Lasso的精确解

夏慧异　著

辽宁大学出版社
Liaoning University Press

资助项目：安徽省高校自然重点项目（KJ2017A578）

Lasso的精确解

◆ 夏慧异　著

辽宁大学出版社
Liaoning University Press

图书在版编目（CIP）数据

Lasso 的精确解/夏慧异著. —沈阳：辽宁大学出版社，2018.6

ISBN 978-7-5610-9190-6

Ⅰ.①L… Ⅱ.①夏… Ⅲ.①有偏估计 Ⅳ.①O121.1

中国版本图书馆 CIP 数据核字（2018）第 084984 号

Lasso 的精确解

Lasso DE JINGQUE JIE

出 版 者：辽宁大学出版社有限责任公司
（地址：沈阳市皇姑区崇山中路 66 号 邮政编码：110036）
印 刷 者：三河市华晨印务有限公司
发 行 者：辽宁大学出版社有限责任公司
幅面尺寸：170mm×240mm
印 张：9.5
字 数：150 千字
出版时间：2018 年 6 月第 1 版
印刷时间：2018 年 6 月第 1 次印刷
责任编辑：刘 葵
封面设计：优盛文化
责任校对：齐 悦

书 号：ISBN 978-7-5610-9190-6
定 价：35.00 元

联系电话：024-86864613
邮购热线：024-86830665
网 址：http://press.lnu.edu.cn
电子邮件：lnupress@vip.163.com

前言

美国科学院院士、斯坦福大学教授 Robert Tibshirani 在 1996 年提出了 lasso（最小绝对收缩和选择算子），美国科学院院士 Efron 等人在 2004 年提出了 lars（最小角回归）算法，并且指出可以用修改后的 lars 算法解决 lasso 问题。由于 lasso 的适用面宽，lars 的计算速度快，使得 lasso 风靡全球，提出 lasso 问题和 lars 算法的这两篇文章是统计学引用率非常高的文章。

Lasso 问题本质上是一个优化问题，Tibshirani 在 1996 年用一种优化的算法（二次规划算法）解决了 lasso 问题，但这种算法的计算速度很慢，因而提出 lasso 问题的文章在发表后，引用率并不高。Efron、Tibshirani 等人在 2004 年提出了 lasso 的 lars 快速算法，使得 lasso 方法迅速被各个学科引用。

Lasso 原有的优化算法（二次规划算法）和 lasso 的快速算法（lars）的关系是具有较高研究价值的问题，如果能找到这两种算法的替代规律，并进行推广，应该会有非常高的学术价值和应用价值。

但是，由于高维数据的复杂性，这两种算法的替代规律很难看清。尽管美国科学院院士 Efron、Tibshirani 等人认为用 Donoho 等人的小波理论可以解释快速算法 lars 的正确性，但小波理论本身就非常抽象，很难理解。

2010 年作者进入上海财经大学攻读博士，当时作者的报考导师博士生导师王小明教授正在研究 lasso 问题。为了搞清 lasso 问题解的结构，Tibshirani 在

文中提出了 lasso 的几何，指出当数据是二维或三维时，lasso 的解可以用几何图形表示。当数据是二维时，lasso 问题就是一个变化的椭圆线和变化的方形区域的交点问题，当变化的椭圆线和变化的方形区域交在方形区域顶点的时候，就进行了变量选择。但是，为什么椭圆线和方形区域一定能交在方形区域的顶点，这个问题难住了我们。我们仔细研读原文，没有找到解决问题的方法。

为了解决 lasso 几何中存在的问题，作者研究了正交设计下，lasso 的几何问题。当数据是二维时，lasso 的几何问题变成一个变化的圆和变化的方形区域的交点问题，解决了这个问题，就是解决了当数据是二维时，正交设计下 lasso 精确解的问题。这一突破非常重要，使得作者找到了一种不用小波理论，直接用数学分析的方法解决 lasso 问题的办法，同时对数据维数超过三维时的几何问题进行了研究，提出用超平面的方法进行研究的猜想。经过进一步的研究，作者运用拉格朗日乘数法，在正交设计下，构造出 lasso 的精确解。

