Optimization Methods **Fall 2020**

Homework 2

Instructor: Lijun Zhang *Name:* 黄彦骁, *StudentId:* 191300020

Notice

- The submission email is: **njuoptfall2019@163.com**.
- Please use the provided LaTeX file as a template. If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: Convex functions

a) Prove that the function $f : \mathbb{R}^n_{++} \to \mathbb{R}$, defined as

$$f(x) = -\sum_{i=1}^{n} \log(x_i),$$

is strictly convex.

Solution :

由于函数 $\log(x)$ 在 R_{++} 上为严格凹函数，所以对 $\forall x, y \in \mathbb{R}^n_{++}, 0 \le \theta \le 1$, 有

$f(\theta x+(1-\theta)y) = -\sum_{i=1}^{n} \log(\theta x_i+(1-\theta)y_i) \le -(\sum_{i=1}^{n}(\theta \log(x_i)+(1-\theta)\log(y_i))) = -(\theta\sum_{i=1}^{n}\log(x_i)+(1-\theta)\sum_{i=1}^{n}\log(y_i)) = \theta f(x) + (1-\theta)f(y)$

且等号能取到当且仅当 $\forall 1 \le i \le n, x_i = y_i$，也即 $x = y$，故函数为严格凸函数。

b) Let f be twice differentiable, with $\mathrm{dom}(f)$ convex. Prove that f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T(x-y) \ge 0,$$

for all x, y.

Solution :

必要性：如果 f 是凸函数，那么有:

$$f(y) \ge f(x) + \nabla f(x)^T(y-x), f(x) \ge f(y) + \nabla f(y)^T(x-y)$$

将两式相加即得 $(\nabla f(x)^T - \nabla f(y)^T)(x-y) \ge 0$

充分性：$(\nabla f(x)^T - \nabla f(y)^T)(x-y) \ge 0$ 表明 ∇f 是单调增函数，考虑任意两点直线上的函数 f，即 $g(t) = f(x+t(y-x))$, 则有 $g^{'}(t) = \nabla f(x+t(y-x))^T(y-x)$ 对于 t 也为单调增函数，故有 $g^{'}(t) > g^{'}(0), for\ t > 0\ and\ t \in \mathrm{dom} g$, 所以有

$$f(y) = g(1) = g(0) + \int_0^1 g^{'}(t)dt \ge g(0) + g^{'}(0) = f(x) + \nabla f(x)^T(y-x)$$

这也就表明 $f(x)$ 为凸函数。

c) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Its *perspective transform* $g : \mathbb{R}^{n+1} \to \mathbb{R}$ is defined by

$$g(x,t) = tf(x/t),$$

with domain $\mathrm{dom}(g) = \{(x,t) \in \mathbb{R}^{n+1} : x \in \mathrm{dom}(f),\ t > 0\}$. Use the definition of convexity to prove that if f is convex, then so is its perspective transform g.

Solution :

$\mathrm{dom} g = \{(x,t)|x/t \in \mathrm{dom} f, t > 0\}$ 为 $\mathrm{dom} f$ 在透视函数下的原象，所以 $domg$ 为凸集，直接利用 $Jensen$ 不等式证明

设有 $t, s > 0, \frac{x}{t}, \frac{y}{s} \in \mathrm{dom} f, 0 \leq \theta \leq 1$

$$\begin{aligned}
g(\theta x + (1-\theta)y, \theta t + (1-\theta)s) &= (\theta t + (1-\theta)s) f(\frac{\theta x + (1-\theta)y}{\theta t + (1-\theta)s}) \\
&= (\theta t + (1-\theta)s) f(\frac{\theta t \frac{x}{t} + (1-\theta)s\frac{y}{s}}{\theta t + (1-\theta)s}) \\
&= \theta t f(\frac{x}{t}) + (1-\theta) s f(\frac{y}{s}) \\
&= \theta g(x,t) + (1-\theta) f(y,s)
\end{aligned}$$

所以 $g(x,t)$ 为凸函数

Problem 2: Concave function

Suppose $p < 1, p \neq 0$. Show that the function

$$f(x) = \left(\sum_{i=1}^{n} x_i^p\right)^{1/p}$$

with dom $f = \mathbb{R}_{++}$ is concave.

