \mu_i v_i = \lim_t \frac{x^t}{|\lambda^t|} = 0.$$

Contradição com a independência linear.

[2]usando a norma da soma

Lema 10 *Para todo $x \in K$ existe $S \subset \{1, \ldots, n\}$ tal que $\{v_i : i \in S\}$ é l.i. e $x \in \text{cone}\{v_i : i \in S\}$.*

Demonstração: Para $x \in K$, seja a representação $x = \sum_{i=1}^{n} \lambda_i v_i$, $\lambda_i \geq 0$ e tal que $S = \{i : \lambda_i > 0\}$ seja tal que $\#S$ é o menor possível. Se $\{v_i : i \in S\}$ não for l.i existe θ_i, $i \in S$ nem todos nulos tal que $\sum_{i \in S} \theta_i v_i = 0$. Sem perda de generalidade pelo menos um $\theta_j > 0$. Seja $t = \min \left\{ \frac{\lambda_i}{\theta_i} : \theta_i > 0 \right\}$. Então $\lambda_i - t\theta_i \geq 0$ para todo i e

$$x = \sum_{i \in S} (\lambda_i - t\theta_i) v_i$$

é uma representação com menos de $\#S$ elementos positivos. Contradição.

Teorema 18 *Seja $K = \text{cone}\{v_1, v_2, \ldots, v_n\}$. Então K é fechado.*

Demonstração: Seja $\mathfrak{S} = \{S \subset \{1, \ldots, n\} : \{v_i : i \in S\}$ é l.i.$\}$. Note que $K = \cup_{S \in \mathfrak{S}} \text{cone}\{v_i : i \in S\}$ e cada $\text{cone}\{v_i : i \in S\}$ é fechado.

Funções convexas

No estudo das funções convexas é conveniente permitir que as funções assumam valores em $[-\infty, \infty]$. O símbolo $\dot{+}$ é ocasionalmente necessário:

$$\begin{aligned} x \dot{+} y &= x + y \text{ se } \{x, y\} \neq \{-\infty, \infty\} \\ -\infty \dot{+} \infty &= \infty \dot{+} -\infty = \infty. \end{aligned}$$

Definição 16 *Seja $C \subset \mathbb{R}^n$ convexo. $f : C \to \mathbb{R}$ é convexa se $f(rc + (1-r)c') \leq rf(c) + (1-r)f(c')$,$\forall c, c' \in C$.*

Comentário 12 *A definição é a usual mas na análise convexa a próxima definição é vantajosa pois deixa C implicíto.*

Definição 17 *$f : \mathbb{R}^n \to [-\infty, \infty]$ é convexa se para todo $x, y \in \mathbb{R}^n$ e $0 < r < 1$,*

$$f(rx + (1-r)y) \leq rf(x) \dot{+} (1-r)f(y). \tag{*}$$

Definição 18 *O domínio efetivo de f, $\text{dom} f = \{x : f(x) < \infty\}$.*

Definição 19 *O epígrafo de f é o conjunto $\text{epi} f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq r\}$.*

Definição 20 *f é própria se não assume o valor $-\infty$ e $\text{dom} f$ é não–vazio.*

Comentário 13 *Se $f : C \to \mathbb{R}$ é convexa pela definição 16 então definindo para $x \in \mathbb{R}^n \setminus C$, $f(x) = \infty$ temos que f é convexa pela definição (*).*

Lema 11 *f é convexa se e somente se* $\operatorname{epi} f$ *for convexo.*

Demonstração: Suponhamos f convexa. E $(x, r), (x', r') \in \operatorname{epi} f$ e $0 < \lambda < 1$. Então como $f(x), f(y) < \infty$,

$$\begin{aligned} f(\lambda x + (1-\lambda) x') &\leq \lambda f(x) + (1-\lambda) f(y) \leq \lambda r + (1-\lambda) r' \\ &\implies \lambda (x, r) + (1-\lambda)(x', r') \in \operatorname{epi} f. \end{aligned}$$

Suponhamos agora que o epígrafo de f é convexo. A desigualdade (*) vale sempre que $f(x) = \infty$ ou $f(y) = \infty$. Suponhamos agora que $f(x) < \infty$ e $f(y) < \infty$. Então se $r > f(x)$ e $s > f(y)$ temos $(x, r), (y, s) \in \operatorname{epi} f$ e então

$$(\lambda x + (1-\lambda) y, \lambda r + (1-\lambda) s) \in \operatorname{epi} f.$$

Logo

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

Fazendo $r \downarrow f(x)$ e $s \downarrow f(y)$ vem $f(rx + (1-r) y) \leq r f(x) \dotplus (1-r) f(y)$.

