MATH 409/509

LECTURE 8

CONVERGENCE OF NEWTON'S METHOD

There exist a $\delta > 0$ such that

* if $x_0 \in B(x_*, \delta)$ where x_* is a root of $F: \mathbb{R}^n \to \mathbb{R}^m$, then the Newton sequence $\{x_k\}$ converges to x_*

Speed of convergence is affected by two quantities

* M such that

$$\|F'(x_k)^{-1}\| \leq M \quad \text{for all } x_k \in B(x_*$$

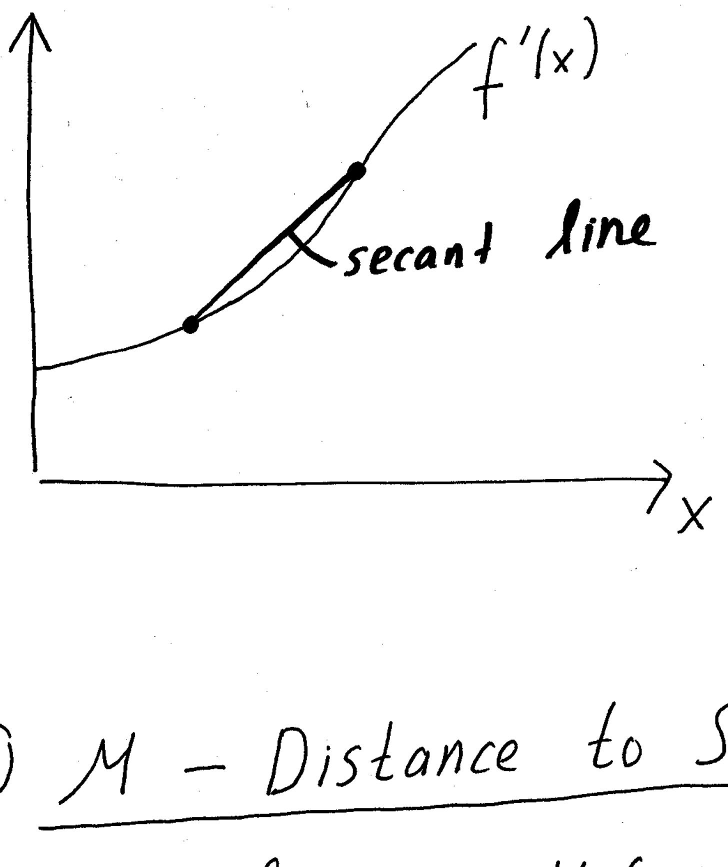
* γ satisfying

$$\|F'(x_k) - F'(x_*)\| \leq \gamma \|x_k - x_*\| \quad \text{for all } x_k \in B(x_*,$$

DEFN (Lipschitz Continuity of Jacobian)
A differentiable function $F: \mathbb{R}^n \to \mathbb{R}^m$ is called Lipschitz continuously differentiable if there exists a $\gamma > 0$ (called the Lipschitz constant) s.t.

$$\| F'(x) - F'(y) \| \leqslant \gamma \| x - y \|$$

for all $x, y \in \mathbb{R}^n$.



Suppose $f: \mathbb{R} \to \mathbb{R}$

Lipschitz continuity of $f'(x)$
Slopes of secant lines on graph of f' never exceeds γ

① M - Distance to Singularity for $F'(x_k)$

The larger M (s.t. $\|F'(x_k)^{-1}\| \leqslant M$) is, the closer $F'(x_k)$ is to singularity.

EXAMPLE

$$F(x) = \begin{bmatrix} 3x_1^2 - x_2 \\ 6x_1 - x_2 - \alpha(x_2^2 - 9) - 3 \end{bmatrix}$$

has a root at $x_* = (1,3)$ for all α.

Jacobian of F at x_*

$$F'(x_*) = \begin{bmatrix} 6 & -1 \\ 6 & -1-6\alpha \end{bmatrix}$$

Newton iterates with $x_0 = (2,2)$

when $\alpha = 10^{-8}$	when $\alpha = \cancel{10^{-8}}\ 1$
$\lVert F(x_1)\rVert \approx 8\times 10^{-1}$	$\lVert F(x_2)\rVert \approx 3\times 10^{-1}$
$\lVert F(x_2)\rVert \approx 2\times 10^{-1}$	$\lVert F(x_3)\rVert \approx 7\times 10^{-3}$
$\lVert F(x_3)\rVert \approx 5\times 10^{-2}$	$\lVert F(x_4)\rVert \approx 5\times 10^{-6}$
$\lVert F(x_4)\rVert \approx 1\times 10^{-2}$	$\lVert F(x_5)\rVert \approx 2\times 10^{-12}$
ALMOST LINEAR	QUADRATIC

$F'(x_*)$ becomes singular as $\alpha \to 0$.
Convergence is slower as $\alpha \to 0$.

(2) γ – Lipschitz Constant for $F'(x_k)$

γ is large $\Longrightarrow$ $F'(x)$ changes rapidl

$\Longrightarrow$ Second derivatives are large (in absolute valu

EXAMPLE

$$g(x) = x^3 + \alpha x^2 + x$$

root at $x_* = 0$ (has two other roots for $\alpha \in [-2,$

$$g'(x) = 3x^2 + 2\alpha x + 1$$

Newton iterates with $x_0 = 2$. (In both cases below $\lim_{k \to \infty} x_k = 0$.)

when $\alpha = \cancel{2000}\ 2$	when $\alpha = 2000$
$x_5 \approx 9 \times 10^{-3}$	$x_5 \approx 6 \times 10^{-2}$
$x_6 \approx 1 \times 10^{-4}$	$x_6 \approx 3 \times 10^{-2}$
$x_7 \approx 4 \times 10^{-8}$	$x_7 \approx 2 \times 10^{-2}$
$x_8 \approx 3 \times 10^{-15}$	$x_8 \approx 8 \times 10^{-3}$
QUADRATIC	ALMOST LINEAR

$g''(x_*)$ is larger as $\alpha \to 0$.
Convergence is slower as $\alpha \to 0$.

