Otimização Contínua: Aspectos teóricos e computacionais

Ademir Alves Ribeiro
Elizabeth Wegner Karas

Capítulo 7 - Otimização com restrições

O problema geral de otimização com restrições

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \leq 0 \end{array}$$

- $f : \mathbb{R}^n \to \mathbb{R}$, $c_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \mathcal{E} \cup \mathcal{I}$: funções $\mathcal{C}^2$

Conjunto viável

$$\Omega = \{x \in \mathbb{R}^n \mid c_{\mathcal{E}}(x) = 0, c_{\mathcal{I}}(x) \leq 0\}$$

CENGAGE Learning

O problema geral de otimização com restrições

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \leq 0 \end{array}$$

- $f : \mathbb{R}^n \to \mathbb{R}$, $c_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \mathcal{E} \cup \mathcal{I}$: funções $\mathcal{C}^2$

Conjunto viável

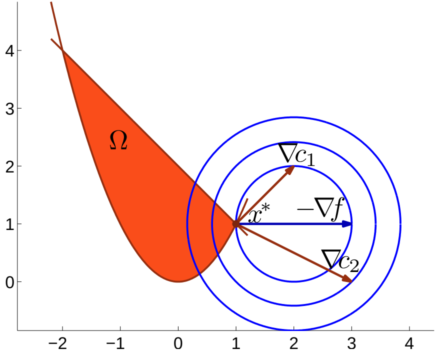
$$\Omega = \{x \in \mathbb{R}^n \mid c_{\mathcal{E}}(x) = 0,\ c_{\mathcal{I}}(x) \leq 0\}$$

CENGAGE Learning

Exemplo

$$\begin{array}{ll} \text{minimizar} & f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{sujeito a} & c_1(x) = x_1 + x_2 - 2 \leq 0 \\ & c_2(x) = x_1^2 - x_2 \leq 0. \end{array}$$

Solução: $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$



Note que $-\nabla f(x^*)$ é uma combinação positiva de $\nabla c_1(x^*)$ e $\nabla c_2(x^*)$

Condições de otimalidade

Teorema (Condições de Karush-Kuhn-Tucker)

Se $x^* \in \Omega$ é um minimizador local para o problema e satisfaz uma condição de qualificação, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I} \mu_i^* \nabla c_i(x^*)$$

$$\mu_i^* \geq 0, \quad i \in I$$

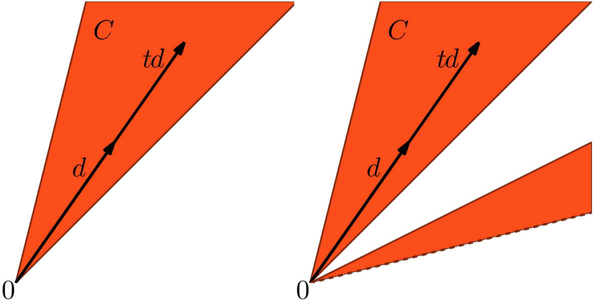
$$\mu_i^* c_i(x^*) = 0, \quad i \in I$$

CENGAGE Learning

Cones: definição e primeiras propriedades

Definição de cone

Um subconjunto $C \subset \mathbb{R}^n$ é um cone quando $td \in C$ para todo $t \geq 0$ e $d \in C$.



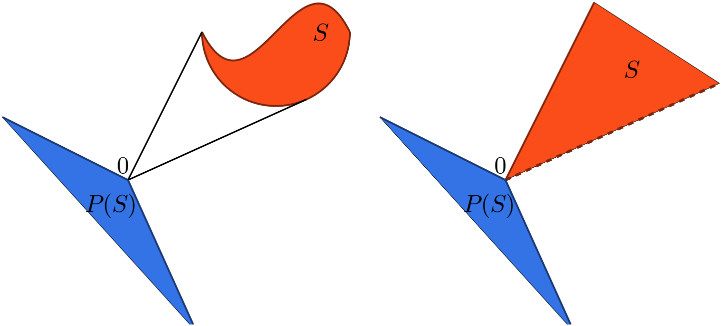
O cone polar

Polar de um conjunto $S \subset \mathbb{R}^n$

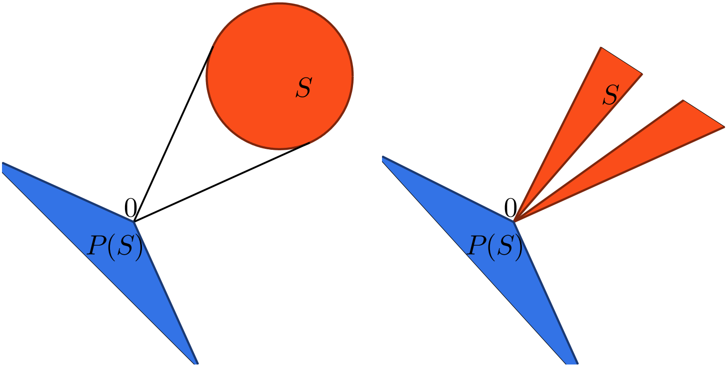
$$P(S) = \left\{ p \in \mathbb{R}^n \mid p^T x \leq 0,\ \forall x \in S \right\}$$

Polar de um conjunto $S \subset \mathbb{R}^n$

$$P(S) = \left\{ p \in \mathbb{R}^n \mid p^T x \le 0, \; \forall x \in S \right\}$$



Outros exemplos de polar



Exercício

Representar geometricamente o polar dos seguintes conjuntos:

- Semireta na origem
- Reta pela origem
- Semiplano $x \geq 0$
- Região com ângulo de 240 graus

Exercício

Considere $A = \begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix}$ e $C = \left\{ d \in \mathbb{R}^2 \mid Ad \leq 0 \right\}$. Pede-se:

- Mostre que C é cone.
- Represente C geometricamente.
- Diga se C pode ser obtido como o polar de algum conjunto.
- Determine $P(C)$.

Propriedades do cone polar

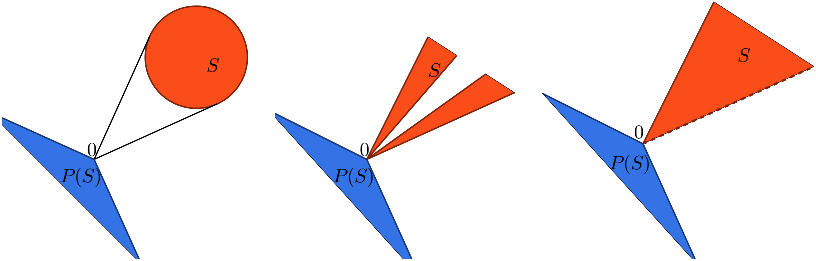
Lema

Dado $S \subset \mathbb{R}^n$, $P(S)$ é cone, convexo e fechado.

Propriedades do cone polar

Lema

Dado $S \subset \mathbb{R}^n$, $P(S)$ é cone, convexo e fechado.



Propriedades do cone polar

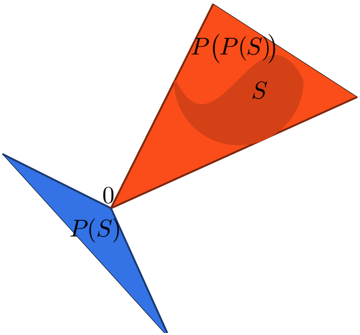
Lema

Dado $S \subset \mathbb{R}^n$, temos $S \subset P\big(P(S)\big)$.

Propriedades do cone polar

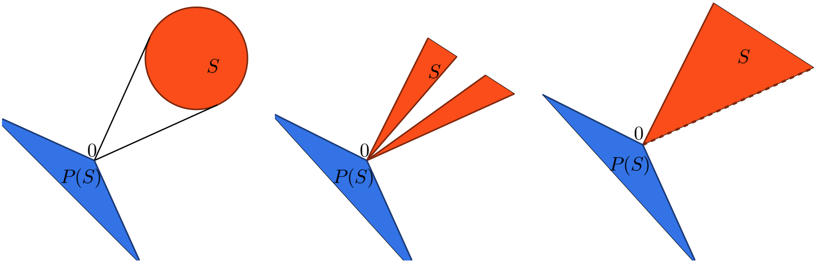
Lema

Dado $S \subset \mathbb{R}^n$, temos $S \subset P\big(P(S)\big)$.



