齐鲁工业大学学报 第40卷 第4期 2026年8月出版
JOURNAL OF QILU UNIVERSITY OF TECHNOLOGY Vol.40 No.4 Aug.2026
DOI:10.16442/j.cnki.qlgydxxb.2026.04.008 【数理与统计科学】

半参数模型稳健性的改进
——基于分位数双重稳健估计

曾 赟
重庆工商大学 数学与统计学院,重庆市 400067

摘要:在半参数模型中,当数据存在非线性趋势或者模型设定错误时,其估计的稳健性结果往往不理想。传统的双重稳健估计极度依赖均值导向的假设和倾向得分模型,模型估计的稳健性得不到保障。基于线性分位数回归估计,联合优化非参数部分 $g(\boldsymbol{X})$ 与参数部分 $\boldsymbol{\beta}(\tau)$,改进半参数分位数双重稳健估计模型,得出有效估计量。通过数值模拟假设 $g(\boldsymbol{X})$、$\boldsymbol{\beta}(\tau)$ 有一个及其以上被错误设定时,对比本文模型和线性分位数回归模型,模型的偏差、均方根误差均显著性低于线性分位数回归模型,且偏差预测降低了45%左右,验证出所提模型的稳健性。在进行双重稳健估计时,采用联合优化非参数部分和参数部分,比现有的均值模型稳健性更高,更适合用于模型错误设定或存在非线性趋势的模型中。

关键词:半参数模型;非线性趋势;模型错误设定;分位数双重稳健估计;稳健性
中图分类号:O212 **文献标志码**:A **文章编号**:2097-2792(2026)04-71-10

Improving robustness in semi-parametric models with quantile double robust estimation

ZENG Yun
School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China

Abstract: In semi-parametric models, estimation robustness is often compromised by nonlinear trends or model misspecification. Traditional doubly robust estimators, which rely heavily on mean-oriented assumptions and propensity score models, frequently fail to deliver reliable inference. This paper develops an enhanced semi-parametric quantile doubly robust estimation model that integrates local linear quantile regression to simultaneously optimize both the nonparametric component $g(\boldsymbol{X})$ and the parametric component $\boldsymbol{\beta}(\tau)$, thereby obtaining efficient estimators. Numerical simulations—conducted under scenarios where at least one of $g(\boldsymbol{X})$ or $\boldsymbol{\beta}(\tau)$ is misspecified—show that the proposed model achieves significantly lower bias and root mean square error compared to conventional linear quantile regression. Specifically, bias is reduced by approximately 45%, highlighting the model's superior robustness. By adopting a joint optimization strategy for both components, the proposed framework exhibits stronger robustness than existing mean-based models, making it particularly suitable for applications involving model misspecification or nonlinear trends.

Key words: semi-parametric model; nonlinear trend; model misspecification; quantile double robust estimation; robustness

收稿日期:2025-10-30 **网络出版时间**:2026-08-31
基金项目:重庆市自然科学基金项目(CSTB2022NSCQ—MSX1370)
作者简介:曾赟,硕士生;研究方向:统计推断;zy011405@163.com

在统计学和计量经济学中半参数模型被广泛应用,它有效克服参数模型无法正确展示研究变量间潜在关系,以及非参数模型估计中可能出现的“维数祸根”问题。

自从 Koenker 等[1]1978 年首次开创性系统提出线性分位数回归思想其模型更是成为研究稳健性重要工具开始,半参数模型有了很多优秀的成果。Boor[2]详细介绍了样条插值的理论知识;Tibshirani 等[3]首次系统性地将非参数平滑方法引入回归框架;Robins 等[4]提出双重稳健估计方法在因果推断和缺失数据处理中展现出独特优势;Laan 等[5]为处理缺失数据和因果推断的半参数模型,建立了一个统一、高效的估计理论框架;Van Der Laan[6]是首批将两种思想整合的,挖掘出缺失数据下分位数估计的稳健性。

真正引入非参成分,则是 Portnoy[7]、Koenker 等[8]对线性分位数回归模型进行扩展,并讨论其渐近性。后续研究中,Stone[9]系统地论证了在各种平滑性假设下,非参数回归估计量能够达到的最优全局收敛速率;Bang 等[10]将双重稳健估计推广到缺失数据问题,为半参数分位数双重稳健估计模型奠定基础;Koenker[11]样条平滑方法从经典的“条件均值回归”领域推广到了“条件分位数回归”领域,为全面刻画响应变量整个条件分布提供了强大的非参数工具;Zou 等[12]提出复合分位数回归方法,显著提升估计效率;Tan 等[13]首次联合分位数回归到双重稳健估计,提出基于逆概率加权的双重稳健估计方法;Zheng 等[14]针对高维协变量,提出半参数双重稳健分位数回归方法;Chernozhukov 等[15]和 Belloni 等[16]在模型中引入工具变量解决其内生性问题,进一步扩展其模型应用场景;Caracciolo 等[17]定义基于分位数的双重稳健估计,揭示处理效应在不同分位数上的异质性。Liu 等[18]倾向得分估计满足一定收敛速率时,部分模型错误设定情况下,其模型仍保持双重稳健性。近期,丁先文等[19]、田玲等[20]、许文超等[21]、张日权[22]、Chernozhukov 等[23]和 Blanchet 等[24]不断优化模型,为提升模型稳健性做出贡献。

尽管半参数双重稳健估计的理论框架逐步完善,并在诸多领域展现出良好的应用前景,但现有方法在面对复杂的真实数据时,仍存在理论局限与实践挑战。大量研究表明,该类方法对模型设定错误及数据结构变化极为敏感。Koenke 等[1]明确指出,当变量间的真实关系偏离线性假设时,传统模型的参数估计将丧失一致性;Bang[9]系统论证了双重稳健估计要求倾向得分模型或结果模型中至少一个必须设定正确,这一苛刻假设在实际研究中往往难以严格满足;Zheng 等[14]也指出传统半参数方法在维度增加时稳健性急剧下降,他提出的自适应方法有所缓解但未完全解决该问题;Caracciolo 等[17]基于意大利劳动力市场的实证研究也表明,不同模型设定导致处理效应估计差异显著。这些局限性制约了半参数双重稳健估计在复杂数据环境下的稳健性与推广价值。

为应对上述挑战,学者们开始探索在模型错误假设或存在复杂非线性趋势时更稳健的估计框架。基于此,本文提出一种改进的分位数双重稳健估计方法,通过半参数建模思想,将非参数部分用分位数回归进行估计,并与参数分量建立耦合关系,最终通过迭代算法实现参数估计。该方法旨在提升半参数模型在复杂数据场景下的适应能力,同时深入挖掘分位数双重稳健估计在半参数建模中的潜力。

