

## ① Reading.

Quarteroni et al.

ch 6. §6.2.2 (The secant method, Newton's method). §6.3

ch 7. §7.1. §7.1.1. §7.1.3 §7.1.4 §7.1.5

§7.2 §7.2.2 §7.2.3

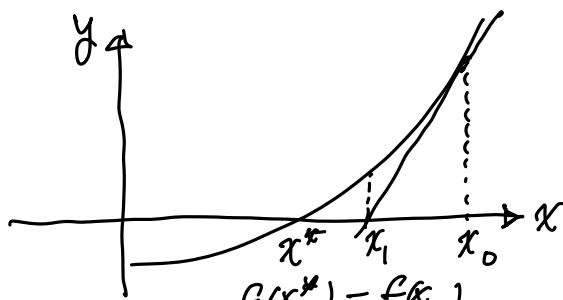
Stoer &amp; Bulirsch: chs. §5.1 - §5.3

## ② HW #1 due Wed. 1/18

## ③ Today: Convergence of Newton's method

General concept of convergence order

Newton's method for systems of equations

Convergence of Newton's method  $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (k=0, 1, \dots)$ Assume  $f(x^*) = 0$ .Denote  $e_k = x_k - x^*$ 

$$e_{k+1} = e_k + \frac{f(x^*) - f(x_k)}{f'(x_k)} = e_k + \frac{f'(x_k)(x^* - x_k) + \frac{1}{2} f''(\xi_k) e_k^2}{f'(x_k)} = \frac{f''(\xi_k)}{2 f'(x_k)} e_k^2.$$

$$|e_{k+1}| \leq C |e_k|^2, \text{ where } C = \frac{1}{2} \frac{\sup |f''|}{\inf |f'|}.$$

Theorem Let  $f \in C^1([a, b])$ ,  $a < x^* < b$ , and  $f(x^*) = 0$ . Assume there exist  $\delta > 0$ ,  $L > 0$ , and  $\rho > 0$  such that

$$|f'(x)| \geq \rho > 0 \quad \forall x \in (x^* - \delta, x^* + \delta),$$

$$|f'(x) - f'(y)| \leq L|x - y| \quad \forall x, y \in (x^* - \delta, x^* + \delta),$$

$$\delta < \rho/L.$$

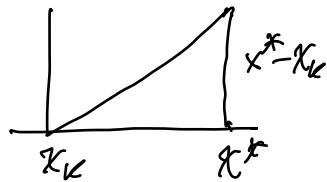
Then, for any  $x_0 \in (x^* - \delta, x^* + \delta)$ , the Newton iteration

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (k=0, 1, \dots)$$

satisfies  $|x_{k+1} - x^*| \leq \frac{L}{2\rho} |x_k - x^*|^2$  ( $k=0, 1, \dots$ ).

Proof

$$\begin{aligned}
 e_{k+1} &= x_{k+1} - x^* = x_k - x^* - \frac{f(x_k)}{f'(x_k)} \\
 &= \frac{1}{f'(x_k)} \left[ (x_k - x^*) f'(x_k) + f(x^*) - f(x_k) \right] \\
 &= \frac{1}{f'(x_k)} \left[ f(x^*) - f(x_k) - (x^* - x_k) f'(x_k) \right] \\
 &= \frac{1}{f'(x_k)} \int_{x_k}^{x^*} [f'(x) - f'(x_k)] dx \\
 |e_{k+1}| &\leq \frac{1}{|f'(x_k)|} \left| \int_{x_k}^{x^*} |f'(x) - f'(x_k)| dx \right| \\
 &\leq \frac{1}{\rho} \left| \int_{x_k}^{x^*} L |x - x_k| dx \right| \\
 &= \frac{L}{2\rho} e_k^2.
 \end{aligned}$$



$k=0$ ,  $x_0 \in (x^* - \delta, x^* + \delta)$ . So,

$$|e_1| = |x_1 - x^*| \leq \frac{L}{2\rho} |x_0 - x^*|^2 \leq \frac{L}{2\rho} \delta \cdot \delta \leq \frac{\delta}{2}.$$

So,  $x_1 \in (x^* - \delta, x^* + \delta)$ ,  $|e_1| \leq \frac{L}{2\rho} e_0^2$

Induction: true for  $k$ .  $\Rightarrow x_k \in (x^* - \delta, x^* + \delta)$  and  $|e_k| \leq \frac{L}{2\rho} e_{k-1}^2$ . Then, same argument  $\Rightarrow$  true for  $k+1$ , i.e.,  $x_{k+1} \in (x^* - \delta, x^* + \delta)$  and  $|e_{k+1}| \leq \frac{L}{2\rho} e_k^2$ . QED

### Fixed-point iteration and concept of convergence

Consider solving  $f(x) = 0$  for  $x \in E \subseteq \mathbb{R}^n$ . Here,

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad E \neq \emptyset, \text{ open.}$$

Construct a function  $\Phi : E \rightarrow E$  such that  $\Phi(x) = x$ ,  $\Leftrightarrow f(x) = 0$ . Thus,

Find  $x \in E$ :  $f(x) = 0 \Leftrightarrow$  Find  $x \in E$ :  $\Phi(x) = x$ .

