凸优化第三次上机作业

黄晃数院 1701210098

2018 年 11 月 26 日

1 问题

原问题为

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

1.1 cvx calling mosek or gurobi

因为primal 1是个凸问题,cvx下可以直接

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

2 算法

2.1 ALM求解对偶问题

2.1.1 对偶问题

首先将原问题转化为对偶问题,做线性变换

$$\begin{aligned} &min\ \frac{1}{2}\|y\|_2^2+\mu\|x\|_1 \\ &s.t\ y=Ax-b. \end{aligned} \tag{2}$$

对应的lagrangian函数为

$$L(x,y,\lambda)=\frac{1}{2}\|y\|_2^2+\mu\|x\|_1+\lambda^T(Ax-b-y)$$

直接可求得

$$g(\lambda) = -\frac{1}{2}\|\lambda\|_2^2 - \lambda^T b,\ \|A^T\lambda\|_\infty \leq \mu$$

所以对偶问题(D)可写作

$$\begin{aligned} min\ &\frac{1}{2}\|\lambda\|_2^2 + \lambda^T b \\ s.t\ &|A^T\lambda\|_\infty \leq \mu. \end{aligned} \tag{3}$$

2.1.2 对偶的增广lagrangian函数

引入变量s,$A^T\lambda = s, \|s\|_\infty \leq \mu$,则对应的增广lagrangian函数为

$$L_t(\lambda, s, x) = \frac{1}{2}\|\lambda\|_2^2 + \lambda^T b + x^T(A^T\lambda - s) + \frac{t}{2}\|A^T\lambda - s\|_2^2,\ s_k \leq \mu.$$

令$C = s\|\|s\|_\infty \leq mu$,那么增广lagrangian函数也可以写作

$$L_t(\lambda, s, x) = \frac{1}{2}\|\lambda\|_2^2 + \lambda^T b + I_C(s) + x^T(A^T\lambda - s) + \frac{t}{2}\|A^T\lambda - s\|_2^2$$

由增广Lagrangian方法,我们需要:对于给定的x,求s, λ极小化$L_t(\lambda, s, x)$,我们利用增广lagrangian函数的第一种表述来进行非精确的求解,即在全空间求极小,然后在C上对s做投影.利用一阶条件,这等价于

$$s_i = \begin{cases} \mu\ \phi(\lambda, x)_i > \mu \\ -\mu\phi(\lambda, x)_i < -\mu \\ phi(\lambda, x) else \end{cases}$$

$$phi(\lambda, x) = A^T\lambda + \frac{x}{t}, \lambda = (I + tAA^T)^{-1}(tAs - Ax - b)$$

显式的按上列顺序进行N次迭代的结果作为s^+, λ^+,然后更新x

$$x^+ = x + t(A^T\lambda - s)$$

2.2 ADMM for DUAL

转化为对于初值x, λ,求s极小化增广lagrangian,然后对x, s^+,求λ极小化增广lagragian,然后更新x.

转化为算法,等价于上面ALM的N次迭代直接取值为1即可.

2.3 ADMM for linearization primal problem

对于原问题,其增广lagrangian函数为

$$L(x,y,\lambda)=\frac{1}{2}\|y\|_2^2+\mu\|x\|_1+\frac{\beta}{2}\|Ax-b-y\|_2^2+z^TAx-z^Ty-z^Tb$$

ADMM迭代为

$$\begin{aligned}x=&arg\min_x \mu\|x\|_1+\frac{\beta}{2}\|Ax-b-y\|_2^2+z^TAx\\ y=&arg\min_y \frac{1}{2}\|y\|_2^2+\frac{\beta}{2}\|Ax-b-y\|_2^2+-z^Ty\\ z=&z+\beta(Ax-y-b).\end{aligned}$$

在求解关于x的极小化问题时,因为有一范数项,选择通过线性化的方式求解,即取

$$g=\beta A^T(Ax-y-b)+z^TA$$

解

$$\begin{aligned}x=&arg\min_x \mu\|x\|_1+x^Tg+\frac{1}{2t}\|x-x_k\|_2^2\\ =&arg\min_y t\mu\|x\|_1+x^Tg+\frac{1}{2}\|x-(x_k-tg)\|_2^2\\ =&Prox_{t\mu\|\cdot\|_1}(x_k-tg)\end{aligned}$$

其中$\frac{1}{t}=\beta\lambda_{max}(A^TA)$为二阶导的近似值.

关于y,z的更新,仍然为

$$\begin{aligned}y=&\frac{\beta(Ax-y-b)+z}{\beta+1}\\ z=&z+\beta(Ax-y-b).\end{aligned}$$

实验证明该方法需要加上同伦才能较好收敛

3 数值结果

method	cpu	err
cvx-call-mosek:	1.12	0.00e+00
ALM for dual problem:	0.08	3.72e-06
ADMM for dual problem:	0.07	3.62e-06
linearization for primal:	0.79	2.41e-06

3.1 结果分析

- 三个算法都达到了要求精度
- ALM与ADMM方法有极好的收敛速度,符合预期
- ALM相较于ADMM方法虽然每次多了10步子迭代,然而对于合适的参数而言,效果并不比后者差
- 对原问题线性化之后收敛速度不如前两种方法,由于线性化损失了精度,需要同伦方法弥补,所以耗时更多是可以预料的