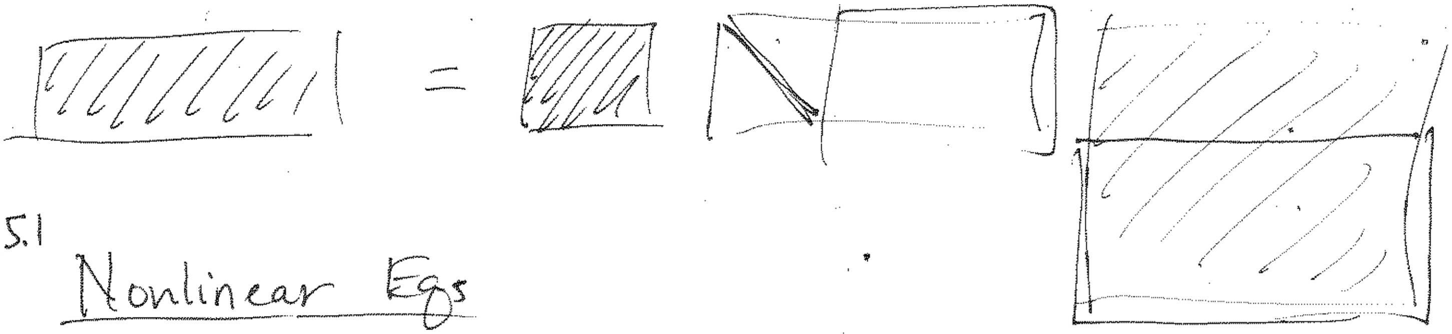


§5.1
Nonlinear Eqs

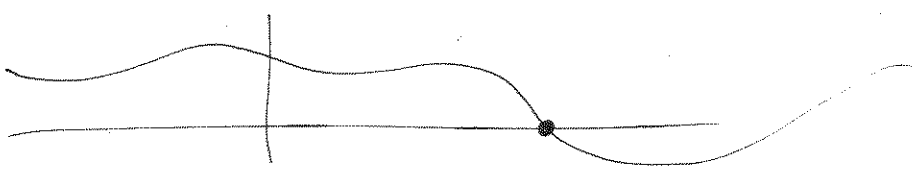
more difficult than <u>linear</u>!

consider scalar case first.

$$g(x) = y$$
$$g(x) - y = 0$$
$$f(x) = 0$$

find "root" or "zero"
Solution = root or zero

~~$f(x) = y$~~
~~$f(x) - y = 0$~~

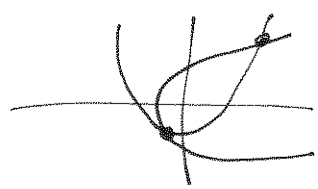


§ 5.2 Existence + Uniqueness

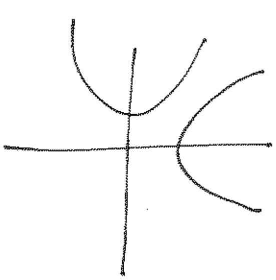
- nonlinear equations can have any # of solutions
- each equation is a hypersurface.

<u>Ex. 5.2</u>

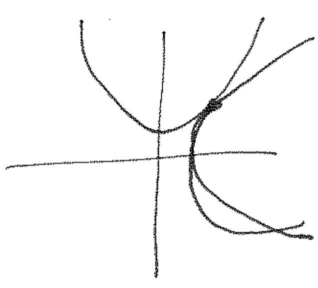
$$x^2 - y + \alpha = 0$$
$$-x + y^2 + \alpha = 0$$



