Problem:

Given $F: \mathbb{R}^n \to \mathbb{R}^n$, find x_* such that $F(x_*) = 0$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{pmatrix}$$

Newton's method

Given an initial x
until termination do

$$x \leftarrow x - F'(x)^{-1} F(x) \left\{ \begin{array}{ll} \text{solve} & F'(x) \cdot s = -F(x) \\ \text{update} & x \leftarrow x + s \end{array} \right.$$

Theorem: Suppose F is continuously differentiable near x_* such that $F(x_*) = 0$ and $F'(x_*)$ nonsingular. Then, $\exists\, \delta > 0$ such that if $\| x - x_* \| \leq \delta$, then Newton iterates $\{x_k\}$ converge to x_*

<u>Superlinearly</u>

- If F is Lipschitz continuously differentiable at x_* i.e,

$$\| F'(x) - F'(x_*) \| \leq L \| x - x_* \|$$

for x near x_*, then the convergence is <u>quadratic</u>

proof: claim: $F'(x)^{-1}$ exist and is continuous for x near x_*

Follows from Long aside below:

have: $x_+ = x - F'(x)^{-1} F(x)$

$$\Rightarrow x_+ - x_* = -F'(x)^{-1} \Big\{ \underbrace{F(x) - \cancel{F(x_*)}^{\,0} - F'(x_*)(x - x_*)}_{1} - \underbrace{\left[F'(x) - F'(x_*) \right](x - x_*)}_{2} \Big\} \quad (*)$$

see: (1) is $o(x-x_*)$ by differentiability property
and (2) is also $o(x-x_*)$ by continuity of $F'(x)$
since $F'(x)$ is bounded near x_*, we should be
able to get superlinear convergence from $(*)$

Choose $\delta > 0$ such that if $\|x - x_*\| \leq \delta$ then

(1) $F'(x)^{-1}$ exists and $\|F'(x)^{-1}\| \leq 2\|F'(x_*)^{-1}\|$

(2) $\omega(\delta) \equiv \max\limits_{\|d\|<\delta} \|F'(x_*+d) - F'(x_*)\|$

is such that $4\|F'(x_*)^{-1}\| \, \|\omega(\delta)\| \equiv \lambda < 1$

remark: $\omega(\delta)$ is the modulus of continuity of F' at x_*

$\omega(\delta) \to 0$ as $\delta \to 0$ since F' is continuous at x_*

Then for $\|x - x_*\| \leq \delta$

$x_+ = x - F'(x)^{-1} F(x)$ is defined

$\|F'(x)^{-1}\| \leq 2\|F'(x_*)^{-1}\|$ (does not depend on x)

$$\|[F'(x) - F'(x_*)](x - x_*)\| \leq \omega(\|x - x_*\|)\, \|x - x_*\|$$

$$\|F(x) - F(x_*) - F'(x_*)(x - x_*)\| \leq \omega(\|x - x_*\|)\, \|x - x_*\|$$

by Handy Lemma (version 1) below

Then $(*)$ gives

$$\|x_+ - x_*\| \leq 4\|F'(x_*)^{-1}\| \, \omega(\|x - x_*\|)\, \|x - x_*\| \quad \cdots \quad (a)$$

$$\leq \underbrace{4\|F'(x_*)^{-1}\| \, \omega(\delta)}_{\leq \lambda < 1} \|x - x_*\| \quad \cdots \quad (b)$$

By (b) have

$$\| x_+ - x_* \| \leq \lambda \| x - x_* \|$$

so if

$\| x_0 - x_* \| \leq \delta$, an easy induction shows that Newton iterates $\{x_k\}$ satisfy $\| x_k - x_* \| < \delta$ and $x_k \to x_*$ <u>linearly</u> with

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

By (a) have

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

where

$$c_k = 4 \| F'(x_*)^{-1} \| \, \omega (\| x_k - x_* \|) \longrightarrow 0$$

$\Rightarrow$ superlinear convergence

Suppose

$$\| F'(x) - F'(x_*) \| \leq L \| x - x_* \| \quad \text{for } x \text{ near } x_*$$

Then

$$\omega (\| x - x_* \|) \equiv \max_{\|d\| \leq \|x - x_*\|} \| F'(x_* - d) - F'(x_*) \|$$

$$\leq \max_{\|d\| \leq \|x - x_*\|} L \|d\| = L \| x - x_* \|$$

so

$$\| x_{k+1} - x_* \| \leq 4 \| F'(x_*)^{-1} \| \, L \, \| x_k - x_* \|^2$$

quadratic convergence

Long Aside

Lemma: Suppose $B \in \mathbb{R}^{n \times n}$ and $\|B\| < 1$

Then $(I - B)$ is invertible and

$$(I-B)^{-1} = \sum_{j=0}^{\infty} B^j \qquad \text{Neuman series}$$

proof:

