arXiv:2305.06234v1 [math.ST] 10 May 2023

利用斯坦方法为多元德尔塔方法中的分布近似建立界

Robert E. Gaunt* 和 Heather Sutcliffe*

摘要

我们为向量统计量（n 个相互独立随机向量的样本均值）的德尔塔方法建立界，以量化其分布近似效果；对于正态和非正态极限分布，我们均使用光滑检验函数进行衡量。对于正态极限分布，我们得到具有最优阶 $n^{-1/2}$ 收敛速度的界；而对于一大类非正态极限分布（其中包括二次型等），我们得到具有更快阶 n^{-1} 收敛速度的界。我们将一般性界应用于多个问题，推导出伯努利方差估计量、若干样本矩统计量的分布近似的显式界；建立一类基于秩的统计量的卡方近似的阶 n^{-1}界；此外，还以高效且独立的方式推导出皮尔逊统计量卡方近似的阶 n^{-1} 界。在建立一般性结果时，我们将近期关于多元正态随机向量函数的斯坦方法结果推广至向量值函数，以及各分量可能存在依赖关系的独立随机向量之和。这些界适用范围广，也具有独立的研究价值。

关键词：Delta 方法；分布逼近；收敛速度；斯坦方法；多元正态随机向量的函数
AMS 2010 学科分类：主要分类 60F05；62E17

1 引言

1.1 Delta 方法及主要结果讨论

设 $\mathrm{X}_1, \ldots, \mathrm{X}_n$ 为均值向量为 0 的独立随机向量。令 $\overline{\mathrm{X}} = n^{-1}\sum_{i=1}^n \mathrm{X}_i$ 表示样本均值。假设 X 的协方差矩阵 $n^{-1}\Sigma \in \mathbb{R}^{d\times d}$ 为非负定矩阵。设函数 $\mathrm{f}\colon \mathbb{R}^d \to \mathbb{R}^m$ 的前 $t+1$ 阶偏导数存在，且满足 $\frac{\partial^t \mathrm{f}}{\prod_{j=1}^t \partial w_{i_j}}(0) \neq 0$，其中某些 $1 \leq i_1, \ldots, i_t \leq d$ 满足该条件；并且（在$t \geq 2$的情况下）f 在均值 0 处的所有低阶偏导数均等于零向量 0。定义 $\mathrm{T}_{t,n} = n^{t/2}(\mathrm{f}(\overline{\mathrm{X}}) - \mathrm{f}(0))$。粗略地说，~~D~~elta 方法

*曼彻斯特大学数学系，Oxford Road，Manchester M13 9PL，英国，
robert.gaunt@manchester.ac.uk;heather.sutcliffe@manchester.ac.uk

断言，对于固定的 $t \geq 1$，统计量 $\mathbf{T}_{t,n}$ 依分布收敛于随机向量

$$\mathbf{Y}_t = \frac{1}{t!} \sum_{i_1,\ldots,i_t=1}^{d} \frac{\partial^t \mathbf{f}}{\prod_{j=1}^{t} \partial w_{i_j}}(\mathbf{0}) Z_{i_1} Z_{i_2} \cdots Z_{i_t}, \tag{1.1}$$

其中，$(Z_1, \ldots, Z_d)^\intercal \sim \mathrm{N}_d(0, \Sigma)$，即均值向量为 0、协方差矩阵为 Σ 的 d 维多元正态分布。在 $t = 1$ 的情况下，极限分布 $\mathbf{Y}_1$ 是一个多元正态随机向量。为便于阐述，我们假设 $\mathbb{E}[\mathbf{X}] = \mathbf{0}$ 。一般情形 $\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$ 的结果可通过简单平移直接得到；例子见第 3 节。

δ 方法是数理统计渐近理论中最基本、应用最广泛的结果之一。关于分布近似的 δ 方法，其表述最早可追溯至 [13] 于 1935 年的一项研究；有关该方法起源的有趣历史讨论，请参阅 [41, 49]。关于 δ 方法的证明、详细论述及应用，可见诸多著作和讲义，例如 [8, 30, 46], ；而 [31, 44] 则论述了将 δ 方法应用于无穷维随机向量的情形。

鉴于德尔塔方法在数理统计中的重要性，文献中却鲜有研究这一自然问题——寻求误差界来量化分布近似的质量——这或许令人意外。不过，近期 [40]（更多结果另见 arXiv 版本 [39] ）开展了一项重要研究，给出了多元德尔塔方法中具有最优阶 $n^{-1/2}$ 的一致与非一致 Berry–Esseen 界（其中 $f\colon\ \mathbb{R}^d \to \mathbb{R}$），针对的是正态近似这一重要情形（$t = 1$）。即便将其专门用于一元德尔塔方法，他们的结果也属全新。[50] 也给出了高维德尔塔方法中的 Berry–Esseen 界。[40] 的界被 [38] 用于建立极大似然估计量正态近似的最优阶 $n^{-1/2}$ 一致与非一致 Berry–Esseen 界；而 [1, 2] 则结合德尔塔方法与斯坦方法[47] ，为可表示为独立随机变量之和函数的极大似然估计量的正态近似，得到了 $O(n^{-1/2})$ 有界 Wasserstein 距离界。近期，[19] 使用斯坦方法，为一元德尔塔方法在情形 $t = 1$、2 下的分布近似推导了界，并采用光滑测试函数进行度量。值得注意的是，[19] 在假设 f 是偶函数（即对所有 $x \in \mathbb{R}$，$f(x) = f(-x)$），且其导数直至四阶均具有多项式增长的条件下，为情形 $t = 2$ 得到了一个阶为 n^{-1} 的界；此时的极限分布为自由度为 1 的卡方分布。

本文利用光滑测试函数度量，获得了多元 delta 方法中分布近似的一般界。我们给出了以下量的界：$|\mathbb{E}[h(\mathbf{T}_{t,\ n})] - \mathbb{E}[h(\mathbf{Y}_t)]|$，其中 $h\colon\ \mathbb{R}^m \to \mathbb{R}$ 是一个取实值且具有适当光滑性的测试函数。关于 $\mathbf{T}_{t,\ n}$ 与 $\mathbf{Y}_t$ 的分布之间距离的界，在积分概率度量 [25, 51] 下可以直接推出；详见备注 2.10。我们的界适用于向量值函数 $\mathbf{f}\colon\ \mathbb{R}^d \to \mathbb{R}^m$ ，其中函数具有充分阶的可微性，且其导数具有多项式增长；样本均值 $\mathbf{X}$ 是相互独立但不一定同分布的 d维随机向量之和；并且这些界对所有 $t \geq 1$都成立。在这一一般情形下，我们得到 $n^{-1/2}$ 阶的界

（见定理 2.1），此外，在某些附加假设下，还可得到具有更快 $O(n^{-1})$ 收敛速率的界。当 $t \geq 2$ 为偶函数，且 f 是偶函数（即对所有 $\mathbf{x} \in \mathbb{R}^d$，均有 $\mathrm{f}(\mathbf{x}) = \mathrm{f}(-\mathbf{x})$），或者对所有$1 \leq i \leq n$、 $1 \leq j$、k、$l \leq d$ 均有 $\mathbb{E}[X_{ij}X_{ik}X_{il}] = 0$（其中 X_{ij} 是随机向量 $\mathbf{X}_i$ 的第 j 个分量）时，会出现这些更快的收敛速率；此外，还需要 f 满足更强的可微性条件以及X_{ij} 满足更强的矩条件（见定理 2.2 和 2.3）。在 delta 方法中识别出这些可导致快于 $O(n^{-1/2})$ 收敛速率的简单充分条件，是本文的主要贡献之一。

[40] 的工作意义重大，因为其中的界是在柯尔莫哥洛夫度量下给出的。这一度量在统计学中十分重要，例如，在该度量下的界可用于构造保守置信区间。然而，在本文中，我们通过使用光滑检验函数获得了多项益处。首先，采用光滑检验函数这一前提对于我们实现快于 $n^{-1/2}$ 阶的收敛速度至关重要；参见备注 2.9。其次，我们可以借助斯坦方法关于多元正态随机向量函数的强大理论，处理正态分布之外的广泛一类极限分布。第三，正如 $[6,$ 第 $151],$ 页所指出的，在理论研究中，针对光滑函数给出的界可能更加自然。我们针对一般 $t \geq 1$ 的单变量 delta 方法所得的 $O(n^{-1/2})$ 界，可直接推出 $O(n^{-1/2})$ Wasserstein 距离界。Wasserstein 距离是一种自然且广泛使用的概率度量，在统计学中有诸多应用（见 $[36]$）。此外，基于光滑检验函数界，我们可以利用 $[21]$ 的一般结果，推导出多变量 delta 方法的柯尔莫哥洛夫距离界（其阶通常并非最优）；详情参见备注 2.10。据我们所知，这是首批适用于一般 $t \geq 1$ 的多变量 delta 方法柯尔莫哥洛夫距离界。

除了补充 $[40],$ 的工作外，我们的结果还大幅推广了 $[19]$ 关于德尔塔方法的结果。在特化到 $t = 2$ 和 $m = 1$的情形时，极限分布为卡方分布。因此，我们的界也补充了 $[20, 22, 23]$ 近期的工作；他们使用斯坦方法，为 Pearson、幂散度和 Friedman 统计量的卡方近似推导出了最优阶 n^{-1} 界。此外，我们的结果也补充了 $[4, 29, 42, 48]$ 的工作；他们使用 Edgeworth 展开，为充分大的 d下 Pearson、幂散度和 Friedman 统计量的 $\chi^2_{(d)}$ 近似推导出了阶为n^{-1} 的 Berry–Esseen 界，但未给出显式常数。值得注意的是，简单应用我们的一般界，就能独立推导出 Friedman 统计量（作为 $[45]$引入的更一般的基于秩的统计量结果的特例）和 Pearson 统计量的显式$O(n^{-1})$ 卡方近似界（见例 3.5 和 3.6）。我们还指出，斯坦方法已被用于推导二次型向量的多元正态近似误差界$[9, 12]$。

1.2 方法讨论与论文结构概述

我们将通过推广$[19, 24]$ 关于多元正态随机向量函数的斯坦方法的近期结果，证明我们关于 delta 方法的一般界。这里我们关注的是如何界定 $g(\mathbf{W}_n)$与 $g(\mathbf{Z})$的分布之间的距离，其中$g:$ $\mathbb{R}^d \to \mathbb{R}$ 是一个连续函数，并且当 $n \to \infty$时，$\mathbf{W}_n$ 依分布收敛于一个。

多元正态随机向量 Z。从现在起，为了简化记号，我们省略 n 在 $\mathbf{W}_n$ 中的依赖关系。在 [19, 24], 中，当 g 具有多项式增长的偏导数，且 $\mathbf{W} = (W_1, \ldots, W_d)^\intercal$，其中 $W_j = n^{-1/2}\sum_{i=1}^{n} X_{ij}$，并且 X_{ij}、$i = 1$、…、n、$j = 1$、…、d 相互独立时，得到了关于量 $|\mathbb{E}[h(g(\mathbf{W}))] - \mathbb{E}[h(g(\mathbf{Z}))]|$ 的一般界。若 g 是偶函数（即，对于所有 $\mathbf{w} \in \mathbb{R}^d$，均有 $g(\mathbf{w}) = g(-\mathbf{w})$），或者若对于所有 i、j，均有 $\mathbb{E}[X_{ij}^3] = 0$，则 [19, 24] 的一般界可得出 $O(n^{-1})$ 收敛速率，此外还需满足通常的一阶和二阶矩假设。若不满足这些附加假设，则可得到阶为 $n^{-1/2}$ 的界。

本文将 [19, 24] 的结果推广到向量值函数 $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^m$，并推广到这样一种情形：对于固定的 i，X_{ij} 可能相互依赖。后一种推广使我们能够为由相互独立的随机向量之和构成的一般样本均值的 delta 方法推导出界，这些随机向量的分量可能相互依赖。我们将这些结果应用于基于秩的统计量和皮尔逊统计量，展示了这种推广的重要性。这些界本身也具有广泛的适用性，并且有独立的研究价值。事实上，针对多元正态随机向量函数的斯坦方法已被用于推导似然比 [3], 幂散度 [20] 和弗里德曼统计量 [23], 的卡方近似界，也被用于推导指数族模型中后验分布 [14] 和最大似然估计量 [1], 的多元正态近似界；本文确立的推广进一步拓宽了该理论的适用范围。有了这些界，我们可以通过应用三角不等式，将问题化简为对形如 $|\mathbb{E}[h(\mathbf{g}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]|$ 的表达式以及另一个仅涉及极限随机变量 $\mathbf{Y}_t$ 的余项进行界定，从而界定感兴趣的量 $|\mathbb{E}[h(\mathbf{T}_{t,\,n})] - \mathbb{E}[h(\mathbf{Y}_t)]|$；后者可以通过初等计算来界定。

本文其余部分安排如下。第 2.1 节给出我们关于 delta 方法中分布近似的一般界，并附上若干讨论这些界的说明。第 2.2 节考察一个值得关注的特殊情形，即 $\mathbf{f}$ 的 $(t+1)$ 阶偏导数为零。这些界将 [19, 24] 近期关于斯坦方法的结果推广到多元正态随机向量的函数、向量值函数，以及分量之间可能相关的独立随机向量之和。在第 3 节，我们给出若干示例，说明如何应用这些一般界。示例涉及伯努利方差估计量的分布近似、若干样本矩统计量，以及对一类基于秩的统计量和皮尔逊统计量进行卡方近似的 n^{-1} 阶界。在第 4 节，我们证明第 2.2 节中的定理；在此过程中，我们得到引理 4.1，正如注释 4.2 中所讨论的，该引理尤为重要，因为它使得可将测试函数为 $h(g(\cdot))$ 时，多元正态斯坦方程解的导数所满足的 [19, 24] 界直接推广到向量值函数 $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^m$。在第 5 节，我们利用第 2.2 节的结果证明第 2.1 节的一般界。在补充材料（SM）中，我们完成示例 3.5 中的一项计算，证明引理 4.3，并给出详细计算，以支持定理 2.1–2.5 证明中的论断。

记号。类 $C_b^t(I)$ 由所有满足 $h^{(t-1)}$ 存在且绝对连续，并且直至第 t 阶的导数均有界的函数 $h : I \subset \mathbb{R} \to \mathbb{R}$ 构成。对于一个

给定函数 $P:\mathbb{R}^d\to\mathbb{R}^+$，类 $C^{t,m}_P(\mathbb{R}^d)$由所有满足以下条件的函数 $\mathbf{g}:\mathbb{R}^d\to\mathbb{R}^m$组成：$\mathbf{g}=(g_1,\dots,g_m)^\intercal$ 的 m 个分量的所有 t阶偏导数均存在，并且对于所有 $\mathbf{w}\in\mathbb{R}^d$，

$$\left|\frac{\partial^k g_l(\mathbf{w})}{\prod_{j=1}^k \partial w_{i_j}}\right|^{t/k}\le P(\mathbf{w}),\quad k=1,\dots,t,\ l=1,\dots,m.$$

对于给定的函数 $P=(P_0,P_1,\dots,P_n)$，类 $\mathcal{C}^{t,n}_{P,m}(\mathbb{R}^d)$ 由所有满足以下条件的函数 $\mathbf{f}:\mathbb{R}^d\to\mathbb{R}^m$ 组成：其各分量 $f_1,\dots,f_m$ 的 $(t+n)$ 阶偏导数均存在，并且对于所有 $\mathbf{w}\in\mathbb{R}^d$，

$$\left|\frac{\partial^{t+k} f_l(\mathbf{w})}{\prod_{j=1}^{t+k}\partial w_{i_j}}\right|\le P_k(\mathbf{w}),\quad k=0,1,\dots,n,\ l=1,\dots,m.$$

我们还记 $h_{p,m}=\sum_{k=1}^p m^k\left\{{p\atop k}\right\}|h|_k$，其中 $|h|_k=\sup_{i_1,\dots,i_k}\left\|\frac{\partial^k h}{\prod_{j=1}^k\partial w_{i_j}}\right\|$，且 $\left\{{n\atop k}\right\}=(1/k!)\sum_{j=0}^k(-1)^{k-j}\binom{k}{j}j^n$ 是第二类斯特林数（见 [35]）。令 $h_p:=h_{p,1}$。有时我们会用记号 $\partial_x=\frac{\partial}{\partial x}$, $\partial_{xy}=\frac{\partial^2}{\partial x\partial y}$ 和 $\partial_x^2=\frac{\partial}{\partial x^2}$ 表示低阶偏导数。对于函数 h 以及随机元素 X 和 Y，我们用 $\Delta_h(X,Y)$ 表示量 $|\mathbb{E}[h(X)]-\mathbb{E}[h(Y)]|$。我们将 $\mathrm{N}(0,\sigma^2)$ 分布的 r 阶绝对矩记为 $\mu_{r,\sigma}=2^{r/2}\sigma^r\Gamma((r+1)/2)/\sqrt{\pi}$。对于 $a,b\in\mathbb{R}$，记作$a\vee b:=\max\{a,b\}$。我们采用约定 $0^0:=1$。

2 主要结果

2.1 Delta 方法的一般界

设 $\mathbf{X}_1,\dots,\mathbf{X}_n$ 为相互独立的零均值随机向量，且 $\mathbf{X}_i=(X_{i,1},\dots,X_{i,d})^\intercal, i=1,\dots,n$。用 $\overline{\mathbf{X}}=n^{-1}\sum_{i=1}^n\mathbf{X}_i$表示样本均值。对于 $j=1,\dots,d$，令$W_j=n^{-1/2}\sum_{i=1}^n X_{ij}$，并记 $\mathbf{W}=(W_1,\dots,W_d)^\intercal$。假设 $\mathbf{W}$ 的协方差矩阵$\Sigma\in\mathbb{R}^{d\times d}$ 是半正定的（因此 $\overline{\mathbf{X}}$ 的协方差矩阵 $n^{-1}\Sigma$ 也为半正定），并令 $\mathbf{Z}\sim\mathrm{N}_d(0,\Sigma)$。令 $\sigma_{jk}=(\Sigma)_{jk}$，并设 $\sigma_j^2=\sigma_{jj}$。假设函数 $\mathbf{f}:\mathbb{R}^d\to\mathbb{R}^m$ 的前 $t+1$ 阶偏导数存在，且满足 $\overline{\partial^t\mathbf{f}/\prod_{j=1}^t\partial w_{i_j}}(\mathbf{0})\ne\mathbf{0}$，其中 $1\le i_1,\dots,i_t\le d$，并且（当 $t\ge2$ 时）$\mathbf{f}$ 在均值 $\mathbf{0}$ 处取值的所有低阶偏导数均等于零向量 $\mathbf{0}$。回顾一下，我们将 $\mathbf{T}_{t,n}=n^{t/2}(\mathbf{f}(\overline{\mathbf{X}})-\mathbf{f}(\mathbf{0}))$ 定义为，并且极限随机向量 $\mathbf{Y}_t$ 按 (1.1) 中的方式定义。我们假设 $\mathbf{f}$ 的各分量的导数具有多项式增长率。为此，设 $\mathbf{f}\in\mathcal{C}^{t,n}_{P,m}(\mathbb{R}^d)$，其中 $P=(P_0,P_1,\dots,P_n)$，且 $P_k(\mathbf{w})=A_{t+k}(1+\sum_{i=1}^d|w_i|^{r_t+k})$，$k=0,1,\dots,n$。令 $a_{n,d,r_t}=d/n^{r_t/2}\vee1$。

定理 2.1. 设随机向量 $\mathbf{X}_1,\dots,\mathbf{X}_n$ 和函数 $\mathbf{f}:\mathbb{R}^d\to\mathbb{R}^m$ 如上所定义。此外，假设 $\mathbf{f}\in\mathcal{C}^{1,2}_{P,m}(\mathbb{R}^d)$ 对于 $t=1$，且 $\mathbf{f}\in\mathcal{C}^{t,1}_{P,m}(\mathbb{R}^d)$ 对于 $t\ge2$。还假设对所有 i,j，均有 $\mathbb{E}|X_{ij}|^{u_{1,t}+3}<\infty$，其中 $u_{1,1}=\max\{3r_1,3r_2/2,r_3\}$, $u_{1,2}=\max\{3(r_2+1)\ r_3\}$，且 $u_{1,t}=3(r_t+t-1)$ 对于 $t\ge3$。令 $C_{1,1}=\max\{4A_1^3,\sqrt{2}A_2^{3/2}/\sqrt{n},A_3/n^{5/6}\}$,

$C_{1,2} = \max\{4A_2^3, \sqrt{2}A_2^{3/2}, A_3/\sqrt{n}\}$ 且 $C_{1,t} = \max\{\frac{4A_t^3}{((t-1)!)^3}, \frac{\sqrt{2}A_t^{3/2}}{((t-2)!)^{3/2}}, \frac{A_t}{(t-3)!}\}$ 对于 $t \geq 3$。假设 $n \geq d^6 \vee 8$。那么，对于 $h \in C_b^3(\mathbb{R}^m)$，

$$\Delta_h(\mathbf{T}_{t,n}, \mathbf{Y}_t) \leq \frac{1}{\sqrt{n}}\big(m|h|_1\mathcal{M}_{1,d} + h_{3,m}\mathcal{M}_{2,d}\big), \tag{2.2}$$

其中

$$\begin{aligned}
\mathcal{M}_{1,d} &= \frac{A_{t+1}d^t}{(t+1)!}\sum_{j=1}^{d}\bigg(\mu_{t+1,\sigma_j} + \frac{d}{n^{r_{t+1}/2}}\mu_{r_{t+1}+t+1,\sigma_j}\bigg),\\
\mathcal{M}_{2,d} &= \frac{C_{1,t}a_{n,d,r_t}^3 d^{3t-2}}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{d}\big((1+2^{u_{1,t}/2}\mu_{u_{1,t},\sigma_k})\mathbb{E}|X_{ij}|^3 + \mathbb{E}|X_{ij}|^{u_{1,t}+3}\\
&\quad + 2^{3u_{1,t}/2}\mathbb{E}|X_{ij}|^3\mathbb{E}|W_k|^{u_{1,t}}\big).
\end{aligned}$$