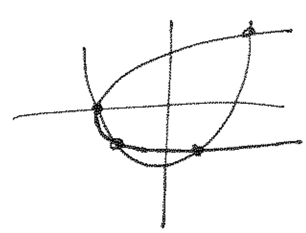
$\alpha = -\frac{1}{2}$



$\alpha = \frac{1}{2}$



$\alpha = \frac{1}{4}$



$\alpha = -1$

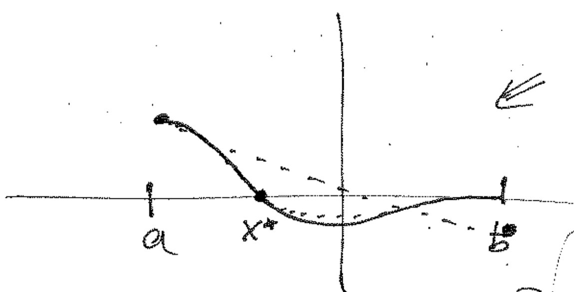
Ex. 8.3

	# solutions
$e^x + 1 = 0$	0
$e^{-x} - x = 0$	1
$x^2 - 4\sin x = 0$	2
$x^3 + 6x^2 + 11x - 6 = 0$	3
$\sin(x) = 0$	inf

Lectures 8 & 9

Nonlinear Equations

Existence - intermediate value theorem

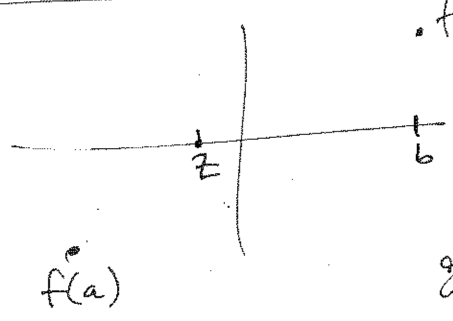


f continuous on $[a,b]$ $\Leftarrow$ for any $c \in [f(b), f(a)]$

$\Rightarrow \exists$ $\exists x^*$ st. $f(x^*) = c$

$\Rightarrow$ $\boxed{\text{if } f(a) \cdot f(b) < 0 \text{ there's a root in } [a, b]}$

$[a, b]$ is a "bracket" for $f(x) = 0$.
solution of



$(b-z) f(b) \geq 0 \iff f(b) \geq 0$

$(a-z) f(a) \geq 0 \iff f(a) \leq 0$

$(x-z) f(x) \geq 0 \quad x = a, b$

generalization $(x-z)^T f(x) \geq 0$

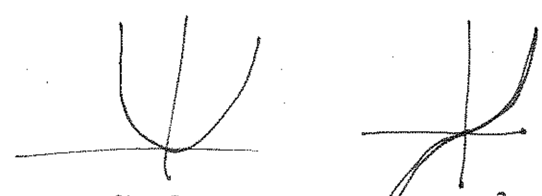
simple root $f'(x^*) \neq 0$ $\quad J_f(x^*)$ invertible.

multiple root $f'(x^*) = 0$.

$f'(x^*) \sim$ cond. #.

