

PPGMA

Exame de qualificação de Otimização - 08/03/2013

Instruções:

- Resolva as 5 questões abaixo (2,0 pontos cada).
- As soluções devem conter o desenvolvimento e/ou justificativas.

Questões:

1. Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe \mathcal{C}^2 e $x^* \in \mathbb{R}^n$ um ponto estacionário de f .
 - (a) Prove que se x^* é um minimizador local de f , então $d^T \nabla^2 f(x^*) d \geq 0$, para todo $d \in \mathbb{R}^n$;
 - (b) Vale a recíproca da afirmação do item anterior? Justifique sua resposta.
2. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita convexa em C quando

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y), \quad (1)$$

para todos $x, y \in C$ e $t \in [0, 1]$.

- (a) Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo. Prove que f é convexa em C se, e somente se,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad (2)$$

para todos $x, y \in C$;

- (b) Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa, diferenciável e $C \subset \mathbb{R}^n$ convexo. Prove que se $\nabla f(x^*) = 0$, então x^* é um minimizador global de f em C ;
- (c) Qual é a interpretação geométrica das desigualdades (1) e (2)? Faça figuras elucidativas.

3. Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe \mathcal{C}^1 .

- (a) Apresente o algoritmo do método do gradiente com busca exata;
- (b) Prove que o algoritmo acima é globalmente convergente.

4. Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica definida positiva.

- (a) Prove que um conjunto qualquer de vetores A -conjugados não nulos é linearmente independente;

- (c) Considere a matriz $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}$. Obtenha uma base de \mathbb{R}^3 formada por vetores

A -conjugados, sendo que um dos vetores é $d^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

5. Considere $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ e os problemas primal e dual de programação linear

$$(P) \quad \begin{array}{ll} \min & c^T x \\ \text{s. a} & Ax = b \\ & x \geq 0 \end{array} \quad \text{e} \quad (D) \quad \begin{array}{ll} \max & b^T y \\ \text{s. a} & A^T y \leq c. \end{array}$$

Suponha que $x^* \in \mathbb{R}^n$ seja um minimizador do primal e $\lambda^* \in \mathbb{R}^m$ o multiplicador de Lagrange associado à restrição de igualdade $b - Ax = 0$.

- (a) Mostre que $b^T y \leq c^T x$, para todos x e y viáveis;
- (b) Escreva as condições de KKT para o problema primal e prove que $c^T x^* = b^T \lambda^*$;
- (c) Prove que λ^* é solução do problema dual;
- (d) Conclua que o valor ótimo primal e dual coincidem.

BOA PROVA

Soluções:

1. (a) Considere $d \in \mathbb{R}^n \setminus \{0\}$ arbitrário. Por Taylor,

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t),$$

com $\lim_{t \rightarrow 0} \frac{r(t)}{t^2} = 0$. Como x^* é estacionário, temos $\nabla f(x^*) = 0$. Portanto, para t suficientemente pequeno,

$$0 \leq f(x^* + td) - f(x^*) = \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t),$$

Dividindo por t^2 e passando o limite quando $t \rightarrow 0$, obtemos $d^T \nabla^2 f(x^*) d \geq 0$.

- (b) A recíproca não vale, pois $f(x) = -x^4$ cumpre $f''(0) = 0$, mas $x = 0$ não é minimizador. Note que se $\nabla^2 f(x)$ for definida positiva, para todo $x \in \mathbb{R}^n$, então vale. De fato, por Taylor (com resto de Lagrange), já usando o fato de x^* ser estacionário, temos que existe $t \in (0, 1)$ tal que

$$f(x^* + d) = f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^* + td) d.$$

Como $\nabla^2 f(x)$ é definida positiva, para todo $x \in \mathbb{R}^n$, o resultado segue.

2. (a) Seja f convexa. Para $x, y \in C$ e $t \in (0, 1]$ quaisquer, definindo $d = y - x$, temos $x + td \in C$ e

$$f(x + td) = f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y).$$

Portanto,

$$f(y) - f(x) \geq \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^T d = \nabla f(x)^T (y - x).$$

Para provar a recíproca, considere $z = (1 - t)x + ty$ e observe que

$$f(x) \geq f(z) + \nabla f(z)^T (x - z) \quad \text{e} \quad f(y) \geq f(z) + \nabla f(z)^T (y - z).$$

Multiplicando a primeira por $(1 - t)$ e a segunda por t obtemos

$$(1 - t)f(x) + tf(y) \geq f((1 - t)x + ty),$$

completando a demonstração.

- (b) Temos

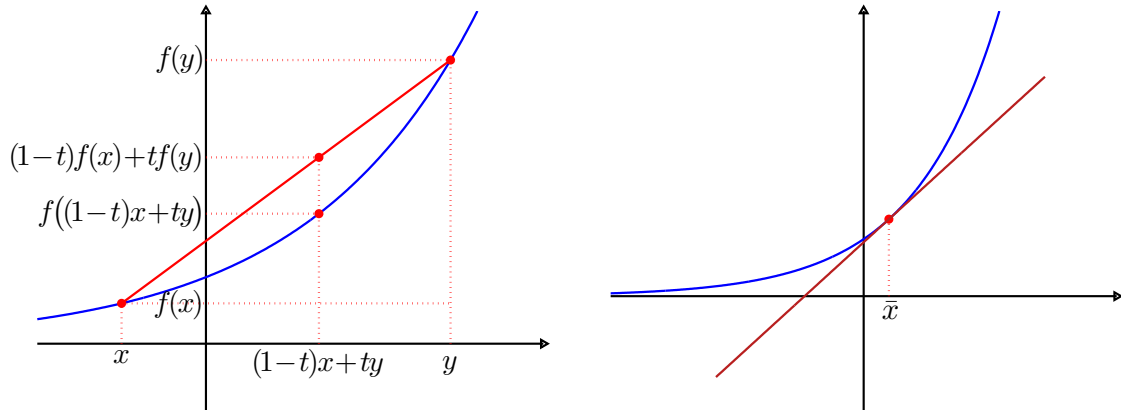
$$f(y) \geq f(x^*) + \nabla f(x^*)^T (y - x^*)$$

para todo $y \in C$. Como $\nabla f(x^*) = 0$, obtemos

$$f(y) \geq f(x^*)$$

para todo $y \in C$.