Solution:

先求 $f(x)$ 的一阶导数：

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

二阶导数为：

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{f(x)}(\frac{f(x)^2}{x_i x_j})^{1-p} (i \neq j)$$

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

我们需要下列不等式成立:

$$y^T \nabla^2 f(x) y = \frac{1-p}{f(x)}((\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}})^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}) \leq 0$$

只需要在柯西施瓦茨不等式 $a^T b \leq \|a\|_2 \|b\|_2$ 中令:

$$a_i = (\frac{f(x)}{x_i})^{-\frac{p}{2}}, b_i = y_i (\frac{f(x)}{x_i})^{1-\frac{p}{2}}$$

所以 f 为凹函数

Problem 3: Convexity

Let $f : W \mapsto \mathbb{R}$ be a convex function and $\lambda_1, \ldots, \lambda_n$ be n positive numbers with $\sum_{i=1}^{n} \lambda_i = 1$. Prove that for any $w_1, \ldots, w_n \in W$,

$$f\left(\sum_{i=1}^{n} \lambda_i w_i\right) \leq \sum_{i=1}^{n} \lambda_i f(w_i). \tag{1}$$

Solution:

采用数学归纳法来证明:

Base: 当 $n = 1, 2$ 时，由 f 为凸函数可知不等式成立

InductiveHypothesis: 假设 $n = k$ 时不等式成立

InductionStep: 当 $n = k + 1$ 时，$f(\sum_{i=1}^{k+1} \lambda_i w_i) = f(\lambda_{k+1} w_i + (1 - \lambda_{k+1}) \sum_{i=1}^{k} \delta_i w_i)$，其中 $\delta_i = \frac{\lambda_i}{1-\lambda_{k+1}}$。同时有 $\sum_{i=1}^{k+1} \lambda_i = 1, \sum_{i=1}^{k} \delta_i = \frac{\sum_{i=1}^{k} \lambda_i}{1-\lambda_{k+1}} = 1$。由 *Base* 有 $f(\sum_{i=1}^{k+1} \lambda_i w_i) = \lambda_{k+1} f(w_{k+1}) + (1 - \lambda_{k+1}) f(\sum_{i=1}^{k} \delta_i w_i)$, 又由归纳假设有 $f(\sum_{i=1}^{k} \delta_i w_i) = \sum_{i=1}^{k} \delta_i f(w_i)$, 所以有 $f(\sum_{i=1}^{k+1} \lambda_i w_i) = \lambda_{k+1} f(w_{k+1}) + (1 - \lambda_{k+1}) f(\sum_{i=1}^{k} \delta_i w_i) \leq \sum_{i=1}^{k+1} \lambda_i f(w_i)$

Problem 4: Projection

For any point y, the projection onto a nonempty and closed convex set X is defined as

$$\Pi_X(y) = \operatorname*{argmin}_{x \in X} \frac{1}{2} \|x - y\|_2^2. \tag{2}$$

a) Prove that $\|\Pi_X(x) - \Pi_X(y)\|_2^2 \leq \langle \Pi_X(x) - \Pi_X(y), x - y \rangle$.

Solution:

设 $\Pi_X(x) = x_1, \Pi_X(y) = x_2$，由 x_1, x_2 使得 x, y 到 X 的距离最小，所以有 $\langle x_1 - x, w - x_1 \rangle \geq 0, \langle x_2 - y, w - x_2 \rangle \geq 0, \forall w \in X$，所以可以有 $\langle x_1 - x, x_2 - x_1 \rangle \geq 0, \langle x_2 - y, x_1 - x_2 \rangle \geq 0$，两式前后调换位置并相加即有 $\langle x_1 - x_2, x - y + x_1 - x_2 \rangle \geq 0$，也即 $\langle x_1 - x_2, x - y \rangle \geq \|x_1 - x_2\|_2^2$，不等式成立

b) Prove that $\|\Pi_X(x) - \Pi_X(y)\|_2 \leq \|x - y\|_2$.

