arXiv:2601.00541v1 [math.ST] 2 Jan 2026

基于投影经验过程和 p值合并的超高维参数回归渐近分布无关检验 *

谭法龙[1], 唐杉[2], 和朱力行[3*]

[1] 湖南大学统计与数据科学系，中国长沙[2] 武汉理工大学数学与统计学院，中国武汉
[3] 北京师范大学珠海校区统计系，中国珠海

摘要

本文提出一种新方法，基于投影经验过程和 p值合并，对稀疏参数回归模型进行拟合优度检验，其中协变量维度可能远远超过样本量。在此类超高维情形下，传统的基于经验过程的检验往往会因维数灾难，或依赖参数估计量的渐近线性和正态性而失效；这些性质在超高维情形下可能并不成立。为克服这些挑战，我们首先在温和条件下将经典鞅变换推广至超高维情形，并针对单位球面上的任意投影，基于经鞅变换的投影残差标记经验过程构造一种 Cram´er-von Mises 型检验。鞅变换使该投影检验具有渐近分布无关性，并使我们仅凭参数估计量的标准收敛速率即可推导其极限分布。在温和条件下，该投影检验对单位球面上几乎所有投影均具有一致性，但针对某些特定投影仍可能出现检验功效损失。因此，我们进一步采用强有力的 p值合并方法（如柯西合并），汇总多个投影对应的 p值，从而增强整体稳健性。此外，鉴于基于经验过程的检验擅长检测低频信号，而局部平滑检验通常更适用于高频备择，我们提出一种新型混合检验，利用柯西合并汇总这两种方法。由此得到的混合检验对低频和高频备择均具有较强检验功效。我们通过详细的模拟研究和两项真实数据分析，说明该方法在超高维情形下的有效性。

关键词： 维数灾难、广义线性模型、混合检验、鞅变换、投影。

*通讯作者。朱力行的研究得到了国家自然科学基金项目（NSFC12131006、NSFC12471276）以及中国中医科学院科技创新工程项目（CI2023C063YLL）的资助。

1 引言

本研究旨在解决超高维回归模型的拟合优度检验问题，其中协变量的维数可能远远超过样本量。考虑如下回归模型:

$$Y = m(X) + \varepsilon, \tag{1.1}$$

其中， $Y \in \mathbb{R}$ 是响应变量， X 是 p维协变量向量， $m(\cdot) = E(Y|X = \cdot)$是未知回归函数， ε 是满足 $E(\varepsilon|X) = 0$的误差项。我们的目标是检验在超高维情形下，均值函数 $m(\cdot)$是否属于某个参数函数类 $\mathcal{M} = \{m(\cdot, \beta) : \beta \in \Theta \subset \mathbb{R}^q\}$ 。

关于低维情形下回归模型拟合优度检验的文献十分丰富，其中维数 p 被视为固定且小于样本量 n。模型检验的一种主要方法称为局部平滑检验，它基于对条件矩限制 $E[\varepsilon(\beta_0)|X]$ 的非参数估计，其中 $\beta_0 \in \Theta$，且 $\varepsilon(\beta_0) = Y - m(X, \beta_0)$。相关例子包括 H¨ardle 和 Mammen (1993)、Zheng (1996)、Dette (1999)、Fan 和 Huang (2001)、Horowitz 和 Spokoiny (2001)、Koul 和 Ni (2004)、Van Keilegom 等 (2008)、Lavergne and Patilea (2008, 2012)、Guo 等 (2016)。这类检验通常是渐近分布无关的，并且对高频备择模型尤为敏感（Horowitz 和 Spokoiny，2001）。然而，由于依赖非参数估计，它们通常会严重受到维数灾难的影响。另一种主要的模型检验方法是基于经验过程构造检验统计量，从而避免对条件矩限制$E[\varepsilon(\beta_0)|X]$进行非参数估计。例如，参见 Bierens (1982)、Stute (1997)、Bierens (1990)、Stute、Gonz´alez Manteiga 和 Presedo Quindimil (1998)、Stute 和 Zhu（2002）、Zhu (2003)、Escanciano (2006b)、Stute 等 (2008)、Escanciano 等 (2018)、Cuesta-Albertos 等 (2019)、Lu 和 Zheng（2020）、Escanciano (2024)。这些基于经验过程的检验通常对低频备择模型具有较强检验力，并且能够以参数速率 $n^{-1/2}$检测局部备择；该速率是假设检验中的最优检测速率。

然而，基于经验过程的检验通常要求参数估计量满足渐近线性或正态性，才能推导其极限零分布。为说明这一点，我们考虑 Stute（1997）的开创性论文。该论文指出，许多回归拟合优度检验都基于经验过程 $\hat{} S_n(t) = n^{-1/2} \sum_{i=1}^n \varepsilon_i(\hat{}\beta) I(X_i \le t)$，其中 $\varepsilon_i(\hat{}\beta) = Y_i - m(X_i, \hat{}\beta)$，$\hat{}\beta$ 是 β 的一致估计量，而 $\{(X_i, Y_i)\}_{i=1}^n$ 是一个与（X，Y）具有相同分布的独立同分布样本。在一些正则条件和固定维度设定下，Stute（1997）证明，在原假设下，

$$\hat{S}_n(t) = S_n^0(t) + \sqrt{n}(\hat{\beta} - \beta_0)^\top M(t) + o_p(1), \tag{1.2}$$

关于 t 一致成立，其中 $M(t) = E[m'(X, \beta_0) I(X \le t)]$ 且 $S_n^0(t) = n^{-1/2} \sum_{i=1}^n \varepsilon_i(\beta_0) I(X_i \le t)$，其中$\varepsilon_i(\beta_0) = Y_i - m(X_i, \beta_0)$。不难看出，要推导 $\hat{} S_n(t)$ 的极限零分布，需要$\sqrt{n}(\hat{}\beta - \beta_0)$ 具有渐近线性展开或满足正态性。然而，在高维情形下，估计参数 β 的这一渐近性质对于 Lasso、post-Lasso 或其变体等方法可能不再成立。因此，若不对相应经验过程作进一步变换，纳入这些估计方法的基于经验过程的检验通常无法直接推广至协变量维度

p 显著超过样本量 n。另一个关键问题是，由于高维空间中的数据稀疏性，现有的大多数基于经验过程的检验都受到维数灾难的影响；关于这一问题的更多详情，参见 Escanciano (2006b) 和 Tan 等（2025）。近年来，Shah 和 B¨uhlmann（2018）以及 Jankov´a 等人（2020）分别针对高维线性模型和广义线性模型，提出了基于残差预测和广义残差预测的两种拟合优度检验，其中协变量维度 p 可以远大于样本量n。然而，Shah 和 B¨uhlmann（2018）研究的是固定设计回归模型的拟合优度检验，这与本文不同。Jankov´a 等人（2020）提出了一种基于投影残差 $w^\top \hat{} R$ 的广义残差预测（GRP）检验，其中 $\hat{} R$ 和 w 分别为估计残差向量和单位球面上的投影。需要注意的是，GRP 检验仅考察模型设定错误的某一特定方面；它可能无法捕捉原假设的所有潜在偏离，并且在高维情形下对某些备择假设缺乏检验效能。

本文旨在为回归模型开发一种新的拟合优度检验，使其能够应用于超高维情形，同时缓解维数灾难。需要指出的是，基于经验过程的检验通常要求参数估计量 β 满足渐近线性或正态性，以推导其极限分布。由于在超高维情形下， β 的这一假设可能不成立，因此相应的基于经验过程的检验也可能无法应用于这些情形。有趣的是，我们发现，模型检验中的经典鞅变换（Stute 等（1998））或可用于解决这一问题。该方法可追溯至 Khmaladze（1981）提出的累积分布函数拟合优度检验；另见 Koul 和 Stute（1999）、Bai 和 Ng（2001）、Koenker 和 Xiao（2002）、Bai（2003）、Khmaladze 和 Koul（2004，2009）、Delgado 和 Stute（2008）、Tan 和 Zhu（2019）、Lu 和 Zheng（2020）以及 Tan 等（2025）等众多文献。例如，鞅变换 T 是一种线性算子，使得基于经鞅变换的经验过程构造的检验能够渐近分布无关。更具体地说，它通过令 $TM(t) \equiv 0$，消除 $S_n(t)$分解中的位移函数，例如$\hat{} M(t)$，并同时确保变换后的过程 $TS_n^0(t)$具有与 $S_n^0(t)$相同的渐近性质。这意味着 $TS_n^0(t)$以及由此得到的 $T\hat{} S_n(t)$将渐近分布无关，其极限为变换后的时间上的布朗运动。请注意，鞅变换消除了函数 $M(t)$，因此$\hat{}$ $S_n(t)$中的平移项$\sqrt{n}(\hat{}\beta - \beta_0)^\top M(t)$也会消失。因此，经鞅变换的经验过程不再涉及估计参数 β，从而可用于超高维模型检验。然而，Stute 等（1998）为模型检验提出的经典鞅变换是针对单变量协变量设计的。Khmaladze 和 Koul（2004）以及 Delgado 和 Stute（2008）进一步将单变量鞅变换推广到固定维数情形下的多变量情形。然而，这两种方法都涉及关于 p维协变量的多重积分（Lu 和 Zheng（2020）），因此难以推广到超高维情形。据我们所知，当协变量维度 p可能大幅超过样本量时，鞅变换能否应用于这类情形仍是文献中的一个未解决问题。

本文受 Stute 和 Zhu（2002）提出的降维检验启发，在温和条件下成功地将经典鞅变换扩展到超高维情形。基于这一扩展，我们提出了一种检验稀疏回归模型拟合优度的新方法，其中协变量维度 p 可能远大于样本量 n。我们的方法包括两个步骤。

首先，对于 $\mathbb{R}^p$中的单位球面上的任意给定投影，我们使用鞅变换投影残差标记经验过程构造检验统计量。在温和条件下，我们建立了鞅变换过程及其相应投影检验统计量的极限零分布，即使维度 p 随样本量 n呈指数增长也成立。在备择假设下，对于单位球面上几乎所有投影，该投影检验均具有一致性，渐近功效为 1。通过将高维协变量 X 投影到一维空间，该投影检验能够检测以参数速率 $n^{-1/2}$ 偏离零假设的局部备择，同时显著缓解维数灾难。此外，所提出的鞅变换仅涉及一元积分，因此在实践中易于计算，即使维度 p远大于样本量 n也是如此。从理论上说，由于鞅变换消除了参数估计所产生的平移项，该投影检验不仅具有渐近分布无关性，而且在推导零假设和备择假设下的渐近性质时，只要求参数估计量达到标准收敛速率，无需满足渐近线性或正态性。

其次，我们采用标准的组合方法，将来自不同投影的投影检验统计量组合起来，构成最终检验。请注意，尽管对于单位球面上几乎所有投影，投影检验都具有一致性，但对于某些选择不当的投影，它们仍可能损失检验功效。因此，为避免特定投影可能造成的功效损失，我们采用柯西组合方法（Liu 和 Xie，2020），汇总不同投影下投影检验所对应的 p-值，以提高检验功效。值得一提的是，对于低频备择模型，基于经验过程的检验通常比局部平滑检验更敏感；而对于高频或振荡模型，局部平滑检验通常更有力（Fan 和 Li，2000；Horowitz 和 Spokoiny，2001）。在实践中，由于回归模型的具体形式通常未知，因此理想的检验程序应能同时有效应对高频和低频备择模型。为此，我们进一步提出一种新颖的混合检验，通过柯西组合方法，将组合后的基于经验过程的检验与 Tan 等人（2025）提出的局部平滑检验进行汇总。由于柯西合并主要受最小的 p-值影响，混合检验继承了基于经验过程的检验和局部平滑检验的优点。模拟研究也表明，该混合检验对于高频和低频备择假设都具有较强的检验功效。

本文其余部分安排如下。第2节构建投影残差标记经验过程，并建立其极限零分布。第3节将经典鞅变换推广至超高维情形，并基于鞅变换投影残差标记经验过程构造投影检验。本节还建立了鞅变换及相应检验统计量的极限零分布。第4节对鞅变换和投影检验统计量进行功效分析。第5节构造组合投影检验，并讨论实际应用中投影的选择。第6节通过模拟研究和两项真实数据分析评估所提检验的有限样本性能。第7节给出总结性评论及未来研究课题。所有理论结果的证明均放在补充材料中。

2 投影残差标记经验过程

本文关注协变量维度 p可能远超样本量 n时广义线性模型（GLM）的拟合优度检验。我们的方法还可扩展至检验更一般模型的适用性，例如准广义线性模型、参数多指标模型等。回顾一下，在广义线性模型设定下，对于某个未知参数 $\beta_0 = (\beta_0^{(1)}, \ldots\ldots, \beta_0^{(p)})^\top \in \mathbb{R}^p$ 以及某个逆链接函数 $\mu(\cdot)$，有 $E(Y|X = x) = \mu(\beta_0^\top X)$和 $var(Y|X = x) = V(\mu(\beta_0^\top x))$。为说明我们的方法，我们将讨论范围限定于检验条件均值函数 $m(x) = E(Y|X = x)$的设定错误。因此，原假设和备择假设为

$$
\begin{aligned}
H_0 &: \mathbb{P}\{m(X) = \mu(\beta_0^\top X)\} = 1, \quad \text{for some } \beta_0 \in \Theta, \\
H_1 &: \mathbb{P}\{m(X) \neq \mu(\beta^\top X)\} > 0, \quad \text{for any } \beta \in \Theta,
\end{aligned}
$$

其中， Θ 是 $\mathbb{R}^p$中的紧集。在 $p \geq n$的高维情形下，正如 Jankov´a 等人（2020）所示，若设计矩阵 $\mathbf{X} = (X_1, \ldots, X_n)^\top$ 满秩，则线性方程组 $m(X_i) = \mu(\beta_0^\top X_i)$（其中 $i = 1, \ldots, n$）总存在一个解 $\beta_0 \in \mathbb{R}^p$。这意味着，在 $p \geq n$时，若不对模型结构作任何假设，广义线性模型在实践中永远不会出现错设。高维情形中常用的一个假设是回归模型的稀疏性。显然，当 $p \geq n$时，对稀疏广义线性模型进行模型检验是合理的。因此，本文在原假设和备择假设下均考虑稀疏回归模型。

下面介绍一些后文将用到的记号。对于向量 $\beta \in \mathbb{R}^p$，令 $\beta^{(j)}$ 表示 β 的第 j 个元素，并令 $\|\beta\|_q = (\sum_{j=1}^p |\beta^{(j)}|^q)^{1/q}$ （其中 $q \in \mathbb{Z}^+$ 且 $\|\beta\|_0$ ）表示β中非零元素的个数。令 $I \subset \{1, 2, \ldots, p\}$，并令 β_I表示仅包含 β中索引属于 I的元素的向量。对于矩阵 $M \in \mathbb{R}^{n\times p}$，令 M_I 表示仅由 M 中索引属于 I的列组成的矩阵，并令 M_{I^c}表示 M中索引属于 I的补集的列。令 $S \subset \{1, \ldots, p\}$表示活跃集，其中包含与响应变量 Y真正相关的协变量 $X = (X^{(1)}, \ldots, X^{(p)})^\top$ 的索引。在原假设 H_0 下，真实回归参数 β_0 是稀疏的，活跃集为$S = \{j : \beta_0^{(j)} \neq 0\}$。

我们用于检验超高维回归模型拟合优度的方法依赖于以下结果。

Pro命题 2.1. (i) 设 $W \in \mathbb{R}$ 和 $X \in \mathbb{R}^p$ 为随机变量。由此可得

$$
\begin{aligned}
E[W|X] = 0 \ a.s. &\iff E[W|\alpha^\top X] = 0 \ a.s. \ \text{ for all } \alpha \in \mathcal{S}^{p-1}, \\
E[W|\alpha^\top X] = 0 \ a.s. &\iff E[WI(\alpha^\top X \leq t)] \equiv 0 \ \text{ for all } t \in \mathbb{R},
\end{aligned}
$$

其中 $\mathcal{S}^{p-1} = \{\alpha \in \mathbb{R}^p : \|\alpha\|_2 = 1\}$。

(ii) 假设 $E|W|^2 < \infty$、 $E\|X\|_2^k < \infty$ 和 $\sum_{k=1}^\infty (E\|X\|_2^k)^{-1/k} = \infty$。若记 $\mathcal{A} = \{\alpha \in \mathbb{R}^p : E[W|\alpha^\top X] = 0 \ a.s.\}$，则

$$
\mathbb{P}\{E[W|X] = 0\} = 1 \quad \Longleftrightarrow \quad \mathcal{A} \text{ has positive Lebesgue measure.}
$$

