Optimização e Algoritmos (2004/2005)

Instituto Superior Técnico – Engenharia Electrotécnica e de Computadores

Série de Problemas 2

Extremos locais/globais, Condições de Optimalidade

Problema 1.[Identificação e classificação de pontos de estacionariedade] Identifique os pontos de estacionariedade das seguintes funções e classifique-os (minimizante local/global, maximizante local/global ou ponto de sela). Determine também se as funções admitem extremos globais.

(a) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = xe^{-x}$
(b) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + (1 + x)^3 y^2$
(c) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^3 - 3xy^2 + y^4$
(d) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = xy + yz + zx$
(e) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{2}ax^2 + e^x$, onde $a \in \mathbb{R}$ é uma constante

Problema 2.[Classificação de pontos de estacionariedade] Para cada uma das seguintes funções $f : \mathbb{R}^2 \to \mathbb{R}$ mostre que $(x, y) = (0, 0)$ é um ponto de estacionariedade e classifique-o (minimizante local/global, maximizante local/global ou ponto de sela).

(a) $f(x, y) = x^4 + y^4 - 2(x - y)^2$
(b) $f(x, y) = x^4 + y^4$
(c) $f(x, y) = x^3 + y^3$

Problema 3.[Pontos de estacionariedade] Considere a função $f : \mathbb{R}^n \to \mathbb{R}$ dada por

$$f(x) = \frac{1}{2}x^T Ax + \phi(b^T x),$$

onde $A \in \mathbb{R}^{n\times n}$ é uma matriz simétrica, $b \in \mathbb{R}^n$ é um vector e $\phi : \mathbb{R} \to \mathbb{R}$ é uma função diferenciável e <u>estritamente</u> crescente.

Mostre que f admite um ponto de estacionariedade só se $b \in \mathsf{Range}\, A$.

[**Nota:** $\mathsf{Range}\, A$ é o subespaço linear gerado pelas colunas de A. Ou seja,

$$\mathsf{Range}\, A = \{Ax : x \in \mathbb{R}^n\}. \quad]$$

Problema 4.[Pontos de estacionariedade/Condições de Optimalidade] Considere a função $f : \mathbb{R}^n \to \mathbb{R}$ dada por

$$f(x) = \frac{1}{2}x^T Ax + b^T x + e^{c^T x},$$

onde

$$A = \begin{bmatrix} 2+\alpha & 2-\alpha & \alpha \\ -2+\alpha & 4+\alpha & -2+\alpha \\ \alpha & -2+\alpha & 2+\alpha \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

A constante $\alpha \in \mathbb{R}$ e o vector $b \in \mathbb{R}^n$ não estão especificados.

(a) Mostre que

$$\nabla f(x) = Ax + b + e^{c^T x} c.$$

(b) Especifique b de modo a tornar $x = 0$ um ponto de estacionariedade de f.

(c) Mostre que

$$\nabla^2 f(x) = A + e^{c^T x} cc^T.$$

(d) Suponha que o vector b é determinado como na alínea (b). Escolha valores para α de modo que $x = 0$ seja: **(i)** um mínimo local estrito, **(ii)** um mínimo local não-estrito e **(iii)** um ponto de sela. [**Pista:** o vector c é um vector próprio de A]

Problema 5.[Funções coercivas] Uma função $f : \mathbb{R}^n \to \mathbb{R}$ diz-se coerciva se

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Mostre: se f é uma função contínua e coerciva, então f tem um minimizante global.

Problema 6.[Algoritmos: coordenadas cíclicas, descida de gradiente e Newton] Considere a função $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = e^{x^T Q x},$$

onde $Q : n \times n$ é uma matriz simétrica definida positiva.

Note que f é uma função coerciva.

(a) Mostre que

$$\begin{aligned} \nabla f(x) &= 2e^{x^T Q x} Qx \\ \nabla^2 f(x) &= 2e^{x^T Q x} \left[Q + 2(Qx)(Qx)^T \right] \end{aligned}$$

(b) Mostre que $x^* = 0$ é o único ponto de estacionariedade e é o único minimizante global de f.

(c) Considere o subproblema de pesquisa em linha

$$\alpha^* = \arg\min_{\alpha \geq 0} f(x_0 + \alpha d), \tag{1}$$

onde $x_0 \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$ é uma direcção de busca não-nula. Mostre que a solução do subproblema (1) é dada por

$$\alpha^* = \begin{cases} -\dfrac{x_0^T Q d}{d^T Q d} & , \quad \text{se } x_0^T Q d \leq 0 \\ 0 & , \quad \text{caso contrário} \end{cases}$$

(d) [Coordenadas cíclicas] Seja

$$Q = Q_a \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix} Q_a^T, \tag{2}$$

onde

$$Q_a = \left[\begin{array}{cc} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{array} \right],$$

com $\theta = \pi/5$.