Teorema 19 (des. Jensen) *Seja f convexa própria. Então*

$$f\left(\sum_{j=1}^{m} \lambda_j x_j\right) \leq \sum_{j=1}^{m} \lambda_j f(x_j),$$

$$\lambda_j \geq 0, \sum_{j=1}^{m} \lambda_j = 1.$$

Comentário 14 *Podemos dizer que essa é a des. de Jensen no caso discreto.*

Exemplo 6 (desigualdade aritmético–geométrica) *Se $x_i > 0$ para $i \leq n$ então*

$$\sqrt[n]{x_1 x_2 \ldots x_n} \leq \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

A função $f(x) = -\log x$ é convexa pois $f''(x) = \frac{1}{x^2} > 0$. Pela desigualdade de Jensen, $f\left(\frac{\sum_{i=1}^n x_i}{n}\right) \leq \sum_{i=1}^n \frac{f(x_i)}{n}$ ou

$$-\log \frac{\sum_{i=1}^n x_i}{n} \leq \sum_{i=1}^{n} \frac{-\log(x_i)}{n} = -\log \sqrt[n]{x_1 x_2 \ldots x_n}.$$

13/03

Teorema 20 *Seja $f : [a,b] \to \mathbb{R}$ diferenciável. Então f é convexa se e somente $f'(x)$ for crescente em $[a,b]$.*

Demonstração: Sejam x, y tais que $a \leq x < y \leq b$. E definamos $g(t) = f((1-t)x + ty) - (1-t)f(x) - tf(y)$. Note que $g(0) = 0 = g(1)$. Então f é convexa se e somente se $g(t) \leq 0$. Suponhamos que $f(\cdot)$ seja convexa. Então $g(t) \leq 0$ sempre e portanto $g'(0) \leq 0$ e $g'(1) \geq 0$. Então de $g'(t) = f'((1-t)x + ty)(y-x) + f(x) - f(y)$ vem

$$g'(0) = f'(x)(y-x) + f(x) - f(y) \leq 0 \leq f'(y)(y-x) + f(x) - f(y)$$

ou seja

$$f'(x) \leq \frac{f(y) - f(x)}{y - x} \leq f'(y)$$

demonstrando que f' é crescente. Para a recíproca, suponhamos agora que f' seja crescente. Queremos demonstrar que $g(t) \leq 0, t \in [0,1]$. Pelo teorema do valor médio, seja $\xi \in (0,1)$ tal que $g'(\xi) = g(1) - g(0)$. Ou

$$f'((1-\xi)x + \xi y) = \frac{f(y) - f(x)}{y - x}.$$

Para $0 \leq t < \xi$,

$$g'(t) = (y-x)\left[f'((1-t)x + ty) - \frac{f(y) - f(x)}{y - x}\right] =$$
$$(y-x)\left[f'((1-t)x + ty) - f'((1-\xi)x + \xi y)\right] \leq 0.$$

Portanto g é decrescente até ξ. E depois é crescente terminando em $g(1) = 0$. Demonstrando que $f(\cdot)$ é convexa.

Corolário 13 *Se f for duas vezes diferenciável, f é convexa se e somente se $f''(x) \geq 0$.*

Exemplo 7 (função indicadora) *Para $C \subset \mathbb{R}^n$ a função indicadora de C é $\delta(\cdot|C)$,*

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

O epígrafo da indicadora é $C \times [0, \infty)$. Portanto C é convexo se e somente se a função indicadora for convexa.

Definição 21 *A função suporte do conjunto convexo C é*

$$\delta^*(x|C) = \sup\{\langle x, y\rangle : y \in C\}.$$

Definição 22 *O gabarito $\gamma(\cdot|C)$ para C não–vazio:*

$$\gamma(x|C) = \inf\{r \geq 0 : x \in rC\}.$$

Definição 23 *A função distância: $d(x, C) = \inf\{|x - y| : y \in C\}$.*

Todas essas funções são convexas: Para a função suporte seja $x, y \in C$ e $r \in (0, 1)$. Então

$$\begin{aligned}\langle rx + (1-r)y, c\rangle = r\langle x, c\rangle + (1-r)\langle y, c\rangle \leq r\delta^*(x)|C) + (1-r)\delta^*(y|C)\\ \Longrightarrow \delta^*(rx + (1-r)y|C) \leq r\delta^*(x)|C) + (1-r)\delta^*(y|C).\end{aligned}$$

Para o gabarito: Sem perda de generalidade suponhamos $\gamma(x|C), \gamma(y|C) < \infty$. Dado $\epsilon > 0$, existe $r < \gamma(x|C) + \epsilon$ tal que $x \in rC$ e exists $s < \gamma(y|C) + \epsilon$ tal que $y \in sC$. Então para $t \in (0, 1)$, $tx + (1-t)y \in trC + (1-t)sC = (tr + (1-t)s)C$ e $\gamma(tx + (1-t)y) \leq tr + (1-t)s < t\gamma(x|C) + (1-t)\gamma(y|c) + \epsilon$. Distância: sejam $c, c' \in C$.

$$\begin{aligned}d(tx + (1-t)y, C) \leq \|tx + (1-t)y - tc + (1-t)c'\| \leq t\|x - c\| + (1-t)\|y - c'\|\\ \Longrightarrow d(tx + (1-t)y, C) \leq td(x, C) + (1-t)d(y, C).\end{aligned}$$

Teorema 21 *Seja f convexa e $\alpha \in [-\infty, \infty]$. Então são convexos:*

$$\{x : f(x) < \alpha\} \text{ e } \{x : f(x) \leq \alpha\}.$$

Demonstração: Se $f(x) \leq \alpha$ e $f(y) \leq \alpha$ então $f(rx + (1-r)y) \leq rf(x) + (1-r)f(y) \leq r\alpha + (1-r)\alpha = \alpha$. Analogamente para a desigualdade estrita.

