Exercise 1

501703

a) $f_{sq}(x) = x^T C x$

$$\nabla f_{sq}(x) = \underline{\underline{2Cx}}$$

$$f_{hole}(x) = \frac{x^T C x}{a^2 + x^T C x}$$

$$\nabla f_{hole}(x) = \frac{2Cx}{a^2 + x^T C x} - \frac{x^T C x \, 2Cx}{(a^2 + x^T C x)^2}$$

$$\cancel{= \frac{2Cx(a^2 + x^T C x)}{(a^2 + x^T C x)^2} - \frac{x^T C x \, 2Cx}{(a^2 + x^T C x)^2}}$$

$$= \left(\frac{2}{a^2 + x^T C x} - \frac{2 x^T C x}{(a^2 + x^T C x)^2} \right) \cdot Cx$$

$$= \left(\frac{2(a^2 + x^T C x) - 2 x^T C x}{(a^2 + x^T C x)^2} \right) Cx$$

$$= \underline{\underline{\frac{2a^2}{(a^2 + x^T C x)^2} Cx}}$$

$$f_{exp}(x) = -\exp\left(-\frac{1}{2} x^T C x\right)$$

$$\nabla f_{exp}(x) = \underline{\underline{+\exp\left(-\frac{1}{2} x^T C x\right) \cdot Cx}}$$

b) $\nabla^2 f_{sq}(x) = \underline{\underline{2C}}$

$$\nabla^2 f_{hole}(x) = \frac{2a^2C}{(a^2+x^TCx)^2} - 2\cdot\frac{2a^2\overset{n\times n}{C}\overset{n\times 1}{x}}{(a^2+x^TCx)^3}\cdot 2\cdot \overset{1\times n}{x^T}\overset{n\times n}{C}$$

$$= \underline{\underline{\frac{2a^2C}{(a^2+x^TCx)^2} - 8\,\frac{a^2Cxx^TC}{(a^2+x^TCx)^3}}}$$

$$\underline{\underline{\nabla^2 f_{exp}(x) = -\exp\left(-\tfrac{1}{2}x^TCx\right)\cdot \overset{n\times n}{C}\overset{n\times 1}{x}\cdot x^TC + \exp\left(-\tfrac{1}{2}x^TCx\right)\cdot C}}$$

501703

2) Converge proof

a) $f: \mathbb{R}^n \to \mathbb{R}$ $\quad f_{min} = \min_x f(x)$ $\quad$ curvatures between m and M

$$f(y) \approx f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) + R(y)$$

rest higher order

$R(y)$: $z = t x + (1-t) y \quad t \in [0,1]$

$$f(y) \overset{\text{exact}}{=} f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \underline{\nabla^2 f(z)}(y-x)$$

with this form we can provide precise upper & lower bound for $f(y)$ near x

- $\nabla^2 f(z) = Q \Lambda Q^T$ $\quad$ SVD

$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ with all $\lambda_i \leq M$, $\lambda_i \geq m$ $\quad$ $\begin{pmatrix} \lambda_i & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

λ_i describe curvature

$$(y-x)^T \nabla^2 f(z)(y-x) = u^T \nabla^2 f(x) \cdot u = w^T \underbrace{Q^T Q}_{I} \Lambda \underbrace{Q^T Q}_{I} w = \underline{\underline{\sum_i \lambda_i w_i^2}}$$

Q is base change

- λ_i is between m and M

~~Smallest~~ lower bound all $\lambda = m$: $m\|w\|^2$

upper bound all $\lambda = M$: $M\|w\|^2$

$\|u\| = \|w\|$:

$$m\|u\|^2 \leq u^T \nabla^2 f(z) u \leq M\|u\|^2$$

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}\|y-x\|^2$$

$$f(y) \leq f(x) + \nabla f(x)^T (y-x) + \frac{M}{2}\|y-x\|^2$$

b) $$f(x) - \frac{1}{2m} |\nabla f(x)|^2 \le f_{min} \le f(x) - \frac{1}{2M} |\nabla f(x)|^2$$

min. of lower bound:

$$\cancel{f(y)} = f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}(y-x)^2$$

lower_bound

$$\cancel{\frac{\partial f(y)}{\partial y} = \nabla f(x)}$$

$$f_{lower\_bound}(y) = f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}(y-x)^2$$

$$\frac{\nabla f(y)}{\partial y} = \nabla f(x)^T + m(y-x) = 0$$

$$y^* = x - \frac{1}{m} \nabla f(x)^T$$

upper bound:
same

$$\cancel{f(y) =} \quad y^* = x - \frac{1}{M} \nabla f(x)^T$$

c) $y = x - \alpha \nabla f(x) \rightarrow y - x = -\alpha \nabla f(x)$

$$f(y) \le f(x) + \nabla f(x)^T (y-x) + \frac{M}{2} \|y-x\|^2$$

$$f(y) \le f(x) + \nabla f(x)^T (-\alpha \nabla f(x)) + \frac{M}{2} \|-\alpha \nabla f(x)\|^2$$

$$f(y) \le f(x) - \alpha \nabla f(x)^2 + \underbrace{\frac{\alpha^2 M}{2} \|\nabla f(x)\|^2}_{\le \frac{\alpha}{2} \|\nabla f(x)\|^2} = f(x) - \alpha \nabla f(x)^2 + \frac{\alpha}{2} \|\nabla f(x)\|^2$$

$$\alpha \le \frac{1}{M} \rightarrow \alpha^2 \le \frac{\alpha}{M}$$

d) $\alpha \leftarrow \beta\alpha$ and $0 < \beta < 1$

$M \geq m$

at some point α will $\frac{1}{M} \leq \alpha \leq \frac{1}{m}$

$\frac{1}{M} \leq \frac{1}{m}$

line search $\nRightarrow$ converge