1996 年，Tibshirani 利用小波理论，在正交设计下，给出了 lasso 的软门限估计，这种估计依赖调节参数和软门限这两个变量，这两个变量的关系很难看清，通过作者构造的 lasso 的精确解，发现调节参数和软门限这两个变量的关系是分段函数的关系。作者构造的 lasso 的精确解只依赖调节参数，更加清楚易懂。从这个角度来说，lasso 的精确解改进了 Tibshirani 在 1996 年提出的 lasso 的软门限估计。

在正交设计下，用 lars 算法结合软门限估计可以得到 lasso 问题的解。作者用 lasso 的精确解和 lasso 的软门限估计进行了比较，发现两者是一致的，进一步验证了作者在正交设计下构造的 lasso 精确解的正确性。

Lasso 问题本质上是一种复杂的向后选择的算法，而 lars 算法是一种向前选择的算法。正交设计下，lasso 问题变得简单，lasso 问题的变量选择只依赖最小二乘估计，在这种情况下，向前选择算法和向后选择算法选择的变量一

致，这时 lars 算法选择的变量和 lasso 选择的变量一致。

Tibshirani 指出，通过投影等理论结合正交设计下的 lasso 问题可以解决一般设计下的 lasso 问题，但这种理论用到 lars 算法理论中显然存在问题，用 lars 算法对房地产数据进行分析，会发现其解决 lasso 问题的效果不理想，因而作者仔细研究 lasso 的文献，在研究过程中发现了斯坦福大学教授的一个思维漏洞。

2013 年 12 月，作者在北京大学参加学术会议时，斯坦福大学终身教授、冯・诺依曼理论奖获得者叶荫宇教授在作学术报告时，提到他的老师、单纯形法算法的创始人、美国科学院院士 G.B.Dantzig 坚信自己的优化算法，不认可替代单纯形法的快速算法，这说明优化问题的算法之间存在争论。而 lars 算法是解决优化问题的一种算法，是一种替代二次规划算法的快速算法，对这种算法的算法理论进行研究非常有价值。

为了让更多的人能看懂作者的研究，说服更多的人相信 lars 算法不能解决 lasso 问题，作者用反证法成功构造出一个反例，通过这个反例可以清楚地看到，用 lars 算法解决 lasso 问题是不正确的。同时，对与 lasso 关系比较近的弹性网（the elastic net）的几何和解的结构进行了研究，指出弹性网问题研究的复杂性。

作者在研究中发现了正交设计的一个缺陷，就是在正交设计下，研究极限性质容易产生共线性，在研究数据计算问题时，必须考虑问题的条件。由于计算机的发展，使得高维数据的计算变成了学术研究的一个热点问题。作者另辟蹊径，用数学分析的方法研究高维数据模型解的结构，解决数据计算中存在的问题。作者用数学分析的方法进一步研究了斯坦福大学教授、美国科学院院士 Tibshirani 在 1996 年提出的 lasso 的几何、lasso 的软门限估计，使得 lasso 的几何理论更加通俗易懂，对正交设计下的 lasso 的精确解进行研究，给出了正交

设计下，lasso 精确解的具体数学表达式，并与 lasso 的软门限估计进行了比较。

这些研究改变了现今 lasso 问题的研究方式，使得 lasso 的研究者们重新重视 lasso 问题的数学化工作。目前，lasso 问题被广泛应用到大数据分析中，很多应用学科运用 lasso 的方法解决问题，有许多相关论文和其他研究成果。由于 lars 算法解决了 lasso 问题的存在问题，因而这些应用学科必须对用到 lasso 的研究成果进行重新思考。有些学科的数据精算只能计算到数十维，但有些学科的数据精算能计算到近千维，这个问题一直困扰着从事数据计算的研究者，作者现在的研究应该对这个问题的解决有很大帮助。

目 录

第一章 Lasso的研究现状

由于“维数灾难”的存在，为了进行高维数据的计算，则必须降低数据的维数。变量选择和模型选择是一种重要的降低数据维数的方法，是统计学研究数据的重要方法。目前，变量选择和模型选择的方法有全局择优法，如 BIC 准则 (Bayesian information criterion)、AIC 准则（Akaike information criterion）、C_p 准则等一些信息准则；变量选择方法有逐步选择法，如前进法（forward selection）、逐步回归法（stepwise regression）、后退法（backward elimination）等。Lasso 就是一种重要的变量选择和模型选择的算法。

