Nonlinear Equations

Find the roots of

$$\left.\begin{array}{l} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \qquad\qquad\qquad\quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{array}\right\} \quad \underline{f}(\underline{x}) = \underline{0}$$

Example:

$$2\underbrace{x_1} - x_2^2 + x_3^4 - 1 = 0$$

$$x_1^2 + \underbrace{x_2} - x_3^3 = 0$$

$$\sin(x_1) + \tan(x_2) + \underbrace{x_3} - 1 = 0$$

Rewrite as

$$\underline{f}(\underline{x}) = \underline{0}$$

① If possible, write as a fixed point iteration

$$\underline{x} = \underline{g}(\underline{x})$$

Example: (continued from above)

$$\Rightarrow \quad x_1 = \frac{1}{2}x_2^2 - \frac{1}{2}x_3^4 - \frac{1}{2}$$

$$x_2 = -x_1^2 + x_3^3$$

$$x_3 = 1 - \sin(x_1) - \tan(x_2)$$

$$\underline{x} = \underline{g}(\underline{x})$$

$$\Rightarrow \quad \underline{x}_{i+1} = \underline{g}(\underline{x}_i)$$

Iterate until $\| \underline{x}_{i+1} - \underline{g}(\underline{x}_{i+1}) \|_p < \varepsilon$

Convergence will depend on character of $\underline{g}(\underline{x})$

and also on $\underline{x}_0$

② Newton-Raphson Method

Focus initially on just the first equation

$$f_1(x_1, x_2, \ldots, x_n) = 0$$

iteration, not power

Given $f_1^{(i)}, x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}$

Write

$$f_1^{(i+1)} \approx f_1^{(i)} + \frac{\partial f_1^{(i)}}{\partial x_1}\left(x_1^{(i+1)} - x_1^{(i)}\right) + \frac{\partial f_1^{(i)}}{\partial x_2}\left(x_2^{(i+1)} - x_2^{(i)}\right)$$

$$+ \ldots + \frac{\partial f_1^{(i)}}{\partial x_n}\left(x_n^{(i+1)} - x_n^{(i)}\right)$$

derivative $\frac{\partial f_1}{\partial x_n}$ evaluated at $x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}$

Since the goal is to have $f_1^{(i+1)} \stackrel{\sim}{=} 0 \Rightarrow$

Since the goal is to have $f_1^{(i+1)} \stackrel{!}{=} 0 \Rightarrow$

$$-f_1^{(i)} = \frac{\partial f_1^{(i)}}{\partial x_1}\left(x_1^{(i+1)} - x_1^{(i)}\right) + \dots + \frac{\partial f_1^{(i)}}{\partial x_n}\left(x_n^{(i+1)} - x_n^{(i)}\right)$$

Apply this idea to every function $f_1(x), f_2(x), \dots, f_n(x)$ in $\underline{f}(\underline{x})$. Then

$$\underbrace{\begin{bmatrix} \frac{\partial f_1^{(i)}}{\partial x_1} & \text{---} & \frac{\partial f_1^{(i)}}{\partial x_n} \\ | & & | \\ \frac{\partial f_n^{(i)}}{\partial x_1} & \text{---} & \frac{\partial f_n^{(i)}}{\partial x_n} \end{bmatrix}}_{\text{Jacobian matrix of } \underline{f}(\underline{x})} \begin{bmatrix} x_1^{(i+1)} - x_1^{(i)} \\ | \\ x_n^{(i+1)} - x_n^{(i)} \end{bmatrix} = \begin{bmatrix} -f_1^{(i)} \\ | \\ -f_n^{(i)} \end{bmatrix}$$

n eqns,
n unknowns

$$\underline{J}^{(i)} \underline{\delta}^{(i)} = -\underline{f}^{(i)}$$

with $\underline{J}^{(i)}$: Jacobian of $\underline{f}(\underline{x})$ evaluated using information at iteration i $(\underline{x}_i)$

iteration i $(\underset{\sim}{x}_i)$

$$\underline{\delta}^{(i)} = \underline{x}^{(i+1)} - \underline{x}^{(i)}$$

$$\underline{f}^{(i)} = \underline{f}(\underline{x}^{(i)}) \quad \text{residual}$$

