Estimation of the directions for unknown parameters in semiparametric models

Han, Jinyue and Wang, Jun and Gao, Wei and Tang, Man-Lai

Northeast Normal University, Yunnan University, Brunel University

13 February 2023

Estimation of the directions for unknown parameters in semiparametric models

Jinyue Han[1] Jun Wang[2] Wei Gao[1*] Man-Lai Tang[3]
1. School of Mathematics and Statistics, Northeat Normal University
2. School of Mathematics and Statistics, Yunnan University
3. Department of Mathematics, Brunel University
* gaow@nenu.edu.cn

Abstract

Semiparametric models are useful in econometrics, social sciences and medicine application. In this paper, a new estimator based on least square methods is proposed to estimate the direction of unknown parameters in semi-parametric models. The proposed estimator is consistent and has asymptotic distribution under mild conditions without the knowledge of the form of link function. simulations show that the proposed estimator is significantly superior to maximum score estimator given by Manski (1975) for binary response variables. When the error term is long-tailed distributions or distribution with no moments, the proposed estimator perform well. Its application is illustrated with data of exportibg participation of manufactures in Guangdong

Key Words: Binary model, direction, least squares estimator, maximum score, semi-parametric models, single index model.

1 Introduction

Considering the problem of estimating the regression model $E(Y \mid \mathbf{X})$, where Y denotes response variable, $\mathbf{X}$ is p-dimensional observable covariates. Model $E(Y \mid \mathbf{X})$ has a significant applications in economics, medicine, and other fields and the estimation of $E(Y \mid \mathbf{X})$ is a key problem. While nonparametric methods are flexible, the price is high: the estimation precision decreases rapidly as p increasing and the estimating results can be hard to interpret when dimension of covariate $\mathbf{X}$ is greater. So to avoids the curse of dimensionality for nonparametric model while still offering flexibility in the functional form of $E(Y \mid \mathbf{X})$, a natural way is to assume that $E[Y \mid \mathbf{X}]$ is a semi-parametric model. A popular semiparametric model is given by

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

where $\beta \in R^p$ and $g : R \to R$, a unknown link function. When $g(\cdot)$ is known, the generalized moment estimation method is used to estimate unknown parameters β. In this paper, here assume that the link function $g(\cdot)$ is unknown and estimate the direction of parameters β.

A special case of models (1) is the binary response model, i.e.,

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

where ϵ is unobservable random variable. The unknown parameters can be estimated via maximum likelihood methods when the conditional distribution of ϵ given $\mathbf{X}$ is known such as logistic Models or Probit Models. When the conditional distribution of ϵ is unknow, Manski (1975) proposed the maximum score estimator for binary response models with a conditional median restriction, i.e., $\text{med}(\epsilon \mid \mathbf{X}) = 0$, where $\text{med}(\epsilon \mid \mathbf{X})$ denotes the conditional median of ϵ given $\mathbf{X}$. Base on results given by Manski, Horowitz (1992) proposed the smoothed maximum score estimator, proved it was asymptotically normally distributed under certain assumptions and the classical bootstrap was applied to make inference. Abrevaya and Huang (2005) showed that the classical bootstrap was inconsistent for the maximum estimator. Patra et al. (2018) proposed a model-based smooth bootstrap process for making statistical inference on the maximum score estimator and proved its consistency. Gao *et al.* (2022) proposed the two-stage maximum score estimator.

Another special form of model (1) is classical single index models

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

There have mainly two kinds of techniques for estimation in single index models. One kind is M-estimation methods, which is based on kernel estimator (Ichimura, 1993),regression splines (Park *et al.*, 2020), local-linear approximation (Zhou *et al.*, 2019), penalized splines (Yu *et al.*, 2002), and smoothing splines (Kuchibhotla and Patra, 2020) to estimate $g(\cdot)$, and minimize some appropriate criterion function such as quadratic loss (Yu and Ruppert, 2002), quantile regression (Wu *et al.*, 2010), estimating function method(Cui*et al.*, 2011), robust L_1 loss (Zou and Zhu, 2014),, quasi-likelihood (Wang and Guo, 2019), profiled likelihood (Patra *et al.*, 2020) and modal regression (Yang *et al.*, 2020)to obtain β. The other kind is direct estimation methods such as maximum rank correlation estimators (Han, 1987) average derivative estimators (Hristache *et al.*, 2001), dimension reduction techniques (Li and Racine, 2007), partial least squares (Naik and Tsai, 2000) and linearized maximum rank correlation estimators(Shen *et al.*, 2022). Recently Kuchibhotla *et al.* (2021) introduced a convex and Lipschitz constrained least-square estimator (CLSE) for both the parametric and the nonparametric components given independent and identically distributed observations.

