凸优化 · 课程项目二

姓名：陈嘉豪
学号：1300010741

2017 年 1 月 9 日

1 LP 问题

LP 问题的标准形式如下

$$
\begin{aligned}
\min_{x\in R^n} \quad & c^T x \\
\text{s.t.} \quad & Ax = b \\
& x \geq 0
\end{aligned}
\tag{1}
$$

等价于对偶形式如下

$$
\begin{aligned}
\min_{y\in R^m, s\in R^n} \quad & -b^T y \\
\text{s.t.} \quad & A^T y + s = c \\
& s \geq 0
\end{aligned}
\tag{2}
$$

2 第一题

2.1 问题描述

(a) 用 Augmented Lagrangian Method 解问题 (2)，使得变量 s 不出现在算法的更新过程中。

(b) 用梯度法极小化子问题中的 Lagrangian 函数

(c) 用 Semi-Smooth Newton Method 极小化子问题中的 Lagrangian 函数 [2]

2.2 (a) Augmented Lagrangian Method

对偶问题 (2) 等价于如下形式 (用 $-y$ 代替 y，消去变量 s)：

$$
\begin{aligned}
\min_{y\in R^m, s\in R^n} \quad & b^T y \\
\text{s.t.} \quad & A^T y + c \geq 0
\end{aligned}
$$

参考 [2](2)，有 Augmented Lagrangian Function 如下：

$$
L_\sigma(y, x) = b^T y + \frac{1}{2\sigma}\left(\left\|\prod\left(x - \sigma(A^T y + c)\right)\right\|^2 - \|x\|^2\right) \tag{3}
$$

其中 $\prod$ 为投影算子，满足 $(\prod(x))_i = \begin{cases} x_i, & x_i > 0 \\ 0, & else \end{cases}$。Augmented Lagrangian Method 按照如下步骤进行迭代得到序列 $x^{(k)}$ 和 $y^{(k)}$ (参照 [2](3))：

$$
\begin{aligned}
& y^{(k)} = \arg\min_y L_{\sigma^{(k-1)}}(y, x^{(k-1)}) \\
& x^{(k)} = \prod\left(x^{(k-1)} - \sigma^{(k-1)}(A^T y + c)\right) \\
& \sigma^{(k)} = \rho\sigma^{(k-1)} \text{ or } \sigma^{(k)} = \sigma^{(k-1)}
\end{aligned}
\tag{4}
$$

2.3 (b) 用梯度下降法求解子问题

为了解决迭代过程中 $y^{(k)}$ 的更新问题，考虑采用梯度下降法。其中 $\phi(y) = L_\sigma(y, x)$ 对 y 的梯度为

$$\nabla\phi(y) = b - A\prod(x - \sigma(A^T y + c)) \tag{5}$$

用梯度下降法得到序列 $z^{(1)} = y^{(k-1)}, z^{(j+1)} = z^{(j)} - \text{step} * \nabla\phi(z^{(j)})$，直到小于给定常数 ϵ。

2.4 (c) 用 Semi-Smooth Newton Method 求解子问题

为了解决迭代过程中 $y^{(k)}$ 的更新问题，考虑采用 Semi-Smooth Newton Conjugate Gradient Method [2]。

2.5 数值实验结果

3 第二题

3.1 问题描述

Semi-Smoothed Method 基于求解一个不动点方程

(a) 用 Douglas-Rachford Splitting (DRS) 解问题 (1)，用 ADMM 解 (2)

(b) 推导上述 DRS 和 ADMM 中的变量的显式关系

(c) 用正则化的 Semi-Smooth Newton Method 求解问题 (2) [1]

3.2 DRS 与 ADMM

给定如下问题

$$\min \quad f(x) = g(x) + h(x)$$

Dougla-Rachford Splitting (DRS) 算法按照如下格式进行迭代

$$\begin{cases} y^+ = \text{prox}_{tg}(x - w) \\ x^+ = \text{prox}_{th}(y^+ + w) \\ w^+ = w + y^+ - x^+ \end{cases} \tag{6}$$

其中 $\text{prox}_h(x) = \arg\min_u \left(h(u) + \frac{1}{2}\|u - x\|_2^2\right)$。

给定如下问题

$$\begin{aligned} \min_{y,z} \quad & f(y) + g(z) \\ \text{s.t} \quad & Ay + Bz = b \end{aligned}$$

和 Augmented Lagrangian Function 如下（x 为乘子）：

$$L(y,z,x)=f(y)+g(z)+\beta x^T(Ay+Bz-b)+\frac{\beta}{2}\|Ay+Bz-b\|_2^2 \tag{7}$$

Alternating Direction Method of Multipliers (ADMM) 算法按照如下格式进行迭代

$$\begin{cases} y^+ & =\arg\min_y L(y,z,x) \\ & =\arg\min_y \left(f(y)+\beta x^T Ay+\frac{\beta}{2}\|Ay+Bz-b\|_2^2\right) \\ z^+ & =\arg\min_z L(y^+,z,x) \\ & =\arg\min_z \left(g(z)+\beta x^T Bz+\frac{\beta}{2}\|Ay^++Bz-b\|_2^2\right) \\ x^+ & =x+(Ay^++Bz^+-b) \end{cases} \tag{8}$$