Set $S_m = \sum_{j=0}^{m} B^j$

norm of product $\leq$ prod. of norms

see

$$\| S_{m+\ell} - S_m \| = \Big\| \sum_{j=m+1}^{m+\ell} B^j \Big\| \leq \sum_{j=m+1}^{m+\ell} \|B\|^j$$

$$= \|B\|^{m+1} \sum_{j=0}^{\ell-1} \|B\|^j \leq \frac{\|B\|^{m+1}}{1 - \|B\|}$$

$$\longrightarrow 0 \quad \text{as} \quad m \to \infty \text{, independent of } \ell$$

Then $\{S_m\}$ is Cauchy and so $S_m \to S \in \mathbb{R}^{n \times n}$

Denote

$$S = \sum_{j=0}^{\infty} B^j$$

have

$$(I-B)S = \lim_{m \to \infty} (I-B) \sum_{j=0}^{m} B^j = \lim_{m \to \infty} I - B^{m+1} = I$$

$$(1-x) \sum_{j=0}^{m} x^j = 1 - x^{m+1}$$

Similarly $S(I-B) = I$

Lemma: Suppose $A \in \mathbb{R}^{n \times n}$ is invertible and $E \in \mathbb{R}^{n \times n}$ is such that $\|A^{-1}E\| < 1$

Then $A+E$ is invertible in

$$(A+E)^{-1} = \left(\sum_{j=0}^{\infty} (-A^{-1}E)^j \right) A^{-1}$$

and

$$\|(A+E)^{-1} - A^{-1}\| = \left\| \left\{ \sum_{j=0}^{\infty} (-A^{-1}E)^j - I \right\} A^{-1} \right\|$$

$$\leq \|A^{-1}\| \frac{\|A^{-1}E\|}{1 - \|A^{-1}E\|}$$

proof: $(A+E) = A(I + A^{-1}E) \Rightarrow (A+E)$ is invertible and

$(A+E)^{-1} = \left(\sum_{j=0}^{\infty} (-A^{-1}E)^j \right) A^{-1}$ by previous lemma

Corollary: If $F : \mathbb{R}^n \to \mathbb{R}^n$ is such that $F'(x)$ is invertible and F' is (Lipschitz) continuous at x, then $F'(y)$ is invertible for y near x and $F'(y)^{-1}$ is (Lipschitz) continuous at x

proof: Apply previous lemma with

$A = F'(x)$, $E = F'(y) - F'(x)$

get: $F'(y)^{-1}$ exists and

$$\|F'(y)^{-1} - F'(x)^{-1}\| \leq \|F'(x)^{-1}\| \frac{\|F'(x)^{-1}(F'(y) - F'(x))\|}{1 - \|F'(x)^{-1}(F'(y) - F'(x))\|}$$

Handy Lemma (Version 1)

Suppose F is continuously differentiable near x. Then for small h

$$\| F(x+h) - F(x) - F'(x)h \| \leqslant \omega(\|h\|)\, \|h\|$$

where

$$\omega(\|h\|) = \max_{\|d\| \leqslant \|h\|} \| F'(x+d) - F'(x) \|$$

proof:

$$F(x+h) - F(x) = \int_0^1 \frac{d}{dt} F(x+th)\, dt = \int_0^1 F'(x+th)h\, dt$$

$$F'(x)\cdot h = \left[\int_0^1 dt \right] F'(x)\cdot h = \int_0^1 F'(x) h\, dt$$

$$\Rightarrow \| F(x+h) - F(x) - F'(x)h \| = \left\| \int_0^1 [F'(x+th) - F'(x)]\, h\, dt \right\|$$

$$\leqslant \int_0^1 \| F'(x+th) - F'(x) \|\, \|h\|\, dt$$

$$\leqslant \int_0^1 \omega(t\|h\|)\, \|h\|\, dt$$

$$\leqslant \omega(\|h\|)\, \|h\|$$

Handy Lemma (version 2)

Suppose F is continuously differentiable near x and Lipschitz continuous at x i.e. $\|F(y) - F(x)\| \leqslant L \| y - x \|$ for y near x. Then for small h

$$\| F(x+h) - F(x) - F'(x)\cdot h \| \leqslant \frac{L}{2} \|h\|^2$$

proof:

$$\| F(x+h) - F(x) - F'(x)\cdot h \| = \left\| \int_0^1 [F'(x+th) - F'(x)]\, h\, dt \right\|$$

$$\leq \int_0^1 \| F'(x+th) - F'(x) \| \, \|h\| \, dt$$

$$\leq \int_0^1 L t \|h\|^2 dt = \frac{L}{2} \|h\|^2$$

What if hypothesis of convergence theorem don't hold?