- (c) A desigualdade (1) significa que o gráfico de uma função convexa fica sempre abaixo (ou sobre) do segmento que une dois pontos, enquanto que (2) significa que o gráfico de uma função convexa fica acima do plano tangente.



3. (a) Dado: $x^0 \in \mathbb{R}^n$

$$k = 0$$

REPITA enquanto $\nabla f(x^k) \neq 0$

$$\text{Defina } d^k = -\nabla f(x^k)$$

Obtenha $t_k > 0$ tal que $f(x^k + t_k d^k) < f(x^k + \bar{t} d^k)$, para todo $t > 0$

$$\text{Faça } x^{k+1} = x^k + t_k d^k$$

$$k = k + 1$$

- (b) Sejam (x^k) uma sequência gerada pelo algoritmo e \bar{x} um ponto de acumulação de (x^k) , digamos $x^k \xrightarrow{\mathbb{N}'} \bar{x}$. Suponha por absurdo que \bar{x} não seja estacionário, isto é, $\nabla f(\bar{x}) \neq 0$. Assim, $\bar{d} = -\nabla f(\bar{x})$ é uma direção de descida, o que garante a existência de $\bar{t} > 0$ tal que $\beta = f(\bar{x}) - f(\bar{x} + \bar{t}\bar{d}) > 0$. Considere $h: \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $h(x) = f(x) - f(\bar{x} + \bar{t}\nabla f(x))$. Como h é contínua, temos que $h(x^k) \xrightarrow{\mathbb{N}'} h(\bar{x}) = \beta$. Portanto,

$$f(x^k) - f(x^k + \bar{t}d^k) = h(x^k) \geq \frac{\beta}{2},$$

para todo $k \in \mathbb{N}'$, suficientemente grande. Deste modo, como t_k foi obtido pela busca exata, podemos concluir que

$$f(x^{k+1}) = f(x^k + t_k d^k) \leq f(x^k + \bar{t}d^k) \leq f(x^k) - \frac{\beta}{2},$$

ou seja,

$$f(x^k) - f(x^{k+1}) \geq \frac{\beta}{2}, \quad (3)$$

para todo $k \in \mathbb{N}'$, suficientemente grande. Por outro lado, pela continuidade de f , temos $f(x^k) \xrightarrow{\mathbb{N}'} f(\bar{x})$. Como a sequência $(f(x^k))_{k \in \mathbb{N}}$ é decrescente, temos $f(x^k) \rightarrow f(\bar{x})$, contradizendo (3).

4. (a) Sejam $d^0, d^1, \dots, d^k \in \mathbb{R}^n \setminus \{0\}$ vetores A -conjugados. Considere constantes $a_0, a_1, \dots, a_k \in \mathbb{R}$ tais que

$$a_0 d^0 + a_1 d^1 + \dots + a_k d^k = 0.$$

Dado $i \in \{0, 1, \dots, k\}$, multiplicando os dois membros da igualdade acima por $(d^i)^T A$, obtemos

$$a_i (d^i)^T A d^i = 0,$$

donde segue que $a_i = 0$, pois A é definida positiva.

(b) Considere $d^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Os vetores d^1 e d^2 devem satisfazer $(d^0)^T A d^j = 0$ e $(d^1)^T A d^2 = 0$.

Como $(d^0)^T A = (1 \ 0 \ -1)$, d^1 e d^2 são da forma $d^1 = \begin{pmatrix} a \\ b \\ a \end{pmatrix}$ e $d^2 = \begin{pmatrix} a' \\ b' \\ a' \end{pmatrix}$. Impondo

$(d^1)^T A d^2 = 0$, obtemos $aa' + bb' = 0$, o que por exemplo vale para os vetores $d^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

e $d^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Note que os vetores d^1 e d^2 não precisam necessariamente ser ortogonais.

De fato, os vetores $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} -2 \\ 1 \\ -2 \end{pmatrix}$ também A -conjugados.

5. (a) Dados x e y viáveis temos

$$b^T y = (Ax)^T y = x^T A^T y \leq x^T c = c^T x.$$

(b) As condições de KKT para o problema primal podem ser escritas como

$$\begin{aligned} -c &= -\mu^* - A^T \lambda^* \\ \mu^* &\geq 0 \\ (\mu^*)^T x^* &= 0. \end{aligned}$$

Além disso, pela viabilidade de x^* , temos $Ax^* = b$ e $x^* \geq 0$. Portanto,

$$c^T x^* = (\mu^* + A^T \lambda^*)^T x^* = b^T \lambda^*.$$

(c) Para ver que λ^* é solução do problema dual, note primeiro que $A^T \lambda^* = c - \mu^* \leq c$, o que significa que λ^* é viável. Considere agora um ponto y viável para o dual. Usando o que já foi provado, temos

$$b^T y \leq c^T x^* = b^T \lambda^*.$$

(d) Pelo que foi provado nos itens anteriores, o valor ótimo primal, $c^T x^*$ coincide com o valor ótimo dual, $b^T \lambda^*$.