文章编号: 2095-2163(2020)04-0294-03　　中图分类号: O229　　文献标志码: A

# 求解 LASSO 问题的广义对称交替方向乘子算法

蒋　峰, 党亚峥, 何泽秀
(上海理工大学 管理学院, 上海 200093)

摘　要: 随着数据规模的增加,有效地解决 LASSO 问题面临巨大挑战。对称交替方向乘子法是求解 LASSO 问题的一种有效方法,其将原问题分解为多个子问题交替求解,很大程度上提升了求解 LASSO 问题的效率。本文提出了一种广义对称交替方向乘法。与对称交替方向乘子法相比,该算法引入了一个半近邻项近似地求解 $x$ 子问题,克服了之前算法的不足。此外,算法中引入了松弛算子进一步提升了算法的效率,数值实验说明了该算法是有效的。

关键词: LASSO 问题; 广义对称交替方向乘子法; 松弛算子

# Generalized symmetric alternating direction method of multipliers for LASSO problem

JIANG Feng, DANG Yazheng, HE Zexiu
(Business school, University of Shanghai for science and technology, Shanghai 20093, China)

【Abstract】As the size of data increases, the effective solution to the Lasso problem faces enormous challenges. The symmetric alternating direction method of multipliers (SADMM) is an effective method for solving the Lasso problem, which decomposes the original problem into several sub-problems. This greatly improves the efficiency of solving Lasso problems. However, in many practical applications, it is costly to solve the sub-problem accurately or it is difficult to obtain the solution of the sub-problem. To this end, this paper proposes a generalized symmetric alternating direction method of multipliers. Compared with the symmetric alternating direction multiplier method, the algorithm introduces a semi-proximal term to approximate the x subproblem, which overcomes the shortcoming of the previous algorithm. In addition, the relaxed factor is introduced in the algorithm to further improve its efficiency. Finally, numerical experiments show that the proposed algorithm is effective.

【Key words】Lasso problem; generalized symmetric ADMM; relaxed factor

## 0　引　言

随着大数据时代的到来,数据收集技术有了巨大的发展,如何有效地从大规模数据中挖掘出潜在的有用信息成为了新的研究热点。线性模型是当前比较常用的数据挖掘手段,其理论丰富,应用广泛,在农业、工程技术、经济、管理、工业、生物等领域取得了长足的发展[1]。线性模型应用的范围越广,模型的准确性就显得越重要。就线性模型而言,其准确性主要取决于变量的选择和回归系数的取值。受 Ridge Regression 和 Nonnegative Garrote 算法的启发,Tibshirani 提出了一种新的变量选择方法,即最小绝对收缩和选择算子(Least absolute shrinkage and selection operator, LASSO)[2]。LASSO 问题包含平方误差的求和和正则化项。L1 正则化项能够避免模型过拟合问题,使得模型更加稳定[3]。

解决 LASSO 问题的方法很多,Efron 提出了最小角回归算法求解 Lasso 问题。该算法通过消除回归系数的异号情况得到 LASSO 问题的解[2]。陈善雄等人考虑各个变量在模型中所占比重不同,提出了多维权重求解算法[3]。该算法在最小回归算法的基础上加入主成分分析、独立权数法、基于 Intercriteria 相关性指标的重要性评价法这 3 种权重评价方法,优化变量的选择。数值实验表明该方法进一步提升了求解 LASSO 问题的准确性。在解决大规模数据问题中,交替方向乘子法(ADMM)得到广泛应用。ADMM 算法将原始的凸优化问题分解为多个子问题进行分别求解,最后再计算拉格朗日乘子[4]。这种方法在很大程度上减少了计算成本,提升了求解 LASSO 问题的效率。

虽然 ADMM 算法求解大数据背景下的 LASSO 问题速度较快,但是在很多实际应用中, 精确地求解子问题难以实现, 或者需要很大代价。因此本文提出了一种广义对称 ADMM 算法。该算法的主要创新之处是在对称交替方向乘子法的 $x$ 极小化子问

基金项目: 河南省科技攻关项目(172102310252)。
作者简介: 蒋　峰(1994-),男,硕士研究生,主要研究方向:机器学习; 党亚峥(1973-),女,副教授,主要研究方向:优化算法、智能电网; 何泽秀(2002-),女,本科生,主要研究方向:金融。
通讯作者: 党亚峥　　Email:jgdyz@ 163.com
收稿日期: 2019-10-27

题中加入半近邻项近似求解此问题。此外,为进一步提高算法的收敛速度,引入了松弛算子。通过具体的数值实验表明,广义对称交替方向乘子法收敛速度比对称交替方向乘子法更快。

## 1　LASSO 问题及广义对称 ADMM 算法

### 1.1　LASSO 问题描述

机器学习中的凸优化问题一般可以表示为

$$\min\left\{P(x)=\sum_{i=1}^{n}\varphi_i(a_i^T x)+ug(x)\right\}. \quad (1)$$

令 $\sum_{i=1}^{n}\varphi_i(a_i^T x)=\frac{1}{2}\|Px-b\|_2^2, g(x)=\|x\|_1$,得到 $l_1$ 线性回归的特殊形式,即 Lasso 问题。

$$\min\left\{\frac{1}{2}\|Px-b\|_2^2+u\|z\|_1 \mid x-z=0\right\}. \quad (2)$$

其中,$x\in R^n, P\in R^{m\times n}$,$u$ 是正则化参数。

### 1.2　广义对称 ADMM 算法

ADMM 算法由 Glowinski 等人提出[5],主要用于求解形如式(3)的凸优化问题

$$\min\{f(x)+g(z) \mid Ax=z, x\in X, z\in Z\}, \quad (3)$$

其中,$A\in R^{m\times n}$,$X\subset R^n$、$Z\subset R^m$ 是给定闭凸集;$f: R^n\to R\cup\{+\infty\}$ 和 $g: R^m\to R\cup\{+\infty\}$ 为凸函数。

