

二次度量回归模型中折叠凹惩罚估计的统计性质

张馨玉*, 杨婧昱

河北工业大学理学院, 天津

收稿日期: 2023年9月19日; 录用日期: 2023年10月12日; 发布日期: 2023年10月20日

摘要

目前, 关于二次度量回归模型的研究受到了广泛关注, 比如相位恢复、动力系统状态估计、无标记的距离几何和各种组合图等问题。本文考虑从高维二次度量回归模型中恢复未知信号。通过采用折叠凹惩罚最小二乘估计方法, 我们得到了真实信号与估计值之间的误差界。并且结果中还表明估计值与真实值有相同的支撑集。另外, 文章中我们主要研究SACD和MCP两种典型的折叠凹惩罚函数。

关键词

二次度量回归模型, SCAD惩罚方法, MCP惩罚方法, 高维统计

Folded Concave Penalized Estimation for Quadratic Measurements Regression

Xinyu Zhang*, Jingyu Yang

School of Science, Hebei University of Technology, Tianjin

Received: Sep. 19th, 2023; accepted: Oct. 12th, 2023; published: Oct. 20th, 2023

Abstract

Recovering an unknown signal from quadratic measurements has gained more attention, such as phase retrieval, power system state estimation and the unlabeled distance geometry problem. In this paper, we reconstruct the unknown signal from high-dimensional quadratic measurements. By employing folded concave penalized least squares method, our main result shows the non-asymptotic error bound between the estimator and the true signal. Our result shows that the es-

*通讯作者。

imator has the same support set as the true signal. In addition, we focus on two typical folding concave penalty functions, SCAD and MCP.

Keywords

Quadratic Measurements, SCAD Penalty, MCP Penalty, High-Dimensional Statistical Analysis

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 研究背景

本文中, 我们考虑二次度量回归模型

$$y_i = \beta^{*T} Z_i \beta^* + \varepsilon_i, i=1, \dots, n, \quad (1.1)$$

其中 $y_i \in \mathbb{R}$ 是可观测的响应变量, $\beta^* \in \mathbb{R}^p$ 是未知的真实信号, n 是样本数目, $Z_i \in \mathbb{R}^{p \times p}$ 是测量矩阵, $\varepsilon_i \in \mathbb{R}$ 为独立同分布的随机误差。目前, 对于二次度量回归模型的研究越来越被重视, 特别是在许多科学和工程领域, 比如众所周知的相位恢复, 电力系统状态估计, 无标记距离几何问题等。并且其主要目标是从二次度量回归模型中恢复未知信号。下面我们给出几个例子进行详细介绍。

1) 相位恢复 相位恢复问题[1] [2] [3]广泛存在于许多科学和工程领域, 包括光学成像、晶体学、中子放射学、电子显微镜、数字通信、天文学和计算生物学等。标准的相位恢复问题是基于强度测量数据 $z_i \in \mathbb{R}^p, i=1, \dots, n$ 恢复未知信号 $\beta \in \mathbb{R}^p$, 其中强度测量数据与未知信号之间满足

$$y_i = |z_i^T \beta|^2 + \varepsilon_i.$$

令矩阵 $Z_i = z_i z_i^T$, 则上述模型就等价于二次度量模型(1.1), 其中 Z_i 为秩 1 矩阵。

2) 电力系统状态估计 智能电网作为一种大型复杂的基础设施, 需要借助大量的测量信息进行实时监测和决策。准确监控系统的运行是电网可靠和经济运行的核心, 从而实施状态估计至关重要[4]。若给定一个总线和传输线路网络, 电力系统状态估计的目标是从测量的功率和电压数据 (y_1, \dots, y_n) , 来估计所有网络总线上的电压 (v_1, \dots, v_n) 。根据交流电功率模型, 可以看出这些测量值是二次依赖于待确定的电压值, 即它们之间满足二次关系(1.1)。

3) 无标记的距离集合问题 无标记的距离集合问题[5]的目标是从一组未标记的成对距离中, 恢复一组点的相对位置。这里对这组点位于一条直线上时进行举例[6]; 假设直线上有 p 个点, 并且点的距离分布函数为 y , 把直线平均分为 n 个线段, 点的密度为 $x = (x_m)_{m=1}^M \in \mathbb{R}^M$, 距离分布函数 y 与点密度 x 之间呈二次关系(1.1)。

