Newton's Method:
The Basic Algorithm

univariate case

a simple problem:

Given a function

$$z \in \mathcal{R} \longrightarrow f(z) \in \mathcal{R} ,$$

we would like to find a REAL root (or roots) Z:

$$\boxed{f(Z) = 0 .}$$

Note z vs. Z and $f(z)$ vs. $f(Z) = 0$.

the fundamental idea: improvement, or update

Given $\tilde{z}$ (near Z),

approximate $f(z)$ near $\tilde{z}$ as

$$f(z) \approx \tilde{f}(z) \equiv \underbrace{f(\tilde{z}) + f'(\tilde{z})(z - \tilde{z})}_{\text{Taylor series about } \tilde{z}} ;$$

find root of $\tilde{f}(z)$:

$$\tilde{f}(\tilde{\tilde{z}}) = 0 \Rightarrow 0 = f(\tilde{z}) + f'(\tilde{z})(\tilde{\tilde{z}} - \tilde{z})$$

$$\Rightarrow \tilde{\tilde{z}} = \tilde{z} - f(\tilde{z})/f'(\tilde{z}) .$$

the conceit: $|\tilde{\tilde{z}} - Z| < |\tilde{z} - Z|$;

linear equation(s) easy to solve (efficiently).

iteration: an example

Consider

$$Z^2 + 2Z = 3 \qquad (Z = -3, Z = 1)$$

$$\Downarrow$$

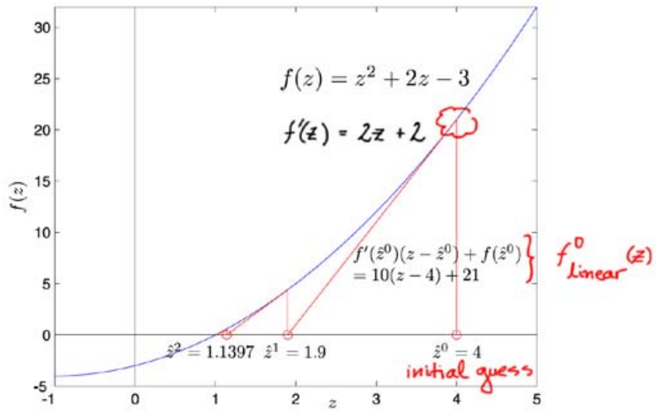
$$Z^2 + 2Z - 3 = 0$$

$$\Downarrow$$

$$\boxed{\begin{array}{c} f(Z) = 0 \\ f(z) = z^2 + 2z - 3 \end{array}}$$

Note: $f'(z) = 2z + 2$.

1st iteration :

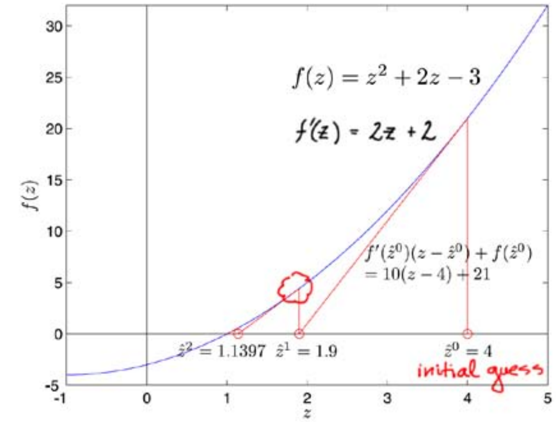


$$f^0_{\text{linear}}(z) \equiv f'(\hat z^0)(z - \hat z^0) + f(\hat z^0) = 10(z-4) + 21$$

$$f^0_{\text{linear}}(\hat z^1) \equiv 10(\hat z^1 - 4) + 21 = 0 \quad \Longrightarrow \quad \hat z^1 = 1.9 = 4 - 21/10$$

$$f'(\hat z^0) \qquad \hat z^1 \qquad \hat z^0 \qquad f(\hat z^0) \qquad \hat z^1 = \hat z^0 - f(\hat z^0)/f'(\hat z^0)$$

2nd iteration :



$$f^1_{\text{linear}}(z) \equiv f'(\hat z^1)(z - \hat z^1) + f(\hat z^1) = 5.8(z - 1.9) + 4.41$$

$$f^1_{\text{linear}}(\hat z^2) = 0$$

$$\hat z^2 = 1.1397 \quad \ldots$$

the algorithm :