If  $x^* \in E$  satisfies  $\Phi(x^*) = x^*$ , then call  $x^*$  a fixed point of  $\Phi: E \rightarrow E$ . 3

A fixed-point iteration.

Given  $x_0 \in E$ ,  $x_{k+1} = \Phi(x_k)$ ,  $k=0, 1, \dots$  (\*)

① If  $\Phi$  is continuous, and  $x_k \rightarrow \hat{x}$ , then

$$\hat{x} = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} \Phi(x_k) = \Phi\left(\lim_{k \rightarrow \infty} x_k\right) = \Phi(\hat{x}),$$

i.e., the limit  $\hat{x}$  is a fixed point of  $\Phi$ .

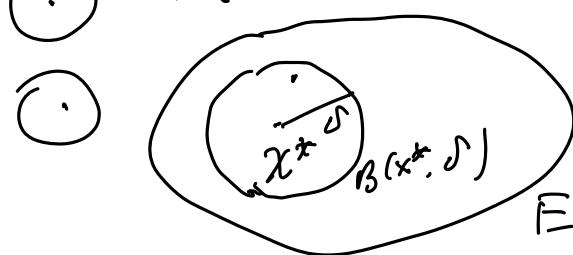
② Newton's method (one-variable)

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \quad \Phi(x) = x - \frac{f(x)}{f'(x)}.$$

Definition Assume  $\Phi: E \rightarrow E$  has a unique fixed point  $x^* \in E$ . The iteration (\*) converges (locally), if there exists  $\delta > 0$  such that the ball  $B(x^*, \delta) \subseteq E$  and  $x_k \rightarrow x^*$  for any  $x_0 \in B(x^*, \delta)$ .

①  $B(x^*, \delta) = \{x \in \mathbb{R}^n : \|x - x^*\| < \delta\}$ .  $\|\cdot\|$ : norm on  $\mathbb{R}^n$ .

②  $x_k \rightarrow x^*$  means  $\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$ .



Note: All norms of  $\mathbb{R}^n$  are equivalent. So, the convergence is indep. of the choice of  $\|\cdot\|$ .

Definition (Order of convergence) Assume  $\Phi: E \rightarrow E$  has a unique fixed point  $x^* \in E$ . Suppose  $p \geq 1$ . If there exist  $C > 0$  for  $p > 1$  or  $C \in (0, 1)$  for  $p = 1$  and  $\delta > 0$  such that for any  $x_0 \in B(x^*, \delta)$  the iteration (\*) satisfies

$$\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|^p \quad (k=0, 1, \dots).$$

then (\*) is (locally) convergent of order (at least)  $p$ .

$p=1$ : linear convergence.  
 $p=2$ : quadratic convergence.

Remarks Denote  $e_k = x_k - x^*$ .

(1) Convergence of order  $p \geq 1 \Rightarrow$  convergence.

$p=1$ .  $0 < c < 1$ .  $\|e_0\| \in C \quad \|e_{k+1}\| \leq \dots \leq C^k \|e_0\| \rightarrow 0$ .

$p > 1$ . choose  $\hat{\delta} = \min(\delta, (\frac{1}{2C})^{\frac{1}{p-1}})$ .

Let  $x_0 \in B(x^*, \hat{\delta}) \subseteq B(x^*, \delta)$ .

Claim:  $\|e_k\| \leq 2^{-k} \|e_0\| \quad (k=1, 2, \dots)$ . So,  $\|e_k\| \rightarrow 0$ .

Induction:

$$\|e_k\| \leq c \|e_0\|^p = 2c \|e_0\|^{p-1} \cdot \frac{1}{2} \|e_0\| \leq 2c \left(\frac{1}{2C}\right) \cdot 2^{-1} \|e_0\|.$$

Assume  $\|e_k\| \leq 2^{-k} \|e_0\|$ . Then  $x_k \in B(x^*, \hat{\delta})$ . since  $x_0 \in B(x^*, \hat{\delta})$ . Thus, same argument leads to

$$\|e_{k+1}\| \leq c \|e_k\| \leq 2^{-1} \|e_k\| \leq 2^{-k-1} \|e_0\|.$$

(2) Let  $1 \leq q < p$ . Then local convergence of order  $p \Rightarrow$  order  $q$ .