Nem sempre vale a igualdade $P\big(P(S)\big) = S$.

- Por não ser cone
- Por não ser convexo
- Por não ser fechado



Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

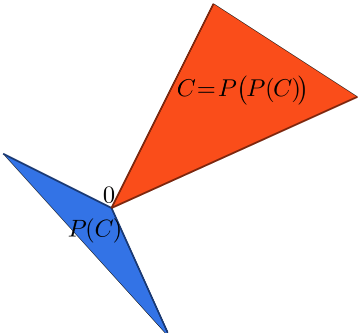
$$P\big(P(C)\big) = C.$$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$



Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\big(P(C)\big)$, $\bar{z} = \text{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \le 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \le 0$
- $z-\bar{z} \in P(C)$
- $(z-\bar{z})^T z \le 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \le 0$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\left(P(C)\right)$, $\bar{z} = \text{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \le 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \le 0$
- $z-\bar{z} \in P(C)$
- $(z-\bar{z})^T z \le 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \le 0$

CENGAGE Learning

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\left(P(C)\right)$, $\bar{z} = \mathrm{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \le 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \le 0$
- $z - \bar{z} \in P(C)$
- $(z-\bar{z})^T z \le 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \le 0$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\left(P(C)\right)$, $\bar{z} = \mathrm{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \le 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \le 0$
- $z - \bar{z} \in P(C)$
- $(z-\bar{z})^T z \le 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \le 0$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\left(P(C)\right)$, $\bar{z} = \mathrm{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \leq 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \leq 0$
- $z - \bar{z} \in P(C)$
- $(z-\bar{z})^T z \leq 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \leq 0$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

- $z \in P\left(P(C)\right)$, $\bar{z} = \text{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \leq 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \leq 0$
- $z - \bar{z} \in P(C)$
- $(z-\bar{z})^T z \leq 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \leq 0$

Lema de Farkas

Lema

Seja $C \subset \mathbb{R}^n$ um cone convexo fechado não vazio. Então

$$P\big(P(C)\big) = C.$$

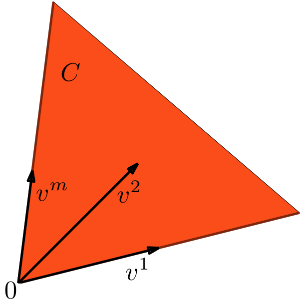
- $z \in P\left(P(C)\right)$, $\bar{z} = \text{proj}_C(z)$
- $(z-\bar{z})^T(x-\bar{z}) \leq 0$, $x = 0$ e $x = 2\bar{z}$
- $(z-\bar{z})^T\bar{z} = 0 \Rightarrow (z-\bar{z})^T x \leq 0$
- $z-\bar{z} \in P(C)$
- $(z-\bar{z})^T z \leq 0$
- $\|z-\bar{z}\|^2 = (z-\bar{z})^T z \leq 0$

Lema

Dados os vetores $v^1, v^2, \ldots, v^m \in \mathbb{R}^n \setminus \{0\}$, o conjunto

$$C = \left\{ \sum_{i=1}^{m} y_i v^i \mid y_i \geq 0,\ i = 1, \ldots, m \right\}$$

é um cone convexo e fechado.



Versão algébrica do Lema de Farkas

Lema

Dados $A \in \mathbb{R}^m n$ e $c \in \mathbb{R}^n$, exatamente um dos dois sistemas abaixo tem solução.

$$Ax \leq 0 \quad \text{e} \quad c^T x > 0 \qquad (1)$$

$$A^T y = c \quad \text{e} \quad y \geq 0 \qquad (2)$$

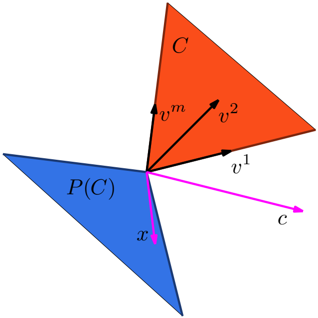
Versão algébrica do Lema de Farkas

Lema

Dados $A \in \mathbb{R}^m n$ e $c \in \mathbb{R}^n$, exatamente um dos dois sistemas abaixo tem solução.

$$Ax \leq 0 \quad \text{e} \quad c^T x > 0 \qquad (1)$$

$$A^T y = c \quad \text{e} \quad y \geq 0 \qquad (2)$$



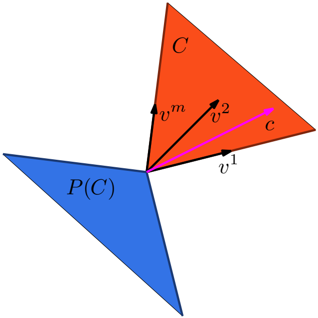
Versão algébrica do Lema de Farkas

Lema

Dados $A \in \mathbb{R}^m n$ e $c \in \mathbb{R}^n$, exatamente um dos dois sistemas abaixo tem solução.

$$Ax \leq 0 \quad \text{e} \quad c^T x > 0 \qquad (1)$$

$$A^T y = c \quad \text{e} \quad y \geq 0 \qquad (2)$$



Condições de Karush-Kuhn-Tucker

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \le 0 \end{array}$$

- $f : \mathbb{R}^n \to \mathbb{R}$, $c_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \mathcal{E} \cup \mathcal{I}$: funções $\mathcal{C}^2$

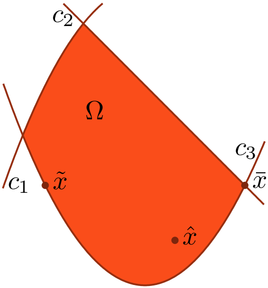
Conjunto viável

$$\Omega = \{x \in \mathbb{R}^n \mid c_{\mathcal{E}}(x) = 0,\, c_{\mathcal{I}}(x) \le 0\}$$

CENGAGE Learning

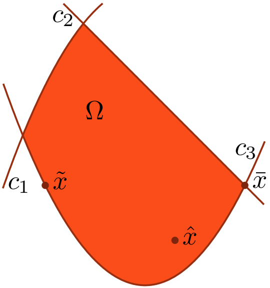
O conjunto ativo

$$I(x) = \{i \in I \mid c_i(x) = 0\}$$



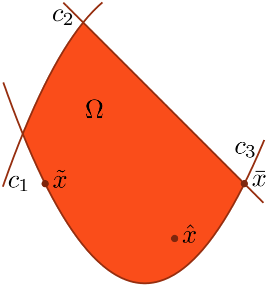
O conjunto ativo

$$I(\hat{x}) = \emptyset$$



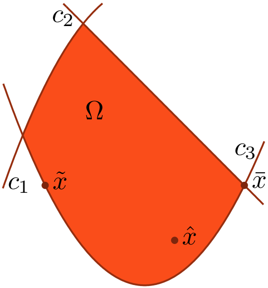
O conjunto ativo

$$I(\tilde{x}) = \{3\}$$



O conjunto ativo

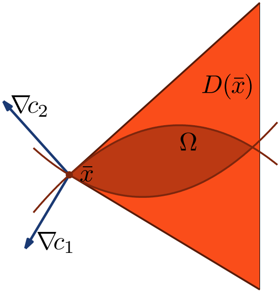
$$I(\bar{x}) = \{2,3\}$$



O cone viável linearizado

Definição

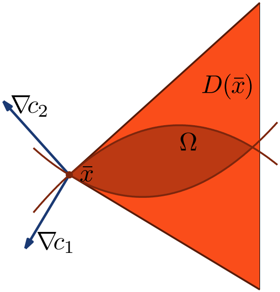
$$D(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \nabla c_i(\bar{x})^T d = 0, \text{ se } i \in \mathcal{E} \text{ e } \nabla c_i(\bar{x})^T d \leq 0, \text{ se } i \in I(\bar{x}) \right\}$$