$J_f(x^*)$

$x^2 - 2x + 1$ $\quad$ $x^3 - 3x^2 + 3x - 1$



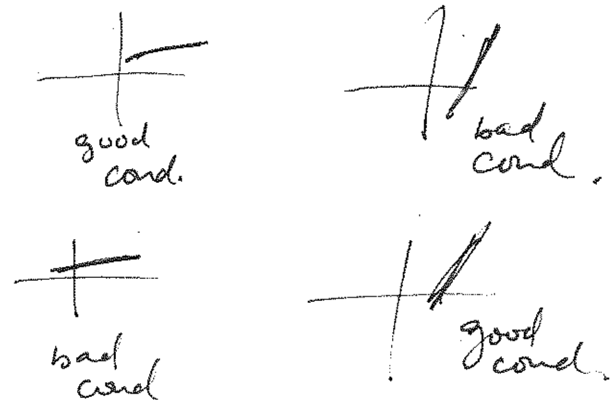
$m = 2$ $\quad$ $m = 3.$

§5.3 Sensitivity & conditioning

opposite sensitivity

evaluation $f(x) = y$

solution (root) $f(x) = y$



cond # for eval $= |f'(x^*)|$

root $= \left|\frac{1}{f'(x^*)}\right|$

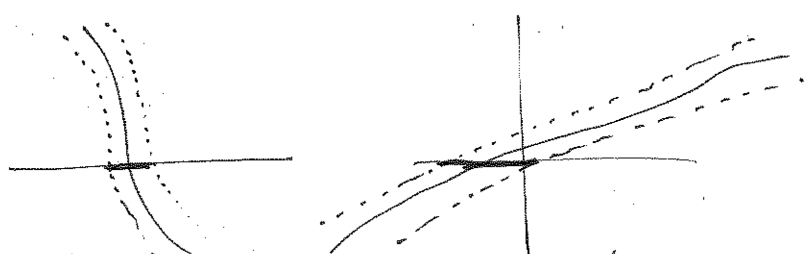
$$|\hat{x} - x^*| \le \left|\frac{1}{f'(x^*)}\right| \varepsilon$$

can be large if $f'(x^*)$ is small

mult.-D

eval : $\| J_f(x^*) \|$

root : $\| J_f^{-1}(x^*) \|$



well-cond ill-cond.

residual $\|f(\hat{x})\|$

error $\|\hat{x} - x^*\|$

small resid $\Rightarrow$ small err only if cond # is small.

§5.4 Conv. Rates + Stopping Criteria

Iterative methods (vs. direct methods)

$$\text{cost} = \frac{\text{cost}}{\text{iter}} \cdot \#\text{iter}$$

Conv. rate

$e_k = x_k - x^*$ error at iter k.

conv w/ rate r if

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|^r} = C, \quad C > 0.$$

- $r = 1$, $C < 1$ $\Rightarrow$ linear — gain constant # of correct digits each iter $-(\log_\beta(C))$
- $r > 1$ $\Rightarrow$ superlinear
- $r = 2$ $\Rightarrow$ quadratic
- $r = 3$ $\Rightarrow$ cubic

} $r\times$ as many correct digits as in prev iter.

Ex. 5.6. Conv rates.

$10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, \ldots$ linear $C = 10^{-1}$

$10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}$ linear $C = 10^{-2}$

$10^{-2}, 10^{-3}, 10^{-5}, 10^{-8}, 10^{-12}$ superlinear, not quad.

$10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}$ quadratic.

Stopping criteria

don't know e_k — a bit difficult to give general

$\|x_{k+1} - x_k\| / \|x_k\|$

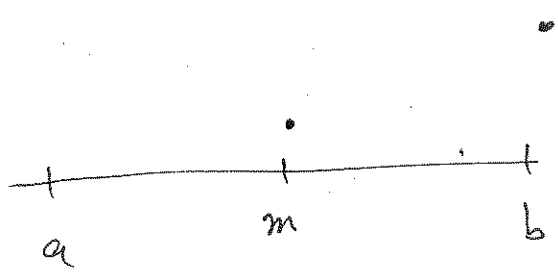
5.5 1D.

cont. function $f: \mathbb{R} \to \mathbb{R}$

find $x^* \in \mathbb{R}$ s.t. $f(x^*) = 0$.

Bisection Method

find short interval $[a, b]$ where $f(x)$ changes sign



while $((b-a) > tol)$ do

  midpoint $= m = a + \left(\frac{b-a}{2}\right)$

  if. sign$(f(a))$ = sign$(f(m))$

    $a = m$

  else

    $b = m$

  end

end

Numerical Pitfalls

~~midpoint = $m = \frac{a+b}{2}$~~

① $[0.67, 0.69]$

$$\begin{array}{r} 0.67 \\ 0.69 \\ \hline 1.36 \end{array}$$

$2 \overline{)1.4} = .7$ ✓

$\frac{.02}{2} = .01$

$0.67 + .01 = .68$

② $a+b$ could overflow

③ $f(a) \cdot f(m) > 0$

as $f \to 0$, this could underflow.