1 预备知识

为保证对本文提出方法有更完整的理解,先介绍一些相关的理论。首先,简要介绍半参数模型与分位数回归的基本设定,阐明二者结合的理论动机。在此基础上,将引出本文的核心框架——分位数双重稳健估计模型(Quantile Doubly Robust Estimation,QDR),该模型通过融合半参数结构的灵活性与分位数回归的稳健性,旨在提升存在模型误设或复杂数据生成机制下的估计精度与可靠性,使其更能广泛适用于经济学、医学等研究领域。

1.1 半参数模型

在提出半参数模型下的分位数双重稳健估计模型前,先回顾传统半参数模型的基本框架。考虑如下形式的半参数模型:

$$Y=g(\boldsymbol{X})+\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}+\varepsilon, \tag{1}$$

其中,Y 为因变量,$\boldsymbol{X}$ 为 m 维自变量,$\boldsymbol{\beta}$ 为 p 维参数向量,$\boldsymbol{Z}$ 是 p 维自变量,ε 为误差项,$g(\cdot)$是非参数部分,$\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}$ 是参数部分。$g(\cdot)$是待估函数,$\boldsymbol{\beta}$ 是待估参数向量。

该模型较于传统线性回归模型增加了非参数部分,一定程度上缓解“维数灾难”。然而,该框架下的传

统估计极度依赖以均值为导向的模型假设,这意味着当模型存在设定错误或数据呈现异方差性时,这些估计量稳健性不强,导致参数估计出现偏误或方差估计不准确。

1.2　条件分位数模型

为了解决传统半参数模型由于均值估计而引发的问题,引入如下的分位数回归模型(Quantile Regression,QR)[22]:

$$Q_\tau(Y|\boldsymbol{Z})=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau),(0<\tau<1), \tag{2}$$

式(2)为线性分位数回归模型,表示在 $\boldsymbol{Z}$ 条件下 $\boldsymbol{Y}$ 的 τ 个条件分位数。其中,τ 表示给定分位数,$\boldsymbol{\beta}(\tau)$ 是分位数 τ 下的回归系数向量。分位数回归模型通过估计条件分位数函数,能够更全面地描述因变量 Y 的条件分布特征。然而,模型仍局限于参数形式设定,难以有效捕捉变量间复杂的非线性关系。

为克服以上两个模型的局限性,本文创新性地将半参数建模思想引入分位数回归框架,扩展半参数分位数回归模型。具体来说,就是对于给定分位数 τ,传统线性分位数回归模型扩展为

$$Q_\tau(Y|\boldsymbol{X},\boldsymbol{Z})=g(\boldsymbol{X})+\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)+Q(\varepsilon|\boldsymbol{X},\boldsymbol{Z}),0<\tau<1, \tag{3}$$

其中,$Q(\varepsilon|\boldsymbol{X},\boldsymbol{Z})$ 表示在给定 $\boldsymbol{X}$、$\boldsymbol{Z}$ 条件下,ε 的第 τ 个条件分位数。

沿用 Koenker 等[1]对于分位数回归的基本假设:$Q(\varepsilon|\boldsymbol{X},\boldsymbol{Z})=0$,即假设误差项的 τ 条件分位数为 0。基于 Koenker[1]、张日权[22]等的理论,当式(3)中满足条件 $g(\boldsymbol{X})=0$ 时,则式(3)退化为式(2)。由此扩展方式,可以将式(1)所示模型函数扩展到分位数回归模型,即有

$$Q_\tau(Y|\boldsymbol{X},\boldsymbol{Z})=g(\boldsymbol{X})+\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}(\tau)\ 。 \tag{4}$$

在此扩展方法下,将传统的条件均值模型扩展到半参数条件分位数模型,保留了半参数模型的灵活性,又结合了分位数模型的稳健性,能同时解决上文提及的半参数模型和分位数回归模型的问题,使得模型能适用于更为复杂数据,也更具稳健性。

2　估计方法

在构建半参数分位数双重稳健模型的理论框架基础上,如何从观测数据中实现对模型参数的有效估计,是理论走向实践的关键环节。在以往均值回归框架下,参数估计通常采用平方损失函数以估计条件期望;而在本文的模型中,为捕捉因变量条件分布的不同分位点特征,可以引入分位数损失函数作为度量工具,用以量化样本点偏离目标分位点的程度,并以此为优化目标进行模型拟合。因此,选择合适的分位数损失函数,是实现半参数分位数双重稳健估计的首要步骤。

在构建分位数损失函数前,为专注于参数部分的估计问题,暂时忽略非参数部分,则有 $Q_\tau(Y|\boldsymbol{Z})=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}$。假设残差 $u=Y-Q_\tau(Y|\boldsymbol{X},\boldsymbol{Z})$,根据分位数基本理论,最优的分位数估计可以对残差进行拆分,再用示性函数整合正、负残差,得到统一表达式

$$\rho_\tau(u)=u[\tau-I(u<0)]=\begin{cases}\tau|u|,\text{if } u\geqslant 0,\\(1-\tau)|u|,\text{if } u<0,\end{cases} \tag{5}$$

并将其称为分位数损失函数,其中 $I(u<0)$ 是示性函数。

建立起分位数损失函数后就可以进行目标函数的估计,采用先初步估计 $g(\boldsymbol{X})$ 和 $\boldsymbol{\beta}(\tau)$,再进行迭代优化得到目标函数估计值的方法。

为方便,先初步估计 $\boldsymbol{\beta}(\tau)$。现在的目标是找到目标值 $\boldsymbol{\beta}$ 使得样本分位数损失函数达到最小,而由于 $\rho_\tau(u)$ 是凸函数,故 $\sum\rho_\tau(u)$ 也为凸函数,就将求解最小值问题转化为求解线性规划问题。于是采用类似 Koenker 等[8]的求解线性规划问题的计算过程,直接拟合线性分位数回归计算得到

$$\hat{\boldsymbol{\beta}}^{(0)}(\tau)=\arg\min_{b}\sum_{i=1}^{n}\rho_\tau(Y_i-\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{b}), \tag{6}$$

其中,$\boldsymbol{b}$ 表示待优化的 $\boldsymbol{\beta}(\tau)$,Y_i 表示第 i 次因变量观测值。

当 $\hat{\boldsymbol{\beta}}^{(0)}(\tau)$ 近似 $\boldsymbol{\beta}(\tau)$ 时,式(6)两边同时减 $\boldsymbol{Z}^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(0)}$ 有

$$Y-\boldsymbol{Z}^{\mathrm{T}}\hat{\beta}^{(0)}=g(\boldsymbol{X})+\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{\beta}(\tau)-\hat{\boldsymbol{\beta}}^{(0)}(\tau))\ 。 \tag{7}$$

由于后续优化可以逐步修正误差项,故先忽略初始误差项$\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{\beta}(\tau)-\hat{\boldsymbol{\beta}}^{(0)}(\tau))$,得到

$$Y-\boldsymbol{Z}^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(0)}(\tau)\approx g(\boldsymbol{X})\ 。\tag{8}$$

