arXiv:2303.05131v1 [stat.ME] 9 Mar 2023

# 半参数模型中未知参数方向的估计

Jinyue Han[1] Jun Wang[2] Wei Gao[1*] Man-LaiTang[3]1. 东北师范大学数学与统计学院 2. 云南大学数学与统计学院 3. 布鲁内尔大学数学系

摘要

半参数模型在计量经济学、社会科学和医学应用中具有重要作用。本文提出一种基于最小二乘法的新估计量，用于估计半参数模型中未知参数的方向。在温和条件下，无需了解链接函数的形式，该估计量即具有相合性和渐近分布。模拟结果表明，对于二元响应变量，所提出的估计量显著优于 Manski（1975）提出的最大得分估计量。当误差项服从重尾分布或具有无穷矩的分布时，所提出的估计量表现良好。本文还通过广东制造业出口参与数据展示了该方法的应用。

关键词：二元模型、方向、最小二乘估计量、最大得分、半参数模型、单指标模型。

## 1 引言

考虑回归模型 $E(Y \mid \mathbf{X})$ 的估计问题，其中 $Y$表示响应变量，$\mathbf{X}$ 是 $p$维可观测协变量。模型 $E(Y \mid \mathbf{X})$ 在经济学、医学及其他领域有着重要应用，其中 $E(Y \mid \mathbf{X})$ 的估计是一个关键问题。尽管非参数方法十分灵活，但代价高昂：随着 $p$ 增加，估计精度会迅速下降；当协变量 $\mathbf{X}$ 的维数较大时，估计结果也难以解释。因此，为了避免非参数模型的维数灾难，同时仍能提供

$E(Y \mid \mathbf{X})$ 函数形式上的灵活性，一种自然的做法是假设 $E[Y \mid \mathbf{X}]$ 是一个半参数模型。一种常用的半参数模型定义如下:

$$E[Y \mid \mathbf{X}] = E[Y \mid \mathbf{X}'\beta] = g(\mathbf{X}'\beta), \tag{1}$$

其中 $\beta \in R^p$ 和 $g : R \to R$ ，是一个未知的链接函数。当 $g(\cdot)$ 已知时，可以使用广义矩估计法估计未知参数 $\beta$。本文假设链接函数 $g(\cdot)$ 未知，并估计参数的方向 $\beta$。

模型（1）的一个特例是二元响应模型，即

$$Y = 1\{\mathbf{X}'\beta + \epsilon > 0\}, \tag{2}$$

其中 $\epsilon$是不可观测随机变量。当给定 X 时 $\epsilon$ 的条件分布已知（例如逻辑斯蒂模型或 Probit 模型），可以通过最大似然法估计未知参数。当 $\epsilon$ 的条件分布未知时，Manski（1975）针对具有条件中位数约束的二元响应模型提出了最大得分估计量，即 $\text{med}(\epsilon \mid \mathbf{X}) = 0$，其中 $\text{med}(\epsilon \mid \mathbf{X})$表示给定 X 时 $\epsilon$的条件中位数。基于 Manski（1975）的研究结果，Horowitz（1992）提出了平滑最大得分估计量，证明在某些假设下该估计量渐近正态，并采用经典自助法进行推断。Abrevaya和 Huang（2005）指出，经典自助法对最大得分估计量不相合。Patra等人（2018）提出了基于模型的平滑自助过程，用于对最大得分估计量进行统计推断，并证明了其相合性。Gao等人（2022）提出了两阶段最大得分估计量。

模型（1）的另一种特殊形式是经典单指标模型

$$Y = g(\mathbf{X}'\beta) + \epsilon. \tag{3}$$

单指标模型的估计方法主要有两类。一类是M估计法，利用核估计量（Ichi mura，1993）、回归样条（Park 等人，2020）、局部线性近似（Zhou 等人，2019）、惩罚样条（Yu 等人，2002）和平滑样条（Kuchibhotla和Patra,

2020）估计 $g(\cdot)$，并最小化适当的准则函数，例如二次损失（Yu 和 Ruppert，2002）、分位数回归（Wu 等人，2010）、估计方程法（Cui 等人，2011）、稳健 $L_1$ 损失（Zou 和 Zhu，2014）、准似然（Wang 和 Guo，2019）、剖面似然（Patra 等人，2020）和众数回归（Liu 等人，2013；Yang 等人，2020），以获得 $\beta$。另一类是直接估计方法，例如最大秩相关估计量（Han，1987；Fan 等人，2020）、平均导数估计量（Chaudhuri 等人，1997；Hristache 等人，2001）、降维技术（Li 和 Duan，1989；Li 和 Racine，2007；Li，2018）、偏最小二乘（Naik 和 Tsai，2000；Zhou 和 He，2008）以及线性化最大秩相关估计量（Shen 等人，2023）。近期，Kuchibhotla 等人（2021）针对独立同分布观测，提出了一种凸利普希茨约束最小二乘估计量（CLSE），用于估计参数分量和非参数分量。

模型（1）包含线性模型，其中提出了一种最小二乘法，用于估计响应变量为连续或离散时半参数模型中参数 $\beta$ 的方向。我们提出的方法计算简单，理论上可靠。模拟结果表明，对于离散响应变量，所提出的估计量显著优于最大得分估计；对于 Probit 模型，其表现与线性化最大秩相关（LMRC）估计和最大似然估计方法相当。当误差项的分布具有厚尾（即 Student t 分布），或分布不存在有限矩（即柯西分布）时，所提出的估计量和 LMRC 估计量的表现优于标准估计量。当协变量维数相对较高时，所提出的方法仍然可行。对于非线性模型，所提出的估计优于线性化最大秩相关估计。

本文组织如下：第 2 节讨论参数 $\beta$方向的估计量，并介绍其理论性质。我们开展了一项模拟研究，结果将在第 3 节报告。第 4 节中，我们将所提出的方法应用于一组真实数据，研究系列出口决定变量对专业化与运输

设备制造商在 2006年中国广东省出口市场参与的影响。第5节将讨论结论。引理和定理的证明将在附录中给出

# 2 所提出的估计量

考虑从以下模型中观测到样本 $(Y_i, \mathbf{X}_i)(i = 1, \cdots, n)$

$$E[Y_i \mid \mathbf{X}_i] = E[Y_i \mid \mathbf{X}_i^T \beta] = g(\mathbf{X}_i^T \beta), \tag{4}$$

