Math 270B HW Answer

6.2. Newton method $x = \phi(x) = x - \dfrac{f(x)}{f'(x)}$

Assume $f(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0$ and $f^{(m)}(\alpha) \neq 0$

$$\phi'(x) = 1 - \frac{f'(x) \cdot f'(x) - f(x) f''(x)}{f'(x)^2} = \frac{f(x) f''(x)}{f'(x)^2}$$

Since $f'(\alpha) = 0$ need to do limitation for estimating $\phi'(\alpha)$

$$\phi'(\alpha) = \lim_{x \to \alpha} \frac{f(x) f''(x)}{f'(x)^2} = \lim_{x \to \alpha} \frac{\left[\frac{f^{(m)}(\alpha)}{m!}(x-\alpha)^m + O\left((x-\alpha)^{m+1}\right)\right] \cdot \left[\frac{f^{(m)}(\alpha)}{(m-2)!}(x-\alpha)^{m-2} + O\left((x-\alpha)^{m-1}\right)\right]}{\left[\frac{f^{(m)}(\alpha)}{(m-1)!}(x-\alpha)^{m-1} + O\left((x-\alpha)^m\right)\right]^2}$$

$$= \lim_{x \to \alpha} \frac{\frac{f^{(m)}(\alpha)}{m!}(x-\alpha)^m \cdot \frac{f^{(m)}(\alpha)}{(m-2)!}(x-\alpha)^{m-2}}{\left[\frac{f^{(m)}(\alpha)}{(m-1)!}(x-\alpha)^{m-1}\right]^2} = \frac{[(m-1)!]^2}{m!\,(m-2)!} = 1 - \frac{1}{m}$$

Modify Newton's method $x_{k+1} = x_k - m \cdot \dfrac{f(x_k)}{f'(x_k)}$

then $\phi(x) = x - m \cdot \dfrac{f(x)}{f'(x)}$ $\qquad \phi'(x) = 1 - m \cdot \dfrac{f'(x) f'(x) - f(x) f''(x)}{f'(x)^2}$

$$\phi'(x) = 1 - m\left(1 - \frac{f(x) f''(x)}{f'(x)^2}\right)$$

then $\lim_{x \to \alpha} \phi'(x) = 0$

thus modified Newton method has order 2 convergence.

6.5 Steffensen's method

$$x_{k+1} = x_k - \frac{f(x_k)}{\varphi(x_k)} \qquad \varphi(x_k) = \frac{f(x_k + f(x_k)) - f(x_k)}{f(x_k)}$$

$$\phi(x) = x - \frac{f(x)}{\varphi(x)} \qquad \phi'(x) = 1 - \frac{f'(x)\cdot\varphi(x) - f(x)\cdot\varphi'(x)}{\varphi^2(x)}$$

$$= 1 - \left[\frac{f'(x)}{\varphi(x)} - \frac{f(x)\,\varphi'(x)}{\varphi^2(x)}\right]$$

$$\lim_{x\to\alpha} \phi'(x) = 1 - \lim_{x\to\alpha} \frac{f'(x)}{\varphi(x)} \qquad \text{notice} \quad \lim_{x\to\alpha} \varphi(x) = f'(\alpha)$$

$$\therefore \lim_{x\to\alpha} \phi'(x) = 0 \qquad \text{Steffensen's method is second order}$$

2.1

(a) $$F(x+h) = F(x) + \int_x^{x+h} F'(t)\,dt$$

Simply Apply Newton–Leibniz formula

(b) $$F(x+h) = F(x) + h\int_0^1 F'(x+\varepsilon h)\,d\varepsilon$$

$$\int_x^{x+h} F'(t)\,dt \;\overset{\text{change of variable}}{\underset{t = x + h\cdot\varepsilon}{=\!=\!=}}\; \int_0^1 F'(x+h\varepsilon)\cdot(h\,d\varepsilon) = h\int_0^1 F'(x+h\varepsilon)\,d\varepsilon$$

(c) $$F(x+h) = F(x) + h\int_0^1 [F'(x+\varepsilon h) - F'(x) + F'(x)]\,d\varepsilon$$

$$= F(x) + hF'(x) + h\int_0^1 [F'(x+\varepsilon h) - F'(x)]\,d\varepsilon$$

(d) $\int_0^1 [F'(x+\varepsilon h) - F'(x)] d\varepsilon$

$$= \int_0^1 \left(\int_x^{x+\varepsilon h} F''(t)\, dt \right) d\varepsilon = \int_0^1 \left(\int_x^{\infty} F''(t) \cdot I_{[x, x+\varepsilon h]}(t)\, dt \right) d\varepsilon$$