美国科学院院士、斯坦福大学教授 Tibshirani 于 1996 年在杂志《Journal of the Royal Statistical Society》上发表了论文《Regression Shrinkage and Selection via the Lasso》，提出了 lasso 方法，这是一种变量选择和系数估计的新方法，但 lasso 方法的引用率并不高。Efron 和 Tibshirani 等人于 2004 年在杂志《The Annals of Statistics》上发表了论文《Least Angle Regression》，提出了 lasso 的快速算法 lars。由于这些杂志很有权威性，斯坦福大学教授、美国科学院院士的影响力又很大，同时，lasso 的应用面宽，使得这种方法被运用到很多应用学科，现在这两篇文章仍然是统计学引用率极高的文章。

2016 年，斯坦福大学教授、美国科学院院士 Hastie 教授在上海财经大学举行的“第四届国际生物统计学大会”上作了《Statistical Learning with Big Data》主题报告，就大数据、系统发育等国际生物前沿提出了问题和解决办法，Hastie 等人把 lars 算法和 lasso 方法运用到大数据的计算中，使得 lasso 方法和 lars 算法对生物学科产生了重要影响。

第一节 Lasso 研究的起源

一、Lasso 问题的算法理论

美国科学院院士、斯坦福大学教授 Robert Tibshirani 在 1996 年提出了 lasso，lasso 问题是在系数的绝对值小于一个常数的条件下使得最小化残差的平方和最小的问题。具体的数学模型如下：

当$\sum_{j=1}^{p}\left|\beta_j\right|\leqslant t$时，

$$\hat{\beta}=\arg\min\left\{\sum_{i=1}^{N}\left(y_i-\alpha-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2\right\}$$

为了解决 lasso 问题，Tibshirani 在正交设计下，运用美国著名数学家 Donoho 和 Johnstone 的小波理论和软门限理论，给出了 lasso 的软门限估计，具体表达式如下：

$$\hat{\beta}_j=\operatorname{sign}\left(\hat{\beta}_j^0\right)\left(\left|\hat{\beta}_j^0\right|-\gamma\right)^+$$

软门限γ是由$\sum\left|\hat{\beta}_j\right|=t$决定的，其中$z^+=\begin{cases}z, & z>0\\ 0, & z\leqslant 0\end{cases}$。

为了解释 lasso 解的结构，Tibshirani 在提出“lasso”的文章中提出了 lasso 的几何，lasso 的几何其实就是当数据是二维和三维时，lasso 问题的解。

为了解决 lasso 问题的计算，Tibshirani 给出了一个优化的算法，具体算法步骤如下：

去掉绝对值号后，lasso 问题变成一个有2^p不等式约束的最小二乘问题。Lawson 和 Hansen 在 1974 年提供了一个程序，解决在一般线性不等式限制$G\beta\leqslant h$下的线性最小二乘问题，G是$m\times p$的矩阵，对应于m个p维向量β的不等式约束，由于$m=2^p$数量很大，直接用编程计算不实际，但可以通过引入不等式约束寻求满足 Kuhn Tucker 条件的可行解来解决这个问题。

具体过程如下：

令 $g(\beta)=\sum_{i=1}^{N}\left(y_i-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2$，令 δ_i， $i=1,2,\cdots,2^p$，令 $\delta_i=(x_1,\cdots,x_p)$，$x_i=1$ 或者 $x_i=-1$，$i=1,\cdots,p$，$\sum_{j=1}^{p}\left|\beta_j\right|\leqslant t$ 等价于 $\delta_i^T\beta\leqslant t,\ i=1,2,\cdots,2^p$。

对给定的 β，令 $E=\left\{i:\delta_i^T\beta=t\right\}$ 和 $S=\left\{i:\delta_i^T\beta<t\right\}$，$G_E$ 表示行为 δ_i，$i\in E$，令向量 l 的元素都是 1，元素的个数和 G_E 行数相等。

下面的算法开始于 $E=\{i_0\}$，这里 $\delta_{i_0}=\operatorname{sign}\left(\hat{\beta}\right)$，$\hat{\beta}$ 是总体的最小二乘估计。先在 $\delta_i^T\beta\leqslant t$ 的条件下解最小二乘问题，然后检查 $\sum\left|\beta_j\right|\leqslant t$ 是否成立，如果成立，计算完成，如果不成立，其他限制加入到 E 中，直到找到满足 $\sum\left|\beta_j\right|\leqslant t$ 的条件为止。这里给出具体算法的步骤：