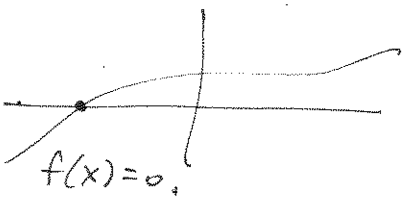
- guaranteed to converge
- slow convergence.
  - linear convergence $(r = 1)$
  - $C = .5$ gain one bit of each iter
- length at iter k

$$\frac{b-a}{2^k}$$

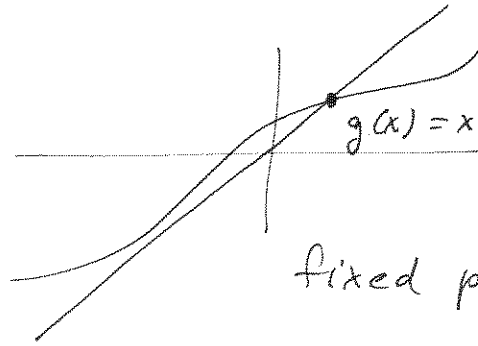
$$\Rightarrow \frac{b-a}{2^k} < tol \Rightarrow \frac{b-a}{tol} < 2^k \Rightarrow \boxed{\log_2\left(\frac{b-a}{tol}\right) < k}$$

§ 5.5.2 Fixed Pt. Iteration

$$x = g(x) \qquad x \text{ is a "fixed pt"}$$



root



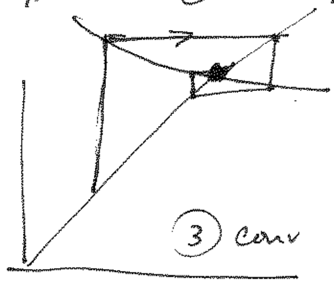
fixed point.

many choices of $g(x)$ for $f(x)=0$.

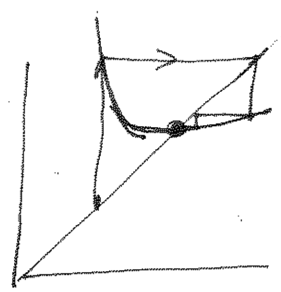
Example 5.8

$$f(x) = x^2 - x - 2 = 0 \qquad \text{sol'n: } x^* = 2,\ x^* = -1$$

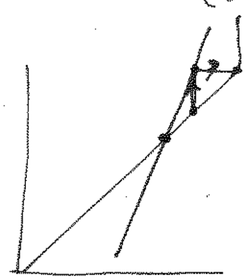
1. $g(x) = x^2 - 2$
2. $g(x) = \sqrt{x+2}$
3. $g(x) = 1 + 2/x$
4. $g(x) = \dfrac{(x^2+2)}{(2x-1)}$



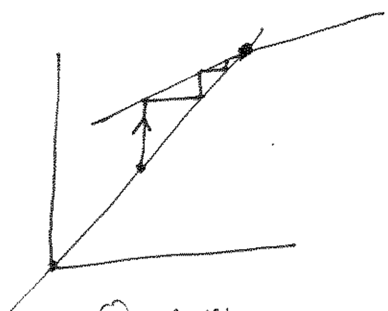
3 conv



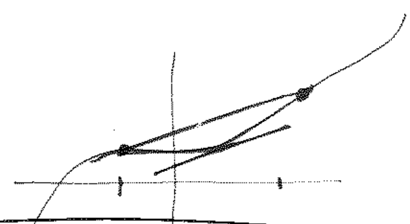
4 conv.



1 diverges



2 . conv



start here ⊗ locally convergent if $|g'(x^*)| < 1$

$$e_{k+1} = x_{k+1} - x^* = g(x_k) - g(x^*)$$

By Mean Value Theorem, $\exists$ ~~~~ θ ~~~~ between x_k & x^* s.t.

$$g'(\theta) = \frac{g(x_k) - g(x^*)}{x_k - x^*} \Rightarrow e_{k+1} = g'(\theta)\, e_k$$

$$\Rightarrow |e_{k}| \le C^k |e_0|, \quad C < 1$$

$$\Rightarrow |e_k| \to 0$$

Fixed Pt. Method Local Convergence.