Algorithm 1 Newton algorithm with storage of intermediate approximations

$k \leftarrow 0$
while $\left|f(\hat z^k)\right| > \text{tol}$
  $\hat z^{k+1} \leftarrow \hat z^k - \frac{f(\hat z^k)}{f'(\hat z^k)}$
  $k \leftarrow k+1$
end
$Z \leftarrow \hat z^k$ $Z \approx \hat z^k$

Algorithm 2 Newton algorithm without storage   and increment form

$\hat z \leftarrow \hat z^0$
while $\left|f(\hat z)\right| > \text{tol}$
  $\delta \hat z \leftarrow \frac{-f(\hat z)}{f'(\hat z)}$
  $\hat z \leftarrow \hat z + \delta \hat z$ current iterate } one iteration
end
$Z \leftarrow \hat z$

stopping criterion : at termination

If $f(z)$ is smooth about Z, $f'(Z) \neq 0$ "0"

$$f(\hat z) = f(Z + (\hat z - Z))$$

$$\approx \underbrace{f(Z)}_{0} + f'(Z)(\hat z - Z) + \underbrace{\cdots}_{\text{h.o.t.}}$$

$$\Downarrow$$

$$\hat z - Z = f(\hat z)/f'(Z)$$

$$\Rightarrow |\hat z - Z| \leq \text{tol}/|f'(Z)| \quad \text{sensitivity}$$

$$\Rightarrow \text{choose tol} = \text{tol}_Z \cdot \underbrace{|f'(Z)|}_{\text{estimate, say } f'(\hat z)}$$

Hence $|f(\hat z)|$ ($\leq$ tol) is a natural error estimator.

2

Convergence rate:

Define error in k^{th} iterate as after k iterations

$$\epsilon^k = \hat{z}^k - Z\ ;$$

then, if

(*i*) $f(z)$ is smooth (e.g., the second derivative exists), ★★

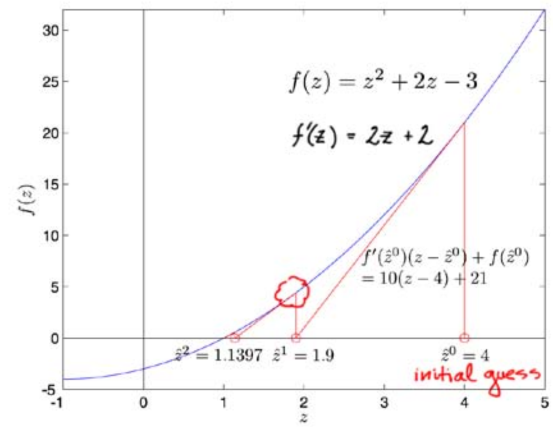
"O" (*ii*) $f'(Z) \neq 0$ (i.e., the derivative at the root is nonzero), and ★

(*iii*) $|\epsilon^0|$ (the error of our initial guess) is sufficiently small, ★★★

we achieve quadratic convergence:

$$\boxed{\epsilon^{k+1} \sim (\epsilon^k)^2 \left(\frac{1}{2}\frac{f''(Z)}{f'(Z)}\right)\ .}$$

evidence: simple example



k	$\hat{z}^k$	$\hat{z}^k - Z$ $(Z = 1)$
iteration	approximation	number of correct digits
0	4	0
1	1.9	1
2	1.1...	1
3	1.004...	3
4	1.000005...	6
5	1.000000000006...	12

(Note here $\frac{1}{2} f''(Z)/f'(Z) = 1/4$.)

> IF Newton's method converges, $\hat{z}$ converges very rapidly to a root Z.

proof: Recall

$$\epsilon^k \equiv \underbrace{\hat{z}^k}_{k^{th}\text{ iterate}} - \underbrace{Z}_{\text{exact solution}}\ ;$$

then

$$\hat{z}^{k+1} = \hat{z}^k - \frac{f(\hat{z}^k)}{f'(\hat{z}^k)} \qquad \text{Newton update}$$

$$\cancel{Z} + \epsilon^{k+1} = \cancel{Z} + \epsilon^k - \frac{f(Z + \epsilon^k)}{f'(Z + \epsilon^k)}$$