① 从 $E=\{i_0\}$ 开始，这里 $\delta_{i_0}=\operatorname{sign}\left(\hat{\beta}^0\right)$，$\hat{\beta}^0$ 是总体的最小二乘估计。

② 找到 $\hat{\beta}$ 最小化 $g(\beta)$，且满足 $G_E\beta\leqslant tl$。

③ 当 $\sum\left|\hat{\beta}_j\right|>t$ 时，增加 i 到集合 E，这里 $\delta_i=\operatorname{sign}\left(\hat{\beta}\right)$，找到 $\hat{\beta}$ 使得最小化 $g(\beta)$，且满足 $G_E\beta\leqslant tl$。

由于这种算法是一种二次规划算法，计算速度很慢，因而美国科学院院士、斯坦福大学教授 Efron、Tibshirani 等人在 2004 年提出了 lars 算法，指出 lars 算法经过修改后，可以解决 lasso 问题的计算。

Lars 被称为最小角回归，它和著名的向前选择法一样，一开始令所有的系数为 0，并找出和因变量的残差相关性最强的自变量，记为 x_{j_1}；然后在这个变量的方向上尽最大可能找出另一个自变量，记为 x_{j_2}，使它与当前的残差有同样的相关性；之后沿着平分前两个变量夹角的方向，找到变量 x_{j_3}，使它满足相关性最强；然后再沿着平分前三个变量夹角的方向找到第四个变量，以此类推。

由于 lars 计算速度非常快，克服了 lasso 问题原有优化算法计算速度慢的缺点，所以使得 lasso 的研究风靡全球。

作者在硕士阶段研究的是统计量的算法理论，硕士毕业论文题目为《双寿命分布中的统计量 J_{1N} 的计算与研究》，文中提出一种统计量 J_{1N} 的优化算法。作者在研究统计量计算时，发现算法理论的研究非常有前景，但其后来被作者用数学方法证明是不可计算的，最后用数学理论找到统计量 J_{1N} 的优化算法。

用数学理论研究算法理论是非常有价值的。因此，作者在做博士论文时，就是用数学方法研究 lasso 的算法理论，对 lars 算法解决 lasso 问题的数学理论进行研究，发现 lasso 问题的快速算法 lars 和原有优化算法的关系存在问题。

Lars 算法本质上是两两选择的算法，不是总体最优的算法，而 lasso 的原有算法是总体最优的算法。这两种算法的计算数量级一个是 p^2，另一个是 2^p，由于是高维数据计算，因而必须研究问题的大样本性质，也就是说，必须搞清楚当 p 趋向正无穷时，问题的性质。但是，当 p 趋向正无穷时，lasso 问题的优化算法的数量级是不可数的，而 lasso 问题的快速算法 lars 的数量级是可数的，如果搞清楚这两种算法的关系，就是搞清了用可数算法解决不可数问题的一种方法，显然其研究价值非同一般。

作者仔细研读相关文献发现，斯坦福大学教授认为用小波等数学理论可以证明 lars 算法解决 lasso 问题，并且给出了证明。由于这种证明牵涉小波理论，作者一时无法看清，因而作者另辟蹊径，先用现代数学分析的方法对 lasso 的几何展开研究，找到当数据是二维和三维时，正交设计下 lasso 的精确解。根据 lasso 的几何，结合拉格朗日乘数法得到在正交设计下 lasso 的精确解。通过正交设计下 lasso 的精确解发现，lasso 问题的变量选择过程是一个向后选择的过程。lasso 问题原有的优化算法是一个向后选择的算法，和 lasso 问题一致，因而没有问题。而 lars 算法是一个向前选择的算法，用向前选择的算法解决一个向后选择的问题，如果找不到特殊的条件，肯定存在问题。

二、Lasso 问题的数学理论研究

2001 年，范剑青等人在杂志《Journal of the American Statistical Association》上发表论文《Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties》，提出 SCAD 估计，给出了一个评价惩罚函数的标准，一个好的惩罚函数得到的估计必须满足以下三个性质：