本文 $\boldsymbol{A}$ 为单位矩阵,其迭代格式如下:

$$\begin{cases} x^{k+1}=\operatorname{argmin}\{f(x)-(y^k)^T(Ax-z^k)+\\ \qquad \frac{\gamma}{2}\|Ax-z^k\|^2\},\\ z^{k+1}=\operatorname{argmin}\{g(y)-(y^k)^T(A x^{k+1}-z)+\\ \qquad \frac{\gamma}{2}\|A x^{k+1}-z\|^2\},\\ y^{k+1}=y^k-\gamma(Ax^{k+1}-z^{k+1}). \end{cases} \quad (4)$$

其中,$y$ 为拉格朗日乘子,$\gamma$ 是惩罚参数。由式(3)可以看出,ADMM 算法先求解 $x$ 最小化问题,再求解 $z$ 最小化问题,最后再更新拉格朗日乘子 $y$。

在文献[6]中,对称 ADMM 算法被用于求解 LASSO 问题。然而,在大数据背景下,精确地求解(3)中的 $x$ 子问题难以实现,或者需要很大代价。因此本文提出了一种广义对称 ADMM 算法。

$$\begin{cases} x^{k+1}=\operatorname{argmin}\{f(x)-(y^k)^T(Ax-z^k)+\\ \qquad \frac{\gamma}{2}\|Ax-z^k\|^2\},\\ y^{k+\frac{1}{2}}=y^k-\gamma(A x^{k+1}-z^k),\\ z^{k+1}=\operatorname{argmin}\{g(z)-(y^{k+\frac{1}{2}})^T(\alpha A x^{k+1}+(1-\\ \qquad \alpha)z^k-z)+\frac{\gamma}{2}\|\alpha A x^{k+1}+(1-\alpha)z^k-z\|^2\},\\ y^{k+1}=y^{k+\frac{1}{2}}-\gamma(\alpha A x^{k+1}+(1-\alpha)z^k-z^{k+1}). \end{cases} \quad (5)$$

其中,$\alpha\in(0,2]$ 是松弛算子。当 $\alpha=1$ 时,(1.5)变为对称 ADMM 算法。

## 2　数值实验

本节将广义对称 ADMM 算法应用于 LASSO 问题。在进行数值实验时,选取矩阵 $\boldsymbol{P}$ 并且规范化 $\boldsymbol{P}$ 中所有列。取 $a$ 和 $\sigma$ 为随机向量,$b=Pa+\sigma$,正则化参数 $u=0.01\|P^Tb\|$。取 $m=2\ 500, \alpha=0.8$,惩罚参数 $\gamma=1$,矩阵 $\boldsymbol{G}=\frac{1}{2}\gamma I_n$。此外,向量 $x$, $y$, $z$ 的初始值皆为零向量。

算法的收敛准则参照文献[6]。定义初始残差 $p^k=\gamma\|z^k-z^{k+1}\|$ 和对偶残差 $d^k=\frac{1}{\gamma}\|u^k-u^{k+1}\|$。对于改进的对称 ADMM 算法,其收敛准则为

$$\max\{r^k, p^k\}\leqslant\sqrt{n}\delta.$$

其中,$\delta=10^{-4}$。

表 1 为应用对称 ADMM 算法和广义对称 ADMM 算法解决该问题的结果,其中 *Iter* 表示迭代次数,*CPU*(s) 表示 CPU 时间(以秒为单位),$n$ 表示矩阵 $\boldsymbol{P}$ 的维数。从表 1 可以看出,本文提出的算法快于对称 ADMM 算法。

**表 1　LASSO 问题数值结果**

**Tab. 1　Numerical results for LASSO**

| 算法 | $n$ | *Iter* | *CPU*/ s |
|---|---|---|---|
| 对称 ADMM | 5 000 | 15 | 4.04 |
| 广义对称 ADMM | 5 000 | 10 | 2.76 |

为了进一步观察 2 种算法的收敛性,比较初始残差和对偶残差随迭代次数的变化情况(纵轴分别是初始残差和对偶残差,横轴是迭代次数)。从图 1、图 2 中可以直观地发现,尽管在算法迭代的某些阶段,对称 ADMM 算法的初始残差、对偶残差减小更快,但是本文提出的算法先于对称 ADMM 算法收敛,因此广义对称 ADMM 算法更高效。

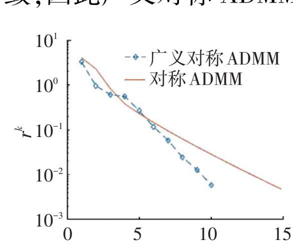


**图 1　对偶残差的变化情况**

**Fig. 1　Evolution of primal residual**

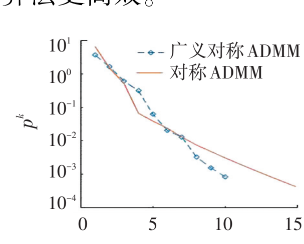


**图 2　初始残差的变化情况**

**Fig. 2　Evolution of dual residual**

## 3　结束语

本文提出了一种求解 LASSO 问题的广义对称 ADMM 算法。该方法的基本思想是在 $x$ 子问题中加

(下转第 298 页)