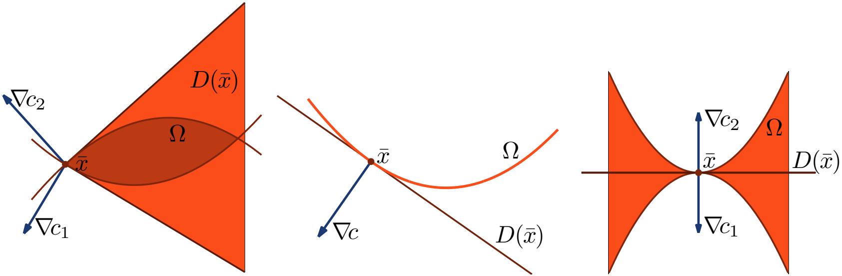
Conjunto viável para a aproximação linear das restrições

Definição

$$D(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \nabla c_i(\bar{x})^T d = 0, \text{ se } i \in \mathcal{E} \text{ e } \nabla c_i(\bar{x})^T d \leq 0, \text{ se } i \in I(\bar{x}) \right\}$$



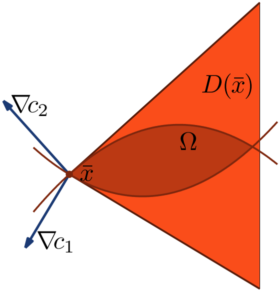
Pode ou não aproximar “bem” o conjunto viável



Propriedade

Lema

O conjunto $D(\bar{x})$ é um cone convexo fechado não vazio



Exercício

Determine o cone viável linearizado em torno de $\bar{x} = 0 \in \mathbb{R}^2$ e represente geometricamente o conjunto viável e sua linearização.

a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \end{cases}$

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

CENGAGE Learning

Resolução

a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid -2d_1 \leq d_2 \leq 2d_1 \right\}$

Resolução

a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \le 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \le 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \le 0$ e $\nabla c_2(\bar{x})^T d \le 0$
- $D(\bar{x}) = \{d \in \mathbb{R}^2 \mid -2d_1 \le d_2 \le 2d_1\}$

CENGAGE Learning

Resolução

a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid -2d_1 \leq d_2 \leq 2d_1 \right\}$

CENGAGE Learning

Resolução

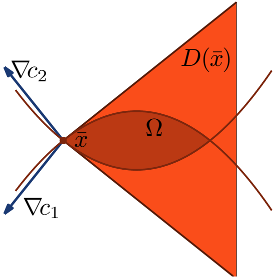
a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid -2d_1 \leq d_2 \leq 2d_1 \right\}$

Resolução

a) $\begin{cases} c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \end{cases}$

- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid -2d_1 \leq d_2 \leq 2d_1 \right\}$



Resolução

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \{d \in \mathbb{R}^2 \mid d_2 = 0\}$

Resolução

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \{d \in \mathbb{R}^2 \mid d_2 = 0\}$

Resolução

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

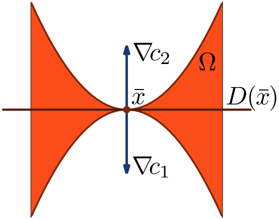
- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \{d \in \mathbb{R}^2 \mid d_2 = 0\}$

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$
- $\nabla c_1(\bar{x})^T d \leq 0$ e $\nabla c_2(\bar{x})^T d \leq 0$
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_2 = 0 \right\}$

b) $\begin{cases} c_1(x) = -x_1^2 + x_2 \leq 0 \\ c_2(x) = -x_1^2 - x_2 \leq 0 \end{cases}$

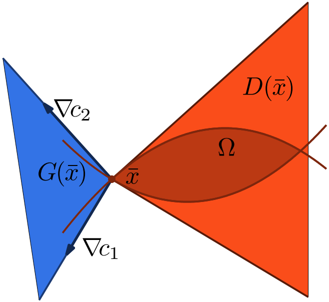
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_2 = 0 \right\}$



O cone gerado pelos gradientes das restrições

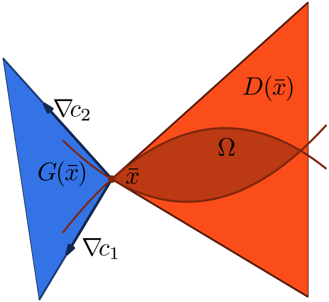
Definição

$$G(\bar{x}) = \left\{ \sum_{i \in \mathcal{E}} \lambda_i \nabla c_i(\bar{x}) + \sum_{i \in I(\bar{x})} \mu_i \nabla c_i(\bar{x}) \mid \mu_i \geq 0, \forall i \in I(\bar{x}) \right\}$$



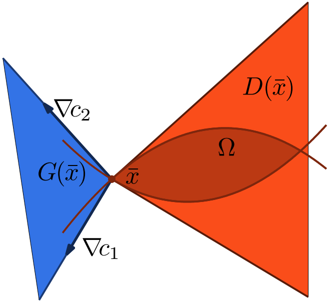
Propriedades

- $D(\bar{x}) = P\big(G(\bar{x})\big)$
- $G(\bar{x})$ é um cone convexo fechado
- $P\big(D(\bar{x})\big) = G(\bar{x})$ (Chave da prova de KKT)



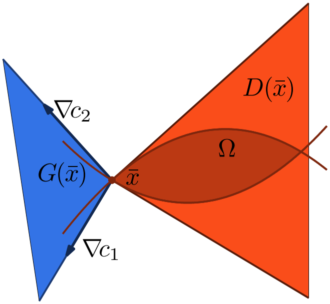
Propriedades

- $D(\bar{x}) = P\big(G(\bar{x})\big)$
- $G(\bar{x})$ é um cone convexo fechado
- $P\big(D(\bar{x})\big) = G(\bar{x})$ (Chave da prova de KKT)



Propriedades

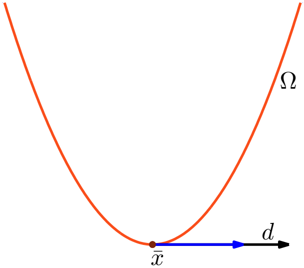
- $D(\bar{x}) = P\big(G(\bar{x})\big)$
- $G(\bar{x})$ é um cone convexo fechado
- $P\big(D(\bar{x})\big) = G(\bar{x})$ (Chave da prova de KKT)



Direções tangentes

$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

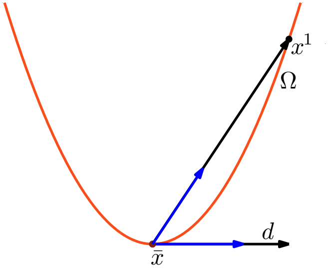
$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



Direções tangentes

$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

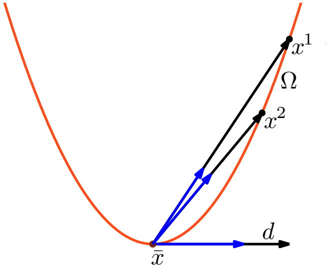
$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



Direções tangentes

$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

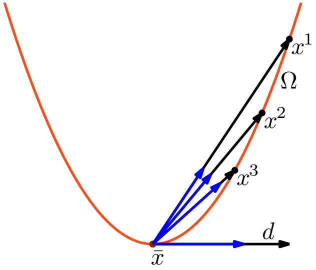
$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



Direções tangentes

$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

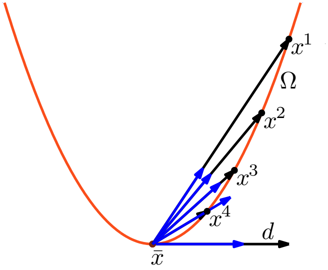
$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