高维数据在科技发展中扮演着越来越重要的角色, 并且已经成为机器学习, 数理统计和信息科学等学科的研究热点。有效的变量选择方法通过筛选掉不重要的变量, 得出简洁的模型, 从而提高模型的预测精度和可靠性。事实上, 如果选择一个太小的子集容易导致信息缺失, 使得结果不准确, 如果选择的变量过多, 则会加重“维数诅咒”。所以, 在建模高维数据时选择正确的变量子集并排除不重要的变量是相当重要的[7]。惩罚方法是我们处理和海量高维数据的最广泛使用的工具之一。近些年来, 尤其是在线性模型中, 利用惩罚方法进行变量选择受到了大量统计学家的关注。惩罚方法不仅可以在变量选择的同时给出参数估计, 还可以减少模型的计算量, 因此和传统的变量选择方法相比, 惩罚估计方法有

很高的优越性。

目前研究最多的惩罚方法是 Lasso [8] 和非凸惩罚方法, 并且在稀疏恢复领域, 非凸惩罚比凸惩罚能够得到更好的稀疏性[9]。特别地, 折叠凹惩罚方法在压缩感知中有了充分研究, 尤其是 SCAD [9] 和 MCP [10] 两种惩罚方法。文献[10]指出基于折叠凹惩罚方法得到的全局解有理想的恢复性能。由于 SCAD 函数的良好性质, 运用 SCAD 惩罚方法进行变量选择既可以有效的降低模型的选择偏差, 又能有效提高模型预测的精准程度[9]。

本文中, 为了实现对高维二次度量回归模型的参数估计以及变量选择问题, 我们研究如下基于折叠凹惩罚最小二乘估计方法,

$$\min_{\beta \in \mathbb{R}^p} L_n(\beta) := \frac{1}{4n} \sum_{i=1}^n (y_i - \beta^T Z_i \beta)^2 + \sum_{j=1}^p P_{\lambda_n}(|\beta_j|), \quad (1.2)$$

其中 $P_{\lambda_n}(|\beta_j|)$ 为惩罚函数, $\lambda_n > 0$ 为惩罚参数。本文中, 我们采用 SCAD 和 MCP 两种典型的、被广泛使用的折叠凹惩罚函数, 其函数解析式表示如下。

SCAD (Smoothly Clipped Absolute Deviation),

$$P_{\lambda_n}(t) = \begin{cases} \lambda_n t, & 0 \leq t < \lambda_n, \\ \frac{(a^2 - 1)\lambda_n^2 - (t - a\lambda_n)^2}{2(a - 1)}, & \lambda_n \leq t < a\lambda_n, \\ \frac{a + 1}{2} \lambda_n^2, & t \geq a\lambda_n. \end{cases}$$

其中 $a > 2$ 为固定的参数。

MCP (Minimax Concave Penalty),

$$P_{\lambda_n}(t) = \begin{cases} \lambda_n t - \frac{t^2}{2b}, & 0 \leq t < b\lambda_n, \\ \frac{1}{2} b\lambda_n^2, & t \geq b\lambda_n. \end{cases}$$

其中 $b > 0$ 为固定参数。

在整篇论文中, 我们使用如下记号, 对任意 p 维向量 $v = (|v_1|, \dots, |v_p|)^T$, 令 $\|v\|_2 = \left(\sum_{i=1}^p v_i^2\right)^{1/2}$,

$\|v\|_1 = \left(\sum_{i=1}^p |v_i|\right)$ 且 $\|v\|_\infty = \max\{|v_1|, \dots, |v_p|\}$ 。对任意 $n \times p$ 矩阵 $Z = (z_{ij})$, 令 $|Z|_\infty = \max_{1 \leq i, j \leq p} |z_{ij}|$ 。 Z^{AB} 为从矩阵 Z 中分别抽取指标集 A 和 B 对应行和列所成子矩阵。 β_A 为从向量 β 直接抽取指标集 A 对应位置的元素所构成的向量。 $e_{d,j}$ 表示 d 阶单位矩阵的第 j 列向量。