1. 稀疏性：估计能够将小的估计值自动地设置为 0，从而起到选择变量，减少模型复杂度的效果。

2. 无偏性：估计几乎是无偏的，特别是对于那些真值的绝对值较大的系数而言。

3. 连续性：估计关于数据本身是连续的，从而减少模型预测的不稳定性。

根据这三个标准，他们提出了 SCAD 估计（the smoothly clipped absolute deviation），提出 lasso 估计没有 oracle 性质的猜想（oracle 性质：估计不仅能够以概率趋向于 1 正确的选择模型，而且对于非零系数的估计，与已知真实模型下的最小二乘估计，具有相同的收敛速度），并且证明了 SCAD 估计有 oracle 性质。SCAD 的数学模型为：

当惩罚函数 $p_\lambda(\theta)$ 满足 $p'_\lambda(\theta)=\lambda\left\{I(\theta\leqslant\lambda)+\frac{(a\lambda-\theta)_+}{(a-1)\lambda}I(\theta>\lambda)\right\}$（$\theta>0,a>2$）时，$\hat{\theta}=\arg\min\left\{\frac{1}{2}(z-\theta)^2+p_\lambda(\theta)\right\}$。此时 θ 的 SCAD 估计为：

$$\hat{\theta}=\begin{cases}\operatorname{sgn}(z)(|z|-\lambda)_+, & |z|\leqslant 2\lambda\\ \{(a-1)-\operatorname{sgn}(z)a\lambda/(a-2)\}, & 2\lambda<|z|\leqslant a\lambda\\ z, & |z|>a\lambda\end{cases}$$

该文中对 lasso 估计提出猜想，这说明 lasso 解的结构存在问题。

2006 年，美国明尼苏达大学教授邹晖在杂志《Journal of American Statistical Association》上发表文章《The Adaptive Lasso and Its Oracle Properties》，提出 the adaptive lasso，证明了 lasso 估计在有些情况下没有 oracle 性质，指出 the adaptive lasso 有 oracle 性质，证明了范剑青等人在 2001 年提出的 lasso 问题没有“Oracle”性质的猜想，这再次证明了 lasso 问题解的结构是不清楚的。

the adaptive lasso 数学模型如下：

当 $\lambda\sum_{j=1}^{p}|\beta_j|\leqslant t$ 时，$\hat{\beta}=\arg\min\left\{\sum_{i=1}^{N}\left(y_i-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2\right\}$

此时得到的 $\hat{\beta}$ 就是 the adaptive lasso 估计。

这是两位著名华人学者的研究成果，也是他们在国内的工作，使得国内统计学领域的统计学家开始关注与 lasso 有关的问题。在范剑青教授对

lasso 问题的性质只能进行猜想，邹晖教授只能对范剑青教授的猜想进行证明的 lasso 问题研究领域，进行科学研究显然是非常有价值的。作者在上海财经大学攻读博士期间，与很多学者进行过交流，特别是在 2014 年 3 月 20 日，和到上海财经大学作学术报告的范剑青教授进行过一次交流。邹晖教授在上海财经大学作过学术报告，但非常遗憾，无缘与他进行交流。因此，本书重点介绍这两人的研究，在本章第二节简单介绍其他学者的一些研究。

第二节 Lasso 研究的发展

Lasso 问题经过 20 多年的发展，lasso 的理论研究和应用研究都达到了很高的水平，这种方法已经被应用到生物、医学、金融、计算机、运筹学等领域，下面介绍国内外的一些研究。

一、运用 Lasso 思想提出的新方法

2010 年，Zhang 在杂志《The Annals of Statistics》上发表论文《Nearly Unbiased Variable Selection under Minimax Concave Penalty》，推广了 lasso 的研究，提出了 MCP(a minimax concave penalty) 方法，即一个最小的凹惩罚和最小的线性无偏的算法。在一般情况下，研究的限制区域为凸区域，运用凹区域进行研究的文章很少，MCP 是研究限制区域是凹区域的一种方法。

2011 年，Belloni 等人在杂志《Biometrika》上发表论文《Square-root lasso : pivotal recovery of sparse signals via conic programming》，提出了平方根 lasso，这是一种估计高维稀疏线性回归模型的重要方法，这里回归总数 p 是大值，可能比 n 更大，而且只有 s 个回归数是显著的。

2009 年，Wasserman and Roeder 在杂志《The Annals of Statistics》上发表论文《High-Dimensional Variable Selection》，探讨了进行高维模型的变量选择的统计方法问题，提出了三种筛选的办法：the lasso，marginal regression 和 forward stepwise regression。