根据式(8)可看出残差近似于$g(\boldsymbol{X})$,记$Y-\boldsymbol{Z}^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(0)}(\tau)$为$\hat{Y}$。现要估计$g(\boldsymbol{X})$,需最小化调整后的残差与$g(\boldsymbol{X})$的分位数损失函数之和,采用局部线性分位数回归(Local Linear Quantile Regression,LLQR)来计算。根据 Koenker 等[1]思想,对于初始的$\hat{\boldsymbol{\beta}}^{(0)}(\tau)$,对第$i$个自变量观测点$X_i$构建局部加权方程

$$\hat{g}^{(0)}(\boldsymbol{X})=\arg\min_{g}\sum_{i=1}^{n}\rho_{\tau}(\hat{Y}_i-g(\boldsymbol{X}_i))\ ,\tag{9}$$

其中,$\hat{Y}_i$为$\hat{Y}$第i次观测值的估计。

为了能有更快的收敛速度,沿用 Boor[2]中的B样条平滑估计逼近$g(\boldsymbol{X})$,即有

$$g(\boldsymbol{X})\approx\sum_{j=1}^{m}\theta_j B_j(X)\ ,\tag{10}$$

其中,$\{B_j(X)\}_{j=1}^{m}$是基函数个数为m的B样条基函数,$\boldsymbol{\theta}=(\theta_1,\cdots,\theta_m)^{\mathrm{T}}$为待估样条系数。

将式(10)代入式(9)可得到

$$\arg\min_{\theta}\sum_{i=1}^{n}\rho_{\tau}(\hat{Y}_i-\sum_{j=1}^{m}\theta_j B_j(X_i))\ 。\tag{11}$$

即上式的解为初步的$g(\boldsymbol{X})$估计值。

得到$g(\boldsymbol{X})$和$\boldsymbol{\beta}(\tau)$初始估计值后,对其进行迭代更新的联合优化。根据 Chernozhukov 等[15]的去偏思想,可进行 Neyman 正交化联合优化。首先固定$\boldsymbol{\beta}(\tau)$,更新$g(\boldsymbol{X})$得到

$$\hat{g}^{(k)}(\boldsymbol{X})=\arg\min_{\theta}\sum_{i=1}^{n}\rho_{\tau}(Y_i-\boldsymbol{Z}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(k-1)}(\tau)-\boldsymbol{B}(X_i)\boldsymbol{\theta})\ 。\tag{12}$$

接下来固定$g(\boldsymbol{X})$,更新$\boldsymbol{\beta}(\tau)$,则只需要求解

$$\hat{\boldsymbol{\beta}}^{(k)}(\tau)=\arg\min_{b}\sum_{i=1}^{n}\rho_{\tau}(\hat{Y}_i-\hat{g}^{(k)}(\boldsymbol{X}_i)-\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{b})\ 。\tag{13}$$

不断重复以上迭代优化步骤,直至满足如下收敛判断法则则停止迭代:

$$\frac{\|\hat{\boldsymbol{\beta}}^{(k)}(\tau)-\hat{\boldsymbol{\beta}}^{(k-1)}(\tau)\|}{1+\|\hat{\boldsymbol{\beta}}^{(k-1)}(\tau)\|}<\gamma,\tag{14}$$

$$\frac{\|\hat{g}^{(k)}(\boldsymbol{X})-\hat{g}^{(k-1)}(\boldsymbol{X})\|}{1+\|\hat{g}^{(k-1)}(\boldsymbol{X})\|}<\gamma,\tag{15}$$

其中,$\|\cdot\|$为L_2范数,γ为给定的收敛阈值。

利用此方法,先消除$g(\boldsymbol{X})$,再初始化$\boldsymbol{\beta}(\tau)$。在此分位数回归中,相较于传统的均值模型,引入条件分位数可更全面描述因变量Y的条件分布,也满足对于需要处理模型错误设定或存在非线性趋势下数据的方便性。

3 渐近性质

现研究估计量的渐近性,在 Koenker 等[1]和 Boor[2]启示下,沿用以下正则性假设条件:

(1)(X_i,Z_i)为 i.i.d.样本,且$\mathrm{E}[\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}}|\boldsymbol{X}]$的最小特征值满足$\lambda_{\min}\geqslant c>0$,$c$为固定常数;

(2)$\boldsymbol{Z}$满足稀疏性条件$\max_j\|\boldsymbol{Z}_j\|_{\psi_2}\ll C$,其中,$\|\cdot\|_{\psi 2}$表示亚高斯范数;

(3)$g_0(\cdot)$属于紧集$X\subseteq\mathbb{R}^d$上的 Hölder 类,即$\Sigma(s,L)$;

(4)$h\to 0$,$nh\to\infty$,$nh^5=O(1)$;

(5)采用B样条来逼近$g(\boldsymbol{X})$时,其中有$m\approx n^{1/(2s+d)}$;

(6)条件函数$f_{Y|X,Z}(y)$满足在$y=Q_{Y(\tau|X,Z)}$连续且有:$\inf_{x,z}f_{Y|X=x,Z=z}(g_0(\boldsymbol{X})+\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau))\geqslant f_{\min}>0$;

(7)在正则性惩罚参数$\lambda\approx\sqrt{\log p/n}$,稀疏性有$s=o(n/\log p)$。

定理 1　在假设(1)~(7)下,有

$$\sqrt{n}\boldsymbol{a}^{\mathrm{T}}(\hat{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau))\xrightarrow{d}\mathrm{N}(0,\frac{\tau(1-\tau)}{f_{\min}^2}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}),\tag{16}$$

其中,$\boldsymbol{a}\in\mathbb{R}^p$ 满足 $\|\boldsymbol{a}\|=1$ 且 $\|a\|_0\leqslant s$,$\|\cdot\|_0$ 表示向量 $\boldsymbol{a}$ 中非零元素个数,

$$\boldsymbol{\Sigma}=\mathrm{E}[\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}}|X],$$

$$f_{\min}=\inf_{x,z}f_{Y|X=x,Z=z}(g_0(x)+\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\beta}_0(\tau))。$$

定理 2　在假设(1)~(7)下,有:

$$\sqrt{nh}(\hat{g}(\boldsymbol{X})-g_0(\boldsymbol{X}))\xrightarrow{d}\boldsymbol{N}(0,\sigma_g^2(\boldsymbol{X})),\tag{17}$$

其中,

$$\sigma_g^2(\boldsymbol{X})=\boldsymbol{B}(X)\boldsymbol{\Gamma}^{-1}\boldsymbol{\Omega}\boldsymbol{\Gamma}^{-1}\boldsymbol{B}(X)^{\mathrm{T}},$$

$$\boldsymbol{\Gamma}=\mathrm{E}[f_{\varepsilon|X}(0)\boldsymbol{B}(X)^{\mathrm{T}}\boldsymbol{B}(X)],$$

$$\boldsymbol{\Omega}=\tau(1-\tau)\boldsymbol{\Sigma}。$$

4　数值模拟

本文采用蒙特卡洛随机模拟进行模拟实验。为进行充分的研究模拟,使用 *RStudio* 软件进行数值模拟。根据式(4)生成半参数分位数双重稳健估计模型假设 3 个场景:

(1)$g(\boldsymbol{X})=2\sin(\pi\boldsymbol{X}_1)+(\boldsymbol{X}_2-0.5)^2$,$\boldsymbol{\beta}(\tau)=(1,-1,0,\cdots,0)^{\mathrm{T}}$;

(2)$g(\boldsymbol{X})=3\boldsymbol{X}_1\mathrm{e}^{-\boldsymbol{X}_2^2}$,$\boldsymbol{\beta}(\tau)=(1.5,-1.2,0.8,0,\cdots,0)^{\mathrm{T}}$;

(3)$g(\boldsymbol{X})=2\sin(\pi\boldsymbol{X}_1)+2(\boldsymbol{X}_1-0.5)^2$,$\boldsymbol{\beta}(\tau)=(1.5,-1.2,0.8,-0.7,0.5,0,\cdots,0)^{\mathrm{T}}$。

以上 3 个场景中,向量 $\boldsymbol{X}$ 均是 [0,1] 三维非线性随机数,即 $\boldsymbol{X}_i\sim\mathrm{U}(0,1)$,$i=1,2,3$,而 $\boldsymbol{Z}$ 是 [0,1] 上的高维均匀随机数,即有 $Z_j\sim\mathrm{U}(0,1)$,$j=1,2,\cdots,20$,误差分布采用重尾分布的 t 分布,即 $\varepsilon\sim t(5)$。

现要验证模型错误设定情况下对模型的影响,即对 3 个场景进行假定:

(1)$g(\boldsymbol{X})$ 被忽略;

(2)$\boldsymbol{\beta}(\tau)$ 被错误设定;

(3)$g(\boldsymbol{X})$、$\boldsymbol{\beta}(\tau)$ 均被错误假定。

对比半参数分位数双重稳健估计模型、线性分位数回归模型(LinearQuantile Regression Model,LQRM),为评价模型的估计表现,对其的偏差 Bias、均方根误差 δ_{RMSE} 进行分析,其中,均方根误差 δ_{RMSE} 公式为

$$\delta_{\mathrm{RMSE}}=\sqrt{\sum_{i=1}^{n}d_i^2/(n-1)},\tag{18}$$

其中,d_i 为样本估计值与真实值的偏差,n 为模拟次数,取 $n=1\,000$。显然,模型的偏差和平方根误差数值越小,估计效果越好。

现取测试集、训练集样本量都为 $n=1\,000$ 情况下对模型进行 1 000 次并行优化模拟,其中,分位数 $\tau=0.5$,最大迭代次数为 30 次,带宽的选择通过交叉验证最小化分位数损失来选择,有

$$h=\arg\min_h\sum_{i=1}^{n}\rho_\tau(Y_i-\hat{g}^{(-i)}(\boldsymbol{X}_i)-\boldsymbol{Z}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(0)}(\tau)),\tag{19}$$

其中,$\hat{g}^{(-i)}(\cdot)$ 为剔除第 i 个样本后的估计量。

得到的半参数分位数双重稳健估计模型(QDR)、线性分位数回归模型(LQR)的偏差 Bias、均方根误差 δ_{RMSE} 结果记录在表 1,并绘制出每个场景下 $g(\boldsymbol{X})$ 的拟合曲线图,如图 1。

表 1　模型估计表现

误设情况	model	Bias	δ_{RMSE}	$g(\boldsymbol{X})$ error	$\boldsymbol{\beta}(\tau)$ error
$g(\boldsymbol{X})$ 被忽略	QDR	−0.000 7	0.662 8	0.128 1	0.072 2
	LQR	−0.001 5	1.202 4	0.418 5	0.122 9
$\boldsymbol{\beta}(\tau)$ 被误设	QDR	0.000 1	0.809 1	0.109 0	−0.000 4
	LQR	0.055 9	1.201 4	0.617 5	0.003 1
$g(\boldsymbol{X})$、$\boldsymbol{\beta}(\tau)$ 均误设	QDR	−0.000 1	0.649 9	−0.019 8	0.002 1
	LQR	−0.001 1	1.202 1	0.399 8	0.003 1

从表 1 和图 1 可得到：

(1)本文提出的模型的偏差、均方根误差均显著性低于线性分位数回归模型，说明通过 Neyman 正交化联合优化非参数部分和参数部分，即使 $g(\boldsymbol{X})$ 或 $\boldsymbol{\beta}(\tau)$ 中有错误设定或者全部错误假设，估计量仍具有一致性。

(2)在错误设定情况下，δ_{RMSE} 比率稳定在 60%左右，意味着相较于线性分位数回归，本文模型的误差预测降低了 45%左右。在面对高维数据，比如本文在 $\boldsymbol{Z}$ 维度为 20 时，$\boldsymbol{\beta}$ 的误差仍能低于 0.001，说明本文模型确实能有效预防“维度灾难”。在估计偏差时，3 个场景的偏差均接近 0，但通过具体数值的比较，能说明 QDR 在模型误设时仍能保持更小的偏差。

(3)图 1 所有场景中，本文模型在定义域边界处均未出现明显的估计不稳定现象，表明样条平滑方法具有较好的边界适应性。

(4)本文模型在 3 种不同模型设定下均能较好地捕捉真实 $g(\boldsymbol{X})$ 的函数形态，估计曲线与真实曲线趋势基本吻合，甚至即使在模型设定均错误的情况下，模型仍能保持相对稳定的估计性能，未出现明显的系统性偏差，说明本文模型确实能有效提升稳健性。

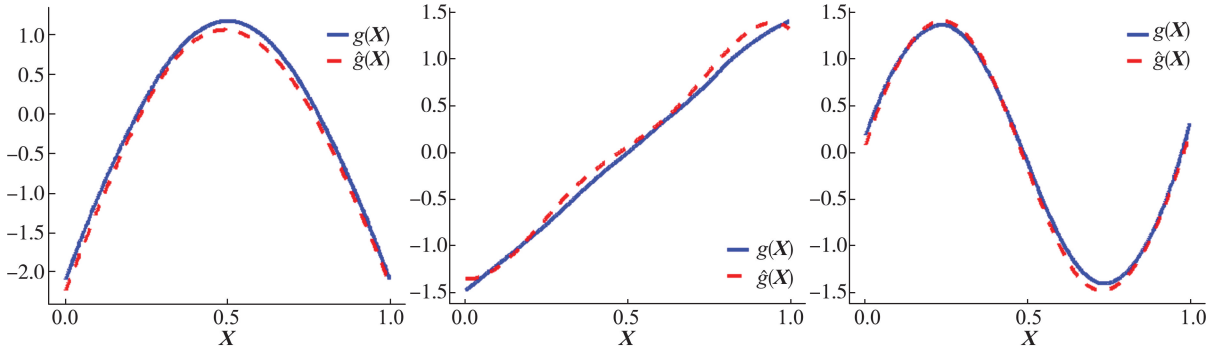


图 1　不同场景下 $g(\boldsymbol{X})$ 的拟合曲线图(从左到右分别是场景 1、2、3)

