HW3　　陈伟杰　　1901111420

1. 若 f 凸，$x\in \mathrm{dom} f$，g 满足下列条件则称为 f 在 x 处的次梯度

$$f(y)\geq f(x)+g^T(y-x)\quad \forall y\in \mathrm{dom} f$$

(a). $f(x)=\|Ax-b\|_2+\|x\|_2$ 是凸函数 （不妨设 $A\neq 0$）

记 $h(y)=\|y\|_2$

则 $f(x)=h(Ax-b)+h(x)$

次梯度：$\partial f(x)=\partial\left(h(Ax-b)+h(x)\right)$

$$=A^T\partial h(Ax-b)+\partial h(x)$$

其中：
$$\partial h(y)=\begin{cases}\left\{\frac{y}{\|y\|_2}\right\}, & y\neq 0\\ \{g\mid \|g\|_2\leq 1\}, & y=0\end{cases}$$

(b). $f(x)=\inf\limits_y \|Ay-x\|_\infty$ 是凸函数，记 $h(x,z)=\|z-x\|_\infty$

A 的列空间为 $C(A)=\{z\mid Ay=z,\ y\in \mathbb{R}^m\}\subset \mathbb{R}^n$

则 $f(x)=\inf\limits_{z\in C(A)}\|z-x\|_\infty$

① 若 $x\in C(A)$，则 $f(x)=0$，故次梯度 $g=0$

② 若 $x\notin C(A)$ 即 $f(x)\neq 0$，记 $\hat{z}=\hat{P}(x)$ 为 x 在 $C(A)$ 的投影.

$$\Rightarrow f(x) = \|\hat{p}(x) - x\|_\infty \qquad \left[\begin{array}{l} \text{i.e. } \|\hat{z}-x\|_\infty \le \|z-x\|_\infty \\ \qquad \forall z \in C(A) \end{array}\right]$$

考虑 L_∞ 的次梯度，由于 $g(x) = \|x\|_\infty = \sup_{\|w\|_1 \le 1} w^T x$

则 $\partial g(x) = \{w \mid \|w\|_1 \le 1,\ x^T w = \|x\|_\infty\}$

$$\Rightarrow \partial f(x) = \{w \mid \|w\|_1 \le 1,\ (\hat{p}(x) - x)^T w = \|\hat{p}(x) - x\|_\infty\}.$$

---

2. $\mathrm{prox}_f(x) = \underset{u}{\mathrm{argmin}} \left(f(u) + \frac{1}{2}\|u - x\|_2^2\right)$

(a). $f(x) = \|x\|_1$, $\mathrm{dom} f = \{x \mid \|x\|_\infty \le 1\} = [-1, 1]^n$

$$\Rightarrow [\mathrm{prox}_f(x)]_i = \underset{u_i}{\mathrm{argmin}} \left[|u_i| + \frac{1}{2}(u_i - x_i)^2\right]$$

若 $u_i \ge 0$, $[\mathrm{prox}_f(x)]_i = \underset{u_i}{\mathrm{argmin}} \left(\frac{1}{2}[u_i - (x_i - 1)]^2\right)$

若 $u_i \le 0$, $[\mathrm{prox}_f(x)]_i = \underset{u_i}{\mathrm{argmin}} \left(\frac{1}{2}(u_i - (x_i + 1))^2\right)$

由于 $-1 \le x_i \le 1$, $\Rightarrow \begin{cases} x_i - 1 \le 0 \\ x_i + 1 \ge 0 \end{cases}$

$$\Rightarrow [\mathrm{prox}_f(x)]_i = 0$$

即 $\mathrm{prox}_f(x) = 0$

(b) $f(x) = \max_k x_k$

$$\text{prox}_f(x) = \arg\min_u \left(\max_k u_k + \frac{1}{2} \|u-x\|_2^2 \right)$$ 不妨记 $u_j = \max_k u_k$

$$\left(\text{prox}_f(x)\right)_i = \arg\min_{u_i} \left[\delta_{ij} u_j + \frac{1}{2}(u_i - x_i)^2 \right]$$

$$= \arg\min_{u_i} \left[\frac{1}{2} \left(u_i - (x_i - \delta_{ij}) \right)^2 \right]$$

设 x_{i_1}，x_{i_2} 分别为 x 的最大元素，第二大元素，且 $x_{i_1} > x_{i_2}$，x_{i_k} 为第 k 大元素.

① 若 $x_{i_1} - 1 \geq x_{i_2}$

则
$$u_i = \begin{cases} x_i - 1, & i = i_1 \\ x_i, & \text{o.w.} \end{cases}$$

② 若 $x_{i_1} - 1 < x_{i_k} \leq \dots \leq x_{i_2}$

在 $k-1$ 个数中选取前 $m-1$ 个数，使得 $\frac{1}{m}\left(\sum_{j=1}^{m} x_{i_j} - 1\right)$ 大于剩下的数

则
$$u_i = \begin{cases} \frac{1}{m}\left(\sum_{j=1}^{m} x_{i_j} - 1\right), & i = i_1, i_2, \dots i_m \\ x_i, & \text{o.w.} \end{cases}$$

(c) $f(x) = \|Ax - b\|_1$，且 $AA^T = D$（对角阵，且对角元 >0）

$$\text{prox}_f(x) = \arg\min_u \left(\|Au - b\|_1 + \frac{1}{2}\|u - x\|_2^2 \right)$$