此外，如果记 $\mathcal{A}_1 = \{\alpha \in \mathcal{S}^{p-1} : E[W|\alpha^\top X] = 0 \ a.s.\}$，那么

$$
\mathbb{P}\{E[W|X] = 0\} = 1 \quad \Longleftrightarrow \quad \mathcal{L}(\mathcal{A}_1) = 1,
$$

$$\mathbb{P}\{E[W|X] \neq 0\} > 0 \quad \Longleftrightarrow \quad \mathcal{L}(\mathcal{A}_1) = 0,$$

其中， $\mathcal{L}$ 表示单位球面上的均匀概率测度 $\mathcal{S}^{p-1}$。

命题 2.1 (i) 已由 Zhu and Li (1998) 的引理 2.1、Escanciano (2006a) 的引理 1，或 Lavergne and Patilea (2008) 的引理 2.1 证明。Patilea 等人 (2016) 和 Cuesta-Albertos 等 (2019) 在函数型数据的背景下，得出了与命题 2.1(ii) 第一部分类似的结果。命题 2.1(ii) 中的条件 $\sum_{k=1}^{\infty}(E\|X\|_2^k)^{-1/k} = \infty$称为 Carleman 条件；若随机向量 X 在零点邻域内具有有限的矩母函数，则该条件可满足；更多详情见 Cuesta-Albertos 等 (2007)。命题 2.1 的详细证明见补充材料。

我们记 $\varepsilon(\beta) = Y - \mu(\beta^\top X)$。由于原假设 H_0 等价于对于某个 $\beta_0 \in \Theta$，有 $\mathbb{P}\{E[\varepsilon(\beta_0)|X] = 0\} = 1$，根据命题 2.1可知，当且仅当对于某个 $\beta_0 \in \Theta$，有 $\mathcal{L}\{\alpha \in \mathcal{S}^{p-1} : E[\varepsilon(\beta_0)|\alpha^\top X] = 0 \ a.s.\} = 1$ 时，H_0 成立。因此，为了检验原假设 H_0，我们可以先选择一个投影 $\alpha \in \mathcal{S}^{p-1}$，然后检验投影原假设

$$H_0^\alpha : \mathbb{P}\{E[\varepsilon(\beta_0)|\alpha^\top X] = 0\} = 1, \quad \text{for some } \beta_0 \in \Theta.$$

这一检验方法背后的原理如下。在原假设 H_0下，投影原假设H_0^α 也成立。在备择假设 H_1下，对于所有 $\beta \in \Theta$，都有 $\mathbb{P}\{E[\varepsilon(\beta)|X] \neq 0\} > 0$，进而$\mathcal{L}\{\alpha \in \mathcal{S}^{p-1} : E[\varepsilon(\beta)|\alpha^\top X] = 0 \ a.s.\} = 0$。这意味着，在备择假设 H_1下，对于 $\mathcal{S}^{p-1}$上的 $\mathcal{L}$-几乎处处的投影，投影原假设 H_0^α均不成立。因此，原假设 H_0 与对于某个 $\beta_0 \in \Theta$，有 H_0^α: $\mathbb{P}\{E[\varepsilon(\beta_0)|\alpha^\top X] = 0\} = 1$ 在 $\mathcal{L}$-几乎处处意义下等价。随后，我们根据投影原假设 H_0^α构造检验统计量。

对于任意给定的 投影 $\alpha \in \mathcal{S}^{p-1}$，由命题 2.1 (i) 可知， H_0^α是等价的 to

$$E[\varepsilon(\beta_0)I(\alpha^\top X \leq t)] \equiv 0 \ \ \forall t \in \mathbb{R}, \quad \text{for some } \beta_0 \in \Theta. \tag{2.1}$$

ˆ设 $\{(X_i, Y_i)\}_{i=1}^n$是一个与（X, Y）具有相同分布的独立同分布样本，并设 β是在广义线性模型设定下 β_0 的估计量，例如惩罚估计量或其变体。受（2.1）的启发，我们提出如下投影残差标记经验过程:

$$\hat{R}_n^\alpha(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i(\hat{\beta}) I(\alpha^\top X_i \leq t), \tag{2.2}$$

其中， $\varepsilon_i(\hat{\beta}) = Y_i - \mu(\hat{\beta}^\top X_i)$。

ˆ投影残差标记经验过程 $R_n^\alpha(t)$ 取决于所选的投影 α。然而，选择适当的投影并非易事。不恰当地选择投影可能会导致备择假设下检验效力显著下降，尤其是在超高维情形下。第5节将详细讨论投影的选择，以确保我们提出的检验能够实现良好的检验效力。

ˆ2.1 $R_n^\alpha(t)$ 的极限零分布

ˆ为了推导超高维情形下 H_0 中 $R_n^\alpha(t)$ 的渐近性质，我们引入一些概念和正则条件。若随机变量 $W \in \mathbb{R}$ 属于

$\tau > 0$ 阶次 Weibull，若

$$\|W\|_{\psi_\tau} := \inf\{\eta > 0 : E\psi_\tau(|W|/\eta) \le 1\} < \infty,$$

其中 $\psi_\tau(x) = \exp(x^\tau) - 1$ 对 $x \ge 0$成立。由于 $\|W\|_{\psi_\tau} = \inf\{\eta > 0\colon E\exp(|W|^\tau/\eta^\tau) \le 2\}$，根据马尔可夫不等式可知，若 W是阶数为 τ的次 Weibull 随机变量，则对所有 $t \ge 0$，都有 $\mathbb{P}(|W| \ge t) \le 2\exp(-t^\tau/\|W\|^\tau_{\psi_\tau}$。容易看出，次高斯和次指数随机变量分别是阶数为 $\tau = 2$ 和 $\tau = 1$的次威布尔分布的特例。注意，次 Weibull 随机变量的均值不必为零。关于次威布尔分布的更多详细结果，参见 Vladimirova 等（2020）和 Kuchibhotla 和 Chakrabortty（2022）。我们进一步定义 $\hat{} S = \{j : \hat{} \beta^{(j)} \ne 0\}$, $\hat{} s = \lceil \hat{} S|$，以及 $s = |S|$。下文中的记号 C 表示一个与 n无关的常数，该常数在不同位置可能取值不同。

(A1) Under H_0, the estimator $\hat{\beta}$ satisfies $\|\hat{\beta} - \beta_0\|_2 = O_p(\sqrt{\frac{s \log p}{n}})$ and $\|\hat{\beta} - \beta_0\|_1 = O_p(\sqrt{\frac{s^2 \log p}{n}})$.

(A2) 随机变量 $\varepsilon(\beta_0)$、 $\mu'(\beta_0^\top X)$ 和 $\mu''(\beta_0^\top X)$ 均为阶数为 $\tau = 2$ 的次 Weibull 随机变量，且 $\max\{\|\varepsilon(\beta_0)\|_{\psi_2}, \|\mu'(\beta_0^\top X)\|_{\psi_2}, \|\mu''(\beta_0^\top X)\|_{\psi_2}\} \le C < \infty$。协变量 $X \in \mathbb{R}^p$ 已中心化，且 $\alpha^\top X$ 为阶数为 $\tau = 2$ 的次 Weibull 随机变量，其中 $\|\alpha^\top X\|_{\psi_2} \le C < \infty$ 对所有 $\alpha \in \mathcal{S}^{p-1}$ 均成立。

(A3) 连接函数 $\mu(\cdot)$ 存在三阶导数，且 $\max\{|\mu''(\beta^\top x)|, |\mu'''(\beta^\top x)|\} \le F(x)$ 对所有 $\beta \in \Theta \subset \mathbb{R}^p$ 均成立，其中 $F(X)$ 为阶数为 $\tau \ge 1/3$ 的次 Weibull 随机变量，且 $\|F(X)\|_{\psi_\tau} \le C < \infty$。

许多常用的惩罚估计量在温和条件下都满足条件（A1），例如广义线性模型 Lasso 估计量、广义线性模型 SCAD 估计量及其变体，前提是底层广义线性模型设定正确（即原假设成立）；例如，参见 B¨uhlmann 和 Van De Geer（2011）。在高维数据分析文献中，通常会采用（A2）中的阶数为 2（次高斯）的次 Weibull 假设（Wainwright，2019）。当维度 p 可能超过样本量 n 时，该假设用于界定分解中余项过程 $\hat{} R_n^\alpha(t)$ 的尾概率。许多实际应用中的广义线性模型都满足条件（A3），例如高斯线性模型、逻辑回归模型和 Probit 回归模型。该条件也用于控制过程 $R_n^\alpha(t)$ 余项的收敛 $\hat{}$ 速率。 $F(X)$ 的次 Weibull 阶数 $\tau \ge 1/3$ 是一个技术条件；如果我们对下文定理 2.1 中 p 的发散速率施加更严格的条件，则可以放宽这一条件。

定理 2.1。设条件（*A1*）–（*A3*）成立。若 $s^2 \log^3(p \vee n) = o(n)$，则在 H_0 下，在 $t \in \mathbb{R}$ 中一致地，

$$\begin{aligned} \hat{R}_n^\alpha(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i(\beta_0) I(\alpha^\top X_i \le t) - \sqrt{n}(\hat{\beta} - \beta_0)^\top M(t) + o_p(1) \\ &=: R_{n1}^\alpha(t) + o_p(1), \end{aligned} \tag{2.3}$$

其中 $p \vee n =$最大值$\{p, n\}$， $\varepsilon_i(\beta_0) = Y_i - \mu(\beta_0^\top X_i)$，以及 $M(t) = E[X\mu'(\beta_0^\top X) I(\alpha^\top X \le t)]$.

$\hat{}$ 值得注意的是，对于某些惩罚估计量 β，例如 Lasso、post-Lasso 及其变体，渐近线性或正态性可能并不成立，这使得难以在高维情形下为这些估计量建立 $R_n^\alpha(t)$ 的极限零分布。因此，无法直接将传统的 Cram´er-von Mises 泛函或 $R_n^\alpha(t)$ 的 Kolmogorov-Smirnov 泛函用于构造回归模型的拟合优度检验。为克服这一困难，我们拓展了

经典的鞅变换方法，使其适用于温和条件下的超高维情形。由于鞅变换会使（2.3）中的平移项 $M(t)$ 消失，即使缺少 $\beta_0 - \beta_0$ 的渐近线性展开，我们仍能推导出所得鞅变换过程的极限零分布。还值得注意的是，Cuesta-Albertos 等（2019）使用投影经验过程检验函数型线性模型的拟合优度。然而，他们的检验统计量基于未经鞅变换的投影残差标记经验过程，因此仍需参数估计量的渐近线性展开来推导其极限分布。

3 超高维情形下的鞅变换

本节构造投影残差标记经验过程 $R_n^\alpha(t)$ 的鞅变换，该变换适用于协变量维度 p 显著大于样本量 n 的情形。由定理 2.1 可知，在 H_0 下，

$$\hat{R}_n^\alpha(t) = R_{n0}^\alpha(t) - \sqrt{n}(\hat{\beta} - \beta_0)^\top M(t) + o_p(1),$$

其中 $R_{n0}^\alpha(t) = n^{-1/2}\sum_{i=1}^n \varepsilon_i(\beta_0) I(\alpha^\top X_i \le t)$，且 $M(t) = E[X\mu'(\beta_0^\top X) I(\alpha^\top X \le t)]$。直接计算可得，在 H_0 下，

$$cov(R_{n0}^\alpha(s), R_{n0}^\alpha(t)) = \psi_n^\alpha(s \wedge t), \tag{3.1}$$

其中 $\psi_n^\alpha(t) = E[\varepsilon^2(\beta_0) I(\alpha^\top X \le t)] = \int_{-\infty}^t \sigma_\alpha^2(u) dF_\alpha(u)$，且 $\sigma_\alpha^2(u) = E[\varepsilon^2(\beta_0)|\alpha^\top X = u]$。注意，对于任意给定的投影 α，函数类 $\{I(\alpha^\top x \le t) : t \in \mathbb{R}\}$ 是一个 VC 类，其 VC 指标为 2。容易看出，经验过程 $R_{n0}^\alpha(t)$ 渐近紧，且 $R_{n0}^\alpha(t)$ 的有限维分布收敛性可通过标准论证证明。由此得到

$$R_{n0}^\alpha(t) \longrightarrow B(\psi(t)), \quad \text{in distribution}$$

在 Skorohod 空间 $D[-\infty, \infty]$ 中，其中 $\psi(t)$ 是 $\psi_n^\alpha(t)$ 的逐点极限，而 $B(t)$ 是标准布朗运动。

ˆ 鞅变换的目的是消除 $R_n^\alpha(t)$ 的平移项中的 $M(t)$，同时将 $R_{n0}^\alpha(t)$ 变换为一个创新过程，使其具有与 $R_{n0}^\alpha(t)$ 相同的极限零分布。鞅变换的原理如下。令 $A^\alpha(t) = {}^{\partial M_{(t)}}{}_{\partial \psi_n^\alpha(t)}$ 为

$$M\ t \qquad \psi_n^\alpha\ t \qquad M\ t\ = E[X\mu'\ \beta_0^\top X\ I\ \alpha^\top X \leq ( )\text{关于}\ (\ )$$
$$t\ ] = \int_{-\infty}^t r_\alpha\ u\ F_\alpha\ du\ , \qquad r_\alpha\ u\ = E[X\mu'\ \beta_0^\top X\ |\alpha^\top X = u] \qquad A^\alpha\ t\ = r_\alpha\ t\ /\sigma_\alpha^2\ t$$

的 Radon-Nikodym导数。回顾 ( ) ( ) ( ) ( ) ( )，其中 ( ) ( ) 。由此可得 ( ) ( ) ( )。然后我们定义一个矩阵为

$$\Gamma^\alpha(t) = \int_t^\infty A^\alpha(u) A^\alpha(u)^\top d\psi_n^\alpha(u) = \int_t^\infty \frac{r_\alpha(u) r_\alpha(u)^\top}{\sigma_\alpha^4(u)} d\psi_n^\alpha(u).$$

假设 对任意 $t \in \mathbb{R}$，g $\Gamma^\alpha(t)$ 均非奇异，鞅变换定义为 a s

$$Tf(t) = f(t) - \int_{-\infty}^t A^\alpha(u)^\top \Gamma^\alpha(u)^{-1} \int_u^\infty A^\alpha(v) df(v) d\psi_n^\alpha(u), \tag{3.2}$$

其中， $f(t)$ 或为有界变差函数，或为随机过程，使得 (3.2) 中的积分有定义。有时它也可以是函数向量。注意到 $\Gamma^{\alpha}(t) = \int_t^{\infty} A^{\alpha}(u)dM(u)^{\top}$and $M(t) = \int_{-\infty}^{t} A(u)d\psi_n^{\alpha}(u)$，不难看出

$$TM(t)^{\top} = M(t)^{\top} - \int_{-\infty}^{t} A^{\alpha}(u)^{\top}\Gamma^{\alpha}(u)^{-1}\int_{u}^{\infty} A^{\alpha}(v)dM(v)^{\top}d\psi_n^{\alpha}(u) \equiv 0.$$

这意味着，鞅变换 T 消除了 $R_n(t)$ 的分解中ˆ的平移项 $M(t)$。另请注意， T 是一个线性算子，因此可得 $TR_{n1}^{\alpha}(t) = TR_{n0}^{\alpha}(t)$。

我们进一步研究原假设下 $TR_{n0}^{\alpha}(t)$ 的渐近性质。回顾 $R_{n0}^{\alpha}(t) = n^{-1/2}\sum_{i=1}^{n}\varepsilon_i(\beta_0)I(\alpha^{\top}X_i \le t)$, 可得

$$TR_{n0}^{\alpha}(t) = R_{n0}^{\alpha}(t) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i(\beta_0)\int_{-\infty}^{t} A^{\alpha}(u)^{\top}\Gamma^{\alpha}(u)^{-1}A^{\alpha}(\alpha^{\top}X_i)I(\alpha^{\top}X_i \ge u)d\psi_n^{\alpha}(u).$$

这意味着 $TR_{n0}^{\alpha}(t)$ 是一个中心化累积和过程，其协方差函数为

$$cov(TR_{n0}^{\alpha}(s), TR_{n0}^{\alpha}(t)) = cov(R_{n0}^{\alpha}(s), R_{n0}^{\alpha}(t)) = \psi_n^{\alpha}(s \wedge t). \tag{3.3}$$