Taylor series

$$\epsilon^{k+1} = \epsilon^k - \frac{\overset{\text{root}}{\cancel{f(Z)}} + \epsilon^k f'(Z) + \frac{1}{2}(\epsilon^k)^2 \overset{\text{smoothness}}{f''(Z)} + \cdots}{f'(Z) + \epsilon^k f''(Z) + \cdots}$$

"O": $|f'(Z)| \leq |\epsilon f''(Z)|$

$$\epsilon^{k+1} = \epsilon^k - \epsilon^k \frac{f'(Z) + \frac{1}{2}\epsilon^k f''(Z) + \cdots}{f'(Z)(1 + \epsilon^k \frac{f''(Z)}{f'(Z)} + \cdots)} \quad \text{if } f'(Z) \neq 0$$

h.o.t. for SMALL ϵ^k

$$\epsilon^{k+1} = \epsilon^k - \epsilon^k \frac{\cancel{f'(Z)}(1 + \frac{1}{2}\epsilon^k \frac{f''(Z)}{f'(Z)} + \cdots}{\cancel{f'(Z)}(1 + \epsilon^k \frac{f''(Z)}{f'(Z)} + \cdots)}\ ;$$

now consider Taylor series of

$$g(y) \equiv \frac{1}{1+y}$$

about $y = 0$:

$$g(\underset{0}{y} + \delta) = g(y) + g'(y)\delta + \cdots$$

$$= \frac{1}{1+y} + \left(\frac{-1}{(1+y)^2}\right)\delta + \cdots$$

$$\underset{(y=0)}{=} 1 - \delta + \cdots$$ h.o.t. for δ small

hence

$$\frac{1}{1 + \underbrace{\epsilon^k \frac{f''(Z)}{f'(Z)}}_{\delta}} = \underline{1 - \epsilon^k \frac{f''(Z)}{f'(Z)}} + \cdots$$

thus

$$\epsilon^{k+1} = \epsilon^k - \epsilon^k \underbrace{\left(1 + \frac{1}{2}\epsilon^k \frac{f''(Z)}{f'(Z)} + \cdots\right)\left(1 - \epsilon^k \frac{f''(Z)}{f'(Z)} + \cdots\right)} + \cdots$$

$$-\epsilon^k\left(1 + \frac{1}{2}\epsilon^k \frac{f''}{f'} - \epsilon^k \frac{f''}{f'} + O((\epsilon^k)^2)\right) = -\epsilon^k + \frac{1}{2}(\epsilon^k)^2 \frac{f''}{f'} + \cdots$$

and

$$\epsilon^{k+1} = \cancel{\epsilon^k} - \cancel{\epsilon^k} + \frac{1}{2}(\epsilon^k)^2 \frac{f''(Z)}{f'(Z)} + \cdots$$

$$\epsilon^{k+1} = \frac{1}{2}\frac{f''(Z)}{f'(Z)}(\epsilon^k)^2 + \cdots$$ QUADRATIC CONVERGENCE

Note if $\underline{f'(Z) = 0}$, then "0"

$$\epsilon^{k+1} = \frac{1}{2}\epsilon^k$$ — much slower convergence.

Newton's Method: The Basic Algorithm

multivariate case

example ($n=2$): 2 equations in 2 unknowns

$$\begin{cases} Z_1^2 + 2Z_2^2 = 22 \\ 2Z_1^2 + Z_2^2 = 17 \end{cases} \Rightarrow Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

Introduce $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$

$$\begin{pmatrix} f_1(z_1,z_2) \\ f_2(z_1,z_2) \end{pmatrix} \equiv \begin{pmatrix} z_1^2 + 2z_2^2 - 22 \\ 2z_1^2 + z_2^2 - 17 \end{pmatrix};$$