models (1) contains the linear model and a least squares method is proposed to estimate the direction of the parameter β in semiparametric models for the response variable being ontinuous or discrete. The proposed method is computationally simple and theoretical reliable. Simulation results shows that the proposed estimator is significantly superior to maximum score estimation for response variables being discrete, and is comparable with linearized maximum rank correlation(LMRC) and them maximum likelihood estimation methods for Probit Models. When the distribution of error term is long-tailed distributions (i.e., Student t) and distributions with no existing moments (i.e., Cauchy), the proposed estimator and LMRC estimator perform better than standard estimator. When the dimension of covariates is relatively high, the proposed method is still feasible. The proposed estimation is superior to the linearized maximum rank correlation estimation with nonlinear models.

This paper is organized as follows. Estimators for the direction of the parameter β are considered and their theoretical properties are presented in Section 2. A simulation study will be conducted, and the results will be reported in Section 3. In Section 4, we apply our methodology to a real data set that studies the influence of series exporting determined variables on the export-market participation of specialized and transport facility manufactures in the province of Guangdong, China in 2006. Conclusions and discussions will be discussed in Section 5. Proofs of lemma and theorems will be presented in the Appendix

2 The Proposed estimator

Consider the samples $(Y_i, \mathbf{X}_i)(i = 1, \cdots, n)$ are observed from the following models

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

where $g(\cdot)$ is a unknown and monotonically increasing function function and β is a parameter in R^p..

In this paper, we mainly consider that the estimation of the direction of parameters β ($\beta \neq 0$), in case of $\beta = 0$ which is simple. Because models (4) includes the linear model, so the estimator for the direction of β in model (4) is obtained via the least squares method

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

where

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

and $\bar{\mathbf{X}}$ is sample mean value of $\mathbf{X}$. The least squares method can be used to estimate the direction of unknown parameters β in semi-parametric model for the response variable being continuous or discrete.

Under some regular conditions, the estimator of direction of β is consistent. The proof is given in the Appendix.

Theorem 1. When $\mathbf{X}$ is distributed by the elliptical distributions with mean μ and covariance $Var(\mathbf{X}) = \Sigma($ positive definite) for Models (4) and $E(g^2(\mathbf{X}'\beta)) < +\infty$, then

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

When the mean of $\mathbf{X}$ is $E\mathbf{X} = \mathbf{0}$, the direction of β can be given by

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

and the following theorem will give the asymptotic distribution of $D^*(\hat{\beta})$.

Theorem 2. When $\mathbf{X}$ is distributed by the elliptical distributions with mean $\mathbf{0}$ and covariance $Var(\mathbf{X}) = \Sigma($ positive definite) for Models (4), $E(g^2(\mathbf{X}'\beta)) < +\infty$ and $E(X_j^4) < +\infty$ for $j = 1, \cdots, p$, then

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

where

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

and

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

Remark. $D^*(\hat{\beta}) - \beta$ is close to 0 with the increase of samples, therefor, $D^*(\hat{\beta})$ is in the tangent plane of β. Therefor, $\frac{1}{||\beta||}\left\{\lambda^{-1}\Sigma^{-1}\mathbf{U} - \Sigma^{-1}\mathbf{V} - \frac{\left[\lambda^{-1}\beta'\Sigma^{-1}\mathbf{U} - \beta'\Sigma^{-1}\mathbf{V}\right]}{||\beta||^2}\beta\right\}$ is a degenerate normal distribution in R^{p-1}.

Similar to Theorem 2, $D(\hat{\beta})$ has the following results.

Corollary 1. When $\mathbf{X}$ is distributed by the elliptical distributions with mean μ and covariance $Var(\mathbf{X}) = \Sigma($ positive definite) for Models (4), $E(g^2(\mathbf{X}'\beta)) < +\infty$ and $E(X_j^4) < +\infty$ for $j = 1, \cdots, p$, then

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

where

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

and

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

3 Simulation studies

In this section, we conduct several simulation studies to evaluate the finite sample performance of the proposed parameter direction estimator $D(\hat{\beta})$ in section 2.

In the first simulation study, we assume that the response variable Y is discrete, and X is $p = 3$ dimensional covariates, i.e., $X \sim N(0, \Sigma = (\sigma_{ij}))$, $\sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, \cdots, p$. We mainly consider that the following data generation scenarios in models (4).

