Solving Nonlinear Systems of Eqs.

Given $\underline{F}: \mathbb{R}^n \to \mathbb{R}^n$, solve $\underline{F}(\underline{x}) = 0$

Many, many applcn's in Scientific Comp.

<u>Ex</u> (1) Fully implicit time-discretizations of ODEs & PDEs. [vortex sheet models of flexible flags]

$$\frac{y^{n+1} - y^n}{\Delta t} = f(y^{n+1}, t_n); \text{ solve for } y^{n+1}$$

(2) Nonlinear elliptic problems

$$\nabla \cdot \left(\mu (|\nabla p|^2) \nabla p \right) = 0$$

$$p|_\Gamma = f$$

(Arises in modeling polymer melt extrusions and glacier movement)

(3) Optimization – finding critical points

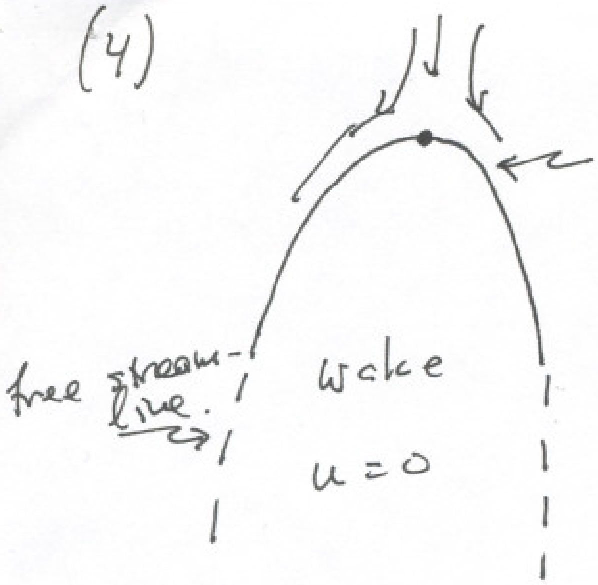
Find a body shape with minimal fluid drag.

$$\mathcal{D} = \left[\int_\Gamma \underline{\underline{\Sigma}} \cdot \hat{n} \, dS \right] \cdot \hat{x} \qquad Re(u \cdot \nabla u) = \nabla \cdot \underline{\underline{\Sigma}}$$

$$\nabla \cdot \underline{u} = 0$$

$$\underline{\underline{\Sigma}} = -p \underline{\underline{I}} + \mu (\nabla u + \nabla u^T)$$

(4)



flexible beam in a flow satisfying

$$\frac{\partial^2 \kappa}{\partial s^2} + \frac{1}{2}\kappa^3 = \eta^2 \cdot P$$

fluid pressure.

$\kappa|_{ends} = \kappa_s|_{ends} = 0$

Alben, Shelley, Zhang
Nature 2002.

$\eta = 0$; rigid

---

(5) Nonlinear least squares or data-fitting.

(6)

Local Methods

Given x_0 suff'ly close to a solution $\bar{x}$, local methods will provide a solution

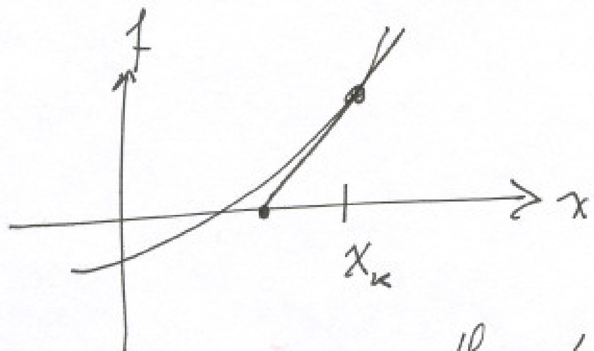
Exs Newton, Secant, Broyden Methods

Global Methods

Attempt to find a solution, even if initial guess is not close to a solution.

Exs Bisection, Steepest Descent, Line Search Methods.