因此，变换后的过程 $TR_{n0}^{\alpha}(t)$ 在 H_0 下与 $R_{n0}^{\alpha}(t)$ 具有相同的协方差结构。断言（3.3）已在补充材料中得到证明。类似于证明 Stute 等（1998）定理 1.2 的论证，我们可以推导出 $TR_{n0}^{\alpha}(t)$ 的渐近紧性和有限维收敛。因此,

$$TR_{n1}^{\alpha}(t) = TR_{n0}^{\alpha}(t) \longrightarrow B(\psi(t)), \quad \text{in distribution}$$

ˆ 在 Skorohod 空间 $D[-\infty, \infty$ 中。此外，在温和条件下，我们可以证明 $TR_n^{\alpha}(t) - \hat{}TR_{n1}^{\alpha}(t) = o_p(1)$ 关于 t 一致地成立。综上可得，鞅变换过程 $TR_n^{\alpha}(t)$ 经时间变换 $z = \psi(t)$ 后，依分布收敛于标准布朗运动$B(t)$。

在实际应用中，鞅变换 T 需要用其经验类比进行估计。为此，回顾

$$A^{\alpha}(t) = \frac{r_{\alpha}(t)}{\sigma_{\alpha}^2(t)} \quad \text{and} \quad \Gamma^{\alpha}(t) = E\left(\frac{r_{\alpha}(\alpha^{\top}X)r_{\alpha}(\alpha^{\top}X)^{\top}}{\sigma_{\alpha}^2(\alpha^{\top}X)}I(\alpha^{\top}X \ge t)\right),$$

其中 $r_{\alpha}(t) = E[X\mu'(\beta_0^{\top}X)|\alpha^{\top}X = t]$ 和 $\sigma_{\alpha}^2(t) = E[\varepsilon^2(\beta_0)|\alpha^{\top}X = t]$。由于除光滑性外，我们对量 $r_{\alpha}(t)$ 和 $\sigma_{\alpha}^2(t)$ 不作任何假设，因此应采用非参数方式对其进行估计。在低维情形下，Nadaraya–Watson 估计量等标准非参数估计量或许适用于此目的。然而，我们注意到 $r_{\alpha}(t) = E[X\mu'(\beta_0^{\top}X)|\alpha^{\top}X = t] \in \mathbb{R}^p$ 是一个 p 维函数。当维数 p 显著超过样本量 n时，对 $r_{\alpha}(t)$ 进行非参数估计，会给推导相应估计鞅变换的渐近性质带来巨大的理论困难。为解决这一难题，受 Stute 和 Zhu（2002）的启发，我们在温和条件下提出了一种新颖的鞅变换变体。

(A4) 协变量 X 满足线性条件均值假设，即对于一个非随机矩阵 $D \in \mathbb{R}^{p\times 2}$，有 $E[X|B^\top X] = DB^\top X$，其中 $B = (\alpha,\ \beta_0) \in \mathbb{R}^{p\times 2}$。此外，$D$ 的 $(i,\ j)$ 元素 D_{ij} 满足 $|D_{ij}| \le C < \infty$，对所有 $1 \le i \le p$ 和 $j \in \{1,2\}$ 均成立。

若协变量 X 的分布椭圆对称（例如服从正态分布），则线性条件均值假设 (A4) 成立。在高维情形下，正如 Hall 和 Li（1993）所证明的，当原始协变量维度 p 较大时，$E[X|B^\top X]$ 可近似表示为 X 的线性函数。

在条件 (A4) 下，我们有

$$M(t) = E\{E[X|B^\top X]\mu'(\beta_0^\top X)I(\alpha^\top X \le t)\} = DE[B^\top X\mu'(\beta_0^\top X)I(\alpha^\top X \le t)].$$

记 $M_0(t) = E[B^\top X\mu'(\beta_0^\top X)I(\alpha^\top X \le t)]$,，于是可得

$$M(t) = DM_0(t) = D\left(\int_{-\infty}^{t} ug_{1\alpha}(u)dF_\alpha(u), \int_{-\infty}^{t} g_{2\alpha}(u)dF_\alpha(u)\right)^\top,$$

其中 $g_{1\alpha}(u) = E[\mu'(\beta_0^\top X)|\alpha^\top X = u]$，且 $g_{2\alpha}(u) = E[\beta_0^\top X\mu'(\beta_0^\top X)|\alpha^\top X = u]$。因此，(2.3) 中的残差标记经验过程 $R_n^\alpha(t)$ 可改写为

$$\hat{R}_n^\alpha(t) = R_{n0}^\alpha(t) - \sqrt{n}(\hat{\beta} - \beta_0)^\top DM_0(t) + o_p(1).$$

为消除平移项 $\sqrt{n}(\hat{\beta} - \beta_0)^\top DM_0(t)$，我们分别将新的 $A^\alpha(t)$ 和 $\Gamma^\alpha(t)$ 定义为

$$A^\alpha(t) = \frac{\partial M_0(t)}{\partial \psi_n^\alpha(t)} = \left(\frac{tg_{1\alpha}(t)}{\sigma_\alpha^2(t)}, \frac{g_{2\alpha}(t)}{\sigma_\alpha^2(t)}\right)^\top$$

and

$$\Gamma^\alpha(t) = \int_t^\infty A^\alpha(u)A^\alpha(u)^\top d\psi_n^\alpha(u) = \int_t^\infty \frac{1}{\sigma_\alpha^4(u)}\begin{pmatrix} u^2g_{1\alpha}(u)^2, & ug_{1\alpha}(u)g_{2\alpha}(u) \\ ug_{1\alpha}(u)g_{2\alpha}(u), & g_{2\alpha}(u)^2 \end{pmatrix} d\psi_n^\alpha(u).$$

需要注意的是，这些新的 $A^\alpha(t)$ 和 $\Gamma^\alpha(t)$ 仅涉及单变量函数 $g_{1\alpha}(t)$、$g_{2\alpha}(t)$ 和 $\sigma_\alpha^2(t)$。随后，我们使用一维 Nadaraya–Watson 估计量估计这些量：

$$\begin{aligned}
\hat{\sigma}_\alpha^2(t) &= \frac{\sum_{i=1}^n \varepsilon_i^2(\hat{\beta})K_h(t - \alpha^\top X_i)}{\sum_{i=1}^n K_h(t - \alpha^\top X_i)}, \\
\hat{g}_{1\alpha}(t) &= \frac{\sum_{i=1}^n \mu'(\hat{\beta}^\top X_i)K_h(t - \alpha^\top X_i)}{\sum_{i=1}^n K_h(t - \alpha^\top X_i)}, \\
\hat{g}_{2\alpha}(t) &= \frac{\sum_{i=1}^n \hat{\beta}^\top X_i\mu'(\hat{\beta}^\top X_i)K_h(t - \alpha^\top X_i)}{\sum_{i=1}^n K_h(t - \alpha^\top X_i)},
\end{aligned}$$

其中 $\varepsilon_i(\hat{\beta}) = Y_i - \mu(\hat{\beta}^\top X_i)$，且 $K_h(\cdot) = h^{-1}K(\cdot/h)$，使用单变量核函数 $K(\cdot)$ 和带宽 $\hat{}\ h$。$A^\alpha(t)$ 和 $\Gamma^\alpha(t)$ 的估计量 $A_n^\alpha(t)$ 和 $\Gamma_n^\alpha(t)$ 分别为

$$\hat{A}_n^\alpha(t) = \left(\frac{t\hat{g}_{1\alpha}(t)}{\hat{\sigma}_\alpha^2(t)}, \frac{\hat{g}_{2\alpha}(t)}{\hat{\sigma}_\alpha^2(t)}\right)^\top \quad \text{and} \quad \hat{\Gamma}_n^\alpha(t) = \int_t^\infty \hat{A}_n^\alpha(u)\hat{A}_n^\alpha(u)^\top d\hat{\psi}_n^\alpha(u), \tag{3.4}$$

其中 $\hat{\psi}_n^{\alpha}(u) = n^{-1}\sum_{i=1}^{n}\varepsilon_i^2(\hat{\beta})I(\alpha^{\top}X_i \le u)$。因此，我们得到鞅变换 T 的经验类比量 $\hat{T}_n$：

$$\hat{T}_n\hat{R}_n^{\alpha}(t) = \hat{R}_n^{\alpha}(t) - \int_{-\infty}^{t}\hat{A}_n^{\alpha}(u)^{\top}\hat{\Gamma}_n^{\alpha}(u)^{-1}\int_{u}^{\infty}\hat{A}_n^{\alpha}(v)d\hat{R}_n^{\alpha}(v)d\hat{\psi}_n^{\alpha}(u), \tag{3.5}$$

其中，$\hat{R}_n^{\alpha}(t) = n^{-1/2}\sum_{i=1}^{n}\varepsilon_i(\hat{\beta})I(\alpha^{\top}X_i \le t)$由（2.2）给出。

3.1 $T_nR_n^{\alpha}(t)$的极限零分布

为了推导超高维情形下鞅变换过程 $T_nR_n^{\alpha}(t)$的渐近性质，我们还需施加一些正则性条件。

（A5）核函数 $K(\cdot)$满足：（i） $K(\cdot)$在 $\mathbb{R}$ 上连续，且在其支撑集 $[-1,1]$上具有连续导数；（ii） $K(x) = K(-x)$和 $K(\cdot)$具有有界变差；（iii） $\int_{-1}^{1}K(u)du = 1$ 且$\int_{-1}^{1}u^iK(u)du = 0$，对 $i = 1,\ldots,k-1$成立。

(A6) 我们记 $\sigma_{1\alpha}^2(t) = \sigma_{\alpha}^2(t)f_{\alpha}(t)$、 $w_{1\alpha}(t) = g_{1\alpha}(t)f_{\alpha}(t)$ 和 $w_{2\alpha}(t) = g_{2\alpha}(t)f_{\alpha}(t)$。函数 $\sigma_{1\alpha}^2(t)$、 $w_{1\alpha}(t)$ 和 $w_{2\alpha}(t)$ 在 t 中具有最高达 $k-1$ 阶的导数。令 $w_{1\alpha}^{(i)}(t) = \frac{d^iw_{1\alpha}(t)}{dt^i}$，$w_{2\alpha}^{(i)}(t) = \frac{d^iw_{2\alpha}(t)}{dt^i}$，且 $(\sigma_{1\alpha}^2)^{(i)}(t) = \frac{d^i\sigma_{1\alpha}^2(t)}{dt^i}$，其中 $i = 1,\ldots,k-1$。函数 $w_{1\alpha}^{(k-1)}(t)$、 $w_{2\alpha}^{(k-1)}(t)$ 和 $(\sigma_{1\alpha}^2)^{(k-1)}(t)$ 满足 Lipschitz 条件：

$$\begin{aligned}
|w_{1\alpha}^{(k-1)}(t+u) - w_{1\alpha}^{(k-1)}(t)| &\le L|u|, \quad \forall\, u \in U,\\
|w_{2\alpha}^{(k-1)}(t+u) - w_{2\alpha}^{(k-1)}(t)| &\le L|u|, \quad \forall\, u \in U,\\
|(\sigma_{1\alpha}^2)^{(k-1)}(t+u) - (\sigma_{1\alpha}^2)^{(k-1)}(t)| &\le L|u|, \quad \forall\, u \in U,
\end{aligned}$$

对于某个零点邻域 U。此外，我们假设对于所有 $\alpha \in \mathcal{S}^{p-1}$，$\sup_{t\in\mathbb{R}}|f_{\alpha}(t)| \le C < \infty$ 且 $\inf_{t\in\mathbb{R}}\sigma_{1\alpha}^2(t) \ge C > 0$。

(A7) 带宽 h 满足 $\sqrt{nh^{2k}} = o(1)$ 和 $\log^4 n = o(nh^4)$，当 $n \to \infty$.

(A8) 矩阵 $\Gamma^{\alpha}(t)$ 满足 $\inf_{t\le t_0}|\det(\Gamma^{\alpha}(t))| > 0$，对任意 $t_0 \in \mathbb{R}$ 和 $\alpha \in \mathcal{S}^{p-1}$，其中 $\det(\Gamma^{\alpha}(t))$ 表示矩阵 $\Gamma^{\alpha}(t)$ 的行列式。

条件 (A5)-(A7) 通常用于高阶非参数估计文献中；例如，参见 Rao (1983) 的第 2 章和第 4 章，以及 Zhu 和 Fang (1996)。条件 (A8) 是保证 $\|\Gamma^{\alpha}(t)^{-1}\|_2$ 一致有界的必要条件。

下一个结果确立了条件 H_0 下鞅变换过程 $T_nR_n^{\alpha}(t)$ 的渐近性质。其证明见补充材料。

定理 3.1. 假设条件 *(A1)-(A8)* 成立，且矩阵 $\Gamma^{\alpha}(t)$ 对所有$t \in \mathbb{R}$均为非奇异矩阵。若 $s^4\log^4(p \vee n) = o(n)$ 且 $\log^5 p = o(n)$ 当 $n \to \infty$ 时，则在 H_0 下，

$$\hat{T}_n\hat{R}_n^{\alpha}(t) \longrightarrow B(\psi(t)), \quad in\ distribution$$

在 *Skorohod* 空间 $D[-\infty,\infty)$。

值得一提的是，当原假设为高斯线性模型时，鞅变换可以具有更理想的结构。我们还要求协变量满足线性条件均值假设 X。

(A4′) 协变量 X 满足线性条件均值假设，即 $E[X|\alpha^\top X] = \alpha^\top X D_0$对于某个非随机向量 $D_0 \in \mathbb{R}^p$，有 $|D_0^{(j)}| \le C < \infty$，其中 $1 \le j \le p$。

回顾在高斯线性模型下 $\mu'(t) \equiv 1$，由 (A4′) 和 (2.3) 可得

$$M(t) = E[XI(\alpha^\top X \le t)] = E\{E[X|\alpha^\top X]I(\alpha^\top X \le t)\} = E[\alpha^\top X I(\alpha^\top X \le t)]D_0.$$

记 $M_0(t) = E[\alpha^\top X I(\alpha^\top X \le t)] = \int_{-\infty}^t u dF_\alpha(u)$。由于在高斯线性模型下，$\sigma_\alpha^2(t) \equiv \sigma^2$ 为常数，因此量 $A^\alpha(t)$ 和 $\Gamma^\alpha(t)$ 可改写为

$$A^\alpha(t) = \frac{\partial M_0(t)}{\partial \psi_n^\alpha(t)} = \frac{t}{\sigma^2}, \text{ and } \Gamma^\alpha(t) = \int_t^\infty A^\alpha(u)^2 d\psi_n^\alpha(u) = \frac{1}{\sigma^4}\int_t^\infty u^2 d\psi_n^\alpha(u),$$

其中 $\psi_n^\alpha(t) = \int_{-\infty}^t \sigma_\alpha^2(u) dF_\alpha(u) = \sigma^2 F_\alpha(t)$。因此，在这一大为简化的情形下，也可以应用鞅变换 T。请注意，在构造估计的鞅变换时，对于高斯线性模型，我们无需对 $\sigma_\alpha^2(u) = E[\varepsilon^2(\beta_0)|\alpha^\top X = u]$、$g_{1\alpha}(u) = E[\mu'(\beta_0^\top X)|\alpha^\top X = u]$ 和$g_{2\alpha}(u) = E[\beta_0^\top X \mu'(\beta_0^\top X)|\alpha^\top X = u]$ 进行非参数估计。因此，在检验超高维高斯线性模型时，所得鞅变换的结构可以简单得多。

ˆ ˆ3.2 基于 $T_n R_n^\alpha(t)$的投影检验统计量

根据命题 2.1 和定理 3.1，我们可以基于 $T_n R_n^\alpha(t)$的任意 ˆ ˆ 泛函构造针对 H_0 的检验，使所得检验渐近分布无关。具体而言，我们采用鞅变换过程 ˆ T_nˆ$R_n^\alpha(t)$的 Cram´er-von Mises 泛函来构造检验统计量。我们将（非正式的）检验统计量定义为:

$$CvM_{n,\alpha}^2 = \int_{-\infty}^{t_0} |\hat{T}_n \hat{R}_n^\alpha(t)|^2 d\hat{\psi}_n^\alpha(t),$$

其中 ˆ $\psi_n^\alpha(t) = n^{-1}\sum_{i=1}^n \varepsilon_i^2(\hat{\ }\beta) I(\alpha^\top X_i \le t)$。根据定理 3.1 和扩展连续映射定理（van der Vaart和 Wellner，1996，定理1.11.1），在 H_0下，我们有:

$$CvM_{n,\alpha}^2 \longrightarrow \int_{-\infty}^{t_0} B^2(\psi(t)) d\psi(t) \quad \text{in distribution},$$