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

When the distributions of error term ϵ is unknown and a conditional median restriction, Manski(1975) proposed the maximum score (MS) estimator for the parameter β. Shen et al. (2022) proposed the linearized maximum rank correlation (LMRC) estimator. Here, we compared the proposed estimator with MS estimators, LMRC estimator and the standard method (probit regression) estimator. The angle between the true direction of the parameter β and the estimation of the direction $D(\hat{\beta})$ is $\mathbf{cos} = \beta * D(\hat{\beta})/(\|\beta\| * \|D(\hat{\beta})\|)$. We obtain that cos and standard error (SE) of estimators of β for different distributions of ϵ with sample sizes equal to 100, 300 and 500 for Case I-III. The results based on 100 repetitions are reported in Table (1)-(3). From Table (1)-(3), we find that as the sample size increases, all SE for cos decrease and the values of **cos** is closer to 1, which is consistent

Table 1: The SE and cos value of proposed estimators with β with sample size n=100, 300, 500 based on 100 repetitions for Case I.

n	ρ	New		MS		LMRC		Standard	
		cos	SE	cos	SE	cos	SE	cos	SE
100	-0.9	0.9601	0.0715	0.8919	0.1726	0.9605	0.0696	0.9637	0.0619
	-0.6	0.9896	0.0132	0.9672	0.0411	0.9893	0.0135	0.9904	0.0127
	-0.3	0.9890	0.0121	0.9704	0.0321	0.9887	0.0125	0.9887	0.0109
	0	0.9883	0.0114	0.9627	0.0344	0.9880	0.0118	0.9888	0.0129
	0.3	0.9805	0.0199	0.9623	0.0399	0.9802	0.0204	0.9834	0.1594
	0.6	0.9647	0.0322	0.9356	0.0679	0.9634	0.0331	0.9634	0.0257
	0.9	0.8543	0.1028	0.7900	0.1461	0.8623	0.0972	0.8924	0.0879
300	-0.9	0.9902	0.0129	0.9582	0.0702	0.9901	0.0128	0.9901	0.0148
	-0.6	0.9969	0.0034	0.9866	0.0139	0.9970	0.0035	0.9971	0.0034
	-0.3	0.9973	0.0027	0.9897	0.0103	0.9973	0.0027	0.9975	0.0026
	0	0.9965	0.0031	0.9866	0.0122	0.9964	0.0030	0.9971	0.0023
	0.3	0.9923	0.0092	0.9732	0.0250	0.9921	0.0095	0.9941	0.0069
	0.6	0.9850	0.0178	0.9591	0.0392	0.9847	0.0180	0.9891	0.0124
	0.9	0.9522	0.0486	0.8815	0.0918	0.9522	0.0488	0.9616	0.0390
500	-0.9	0.9946	0.0072	0.9620	0.0738	0.9945	0.0073	0.9948	0.0075
	-0.6	0.9963	0.0041	0.9865	0.0129	0.9962	0.0042	0.9969	0.0034
	-0.3	0.9920	0.0094	0.9728	0.0303	0.9919	0.0093	0.9936	0.0074
	0	0.9978	0.0181	0.9894	0.0109	0.9978	0.0020	0.9982	0.0020
	0.3	0.9986	0.0015	0.9909	0.0081	0.9986	0.0015	0.9987	0.0014
	0.6	0.9850	0.0014	0.9923	0.0090	0.9985	0.0013	0.9986	0.0014
	0.9	0.9634	0.0331	0.9093	0.0789	0.9640	0.0336	0.9744	0.0389

with Theorem 1. Besides, we observe that the proposed estimator is significantly better than MS estimator in terms of the values of cos and SE. The proposed estimator and LMRC estimator are comparable, the cos value and SE obtained by the proposed method and LMRC method are almost the same for the sample size is large enough. When the distribution of error term is long-tailed distribution (i.e., Student t) and distribution with no moments (i.e., Cauchy), the proposed estimator and LMRC estimator perform better than standard estimator.

A main drawback of the maximum score estimator is its computational difficulty, because the objective function of optimization is non-convex and non-smooth, which makes it a difficult task to find the global optimal solution. In addition, the calculation difficulty is more serious with the dimension of X being larger (Khan *et al.*, 2021). To consider the influence of the increase in the dimension of covariates on the estimation results, we consider the following data generation scenarios, i.e., $X \sim N(0, \Sigma = (\sigma_{ij}))$, $\sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, \cdots, p$. $Y = 1\{X^T\beta + \epsilon > 0\}$, where $p = 10, 15$, $\epsilon \sim N(0, 1)$ and $\beta = 1_p$. Compared the proposed method with the LMRC method and standard method with the dimension of X is relatively large. Simulation results are listed in Table (4) with n=500 and 100 repetitions for two different p values. Results show that our proposed method is still feasible when the dimension of X is relatively high, e.g., $p = 10, 15$. The proposed estimator is comparable with