当原问题可分，

$$\begin{array}{ll} \min & f_1(x_1)+f_2(x_2) \\ \text{s.t.} & A_1x_1+A_2x_2=b \end{array}$$

转化为对偶问题即

$$\max \quad -b^Tz-f_1^*(-A_1^Tz)-f_2^*(-A_2^Tz)$$

令 $g(z)\triangleq b^Tz+f_1^*(-A_1^Tz), h(z)\triangleq f_2^*(-A_2^Tz)$，对该对偶问题使用 DRS (6)，等价于如下形式，

$$\begin{aligned} //y^+ &= \operatorname{prox}_{tg}(z-w) \\ \hat{x_1}^+ &= \arg\min_{x_1}\left(f_1(x_1)+z^T(A_1x_1-b)+\frac{t}{2}\|A_1x_1+A_2\hat{x_2}-b\|_2^2\right) \\ y^+ &= z-w+t(A_1\hat{x_1}^+-b) \\ //z^+ &= \operatorname{prox}_{th}(y^++w)=\operatorname{prox}_{th}(z+t(A_1\hat{x_1}^+-b)) \\ \hat{x_2}^+ &= \arg\min_{x_2}\left(f_2(x_2)+z^TA_2x_2+\frac{t}{2}\|A_1\hat{x_1}^++A_2x_2-b\|_2^2\right) \\ z^+ &= z+t(A_1\hat{x_1}^++A_2\hat{x_2}^+-b) \\ //w^+ &= w+y^+-z^+ \\ w^+ &= -tA_2\hat{x_2}^+ \end{aligned} \tag{9}$$

将上述迭代与 (8) 比较可知，将 DRS 用于对偶问题等价于将 ADMM 用于原问题。而且 $\beta=t$，上式中的 $\hat{x}_1,\hat{x}_2$ 分别等于 ADMM 中的两个变量 y,z；目标变量也是一致的，上式中 z 的更新形式与 (8) 中 x 的更新形式完全一致。

3.3 用 ADMM 解对偶问题

首先给出 ADMM 用于解对偶问题的方式，将对偶问题 (2) 转化为如下形式

$$\begin{array}{ll} \min_{y,s} & -b^Ty+I_{s\geq 0} \\ \text{s.t.} & A^Ty+s=c \end{array} \tag{10}$$

将 ADMM 算法 (8) 应用到该问题中，得到如下迭代。

$$\begin{cases} y^+ & = (AA^T)^{-1}(\frac{1}{\beta}b + A(c - s - x)) \\ s^+ & = \prod(c - A^T y^+ - x) \\ x^+ & = x + (A^T y^+ + s^+ - c) \\ & = s^+ - (c - A^T y^+ - x) \geq 0 \end{cases} \tag{11}$$

3.4 用 DRS 解原问题

为了给出 DRS 和 ADMM 变量之间的显式关系，参考3.2中的迭代形式9，有如下的对照表

$f_1(x)$	$f_2(x)$	x_1	x_2	A_1	A_2	b	β
$-b^T x_1$	$I_{x\geq 0}$	$y_{(\text{ADMM})}$	$s_{(\text{ADMM})}$	A^T	I	c	t

表 1: 对照表

依照该表得到对应的 g, h 如下

$$\begin{aligned} g(x) &= b^T x + f_1^*(-A_1^T x) \\ &= c^T x + I_{Ax=b} \\ h(x) &= f_2^*(-A_2^T x) \\ &= I_{x\geq 0} \end{aligned} \tag{12}$$

其中 $f_1^*(y) = \sup_x(y^T x + b^T x) = I_{y=-b}$，$f_2^*(y) = \sup_x(y^T x - I_{x\geq 0}) = I_{y\leq 0}$。于是有原问题的等价形式如下

$$\min \quad g(x) + h(x) = c^T x + I_{Ax=b} + I_{x\geq 0}$$

对该形式下的原问题使用 DRS 等价于将 ADMM 用于对偶问题，参照迭代格式 (9) 与表 (1) 可知，DRS 的迭代格式、变量与 (11) 中变量的关系如下

$$\begin{cases} //\hat{x}_1 & = y_{(\text{ADMM})} \\ //\hat{x}_2 & = s_{(\text{ADMM})} \\ y^+ & = x - w + t(A^T \hat{x}_1{}^+ - c) \\ x^+ & = x + t(A^T \hat{x}_1{}^+ + \hat{x}_2{}^+ - c) \\ w^+ & = -\beta \hat{x}_2{}^+ \end{cases} \tag{13}$$

当 $\beta = t = 1$ 时，可以看到 (11) 和 (13) 中的目标变量的更新完全一致。

3.5 数值实验结果

4 数值实验结果对比

参考文献

[1] Xiantao Xiao, Yongfeng Li, Zaiwen Wen, and Liwei Zhang. A regularized semi-smooth newton method with projection steps for composite convex programs, 2016.

[2] Xin Yuan Zhao, Defeng Sun, and Kim Chuan Toh. A newton-cg augmented lagrangian method for semidefinite programming. *Siam Journal on Optimization*, 20(4):1737–1765, 2010.