Example: for $f : \mathbb{R}^1 \to \mathbb{R}^1$ given by $f(x) = x^2$

$$x_+ = x - \frac{x^2}{2x} = \frac{1}{2} x$$

$\Rightarrow \quad x_k \to 0$ linearly

proposition. suppose $f : \mathbb{R}^1 \to \mathbb{R}^1$ is $(m+1)$ times continuously diff'ble near x_* and

$$f(x_*) = f'(x_*) = \dots = f^{(m)}(x_*) = 0$$

but $f^{(m+1)}(x) \neq 0$, Then for x_0 near x_* (but $x_0 \neq x_*$) Newton iterates $\{x_k\}$ are well-defined and

$$\lim \frac{|x_{k+1} - x_*|}{|x_k - x_*|} = \frac{m}{m+1}$$

proof sketch:

$$f(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!} (x - x_*)^{m+1}, \quad \text{some } \xi \text{ between } x \text{ and } x_*$$

$$f'(x) = \frac{f^{(m+1)}(\eta)}{m!} (x - x_*)^{m}, \quad \text{some } \eta \text{ between } x \text{ and } x_*$$

so

$$x_+ = x - \frac{\dfrac{f^{(m+1)}(\xi)}{(m+1)!}(x-x_*)^{m+1}}{\dfrac{f^{(m+1)}(\eta)}{m!}(x-x_*)^{m}}$$

$$\Rightarrow \quad x_+ - x_* = \underbrace{\left\{ 1 - \frac{1}{m+1}\,\frac{f^{(m+1)}(\xi)}{f^{(m+1)}(\eta)} \right\}}(x-x_*)$$

$$\frac{m}{m+1} + \frac{1}{m+1}\left(1 - \frac{f^{(m+1)}(\xi)}{f^{(m+1)}(\eta)}\right) \approx \frac{m}{m+1}(x-x_*) \quad \text{for } x \text{ near } x_*$$

<u>Example</u>: $f(x) = x + \frac{2}{3}|x|^{3/2}\operatorname{sign}(x) = x\left(1+\frac{2}{3}|x|^{1/2}\right)$

$$f'(x) = 1 + |x|^{1/2}$$

so f is continuously differentiable at $x_* = 0$ but not Lipschitz continuously differentiable at $x_* = 0$

$$x_+ = x - \frac{x + \frac{2}{3}|x|^{3/2}\operatorname{sign}(x)}{1+|x|^{1/2}}$$

$$= \left(1 - \frac{1+\frac{2}{3}|x|^{1/2}}{1+|x|^{1/2}}\right)x = \left(\frac{\cancel{1} - |x|^{1/2} - \cancel{1} - \frac{2}{3}|x|^{1/2}}{1+|x|^{1/2}}\right)x$$

$$= -\frac{1}{3}\,\frac{1}{1+|x|^{1/2}}\,|x|^{1/2}\,x$$

$$|x_+ - \underset{=0}{x_*}| \le c\,|x-0|^{3/2}$$

$\Rightarrow$ order $\frac{3}{2}$ convergence

More generally:

Def: For $0 < p \leq 1$, g is Hölder continuous at x with exponent p if

$$\| g(y) - g(x) \| \leq \gamma \| y - x \|^{p}$$

(If $p = 1$ → Lipschitz continuous ⇒ generalization of Lipschitz continuity)

Example:

$$g(x) = |x|^{1/2}$$

proposition: Suppose F is continuously differentiable near x_* and Hölder continuously differentiable with exponent p at x_* such that $F(x_*) = 0$, $F'(x_*)$ nonsingular

If $\{x_k\}$ is a Newton sequence that converges to x_* then

$$\| x_{k+1} - x_* \| \leq \gamma \| x_k - x_* \|^{1+p}$$

for some γ independent of k

proof: slight modification of proof of theorem

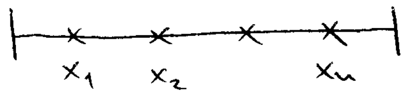
For 2^{nd} Homework

- Must solve a discretized BVP

$$u'' + \mu u(1-u) = 0 \qquad , \quad u(0) = u(1) = 0$$

steady state version of

$$\left.\begin{aligned} u_t &= u_{xx} + \mu u(1-u) \\ u(0,t) &= u(1,t) = 0 \\ u(x,0) &= f(x) \end{aligned}\right\} \text{ Fischer equation}$$



n – number of interior mesh points

$$h = \frac{1}{n+1}$$

$$x_i = ih \quad , \quad 1 \leq i \leq n$$

$$u_i \approx u(x_i)$$

$$\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} = u''(x_i) + O(h^2)$$

(from Taylor series expansions)

suggest

$$u''(x_i) \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

Then

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + \mu u_i (1 - u_i) = 0 \qquad \text{for } i = 1, \dots, n$$

with $u_0 = u_{n+1} = 0$

This gives

$$F(u) = 0 \qquad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

see

$$F(u) = Lu + \mu \begin{pmatrix} u_1(1-u_1) \\ \vdots \\ u_1(1-u_n) \end{pmatrix}$$

$$F'(u) = L +$$

where $L = -\frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$