Table 2: The SE and cos value of proposed estimators with β with sample size n=100, 300, 500 based on 100 repetitions for Case II.

n	ρ	New		MS		LMRC		Standard	
		cos	SE	cos	SE	cos	SE	cos	SE
100	-0.9	0.8926	0.1967	0.8467	0.2595	0.8839	0.2108	0.8921	0.1992
	-0.6	0.9748	0.0333	0.9482	0.0842	0.9734	0.0366	0.9743	0.0352
	-0.3	0.9792	0.0216	0.9662	0.0299	0.9790	0.0220	0.9791	0.0233
	0	0.9732	0.0292	0.9498	0.0657	0.9731	0.0293	0.9734	0.0285
	0.3	0.9572	0.0392	0.9267	0.0615	0.9579	0.0383	0.9563	0.0393
	0.6	0.9228	0.0671	0.8749	0.1071	0.9243	0.0649	0.9225	0.0642
	0.9	0.7582	0.1570	0.7713	0.1721	0.7555	0.1524	0.7569	0.1552
300	-0.9	0.9815	0.0225	0.9021	0.1990	0.9810	0.0232	0.9810	0.0239
	-0.6	0.9941	0.0069	0.9792	0.0284	0.9941	0.0072	0.9941	0.0071
	-0.3	0.9938	0.0060	0.9835	0.0166	0.9938	0.0060	0.9937	0.0062
	0	0.9932	0.0076	0.9750	0.0231	0.9931	0.0075	0.9927	0.0080
	0.3	0.9878	0.0124	0.9656	0.0359	0.9881	0.0118	0.9872	0.0130
	0.6	0.9755	0.0271	0.9464	0.0497	0.9759	0.0258	0.9739	0.0288
	0.9	0.8875	0.0968	0.8420	0.1126	0.8883	0.0970	0.8847	0.0980
500	-0.9	0.9848	0.0176	0.9650	0.0649	0.9847	0.0177	0.9842	0.0186
	-0.6	0.9957	0.0048	0.9834	0.0193	0.9957	0.0047	0.9955	0.0051
	-0.3	0.9964	0.0037	0.9909	0.0079	0.9963	0.0037	0.9962	0.0041
	0	0.9965	0.0039	0.9845	0.0131	0.9965	0.0039	0.9969	0.0032
	0.3	0.9911	0.0087	0.9752	0.0254	0.9911	0.0087	0.9906	0.0095
	0.6	0.9817	0.0178	0.9601	0.0373	0.9817	0.0177	0.9807	0.0191
	0.9	0.9229	0.0734	0.8851	0.0988	0.9230	0.0734	0.9205	0.0764

Table 3: The SE and cos value of proposed estimators with β with sample size n=100, 300, 500 based on 100 repetitions for Case III.

n	ρ	New		MS		LMRC		Standard	
		cos	SE	cos	SE	cos	SE	cos	SE
100	-0.9	0.9773	0.0289	0.8837	0.2288	0.9786	0.0275	0.9793	0.0287
	-0.6	0.9934	0.0071	0.9819	0.0165	0.9935	0.0072	0.9944	0.0066
	-0.3	0.9929	0.0072	0.9850	0.0156	0.9928	0.0075	0.9945	0.0061
	0	0.9906	0.0089	0.9785	0.0222	0.9905	0.0093	0.9930	0.0097
	0.3	0.9840	0.0169	0.9650	0.0355	0.9837	0.0172	0.9880	0.0121
	0.6	0.9649	0.0338	0.9388	0.0572	0.9649	0.0326	0.9761	0.0217
	0.9	0.8800	0.1035	0.8320	0.1176	0.8828	0.1017	0.9103	0.0806
300	-0.9	0.9938	0.0060	0.9747	0.0342	0.9936	0.0064	0.9943	0.0060
	-0.6	0.9981	0.0018	0.9919	0.0083	0.9981	0.0018	0.9982	0.0019
	-0.3	0.9982	0.0019	0.9928	0.0094	0.9982	0.0019	0.9984	0.0016
	0	0.9975	0.0023	0.9908	0.0093	0.9974	0.0024	0.9983	0.0017
	0.3	0.9946	0.0050	0.9854	0.0144	0.9945	0.0050	0.9966	0.0027
	0.6	0.9889	0.0123	0.9700	0.0260	0.9885	0.0123	0.9931	0.0076
	0.9	0.9496	0.0495	0.9119	0.0695	0.9495	0.0502	0.9670	0.0314
500	-0.9	0.9963	0.0048	0.9797	0.0208	0.9962	0.0048	0.9965	0.0042
	-0.6	0.9991	0.0011	0.9945	0.0058	0.9991	0.0011	0.9990	0.0011
	-0.3	0.9989	0.0009	0.9945	0.0056	0.9989	0.0009	0.9990	0.0009
	0	0.9985	0.0016	0.9929	0.0064	0.9983	0.0016	0.9988	0.0013
	0.3	0.9970	0.0029	0.9894	0.0130	0.9970	0.0029	0.9982	0.0018
	0.6	0.9931	0.0072	0.9800	0.0197	0.9932	0.0070	0.9963	0.0040
	0.9	0.9653	0.0337	0.9325	0.0554	0.9662	0.0339	0.9829	0.0181