Premier Method – Newton's Method – is local.



Approximate f w. ftn that is easy to solve:

$$M(x; x_k) = f(x_k) + f'(x_k)(x - x_k)$$

the tangent line approx. is an affine ~~approx.~~ ftn satisfying 2 interpolation conditions.

$$M(x_k; x_k) = f(x_k)$$
$$M'(x_k; x_k) = f'(x_k)$$

Determine x_{k+1} thru condition

$$M(x_{k+1}; x_k) = 0 \Rightarrow f'(x_k)(x_{k+1} - x_k) = -f(x_k)$$
$$\Rightarrow x_{k+1} = x_k - f(x_k)/f'(x_k)$$

Accuracy of the Affine approximation:

Let f' be Lipschitz continuous, i.e. $\exists\, \gamma$ s.t. $|f'(x) - f'(y)| \le \gamma |x-y| \quad \forall\, x,y$ in some set.

$$f(x+p) - M(x;p) = f(x+p) - [f(x) + f'(x)\cdot p]$$

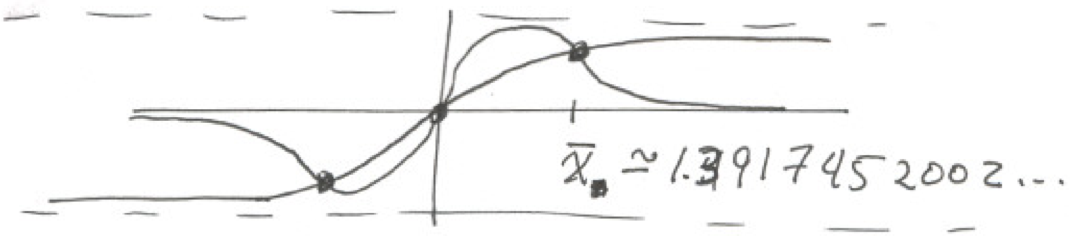
$$= \int_0^1 dt\, f'(x+tp)p - f'(x)p$$

$$= \int_0^1 dt\, [f'(x+tp) - f'(x)]p$$

$$|f(x+p) - M(x+p)| \le \int_0^1 |f'(x+tp) - f'(x)|\cdot|p|\, dt$$

$$\le \gamma \int_0^1 |tp|\cdot|p| = \frac{\gamma}{2}|p|^2 \quad // \text{ 2nd order}$$

Ex Solve $f(x) = \arctan x - \frac{2x}{1+x^2} = 0$



3 solutions

k	x_k	# digits
0	$\underline{1}.0$	1
1	$\underline{1.4}29\ldots$	2
2	$\underline{1.3916}1\ldots$	5
3	$\underline{1.39174519}$	9
4	$1.$ ______	16

N.M. doubles the # of correct digits at each step, once x_k is "close enough" to $\bar{x}$.
Hallmark of quadratic convergence.

Ad hoc fixed point iteration:

$$x = \frac{1}{2}(1+x^2)\arctan x = g(x)$$

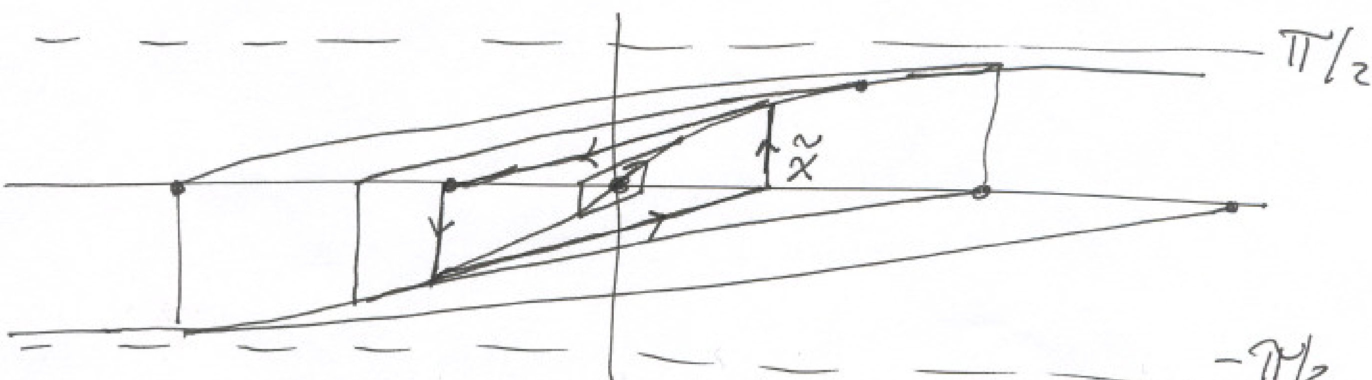
$$x_{k+1} = g(x_k)$$

Properties $|x_0| < \bar{x}$, $x_k \to 0$

$|x_0| > \bar{x}$, $|x_k| \to \infty$

Newton's Method is guaranteed to converge to $\bar{x}$ if x_0 is suff'ly close, that is, guaranteed <u>local convergence</u>. No guarantees ~~on local~~ when x_0 is far from $\bar{x}$.

<u>Another example</u> $f(x) = \arctan x$, only <u>one</u> solution, $x = 0$.



<u>Note:</u> The Newton step always points in the direction of decreasing f.

$$x_{k+1} = x_k - (1 + x_k^2)\, \text{Arctan}\, x_k$$

Condition for 2-cycle (in this case; generally $x_{k+2} = x_k$): $\boxed{-x = x - (1+x^2)\,\text{atan}\, x}$

3 solutions: $\pm\tilde{x}, 0$.

If $\arctan|x_0| < \dfrac{2|x_0|}{1+|x_0|^2}$; $x_k \to 0$

$=$ " ; 2-cycle

$>$ " ; $|x_k| \to \infty$.

Newton's Method in $\mathbb{R}^n$

Given $F : \mathbb{R}^n \to \mathbb{R}^n$, F sufficiently continuous, find $\bar{x} \in \mathbb{R}^n$ s.t. $F(\bar{x}) = 0$.

In $\mathbb{R}^n$ use the affine approximation

$$\underline{M}(x_k) = \underline{F}(x_k) + \underline{\underline{J}}_F(x_k)(\underline{x} - \underline{x}_k)$$

w. $[J_F]_{ij} = \dfrac{\partial F_i}{\partial x_j}$

Again $M(x; x_k)$ satisfies $\begin{cases} M(x_k; x_k) = F(x_k) \\ J_M(x_k; x_k) = J_F(x_k) \end{cases}$

and if J_F is Lipschitz continuous will satisfy

$$\| f(x) - f(y) - J_F(y)(x-y) \| \le \frac{\gamma}{2} \|x-y\|^2$$

<u>Newton's Method:</u> $M(x_{k+1}; x_k) = 0$

$$J_F(x_k)(x_{k+1} - x_k) = -F(x_k) \qquad (1)$$

or

$$\underline{\underline{J}}_F(x_k)\, s_k = -F(x_k) \qquad (2)$$

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

Now a matrix inversion problem in $\mathbb{R}^n$.

Example Find the fixed points of the

Lorenz Eqs. (E.N. Lorenz, 1963, Deterministic nonperiodic flow; J. Atmos Sci.)

$$\dot{x} = \sigma(y - x)$$

$$\dot{y} = \rho x - y - xz$$

$$\dot{z} = -\beta z + xy$$

Fixed points satisfy $F(\bar{x},\bar{y},\bar{z}) = \begin{pmatrix} \sigma(\bar{y}-\bar{x}) \\ \rho\bar{x} - \bar{y} - \bar{x}\bar{z} \\ -\beta\bar{z} + \bar{x}\bar{y} \end{pmatrix} = \underset{\sim}{0}$

