CM106 - Otimização I

14 de Junho de 2018 - Prova 2

Gabarito

Questão 1 .... $\boxed{30}$

Encontre e classifique os pontos críticos de $f(x,y) = x^2 - y^2$ sujeito à $h(x,y) = 1 - x^2 - y^2 \geq 0$, usando KKT e condições de segunda ordem.

Solution: KKT:

$$\begin{cases} 2x &= -2x\lambda, \\ -2y &= -2y\lambda, \\ (1 - x^2 - y^2)\lambda &= 0, \\ x^2 + y^2 &\leq 1. \end{cases}$$

Da terceira equação, tiramos $\lambda = 0$ ou $x^2 + y^2 = 1$.

- Se $\lambda = 0$, então $x = 0 = y$.
- Se $x^2 + y^2 = 1$, da primeira equação tiramos $x = 0$ ou $\lambda = -1$.
  - Se $x = 0$, então $y = \pm 1$ e $\lambda = -1$.
  - Se $\lambda = -1$, então $y = 0$ e $x = \pm 1$ e $\lambda = 1$.

Segunda ordem:

$$\nabla^2 f(x,y) - \lambda \nabla^2 h(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} - \lambda \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 + 2\lambda & 0 \\ 0 & -2 + 2\lambda \end{bmatrix}.$$

Jacobiana: $[-2x \ \ -2y]$.

- $(x, y, \lambda) = (0, 0, 0)$: Hessiana $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$. Indefinida, logo ponto de sela.
- $(x, y, \lambda) = (0, \pm 1, 1)$: Jacobiana: $[0 \ \ \mp 2]$, de modo que o núcleo é gerado por $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

  Hessiana: $\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$. Daí,

  $$\begin{bmatrix} 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 4 > 0,$$

  logo o ponto é um minimizador local.

- $(x, y, \lambda) = (\pm 1, 0, -1)$: Jacobiana: $[\mp 2\ 0]$, de modo que o núcleo é gerado por $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

  Hessiana: $\begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix}$. Daí

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -4 < 0,$$

  logo o ponto é um maximizador local.

Questão 2 .... $\boxed{20}$

Considere os seguintes problemas:

$$\min_{x,r} \frac{1}{2} \|r\|^2 \qquad \text{suj. a} \qquad Ax + r = b,$$

e

$$\min_{y,t} \frac{1}{2} \|y\|^2 \qquad \text{suj. a} \qquad A^T y = c,$$

onde $A \in \mathbb{R}^{m \times n}$, $m > n$, onde A não necessariamente tem posto coluna completo, $b, r, y \in \mathbb{R}^m$, $x, c \in \mathbb{R}^n$. Sejam $x = \overline{x}$, $r = \overline{r}$ e $y = \overline{y}$ as soluções dos problemas acima.

(a) $\boxed{10}$ Escreva as condições de otimalidade de primeira ordem de cada problema.

Solution: Primeiro:

$$\begin{bmatrix} 0 \\ r \end{bmatrix} = \begin{bmatrix} A^T \\ I \end{bmatrix} \lambda \qquad Ax + r = b,$$

ou seja

$$A^T r = 0 \qquad Ax + r = b.$$

Segundo:

$$y = A\lambda \qquad A^T y = c.$$

(b) $\boxed{5}$ Verifique que $\overline{x}^T c = \overline{y}^T b$.

Solution: Temos

$$x^T c = x^T A^T y = x^T A^T A \lambda = b^T A \lambda = b^T y,$$

logo $x^T c = b^T y$.

(c) $\boxed{5}$ Mostre que se $c = A^T b$, então $Ax - y \in \mathrm{Nu}(A^T)$.

Solution: Temos

$$A^T (Ax - y) = A^T A x - A^T y = A^T b - c = 0.$$

Questão 3 .... $\boxed{20}$

Resolva o seguinte problema de otimização:

$$\begin{aligned} \max \quad & P(x) = x_1 x_2 \dots x_n \\ \text{suj. a} \quad & x_1 + x_2 + \dots x_n = c, \\ & x \geq 0, \end{aligned}$$

utilizando KKT e justificando o valor máximo.

Solution: KKT:

$$\prod_{j=1, j\neq i}^{n} x_j = \lambda + \mu_i,$$

$$\sum_{j=1}^{n} x_j = c,$$

$$\mu_i x_i = 0.$$

Multiplicando a primeira equação por x_i, e usando a terceira equação, temos

$$P(x) = \lambda x_i.$$

Se algum $x_i = 0$, então $P(x) = 0$. Como $x_1 + \dots + x_n = c$, supondo $c > 0$, teremos algum $x_j \neq 0$. Daí, $\lambda = 0$. Todo ponto satisfazendo isso será minimizador, que não nos interessa. Vamos supor $x_i > 0$ então. Temos

$$x_i = \frac{P(x)}{\lambda},$$

ou seja $x_i = x_j$ para qualquer i, j. Sendo assim $x_1 + \dots + x_n = c \Rightarrow x_i = \dfrac{c}{n}$. O máximo será $P(x) = \left(\dfrac{c}{n}\right)^n$. Sendo o único ponto crítico com $x_i \neq 0$, será o maximizador global, já que a região factível é fechada e limitada.