定理 2.2。设 $\mathbf{X}1, \ldots, \mathbf{X}_n$ 和 $\mathbf{f}: \mathbb{R}^d \to \mathbb{R}^m$ 如上定义。此外，假设 $t \geq 2$ 为偶数，$\mathbf{f}$ 是偶函数（即对所有 $\mathbf{x} \in \mathbb{R}^d$，都有 $\mathbf{f}(\mathbf{x}) = \mathbf{f}(-\mathbf{x})$），且对 $t = 2$，$\mathbf{f} \in \mathcal{C}^{2,4}_{P^m}(\mathbb{R}^d)$，对 $t \geq 4$，$\mathbf{f} \in \mathcal{C}^{t,2}_{P^m}(\mathbb{R}^d)$。还假设对所有 i, j，均有 $\mathbb{E}|X_{ij}|^{u_{2,t}+4} < \infty$，其中 $u_{2,2} = \max\{6(r_2+1)\ 2r_3, 3/2r_4, 6/5r_5, r_6\}$，$u_{2,4} = \max\{6(r_4+3)\ 6r_5/5, r_6\}$，且 $u_{2,t} = 6(r_t + t - 1)$，$t \geq 6$.。令

$$\begin{aligned}
C_{2,2} &= \max\{32A_2^6, 4A_2^3, \tfrac{4A_3^2}{n}, \tfrac{\sqrt{2}A_4^{3/2}}{n^{3/2}}, \tfrac{2^{1/5}A_5^{6/5}}{n^{9/5}}, \tfrac{A_6}{n^2}\},\\
C_{2,4} &= \max\{\tfrac{A_4^6}{1458}, \tfrac{A_4^3}{2}, 2A_4^2, \sqrt{2}A_4^{3/2}, \tfrac{2^{1/5}A_5^{6/5}}{n^{3/5}}, \tfrac{A_6}{n}\},\\
C_{2,t} &= \max\big\{\tfrac{32A_t^6}{((t-1)!)^6}, \tfrac{4A_t^3}{((t-2)!)^3}, \tfrac{2A_t^2}{((t-3)!)^2}, \tfrac{\sqrt{2}A_t^{3/2}}{((t-4)!)^{3/2}}, \tfrac{2^{1/5}A_t^{6/5}}{((t-5)!)^{6/5}}, \tfrac{A_t}{(t-6)!}\big\}, \quad t \geq 6.
\end{aligned}$$

假设 $n \geq 12$。那么，对于 $h \in C_b^6(\mathbb{R}^m)$，

$$\Delta_h(\mathbf{T}_{t,n}, \mathbf{Y}_t) \leq \frac{1}{n}\big(m|h|_1\mathcal{K}_{1,d} + m^2|h|_2\mathcal{K}_{2,d} + h_{4,m}\mathcal{K}_{3,d} + h_{6,m}\mathcal{K}_{4,d}\big), \tag{2.3}$$

其中

$$\begin{aligned}
\mathcal{K}_{1,d} &= \frac{A_{t+2}d^{t+1}}{(t+2)!}\sum_{j=1}^{d}\bigg(\mu_{t+2,\sigma_j} + \frac{d}{n^{r_{t+2}/2}}\mu_{r_{t+2}+t+2,\sigma_j}\bigg),\\
\mathcal{K}_{2,d} &= \frac{d^{2t+1}}{((t+1)!)^2}\sum_{j=1}^{d}\bigg(A_{t+1}^2\mu_{2(t+1),\sigma_j} + \frac{2A_{t+2}^2 d^2}{(t+2)^2 n}\big(\mu_{2(t+2),\sigma_j} + d^2\mu_{2(r_{t+2}+t+2),\sigma_j}\big)\bigg),\\
\mathcal{K}_{3,d} &= \frac{13C_{2,t}a_{n,d,r_t}^6 d^{6t-4}}{12n}\sum_{i=1}^{n}\sum_{j,k=1}^{d}\big((1+2^{u_{2,t}/2}\mu_{u_{2,t},\sigma_k})\mathbb{E}[X_{ij}^4] + \mathbb{E}|X_{ij}|^{u_{2,t}+4}\\
&\quad + 2^{3u_{2,t}/2}\mathbb{E}[X_{ij}^4]\mathbb{E}|W_k|^{u_{2,t}}\big),\\
\mathcal{K}_{4,d} &= \frac{C_{2,t}a_{n,d,r_t}^6 d^{6t-5}}{12n^2}\sum_{i,\alpha=1}^{n}\sum_{j,k,l,a,q=1}^{d}|\mathbb{E}[X_{ij}X_{ik}X_{il}]|\big((1+2\cdot 3^{u_{2,t}/2}\mu_{u_{2,t},\sigma_q})\mathbb{E}|X_{\alpha a}|^3\\
&\quad + \mathbb{E}|X_{\alpha a}|^{u_{2,t}+3} + 12^{u_{2,t}/2}\mathbb{E}|X_{\alpha a}|^3\mathbb{E}|W_q|^{u_{2,t}}\big).
\end{aligned}$$

定理 2.3. 设 X1, . . ., X$_n$ 和 f : $\mathbb{R}^d \to \mathbb{R}^m$（不一定是偶函数）如上定义。进一步假设 $t \geq 2$ $\in \mathcal{C}_P^{2\,2\,m}\,\mathbb{R}^d$ $t =$ 为偶函数, 2 $\in \mathcal{C}_P^{t\,2\,m}\,\mathbb{R}^d$ $t \geq 4$ $\mathbb{E}[X_{ij}X_{ik}X_{il}] = 0$ 且对于 和, f , , ( ); 对于, f , , ( )。还假设对所有 i, j, k, l, 均有 , 且对所有 i, j, 均有 $\mathbb{E}|X_{ij}|^{u_3}$, $_{t+4} < \infty$, 其中 $u_{3,2} =$ max$\{4(r_2 + 1)\ 4r_3/3, r_4\}$ 和 $u_{3,t} = 4(r_t + t - 1)$, $t \geq 4$。令 $C_{3,2} =$ max$\{8A_2^4, 2A_2^2, \underline{2^{1/3}A_3^{4/3}}/\underline{n^{2/3}}, \underline{A_4}/n\}$ 和 $C_{3,t} =$ max$\{$ $^{8A_t^4}((t-1)!)^4$, $^{2A_t^2}((t-2)!)^2$, $\frac{2^{1/3}A_t^{4/3}}{((t-3)!)^{4/3}}$, $^{A_t}(t-4)!\}$, 其中 $t \geq 4$。假设 $n \geq 8$。那么，对于$h \in C_b^4(\mathbb{R}^m)$,

$$\Delta_h(\mathbf{T}_{t,n}, \mathbf{Y}_t) \leq \frac{1}{n}\big(m|h|_1\mathcal{K}_{1,d} + m|h|_2\mathcal{K}_{2,d} + h_{4,m}\mathcal{K}_{5,d}\big), \tag{2.4}$$

其中 $\mathcal{K}_{1,d}$ 和 $\mathcal{K}_{2,d}$ 在定理 2.2 的陈述中定义，且

$$\mathcal{K}_{5,d} = \frac{5C_{3,t}a_{n,d,r_t}^4 d^{4t-2}}{6n}\sum_{i=1}^{n}\sum_{j,k=1}^{d}\big((1 + 2^{u_{3,t}/2}\mu_{u_{3,t},\sigma_k})\mathbb{E}[X_{ij}^4] + \mathbb{E}|X_{ij}|^{u_{3,t}+4} + 2^{3u_{3,t}/2}\mathbb{E}[X_{ij}^4]\mathbb{E}|W_k|^{u_{3,t}}\big).$$

以下定理给出了单变量 delta 方法（$d = m = 1$）中更优的界。这些界适用于更广泛的检验函数类 h，并且所需的矩条件更弱，对函数导数的假设也更弱 f： $\mathbb{R} \to \mathbb{R}$。我们将符号简化如下。对于均值为零的独立随机变量 X_1、.....、 X_n，用 $\overline{X} = n^{-1}\sum_{i=1}^n X_i$ 表示其样本均值，并令 $W = n^{-1/2}\sum_{i=1}^n X_i$。我们假设 Var($W$) $= \sigma^2 \in (0, \infty)$。关于 f 导数的条件： $\mathbb{R} \to \mathbb{R}$ 在均值 0 处现在只需满足 $f^{(t)}(0) \neq 0$, 且（在 $t \geq 2$ 的情况下） $f^{(j)}(0) = 0$, 其中 $j = 1$、.....、 $t - 1$。

定理 2.4. 设随机变量 $X_1, \ldots, X_n$ 和函数 f: $\mathbb{R} \to \mathbb{R}$ 如上所定义。假设对于 $t \geq 1$, $f \in \mathcal{C}^t$, $_{P^{1\,1}}$,($\mathbb{R}$) 和 $\mathbb{E}|X_i|^{u_{4,t}+3} < \infty$, 对所有 i 均成立，其中 $u_{4,t} = r_t + t - 1$。假设 $n \geq 8$。那么，对于 $h \in C_b^1(\mathbb{R})$,

$$\Delta_h(T_{t,n}, Y_t) \leq \frac{\|h'\|}{\sqrt{n}}(\mathcal{M}_{1,1} + \mathcal{M}_3), \tag{2.5}$$

其中 $\mathcal{M}_{1,1}$ 的定义与定理 2.1 的陈述相同，并且

$$\mathcal{M}_3 = \frac{3A_t}{(t-1)!\sigma^2 n}\sum_{i=1}^{n}\big((\alpha_{u_{4,t}} + 2^{u_{4,t}/2}\gamma_{u_{4,t},\sigma})\mathbb{E}|X_i|^3 + 2^{3u_{4,t}/2}\beta_{u_{4,t}}\mathbb{E}|X_i|^3\mathbb{E}|W|^{u_{4,t}} + \beta_{u_{4,t}}\mathbb{E}|X_i|^{u_{4,t}+3}\big),$$

若 $0 \leq r \leq 1$, 则 $(\alpha_r, \beta_r, \gamma_{r,\sigma}) = (4, 4, 2\mu_{r,\sigma})$; 若 $r > 1$, 则 $(\alpha_r, \beta_r, \gamma_{r,\sigma}) = (r + 3, r + 5, \sigma^{-1}$ $(r + 1)\mu_{r+1,\sigma})$。

定理 2.5. 设 X_1, ..., X_n 和 f: $\mathbb{R} \to \mathbb{R}$ 如上定义。假设 $t \geq 2$ 是偶函数，且 f 是偶函数。进一步假设 $f \in \mathcal{C}^t$, $_{P^{2\,1}}$, ($\mathbb{R}$) 和 $\mathbb{E}|X_i|^{u_5,\ t+4} < \infty$

对所有 i，其中 $u_{5,t}=2(r_t+t-1)$。令 $C_{4,t}=\max\{A_t/(t-2)!\ 2A_t^2/((t-1)!)^2\}$。假设$n\geq 12$。那么，对于 $h\in C_b^2(\mathbb{R})$,

$$\Delta_h(T_{t,n},Y_t)\leq\frac{1}{n}\Big(\|h'\|\mathcal{K}_{1,1}+\|h''\|\mathcal{K}_{2,1}+(\|h'\|+\|h''\|)\Big(\frac{13}{10}\mathcal{K}_6+\mathcal{K}_7\Big)\Big),\tag{2.6}$$

其中 $\mathcal{K}_{1,1}$ 和 $\mathcal{K}_{2,1}$ 的定义与定理 2.2 的陈述中相同，且

$$\begin{aligned}\mathcal{K}_6&=\frac{10C_{4,t}}{3\sigma^2n}\sum_{i=1}^n\big((\alpha_{u_{5,t}}+2^{u_{5,t}/2}\gamma_{u_{5,t},\sigma})\mathbb{E}[X_i^4]+2^{3u_{5,t}/2}\beta_{u_{5,t}}\mathbb{E}[X_i^4]\mathbb{E}|W|^{u_{5,t}}\\&\quad+\beta_{u_{5,t}}\mathbb{E}|X_i|^{u_{5,t}+4}\big),\\\mathcal{K}_7&=\frac{3C_{4,t}}{2\sigma^4n^2}\sum_{i,\alpha=1}^n|\mathbb{E}[X_i^3]|\big((\tilde\alpha_{u_{5,t}}+3^{u_{5,t}/2}\tilde\gamma_{u_{5,t},\sigma})\mathbb{E}|X_\alpha|^3+12^{u_{5,t}/2}\tilde\beta_{u_{5,t}}\mathbb{E}|X_\alpha|^3\mathbb{E}|W|^{u_{5,t}}\\&\quad+\tilde\beta_{u_{5,t}}\mathbb{E}|X_\alpha|^{u_{5,t}+3}\big),\end{aligned}$$

其中，$(\tilde\alpha_r,\tilde\beta_r,\tilde\gamma_{r,\sigma})=(10,10,10\mu_{r+1,\sigma})$，当 $0\leq r\leq 1$ 时成立；而 $(\tilde\alpha_r,\tilde\beta_r,\tilde\gamma_{r,\sigma})=(r^2+r+8,r^2+2r+18,\sigma^{-1}(2r^2+r+5)\mu_{r+1,\sigma})$，当 $r>1$ 时成立。

定理 2.6。令 X_1、…、X_n 和 $f\colon\ \mathbb{R}\to\mathbb{R}$（不一定是偶函数）如上定义。进一步假设 $t\geq 2$ 是偶函数，且 $f\in\mathcal{C}^t,\ {}_{2\,1}$,$(\mathbb{R})$。假设对所有 $1\leq i\leq n$，$\mathbb{E}[X_i^3]=0$，且对所有 i，$\mathbb{E}|X_i|^{u_{5,\ t}+4}<\infty$。假设 $n\geq 8$。那么，对于$h\in C_b^2(\mathbb{R})$,

$$\Delta_h(T_{t,n},Y_t)\leq\frac{1}{n}\big(\|h'\|\mathcal{K}_{1,1}+\|h''\|\mathcal{K}_{2,1}+(\|h'\|+\|h''\|)\mathcal{K}_6\big).\tag{2.7}$$

注 2.7. 如果只要求得到显式依赖于样本量 n和维度 d 与 m、但不含显式常数的界，那么在满足定理 2.1–2.6 以及下一节定理 2.11–2.14 的假设时，均可得到简化的界。例如，在定理 2.1 的假设下，存在一个不依赖于 n、d 和 m 的常数 C，使得

$$\Delta_h(\mathbf{T}_{t,n},\mathbf{Y}_t)\leq\frac{Ca_{n,d,r_t}^3d^{3t-1}h_{3,m}}{n^{3/2}}\sum_{i=1}^n\sum_{j=1}^d(1+\mathbb{E}|X_{ij}|^{u_{1,t}+3}).$$

这一简化基于以下几点。首先，由于基本不等式 $|x|^a\leq 1+|x|^b$适用于 $0\leq a<b$，我们有界 $\mathbb{E}|X_{ij}|^3\leq 1+\mathbb{E}|X_{ij}|^{u_{1,t}+3}$。其次，对于一个不依赖于 n的常数 K，有 $\mathbb{E}|W_k|^r\leq K(1+n^{-1}\sum_{i=1}^n\mathbb{E}|X_{ik}|^r)$。对于 $r\geq 2$，这可由 Marcinkiewicz-Zygmund 不等式 *[33]* $\mathbb{E}|\sum_{i=1}^n Y_i|^r\leq C_r\{\sum_{i=1}^n\mathbb{E}|Y_i|^r+(\sum_{i=1}^n\mathbb{E}[Y_i^2])^{1/2}\}$ 以及不等式 $\mathbb{E}|Y|^a\leq 1+\mathbb{E}|Y|^b$推出，其中任意 $0\leq a<b$均成立。对于 $0\leq r<2$，我们只需应用 Hölder 不等式，$\mathbb{E}|W_k|^r\leq(\mathbb{E}[W_k^2])^{r/2}=\sigma_k^r$。第三，根据 Hölder 不等式，我们有 $\mathbb{E}|X_{ij}|^a\mathbb{E}|X_{ik}|^b\leq(\mathbb{E}|X_{ij}|^{a+b})^{a/(a+b)}(\mathbb{E}|X_{ik}|^{a+b})^{b/(a+b)}\leq\max\{\mathbb{E}|X_{ij}|^{a+b},\mathbb{E}|X_{ik}|^{a+b}\}$，其中 a、$b\geq 1$。

注 2.8. 在定理 2.1 中，我们假设 $n\geq d^6$，以便得到更紧凑的最终界。这个假设非常宽松，因为要使定理 2.1 的界 (2.2) 趋于零，我们要求 $n/(d^{6t}m^6)\to\infty$。

若 $r_t \geq 1/(3t)$（因此，对于 $n \geq d^{6t}$，有 $a_{n,d,r_t} = 1$），则当 $n/(d^{6t}m^6) \to \infty$ 时，界 (2.2) 趋于零。若 $0 \leq r_t < 1/(3t)$，则维数的影响会变得更差。例如，当 $r_t = 0$ 时，我们有 $a_{n,d,0} = d$，而要使界 (2.2) 趋于零，我们要求 $n/(d^{6t+6}m^6) \to \infty$。在大多数应用中（第 3 节的所有例子均属此类），我们有 $r_t > 1/(3t)$ 或 $r_t = 0$。在后一种情况下，我们可以应用第 2.2 节的界。此外，若我们能找到 $A_t > 0$ 和 $0 \leq r_t < 1/(3t)$，使得对所有 $\mathbf{w} \in \mathbb{R}^d$ 都有 $|\frac{\partial^t f_t(\mathbf{w})}{\prod_{j=1}^t \partial w_{i_j}}| \leq A_t(1 + \sum_{i=1}^d |w_i|^{r_t})$，那么我们也能找到 $A_t' > 0$和 $r_t' \geq 1/(3t)$，使得对所有 $\mathbf{w} \in \mathbb{R}^d$ 都有 $|\frac{\partial^t f_t(\mathbf{w})}{\prod_{j=1}^t \partial w_{i_j}}| \leq A_t'(1 + \sum_{i=1}^d |w_i|^{r_t'})$，因为 $|w_i|^{r_t} \leq 1 + |w_i|^{r_t'}$。因此，对于当前问题，若维数 d 的影响比矩条件更重要，我们可以取 $r_t \geq 1/(3t)$。定理 2.2 和 2.3 的界也有类似的结论；特别是，即使允许维数 d 和 m 随 n 增长，定理 2.2 和 2.3 中的界仍趋于零，前提是 $n/(d^{6t}m^6) \to \infty$（若 $r_t \geq 1/(3t)$，因此，对于 $n \geq d^{6t}$，有 $a_{n,d,r_t} = 1$）且$n/(d^{4t}m^4) \to \infty$（若 $r_t \geq 1/(2t)$，因此，对于 $n \geq d^{4t}$，有 $a_{n,d,r_t} = 1$）。

注 2.9. 光滑测试函数这一前提至关重要，因为对于非光滑测试函数，一般不成立类似定理 2.2、2.3、2.5 和 2.6 中那样的 $O(n^{-1})$界。为说明这一点，考虑单点测试函数 $h \equiv \chi_{\{0\}}$，并假设$n = 2k$。考虑上述定理中的设定，其中 $f(x) = x^t$，$T_{t,\,n} = n^{t/2}(f(\overline{X}) - f(0)) = n^{t/2}\overline{X}^t$，$Y_t = Z^t$（其中 $Z \sim \mathrm{N}(0,1)$），并令 $X_1, \ldots, X_n$为独立同分布随机变量，且 $\mathbb{P}(X_1 = -1) = \mathbb{P}(X_1 = 1) = 1/2$. 那么

$$\mathbb{E}\big[h\big(n^{t/2}\overline{X}^t\big)\big] = \mathbb{P}\big(n^{t/2}\overline{X}^t = 0\big) = \mathbb{P}\Big(\sum_i X_i = 0\Big) = \binom{2k}{k}\Big(\frac{1}{2}\Big)^{2k} \approx \frac{1}{\sqrt{\pi k}} = \sqrt{\frac{2}{\pi n}},$$

这里用到了斯特林近似。由于 $\mathbb{E}[h(Z^t)] = \mathbb{P}(Z^t = 0) = 0$，$n^{t/2}\overline{X}^t$ 和 Z^t 的分布之间的 Kolmogorov 距离不可能比 $n^{-1/2}$具有更小的阶。

备注 2.10。定理 2.1–2.6 的界可以用决定收敛的积分概率度量（IPM）表示如下。令

$$\mathcal{H}_p = \{h : \mathbb{R}^m \to \mathbb{R} \,:\, h^{(p-1)} \text{ is Lipschitz with } |h|_j \leq 1,\, 1 \leq j \leq p\}.$$

然后，我们将 $\mathbb{R}^m$ 值随机向量 X 和 Y 的分布之间的 IPM 定义为 $d_p(\mathbf{X}, \mathbf{Y}) = \sup_{h \in \mathcal{H}_p} |\mathbb{E}[h(\mathbf{X})] - \mathbb{E}[h(\mathbf{Y})]|$（这里采用了标准的记号简化）。注意，$d_1$对应于 Wasserstein 距离，我们将其记为 d_{W}。此时，由定理 2.4 的界 (2.5) 可得 Wasserstein 距离界

$$d_{\mathrm{W}}(T_{t,n}, Y_t) \leq (\mathcal{M}_{1,1} + \mathcal{M}_3)/\sqrt{n}. \tag{2.8}$$

类似地，在定理 2.1–2.3 的界中令 $|h|_k = 1$, $1 \leq k \leq 6$，使得$h_{3,\,m} = \sum_{k=1}^3 m^k \left\{ {3 \atop k} \right\} = m + 3m^2 + m^3$, $h_{4,\,m} = m + 7m^2 + 6m^3 + m^4$ 和 $h_{6,\,m} = m + 31m^2 + 90m^3 + 65m^4 + 15m^5 + m^6$，即可分别得到 d_3、d_6 和 d_4 度量下的界。

Kolmogorov 距离是由检验函数类 $\mathcal{H}_{\mathrm{K}} = \{\mathbf{1}_{\cdot \leq z} : z \in \mathbb{R}^m\}$ 诱导的积分概率度量（IPM）。我们可以从中提取 Kolmogorov 距离界（通常具有次优阶）。

通过应用 *[21]* 的界，可以得到定理 2.1–2.6 中的界。首先考虑单变量情形。对于 $t \geq 1$，回顾一下，$Y_t =_d f^{(t)}(0)Z^t/t!$，其中 $Z \sim \mathrm{N}(0, \sigma^2)$。记 Y_t 的密度函数为 p_t。随后可以很容易地验证，存在一个常数 $C > 0$，使得 $p_t(x) \leq C|x|^{1/t-1}$，对所有 $x \in \mathbb{R}$ 均成立（例如，参见 *[17, 第 3112]* 页）。因此，应用 *[21]* 的命题 2.1 的第 (iii) 部分可知，存在常数 C' 和 C''，使得

$$d_{\mathrm{K}}(T_{t,n}, Y_t) \leq C'\big(d_{\mathrm{W}}(T_{t,n}, Y_t)\big)^{1/(1+t)}, \tag{2.9}$$

$$d_{\mathrm{K}}(T_{t,n}, Y_t) \leq C''\big(d_2(T_{t,n}, Y_t)\big)^{1/(1+2t)} \tag{2.10}$$