Direções tangentes

$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

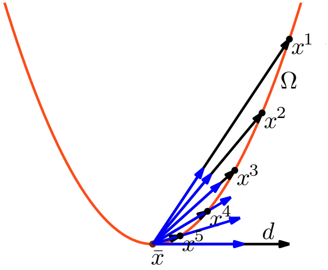
$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



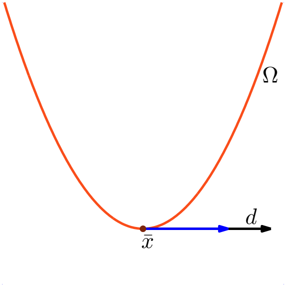
Direções tangentes

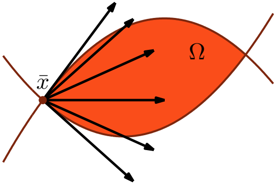
$d \in \mathbb{R}^n$ é tangente a Ω em $\bar{x}$ quando existe $(x^k) \subset \Omega$ tal que

$$x^k \to \bar{x} \qquad \text{e} \qquad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}$$



Direções tangentes

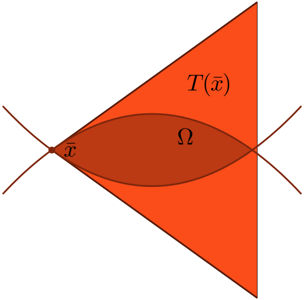


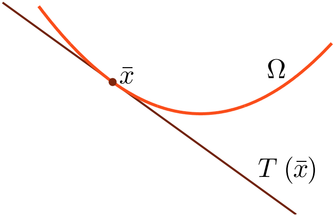


O cone tangente

Definição

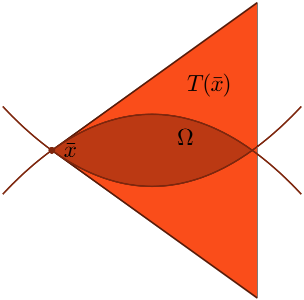
$$T(\bar{x}) = \{0\} \cup \left\{ d \in \mathbb{R}^n \mid \exists (x^k) \subset \Omega \, ; \, x^k \to \bar{x} \, , \, \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|} \right\}$$

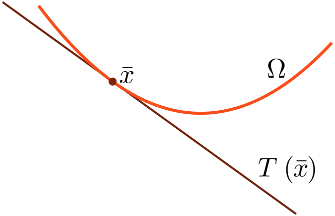




Propriedades

- $T(\bar{x}) \subset D(\bar{x})$
- $T(\bar{x})$ é um cone fechado (não necessariamente convexo)





Ideia para provar que $T(\bar{x}) \subset D(\bar{x})$

- $c_\ell(x^k) = c_\ell(\bar{x}) + \nabla c_\ell(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \quad \ell \in \mathcal{E} \cup I(\bar{x})$
- $\nabla c_i(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0, \quad i \in \mathcal{E}$
- $\nabla c_j(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \leq 0, \quad j \in I(\bar{x})$
- Passando o limite, $\nabla c_i(\bar{x})^T \dfrac{d}{\|d\|} = 0$ e $\nabla c_j(\bar{x})^T \dfrac{d}{\|d\|} \leq 0$

Ideia para provar que $T(\bar{x}) \subset D(\bar{x})$

- $c_\ell(x^k) = c_\ell(\bar{x}) + \nabla c_\ell(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \quad \ell \in \mathcal{E} \cup I(\bar{x})$
- $\nabla c_i(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0, \quad i \in \mathcal{E}$
- $\nabla c_j(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \le 0, \quad j \in I(\bar{x})$
- Passando o limite, $\nabla c_i(\bar{x})^T \dfrac{d}{\|d\|} = 0$ e $\nabla c_j(\bar{x})^T \dfrac{d}{\|d\|} \le 0$

Ideia para provar que $T(\bar{x}) \subset D(\bar{x})$

- $c_\ell(x^k) = c_\ell(\bar{x}) + \nabla c_\ell(\bar{x})^T(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \quad \ell \in \mathcal{E} \cup I(\bar{x})$
- $\nabla c_i(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0, \quad i \in \mathcal{E}$
- $\nabla c_j(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \leq 0, \quad j \in I(\bar{x})$
- Passando o limite, $\nabla c_i(\bar{x})^T \dfrac{d}{\|d\|} = 0$ e $\nabla c_j(\bar{x})^T \dfrac{d}{\|d\|} \leq 0$

Ideia para provar que $T(\bar{x}) \subset D(\bar{x})$

- $c_\ell(x^k) = c_\ell(\bar{x}) + \nabla c_\ell(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \quad \ell \in \mathcal{E} \cup I(\bar{x})$
- $\nabla c_i(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0, \quad i \in \mathcal{E}$
- $\nabla c_j(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \leq 0, \quad j \in I(\bar{x})$
- Passando o limite, $\nabla c_i(\bar{x})^T \dfrac{d}{\|d\|} = 0$ e $\nabla c_j(\bar{x})^T \dfrac{d}{\|d\|} \leq 0$

Ideia para provar que $T(\bar{x}) \subset D(\bar{x})$

- $c_\ell(x^k) = c_\ell(\bar{x}) + \nabla c_\ell(\bar{x})^T(x^k - \bar{x}) + o(\|x^k - \bar{x}\|), \quad \ell \in \mathcal{E} \cup I(\bar{x})$
- $\nabla c_i(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0, \quad i \in \mathcal{E}$
- $\nabla c_j(\bar{x})^T \dfrac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \dfrac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \leq 0, \quad j \in I(\bar{x})$
- Passando o limite, $\nabla c_i(\bar{x})^T \dfrac{d}{\|d\|} = 0$ e $\nabla c_j(\bar{x})^T \dfrac{d}{\|d\|} \leq 0$

Exemplo onde $T(\bar{x}) \neq D(\bar{x})$ e $T(\bar{x})$ não é convexo

$$\Omega : \begin{cases} c_1(x) = x_1 x_2 = 0 \\ c_2(x) = -x_1 \leq 0 \\ c_3(x) = -x_2 \leq 0 \end{cases}$$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\nabla c_2(\bar{x}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, e $\nabla c_3(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$,
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \geq 0,\ d_2 \geq 0 \right\}$
- $G(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \leq 0,\ d_2 \leq 0 \right\}$
- $T(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \geq 0,\ d_2 \geq 0,\ d_1 d_2 = 0 \right\}$

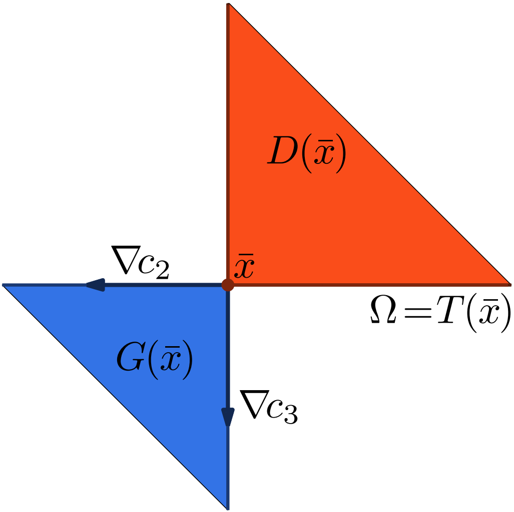
Exemplo onde $T(\bar{x}) \neq D(\bar{x})$ e $T(\bar{x})$ não é convexo

$$\Omega : \begin{cases} c_1(x) = x_1 x_2 = 0 \\ c_2(x) = -x_1 \leq 0 \\ c_3(x) = -x_2 \leq 0 \end{cases}$$

- $\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\nabla c_2(\bar{x}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, e $\nabla c_3(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$,
- $D(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \geq 0,\ d_2 \geq 0 \right\}$
- $G(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \leq 0,\ d_2 \leq 0 \right\}$
- $T(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 \geq 0,\ d_2 \geq 0,\ d_1 d_2 = 0 \right\}$

Exemplo onde $T(\bar{x}) \neq D(\bar{x})$ e $T(\bar{x})$ não é convexo

$$\Omega : \begin{cases} c_1(x) = x_1 x_2 = 0 \\ c_2(x) = -x_1 \leq 0 \\ c_3(x) = -x_2 \leq 0 \end{cases}$$



Uma condição necessária de otimalidade

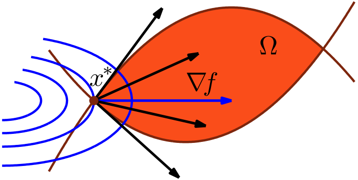
Lema

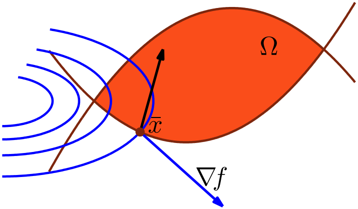
Se $x^* \in \Omega$ é uma solução local do problema de otimização, então $\nabla f(x^*)^T d \geq 0$, para todo $d \in T(x^*)$.

