Optimization Methods **Fall 2020**

Homework 2

Instructor: Lijun Zhang Name: 左之睿, StudentId: 191300087

Problem 1:

a). $\frac{\partial f}{\partial x_i} = -\frac{1}{x_i}$ $\frac{\partial^2 f}{\partial x_i \partial x_j} = 0 \cdot (i \neq j)$. $\frac{\partial^2 f}{\partial x_i^2} = \frac{1}{x_i^2}$.

故 f 的 Hessen 矩阵

$$H = \begin{bmatrix} \frac{1}{x_1^2} & 0 & 0 & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & & \frac{1}{x_n^2} \end{bmatrix}$$

是正定的、故 $f(x)$ 是严格凸的

b).

$\Rightarrow$ 若 f 为凸函数.

则 $\nabla f(x)^T (x-y) \geq f(x) - f(y)$. ①

$\begin{cases} y = x \\ x = y \end{cases}$ 有 $\nabla f(y)^T (y-x) \geq f(y) - f(x)$.

即 $\nabla f(y)^T (x-y) \leq f(x) - f(y)$. ②

由①② $\nabla f(x)^T (x-y) \geq \nabla f(y)^T (x-y)$.

即 $(\nabla f(x) - \nabla f(y))^T (x-y) \geq 0$.

⇐. 已知 $(\nabla f(x) - \nabla f(y))^T(\cancel{\ } x-y) \geq 0$.

即 $\nabla f(x)^T (x-y) \geq \nabla f(y)^T (x-y)$.

令 $x = y + \Delta y$. $(\Delta y \to 0^+)$.

由 梯度定义

$$\nabla f(y) = \frac{f(x) - f(y)}{x-y} \quad (\Delta y \to 0).$$

故 $\nabla f(x)^T \cdot (x-y) \geq f(x) - f(y)$. $\to$ $f(x)$ 为凸函数.

(c). $\forall x, y \in \mathrm{dom}(f)$, $\theta \in [0,1]$

$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$.

$\forall\ s, t > 0$.

$$g(\theta x + (1-\theta) y, \cancel{\theta t}\ \theta t + (1-\theta) s)$$

$$= (\theta t + (1-\theta) s) \cdot f\left(\frac{\theta x + (1-\theta) y}{\theta t + (1-\theta) s}\right)$$

$$= (\theta t + (1-\theta) s) \cdot f\left(\frac{\theta t \cdot \frac{x}{t}}{\theta t + (1-\theta) s} + \frac{(1-\theta) s \cdot \frac{y}{s}}{\theta t + (1-\theta) s}\right)$$

$$\leq (\theta t + (1-\theta) s) \cdot \left[\frac{\theta t}{\theta t + (1-\theta) s} \cdot f\left(\frac{x}{t}\right) + \frac{(1-\theta) s}{\theta t + (1-\theta) s} \cdot f\left(\frac{y}{s}\right)\right]$$

$$= \theta \cdot t f\left(\frac{x}{t}\right) + (1-\theta) \cdot s f\left(\frac{y}{s}\right)$$

$$= \theta \cdot g(x, t) + (1-\theta)\, g(y, s).$$

Problem 2:

$$\frac{\partial f(x)}{\partial x_i} = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1-p}{p}} \cdot x_i^{p-1} = \left(\frac{f(x)}{x_i}\right)^{1-p}$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{x_i} \cdot \left(\frac{f(x)}{x_i}\right)^{-p} \left(\frac{f(x)}{x_j}\right)^{1-p} = \frac{1-p}{f(x)} \cdot \left(\frac{f^2(x)}{x_i x_j}\right)^{1-p}$$

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{1-p}{f(x)} \cdot \left(\frac{f^2(x)}{x_i^2}\right)^{1-p} - \frac{1-p}{x_i} \cdot \left(\frac{f(x)}{x_i}\right)^{1-p}$$

只需证明 $y^T \nabla^2 f(x)\, y = \frac{1-p}{f(x)} \cdot \left(\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) \le 0.$

考虑 Cauchy-Schwarz 不等式

$a^T b \le \|a\|_2 \cdot \|b\|_2.$ $\quad a=(a_1,\cdots a_n) \quad b=(b_1,\cdots b_n)$

令 $a_i = \left(\frac{f(x)}{x_i}\right)^{-\frac{p}{2}} \quad b_i = y_i \left(\frac{f(x)}{x_i}\right)^{1-\frac{p}{2}}$

上式得证. 故 $f(x)$ 是 concave 的.