其中，$B(t)$ 是标准布朗运动。由于 $B(t\psi(t_0))/\sqrt{\psi(t_0)} = B(t)$ 服从相同的分布，因此可得

$$\int_{-\infty}^{t_0} B^2(\psi(t)) d\psi(t) = \psi^2(t_0)\int_0^1 B^2(t) dt, \quad \text{in distribution.}$$

因此，基于给定投影 α 的最终检验统计量为

$$TCvM_{n,\alpha}^2 = \frac{1}{\hat{\psi}_n^\alpha(t_0)^2}\int_{-\infty}^{t_0} |\hat{T}_n \hat{R}_n^\alpha(t)|^2 d\hat{\psi}_n^\alpha(t), \tag{3.6}$$

其中，$\hat{\psi}_n^{\alpha}(t_0) = n^{-1}\sum_{i=1}^{n}[Y_i - \mu(\hat{\beta}^{\top}X_i)]^2 I(\alpha^{\top}X_i \le t_0)$ 是 $\psi(t_0)$ 的经验类比。再次应用定理 3.1 和扩展连续映射定理，我们很容易得到超高维情形下 $TCvM_{n,\alpha}^2$ 的极限零分布。

推论 3.1. 假设定理 *3.1* 的条件成立。那么，在 H_0 下，

$$TCvM_{n,\alpha}^2 \longrightarrow \int_0^1 B^2(t)dt \quad \text{in distribution.} \tag{3.7}$$

推论 3.1 表明，基于鞅变换过程 $\hat{T}_n\hat{R}_n^{\alpha}(t)$ 的投影检验 $TCvM_{n,\alpha}^2$ 渐近分布无关，因此其临界值可以制成表格。

在同方差情形下，例如高斯线性模型，我们有 $\sigma_{\alpha}^2(t) \equiv \sigma^2$，它与 t 相互独立。回想一下，$\sigma_{\alpha}^2(t) = E[\varepsilon^2(\beta_0)|\alpha^{\top}X = t]$，且 $\psi_n^{\alpha}(t) = E[\varepsilon^2(\beta_0)I(\alpha^{\top}X \le t)] = \sigma^2 F_{\alpha}(t)$。因此，$\psi_n^{\alpha}(t)$ 可以用 $\hat{\sigma}^2\hat{F}_{\alpha}(t)$ 估计，其中 $\hat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}[Y_i - \mu(\hat{\beta}^{\top}X_i)]^2$，而 $\hat{F}_{\alpha}(t)$ 是 $\{\alpha^{\top}X_i$ 的经验分布：$i = 1,\ldots,n\}$。因此，检验统计量 $TCvM_{n,\alpha}^2$ 变为

$$TCvM_{n,\alpha}^2 = \frac{1}{\hat{\sigma}^2\hat{F}_{\alpha}(t_0)^2}\int_{-\infty}^{t_0}|\hat{T}_n\hat{R}_n^{\alpha}(t)|^2 d\hat{F}_{\alpha}(t).$$

对于 t_0，在实际应用中，我们采用 F_{α} 的99% 分位数，正如 Stute 等（1998）以及 Stute 和 Zhu（2002）所建议的那样。

4 功效分析

在本节中，我们研究在各种备择假设下，鞅变换过程 $\hat{T}_n\hat{R}_n^{\alpha}(t)$ 和投影检验 $TCvM_{n,\alpha}^2$的渐近性质。考虑以下以速率 $\gamma_n = n^{-c}$趋近于原假设 H_0 的备择假设：

$$H_{1n} : m(x) = E(Y|X = x) = \mu(\beta_0^{\top}x) + \gamma_n L(x),$$

其中，$c \in [0, 1/2]$，且 $L(\cdot)$是满足 $\mathbb{P}\{L(X) = 0\} < 1$的非恒定函数。这里，$c = 0$ 对应于全局备择假设，而 $c > 0$ 对应于局部备择假设。我们还假设回归函数 $m(\cdot)$具有稀疏性。

为了推导在超高维情形下各种备择假设下 $T_nR_n^{\alpha}(t)$的渐近性质，我们提出一些额外条件。令 $\tilde{\sigma}_{\alpha}^2(t) = E[\varepsilon^2(\tilde{\beta}_0)|\alpha^{\top}X = t]$，$\tilde{g}_{1\alpha}(t) = E[\mu'(\tilde{\beta}_0^{\top}X)|\alpha^{\top}X = t]$，$\tilde{g}_{2\alpha}(t) = E[\tilde{\beta}_0^{\top}X\mu'(\tilde{\beta}_0^{\top}X)|\alpha^{\top}X = t]$，以及 $l_{2\alpha}(t) = E[L(X)^2|\alpha^{\top}X = t]$。

(A9) (i) 在全局备择假设 H_1 下，存在一个参数 $\tilde{\beta}_0 = (\tilde{\beta}_0^{(1)},\ldots,\tilde{\beta}_0^{(p)})^{\top} \in \Theta$，使得

$$\|\hat{\beta} - \tilde{\beta}_0\|_1 = O_p(\tilde{s}\sqrt{\frac{\log p}{n}}) \quad \text{and} \quad \|\hat{\beta}_0 - \tilde{\beta}_0\|_2 = O_p(\sqrt{\frac{\tilde{s}\log p}{n}}),$$

其中 $\tilde{s} = |\tilde{S}|$和 $\tilde{S} = \{j : \tilde{\beta}_0^{(j)} \neq 0\}$。**(ii)** 在局部备择假设 H_{1n} 下，当 $c \in (0, 1/2]$时，我们有 $\tilde{\beta}_0 - \beta_0 = r_n M^L + o_p(r_n)$，其中 $M_{\tilde{S}_1^c}^L = 0$和 $\|M^L\|_2 = O(1)$，并且

$$\|\hat{\beta}_0 - \beta_0\|_1 = O_p(\tilde{s}\sqrt{\frac{\log p}{n}} + \gamma_n\sqrt{\tilde{s}_1}) \quad \text{and} \quad \|\hat{\beta}_0 - \beta_0\|_2 = O_p(\tilde{s}\sqrt{\frac{\log p}{n}} + \gamma_n),$$

其中 $\tilde{s}_1 = |\tilde{S}_1|$、$\tilde{s} = |\tilde{S}|$ 和 $\tilde{S}_1 = S \cup \tilde{S}$ 满足 $S = \{j : \beta_0^{(j)} \neq 0\}$。

(A10) 随机变量 $\varepsilon(\tilde{\beta}_0)$、$\mu'(\tilde{\beta}_0^\top X)$ 和 $\mu''(\tilde{\beta}_0^\top X)$ 是阶数为 $\tau = 2$ 的次 Weibull 随机变量，且其最大值为 $\{\|\varepsilon(\tilde{\beta}_0)\|_{\psi_2}, \|\mu'(\tilde{\beta}_0^\top X)\|_{\psi_2}, \|\mu''(\tilde{\beta}_0^\top X)\|_{\psi_2}\} \leq C < \infty$。随机变量 $L(X)$ 是阶数为 $\tau \geq 2/3$ 的次 Weibull 随机变量，且 $\|L(X)\|_{\psi_\tau} \leq C < \infty$。

(A11) 记 $\tilde{\sigma}_{1\alpha}^2(t) = \tilde{\sigma}_\alpha^2(t) f_\alpha(t)$、$\tilde{w}_{1\alpha}(t) = \tilde{g}_{1\alpha}(t) f_\alpha(t)$ 和 $\tilde{w}_{2\alpha}(t) = \tilde{g}_{2\alpha}(t) f_\alpha(t)$。函数$\tilde{\sigma}_{1\alpha}^2(t)$、$\tilde{w}_{1\alpha}(t)$ 和 $\tilde{w}_{2\alpha}(t)$ 存在关于 t 的 $k-1$ 阶导数，并令 $(\tilde{\sigma}_{1\alpha}^2)^{(i)}(t) = {}^{d^i\tilde{\sigma}_{1\alpha}^2(t)}{}_{dt^i}$，$\tilde{w}_{1\alpha}^{(i)}(t) = {}^{d^i\tilde{w}_{1\alpha}(t)}{}_{dt^i}$，以及 $\tilde{w}_{2\alpha}^{(i)}(t) = {}^{d^i\tilde{w}_{2\alpha}(t)}{}_{dt^i}$，其中 $i = 1, 2, \ldots, k-1$。函数 $(\tilde{\sigma}_{1\alpha}^2)^{(k-1)}(t)$ $w_{1\alpha}^{(k-1)}(t)$、$w_{2\alpha}^{(k-1)}(t)$ 和 $l_{2\alpha}(t)$ 满足 Lipschitz 条件

$$\begin{aligned}
|w_{1\alpha}^{(k-1)}(t+u) - w_{1\alpha}^{(k-1)}(t)| &\leq L|u| \quad \forall\, u \in U, \\
|w_{2\alpha}^{(k-1)}(t+u) - w_{2\alpha}^{(k-1)}(t)| &\leq L|u| \quad \forall\, u \in U, \\
|(\tilde{\sigma}_{1\alpha}^2)^{(k-1)}(t+u) - (\tilde{\sigma}_{1\alpha}^2)^{(k-1)}(t)| &\leq L|u| \quad \forall\, u \in U, \\
|l_{2\alpha}(t+u) f_\alpha(t+u) - l_{2\alpha}(t) f_\alpha(t)| &\leq L|u| \quad \forall\, u \in U,
\end{aligned}$$

对于零点的某个邻域 内 U。此外，我们假设对于所有 $\alpha \in \mathcal{S}^{p-}$，都有 $\inf_{t\in\mathbb{R}} \tilde{\sigma}_{1\alpha}^2(t) \geq C > 0$ [1].

(A12) 令 $\tilde{\Gamma}^\alpha(t) = \int_t^\infty \tilde{A}^\alpha(u)\, \tilde{A}^\alpha(u)^\top d\tilde{\psi}_n^\alpha(u)$，其中 $\tilde{\psi}_n^\alpha(v) = E[\varepsilon(\tilde{\beta}_0)^2 I(\alpha^\top X \leq v)]$ 且 $\tilde{A}^\alpha(t) = \begin{pmatrix} t\tilde{g}_{1\alpha}(t) & \tilde{g}_{2\alpha}(t) \\ \tilde{\sigma}_\alpha^2 & \end{pmatrix}^\top$。矩阵 $\tilde{\Gamma}^\alpha(t)$ 满足对于任意 $t_0 \in \mathbb{R}$，$\inf_{t\leq t_0} |\det(\tilde{\Gamma}^\alpha(t))| \geq C > 0$，其中 $\det(\Gamma^\alpha(t))$ 表示矩阵 $\Gamma^\alpha(t)$ 的行列式。

我们在补充材料中证明，在局部和全局备择假设（错误设定模型）下，广义线性模型 Lasso 估计量均满足条件（A9）。还值得注意的是，B¨uhlmann 和 van de Geer（2015）证明，在错设线性模型下，若协变量 X 服从具有正定协方差矩阵的高斯分布，则 的支撑集 S β_0 满足 $S \subset S$。对于维数固定为 p 的错误设定广义线性模型，Lu等人（2012）证明，若真实的底层模型也是广义线性模型，且其链接函数设定有误，并且对 X满足线性条件均值假设，即对所有 $\beta \in \mathbb{R}^p$ 和 a，$b \in \mathbb{R}$，$S = S$。这两项结果都为条件（A9）提供了依据。条件（A10）和（A11）与第3节中的（A2）和（A3）类似，用于控制备择假设下 $T_n R_n^\alpha(t)$余项的收敛速率。条件（A12）用于确保 $\|\tilde{\Gamma}^\alpha(t)^{-1}\|_2$ 一致有界，不趋于无穷大。

下一个定理给出了 $T_n R_n^\alpha(t)$ 在各种备择假设下的渐近性质。其证明见补充材料。记 $\tilde{G}^\alpha(t) = E[\varepsilon(\tilde{\beta}_0) I(\alpha^\top X \leq t)]$,

$S_L^\alpha\ t\ = E[L\ X\ I\ \alpha^\top X \leq t\ ]\ \tilde{T}\tilde{G}^\alpha\ t\ = \tilde{G}^\alpha\ t\ - \int_{-\infty}^t \tilde{A}^\alpha\ u\ ^\top \tilde{\Gamma}^\alpha\ u\ ^{-1} \int_u^\infty \tilde{A}^\alpha\ v\ d\tilde{G}^\alpha\ v\ d\tilde{\psi}_n^\alpha\ u$ ( ) ( ) ( ) , ( ) ( ) ( ) ( ) ( ) ( ) ( ), and$TS_L^\alpha(t) = S_L^\alpha(t) - \int_{-\infty}^t A^\alpha(u)^\top \Gamma^\alpha(u)^{-1} \int_u^\infty A^\alpha(v) dS_L^\alpha(v) d\psi_n^\alpha(u)$。

定理4.1。 假设条件（*A2*）-（*A12*）成立。
(1)在 H_1 下， 若对于所有 $t \in \mathbb{R}$， $\tilde{\Gamma}^\alpha(t)$ 均非奇异，且当 $n \to \infty$ 时， $\tilde{s}^2 \log^4(p \vee n) = o(n)$，则 n

$$n^{-1/2} \hat{T}_n \hat{R}_n^\alpha(t) \longrightarrow L_1(t), \quad in\ probability,$$

其中， $L_1(t)$ 是 $TG^\alpha(t)$ 的逐点极限。
(2) 在 H_{1n} 且 $r_n = n^{-a}$ 和 $a \in (0, 1/2)$的条件下，若对所有 $t \in \mathbb{R}$， $\Gamma^\alpha(t)$ 均为非奇异矩阵，则 $\gamma_n \tilde{s}_1^2 \log^2(p \vee n) =$

$o(1)$、$\log^5 p = o(n)$和 $\sqrt{n}\gamma_n^2 = o(h)$，则

$$(nr_n^2)^{-1/2}\hat{T}_n\hat{R}_n^{\alpha}(t) \longrightarrow L_2(t), \quad in\ probability,$$

其中，$L_2(t)$ 是$TS_L^{\alpha}(t)$的逐点极限。

(3) 在 H_{1n} 且 $r_n = n^{-1/2}$的条件下，如果对所有 $t \in \mathbb{R}$，$\Gamma^{\alpha}(t)$ 均非奇异，且 $\tilde{s}_1^2\log^2(p \vee n) = o(\sqrt{n})$和$\overline{\log}^5 p = o(n)$，则

$$\hat{T}_n\hat{R}_n^{\alpha}(t) \longrightarrow B(\psi(t)) + L_2(t), \quad in\ distribution,$$

在 *Skorohod* 空间 $D[-\infty,\infty)$中，其中 $B(\psi(t))$ 由定理 *3.1* 给出。

以下关于投影检验统计量$TCvM_{n,\alpha}^2$的渐近结果是定理 4.1 和扩展连续映射定理（van der Vaart和 Wellner，1996，定理1.11.1）的推论。我们记 $\tilde{\psi}_n^{\alpha}(t) = E[\varepsilon^2(\tilde{\beta}_0)I(\alpha^\top X \le t)] = E[\tilde{\sigma}^2(\alpha^\top X)I(\alpha^\top X \le t)]$。

推论 4.1。假设条件（*A2*）-（*A12*）成立。

(1) 在 H_1 下，若对所有 $t \in \mathbb{R}$，$\tilde{\Gamma}^{\alpha}(t)$ 均非奇异，且当 $n \to \infty$ 时，$\tilde{s}^2\log^4(p \vee n) = o(n)$，则

$$\frac{1}{n}TCvM_{n,\alpha}^2 \longrightarrow \frac{1}{\tilde{\psi}(t_0)}\int_{-\infty}^{t_0}|L_1(t)|^2 d\tilde{\psi}(t) \quad in\ probability,$$

这里，$\tilde{\psi}(t)$和 $L_1(t)$分别是 $\tilde{\psi}_n^{\alpha}(t)$和 $TG^{\alpha}(t)$ 的逐点极限。

(2) 在满足 $r_n = n^{-c}$ 且 $c \in (0, 1/2)$ 的 H_{1n} 下，若对所有 $t \in \mathbb{R}$，$\Gamma^{\alpha}(t)$ 均非奇异，且 $\gamma_n\tilde{s}_1^2\log^2(p \vee n) = o(1)$、$\log^5 p = o(n)$ 和 $\sqrt{n}\gamma_n^2 = o(h)$，则

$$\frac{1}{nr_n^2}TCvM_{n,\alpha}^2 \longrightarrow \frac{1}{\psi(t_0)}\int_{-\infty}^{t_0}|L_2(t)|^2 d\psi(t) \quad in\ probability,$$