（通过更仔细的分析，可以给出明确常数来代替 C' and C''）。我们注意到，将 $O(n^{-1/2})$ 界 (2.8) 代入不等式 (2.9)，可得到阶为 $n^{-1/(2+2t)}$ 的 Kolmogorov 距离界；而将定理 2.5 和 2.6 的 $O(n^{-1})$ 界（其中 $\|h'\| = \|h''\| = 1$）代入不等式 (2.10)，可得到阶为 $n^{-1/(1+2t)}$ 的 Kolmogorov 距离界。

在一般的多元情形下，若 $t \geq 1$ 且 $d > 1$，我们并不知道极限随机向量 $\mathbf{Y}_t$ 的密度界，因此在如此一般的情形下，无法应用 *[21]* 的界来推出 Kolmogorov 距离界。不过，在 $t = 1$ 的情形下，也就是多元正态近似时，我们可以应用 *[21]* 的命题 2.6，得到以下界。令 $\sigma^2 = \min_{1 \leq j \leq d} \sigma_j^2$。那么，只要 $d_3(\mathbf{T}_{1,n}, \mathbf{Y}_1) \leq (2 + \sqrt{2\log d})/(2\sigma^2)$，便有

$$d_{\mathrm{K}}(\mathbf{T}_{1,n}, \mathbf{Y}_1) \leq 6.17\left(\frac{\sqrt{\log d} + \sqrt{2}}{\sigma^2}\right)^{3/4}\big(d_3(\mathbf{T}_{1,n}, \mathbf{Y}_1)\big)^{1/4}. \tag{2.11}$$

将定理 2.1 的 $O(n^{-1/2})$ 界应用于不等式 (2.11)，可得到阶为 $n^{-1/8}$ 的 Kolmogorov 距离界。我们注意到，若 $(\log d)^3 d^{12}/n \to 0$，则该 Kolmogorov 距离界将趋于零；这与 $t = 1$ 时定理 2.1 的界趋于零所需的条件相同，仅相差一个对数因子。

2.2 $(t + 1)$ 阶偏导数消失的情形

令 \(g\{v1\}\)$_{,n}$(w) $= n^{t/2}$(f($n^{-1/2}$w) $-$ f(0))，并令 \(W $=$ $\sqrt{n}$X\)，则有 \(T$_t$\)$_{,n}$ $=$ g$_{t,n}$(W)。此外，令 \(Z $\sim$ N$_d$(0, Σ)\)。假设 f 的所有 $(t + 1)$ 阶偏导数均等于零向量 0，且满足第 2.1 节中的基本假设，即 ∂^tf$/\overline{\prod_{j=1}^t \partial w_{i_j}}$(0) $\neq$ 0，其中某些 $1 \leq i_1, \ldots, i_t \leq d$ 满足条件（在 $t \geq 2$ 的情况下，f 的所有低阶偏导数在均值 0 处的取值均等于零向量 0）。在这种情况下，g$_{t,n}$(Z) $=_d$ Y$_t$ （这一点可由定理 2.1 的证明直接得出）。因此

$$\Delta_h(\mathbf{T}_{t,n}, \mathbf{Y}_t) = \Delta_h(\mathbf{g}_{t,n}(\mathbf{W}), \mathbf{g}_{t,n}(\mathbf{Z})), \tag{2.12}$$

右端的 (2.12) 式中的量可以通过应用下述定理 2.11–2.14 的界来估计。这些定理中的界用于我们对定理 2.1–2.6 的证明。这些界还具有广泛的适用性，并且本身也具有独立的研究意义。例如，在多元情形 $(d, m \geq 1)$ 下，应用本节中的界，可能会得到对维数 d 具有更优依赖关系的界，而不是应用第 2.1 节中的界所得到的结果；示例 3.5 对此作了说明。

与第 2.1 节一样，设 $\mathbf{X}_1$、……、$\mathbf{X}_n$ 为相互独立的零均值随机向量，其中 $\mathbf{X}_i = (X_{i,1}$、……、$X_{i,d})^{\intercal}$, $i = 1$、……、n。对于 $j = 1$、……、d，令 $W_j = n^{-1/2}\sum_{i=1}^{n} X_{ij}$，并记 $\mathbf{W} = (W_1$、……、$W_d)^{\intercal}$。假设 $\mathbf{W}$ 的协方差矩阵 Σ 为半正定。我们将假设函数 $\mathbf{g}: \mathbb{R}^d \to \mathbb{R}^m$ 的导数具有多项式增长。为此，设 $\mathbf{g} \in C^{t}_{,P^m}(\mathbb{R}^d)$，其中 $P(\mathbf{w}) = A + B\sum_{i=1}^{d} |w_i|^r$，且$A$、$B \geq 0$ 和 $r \geq 0$。

以下定理 2.11–2.14 将定理 3.1 的界推广到向量值函数 $\mathbf{g}: \mathbb{R}^d \to \mathbb{R}^m$，以及各分量可能相依的随机向量。定理 2.12 给出了简单的充分条件，在这些条件下，对于光滑测试函数，可以得到 $O(n^{-1})$ 界。

定理 2.11. 假设上述记号和假设均成立。另假设对所有 i, j，$\mathbb{E}|X_{ij}|^{r+3} < \infty$，且 $\mathbf{g} \in C^{3}_{,P^m}(\mathbb{R}^d)$。设 $n \geq 8$。则对于$h \in C_b^3(\mathbb{R}^m)$,

$$\Delta_h(\mathbf{g}(\mathbf{W}), \mathbf{g}(\mathbf{Z})) \leq \frac{d^2 h_{3,m}}{2n^{3/2}} \sum_{i=1}^{n}\sum_{j,k=1}^{d} \Big[\Big(\frac{A}{d} + 2^{r/2}\mu_{r,\sigma_k} B\Big)\mathbb{E}|X_{ij}|^3 + B\mathbb{E}|X_{ij}|^{r+3} + 2^{3r/2} B\mathbb{E}|X_{ij}|^3 \mathbb{E}|W_k|^r\Big]. \tag{2.13}$$

定理 2.12. 设 X_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, d$ 的定义同定理 2.11，但采用更强的假设：对所有 i, j，$\mathbb{E}|X_{ij}|^{r+4} < \infty$。

(i) 假设 $\mathbf{g} \in C^{6}_{,P^m}(\mathbb{R}^d)$ 是偶函数，且 $n \geq 12$。那么，对于 $h \in C_b^6(\mathbb{R}^m)$,

$$\Delta_h(\mathbf{g}(\mathbf{W}), \mathbf{g}(\mathbf{Z})) \leq \frac{1}{n}\Big(\frac{13}{10} h_{4,m} K_1 + h_{6,m} K_2\Big). \tag{2.14}$$

(ii) 假设 $\mathbf{g} \in C^{4}_{,P^m}(\mathbb{R}^d)$（不一定是偶函数），且对所有 $1 \leq i \leq n$ 和所有 $1 \leq j$、k、$l \leq d$，$\mathbb{E}[X_{ij}X_{ik}X_{il}] = 0$。假设 $n \geq 8$。那么，对于 $h \in C_b^4(\mathbb{R}^m)$,

$$\Delta_h(\mathbf{g}(\mathbf{W}), \mathbf{g}(\mathbf{Z})) \leq \frac{h_{4,m} K_1}{n}. \tag{2.15}$$

Here

$$K_1 = \frac{5d^3}{12n}\sum_{i=1}^{n}\sum_{j,k=1}^{d}\Big[\Big(\frac{A}{d} + 2^{r/2}\mu_{r,\sigma_k} B\Big)\mathbb{E}[X_{ij}^4] + B\mathbb{E}|X_{ij}|^{r+4} + 2^{3r/2} B\mathbb{E}[X_{ij}^4]\mathbb{E}|W_k|^r\Big],$$

$$K_2 = \frac{d^2}{24n^2}\sum_{i,\alpha=1}^{n}\sum_{j,k,l,a,q=1}^{d} |\mathbb{E}[X_{ij}X_{ik}X_{il}]|\Big[\Big(\frac{A}{d} + 2\cdot 3^{r/2}\mu_{r,\sigma_q} B\Big)\mathbb{E}|X_{\alpha a}|^3 + B\mathbb{E}|X_{\alpha a}|^{r+3} + 12^{r/2} B\mathbb{E}|X_{\alpha a}|^3\mathbb{E}|W_q|^r\Big].$$

在单变量情形下（$d = m = 1$），我们可以在对函数 g 和 h的假设较弱的条件下，得到相应的界。我们简化记号，记为 $W = n^{-1/2}\sum_{i=1}^{n} X_i$，其中 $X_1, \ldots, X_n$ 是相互独立的随机变量，满足 $\mathbb{E}[X_i] = 0$, $i = 1, \ldots, n$，且 $\mathbb{E}[W^2] = \sigma^2 \in (0, \infty)$。控制函数也具有更简单的形式 $P(w) = A + B|w|^r$。

定理 2.13. 设 $X_1, \ldots, X_n$ 如上所定义，并假设 $\mathbb{E}|X_i|^{r+3} < \infty$, $1 \le i \le n$。此外，假设 $g \in C^{1,1}_P(\mathbb{R})$，且 $n \ge 8$。那么，对于 $h \in C_b^1(\mathbb{R})$,

$$\Delta_h(g(W), g(Z)) \le \frac{3\|h'\|}{2\sigma^2 n^{3/2}} \sum_{i=1}^{n} \big[(A\alpha_r + 2^{r/2} B\gamma_{r,\sigma})\mathbb{E}|X_i|^3 + 2^{3r/2} B\beta_r \mathbb{E}|X_i|^3 \mathbb{E}|W|^r + B\beta_r \mathbb{E}|X_i|^{r+3}\big], \tag{2.16}$$

其中常数 α_r、β_r、$\gamma_{r,\sigma}$、$r \ge 0$的定义如定理 2.4 的陈述所述 .

定理 2.14。令 X_1, ..., X_n 如上定义，并假设 $\mathbb{E}|X_i|^{r+4} < \infty$、$1 \le i \le n$。此外，假设 $g \in C^{2,1}_P(\mathbb{R})$。

(i) 进一步假设 g 是偶函数，且 $n \ge 12$。那么，对于 $h \in C_b^2(\mathbb{R})$,

$$\Delta_h(g(W), g(Z)) \le \frac{1}{n}(\|h'\| + \|h''\|)\Big(\frac{13}{10}K_3 + K_4\Big). \tag{2.17}$$

（ii）进一步假设对所有 $1 \le i \le n$ 均有 $\mathbb{E}[X_i^3] = 0$，且 $n \ge 8$ (g 不一定是偶函数)。那么，对于 $h \in C_b^2(\mathbb{R})$,

$$\Delta_h(g(W), g(Z)) \le \frac{1}{n}(\|h'\| + \|h''\|)K_3. \tag{2.18}$$

在此，对于 $\tilde{\alpha}_r$、β_r、$\tilde{\gamma}_{r,\sigma}$、$r \ge 0$，其定义如定理 2.5 的陈述所示,

$$K_3 = \frac{5}{3\sigma^2 n} \sum_{i=1}^{n} \big[(A\alpha_r + 2^{r/2} B\gamma_{r,\sigma})\mathbb{E}[X_i^4] + 2^{3r/2} B\beta_r \mathbb{E}[X_i^4]\mathbb{E}|W|^r + B\beta_r \mathbb{E}|X_i|^{r+4}\big],$$

$$K_4 = \frac{3}{4\sigma^4 n^2} \sum_{i,\alpha=1}^{n} |\mathbb{E}[X_i^3]|\big[(A\tilde{\alpha}_r + 3^{r/2} B\tilde{\gamma}_{r,\sigma})\mathbb{E}|X_\alpha|^3 + 12^{r/2} B\tilde{\beta}_r \mathbb{E}|X_\alpha|^3 \mathbb{E}|W|^r + B\tilde{\beta}_r \mathbb{E}|X_\alpha|^{r+3}\big].$$

注 2.15. 如果允许维度 d 和 m 随着 n 增长，且满足 $n/(d^8 m^6) \to \infty$，则定理 2.11 的界 (2.13) 趋于零；而如果分别满足 $n/(d^7 m^6) \to \infty$ 和 $n/(d^5 m^4) \to \infty$，则定理 2.12 的界 (2.14) 和 (2.15) 趋于零。

3 个示例

本节在具体情形中举例说明第 2 节一般性界的应用。例 3.1–3.3 展示了参数值的变化可能导致不同的极限分布，以及具有更快 $O(n^{-1})$ 收敛速率的界的情形。例 3.4 简要说明了我们所得界在向量值统计量中的应用。在例 3.5 中，我们为 [45], 引入的一族基于秩的统计量的卡方逼近，得到了最优阶 n^{-1} 界；作为特例，我们还独立推导出弗里德曼统计量的一个 n^{-1} 阶界。在例 3.6 中，我们也独立推导出皮尔逊统计量卡方逼近的最优阶 n^{-1} 界。为满足第 2 节定理的一项温和要求，我们假设所有例子中 $n \ge 12$ 。

示例 3.1（估计伯努利方差）。考虑参数为 p 的独立同分布伯努利数据 $X_1, \ldots, X_n$。p 的最大似然估计量（MLE）为 $\overline{X} = n^{-1}\sum_{i=1}^n X_i$，因此方差 $p(1-p)$ 的最大似然估计量为 $\overline{X}(1-\overline{X})$。对于 $i = 1, \ldots, n$，令 $V_i = X_i - p$，则 $V_1, \ldots, V_n$ 是均值为零的独立同分布随机变量。记 $\overline{V} = n^{-1}\sum_{i=1}^n V_i$。定义 $f(v) = (p+v)(1-p-v)$，使得 $f(\overline{V}) = \overline{X}(1-\overline{X})$。

首先，假设 $p \neq 1/2$。我们有 $f'(0) = 1 - 2p \neq 0$，因此可以在 $t = 1$ 的条件下应用定理 2.4。此外，对所有 $v \in \mathbb{R}$，$|f'(v)| \leq |1-2p| + 2|v| \leq 2(1+|v|)$ 且 $|f''(v)| = 2 = 1 + |v|^0$，所以我们可以取 $A_1 = 2$，$A_2 = 1$，$r_1 = 1$ 和 $r_2 = 0$。令 $T_{1,n} = \sqrt{n}(\overline{X}(1-\overline{X}) - p(1-p))$，并令 $Y_1 \sim \mathsf{N}(0, \sigma_p^2)$，其中 $\sigma_p^2 = p(1-p)(1-2p)^2$。对于 $m \geq 2$，我们有 $\mathbb{E}|V_i|^m \leq p(1-p)$；由柯西-施瓦茨不等式，$\mathbb{E}|W| \leq (\mathbb{E}[W^2])^{1/2} = \sigma = \sqrt{p(1-p)}$。此外，我们还有 $\mu_{1,\sigma} = \sqrt{2/\pi}\sqrt{p(1-p)}$ 和 $\mu_{2,\sigma} = p(1-p)$。现在，应用界 (2.5) 可得 $h \in C_b^1(\mathbb{R})$，

$$\Delta_h(T_{1,n}, Y_1) \leq 89\|h'\|/\sqrt{n}, \tag{3.19}$$

其中，为了得到一个紧凑的最终界，我们使用了基本不等式 $p(1-p) \leq 1/4$ for $0 < p < 1$，并将最终数值常数向上取整到最近的整数。现在定义 $T'_{1,n} = T_{1,n}/\sigma_p$ 和 $Y'_1 \sim N(0,1)$，使得 $\mathrm{Var}(T'_{1,n}) = 1$。于是我们立即由 (3.19) 推得该界

$$\Delta_h(T'_{1,n}, Y'_1) \leq \frac{89\|h'\|}{\sqrt{np(1-p)(1-2p)^2}},$$

其中 b 在预期条件下，当 p 趋近于 0、$1/2$ 或 1 时上升。

假设现在 $p = 1/2$。那么 $f(v) = (1/4 - v^2)$。我们有 $f'(0) = 0$, $f''(0) = -2$并且对于所有 $1 \leq i \leq n$，$\mathbb{E}[V_i^3] = 0$，因此可以应用定理 2.6，其中 $t = 2$。同样，我们可以取 $A_2 = 1$和 $r_2 = 0$，以及 $A_3 = A_4 = r_3 = r_4 = 0$ ($f_{(3)}(v) = f^{(4)}(v) = 0$)。令 $T_{2,n} = n(\overline{X}(1-\overline{X}) - 1/4)$，并令 $Y_2 = -Z^2/4$ 服从 N(0, 1) 分布。我们有 $\mathbb{E}|V_i|^m = 2^{-m}$ 对于$m \geq 1$，且 $\mathbb{E}[W^2] = 1/4$。然后，应用界 (2.7)，可得，对于 $h \in C_b^2((-\infty, 0))$，

$$\Delta_h(T_{2,n}, Y_2) \leq 39(\|h'\| + \|h''\|)/n.$$

该界具有更快的 $O(n^{-1})$ 收敛速率。我们也可以对$t = 2$ 应用定理 2.4，从而得到一个阶为 $n^{-1/2}$ 的界，该界适用于更广泛的 $C_b^1((-\infty, 0))$ 测试函数类。

例 3.2（$\overline{X}^p$的渐近分布）。设 X_1，.....，X_n是均值为 μ、方差为 $\sigma^2 \in (0, \infty)$的独立同分布随机变量。定义 $\overline{X} = n^{-1}\sum_{i=1}^n X_i$，并令 $p \geq 2$。对于 $i = 1$，.....，n，令 $V_i = X_i - \mu$，于是 V_1，.....，V_n是均值为零的独立同分布随机变量。记$\overline{V} = n^{-1}\sum_{i=1}^n V_i$。定义 $f(v) = (\mu + v)^p$，于是 $f(\overline{V}) = \overline{X}^p$。

首先，设 $\mu \neq 0$。我们有 $f'(0) = p\mu^{p-1} \neq 0$，因此可以对 $t = 1$应用定理 2.4。此外，对于 $v \in \mathbb{R}$,

$$|f'(v)| = p|(\mu+v)^{p-1}| \leq 2^{p-2}p(|\mu|^{p-1} + |v|^{p-1}) \leq b_{\mu,p}(1 + |v|^{p-1}),$$

其中 $b_{\mu,p} := 2^{p-2}p(1 \vee |\mu|^{p-1})$，类似地，我们可以得到界 $|f''(v)| \leq c_{\mu,p}(1 + |v|^{p-2})$，$v \in \mathbb{R}$，其中 $c_{\mu,p} := 2^{p-3}p(p-1)(1 \vee |\mu|^{p-2})$。因此，我们可以取 $A_1 = b_{\mu,p}$,

$A_2 = c_{\mu,p}$, $r_1 = p-1$ 和 $r_2 = p-2$。令 $T_{1,n} = \sqrt{n}(\overline{X}^p - \mu^p)$，且 $Y_1 \sim$ 服从 N(0, $(p\sigma)^2\mu^{2p-2}$)。于是，应用不等式 (2.5)，对于 $h \in C_1^b(\mathbb{R})$，我们得到

$$\Delta_h(T_{1,n}, Y_1) \le \frac{\|h'\|}{\sqrt{n}}\left[\frac{c_{\mu,p}}{2}\left(\sigma^2 + \frac{\mu_{p,\sigma}}{n^{p/2-1}}\right) + \frac{3(p+4)b_{\mu,p}}{\sigma^2}\left(\left(1 + \frac{2^{(p-1)/2}\mu_{p,\sigma}}{\sigma}\right)\mathbb{E}|V_1|^3 + 2^{3(p-1)/2}\mathbb{E}|V_1|^3\mathbb{E}|W|^{p-1} + \mathbb{E}|V_1|^{p+2}\right)\right].$$

现在假设 $\mu = 0$。那么 $f(v) = v^p$。我们有 $f'(0) = \cdots = f_{(p-1)}(0) = 0$ 且 $f_{(p)}(0) = p! \neq 0$，因此可以在 $t = p$ 的条件下应用定理 2.4。我们还有 $f^{(p)}(v) = p! = (p!/2)(1 + |v|^0)$，且 $f^{(p-1)}(v) = 0$，因此可以取 $A_p = p!/2$, $A_{p+1} = r_p = r_{p+1} = 0$。令 $T_{p,n} = n^{p/2}\overline{X}^p$ 且 $Y_p = Z^p$，其中 $Z \sim \mathrm{N}(0, \sigma^2)$。于是，我们得到

$$\Delta_h(T_{p,n}, Y_p) \le \frac{3p(p+4)\|h'\|}{2\sigma^2\sqrt{n}}\left[\left(1 + \frac{2^{(p-1)/2}\mu_{p,\sigma}}{\sigma}\right)\mathbb{E}|X_1|^3 + 2^{3(p-1)/2}\mathbb{E}|X_1|^3\mathbb{E}|W|^{p-1} + \mathbb{E}|X_1|^{p+2}\right].$$

如果我们进一步假设 $p \ge 2$ 是偶数，那么在对测试函数 h 作出更强假设的前提下，可以得到一个具有更快 $O(n^{-1})$ 收敛速度的界。这是因为，如果 $p \ge 2$ 是偶数，那么 $f(v) = v^p$ 是偶函数，因此可以在 $t = p$ 的条件下应用定理 2.5。我们可以取 $A_p = p!/2$ 和 $A_{p+1} = A_{p+2} = r_p = r_{p+1} = r_{p+2} = 0$。现在应用不等式 (2.6)，可得对于 $h \in C_b^2(\mathbb{R}^+)$,

$$\Delta_h(T_{p,n}, Y_p) \le \frac{p^2 h_2}{\sigma^2 n}\left\{\frac{13(2p+3)}{6}\left[\left(1 + \frac{2^{p-1}\mu_{2p-1,\sigma}}{\sigma}\right)\mathbb{E}[X_1^4] + 2^{3(p-1)}\mathbb{E}[X_1^4]\mathbb{E}|W|^{2p-2} + \mathbb{E}|X_1|^{2p+2}\right] + \frac{3(4p^2-7p+12)}{2\sigma^2}|\mathbb{E}[X_1^3]|\left[\left(1 + \frac{3^{p-1}\mu_{2p-1,\sigma}}{\sigma}\right)\mathbb{E}|X_1|^3 + 12^{p-1}\mathbb{E}|X_1|^3\mathbb{E}|W|^{2p-2} + \mathbb{E}|X_1|^{2p+1}\right]\right\}.$$