Teorema 22 *Seja $f : \mathbb{R}^n \to (-\infty, \infty]$ convexa e $\phi : (-\infty, \infty] \to (-\infty, \infty]$ convexa e não–decrescente. Então $\phi \circ f$ é convexa.*

Demonstração: De $f(rx + (1-r)y) \leq rf(x) + (1-r)f(y)$ obtemos aplicando ϕ:

$$\phi(f(rx + (1-r)y)) \leq \phi(rf(x) + (1-r)f(y)) \leq r\phi(f(x)) + (1-r)\phi(f(y)).$$

Exemplo 8 *1. Se f convexa própria, então $e^{f(x)}$ é convexa própria.*

2. Se f for convexa não–negativa e $p > 1$, $(f(x))^p$ é convexa.

Proposição 2 *A soma de duas convexas próprias é convexa. A soma é própria se uma das funções for finita sempre.*

A demonstração é imediata.

Proposição 3 *Seja* $F \subset \mathbb{R}^{n+1}$ *convexo não–vazio. A seguinte função é convexa:*

$$f(x) = \inf \{\mu : (x, \mu) \in F\}.$$

Demonstração: Note que $\inf \emptyset = \infty$. Então se $rf(x) + (1-r) f(y) < \infty$ existe $(x, \mu) \in F$ e $(y, \nu) \in F$. Portanto $(rx + (1-r) y, r\mu + (1-r) \nu) \in F$ e logo $f(rx + (1-r) y) \leq r\mu + (1-r) \nu$.

Proposição 4 (convolução) *Sejam* $f_1, \ldots, f_m$ *convexas próprias. E seja*

$$f(x) = \inf \{f_1(x_1) + \ldots + f_m(x_m) : x_1 + \ldots + x_m = x\}.$$

Então f *é convexa.*

Demonstração: Seja $F_i = \operatorname{epi} f_i$, $1 \leq i \leq m$ e $F = F_1 + F_2 + \ldots + F_m$ é convexo de $\mathbb{R}^{n+1}$. Então $(x, \mu) \in F$ se existirem $(x_i, \mu_i) \in F_i$, $\mu = \sum_i \mu_i$, $x = \sum_i x_i$, $f_i(x_i) \leq \mu_i$. Portanto f é convexa.

Teorema 23 *O supremo de uma família de funções convexas é uma função convexa.*

Demonstração: Seja $f(x) = \sup_{i \in I} f_i(x)$. Então $f_i(rx + (1-r) y) \leq rf_i(x) \dot{+} (1-r) f_i(y) \leq rf(x) \dot{+} (1-r) f(y)$. Outra demonstração: $\operatorname{epi} f = \cap_{i \in I} \operatorname{epi} f_i$.

Continuidade

Definição 24 *Uma função entre espaços métricos,* $f : X \to Y$ *é Lipschitz se existir* $k > 0$ *tal que*

$$d(f(x), f(y)) \leq kd(x, y), x, y \in X.$$

Comentário 15 *Se quisermos especificar a constante dizemos que* f *é* k*-lipschitz.*

Definição 25

1. *Uma função* $f : \mathbb{R}^n \to \mathbb{R}$ *é lipschitz numa vizinhança de* $x \in \mathbb{R}^n$ *se exister* $\epsilon > 0$ *e* $K > 0$ *tais que*

$$|f(y) - f(z)| \leq K |y - z|, y, z \in B(x, \epsilon).$$

2. *E f é localmente lipschitz no aberto U se para todo $x \in U$, f é lipschitz numa vizinhança de x contida em U.*

Comentário 16 *Note que essa condição implica a continuidade de f em U.*

Teorema 24 *Seja f convexa própria. Seja $U = \operatorname{int}(\operatorname{dom} f)$. Suponha que existe bola aberta, $x_0 + \epsilon B \subset U$ e $M < \infty$ tais que $f(x) \leq M$ para todo $x \in x_0 + \epsilon B$. Então f é localmente lipschitz em U.*

Demonstração: Sem perda de generalidade[3], $x_0 = 0$. Assim $f(u) \leq M$ se $|u| < \epsilon$. Seja $x \in U$. Vou demonstrar que f é limitada numa vizinhança de x. Seja $\rho > 1$ tal que $y := \rho x \in U$.

Seja $\lambda = \frac{1}{\rho}$. Então V a seguir é uma vizinhança de x:

$$V = \{v : v = (1-\lambda)x' + \lambda y, |x'| < \epsilon\} = x + (1-\lambda)\epsilon B.$$

Por convexidade, $f(v) \leq (1-\lambda) f(x') + \lambda f(y) \leq M + \lambda f(y)$. Então f é limitada superiormente numa vizinhança de x. Se $z \in V$ existe $z' \in V$, $x = \frac{z+z'}{2}$. Então

$$f(x) \leq \frac{f(z) + f(z')}{2} \implies f(z) \geq 2f(x) - f(z') \geq 2f(x) - M - \lambda f(y).$$

Assim f é limitada numa vizinhança de x. Seja N uma cota superior para $|f|$ em $x + 2\delta B, \delta > 0$. Para $x_1 \neq x_2$ em $x + \delta B$, seja $x_3 = x_2 + \frac{\delta}{\alpha}(x_2 - x_1)$ sendo $\alpha = |x_2 - x_1|$. Note que $x_3 \in x + 2\delta B$. Resolvendo para x_2:

$$\begin{aligned} x_2 = \frac{\delta}{\alpha+\delta}x_1 + \frac{\alpha}{\alpha+\delta}x_3 &\implies f(x_2) \leq \frac{\delta}{\alpha+\delta} f(x_1) + \frac{\alpha}{\alpha+\delta} f(x_3) \\ &\implies f(x_2) - f(x_1) \leq \frac{\alpha}{\alpha+\delta}[f(x_3) - f(x_1)] \leq \frac{\alpha}{\delta}2N = \frac{2N}{\delta}|x_2 - x_1|. \end{aligned}$$

Trocando x_2 com x_1 obtemos $|f(x_2) - f(x_1)| \leq \frac{2N}{\delta}|x_2 - x_1|$.