5　定理证明

5.1　3 个引理

为进行定理 1 和定理 2 的证明，先给出 3 个引理。

引理 1　在假设(1)~(7)下，对于任意 $u,v\in R$ 有

$$\rho_\tau(u-v)-\rho_\tau(u)=-v\psi_\tau(u)+\int_0^v(I(u\leqslant s)-I(u\leqslant 0))\mathrm{d}s。\tag{20}$$

证明　将上式的左边记为 $H(v)$，其中，$\psi_\tau(u)=\tau-I(u<0)$ 是得分函数。对 $H(v)$ 求右导数得 $\frac{\partial^+H}{\partial v}=-\frac{\partial^+\rho_\tau}{\partial u}\big|_{u=u-v}+\psi_\tau(u)$，在 $u=u_0$ 处对其求右导数得：

$$\frac{\partial^+\rho_\tau}{\partial u}\big|_{u=u_0}=\begin{cases}\lim\limits_{h\to0^+}\dfrac{\tau(u_0+h)-\tau u_0}{h},\text{if } u>0,\\ \lim\limits_{h\to0^+}\dfrac{\rho_\tau(h)-\rho_\tau(0)}{h},\text{if } u=0,\\ \lim\limits_{h\to0^+}\dfrac{(1-\tau)|u_0+h|-(1-\tau)|u_0|}{h},\text{if } u<0\end{cases}=\begin{cases}\tau,\text{if } u\geqslant0,\\ \tau-1,\text{if } u<0,\end{cases}\tag{21}$$

将 $u_0=u-v$ 代入式(21)即有 $\frac{\partial^+\rho_\tau}{\partial u}\big|_{u=u-v}=\tau-I(u-v<0)$。则 $H(v)$ 的右导数有

$$\frac{\partial^+H}{\partial v}=-(\tau-I(u-v<0))+(\tau-I(u\leqslant0))=I(u-v<0)-I(u\leqslant0)。\tag{22}$$

利用微积分基本定理则计算可得

$$H(v)-H(0)=\int_0^v\frac{\partial^+H}{\partial s}\mathrm{d}s=\int_0^v(I(u-s<0)-I(u\leqslant0))\mathrm{d}s。\tag{23}$$

$H(v)=0$，经过计算得到引理 1。

引理 2　在假设(1)~(7)下,有

$$\| g^{*}-g_0 \|_{\infty} \leqslant Cm^{-s}。\tag{24}$$

证明　对于每个点 $x \in X$,构造 Taylor 多项式 $P_x(y)=\sum_{|\alpha| \leqslant k} \frac{D^{\alpha} g_o(x)}{\alpha!}(y-x)^{\alpha}$,$k=\lfloor s \rfloor$,当 $C(X) \to G_m$,再定义一个拟插值算子

$$Qg_0=\sum_{i=1}^{m} \lambda_j(g_o) B_j(x),\tag{25}$$

其中,G_m 为样本空间,其包含所有 $k=\lfloor s \rfloor$ 阶多项式,泛函 λ_j 满足 $|\lambda_j(g_0)| \leqslant C \sup_{x \in \operatorname{supp}(B_j)} |g_0(x)|$,且对多项式 p 有 $Qp=p$。

当 $y \in I_j$,有

$$\begin{aligned}
&|g_0(y)-Qg_0(y)| \\
&\leqslant |g_0(y)-P_{cj}(y)|+|P_{cj}(y)-QP_{cj}(y)|+|Q(P_{cj}(y)-g_0(y))| \\
&\leqslant L \| y-cj \|^{s}+0+\sum_{i} |\lambda_i(P_{cj}(y)-g_0(y))| \| B_i(y) \\
&\leqslant L(d^{1/2}h)^{s}+C \sup_{z \in X} |P_{cj}(z)-g_0(z)| \\
&\leqslant CLh^{s}。
\end{aligned}\tag{26}$$

将式(26)代入 $h \approx m^{-1}$ 则有 $| \| g_0-Qg_0 \|_{\infty} \leqslant Cm^{-s}$,将 $g^{*}=Qg_0$ 代入计算后引理 2 即得证。

引理 3　在假设(1)~(7)下,有

$$\mathrm{E}\left[\sup_{\boldsymbol{\beta} \in B}\left|\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_i \psi_{\tau}(Y_i-\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{\beta})\right|\right] \lesssim C\sqrt{\frac{s \log p}{n}}。\tag{27}$$

证明　首先引入 Rademacher 变量 $\{\sigma_i\}_{i=1}^{n}$,通过对称化不等式为

$$\mathrm{E}\left[\sup_{\boldsymbol{\beta} \in B}\left|\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_i \psi_{\tau}(\varepsilon_i^{\boldsymbol{\beta}})-\mathrm{E}[\boldsymbol{Z}\psi_{\tau}(\varepsilon^{\boldsymbol{\beta}})]\right|\right] \leqslant 2\mathrm{E}\left[\mathrm{E}_{\sigma}\left[\sup_{\boldsymbol{\beta} \in B}\left|\frac{1}{n}\sum_{i=1}^{n} \sigma_i \boldsymbol{Z}_i \psi_{\tau}(\varepsilon_i^{\boldsymbol{\beta}})\right|\right]\right],\tag{28}$$

其中,$\varepsilon_i^{\boldsymbol{\beta}}=Y_i-\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau)\sigma_i$,$\sigma_i$ 为 i.i.d.变量且 $P(\sigma_i=\pm 1)=1/2\sigma_i$。

定义 $\xi_i^{\boldsymbol{\beta}}=\mathrm{E}[I(\varepsilon_i^{\boldsymbol{\beta}} \leqslant 0) \mid Z_i]-I(\varepsilon_i^{\boldsymbol{\beta}} \leqslant 0)$,且对稀疏类 $B=\{\boldsymbol{\beta} \in \mathbb{R}^{p}: \| \boldsymbol{\beta} \|_0 \leqslant s\}$,构造 ε-网 N_{ε}:$\log |N_{\varepsilon}| \leqslant s\log(ep/\varepsilon s)$。设 $\| Z_i \|_{\psi_2} \leqslant K$,固定 $\boldsymbol{\beta}$、$\boldsymbol{Z}_i\xi_i^{\boldsymbol{\beta}}$ 为亚高斯随机向量,利用 Dudley 熵积分,对式(28)左边的期望拆分为:

$$\begin{aligned}
\mathrm{E}_{\sigma}\left[\sup_{\boldsymbol{\beta}}\left|\frac{1}{n}\sum_{i=1}^{n} \sigma_i \boldsymbol{Z}_i \psi_{\tau}(\varepsilon_i^{\boldsymbol{\beta}})\right|\right] &\leqslant \mathrm{E}_{\sigma}\left[\sup_{\boldsymbol{\beta}}\left|\frac{1}{n}\sum_{i=1}^{n} \sigma_i \boldsymbol{Z}_i \psi_{\tau}(\xi_i^{\boldsymbol{\beta}})\right|\right]+R_n \\
&\leqslant \frac{1}{\sqrt{n}}\int_0^{D} \sqrt{\log N(\varepsilon, B, d_n)}\, \mathrm{d}\varepsilon。
\end{aligned}\tag{29}$$