其中，$\psi(t)$ 和 $L_2(t)$ 分别是 ψ_n^{α} 和 $TS_L^{\alpha}(t)$ 的逐点极限。

(3) 在 H_{1n} 且 $r_n = n^{-1/2}$ 的条件下，若对所有 $t \in \mathbb{R}$，$\Gamma^{\alpha}(t)$ 均为非奇异矩阵，且 $\tilde{s}_1^2\log^2(p \vee n) = o(\sqrt{n})$，以及$\log^5 p = o(n)$，则

$$TCvM_{n,\alpha}^2 \longrightarrow \int_0^1 |B(t) + L_2(\psi^{-1}(t\psi(t_0)))/\sqrt{\psi(t_0)}|^2 dt, \quad in\ distribution,$$

其中 $B(t)$ 是标准布朗运动。

推论 4.1 表明，若 $L_1(t)$ 和 $L_2(t)$ 是非零函数，则投影检验$TCvM_{n,\alpha}^2$在全局备择假设下是一致的，并且即使协变量维度 p 随样本量 n 指数增长，也能以参数速率 $n^{-1/2}$ 检测出与零假设不同的局部备择。

5 实际应用中的检验统计量

5.1 组合投影检验统计量

注意，可以基于不同的投影构造一系列检验统计量。根据命题 2.1 和推论 4.1，若极限函数 $L_1(t)$ 非零，则投影检验

$TCvM^2_{n,\ \alpha}$相对于均匀测度 $\mathcal{L}$，在几乎所有投影 $\alpha \in \mathcal{S}^{p-1}$下均为 H_1一致的。然而，对于某些不恰当的投影选择，该检验仍可能遭受显著的功效损失，在超高维情形下尤其如此。另一个潜在问题是，投影检验统计量$TCvM^2_{n,\ \alpha}$的取值可能因投影不同而异，若仅依赖单个投影，可能导致统计功效表现不稳定。为解决这些局限，我们提出结合多个投影检验统计量$TCvM^2_{n,\ \alpha}$，构造最终检验统计量，从而提升整体统计功效。

文献中有多种方法可用于合并检验统计量或其对应的p值，例如经典的费希尔组合方法（Fisher，1925年）以及较新的柯西组合方法（Liu 和 Xie，2020年）。对于一组给定的投影 $\{\alpha_i \in \mathcal{S}^{p-1}, i = 1, \ldots, k_1\}$，投影检验$TCvM^2_{n,\alpha_i}$的渐近$p$值为$\hat{p}_{1\alpha_i} = 1 - \Psi(TCvM^2_{n,\alpha_i})$，其中 $\Psi(\cdot)$表示随机变量 $\int_0^1 B^2(t)dt$ 的累积分布函数。与定理 3.1 的论证类似，并应用连续映射定理，在H_0条件下，

$$(\hat{p}_{1\alpha_1}, \ldots, \hat{p}_{1\alpha_{k_1}}) \longrightarrow (p_1, \ldots, p_{k_1}), \quad \text{in distribution}, \tag{5.1}$$

其中 $\{p_i\colon\ i = 1, \ldots, k_1\}$ 是 $(0, 1)$ 上的均匀随机变量。注意，这些 p值 $p_1, \ldots, p_{k_1}$可能相互依赖。因此，经典的费希尔组合方法无法在此使用，因为该方法要求这些 p值相互独立。相比之下，柯西组合方法对 p值之间的依赖具有稳健性；有关此问题的更多细节，请参见 Liu 和 Xie（2020年）。因此，我们采用该方法合并这些 p值，以构造最终的检验统计量。基于柯西组合方法的检验统计量定义为

$$TCvM^2_C = \sum_{i=1}^{k_1} w_i \tan\{(\frac{1}{2} - \hat{p}_{1\alpha_i})\pi\} = \sum_{i=1}^{k_1} w_i \tan\{[\Psi(TCvM^2_{n,\alpha_i}) - \frac{1}{2}]\pi\}, \tag{5.2}$$

其中 w_i 为满足条件 $\sum_{i=1}^{k_1} w_i = 1$的非负权重。本文简单地采用等权重，即对于 $i = 1, \ldots, k_1$，取 $w_i = 1/k_1$，这在我们的模拟研究中表现良好。由 (5.1) 并再次应用连续映射定理可得，在 H_0下，

$$TCvM^2_C \longrightarrow \sum_{i=1}^{k_1} w_i \tan\{(\frac{1}{2} - p_i)\pi\}, \quad \text{in distribution}.$$

令 $P_w = \sum_{i=1}^{k_1} w_i \tan\{(\frac{1}{2} - p_i)\pi\}$。若 $p_1, \ldots, p_{k_1}$是定义在 $(0, 1)$ 上的独立同分布均匀随机变量，则易知 P_w 服从标准柯西分布， $Cauchy(0,1)$。此外，即使 p-值 $p_1, \ldots, p_{k_1}$ 相互依赖，Liu 和 Xie (2020) 证明了在原假设下，$\lim_{t\to\infty} \mathbb{P}(P_w > t)/\mathbb{P}(P_0 > t) = 1$ ，其中 P_0 表示一个标准柯西随机变量。这意味着，即使 p-值之间存在依赖关系，P_w 的尾概率也可以用标准柯西分布的尾概率近似。因此，基于柯西合并的检验 $TCvM^2_C$的渐近临界值也可以用标准柯西分布的分位数近似。

在模型检验中，众所周知，基于经验过程的检验（例如$TCvM^2_{n,\ \alpha}$）对低频备择模型更为敏感，而局部平滑检验通常对高频备择模型更为敏感，检验效能也更高。然而在实践中，若没有先验

信息，研究者通常无法预先知道底层回归模型属于哪一类。因此，提出一种同时对低频和高频备择假设都敏感的检验统计量至关重要。请注意，柯西合并主要由最小的 p 值决定（Liu 和 Xie，2020年）；因此，一个自然的想法是采用这种合并方法，将基于经验过程的检验与局部平滑检验结合起来，以实现这一目标。最近，Tan 等（2025）提出了一种通过投影针对超高维回归模型的新型局部平滑检验，该检验对高频备择模型也表现出很好的检验效能。基于给定投影 α 的检验统计量为

$$PLS_{n,\alpha} = \frac{\sum_{1\le i\neq j\le n} \varepsilon_i(\hat{\beta})\varepsilon_j(\hat{\beta})K(\frac{\alpha^\top X_i - \alpha^\top X_j}{h})}{\left(2\sum_{1\le i\neq j\le n} \varepsilon_i^2(\hat{\beta})\varepsilon_j^2(\hat{\beta})K^2(\frac{\alpha^\top X_i - \alpha^\top X_j}{h})\right)^{1/2}}, \tag{5.3}$$

其中，$K(\cdot)$ 是一元核函数，h 是带宽。根据 Tan 等（2025）的推论 3.1，对于任意给定的投影 α，$PLS_{n,\ \alpha}$ 在 H_0 下依分布收敛于标准正态分布。令 $\hat{p}_{2\alpha_i} = 1 - \Phi(PLS_{n,\ \alpha_i})$ 为 $PLS_{n,\ \alpha_i}$ 的渐近 p 值，其中 $i = 1, \ldots, k_2$，$\Phi(\cdot)$ 表示标准正态分布的累积分布函数。接下来，我们提出一种基于柯西组合方法的混合检验统计量：

$$Hybrid_C = \sum_{i=1}^{k_1} w_i \tan\{(\frac{1}{2} - \hat{p}_{1\alpha_i})\pi\} + \sum_{j=1}^{k_2} v_j \tan\{(\frac{1}{2} - \hat{p}_{2\alpha_j})\pi\}, \tag{5.4}$$

其中，权重 w_i 和 v_j 满足 $\sum_{i=1}^{k_1} w_i + \sum_{j=1}^{k_2} v_j = 1$。$Hybrid_C$ 的渐近临界值也可以根据标准柯西分布的分位数确定。

注 1. 组合检验 $TCvM_C^2$ 和 $Hybrid_C$ 均为渐近分布无关，因此我们无需采用野生自助法等重抽样方法来近似极限零分布。因此，这些检验在实践中易于实施，尤其适用于超高维情形。在全局备择假设 H_1 和温和条件下，根据 *Tan* 等（*2025*）的推论 *4.1* 和定理 *3.3*，渐近 p 值 $\hat{p}_{1\alpha_i}$ 和 $\hat{p}_{2\alpha_i}$ 对于几乎所有投影 $\alpha_i \in \mathcal{S}^{p-1}$ 都收敛于零。因此，随着样本量 $n \to \infty$，组合检验统计量 $TCvM_C^2$ 和 $Hybrid_C$ 对于几乎所有投影都趋于无穷大。*Liu* 和 *Xie*（*2020*）从理论上证明，p 值的柯西合并对相关的 p 值具有稳健性，且主要受最小的 p 值影响。因此，即使基于单个投影 α_i 的检验统计量 $TCvM_{n,\ \alpha_i}^2$ 或 $PLS_{n,\ \alpha_i}$ 可能不具一致性，组合检验统计量 $TCvM_C^2$ 和 $Hybrid_C$ 在备择假设下仍能表现出稳健的检验功效。我们还注意到，基于经验过程的检验 $TCvM_{n,\ \alpha}^2$ 对低频备择更敏感，而局部平滑检验 $PLS_{n,\ \alpha}$ 对高频模型的检验功效更高。由于 p 值的柯西合并主要由最小的 p 值决定，预计混合检验 $Hybrid_C$ 对低频和高频备择模型均具有较高的检验功效。第 *6* 节的模拟研究验证了这些理论结论。

5.2 投影的选择

在实际应用中，在超高维情形下，组合检验 $TCvM_C^2$和 $Hybrid_C$的检验效能会受到投影选择的显著影响。为说明这一点，我们考虑

高斯线性模型拟合优度检验问题。不失一般性地，假设数据已标准化，因此有 $E(X) = 0$ 和 $E(Y) = 0$。根据推论 4.1，对于给定的投影 $\alpha \in \mathcal{S}^{p-1}$，投影检验 $TCvM^2_{n,\alpha}$ 只有在某个 $t \in \mathbb{R}$ 满足 $E[\varepsilon(\tilde{\beta}_0)I(\alpha^\top X \le t)] \neq 0$ 时，才能在备择假设下具有检验效能，其中 $\varepsilon(\tilde{\beta}_0) = Y - \tilde{\beta}_0^\top X$。考虑一个极端情形：对于某个投影 $\alpha \in \mathcal{S}^{p-1}$，$\varepsilon(\tilde{\beta}_0)$ 与 $\alpha^\top X$ 均值独立，即 $E[\varepsilon(\tilde{\beta}_0)|\alpha^\top X] = E[\varepsilon(\tilde{\beta}_0)]$。若 $X \sim N(0, I_p)$，ε 与 X 相互独立，且 $\alpha_S = 0$，其中 S 为真实活跃集，则可实现这一情形。事实上，若 $X \sim N(0, I_p)$，根据 Bühlmann 和 van de Geer（2015），β_0 的支撑集 S 满足 $\tilde{S} \subset S$。注意，$\varepsilon(\tilde{\beta}_0) = m(X) + \varepsilon - \tilde{\beta}_0^\top X = m(X_S) + \varepsilon - \tilde{\beta}_{0S}^\top X_S$ 且 $\alpha^\top X = \alpha_{S^c}^\top X_{S^c}$，其中 S^c 表示 S 的补集。这意味着 $\varepsilon(\tilde{\beta}_0)$ 与 $\alpha^\top X$ 独立，因此 $E[\varepsilon(\tilde{\beta}_0)|\alpha^\top X] = E[\varepsilon(\tilde{\beta}_0)]$。回顾 $E(X) = 0$ 以及 $E(Y) = 0$，可得 $E[\varepsilon(\tilde{\beta}_0)|\alpha^\top X] = 0$，从而 $E[\varepsilon(\tilde{\beta}_0)I(\alpha^\top X \le t)] \equiv 0$。这意味着在此类情形下，投影检验 $TCvM^2_{n,\alpha}$ 可能不具备检验效能。因此，为确保良好的检验效能，我们应选择使 $\alpha^\top X$ 与误差项 $\varepsilon(\tilde{\beta}_0)$ 高度相关的投影。为此，我们提出一种数据驱动程序来选择投影，以确保所提出的检验 $TCvM^2_C$ 和 $Hybrid_C$ 在备择假设下具有良好的检验效能。

回顾一下，$\varepsilon(\tilde{\beta}_0) = m(X) - \mu(\tilde{\beta}_0^\top X) + \varepsilon$，并且我们假设零假设和备择假设下的回归模型均具有稀疏性。受此启发，不失一般性地假设误差项 $\varepsilon(\tilde{\beta}_0)$ 具有多指标模型结构，即 $\varepsilon(\tilde{\beta}_0) = g(\tilde{\vartheta}, \tilde{\theta}_1^\top X, \ldots, \tilde{\theta}_d^\top X, \varepsilon)$，其中 $\tilde{\vartheta} \in \mathbb{R}$，而投影 $\tilde{\theta}_1, \ldots, \tilde{\theta}_d \in \mathcal{S}^{p-1}$ 是潜在参数。如果该模型中不存在降维结构，那么我们将有 $d = p$ 和 $\tilde{\theta}_i = e_i$，其中 $e_i \in \mathcal{S}^{p-1}$ 在第 i 个分量上为 1，其余分量为 0。然而，鉴于零假设和备择假设下均假设具有稀疏性，这种情况不太可能发生。显然，投影变量 $\tilde{\theta}_1^\top X, \cdots \tilde{\theta}_d^\top X$ 与误差（$\varepsilon\ \tilde{\beta}_0$）高度相关。此外，具有 $\theta_0 = \beta_0/\|\beta_0\|$ 的投影预测变量也可能与误差高度相关（Stute 和 Zhu，2002）。因此，一个自然的想法是基于这些潜在投影 $\theta_0, \theta_1, \ldots, \theta_d$ 构造基于柯西组合的检验统计量。

然而，所有这些潜在投影都是未知的，在实际应用中必须进行估计。如果使用完整样本来估计潜在投影，估计得到的投影与检验统计量之间就会产生依赖关系，从而使渐近分布的推导更加复杂。为解决这一问题，我们将数据随机分为两部分，$\mathcal{D}_1$ 和 $\mathcal{D}_2$，且大小相等。这里不失一般性地假设样本量 n 为偶数。我们使用数据的第一部分，即 $\mathcal{D}_1$，来估计投影 $\theta_0, \theta_1, \ldots, \theta_d$，随后基于第二部分 $\mathcal{D}_2$ 构造检验统计量。分别使用 $\mathcal{D}_1$ 得到的 $\tilde{\theta}_0, \tilde{\theta}_1, \ldots, \tilde{\theta}_d$ 的估计量为 $\hat{\theta}_0^{(1)}, \hat{\theta}_1^{(1)}, \ldots, \hat{\theta}^{(1)}{}_{\hat{d}^{(1)}}$。由此得到的柯西组合检验统计量定义如下：

$$TCvM_C^2 = \sum_{i=0}^{\hat{d}^{(1)}} w_i \tan\{(\frac{1}{2} - \hat{p}^{(2)}_{1\hat{\theta}_i^{(1)}})\pi\}, \tag{5.5}$$

其中 $\sum_{i=0}^{\hat{d}^{(1)}} w_i = 1$，而 $\hat{p}^{(2)}_{1\hat{\theta}_i^{(1)}} = 1 - \Psi(TCvM^{2(2)}_{n,\hat{\theta}_i^{(1)}})$ 是使用 $\mathcal{D}_2$ 和估计的投影 $\hat{\theta}_i^{(1)}$ 计算得到的。

式（5.5）中的检验 $TCvM_C^2$ 仅使用一半数据来构造检验统计量，这可能导致在备择假设下检验效能损失。为解决这一问题，我们采用交叉拟合策略来提高检验效能。具体而言，我们交换两个子样本的角色，使用 $\mathcal{D}_2$

估计潜在投影，而使用 $\mathcal{D}_1$ 构造检验统计量。由此得到的交叉拟合检验统计量为

$$TCvM_{CF}^2 = \sum_{i=0}^{\hat{d}^{(1)}} w_i \tan\{(\frac{1}{2} - \hat{p}^{(2)}_{1\hat{\theta}_i^{(1)}})\pi\} + \sum_{j=0}^{\hat{d}^{(2)}} w_j \tan\{(\frac{1}{2} - \hat{p}^{(1)}_{1\hat{\theta}_j^{(2)}})\pi\}, \tag{5.6}$$