Comentário 17 *A demonstração acima vale–sem modificações– para $f : B \to [-\infty, \infty]$, B normado, f convexa própria..*

Comentário 18 *Para $f : \mathbb{R}^n \to (-\infty, \infty]$ própria podemos supor f finita num aberto e ainda obter a continuidade no interior de* $\operatorname{dom} f$.

[3] Basta considerar $g(x) = f(x_0 + x)$.

18/03

Derivada direcional

Seja $f : \mathbb{R}^n \to \mathbb{R}$ convexa, $h \in \mathbb{R}^n$. Seja $g(\lambda) = \frac{f(x+\lambda h) - f(x)}{\lambda}, \lambda > 0$. Suponhamos $\lambda > \mu > 0$. Seja $r = \mu/\lambda$. Então

$$\begin{aligned} f(x + \mu h) = f(x + r\lambda h) = f(r(x + \lambda h) + (1-r)x) &\leq rf(x + \lambda h) + (1-r) f(x) \\ \implies f(x+\mu h) - f(x) &\leq r(f(x+\lambda h) - f(x)) \\ \implies g(\mu) &\leq g(\lambda). \end{aligned}$$

Então $f'(x,h) = \lim_{\lambda \downarrow 0} g(\lambda) = \inf_{\lambda > 0} g(\lambda)$ existe. Note que $f'(x, 0) = 0$. E $f'(x, rh) = rf'(x,h)$ se $r > 0$.

Lema 12 $-f'(x, -h) \leq f'(x, h)$.

Demonstração: Note que $\frac{f(x+rh)+f(x-rh)}{2} \geq f(x)$ e então

$$\frac{f(x+rh) - f(x)}{r} \geq -\frac{f(x-rh) - f(x)}{r} \implies f'(x,h) \geq -f'(x,-h).$$

Definição 26 *Um funcional $f : V \to \mathbb{R}$ é sub–aditivo se $f(x+y) \leq f(x) + f(y)$.*

Subgradiente

Seja f convexa. O vector x^* é um subgradiente de f em x se para todo $z \in \mathbb{R}^n$,

$$f(z) - f(x) \geq < x^*, z - x > .$$

O conjunto dos subgradientes em x é denotado $\partial f(x)$. É imediato da definição que $\partial f(x)$ é convexo e fechado.

Exemplo 9 (subgradiente da indicadora) *Seja C convexo não–vazio. Então $x^* \in \partial\delta(x|C)$ se e somente se*

$$\begin{aligned} \delta(z|C) &\geq \delta(x|C) + < x^*, z - x >, \forall z \\ &\implies x \in C \text{ e } 0 \geq < x^*, z - x >, z \in C. \end{aligned}$$

Comentário 19 *x^* nesse caso é normal a C em x. (Não entraremos nesse assunto.)*

Teorema 25 *Seja f convexa própria. Então $\partial f(x) \neq \emptyset$ para todo $x \in \operatorname{int}(\operatorname{dom} f)$.*

Demonstração: Seja $x \in \operatorname{int}(\operatorname{dom} f)$. Seja $f(x) < \mu < \infty$. Pela continuidade de f existe $W \ni x$ tal que $f(W) \subset \left(-\infty, \frac{\mu + f(x)}{2}\right)$. Então $W \times \left(\frac{\mu+f(x)}{2}, \infty\right) \subset \operatorname{epi} f$ e portanto $(x, \mu) \in \operatorname{int}(\operatorname{epi} f)$.

E $(x, f(x)) \notin \operatorname{int}(\operatorname{epi} f)$. Existe então um hiperplano não–trivial, suporte de $\operatorname{epi} f$ e contém $(x, f(x))$. Seja $(b, \lambda) \in (\mathbb{R}^n \times \mathbb{R}) \setminus \{0\}$ tal que

$$\begin{aligned} \langle b, x\rangle + \lambda f(x) &\leq < b, y > + \lambda\mu, \\ & y \in \mathbb{R}^n, \mu > f(y). \end{aligned}$$

Se $\lambda = 0$ temos $\langle b, x\rangle \leq \langle b, y\rangle$ para todo $y \in \operatorname{int}(\operatorname{dom} f)$ e portanto $b = 0$ contradição com $(b, \lambda) \neq 0$. Logo $\lambda \neq 0$. Fazendo $\mu \to \infty$ podemos concluir que $\lambda > 0$. Sem perda de generalidade $\lambda = 1$:

$$\begin{aligned} \langle b, x\rangle + f(x) &\leq < b, y > + \mu, f(y) < \mu \\ &\Longrightarrow \langle b, x - y\rangle \leq f(y) - f(x). \end{aligned}$$

Seja $x^* = -b$. Assim $x^* \in \partial f(x)$.

Comentário 20 *O resultado vale mais geralmente se* $x \in \operatorname{ri}(\operatorname{dom} f)$.

