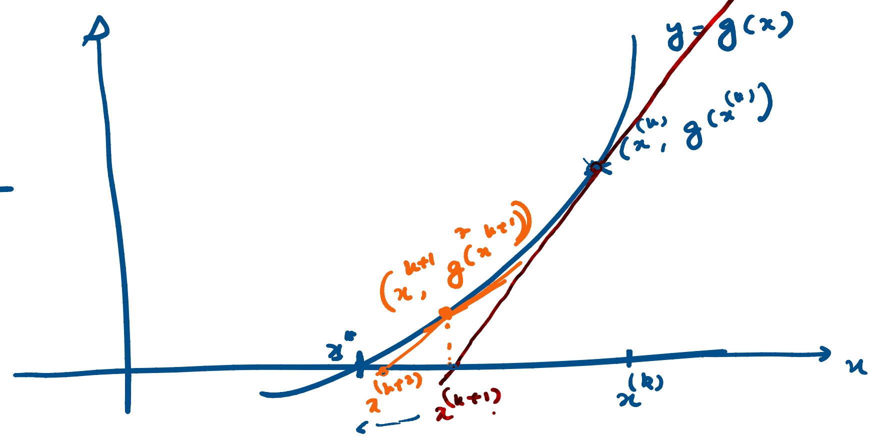
Lecture 18

Thursday, 10 March 2022 1:33 PM

(1D case)

Newton's



min $f(x)$

$$\underline{f'(x^*) = 0}$$

$$\Downarrow$$

$$g(x^*) = 0$$

$$y - g(x^{(k)}) = g'(x^{(k)})\,(x - x^{(k)})$$

$$x = x^{(k+1)}$$

$$0 = y = g(x^k) + g'(x^{(k)})\,(x - x^{(k)})$$

$$\boxed{x^{(k+1)} = x^{(k)} - \left(g'(x^{(k)})\right)^{-1} g(x^{(k)})}$$

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

multivariable

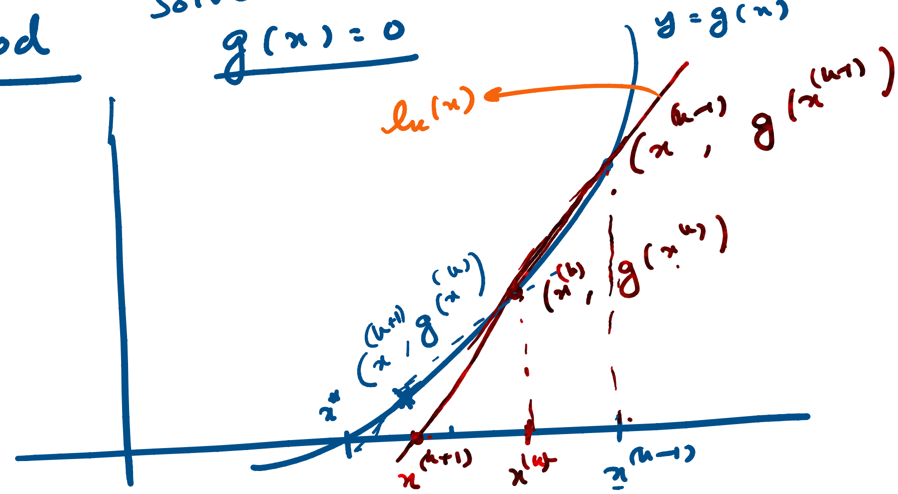
$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \left[\nabla g(\underline{x}^{(k)})\right]^{-1} g(\underline{x}^{(k)})$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \left[D^2 f(\underline{x}^{(k)})\right]^{-1} \nabla f(x^{(k)})$$

Secant' method

Solve.

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



$$x^{(k+1)} = x^{(k)} - \left[\frac{x^{(k)} - x^{(k-1)}}{g(x^{(k)}) - g(x^{(k-1)})}\right] g(x^{(k)}).$$

$l_k(x)$ passes

$$\left[g(x^{(k)}) - g(x \quad) \right.$$

$$\ell_k(x) = g(x^{(k)}) + B_k (x - x^{k})$$

$\ell_k(x)$ passes thro' $(x^{(k+1)}, g(x^{(k+1)}))$ and $(x^{(k)}, g(x^{(k)}))$

$$\rightarrow \ell_{k+1}(x) = g(x^{(k+1)}) + B_{k+1}(x - x^{(k+1)})$$

$\ell_{k+1}(x)$ passes thro' $(x^{(k+1)}, g(x^{(k+1)}))$ and $(x^{(k)}, g(x^{(k)}))$

$$\ell_{k+1}(x^{(k)}) = g(x^{(k+1)}) + B_{k+1}(x^{(k)} - x^{(k+1)})$$

$$g(x^{(k)}) - g(x^{(k+1)}) = B_{k+1}(x^{(k)} - x^{(k+1)})$$

$$\Delta g(x^{k}) = B_{k+1} \Delta x^{(k)}.$$

Undetermined system $\leftarrow$ $\boxed{(B_{k+1})_{n \times n} \Delta \underline{x}^{(k)} = \Delta g(\underline{x}^{(k)})}$ Secant condition