$$e_{k+1} = x_{k+1} - x^* = g(x_k) - x^* = g(x_k) - g(x^*)$$

$$e_k = x_k - x^*$$

$$\frac{e_{k+1}}{e_k} = \frac{g(x_k) - g(x^*)}{x_k - x^*}$$

$$g(x_k) = g(x^*) + g'(x^*)(x_k - x^*) + \frac{1}{2} g''(x^*)(x_k - x^*)^2 + \cdots$$

$$\Rightarrow g(x_k) - g(x^*) = g'(x^*)(x_k - x^*) + \frac{1}{2} g''(x^*)(x_k - x^*)^2 + \cdots$$

$$\frac{e_{k+1}}{e_k} = g'(x^*) \frac{\cancel{(x_k - x^*)}}{\cancel{(x_k - x^*)}} + \frac{1}{2} g''(x^*) \frac{(x_k - x^*)^{\cancel{2}}}{\cancel{(x_k - x^*)}} + O((x_k - x^*)^3)$$

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

if $g'(x^*) = 0$,

$$\frac{e_{k+1}}{e_k} = \frac{1}{2} g''(x^*)(x_k - x^*) + \frac{1}{3!} g'''(x^*)(x_k - x^*)^2 + \cdots$$

$$\Rightarrow \frac{e_{k+1}}{e_k^2} = \frac{1}{2} g''(x^*) + \frac{1}{3!} g'''(x^*)(x_k - x^*) + \cdots$$

$$\lim_{k \to \infty} \frac{e_{k+1}}{e_k^2} = \frac{1}{2} g''(x^*)$$ quadratic convergence.

when fixed pt. conv., asymptotic conv rate is linear with $C = |g'(x^*)|$

$$\boxed{\begin{aligned} &\text{Ideally have } g'(x^*) = 0 \\ &\Rightarrow \quad g(x_k) - g(x^*) = \frac{g''(\xi_k)}{2}(x_k - x^*)^2 \\ &\Rightarrow \quad \frac{g(x_k) - g(x^*)}{(x_k - x^*)^2} = \frac{g''(\xi_k)}{2} = \frac{e_{k+1}}{e_k^2} \\ &\qquad \lim_{k \to \infty} \frac{\|e_{k+1}\|}{\|e_k\|^2} = \frac{g''(x^*)}{2} \quad \text{quadratic convergence.} \end{aligned}}$$

① $g'(x) = 2x$

$g'(2) = 4 \Rightarrow$ diverges

② $g'(x) = \frac{1}{2}(x+2)^{-1/2}$

$g'(2) = \frac{1}{2}(4)^{-1/2} = \frac{1}{4} \Rightarrow$ converges $C = \frac{1}{4}$

+tive sign $\Rightarrow$ iteration approaches from one side.

③ $g'(x) = -2x^{-2}$

$g'(2) = -\frac{2}{2^2} = -\frac{1}{2} \Rightarrow$ converges $C = +\frac{1}{2}$

−ive sign $\Rightarrow$ spiral

④. $g'(x) = \frac{2x^2 - 2x - 4}{(2x-1)^2} \Rightarrow$ converges quadratically

$g'(2) = 0$

$$g(x_k) = g(x^*) + \frac{(x_k - x^*)}{1} g'(x^*) + \frac{(x_k - x^*)^2}{2} g''(x^*) + \cdots$$

Taylor's theorem

$$g(x_k) = g(x^*) + \cancel{(x_k - x^*) g'(x^*)}^{\,0} + \frac{(x_k - x^*)^2}{2} g''(\theta)$$

§ 5.5.3 Newton's Method

$$f(x+h) = f(x) + h f'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{3!} f'''(x) + \dots$$