2. 主要结果与证明

本节将讨论基于 SCAD 和 MCP 惩罚最小二乘估计量的稀疏性和一致性, 并给出了合理的证明过程。

2.1. 估计量的统计性质

为了研究估计量的统计性质, 我们首先给出一些关于矩阵 Z_i 和噪声 ε_i 的假设条件。令 β^* 为真实值, 并且 $\Gamma^* = \text{supp}(\beta^*) := \{j : e_{p,j}^T \beta^* \neq 0, j = 1, \dots, p\}$ 。像文献[11]一样, 我们假设存在常数 $0 < c_1 \leq c_2 < \infty$ 满足

$$c_1 \leq \min\{|e_{p,j}^T \beta^*|, j \in \Gamma^*\} \leq \max\{|e_{p,j}^T \beta^*|, j \in \Gamma^*\} \leq c_2. \quad (2.1)$$

进一步, 像文献[11] [12]一样, 我们假设数据已经被标准化

$$\sum_{i=1}^n y_i = 0 \text{ 和 } \sum_{i=1}^n |Z_i|_{\infty}^2 = n.$$

假设 1 对任意 $x, y \in \mathbb{R}^s$, 存在常数 $0 < c_3 \leq c_4$, $c_5 > 0$ 满足

$$c_3 \|x\|^2 \|y\|^2 \leq \frac{1}{n} \sum_{i=1}^n \left(x^T Z_i \Gamma^* \Gamma^* y \right)^2 \leq c_4 \|x\|^2 \|y\|^2,$$

且

$$\left| \sum_{i=1}^n Z_i \Gamma^* \Gamma^* \otimes Z_i^{*c} \Gamma^{*c} \right|_{\infty} \leq c_5 \sqrt{n},$$

$$\left| \sum_{i=1}^n Z_i \Gamma^* \Gamma^* \otimes Z_i \Gamma^{*c} \Gamma^{*c} \right|_{\infty} \leq c_5 \sqrt{n}$$

其中 \otimes 为 Kronecker 积。

在标准相位恢复问题中, 第一个不等式可以简化为文献[13] [14]中类似的条件, 并且文献中已经证明了向量 z_i 的可行性。此外, 后两个不等式可以看作是线性模型中正交性条件的扩展[15]。并且在二次度量回归模型中, 文献[12]已经使用了相同假设。

假设 2 假设噪声 $\{\varepsilon_1, \dots, \varepsilon_n\}$ 服从参数为 σ 的高斯分布, 并且 ε_i 均值为 0, 方差为 $E\varepsilon_i^2$ 。

我们令 $\hat{\beta}$ 为二次测量模型的最优解, 容易看出 $-\beta^*$ 也能够满足二次模型(1.1)。从而我们希望最优解 $\hat{\beta}$ 可以充分接近真实值 β^* 或 $-\beta^*$, 所以我们用如下式子去描述真实值与估计值之间的界

$$\min \left\{ \|\hat{\beta} - \beta^*\|, \|\hat{\beta} + \beta^*\| \right\}.$$

下面我们给出本文中主要的结论。

定理 对于二次测量模型(1.1), 假定假设 1-2 成立, 惩罚函数为 SCAD 或者 MCP, 并且惩罚参数 λ_n 当 $n \rightarrow \infty$ 时满足

$$\lambda_n \rightarrow 0 \text{ 和 } \lambda_n^{-2} \sqrt{\frac{\ln p}{n}} \rightarrow 0, \quad (2.2)$$

则最优化问题(1.2)存在局部最优解 $\hat{\beta} = (\hat{\beta}_{\Gamma^*}^T, \hat{\beta}_{\Gamma^{*c}}^T)^T$ 使得对于足够大的 n 满足

$$(i) \quad P(\hat{\beta}_{\Gamma^{*c}} = 0) \geq 1 - \frac{1}{n},$$

$$(ii) \quad P\left(\min \left\{ \|\hat{\beta} + \beta^*\|, \|\hat{\beta} - \beta^*\| \right\} \leq \frac{8\sigma\sqrt{c_4}}{c_1 c_3} \sqrt{\frac{\ln(1+2n)}{n}} + \frac{4}{c_1 c_3 \sqrt{s}} \lambda_n \right) \geq 1 - \frac{1}{n}.$$

