

# 凸优化第二次上机作业

黄晃数院 1701210098

2017 年 12 月 1 日

## 1 问题

原问题为

$$\min \frac{1}{2} \|Ax - b\|_2^2 + \mu \|x\|_1 \quad (1)$$

### 1.1 cvx calling mosek or gurobi

因为primal 1是个凸问题,cvx下可以直接

$$\text{minimize}(1/2\text{sum\_square}(A * x - b) + \mu * \text{norm}(x, 1))$$

## 2 smoothed primal problem

一共使用了三种光滑函数:*huber*, *log - sum - exp*, *sqrt*,光滑化的程度由参数lambda给出.

### 2.1 gradient method

使用bb步长的梯度方法求解,光滑函数选择为sqrt,加上对 $\mu$ 的同伦: $\mu = [10, 1, 0.1, 0.01, 0.001]$ ,对应的参数 $\text{lambda} = [1e-1, 1e-2, 1e-3, 1e-5, 1e-6]$ ,因为 $\mu$ 较大时的解只是作为之后的初值,所以为了加速收敛,相应的使用较大的lambda.

## 2.2 fast gradient method

使用Nesterov加速,因为作用在光滑之后的函数上,所以可以视作

$$f(x) = g(x) + h(x), \quad h(x) = 0;$$

所以有

$$\text{prox}_h(x) = x$$

所以迭代格式为

$$\begin{aligned} y &= x_{k-1} + \frac{k-2}{k+1}(x_{k-1} - x_{k-2}) \\ x_k &= y - t \nabla f(y) \end{aligned} \quad (2)$$

**参数** 参数选择为  $\mu = [10, 0.1, 0.001]$ , 对应的参数  $\text{lambda} = [1e-2, 1e-4, 1e-6]$ . 步长选为恒定步长  $t = \frac{1}{L}$ , 为使得  $L$  满足lipschitz, 令  $L = A^T A + 2\mu$

## 3 Proximal gradient method

问题为

$$\begin{aligned} f(x) &= g(x) + h(x) \\ g(x) &= \|Ax - b\|_2^2 \\ h(x) &= \mu \|x\|_1 \end{aligned} \quad (3)$$

所以  $h(x)$  的近似点逼近为

$$\text{prox}_{th}(x)_i = \begin{cases} x_i - \mu t & x_i \geq \mu t \\ 0 & -\mu t \leq x_i \leq \mu t \\ x_i + \mu t & x_i \leq -\mu t \end{cases}$$

### 3.1 proximal gradient method for primal problem

迭代格式为

$$\begin{aligned} y &= x_{k-1} - t \nabla g(x_{k-1}) \\ x(x) &= \text{prox}_{th}(y) \end{aligned} \quad (4)$$

**参数** 加上同伦  $\mu = [10, 1, 0.1, 0.01, 1e-3]$ , 步长  $t_k = 1/L$ ,  $L = \lambda_{\max}(A^T A)$

### 3.2 fast proximal gradient method

迭代格式为

$$\begin{aligned} u &= x_{k-1} + \frac{k-2}{k+1}(x_{k-1} - x_{k-2}) \\ y &= x_{k-1} - t \nabla g(u) \\ x(x) &= \text{prox}_{th}(y) \end{aligned} \tag{5}$$

参数 加上同伦  $mu = [10, 1, 0.1, 0.01, 1e-3]$ , 步长  $t_k = 1/L$ ,  $L = \lambda_{\max}(A^T A)$

## 4 计算结果

method	cpu	err
cvx-call-mosek:	1.35	0.00e+00
gradient for smoothed problem:	1.68	1.70e-07
fast gradient for smoothed problem:	1.55	2.02e-06
proximal gradient for primal:	4.00	3.21e-06
fast proximal gradient method for primal:	2.32	2.08e-06

### 4.1 结果分析

- 四个算法都达到了要求精度
- 对同一方法,使用了Nesterovs加速技巧后,收敛速度得到了提升.尤其对于proximal gradient方法,加速效果明显
- 相较于直接求解原问题,在对原问题进行了适当的光滑化之后的求解速度更快