其中， $g(\cdot)$ 是一个未知单调递增函数， $\beta$ 是 $R^p$.. 中的一个参数。

本文主要考虑参数 $\beta$ ($\beta \neq 0$) 的方向估计，其中 $\beta = 0$ 较为简单。由于模型（4）包括线性模型，因此模型（4）中 $\beta$ 的方向估计量可通过最小二乘法得到

$$D(\hat{\beta}) = \frac{\hat{\Sigma}_X^{-1} \times \sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}})}{\left|\left| \hat{\Sigma}_X^{-1} \times \sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}}) \right|\right|} \tag{5}$$

其中

$$\hat{\Sigma}_X = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})',$$

¯，而 X 是 X 的样本均值。对于响应变量为连续型或离散型的情形，均可使用最小二乘法估计半参数模型中的未知参数 $\beta$ 的方向。

在某些正则条件下， $\beta$ 的方向估计量是相合的。证明见附录。

定理 1. 当 X 服从均值为 $\mu$、协方差为 $Var$(X) $= \Sigma$(正定) 的椭圆分布时，对于模型（4）和 $E(g^2(\mathbf{X}'\beta)) < +\infty$，则

$$D(\hat{\beta}) \xrightarrow{P} \frac{\beta}{\|\beta\|}.$$

当 X 的均值为 $E\mathbf{X} = \mathbf{0}$ 时， $\beta$ 的方向可表示为

$$D^*(\hat{\beta}) = \frac{\hat{\Sigma}_X^{-1} \times \sum_{i=1}^{n} Y_i \mathbf{X}_i}{\left|\left|\hat{\Sigma}_X^{-1} \times \sum_{i=1}^{n} Y_i \mathbf{X}_i\right|\right|}, \tag{6}$$

a ˆ 以下定理将给出 $D^*(\beta)$的渐近分布。

定理2. 对于模型（4），当 X 服从均值为 0、协方差为 $Var(\mathbf{X}) = \Sigma$( 正定) 的椭圆分布，且 $E(g^2(\mathbf{X}'\beta)) < +\infty$和 $E(X_j^4) < +\infty$ 对于 $j = 1, \cdots, p$ 时，则

$$\sqrt{n}\left(D^*(\hat{\beta}) - \frac{\beta}{\|\beta\|}\right) \xrightarrow{d} \frac{1}{\|\beta\|}\left\{\lambda^{-1}\Sigma^{-1}\mathbf{U} - \Sigma^{-1}\mathbf{V} - \frac{\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U} - \beta'\Sigma^{-1}\mathbf{V}\right]}{\|\beta\|^2}\beta\right\}$$

其中

$$\lambda = E(Y\beta'\Sigma^{-1}\mathbf{X})/\|\beta\|^2$$

and

$$\begin{pmatrix}\mathbf{U}\\ \mathbf{V}\end{pmatrix} \sim N\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \Omega\right), \quad \Omega = Var\begin{pmatrix}Y\mathbf{X}\\ \mathbf{X}\mathbf{X}'\beta\end{pmatrix}.$$

ˆ 注. 随着样本数量增加， $D^*(\beta) - \beta$ 趋近于 0，因此， $D^*(\beta)$ 位于 $\beta$的切平面上。因此， $\frac{1}{\|\beta\|}\left\{\lambda^{-1}\Sigma^{-1}\mathbf{U} - \Sigma^{-1}\mathbf{V} - \frac{\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U} - \beta'\Sigma^{-1}\right]}{\|\beta\|^2}\beta\right\}$ $\mathbf{v}$在 $R^{p-1}$ 中服从退化正态分布。

ˆ 与定理2类似， $D(\beta)$有以下结果。

推论1. 对于模型（4），当 X 服从均值$\mu$ 、协方差为 $Var(\mathbf{X}) = \Sigma$( 正定) 的椭圆分布，且 $E(g^2(\mathbf{X}'\beta)) < +\infty$和 $E(X_j^4) < +\infty$ 对于 $j = 1, \cdots, p$ 时，则

$$\begin{aligned}\sqrt{n}\left(D(\hat{\beta}) - \frac{\beta}{\|\beta\|}\right) \xrightarrow{d} \frac{1}{\|\beta\|}\Bigg\{&\lambda^{-1}\Sigma^{-1}\mathbf{U} - \Sigma^{-1}\mathbf{V} - \lambda^{-1}\gamma\Sigma^{-1}\mathbf{Z}\\ &- \frac{\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U} - \beta'\Sigma^{-1}\mathbf{V} - \lambda^{-1}\gamma\beta'\Sigma^{-1}\mathbf{Z}\right]}{\|\beta\|^2}\beta\Bigg\}\end{aligned}$$

其中

$$\lambda = E(Y\beta'\Sigma^{-1}\mathbf{X})/\|\beta\|^2, \ \gamma = EY$$

and

$$\begin{pmatrix}\mathbf{Z}\\ \mathbf{U}\\ \mathbf{V}\end{pmatrix} \sim N\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{pmatrix}, \Xi\right), \quad \Xi = Var\begin{pmatrix}\mathbf{X}\\ Y\mathbf{X}\\ \mathbf{X}\mathbf{X}'\beta\end{pmatrix}.$$

基于推论1或定理2，我们考虑以下用于检验 $H_0$：$\beta = \beta_0$方向的Wald统计量，

$$W = \sqrt{n}\left(D(\hat{\beta}) - \frac{\beta}{\|\beta\|}\right)' \Sigma_\beta^+ \sqrt{n}\left(D(\hat{\beta}) - \frac{\beta}{\|\beta\|}\right),$$

其中， $\Sigma_\beta$是 $\sqrt{n}\left(D(\hat{\beta}) - \frac{\beta}{\|\beta\|}\right)$的协方差， $A^+$表示矩阵 $A$的Moore-Penrose逆。由于 $\|\beta\|$ 涉及协方差矩阵 $\Sigma_\beta$，因此我们对估计协方差矩阵 $\Sigma_\beta$施加约束 $\|\beta\| = 1$ 。在 $H_0$ 和$\|\beta_0\| = 1$条件下，Wald统计量为

$$W^* = n\left(D(\hat{\beta}) - \beta_0\right)' \Sigma_{D(\hat{\beta})}^+ \left(D(\hat{\beta}) - \beta_0\right) \sim \chi^2(m),$$