其中，$\sum_{i=0}^{\hat{d}^{(1)}} w_i + \sum_{j=0}^{\hat{d}^{(2)}} w_j = 1$、$\{\hat{\theta}_j^{(2)}\}_{j=0}^{\hat{d}^{(2)}}$使用 $\mathcal{D}_2$进行估计，而 $\hat{p}^{(1)}_{1\hat{\theta}_j^{(2)}} = 1 - \Psi(TCvM_n^{2(1)}, \hat{\theta}_j^{(2)})$ 使用 $\mathcal{D}_1$ 和投影 $\hat{\theta}_j^{(2)}$ 计算得到。类似地，所得的交叉拟合混合检验统计量为

$$\begin{aligned} Hybrid_{CF} = & \sum_{i=0}^{\hat{d}^{(1)}} w_i \tan\{(\frac{1}{2} - \hat{p}^{(2)}_{1\hat{\theta}_i^{(1)}})\pi\} + \sum_{j=0}^{\hat{d}^{(2)}} w_j \tan\{(\frac{1}{2} - \hat{p}^{(1)}_{1\hat{\theta}_j^{(2)}})\pi\} \\ & + \sum_{i=1}^{\hat{d}^{(1)}} v_i \tan\{(\frac{1}{2} - \hat{p}^{(2)}_{2\hat{\theta}_i^{(1)}})\pi\} + \sum_{j=1}^{\hat{d}^{(2)}} v_j \tan\{(\frac{1}{2} - \hat{p}^{(1)}_{2\hat{\theta}_j^{(2)}})\pi\}, \end{aligned} \tag{5.7}$$

其中，权重 w_i，w_j，v_i，v_j 满足 $\sum_{i=0}^{\hat{d}^{(1)}} w_i + \sum_{j=0}^{\hat{d}^{(2)}} w_j + \sum_{i=1}^{\hat{l}^{(1)}} v_i + \sum_{j=1}^{\hat{l}^{(2)}} v_j = 1$；$\hat{p}^{(2)}_{2\hat{\theta}_i^{(1)}} = 1 - \Phi(PLS_n^{(2)}, \hat{\theta}_i^{(1)})$ 使用 $\mathcal{D}_2$和来自 $\mathcal{D}_1$的投影 $\hat{\theta}_i^{(1)}$ 计算得到，而 $\hat{p}^{(1)}_{2\hat{\theta}_j^{(2)}} = 1 - \Phi(PLS_n^{(1)}, \hat{\theta}_j^{(2)})$ 使用 $\mathcal{D}_1$ 和来自 $\mathcal{D}_2$的投影 $\hat{\theta}_i^{(2)}$ 计算得到。

6 数值研究

6.1 模拟

本节通过模拟研究评估所提出的检验 $TCvM_C^2$、$TCvM_{CF}^2$ 和 $Hybrid_{CF}$ 在协变量维度 p 可能远大于样本量 n 时的有限样本性能。由于我们的检验依赖于参数估计，参数估计不准确可能会影响检验的有限样本性能。因此，我们采用 β_0 的后Lasso估计量来构造检验统计量。post-Lasso 方法对 Lasso 估计量所选定的模型进行最小二乘拟合。根据 Belloni 和 Chernozhukov（2013），就收敛速率而言，后Lasso估计量的表现至少不逊于 Lasso，并且还具有偏差较小的优势。这意味着后Lasso估计量 β_0 满足条件（A1），由此得到的检验渐近分布无关。我们还使用 Lasso 估计量对所提出的检验进行了模拟研究。未报告的结果表明，尽管在检验高斯线性模型时，所提出的检验能够控制经验尺寸，但在检验逻辑回归模型时却无法维持显著性水平。这可能表明，在高维情形下，回归模型拟合优度检验的有限样本性能也会受到参数估计方法的影响。

为计算检验 $TCvM_C^2$、$TCvM_{CF}^2$ 和 $Hybrid_{CF}$，我们将数据随机分成两部分，$\mathcal{D}_1$ 和 $\mathcal{D}_2$，然后根据 (5.5)、(5.6) 和 (5.7) 构造检验统计量。为

~ ~ ~准确估计潜在投影 θ_0，θ_1、.....、θ_d，我们采用距离相关性确定独立性筛选方法（Li 等，2012）选取排名前 $^{n/2}{}_{\log(n/2)}$ 的变量，然后应用稀疏充分降维技术 LassoSIR（Lin 等，2019）构造 $i = 0$、1、.....、d对应的估计投影 $\hat{}\theta_i^{(1)}$ 和 $\hat{}\theta_i^{(2)}$。此外，由于所提出的检验涉及非参数估计，我们使用 Epanechnikov 核 $K(x) = (3/4)(1 - x^2)I(|x| \leq 1)$ 进行计算，并通过交叉验证自动选择带宽。

我们将我们的检验与 Shah 和 B¨uhlmann（2018）的 RP_n 检验、Jankov´a 等人（2020）的 GRP_n 检验，以及 Tan 等人（2025）提出的局部平滑检验 $\hat{}T^C_{\text{Fisher}}$ 进行比较。后者基于结合来自不同投影的投影检验 $PLS_{n,\alpha}$。关于这些检验的具体形式，详情请参阅相应论文。在以下模拟中，参数 $a = 0$ 对应于原假设，而 $a \neq 0$ 对应于备择假设。模拟结果基于 1000 次重复试验在名义水平 $\tau = 0.05$下的平均值。RP_n 检验的模拟结果使用 R 软件包 RPtests 计算，而 GRP_n的结果则通过运行 Jankov´a 等人（2020）提供的、网址为 https://github.com/jankova/GRPtests 的网站上提供的代码计算。我们使用 Epanechnikov 核 $K(x) = (3/4)(1 - x^2)I(|x| \leq 1)$计算检验 $\hat{}T^C_{\text{Fisher}}$，并采用 Tan 等人（2025）建议的带宽 $h = n^{-2/9}$。此外，尽管以下所有原假设模型的截距均为零，但在模拟研究中我们仍将其作为未知参数进行估计。

在第一个模拟研究中，我们考虑高斯线性模型的检验问题，其中协变量维度 p 可能远大于样本量 n。

研究*1*。 根据以下模型生成数据:

$$\begin{aligned} H_{11} : Y &= \beta_0^\top X + 0.1a(\beta_0^\top X)^2 + \varepsilon; \\ H_{12} : Y &= \beta_0^\top X + a\cos(0.6\pi\beta_0^\top X) + \varepsilon; \\ H_{13} : Y &= \beta_0^\top X + a\exp(0.25\beta_1^\top X) + \varepsilon; \end{aligned}$$

其中 $\beta_0 = (1, 1, 1, 1, 1, 0, ..., 0)^\top$ 和 $\underbrace{\beta_1}$ $= (1, ..., 1, 0, ..., 0)$ 满足 $p_1 = 10$。协变量 $X \in \mathbb{R}^{p_{p_1}}$
与标准高斯误差 ε 相互独立，其中 $N(0, \Sigma)$ $\Sigma = I_p$ 或 $\Sigma = (\rho^{|i-j|})_{p\times p}$，且 $\rho = 0.4$ 和 $\rho = 0.8$。这里，H_{11} 和 H_{13} 是低频模型，而 H_{12} 是在备择假设下的高频模型。我们考虑样本量 $n = 300$，其中协变量维度为 $p \in \{50, 100, 300, 600, 900, 1200\}$。

模拟结果见表1–3。我们观察到，在所有样本量和维度的情况下，所有检验都能很好地控制显著性水平。Jankov´a 等人（2020）的 GRP_n 检验略显保守，经验尺寸较小。对于低频模型 H_{11} 和 H_{13}，我们的检验$TCvM^2_C$、$TCvM^2_{CF}$和 $Hybrid_{CF}$在各种设定下通常比另外三种竞争检验具有更高的检验功效。对于高频模型 H_{12}，检验 RP_n、GRP_n、$TCvM^2_C$和$TCvM^2_{CF}$几乎没有检验功效，即使在低维设定（$p = 50$）下也是如此。相比之下，局部平滑检验$\hat{}T^C_{\text{Fisher}}$ 和我们的检验$Hybrid_{CF}$ 的检验功效显著高于另外四种检验。这一结果与低维模型检验中的传统发现一致：局部平滑检验通常对高频模型更敏感，而基于经验过程的检验对低频模型更有检验功效。有趣的是，我们的混合检验 $Hybrid_{CF}$ 兼具局部平滑检验和基于经验过程的检验的优点，能够保持较高的检验功效

对于低频和高频备择假设均如此。

表1：研究1中针对 H_{11} 的检验$TCvM_C^2$、 $TCvM_{CF}^2$、$\hat{T}^C_{\text{Fisher}}$、 $Hybrid_{CF}$、 RP_n和GRP_n 的经验尺寸和检验功效。

	a	n=300 p=50	n=300 p=100	n=300 p=300	n=300 p=600	n=300 p=900	n=300 p=1200
$TCvM_C^2,\ \Sigma = I_p$	0.0	0.052	0.040	0.042	0.051	0.046	0.047
	0.5	0.664	0.581	0.401	0.357	0.334	0.279
	1.0	0.991	0.953	0.831	0.716	0.628	0.608
$TCvM_{CF}^2,\ \Sigma = I_p$	0.0	0.047	0.035	0.035	0.056	0.048	0.049
	0.5	0.847	0.731	0.547	0.474	0.432	0.396
	1.0	1.000	0.995	0.929	0.875	0.791	0.773
$\hat{T}^C_{\text{Fisher}},\ \Sigma = I_p$	0.0	0.021	0.034	0.026	0.036	0.026	0.034
	0.5	0.195	0.143	0.111	0.104	0.084	0.073
	1.0	0.917	0.791	0.534	0.419	0.343	0.300
$Hybrid_{CF},\ \Sigma = I_p$	0.0	0.057	0.050	0.058	0.065	0.054	0.063
	0.5	0.772	0.679	0.478	0.405	0.362	0.336
	1.0	0.999	0.993	0.917	0.828	0.744	0.730
$RP_n,\ \Sigma = I_p$	0.0	0.038	0.039	0.034	0.024	0.028	0.041
	0.5	0.094	0.086	0.058	0.054	0.064	0.057
	1.0	0.176	0.149	0.144	0.127	0.110	0.114
$GRP_n,\ \Sigma = I_p$	0.0	0.034	0.020	0.018	0.017	0.014	0.012
	0.5	0.085	0.049	0.025	0.019	0.018	0.025
	1.0	0.291	0.182	0.079	0.054	0.040	0.055
$TCvM_C^2,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.039	0.051	0.045	0.044	0.030	0.037
	0.5	0.994	0.984	0.903	0.857	0.821	0.808
	1.0	1.000	1.000	0.982	0.951	0.924	0.904
$TCvM_{CF}^2\ \ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.043	0.048	0.046	0.046	0.037	0.037
	0.5	0.999	1.000	0.981	0.959	0.947	0.937
	1.0	1.000	1.000	0.997	0.990	0.982	0.981
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.025	0.029	0.031	0.026	0.033	0.034
	0.5	0.896	0.779	0.627	0.574	0.529	0.487
	1.0	1.000	1.000	0.985	0.918	0.887	0.874
$Hybrid_{CF},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.055	0.058	0.057	0.053	0.047	0.050
	0.5	0.999	1.000	0.975	0.950	0.929	0.908
	1.0	1.000	1.000	0.996	0.983	0.980	0.978
$RP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.029	0.037	0.035	0.037	0.039	0.042
	0.5	0.470	0.391	0.340	0.314	0.327	0.318
	1.0	0.601	0.534	0.481	0.436	0.445	0.485
$GRP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.026	0.031	0.030	0.025	0.019	0.019
	0.5	0.785	0.604	0.280	0.162	0.139	0.121
	1.0	0.998	0.981	0.700	0.582	0.542	0.475
$TCvM_C^2,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.047	0.050	0.039	0.030	0.048	0.047
	0.5	1.000	1.000	0.998	0.997	0.988	0.977
	1.0	1.000	1.000	1.000	0.994	0.990	0.976
$TCvM_{CF}^2\ \ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.040	0.046	0.037	0.035	0.056	0.049
	0.5	1.000	1.000	1.000	1.000	0.999	1.000
	1.0	1.000	1.000	1.000	0.999	0.999	0.997
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.028	0.028	0.025	0.032	0.041	0.038
	0.5	1.000	1.000	0.997	0.989	0.970	0.968
	1.0	1.000	1.000	1.000	0.998	0.992	0.984
$Hybrid_{CF},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.050	0.055	0.040	0.055	0.068	0.069
	0.5	1.000	1.000	1.000	1.000	0.998	0.998
	1.0	1.000	1.000	1.000	0.999	0.999	0.998
$RP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.039	0.029	0.039	0.032	0.035	0.034
	0.5	0.746	0.654	0.578	0.574	0.582	0.541
	1.0	0.741	0.653	0.603	0.561	0.547	0.552
$GRP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.043	0.041	0.027	0.024	0.025	0.037
	0.5	1.000	1.000	0.987	0.969	0.961	0.950
	1.0	1.000	1.000	1.000	0.997	0.995	0.999

接下来，我们研究所提出的逻辑回归模型拟合优度检验的有限样本性能。

表2：研究1中检验$TCvM_C^2$、 $TCvM_{CF}^2$、 $\hat{T}^C_{\text{Fisher}}$、 $Hybrid_{CF}$、 RP_n和GRP_n 对 H_{12} 的经验尺寸和检验功效。

	a	n=300 p=50	n=300 p=100	n=300 p=300	n=300 p=600	n=300 p=900	n=300 p=1200
$TCvM_C^2,\ \Sigma = I_p$	0.0	0.033	0.045	0.052	0.042	0.048	0.031
	0.5	0.055	0.052	0.039	0.033	0.046	0.061
	1.0	0.046	0.059	0.035	0.044	0.043	0.035
$TCvM_{CF}^2,\ \Sigma = I_p$	0.0	0.048	0.049	0.042	0.045	0.046	0.038
	0.5	0.054	0.055	0.043	0.048	0.032	0.051
	1.0	0.052	0.052	0.037	0.051	0.048	0.043
$\hat{T}^C_{\text{Fisher}},\ \Sigma = I_p$	0.0	0.020	0.017	0.031	0.036	0.037	0.037
	0.5	0.320	0.216	0.085	0.074	0.053	0.053
	1.0	0.838	0.650	0.227	0.124	0.091	0.095
$Hybrid_{CF},\ \Sigma = I_p$	0.0	0.050	0.055	0.050	0.059	0.059	0.050
	0.5	0.274	0.182	0.096	0.075	0.052	0.059
	1.0	0.689	0.413	0.178	0.116	0.083	0.085
$RP_n,\ \Sigma = I_p$	0.0	0.036	0.030	0.039	0.040	0.038	0.032
	0.5	0.045	0.037	0.032	0.042	0.029	0.040
	1.0	0.033	0.058	0.032	0.049	0.042	0.036
$GRP_n,\ \Sigma = I_p$	0.0	0.037	0.033	0.018	0.010	0.016	0.012
	0.5	0.040	0.034	0.027	0.019	0.014	0.015
	1.0	0.069	0.066	0.045	0.027	0.032	0.027
$TCvM_C^2,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.048	0.052	0.041	0.037	0.041	0.058
	0.5	0.042	0.043	0.039	0.038	0.049	0.057
	1.0	0.052	0.046	0.037	0.037	0.039	0.043
$TCvM_{CF}^2\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.044	0.053	0.047	0.038	0.044	0.040
	0.5	0.037	0.043	0.043	0.042	0.037	0.057
	1.0	0.052	0.055	0.044	0.042	0.035	0.043
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.031	0.027	0.030	0.031	0.037	0.037
	0.5	0.350	0.279	0.155	0.142	0.114	0.106
	1.0	0.924	0.837	0.567	0.404	0.308	0.258
$Hybrid_{CF},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.054	0.060	0.058	0.059	0.058	0.068
	0.5	0.289	0.197	0.123	0.121	0.094	0.094
	1.0	0.753	0.563	0.315	0.210	0.160	0.153
$RP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.036	0.039	0.033	0.041	0.042	0.041
	0.5	0.040	0.051	0.019	0.039	0.041	0.045
	1.0	0.035	0.037	0.037	0.041	0.048	0.041
$GRP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.033	0.034	0.021	0.030	0.022	0.016
	0.5	0.043	0.053	0.032	0.018	0.025	0.021
	1.0	0.075	0.056	0.062	0.036	0.027	0.038
$TCvM_C^2,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.032	0.047	0.049	0.038	0.043	0.045
	0.5	0.042	0.053	0.042	0.048	0.059	0.052
	1.0	0.051	0.048	0.039	0.040	0.045	0.034
$TCvM_{CF}^2\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.036	0.046	0.040	0.041	0.054	0.045
	0.5	0.037	0.046	0.046	0.049	0.042	0.041
	1.0	0.047	0.044	0.031	0.052	0.049	0.032
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.024	0.033	0.032	0.034	0.028	0.040
	0.5	0.307	0.247	0.246	0.178	0.150	0.169
	1.0	0.911	0.872	0.773	0.693	0.633	0.581
$Hybrid_{CF},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.054	0.056	0.066	0.057	0.051	0.069
	0.5	0.332	0.267	0.216	0.153	0.135	0.123
	1.0	0.881	0.767	0.536	0.399	0.403	0.306
$RP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.035	0.052	0.051	0.049	0.046	0.035
	0.5	0.050	0.029	0.038	0.042	0.037	0.032
	1.0	0.055	0.034	0.036	0.044	0.040	0.033
$GRP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.038	0.032	0.032	0.034	0.027	0.035
	0.5	0.065	0.039	0.047	0.038	0.037	0.043
	1.0	0.074	0.068	0.061	0.064	0.059	0.057

