2022年秋季《最优化方法》期末大作业

要求：请写出详细的推导过程或证明过程。

1. 考虑优化问题

$$\min_x \quad f(x) = x_1^2 + 2x_2^2 + 4x_1 + 4x_2.$$

   (a) 证明如果使用精确线搜索的最速下降梯度法从初始点$x^1 = (0,0)^\top$求解该问题，产生的点列为

$$x^{k+1} = \left(\frac{2}{3^k} - 2, \left(-\frac{1}{3}\right)^k - 1\right)^\top.$$

   (b) 这个迭代点列收敛到哪一个点？收敛速度是什么？

2. 设可微函数$f(x)$的定义域$\mathbf{dom} f = \mathbb{R}^n$，如果梯度$\nabla f(x)$是$L$-利普希茨连续，即存在$L > 0$有

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|, \quad \forall\, x, y \in \mathbf{dom} f, \tag{1}$$

   证明函数$f(x)$有二次上界：

$$f(y) \le f(x) + \nabla f(x)^{\mathrm{T}}(y - x) + \frac{L}{2}\|y - x\|^2, \quad \forall\, x, y \in \mathbf{dom} f. \tag{2}$$

3. 假设可微函数$f(x)$的梯度$\nabla f(x)$是L-利普希茨连续，且f是可微强凸函数，即存在常数$m > 0$使得对任意$x, y \in \mathbb{R}^n$和$\theta \in (0,1)$，有：

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y) - \frac{m}{2}\theta(1-\theta)\|x - y\|^2.$$

   证明:

$$(\nabla f(x) - \nabla f(y))^{\mathrm{T}}(x - y) \ge \frac{mL}{m+L}\|x - y\|^2 + \frac{1}{m+L}\|\nabla f(x) - \nabla f(y)\|^2.$$

4. 假设函数$f(x)$可微且梯度$\nabla f(x)$是L-利普希茨连续. 从初始步长$t = \hat{t} > 0$开始考虑线搜索回退法. 如果不等式

$$f(x - t\nabla f(x)) < f(x) - \frac{t}{2}\|\nabla f(x)\|_2^2 \tag{3}$$

   不满足，则缩小步长$t = \beta t$，其中$\beta \in (0,1)$，然后继续验证(3)是否满足. 证明回退法生成的步长满足$t \ge \min\{\hat{t}, \beta/L\}$.

5. 假设函数$f(x)$可微且梯度$\nabla f(x)$是L-利普希茨连续. 考虑无约束优化问题：

$$\min_{x \in \mathbb{R}^n} f(x).$$

   - 写出求解该问题的梯度法。
   - 假设$f(x)$为非凸函数，明确给出梯度法步长的取法并写出收敛性结果.
   - 假设$f(x)$为凸函数，明确给出梯度法步长的取法并写出的收敛性(复杂度分析)结果.
   - 假设$f(x)$为强凸函数，明确给出梯度法步长的取法并写出的收敛性(复杂度分析)结果.

6. 给定实对称矩阵$A \in \mathcal{S}^n$, $b \in \mathbb{R}^n$, $\Delta > 0$。考虑优化问题

$$\min_{x \in \mathbb{R}^n} \quad x^T A x + 2b^T x, \quad \text{s.t.} \quad \|x\|_2 \le \Delta.$$

给出并证明点$z \in \mathbb{R}^n$ 是该问题最优解的充要条件。

7. 考虑BFGS公式：

$$B_{k+1} = B_k - \frac{B_k s_k s_k^\top B_k}{s_k^\top B_k s_k} + \frac{y_k y_k^\top}{y_k^\top s_k}.$$

   (a) 证明:

$$\text{trace}(B_{k+1}) = \text{trace}(B_k) - \frac{\|B_k s_k\|_2^2}{s_k^\top B_k s_k} + \frac{\|y_k\|_2^2}{y_k^\top s_k}.$$

   (b) 证明：

$$\det(B_{k+1}) = \det(B_k) \frac{y_k^\top s_k}{s_k^\top B_k s_k}.$$

8. 假设x^*是最优化问题$\min_{x \in \mathbb{R}^n} \ f(x)$的最优解。假设函数$f(x)$二阶可微，海瑟矩阵$\nabla^2 f(x)$在$x^*$的一个领域内利普希茨连续且$\nabla^2 f(x^*)$严格正定。考虑牛顿算法

$$x_{k+1} = x_k + p, \quad p = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k).$$

如果初始点x_0充分靠近x^*，证明迭代序列$\{x_k\}$收敛到x^*且收敛速度为二阶。

9. 给定向量$g \in \mathbb{R}^n$, 对称矩阵$B \in \mathbb{R}^{n \times n}$ 和信赖域半径Δ，考虑信赖域子问题：

$$\min_{p \in \mathbb{R}^n} \quad m(p) = g^\top p + \frac{1}{2} p^\top B p, \quad \text{s.t.} \quad \|p\|_2 \le \Delta.$$

   (a) 写出下面问题的最优解p^s：

$$p^s = \arg\min_{p \in \mathbb{R}^n} \quad g^\top p, \quad \text{s.t.} \quad \|p\|_2 \le \Delta.$$

   (b) 写出下面问题的最优解τ：

$$\tau = \arg\min_{\tau \ge 0} \quad m(\tau p^s), \quad \text{s.t.} \quad \|\tau p^s\|_2 \le \Delta.$$

   (c) 令科西步为$p^c = \tau p^s$。证明如下不等式：

$$m(0) - m(p^c) \ge \frac{1}{2} \|g\|_2 \min \left\{ \Delta, \frac{\|g\|_2}{\|B\|_2} \right\}.$$

10. 给定矩阵$A \in \mathbb{R}^{m \times n}$和向量$b \in \mathbb{R}^m$，$\mu > 0$，考虑优化问题:

$$\min_x \quad \psi(x) = \mu \|x\|_1 + \frac{1}{2} \|Ax - b\|_2^2. \tag{4}$$

   (a) 写出并验证$\psi(x)$的次梯度表达式。给出并证明问题(4)的最优性条件。
   (b) 给出求解问题(4)的次梯度算法及其步长的取法，写出收敛性(复杂度分析)结果。
   (c) 写出邻近算子的具体表达式:

$$\text{prox}_t(y) = \arg\min_x \ t\|x\|_1 + \frac{1}{2} \|x - y\|_2^2.$$

   (d) 给出求解问题(4)的近似点梯度算法及其步长的取法，写出收敛性(复杂度分析)结果。
   (e) 给出求解问题(4)的Nesterov加速算法及其步长的取法，写出收敛性(复杂度分析)结果。

(f) 引入合适的拆分写出问题(4)的等价形式，写出求解该等价形式的增广拉格朗日函数法。要求明确写出如何求解增广拉格朗日函数子问题。

(g) 继上一小问: 写出求解该等价形式的交替方向乘子法(ADMM)。要求推导并明确写出每一个子问题的显式解。

(h) 写出问题(4)的对偶问题，写出求解该对偶问题的增广拉格朗日函数法。要求明确写出如何求解增广拉格朗日函数子问题。

(i) 继上一小问: 写出求解该对偶问题的交替方向乘子法(ADMM)。要求推导并明确写出每一个子问题的显式解。