其中,$d_n(\boldsymbol{\beta},\boldsymbol{\beta}')=\frac{1}{n}\sum_{i=1}^{n} \| \boldsymbol{Z}_i \| \cdot |\boldsymbol{Z}_i^{\mathrm{T}}(\boldsymbol{\beta}-\boldsymbol{\beta}')|^2$,$\log N(\varepsilon, B, d_n) \leqslant s\log(Cp/\varepsilon\sqrt{s})$。

取 $D=\sup_{\boldsymbol{\beta},\boldsymbol{\beta}'} \mathrm{d}_{\mathrm{n}}(\boldsymbol{\beta},\boldsymbol{\beta}') \leqslant 2K\sqrt{s}$,计算得 $\int_0^{D} \sqrt{s\log(\frac{Cp}{\varepsilon\sqrt{s}})}\, \mathrm{d}\varepsilon \leqslant C'\sqrt{\frac{s\log p}{n}}$。

代入计算则引理 3 得证。

5.2　定理 1 证明

对式(5)设:$\varepsilon_i(\boldsymbol{\beta})=Y_i-g(\boldsymbol{X}_i)-\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{\beta}$、$L(\boldsymbol{\beta})=\rho_{\tau}(\varepsilon_i)$。

由于分位数损失函数在 0 点不可导,参考引理 1 可以换为求 $L(\boldsymbol{\beta})$ 次梯度求导,有

$$\frac{\partial}{\partial \boldsymbol{\beta}}(L(\boldsymbol{\beta}))=-\sum_{i=1}^{n} \boldsymbol{Z}_i \psi_{\tau}(\varepsilon_i(\boldsymbol{\beta}))。\tag{30}$$

令 $\boldsymbol{S}(\boldsymbol{\beta})=\sum_{i=1}^{n} \boldsymbol{Z}_i \psi_{\tau}(\varepsilon_i(\boldsymbol{\beta}))$,在 $\hat{\boldsymbol{\beta}}$ 处有 $\boldsymbol{S}(\hat{\boldsymbol{\beta}})=0$,$\boldsymbol{S}(\boldsymbol{\beta})$ 在 $\hat{\boldsymbol{\beta}}$ 处一阶泰勒展开有

$$\boldsymbol{S}(\hat{\boldsymbol{\beta}})=\boldsymbol{S}(\beta_0)+\frac{\partial \boldsymbol{S}(\boldsymbol{\beta}^{*})}{\partial \boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+R_n=0,\tag{31}$$

为进一步求解,利用 Jacobian 矩阵得到

$$\frac{\partial \boldsymbol{S}(\boldsymbol{\beta}^*)}{\partial \boldsymbol{\beta}}=-\sum_{i=1}^{n}\boldsymbol{Z}_i\frac{\partial \psi_\tau(\varepsilon_i(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}}=-\sum_{i=1}^{n}\boldsymbol{Z}_i\frac{\partial \psi_\tau}{\partial \boldsymbol{\beta}}\cdot Z_i^{\mathrm{T}}。\tag{32}$$

由于 $\psi_\tau(u)$ 在 $u=0$ 处不可导,参考引理 3 可以参考其期望意义,即对其求期望有

$$\begin{aligned}\mathrm{E}\left[\frac{\partial \psi_\tau(\varepsilon_i(\boldsymbol{\beta}))}{\partial}\right]&=\mathrm{E}\left[\frac{\partial}{\partial \boldsymbol{\beta}}\mathrm{E}\left[\psi_\tau(\varepsilon_i(\boldsymbol{\beta}))\mid \boldsymbol{X}_i,\boldsymbol{Z}_i\right]\right]\\&=\mathrm{E}\left[\frac{\partial}{\partial \boldsymbol{\beta}}(\tau-\mathrm{E}\left[I(\varepsilon_i(\boldsymbol{\beta})\leqslant 0)\mid \boldsymbol{X}_i,\boldsymbol{Z}_i\right])\right]\\&=\mathrm{E}\left[\frac{\partial}{\partial \boldsymbol{\beta}}(\tau-F_{\varepsilon\mid X,Z}(0;\boldsymbol{\beta}))\right],\end{aligned}\tag{33}$$

再由 Leibniz 法则得到$\frac{\partial}{\partial \boldsymbol{\beta}}F_{\varepsilon\mid X,Z}(0;\boldsymbol{\beta})=\frac{\partial}{\partial \boldsymbol{\beta}}P(\varepsilon_i(\boldsymbol{\beta})\leqslant 0\mid \boldsymbol{X}_i,\boldsymbol{Z}_i)=-f_{\varepsilon\mid X,Z}(0;\boldsymbol{\beta})\boldsymbol{Z}_i$。在$\boldsymbol{\beta}_0$ 处有 $\varepsilon_i(\boldsymbol{\beta}_0)=\varepsilon_i$,满足 $P(\varepsilon_i\leqslant 0\mid \boldsymbol{X}_i,\boldsymbol{Z}_i)=\tau$ 且条件密度为$f_{\varepsilon\mid X,Z}(0)$。进而有

$$\mathrm{E}\left[\frac{\partial \psi_\tau(\varepsilon_i(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}}\right]=\mathrm{E}\left[f_{\varepsilon\mid X,Z}(0)\boldsymbol{Z}_i\right]。\tag{34}$$

结合一致性$\boldsymbol{\beta}^*\xrightarrow{p}\boldsymbol{\beta}_0$、大数定律有

$$\frac{1}{n}\frac{\partial \boldsymbol{S}(\hat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}}\xrightarrow{p}\mathrm{E}\left[\frac{\partial S(\boldsymbol{\beta}_0)}{\partial \boldsymbol{\beta}}\right]=-\boldsymbol{\Gamma}。\tag{35}$$

将式(35)代入式(31)计算得

$$\begin{aligned}\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0&=-\left(\frac{\partial \boldsymbol{S}(\boldsymbol{\beta}^*)}{\partial \boldsymbol{\beta}}\right)^{-1}\cdot\left[\boldsymbol{S}(\boldsymbol{\beta}_0)+R_n\right]\\&=\left(\frac{1}{n}\boldsymbol{\Gamma}^{-1}+o_p\left(\frac{1}{n}\right)\right)\cdot\left[\boldsymbol{S}(\boldsymbol{\beta}_0)+R_n\right]。\end{aligned}\tag{36}$$