Proposição 5 *O subgradiente é monótono.*

Demonstração: Sejam $x, y \in \operatorname{ri} \operatorname{dom} f$ e $x^* \in \partial f(x)$, $y^* \in \partial f(y)$. Então

$$\begin{aligned} f(y) - f(x) &\geq < x^*, y - x > \\ f(x) - f(y) &\geq < y^*, x - y > . \end{aligned}$$

Somando, $0 \geq \langle x^* - y^*, y - x\rangle$.

$$< y^* - x^*, y - x > \geq 0.$$

Teorema 26 *Sejam $f_1, \ldots, f_m$ convexas próprias. Então para $f = \sum_{j=1}^m f_j$,*

$$\partial f(x) \supseteq \partial f_1(x) + \ldots + \partial f_m(x).$$

Além disso se $\cap_{i=1}^m \operatorname{ri} \operatorname{dom} f_i \neq \emptyset$,

$$\partial f(x) = \partial f_1(x) + \ldots + \partial f_m(x).$$

Se $x_j^* \in \partial f_j(x), 1 \leq j \leq m$ então

$$\begin{aligned} &\left\langle x_j^*, y - x\right\rangle \leq f_j(y) - f_j(x), j \leq m \\ &\qquad \Longrightarrow \left\langle \sum_{j=1}^m x_j^*, y - x\right\rangle \leq \left(\sum_{j=1}^m f_j\right)(y) - \left(\sum_{j=1}^m f_j\right)(x). \end{aligned}$$

A demonstração da igualdade será omitida por falta de tempo. A demonstração da próxima proposição é imediata.

Proposição 6 $0 \in \partial f(x) \iff x$ *é ponto de mínimo de* f *convexa própria.*

Proposição 7 x *é ponto de mínimo de* f *se e somente se* $f'(x,h) \geq 0$ *para todo* h.

Multiplicadores de Kuhn Tucker

Definição 27 ((função afim)) *Uma função* $f : \mathbb{R}^n \to \mathbb{R}$ *é afim se for da forma*

$$f(x) = \langle x, v \rangle + \beta, v \in \mathbb{R}^n, \beta \text{ real.}$$

Teorema 27 *Seja* C *convexo. Sejam* $f_1, \ldots, f_m$ *convexas próprias com* $\operatorname{dom} f_i \supset \operatorname{ri} C$, $1 \leq i \leq m$. *Então somente uma das alternativas a seguir é válida:*

a) *Existe* $x \in C$, $(f_1(x), \ldots, f_m(x)) << 0$;

b) *Existe* $\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0$, $\lambda \neq 0$,

$$\sum_{i=1}^{m} \lambda_i f_i(x) \geq 0, \forall x \in C.$$

Demonstração: Seja $g(x) = (f_1(x), \ldots, f_m(x))$. É imediato que se $f_i(x) < 0$ então $\lambda_i f_i(x) \leq 0$ e pelo menos para um $i \leq m$, é < 0. Logo (b) não vale. Suponhamos que (a) seja falso. Devemos demonstrar que vale (b). Seja

$$C_1 = \{z \in \mathbb{R}^m : \exists x \in C, g(x) << z\} = g(C) + \mathbb{R}^m_{++}.$$

Então $C_1 \cap \left(-\mathbb{R}^m_+\right) = \emptyset$. Podemos então, pelo teorema 14, separar C_1 e $-\mathbb{R}^m_+$ propriamente. Existe $\lambda = (\lambda_1, \ldots, \lambda_m) \neq 0$, e α real,

$$\begin{aligned} &\alpha \leq \lambda \cdot z = \lambda_1 z_1 + \ldots + \lambda_m z_m, z \in C_1; \\ &\alpha \geq \lambda \cdot z, z \in -\mathbb{R}^m_+. \end{aligned}$$

Logo $\alpha \geq 0$ e $\lambda \geq 0$. Portanto $0 \leq \lambda \cdot z, z \in C_1$. Então para $x \in D = C \cap \cap_{i=1}^m \operatorname{dom} f_i \supset \operatorname{ri} C$,

$$\begin{aligned} &0 \leq \lambda_1 (f_1(x) + \epsilon) + \ldots + \lambda_m (f_m(x) + \epsilon), \\ &\epsilon \downarrow 0 \implies 0 \leq \lambda_1 f_1(x) + \ldots + \lambda_m f_m(x), x \in D. \end{aligned}$$

Então a desigualdade vale para $x \in \overline{D}$ (ver cor. 15 abaixo) e então vale para $x \in C$ pois $C \subset \overline{\operatorname{ri} C} \subset \overline{D}$.

Exemplo 10 *A hipótese* $\operatorname{dom} f_i \supset \operatorname{ri} C$ *é necessária. Por exemplo seja*

$$f_1(x) = \begin{cases} -\sqrt{x} & se \quad x \geq 0 \\ \infty & se \quad x < 0. \end{cases}$$

E $f_2(x) = x$ *para* $x \in C := \mathbb{R}$. *Então (a) acima não vale. Mas* $\lambda_1 f_1(x) + \lambda_2 f_2(x) \geq 0$ *implica para* $x > 0$,

$$-\lambda_1\sqrt{x} + \lambda_2 x \geq 0 \implies -\lambda_1 + \lambda_2\sqrt{x} \geq 0 \implies -\lambda_1 \geq 0 \implies \lambda_1 = 0 \implies \lambda_2 = 0.$$

Lema 13 *Seja* f *convexa e* $x \in \operatorname{dom} f$ *tal que* $f(x) < \alpha$. *Então existe* $x' \in \operatorname{ri}\operatorname{dom} f$ *tal que* $f(x') < \alpha$.