我们可以发现条件(2.2)蕴含了样本量 n 与维数 p 之间的关系。如果取 $\lambda_n = n^a$, 则对于某个 $a \in (-1/8, 0)$, 条件(2.2)可以简化为

$$\frac{\ln p}{\sqrt{n}} \rightarrow 0.$$

通过定理中的结论(i), 我们可以看出折叠凹惩罚方法得到的估计量具有相合性。通过结论(ii), 我们可以看出折叠凹惩罚估计能够以趋于 1 的概率正确选择出非零变量, 其中这个概率的收敛速度时线性的。

2.2. 定理证明

为了证明叙述中记号的简洁, 我们不妨设 $\Gamma^* = \{1, \dots, s\}$, $\beta^* = (\beta_1^{*T}, 0^T)$ 以及 $Z_i = \begin{bmatrix} Z_i^{11} & Z_i^{12} \\ Z_i^{21} & Z_i^{22} \end{bmatrix}$. 记

$$\min \tilde{L}_n(\beta_1) := \frac{1}{4n} \sum_{i=1}^n (y_i - \beta_1^T Z_i^{11} \beta_1)^2 + \sum_{j=1}^s P_{\lambda_n}(|\beta_j|). \quad (2.3)$$

我们首先证明优化问题(2.3)最优解的存在性. 令 $\hat{\beta}_1 = \arg \min_{\beta_1 \in \mathbb{R}^s} \tilde{L}_n(\beta_1)$. 考虑水平集 $M = \{\beta_1 \in \mathbb{R}^s : \tilde{L}_n(\beta_1) \leq \tilde{L}_n(\beta_1^*)\}$, 其中 M 为非空闭集, 并且 $\inf_{\beta_1 \in \mathbb{R}^s} \tilde{L}_n(\beta_1) = \inf_{\beta_1 \in M} \tilde{L}_n(\beta_1)$. 因为 $\tilde{L}_n(\cdot)$ 是连续函数, 所以 $\tilde{L}_n(\cdot)$ 在水平集中至少有一个最优解. 从而可以证明 $\hat{\beta}_1$ 的存在性.

下面我们证明 $\hat{\beta}_1$ 与 β_1^* 之间的误差界. 利用假设 1 和 $\hat{\beta}_1$ 的定义, 我们可以得到对任意的 $\beta_1 \in \mathbb{R}^s$, 有

$$\tilde{L}_n(\hat{\beta}_1) \leq \tilde{L}_n(\beta_1).$$

从而

$$\begin{aligned} 0 &\geq \tilde{L}_n(\hat{\beta}_1) - \tilde{L}_n(\beta_1^*) \\ &= \frac{1}{4n} \sum_{i=1}^n \left[(\hat{\beta}_1 - \beta_1^*)^T Z_i^{11} (\hat{\beta}_1 + \beta_1^*) \right]^2 + \frac{1}{2n} \sum_{i=1}^n (\beta_1^{*T} Z_i^{11} \beta_1^* - \hat{\beta}_1^T Z_i^{11} \hat{\beta}_1) \varepsilon_i + P_{\lambda_n}(|\hat{\beta}_j|) - P_{\lambda_n}(|\beta_j^*|) \\ &\geq \frac{c_3}{4} \|\hat{\beta}_1 - \beta_1^*\|^2 \|\hat{\beta}_1 + \beta_1^*\|^2 - \frac{1}{2n} \sum_{i=1}^n (\hat{\beta}_1 - \beta_1^*)^T Z_i^{11} (\hat{\beta}_1 + \beta_1^*) \varepsilon_i + P_{\lambda_n}(|\hat{\beta}_j|) - P_{\lambda_n}(|\beta_j^*|). \end{aligned}$$

令 $u = \frac{\hat{\beta}_1 - \beta_1^*}{\|\hat{\beta}_1 - \beta_1^*\|_2}$ 和 $v = \frac{\hat{\beta}_1 + \beta_1^*}{\|\hat{\beta}_1 + \beta_1^*\|_2}$. 由上述式子我们可以得到

$$\frac{c_3}{4} \|\hat{\beta}_1 - \beta_1^*\|^2 \|\hat{\beta}_1 + \beta_1^*\|^2 \leq \frac{1}{2n} \|\hat{\beta}_1 - \beta_1^*\| \|\hat{\beta}_1 + \beta_1^*\| \sum_{i=1}^n u^T Z_i^{11} v \varepsilon_i + \sum_{j=1}^s P_{\lambda_n}(|\beta_j^*|) - P_{\lambda_n}(|\hat{\beta}_j|). \quad (2.4)$$