Choose  $\hat{\delta} = \min \left\{ \delta, \left(\frac{1}{2C}\right)^{\frac{1}{p-1}} \right\}$ .  $x_0 \in B(x^*, \hat{\delta})$ .

Already proved:  $\|e_k\| \leq 2^{-k} \|e_0\| \Rightarrow x_k \in B(x^*, \hat{\delta}) \quad (k=0, 1, \dots)$ .

$$\begin{aligned} \|e_{k+1}\| &\leq c \|e_k\|^p = c \|e_k\|^{p-q} \|e_k\|^q \\ &\leq c \hat{\delta}^{p-q} \|e_k\|^q = \tilde{c} \|e_k\|^q, \quad \tilde{c} = c \hat{\delta}^{p-q} > 0. \end{aligned}$$

Theorem Newton's method converges quadratically.

Newton's method for systems of equations

---


$$f(x) = 0, \quad f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in E, \quad E \subseteq \mathbb{R}^n, \quad E: \text{open}$$

$x^* \in E$ .  $f(x^*) = 0$ .

[5]

The Jacobian matrix of  $f$  (if all partial derivatives  $\frac{\partial f_i}{\partial x_j}$  exist) is  $Df(x) = \left[ \frac{\partial f_i}{\partial x_j} \right] (= Jf(x))$

$$y = f(x) = f(x^{(k)}) + Df(x^{(k)})(x - x^{(k)}) + \dots$$

$$y=0 \Rightarrow x=x^{(k+1)}. 0 = f(x^{(k)}) + Df(x^{(k)}) (x^{(k+1)} - x^{(k)}).$$

$$Df(x^{(k)}) (x^{(k+1)} - x^{(k)}) = -f(x^{(k)})$$

$$x^{(k+1)} - x^{(k)} = -[Df(x^{(k)})]^{-1} f(x^{(k)})$$

$$\boxed{x^{(k+1)} = x^{(k)} - [Df(x^{(k)})]^{-1} f(x^{(k)}), k=0, 1, \dots}$$

We shall denote by  $\|\cdot\|$  a norm on  $\mathbb{R}^n$  and also the induced matrix norm on  $\mathbb{R}^{n \times n}$ , the space of all  $n \times n$  real matrices. Recall  $\overline{B(x^*, \delta)} = \{x \in \mathbb{R}^n : \|x - x^*\| \leq \delta\}$ .

Theorem Let  $E \neq \emptyset$  be an open subset of  $\mathbb{R}^n$ . Assume  $f \in C^1(E, \mathbb{R}^n)$  and  $f$  has a unique root  $x^* \in E$ :  $f(x^*) = 0$ . Assume there exist  $\delta > 0$ ,  $p > 0$ , and  $L > 0$  satisfying:

- (1) The closed ball  $\overline{B(x^*, \delta)} \subseteq E$  and  $\delta \leq \frac{p}{L}$ ;
- (2)  $Df(x)$  is nonsingular and  $\|[Df(x)]^{-1}\| \leq \frac{1}{p}$  for any  $x \in \overline{B(x^*, \delta)}$ ; and
- (3)  $\|Df(x) - Df(y)\| \leq L \|x - y\| \quad \forall x, y \in \overline{B(x^*, \delta)}$ . Then, for any  $x^{(0)} \in \overline{B(x^*, \delta)}$ . Newton's iteration

$$x^{(k+1)} = x^{(k)} - [Df(x_k)]^{-1} f(x_k) \quad (k=0, 1, \dots)$$

Satisfies

$$\|x^{(k+1)} - x^{(0)}\| \leq \frac{L}{2p} \|x^{(k)} - x^{(0)}\|^2 \quad (k=0, 1, \dots)$$

The proof of theorem is similar to that for single equations. But, here are a few points that are different. (6)

①  $f(x) - f(y) = \int_0^1 g'(t) dt = \int_0^1 Df(tx + (1-t)y)(x-y) dt$

where  $g(t) = f(tx + (1-t)y)$ .

- ② If  $A_1, A_2$  are  $n \times n$  matrices and  $x \in \mathbb{R}^n$  then  $\|A_1 A_2\| \leq \|A_1\| \|A_2\|$  and  $\|A_1 x\| \leq \|A_1\| \|x\|$  since the matrix norm is induced from the vector norm.