根据稀疏性可得 $\|R_n\|_\infty\leqslant C\mathrm{slogp}\cdot\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\|=o_p(1/\sqrt{n})$,可进一步根据正交性条件计算得到 $\hat{\boldsymbol{S}}(\boldsymbol{\beta})=\sum_{i=1}^{n}\boldsymbol{Z}_i\psi_\tau(\varepsilon_i)-\eta(\boldsymbol{X}_i)(g(\boldsymbol{X}_i)-g_0(\boldsymbol{X}_i))$。其中 $\eta(\boldsymbol{X}_i)=\mathrm{E}\left[\boldsymbol{Z}_i f_{\varepsilon\mid X,Z}(0)\mid \boldsymbol{X}_i\right]$,有 $\mathrm{E}[\eta(\boldsymbol{X}_i)(\hat{g}(\boldsymbol{X}_i)-g_0(\boldsymbol{X}_i))]=0$,又有 $\|\hat{g}-g_0\|=O_p(m^{-s})=o_p(n^{-1/4})$,故

$$\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)=\boldsymbol{S}(\boldsymbol{\beta}_0)+o_p(1/\sqrt{n})。\tag{37}$$

将稀疏性结果和式(37)代入式(36)计算得

$$\begin{aligned}\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0&=\frac{1}{n}\boldsymbol{\Gamma}^{-1}\left[\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)+o_p(1/\sqrt{n})\right]+o_p(1/\sqrt{n})\\&=\frac{1}{n}\boldsymbol{\Gamma}^{-1}\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)+\frac{1}{n}\boldsymbol{\Gamma}^{-1}o_p(1/\sqrt{n})+o_p(1/\sqrt{n})\\&=\frac{1}{n}\boldsymbol{\Gamma}^{-1}\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)+o_p(1/\sqrt{n})。\end{aligned}\tag{38}$$

为方便后续计算,对式(38)进行标准化

$$\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)=\boldsymbol{\Gamma}^{-1}\cdot\frac{1}{\sqrt{n}}\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)+o_p(1),\tag{39}$$

其中,$\boldsymbol{\Gamma}$ 是信息矩阵,所以 $\boldsymbol{\Gamma}$ 是对称矩阵,故$\boldsymbol{\Gamma}^{-1}$也是对称矩阵,即$(\boldsymbol{\Gamma}^{-1})^{\mathrm{T}}=\boldsymbol{\Gamma}^{-1}$。

显然,$\mathrm{E}[\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)]=0$,$\mathrm{Var}(\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0))=\tau(1-\tau)\boldsymbol{\Sigma}$。根据多元中心极限定理有$\frac{1}{\sqrt{n}}\boldsymbol{S}(\boldsymbol{\beta}_0)\xrightarrow{d}\mathrm{N}_p(0,\tau(1-\tau)\boldsymbol{\Sigma})$,将其代入到式(39)

$$\frac{1}{\sqrt{n}}\hat{\boldsymbol{S}}(\boldsymbol{\beta}_0)\xrightarrow{d}\mathrm{N}_p(0,\tau(1-\tau)\boldsymbol{\Sigma})。\tag{40}$$

将式(40)代入式(39),并结合 Slutsky 定理得:

$$\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\xrightarrow{d}\mathrm{N}(0,\tau(1-\tau)\boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1})\ 。\tag{41}$$

再结合$\boldsymbol{\Gamma}^{-1}\prec f_{\min}^{-1}\boldsymbol{\Sigma}^{-1}$、Slutsky 定理,对于任意方向 $\boldsymbol{a}$ 有

$$\sqrt{n}\boldsymbol{a}^{\mathrm{T}}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\xrightarrow{d}\mathrm{N}(0,\tau(1-\tau)\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1}\boldsymbol{a})\prec\mathrm{N}(0,\frac{\tau(1-\tau)}{f_{\min}^2}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{a})\ 。\tag{42}$$

定理 1 即证得。

5.3　定理 2 证明

对式(9)定义:$\boldsymbol{\epsilon}_i=Y_i-\boldsymbol{B}(\boldsymbol{X}_i)\boldsymbol{\theta}-Z_i^{\mathrm{T}}\boldsymbol{\beta}$,$L(\boldsymbol{\theta})=\rho_\tau(\boldsymbol{\epsilon}_i(\boldsymbol{\theta}))$。根据引理 2 可知 $\hat{g}(\boldsymbol{X})=\boldsymbol{B}(X)\hat{\boldsymbol{\theta}}$。

由引理 1、定理 1 对 $L(\boldsymbol{\theta})$求次梯度有$\dfrac{\partial L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=-\sum_{i=1}^{n}\boldsymbol{B}(\boldsymbol{X}_i)^{\mathrm{T}}\psi_\tau(\boldsymbol{\epsilon}_i(\boldsymbol{\theta}))$。

令 $\boldsymbol{T}(\boldsymbol{\theta})=-\sum_{i=1}^{n}\boldsymbol{B}(\boldsymbol{X}_i)^{\mathrm{T}}\psi_\tau(\boldsymbol{\epsilon}_i(\boldsymbol{\theta}))$,对于 $\boldsymbol{T}(\boldsymbol{\theta})$在$\boldsymbol{\theta}_0$ 处一阶泰勒展开得

$$\boldsymbol{T}(\hat{\boldsymbol{\theta}})=\boldsymbol{T}(\boldsymbol{\theta}_0)+\frac{\partial\boldsymbol{T}(\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)+R_n=0\ 。\tag{43}$$

类似定理 1 中的证明结果,直接计算得到 $\boldsymbol{\Gamma}=\mathrm{E}[f_{\epsilon|X}(0)\boldsymbol{B}(X)^{\mathrm{T}}\boldsymbol{B}(X)]$,$\boldsymbol{\Omega}=\tau(1-\tau)\mathrm{E}[\boldsymbol{B}(X)^{\mathrm{T}}\boldsymbol{B}(X)]=\tau(1-\tau)\boldsymbol{\Sigma}$,由大数定律得

$$\frac{1}{n}\frac{\partial\boldsymbol{T}(\hat{\boldsymbol{\theta}})}{\partial\theta}\xrightarrow{p}-\boldsymbol{\Gamma}。\tag{44}$$

参考 Koenker[1] 中,定义交叉项 $\Lambda=\mathrm{E}[f_{\epsilon|X}(0)\boldsymbol{B}(X)^{\mathrm{T}}\boldsymbol{Z}]=0$,则 $\boldsymbol{T}(\boldsymbol{\theta})$在$\boldsymbol{\theta}_0$ 处有

$$\boldsymbol{T}(\boldsymbol{\theta}_0)=\sum_{i=1}^{n}\boldsymbol{B}(\boldsymbol{X}_i)^{\mathrm{T}}\psi_\tau(\boldsymbol{\epsilon}_i)+o_p(1/\sqrt{n})\ ,\tag{45}$$