根据 SCAD 惩罚函数的定义, 对任意的 $t_1, t_2 > 0$, 有 $P_{\lambda_n}(t_1) - P_{\lambda_n}(t_2) \leq \lambda_n |t_2 - t_1|$, 即

$$\sum_{j=1}^s P_{\lambda_n}(|\beta_j^*|) - P_{\lambda_n}(|\hat{\beta}_j|) \leq \lambda_n \|\hat{\beta}_1 - \beta_1^*\|.$$

联合上式, (2.4) 以及 $\|\beta_1\|_1 \leq \sqrt{s} \|\beta_1\|_2$, 我们有

$$\frac{c_3}{4} \|\hat{\beta}_1 - \beta_1^*\|^2 \|\hat{\beta}_1 + \beta_1^*\|^2 \leq \frac{1}{2n} \|\hat{\beta}_1 - \beta_1^*\| \|\hat{\beta}_1 + \beta_1^*\| \sum_{i=1}^n u^T Z_i^{11} v \varepsilon_i + \lambda_n \sqrt{s} \|\hat{\beta}_1 - \beta_1^*\|_2. \quad (2.5)$$

在下面证明中我们不妨设 $\|\hat{\beta}_1 - \beta_1^*\| \leq \|\hat{\beta}_1 + \beta_1^*\|$, 容易验证

$$\beta^* \leq \|\hat{\beta}_1 + \beta_1^*\|. \quad (2.6)$$

利用 Bernstein 不等式和矩阵 Z_i 的性质, 我们可以得到

$$\frac{1}{n} \sum_{i=1}^n u^T Z_i^{11} v \varepsilon_i \leq 4\sigma \sqrt{c_4} \sqrt{\frac{s \ln(1+2n)}{n}}.$$

利用上式, (2.5), (2.6) 和真实值 β^* 的界(2.1), 我们可以验证

$$\|\hat{\beta}_1 - \beta_1^*\| \leq \frac{8\sigma\sqrt{c_4}}{c_3\|\beta_1^*\|} \sqrt{\frac{s \ln(1+2n)}{n}} + \frac{4}{c_1 c_3 \sqrt{s}} \lambda_n. \tag{2.7}$$

下面我们证明 $\hat{\beta} = (\hat{\beta}_1^T, 0^T)^T$ 是优化问题(1.2)在邻域 $\{\beta^* + u, \|u\| \leq C\}$ 内的最优解, 其中 C 是足够大的常数。记 $u = (u_1^T, u_2^T)^T \in \mathbb{R}^p$ 和 $u_1 \in \mathbb{R}^s$, 我们需要证明 $L_n(\hat{\beta}, 0) \leq L_n(\beta_1^* + u_1, u_2)$ 。因为 $\hat{\beta}_1$ 是最优化问题(2.3)的最优解, 则对任意的 $\beta_1 \in \mathbb{R}^s$, 有 $L_n(\hat{\beta}_1, 0) \leq L_n(\beta_1, 0)$ 。从而, 我们只需要证明下面不等式成立

$$P(L_n(\beta_1^* + u_1, u_2) \geq L_n(\beta_1^* + u_1, 0)) \geq 1 - \frac{1}{n}.$$

