Homework 5 Haocong Wang

1. $f(x) = \max\left((x+1)^2, (x-3)^2 \right)$. let $g(x) = (x+1)^2$, $h(x) = (x-3)^2$

$\Rightarrow f(x) = \begin{cases} (x+1)^2, & x > 1 \\ 4, & x = 1 \\ (x-3)^2, & x < 1 \end{cases}$. $g(x)$, $h(x)$ are convex & differentiable

$\Rightarrow \partial f(x) : \begin{cases} \partial g(x) = 2x+2, & x > 1 \\ \text{any point on the line segment between } \partial g(1) = 4 \text{ and } \partial h(1) = -4, & x = 1 \\ \partial h(x) = 2x-6, & x < 1 \end{cases}$

2. $\partial f(x) = \{ g : f(y) \geq f(x) + g^T (y-x) \}$

$f(x)$ is convex $\Rightarrow f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

Let $x = \lambda x + (1-\lambda) y \Rightarrow f(y) \geq f(\lambda x + (1-\lambda) y) + g^T (y - \lambda x - (1-\lambda) y) = f(\lambda x + (1-\lambda) y) + \lambda g^T (y-x)$ ①

We also have $f(x) \geq f(\lambda x + (1-\lambda) y) + (\lambda - 1) g^T (y-x)$ ②

② $\times (\lambda - 1) \Rightarrow (\lambda - 1) f(y) \geq (\lambda - 1) f(\lambda x + (1-\lambda) y) + \lambda(\lambda-1) g^T (y-x)$ ③

③ $\times \lambda \Rightarrow \lambda f(x) \geq \lambda f(\lambda x + (1-\lambda) y) + \lambda(\lambda-1) g^T (y-x)$ ④

① - ④ $\Rightarrow \lambda f(x) + (1-\lambda) f(y) \geq (\lambda - (\lambda-1)) f(\lambda x + (1-\lambda) y) = f(\lambda x + (1-\lambda) y)$

$\Rightarrow \partial f(x)$ is convex

3. $x^+ = \mathrm{prox}_{t}(x) = \arg\min_z \left(\|z\|_* + \frac{1}{2t} \|z - x\|_F^2 \right)$. if and only if $x - x^+ \in t \partial \|x^+\|_*$

$\Rightarrow \frac{x - x^+}{t} = \partial \|x^+\|_*$ ①

If $x^+ = U \Sigma^+ V^T$, where $\Sigma^+[i,i] = \max(\sigma_i - t, 0)$, $\frac{x - x^+}{t} = U \left(\frac{\Sigma - \Sigma^+}{t} \right) V^T$

If $\Sigma_{ii} \geq t$, $\left(\frac{\Sigma - \Sigma^+}{t} \right)_{ii} = 1$; If $\Sigma_{ii} < t$, $0 < \left(\frac{\Sigma - \Sigma^+}{t} \right)_{ii} = \left(\frac{\Sigma}{t} \right)_{ii} < 1$

So we have the original matrix $\frac{\Sigma - \Sigma^+}{t}$ as $\begin{bmatrix} I & 0 \\ 0 & \Sigma_1 \end{bmatrix}$

$\Rightarrow \frac{x - x^+}{t} = U \left(\frac{\Sigma - \Sigma^+}{t} \right) V^T = U \begin{bmatrix} I & 0 \\ 0 & \Sigma_1 \end{bmatrix} V^T = U_0 V_0^T + U_1 \Sigma_1 V_1^T$, $U = [U_0 U_1]$, $V = [V_0 V_1]$

$\because U_0 U_1$ orthonormal $\therefore U_0^T U_1 = 0$. let $W = U_1 \Sigma_1 V_1^T$ $\therefore U^T W = 0 \Rightarrow W \cdot V = 0$

$\because$ Maximum singular value of Σ_1 is less than 1

$\therefore \frac{x - x^+}{t} = U_0 V_0^T + W$, $W = U_1 \Sigma_1 V_1^T$, $U = [U_0 U_1]$, $V = [V_0 V_1]$, $U^T W = 0$, $WV = 0$, $\|W\|_2 < 1$

$\Rightarrow \frac{x - x^+}{t} \in \partial \|x^+\|_* \Rightarrow$ We can compute the prox operator by SVT.