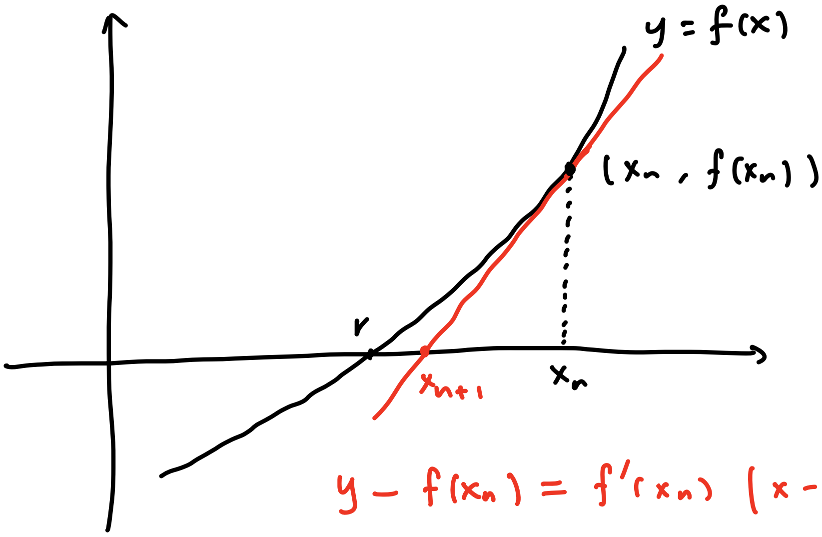
3.2 Newton's method



$$y - f(x_n) = f'(x_n)(x - x_n)$$

Set $y = 0 \Rightarrow \quad -f(x_n) = f'(x_n)(x - x_n)$

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

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$ Newton's method

- Error analysis (heuristic)

Denote the error $e_n = x_n - r$

$$e_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - r$$

$$= e_n - \frac{f(x_n)}{f'(x_n)}$$

$$= \frac{f'(x_n)e_n - f(x_n)}{f'(x_n)}$$

$$= \frac{f''(\xi_n)}{2f'(x_n)} e_n^2$$

$$0 = f(r) = f(x_n - e_n)$$

$$= f(x_n) - f'(x_n)e_n + \frac{1}{2}f''(\xi_n)e_n^2$$

$\uparrow$ between x_n and r

$\Rightarrow$ expect quadratic convergence.

Thm Assume $f \in C^2(\mathbb{R})$, and r is a simple root of f (i.e. $f(r) = 0$, $f'(r) \neq 0$). Then there exists $\delta > 0$ s.t. if initial value $x_0 \in [r-\delta, r+\delta]$, then Newton's method converges w/ error estimate

$$|x_{n+1} - r| \leq C |x_n - r|^2 \qquad n = 0, 1, 2, \dots$$

for some $C > 0$.

Proof Since f', f'' are continuous and $f'(r) \neq 0$, we have

$\min_{x \in [r-\delta, r+\delta]} |f'(x)| > 0$ for sufficiently small $\delta > 0$, and

$$c(\delta) = \frac{\max_{x \in [r-\delta, r+\delta]} |f''(x)|}{2 \min_{x \in [r-\delta, r+\delta]} |f'(x)|}$$

We also have $\displaystyle\lim_{\delta \to 0} c(\delta) = \frac{|f''(r)|}{2|f'(r)|}$

Since $e_{n+1} = \dfrac{f''(\xi_n)}{2 f'(x_n)} e_n^2$ by previous calculation,

we have $|e_{n+1}| \leq c(\delta) |e_n|^2$ provided $x_n \in [r-\delta, r+\delta]$.
(i.e. $|e_n| \leq \delta$)
$\leq c(\delta)\, \delta\, |e_n|$

Since $c(\delta)\delta \to 0$ as $\delta \to 0$, we have $c(\delta)\delta \leq \frac{1}{2}$ for sufficiently small $\delta > 0$. Then $|e_{n+1}| \leq \frac{1}{2} |e_n|$ provided $|e_n| \leq \delta$.