Uma condição necessária de otimalidade

Lema

Se $x^* \in \Omega$ é uma solução local do problema de otimização, então $\nabla f(x^*)^T d \geq 0$, para todo $d \in T(x^*)$.





O teorema de KKT

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in \mathcal{I}} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in \mathcal{I},$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in \mathcal{I}.$$

- $-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*)$
- $-\nabla f(x^*) \in P\big(T(x^*)\big)$
- $-\nabla f(x^*) \in P\big(D(x^*)\big)$
- $-\nabla f(x^*) \in G(x^*)$

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in I,$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in I.$$

- $-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*)$
- $-\nabla f(x^*) \in P\big(T(x^*)\big)$
- $-\nabla f(x^*) \in P\big(D(x^*)\big)$
- $-\nabla f(x^*) \in G(x^*)$

CENGAGE Learning

O teorema de KKT

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in \mathcal{I}} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in \mathcal{I},$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in \mathcal{I}.$$

- $-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*)$
- $-\nabla f(x^*) \in P\big(T(x^*)\big)$
- $-\nabla f(x^*) \in P\big(D(x^*)\big)$
- $-\nabla f(x^*) \in G(x^*)$

CENGAGE Learning

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in I,$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in I.$$

- $-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*)$
- $-\nabla f(x^*) \in P\big(T(x^*)\big)$
- $-\nabla f(x^*) \in P\big(D(x^*)\big)$
- $-\nabla f(x^*) \in G(x^*)$

CENGAGE Learning

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in I,$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in I.$$

- $-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*)$
- $-\nabla f(x^*) \in P\big(T(x^*)\big)$
- $-\nabla f(x^*) \in P\big(D(x^*)\big)$
- $-\nabla f(x^*) \in G(x^*)$

O teorema de KKT

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in I,$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in I.$$

$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i \nabla c_i(x^*) + \sum_{i \in I(x^*)} \mu_i \nabla c_i(x^*)$$

CENGAGE Learning

Teorema

Se $x^* \in \Omega$ é solução local do problema e $P\big(T(x^*)\big) = P\big(D(x^*)\big)$, então existem vetores λ^* e μ^* tais que

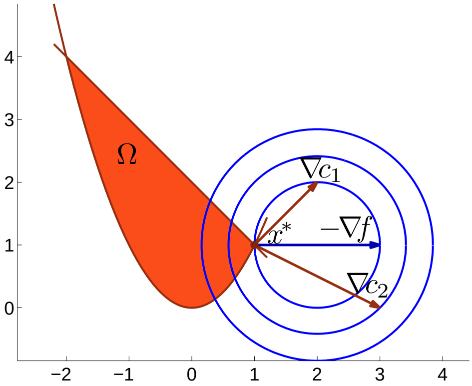
$$-\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in \mathcal{I}} \mu_i^* \nabla c_i(x^*),$$

$$\mu_i^* \geq 0, \quad i \in \mathcal{I},$$

$$\mu_i^* c_i(x^*) = 0, \quad i \in \mathcal{I}.$$

x^* é dito ponto estacionário.

$$\begin{array}{ll} \text{minimizar} & f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{sujeito a} & c_1(x) = x_1 + x_2 - 2 \leq 0 \\ & c_2(x) = x_1^2 - x_2 \leq 0. \end{array}$$



Voltando ao exemplo

- $x_1^2 \leq x_2 \leq 2 - x_1 \Rightarrow -2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \geq 0$ e $\mu_i c_i(x) = 0,\ i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \leq -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2$: $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\bar{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \le x_2 \le 2 - x_1 \Rightarrow -2 \le x_1 \le 1$ e $0 \le x_2 \le 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \ge 0$ e $\mu_i c_i(x) = 0,\ i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \le -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2 : \quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \bar{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \le x_2 \le 2 - x_1 \Rightarrow -2 \le x_1 \le 1$ e $0 \le x_2 \le 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \ge 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \le -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2$: $\quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \leq x_2 \leq 2 - x_1 \Rightarrow -2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \geq 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \leq -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2$: $\quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \leq x_2 \leq 2 - x_1 \Rightarrow -2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$
- $T(x) = D(x) \Rightarrow -2 \begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \geq 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \leq -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2 : \quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \le x_2 \le 2 - x_1 \Rightarrow -2 \le x_1 \le 1$ e $0 \le x_2 \le 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \ge 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \le -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2$: $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \leq x_2 \leq 2 - x_1 \Rightarrow -2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \geq 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \leq -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2 : \quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Voltando ao exemplo

- $x_1^2 \leq x_2 \leq 2 - x_1 \Rightarrow -2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$
- $T(x) = D(x) \Rightarrow -2\begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$
- $\mu_i \geq 0$ e $\mu_i c_i(x) = 0$, $i = 1, 2$
- $\mu_1 = 0$ e $\mu_2 = 0 \Rightarrow x_1 = 2$
- $\mu_1 = 0$ e $\mu_2 > 0 \Rightarrow x_2 > 1$ e $x_2 = x_1^2 \Rightarrow x_1 \leq -1$
- $\mu_1 > 0$ e $\mu_2 = 0 \Rightarrow x_1 + x_2 = 2$ e $x_1 - 2 = x_2 - 1 \Rightarrow x_1 = \frac{3}{2}$
- $x_1 + x_2 = 2$ e $x_1^2 = x_2$: $\quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

Solução:

$$x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$$

Caracterizações alternativas de estacionariedade

Considere $\bar{x} \in \mathbb{R}^n$ e a aproximação linear do conjunto viável

$$\mathcal{L}(\bar{x}) = \{\bar{x} + d \in \mathbb{R}^n \mid c_{\mathcal{E}}(\bar{x}) + A_{\mathcal{E}}(\bar{x})d = 0\,,\ c_{\mathcal{I}}(\bar{x}) + A_{\mathcal{I}}(\bar{x})d \leq 0\}\,,$$

onde $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$ denotam as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$, respectivamente.
Note que $\mathcal{L}(\bar{x})$ é um conjunto convexo e fechado.

Direção do gradiente projetado

$$d^c(\bar{x}) = \mathrm{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}.$$

Caracterizações alternativas de estacionariedade

Considere $\bar{x} \in \mathbb{R}^n$ e a aproximação linear do conjunto viável

$$\mathcal{L}(\bar{x}) = \{\bar{x} + d \in \mathbb{R}^n \mid c_{\mathcal{E}}(\bar{x}) + A_{\mathcal{E}}(\bar{x})d = 0 \, , \; c_{\mathcal{I}}(\bar{x}) + A_{\mathcal{I}}(\bar{x})d \leq 0\} \, ,$$

onde $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$ denotam as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$, respectivamente.
Note que $\mathcal{L}(\bar{x})$ é um conjunto convexo e fechado.