例 3.3（两个样本矩乘积的渐近分布）。设 $X_{1,j}, \ldots, X_{n,j}$, $j = 1, 2$, 为均值为 $\mathbb{E}[X_{1j}] = \mu_j$ 且方差 $\mathrm{Var}(X_{1j}) = \sigma_j^2 \in (0, \infty)$ 的独立同分布随机变量。定义 $\overline{X}_j = n^{-1}\sum_{i=1}^n X_{ij}$, $j = 1, 2$。对于 $i = 1, \ldots, n$，令 $V_{ij} = X_{ij} - \mu_j$，则 $V_{1,j}, \ldots, V_{n,j}$, $j = 1, 2$, 是均值为零的独立同分布随机变量。记作 $\overline{V}_j = n^{-1}\sum_{i=1}^n V_{ij}$, $j = 1, 2$，并令 $\mathbf{V} = (\overline{V}_1, \overline{V}_2)^\intercal$。定义 $f(\mathbf{v}) = (\mu_1 + v_1)(\mu_2 + v_2)$，于是 $f(\mathbf{V}) = X_1X_2$。

首先，假设 $(\mu_1, \mu_2)^\intercal \neq \mathbf{0}$。我们有 $(\partial_{v_1}f(\mathbf{0})\ \partial_{v_2}f(\mathbf{0}))^\intercal = (\mu_2, \mu_1)^\intercal \neq \mathbf{0}$，因此可以取 $t = 1$ 应用定理 2.1。此外，$\partial_{v_1}f(\mathbf{v}) = \mu_2 + v_2$，$\partial_{v_2}f(\mathbf{v}) = \mu_1 + v_1$，且 $\partial_{v_1v_2}f(\mathbf{v}) = 1$，其余偏导数均为零。对于 $j = 1, 2$，我们有 $|\partial_{v_j}f(\mathbf{v})| \le c_{\mu_1,\mu_2}(1 + |v_1| + |v_2|)$，其中 $\mathbf{v} \in \mathbb{R}^2$，且 $c_{\mu_1,\mu_2} = \max\{1, |\mu_1|, |\mu_2|\}$。我们还有 $|\partial_{v_jv_k}f(\mathbf{v})| \le 1 = (1/3)(1 + |v_1|^0 + |v_2|^0)$，$1 \le j, k \le 2$。因此，我们可以取 $A_1 = c_{\mu_1,\mu_2}$, $A_2 = 1/3$, $r_1 = 1$ 和 $A_3 = r_2 = r_3 = 0$。令 $T_{1,n} = \sqrt{n}(\overline{X}_1\overline{X}_2 - \mu_1\mu_2)$，并且

$Y_1 \sim \mathrm{N}(0, (\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2)$。现在应用不等式 (2.2)，可得对于 $h \in C_b^3(\mathbb{R})$，

$$\Delta_h(T_{1,n}, Y_1) \le \frac{1}{\sqrt{n}}\Bigg\{ \|h'\|(\sigma_1^2+\sigma_2^2) + 8c_{\mu_1,\mu_2}^3 h_3 \sum_{j,k=1}^{2}\Bigg[\left(1+\frac{8\sigma_k^3}{\sqrt{\pi}}\right)\mathbb{E}|V_{1j}|^3 + \mathbb{E}[V_{1j}^6] + 16\sqrt{2}\mathbb{E}|V_{1j}|^3\left(3\sigma_k^4 + \frac{\mathbb{E}[V_{1k}^4]}{n}\right)^{3/4}\Bigg]\Bigg\},$$

其中，对于 $W_k = n^{-1/2}\sum_{i=1}^n V_{ik}$，我们使用 Hölder 不等式对 $\mathbb{E}|W_k|^3 \le (\mathbb{E}[W_k^4])^{3/4} \le (3\sigma_k^4 + \mathbb{E}[V_{1k}^4]/n)^{3/4}$ 进行界定。最后一个不等式由一项标准计算得出，该计算利用了以下事实：$V_{1,k}, \ldots, V_{n,k}$ 是均值为零、方差为 σ_k^2 的独立同分布随机变量。

现在假设 $\mu_1 = \mu_2 = 0$。那么 $f(\mathbf{v}) = v_1 v_2$ 是一个偶函数。我们还有 $(\partial_{v_1} f(\mathbf{0})\ \partial_{v_2} f(\mathbf{0}))^\intercal = \mathbf{0}$ 且 $\partial_{v_1 v_2} f(\mathbf{0}) = 1 \neq 0$，因此可以在 $t = 2$ 条件下应用定理 2.2。我们可以取 $A_2 = 1/3$, $A_3 = \cdots = A_6 = 0$ 和 $r_2 = \cdots = r_6 = 0$（请注意，所有二阶以上的偏导数都等于零）。令 $T_{2,n} = n\overline{X}_1\overline{X}_2$ 和 $Y_2 = Z_1 Z_2/2$，其中 $Z_1 \sim \mathrm{N}(0, \sigma_1^2)$ 和 $Z_2 \sim \mathrm{N}(0, \sigma_2^2)$ 相互独立。极限随机变量 Y_2 服从方差伽马分布，其密度为 $p(x) = (2/(\pi s))K_0(2|x|/s)$，$x \in \mathbb{R}$，其中 $s = \sigma_1\sigma_2$ 且 $K_0(x) = \int_0^\infty \mathrm{e}^{-x\cosh(t)}\,\mathrm{d}t$ 是第二类修正贝塞尔函数（参见 *[18, 定理 1]*）。令 $W_k = n^{-1/2}\sum_{i=1}^n X_{ik}$, $k = 1, 2$。现在应用界 (2.3)，可得对于 $h \in C_b^6(\mathbb{R})$，

$$\Delta_h(T_{2,n}, Y_2) \le \frac{1}{n}\Bigg[2630h_4 \sum_{j,k=1}^{2}\left((1+120\sigma_k^6)\mathbb{E}[X_{1j}^4] + \mathbb{E}[X_{1j}^{10}] + 512\mathbb{E}[X_{1j}^4]\mathbb{E}[W_k^6]\right) + 102h_6 \sum_{j,a,q=1}^{2}|\mathbb{E}[X_{1j}^3]|\left((1+810\sigma_q^6)\mathbb{E}|X_{1a}|^3 + \mathbb{E}|X_{1a}|^9 + 1728\mathbb{E}|X_{1a}|^3\mathbb{E}[W_q^6]\right)\Bigg].$$

如果我们进一步假设，对所有 i、j，$\mathbb{E}[X_{ij}^3] = 0$，那么可以应用定理 2.3，得到一个在较弱假设 $h \in C_b^4(\mathbb{R})$ 下仍成立的界，并且只要求 $X_{1,j}$, $j = 1$ 的八阶矩存在，2。

例 3.4（样本均值与样本方差的渐近联合分布）。设 $X_1, \ldots, X_n$ 为均值为 μ、方差为 $\sigma^2 \in (0, \infty)$ 的独立同分布随机变量。对于 $i = 1, \ldots, n$，令 $V_{i,1} = X_i - \mu$，使得 $V_{1,1}, \ldots, V_{n,1}$ 的均值为零。此外，对于 $i = 1, \ldots, n$，令 $V_{i,2} = V_{i,1}^2 - \sigma^2$，使得 $V_{1,2}, \ldots, V_{n,2}$ 的均值为零。记 $\overline{V}_j = n^{-1}\sum_{i=1}^n V_{ij}$, $j = 1, 2$，并令 $\mathbf{V} = (\overline{V}_1, \overline{V}_2)^\intercal$。令 $\mathbf{f}(\mathbf{v}) = (f_1(\mathbf{v})\ f_2(\mathbf{v}))^\intercal$，其中 $f_1(\mathbf{v}) = \mu + v_1$ 且 $f_2(\mathbf{v}) = \sigma^2 + v_2 - v_1^2$，于是 $\mathbf{f}(\mathbf{V}) = (\overline{X}, n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2)^\intercal$。我们有 $(\partial_{v_1} f_1(\mathbf{0})\ \partial_{v_2} f_1(\mathbf{0}))^\intercal = (1, 0)^\intercal$ 且 $(\partial_{v_1} f_2(\mathbf{0})\ \partial_{v_2} f_2(\mathbf{0}))^\intercal = (0, 1)^\intercal$，因此可以取 $t = 1$ 应用定理 2.1。此外，$\partial_{v_1} f_1(\mathbf{v}) = 1$, $\partial_{v_2} f_1(\mathbf{v}) = 0$, $\partial_{v_1} f_2(\mathbf{v}) = -2v_1$, $\partial_{v_2} f_2(\mathbf{v}) = 1$，且 $\partial_{v_1}^2 f_2(\mathbf{v}) = -2$，其余偏导数均为零。按照通常的论证，我们因此可以取 $A_1 = 2$, $A_2 = 2/3$, $r_1 = 1$ 和 $A_3 = r_2 = r_3 = 0$。令 $\mathbf{T}_{1,n} = \sqrt{n}(\overline{X} - \mu, n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2 - \sigma^2)^\intercal$ 且 $\mathbf{Y}_1 \sim \mathrm{N}_2(\mathbf{0}, \Sigma)$，其中 $(\Sigma)_{ij} = \sigma_{ij}$，满足 $\sigma_{11} = \sigma^2$, $\sigma_{12} = \sigma_{21} = \mathbb{E}[(X_1 - \mu)^3] = \mu_3$ 且 $\sigma_{22} = \mathrm{Var}((X_1 - \mu)^2) = \mu_4 - \sigma^4$；$\mu_3$ 和 μ_4 分别表示 X_1 的三阶和四阶中心矩。于是，应用不等式 (2.2) 可得，对于 $h \in C_b^3(\mathbb{R}^2)$，

$$\Delta_h(\mathbf{T}_{1,n}, \mathbf{Y}_1) \le \frac{4}{\sqrt{n}}\Bigg\{ |h|_1(\mu_4 + \sigma^2 - \sigma^4) + 16h_{3,2}\sum_{j,k=1}^{2}\Bigg[\left(1+\frac{8\sigma_{kk}^{3/2}}{\sqrt{\pi}}\right)\mathbb{E}|V_{1j}|^3$$

$$+\mathbb{E}[V_{1j}^6]+16\sqrt{2}\mathbb{E}|V_{1j}|^3\left(3\sigma_{kk}^2+\frac{\mathbb{E}[V_{1k}^4]}{n}\right)^{3/4}\Bigg]\Bigg\},$$

其中，我们采用与例 3.3 中类似的方法对该界进行了简化。

我们指出，本例中使用的方法可用于获得矩估计量$n^{-1}\sum_{i=1}^n(X_i-\overline{X})^p$的正态近似的 $O(n^{-1/2})$ Wasserstein 距离界 $p\geq 2$，以及各种矩估计量联合分布的多元正态近似的 $O(n^{-1/2})$ 界。

例 3.5（基于秩的统计量）。假设我们有 $r\geq 2$ 个处理，分布在 n 个相互独立的试验中；对于第 i 次试验，排名为 $\pi_i(1),\ldots,\pi_i(r)$，其中 $\pi_i(j)\in\{1,\ldots,r\}$，涵盖这 r 个处理。在原假设下，这些排名是相互独立的排列 $\pi_1,\ldots,\pi_n$，且每个排列出现的可能性相同。固定一个函数$J\colon\{1,\ldots,r\}\to\mathbb{R}$，并令

$$\overline{J}=\frac{1}{r}\sum_{k=1}^r J(k),\quad \sigma_J^2=\frac{1}{r-1}\sum_{k=1}^r(J(k)-\overline{J})^2.$$

对于 $i=1,\ldots,n$ 和 $j=1,\ldots,r$，令 $X_{ij}^J=(J(\pi_i(j))-\overline{J})/\sigma_J$ 和 $W_j^J=n^{-1/2}\sum_{i=1}^n X_{ij}^J$。随后，Sen *[45]* 提出了以下检验统计量

$$S_{n,r}^J=\sum_{j=1}^r(W_j^J)^2,$$

并证明在原假设下，它渐近地服从 $\chi^2_{(r-1)}$ 分布。

弗里德曼统计量 *[15]*，记为 $F_{n,\,r}$，对应于情形 $J(k)=k$，在此情形下$\overline{J}=(r+1)/2$ 且 $\sigma_J^2=r(r+1)/12$。另一个特殊情形是 Brown-Mood 统计量，由 *[7]* 提出，作为弗里德曼统计量的替代。对于固定的 $a\in\{1,\ldots,r-1\}$，该统计量对应于情形 $J(k)=\mathbf{1}(k\leq a)$，并且有$\overline{J}=a/r$ 和 $\sigma_J^2=a(r-a)/(r(r-1))$。

*[42]*推导了 Kolmogorov 距离界，用于量化 Sen 的基于秩的统计量族的卡方近似，推导时假设对所有 $k=1$，.....，r，都有 $|J(k)-\overline{J}|\leq B$，其中 B 为某个常数（可能依赖于 r）。请注意，Friedman 统计量和 Brown-Mood 统计量都满足这一假设。对于情形 $r\geq 6$，*[42]* 推导了一个阶为 n^{-1}（不显式依赖于 r）的 Kolmogorov 距离界，并为 $2\leq r\leq 5$ 推导了一个收敛速率阶为 $n^{-1+1/r}$ 的界。在本例中，我们为光滑测试函数得到了 $O(n^{-1})$ 界，且该界对所有 $r\geq 2$ 均成立。

定义 $\overline{X}_j=n^{-1}\sum_{i=1}^n X_{ij}$，其中我们在记号中省略了上标 J。对于固定的 $j=1,\ldots,r$，随机变量 $\pi_1(j),\ldots,\pi_n(j)$ 是具有 $\{1,\ldots,r\}$ 上均匀分布的独立同分布随机变量。因此，对于固定的 j，随机变量 $X_{1j},\ldots,X_{nj}$ 也是独立同分布的。注意，我们有如下表示：$S_{n,\,r}^J=nf(\mathbf{X})$，其中 $f(\mathbf{x})=\sum_{j=1}^r x_j^2$（且 $f(\mathbf{0})=0$）。对于 $j=1,\ldots,r$，我们还有 $\partial_{x_j}f(\mathbf{x})=2x_j$ 以及 $\partial^2_{x_j}f(\mathbf{x})=2$，其余偏导数均为零。特别地，对于 $j=1,\ldots,r$，$\partial_{x_j}f(\mathbf{0})=0$ 且 $\partial^2_{x_j}f(\mathbf{0})=2$。此外，$\mathbf{W}=(W_1,\ldots,W_r)^\intercal$ 的协方差矩阵是半正定的，其元素为 $\sigma_{jj}=(r-1)/r$ 和 $\sigma_{jk}=-1/r$, $j\neq k$（见补充材料）。此外，f 是一个

偶函数。因此，我们可以应用定理 2.2，其中 $t = 2$。显然，可以取 $A_2 = 2$, $r_2 = 1/6$, $A_3 = \cdots A_6 = r_3 = \cdots = r_6 = 0$（我们取 $r_2 = 1/6$，以获得对参数 r 依赖性更优的界；见注 2.8）。我们立即得到界 $\mathbb{E}|X_{i,j}|^k \leq (B\sigma_J)^k$，对所有 i, j 和 $k \geq 1$ 均成立；再应用 Marcinkiewicz-Zygmund 不等式 *[33]*，可得对于某个常数 K，$\mathbb{E}|W_j|^k \leq K(B/\sigma_J)^k$。令 $V_r \sim \chi^2_{(r-1)}$。现在，应用不等式 (2.3)，对于 $h \in C_b^6(\mathbb{R}^+)$,

$$\Delta_h(S^J_{n,r}, V_r) \leq Cr^{12}(B/\sigma_J)^{13}h_6/n, \tag{3.20}$$

对于某个通用常数 $C > 0$，为得到紧凑的最终界，我们利用了$1 \leq B/\sigma_J$。

由于 f 的所有三阶偏导数都等于零，我们还可以应用定理 2.12 的第 (i) 部分，得到另一个界。我们可以写成 $S^J_{n,\,r} = g(\mathbf{w})$，其中 $g(\mathbf{w}) = \sum_{j=1}^r w_j^2$。对于 $k = 1, \ldots, r$，有 $\partial_{w_k} g(\mathbf{w}) = 2w_k$, $\partial^2_{w_k} g(\mathbf{w}) = 2$，且所有其他导数均等于 0。现在，$|\partial_{w_k} g(\mathbf{w})|^6 = 64w_k^6$ 且 $|\partial^2_{w_k} g(\mathbf{w})|^3 = 8$，这意味着我们可以取 $P(\mathbf{w}) = 8 + 64\sum_{j=1}^r w_j^6$作为控制函数。将界 (2.14) 应用于$A = 8$, $B = 64$ 和 $r = 6$，可得，对于 $h \in C_b^6(\mathbb{R}^+)$,

$$\Delta_h(S^J_{n,r}, V_r) \leq C'r^7(B/\sigma_J)^{12}h_6/n, \tag{3.21}$$

对于某个通用常数 $C' > 0$。与界（3.20）相比，界（3.21）对处理数量 r 的依赖性更优。

对于弗里德曼统计量，随机变量 X_{ij} 的分布关于 0 对称，因此对于所有 $1 \leq i \leq n$，$\mathbb{E}[X_{ij}X_{ik}X_{il}] = 0$，且 $1 \leq j,\ k,\ l \leq r$。因此，我们可以应用定理 2.12 的第 (i) 部分，得到一个更优的界。我们有 $|\partial_{w_k} g(\mathbf{w})|^4 = 16w_k^4$ 和 $|\partial^2_{w_k} g(\mathbf{w})|^2 = 4$，因此可以取 $P(\mathbf{w}) = 4 + 16\sum_{j=1}^r w_j^4$ 作为控制函数。将 $A = 4$、$B = 16$ 和 $r = 4$ 代入不等式 (2.15)，可得，对于 $h \in C_b^4(\mathbb{R}^+)$,

$$\Delta_h(F_{n,r}, V_r) \leq C''r^5h_4/n,$$

对于某个通用常数 $C'' > 0$。这优于将界（3.21）特化到弗里德曼统计量情形时所得到的界 Cr^7h_6/n。然而，其对 r 的依赖性劣于 *[23,* 定理1中的界*1]*，后者对于光滑检验函数同样具有最优的$O(n^{-1})$收敛速率，但仅在满足最优条件$n/r \to \infty$时才趋于零。可以预期，*[23]* 专门针对弗里德曼统计量的卡方近似所采用的方法能够得到更好的界，尽管值得注意的是，作为我们一般性界的推论，我们能够以如此高的效率推导出具有最优 $O(n^{-1})$ 收敛速率的界。

例 3.6（皮尔逊统计量）。考虑 n 次相互独立的试验，每次试验的结果都被唯一归入 r 个类别之一。设 $p_1, \ldots, p_r$ 为非零的分类概率，并令 $(U_1, \ldots, U_r)$ 表示各类别中观测到的数量。则用于拟合优度检验的皮尔逊卡方统计量 *[37],* 为

$$\chi^2 = \sum_{j=1}^r \frac{(U_j - np_j)^2}{np_j},$$

在原假设下渐近服从 $\chi^2_{(r-1)}$。

单元格计数 $U_j \sim \mathrm{Bin}(n, p_j)$，$1 \le j \le r$ 是满足 $\sum_{j=1}^r U_j = n$ 的相关随机变量。令 $I_{ij} \sim \mathrm{Ber}(p_j)$ 表示第 i 次试验落入第 j 个单元格的指示变量。通过令 $X_{ij} = (I_{ij} - p_j)/\sqrt{p_j}$ and $W_j = n^{-1/2}\sum_{i=1}^n X_{ij}$，我们可以将 $\chi^2 = g(\mathbf{W}) = \sum_{j=1}^r W_j^2$ 写成，其中 $g(\mathbf{w}) = \sum_{j=1}^r w_j^2$。与基于秩的统计量的情况一样，函数 g 是偶函数，并且具有多项式增长的导数。此外，由于假设各次试验相互独立，因此对于固定的 j，随机变量 $X_{1,\,j}, \ldots, X_{n,\,j}$ 相互独立。$\mathbf{W} = (W_1, \ldots, W_r)^\intercal$ 的协方差矩阵是半正定的，其元素为 $\sigma_{jj} = 1 - p_j$ 和 $\sigma_{jk} = -\sqrt{p_j p_k}$, $j \ne k$ （参见 *[22]*）。因此，皮尔逊统计量属于定理 2.12 第 (i) 部分所涵盖的统计量类别。

与基于秩的统计量的情形一样，$|\partial_{w_k} g(\mathbf{w})|^6 = 64 w_k^6$ 和 $|\partial^2_{w_k} g(\mathbf{w})|^3 = 8$，因此我们可以取 $P(\mathbf{w}) = 8 + 64\sum_{j=1}^r w_j^6$ 作为控制函数。将不等式（2.14）应用于 $A = 8$、$B = 64$ 和 $r = 6$，可得到如下形式的界：

$$\Delta_h(\chi^2, V_r) \le C n^{-1} h_6,$$

其中，$C > 0$ 是一个依赖于 r 和 p_1、……、p_r，但不依赖于 n 的常数。我们不明确求出这样的 C，因为 *[22]* 得到了一个形式为 $Kr(np_*)^{-1}\sum_{k=0}^5 \|h^{(k)}\|$ 的更优上界，其中 $p_* = \min_{1\le j\le r} p_j$ 和 $h^{(0)} \equiv h$。与应用定理 2.12 所得到的界相比，这个界对 r 和 p_1、……、p_r 的依赖性要好得多。与弗里德曼统计量的情形一样，预计 *[22]* 的界会优于我们的界，因为他们的方法专门针对皮尔逊统计量，而非本文所考虑的更一般的统计量类别。

4 多元正态随机向量函数的斯坦方法

本节首先概述如何运用斯坦方法证明逼近定理，其中极限分布可以表示为多元正态随机向量的函数（更多详情见 [19]）。设 $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^m$ 连续，且 $\mathbf{Z} \sim \mathrm{N}_d(\mathbf{0}, \Sigma)$，其中协方差矩阵 Σ 为半正定。假设我们想要界定 $\mathbf{g}(\mathbf{W})$ 与 $\mathbf{g}(\mathbf{Z})$ 的分布之间的距离。考虑测试函数为 $h(\mathbf{g}(\cdot))$ 的多元正态斯坦方程 [5, 28]:

$$\nabla^\intercal \Sigma \nabla f(\mathbf{w}) - \mathbf{w}^\intercal \nabla f(\mathbf{w}) = h(\mathbf{g}(\mathbf{w})) - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]. \tag{4.22}$$

因此，我们可以通过求解 (4.22) 得到 f，再对期望进行界定，从而界定所关注的量 $|\mathbb{E}[h(\mathbf{g}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]|$

$$\mathbb{E}[\nabla^\intercal \Sigma \nabla f(\mathbf{W}) - \mathbf{W}^\intercal \nabla f(\mathbf{W})]. \tag{4.23}$$

人们已经发展出许多耦合技巧，用于界定此类期望（见 [10, 26, 27, 34, 43]）。这些论文还给出了这一量的一般代入界，不过这些界仅适用于经典情形，即检验函数（此处为 $h(\mathbf{g}(\cdot))$）的导数有界；在这种情况下，可以应用解的导数的标准界