$$I_{[x,x+\varepsilon h]}(t) = \begin{cases} 1 & t \in [x, x+\varepsilon h] \\ 0 & \text{else} \end{cases}$$

indicator function of Interval $[x, x+\varepsilon h]$

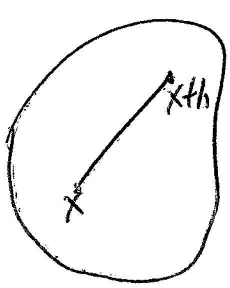
then Apply Fubini's theorem to exchage order of integration

$$= \int_x^{\infty} \left(\int_0^1 F''(t) \cdot I_{[x,x+\varepsilon h]}(t)\, d\varepsilon \right) dt = \int_x^{\infty} F''(t) \cdot \left(\int_0^1 I_{[x,x+\varepsilon h]}(t)\, d\varepsilon \right) dt$$

$$\int_0^1 I_{[x,x+\varepsilon h]}(t)\, d\varepsilon = \begin{cases} 0 & t \notin [x, x+h] \\ 1 - \frac{t-x}{h} & t \in [x, x+h] \end{cases}$$

$$= \int_x^{x+h} F''(t) \cdot \left(1 - \frac{t-x}{h}\right) dt \xlongequal{t = x+h\varepsilon} h \int_0^1 F''(x+h\varepsilon) \cdot (1-\varepsilon)\, d\varepsilon$$

2.2



connect x and $x+h$ with segment $x + ht$ $(0 \le t \le 1)$

define $\varphi(t) = F(x+ht)$ $\quad \varphi'(t) = F'(x+ht)h$

$$\varphi(1) = F(x+h) = \varphi(0) + \int_0^1 \varphi'(t)\, dt$$

$$= F(x) + \int_0^1 F'(x+ht) h\, dt$$

$$= F(x) + F'(x)h + \int_0^1 [F'(x+ht) - F'(x)] h\, dt$$

2.4. (a) $f(x) = \frac{1}{2}x^T H x = \frac{1}{2}\sum_{i,j} h_{ij} x_i x_j$

$$\frac{\partial f}{\partial x_s} = \frac{\partial}{\partial x_s}\left(\frac{1}{2}\sum_{i,j} h_{ij} x_i x_j\right)$$

$$= \frac{1}{2}\sum_{i,j} h_{ij} \frac{\partial}{\partial x_s}(x_i x_j) = \frac{1}{2}\sum_{i,j} h_{ij}\left(\frac{\partial x_i}{\partial x_s} x_j + x_i \cdot \frac{\partial x_j}{\partial x_s}\right)$$

$$= \frac{1}{2}\sum_{i,j} h_{ij}\left(\delta_s^i x_j + x_i \delta_s^j\right)$$

$$= \frac{1}{2}\sum_j \left(h_{sj} x_j + h_{is} x_i\right) \qquad \therefore \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \frac{1}{2}(H + H^T)x$$

$f'(x) = \nabla f(x)^T$. $\qquad \nabla^2 f(x) = \frac{1}{2}(H + H^T)$

especially when H is symmetric $\quad \nabla f(x) = Hx \quad \nabla^2 f(x) = H$

(b) $f(x) = b^T A x - \frac{1}{2}x^T A^T A x \qquad A^T A$ is symmetric

$$\nabla f(x) = \nabla(b^T A x) - \nabla\left(\frac{1}{2}x^T A^T A x\right) = A^T b - A^T A x$$

$$\nabla^2 f(x) = -A^T A$$

(c) $f(x) = \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots x_n^2}$

$$\frac{\partial f}{\partial x_i} = \frac{1}{2}\left(x_1^2 + x_2^2 + \cdots x_n^2\right)^{-\frac{1}{2}} \cdot (2x_i) = \frac{x_i}{\|x\|} \qquad \nabla f(x) = \frac{x}{\|x\|}$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\frac{\partial x_i}{\partial x_j} \cdot \|x\| - x_i \cdot \frac{\partial(\|x\|)}{\partial x_j}}{\|x\|^2} = \frac{\delta_j^i}{\|x\|} - \frac{x_i}{\|x\|^2} \cdot \frac{x_j}{\|x\|} \qquad \therefore \text{Hessian}(H) = \frac{1}{\|x\|} \cdot I - \frac{xx^T}{\|x\|^3}$$

2.10 $L_k(x) = F(x_k) + F'(x_k)(x - x_k)$

$$F(x) - L_k(x) = F(x) - [F(x_k) + F'(x_k)(x-x_k)]$$

$$= \int_0^1 F'((1-t)x_k + t\cdot x)\cdot(x-x_k)\,dt - F'(x_k)(x-x_k)$$

$$= \int_0^1 [F'((1-t)x_k + t\cdot x) - F'(x_k)](x-x_k)\,dt$$

$$\|F(x) - L_k(x)\| \le \int_0^1 \|F'((1-t)x_k + tx) - F'(x_k)\|\cdot\|x-x_k\|\,dt$$