Then,

$$\underline{\delta}^{(i)} = -\left(\underline{J}^{(i)}\right)^{-1} \underline{f}^{(i)}$$

$$\boxed{\underline{x}^{(i+1)} = \underline{x}^{(i)} - \left(\underline{J}^{(i)}\right)^{-1} \underline{f}^{(i)}}$$

Note: This is a set of linear algebraic equations

However, a solution is required at each iteration, because $\underline{J}$ varies between iterations

Example:

$$\left.\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1 \\ f_2(x_1, x_2) &= x_1^2 - x_2 \end{aligned}\right\} \quad \underline{f}(\underline{x}) = \underline{0}$$

$$\underline{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \cdots \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \\ \cdots \end{bmatrix}$$

$$\underset{\sim}{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{bmatrix}$$

Let $\underline{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then $\underset{\sim}{J}^{(0)} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$

Note: Here $\underset{\sim}{J}$ is symmetric, but need not be

matrix of constants

Note: If the initial guess $\underline{x}^{(0)}$ is not close to the solution $\underline{x}^*$, it might not converge (just as in the single equation Newton-Raphson method)

Also, sometimes a damped Newton method is needed, such as

$$\underline{\delta}^{(i)} = -\left(\underset{\sim}{J}^{(i)}\right)^{-1} \underline{f}^{(i)}$$

but then

$$\underline{x}^{(i+1)} = \underline{x}^{(i)} + \alpha^{(i)} \underline{\delta}^{(i)} \text{ with } \alpha^{(i)} \in (0, 1]$$

$$\underline{x}^{(i+1)} = \underline{x}^{(i)} + \alpha^{(i)}\underline{\delta}^{(i)} \text{ with } \alpha^{(i)} \in (0, 1]$$

that moves $\underline{x}^{(i+1)}$ closer to $\underline{x}^*$

Example:

$$\underline{f}(\underline{x}) = \begin{bmatrix} x_1^2 + x_1 x_2 + x_1 - 1 \\ x_1 x_2 + x_2 + x_3^2 - 0.25 \\ x_1^2 + x_2^2 - 4x_3 \end{bmatrix} = \underline{0}$$

Find roots

Want $\underline{f}(\underline{x}) = \underline{0}$

For N-R, need $\frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \frac{\partial f_1}{\partial x_3}$

$\vdots \quad \vdots \quad \vdots$

$\frac{\partial f_3}{\partial x_1}, \dots, \frac{\partial f_3}{\partial x_3}$

Fixed point method

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -x_1^2 - x_1 x_2 + 1 \\ -x_1 x_2 - x_3^2 + 0.25 \\ 0.25\,(x_1^2 + x_2^2) \end{bmatrix}}_{\underline{g}(\underline{x})}$$

Define

$$\varepsilon^i = \| \underline{f}(\underline{x}^i) \|_\infty$$

Try $\underline{x}^{(0)} = \begin{bmatrix} 1/2 \\ -1 \\ 0 \end{bmatrix}$

i	x_1	x_2	x_3	ε

1	1.25	0.75	0.3125	2.75
2	-1.5	-0.78	0.53	0.927
3	-2.42	-1.21	0.72	5.4
4	-7.83	-3.2	1.84	~77

$\downarrow$

∞

Diverges → No solution obtained

Try $\underline{x}^{(0)} = \begin{bmatrix} -1/2 \\ -1 \\ 0 \end{bmatrix}$

i	x_1	x_2	x_3	ε
1	0.25	-0.25	0.3125	1.125
2	1	0.21	0.03	1.214
3	-0.21	0.034	0.26	1.17
4	0.961	0.188	0.01	1.06