Table 4: The SE and cos value of proposed estimators with β with sample size n=500 based on 100 repetitions for different dimensions of X .

Dimension	ρ	New		LMRC		Standard	
		cos	SE	cos	SE	cos	SE
10	-0.9	0.9708	0.0202	0.9709	0.0195	0.9703	0.0205
	-0.6	0.9862	0.0070	0.9862	0.0071	0.9858	0.0071
	-0.3	0.9889	0.0058	0.9889	0.0059	0.9886	0.0057
	0	0.9865	0.0064	0.9864	0.0065	0.9857	0.0069
	0.3	0.9757	0.0122	0.9757	0.0120	0.9759	0.0127
	0.6	0.9396	0.0283	0.9394	0.0290	0.9417	0.0290
	0.9	0.6999	0.0964	0.6983	0.0967	0.7048	0.1026
15	-0.9	0.9569	0.0205	0.9568	0.0201	0.9560	0.0211
	-0.6	0.9828	0.0080	0.9829	0.0077	0.9818	0.0085
	-0.3	0.9852	0.0066	0.9852	0.0066	0.9846	0.0072
	0	0.9804	0.0085	0.9802	0.0087	0.9799	0.0085
	0.3	0.9661	0.0152	0.9661	0.0152	0.9648	0.0158
	0.6	0.9133	0.0341	0.9133	0.0333	0.9113	0.0386
	0.9	0.5758	0.0893	0.5745	0.0900	0.5851	0.0941

LMRC estimator, and performs better than standard estimator.

Finally, we assume that the response variable Y is continuous and $X \sim N(0, \Sigma = (\sigma_{ij}))$, $\sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, \cdots, p = 3$. and following data generation scenarios are considered.

- Case 1. $Y = \mathbf{X}'\beta + \epsilon$, where $\epsilon \sim t(1)$ and $\beta = (1, 1, 1)$.

- Case 2. $Y = \Phi(\mathbf{X}'\beta) + \epsilon$, where $\epsilon \sim N(0, 1)$ and $\beta = (1, 1, 1)$.

- Case 3. $Y = \log(1 + \exp(\mathbf{X}'\beta)) + \epsilon$, where $\epsilon \sim N(0, 1)$ and $\beta = (1, 1, 1)$.

- Case 4. $Y = \frac{1}{1+\exp(-50*\mathbf{X}'\beta)} + \epsilon$, where $\epsilon \sim N(0, 1)$ and $\beta = (1, 1, 1)$.

We compare the proposed estimator with LMRC estimator. The SE and cos value of proposed estimators with β for different values of ρ with different distributions of ϵ and simple size n=100, 300, 500 based on 100 repetitions are reported in Table (5). From Table (5), we observe that as the sample size increases, all SE for cos decrease and the values of **cos** is closer to 1. Although the proposed estimator is slightly inferior to LMRC estimator When the model is lineal model, the proposed estimator performs better than LMRC estimator when the model is nonlinear.

4 Real data

In this section, we apply our proposed methodology to a real dataset that studies the influence of series exporting determined variables on the export-market participation of specialized and transport

Table 5: The SE and cos value of proposed estimators with β with sample size n=100, 300, 500 based on 100 repetitions for Case 1-4.