Considere a utilização do método de coordenadas cíclicas para minimizar f. Ou seja, designando por x^k a k-ésima iteração, temos

$$x^{k+1} \quad = \quad x^k + \alpha^k d^k \tag{3}$$

onde $d^0 = e_1$, $d^1 = e_2$, $d^2 = e_1$, $d^3 = e_2$, $d^4 = e_1$, etc com

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

e

$$\alpha^k \quad = \quad \arg\min_{\alpha \in \mathbb{R}} f\left(x^k + \alpha\, d^k\right).$$

Implemente a iteração em (3) em MATLAB, tomando como condição inicial $x^0 = (0.3, 0.2)$. Qual é o primeiro k que satisfaz $\left\|x^k - x^*\right\| < 10^{-10}$?

(e) [Descida de gradiente] Seja Q dado por (2). Considere a utilização do método de descida de gradiente com pesquisa em linha exacta para minimizar f. Ou seja, temos

$$x^{k+1} = x^k + \alpha^k d^k, \tag{4}$$

onde $d^k = -g^k$, $g^k = \nabla f\left(x^k\right)$, e

$$\alpha^k = \arg\min_{\alpha \geq 0} f\left(x^k + \alpha d^k\right).$$

Implemente a iteração em (4) em MATLAB. Tomando como condição inicial $x^0 = (0.3, 0.2)$ qual é o primeiro k que satisfaz $\left\|x^k - x^*\right\| < 10^{-10}$?

(f) [Método de Newton puro] Seja Q dado por (2). Considere a utilização do método de Newton puro para minimizar f. Ou seja, temos

$$x^{k+1} = x^k - \left(H^k\right)^{-1} g^k, \tag{5}$$

onde $g^k = \nabla f\left(x^k\right)$ e $H^k = \nabla^2 f(x^k)$.

Implemente a iteração em (5) em MATLAB, tomando como condição inicial $x^0 = (0.3, 0.2)$. Qual é o primeiro k que satisfaz $\left\|x^k - x^*\right\| < 10^{-10}$?

Problema 7.[Conjuntos compactos/Matrizes ortogonais] Seja

$$\mathsf{O}(n) = \left\{Q \in \mathbb{R}^{n\times n} \,:\, Q^T Q = QQ^T = I_n\right\}$$

o conjunto das matrizes ortogonais $n \times n$. (Nota: I_n é a matriz identidade $n \times\ n$)

(a) Prove que $\mathsf{O}(n)$ é um subconjunto compacto de $\mathbb{R}^{n\times n}$. [**Pista:** mostre que $\mathsf{O}(n)$ é limitado e fechado]

(b) Mostre que as matrizes da forma

$$Q(\theta) = \left[\begin{array}{cc} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{array}\right], \tag{6}$$

onde $\theta \in \mathbb{R}$, pertencem a $\mathsf{O}(2)$.

(c) Seja

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

uma matriz simétrica 2×2. Mostre que A é ortogonalmente equivalente a uma matriz simétrica com os elementos na diagonal iguais. Ou seja, mostre que existe $Q \in \mathsf{O}(2)$ e $d, e \in \mathbb{R}$ tal que

$$Q^T A Q = \begin{bmatrix} d & e \\ e & d \end{bmatrix}.$$

[**Pista:** Comece por considerar o caso A diagonal, ou seja,

$$A = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix},$$

e tente encontrar uma matriz $Q(\theta)$ da forma (6) tal que $Q(\theta)^T A Q(\theta)$ tenha os elementos da diagonal iguais. Para tratar o caso de A geral, relembre-se que, pelo teorema da decomposição em valores próprios, existe uma matriz $V \in \mathsf{O}(2)$ e uma matriz diagonal

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

que contém os valores próprios de A, tal que

$$A = V \Lambda V^T.$$

Agora, fazendo V^TAV, reduzimos o problema ao caso anterior.]

(d) Demonstre a alínea anterior para o caso geral $n \times n$. Ou seja, prove que para qualquer matriz A simétrica $n \times n$, existe $Q \in \mathsf{O}(n)$ tal que Q^TAQ tem os elementos na diagonal iguais. [**Pista:** Comece por provar que, dado A com $a_{ii} \neq a_{jj}$, existe um $R \in \mathsf{O}(n)$ tal que $B = R^TAR$ verifica $b_{ii} = b_{jj}$; para descobrir este R pense numa generalização de $Q(\theta)$ em (6) para o caso $n \times n$. Depois, defina a função $f : \mathbb{R}^{n\times n} \to \mathbb{R}$,

$$f(X) = \sum_{i\neq j}(y_{ii} - y_{jj})^2, \text{ onde } Y = X^TAX,$$

e note que f é uma função contínua. Usando os resultados obtidos até agora, prove que f tem um minimizante global Q^* em $\mathsf{O}(n)$ e que $f(Q^*) = 0$.]