$\downarrow$

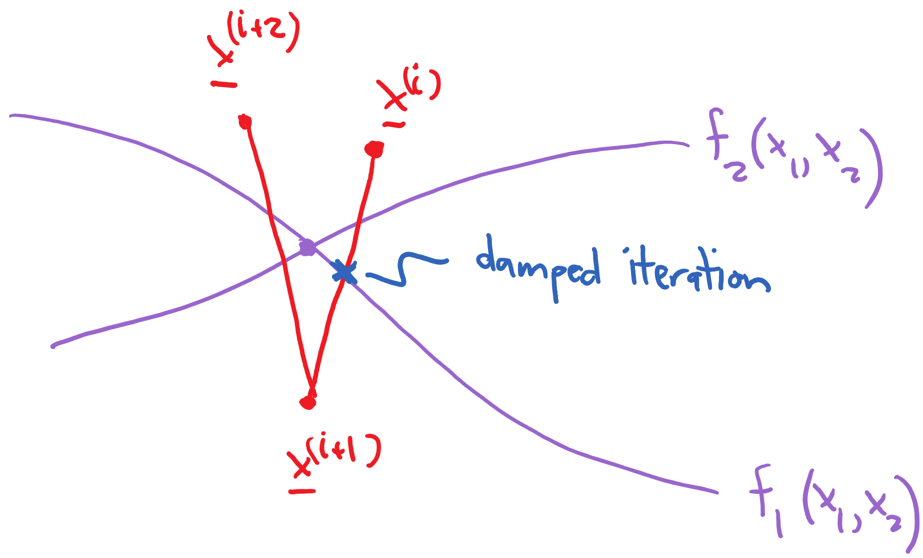
~ 1

Does not diverge or converge

Try a damped iteration $(\alpha = \frac{1}{2})$

$$\hat{\underline{x}}^{(i+1)} = \underline{g}(\underline{x}^{(i)}) \Rightarrow \underline{x}^{(i+1)} = \frac{1}{2}\underline{x}^{(i)} + \frac{1}{2}\hat{\underline{x}}^{(i+1)}$$

i	x_1	x_2	x_3	ε
10	0.577	0.513	0.0896	1.2×10^{-3}
27	$\sim$0.577	$\sim$0.513	$\sim$0.0896	1.3×10^{-8}



Now try Newton-Raphson

$$\underset{\sim}{J} = \begin{bmatrix} 1+2x_1+x_2 & x_1 & 0 \\ x_2 & 1+x_1 & 2x_3 \\ 2x_1 & 2x_2 & -4 \end{bmatrix}$$

Note: $\underset{\sim}{J}$ here is not symmetric

Try

$$\underset{\sim}{x}^{(0)} = \begin{bmatrix} -1/2 \\ -1 \\ 0 \end{bmatrix}$$

i	x_1	x_2	x_3	ε
1	-1.25	-1	0.5	0.563
4	-1.0465	-0.99	0.475	0.186
6	-1	-0.99	0.49	1.3×10^{-5}

[illegible]	-1.0465	-0.[illegible]	0.[illegible]15	0.106
6	-1	-0.99	0.49	1.3×10^{-5}
7	$\boxed{-1}$	$\boxed{-1}$	$\boxed{0.5}$	3×10^{-9}

Different root!

Try $\underline{x}^{(0)} = \begin{bmatrix} 1/2 \\ -1 \\ 0 \end{bmatrix}$

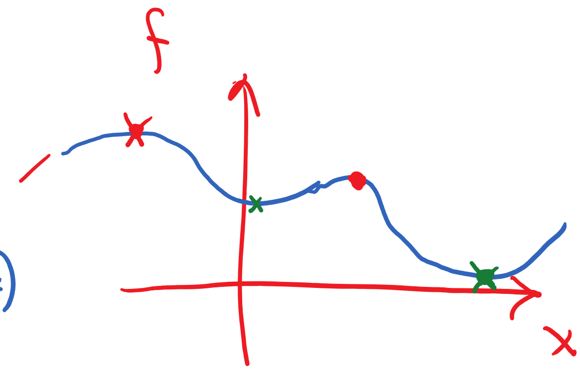
$i = 5$

$$\underline{x}^{(5)} = \begin{bmatrix} 0.5773 \\ 0.153 \\ 0.0893 \end{bmatrix}$$