n	ρ		Case 1		Case 2		Case 3		Case 4	
			cos	SE	cos	SE	cos	SE	cos	SE
100	-0.3	New	0.9958	0.0048	0.9307	0.0832	0.9840	0.0167	0.9544	0.0528
		LMRC	0.99575	0.0052	0.9249	0.0882	0.9826	0.0195	0.9515	0.0569
	0	New	0.9949	0.0050	0.8998	0.0970	0.9795	0.0184	0.9220	0.0762
		LMRC	0.9952	0.0046	0.8937	0.1039	0.9792	0.0196	0.9189	0.0806
	0.3	New	0.9940	0.0058	0.8543	0.1377	0.9716	0.0270	0.8710	0.1205
		LMRC	0.9944	0.0051	0.8439	0.1491	0.9724	0.0298	0.8648	0.1268
300	-0.3	New	0.9989	0.0010	0.9831	0.0166	0.9957	0.0042	0.9880	0.0120
		LMRC	0.9990	0.0010	0.9829	0.0174	0.9959	0.0040	0.9883	0.0123
	0	New	0.9988	0.0012	0.9744	0.0257	0.9944	0.0065	0.9796	0.0215
		LMRC	0.9987	0.0014	0.9729	0.0267	0.9943	0.0065	0.9791	0.0221
	0.3	New	0.9982	0.0020	0.9559	0.0429	0.9910	0.0082	0.9630	0.0372
		LMRC	0.9983	0.0016	0.9556	0.0450	0.9914	0.0086	0.9636	0.0374
500	-0.3	New	0.9993	0.0007	0.9883	0.0139	0.9970	0.0029	0.9918	0.0090
		LMRC	0.9993	0.0007	0.9881	0.0134	0.9970	0.0031	0.9920	0.0087
	0	New	0.9914	0.0010	0.9815	0.0242	0.9962	0.0044	0.9858	0.0179
		LMRC	0.9918	0.0009	0.9808	0.0243	0.9963	0.0042	0.9855	0.0177
	0.3	New	0.9987	0.0019	0.9670	0.0426	0.9950	0.0054	0.9721	0.0366
		LMRC	0.9987	0.0016	0.9672	0.0430	0.9950	0.0070	0.9729	0.0360

facility manufactures in the province of Guangdong, China in 2006 (Baltagi *et al.* 2022). The data is available on the National Bureau of Statistics of China (NBS).

In the subsequent analyses, the variables we mainly consider include **expd-ford** (=1 if the company is an exporter; 0 otherwise), **lemp** (log firm sizes), **lprod** (log output per worker), **lcapint** (capital divided by total sales), **intastr** (intangible assets over total assets), **cmp** (log sales over operating profits), $\mathbf{cmp^2}$ (Square of **cmp**), **ltastx** (Fixed export costs), and **sez** (=1 if the firm is located in the Special Economic Zone; =0 otherwise). Finally, we obtained that a total of 1614 companies with total annual sales of at least 5 mn. RMB (about 700,000 US dollars) in 2006. As mentioned in Khan *et al.* (2021), the maximum score method was extremely difficult to calculate for the dimension of the covariate being large. There are a total of 8 covariates in this analysis, therefore the maximum score estimator is ignored here.

To illustrate our methodologies, we let the response variable (i.e., Y) be **expd-ford** and X includes the remaining variables. The results estimated by the proposed method, LMRC method and standard method are presented in Table (6). From Table (6), we can explain some economic hypotheses. Obviously, a simple test $T = \hat{\beta}_i/\hat{\sigma}_i, i, \cdots, 8$ suggests that **lemp**, **lcapint**, **ltastx**, **sez** have a significant impact on response variable at a significance level of 0.05 for the proposed method. Large size of firms where fixed costs are more important (through a higher capital intensity) tend to be more

Table 6: The influence of series exporting determined variables on the export-market participation of specialized and transport facility manufactures based on various methods.

Parameter	Proposed method	Probit	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)

likely to export. Higher the Fixed export costs is, lower the profitability of exporting will be, which explain the decrease influence of export. Compared with the LMRC, the proposed method finds that **sez** is a new influencing factors on response variables. Since the distance between firms and Special Economic Zone should also affect the exporting, the binary variable do have a significant influence, hence our proposed method is more reasonable.

5 Conclusion and discussion

In this paper, we consider that the estimation of unknown parameter direction in semi-parametric models for response variables, which can be continuous or discrete. The least square method are proposed to estimate the direction of unknown parameters in semi-parametric models. The proposed estimator is computationally simple and has a closed-form expression. It is proved that the proposed estimator is consistent and asymptotically normal. The proposed estimation is significantly superior to the maximum score estimation for binary response variables, is comparable with the linearized maximum rank correlation and the Probit estimation. When the distribution of error term is long-tailed distributions (i.e., Student t) and distributions with no existing moments (i.e., Cauchy), the proposed estimator and LMRC estimator perform better than the Probit estimator. The proposed estimation is superior to the linearized maximum rank correlation estimation for continuous response variable with nonlinear models. Furthermore if one ise interested in the estimation of the link function $g(\cdot)$, it can be directly estimated by non-parametric methods with $E[Y \mid X^T\hat{\beta}]$.