Demonstração: Seja $x^0 \in \operatorname{ri}\operatorname{dom} f$. Então $f((1-r)x^0 + rx) \leq (1-r) f(x^0) + rf(x) < \alpha$ se $r < 1$ estiver suficientemente próximo de 1. Então $x' = (1-r)x^0 + rx \in \operatorname{ri}\operatorname{dom} f$ pelo teorema 9.

Corolário 14 *Seja* f *convexa. Seja* C *convexo tal que* $\operatorname{ri} C \subset \operatorname{dom} f$. *Se* $f(x) < \alpha$ *para* $x \in \overline{C}$ *existe* $x' \in \operatorname{ri} C$ *tal que* $f(x') < \alpha$.

Demonstração: Seja $x^0 \in \operatorname{ri} C$. Então $f((1-r)x^0 + rx) \leq (1-r) f(x^0) + rf(x) < \alpha$ se r estiver próximo de 1. Então $x' = (1-r)x^0 + rx \in \operatorname{ri} C$ e $f(x') < \alpha$.

Corolário 15 *Seja* f *convexa e* $C \subset \operatorname{dom} f$ *convexo. Se* $f(x) \geq \alpha$ *para todo* $x \in C$ *então* $f(x) \geq \alpha$ *para* $x \in \overline{C}$.

20/03

Programa convexo

Um programa convexo[4], (P), é definido pela $m+2$ upla $(C, f_0, \ldots, f_m)$ sendo

1. C convexo não–vazio;
2. f_j convexa própria, $\operatorname{dom} f_j \supset C$, $0 \leq j \leq m$.

E queremos minimizar $f_0(x)$, $x \in C$ com as restrições:

$$f_j(x) \leq 0, 1 \leq j \leq m. \qquad (*)$$

Em geral supomos:

[4]No livro do Rockafellar a definição é mais geral permitindo restrições de igualdade com funções afim.

(a) $\operatorname{dom} f_0 = C$,

(b) $\operatorname{ri} \operatorname{dom} f_j \supset \operatorname{ri} C$.

Definição 28 *O vetor* $x \in \mathbb{R}^n$ *é factível se* $x \in C$ *e* $f_j(x) \leq 0, 1 \leq j \leq m$.

Seja $C_0 = C \cap C_1 \cap \ldots \cap C_m$ sendo $C_j = [f_j \leq 0]$, $j = 1, \ldots, m$. A função objetivo é $f(x) = f_0(x) + \delta(x|C_0)$. O ínfimo de f é o valor ótimo do problema (P). E os pontos nos quais o ínfimo é alcançado são soluções ótimas de (P).

Definição 29 *Um multiplicador de Kuhn–Tucker (KT) do problema (P) é um vetor* $\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0$ *e tal que*

$$\inf\{f(x) + \lambda_1 f_1(x) + \ldots + \lambda_m f_m(x) : x \in \mathbb{R}^n\}$$

é finito e igual ao valor ótimo de (P).

Teorema 28 *Seja (P) um programa convexo. E* λ *multiplicador de KT de (P). Seja* $h = f_0 + \lambda_1 f_1 + \ldots + \lambda_m f_m$ *e* $D = \{x : h(x) = \inf h(\mathbb{R}^n)\}$. *Seja* $I = \{j \leq m : \lambda_j = 0\}$, $J = \{j \leq m : \lambda_j > 0\}$. *Seja* D_0 *os pontos* $\bar{x} \in D$ *tais que*

$$\begin{aligned} f_i(\bar{x}) &= 0, i \in J \\ f_i(\bar{x}) &\leq 0, i \in I. \end{aligned}$$

Então D_0 *é o conjunto das soluções ótimas de (P).*

Demonstração: Por hipótese, $\inf h(\mathbb{R}^n) = \inf f\ (\mathbb{R}^n)$ é finito. Se $x \in C_0$, $\lambda_i f_i(x) \leq 0$, $1 \leq i \leq m$. E portanto

$$f_0(x) + \lambda_1 f_1(x) + \ldots + \lambda_m f_m(x) \leq f_0(x) = f(x).$$

Então $h(x) \leq f(x)$ para todo x com igualdade se e somente se x é factível e $\lambda_i f_i(x) = 0, i \leq m$. Assim se $i \in J$, $i \leq m$ temos $\lambda_i > 0$ o que implica $f_i(\bar{x}) = 0$. Então o mínimo de f está contido no mínimo de h e é D_0.

Comentário 21 $D_0 \neq D$ *é possível. Por exemplo se* $C = \mathbb{R}^n$ *e cada* f_i *afim. Nesse caso* h *é afim e tendo ínfimo finito é constante,* $D = \mathbb{R}^n$. *Mas* $D_0 \subset C_0$.

