CONVERGÊNCIA DO MÉTODO DE NEWTON

NORMA DE MATRIZES: A MATRIZ $m \times m$,

$\| \cdot \|$ NORMA EM $\mathbb{R}^m$.

$$\|A\| := \sup_{\|x\|=1} \|Ax\|.$$

EXERCÍCIOS:

1) MOSTRE QUE $\|A\|$ É UMA NORMA NO ESPAÇO DAS MATRIZES.

2) MOSTRE QUE $\|Ax\| \leq \|A\| \cdot \|x\|$.

LEMA: SEJA A_* NÃO SINGULAR. SE $\|A - A_*\| \leq \dfrac{1}{\|A_*^{-1}\|}$,

ENTÃO A É NÃO SINGULAR E $\|A^{-1}\| \leq 2\|A_*^{-1}\|$.

PROVA: VER LIVROS DE ANÁLISE MATRICIAL (GOLUB, WATCHKINS)

HIPÓTESE COMUM:

H1: A FUNÇÃO $\nabla^2 f(x)$ É LIPSCHITZ, ISTO É,

EXISTE $L > 0$ TAL QUE

$$\|\nabla^2 f(\tilde{x}) - \nabla^2 f(x)\| \leq L \|\tilde{x} - x\|.$$

TEOREMA (CONVERGÊNCIA DE NEWTON)

SUPONHA QUE f TENHA DERIVADAS ATÉ A SEGUNDA ORDEM CONTÍNUAS. SEJA x^* TAL QUE $\nabla f(x^*) = 0$. SUPONHA QUE $\nabla^2 f(x^*)$ É NÃO SINGULAR. EXISTE ENTÃO $\varepsilon > 0$ TAL QUE, SE $\|x^0 - x^*\| \leq \varepsilon$ TEMOS

(i) A SEQUÊNCIA DEFINIDA POR $x^{k+1} = x^k - (\nabla f(x^k))^{-1} \nabla f(x^k)$

ESTÁ BEM DEFINIDA.

(O PASSO NEWTONIANO É POSSÍVEL)

(ii) $\lim x^k = x^*$ COM ORDEM SUPERLINEAR.

(iii) SE VALE H1 (ISTO É, SE $\nabla^2 f(x)$ É LIPSCHITZ) ENTÃO A ORDEM DE CONVERGÊNCIA É QUADRÁTICA.

PROVA:

(ii) Supomos por hora que a sequência esteja bem definida para todo k. Temos

$$\|x^{k+1} - x^*\| = \| x^k - x^* - (\nabla^2 f(x^k))^{-1} \nabla f(x^k) \|$$

$$= \| (\nabla^2 f(x^k))^{-1} [-\nabla^2 f(x^k)(x^k - x^*) + \nabla f(x^k) - \nabla f(x^*)] \|$$

$$\Longrightarrow \|x^{k+1} - x^*\| \leq \|\nabla^2 f(x^k)^{-1}\| \cdot \|\nabla f(x^k) - \nabla f(x^*) - \nabla^2 f(x^k)(x^k - x^*)\|$$

CONSIDERAMOS A FUNÇÃO

$$\varphi(t) = \nabla f(t x^k + (1-t) x^*) \quad , \quad t \in [0,1].$$

TEMOS

$$\varphi'(t) = \nabla^2 f(t x^k + (1-t) x^*)(x^k - x^*).$$

PELO O TEO. DO VALOR MÉDIO EXISTE $\bar{t} \in (0,1)$

TAL QUE

$$\varphi'(\bar{t}) = \frac{\varphi(1) - \varphi(0)}{1-0} = \nabla f(x^k) - \nabla f(x^*).$$

DAÍ, PARA TODO K SUFICIENTEMENTE GRANDE,

$$\|x^{k+1} - x^*\| \leq \|\nabla^2 f(x^k)^{-1}\| \cdot \|\nabla^2 f(\bar{t}x^k + (1-\bar{t})x^*)(x^k - x^*) - \nabla^2 f(x^*)(x^k - x^*)\|$$

$$\leq \|\nabla^2 f(x^k)^{-1}\| \, \|x^k - x^*\| \cdot \|\nabla^2 f(\bar{t}x^k + (1-\bar{t})x^*) - \nabla^2 f(x^*)\|$$

$$\leq \|\nabla^2 f(x^k)^{-1}\| \cdot \|x^k - x^*\| \sup_{t \in [0,1]} \|\nabla^2 f(tx^k + (1-t)x^*) - \nabla^2 f(x^*)\|$$