Lecture 17 $\rightarrow$ $(B_{k+1} = H^{-1}_{k+1})$.

$$\nabla f(\underline{x}^{(k+1)}) - \nabla f(\underline{x}^{(k)}) = \boxed{B_{k+1}} (\underbrace{\underline{x}^{(k+1)} - \underline{x}^{(k)}}_{\Delta x^{(k)}}).$$

$$[\ ]_{n\times 1} = [\ ]_{n \times n} [\ ]_{n \times 1}$$

$n \times n$

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

We want

$$\boxed{\begin{aligned} &\min \quad \| B_{k+1} - B_k \| \\ &\text{s.t.} \quad B_{k+1}^T = B_{k+1} \\ &\qquad B_{k+1} \Delta x^{(k)} = \Delta g^{(k)}. \end{aligned}}$$

$$\boxed{\begin{aligned} &\min \quad \| B_{k+1}^{-1} - B_k^{-1} \| \\ &\text{s.t.} \quad \left\{ \begin{aligned} &(B_{k+1}^{-1})^T = B_{k+1}^{-1} \\ &\Delta x^{(k)} = B_{k+1}^{-1} \Delta g^{(k)} \end{aligned} \right. \end{aligned}}$$

$\rightarrow$ (Motivation for Lagrange Method)

Quasi-Newton

- $d^{(k)} = - H_k\, g^{(k)}$ (Hessian inverse in Newton)
- $\alpha^k = \arg\min_{\alpha \geq 0} f(\underline{x}^{(k)} + \alpha d^{(k)})$
- $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^{(k)} \underline{d}^{(k)}$

(Many options)

constants

$$\dots u u^T + \beta v v^T$$

Construct updates (Many options)

$$B_{k+1} = B_k + \alpha\, \underline{u}\,\underline{u}^T + \beta\, \underline{v}\,\underline{v}^T$$

(Rank-two update)

α, β: constants

(BFGS Method)

Broyden-Fletcher-Goldfarb-Shanno (1970)

$$\underline{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \qquad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$\underline{u}, \underline{v}$ l.i.

Rank 1 $\quad \underline{u}\,\underline{u}^T = \begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix}$

$$\boxed{B_{k+1}\, \Delta \underline{x}^{(k)} = \Delta g^{(k)}.}$$

$$B_{k+1}^T = B_{k+1}$$

$$B_{k+1}\, \Delta \underline{x}^{(k)} = B_k\, \Delta \underline{x}^{(k)} + \alpha\, \underline{u}\,\underline{u}^T\, \Delta \underline{x}^{(k)} + \beta\, \underline{v}\,\underline{v}^T \Delta x^{k} = \Delta g^{(k)}$$

Choose
$$\underline{u} = \Delta g^{(k)}$$
$$\underline{v} = B_k\, \Delta \underline{x}^{(k)}$$

[Qn: Are they l.i ? Exercise. (Prove)]

$$B_k\, \Delta \underline{x}^{(k)} + \alpha\, \underbrace{\Delta g^{(k)} (\Delta g^{(k)})^T \Delta \underline{x}^{(k)}} + \beta\, B_k\, \Delta \underline{x}^{(k)} (\Delta \underline{x}^{(k)})^T B_k^T \Delta x^{(k)} = \Delta g^{(k)}$$

$$\Rightarrow \underbrace{\Delta g^{(k)}}_{u} \Big[1 - \underbrace{\alpha\, (\Delta g^{(k)})^T \Delta x^{(k)}}_{\text{Constant}} \Big] = \underbrace{B_k\, \Delta \underline{x}^{(k)}}_{v} \Big[1 + \underbrace{\beta\, (\Delta x^{(k)})^T B_k \Delta x^{(k)}}_{\text{Constant}} \Big]$$

$$\Rightarrow \boxed{\alpha = \frac{1}{(\Delta g^{(k)})^T \Delta x^{(k)}} \qquad \beta = -\frac{1}{(\Delta x^{(k)})^T B_k (\Delta \underline{x}^{(k)})}}$$

$$\boxed{B_{k+1} = B_k + \frac{(\Delta g^{(k)}) (\Delta g^{(k)})^T}{(\Delta g^{(k)})^T \Delta x^{(k)}} - \frac{B_k\, \Delta x^{(k)} (\Delta x^{(k)})^T B_k}{(\Delta x^{(k)})^T B_k (\Delta x^{(k)})}}$$

Rank-two update.