Teorema 29 *Seja (P) programa convexo. Suponhamos que o valor ótimo de (P) é finito e existe* $x \in \operatorname{ri} C$, *factível, que satisfaz as restrições com desigualdade estrita para cada* $i \leq m$. *Então existe pelo menos um multiplicador de Kuhn-Tucker para (P).*

Demonstração: Seja α o valor ótimo de (P). Existe uma solução, $x \in \operatorname{ri} C$, de

$$f_i(x) < 0, 1 \leq i \leq m.$$

Pela definição de α, o sistema

$$f_0(x) - \alpha < 0, f_1(x) < 0, \ldots, f_m(x) < 0 \qquad (*)$$

não tem solução em C. As desigualdades (*) satisfazem as hipóteses do teorema 27. Existem então $\lambda_i \geq 0, 0 \leq i \leq m$, $(\lambda_0, \ldots, \lambda_m) \neq 0$, tais que

$$\lambda_0 (f_0(x) - \alpha) + \lambda_1 f_1(x) + \ldots + \lambda_m f_m(x) \geq 0, x \in C.$$

Necessariamente, $\lambda_0 > 0$ (pois (*) e $(\lambda_0, \ldots, \lambda_k) \neq 0$) e então sem perda de generalidade, $\lambda_0 = 1$. Portanto

$$h = f_0 + \lambda_1 f_1 + \ldots + \lambda_m f_m \geq \alpha.$$

Mas $h \leq f_0$ nos pontos factíveis, e portanto $\inf h = \alpha$ e $\lambda_1, \ldots, \lambda_m$ são multiplicadores de KT.

Ponto de sela e Lagrangeano

O Lagrangiano do programa convexo (P) é a função $L : \mathbb{R}^m \times \mathbb{R}^n \to [-\infty, \infty]$,

$$L(u^*, x) = \begin{cases} f_0(x) + \sum_{i=1}^m v_i^* f_i(x) & \text{se} \quad u^* \in \mathbb{R}_+^m, x \in C \\ -\infty & \text{se} \quad u^* \notin \mathbb{R}_+^m, x \in C \\ \infty & \text{se} \quad x \notin C, \end{cases}$$

sendo $u^* = (v_i^*)_{i \leq m}$ multiplicador de KT. Temos L côncava em u^* e convexa em x. O par $(\bar{u}^*, \bar{x})$ é ponto de sela de L (com respeito a maximizar em u^* e minimizar em x) se

$$L(u^*, \bar{x}) \leq L(\bar{u}^*, \bar{x}) \leq L(\bar{u}^*, x), \forall u^*, \forall x.$$

Teorema 30 *$\bar{u}^*$ é vetor de KT para (P) e $\bar{x}$ solução ótima para (P) se e somente se $(\bar{u}^*, \bar{x})$ for ponto de sela para o Lagrangiano. Além disso essa condição vale se e somente se $\bar{x}$ e as componentes λ_i de $\bar{u}^*$ satisfazem*

a) $\lambda_i \geq 0, f_i(\bar{x}) \leq 0$ *e* $\lambda_i f_i(\bar{x}) = 0, 1 \leq i \leq m$

b) $0 \in \partial f_0(\bar{x}) + \lambda_1 \partial f_1(\bar{x}) + \ldots + \lambda_m \partial f_m(\bar{x})$.

Comentário 22 *Se as funções f_i forem diferenciáveis em $\bar{x} \in \operatorname{int} C$ então temos* $0 = \nabla f_0(\bar{x}) + \lambda_1 \nabla f_1(\bar{x}) + \ldots + \lambda_m \nabla f_m(\bar{x})$.

Comentário 23 *Note que em geral na teoria do consumidor podemos ter soluções na fronteira de C e então o subgradiente de (c) não precisa coincidir com o gradiente. Esta situação acontece no exemplo abaixo.*

Exemplo 11 *Seja $u(x,y) = 2x + y$ definida para $(x,y) \geq 0$. Sejam $p > 0, q > 0$. O problema (do consumidor) é*

$$\begin{aligned} \max u(x,y), (x,y) &\geq 0, \\ px + qy &\leq 1. \end{aligned} \tag{\$}$$

Para colocar o problema do consumidor no contexto do programa convexo devemos escolher C. Temos $f_0(x,y) = -2x - y + \delta((x,y)|C)$, $f_1(x,y) = px + qy - 1$, $\operatorname{dom} f_1 = \mathbb{R}^2$. *Se $C = \mathbb{R}^2$ temos ainda $f_2(x,y) = -x$ e $f_3(x,y) = -y$. Mas se $C = \mathbb{R}^2_+$ não precisamos de f_2 e f_3. Note para uso posterior que $\partial f_2(x,y) = (-1,0)$ e $\partial f_3(x,y) = (0,-1)$.*

1. *Primeiro ataque: $C = \mathbb{R}^2$. Nesse caso temos três restrições de desigualdade e o problema é*

$$\begin{aligned} \min -2x - y& \\ -x &\leq 0 \\ -y &\leq 0 \\ px + qy - 1 &\leq 0. \end{aligned}$$

O problema pode ser reescrito da forma usual:

$$\begin{aligned} \max 2x + y& \\ x &\geq 0 \\ y &\geq 0 \\ px + qy - 1 &\leq 0. \end{aligned}$$

