PROGRAMAÇÃO NÃO LINEAR

2º Semestre de 2009 — MS629A

Profa.: Sandra Augusta Santos — sala IM111

Nome: ______________________ **RA:** ____________

PRIMEIRA PROVA
14/10/2009

Questão	
Questão 1	
Questão 2	
Questão 3	
Questão 4	
Questão 5	
Total	

Seja cuidadoso no encaminhamento do seu raciocínio.
A clareza da sua argumentação também será avaliada.
Cada questão vale 2.0 pontos.

1. Considere a função quadrática $q(x) = \frac{1}{2}x^T Hx + b^T x + c$, em que $H \in I\!R^{n\times n}$ é simétrica, $b \in I\!R^n$ e $c \in I\!R$. Seja $x^\star$ minimizador local de q. Prove que $x^\star$ é minimizador global de q.

2. Considere um modelo algorítmico geral para minimização irrestrita com direções de descida. Utilizando **argumentos algébricos, analíticos e geométricos**, explique a diferença entre fazer uma **busca linear exata** e trabalhar com a **condição de Armijo**. Especifique também as situações em que cada uma destas estratégias é mais indicada.

3. Seja $f : I\!R^n \to I\!R$ uma função diferenciável em $\tilde{\boldsymbol{x}}$ e sejam $\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^n \in I\!R^n$ vetores linearmente independentes. Suponha que o mínimo de $f(\tilde{\boldsymbol{x}}+\lambda\boldsymbol{d}^j)$ com $\lambda \in I\!R$ ocorra em $\lambda = 0$ para todo $j = 1, \ldots, n$.

   (a) Prove que $\nabla f(\tilde{\boldsymbol{x}}) = 0$.

   (b) Isto implica que $\tilde{\boldsymbol{x}}$ é **minimizador local** de f? Justifique.

4. Ao minimizar a função $f(\boldsymbol{x}) = 10x_1^2 + x_2^2$ pelo **método de máxima descida com busca exata** a partir do ponto inicial $(1, 1)^T$, quantas iterações você espera realizar para atingir a solução? Justifique.

5. Seja $f : I\!R^n \to I\!R$ definida como $f(\boldsymbol{x}) = \sum_{i=1}^{n} \left(\frac{1}{2}a_i x_i^2 + b_i x_i\right)$ onde $a_i, b_i, i = 1, \ldots, n$ são números reais. Exiba **condições suficientes** para que a **direção de Newton** esteja bem definida e seja de descida para qualquer $\boldsymbol{x}$ tal que $\nabla f(\boldsymbol{x}) \neq 0$. Explique como obteve tais condições.

Primeira Prova - MS629 14/out/09

1)

Temos $q(x) = \frac{1}{2}x^T H x + b^T x + c$ $\quad\left\{ q \in C^\infty \right.$ (quadrática)

$\nabla q(x) = Hx + b$ e $\nabla^2 q(x) = H$.

Se x^* é minimizador local de q, pelas CN2, $\boxed{\nabla q(x^*) = 0}$ ① $\Rightarrow Hx^* + b = 0$ e

$\boxed{\nabla^2 q(x^*) = H \geq 0}$ ②

um jeito:

Como a matriz hessiana de q é constante e semi-definida positiva, então q é convexa em $\mathbb{R}^n$.

Mas então, x^* minimizador local de q em $\mathbb{R}^n$ é minimizador global de q em $\mathbb{R}^n$.

outro jeito:

Escrevendo a expansão de Taylor p/ q em torno de x^*, ela é exata pois q é quadrática:

$$q(x) = q(x^*) + \nabla q(x^*)^T (x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 q(x^*)(x - x^*)$$

$$q(x) \overset{①}{=} q(x^*) + \frac{1}{2}(x - x^*)^T H (x - x^*) \underset{②}{\geq} q(x^*)$$

$\therefore q(x^*) \leq q(x), \forall x \in \mathbb{R}^n$

ou seja, x^* é minimizador global de q em $\mathbb{R}^n$.

2) Problema de minimização irrestrita: $\boxed{\begin{array}{l} \min f(x) \\ \text{sa } x \in \mathbb{R}^n \end{array}}$

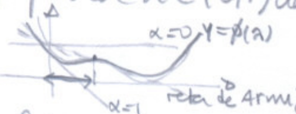
Um modelo algorítmico para esse problema baseia-se na construção de uma seqüência $\{x^k\} \subset \mathbb{R}^n$ na qual $x^{k+1} = x^k + \lambda_k d^k$, d^k é direção de descida e o passo λ_k pode ser calculado minimizando-se exatamente $\phi(\lambda) = f(x^k + \lambda d^k)$ ou apenas impondo-se uma condição de decréscimo suficiente para f.

No primeiro caso, como $\phi'(\lambda) = \nabla f(x^k + \lambda d^k)^T d^k = 0$ então x^{k+1} é tal que $\nabla f(x^{k+1}) \perp d^k$. Geometricamente, isso é fácil de ser visualizado e obtido graficamente para $f: \mathbb{R}^2 \to \mathbb{R}$ com o mapa de curvas de nível. Algebricamente, no entanto, o cálculo efetivo do passo ótimo da busca linear exata só é simples para f quadrática, em que temos uma fórmula fechada: $\lambda^* = -\dfrac{(d^k)^T \nabla f(x^k)}{(d^k)^T \nabla^2 f(x^k) d^k}$.

Para f não linear genérica, o cálculo de λ^* envolve encontrar o zero de uma função real não linear de uma variável e em geral este esforço não compensa computacionalmente o benefício de se encontrar um minimizador exato de f ao longo da direção d^k.