Direção do gradiente projetado

$$d^c(\bar{x}) = \text{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}.$$

Caracterizações alternativas de estacionariedade

Considere $\bar{x} \in \mathbb{R}^n$ e a aproximação linear do conjunto viável

$$\mathcal{L}(\bar{x}) = \{\bar{x} + d \in \mathbb{R}^n \mid c_{\mathcal{E}}(\bar{x}) + A_{\mathcal{E}}(\bar{x})d = 0\ ,\ c_{\mathcal{I}}(\bar{x}) + A_{\mathcal{I}}(\bar{x})d \leq 0\},$$

onde $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$ denotam as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$, respectivamente.
Note que $\mathcal{L}(\bar{x})$ é um conjunto convexo e fechado.

Direção do gradiente projetado

$$d^c(\bar{x}) = \mathrm{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}.$$

Teorema

Considere um ponto viável $\bar{x} \in \Omega$ e a direção do gradiente projetado $d^c(\bar{x}) = \text{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}$.

Teorema

- $\bar{x}$ cumpre as condições de KKT se, e somente se, $d^c(\bar{x}) = 0$.
- Se $\bar{x}$ não é estacionário, então $\nabla f(\bar{x})^T d^c(\bar{x}) < 0$.

Considere um ponto viável $\bar{x} \in \Omega$ e a direção do gradiente projetado $d^c(\bar{x}) = \text{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}$.

Teorema

- $\bar{x}$ cumpre as condições de KKT se, e somente se, $d^c(\bar{x}) = 0$.
- Se $\bar{x}$ não é estacionário, então $\nabla f(\bar{x})^T d^c(\bar{x}) < 0$.

Teorema

Considere um ponto viável $\bar{x} \in \Omega$ e a direção do gradiente projetado $d^c(\bar{x}) = \text{proj}_{\mathcal{L}(\bar{x})}\big(\bar{x} - \nabla f(\bar{x})\big) - \bar{x}$.

Teorema

- $\bar{x}$ cumpre as condições de KKT se, e somente se, $d^c(\bar{x}) = 0$.
- Se $\bar{x}$ não é estacionário, então $\nabla f(\bar{x})^T d^c(\bar{x}) < 0$.

$d^c(\bar{x}) = 0$ não é uma condição necessária de otimalidade

Exemplo em $\mathbb{R}^2$

Minimizar $f(x) = x_1 + 2x_2$ no conjunto

$$\Omega = \left\{ x \in \mathbb{R}^2 \mid c_1(x) = 0\,,\; c_2(x) \le 0 \right\},$$

onde $c_1(x) = x_1 x_2$, $c_2(x) = -x_1 - x_2$.

Verifique que o ponto $\bar{x} = 0$ é uma solução global, mas $d^c(\bar{x}) \neq 0$.

Solução

- Qualquer ponto viável, que não seja $\bar{x}$, tem uma componente nula e a outra positiva. Portanto, $\bar{x} = 0$ é o minimizador global de f em Ω.
- Além disso, temos

$$\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\; \nabla c_2(\bar{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla f(\bar{x}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Assim, $\mathcal{L}(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 + d_2 \ge 0 \right\}$ e

$$\bar{z} = \mathrm{proj}_{\mathcal{L}(\bar{x})}\left(\bar{x} - \nabla f(\bar{x})\right) = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \neq \bar{x}.$$

CENGAGE Learning

$d^c(\bar{x}) = 0$ não é uma condição necessária de otimalidade

Exemplo em $\mathbb{R}^2$

Minimizar $f(x) = x_1 + 2x_2$ no conjunto

$$\Omega = \left\{ x \in \mathbb{R}^2 \mid c_1(x) = 0 \ , \ c_2(x) \le 0 \right\},$$

onde $c_1(x) = x_1 x_2$, $c_2(x) = -x_1 - x_2$.

Verifique que o ponto $\bar{x} = 0$ é uma solução global, mas $d^c(\bar{x}) \neq 0$.

Solução

- Qualquer ponto viável, que não seja $\bar{x}$, tem uma componente nula e a outra positiva. Portanto, $\bar{x} = 0$ é o minimizador global de f em Ω.
- Além disso, temos

$$\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \nabla c_2(\bar{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla f(\bar{x}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Assim, $\mathcal{L}(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 + d_2 \ge 0 \right\}$ e

$$\bar{z} = \text{proj}_{\mathcal{L}(\bar{x})}\left(\bar{x} - \nabla f(\bar{x})\right) = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \neq \bar{x}.$$

CENGAGE Learning

$d^c(\bar{x}) = 0$ não é uma condição necessária de otimalidade

Exemplo em $\mathbb{R}^2$

Minimizar $f(x) = x_1 + 2x_2$ no conjunto

$$\Omega = \left\{ x \in \mathbb{R}^2 \mid c_1(x) = 0\ ,\ c_2(x) \leq 0 \right\},$$

onde $c_1(x) = x_1 x_2$, $c_2(x) = -x_1 - x_2$.

Verifique que o ponto $\bar{x} = 0$ é uma solução global, mas $d^c(\bar{x}) \neq 0$.

Solução

- Qualquer ponto viável, que não seja $\bar{x}$, tem uma componente nula e a outra positiva. Portanto, $\bar{x} = 0$ é o minimizador global de f em Ω.
- Além disso, temos

$$\nabla c_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \nabla c_2(\bar{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla f(\bar{x}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Assim, $\mathcal{L}(\bar{x}) = \left\{ d \in \mathbb{R}^2 \mid d_1 + d_2 \geq 0 \right\}$ e

$$\bar{z} = \operatorname{proj}_{\mathcal{L}(\bar{x})}\left(\bar{x} - \nabla f(\bar{x})\right) = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \neq \bar{x}.$$

CENGAGE Learning

$d^c(\bar{x}) = 0$ não é uma condição necessária de otimalidade

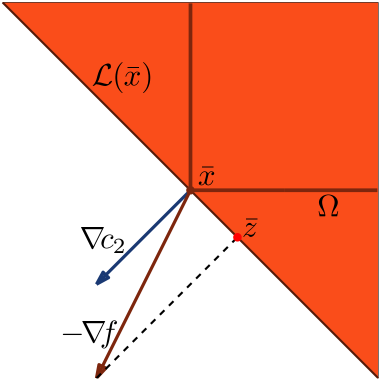
Exemplo em $\mathbb{R}^2$

Minimizar $f(x) = x_1 + 2x_2$ no conjunto

$$\Omega = \left\{ x \in \mathbb{R}^2 \mid c_1(x) = 0 \ , \ c_2(x) \leq 0 \right\},$$

onde $c_1(x) = x_1 x_2$, $c_2(x) = -x_1 - x_2$.

Verifique que o ponto $\bar{x} = 0$ é uma solução global, mas $d^c(\bar{x}) \neq 0$.



Outra medida de estacionariedade

Medida de estacionariedade de um ponto viável:

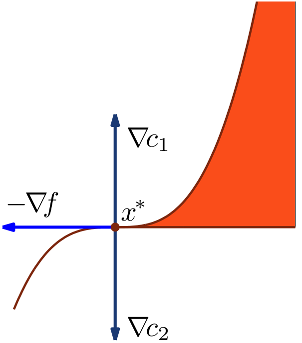
$$\chi(x) = \left| \min_{x+d \in \mathcal{L}(x),\, \|d\| \leq 1} \nabla f(x)^T d \right|.$$

Teorema

Dado $\bar{x} \in \Omega$, temos $\chi(\bar{x}) \leq \left\| \mathrm{proj}_{D(\bar{x})}\left(-\nabla f(\bar{x})\right) \right\|$.
Em particular, se $\bar{x}$ satisfaz KKT, então $\chi(\bar{x}) = 0$.

Importância de uma condição de qualificação

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_2 \leq 0.
\end{array}
$$



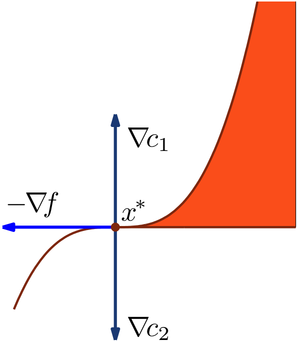
- $x^* = 0$ é o minimizador deste problema
  de $0 \leq x_2 \leq x_1^3$, segue que $f(x) = x_1 \geq 0 = f(x^*)$, para todo x viável.
- $x^* = 0$ não satisfaz as condições de KKT

$$\nabla f(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

- $P(T(x^*)) \neq P(D(x^*))$.