对于 (4.22)，可以应用这些标准界（见 [16, 27, 34]）。然而，一般而言，检验函数 h (g(·)) 的导数并无界（例如，考虑 $g(w) = w^2$），因此解的导数（见 [5, 27]）

$$f(\mathbf{w}) = -\int_0^\infty \{\mathbb{E}[h(\mathbf{g}(\mathrm{e}^{-s}\mathbf{w} + \sqrt{1-\mathrm{e}^{-2s}}\mathbf{Z}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]\}\,\mathrm{d}s \tag{4.24}$$

通常也会无界。对于一大类函数 $g : \mathbb{R}^d \to \mathbb{R}$, [19, 24] 给出了解 (4.24) 的偏导数界；特别是，当 g 的偏导数具有多项式增长时，也给出了相应的界。利用解的这类界以及为多元正态近似而提出的耦合策略，原则上可以对期望 (4.23) 给出界。本节余下部分将采用这一方法。

4.1 g : $\mathbb{R}^d \to \mathbb{R}^m$ 的 Stein 方程解的界

对于实值函数 $g : \mathbb{R}^d \to \mathbb{R}$, [19, 24] 给出了 Stein 方程 (4.24) 解的界。本节将这些界推广到向量值函数 g : $\mathbb{R}^d \to \mathbb{R}^m$。这一推广的关键是下面的引理 4.1，其中我们得到了 ϕ(w) = h(g(w)) 的 n 阶偏导数的一个简单界。为得到此界，我们使用了由 [11] 提出的 Fa`a di Bruno 公式的推广形式，该形式适用于任意 $d$、$m \geq 1$。文献 [19] 使用了 [32] 的一个公式，将经典 Fa`a di Bruno 公式推广到任意 $d \geq 1$ 和 $m = 1$。

在陈述 [11], 的公式之前，我们先回顾一些多元记号，以便简化表达式。令 $\mathbb{N}$ 为自然数集合。对于 $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \mathbb{N}^d$，且 z $\in (z_1, \ldots, z_d) \in \mathbb{R}^d$，我们记作

$$|\boldsymbol{\nu}| = \sum_{i=1}^d \nu_i, \quad \boldsymbol{\nu}! = \prod_{i=1}^d (\nu_i!), \quad D_{\mathbf{x}}^{\boldsymbol{\nu}} = \frac{\partial^{|\boldsymbol{\nu}|}}{\partial x_1^{\nu_1} \cdots \partial x_d^{\nu_d}}, \quad \text{for } |\boldsymbol{\nu}| > 0,$$

$$D_{\mathbf{x}}^{\mathbf{0}} = \text{identity operator}, \quad \mathbf{z}^{\boldsymbol{\nu}} = \prod_{i=1}^d z_i^{\nu_i}.$$

若$(l_1, \ldots, l_d) \in \mathbb{N}^d$，则当 $l_i \leq \nu_i$对 $i = 1, \ldots, d$成立时，我们记$\mathbf{l} \leq \boldsymbol{\nu}$。这使我们能够在 $\mathbb{N}^d$上引入一个线性序。若 $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$和 $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$属于 $\mathbb{N}^d$，则当以下条件之一成立时，我们记作 $\boldsymbol{\mu} \prec \boldsymbol{\nu}$：(i) $|\boldsymbol{\mu}| < |\boldsymbol{\nu}|$；(ii) $|\boldsymbol{\mu}| = |\boldsymbol{\nu}|$ 且 $\mu_1 < \nu_1$；或(iii) $|\boldsymbol{\mu}| = |\boldsymbol{\nu}|$, $\mu_1 = \nu_1, \ldots, \mu_k = \nu_k$，且对某个 $1 \leq k \leq d-1$, $\mu_{k+1} < \nu_{k+1}$。

最后，我们令 $\phi_{\boldsymbol{\nu}} = D_{\mathbf{x}}^{\boldsymbol{\nu}}\phi$(x)、$h_{\boldsymbol{\lambda}} = D_{\mathbf{y}}^{\boldsymbol{\lambda}}$(y)、$g_{\boldsymbol{\mu}}^{(i)} = D_{\mathbf{x}}^{\boldsymbol{\mu}} g^{(i)}$(x) 和 g$_{\boldsymbol{\mu}}$ = $(g_{\boldsymbol{\mu}}^{(1)}, \ldots, g_{\boldsymbol{\mu}}^{(m)})$。采用这一记号，我们可以给出广义 Fa`a di Bruno 公式（[11, 定理 2.1]）：

$$\phi_{\boldsymbol{\nu}} = \sum_{1 \leq |\boldsymbol{\lambda}| \leq n} h_{\boldsymbol{\lambda}} \sum_{s=1}^n \sum_{p_s(\boldsymbol{\nu}, \boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^s \frac{[\mathbf{g}_{\mathbf{l}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\mathbf{l}_j!]^{|\mathbf{k}_j|}}, \tag{4.25}$$

其中 $n = |\boldsymbol{\nu}|$ 和

$$p_s(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \{(\mathbf{k}_1, \ldots, \mathbf{k}_s; \mathbf{l}_1, \ldots, \mathbf{l}_s) : |\mathbf{k}_i| > 0,$$

$$\mathbf{0} \prec \mathbf{l}_1 \prec \cdots \prec \mathbf{l}_s,\ \sum_{i=1}^{s} \mathbf{k}_i = \boldsymbol{\lambda} \text{ and } \sum_{i=1}^{s} |\mathbf{k}_i|\mathbf{l}_i = \boldsymbol{\nu}\}.$$

在上述内容中，向量 $\mathbf{k}$ 是 m 维的，向量 $\mathbf{l}$ 是 d 维的，并且我们始终令 $0^0 = 1$。我们还将使用以下公式 [11, 推论 2.9]:

$$\sum_{|\boldsymbol{\lambda}|=k} \sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{s} \frac{1}{(\mathbf{k}_j!)[\mathbf{l}_j!]^{|\mathbf{k}_j|}} = m^k \left\{ {n \atop k} \right\}, \tag{4.26}$$

其中，我们回顾一下，$\left\{ {n \atop k} \right\}$是第二类斯特林数。

利用公式 (4.25) 和 (4.26)，我们能够得到 n 阶偏导数的一个简单界。对于给定函数 $P\colon\ \mathbb{R}^d \to \mathbb{R}^+$，我们回顾一下，当 $\mathbb{R}^d \to \mathbb{R}^m$ 的所有 n 阶偏导数均存在，且满足对于 $\mathbf{w} \in \mathbb{R}^d$，函数 $\mathbf{g}: \mathbb{R}^d \to \mathbb{R}^m$ 属于类 $C^{n}{}_{,P}{}^{m}(\mathbb{R}^d)$ 的条件是：$g^{(1)}, \ldots, g^{(m)}$。

$$|g_{\mathbf{l}_j}^{(i)}(\mathbf{w})|^{n/|\mathbf{l}_j|} \le P(\mathbf{w}), \quad |\mathbf{l}_j| = 1, \ldots, n,\ i = 1, \ldots, m.$$

引理 4.1。假设 $h \in C_b^n(\mathbb{R}^m)$ 且 $\mathbf{g} \in C^{n}{}_{,P}{}^{m}(\mathbb{R}^d)$。那么，对于 $\boldsymbol{\nu} \in \mathbb{N}^d$ 满足$|\boldsymbol{\nu}| = n$ 的情形，以及 $\mathbf{w} \in \mathbb{R}^d$,

$$|\phi_{\boldsymbol{\nu}}(\mathbf{w})| \le h_{n,m} P(\mathbf{w}). \tag{4.27}$$

备注 4.2。引理 4.1 以非常简单的方式将 *[19]* 中的引理 2.1 从 $m = 1$ 推广到 $m \ge 1$。(4.27) 中关于 $m \ge 1$ 的界与 *[19],* 中的不等式 (2.1) 形式相同，唯一的区别是其中出现了因子 $h_{n,m}$，而非因子$h_n := h_{n,1}$。*[19]* 中引理 2.1 的不等式 (2.1) 是对解 (4.24) 的偏导数进行界定的起点（在 $m = 1$ 的情形下）。由于我们的推广界仅相差一个与 $\mathbf{w}$ 无关的乘法因子，因此只需将因子 $h_i = h_{i,1}$ 替换为 $h_{i,m}$，即可将 *[19]* 第 2 节和 *[24]* 第 2 节给出的所有界推广到 $m \ge 1$ 的情形。

证明。由于 $\mathbf{g} \in C^{n}{}_{,P}{}^{m}(\mathbb{R}^d)$，我们有

$$\prod_{j=1}^{s} |\mathbf{g}_{\mathbf{l}_j}(\mathbf{w})|^{\mathbf{k}_j} = \prod_{j=1}^{s} \prod_{i=1}^{m} |\mathbf{g}_{\mathbf{l}_j}^{(i)}(\mathbf{w})|^{(\mathbf{k}_j)_i} \le \prod_{j=1}^{s} \prod_{i=1}^{m} \left((P(\mathbf{w}))^{|\mathbf{l}_j|/n} \right)^{(\mathbf{k}_j)_i},$$

其中 $(\mathbf{k}j)_i$ 是 $\mathbf{k}j$ 的第 i 个分量。但

$$\sum_{j=1}^{s} \sum_{i=1}^{m} |\mathbf{l}_j|(\mathbf{k}_j)_i = \sum_{j=1}^{s} |\mathbf{l}_j| \sum_{i=1}^{m} (\mathbf{k}_j)_i = \sum_{j=1}^{s} |\mathbf{l}_j||\mathbf{k}_j| = n,$$

因此，对于 $\mathbf{w} \in \mathbb{R}^d$,

$$\prod_{j=1}^{s} |\mathbf{g}_{\mathbf{l}_j}(\mathbf{w})|^{\mathbf{k}_j} \le P(\mathbf{w}). \tag{4.28}$$

回顾公式（4.25），我们得到

$$|\phi_{\boldsymbol{\nu}}(\mathbf{w})| = \left| \sum_{1 \le |\boldsymbol{\lambda}| \le n} h_{\boldsymbol{\lambda}}(\mathbf{g}(\mathbf{w})) \sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{s} \frac{[\mathbf{g}_{\mathbf{l}_j}(\mathbf{w})]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\mathbf{l}_j!]^{|\mathbf{k}_j|}} \right|$$

$$
\begin{aligned}
&\leq \sum_{1\leq|\boldsymbol{\lambda}|\leq n} |h|_{|\boldsymbol{\lambda}|} \sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{s} \frac{|\mathbf{g}_{\mathbf{l}_j}(\mathbf{w})|^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\mathbf{l}_j!]^{|\mathbf{k}_j|}} \\
&\leq P(\mathbf{w}) \sum_{1\leq|\boldsymbol{\lambda}|\leq n} |h|_{|\boldsymbol{\lambda}|} \sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{s} \frac{1}{(\mathbf{k}_j!)[\mathbf{l}_j!]^{|\mathbf{k}_j|}} \\
&= P(\mathbf{w}) \sum_{1\leq|\boldsymbol{\lambda}|\leq n} m^{|\boldsymbol{\lambda}|} \left\{ {n \atop |\boldsymbol{\lambda}|} \right\} |h|_{|\boldsymbol{\lambda}|},
\end{aligned}
$$

其中，我们使用不等式（4.28）得到了最后一个不等式，并使用（4.26）得到了最后一个等式。证毕。 □

4.2 定理 2.11–2.14 的证明

下面的引理给出了若干期望的界，这些期望将在定理 2.11 和 2.12 的证明中出现。证明见补充材料。在该引理中，f 表示 Stein 方程 (4.22) 的解 (4.24)，ψ_{jkl} 则是下面这个偏微分方程的解，该方程由 [19] 引入：

$$
\nabla^{\intercal}\Sigma\nabla\psi_{jkl}(\mathbf{w}) - \mathbf{w}^{\intercal}\nabla\psi_{jkl}(\mathbf{w}) = \frac{\partial^3 f}{\partial w_j \partial w_k \partial w_l}(\mathbf{w}). \tag{4.29}
$$

对于 $i = 1$、……、n 和 $j = 1$、……、d，我们引入随机变量 $W_{(i)j} = W_j - n^{-1/2}X_{ij}$，使得 $W_{(i)j}$ 与 X_{ij} 相互独立。我们还记 $\mathbf{W}^{(i)} = (W_1^{(i)}$、……、$W_{(i)d})^{\intercal}$。

引理 4.3. 令 $\mathbf{X}_i$ 表示随机向量 $(X_{i,1}, \ldots\ldots, X_{i,d})^{\intercal}$，对于 $r \geq 0$，令 $u:\mathbb{R}^d \to \mathbb{R}^+$ 满足对所有 $i = 1$、……、n 以及 $j = 1$、……、d，均有 $\mathbb{E}|X_{ij}^r u(\mathbf{X}_i)| < \infty$。假设 Σ 半正定。那么，对于 $\theta \in \in (0, 1)$，

$$
\begin{aligned}
\mathbb{E}\left| u(\mathbf{X}_i) \frac{\partial^t f}{\prod_{j=1}^t \partial w_{i_j}}(\mathbf{W}_\theta^{(i)}) \right| \leq \frac{h_{t,m}}{t} \Big[A\mathbb{E}[u(\mathbf{X}_i)] + B \sum_{j=1}^{d} \big(& 2^{3r/2}\mathbb{E}[u(\mathbf{X}_i)]\mathbb{E}|W_j|^r \\
& + \mathbb{E}|X_{ij}^r u(\mathbf{X}_i)| + 2^{r/2}\mu_{r,\sigma_j}\mathbb{E}[u(\mathbf{X}_i)] \big) \Big], \qquad (4.30) \\
\mathbb{E}\left| u(\mathbf{X}_i) \frac{\partial^3 \psi_{abc}}{\partial w_j \partial w_k \partial w_l}(\mathbf{W}_\theta^{(i)}) \right| \leq \frac{h_{6,m}}{18} \Big[A\mathbb{E}[u(\mathbf{X}_i)] + B \sum_{j=1}^{d} \big(& 12^{r/2}\mathbb{E}[u(\mathbf{X}_i)]\mathbb{E}|W_j|^r \\
& + \mathbb{E}|X_{ij}^r u(\mathbf{X}_i)| + 2 \cdot 3^{r/2}\mu_{r,\sigma_j}\mathbb{E}[u(\mathbf{X}_i)] \big) \Big], \qquad (4.31)
\end{aligned}
$$

其中，这些不等式分别适用于类 $C^{t,m}_{,P}$ $(\mathbb{R}^d)$ 和 $C^{6,m}_{,P}$ $(\mathbb{R}^d)$ 中的 g，并且 $\mathbf{W}_\theta^{(i)} = \mathbf{W}^{(i)} + \theta n^{-1/2}\mathbf{X}_i$. 不等式 (4.30) 对 $n \geq 8$ 有效，而不等式 (4.31) 对 $n \geq 12$ 成立。

定理 2.11 的证明。将 $\frac{\partial f}{\partial w_j}(\mathbf{W})$ 在 $\mathbf{W}^{(i)}$ 处进行泰勒展开，利用独立性以及对所有 i, j 均有 $\mathbb{E}[X_{ij}] = 0$，可得

$$
\sum_{j=1}^{d} \mathbb{E}\left[W_j \frac{\partial f}{\partial w_j}(\mathbf{W}) \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbb{E}\left[X_{ij} \frac{\partial f}{\partial w_j}(\mathbf{W}) \right]
$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbb{E}[X_{ij}] \mathbb{E}\left[\frac{\partial f}{\partial w_j}(\mathbf{W}^{(i)}) \right] + \frac{1}{n} \sum_{i=1}^{n} \sum_{j,k=1}^{d} \mathbb{E}[X_{ij} X_{ik}] \mathbb{E}\left[\frac{\partial^2 f}{\partial w_j \partial w_k}(\mathbf{W}^{(i)}) \right] + R_1$$
$$= \sum_{j,k=1}^{d} \sigma_{jk} \mathbb{E}\left[\frac{\partial^2 f}{\partial w_j \partial w_k}(\mathbf{W}) \right] + R_1 + R_2,$$

其中

$$|R_1| \le \frac{1}{2n^{3/2}} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} \mathbb{E}\left| X_{ij} X_{ik} X_{il} \frac{\partial^3 f}{\partial w_j \partial w_k \partial w_l}(\mathbf{W}_{\theta_1}^{(i)}) \right|,$$
$$|R_2| \le \frac{1}{n^{3/2}} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} |\mathbb{E}[X_{ij} X_{ik}]| \mathbb{E}\left| X_{il} \frac{\partial^3 f}{\partial w_j \partial w_k \partial w_l}(\mathbf{W}_{\theta_2}^{(i)}) \right|,$$

且 $\mathbf{W}_{\theta}^{(i)} = \mathbf{W}^{(i)} + \theta n^{-1/2} \mathbf{X}_i$，其中某个 $\theta \in (0, 1)$。这里我们用到了 $n^{-1} \sum_{i=1}^{n} \mathbb{E}[X_{ij} X_{ik}] = \mathbb{E}[W_j W_k] = \sigma_{jk}$。利用引理 4.3 对 $|R_1|$ 和 $|R_2|$ 中的期望取界，可得

$$|\mathbb{E}[h(\mathbf{g}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]| \le |\mathbb{E}[\nabla^\intercal \Sigma \nabla f(\mathbf{W}) - \mathbf{W}^\intercal \nabla f(\mathbf{W})]| \le |R_1| + |R_2|$$
$$\le \frac{h_{3,m}}{6n^{3/2}} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} \Bigg\{ A\mathbb{E}|X_{ij} X_{ik} X_{il}| + B \sum_{t=1}^{d} \big(2^{3r/2} \mathbb{E}|X_{ij} X_{ik} X_{il}| \mathbb{E}|W_t|^r$$
$$+ \mathbb{E}|X_{ij} X_{ik} X_{il} X_{it}^r| + 2^{r/2} \mu_{r,\sigma_t} \mathbb{E}|X_{ij} X_{ik} X_{il}| \big) + 2|\mathbb{E}[X_{ij} X_{ik}]| \Bigg[A\mathbb{E}|X_{il}|$$
$$+ B \sum_{t=1}^{d} \big(2^{3r/2} \mathbb{E}|X_{il}| \mathbb{E}|W_t|^r + \mathbb{E}|X_{il} X_{it}^r| + 2^{r/2} \mu_{r,\sigma_t} \mathbb{E}|X_{il}| \big) \Bigg] \Bigg\}.$$

简化后的界 (2.13) 通过以下考虑得到。我们使用乘积的杨氏不等式（$ab \le a^p/p + b^q/q$、a、$b \ge 0$、p、$q > 1$，使得 $p^{-1} + q^{-1} = 1$）对诸如 $\mathbb{E}|X_{ij} X_{ik} X_{il} X_{it}^r| \le c_1 \mathbb{E}|X_{ij}|^{r+3} + c_2 \mathbb{E}|X_{ik}|^{r+3} + c_3 \mathbb{E}|X_{il}|^{r+3} + c_4 \mathbb{E}|X_{it}|^{r+3}$ 这样的期望取界，其中 $c_1 + c_2 + c_3 + c_4 = 1$。我们还使用乘积的杨氏不等式对诸如 $|\mathbb{E}[X_{ij} X_{ik}]| \mathbb{E}|X_{il} X_{it}^r| \le \mathbb{E}|X_{ij} X_{ik}| \mathbb{E}|X_{il} X_{it}^r| \le c_1' \mathbb{E}|X_{ij}|^{r+3} + c_2' \mathbb{E}|X_{ik}|^{r+3} + c_3' \mathbb{E}|X_{il}|^{r+3} + c_4' \mathbb{E}|X_{it}|^{r+3}$) 这样的量取界，其中 $c_1' + c_2' + c_3' + c_4' = 1$。□

定理 2.12 的证明。(i) 假设 g 是偶函数。采用与定理 2.11 证明中类似的论证，

$$\sum_{j=1}^{d} \mathbb{E}\left[W_j \frac{\partial f}{\partial w_j}(\mathbf{W}) \right] = \frac{1}{n} \sum_{i=1}^{n} \sum_{j,k=1}^{d} \mathbb{E}[X_{ij} X_{ik}] \mathbb{E}\left[\frac{\partial^2 f}{\partial w_j \partial w_k}(\mathbf{W}) \right] + N_1 + N_2 + R_1 + R_2,$$

其中

$$N_1 = \frac{1}{2n^{3/2}} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} \mathbb{E}[X_{ij} X_{ik} X_{il}] \mathbb{E}\left[\frac{\partial^3 f}{\partial w_j \partial w_k \partial w_l}(\mathbf{W}^{(i)}) \right],$$
$$N_2 = -\frac{1}{n^{3/2}} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} \mathbb{E}[X_{ij} X_{ik}] \mathbb{E}\left[X_{il} \frac{\partial^3 f}{\partial w_j \partial w_k \partial w_l}(\mathbf{W}) \right],$$

$$|R_1| \le \frac{1}{6n^2}\sum_{i=1}^{n}\sum_{j,k,l,t=1}^{d}\mathbb{E}\left|X_{ij}X_{ik}X_{il}X_{it}\frac{\partial^4 f}{\partial w_j\partial w_k\partial w_l\partial w_t}(\mathbf{W}_{\theta_1}^{(i)})\right|,$$

$$|R_2| \le \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j,k,l,t=1}^{d}|\mathbb{E}[X_{ij}X_{ik}]|\mathbb{E}\left|X_{il}X_{it}\frac{\partial^4 f}{\partial w_j\partial w_k\partial w_l\partial w_t}(\mathbf{W}_{\theta_2}^{(i)})\right|.$$

我们可以将 N_1 和 N_2 写作

$$N_1 = \frac{1}{2n^{3/2}}\sum_{i=1}^{n}\sum_{j,k,l=1}^{d}\mathbb{E}[X_{ij}X_{ik}X_{il}]\mathbb{E}\left[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{W})\right] + R_3,$$

$$N_2 = -\frac{1}{n^{3/2}}\sum_{i=1}^{n}\sum_{j,k,l=1}^{d}\mathbb{E}[X_{ij}X_{ik}]\mathbb{E}[X_{il}]\mathbb{E}\left[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{W}^{(i)})\right] + R_4 = R_4,$$

其中

$$|R_3| \le \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j,k,l,t=1}^{d}|\mathbb{E}[X_{ij}X_{ik}X_{il}]|\mathbb{E}\left|X_{it}\frac{\partial^4 f}{\partial w_j\partial w_k\partial w_l\partial w_t}(\mathbf{W}_{\theta_3}^{(i)})\right|,$$

$$|R_4| \le \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j,k,l,t=1}^{d}|\mathbb{E}[X_{ij}X_{ik}]|\mathbb{E}\left|X_{il}X_{it}\frac{\partial^4 f}{\partial w_j\partial w_k\partial w_l\partial w_t}(\mathbf{W}_{\theta_4}^{(i)})\right|.$$