研究*2.* 数据根据逻辑回归模型生成，具体如下：

$$Y|X \sim Bernoulli\{\mu(\beta_0^\top X + ag(X))\},$$

表3：研究1中检验$TCvM_C^2$、 $TCvM_{CF}^2$、 $\hat{T}^C_{\text{Fisher}}$、 $Hybrid_{CF}$、 RP_n和GRP_n 针对 H_{13} 的经验尺寸和检验功效。

	a	n=300 p=50	n=300 p=100	n=300 p=300	n=300 p=600	n=300 p=900	n=300 p=1200
$TCvM_C^2,\ \Sigma = I_p$	0.0	0.047	0.043	0.046	0.043	0.038	0.040
	0.5	0.312	0.211	0.135	0.150	0.115	0.104
	1.0	0.776	0.668	0.401	0.310	0.267	0.243
$TCvM_{CF}^2,\ \Sigma = I_p$	0.0	0.061	0.037	0.045	0.039	0.037	0.045
	0.5	0.397	0.294	0.189	0.177	0.132	0.120
	1.0	0.920	0.834	0.545	0.417	0.387	0.310
$\hat{T}^C_{\text{Fisher}},\ \Sigma = I_p$	0.0	0.025	0.021	0.025	0.022	0.025	0.031
	0.5	0.056	0.038	0.037	0.038	0.026	0.029
	1.0	0.189	0.095	0.061	0.063	0.051	0.050
$Hybrid_{CF},\ \Sigma = I_p$	0.0	0.055	0.051	0.061	0.044	0.044	0.052
	0.5	0.334	0.226	0.164	0.136	0.111	0.110
	1.0	0.892	0.778	0.483	0.365	0.332	0.258
$RP_n,\ \Sigma = I_p$	0.0	0.043	0.044	0.033	0.044	0.034	0.048
	0.5	0.054	0.045	0.061	0.046	0.042	0.050
	1.0	0.072	0.061	0.069	0.058	0.065	0.053
$GRP_n,\ \Sigma = I_p$	0.0	0.026	0.048	0.019	0.010	0.008	0.014
	0.5	0.065	0.056	0.022	0.024	0.016	0.018
	1.0	0.108	0.097	0.040	0.040	0.028	0.029
$TCvM_C^2,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.037	0.043	0.052	0.045	0.043	0.042
	0.5	0.943	0.866	0.741	0.647	0.613	0.574
	1.0	0.998	0.996	0.943	0.869	0.798	0.803
$TCvM_{CF}^2\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.038	0.040	0.045	0.038	0.041	0.047
	0.5	0.990	0.976	0.895	0.820	0.790	0.758
	1.0	1.000	1.000	0.992	0.974	0.946	0.945
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.031	0.026	0.028	0.033	0.030	0.037
	0.5	0.437	0.306	0.187	0.143	0.129	0.138
	1.0	0.954	0.831	0.586	0.409	0.350	0.305
$Hybrid_{CF},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.044	0.060	0.061	0.057	0.056	0.056
	0.5	0.981	0.959	0.855	0.781	0.735	0.716
	1.0	1.000	1.000	0.986	0.960	0.924	0.918
$RP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.037	0.039	0.037	0.042	0.041	0.033
	0.5	0.226	0.214	0.160	0.166	0.159	0.179
	1.0	0.284	0.237	0.214	0.203	0.236	0.228
$GRP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.030	0.030	0.015	0.021	0.020	0.021
	0.5	0.642	0.462	0.211	0.155	0.134	0.129
	1.0	0.933	0.800	0.430	0.312	0.320	0.269
$TCvM_C^2,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.041	0.035	0.042	0.039	0.040	0.037
	0.5	1.000	0.999	0.989	0.964	0.954	0.941
	1.0	1.000	0.999	0.987	0.954	0.934	0.921
$TCvM_{CF}^2\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.048	0.048	0.044	0.044	0.040	0.034
	0.5	1.000	1.000	1.000	0.999	0.991	0.996
	1.0	1.000	1.000	0.998	0.997	0.990	0.986
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.036	0.024	0.034	0.032	0.034	0.042
	0.5	1.000	1.000	0.991	0.969	0.928	0.887
	1.0	1.000	1.000	0.998	0.983	0.962	0.935
$Hybrid_{CF},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.053	0.052	0.047	0.053	0.066	0.057
	0.5	1.000	1.000	1.000	0.998	0.990	0.990
	1.0	1.000	1.000	0.998	0.998	0.988	0.978
$RP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.043	0.042	0.048	0.032	0.038	0.036
	0.5	0.663	0.622	0.513	0.506	0.521	0.482
	1.0	0.687	0.603	0.557	0.523	0.525	0.531
$GRP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.053	0.037	0.032	0.024	0.029	0.032
	0.5	1.000	1.000	0.993	0.985	0.974	0.975
	1.0	1.000	1.000	1.000	0.998	0.996	0.989

其中 $\mu(z) = 1/(1+\exp(-z))$。我们考虑两种不同的模型设定错误情形： $g(X)$:

$$H_{21} : g(X) \quad = \quad 0.2(\beta_0^\top X)^2,$$

$$H_{22}: g(X) \quad = \quad X^{(1)}X^{(2)} + X^{(2)}X^{(3)} + X^{(3)}X^{(4)} + X^{(4)}X^{(5)},$$

其中，参数 $\beta_0 =$(1，1，1，1，1，0，.....，0)$^\top$ 和协变量 X 与研究1中相同，样本量为 $n = 600$，协变量维度为 $p \in \{50, 100, 300, 600, 900, 1200\}$。

由于 Shah 和 B¨uhlmann（2018）提出的 RP_n 检验不能用于广义线性模型的检验，我们仅将自己的检验与 Tan 等（2025）和 Jankov´a 等（2020）分别提出的检验 GRP_n 和 $\hat{}T^C_{\text{Fisher}}$ 进行比较。模拟结果见表4-5。可以看出，我们的检验 $TCvM^2_C$、 $TCvM^2_{CF}$、 $Hybrid_{CF}$以及局部平滑检验 $\hat{}T^C_{\text{Fisher}}$在大多数情况下都能控制经验尺寸。检验 $TCvM^2_C$和 $TCvM^2_{CF}$通常较为保守，在高维情形下表现出较小的经验尺寸。检验 $\hat{}T^C_{\text{Fisher}}$ 在高相关性（$\rho = 0.8$）和高维度（$p = 1200$）情形下变得偏宽松，经验尺寸较大。然而，当协变量相关性 $\rho = 0.4$ 或 $\rho = 0.8$时， GRP_n 的经验尺寸与显著性水平相差甚远。就经验功效而言，我们的检验 $TCvM^2_{CF}$和 $Hybrid_{CF}$通常比其他竞争方法具有更高的检验功效。此外，随着协变量相关性X 增大，所有检验的经验功效均有所提高。

6.2 真实数据集

在本小节中，我们使用两个真实数据集评估所提出的检验：社区与犯罪数据以及急性髓系白血病（AML）数据（Bottomly 等，2022）。社区与犯罪数据集可从 UCI机器学习知识库（ https://archive.ics. uci.edu/dataset/183/communities+and+crime）获取，其中包含 1994 个观测值、一个响应变量（人均暴力犯罪率）以及 99 个描述人口统计特征和执法特征的预测变量。令 Y 表示人均暴力犯罪率， $X = (X^{(1)}, \ldots, X^{(99)})^\top$ 表示预测变量向量。首先，我们使用所提出的检验评估稀疏线性回归模型 $Y = a_0 + \beta_0^\top X + \varepsilon$ 对该数据集的适用性。核函数和带宽的选择与模拟研究中的相同。 p 的 $TCvM^2_C$、 $TCvM^2_{CF}$ 和 $Hybrid_{CF}$ 值分别约为 0.0037、0.0053 和 0.0001。这些结果强烈拒绝了原假设，表明 Y 与 X 之间的线性关系不足以拟合该数据集。图1给出了响应变量 Y 与 $\hat{}\beta_0^\top X$ 的散点图，其中 $\hat{}\ \beta_0$ 是线性模型中的后Lasso估计量。该图同样表明， Y 与 X 之间的线性关系可能并不合理。此外，该图还提示 Y 与 $\hat{}\beta_0^\top X$ 之间可能存在二次关系。

为确定 Y 与 X 之间更合适的关系，我们通过纳入协变量的二次项和交互项来扩展模型，从而得到以下二次多项式模型

$$Y = a_0 + \beta_0^\top X + \sum_{i,j=1}^{p} \beta_1^{(ij)} X^{(i)} X^{(j)} + \varepsilon. \tag{6.1}$$

将我们的检验应用于这一二次多项式模型时，检验所得的 p值$TCvM^2_C$、 $TCvM^2_{CF}$和 $Hybrid_{CF}$分别为 0.4759、0.6510 和 0.5989。这些结果未能拒绝原假设，表明多项式回归模型 (6.1) 可能适合拟合该数据集。为进一步直观展示这一拟合结果，我们给出了残差相对于

表4：研究2中检验$TCvM_C^2$、 $TCvM_{CF}^2$、$\hat{T}^C_{\text{Fisher}}$、 $Hybrid_{CF}$和 GRP_n在 H_{21} 下的经验尺寸和检验功效。

	a	n=600 p=50	n=600 p=100	n=600 p=300	n=600 p=600	n=600 p=900	n=600 p=1200
$TCvM_C^2,\ \Sigma = I_p$	0.0	0.068	0.065	0.033	0.013	0.029	0.027
	0.5	0.296	0.209	0.084	0.043	0.038	0.036
	1.0	0.800	0.630	0.338	0.226	0.156	0.155
$TCvM_{CF}^2,\ \Sigma = I_p$	0.0	0.087	0.065	0.016	0.011	0.009	0.011
	0.5	0.348	0.264	0.085	0.045	0.024	0.019
	1.0	0.944	0.816	0.493	0.266	0.165	0.153
$\hat{T}^C_{\text{Fisher}},\ \Sigma = I_p$	0.0	0.028	0.028	0.026	0.032	0.040	0.049
	0.5	0.084	0.077	0.032	0.034	0.034	0.045
	1.0	0.681	0.489	0.203	0.101	0.068	0.058
$Hybrid_{CF},\ \Sigma = I_p$	0.0	0.080	0.053	0.041	0.024	0.028	0.028
	0.5	0.278	0.236	0.088	0.066	0.041	0.033
	1.0	0.924	0.787	0.445	0.257	0.169	0.144
$GRP_n,\ \Sigma = I_p$	0.0	0.084	0.065	0.030	0.030	0.018	0.020
	0.5	0.068	0.048	0.034	0.027	0.027	0.023
	1.0	0.077	0.042	0.029	0.025	0.024	0.024
$TCvM_C^2,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.093	0.076	0.042	0.023	0.011	0.011
	0.5	0.566	0.444	0.259	0.161	0.102	0.085
	1.0	0.960	0.902	0.711	0.537	0.449	0.408
$TCvM_{CF}^2\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.088	0.050	0.033	0.006	0.008	0.005
	0.5	0.673	0.562	0.313	0.174	0.102	0.075
	1.0	1.000	0.992	0.950	0.776	0.677	0.581
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.027	0.023	0.033	0.030	0.056	0.076
	0.5	0.215	0.224	0.146	0.090	0.076	0.090
	1.0	0.996	0.961	0.810	0.682	0.588	0.519
$Hybrid_{CF},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.087	0.054	0.049	0.038	0.037	0.034
	0.5	0.620	0.503	0.298	0.187	0.112	0.101
	1.0	0.999	0.991	0.948	0.809	0.714	0.630
$GRP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.225	0.168	0.114	0.105	0.095	0.078
	0.5	0.151	0.103	0.057	0.036	0.036	0.038
	1.0	0.776	0.478	0.167	0.104	0.077	0.072
$TCvM_C^2,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.088	0.073	0.051	0.037	0.031	0.028
	0.5	0.924	0.923	0.833	0.757	0.644	0.591
	1.0	0.980	0.965	0.860	0.771	0.724	0.664
$TCvM_{CF}^2\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.088	0.055	0.031	0.018	0.012	0.008
	0.5	0.991	0.981	0.950	0.877	0.776	0.702
	1.0	1.000	1.000	0.988	0.971	0.946	0.914
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.034	0.031	0.026	0.058	0.079	0.130
	0.5	0.935	0.921	0.874	0.872	0.823	0.757
	1.0	1.000	0.998	0.954	0.910	0.879	0.854
$Hybrid_{CF},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.081	0.058	0.035	0.048	0.049	0.063
	0.5	0.991	0.983	0.958	0.912	0.858	0.801
	1.0	1.000	0.999	0.988	0.971	0.957	0.933
$GRP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.410	0.393	0.357	0.322	0.310	0.305
	0.5	0.830	0.769	0.606	0.461	0.366	0.305
	1.0	1.000	1.000	0.964	0.970	0.938	0.893

ˆ 二次多项式模型的拟合值 Y 的散点图（图 2）。残差与拟合值之间不存在明显趋势，进一步支持了该模型的适用性。

接下来，我们将所提出的检验应用于 AML 数据集，该数据集取自 cBioPortal (https://www.cbioportal.org/study/summary?id=aml_ohsu_2022)。该数据集包含 444 名患者的 RNA 测序表达谱；根据 ELN2017 标准，这些患者被分为高危组（319 人）和非高危组（125 人）。表达数据包含 22834 个基因，为评估提供了一个超高维场景。随后，我们评估稀疏线性逻辑回归模型对 AML 分类任务的适用性，并应用所提出的检验来检验函数形式是否