$$\begin{pmatrix} f_1(z) \\ f_2(z) \end{pmatrix}$$

$f(z)$

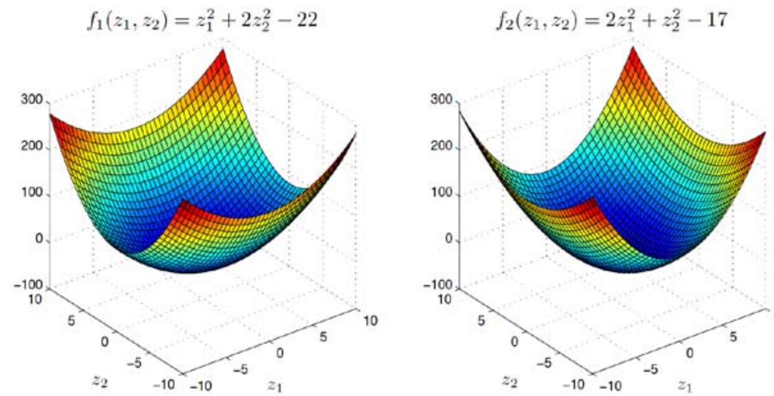
a vector of functions — f (vector) — z

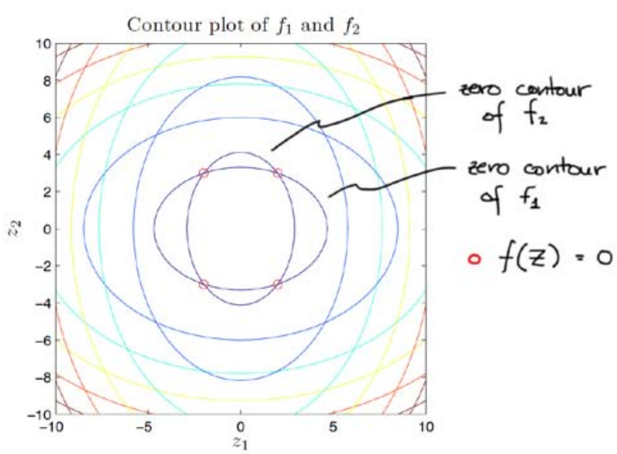
then

$$f(\mathbf{Z}) = 0 \qquad \text{2 equations in 2 unknowns}$$

$$f\left(\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}\right) = 0$$

$$\left.\begin{array}{lll} f_1\left(\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}\right) = 0 & \Rightarrow & Z_1^2 + 2Z_2^2 - 22 = 0 \\ f_2\left(\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}\right) = 0 & \Rightarrow & 2Z_2^2 + Z_1^2 - 17 = 0 \end{array}\right\} .$$





the general case:

Define

z : an n-vector, $(z_1, \ldots, z_n)^T$ $\qquad \mathbb{R}^n$

f : an n-vector of functions of z

$(f_1(z), f_2(z), \ldots, f_n(z))^T$;

then we wish to find a root Z such that

$$f(Z) = 0 \quad \Longleftrightarrow \quad \left.\begin{array}{l} f_1((Z_1, \ldots, Z_n)^T) = 0 \\ f_2((Z_1, \ldots, Z_n)^T) = 0 \\ \vdots \\ f_n((Z_1, \ldots, Z_n)^T) = 0 \end{array}\right\} .$$

the Newton update:

Recall for $n = 1$: to find $\hat{z}^{k+1}$

$$f^k_{\text{linear}}(z) \equiv f'(\hat{z}^k)(z - \hat{z}^k) + f(\hat{z}^k) \quad ;$$

$$f^k_{\text{linear}}(\hat{z}^{k+1}) = f'(\hat{z}^k)\underbrace{(\hat{z}^{k+1} - \hat{z}^k)}_{\delta z^k} + f(\hat{z}^k) = 0$$

$$\Downarrow$$

$$\boxed{\begin{aligned} \delta z^k &= -f(\hat{z}^k)/f'(\hat{z}^k) \\ \hat{z}^{k+1} &= \hat{z}^k + \delta z^k \quad . \end{aligned}}$$

Now for general n:

$$f^k_{i,\text{linear}}(z) \equiv \left.\frac{\partial f_i}{\partial z_1}\right|_{\hat{z}^k}(z_1 - \hat{z}^k_1) + \left.\frac{\partial f_i}{\partial z_2}\right|_{\hat{z}^k}(z_2 - \hat{z}^k_2) + \cdots + \left.\frac{\partial f_i}{\partial z_n}\right|_{\hat{z}^k}(z_n - \hat{z}^k_n) + f_i(\hat{z}^k)$$

for $1 \le i \le n$;

then

$$f^k_{1,\text{linear}}(\hat{z}^{k+1}) \equiv \left.\frac{\partial f_1}{\partial z_1}\right|_{\hat{z}^k}(\hat{z}^{k+1}_1 - \hat{z}^k_1) + \left.\frac{\partial f_1}{\partial z_2}\right|_{\hat{z}^k}(\hat{z}^{k+1}_2 - \hat{z}^k_2) + \cdots + \left.\frac{\partial f_1}{\partial z_n}\right|_{\hat{z}^k}(\hat{z}^{k+1}_n - \hat{z}^k_n) + f_1(\hat{z}^k) = 0\,,$$