Lipschitz Continuous of F'

$$\le \int_0^1 \gamma\cdot t\cdot\|x-x_k\|\,dt \cdot \|x-x_k\| = \frac{1}{2}\gamma\|x-x_k\|^2$$

2.12 $x_{k+1} = x_k - [F'(x_k)]^{-1}F(x_k)$

$$x_{k+1} - x^* = x_k - [F'(x_k)]^{-1}F(x_k) - x^*$$

$$= [F'(x_k)]^{-1}\cdot[F'(x_k)(x_k - x^*) - F(x_k)]$$

$$= [F'(x_k)]^{-1}\cdot(F(x^*) - F(x_k) - F'(x_k)(x^*-x_k))$$

$$= [F'(x_k)]^{-1}\cdot\int_0^1 [F'((1-\varepsilon)x_k + \varepsilon x^*) - F'(x_k)]\,d\varepsilon\cdot(x^*-x_k)$$

$$\|x_{k+1} - x^*\| \le \|F'(x_k)^{-1}\|\cdot\int_0^1 \|F'((1-\varepsilon)x_k + \varepsilon x^*) - F'(x_k)\|\,d\varepsilon\cdot\|x^*-x_k\|$$

when Newton's method start to Converge

superlinear

$$\lim_{k\to\infty} \|F'(x_k)^{-1}\|\cdot\int_0^1 \|F'((1-\varepsilon)x_k + \varepsilon x^*) - F'(x_k)\|\,d\varepsilon = 0$$

when F' is Lipschitz, (2.10) shows second order Convergence

2.15. (a) $q(x) = c^T x + \frac{1}{2} x^T H x$ $\quad$ H is symmetric

$\nabla q(x) = c + Hx$

(b) $\text{Hessian}\, q(x) = H$

$q(x_0+p) = q(x_0) + \nabla q(x_0)^T \cdot p + \frac{1}{2} p^T \text{Hessian}\, q(x_0)\, p = q(x_0) + (c+Hx_0)^T p + \frac{1}{2} p^T H p$

(c) $q(x_0+\alpha p) = q(x_0) + \alpha (c+Hx_0)^T p + \alpha^2 (\frac{1}{2} p^T H p)$

Since H is positive definite $\frac{1}{2} p^T H p > 0$

$q(x_0+\alpha p)$ is a concave up quadratic function

the minimizer $\alpha^* = \dfrac{-(c+Hx_0)^T p}{2(\frac{1}{2} p^T H p)} = \dfrac{-\nabla q(x_0)^T p}{p^T H p} > 0$

2.19. minimize $x_1^2 + 2x_2^2$ $\quad$ $f(x_1,x_2) = x_1^2 + 2x_2^2$

s.t $x_1 + x_2 - 1 = 0$ $\quad$ $g(x_1,x_2) = x_1 + x_2 - 1$

(a) only equality condition. KKT condition is reduced to lagrange multiplier.

$L(x_1, x_2, \lambda) = x_1^2 + 2x_2^2 + \lambda(x_1 + x_2 - 1)$

$$\begin{cases} 2x_1 + \lambda = 0 \\ 4x_2 + \lambda = 0 \\ x_1 + x_2 - 1 = 0 \end{cases}$$

unique solution $\begin{cases} x_1 = \frac{2}{3} \\ x_2 = \frac{1}{3} \\ \lambda = -\frac{4}{3} \end{cases}$

$\nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} > 0$

$\therefore (\frac{2}{3}, \frac{1}{3})$ is optimal.

(b) minimize $x_1^3 + x_2^3$

s.t $x_1 + x_2 - 1 = 0$

$$L(x_1, x_2, \lambda) = x_1^3 + x_2^3 + \lambda(x_1 + x_2 - 1)$$

$$\begin{cases} 3x_1^2 + \lambda = 0 \\ 3x_2^2 + \lambda = 0 \\ x_1 + x_2 - 1 = 0 \end{cases} \qquad \begin{cases} x_1 = \frac{1}{2} \\ x_2 = \frac{1}{2} \\ \lambda = -\frac{3}{4} \end{cases}$$

$$f(x_1, x_2) = x_1^3 + x_2^3 \qquad \nabla^2 f = \begin{pmatrix} 6x_1 & 0 \\ 0 & 6x_2 \end{pmatrix}\Bigg|_{(\frac{1}{2}, \frac{1}{2})} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} > 0$$

local minimizer