其中，自由度 $m$是矩阵$\Sigma_{D(\hat{\beta})}$的秩，该矩阵是 $\Sigma_\beta$的一个估计量，且 $m = p - 1$。因此，置信水平为 $1 - \alpha$ 时的接受域为

$$p\left(W^* \le \chi^2_{1-\alpha}(m)\right) = 1 - \alpha.$$

# 3 模拟研究

本节中，我们开展了若干模拟研究，以评估第 2 节中提出的参数方向估计量 $D(\beta)$的有限样本性能。

在第一个模拟研究中，我们假设响应变量 $Y$ 是离散的，且 $X$的维数为 $p = 3$，即 $X \sim N(0, \Sigma = (\sigma_{ij}))$， $\sigma_{ij} = \rho^{|i-j|}$， $i$， $j = 1$， 2， 3， 以及 $\beta = (\beta_1, \beta_2, \beta_3)' = (s_1, s_2, s_3)'/\|s\|$， $s_i \sim U(-1, 1)$， $i = 1$， 2， 3。我们主要考虑模型（4）中的以下数据生成情景。

- 情形 I. $Y = 1\{\mathbf{X}'\beta + \epsilon > 0\}$，其中 $\epsilon \sim N(0, 1)$。
- 情形 II. $Y = 1\{\mathbf{X}'\beta + \epsilon > 0\}$，其中 $\epsilon \sim t(1)$。
- 情形 III. $Y = 1\{\mathbf{X}'\beta + \epsilon > 0\}$，其中 $\epsilon \sim 0.4 * N(-3, 1) + 0.6 * N(2, 2$ ).

当误差项 $\epsilon$ 的分布未知且满足条件中位数约束时，Manski（1975）提出了参数$\beta$的最大得分（MS）估计量。Shen 等人（2023）提出了线性化最大秩相关（LMRC）估计量。此处，我们将所提出的估计量与 MS 估计量、LMRC估计量以及标准方法（Probit 回归）估计量进行比较。为了衡量参数$\beta$的真实方向 $\beta$ 与估计方向 $\beta$ 之间的距离，我们使用两个方向之间的夹角 $\theta$ 的余弦值（Bandiera 等人，2007；Rong 等人，2021），即

$$\cos(\theta) = \frac{\beta * D(\hat{\beta})}{\|\beta\| * \|D(\hat{\beta})\|}.$$

我们得到在情形 I–III 下，样本量分别为 100、300 和 500 时， $\beta$ 的估计量在 $\epsilon$ 不同分布下的余弦值和标准误（标准误）。基于 100 次重复以及不同 $\rho$ 的结果报告于表（1）中。

最大得分估计量的一个主要缺点是计算困难，因为优化的目标函数是非凸且非光滑的，因此很难找到全局最优解。此外，随着 $X$ 的维数增大，计算难度会更加突出（Khan 等人，2021年）。为考察协变量维数增加对估计结果的影响，我们考虑以下数据生成情景，即 $X \sim N(0, \Sigma = (\sigma_{ij}))$,$\sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, \cdots, p$。$Y = 1\{X^T\beta + \epsilon > 0\}$，其中 $p = 10$、15、 $\epsilon \sim t(1)$和$\beta = (\beta_1, \cdots, \beta_p)' = (s_1, \cdots, s_p)'/\|s\|$, $s_i \sim U(-1, 1)$, $i = 1, \cdots, p$。当 $X$ 的维数相对较大时，将提出的方法与 LMRC 方法和标准方法进行比较。对于两种不同的 $p$ 值，在 n=500 且重复 100 次的条件下，模拟结果列于表（2）。结果表明，当 X 的维数相对较高时（例如 $p = 10$、15），我们提出的方法仍然可行。所提出的估计量与 LMRC估计量相当，且优于标准估计量。

最后，我们假设响应变量 $Y$是连续的，并且 $X \sim N(0, \Sigma = (\sigma_{ij}))$, $\sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, 2, 3$, $\beta = (\beta_1, \beta_2, \beta_3)' = (s_1, s_2, s_3)'/\|s\|$, $s_i \sim U(-1, 1)$,$i = 1, 2, 3$, $\epsilon \sim N(0, 1)$。并考虑以下数据生成情景。