By induction, if $|e_0| \leq \delta$ (i.e. $x_0 \in [r-\delta, r+\delta]$) then $|e_n| \leq \delta$ $\forall n$, and $\displaystyle\lim_{n \to \infty} e_n = 0$

We also have $|e_{n+1}| \leq c(\delta) |e_n|^2$, i.e. $|x_{n+1} - r| \leq c(\delta) |x_n - r|^2$

- Stopping criterion : $\begin{cases} |x_{n+1} - x_n| < TOL_{(1)} \\ |f(x_n)| < TOL_{(2)} \end{cases}$

- Multi-dimensional generalization

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases} \qquad \text{Solve for } x_1, x_2$$

Suppose the current iteration is $x_1^{(k)}, x_2^{(k)}$

Linearize :

$$f_1(x_1,x_2) \approx f_1(x_1^{(k)}, x_2^{(k)}) + \frac{\partial f_1}{\partial x_1}(x_1^{(k)}, x_2^{(k)})(x_1 - x_1^{(k)}) + \frac{\partial f_1}{\partial x_2}(x_1^{(k)}, x_2^{(k)})(x_2 - x_2^{(k)}) = 0$$

$$f_2(x_1,x_2) \approx f_2(x_1^{(k)}, x_2^{(k)}) + \frac{\partial f_2}{\partial x_1}(x_1^{(k)}, x_2^{(k)})(x_1 - x_1^{(k)}) + \frac{\partial f_2}{\partial x_2}(x_1^{(k)}, x_2^{(k)})(x_2 - x_2^{(k)}) = 0$$

vector notations : $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \vec{F}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{pmatrix}$

$$\nabla \vec{F}(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix}$$

$$\vec{F}(\vec{x}) \approx \vec{F}(\vec{x}^{(k)}) + \nabla \vec{F}(\vec{x}^{(k)})(\vec{x} - \vec{x}^{(k)}) = 0$$

$$\nabla \vec{F}(\vec{x}^{(k)})(\vec{x} - \vec{x}^{(k)}) = -\vec{F}(\vec{x}^{(k)})$$

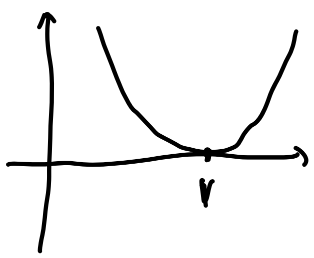
$$\boxed{\vec{x}^{(k+1)} = \vec{x}^{(k)} - \left(\nabla \vec{F}(\vec{x}^{(k)})\right)^{-1} \vec{F}(\vec{x}^{(k)})}$$

For dim $= n$ problem, in each iteration you need n^2 derivatives and solving $n \times n$ linear system.

- Evaluate Newton's method
  - Accuracy: very good (quadratic for simple roots)
  - Stability: only guaranteed when close enough to a root.
  - Regularity / info: need to calculate f'. need C^2 to achieve quadratic convergence.
  - Multi-dim: can be done w/ reasonable cost.

- Convergence rate of Newton's for multiple roots

Suppose r is a double root of f (i.e. $f(r) = f'(r) = 0$, $f''(r) \neq 0$)



$$e_{n+1} = \frac{f'(x_n) e_n - f(x_n)}{f'(x_n)} = \frac{\frac{1}{2} f''(\xi_n) e_n^2}{f''(\eta_n) e_n} = \frac{1}{2} \cdot \frac{f''(\xi_n)}{f''(\eta_n)} e_n \approx \frac{1}{2} e_n$$

$$f'(x_n) = f'(r) + f''(\eta_n)(x_n - r) = f''(\eta_n) e_n$$

$\Rightarrow$ expect linear convergence w/ $\lambda = \frac{1}{2}$.