$$f^k_{2,\text{linear}}(\hat{z}^{k+1}) \equiv \left.\frac{\partial f_2}{\partial z_1}\right|_{\hat{z}^k}(\hat{z}^{k+1}_1 - \hat{z}^k_1) + \left.\frac{\partial f_2}{\partial z_2}\right|_{\hat{z}^k}(\hat{z}^{k+1}_2 - \hat{z}^k_2) + \cdots + \left.\frac{\partial f_2}{\partial z_n}\right|_{\hat{z}^k}(\hat{z}^{k+1}_n - \hat{z}^k_n) + f_2(\hat{z}^k) = 0\,,$$

$$f^k_{n,\text{linear}}(\hat{z}^{k+1}) \equiv \left.\frac{\partial f_n}{\partial z_1}\right|_{\hat{z}^k}(\hat{z}^{k+1}_1 - \hat{z}^k_1) + \left.\frac{\partial f_n}{\partial z_2}\right|_{\hat{z}^k}(\hat{z}^{k+1}_2 - \hat{z}^k_2) + \cdots + \left.\frac{\partial f_n}{\partial z_n}\right|_{\hat{z}^k}(\hat{z}^{k+1}_n - \hat{z}^k_n) + f_n(\hat{z}^k) = 0 \quad ;$$

hence

$$f^k_{\text{linear}}(\hat{z}^{k+1}) \equiv \underbrace{\left.\begin{bmatrix} \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial z_2} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial z_2} & \cdots & \frac{\partial f_2}{\partial z_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial z_1} & \frac{\partial f_n}{\partial z_2} & \cdots & \frac{\partial f_n}{\partial z_n} \end{bmatrix}\right|_{\hat{z}^k}}_{J(\hat{z}^k)} \underbrace{\begin{bmatrix} (\hat{z}^{k+1}_1 - \hat{z}^k_1) \\ (\hat{z}^{k+1}_2 - \hat{z}^k_2) \\ \vdots \\ (\hat{z}^{k+1}_n - \hat{z}^k_n) \end{bmatrix}}_{\delta\hat{z}^k} + \underbrace{\begin{bmatrix} f_1(\hat{z}^k) \\ f_2(\hat{z}^k) \\ \vdots \\ f_n(\hat{z}^k) \end{bmatrix}}_{f(\hat{z}^k)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$\boxed{J(\hat{z}^k)\,\delta\hat{z}^k = -f(\hat{z}^k) \quad ;}$$

n equations in n unknowns
LINEAR

then

$$\boxed{\hat{z}^{k+1} = \hat{z}^k + \delta\hat{z}^k \quad .}$$

Note

$$J(z) \equiv \left.\begin{bmatrix} \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial z_2} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial z_2} & \cdots & \frac{\partial f_2}{\partial z_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial z_1} & \frac{\partial f_n}{\partial z_2} & \cdots & \frac{\partial f_n}{\partial z_n} \end{bmatrix}\right|_{z} ,$$

is the $n \times n$ Jacobian matrix:

univariate $\underbrace{\frac{\partial f_1}{\partial z_1}}_{f'} \longrightarrow$ multivariate $J_{ij} = \frac{\partial f_i}{\partial z_j}$, $1 \le i, j \le n$.

the algorithm

Algorithm 3 Multivariate Newton algorithm without storage

$\hat{\boldsymbol{z}} \leftarrow \hat{\boldsymbol{z}}^0$

while $\|\boldsymbol{f}(\hat{\boldsymbol{z}})\| > \mathrm{tol}$

$\quad \boldsymbol{J}(\hat{\boldsymbol{z}})\delta\hat{\boldsymbol{z}} = -\boldsymbol{f}(\hat{\boldsymbol{z}})$ {Solve the linearized system for $\delta\hat{\boldsymbol{z}}$.} J\(-f)

$\quad \hat{\boldsymbol{z}} \leftarrow \hat{\boldsymbol{z}} + \delta\hat{\boldsymbol{z}}$

end

$\boldsymbol{Z} \leftarrow \hat{\boldsymbol{z}}$ $\quad Z \approx \hat{z}$

MIT OpenCourseWare
http://ocw.mit.edu

2.086 Numerical Computation for Mechanical Engineers
Spring 2013

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.