表 1：基于 100 次重复，在情形 I–III 中，样本量为 n=100、300、500 时所提出的估计量的标准误和余弦值（$\beta$）

| | n | $\rho$ | New | | MS | | LMRC | | S标准 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | cos | SE | cos | SE | cos | SE | cos | SE |
| 情形 I | 100 | -0.6 | 0.9612 | 0.0372 | 0.7326 | 0.3018 | 0.9602 | 0.0382 | 0.9629 | 0.0369 |
| | | -0.3 | 0.9705 | 0.0307 | 0.7545 | 0.2930 | 0.9700 | 0.0314 | 0.9720 | 0.0292 |
| | | 0 | 0.9805 | 0.0173 | 0.7879 | 0.2580 | 0.9807 | 0.0175 | 0.9810 | 0.0186 |
| | | 0.3 | 0.9713 | 0.0339 | 0.7546 | 0.2747 | 0.9711 | 0.0342 | 0.9713 | 0.0341 |
| | | 0.6 | 0.9733 | 0.0276 | 0.7402 | 0.3136 | 0.9730 | 0.0276 | 0.9751 | 0.0255 |
| | 300 | -0.6 | 0.9877 | 0.0148 | 0.7819 | 0.2633 | 0.9877 | 0.0149 | 0.9886 | 0.0127 |
| | | -0.3 | 0.9918 | 0.0087 | 0.7799 | 0.2593 | 0.9918 | 0.0086 | 0.9922 | 0.0084 |
| | | 0 | 0.9922 | 0.0069 | 0.7966 | 0.2457 | 0.9922 | 0.0070 | 0.9922 | 0.0071 |
| | | 0.3 | 0.9929 | 0.0073 | 0.7821 | 0.2680 | 0.9929 | 0.0073 | 0.9928 | 0.0074 |
| | | 0.6 | 0.9883 | 0.0141 | 0.7528 | 0.2918 | 0.9883 | 0.0142 | 0.9880 | 0.0148 |
| | 500 | -0.6 | 0.9932 | 0.0073 | 0.7902 | 0.3238 | 0.9931 | 0.0073 | 0.9932 | 0.0072 |
| | | -0.3 | 0.9963 | 0.0035 | 0.8358 | 0.2370 | 0.9962 | 0.0035 | 0.9961 | 0.0036 |
| | | 0 | 0.9958 | 0.0036 | 0.8464 | 0.2280 | 0.9958 | 0.0036 | 0.9960 | 0.0035 |
| | | 0.3 | 0.9957 | 0.0036 | 0.8848 | 0.2192 | 0.9957 | 0.0036 | 0.9958 | 0.0038 |
| | | 0.6 | 0.9930 | 0.0082 | 0.8229 | 0.2367 | 0.9930 | 0.0082 | 0.9932 | 0.0077 |
| 情形 II | 100 | -0.6 | 0.9142 | 0.1054 | 0.6699 | 0.3730 | 0.9144 | 0.1048 | 0.9133 | 0.1059 |
| | | -0.3 | 0.9449 | 0.0780 | 0.7181 | 0.3208 | 0.9449 | 0.0773 | 0.9449 | 0.0785 |
| | | 0 | 0.9478 | 0.0638 | 0.7513 | 0.2626 | 0.9472 | 0.0648 | 0.9475 | 0.0632 |
| | | 0.3 | 0.9420 | 0.0708 | 0.7228 | 0.2859 | 0.9427 | 0.0699 | 0.9419 | 0.0702 |
| | | 0.6 | 0.9186 | 0.0843 | 0.6363 | 0.3763 | 0.9194 | 0.0832 | 0.9183 | 0.0857 |
| | 300 | -0.6 | 0.9753 | 0.0319 | 0.7289 | 0.3403 | 0.9752 | 0.0323 | 0.9749 | 0.0328 |
| | | -0.3 | 0.9832 | 0.0172 | 0.7902 | 0.2512 | 0.9831 | 0.0174 | 0.9829 | 0.0178 |
| | | 0 | 0.9559 | 0.0467 | 0.7826 | 0.2481 | 0.9558 | 0.0468 | 0.9559 | 0.0459 |
| | | 0.3 | 0.9813 | 0.0228 | 0.7919 | 0.2321 | 0.9814 | 0.0228 | 0.9811 | 0.0229 |
| | | 0.6 | 0.9708 | 0.0353 | 0.7514 | 0.2748 | 0.9711 | 0.0351 | 0.9709 | 0.0353 |
| | 500 | -0.6 | 0.9862 | 0.0172 | 0.7942 | 0.2779 | 0.9861 | 0.0174 | 0.9860 | 0.0169 |
| | | -0.3 | 0.9909 | 0.0093 | 0.8342 | 0.2266 | 0.9909 | 0.0093 | 0.9907 | 0.0094 |
| | | 0 | 0.9899 | 0.0086 | 0.8421 | 0.2118 | 0.9899 | 0.0086 | 0.9895 | 0.0091 |
| | | 0.3 | 0.9902 | 0.0119 | 0.8315 | 0.2081 | 0.9903 | 0.0119 | 0.9900 | 0.0117 |
| | | 0.6 | 0.9851 | 0.0181 | 0.7999 | 0.2247 | 0.9851 | 0.0181 | 0.9846 | 0.0178 |
| 情形 III | 100 | -0.6 | 0.9454 | 0.0600 | 0.6863 | 0.3439 | 0.9441 | 0.0649 | 0.9463 | 0.0630 |
| | | -0.3 | 0.9664 | 0.0365 | 0.7421 | 0.2865 | 0.9657 | 0.0396 | 0.9672 | 0.0363 |
| | | 0 | 0.9704 | 0.0342 | 0.7461 | 0.2860 | 0.9708 | 0.0343 | 0.9713 | 0.0349 |
| | | 0.3 | 0.9712 | 0.0268 | 0.7338 | 0.2931 | 0.9705 | 0.0276 | 0.9717 | 0.0280 |
| | | 0.6 | 0.9518 | 0.0508 | 0.6830 | 0.3096 | 0.9511 | 0.0501 | 0.9533 | 0.0493 |
| | 300 | -0.6 | 0.9824 | 0.0185 | 0.7791 | 0.2321 | 0.9823 | 0.0186 | 0.9829 | 0.0185 |
| | | -0.3 | 0.9875 | 0.0129 | 0.8214 | 0.2012 | 0.9875 | 0.0129 | 0.9879 | 0.0126 |
| | | 0 | 0.9886 | 0.0136 | 0.8241 | 0.2285 | 0.9886 | 0.0136 | 0.9886 | 0.0142 |
| | | 0.3 | 0.9886 | 0.0119 | 0.8193 | 0.2012 | 0.9887 | 0.0118 | 0.9887 | 0.0123 |
| | | 0.6 | 0.9825 | 0.0206 | 0.7812 | 0.2299 | 0.9826 | 0.0206 | 0.9826 | 0.0213 |
| | 500 | -0.6 | 0.9908 | 0.0098 | 0.8249 | 0.2252 | 0.9908 | 0.0098 | 0.9910 | 0.0092 |
| | | -0.3 | 0.9942 | 0.0051 | 0.8329 | 0.2004 | 0.9942 | 0.0051 | 0.9944 | 0.0047 |
| | | 0 | 0.9941 | 0.0065 | 0.8101 | 0.2230 | 0.9941 | 0.0065 | 0.9942 | 0.0062 |
| | | 0.3 | 0.9940 | 0.0053 | 0.7729 | 0.2847 | 0.9940 | 0.0053 | 0.9942 | 0.0051 |
| | | 0.6 | 0.9906 | 0.0130 | 0.7424 | 0.3193 | 0.9906 | 0.0131 | 0.9907 | 0.0123 |

表2：基于100次重复，在不同的 $X$ 维度下，样本量为n=500时所提出的估计量的标准误和余弦值（ $\beta$ ）。

| 维数 | $\rho$ | New | | LMRC | | 标准 | |
|---|---|---|---|---|---|---|---|
| | | cos | SE | cos | SE | cos | SE |
| 10 | -0.6 | 0.9265 | 0.0397 | 0.9263 | 0.0397 | 0.9259 | 0.0395 |
| | -0.3 | 0.9490 | 0.0275 | 0.9489 | 0.0277 | 0.9483 | 0.0280 |
| | 0 | 0.9588 | 0.0195 | 0.9589 | 0.0194 | 0.9581 | 0.0200 |
| | 0.3 | 0.9506 | 0.0284 | 0.9506 | 0.0283 | 0.9500 | 0.0295 |
| | 0.6 | 0.9150 | 0.0504 | 0.9147 | 0.0507 | 0.9141 | 0.0514 |
| 15 | -0.6 | 0.8950 | 0.0516 | 0.8951 | 0.0518 | 0.8942 | 0.0525 |
| | -0.3 | 0.9283 | 0.0327 | 0.9283 | 0.0326 | 0.9273 | 0.0330 |
| | 0 | 0.9351 | 0.0215 | 0.9350 | 0.0216 | 0.9348 | 0.0218 |
| | 0.3 | 0.9264 | 0.0332 | 0.9263 | 0.0332 | 0.9255 | 0.0334 |
| | 0.6 | 0.8883 | 0.0438 | 0.8882 | 0.0439 | 0.8858 | 0.0455 |