令 $\zeta_i = 2Z_i^{21}(\beta_1^* + r_n u_1) + \tilde{r}_n Z_i^{22} u_2$ 和 $\eta_i = Z_i^{11}(2\beta_1^* + r_n u_1)$ 。通过简单计算, 我们可以得到

$$\begin{aligned} & L_n(\beta_1^* + u_1, u_2) - L_n(\beta_1^* + u_1, 0) \\ &= \frac{1}{n} \sum_{i=1}^n |u_2^T \zeta_i|^2 + \frac{2}{n} \sum_{i=1}^n u_1^T \eta_i u_2^T \zeta_i - \frac{2}{n} \sum_{i=1}^n u_2^T \zeta_i \varepsilon_i + \sum_{j=s+1}^p P_{\lambda_n}(|u_{2j}|) \\ &\geq -\frac{2}{n} r_n^2 \|u_1\| \|u_2\| \left\| \sum_{i=1}^n \zeta_i \eta_i^T \right\|_{\infty} - \frac{2}{n} \|u_2\| \left\| \sum_{i=1}^n \zeta_i \varepsilon_i \right\|_{\infty} + \sum_{j=s+1}^p P_{\lambda_n}(|u_{2j}|). \end{aligned}$$

利用 Bernstein 不等式和无穷范数的定义, 我们可以计算得到至少以 $1 - n^{-1}$ 概率有以下式子成立

$$\frac{1}{n} \left\| \sum_{i=1}^n \zeta_i \varepsilon_i \right\|_{\infty} \leq 2\sqrt{6}\sigma c_2 \sqrt{\frac{s \ln p}{n}}. \tag{2.8}$$

利用假设 1 中后两个不等式以及(2.1), 我们可以得到

$$\frac{1}{n} \left\| \sum_{i=1}^n \zeta_i \eta_i \right\|_{\infty} \leq \frac{c}{\sqrt{n}}, \tag{2.9}$$

其中 c 是一个大于零的常数。

根据 SCAD 惩罚函数的定义, 对任意的 $u_j \in \mathbb{R}$, 可以证明有下式成立

$$P_{\lambda_n}(|u_j|) \geq C' \lambda_n^2 |u_j|, \tag{2.10}$$

其中常数 $C' = 2^{-1} \min\{\lambda_n^{-1}, (a+1)a^{-1}\}$, 下面我们分三种情况对上式进行证明。

(i) 当 $|u_j| \leq \lambda_n$ 时, 有 $P_{\lambda_n}(|u_j|) = \lambda_n |u_j|$ 。取 $C' = \lambda_n^{-1}$, 则(2.10)显然成立。

(ii) 当 $\lambda_n < |u_j| < a\lambda_n$ 时, 有 $P_{\lambda_n}(|u_j|) = (2a-2)^{-1}(-|u_j|^2 + 2a\lambda_n |u_j| - \lambda_n^2)$ 。由 $|u_j|$ 的上、下界, 有

$$P_{\lambda_n}(|u_j|) \geq \frac{a\lambda_n}{2(a-1)} |u_j| - \frac{1}{2(a-1)} \lambda_n |u_j| = \frac{1}{2} \lambda_n |u_j|,$$

取 $C' = (2\lambda_n)^{-1}$, 则(2.10)显然成立。

(iii) 当 $|u_j| \geq a\lambda_n$ 时, 有 $P_{\lambda_n}(|u_j|) = \frac{a+1}{2} \lambda_n^2$ 。由 $\|u\| \leq C$, 有 $|u_j| \leq C$ 。我们可以计算得到

$$P_{\lambda_n}(|u_j|) = \frac{a+1}{2C} \lambda_n^2 |u_j|,$$

其中 $C' = (2C)^{-1}(a+1)$ 。从而(2.10)得证。

结合上述不等式(2.8), (2.9)以及(2.10), 我们可以证得

$$\begin{aligned}
& L_n(\beta_1^* + u_1, u_2) - L_n(\beta_1^* + u_1, 0) \\
& \geq -2\sqrt{6}\sigma c_2 \sqrt{\frac{s \ln p}{n}} - \frac{c}{\sqrt{n}} + C' \lambda_n^2 |u_j| \\
& = \lambda_n^2 |u_j| \left(-2\sqrt{6}\sigma c_2 \lambda_n^{-2} \sqrt{\frac{s \ln p}{n}} - \lambda_n^{-2} \frac{c}{\sqrt{n}} + C' \right).
\end{aligned}$$