Os multiplicadores de KT do problema $(\lambda_1, \lambda_2, \lambda_3) \geq 0$, são tais que se $(\bar{x}, \bar{y})$ é solução do problema ótimo,

$$\min -2x - y + \lambda_1(-x) + \lambda_2(-y) + \lambda_3(px + qy - 1)$$

tem solução $(\bar{x}, \bar{y})$ tal que (a,b,c) acima:

$$\begin{aligned} f_i(\bar{x}, \bar{y}) \leq 0 \text{ e } \lambda_i f_i(\bar{x}, \bar{y}) = 0, i = 1,2,3& \\ 0 \in \partial f_0(\bar{x}, \bar{y}) + \lambda_1 \partial f_1(\bar{x}, \bar{y}) + \ldots + \lambda_3 \partial f_3(\bar{x}, \bar{y})&. \end{aligned}$$

Escrevendo como um problema de maximização:

$$\max 2x + y + \lambda_1 x + \lambda_2 y - \lambda_3 (px + qy - 1).$$

E

$$\begin{aligned}
&0 = (-2,-1) + \lambda_1(-1,0) + \lambda_2(0,-1) + \lambda_3(p,q) \implies \\
&0 = -2 - \lambda_1 + \lambda_3 p \\
&0 = -1 - \lambda_2 + \lambda_3 q.
\end{aligned}$$

E $\lambda_1 \bar{x} = 0$, $\lambda_2 \bar{y} = 0$, $\lambda_3 (p\bar{x} + q\bar{y} - 1) = 0$. *É imediato que* $\lambda \neq 0$ *pois senão o ínfimo de* f_0 *seria* $-\infty$. *É necessário considerar casos.*

(a) $\bar{x} > 0$, $\bar{y} > 0$. *Então* $\lambda_1 = 0 = \lambda_2$ *e* $\lambda_3 p = 2 = 2\lambda_3 q \implies p = 2q$ *e* $\lambda_3 = \frac{1}{q}$.

(b) $\bar{x} > 0$ *e* $\bar{y} = 0$. *Então* $\lambda_1 = 0$. *Logo* $\lambda_3 = \frac{2}{p}$ *e* $\lambda_2 = -1 + \frac{2q}{p} \geq 0 \implies 2q \geq p$.

(c) $\bar{x} = 0$ *e* $\bar{y} > 0$. *Então* $\lambda_2 = 0$ *e* $\lambda_3 = \frac{1}{q}$, $\lambda_1 = -2 + \frac{p}{q} \geq 0 \implies p \geq 2q$.

2. *Segundo ataque: Seja* $C = \mathbb{R}^2_+$. $f_0(x,y) = -2x - y$. *E* $f_1(x,y) = px + qy - 1$ *se* $(x,y) \geq 0$. *Temos* $f_1(0,0) = -1 < 0$. *Existe então* $\lambda > 0$ *tal que*

$$\min_{(x,y)\geq 0} -2x - y + \lambda(px + qy - 1)$$

é o ótimo do programa convexo associado. Escrevendo em termos de maximização, se $(\bar{x}, \bar{y}) \geq 0$ *é solução de (\$):*

$$\begin{aligned}
&\max 2x + y - \lambda(px + qy - 1) \\
&(x,y) \geq 0
\end{aligned}$$

e

$$\begin{aligned}
\bar{x} \geq 0, \bar{y} \geq 0, \\
p\bar{x} + q\bar{y} - 1 = 0.
\end{aligned}$$

Note que $2x + y - \lambda(px + qy - 1) = x(2 - \lambda p) + y(1 - \lambda q)$. *Portanto para ter um ótimo é necessário que* $2 \leq \lambda p$ *e* $1 \leq \lambda q$. *Logo* $\max\left\{\frac{2}{p}, \frac{1}{q}\right\} \leq \lambda$. *Mas se tivermos a desigualdade estrita,* $\bar{x} = \bar{y} = 0$ *que não é ótimo. Assim* $\lambda = \max\left\{\frac{2}{p}, \frac{1}{q}\right\}$. *Se* $\frac{2}{p} > \frac{1}{q}$ *então* $\bar{y} = 0$ *e* $\bar{x} = \frac{1}{p}$. *Caso* $\frac{1}{q} > \frac{2}{p}$ *vem* $\bar{x} = 0$ *e* $\bar{y} = \frac{1}{q}$. *Se* $\frac{2}{p} = \frac{1}{q}$ *qualquer combinação* $(\bar{x}, \bar{y}) \geq 0$ *com* $p\bar{x} + q\bar{y} = 1$ *é um ótimo.*

Para escrever em termos de (c) acima devemos calcular o subgradiente de f_0 nos pontos (x, y) com $xy = 0$. (exerc.)

Para usar o item (c) acima precisamos de calcular o subgradiente de f_0 nos pontos (x, y) com $xy = 0$. Por exemplo $z^* \in \partial f_0(0, \bar{y})$ se e somente se

$$-2x - y \geq -\bar{y} + z_1^* x + z_2^* (y - \bar{y}), (x, y) \geq 0 \iff z_2^* = -1 \text{ e } z_1^* \leq -2.$$

Então (c) diz caso $\bar{x} = 0$, $\bar{y} > 0$ (implica $\lambda_2 = 0$).

$$\begin{aligned} 0 &= (z_1^*, -1) + \lambda_1(-1, 0) + \lambda_2(0, -1) + \lambda_3(p, q), z_1^* \leq -2 \\ -z_1^* + \lambda_1 &= \lambda_3 p \\ 1 &= 1 + \lambda_2 = \lambda_3 q \implies \lambda_1 = z_1^* + \frac{p}{q} \implies p \geq 2q. \end{aligned}$$

Fim!