结合各项界可得:

$$|\mathbb{E}[h(\mathbf{g}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]| \le \frac{1}{2n^{3/2}}\sum_{i=1}^{n}\sum_{j,k,l=1}^{d}|\mathbb{E}[X_{ij}X_{ik}X_{il}]|\left|\mathbb{E}\left[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{W})\right]\right| + |R_1| + |R_2| + |R_3| + |R_4|. \tag{4.32}$$

为了得到所需的 $O(n^{-1})$ 界，我们需要证明 $\mathbb{E}[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{W})]$ 的阶为 $n^{-1/2}$。为此，我们考虑 Stein 方程。

$$\nabla^\intercal\Sigma\nabla\psi_{jkl}(\mathbf{w}) - \mathbf{w}^\intercal\nabla\psi_{jkl}(\mathbf{w}) = \frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{w}) - \mathbb{E}\left[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{Z})\right],$$

其中 $\mathbf{Z} \sim \mathrm{N}_d(\mathbf{0},\Sigma)$。由于 g 是偶函数，因此解 f（如式 (4.24) 所示）也是偶函数（这是对 [19, 引理 3.2]的简单推广）。因此$\mathbb{E}[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{Z})] = 0$，于是

$$\begin{aligned}\mathbb{E}\left[\frac{\partial^3 f}{\partial w_j\partial w_k\partial w_l}(\mathbf{W})\right] &= \mathbb{E}[\nabla^\intercal\Sigma\nabla\psi_{jkl}(\mathbf{W}) - \mathbf{W}^\intercal\nabla\psi_{jkl}(\mathbf{W})] \\ &\le \frac{1}{n^{3/2}}\sum_{\alpha=1}^{n}\sum_{a,b,c=1}^{d}\left\{\frac{1}{2}\mathbb{E}\left|X_{\alpha a}X_{\alpha b}X_{\alpha c}\frac{\partial^3\psi_{jkl}}{\partial w_a\partial w_b\partial w_c}(\mathbf{W}_{\theta_5}^{(\alpha)})\right|\right. \\ &\quad \left. + |\mathbb{E}[X_{\alpha a}X_{\alpha b}]|\mathbb{E}\left|X_{\alpha c}\frac{\partial^3\psi_{jkl}}{\partial w_a\partial w_b\partial w_c}(\mathbf{W}_{\theta_6}^{(\alpha)})\right|\right\},\end{aligned} \tag{4.33}$$

其中，我们通过类似于定理 2.11 证明中的论证得到了该不等式。利用引理 4.3 中的不等式对 $|R_1|, |R_2|, |R_3|, |R_4|$ 和 (4.33) 进行界定，并将这些界应用于 (4.32)，即可得到以下界

$$|\mathbb{E}[h(\mathbf{g}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}(\mathbf{Z}))]| \le K_1' + K_2',$$

其中

$$\begin{aligned}
K_1' =& \frac{h_{4,m}}{24n^2} \sum_{i=1}^{n} \sum_{j,k,l,t=1}^{d} \Big\{ A\mathbb{E}|X_{ij}X_{ik}X_{il}X_{it}| + B\sum_{q=1}^{d} \big(2^{3r/2}\mathbb{E}|X_{ij}X_{ik}X_{il}X_{it}|\mathbb{E}|W_q|^r \\
&+ \mathbb{E}|X_{ij}X_{ik}X_{il}X_{it}X_{iq}^r| + 2^{r/2}\mu_{r,\sigma_q}\mathbb{E}|X_{ij}X_{ik}X_{il}X_{it}|\big) + 9|\mathbb{E}[X_{ij}X_{ik}]|\Big[A\mathbb{E}|X_{il}X_{it}| \\
&+ B\sum_{q=1}^{d} \big(2^{3r/2}\mathbb{E}|X_{il}X_{it}|\mathbb{E}|W_q|^r + \mathbb{E}|X_{il}X_{it}X_{iq}^r| + 2^{r/2}\mu_{r,\sigma_q}\mathbb{E}|X_{il}X_{it}|\big)\Big] \\
&+ 3|\mathbb{E}[X_{ij}X_{ik}X_{il}]|\Big[A\mathbb{E}|X_{it}| + B\sum_{q=1}^{d} \big(2^{3r/2}\mathbb{E}|X_{it}|\mathbb{E}|W_q|^r + \mathbb{E}|X_{it}X_{iq}^r| \\
&+ 2^{r/2}\mu_{r,\sigma_q}\mathbb{E}|X_{it}|\big)\Big]\Big\}, \\
K_2' =& \frac{h_{6,m}}{72n^3} \sum_{i=1}^{n} \sum_{j,k,l=1}^{d} |\mathbb{E}[X_{ij}X_{ik}X_{il}]| \sum_{\alpha=1}^{n} \sum_{a,b,c=1}^{d} \Big\{ A\mathbb{E}|X_{\alpha a}X_{\alpha b}X_{\alpha c}| \\
&+ B\sum_{q=1}^{d} \big(12^{r/2}\mathbb{E}|X_{\alpha a}X_{\alpha b}X_{\alpha c}|\mathbb{E}|W_q|^r + \mathbb{E}|X_{\alpha a}X_{\alpha b}X_{\alpha c}X_{\alpha q}^r| \\
&+ 2\cdot 3^{r/2}\mu_{r,\sigma_q}\mathbb{E}|X_{\alpha a}X_{\alpha b}X_{\alpha c}|\big) + 2|\mathbb{E}[X_{\alpha a}X_{\alpha b}]|\Big[A\mathbb{E}|X_{\alpha c}| \\
&+ B\sum_{q=1}^{d} \big(12^{r/2}\mathbb{E}|X_{\alpha c}|\mathbb{E}|W_q|^r + \mathbb{E}|X_{\alpha c}X_{\alpha q}^r| + 2\cdot 3^{r/2}\mu_{r,\sigma_q}\mathbb{E}|X_{\alpha c}|\big)\Big]\Big\}.
\end{aligned}$$

现在，我们像在定理 2.11 的证明中所做的那样，进一步简化这个界。

(ii) 请注意，当 $1 \le i \le n$ 且 $1 \le j$、k、$l \le d$ 时，界 (4.32) 化简为 $|R_1|+|R_2|+|R_3|+|R_4|$ $\mathbb{E}[X_{ij}X_{ik}X_{il}] = 0$。因此，当我们施加这一条件时，$K_2' = 0$，而 K_1' 中涉及 $\mathbb{E}[X_{ij}X_{ik}X_{il}]$ 倍数的最后一项消失。因此，我们不再要求 g 是偶函数，在这种情况下，g 和 h 的函数类放宽为 g $\in C^4{}_P{}^m$ ($\mathbb{R}^d$) 和 $h \in C_b^4(\mathbb{R}^m)$。 □

定理 2.13 和定理 2.14 的证明。对于定理 2.13，我们在定理 2.11 的证明中给出的$|R_1|$ 和 $|R_2|$ 的界中设定 $d = m = 1$，并使用 [24] 的引理 3.6 中的一元界来估计期望（将单位方差简单推广到一般方差 Var(W) $= \sigma^2$）。类似地，对于定理 2.14，我们在合并不等式 (4.32) 和 (4.33) 所得到的界中设定 $d = m = 1$，并使用 [24] 的引理 3.6 中的一元界来估计期望（采用同样的轻微

推广）。利用假设$n \geq 8$ 和 $n \geq 12$对所得界进行一些简化后，即可得到所需的界。

□

5 定理 2.1–2.6 的证明

定理 2.1 的证明。令 $g_{t,n}(w) = n^{t/2}(f(n^{-1/2}w) - f(0))$，并令 $W = \sqrt{n}X$，使得 $T_{t,n} = g_{t,n}(W)$。令 $Z = (Z_1, \dots, Z_d)^T \sim N_d(0, \Sigma)$。于是，根据三角不等式，

$$\begin{aligned}\Delta_h(\mathbf{T}_{t,n}, \mathbf{Y}_t) &\leq |\mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{W}))] - \mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{Z}))]| + |\mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{Z}))] - \mathbb{E}[h(\mathbf{Y}_t)]| \\ &:= R_1 + R_2. \end{aligned} \tag{5.34}$$

我们首先对 R_2 进行界定。将 $f(n^{-1/2}w)$ 在 0 处作泰勒展开，并利用对所有 $1 \leq i_1, \dots, i_{t-1} \leq d$ 都有 $\partial_f/\partial w_{i_1}(0) = \cdots = \partial^{t-1}f/\prod_{j=1}^{t-1}\partial w_{i_j}(0) = 0$（在 $t \geq 2$ 的情形下），可得

$$\mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{Z}))] = \mathbb{E}\bigg[h\bigg(\mathbf{Y}_t + \frac{1}{(t+1)!\sqrt{n}} \sum_{i_1,\dots,i_{t+1}=1}^{d} Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1}\mathbf{f}}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg)\bigg],$$

对于某个 $\theta \in (0, 1)$。于是，根据均值定理和三角不等式，

$$\begin{aligned} R_2 &\leq \frac{|h|_1}{(t+1)!\sqrt{n}} \sum_{k=1}^{m} \sum_{i_1,\dots,i_{t+1}=1}^{d} \mathbb{E}\bigg|Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg| \\ &\leq \frac{A_{t+1}m|h|_1}{(t+1)!\sqrt{n}}\mathbb{E}\bigg[\bigg(1 + \sum_{j=1}^{d}\bigg|\frac{Z_j}{\sqrt{n}}\bigg|^{r_{t+1}}\bigg) \sum_{i_1,\dots,i_{t+1}=1}^{d} |Z_{i_1}\cdots Z_{i_{t+1}}|\bigg] \\ &\leq \frac{A_{t+1}m|h|_1}{(t+1)!\sqrt{n}}\mathbb{E}\bigg[\bigg(1 + \sum_{j=1}^{d}\bigg|\frac{Z_j}{\sqrt{n}}\bigg|^{r_{t+1}}\bigg) d^t \sum_{\ell=1}^{d} |Z_\ell|^{t+1}\bigg] \\ &\leq \frac{A_{t+1}md^t|h|_1}{(t+1)!\sqrt{n}} \sum_{j=1}^{d}\bigg(\mathbb{E}|Z_j|^{t+1} + \frac{d}{n^{r_{t+1}/2}}\mathbb{E}|Z_j|^{r_{t+1}+t+1}\bigg). \end{aligned}$$

我们利用定理 2.11 的界 (2.13) 来界定 R_1。要应用该定理，我们需要找到一个控制函数 P，使得 $g_{t,n} \in C^3, P^m(\mathbb{R}^d)$。假设对于 $t = 1$，$f \in \mathcal{C}^1, 2, P^m(\mathbb{R}^d)$；对于 $t \geq 2$，$f \in \mathcal{C}^t, 1, P^m(\mathbb{R}^d)$。在这些假设下，我们可以证明，一个合适的控制函数为 $P(w) = 2C_{1,t}a_{n,d,r_t}^3 d^{3t-4}(d + \sum_{i=1}^{d}|w_i|^{u_{1,t}})$，其中常数 $C_{1,t}$ 和 $u_{1,t}$ 的定义与定理陈述中的相同。利用该控制函数，我们可得界 $R_1 \leq h_{3,m}\mathcal{M}_{2,d}/\sqrt{n}$，其中 $\mathcal{M}_{2,d}$ 的定义与定理陈述中的相同。将关于 R_1 和 R_2 的界相加，即得所需的界 (2.2)。接下来只需证明 $g_{t,n} \in C^3, P^m(\mathbb{R}^d)$，其中 P 如上所定义。这一点可在补充材料中验证。

□

定理 2.2 的证明。我们先对 (5.34) 中的 R_2 求界。将 $f(n^{-1/2}w)$ 在 0 处作泰勒展开，并利用对于所有 $1 \leq, \partial_f/\partial w_{i_1}(0) = \cdots = \partial^{t-1}f/\prod_{j=1}^{t-1}\partial w_{i_j}(0) = 0$

$i_1, \ldots, i_{t-1} \le d$（在 $t \ge 2$ 的情形下），我们有

$$\mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{Z}))] = \mathbb{E}\bigg[h\bigg(\mathbf{Y}_t + \frac{1}{(t+1)!\sqrt{n}} \sum_{i_1,\ldots,i_{t+1}=1}^{d} Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1}\mathbf{f}}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}(\mathbf{0}) \\ + \frac{1}{(t+2)!n} \sum_{i_1,\ldots,i_{t+2}=1}^{d} Z_{i_1}\cdots Z_{i_{t+2}} \frac{\partial^{t+2}\mathbf{f}}{\partial w_{i_1}\cdots\partial w_{i_{t+2}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg)\bigg],$$

对于某个 $\theta \in (0,1)$。然后，对 $\mathbf{Y}_t$ 处的 h 进行泰勒展开，可得

$$|R_2| = |\mathbb{E}[h(\mathbf{g}_{t,n}(\mathbf{Z}))] - \mathbb{E}[h(\mathbf{Y}_t)]| = |R_3 + R_4|,$$

其中

$$R_3 = \mathbb{E}\bigg[\sum_{k=1}^{m} \frac{\partial h}{\partial w_k}(\mathbf{Y}_t)\bigg(\frac{1}{\sqrt{n}(t+1)!} \sum_{i_1,\ldots,i_{t+1}=1}^{d} Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}(\mathbf{0}) \\ + \frac{1}{n(t+2)!} \sum_{i_1,\ldots,i_{t+2}=1}^{d} Z_{i_1}\cdots Z_{i_{t+2}} \frac{\partial^{t+2} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+2}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg)\bigg],$$

$$|R_4| \le \frac{1}{2}|h|_2 \mathbb{E}\bigg[\bigg(\sum_{k=1}^{m}\bigg|\frac{1}{\sqrt{n}(t+1)!} \sum_{i_1,\ldots,i_{t+1}=1}^{d} Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}(\mathbf{0}) \\ + \frac{1}{n(t+2)!} \sum_{i_1,\ldots,i_{t+2}=1}^{d} Z_{i_1}\cdots Z_{i_{t+2}} \frac{\partial^{t+2} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+2}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg|\bigg)^2\bigg].$$

注意到

$$\mathbb{E}\bigg[\sum_{k=1}^{m} \frac{\partial h}{\partial w_k}(\mathbf{Y}_t)\bigg(\frac{1}{\sqrt{n}(t+1)!} \sum_{i_1,\ldots,i_{t+1}=1}^{d} Z_{i_1}\cdots Z_{i_{t+1}} \frac{\partial^{t+1} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+1}}}(\mathbf{0})\bigg)\bigg] = 0,$$

因为该期望的形式为 $\mathbb{E}[\tau(\mathbf{Z})],$，其中 τ 是奇函数，而 $\mathbf{Z} \sim \mathbf{N}_d(\mathbf{0}, \Sigma)$ 关于原点对称分布（即 $\mathbf{Z} =_d -\mathbf{Z}$），这意味着该期望必定等于零。因此

$$|R_3| = \frac{1}{(t+2)!n}\bigg|\mathbb{E}\bigg[\sum_{k=1}^{m} \frac{\partial h}{\partial w_k}(\mathbf{Y}_t) \sum_{i_1,\ldots,i_{t+2}=1}^{d} Z_{i_1}\cdots Z_{i_{t+2}} \frac{\partial^{t+2} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+2}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg]\bigg| \\ \le \frac{|h|_1}{(t+2)!n} \sum_{k=1}^{m} \sum_{i_1,\ldots,i_{t+2}=1}^{d} \mathbb{E}\bigg|Z_{i_1}\cdots Z_{i_{t+2}} \frac{\partial^{t+2} f_k}{\partial w_{i_1}\cdots\partial w_{i_{t+2}}}\bigg(\frac{\theta\mathbf{Z}}{\sqrt{n}}\bigg)\bigg| \\ \le \frac{A_{t+2} m |h|_1}{(t+2)!n} \mathbb{E}\bigg[\bigg(1 + \sum_{j=1}^{d}\bigg|\frac{Z_j}{\sqrt{n}}\bigg|^{r_{t+2}}\bigg) \sum_{i_1,\ldots,i_{t+2}=1}^{d} \big|Z_{i_1}\cdots Z_{i_{t+2}}\big|\bigg] \\ \le \frac{A_{t+2} m d^{t+1} |h|_1}{(t+2)!n} \sum_{j=1}^{d}\bigg(\mathbb{E}|Z_j|^{t+2} + \frac{d}{n^{r_{t+2}/2}}\mathbb{E}|Z_j|^{r_{t+2}+t+2}\bigg).$$

现在我们来界定 R_4:

$$
\begin{aligned}
|R_4| &\le \frac{m^2|h|_2}{2((t+1)!)^2 n}\mathbb{E}\Bigg[\Bigg(A_{t+1}\sum_{i_1,\dots,i_{t+1}=1}^{d}\left|Z_{i_1}\cdots Z_{i_{t+1}}\right| \\
&\quad + \frac{A_{t+2}}{(t+2)\sqrt{n}}\Bigg(1+\sum_{j=1}^{d}|Z_j|^{r_{t+2}}\Bigg)\sum_{i_1,\dots,i_{t+2}=1}^{d}\left|Z_{i_1}\cdots Z_{i_{t+2}}\right|\Bigg)^2\Bigg] \\
&\le \frac{m^2|h|_2}{2((t+1)!)^2 n}\mathbb{E}\Bigg[\Bigg(A_{t+1}d^t\sum_{j=1}^{d}|Z_j|^{t+1} \\
&\quad + \frac{A_{t+2}d^{t+1}}{(t+2)\sqrt{n}}\Bigg(\sum_{j=1}^{d}|Z_j|^{t+2}+d\sum_{j=1}^{d}|Z_j|^{r_{t+2}+t+2}\Bigg)\Bigg)^2\Bigg] \\
&\le \frac{m^2 d^{2t+1}|h|_2}{((t+1)!)^2 n}\sum_{j=1}^{d}\Bigg[A_{t+1}^2\mathbb{E}|Z_j|^{2(t+1)}+\frac{2A_{t+2}^2 d^2}{(t+2)^2 n}\Bigg(\mathbb{E}|Z_j|^{2(t+2)}+d^2\mathbb{E}|Z_j|^{2(r_{t+2}+t+2)}\Bigg)\Bigg].
\end{aligned}
$$

现在，我们利用定理 2.12 第 (i) 部分的界 (2.14)，对式 (5.34) 中的 R_1 取界。要应用该界，我们需要找到一个控制函数 P，使得 $\mathsf{g}_{t,n}\in C^6_{,P^m}(\mathbb{R}^d)$。假设对于 $t=2$，$\mathsf{f}\in\mathcal{C}^2$，$4_{,P^m}(\mathbb{R}^d)$；对于 $t\ge 4$，$\mathsf{f}\in\mathcal{C}^t,2_{,P^m}(\mathbb{R}^d)$。在这些假设下，一个合适的控制函数为 $P(\mathsf{w})=2C_{2,t}a^6_{n,d,r_t}d^{6t-7}(d+\sum_{i=1}^d|w_i|^{u_{2,t}})$，其中，对于 $t\ge 2$，常数 $C_{2,t}$ 和 $u_{2,t}$ 按照定理陈述中的定义给出。SM 中验证了这一点。用该控制函数应用界 (2.14)，可得界 $R_1\le(h_{4,m}\mathcal{K}_{3,d}+h_{6,m}\mathcal{K}_{4,d})/n$。将我们对 R_1,R_3 和 R_4 所得的界相加，即得所需的界 (2.3)。

□

定理 2.3 的证明。对于 R_2，我们得到的界与定理 2.2 中相同，因为在得到该界时并未用到 f 是偶函数这一假设。为了界定 R_1，我们使用定理 2.12 的第 (ii) 部分中的界 (2.15)。要应用该定理，我们需要找到一个控制函数 P，使得 $\mathsf{g}_{t,n}\in C^4_{,P^m}(\mathbb{R}^d)$。假设对于 $t=2$，$\mathsf{f}\in\mathcal{C}^2,2_{,P^m}(\mathbb{R}^d)$；对于 $t\ge 4$，$\mathsf{f}\in\mathcal{C}^t,2_{,P^m}(\mathbb{R}^d)$。在这些假设下，一个合适的控制函数为 $P(\mathsf{w})=2C_{3,t}a^4_{n,d,r_t}d^{4t-5}(d+\sum_{i=1}^d|w_i|^{u_{3,t}})$，其中 $C_{3,t}$ 和 $u_{3,t}$ 的定义与定理陈述中相同。该结论在补充材料中得到验证。利用这个控制函数，我们得到界 $R_1\le h_{4,m}\mathcal{K}_{5,d}/n$。将关于 R_1 和 R_2 的界相加，即得界 (2.4)。

□

定理 2.4 的证明。与多元情形一样，我们对 (5.34) 中的 R_1 和 R_2 进行界定。通过在界 $m|h|_1\mathcal{M}_{1,d}/\sqrt{n}$ 中令 $d=m=1$，我们可以立即对 R_2 进行界定。我们利用定理 2.13 的界 (2.16) 对 R_1 进行界定。为了应用该界，我们需要找到一个控制函数 P，使得 $g_{t,n}\in C^1_{,P^1}(\mathbb{R})$。对于 $f\in\mathcal{C}^t_{,P^{1\,1}}(\mathbb{R})$，一个合适的控制函数为 $P(w)=2A_t(1+|w|^{r_t+t-1})/(t-1)!$。这一点已在补充材料（SM）中得到验证。利用该控制函数应用不等式 (2.16) 对 R_1 进行界定，即可得到界 (2.5)。

□

定理 2.5 的证明。与定理 2.4 的证明相同，我们可以立即将 R_2界定为 $(\|h'\|\mathcal{K}_{1,1}+\|h''\|\mathcal{K}_{2,1})/n$。我们通过对控制函数 $P(w)=2C_{4,t}(1+|w|^{2(r_t+t-1)})$ 应用定理 2.14 第 (i) 部分中的不等式 (2.17)，来界定 R_1，由此可得

$g_{t,n} \in C^2_{,P^1}(\mathbb{R})$ 对于 $f \in \mathcal{C}^t_{,\,P^{2\,1}},(\mathbb{R})$；验证过程见补充材料。将 R_1 和 R_2 的界相加，即得所需的界。 □

定理 2.6 的证明。我们对 R_2 得到的界与定理 2.5 中相同，因为在推导该界时并未用到 f 是偶函数这一假设。我们也使用相同的控制函数 $P(w) = 2C_{4,\,t}(1 + |w|^{2(r_t+t-1)})$，但改用定理 2.14 第 (ii) 部分中的界 (2.18) 来界定 R_{1}。合并 R_1 和 R_2 的界即可得到所需的界。 □

致谢

HS 获得 EPSRC 博士研究奖学金的资助。

参考文献

[1] Anastasiou, A. 和 Gaunt, R. E. 通过 Delta 方法对最大似然估计量进行多元正态近似。Braz. J. Probab. Stat. 34 (2020), 136–149.[2] Anastasiou, A. 和 Ley, C. 使用 Delta 方法得到最大似然估计量渐近正态性的界。ALEA Lat. Am. J. Probab. Math. Stat. 14 (2017), 153–171.[3] Anastasiou, A. 和 Reinert, G. 似然比渐近分布的界。Ann. Appl. Probab. 30 (2020), 608–643.[4] Asylebekov, Z. A.、Zubov, V. N. 和 Ulyanov, V. V. 关于三维离散数据情形下若干拟合优度检验统计量的近似。Siberian Math. J. 52 (2011), 571–584.[5] Barbour, A. D. 用于扩散近似的斯坦方法。Probab. Theory Rel. 84 (1990), 297–322.[6] Barbour, A. D. 和 Hall, P. 关于中心极限定理中收敛速率的界。Bull. Lond. Math. Soc. 17 (1985), 151–156.[7] Brown, G. 和 Mood, A. 线性假设的中位数检验。Berkeley Symposium on Mathematical Statistics and Probability (1951), 159–166.[8] Bishop, Y. M. M.、Fienberg, S. E. 和 Holland, P. W.《离散多元分析》。马萨诸塞州剑桥：MIT Press, 1975.[9] Chatterjee, S. 一种新的正态近似方法。Ann. Probab. 36 (2008), 1584–1610.[10] Chatterjee, S. 和 Meckes, E. 使用可交换对进行多元正态近似。ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008), 257–283.[11] Constantine, G. M. 和 Savits, T. H. 一个多元 Faà di Bruno 公式及其应用。T. Amer. Math. Soc. 348 (1996), 503–520.[12] Dicker, L. H. 和 Erdogdu, M. A. 二次型的灵活结果及其在方差分量估计中的应用。Ann. Stat. 45 (2017), 386–414.[13] Doob, J. L. 某些统计量的极限分布。Ann. Math. Stat. 6 (1935), 160–169.[14] Fischer, A.、Gaunt, R. E.、Reinert, G. 和 Swan, Y. 指数族中后验分布的正态近似。arXiv:2209.08806, 2022.[15] Friedman, M. 使用秩以避免方差分析中隐含的正态性假设。J. Amer. Stat. Assoc. 32 (1937), 675–701.[16] Gaunt, R. E. 通过 Stein 方程解的新界，在矩条件下得到正态近似的收敛速率。J. Theoret. Probab. 29 (2016), 231–247.[17] Gaunt, R. E. 关于正态随机变量乘积的斯坦方法及零偏耦合。Bernoulli 23 (2017), 3311–3345.[18] Gaunt, R. E. 关于零均值相关正态随机变量乘积的分布的一则注记。Stat. Neerl. 73 (2019), 176–179.[19] Gaunt, R. E. 多元正态随机变量函数的斯坦方法。Ann. I. H. Poincare-PR 56 (2020), 1484–1513.