Na prática, é mais interessante buscar λ que satisfaça uma condição de descida suficiente, tipo Armijo: queremos λ tal que, para $\alpha \in (0,1)$ dado,

$$f(x^k + \lambda d^k) \leq f(x^k) + \alpha \lambda \nabla f(x^k)^T d^k.$$



Esta condição é indicada para f não linear geral, e juntamente com as condições de tamanho do passo e cond. do ângulo entre d^k e $\nabla f(x^k)$, produz alg. globalmente convergente.

3) a) Seja $\varphi_j(\lambda) = f(\tilde{x} + \lambda d^j)$

então $\varphi_j'(\lambda) = \nabla f(\tilde{x} + \lambda d^j)^T d^j$

Como $\lambda = 0$ minimiza $\varphi_j(\lambda)$, $\forall j = 1, \dots, n$ segue que

$$\varphi_j'(0) = \nabla f(\tilde{x})^T d^j = 0, \quad \forall j = 1, \dots, n.$$

Sendo $\{d^j\}_{j=1}^n$ um conjunto de vetores LI's,

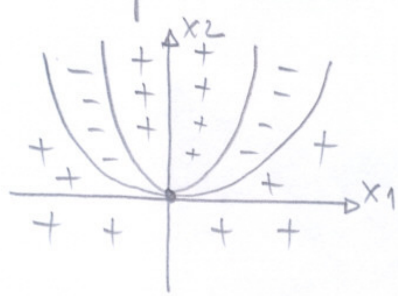
o sistema homogêneo

$$\begin{pmatrix} (d^1)^T \\ \vdots \\ (d^n)^T \end{pmatrix} y = 0$$ Admite apenas a solução trivial.

Portanto $\nabla f(\tilde{x}) = 0$.

b) Embora f seja minimizada ao longo de n direções LI's em $\mathbb{R}^n$, a partir de $\tilde{x}$, o ponto $\tilde{x}$ pode ser um ponto de sela

Ex:

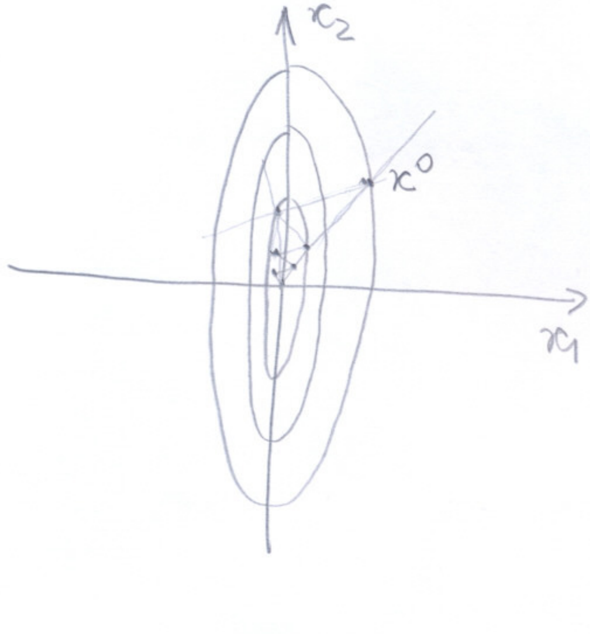


Em $\mathbb{R}^2$, a origem é um ponto de sela, embora seja um ponto de mínimo ao longo das direções $d^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $d^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

4) A quadrática $f(x) = 10x_1^2 + x_2^2 = \frac{1}{2}(x_1 \; x_2)\begin{pmatrix}20 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}$ é uma função estritamente convexa, com minimizador global na origem. Sua hessiana $\nabla^2 f(x) = \begin{pmatrix}20 & 0\\ 0 & 2\end{pmatrix} = H$ é uma matriz positiva definida e diagonal, portanto as curvas de nível de f são elipses, centradas na origem. Os autovetores de H são os vetores canônicos: $v_1 = \begin{pmatrix}1\\0\end{pmatrix}$ e $v_2 = \begin{pmatrix}0\\1\end{pmatrix}$.

O método da máxima descida com busca linear exata só possui convergência finita se $x^0 - x^* = \mu v$, onde x^* é a solução, v é autovetor de H e $\mu \in \mathbb{R}$.

Com $x^0 = \begin{pmatrix}1\\1\end{pmatrix}$, portanto, MMD com BLE necessitará de <u>infinitas</u> iterações para chegar na solução.



$$10x_1^2 + x_2^2 = cte$$

$$\frac{x_1^2}{1/10} + x_2^2 = cte$$

$\boxed{5}$ $f: \mathbb{R}^n \to \mathbb{R}$

$$f(x) = \sum_{i=1}^{n} \left(\frac{1}{2} a_i x_i^2 + b_i x_i\right), \quad a_i, b_i \in \mathbb{R}, \; i = 1 \dots n$$

$$\nabla f(x) = \begin{pmatrix} a_1 x_1 + b_1 \\ a_2 x_2 + b_2 \\ \vdots \\ a_n x_n + b_n \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{pmatrix} = A$$

Direcção de Newton:

$$\nabla^2 f(x^k)\, d^k = -\nabla f(x^k)$$

d^k estará bem definida e será de descida se $A > 0$. Para que isso ocorra é preciso que $a_i > 0$, $\forall i = 1 \dots n$.

→ Note que $\nabla^2 f(x^k) = A$ e $d^k = -A^{-1} \nabla f(x^k)$

$\underline{a_i \neq 0}$ $\forall i$ garantem a inversibilidade de A. Para que d^k seja de descida,

$$\nabla f(x^k)^T d^k = -\nabla f(x^k)^T A^{-1} \nabla f(x^k) < 0$$

sempre que $A > 0$, ou seja $\underline{\underline{a_i > 0, \; \forall i = 1 \dots n}}$ são as condições suficientes procuradas.