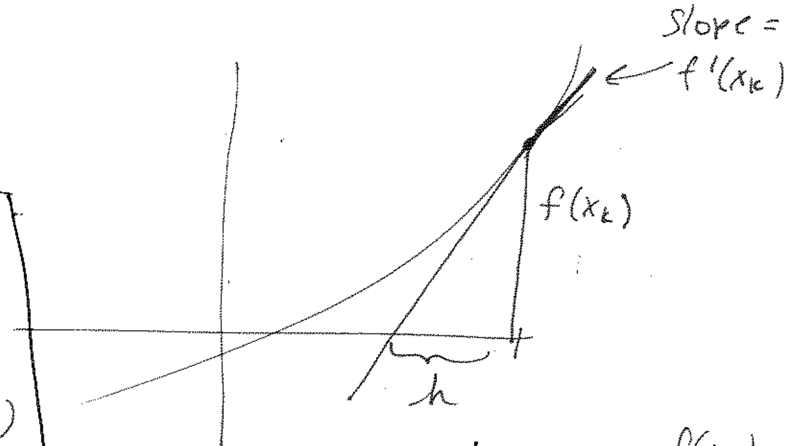
truncate

$$f(x+h) \cong f(x) + h f'(x)$$

$$\Rightarrow \quad h = \frac{f(x+h) - f(x)}{f'(x)}$$

want $f(x+h) = 0$

$$\Rightarrow \quad h = -\frac{f(x)}{f'(x)}$$



x_0 = initial guess
for $k = 1, 2, \dots$
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$
end

$$f'(x_k) = \frac{f(x_k)}{h}$$

transforms into a fixed pt iteration

$$g(x) = x - \frac{f(x)}{f'(x)}$$

$$= x - f(x)\left(f'(x)\right)^{-1}$$

$$g'(x) = 1 - f'(x)(f'(x))^{-1} + f(x)\, (f'(x))^{-2} f''(x)$$

$$= \cancel{1 - 1} + \frac{f(x) f''(x)}{f'(x)^2}$$

if x^* is a simple root $\begin{cases} f(x) = 0 \\ f'(x) \,!= 0 \end{cases}$ $\Rightarrow$ $g'(x) = 0$ $\Rightarrow$ quadratic convergence.

if x^* is a multiple root (multiplicity m) $\Rightarrow$ linear conv

$$C = 1 - \frac{1}{m}$$

5.5.4 Secant Method

Drawback of Newton's Method: need an expression for $f'(x)$

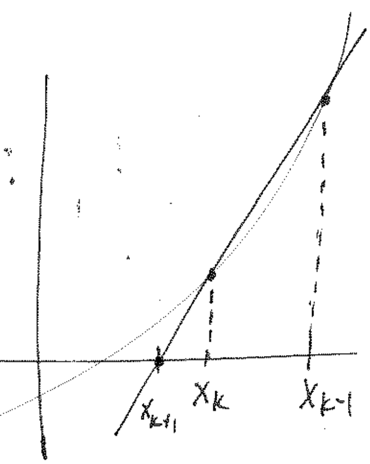
- may not be available
- may be expensive

use finite difference approximation instead.
use successive iterates – no need for extra function evaluations

$$f'(x_k) \cong \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

$\boxed{r \approx 1.618}$ convergence rate.

$$x_{k+1} = x_k - \frac{f(x_k)}{\left[\dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}\right]}$$



– need two starting guesses.
– a few more iterations are needed than Newton's method
+ but ½ function evaluations per iteration than Newton's method.

→ often better than Newton's method for total cost.

~~5.5.5~~ ~~5.5.6~~

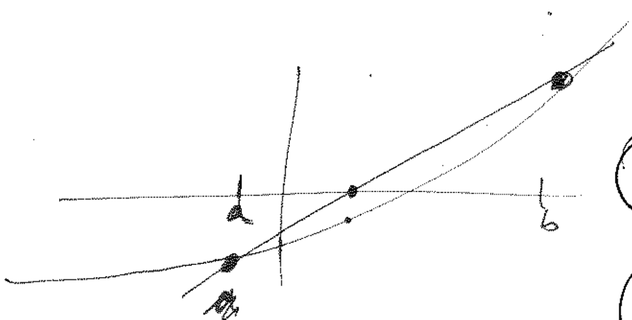
§5.5.7 Safeguarded Methods

- Newton's method, + secant method <u>unsafe</u>
  — need to start close to solution
- Bisection <u>safe</u> slow

<u>Hybrid</u>

- if fast method gives iterate <u>outside bracket</u> do one iteration w/ safe method.
- go back to fast method.