做定理 1 中类似证明,有 $\|R_n\|=O_p(\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|^2)$。由$|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0|=O_p(1)$得 $R_n=O_p(1)$ 。

则将式(45)、(44)代入式(43),进行计算有

$$\begin{aligned}\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0&=-\left(\frac{\partial T(\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}}\right)^{-1}\cdot[\boldsymbol{T}(\boldsymbol{\theta}_0)+R_n]\\&=-\left[\frac{1}{n}\boldsymbol{\Gamma}^{-1}+o_p\left(\frac{1}{n}\right)\right][\boldsymbol{T}(\boldsymbol{\theta}_0)+O_p(\sqrt{n})]\\&=\frac{1}{n}\boldsymbol{\Gamma}^{-1}\boldsymbol{T}(\boldsymbol{\theta}_0)+o_p(1/\sqrt{n})\ 。\end{aligned}\tag{46}$$

为方便后续计算,对式(46)进行标准化得

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)=\boldsymbol{\Gamma}^{-1}\cdot\frac{1}{\sqrt{n}}\boldsymbol{T}(\boldsymbol{\theta}_0)+o_p(1)\ ,\tag{47}$$

由多元中心极限定理有

$$\frac{1}{\sqrt{n}}\boldsymbol{T}(\boldsymbol{\theta}_0)\xrightarrow{d}\mathrm{N}_m(0,\Omega)\ ,\tag{48}$$

将式(48)代入式(47),结合 Slutsky 定理、多元中心极限定理计算得到

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)=\boldsymbol{\Gamma}^{-1}\cdot\frac{1}{\sqrt{n}}\boldsymbol{T}(\boldsymbol{\theta}_0)+o_p(1)\xrightarrow{d}\mathrm{N}_m(0,\boldsymbol{\Gamma}^{-1}\boldsymbol{\Omega}\boldsymbol{\Gamma}^{-1})\ 。\tag{49}$$

在点态估计中,估计误差分解有 $\hat{g}(\boldsymbol{X})-g_0(\boldsymbol{X})=\boldsymbol{B}(X)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)-\boldsymbol{b}(\boldsymbol{X})$,对其标准化处理后为

$$\sqrt{n}(\hat{g}(\boldsymbol{X})-g_0(\boldsymbol{X}))=\boldsymbol{B}(X)\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)-\sqrt{n}\boldsymbol{b}(\boldsymbol{X})\ 。\tag{50}$$

根据假设 5 有$\sqrt{n}\boldsymbol{b}(X)=O(\sqrt{n}m^{-s})\xrightarrow{p}0$,将式(50)代入式(49),由连续映射定理即可推出定理 2 的结论。

6　结　语

本文通过理论研究与数值模拟研究,构建出基于分位数模型下的双重稳健估计框架,基于 1 000 次重复

模拟实验,引入正则化约束,用偏差和均方根误差对比分析 2 个模型在估计中的表现,均发现该模型在存在模型错误设定和异常值时仍能保持一致性,验证了模型错误设定的鲁棒性,其稳健性能良好,能更有效解决高维变量的“维数灾难”问题,有望运用于医学、经济学、金融学等领域。

尽管模型具有显著效果,但仍存在一些需要进一步研究的问题。比如,其模型在面对更为复杂的半参数模型时的效率问题有待解决,另外,能否将这一框架融入最新的机器学习算法也是一个得研究的重要方向。

参考文献:

[1] Koenker R,Bassett G J.Regression quantiles[J].Econometrica,1978,46(1):33-50.

[2] de Boor C.A practical guide to splines[M].New York:Springer,1978.

[3] Hastie T,Tibshirani R.Generalized additive models:some applications[J].Journal of the American Statistical Association,1987,82(398):371-386.

[4] Robins J M,Rotnitzky A.Semiparametric efficiency in multivariate regression models with missing data[J].Journal of the American Statistical Association,1995,90(429):122-129.

[5] Laan M J,Robins J M.Unified methods for censored longitudinal data and causality[M].New York:Springer,2003.

[6] van der Vaart A W.Asymptotic statistics[M].Cambridge:Cambridge University Press,2000.

[7] Portnoy S.Local asymptotics for quantile smoothing splines[J].The Annals of Statistics,1997,25(1):414-434.

[8] Koenker R,Ng P,Portnoy S.Quantile smoothing splines[J].Biometrika,1994,81(4):673-680.

[9] Stone C J.Optimal global rates of convergence for nonparametric regression[J].The Annals of Statistics,1982,10(4):1040-1053.

[10] Bang H,Robins J M.Doubly robust estimation in missing data and causal inference models[J].Biometrics,2005,61(4):962-973.

[11] Koenker R.Quantile regression[M].Cambridge:Cambridge University Press,2005.

[12] Zou Hui,Yuan Ming.Composite quantile regression and the oracle model selection theory[J].The Annals of Statistics,2008,36(3):1108-1126.

[13] Tan Zhiqiang.Bounded,efficient and doubly robust estimation with inverse weighting[J].Biometrika,2010,97(3):661-682.

[14] Zheng Qi,Peng Limin,He Xuming.Globally adaptive quantile regression with ultra-high dimensional data[J].The Annals of Statistics,2015,43(5):2225-2258.

[15] Chernozhukov V,Hansen C.Instrumental variable quantile regression:A robust inference approach[J].Journal of Econometrics,2008,142(1):379-398.

[16] Belloni A,Chernozhukov V,Chetnerikov D,et al.Uniformly valid post-regularization confidence regions for many functional parameters in z-estimation framework[J].The Annals of Statistics,2018,46(6B):3643-3675.

[17] Caracciolo F,Furno M.Quantile treatment effect and double robust estimators:an appraisal on the Italian labor market[J].Journal of Economic Studies,2017,44(4):585-604.

[18] Liu Molei,Zhang Yi,Zhou Doudou.Double/debiased machine learning for logistic partially linear model[J].The Econometrics Journal,2021,24(3):559-588.

[19] 丁先文,张文,袁红,等.含缺失数据的半参数模型的稳健估计[J].统计与决策,2022,38(1):25-28.

[20] 田玲,孙宁,杨琛.基于半参数分位数回归法的地震损失研究[J].统计与决策,2019,35(9):11-14.

[21] 许文超,张日权,梁华.两样本广义函数型线性模型的估计和半参数有效性[J].中国科学:数学,2025,55(6):1123-1158.

[22] 张日权.现代非参数统计方法[M].北京:科学出版社,2012.

[23] Chernozhukov V,Escanciano J C,Ichimura H,et al.Locally robust semiparametric estimation[J].Econometrica,2022,90(4):1501-1535.

[24] Blanchet J,Li Jiajin,Lin Sirui,et al.Distributionally robust optimization and robust statistics[J].Statistical Science,2025,40(3):351-377.

(责任编辑:孙洪清;校 对:徐学友)