根据条件中的极限 $\lambda_n^{-2} n^{-1/2} (\ln p)^{1/2} \rightarrow 0$, 其蕴含了 $\lambda_n^{-2} n^{-1/2} \rightarrow 0$, 显然

$$L_n(\beta_1^* + u_1, u_2) - L_n(\beta_1^* + u_1, 0) \geq 0,$$

从而完成该定理证明。

3. 结论

本文研究的高维二次度量回归模型, 其在相位恢复、电力系统状态估计等问题中有广泛的应用价值。本文基于折叠凹惩罚最小二乘估计方法对二次度量回归模型进行参数估计, 该方法还能够对参数进行变量选择。其中折叠凹惩罚方法是本文的创新点, 并且该方法具有良好的优势。另外, 本文还给出了在一定条件下真实值与估计值之间的误差界。

参考文献

- [1] Wang, Y. and Xu, Z. (2019) Generalized Phase Retrieval: Measurement Number, Matrix Recovery and Beyond. *Applied and Computational Harmonic Analysis*, **47**, 423-446. <https://doi.org/10.1016/j.acha.2017.09.003>
- [2] Candes, E., Eldar, Y., Strohmer, T. and Vershynina, V. (2013) Phase Retrieval via Matrix Completion. *SIAM Review*, **6**, 199-225. <https://doi.org/10.1137/110848074>
- [3] Shechtman, Y., Beck, A. and Eldar, Y. (2014) GESPAR: Efficient Phase Retrieval of Sparse Signals. *IEEE Transactions on Signal Processing*, **62**, 928-938. <https://doi.org/10.1109/TSP.2013.2297687>
- [4] Wang, G., Zamzam, A.S., Giannakis, G.B. and Sidiropoulos, N.D. (2018) Power System State Estimation via Feasible Point Pursuit: Algorithms and Cramér Rao bound. *IEEE Transactions on Signal Processing*, **66**, 1649-1658. <https://doi.org/10.1109/TSP.2018.2791977>
- [5] Huang, S. and Dokmani, I. (2021) Reconstructing Point Sets from Distance Distributions. *IEEE Transactions on Signal Processing*, **69**, 1811-1827. <https://doi.org/10.1109/TSP.2021.3063458>
- [6] Duxnury, P.M., Granlund, L., Gujarathi, S.R., Juhas, P. and Billinge, S.J.L. (2016) The Unassigned Distance Geometry Problem. *Discrete Applied Mathematics*, **204**, 117-132. <https://doi.org/10.1016/j.dam.2015.10.029>
- [7] Donoho, D.L. (2000) High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. *Math Challenges Lecture*, 1-32. https://www.researchgate.net/publication/220049061_High-Dimensional_Data_Analysis_The_Curses_and_Blessings_of_Dimensionality
- [8] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [9] Fan, J. and Li, G. (2011) Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. *Journal of the American Statistical Association*, **96**, 1348-1360. <https://doi.org/10.1198/016214501753382273>
- [10] Zhang, C.H. (2010) Nearly Unbiased variable Selection under Minimax Concave Penalty. *The Annals of Statistics*, **38**, 894-942. <https://doi.org/10.1214/09-AOS729>
- [11] Huang, J., Horowitz, J.L. and Ma, S. (2008) Asymptotic Properties of Bridge Estimators in Sparse High-Dimensional Regression Models. *The Annals of Statistics*, **36**, 587-613. <https://doi.org/10.1214/009053607000000875>
- [12] Fan, J., Kong, L., Wang, L. and Xiu, N. (2018) Variable Selection in Sparse Regression with Quadratic Measurements. *Statistica Sinica*, **28**, 1157-1178. <https://doi.org/10.5705/ss.202015.0335>
- [13] Candes, E.J., Li, X. and Soltanolkotabi, M. (2015) Phase Retrieval via Wirtinger Flow: Theory and Algorithms. *IEEE Transactions on Information Theory*, **61**, 1985-2007. <https://doi.org/10.1109/TIT.2015.2399924>
- [14] Cai, T., Li, X. and Ma, Z. (2016) Optimal Rates of Convergence for Noisy Sparse Phase Retrieval via Thresholded

- Wirtinger Flow. *The Annals of Statistics*, **44**, 2221-2251. <https://doi.org/10.1214/16-AOS1443>
- [15] Zhang, C. and Zhang, T. (2012) A General Theory of Concave Regularization for High-Dimensional Sparse Estimation Problems. *Statistical Science*, **27**, 576-593. <https://doi.org/10.1214/12-STS399>