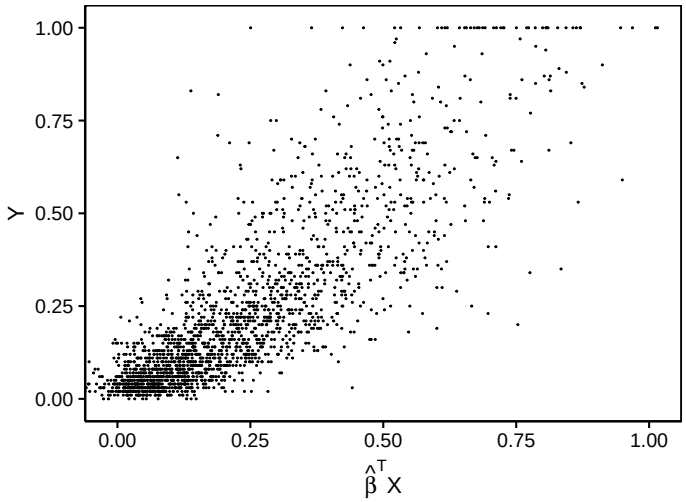


图1： Y 与$\hat{\beta}^{\top} X$的散点图。

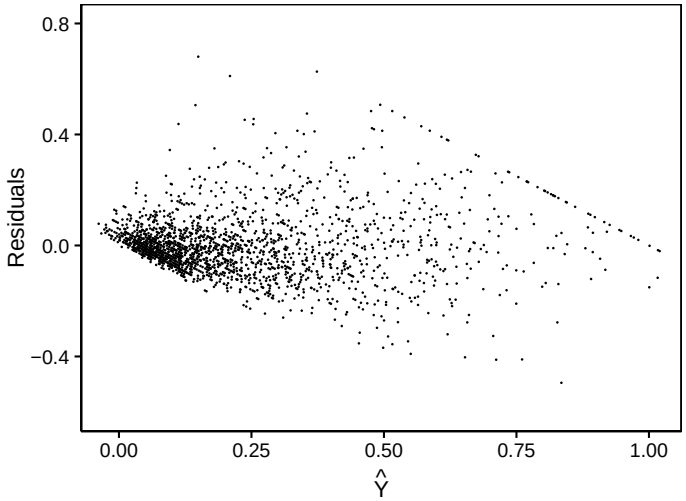


图 2：二次多项式模型的残差与拟合值$\hat{} Y$的散点图。

表5：研究2中，检验$TCvM_C^2$、 $TCvM_{CF}^2$、$\hat{T}^C_{\text{Fisher}}$、 $Hybrid_{CF}$和 GRP_n在 H_{22} 下的经验尺寸和检验功效。

	a	n=600 p=50	n=600 p=100	n=600 p=300	n=600 p=600	n=600 p=900	n=600 p=1200
$TCvM_C^2,\ \Sigma = I_p$	0.0	0.090	0.073	0.026	0.023	0.019	0.016
	0.5	0.169	0.129	0.064	0.045	0.031	0.025
	1.0	0.433	0.357	0.141	0.087	0.079	0.070
$TCvM_{CF}^2,\ \Sigma = I_p$	0.0	0.093	0.066	0.011	0.017	0.005	0.007
	0.5	0.182	0.121	0.048	0.031	0.021	0.011
	1.0	0.536	0.425	0.155	0.094	0.074	0.063
$\hat{T}^C_{\text{Fisher}},\ \Sigma = I_p$	0.0	0.022	0.027	0.033	0.046	0.033	0.043
	0.5	0.039	0.046	0.031	0.043	0.045	0.042
	1.0	0.241	0.195	0.064	0.046	0.043	0.041
$Hybrid_{CF},\ \Sigma = I_p$	0.0	0.082	0.054	0.033	0.043	0.026	0.028
	0.5	0.162	0.121	0.057	0.037	0.049	0.023
	1.0	0.489	0.367	0.153	0.103	0.095	0.066
$GRP_n,\ \Sigma = I_p$	0.0	0.076	0.065	0.035	0.029	0.025	0.020
	0.5	0.130	0.078	0.033	0.036	0.020	0.022
	1.0	0.235	0.094	0.037	0.031	0.027	0.019
$TCvM_C^2,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times\ p}$	0.0	0.092	0.076	0.038	0.019	0.019	0.015
	0.5	0.451	0.351	0.185	0.100	0.081	0.082
	1.0	0.902	0.810	0.570	0.439	0.334	0.334
$TCvM_{CF}^{2D},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times\ p}$	0.0	0.088	0.046	0.020	0.009	0.010	0.003
	0.5	0.506	0.416	0.193	0.102	0.072	0.055
	1.0	0.989	0.965	0.858	0.705	0.574	0.488
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.023	0.027	0.034	0.049	0.061	0.090
	0.5	0.154	0.136	0.090	0.067	0.079	0.063
	1.0	0.906	0.835	0.627	0.487	0.384	0.327
$Hybrid_{CF},\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.071	0.051	0.036	0.041	0.038	0.028
	0.5	0.491	0.376	0.183	0.114	0.098	0.069
	1.0	0.985	0.953	0.866	0.710	0.585	0.531
$GRP_n,\ \Sigma = (0.4^{\lvert i-j\rvert})_{p\times p}$	0.0	0.195	0.173	0.110	0.072	0.098	0.078
	0.5	0.280	0.181	0.107	0.075	0.065	0.048
	1.0	0.944	0.790	0.336	0.191	0.152	0.109
$TCvM_C^2,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times\ p}$	0.0	0.115	0.098	0.054	0.056	0.029	0.022
	0.5	0.778	0.699	0.582	0.475	0.386	0.344
	1.0	0.985	0.969	0.908	0.791	0.720	0.694
$TCvM_{CF}^2\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.089	0.077	0.028	0.024	0.012	0.007
	0.5	0.906	0.833	0.703	0.581	0.455	0.398
	1.0	1.000	1.000	0.996	0.982	0.935	0.912
$\hat{T}^C_{\text{Fisher}},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.020	0.029	0.035	0.047	0.093	0.119
	0.5	0.525	0.519	0.479	0.505	0.423	0.391
	1.0	1.000	0.999	0.975	0.937	0.906	0.886
$Hybrid_{CF},\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.082	0.068	0.034	0.037	0.061	0.063
	0.5	0.891	0.824	0.699	0.632	0.521	0.479
	1.0	1.000	1.000	0.994	0.983	0.958	0.946
$GRP_n,\ \Sigma = (0.8^{\lvert i-j\rvert})_{p\times p}$	0.0	0.409	0.392	0.368	0.357	0.345	0.323
	0.5	0.566	0.499	0.341	0.282	0.228	0.195
	1.0	1.000	1.000	0.962	0.996	0.981	0.944

条件期望， $E[Y|X] = \frac{\exp(\beta^\top X)}{1+\exp(\beta^\top X)}$，是合理的。检验 $TCvM_C^2$、$TCvM_{CF}^2$ 和 $Hybrid_{CF}$ 的 p 值分别约为 0.9877 、 0.9961 和 0.9921。普遍较高的 p 值使我们无法拒绝原假设，这表明稀疏线性逻辑回归模型适用于该数据集。我们进一步通过进行 20 次 5折交叉验证，计算线性逻辑回归模型的模型预测准确率。所得平均预测准确率和 ROC曲线下面积（AUROC）分别为 0.8514 和 0.9292。这些指标进一步证实，线性逻辑回归模型适用于该数据集。

7 讨论

本文提出了一种两步法，用于检验稀疏参数回归模型的拟合优度，适用于协变量维度 p 可能显著大于样本量n的情形。第一步，我们基于投影残差标记经验过程的鞅变换，构造 Cram´er-von Mises 型检验。在原假设下，我们的投影检验具有渐近分布无关性。在备择假设及温和条件下，对于单位球面上几乎所有投影，投影检验均具有一致性，渐近功效为 1，并且能够以 $O(n^{-1/2})$的参数速率检测偏离原假设的局部备择。第二步，我们采用柯西组合方法和数据拆分，将各投影检验合并，构成最终检验，从而提高检验功效。此外，由于基于经验过程的检验通常对低频备择更敏感，而局部平滑检验对高频备择更有功效，我们进一步提出一种混合检验，将基于经验过程的检验与 Tan 等（2025）提出的局部平滑检验相结合。模拟结果表明，该混合检验对低频和高频备择模型均表现出色。需要指出的是，我们的方法在构造检验统计量时需要进行数据拆分，这会导致检验统计量的取值存在变异。一个值得探讨的问题是，能否完全避免数据拆分策略。对于观测值之间存在依赖关系的模型检验而言，这类不需要数据拆分的替代方法尤其有用。我们还指出，条件均值函数的模型检验是检验条件矩限制的一个特例。将我们的方法扩展到超高维情形下的一般条件矩限制检验，具有重要意义。

参考文献

Bai, J. (2003). 动态模型的参数条件分布检验。《经济学与统计学评论》*85*(3), 531–549. 1 Bai, J. and S. Ng (2001). 时间序列模型中条件对称性的相合检验。《计量经济学杂志》*103*(1), 225–258. 估计与检验研究。1 Belloni, A. and V. Chernozhukov (2013). 高维稀疏模型中模型选择后的最小二乘法。*Bernoulli 19*(2), 521 – 547. 6.1 Bierens, H. J. (1982). 相合的模型设定检验。*J. Economet.20*, 105–134. 1 Bierens, H. J. (1990). 函数形式的相合条件矩检验。*Econometrica 58*, 1443–1458. 1 Bottomly, D., N. Long, A. R. Schultz, S. E. Kurtz, C. E. Tognon, K. Johnson, M. Abel, A. Agarwal, S. Avaylon, E. Benton, 等. (2022). 急性髓系白血病中药物反应与临床结局的综合分析。*Cancer cell 40*(8), 850–864. 6.2 B¨uhlmann, P. and S. Van De Geer (2011). 《高维数据统计：方法、理论与应用》. Springer Science & Business Media. 2.1

B¨uhlmann, P. and S. van de Geer (2015). 错设线性模型中的高维推断。《电子统计学杂志》*9*(1), 1449 – 1473. 4, 5.2 Cuesta-Albertos, J. A., R. Fraiman, and T. Ransford (2007). Cramér–Wold 定理的一个锐利形式。《理论概率杂志》*20*(2), 201–209. 2 Cuesta-Albertos, J. A., E. Garc´ıa-Portugu´es, M. Febrero-Bande, and W. Gonz´alez-Manteiga (2019). 基于随机投影经验过程的函数型线性模型拟合优度检验。《统计年鉴》*47*(1), 439 – 467. 1, 2, 2.1 Delgado, M. A. and W. Stute (2008). 条件模型的分布无关设定检验。《计量经济学杂志》*143*(1), 37–55. 1 Dette, H. (1999). 基于方差估计量之差的回归函数形式相合检验。统计年刊 *27*(3), 1012–1040. 1 Escanciano, J. C. (2006a). 使用投影的回归模型相合诊断检验。*Econome. Theo. 22*(06), 1030–1051. 2 Escanciano, J. C. (2006b). 线性与非线性时间序列模型的拟合优度检验。*J. Am.Statist. Assoc. 101*(474), 531–541. 1, 1 Escanciano, J. C. (2024). 模型检验的高斯过程方法。《统计年鉴》*52*(5), 2456–2481. 1 Escanciano, J. C., J. C. Pardo-Fern´andez, and I. Van Keilegom (2018). 相关数据下半参数回归的渐近分布无关检验。《统计年鉴》*46*(3), 1167– 1196. 1 Fan, J. and L. Huang (2001). 参数回归模型的拟合优度检验。*J. Am. Statist.Assoc. 96*(454), 640–652. 1 Fan, Y. and Q. Li (2000). 相合的模型设定检验：基于核的检验与 Bierens 的 ICM 检验。*Econometric Theory 16*(6), 1016–1041. 1 Fisher, R. A. (1925). 《研究人员统计方法》. Springer. 5.1 Guo, X., T. Wang, and L. Zhu (2016). 广义线性模型的模型检验：一种降维的模型自适应方法。英国皇家统计学会会刊：B 辑 *78*, 1013–1035. 1 Hall, P. and K.-C. Li (1993). 高维数据低维投影的近似线性性。《统计年鉴》, 867–889. 3 H¨ardle, W. K. and E. Mammen (1993). 非参数回归拟合与参数回归拟合的比较。统计年刊 *21*(4), 1926–1947. 1 Horowitz, J. L. and V. Spokoiny (2001). 针对非参数备择检验参数均值回归模型的自适应速率最优检验。*Econometrica 69*(3), 599–631. 1, 1 Jankov´a, J., R. D. Shah, P. B uhlmann, and R. J. Samworth (2020). 高维广义线性模型的拟合优度检验。英国皇家统计学会会刊：*B*辑（统计方法学）*82*(3), 773–795. 1, 2, 6.1, 6.1

Khmaladze, E. V. (1981). 拟合优度检验理论中的鞅方法. 概率论及其应用 *26*, 240–257. 1 Khmaladze, E. V. and H. L. Koul (2004). 回归模型中的鞅变换拟合优度检验. 统计年刊 *32*(3), 995–1034. 1 Khmaladze, E. V. and H. L. Koul (2009). 非参数回归误差的拟合优度问题：分布无关方法. 《统计年鉴》 *37*(6A), 3165–3185. 1 Koenker, R. and Z. Xiao (2002). 分位数回归过程的推断. *Econometrica 70*(4), 1583–1612. 1 Koul, H. L. and P. Ni (2004). 最小距离回归模型检验. *J. Statist. Plan. In-fer. 119*(1), 109–141. 1 Koul, H. L. and W. Stute (1999). 时间序列的非参数模型检验. 《统计年鉴》 *27*(1), 204–236. 1 Kuchibhotla, A. K. and A. Chakrabortty (2022, 06). 高维统计学中超越次高斯性：在协方差估计和线性回归中的应用. 信息与推断：*IMA*期刊 *11*(4), 1389–1456. 2.1 Lavergne, P. and V. Patilea (2008). 打破非参数检验中的维数灾难.*J. Economet.143*, 103–122. 1, 2 Lavergne, P. and V. Patilea (2012). 一人为众，众人为一：具有众多回归变量的回归检验.*J. Busi. Econom. Statist. 30*(1), 41–52. 1 Li, R., W. Zhong, and L. Z. and (2012). 通过距离相关学习进行特征筛选. 美国统计学会会刊 *107*(499), 1129–1139. 6.1 Lin, Q., Z. Zhao, and J. S. Liu (2019). 通过 Lasso 进行稀疏切片逆回归. 美国统计学会会刊 *114*(528), 1726–1739. 6.1 Liu, Y. and J. Xie (2020). 柯西组合检验：一种强效检验，可在任意依赖结构下解析计算p值. 美国统计学会会刊 *115*(529), 393–402. 1, 5.1, 5.1, 5.1, 1 Lu, W., Y. Goldberg, and J. P. Fine (2012). 自适应套索回归对模型设定错误的稳健性. 生物统计学 *99*(3), 717–731. 4 Lu, X. and X. Zheng (2020). 基于鞅变换的 Copula 拟合优度检验.《计量经济学杂志》 *215*(1), 84–117. 1, 1 Patilea, V., C. S´anchez-Sellero, and M. Saumard (2016). 函数型响应的预测变量效应检验. 美国统计学会会刊 *111*(516), 1684–1695. 2 Rao, B. P. (1983). 非参数函数估计. 学术出版社. 3.1 Shah, R. D. and P. B¨uhlmann (2018). 高维线性模型的拟合优度检验. 英国皇家统计学会会刊：*B*辑（统计方法学） *80*(1), 113–135. 1, 6.1, 6.1

Stute, W. (1997). 回归的非参数模型检验. 统计年刊 *25*(2), 613–641. 1 Stute, W., W. Gonz´ález Manteiga, and M. Presedo Quindimil (1998). 回归模型检验中的自助法近似. *J. Am. Statist. Assoc. 93*(441), 141–149. 1 Stute, W., S. Thies, and L. Zhu (1998). 回归模型检验：创新过程方法.统计年刊 *26*(5), 1916–1934. 1, 3, 3.2 Stute, W., W. L. Xu, and L. Zhu (2008). 高维空间中参数回归的模型诊断. 生物统计学 *95*(2), 451–467. 1 Stute, W. and L. Zhu (2002). 广义线性模型的模型检验. 《斯堪的纳维亚统计学杂志》 *29*(3), 535–545. 1, 1, 3, 3.2, 5.2 Tan, F., X. Guo, and L. Zhu (2025). 参数数量发散的回归中的加权残差经验过程、鞅变换和模型设定检验. 《计量经济学杂志》 *252*, 106113. 1 Tan, F., J. Liu, H. Peng, and L. Zhu (2025). 超高维稀疏回归的两步投影拟合优度检验. 1, 5.1, 5.1, 1, 6.1, 6.1, 7 Tan, F. and L. Zhu (2019). 预测变量数量发散的回归的自适应模型检验. 统计年刊 *47*(4), 1960–1994. 1 van der Vaart, A. W. and J. Wellner (1996). 弱收敛与经验过程：及其在统计学中的应用. Springer Science & Business Media. 3.2, 4 Van Keilegom, I., W. Gonz´ález Manteiga, and C. s. S´ánchez Sellero (2008). 基于误差分布估计的参数回归拟合优度检验. *Test 17*, 401–415. 1 Vladimirova, M., S. Girard, H. Nguyen, and J. Arbel (2020). 次威布尔分布：将次高斯和次指数性质推广到重尾分布. *Stat 9*(1), e318. 2.1 Wainwright, M. J. (2019). 高维统计学：非渐近视角. 统计与概率数学剑桥丛书. Cambridge University Press. 2.1 Zheng, J. X. (1996). 通过非参数估计技术进行函数形式的相合检验. *J.Economet. 75*(2), 263–289. 1 Zhu, L. (2003). 回归的降维类型模型检验. *Statist. Sinica 13*, 283–296. 1 Zhu, L. and K. Fang (1996). 切片逆回归核估计的渐近性质. 《统计年鉴》 *24*(3), 1053–1068. 3.1 Zhu, L. and R. Li (1998). 随机回归模型线性性的降维类型检验.*Acta Mathematicae Applicatae Sinica 14*(2), 165–175. 2