[20] Gaunt, R. E. 幂散度统计量族的卡方近似界。J. Appl. Probab. 59 (2022), 1059–1080.[21] Gaunt, R. E. 和 Li, S. 通过 Wasserstein 距离及相关积分概率度量界定 Kolmogorov 距离。J. Math. Anal. Appl. 522 (2023), Art. 126985.[22] Gaunt, R. E.、Pickett, A. 和 Reinert, G. 使用斯坦方法进行卡方近似及其在皮尔逊统计量中的应用。Ann. Appl. Probab. 27 (2017), 720–756.[23] Gaunt, R. E. 和 Reinert, G. 通过斯坦方法得到弗里德曼统计量卡方近似的界。Bernoulli 29 (2023), 2008–2034.[24] Gaunt, R. E. 和 Sutcliffe, H. 多元正态随机向量函数的斯坦方法的改进界。即将发表于 J. Theoret. Probab.，2023+.[25] Gibbs, A. L. 和 Su, F. E. 关于概率度量的选择与界定。Int. Stat. Rev. 70 (2002), 419–435.[26] Goldstein, L. 和 Reinert, G. 一维及更高维中的零偏化及其应用。载于《斯坦方法及其应用》。Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. 5 (2005), 1–18, Singapore Univ. Press, Singapore.[27] Goldstein, L. 和 Rinott, Y. 通过斯坦方法和大小偏倚耦合进行多元正态近似。J. Appl. Probab. 33 (1996), 1–17.[28] G¨otze, F. 关于多元中心极限定理的收敛速率。Ann. Probab. 19 (1991), 724–739.[29] G¨otze, F. 和 Ulyanov, V. V. χ^2型统计量的渐近分布。预印本 03–033，Bielefeld 大学谱分析、渐近分布与随机动力学研究组，Bielefeld，2003 年。[30] Hunter, D. R. 统计学研究生层次渐近学课程讲义。宾夕法尼亚州立大学，2014 年。[31] Kosorok, M. R.《经验过程与半参数推断导论》。统计学丛书。Springer, New York, 2008.[32] Ma, T. W. 通过组合学证明的高阶链式公式。Electron. J. Comb. 16, Issue 1 (2009), N21.[33] Marcinkiewicz, J. 和 Zygmund, A. 关于相互独立函数。Fundam. Math. 29 (1937), 60–90.[34] Meckes, E. 关于多元正态近似的斯坦方法。IMS Collect. 5 (2009), 153–178.[35] Olver, F. W. J.、Lozier, D. W.、Boisvert, R. F. 和 Clark, C. W.《NIST 数学函数手册》。Cambridge University Press, 2010.[36] Panaretos, V. M. 和 Zemel, Y. Wasserstein 距离的统计学方面。Annu. Rev. Stat. Appl. 6 (2019), 405–431.[37] Pearson, K. 关于这样一个判据：给定的偏差系统可合理地认为源于随机抽样。Phil. Mag. 50 (1900), 157–175.[38] Pinelis, I. 最大似然估计量趋于正态的收敛速率的最优阶一致界与非一致界。Electron. J. Stat. 11 (2017), 1160–1179.[39] Pinelis, I. 和 Molzon, R. 多元 Delta 方法中趋于正态的收敛速率的最优阶界。arXiv:0906.0177v5, 2016.[40] Pinelis, I. 和 Molzon, R. 多元 Delta 方法中趋于正态的收敛速率的最优阶界。Electron. J. Stat. 10 (2016), 1001–1063.[41] Portnoy, S. (2013). 谁发明了 Delta 方法？来信。Am. Stat. 67 (2013), 190–190.[42] Puchkin, N. 和 Ulyanov, V. 通过随机化检验统计量进行推断。即将发表于 Ann. I. H. Poincare-PR，2023+.[43] Reinert, G. 和 R¨ollin, A. 在一般线性条件下使用可交换对的斯坦方法进行多元正态近似。Ann. Probab. 37 (2009), 2150–2173.[44] R¨omisch, W. Delta 方法，无限维情形。《统计科学百科全书》，2006.[45] Sen, P. K. 通过 n 秩方法构造渐近有效检验。J. Roy. Statist. Soc. B 30 (1968), 312–317.[46] Small, C. G. 统计量的展开与渐近理论。统计与应用概率专著 115。CRC Press, Boca Raton, FL, 2010.[47] Stein, C. 关于和的正态近似中误差的一个界

相关随机变量。载于《第六届伯克利数学统计与概率研讨会论文集》（1972），第2卷，加州大学出版社，伯克利，583–602。[48] Ulyanov, V. V. 和 Zubov, V. N. 关于一类拟合优度统计量收敛到卡方分布的改进。Hiroshima Math. J. 39 (2009), 133–161。[49] Ver Hoef, J. M. 谁发明了 delta 方法? Am. Stat. 66 (2012), 124–127。[50] Wasserman, L., Kolar, M. 和 Rinaldo, A. 估计无向图的 Berry–Esseen 界。Electron. J. Stat. 8 (2014), 1188–1224。[51] Zolotarev, V. M. 概率度量。Teor. Veroyatn. Primen. 28 (1983), 264–287。

A 补充证明与计算

A.1 C从例 3.5 计算协方差矩阵

对于固定的 $j = 1$, ..., r, $\pi_1(j)$ ....., $\pi_n(j)$ 是相互独立的 Unif$\{1, \ldots, r\}$ 随机变量，因此我们有

$$\sigma_{jj} = \mathrm{Var}(W_j) = \frac{1}{\sigma_J^2 n} \sum_{i=1}^{n} \mathrm{Var}(J(\pi_i(j)) = \frac{1}{\sigma_J^2} \sum_{k=1}^{r} (J(k) - \overline{J})^2 \frac{1}{r} = \frac{r-1}{r}.$$

现在假设 $j \neq k$。由于 $\sum_{j=1}^{r} W_j = 0$ 和 W_1, ..., W_r 同分布，我们有

$$0 = \mathbb{E}\left[W_j \sum_{l=1}^{r} W_l\right] = \mathbb{E}[W_j^2] + \sum_{j \neq l} \mathbb{E}[W_j W_l] = \mathbb{E}[W_j^2] + (r-1)\mathbb{E}[W_j W_k].$$

由于 $\mathbb{E}[W_j] =$(r − 1)/r，重新整理可得 $\mathbb{E}[W_j W_k] = -1/r$，其中 $j \neq k$。由于 $\mathbb{E}[W_j] = 0$，因此可得 $\sigma_{jk} =$ Cov$(W_j, W_k) = \mathbb{E}[W_j W_k] = -1/r$。 □

A.2 引理4.3的证明

我们首先陈述以下引理，其中包含 [24] 给出的界。通过应用引理 4.1 和备注 4.2 中的讨论，我们将这些界从 $m = 1$ 情形推广到了一般的 $m \geq 1$ 。

引理 A. 1 [24], 2 1 2 2 $P = A +$ $B \sum_{i=1}^{d} |w_i|^r$ $r \geq 0$ $\sigma_{i\,i} = \Sigma_{i\,i}$ $i = 1 \ldots d$ $f = f_h$（Gaunt 和 Sutcliffe 的命题 . 和 .）。设 (w) , 其中 . 设 ( ), , , . 令 ( ) 表示多元正态斯坦方程 (4.22) 的解 (4.24)，并令 ψ_{jkl} 表示 (4.29) 的解。假设 Σ 半正定。

(i) 设 $h \in C_b^t(\mathbb{R}^m)$ 且 $g \in C^{t,}{}_P{}^m(\mathbb{R}^d)$，其中 $t \geq 1$。那么，对于 w $\in \mathbb{R}^d$,

$$\left| \frac{\partial^t f(\mathbf{w})}{\prod_{j=1}^{t} \partial w_{i_j}} \right| \leq \frac{h_{t,m}}{t} \left[A + \sum_{i=1}^{d} 2^{r/2} B\big(|w_i|^r + \mu_{r,\sigma_i}\big) \right]. \tag{A.35}$$

(ii) 设 $h \in C_b^6(\mathbb{R}^m)$ 和 $g \in C^{6,}{}_P{}^m(\mathbb{R}^d)$。那么，对于 w $\in \mathbb{R}^d$,

$$\left| \frac{\partial^3 \psi_{jkl}(\mathbf{w})}{\partial w_a \partial w_b \partial w_c} \right| \leq \frac{h_{6,m}}{18} \left[A + \sum_{i=1}^{d} 3^{r/2} B\big(|w_i|^r + 2\mu_{r,\sigma_i}\big) \right]. \tag{A.36}$$

引理 4.3 的证明。我们来证明第一个不等式。由不等式 (A.35) 可得

$$\mathbb{E}\left|u(\mathbf{X}_i)\frac{\partial^t f}{\prod_{j=1}^t \partial w_{i_j}}(\mathbf{W}_\theta^{(i)})\right| \le \frac{h_{t,m}}{t}\left[A\mathbb{E}[u(\mathbf{X}_i)] + \sum_{j=1}^d 2^{r/2}B\Big(\mathbb{E}|u(\mathbf{X}_i)(W_{j,\theta}^{(i)})^r| + \mu_{r,\sigma_j}\mathbb{E}|Z|^r\mathbb{E}[u(\mathbf{X}_i)]\Big)\right],$$

其中，$W_{j\,\theta}^{(i)}$ 是 $\mathbf{W}_\theta^{(i)}$ 的第 j 个分量。利用对任意 $s \ge 0$ 均成立的粗略不等式 $|a+b|^s \le 2^s(|a|^s + |b|^s)$，以及 X_{ij} 和 $W_j^{(i)}$ 的独立性，我们得到

$$\begin{aligned}2^{r/2}\mathbb{E}|u(\mathbf{X}_i)(W_{j,\theta}^{(i)})^r| &\le 2^{3r/2}\mathbb{E}\left|u(\mathbf{X}_i)\left(|W_j^{(i)}|^r + \frac{\theta^r}{n^{r/2}}|X_{ij}|^r\right)\right| \\ &\le 2^{3r/2}\mathbb{E}[u(\mathbf{X}_i)]\mathbb{E}|W_j^{(i)}|^r + \mathbb{E}|X_{ij}^r u(\mathbf{X}_i)|,\end{aligned}$$

其中我们利用假设 $n \ge 8$ 简化了界。利用 $\mathbb{E}|W_j^{(i)}|^r \le \mathbb{E}|W_j|^r$（见 [19, 第 1501] 页）可得第一个不等式。另一个不等式的证明类似；我们只需用不等式 (A.36) 替代不等式 (A.35)。我们还利用假设 $n \ge 12$ 简化了界，因此 $3^{r/2}\cdot 2^r/n^{r/2} = 12^{r/2}/n^{r/2} \le 1$。

□

A.3 控制函数的验证

定理 2.1 的证明验证。为简化记号，我们在假设 $d/n^{r_t/2} \le 1$ 的条件下证明该结论，其中 $a_{d,\ n,\ r_t} = 1$；容易看出，在一般情形下，$n, d \ge 1$, $r_t \ge 0$ 再乘上一个额外因子 $a^3_{n,\ d,\ r_t}$，即可得到所需的控制函数。我们分别考虑 $t = 1$、$t = 2$ 和 $t \ge 3$ 这几种情形。先考虑 $t \ge 3$ 的情形。将 $\partial_{\mathbf{f}}\overline{\partial w_{i_1}}(n^{-1/2}\mathbf{w})$ 在 0 处作泰勒展开，并利用对所有 $1 \le i_1$，都有 $\partial_{\mathbf{f}}\overline{\partial w_{i_1}}(0) = \cdots = \partial^{t-1}{}_{\mathbf{f}}\overline{\prod_{j=1}^{t-1}\partial w_{i_j}}(0) = 0$（在 $t \ge 2$ 的情形下），可得】【。

$$\frac{\partial \mathbf{g}_{t,n}}{\partial w_{i_1}}(\mathbf{w}) = \frac{1}{(t-1)!}\sum_{i_2,\ldots,i_t=1}^d w_{i_2}\cdots w_{i_t}\frac{\partial^t \mathbf{f}}{\partial w_{i_1}\cdots\partial w_{i_t}}\left(\frac{\theta\mathbf{w}}{\sqrt{n}}\right),$$

对于某个 $\theta \in (0,1)$（该值可能随行而变化）。由于 $\mathbf{f} \in \mathcal{C}^{t}{}_{,1,P}{}^{m}(\mathbb{R}^d)$，我们得到界

$$\begin{aligned}\left|\frac{\partial g_{t,j}}{\partial w_{i_1}}(\mathbf{w})\right| &\le \frac{A_t}{(t-1)!}\left(1 + \sum_{i=1}^d\left|\frac{w_i}{\sqrt{n}}\right|^{r_t}\right)\sum_{i_2,\ldots,i_t=1}^d |w_{i_2}\cdots w_{i_t}| \\ &\le \frac{A_t d^{t-2}}{(t-1)!}\sum_{i=1}^d\left(|w_i|^{t-1} + \frac{d|w_i|^{r_t+t-1}}{n^{r_t/2}}\right) \\ &\le \frac{A_t d^{t-2}}{(t-1)!}\sum_{i=1}^d\left(|w_i|^{t-1} + |w_i|^{r_t+t-1}\right), \qquad (A.37)\end{aligned}$$

其中 $g_{t,j}$ 表示函数 $\mathrm{g}_{t,n}$ 的第 j 个分量；在推导最后一步时，我们使用了假设 $d/n^{r_t/2} \le 1$。类似地，我们得到

$$\begin{aligned}\left|\frac{\partial^2 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\right| &= \frac{1}{(t-2)!}\left|\sum_{i_3,\ldots,i_t=1}^{d} w_{i_3}\cdots w_{i_t}\frac{\partial^t f_j}{\partial w_{i_1}\cdots\partial w_{i_t}}\left(\frac{\theta\mathbf{w}}{\sqrt{n}}\right)\right| \\ &\le \frac{A_t}{(t-2)!}\left(1+\sum_{i=1}^{d}\left|\frac{w_i}{\sqrt{n}}\right|^{r_t}\right)\sum_{i_3,\ldots,i_t=1}^{d}|w_{i_3}\cdots w_{i_t}| \\ &\le \frac{A_t d^{t-3}}{(t-2)!}\sum_{i=1}^{d}\big(|w_i|^{t-2}+d|w_i|^{r_t+t-2}\big). \end{aligned} \tag{A.38}$$

我们以类似的方式处理三阶偏导数：

$$\begin{aligned}\left|\frac{\partial^3 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\right| &\le \frac{1}{(t-3)!}\left|\sum_{i_4,\ldots,i_t=1}^{d} w_{i_4}\cdots w_{i_t}\frac{\partial^t f_j}{\partial w_{i_1}\cdots\partial w_{i_t}}\left(\frac{\theta\mathbf{w}}{\sqrt{n}}\right)\right| \\ &\le \frac{A_t}{(t-3)!}\left(1+\sum_{i=1}^{d}\left|\frac{w_i}{\sqrt{n}}\right|^{r_t}\right)\sum_{i_4,\ldots,i_t=1}^{d}|w_{i_4}\cdots w_{i_t}| \\ &\le \frac{A_t d^{t-4}}{(t-3)!}\sum_{i=1}^{d}\big(|w_i|^{t-3}+d|w_i|^{r_t+t-3}\big) \qquad (A.39) \\ &\le 2C_{1,t}d^{3t-4}\left(d+\sum_{i=1}^{d}|w_i|^{u_{1,t}}\right), \qquad (A.40)\end{aligned}$$

其中，最后一步使用了基本不等式 $|w_i|^{t-3} \le 1+|w_i|^{u_{1,\,t}}$ 和 $|w_i|^{r_t+t-3} \le 1+|w_i|^{u_{1,\,t}}$，因为 $t-3 \le r_t+t-3 \le u_{1,\,t}$。将不等式 $(|a_1|+\ldots+|a_q|)^s \le q^{s-1}(|a_1|^s+\ldots+|a_q|^s)$，$q, s \ge 1$ 应用于界 (A.37)，得到界

$$\left|\frac{\partial g_{t,j}}{\partial w_{i_1}}(\mathbf{w})\right|^3 \le \frac{4A_t^3 d^{3t-4}}{((t-1)!)^3}\sum_{i=1}^{d}\big(|w_i|^{3t-3}+|w_i|^{3r_t+3t-3}\big) \le \frac{8A_t^3 d^{3t-4}}{((t-1)!)^3}\left(d+\sum_{i=1}^{d}|w_i|^{u_{1,t}}\right). \tag{A.41}$$

同样地，由（A.38）可得

$$\begin{aligned}\left|\frac{\partial^2 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\right|^{3/2} &\le \frac{\sqrt{2}A_t^{3/2}d^{3t/2-4}}{((t-2)!)^{3/2}}\sum_{i=1}^{d}\big(|w_i|^{3t/2-3/2}+d^{3/2}|w_i|^{3(r_t+t-2)/2}\big) \\ &\le \frac{2\sqrt{2}A_t^{3/2}d^{3t-4}}{((t-2)!)^{3/2}}\left(d+\sum_{i=1}^{d}|w_i|^{u_{1,t}}\right). \end{aligned} \tag{A.42}$$

由不等式 (A.40)、(A.41) 和 (A.42) 可得，对于 $t \ge 3$，$\mathrm{g}_{t,n} \in C^3_{,P^m}(\mathbb{R}^d)$】【。

现在考虑情形 $t = 2$。不等式 (A.41) 对于 $t = 2$仍然成立，但不等式 (A.40) 和 (A.42) 是在假设 $t \ge 3$ 的条件下推导出来的。我们有

$\mathsf{g}_{2,\ n}(\mathsf{w}) = n(\mathsf{f}(n^{-1/2}\mathsf{w}) - \mathsf{f}(0))$，因此

$$\left|\frac{\partial^2 g_{2,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\right| = \left|\frac{\partial^2 f_j}{\partial w_{i_1}\partial w_{i_2}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right| \le A_2\left(1+\sum_{i=1}^d\left|\frac{w_i}{\sqrt{n}}\right|^{r_2}\right) \le A_2\left(1+\sum_{i=1}^d|w_i|^{r_2}\right),$$

由此我们得出，不等式 (A.38) 对 $t=2$ 也成立，进而意味着界 (A.42) 对 $t=2$ 有效。我们还有以下关于三阶偏导数的界：

$$\left|\frac{\partial^3 g_{2,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\right| = \left|\frac{1}{\sqrt{n}}\frac{\partial^3 f_j}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right| \le \frac{A_3}{\sqrt{n}}\left(1+\sum_{i=1}^d|w_i|^{r_3}\right). \tag{A.43}$$

根据不等式 (A.41)、(A.42) 和 (A.43)，并结合与处理 $t\ge 3$ 情形时类似的推理，我们得出 $\mathsf{g}_{2,\ n} \in C^3_{,P}{}^m(\mathbb{R}^d)$。

最后，我们考虑 $t=1$ 情形。与证明不等式 (A.42) 对 $t=2$ 也成立的方法类似，可以证明不等式 (A.41) 对 $t=1$ 也成立；细节从略。对 $\mathsf{g}_{1,n}(\mathsf{w}) = \sqrt{n}(\mathsf{f}(n^{-1/2}\mathsf{w}) - \mathsf{f}(0))$ 求导，并利用 $\mathsf{f} \in \mathcal{C}^1{}_{,2,P}{}^m(\mathbb{R}^d)$，可得到以下关于二阶和三阶偏导数的界：

$$\begin{aligned}\left|\frac{\partial^2 g_{1,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\right| &= \left|\frac{1}{\sqrt{n}}\frac{\partial^2 f_j}{\partial w_{i_1}\partial w_{i_2}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right| \\ &\le \frac{A_2}{\sqrt{n}}\left(1+\sum_{i=1}^d|w_i|^{r_2}\right) \le \frac{A_2}{n^{1/3}}\left(1+\frac{1}{d}\sum_{i=1}^d|w_i|^{r_2}\right), \qquad (A.44)\\ \left|\frac{\partial^3 g_{1,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\right| &= \left|\frac{1}{n}\frac{\partial^3 f_j}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right| \\ &\le \frac{A_3}{n}\left(1+\sum_{i=1}^d|w_i|^{r_3}\right) \le \frac{A_3}{n^{5/6}}\left(1+\frac{1}{d}\sum_{i=1}^d|w_i|^{r_3}\right), \qquad (A.45)\end{aligned}$$