E.g., bisection for safety
secant for speed.



let $m = x_{k+1}$ for updating bracket

1. bracketing interval $[a, b]$
   $x_0 = a, \quad x_1 = b.$
2. Secant method x_{k+1}
3. if $x_{k+1} \in [a, b]$, $m = x_{k+1}$
   update bracket ~~bisection~~
4. if $x_{k+1} \notin [a, b]$, or $f(b) - f(a)$ too small to apply secant method, use <u>bisection</u>

only one evaluation
$f(x_k)$
per iteration

~~5.5.8~~ <u>Zeros of Polynomial</u>

- methods above + deflate $\frac{p(x)}{(x - x_1)}$
- companion matrix MATLAB
  reliable, less efficient
- all roots
  - Laguerre, Bairstow
  - Jenkins, Traub

§ 5.6. Systems of Nonlinear Equations $\boxed{\text{Lecture 9}}$

- wide range of behaviors is possible
  - theoretical analysis more complex
- not simple to bracket solution
- computationally expensive.

In multi-D, Jacobian plays role of f'

$$J(x) = \frac{\partial F}{\partial x}(x) \qquad J_{ij} = \frac{\partial F_i}{\partial x_j}$$

$$F(\vec{x}) = \begin{pmatrix} F_1(\vec{x}) \\ F_2(\vec{x}) \\ \vdots \\ F_m(\vec{x}) \end{pmatrix} = \begin{pmatrix} F_1(x_1, x_2, \ldots, x_n) \\ F_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ F_m(x_1, x_2, \ldots, x_n) \end{pmatrix}$$

e.g.) $F_1(x_1, x_2) = x_1^2 + \sin x_2 + 5$

$F_2(x_1, x_2) = x_1 + x_2^3$

$$\frac{\partial F_1}{\partial x_1} = 2x_1, \quad \frac{\partial F_1}{\partial x_2} = \cos x_2$$

$$\frac{\partial F_2}{\partial x_1} = 1, \quad \frac{\partial F_2}{\partial x_2} = 3x_2^2$$

$$\Rightarrow \boxed{J(x_1, x_2) = \begin{pmatrix} 2x_1 & \cos x_2 \\ 1 & 3x_2^2 \end{pmatrix}}$$

fixed pt. iter

local conv. cond $\quad |g'(x^*)| < 1$

analogous cond $\quad \rho(J(x^*)) < 1 \quad$ spectral radius.

(don't necessarily need to compute eigenvalues)

$\rho(A) \leq \|A\| \quad \forall$ vector-norm induced matrix norms

$J(x^*) = 0 \Rightarrow$ quadratic conv.

§ 5.6.2 Newton's Method

truncated Taylor series

$$f(x+s) \cong f(x) + J(x)s$$

Solve: $J(x)s = -f(x)$ for s.

$$x_{k+1} = x_k + s_k$$

Ex.

$$F(x) = \begin{pmatrix} x_1 + 2x_2 - 2 \\ x_1^2 + 4x_2^2 - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$x_0 = (1,2)^T$$

Notes: Must repeatedly solve linear systems!

- computing $J(x_k)$ may be expensive
  dense $\leadsto$ n^2 function evaluations
- Solving $J(x_k)s = -f(x_k)$ by LU $O(n^3)$ operations

"Quasi – Newton Methods"

cut some corners: e.g.,
- don't reevaluate J each iteration
- don't solve eq. exactly $Js = f$

Secant like Methods: Secant updating Methods:

Broyden's Method
- build up J incrementally
- update factorizations of J to save

$$B_{k+1} = B_k\left(I - \frac{s_k s_k^T}{s_k^T s_k}\right) + \frac{(f(x_{k+1}) - f(x_k))\, s_k^T}{s_k^T s_k}$$

---

damped Newton method $\alpha_k \le 1$