Solution:

首先我们知道，对于欧几里德范数而言，$\langle x, y \rangle = \langle y, x \rangle$，和第一问一样设 $\Pi_X(x) = x_1, \Pi_X(y) = x_2$，现有:

$$\|x - y\|_2^2 - \|x_1 - x_2\|_2^2 = \langle (x - y) + (x_1 - x_2), (x - y) - (x_1 - x_2) \rangle$$

由第一问知:

$$\langle x_1 - x_2, (x - y) - (x_1 - x_2) \rangle \geq 0$$

而又有

$$\langle x - y, (x - y) - (x_1 - x_2) \rangle - \langle x_1 - x_2, (x - y) - (x_1 - x_2) \rangle = \|(x - y) - (x_1 - x_2)\|_2^2 \geq 0$$

所以 $\langle x - y, (x - y) - (x_1 - x_2) \rangle \geq 0$, 因此可以得到:

$$\|x - y\|_2^2 - \|x_1 - x_2\|_2^2 = \langle x - y, (x - y) - (x_1 - x_2) \rangle + \langle x_1 - x_2, (x - y) - (x_1 - x_2) \rangle \geq 0$$

所以:

$$\|x - y\|_2^2 \geq \|x_1 - x_2\|_2^2$$

Problem 5: Convexity

Let $\psi : \Omega \mapsto \mathbb{R}$ be a strictly convex and continuously differentiable function. We define

$$\Delta_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y\rangle, \quad \forall x, y \in \Omega.$$

a) Prove that $\Delta_\psi(x, y) \geq 0, \forall x, y \in \Omega$ and the equality holds only when $x = y$.

Solution:

$\Delta_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y\rangle = \psi(x) - (\psi(y) + \nabla\psi(y)^T(x - y))$, 因为 ψ 为严格凸函数且连续可微，所以有 $\psi(x) \geq \psi(y) + \nabla\psi(y)^T(x - y)$，等号成立当且仅当 $x = y$，所以有 $\Delta_\psi(x, y) \geq 0$，等号成立当且仅当 $x = y$。

b) Let L be a convex and differentiable function defined on Ω and $C \subset \Omega$ be a convex set. Let $x_0 \in \Omega - C$ and define

$$x^* = \underset{x \in C}{\arg\min}\ L(x) + \Delta_\psi(x, x_0).$$

Prove that for any $y \in C$,

$$L(y) + \Delta_\psi(y, x_0) \geq L(x^*) + \Delta_\psi(x^*, x_0) + \Delta_\psi(y, x^*). \tag{3}$$

Solution:

x^* 在 C 上最小化了 $g(x) = L(x) + \Delta_\psi(x, x_0)$，所以有 $\nabla_{x=x^*} g(x)^T(y - x) \geq 0, \forall y \in C$, 也即为:

$$\langle L^{'}(x^*) + \nabla\psi(x^*) - \nabla\psi(x_0), y - x\rangle \geq 0$$

然后再利用 L 为可微凸函数，有:

$$\begin{aligned}
L(y) &\geq L(x^*) + \langle L^{'}(x^*), y - x^*\rangle \\
&\geq L(x^*) + \langle \nabla\psi(x_0) - \nabla\psi(x^*), y - x^*\rangle \\
&= L(x^*) - \langle \nabla\psi(x_0), x^* - x_0\rangle + \psi(x^*) - \psi(x_0) + \\
&\quad \langle \nabla\psi(x_0), y - x_0\rangle - \psi(y) + \psi(x_0) - \langle \nabla\psi(x^*), y - x^*\rangle + \psi(y) - \psi(x^*) \\
&= L(x^*) + \Delta_\psi(x^*, x_0) - \Delta_\psi(y, x_0) + \Delta_\psi(y, x^*)
\end{aligned}$$

移项即可得，证明完成