其中我们使用了假设 $n\ge d^6$。由不等式 (A.41)、(A.44) 和 (A.45)，并采用本证明中一直使用的类似论证，可直接得出 $\mathsf{g}_{1,n}\in C^3_{,P}{}^m(\mathbb{R}^d)$。至此，我们最终证明了对于 $t\ge 1$，$\mathsf{g}_{t,n}\in C^3_{,P}{}^m(\mathbb{R}^d)$，证明至此完成。 □

定理 2.2 证明的验证。与定理 2.1 证明的验证一样，我们假设 $d/n^{r_t/2}\le 1$ 以简化记号，一般情形可轻易推出。我们首先考虑 $t\ge 6$ $\left(|a_1|+\cdots+|a_q|\right)^s \le$ 的情形。 $q^{s-1}\left(|a_1|^s+\cdots+|a_q|^s\right)$ q $s\ge 1$ 37 将不等式，对于，，应用于界 (A.)，得到

$$\begin{aligned}\left|\frac{\partial g_{t,j}}{\partial w_{i_1}}(\mathbf{w})\right|^6 &\le \frac{A_t^6}{((t-1)!)^6}d^{6(t-2)}\left(\sum_{i=1}^d\left(|w_i|^{t-1}+|w_i|^{r_t+t-1}\right)\right)^6 \\ &\le \frac{32A_t^6}{((t-1)!)^6}d^{6t-7}\sum_{i=1}^d\left(|w_i|^{6(t-1)}+|w_i|^{6(r_t+t-1)}\right)\end{aligned}$$

$$\leq 2C_{2,t}d^{6t-7}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\bigg). \tag{A.46}$$

通过对界 (A.38) 作类似的考虑，我们得到

$$\begin{aligned}\left|\frac{\partial^2 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\right|^3 &\leq \frac{A_t^3}{((t-2)!)^3}d^{3(t-3)}\bigg(\sum_{i=1}^{d}\big(|w_i|^{t-2}+d|w_i|^{r_t+t-2}\big)\bigg)^3\\ &\leq \frac{4A_t^3}{((t-2)!)^3}d^{3t-7}\sum_{i=1}^{d}\big(|w_i|^{3(t-2)}+d^3|w_i|^{3(r_t+t-2)}\big)\\ &\leq 2C_{2,t}d^{6t-7}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\bigg). \end{aligned}\tag{A.47}$$

再次，对界 (A.39) 进行类似的分析可得

$$\begin{aligned}\left|\frac{\partial^3 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\right|^2 &\leq \frac{A_t^2}{((t-3)!)^2}d^{2(t-4)}\bigg(\sum_{i=1}^{d}\big(|w_i|^{t-3}+d|w_i|^{r_t+t-3}\big)\bigg)^2\\ &\leq \frac{2A_t^2}{((t-3)!)^2}d^{2t-7}\bigg(|w_i|^{2(t-3)}+d^2|w_i|^{2(r_t+t-3)}\bigg)\\ &\leq 2C_{2,t}d^{6t-7}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\bigg). \end{aligned}\tag{A.48}$$

现在我们得到 $g_{t,j}$ 的四阶、五阶和六阶偏导数的界。通过将 $^{\partial^4}\mathbf{f}\overline{\partial w_{i_1}\cdots\partial w_{i_4}}(n^{-1/2}\mathbf{w})$ 在 0 处作泰勒展开，并利用 $^{\partial^4}\mathbf{f}\overline{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{0}) = \cdots = {}^{\partial^{t-1}}\mathbf{f}\overline{\prod_{j=1}^{t-1}\partial w_{i_j}}(\mathbf{0}) = 0$ 对所有 $1\leq i_1,\ldots,i_{t-1}\leq d$ 成立，我们得到，对于 $t\geq 5$,

$$\begin{aligned}\left|\frac{\partial^4 g_{t,j}}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\right| &\leq \frac{1}{(t-4)!}\sum_{i_5,\ldots,i_t=1}^{d}\left|w_{i_5}\cdots w_{i_t}\frac{\partial^t f_j}{\partial w_{i_1}\cdots\partial w_{i_t}}\left(\frac{\theta\mathbf{w}}{\sqrt{n}}\right)\right|\\ &\leq \frac{A_t}{(t-4)!}\bigg(1+\sum_{i=1}^{d}\left|\frac{w_i}{\sqrt{n}}\right|^{r_t}\bigg)\sum_{i_5,\ldots,i_t=1}^{d}|w_{i_5}\cdots w_{i_t}|\\ &\leq \frac{A_t}{(t-4)!}d^{t-5}\sum_{i=1}^{d}\big(|w_i|^{t-4}+d|w_i|^{r_t+t-4}\big). \end{aligned}\tag{A.49}$$

注意，不等式 (A.49) 对于 $t=4$ 也成立；这一点可以通过类似于定理 2.1 的证明中用于说明不等式 (A.42) 对于 $t=2$ 成立的论证得出。采用类似的方法，我们得到，对于 $t\geq 6$,

$$\left|\frac{\partial^5 g_{t,j}}{\partial w_{i_1}\cdots\partial w_{i_5}}(\mathbf{w})\right| \leq \frac{A_t}{(t-5)!}d^{t-6}\sum_{i=1}^{d}\big(|w_i|^{t-5}+d|w_i|^{r_t+t-5}\big). \tag{A.50}$$

我们以类似的方式处理六阶偏导数:

$$\begin{aligned}\left|\frac{\partial^6 g_{t,j}}{\partial w_{i_1}\cdots\partial w_{i_6}}(\mathbf{w})\right| &\le \frac{1}{(t-6)!}\sum_{i_7,\ldots,i_t=1}^{d}\left|w_{i_7}\cdots w_{i_t}\frac{\partial^t f_j}{\partial w_{i_1}\cdots\partial w_{i_t}}\left(\frac{\theta\mathbf{w}}{\sqrt{n}}\right)\right| \\ &\le \frac{A_t}{(t-6)!}\left(1+\sum_{i=1}^{d}\left|\frac{w_i}{\sqrt{n}}\right|^{r_t}\right)\sum_{i_7,\ldots,i_t=1}^{d}|w_{i_7}\cdots w_{i_t}| \\ &\le \frac{A_t}{(t-6)!}d^{t-7}\sum_{i=1}^{d}\left(|w_i|^{t-6}+d|w_i|^{r_t+t-6}\right) \\ &\le 2C_{2,t}d^{6t-7}\left(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\right), \end{aligned} \tag{A.51}$$

对于某个 $\theta\in$ (0,1)，再次注意到界 (A.51) 对 $t\ge 6$ 有效。将基本不等式 $\left(|a_1|+\cdots+|a_q|\right)^s \le q^{s-1}\left(|a_1|^s+\cdots+|a_q|^s\right)$（对于 $q, s\ge 1$）应用于不等式 (A.49)，得到界

$$\begin{aligned}\left|\frac{\partial^4 g_{t,j}}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\right|^{3/2} &\le \frac{\sqrt{2}A_t^{3/2}}{((t-4)!)^{3/2}}d^{3t/2-7}\sum_{i=1}^{d}\left(|w_i|^{3(t-4)/2}+d^{3/2}|w_i|^{3(r_t+t-4)/2}\right) \\ &\le 2C_{2,t}d^{6t-7}\left(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\right). \end{aligned} \tag{A.52}$$

类似地，由 (A.50) 可得

$$\begin{aligned}\left|\frac{\partial^5 g_{t,j}}{\partial w_{i_1}\cdots\partial w_{i_5}}(\mathbf{w})\right|^{6/5} &\le \frac{2^{1/5}A_t^{6/5}}{((t-5)!)^{6/5}}d^{6t/5-7}\sum_{i=1}^{d}\left(|w_i|^{6(t-5)/5}+d^{6/5}|w_i|^{6(r_t+t-5)/5}\right) \\ &\le 2C_{2,t}d^{6t-7}\left(d+\sum_{i=1}^{d}|w_i|^{u_{2,t}}\right). \end{aligned} \tag{A.53}$$

由 (A.46)、(A.47)、(A.48)、(A.51)、(A.52) 和 (A.53)，我们推得 $\mathrm{g}_{t,\ n}\in C^6_{,\ p^m}(\mathbb{R}^d)$，对于 $t\ge 6$。

现在考虑 $t=4$ 的情形。不等式 (A.46)、(A.47)、(A.48) 和 (A.52) 对 $t=4$ 均成立，但不等式 (A.51) 和 (A.53) 是在假设 $t\ge 6$ 的条件下推导的。我们有 $\mathrm{g}_{4,\ n}(\mathrm{w}) = n^2(\mathrm{f}(n^{-1/2}\mathrm{w}) - \mathrm{f}(0))$，因此

$$\left|\frac{\partial^5 g_{4,j}}{\partial w_{i_1}\cdots\partial w_{i_5}}(\mathbf{w})\right| \le \left|\frac{1}{\sqrt{n}}\frac{\partial^5 f_j}{\partial w_{i_1}\cdots\partial w_{i_5}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right| \le \frac{A_5}{\sqrt{n}}\left(1+\sum_{i=1}^{d}|w_i|^{r_5}\right).$$

那么我们有

$$\left|\frac{\partial^5 g_{4,j}}{\partial w_{i_1}\cdots\partial w_{i_5}}(\mathbf{w})\right|^{6/5} \le \left|\frac{1}{\sqrt{n}}\frac{\partial^5 f_j}{\partial w_{i_1}\cdots\partial w_{i_5}}\left(\frac{\mathbf{w}}{\sqrt{n}}\right)\right|^{6/5} \le \frac{A_5^{6/5}}{n^{3/5}}\left(1+\sum_{i=1}^{d}|w_i|^{r_5}\right)^{6/5}$$

$$\leq \frac{2^{1/5}A_5^{6/5}}{n^{3/5}}d^{1/5}\bigg(1+\sum_{i=1}^{d}|w_i|^{6r_5/5}\bigg)$$
$$\leq 2C_{2,4}d^{17}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{2,4}}\bigg). \tag{A.54}$$

我们还有关于六阶偏导数的如下界：

$$\bigg|\frac{\partial^6 g_{4,j}}{\partial w_{i_1}\cdots\partial w_{i_6}}(\mathbf{w})\bigg| \leq \bigg|\frac{1}{n}\frac{\partial^6 f_j}{\partial w_{i_1}\cdots\partial w_{i_6}}\bigg(\frac{\mathbf{w}}{\sqrt{n}}\bigg)\bigg| \leq \frac{A_6}{n}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_6}\bigg)$$
$$\leq 2C_{2,4}d^{17}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{2,4}}\bigg). \tag{A.55}$$

由不等式 (A.46)、(A.47)、(A.48)、(A.52)、(A.54) 和 (A.55)，我们得出 $\mathrm{g}_{4,n}\in {C^6_{,P}}^m(\mathbb{R}^d)$。

最后，我们考虑情形 $t=2$。对于 $t=2$，不等式 (A.46) 仍然成立。在定理 2.1 的证明中已经证明，(A.38) 对 $t=2$ 成立，因此这又意味着 (A.47) 对 $t=2$ 也成立。对 $\mathrm{g}_{2,n}(\mathrm{w})=n(\mathrm{f}(n^{-1/2}\mathrm{w})-\mathrm{f}(0))$ 求导，并利用 $\mathrm{f}\in {\mathcal{C}^2_{,4,P}}^m(\mathbb{R}^d)$，可得以下关于四阶、五阶和六阶偏导数的界：

$$\bigg|\frac{\partial^4 g_{2,j}}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\bigg| = \bigg|\frac{1}{n}\frac{\partial^4 f_j}{\partial w_{i_1}\cdots\partial w_{i_4}}\bigg(\frac{\mathbf{w}}{\sqrt{n}}\bigg)\bigg| \leq \frac{A_4}{n}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_4}\bigg), \tag{A.56}$$
$$\bigg|\frac{\partial^5 g_{2,j}}{\partial w_{i_1}\cdots\partial w_{i_5}}(\mathbf{w})\bigg| = \bigg|\frac{1}{n^{3/2}}\frac{\partial^5 f_j}{\partial w_{i_1}\cdots\partial w_{i_5}}\bigg(\frac{\mathbf{w}}{\sqrt{n}}\bigg)\bigg| \leq \frac{A_5}{n^{3/2}}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_5}\bigg), \tag{A.57}$$
$$\bigg|\frac{\partial^6 g_{2,j}}{\partial w_{i_1}\cdots\partial w_{i_6}}(\mathbf{w})\bigg| = \bigg|\frac{1}{n^2}\frac{\partial^6 f_j}{\partial w_{i_1}\cdots\partial w_{i_6}}\bigg(\frac{\mathbf{w}}{\sqrt{n}}\bigg)\bigg| \leq \frac{A_6}{n^2}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_6}\bigg). \tag{A.58}$$

由不等式 (A.46) 和 (A.47)，并将本证明中一贯使用的类似推理应用于不等式 (A.43)、(A.56)、(A.57) 和 (A.58)，可直接得出 $\mathrm{g}_{2,n}\in {C^6_{,P}}^m(\mathbb{R}^d)$。至此，我们最终证明了，对于$t\geq 2$，$\mathrm{g}_{t,n}\in {C^6_{,P}}^m(\mathbb{R}^d)$。 □

定理 2.3 的证明验证。为简化记号，我们再次假设 $d/n^{r_t/2}\leq 1$，一般情形可轻易推得。我们先考虑 $t\geq 4$ 的情形。将不等式 $\big(|a_1|+\cdots+|a_q|\big)^s\leq q^{s-1}\big(|a_1|^s+\cdots+|a_q|^s\big)$（其中 q、$s\geq 1$）应用于界 (A.37)，可得

$$\bigg|\frac{\partial g_{t,j}}{\partial w_{i_1}}(\mathbf{w})\bigg|^4 \leq \frac{A_t^4}{((t-1)!)^4}d^{4(t-2)}\bigg(\sum_{i=1}^{d}\big(|w_i|^{t-1}+|w_i|^{r_t+t-1}\big)\bigg)^4$$
$$\leq \frac{8A_t^4}{((t-1)!)^4}d^{4t-5}\sum_{i=1}^{d}\big(|w_i|^{4(t-1)}+|w_i|^{4(r_t+t-1)}\big)$$

$$\leq 2C_{3,t}d^{4t-5}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{3,t}}\bigg). \tag{A.59}$$

通过对界 (A.38)、(A.39) 和 (A.49) 应用类似的考量，我们得到

$$\begin{aligned}
\bigg|\frac{\partial^2 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}}(\mathbf{w})\bigg|^2 &\leq \frac{A_t^2}{((t-2)!)^2}d^{2(t-3)}\bigg(\sum_{i=1}^{d}\big(|w_i|^{t-2}+d|w_i|^{r_t+t-2}\big)\bigg)^2\\
&\leq \frac{2A_t^2}{((t-2)!)^2}d^{2t-5}\sum_{i=1}^{d}\big(|w_i|^{2(t-2)}+d^2|w_i|^{2(r_t+t-2)}\big)\\
&\leq 2C_{3,t}d^{4t-5}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{3,t}}\bigg), \qquad (A.60)\\
\bigg|\frac{\partial^3 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\bigg|^{4/3} &\leq \frac{A_t^{4/3}}{((t-3)!)^{4/3}}d^{4(t-4)/3}\bigg(\sum_{i=1}^{d}\big(|w_i|^{t-3}+d|w_i|^{r_t+t-3}\big)\bigg)^{4/3}\\
&\leq \frac{2^{1/3}A_t^{4/3}}{((t-3)!)^{4/3}}d^{4t/3-5}\sum_{i=1}^{d}\big(|w_i|^{4(t-3)/3}+d^{4/3}|w_i|^{4(r_t+t-3)/3}\big)\\
&\leq 2C_{3,t}d^{4t-5}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{3,t}}\bigg), \qquad (A.61)\\
\bigg|\frac{\partial^4 g_{4,j}}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\bigg| &\leq \frac{A_t}{(t-4)!}d^{t-5}\sum_{i=1}^{d}\big(|w_i|^{t-4}+d|w_i|^{r_t+t-4}\big)\\
&\leq 2C_{3,t}d^{4t-5}\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{3,t}}\bigg). \qquad (A.62)
\end{aligned}$$

由 (A.59)、(A.60)、(A.61) 和 (A.62)，我们可以推导出，对于 $t\geq 4$，$\mathsf{g}_{t,\,n}\in C^4_{,\,P}{}^m(\mathbb{R}^d)$。

现在考虑 $t=2$ 的情形。不等式 (A.59) 对 $t=2$ 成立，但不等式 (A.60)、(A.61) 和 (A.62) 是在假设 $t\geq 4$ 的条件下推导的。在定理 2.1 的证明中已证明，(A.38) 对 $t=2$ 成立，而这又意味着界 (A.60) 对 $t=2$ 成立。对 $\mathsf{g}_{2,\,n}(\mathsf{w})=n\big(\mathsf{f}(n^{-1/2}\mathsf{w})-\mathsf{f}(0)\big)$ 求导，我们得到

$$\bigg|\frac{\partial^3 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\bigg|=\bigg|\frac{1}{\sqrt{n}}\frac{\partial^3 f_j}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}\bigg(\frac{\mathbf{w}}{\sqrt{n}}\bigg)\bigg|\leq\frac{A_3}{\sqrt{n}}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_3}\bigg),$$

由此得到

$$\begin{aligned}
\bigg|\frac{\partial^3 g_{t,j}}{\partial w_{i_1}\partial w_{i_2}\partial w_{i_3}}(\mathbf{w})\bigg|^{4/3} &\leq \frac{A_3^{4/3}}{n^{2/3}}\bigg(1+\sum_{i=1}^{d}|w_i|^{r_3}\bigg)^{4/3}\leq\frac{2^{1/3}A_3^{4/3}}{n^{2/3}}\bigg(1+d^{1/3}\sum_{i=1}^{d}|w_i|^{4r_3/3}\bigg)\\
&\leq 2C_{3,2}d^3\bigg(d+\sum_{i=1}^{d}|w_i|^{u_{3,2}}\bigg). \qquad (A.63)
\end{aligned}$$

此外，四阶偏导数还满足以下界：

$$\left|\frac{\partial^4 g_{4,j}}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\right| = \left|\frac{1}{n}\frac{\partial^4 f_j}{\partial w_{i_1}\cdots\partial w_{i_4}}(\mathbf{w})\right| \le \frac{A_4}{n}\left(1+\sum_{i=1}^{d}|w_i|^{r_4}\right)$$
$$\le 2C_{3,2}d^3\left(d+\sum_{i=1}^{d}|w_i|^{u_{3,2}}\right). \tag{A.64}$$

由不等式 (A.59)、(A.60)、(A.63) 和 (A.64)，可得 $g_{2,n}\in C^4_{,p}{}^m(\mathbb{R}^d)$。至此，我们最终证明了对于 $t\ge 2$，$g_{t,n}\in C^4_{,p}{}^m(\mathbb{R}^d)$。 □

定理 2.4 证明的验证。先将 $f(n^{-1/2}w)$ 在 0 处作泰勒展开，并利用 $f'(0) = \cdots = f^{(t-2)}(0) = 0$（在 $t\ge 2$ 的情况下），然后求导，可得

$$g'_{t,n}(w) = \frac{w^{t-1}}{(t-1)!}f^{(t)}\left(\frac{\theta w}{\sqrt{n}}\right),$$

其中 $\theta\in(0,1)$。由于 $f\in\mathcal{C}^t_{,P,1\,1,}(\mathbb{R})$，我们得到界

$$|g'_{t,n}(w)| \le \frac{A_t}{(t-1)!}\left(|w|^{t-1}+\frac{1}{n^{r_t/2}}|w|^{r_t+t-1}\right) \le \frac{2A_t}{(t-1)!}(1+|w|^{r_t+t-1}). \tag{A.65}$$

不等式 (A.65) 是在假设 $t\ge 2$ 的条件下推导的，但我们可以证明它对于 $t=1$ 也成立。对 $g_{1,n}(w) = \sqrt{n}(f(n^{-1/2}w) - f(0))$ 求导可得

$$|g'_{1,n}(w)| = |f'(n^{-1/2}w)| \le A_1\left(1+\frac{1}{n^{r_1/2}}|w|^{r_1}\right) \le 2A_1(1+|w|^{r_1}),$$

因此，(A.65) 对于 $t\ge 1$ 成立。因此，$g_{t,\,n}\in C^1_{,\,P}{}^1(\mathbb{R})$，如所要求的那样。 □

定理 2.5 证明的验证。将不等式 $(a+b)^2 \le 2(a^2+b^2)$ 应用于 (A.65)，并结合我们通常的考虑，可得

$$|g'_{t,n}(w)|^2 \le \frac{2A_t^2}{((t-1)!)^2}(|w|^{2(t-1)}+|w|^{2(r_t+t-1)}) \le \frac{4A_t^2}{((t-1)!)^2}(1+|w|^{2(r_t+t-1)}). \tag{A.66}$$

现在只需界定 $g_{t,n}$ 的二阶导数。首先将 $f(n^{-1/2}w)$ 在 0 处作泰勒展开，并利用 $f'(0) = \cdots = f_{(t-3)}(0) = 0$，然后求导，可得对于 $t\ge 3$,

$$g''_{t,n}(w) = \frac{w^{t-2}}{(t-2)!}f^{(t)}\left(\frac{\theta w}{\sqrt{n}}\right).$$

由于 $f\in\mathcal{C}^t_{,P,2\,1,}(\mathbb{R})$，我们得到该界。

$$|g''_{t,n}(w)| \le \frac{A_t}{(t-2)!}\left(|w|^{t-2}+\frac{1}{n^{r_t/2}}|w|^{r_t+t-2}\right). \tag{A.67}$$

根据我们通常的论证，可以看出不等式 (A.67) 对于 $t=2$ 也成立。将基本不等式 $|w|^{t-2}\le 1+|w|^{2(r_t+t-1)}$ 和 $|w|^{r_t+t-2}\le 1+|w|^{2(r_t+t-1)}$ 应用于 (A.67)，可得

$$|g''_{t,n}(w)| \le \frac{2A_t}{(t-2)!}(1+|w|^{2(r_t+t-1)}). \tag{A.68}$$

由不等式 (A.66) 和 (A.68)，我们推得 $g_{t,n}\in C^2_{,P}{}^1(\mathbb{R})$，对于 $t\ge 2$。 □