Questão 4 .... $\boxed{20}$

No problema de minimizar $f(x)$ sujeito à $\ell \leq x \leq u$, com $\ell_i < u_i$, seja $g = \nabla f(x)$ e d definida por

$$d_i = \begin{cases} 0, & \text{se } x_i = \ell_i \text{ e } g_i \geq 0, \\ 0, & \text{se } x_i = u_i \text{ e } g_i \leq 0, \\ -g_i, & \text{caso contrário.} \end{cases}$$

(a) $\boxed{10}$ Mostre que se $d \neq 0$, então d é uma direção factível e de descida a partir de x.

Solution: Se $d \neq 0$, existe i tal que $d_i = -g_i \neq 0$. Para todo $d_i \neq 0$, temos 3 opções

- $x_i = \ell_i$ e $g_i < 0$, de modo que $d_i > 0$ e $\ell_i < x_i + \alpha d_i < u_i$ para α suficientemente pequeno.

- $x_i = u_i$ e $g_i > 0$, de modo que $d_i < 0$ e $\ell_i < x_i + \alpha d_i < u_i$ para α suficientemente pequeno.

- $\ell_i < x_i < u_i$, de modo que, independentemente do sinal de d_i, $\ell_i < x_i + \alpha d_i < u_i$ para α suficientemente pequeno.

Isso quer dizer que d é factível.
Para mostrar que é descida, é ainda mais fácil

$$d^T g = \sum_{i=1}^{n} d_i g_i = \sum_{i:d_i \neq 0} d_i g_i = \sum_{i:d_i \neq 0} -g_i^2 < 0,$$

pois existe ao menos um i tal que $d_i \neq 0$.

(b) $\boxed{10}$ Mostre que se x é um minimizador local do problema, então $d = 0$.

Solution: Seja x minimizador local do problema, e suponha que $d \neq 0$. Mas então, pelo exercício anterior, $g^T d = -\sum_{i:d_i \neq 0} g_i^2 < 0$. Mas pelo Teorema de otimalidade para direções factíveis, $g^T d \geq 0$. Absurdo. Logo $d = 0$.
Outra maneira, é usar as condições de otimalidade

$$\begin{cases} g & = \lambda_L - \lambda_U \\ (x_i - \ell_i)\lambda_{L_i} & = 0 \\ (u_i - x_i)\lambda_{U_i} & = 0 \\ \lambda_L, \lambda_U & \geq 0 \end{cases}$$

e $\ell \leq x \leq u$. Daqui, temos

- Se $x_i = \ell_i$, então $g_i = \lambda_{L_i} \geq 0$, de modo que $d_i = 0$.

- Se $x_i = u_i$, então $g_i = -\lambda_{U_i} \leq 0$, de modo que $d_i = 0$.

- Se $\ell_i < x_i < u_i$, então $d_i = -g_i = 0$.

Logo, $d = 0$.

Questão 5 .... $\boxed{20}$

Considere o esboço abaixo, que representa a minimização de uma função quadrática f com 5 restrições lineares de desigualdade ($a_i^T x \geq b_i,\ i = 1, \ldots, 5$). O método de restrições ativas com gradiente projetado pode ser descrito da seguinte maneira:

1. Dado x, e $\mathcal{W} = \mathcal{A}(x)$,
2. Tente resolver o sistema $\nabla f(x) = \sum_{i \in \mathcal{W}} a_i \lambda_i$. Se não for possível, vá ao passo 3, se for possível vá ao passo 7.,
3. Calcule d a projeção de $-\nabla f(x)$ sobre as restrições em $\mathcal{W}$.
4. Calcule o minimizador de $f(x + \alpha d)$, $\alpha \geq 0$ e $a_i^T(x + \alpha d) \geq b_i$.
5. Se $x + \alpha d$ encontrou uma ou mais restrições, então adicione essas restrições adicionais em $\mathcal{W}$.
6. Volte ao passo 2.
7. Se $\lambda \geq 0$, FIM
8. Se algum $\lambda_i < 0$, escolha **uma** restrição com $\lambda_i < 0$ e remova de $\mathcal{W}$.
9. Volte ao passo 2.

Aplique o algoritmo acima no problema esboçado abaixo, a partir do ponto $x_0 = (0, 1)$, **graficamente**. Numere as retas, explique os passos tomados, e em cada ponto, esboce o que seria o gradiente da função e das restrições.