~~Fixed~~

$$J_F = \begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - z & -1 & -x \\ y & x & -\beta \end{pmatrix}$$

N.M.:

$$\begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - z_k & -1 & -x_k \\ y_k & x_k & -\beta \end{pmatrix} \begin{pmatrix} x_{k+1} - x_k \\ y_{k+1} - y_k \\ z_{k+1} - z_k \end{pmatrix} = -\begin{pmatrix} \sigma(y_k - x_k) \\ \rho x_k - y_k - x_k z_k \\ -\beta z_k + x_k y_k \end{pmatrix}$$

Define $N(\underline{x}, \underline{r}) = \{ y \in \mathbb{R}^n : \|y - x\| < r \}$ w. $\|\cdot\|$ some norm on $\mathbb{R}^n$ w. induced matrix norm

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Thm (Local Convergence of N.M. in $\mathbb{R}^n$)

Let $F: \mathbb{R}^n \to \mathbb{R}^n$ be cont'ly diff'ble in an open convex set $D \subseteq \mathbb{R}^n$. Assume $\exists\ \bar{x} \in \mathbb{R}^n$ and $r, \beta > 0$ s.t.

(a) $N(\bar{x}, r) \subseteq D$ w. $F(\bar{x}) = 0$ $\left[\dfrac{\text{Solution inside } D}{\text{Isolated, simple zero.}}\right]$

(b) $J_F(\bar{x})^{-1}$ exists w. $\|J(\bar{x})\|^{-1} \leq \beta$ for J_F Lipschitz continuous in $N(\bar{x}, r)$ w. L. constant γ.

Then $\exists\ \varepsilon > 0$ s.t. $\forall\ x_0 \in N(\bar{x}, \varepsilon)$

(i) $x_k \to \bar{x}$ w. $x_{k+1} = x_k - J_F(x_k)^{-1} F(x_k)$ $\quad k = 0, 1, \ldots$

(Must show $J_F(x)^{-1}$ exists in $N(\bar{x}, \varepsilon)$ & is well-defined)

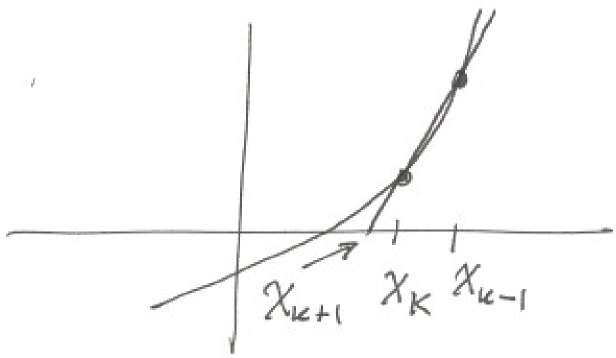
(ii) $\|x_{k+1} - \bar{x}\| \leq \beta \gamma \|x_k - \bar{x}\|^2$; quadratic convergence.

Newton-like methods

Jacobians are often difficult to calculate.

F could be a black box, extremely complicated, etc.

Good method for $n=1$ is the secant method



Replace $f'(x_k)$ w. $\dfrac{f(x_k)-f(x_{k-1})}{x_k - x_{k-1}} = D(x_k, x_{k-1})$

which is a good approx. ~~for~~ near $\bar{x}$.

Requires no add'l ftn evaluations.

Secant Method

$$x_k = x_k - f(x_k)/D(x_k, x_{k-1})$$

Thm Under the same conditions (or nearly so), the secant method will converge, with

$$|x_{k+1} - \bar{x}| \le C_k |x_k - \bar{x}|, \text{ w. } C_k \to 0 \text{ as } k \to \infty.$$

Called superlinear convergence

$$f(x) = \arctan x - \frac{2x}{1+x^2} \qquad x_0 = 1$$

k	N.M.	S.M
		1
0	1	1
1	2	1
2	5	2
3	9	4
4	16	6
5	—	11
6	—	16

Still very rapid convergence.