Problem 3:

① $n=1$ 或 2 时 显然成立.

② 假设 $n=k$ 时 $f\left(\sum_{i=1}^{k} \lambda_i w_i\right) \le \sum_{i=1}^{k} \lambda_i f(w_i)$ 成立.

当 $n=k+1$ 时

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f\left(\lambda_{k+1} x_{k+1} + \sum_{i=1}^{k} \lambda_i x_i\right)$$
$$= f\left(\lambda_{k+1} x_{k+1} + (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i x_i}{1-\lambda_{k+1}}\right)$$
$$\le \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i x_i}{1-\lambda_{k+1}}\right)$$

$\because \sum_{i=1}^{k+1} \lambda_i = 1 \rightarrow \sum_{i=1}^{k} \lambda_i = 1-\lambda_{k+1} \rightarrow \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1$

故 $\cancel{f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) \le \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1})}$

故 $f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) \le \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1}) \cdot \sum_{i=1}^{k} \frac{\lambda_i f(x_i)}{1-\lambda_{k+1}}$

$$= \sum_{i=1}^{k+1} \lambda_i f(x_i)$$

综上. 得证.

Problem 4:

Problem 4

a) 设 $\Pi_x(x)=a$ $\quad \sum(a_i-x_i)^2$ 最小

$\Pi_x(y)=b$ $\quad \sum(b_i-y_i)^2$ 最小

$$\|\Pi_x(x)-\Pi_x(y)\|_2^2 - \langle \Pi_x(x)-\Pi_x(y),\ x-y \rangle$$

$$= \sum (a_i-b_i)^2 - \sum(a_i-b_i)(x_i-y_i) \le 0$$

故 $\|\Pi_x(x)-\Pi_x(y)\|_2^2 \le \langle \Pi_x(x)-\Pi_x(y), x-y \rangle$

b). $\sum(a_i-b_i)^2 - \sum(x_i-y_i)^2$

$$\le \sum(a_i-b_i)(x_i-y_i) - \sum(x_i-y_i)^2 \le 0$$

即 $\|\Pi_x(x)-\Pi_x(y)\|_2^2 \le \|x-y\|_2^2$

$\rightarrow \quad \|\Pi_x(x)-\Pi_x(y)\|_2 \le \|x-y\|_2$

Problem 5:

a). ∵ $\varphi(x)$ 是凸函数.

故 ~~$\nabla\varphi^T x$~~ $\nabla\varphi^T(x)(x-y) \geq \varphi(x)-\varphi(y)$.

令 $\begin{cases} x=y \\ y=x \end{cases}$

有 $\nabla\varphi^T(y)(y-x) \geq \varphi(y)-\varphi(x)$.

故 $\varphi(x)-\varphi(y)-\nabla\varphi^T(y)(x-y) \geq 0$.

即 $\Delta\varphi(x,y) \geq 0$.

且当 $x=y$ 时取到 0.

b). 设 $\underset{x\in C}{\arg\min} L(x) = t$

$x^* = t + \varphi(x) - \varphi(x_0) - \nabla\varphi(x_0)^T(x-x_0)$

欲证 $L(y) + \Delta_\varphi(y, x_0) \geq L(x^*) + \Delta\varphi(x^*, x_0) + \Delta_\varphi(y, x^*)$

即证 $L(y)+\varphi(y)-\varphi(x_0)-\varphi(x_0)^T(y-x_0)$

$\geq L(x^*)+\varphi(x^*)-\varphi(x_0)-\varphi(x_0)^T(x^*-x_0)$

$+\varphi(y)-\varphi(x^*)-\varphi(x^*)^T(y-x^*)$

即 $L(y)-\varphi(x_0)^T(y-x^*) \geq L(x^*)-\varphi(x^*)^T(y-x^*)$

∵ L 是 convex 且可微. 故上式成立