- 情形1。 $Y = \mathbf{X}'\beta + \epsilon$。
- 情形2。 $Y = \Phi(\mathbf{X}'\beta) + \epsilon$。
- 情形3。 $Y = \log(1 + \exp(\mathbf{X}'\beta)) + \epsilon$。
- 情形4。 $Y = \frac{1}{1+\exp(-\mathbf{X}'\beta)} + \epsilon$。

我们将所提出的估计量与LMRC估计量进行比较。表（3）报告了基于100次重复、样本量为=100、300、500时，不同 $\epsilon$分布下，所提出的估计量在 $\beta$ 和不同 $\rho$取值下的余弦值和标准误。从表（3）可以看出，随着样本量增加，所有余弦值的标准误均减小，且余弦值更接近1。尽管当模型为线性模型时，所提出的估计量略逊于LMRC估计量，但当模型为非线性模型时，所提出的估计量优于LMRC估计量。

# 4 真实数据

在本节中，我们将提出的方法应用于一组真实数据，该数据研究了2006年中国广东省专业设备和运输设备制造商的系列出口决定变量对出口市场参与的影响（Baltagi等人，2022年）。数据可从中国国家统计局（NBS）获取。

表3：基于100次重复，对情形1–4中样本量为 n=100、300、500 的所提出的估计量的标准误和余弦值 $\beta$。

| n | $\rho$ | | 情形1 | | 情形2 | | 情形3 | | 情形4 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | cos | SE | cos | SE | cos | SE | cos | SE |
| 100 | -0.3 | New | 0.9893 | 0.0109 | 0.8640 | 0.1500 | 0.9558 | 0.0492 | 0.7680 | 0.2516 |
| | | LMRC | 0.9880 | 0.0117 | 0.8621 | 0.1537 | 0.9513 | 0.0507 | 0.7633 | 0.2666 |
| | 0 | New | 0.9902 | 0.0110 | 0.8841 | 0.1186 | 0.9590 | 0.0465 | 0.7844 | 0.2522 |
| | | LMRC | 0.9896 | 0.0108 | 0.8813 | 0.1298 | 0.9565 | 0.0467 | 0.7858 | 0.2562 |
| | 0.3 | New | 0.9899 | 0.0097 | 0.8855 | 0.1264 | 0.9607 | 0.0432 | 0.7913 | 0.2719 |
| | | LMRC | 0.9882 | 0.0126 | 0.8790 | 0.1479 | 0.9570 | 0.0494 | 0.7834 | 0.2865 |
| 300 | -0.3 | New | 0.9960 | 0.0046 | 0.9487 | 0.0677 | 0.9828 | 0.0216 | 0.9046 | 0.1164 |
| | | LMRC | 0.9953 | 0.0052 | 0.9450 | 0.0670 | 0.9808 | 0.0242 | 0.8939 | 0.1311 |
| | 0 | New | 0.9967 | 0.0031 | 0.9595 | 0.0375 | 0.9867 | 0.0127 | 0.9252 | 0.0704 |
| | | LMRC | 0.9964 | 0.0034 | 0.9586 | 0.0412 | 0.9860 | 0.0133 | 0.9245 | 0.0777 |
| | 0.3 | New | 0.9965 | 0.0038 | 0.9561 | 0.0491 | 0.9855 | 0.0169 | 0.9198 | 0.0875 |
| | | LMRC | 0.9960 | 0.0044 | 0.9545 | 0.0509 | 0.9838 | 0.0187 | 0.9152 | 0.0964 |
| 500 | -0.3 | New | 0.9979 | 0.0020 | 0.9739 | 0.0261 | 0.9912 | 0.0085 | 0.9512 | 0.0514 |
| | | LMRC | 0.9978 | 0.0023 | 0.9733 | 0.0260 | 0.9907 | 0.0091 | 0.9500 | 0.0487 |
| | 0 | New | 0.9981 | 0.0017 | 0.9775 | 0.0198 | 0.9923 | 0.0070 | 0.9589 | 0.0357 |
| | | LMRC | 0.9980 | 0.0020 | 0.9763 | 0.0209 | 0.9917 | 0.0076 | 0.9559 | 0.0384 |
| | 0.3 | New | 0.9979 | 0.0019 | 0.9726 | 0.0284 | 0.9914 | 0.0088 | 0.9488 | 0.0553 |
| | | LMRC | 0.9977 | 0.0021 | 0.9718 | 0.0280 | 0.9905 | 0.0092 | 0.9468 | 0.0532 |

在后续分析中，我们主要考虑的变量包括 expd-ford（=1 如果公司是出口商；否则为 0）、lemp（企业规模的对数）、lprod（每名工人产出的对数）、lcapint（资本除以总销售额）、intastr（无形资产占总资产的比例）、cmp（销售额与营业利润之比的对数）、cmp² （cmp 的平方）、ltastx（固定出口成本）以及 sez（=1 如果企业位于经济特区； =0 否则）。最终，我们得到一个包含1614家公司的样本，这些公司在2006年的年度总销售额均至少为500万元人民币（约合70万美元）。正如 Khan 等人（2021年）所述，当协变量维数较大时，最大得分法极难计算。本分析共有8个协变量，因此这里不考虑最大得分估计量。

为说明我们的方法，我们令响应变量（即 Y）为 expd- ford， $X$ 包括其余变量。提出的方法、LMRC 方法和标准方法的估计结果见表（4）。根据表（4），我们可以解释一些经济假设。显然，简单检验 $T = \hat{\beta}_i / \hat{\sigma}_i, i, \cdots, 8$ 表明，对于提出的方法而言，在 0.05 的显著性水平下，lemp、lcapint、ltastx、sez 对响应变量有显著影响。规模较大的

表4：采用不同方法估计的系列出口决定变量对专业设备和运输设备制造商出口市场参与的影响。

| | 参数 | 提出的方法 | Probit模型 | LMRC |
|---|---|---|---|---|
| | lemp | 0.7621(0.0911) | 0.8785(0.1210) | 0.8934(0.0885) |
| | lprod | 0.0236(0.0586) | 0.0231(0.0677) | 0.0187(0.0668) |
| | lcapint | 0.1463(0.0592) | 0.1796(0.0703) | 0.1957(0.0671) |
| | intastr | -0.0844(0.0456) | -0.0960(0.0524) | -0.1032(0.0514) |
| | Cmp | -0.1228(0.2357) | -0.2154(0.3172) | -0.2823(0.2461) |
| | $Cmp^2$ | 0.0846(0.2393) | 0.1478(0.3494) | 0.1878(0.2540) |
| | ltastx | -0.1766(0.0552) | -0.1981(0.0641) | -0.1925(0.0601) |
| | sez | 0.5802(0.1081) | 0.2806(0.0623) | 0.0228(0.0364) |