Generalizing to $\mathbb{R}^n$

To approximate $\underline{\underline{J}}(\vec{x}_k) = \left[\frac{\partial f_i}{\partial x_j}\right]\Big|_{\vec{x} = \vec{x}_k} \in \mathbb{R}^{n\times n}$

we need to approximate n^2 derivatives. A single previous iterate is apparently insufficient.

Could use finite-difference methods, i.e.

$$x_{k+1} = x_k - A_k^{-1} F(x_k)$$

$$\text{w.} \quad [A_k]_{\cdot j} = \frac{F(x_k + h_k \hat{e}_j) - F(x_k)}{h_k}$$

(j^{th} column)

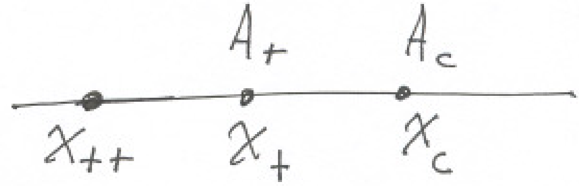
$$j = 1(1)n$$

Requires n add'l evaluations of vector fn F. Could be quite expensive. Still there is a local convergence theory for such methods.

Is there a secant method in $\mathbb{R}^n$? req'ing no new fn evaluations?

Back to affine approximations –

$f: \mathbb{R} \to \mathbb{R}$ $(n=1)$



$\boxed{\text{Given } x_c \text{ \& } x_+}$

the affine model of f was

$$m_+(x) = f(x_+) + a_+(x - x_+)$$

which satisfies $\boxed{m_+(x_+) = f(x_+)}$ indep'ly of a_+.

N.M. Req. $m_+'(x_+) = f'(x_+) = a_+$

S.M. Req. $m_+(x_c) = f(x_c)$; Secant eqn.

$$\Rightarrow \quad a_+ = \frac{f(x_+) - f(x_c)}{x_+ - x_c}$$

Next iterate: Solve $m_+(x_{++}) = 0$

$$\Rightarrow \quad x_{++} = x_+ - f(x_+)/a_+$$

In $\mathbb{R}^n$: $\quad M_+(x) = F(x_+) + A_+(x - x_+)$

$A_+ \in \mathbb{R}^{n\times n}$

$M_+(x_+) = F(x_+)$

Req'ing that $M_+(x_c) = F(x_c)$

$$\Rightarrow \boxed{A_+(x_+ - x_c) = F(x_+) - F(x_c)} \quad (*)$$

Called the secant equation, and it incompletely specifies A_+ if $n > 1$.

$$s_c = x_+ - x_c, \quad y_c = F(x_+) - F(x_c)$$

$$\Rightarrow \boxed{A_+ s_c = y_c}; \quad n \text{ eqns for } n^2 \text{ unknowns.}$$

A Possible solns

- Determine A_+ by interpolation of affine model at $(n-1)$ previous iterates:

$$M_+(x_{-i}) = F(x_{-i}) \Leftrightarrow A_+(x_+ - x_{-i}) = F(x_+) - F(x_{-i})$$

$$i = 1(1)(n-1)$$

Badly conditioned & expensive.

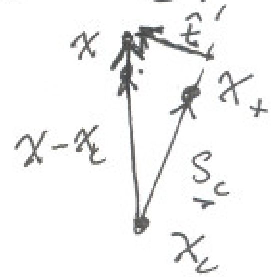
Instead, given A_c, choose A_+ to minimize the change in the affine model, subject to the secant equation.