DEFINIMOS

$$r_k = \|\nabla f(x^k)^{-1}\| \cdot \sup_{t \in [0,1]} \|\nabla^2 f(tx^k + (1-t)x^*) - \nabla^2 f(x^*)\|.$$

PARA $t \in [0,1]$ TEMOS

$$\| t\tilde{x} + (1-t)x^* - \tilde{x} \| = (1-t)\|\tilde{x} - x^*\| \leq \|\tilde{x} - x^*\|$$

ASSIM, PELA CONTINUIDADE DE $\nabla^2 f$ e pelo Lema, existe

$\varepsilon > 0$ TAL QUE $\|\tilde{x} - x^*\| \leq \varepsilon \Rightarrow \|\nabla f(\tilde{x})^{-1}\| \leq 2\|\nabla f(x^*)^{-1}\|$

e $2\|\nabla f(x^*)^{-1}\| \sup_{t\in[0,1]} \|\nabla^2 f(t\tilde{x} + (1-t)x^*) - \nabla^2 f(\tilde{x})\| \leq 1/2$

ASSIM, $\|x^0 - x^*\| \leq \varepsilon \Longrightarrow r_0 \leq 1/2$.

COMO

$$\|x^{k+1} - x^*\| \leq r_k \|x^k - x^*\|,$$

SE $\|x^0 - x^*\| \leq \varepsilon \Rightarrow \|x^1 - x^*\| \leq \varepsilon/2$. NOVAMENTE,

$$\|x^1 - x^*\| \leq \frac{\varepsilon}{2} < \varepsilon \Rightarrow r_1 \leq 1/2 .$$

DAÍ,

$$\|x^2 - x^*\| \leq \frac{1}{2} \cdot \frac{\varepsilon}{2} = \frac{\varepsilon}{4} .$$

REPETINDO O ARGUMENTO, TEMOS

$$\|x^k - x^*\| \leq \frac{\varepsilon}{2^k} . \qquad (1)$$

PELA CONTINUIDADE DE $\nabla^2 f$, TEMOS

$$\|x^k - x^*\| \to 0 \Rightarrow 2\|\nabla^2 f(x^*)^{-1}\| \sup_{t \in [0,1]} \|\nabla^2 f(tx^k + (1-t)x^*) - \nabla^2 f(x^k)\| \longrightarrow 0 .$$

OU SEJA, $r_k \to 0$ (A CONVERGÊNCIA É SUPERLINEAR).

(i) Voltamos à boa definição da sequência $\{x^k\}$: pelo Lema e pela continuidade da Hessiana, x^1 está bem definido se tomarmos x^0 próximo a x^*. Da expressão (1), x^1 está próximo à x^*, e logo x^2 também está bem definido. A boa definição x^k segue de forma indutiva, usando o Lema e (1).

(iii) AQUI, ESTAMOS SUPONDO QUE $\nabla^2 f$ É LIPSCHITZ, ISTO É,

$$\|\nabla^2 f(\tilde{x}) - \nabla^2 f(x)\| \leq L\|\tilde{x} - x\|, \quad L > 0.$$

TEMOS

$$\|x^{k+1} - x^*\| \leq r_k \|x^k - x^*\|$$

$$= 2\|\nabla^2 f(x^*)^{-1}\| \sup_{t\in[0,1]} \|\nabla^2 f(tx^k + (1-t)x^*) - \nabla^2 f(x^k)\| \cdot \|x^k - x^*\|$$

$$\leq 2\|\nabla^2 f(x^*)^{-1}\| \sup_{t\in[0,1]} L \underbrace{\| tx^k + (1-t)x^* - x^k\|}_{(1-t)\|x^k - x^*\|} \cdot \|x^k - x^*\|$$

$$= 2\|\nabla f(x^*)^{-1}\| L \underbrace{\sup_{t\in[0,1]}(1-t)}_{1} \cdot \|x^k - x^*\|^2$$

$$\Longrightarrow \quad \|x^{k+1} - x^*\| \leq \left[2L \|\nabla f(x^*)^{-1}\| \right] \cdot \|x^k - x^*\|^2 .$$

(A CONVERGÊNCIA É QUADRÁTICA). $\blacksquare$.

<u>OBS</u>: O MÉTODO DO GRADIENTE ($d^k = -\nabla f(x^k)$) CONVERGE NO MÁXIMO EM ORDEM LINEAR. OU SEJA, LONGE DA SOLUÇÃO $\Rightarrow$ CONVERG. LENTA (COM GRAD.); PERTO DA SOL $\Rightarrow$ CONV. RÁPIDA (NEWTON).