固定成本（通过更高的资本密集度）更重要的企业，往往更有可能出口。固定出口成本越高，出口盈利能力就越低，这解释了出口影响的减弱。与 LMRC 相比，提出的方法发现 sez 是影响响应变量的一个新因素。由于企业与经济特区之间的距离也会影响出口，因此该二元变量确实有显著影响，所以我们提出的方法更为合理。

# 5 结论与讨论

本文研究响应变量为连续型或离散型时，半参数模型中未知参数方向的估计问题。我们提出使用最小二乘法估计半参数模型中的未知参数方向。所提出的估计量计算简单，并具有闭式表达式。我们证明了所提出的估计量具有一致性且渐近正态。对于二元响应变量，所提出的估计显著优于最大得分估计；对于 Probit 模型，其表现与线性化最大秩相关（LMRC）估计和最大似然估计相当。当误差项的分布具有厚尾（即 Student t分布），或具有无穷矩（即柯西分布）时，所提出的估计量表现良好。对于非线性模型中的连续响应变量，所提出的估计优于线性化最大秩

相关估计。此外，若关注链接函数 $g(\cdot)$ 的估计，则可通过非参数方法结合 $E[Y \mid X^T \hat{\beta}]$ 直接估计。

本文主要考虑链接函数 $g(\cdot)$ 单调递增的情况。当链接函数单调递减时，可以用类似的方法得到未知参数方向的估计量；此时，未知参数方向的估计与参数的真实方向相反。在实际应用中，由于各种原因，可观测协变量通常具有高维性，因此我们将提出的方法扩展到含有高维协变量的模型（4），以估计参数方向。我们可以通过最小化损失函数与惩罚函数之和来估计参数方向 $\beta$。这一研究问题将在未来加以探讨。

# 附录

Brillinger（1983）、Li 和 Duan（1989）以及 Li（2018，定理 8.3）也给出了以下类似结果，本文则提供了另一种证明方法。

引理 A. 对于模型（4），若 X 服从均值为 0、协方差为 $\mathrm{Var}(\mathbf{X}) = \Sigma$（正定）的椭圆分布，则

$$E(Y\mathbf{X}) = \lambda \Sigma \beta,$$

其中 $\lambda > 0$。

证明：

$$E(Y\mathbf{X}) = E[g(\mathbf{X}'\beta)\mathbf{X}] = E[(g(\mathbf{X}'\beta) - g(0))\mathbf{X}].$$

(I) 在 $\beta = 0$ 的情况下，这显然可以证明。(II) 在 $\beta \neq 0$ 和 $\Sigma = \sigma^2 I$ 的情况下。设 $A$ 为第一行为 $\beta'/\|\beta\|$ 的正交矩阵，并且

$$\mathbf{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_p \end{pmatrix} = A\mathbf{X}.$$

那么根据 W 的定义，可得 $W_1 = \beta'/\|\beta\|$X，并且

$$E(Y\mathbf{X}) = E[(g(\mathbf{X}'\beta) - g(0))\mathbf{X}] = A'E[(g(\mathbf{X}'\beta) - g(0))A\mathbf{X}] = A'E[(g(W_1\|\beta\|) - g(0))\mathbf{W}].$$

由于 X 服从均值为 0、方差为 Var(X) $=\sigma^2 I$ 的椭圆分布，W $= A$X 服从均值为 0、方差为$\sigma^2 AA' = \sigma^2 I$ 的椭圆分布，因此

$$E(W_i \mid W_1) = 0, i = 2, \cdots, p,$$

根据 Frahm（2004）给出的定理6。因此

$$E[(g(\|\beta\|W_1) - g(0))W_i] = 0, i = 2, \cdots, p,$$

$$E[(g(\|\beta\|W_1) - g(0))W_1] > 0$$

由 $g$(·) 严格递增，且

$$\begin{aligned}
E(Y\mathbf{X}) =& A'E[(g(\|\beta\|W_1) - g(0))W] \\
=& A' \begin{pmatrix} E[(g(\|\beta\|W_1) - g(0))W_1] \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
=& \frac{\beta}{\|\beta\|} \times E[(g(\|\beta\|W_1) - g(0))W_1] \\
=& \lambda\sigma^2 I\beta \\
=& \lambda\Sigma\beta,
\end{aligned}$$

其中

$$\lambda = \frac{E[(g(\|\beta\|W_1) - g(0))W_1]}{\sigma^2\|\beta\|} > 0.$$

(III) 一般而言，$\beta \neq 0$且 Var($X$) $= \Sigma$。令 X$^*$ $= \Sigma^{-1/2}$X 且 $\beta^* = \Sigma^{1/2}\beta$。则 X$^*$ 服从均值为 0、方差为 $I$ 的椭圆分布。根据模型（4），可得

$$E[Y \mid \mathbf{X}] = E[Y \mid \mathbf{X}^*] = g(\mathbf{X}^{*'}\beta^*).$$

因此

$$E(Y\mathbf{X}) = \Sigma^{1/2}E(Y\mathbf{X}^*) = \Sigma^{1/2} \times \lambda I\beta^* = \lambda\Sigma\beta$$

根据情形（II）和 $\lambda > 0$。

定理 1 的证明：根据大数定律，我们有

$$\frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}}) = \frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{X}_i - \mu) + (\mu - \bar{\mathbf{X}})\bar{Y} \xrightarrow{p} E(Y\mathbf{X}) = \lambda\Sigma\beta.$$

根据引理 A 和

$$\hat{\Sigma} \xrightarrow{p} \Sigma.$$

因此，我们可以得到

$$\hat{\Sigma}_X^{-1} \times \frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}}) \xrightarrow{p} \Sigma^{-1} \times \lambda\Sigma\beta = \lambda\beta, \lambda > 0.$$

也就是说

$$D(\hat{\beta}) = \frac{\hat{\Sigma}_X^{-1} \times \frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}})}{\|\hat{\Sigma}_X^{-1} \times \frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{X}_i - \bar{\mathbf{X}})\|} \xrightarrow{p} \frac{\lambda\beta}{\|\lambda\beta\|} = \frac{\beta}{\|\beta\|}.$$