In this paper, we mainly consider that the link function $g(\cdot)$ is monotonically increasing. The estimator of unknown parameter direction can be obtained by a similar method when the link function is monotonically decreasing, the direction estimation of unknown parameters is opposite to the true direction of parameters. In practice, the observable covariates are high-dimensional for various reasons. we extend the proposed method to handle parameter direction estimation for model (4) with high-dimensional covariates. We can estimate the direction of parameters β via minimize the following objective function

$$\hat{\beta}_n = \arg\min_{\|\beta\|=1} \left\{ \frac{1}{n} \sum_{i=1}^{n} (Y_i - c - \mathbf{X}_i'\beta)^2 + P(\beta) \right\}, \tag{7}$$

where $P(\beta)$ is the penalty function, e.g., lasso, SCAD. This research problem will be considered in the future.

Appendix

Lemma A. When $\mathbf{X}$ is distributed by the elliptical distributions with mean $\mathbf{0}$ and covariance $\mathrm{Var}(\mathbf{X}) = \Sigma$(positive definite) for Models (4), then

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

where $\lambda > 0$.

Proof:

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

(I) in case $\beta = 0$, it is obviously proved. (II) in case $\beta \neq 0$ and $\Sigma = \sigma^2 I$. Let A be the orthogonal matrix with the first row $\beta/\|\beta\|$ and

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

Then from definition of $\mathbf{W}$, one has $W_1 = \beta' X/\|\beta\|$ and

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

Since $\mathbf{X}$ is distributed by the elliptical distributions with mean $\mathbf{0}$ and variance $\mathrm{Var}(\mathbf{X}) = \sigma^2 I$, $\mathbf{W} = A\mathbf{X}$ is distributed by the elliptical distributions with mean $\mathbf{0}$ and variance $\sigma^2 AA' = \sigma^2 I$ and then

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

by Theorem 6 given by Frahm (2004). So

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

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

by $g(\cdot)$ is strictly increasing and

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

where

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

(III) in general $\beta \neq 0$ and $\mathrm{Var}(X) = \Sigma$. Let $\mathbf{X}^* = \Sigma^{-1/2}\mathbf{X}$ and $\beta^* = \Sigma^{1/2}\beta$. Then $\mathbf{X}^*$ is distributed by the elliptical distributions with mean $\mathbf{0}$ and variance I. By Models (4), one has

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

and so

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

by the case (II) and $\lambda > 0$.

Proof of Theorem 1: By the Law of Large Number, we have

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

According to the Lemma A and

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

Hence, we can obtain that

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

that is

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

Therefor, the proposed estimator $D(\hat{\beta})$ is consistent.

Proof of Theorem 2: Let

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

By Lemma A and the Central Limit Theorem, one know $E\mathbf{U}_n = \lambda\Sigma\beta$ and

$$\sqrt{n}\begin{pmatrix}\mathbf{U}_n - E[U_n] \\ \mathbf{V}_n - \Sigma\beta\end{pmatrix} = \begin{pmatrix}\mathbf{U}_n - \lambda\Sigma\beta \\ \mathbf{V}_n - \Sigma\beta\end{pmatrix} \xrightarrow{d} \begin{pmatrix}\mathbf{U} \\ \mathbf{V}\end{pmatrix} \sim N\left(\begin{pmatrix}\mathbf{0} \\ \mathbf{0}\end{pmatrix}, \Omega\right), \tag{8}$$

where

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

According to the form of the proposed estimator, we have

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

In order to obtain the asymptotic distribution of $D^*(\hat{\beta}) - \beta$, we first require to prove the asymptotic distribution of $\sqrt{n}(\mathbf{U}_n - E\mathbf{U}_n)$, $\sqrt{n}\left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n - \lambda\beta\right)$ and $\sqrt{n}\left(\frac{\lambda\beta}{||\hat{\Sigma}_X^{-1}\mathbf{U}_n||} - \frac{\beta}{||\beta||}\right)$. From (8), we have $\sqrt{n}(\mathbf{U}_n - E\mathbf{U}_n) \xrightarrow{d} \mathbf{U} \sim N(0, Var(Y\mathbf{X})$. Next, we will prove the asymptotic properties of $\sqrt{n}\left(\hat{\Sigma}_X^{-1}E\mathbf{U}_n - \lambda\beta\right)$ and $\sqrt{n}\left(\frac{\lambda\beta}{||\hat{\Sigma}_X^{-1}\mathbf{U}_n||} - \frac{\beta}{||\beta||}\right)$ respectively.