2008 年，Fan and Lv 在杂志《Journal of the Royal Statistical Society》上发表论文《Sure independence screening for ultrahigh dimensional feature space》，提出了 SIS（Sure Independence Screening）方法，对超高维数据进行降维，在一定条件下，证明了 SIS 方法具有 Sure Screening 性质，即能以概率趋向于 1，确保降维后的模型保留了所有的重要变量。对筛选后的变量，再使用 lasso，SCAD 等变量选择方法，得到最终的简化模型。

2005 年，Zhou 等人在杂志《Journal of Royal Statistics Society》上发表论文《Regularition and variable selection via the elastic net》，提出了 the elastic net 算法，这种算法有很强的收敛性。

2011 年，Tibshirani 应邀在杂志《Journal of the Royal Statistical Society》上发表文章《Regression shrinkage and selection via the lasso: a retrospective》，对 lasso 进行了回顾，文章指出 Yuan and Lin(2007a) 提出了 Group lasso，Zou and Hastie(2005) 提出了 the elastic net，Tibshirani 等人 (2005) 提出了 Fused lasso，Yuan and Lin(2007b) 和 Friedman 等人 (2007) 提出了 Graphical lasso，Candes and Tao(2007) 提出了 Dantzig selector，Tibshiani 等人 (2010) 提出了 Near isotonic regularization，Candes and Tao(2009) 和 Mazumder 等人 (2010) 提出了 Matrix completion，Donoho(2004) 和 Candes(2006) 提出了 Compressive sensing，Jolliffe et al.(2003) 和 Wrtten 等人 (2009) 提出了 Multivariate methods。

这些研究根据 lasso 的思想提出了新的模型和新的方法，进一步拓宽了 lasso 的应用范围。lasso 的研究成果很多，下面介绍一些国内外学者的研究。

二、外国专家的一些研究

2009 年，Zou and Zhang 在杂志《The Annals of Statistics》上发表论文《On the adaptive elastic-net with a diverging number of parameters》，研究了参数个数和样本数量同时发散的模型选择和参数估计，提出了 the adaptive elastic-net 可以解决共线性问题，并且证明 the adaptive elastic-net 有 oracle 性质。

2008 年，Kim 等人在杂志《Journal of American Statistical Association》

上发表论文《Smooth Clipped Absolute Deviation on High Dimensions》，对 SCAD 估计在高维的情况进行了研究，找到了一个更有效的算法，证明了 SCAD 估计在高维有 oracle 性质。

2000 年，Knight and Fu 在杂志《The Annals of Statistics》上发表论文《Asymptotics for Lasso-Type Estimators》，考虑了带惩罚部分 $\sum|\beta|^{\gamma}$ 且 $\gamma>0$ 最小化残差平方和估计的渐进性质，得到一些有用的定理，当 $\gamma=1$ 时，就是 lasso 估计。

2008 年，Huang 等人在杂志《The Annals of Statistics》上发表论文《Asymptotic Properties of Bridge Estimators in Sparse High-Dimensional Regression Models》，研究了当协变量数目和样本量同时趋向无穷时，稀疏、高维的线性模型的桥回归（bridge regression）的渐进性质，指出在合适的条件下，桥回归有 oracle 性质。

2007 年，Yuan and Lin 在杂志《Journal of Royal Statistics Society》上发表论文《On the non-negative garrote estimate》，从一致性、计算和灵活性三个方面研究了 non-negative garotte 估计，同时研究了 non-negative garotte 估计的其他性质。

Friedman 等人在 2010 年给出“The glmnet R language package”，Hastie 和 Efron 在 2013 年 4 月发布 lars1.20，对 lasso 型问题的算法进行总结和更新，这些研究被应用于各个学科的科学研究中。

上面介绍的是国外的一些研究，说明 lasso 的研究是主流研究，下面介绍国内的一些相关研究。

三、国内专家的一些研究

2017 年，曾津和周建军在《数理统计与管理》杂志上发表论文《高维数据变量选择方法综述》，指出 lasso 方法的提出催生了大量有关算法研究的文献。从最初既不具有规模性又缺乏透明性的二次规划的方法到 Coordinate 算法，真正使 lasso 广为流行的算法当属 lars 算法。

2012 年，孙燕在《数量经济技术经济研究》杂志上发表论文《随机效应 Logit 计量模型的自适应 Lasso 变量选择方法研究》，重点研究了计