ˆ 因此，所提出的估计量 $D(\beta)$ 是相合的。

定理2 的证明：令

$$\mathbf{U}_n = \frac{1}{n}\sum_{i=1}^{n} Y\mathbf{X}_i, \quad \mathbf{V}_n = \hat{\Sigma}_X\beta = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'\beta.$$

由引理 A 和中心极限定理可知 $E\mathbf{U}_n = \lambda\Sigma\beta$ 且

$$\sqrt{n}\begin{pmatrix}\mathbf{U}_n - E[U_n] \\ \mathbf{V}_n - \Sigma\beta\end{pmatrix} = \begin{pmatrix}\mathbf{U}_n - \lambda\Sigma\beta \\ \mathbf{V}_n - \Sigma\beta\end{pmatrix} \xrightarrow{d} \begin{pmatrix}\mathbf{U} \\ \mathbf{V}\end{pmatrix} \sim N\left(\begin{pmatrix}\mathbf{0} \\ \mathbf{0}\end{pmatrix}, \Omega\right), \tag{7}$$

其中

$$\Omega = Var\begin{pmatrix}Y\mathbf{X} \\ \mathbf{X}\mathbf{X}'\beta\end{pmatrix}.$$

根据所提出的估计量的形式，我们有

$$\begin{aligned} D^*(\hat{\beta}) = \frac{\hat{\Sigma}_X^{-1}\mathbf{U}_n}{\left\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right\|} &= \frac{\hat{\Sigma}_X^{-1}\left(\mathbf{U}_n - E\mathbf{U}_\mathbf{n}\right) + \left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n - \lambda\beta\right) + \lambda\beta}{\left\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right\|} \\ &= \frac{\hat{\Sigma}_X^{-1}\left(\mathbf{U}_n - E\mathbf{U}_n\right) + \left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n - \lambda\beta\right)}{\left\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right\|} + \frac{\lambda\beta}{\left\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right\|}. \end{aligned}$$

ˆ 为了获得 $D^*(\beta)-\beta$ 的渐近分布，我们首先需要证明 $\sqrt{n}(\mathbf{U}_n-E\mathbf{U}_n)$、$\sqrt{n}\left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n-\lambda\beta\right)$ 和 $\sqrt{n}\left(\frac{\lambda\beta}{\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\|}-\frac{\beta}{\|\beta\|}\right)$ 的渐近分布。由（7）可得 $\sqrt{n}(\mathbf{U}_n-E\mathbf{U}_n)\xrightarrow{d}\mathbf{U}\sim N(\mathbf{0},Var(\mathbf{X}))$。接下来，我们将分别证明 $\sqrt{n}\left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n-\lambda\beta\right)$ 和 $\sqrt{n}\left(\frac{\lambda\beta}{\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\|}-\frac{\beta}{\|\beta\|}\right)$ 的渐近性质。

$$\begin{aligned}\sqrt{n}\left(\hat{\Sigma}_X^{-1}E[\mathbf{U}_n]-\lambda\beta\right) &= \sqrt{n}\left(\hat{\Sigma}_X^{-1}\lambda\Sigma\beta-\lambda\beta\right)\\ &= \lambda\hat{\Sigma}_X^{-1}\sqrt{n}\left(\Sigma\beta-\hat{\Sigma}_X\beta\right)\\ &= -\lambda\hat{\Sigma}_X^{-1}\sqrt{n}(\mathbf{V}_n-\Sigma\beta)\\ &\xrightarrow{d} -\lambda\Sigma^{-1}\mathbf{V}.\end{aligned}$$

Let

$$\mathbf{S}_n(\mathbf{U}_n,\Sigma_X^{-1})=\sqrt{n}\left(\frac{\lambda\beta}{\left\|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right\|}-\frac{\beta}{\|\beta\|}\right)=\sqrt{n}\left(\frac{\lambda}{\sqrt{\mathbf{U}_n'\hat{\Sigma}_X^{-2}\mathbf{U}_n}}-\frac{1}{\|\beta\|}\right)\beta.$$

并在 $(\lambda\Sigma\beta,\Sigma^{-1})$ 处展开 $\mathbf{S}_n(\mathbf{U}_n,\Sigma_X^{-1})$，我们可以得到

$$\begin{aligned}\mathbf{S}_n(\mathbf{U}_n,\Sigma_X^{-1}) &= \sqrt{n}\left[-\lambda^{-1}\beta'\Sigma^{-1}(\mathbf{U}_n-E\mathbf{U}_n)-\beta'\Sigma\Sigma^{-1}(\hat{\Sigma}_X^{-1}-\Sigma^{-1})\Sigma\beta+o_p(n^{-1/2})\right]\frac{\beta}{\|\beta\|^3}\\ &= \sqrt{n}\left[-\lambda^{-1}\beta'\Sigma^{-1}(\mathbf{U}_n-E\mathbf{U}_n)-\beta'(\hat{\Sigma}_X^{-1}-\Sigma^{-1})\Sigma\beta+o_p(n^{-1/2})\right]\frac{\beta}{\|\beta\|^3}\\ &= \sqrt{n}\left[-\lambda^{-1}\beta'\Sigma^{-1}(\mathbf{U}_n-E\mathbf{U}_n)-\beta'\hat{\Sigma}_X^{-1}(\Sigma-\hat{\Sigma}_X)\beta+o_p(n^{-1/2})\right]\frac{\beta}{\|\beta\|^3}\\ &= \sqrt{n}\left[-\lambda^{-1}\beta'\Sigma^{-1}(\mathbf{U}_n-E\mathbf{U}_n)-\beta'\Sigma_n^{-1}(\Sigma\beta-\mathbf{V}_n)+o_p(n^{-1/2})\right]\frac{\beta}{\|\beta\|^3}\\ &\xrightarrow{d} -\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U}-\beta'\Sigma^{-1}\mathbf{V}\right]\frac{\beta}{\|\beta\|^3}.\end{aligned}$$