Importância de uma condição de qualificação

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_2 \leq 0.
\end{array}
$$



- $x^* = 0$ é o minimizador deste problema

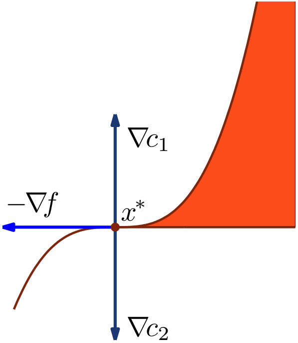
  de $0 \leq x_2 \leq x_1^3$, segue que $f(x) = x_1 \geq 0 = f(x^*)$, para todo x viável.
- $x^* = 0$ não satisfaz as condições de KKT

$$\nabla f(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

- $P(T(x^*)) \neq P(D(x^*))$.

Importância de uma condição de qualificação

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_2 \leq 0.
\end{array}
$$



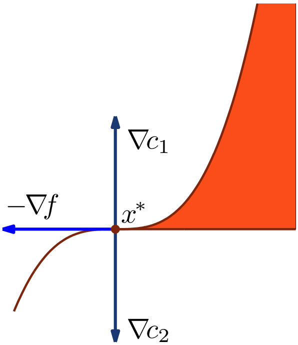
- $x^* = 0$ é o minimizador deste problema
  de $0 \leq x_2 \leq x_1^3$, segue que $f(x) = x_1 \geq 0 = f(x^*)$, para todo x viável.
- $x^* = 0$ não satisfaz as condições de KKT

$$
\nabla f(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.
$$

- $P(T(x^*)) \neq P(D(x^*))$.

Importância de uma condição de qualificação

$$\begin{array}{ll} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = -x_1^3 + x_2 \le 0 \\ & c_2(x) = -x_2 \le 0. \end{array}$$



- $x^* = 0$ é o minimizador deste problema
  de $0 \le x_2 \le x_1^3$, segue que $f(x) = x_1 \ge 0 = f(x^*)$, para todo x viável.
- $x^* = 0$ não satisfaz as condições de KKT

$$\nabla f(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

- $P(T(x^*)) \ne P(D(x^*))$.

CENGAGE Learning

Condições de Qualificação

Definição

As restrições $c_{\mathcal{E}}(x) = 0$ e $c_{\mathcal{I}}(x) \leq 0$ cumprem uma condição de qualificação em $x^* \in \Omega$ quando, dada qualquer função diferenciável f, que tenha mínimo em x^*, relativamente a Ω, sejam satisfeitas as condições de otimalidade de KKT.

minimizador local + condição de qualificação $\Longrightarrow$ KKT.

Condições de Qualificação

Definição

As restrições $c_{\mathcal{E}}(x) = 0$ e $c_I(x) \leq 0$ cumprem uma condição de qualificação em $x^* \in \Omega$ quando, dada qualquer função diferenciável f, que tenha mínimo em x^*, relativamente a Ω, sejam satisfeitas as condições de otimalidade de KKT.

minimizador local + condição de qualificação $\Longrightarrow$ KKT.

Condição de Qualificação de Indepenência Linear

LICQ

A condição de qualificação de independência linear (LICQ) é satisfeita em $\bar{x}$ quando o conjunto formado pelos gradientes das restrições de igualdade e das restrições de desigualdade ativas é linearmente independente, isto é,

$$\{\nabla c_i(\bar{x}) \mid i \in \mathcal{E} \cup I(\bar{x})\} \text{ é LI.}$$

Considere duas restrições de desigualdades definidas por $c_1, c_2 : \mathbb{R}^2 \to \mathbb{R}$, onde $c_1(x) = x_1^2 - 2x_1 - x_2$ e $c_2(x) = x_1^2 - 2x_1 + x_2$.
Verifique que o ponto $\bar{x} = 0$ cumpre LICQ.

Solução:

As duas restrições são ativas em $\bar{x}$ e os vetores $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ são linearmente independentes.

CENGAGE Learning

Condição de Qualificação de Indepenência Linear

Considere duas restrições de desigualdades definidas por $c_1, c_2 : \mathbb{R}^2 \to \mathbb{R}$, onde $c_1(x) = x_1^2 - 2x_1 - x_2$ e $c_2(x) = x_1^2 - 2x_1 + x_2$.
Verifique que o ponto $\bar{x} = 0$ cumpre LICQ.

Solução:

As duas restrições são ativas em $\bar{x}$ e os vetores $\nabla c_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla c_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ são linearmente independentes.

Exemplo em que vale KKT sem que LICQ seja satisfeita

$$\begin{array}{ll} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ & c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \\ & c_3(x) = -x_1 \leq 0. \end{array}$$

O ponto $x^* = 0$ é o minimizador deste problema, cumpre as condições de KKT mas não satisfaz LICQ.

Solução:

- As três restrições são ativas em x^* e os vetores

$$\nabla c_1(x^*) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

  são linearmente dependentes.

- Além disso, $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$, ou seja, vale KKT.

CENGAGE Learning

Exemplo em que vale KKT sem que LICQ seja satisfeita

$$\begin{array}{ll} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ & c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \\ & c_3(x) = -x_1 \leq 0. \end{array}$$

O ponto $x^* = 0$ é o minimizador deste problema, cumpre as condições de KKT mas não satisfaz LICQ.

Solução:

- As três restrições são ativas em x^* e os vetores

$$\nabla c_1(x^*) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

  são linearmente dependentes.

- Além disso, $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$, ou seja, vale KKT.

Exemplo em que vale KKT sem que LICQ seja satisfeita

$$\begin{array}{ll} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ & c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \\ & c_3(x) = -x_1 \leq 0. \end{array}$$

O ponto $x^* = 0$ é o minimizador deste problema, cumpre as condições de KKT mas não satisfaz LICQ.

Solução:

- As três restrições são ativas em x^* e os vetores

$$\nabla c_1(x^*) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

  são linearmente dependentes.

- Além disso, $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$, ou seja, vale KKT.

Condição de Qualificação de Mangasarian-Fromovitz

MFCQ

A condição de qualificação de Mangasarian-Fromovitz (MFCQ) é satisfeita em $\bar{x}$ quando os gradientes das restrições de igualdade são linearmente independentes e existir um vetor $d \in \mathbb{R}^n$ tal que

$$\nabla c_i(\bar{x})^T d = 0 \quad \text{e} \quad \nabla c_j(\bar{x})^T d < 0,$$

para todos $i \in \mathcal{E}$ e $j \in I(\bar{x})$.

Exemplo em que não vale LICQ mas vale MFCQ

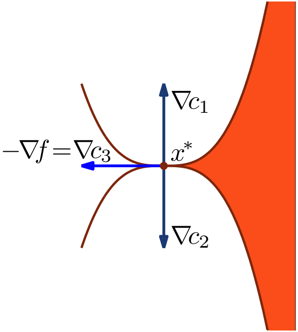
$$\begin{array}{ll} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = x_1^2 - 2x_1 - x_2 \leq 0 \\ & c_2(x) = x_1^2 - 2x_1 + x_2 \leq 0 \\ & c_3(x) = -x_1 \leq 0. \end{array}$$

Solução:

As restrições cumprem MFCQ no ponto $\bar{x} = 0$, pois o vetor $d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ satisfaz $\nabla c_i(\bar{x})^T d < 0$, $i = 1, 2, 3$.