$$M_+(x) - M_c(x) = \left[\underline{F(x_+)} + A_+(x - \underline{x_+})\right] - \left[F(x_c) + A_c(x - x_c)\right]$$

$$= \left[F(x_+) - F(x_c) - A_+(x_+ - x_c)\right] \Big\} = 0 \text{ by Sec. Eq. } (*)$$

$$+ A_+(x_+ - x_c) + A_+(x - x_+) - A_c(x - x_c)$$

$$= (A_+ - A_c)(x - x_c) \checkmark$$



For any $x \in \mathbb{R}^n$, $\exists$ α, t s.t. ($\alpha \in \mathbb{R}$)

$$x - x_c = \alpha s_c + t \quad \text{w.} \quad t \cdot s_c = 0$$

$$\Rightarrow M_+(x) - M_c(x) = \alpha (A_+ - A_c) s_c + \underbrace{(A_+ - A_c) t}_{\text{think of this as an error term}} \quad (**)$$

<u>Condition</u> Choose A_+ s.t. $(A_+ - A_c) t = 0$

$\forall t$ s.t. $\hat{t} \cdot s_c = 0$. Since t lies in an $(n-1)$ dim'l subspace, this gives $(n-1)$ add'l eqns.

$(A_+ - A_c)$ has an $(n-1)$ dim'l null space

$\Rightarrow$ $A_+ - A_c$ is a rank-one matrix of the form $\boxed{A_+ - A_c = u s_c^T, \quad u \in \mathbb{R}^n.}$ $(***)$

Return to the secant equation (*), rewritten as:

$$(A_+ - A_c)s_c = y_c - A_c s_c$$

$$\Rightarrow \quad u\, s_c^T s_c = y_c - A_c s_c$$

$$\Rightarrow \quad u = \frac{1}{s_c^T s_c}(y_c - A_c s_c)$$

$$\Rightarrow \quad A_+ = A_c + \underbrace{\frac{1}{s_c s_c^T}(y_c - A_c s_c) s_c^T}_{\text{A rank-one update of } A_c}$$

Broyden's Method

Given $F: \mathbb{R}^n \to \mathbb{R}^n$; $x_0 \in \mathbb{R}^n$, $A_0 \in \mathbb{R}^{n \times n}$

Solve: $A_k s_k = -F(x_k)$ for s_k

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

$$y_k = F(x_{k+1}) - F(x_k)$$

$$A_{k+1} = A_k + \frac{1}{s_k^T s_k}(y_k - A_k s_k) s_k^T$$

Need A_0. Use $J_F(x_0)$ or a finite-diff. approx.

Uses 1 function evaluation/step.

<u>Thm</u> Let all the assumptions of the previous theorem hold (local convergence of N.M). Then $\exists\ \varepsilon, \delta > 0$ s.t. if $\|x_0 - \bar{x}\| \leq \varepsilon$, $\|A_0 - J(\bar{x})\| \leq \delta$, then $\{x_k\}$ is well defined and converges <u>superlinearly</u> to $\bar{x}$.

~~Example~~

Example

$$F(x) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix}, \quad \text{solns } \begin{bmatrix} 0 \\ 3 \end{bmatrix} \text{ \& } \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$J = \begin{bmatrix} 1 & 1 \\ 2x_1 & 2x_2 \end{bmatrix}$$

Let $A_0 = J(x_0)$, $x_0 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$

$x_k \to \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, show only second component

k	B.M.	~~N.M~~	N.M	
0	5	0	5	0
1	3.6	1	3.6	1
2	3.07	2	3.09	2
3	3.01	3	3.002	4
4	3.0003	4	3.000002	7
5	3.000001	7	3.000000000002	13
6	3.0000000001	11	3.0	all
7	3.0 to 14 digits	14		

Note: $J(\bar{x}) = \begin{bmatrix} 1 & 1 \\ 0 & 6 \end{bmatrix}$, $\lim_{k \to \infty} A_k \begin{bmatrix} 1 & 1 \\ 1.5 & 7.5 \end{bmatrix}$

~~Even when~~

Convergence of A_k to J is not nec. for convergence, only convergence of the step direction.

Further details, see Dennis & Schnabel 'Numerical Methods for Unconstrained Optimization and Nonlinear equations.