由 $AA^T = D \Rightarrow A$ 可逆；令 $y = Au - b \Rightarrow u = A^{-1}(y+b)$

$$\Rightarrow \text{prox}_f(x) = \arg\min_y \left[\|y\|_1 + \frac{1}{2}\left(A^{-1}(y+b) - x\right)^T \left(A^{-1}(y+b) - x\right) \right]$$

$$= \arg\min_y \left[\|y\|_1 + \frac{1}{2} y^T D^{-1} y + \left(b^T D^{-1} - x^T A^{-1}\right) y \right]$$

记 $z = D^{-1}b - A^{-T}x$

$$\text{prox}_f(x) = \arg\min_y \left(\|y\|_1 + \frac{1}{2} y^T D^{-1} y + z^T y \right)$$

由于 D 对角，故可以拆成对每个 y_i 的 argmin

$$\Rightarrow [\text{prox}_f(x)]_i = \arg\min_{y_i} \left[|y_i| + \frac{1}{2}\frac{1}{D_{ii}} y_i^2 + z_i y_i \right]$$

若 $y_i \geq 0$，$[\text{prox}_f(x)]_i = \arg\min_{y_i} \left[\frac{1}{2D_{ii}} \left(y_i + (z_i+1)D_{ii} \right)^2 \right]$

若 $y_i \leq 0$，$[\text{prox}_f(x)]_i = \arg\min_{y_i} \left[\frac{1}{2D_{ii}} \left(y_i + (z_i-1)D_{ii} \right)^2 \right]$

$$\Rightarrow [\text{prox}_f(x)]_i = \begin{cases} -D_{ii}(z_i+1), & z_i+1 < 0 \\ 0 & \text{o.w.} \\ -D_{ii}(z_i-1), & z_i-1 > 0 \end{cases}$$ 其中 $z_i = (D^{-1}b - A^{-T}x)_i$

3.

$$\min_{t,y} \phi(t,y) \quad \text{其中} \quad \phi(t,y) = t + \frac{1}{(1-\alpha)n} \sum_{i=1}^n (y_i - t)_+ + \frac{\sigma}{2} \|y - w\|_2^2$$

$w \in \mathbb{R}^n$, $\sigma, \alpha > 0$

$$\Rightarrow \phi(t,y) = t + \sum_{i=1}^n \left(\frac{1}{(1-\alpha)n} (y_i - t)_+ + \frac{\sigma}{2} (y_i - x_i)^2 \right)$$

使用交替极小化的方法，记 $f(s) = \frac{1}{(1-\alpha)\sigma n}(s-t)_+$

对于 y:

$$\mathrm{prox}_f(x_i) = \underset{y_i}{\mathrm{argmin}}\left(\frac{1}{(1-\alpha)n\sigma}(y_i-t)_+ + \frac{1}{2}(y_i-x_i)^2\right)$$

若 $y_i > t$， $y_i^+ = \mathrm{prox}_f(x_i) = \max\left(x_i - \frac{1}{(1-\alpha)n\sigma}, t\right)$

若 $y_i \le t$， $y_i^+ = \mathrm{prox}_f(x_i) = \min(x_i, t)$

对于 t:

$$\min_t \phi(t, y^+) = \min_t \left[t + \sum_{i=1}^{n}\frac{1}{(1-\alpha)n}(y_i-t)_+\right]$$

若 $\alpha \in (0,1)$， $\frac{\partial \phi(t,y^+)}{\partial t} = 1 + \sum_{i=1}^{n}\frac{1}{(1-\alpha)n}\frac{\partial}{\partial t}(y_i-t)_+$

若 $y_i > t$ ， $\frac{\partial}{\partial t}(y_i-t)_+ = -1$

若 $y_i \le t$ ， $\frac{\partial}{\partial t}(y_i-t)_+ = 0$

$\Rightarrow$ $t^+ = t - \gamma \frac{\partial}{\partial t}\phi(t, y^+)$ ，ε 为步长.

因此算法如下： $y=0,\ t=0,\ y^+=\infty,\ t^+=\infty$

Loop: until $|y^+-y| < \varepsilon_1$， $|t^+-t| < \varepsilon$

① $y_i^+ = \begin{cases} \max\left(x_i - \frac{1}{(1-\alpha)n\sigma}, t\right) \\ \min(x_i, t) \end{cases}$

② $t^+ = t - \gamma \frac{\partial}{\partial t}\phi(t, y^+)$

---

4. $\mathrm{prox}_f(X) = \underset{U}{\mathrm{argmin}} \left(f(U) + \frac{1}{2} \|U - X\|_F^2 \right)$, $X \in S^n$

$f(U) = -\log\det U$, $U \in S_{++}^n$ 为凸函数

由于 $\|X\|_F^2 \equiv Tr(X^T X) = Tr(X^2)$

若 $U = \mathrm{prox}_f(X)$，则 $\frac{d}{dU}\left(f(U) + \frac{1}{2}\|U - X\|_F^2\right) = 0$

$\Rightarrow \quad 0 = -U^{-1} + U - X$

设 U 与 X 可同时对角化，则有 $UX = XU$

上式可改写：$U^2 - XU - I = 0$

$\Rightarrow \quad U = \frac{1}{2}\left[(X^2 + 4I)^{1/2} + X \right]$

即 $\mathrm{prox}_f(X) = \frac{1}{2}\left[(X^2 + 4I)^{1/2} + X \right]$