THM (Convergence of Newton's Method)
Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable, and x_* be a root of F such that $F'(x_*)$ is not singular. Then the following holds for the Newton sequence $\{x_k\}$.

(i) There exists a $\delta > 0$ such that for all $x_0 \in B(x_*, \delta)$

$$\lim_{k \to \infty} x_k = x_*.$$

(ii) Indeed for all $x_0 \in B(x_*, \delta)$

$$\lim_{k \to \infty} \frac{\| x_{k+1} - x_* \|}{\| x_k - x_* \|} = 0.$$

(iii) Additionally if there exists a $\gamma > 0$ s.t.

$$\| F'(x) - F'(\tilde{x}) \| \leq \gamma \| x - \tilde{x} \|,$$

for all $\tilde{x}, x \in B(x_*, \delta)$

then

$$\lim_{k \to \infty} \frac{\| x_{k+1} - x_* \|}{\| x_k - x_* \|^2} = c$$

for some constant $c \geq 0$.

· PROOF

(i) $\boxed{\text{CONVERGENCE}}$

Due to

* continuity of the Jacobian, and
* the assumption that $F'(x_*)$ is invertible,

there exists a $\delta > 0$ s.t. for some M, θ satisfying $M\theta < 1$ we have

$$\|F'(x)^{-1}\| \leqslant M \quad \text{and} \quad \|F'(x) - F'(\tilde{x})\| \leqslant \theta$$

for all $x, \tilde{x} \in B(x_*, \delta)$.

It follows that

$$\begin{aligned} x_{k+1} - x_* &= x_k + p_k - x_* \\ &= x_k - F'(x_k)^{-1} F(x_k) - x_* \\ &= F'(x_k)^{-1}\big(\underbrace{F(x_*) - F(x_k) - F'(x_k)(x_* - x_k)}_{F(x_*) - L(x_*)}\big) \end{aligned}$$

$\Longrightarrow$

$$(*) \quad \|x_{k+1} - x_*\| \underset{\substack{(\text{SUBMULTIPLICATIVE} \\ \text{PROPERTY})}}{\leqslant} \|F'(x_k)^{-1}\| \, \|F(x_*) - L(x_*)\|$$

$$\leqslant M \|F(x_*) - L(x_*)\|$$

Furthermore

$$(\tilde{*}) \quad \|F(x_*) - L(x_*)\| = \Big\| F(x_k) + \int_0^1 F'(x_k + t(x_* - x_k))(x_* - x_k)\,dt$$

$$- F(x_k) - F'(x_k)(x_* - x_k)\Big\|$$

$$= \left\| \int_0^1 \left(F'(x_k + t(x_* - x_k)) - F'(x_k)\right)(x_* - x_k)\, dt \right.$$

$$\leq \int_0^1 \| F'(x_k + t(x_* - x_k)) - F'(x_k) \| \, \| x_* - x_k \| \, d$$

Now suppose $x_k \in B(x_*, \delta)$. Then $\underbrace{x_k + t(x_* - x_k)}_{x_k(1-t) + t x_*} \in B($

Consequently

$$(**) \quad \| F(x_*) - L(x_*) \| \leq \underbrace{\int_0^1 \theta \, \| x_* - x_k \| \, dt}_{\theta \| x_* - x_k \|}$$

Combining $(*)$ and $(**)$ we obtain

$$\| x_{k+1} - x_* \| \leq M\theta \, \| x_k - x_* \|$$

which implies $x_{k+1} \in B(x_*, \delta)$ and

$$\| x_{k+1} - x_* \| \leq M^{k+1} \theta^{k+1} \| x_k - x_* \|$$

$$\Longrightarrow$$

$$\lim_{k \to \infty} \| x_k - x_* \| = 0 \qquad \square$$

(ii) $\boxed{\text{SUPERLINEAR CONVERGENCE}}$

~~Once again combine (*) and (**) to obtain~~

$$\cancel{\| x_{k+1} - x_* \| \leq}$$

~~From ($\tilde{*}$) we have~~

- Combine $(*)$ and $(\tilde{*})$ to obtain

$$\|x_{k+1} - x_*\| \leqslant M \int_0^1 \|F'(x_k + t(x_* - x_k)) - F'(x_k)\| dt \, \|x_k - x$$

$$\Longrightarrow$$

$$(***) \quad \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leqslant M \int_0^1 \|F'(x_k + t(x_* - x_k)) - F'(x_k)\| d$$

Since $x_k \to x_*$ as $k \to \infty$, by the continuity of the Jacobian the integral on right goes to 0 implying

$$\overline{\lim_{k \to \infty}} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

□

(iii) $\boxed{\text{QUADRATIC CONVERGENCE}}$

When $F'(x)$ is Lipschitz continuous with Lipschitz constant γ $(***)$ implies

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leqslant M \int_0^1 \gamma \|t(x_k - x_*)\| dt$$

$$= \frac{M\gamma t^2}{2}\Big|_0^1 \|x_k - x_*\|$$

$$\Longrightarrow$$

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2} \leq \frac{M\gamma}{2}$$

□