作业二

最优化方法

截止时间：11 月 1 日 23:59（周三晚）

请在规定时间前提交到大夏学堂，超时得分将有折扣。
计算、证明题提交 pdf 电子版，编程题提交 Python 代码、结果及必要的解释。

1 带 ℓ_2 惩罚的部分优化问题

考虑问题

$$\min_{\beta,\,\sigma\geq 0}\ f(\beta)+\frac{\lambda}{2}\sum_{i=1}^{n}g(\beta_i,\sigma_i), \tag{1}$$

其中，f 为定义在 $\mathbb{R}^n$ 上的凸函数，$\lambda\geq 0$，且

$$g(x,y)=\begin{cases}x^2/y+y & \text{if } y>0\\ 0 & \text{if } x=0,\, y=0\\ \infty & \text{else.}\end{cases}$$

a. 证明 g 是凸函数，即上述问题为凸优化问题。（后面我们可根据此进行部分优化，且部分优化后的函数也是凸函数）

b. 证明：

$$\min_{y\geq 0}\ g(x,y)=2|x|$$

c. 证明(1)中对于 $\sigma\geq 0$ 的优化可得 ℓ_1 惩罚问题

$$\min_{\beta}\ f(\beta)+\lambda\|\beta\|_1.$$

2 Lipschitz 梯度与强凸性

令 f 为二次连续可微的凸函数

a. 证明以下命题等价：

i. ∇f 为 L-Lipschitz 函数
（即存在常数 $L>0$，使得 $\|f(x)-f(y)\|_2\leq L\|x-y\|_2$, 对任意 x,y.）

ii. 对任意 x,y，$(\nabla f(x)-\nabla f(y))^\top(x-y)\leq L\|x-y\|_2^2$

iii. 对任意 x，$\nabla^2 f(x)\preceq LI$

iv. 对任意 x,y，$f(y)\leq f(x)+\nabla f(x)^\top(y-x)+\frac{L}{2}\|y-x\|_2^2$

循环证明 $i \Rightarrow ii, ii \Rightarrow iii, iii \Rightarrow iv, iv \Rightarrow ii, iii \Rightarrow i$

b. 证明以下命题等价：

   i. f 为 m-强凸函数（即 $f(x) - \frac{m}{2}\|x\|_2^2$ 为凸函数）
   ii. 对任意 x, y，$(\nabla f(x) - \nabla f(y))^\top (x-y) \geq m\|x-y\|_2^2$
   iii. 对任意 x，$\nabla^2 f(x) \succeq mI$
   iv. 对任意 x, y，$f(y) \geq f(x) + \nabla f(x)^\top (y-x) + \frac{m}{2}\|y-x\|_2^2$

   循环证明 $i \Rightarrow ii, ii \Rightarrow iii, iii \Rightarrow iv, iv \Rightarrow i$

3 实践：使用 Pytorch 编写带回溯线搜索的梯度下降算法

- 确保你已经配置好 Python 环境，并已安装 Pytorch。
- Pytorch: https://pytorch.org
- 根据所给模板，编写带回溯线搜索的梯度下降函数，并用该函数解决以下非约束光滑优化问题：
  - Rosenbrock 函数
  $$\min_{x \in \mathbb{R}^2} (1-x_1)^2 + 100(x_2 - x_1^2)^2. \tag{2}$$
  - Beale 函数
  $$\min_{x \in \mathbb{R}^2} (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2 + (2.625 - x_1 + x_1 x_2^3)^2. \tag{3}$$
  - 线性回归问题（已给定模拟数据）
  $$\min_{\beta \in \mathbb{R}^p} \frac{1}{2n} \sum_{i=1}^n (y_i - \beta^\top x_i)^2. \tag{4}$$
  - 逻辑回归问题（已给定模拟数据）
  $$\min_{\beta \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^n \log\big(1 + \exp\big(-y_i \beta^\top x_i\big)\big). \tag{5}$$