Exemplo em que vale KKT mas MFCQ não é satisfeita

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_1^3 - x_2 \leq 0 \\
 & c_3(x) = -x_1 \leq 0.
\end{array}
$$



$x^* = 0$ é o minimizador, satisfaz KKT, mas não cumpre MFCQ. De fato:

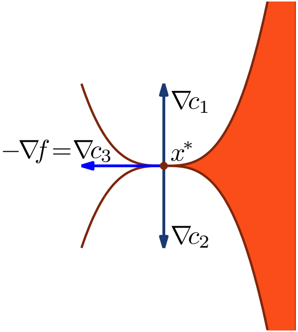
- As três restrições são ativas em x^* e

$$\nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

- Não existe um vetor $d \in \mathbb{R}^2$ tal que $\nabla c_i(\bar{x})^T d < 0$ para $i = 1,2,3$.
- Vale KKT, pois $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$.

Exemplo em que vale KKT mas MFCQ não é satisfeita

$$\begin{array}{rl} \text{minimizar} & f(x) = x_1 \\ \text{sujeito a} & c_1(x) = -x_1^3 + x_2 \le 0 \\ & c_2(x) = -x_1^3 - x_2 \le 0 \\ & c_3(x) = -x_1 \le 0. \end{array}$$



$x^* = 0$ é o minimizador, satisfaz KKT, mas não cumpre MFCQ. De fato:

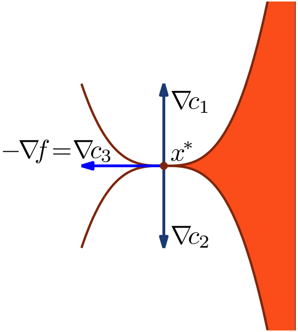
- As três restrições são ativas em x^* e

$$\nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

- Não existe um vetor $d \in \mathbb{R}^2$ tal que $\nabla c_i(\bar{x})^T d < 0$ para $i = 1, 2, 3$.
- Vale KKT, pois $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$.

Exemplo em que vale KKT mas MFCQ não é satisfeita

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_1^3 - x_2 \leq 0 \\
 & c_3(x) = -x_1 \leq 0.
\end{array}
$$



$x^* = 0$ é o minimizador, satisfaz KKT, mas não cumpre MFCQ. De fato:

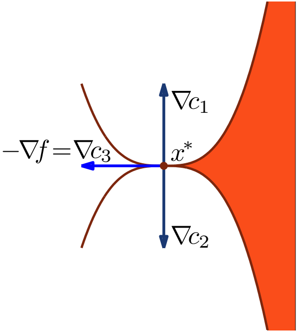
- As três restrições são ativas em x^* e

$$\nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

- Não existe um vetor $d \in \mathbb{R}^2$ tal que $\nabla c_i(\bar{x})^T d < 0$ para $i = 1, 2, 3$.
- Vale KKT, pois $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$.

Exemplo em que vale KKT mas MFCQ não é satisfeita

$$
\begin{array}{ll}
\text{minimizar} & f(x) = x_1 \\
\text{sujeito a} & c_1(x) = -x_1^3 + x_2 \leq 0 \\
 & c_2(x) = -x_1^3 - x_2 \leq 0 \\
 & c_3(x) = -x_1 \leq 0.
\end{array}
$$



$x^* = 0$ é o minimizador, satisfaz KKT, mas não cumpre MFCQ. De fato:

- As três restrições são ativas em x^* e

$$\nabla c_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{e} \quad \nabla c_3(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

- Não existe um vetor $d \in \mathbb{R}^2$ tal que $\nabla c_i(\bar{x})^T d < 0$ para $i = 1, 2, 3$.
- Vale KKT, pois $-\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \nabla c_3(x^*)$.

Implicações

Considere $\bar{x}$ um ponto viável. Então:

$$\text{LICQ} \quad \Rightarrow \quad \text{MFCQ} \quad \Rightarrow \quad T(\bar{x}) = D(\bar{x})$$

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \leq 0 \end{array}$$

Lagrangiano

$$\ell(x,\lambda,\mu) = f(x) + \lambda^T c_{\mathcal{E}}(x) + \mu^T c_{\mathcal{I}}(x)$$

Indicando as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$ por $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$, respectivamente, temos

$$\nabla_x \ell(x,\lambda,\mu) = \nabla f(x) + A_{\mathcal{E}}(x)^T \lambda + A_{\mathcal{I}}(x)^T \mu$$

e

$$\nabla_{xx}^2 \ell(x,\lambda,\mu) = \nabla^2 f(x) + \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 c_i(x) + \sum_{i \in \mathcal{I}} \mu_i \nabla^2 c_i(x).$$

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \leq 0 \end{array}$$

Lagrangiano

$$\ell(x,\lambda,\mu) = f(x) + \lambda^T c_{\mathcal{E}}(x) + \mu^T c_{\mathcal{I}}(x)$$

Indicando as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$ por $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$, respectivamente, temos

$$\nabla_x \ell(x,\lambda,\mu) = \nabla f(x) + A_{\mathcal{E}}(x)^T \lambda + A_{\mathcal{I}}(x)^T \mu$$

e

$$\nabla^2_{xx} \ell(x,\lambda,\mu) = \nabla^2 f(x) + \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 c_i(x) + \sum_{i \in \mathcal{I}} \mu_i \nabla^2 c_i(x).$$

CENGAGE Learning

Condições de otimalidade de segunda ordem

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \leq 0 \end{array}$$

Lagrangiano

$$\ell(x,\lambda,\mu) = f(x) + \lambda^T c_{\mathcal{E}}(x) + \mu^T c_{\mathcal{I}}(x)$$

Indicando as Jacobianas de $c_{\mathcal{E}}$ e $c_{\mathcal{I}}$ por $A_{\mathcal{E}}$ e $A_{\mathcal{I}}$, respectivamente, temos

$$\nabla_x \ell(x,\lambda,\mu) = \nabla f(x) + A_{\mathcal{E}}(x)^T \lambda + A_{\mathcal{I}}(x)^T \mu$$

e

$$\nabla^2_{xx} \ell(x,\lambda,\mu) = \nabla^2 f(x) + \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 c_i(x) + \sum_{i \in \mathcal{I}} \mu_i \nabla^2 c_i(x).$$

Condições de otimalidade de segunda ordem

Seja $x^* \in \Omega$ satisfazendo a condição de qualificação de independência linear.

Condições necessárias de 2ª ordem

Suponha que x^* é um minimizador local do problema. Considere os multiplicadores λ^* e μ^*, que satisfazem as condições de KKT. Então,

$$d^T \nabla^2_{xx} \ell(x^*, \lambda^*, \mu^*) d \geq 0,$$

para todo $d \in \mathcal{N}\big(A_{\mathcal{E}}(x^*)\big) \cap \mathcal{N}\big(A_{I(x^*)}(x^*)\big)$.

Suponha que existem $\lambda^* \in \mathbb{R}^m$ e $\mu^* \in \mathbb{R}^q_+$ tais que $(\mu^*)^T c_I(x^*) = 0$ e

$$\nabla f(x^*) + A_{\mathcal{E}}(x^*)^T \lambda^* + A_I(x^*)^T \mu^* = 0.$$

Considere $I^+ = \{i \in I(x^*) \mid \mu_i^* > 0\}$. Se

$$d^T \nabla^2_{xx} \ell(x^*, \lambda^*, \mu^*) d > 0,$$

para todo $d \in \mathcal{N}\big(A_{\mathcal{E}}(x^*)\big) \cap \mathcal{N}\big(A_{I^+}(x^*)\big)$, então existem $\delta > 0$ e uma vizinhança V de x^* tal que

$$f(x) - f(x^*) \geq \delta \|x - x^*\|^2,$$

para todo ponto viável $x \in V$. Em particular, segue que x^* é um minimizador local estrito do problema.

Principais Referências Bibliográficas

R. Andreani, J. M. Martínez, and M. L. Schuverdt.
On the relation between constant positive linear dependence condition and quasinormality constraint qualification.
Journal of Optimization theory and Applications, 125(2):473–485, 2005.

A. Izmailov and M. Solodov.
Otimização, vol. 1, *IMPA, 2005.*

R. G. Eustáquio
Condições de Otimalidade e de Qualificação para Problemas de Programação Não Linear. Dissertação de Mestrado, *Universidade Federal do Paraná, 2007.*