因此,

$$\begin{aligned}\sqrt{n}\left(D^*(\hat{\beta})-\frac{\beta}{\|\beta\|}\right) &= \frac{\hat{\Sigma}_X}{\|\hat{\Sigma}_X\mathbf{U}_n\|}\sqrt{n}(\mathbf{U}_n-E[U_n])+\frac{1}{\|\hat{\Sigma}_X\mathbf{U}_n\|}\sqrt{n}(\left(\hat{\Sigma}_X^{-1}E[\mathbf{U}_n]-\lambda\beta\right))\\ &\quad+\sqrt{n}\left(\frac{\lambda\beta}{\left|\hat{\Sigma}_X^{-1}\mathbf{U}_n\right|}-\frac{\beta}{\|\beta\|}\right)\\ &\quad\xrightarrow{d}\frac{1}{\|\beta\|}\left\{\lambda^{-1}\Sigma^{-1}\mathbf{U}-\Sigma^{-1}\mathbf{V}-\frac{\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U}-\beta'\Sigma^{-1}\mathbf{V}\right]}{\|\beta\|^2}\beta\right\},\end{aligned}$$

其中

$$\lambda=E(Y\beta'\Sigma^{-1}\mathbf{X})/\|\beta\|^2,\quad\begin{pmatrix}\mathbf{U}\\\mathbf{V}\end{pmatrix}\sim N\left(\begin{pmatrix}\mathbf{0}\\\mathbf{0}\end{pmatrix},\Omega\right),\quad\Omega=Var\begin{pmatrix}Y\mathbf{X}\\\mathbf{X}\mathbf{X}'\beta\end{pmatrix}.$$

# 参考资料


Abrevaya, J., Huang, J.（2005）。关于最大得分估计量的自助法。*Econometrica*, 73(4), 1175-1204。

Bandiera, F., De Maio, A., Ricci, G.（2007）。采用圆锥拒斥的自适应恒虚警率雷达检测。 *IEEE Transactions on Signal Processing*, 55(6), 2533-2541。

Baltagi, B. H., Egger, P. H., Kesina, M.（2022年）。具有高阶网络相互依赖的多元面板 Probit 模型的贝叶斯估计，以及其在广东企业的全球市场参与中的应用。 *Journal of Applied Econo-metrics*, 37(7), 1356-1378。

Brillinger, D. R.（1983），“具有‘高斯’回归变量的广义线性模型。”载于《Erick L. Lehmann纪念文集》，加利福尼亚州贝尔蒙特：Wadsworth，第97-114页。

Chaudhuri, P., Doksum, K., Samarov, A.（1997）。论平均导数分位数回归。 *The Annals of Statistics*，25(2)，715-744。

Cui, X., H¨ardle, W. K., Zhu, L. (2011). 单指标模型的 EFM 方法。 *The Annals of Statistics*, 39(3), 1658-1688。

Fan, Y., Han, F., Li, W., Zhou, X. H. (2020). 关于递增维度下的秩估计量。 *Journal of Econometrics*, 214(2), 379-412。Frahm, G. (2004). 广义椭圆分布：理论与应用（博士论文，Universit¨at zu K¨oln）。Gao, W. Y., Xu, S. (2022). 两阶段最大得分估计量。arXiv 预印本 arXiv:2009.02854。Han, A. K. (1987). 广义回归模型的非参数分析：最大秩相关估计量。 *Journal of Econometrics*, 35(2-3), 303-316。Horowitz, J. L. (1992). 二元响应模型的平滑最大得分估计量。 *Econometrica*, 505-531。Hristache, M., Juditsky, A., Spokoiny, V. (2001). 单指标模型中指数系数的直接估计。 *Annals of Statistics*, 595-623。Ichimura, H. (1993). 单指标模型的半参数最小二乘法（SLS）与加权 SLS 估计。 *Journal of econometrics*, 58(1-2), 71-120。Khan, S., Lan, X., Tamer, E. (2021). 通过迭代凸优化估计高维单调指数模型1。arXiv 预印本 arXiv:2110.04388。Kuchibhotla, A. K., Patra, R. K. (2020). 通过平滑样条实现单指标模型的有效估计。 *Bernoulli*, 26(2), 1587-1618。Kuchibhotla, A. K., Patra, R. K., Sen, B. (2021). 凸性约束单指数模型中的半参数效率。 *Journal of the American Statistical Association*, 1-15。Li, B. (2018). 充分降维：方法及其在 R 中的应用。 Chapman & Hall/CRC。Li, K.-C. and Duan, N.(1989). 链接函数设定错误下的回归分析。 *Ann. Stat.*, 17(3), 1009-1052。

Li, Q. 和 Racine, J.S.（2007）。《非参数计量经济学：理论与实践》。新泽西州普林斯顿：普林斯顿大学出版社。Liu, J., Zhang, R., Zhao, W., Lv, Y.（2013）。一种针对单指标模型的稳健有效估计方法。 *Journal of Multivariate Analysis*, 122，226-238。Manski, C. F.（1975）。随机效用选择模型的最大得分估计。 *Journal of econometrics*，3(3)，205-228。

Naik, P., Tsai, C. L.（2000）。单指标模型的偏最小二乘估计量。 *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4)，763-771。

Park, H., Petkova, E., Tarpey, T., Ogden, R. T.（2020年）。带有多重链接函数的单指标模型。 *Journal of statistical planning and inference*，205，115-128。

Patra, R. K., Seijo, E., Sen, B. (2018年). 最大得分估计量的相合自助法程序. *Journal of Econometrics*, 205(2), 488-507.

Rong, Y., Aubry, A., De Maio, A., Tang, M. (2021年). 基于不变性理论的低秩异质杂波自适应雷达检测. *IEEE Transactions onSignal Processing*, 69, 1492-1506.

Shen, G., Chen, K., Huang, J., Lin, Y. (2023). 线性化最大秩相关估计。 Biometrika. 110 (1), 187-203, Wang, L., Cao, G. (2018). 广义部分线性单指标模型的有效估计。 *Bernoulli*, 24, 1101-1127. Wu, T. Z., Yu, K., Yu, Y. (2010). 单指标分位数回归。*Journal ofMultivariate Analysis*, 101(7), 1607-1621. Yang, J., Tian, G., Lu, F., Lu, X. (2020). 基于外积梯度的单指标众数回归。 *Computational Statistics and Data Analysis*, 144, 106867.

Yu, Y., Ruppert, D. (2002). 部分线性单指标模型的惩罚样条估计。
*Journal of the American Statistical Association*, 97(460), 1042- 1054.

Zhou, J. 和 He, X. (2008)。基于约束典型相关和变量筛选的降维。
*Annals of Statistics*, 36, 1649-1668.

Zhou, L., Lin, H., Chen, K., Liang, H. (2019)。广义半／非参数回归模型中参数和非参数函数的有效估计与计算。 *Journal of Econometrics*, 213(2), 593-607.

Zou, Q. 和 Zhu, Z. (2014)。使用B样条的单指数模型M估计量。*Metrika*, 77, 225-246.