Same root as with damped fixed point method

$$\varepsilon^{(5)} = 1.6 \times 10^{-11}$$

Minimization

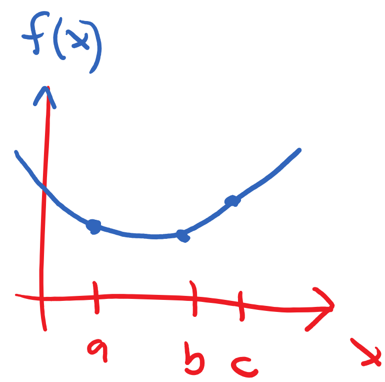


Closely related to root finding

Start w/ one variable functions $f(x)$

Find the minimum (or maximum) of $f(x)$

① Brent's Method



Let (a, b, c) be a triplet such that

$a < b < c$ and

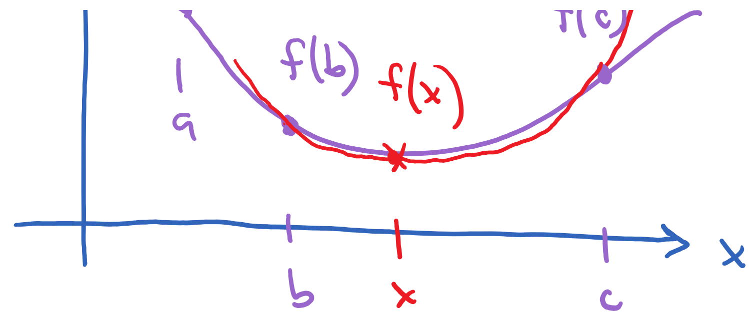
$f(a) > f(b)$ and $f(b) < f(c)$

$\Rightarrow$ A minimum must exist in $[a, c]$

Construct a 2^{nd} order, quadratic polynomial through

$(a, f(a)), (b, f(b)), (c, f(c)),$

then find where the derivative is zero



Let that minimum be at x

Choose (a, x, b) or (b, x, c) as appropriate

Choose this in above case

Repeat until convergence

Need to keep an eye on step size,

need to make sure that

$$f(a) > f(x) \text{ \& } f(x) < f(b) \text{ for example}$$

② Newton's Method for minimization

Let x_n be the current approximate solution

to the true minimum location x^*

Write Taylor's Series for $f'(x^*)$

$$f'(x_n+\Delta x) = f'(x_n) + \Delta x\, f''(x_n) + \frac{1}{2}\Delta x^2 f'''(x_n) + \mathcal{O}(\Delta x^3)$$

$\Rightarrow$ find Δx, such that $f'(x_n+\Delta x) = 0$

Need derivative to be zero for minimum

Then, for zero slope

$$0 = f'(x_n) + \Delta x\, f''(x_n) + \mathcal{O}(\Delta x^2)$$

$$\Rightarrow \Delta x \approx -\frac{f'(x_n)}{f''(x_n)}$$

Let

$$\Delta x = x_{n+1} - x_n = -\frac{f'(x_n)}{f''(x_n)}$$

$$\Rightarrow \underbrace{x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}}$$

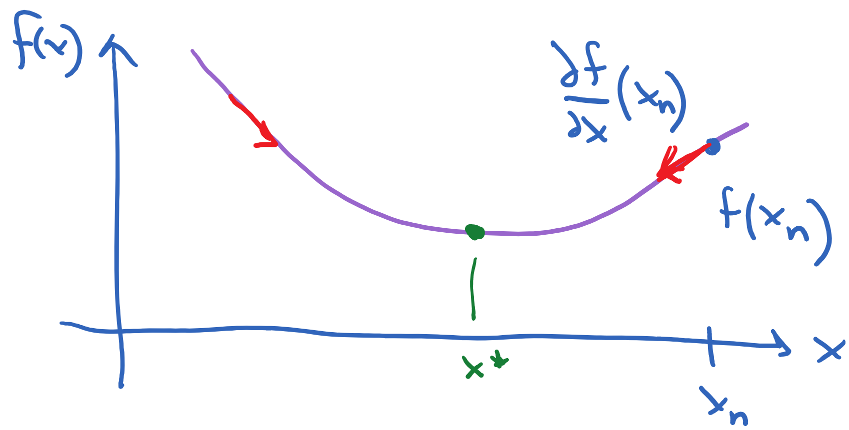
Newton iteration to find the root $f'(x) = 0$

(3) Steepest Gradient Descent

③ Steepest Gradient Descent

More useful if # of variables is > 1

Consider



Minimum is in the negative gradient direction

Construct an iteration that takes a step in $-\frac{\partial f}{\partial x}$, such that

$$x_{n+1} = x_n - \alpha_n \frac{\partial f(x_n)}{\partial x}$$

Where α_n is chosen every step to assure $|f(x_n - \alpha_n f'(x_n))| < |f(x_n)|$

—

These are called line-search methods,

as the minimization problem now

becomes about how to choose α_n