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

Let

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

and expand $\mathbf{S}_n(\mathbf{U}_n, \Sigma_X^{-1})$ at $(\lambda\Sigma\beta, \Sigma^{-1})$, we can obtain that

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

Therefore,

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

where

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

Reference

Abrevaya, J., Huang, J. (2005). On the bootstrap of the maximum score estimator. *Econometrica*, **73**(4), 1175-1204.

Balabdaoui, F., Durot, C., Jankowski, H. (2019). Least squares estimation in the monotone single index model. *Bernoulli*, **25**(4B), 3276-3310.

Baltagi, B. H., Egger, P. H., Kesina, M. (2022). Bayesian estimation of multivariate panel probits with higher-order network interdependence and an application to firms' global market participation in Guangdong. *Journal of Applied Econometrics*, **37**(7), 1356-1378.

Cui, X., Härdle, W. K., Zhu, L. (2011). The EFM approach for single-index models. *The Annals of Statistics*, **39**(3), 1658-1688.

Frahm, G. (2004). Generalized elliptical distributions: theory and applications (Doctoral dissertation, Universität zu Köln).

Gao, W. Y., Xu, S. (2022). Two-Stage Maximum Score Estimator. arXiv preprint arXiv:2009.02854.

Han, A. K. (1987). Non-parametric analysis of a generalized regression model: the maximum rank correlation estimator. *Journal of Econometrics*, 35(2-3), 303-316.

Härdle, W., Hall, P., Ichimura, H. (1993). Optimal smoothing in single-index models. *The annals of Statistics*, **21**(1), 157-178.

Härdle, W., Stoker, T. M. (1989). Investigating smooth multiple regression by the method of average derivatives. *Journal of the American statistical Association*, **84**(408), 986-995.

Horowitz, J. L. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica*, 505-531.

Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Annals of Statistics*, 595-623.

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of econometrics*, **58**(1-2), 71-120.

Khan, S., Lan, X., Tamer, E. (2021). Estimating High Dimensional Monotone Index Models by Iterative Convex Optimization1. arXiv preprint arXiv:2110.04388.

Kuchibhotla, A. K., Patra, R. K. (2020). Efficient estimation in single index models through smoothing splines. *Bernoulli*, **26**(2), 1587-1618.

Kuchibhotla, A. K., Patra, R. K., Sen, B. (2021). Semiparametric Efficiency in Convexity Constrained Single-Index Model. *Journal of the American Statistical Association*, 1-15.

Li, Q. and Racine, J.S. (2007). Nonparametric Econometrics: Theory and Practice. Princeton, NJ: Princeton Univ. Press. MR2283034.

Liu, J., Zhang, R., Zhao, W., Lv, Y. (2013). A robust and efficient estimation method for single index models. *Journal of Multivariate Analysis*, **122**, 226-238.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of econometrics*, **3**(3), 205-228.

Manski, C. F. (1988). Identification of binary response models. *Journal of the American statistical Association*, **83**(403), 729-738.

Naik, P., Tsai, C. L. (2000). Partial least squares estimator for single-index models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**(4), 763-771.

Park, H., Petkova, E., Tarpey, T., Ogden, R. T. (2020). A single-index model with multiple-links. *Journal of statistical planning and inference*, **205**, 115-128.

Patra, R. K., Seijo, E., Sen, B. (2018). A consistent bootstrap procedure for the maximum score estimator. *Journal of Econometrics*, **205**(2), 488-507.

Shen, G., Chen, K., Huang, J., Lin, Y. (2022). Linearized maximum rank correlation estimation. Biometrika.

Song, X., Ma, S., Huang, J., Zhou, X. H. (2007). A semiparametric approach for the nonparametric transformation survival model with multiple covariates. *Biostatistics*, 8(2), 197-211.

Wu, T. Z., Yu, K., Yu, Y. (2010). Single-index quantile regression. *Journal of Multivariate Analysis*, **101**(7), 1607-1621.

Yang, J., Tian, G., Lu, F., Lu, X. (2020). Single-index modal regression via outer product gradients. *Computational Statistics and Data Analysis*, **144**, 106867.

Yu, Y., Ruppert, D. (2002). Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association*, **97**(460), 1042-1054.

Zhou, J. and He, X. (2008). Dimension reduction based on constrained canonical correlation and variable filtering. *Annals of Statistics*, **36**, 1649C1668.

Zhou, L., Lin, H., Chen, K., Liang, H. (2019). Efficient estimation and computation of parameters and nonparametric functions in generalized semi/non-parametric regression models. *Journal of Econometrics*, **213**(2), 593-607.