arXiv:2603.04576v1 [stat.ME] 4 Mar 2026

调查中线性回归插补的变量选择

Ziming An[1], **MehdiDagdoug**[2] **和 David Haziza**[1]

[1]渥太华大学数学与统计系，加拿大渥太华，电子邮件：
zan032@uottawa.ca；dhaziza@uottawa.ca

[2]麦吉尔大学数学与统计系，加拿大蒙特利尔，电子邮件：
mehdi.dagdoug@mcgill.ca

摘要：调查抽样关注有限总体参数的估计。实际上，调查数据常存在项目无应答，通常通过插补来处理，即用预测值替换缺失值。因此，所得插补估计量的性质在很大程度上取决于所采用预测方法的性质。而预测方法本身又取决于拟合插补模型时所选变量和调优参数。本文研究线性回归插补中的变量选择问题。尽管变量选择已在许多领域得到广泛研究，主要用于识别或预测，但其在调查数据插补中的作用受到的关注相对较少。我们通过预言损失函数定义最优插补模型，并证明当样本量趋于无穷时，该模型以趋近于 1 的概率与真实模型一致。我们还考察了使用设定错误的模型——遗漏相关协变量或纳入无关协变量——对一致性和渐近方差的影响。随后，我们构建了一套完整的方法论框架，用于在模型选择后构造置信区间。研究表明，所提出的置信区间在渐近意义下有效，并且在所有候选模型中最优。模拟研究表明，所提出的方法论在有限样本中表现良好。

关键词和短语：调查抽样、缺失数据、插补、变量选择、方差估计、渐近有效置信区间。

1. 引言

无应答是官方统计中的一项重大挑战，因为它影响大多数调查。如果忽略缺失数据，所得估计可能存在偏差且不一致。在调查抽样中，通常区分个体无应答与项目无应答。前者对应于未对抽样个体收集任何信息的情形，而后者则发生在某些变量缺失但其他变量被观测到时。项目无应答最常通过插补来处理，即用预测值替换缺失值的程序。只要插补模型设定正确，且方差估计恰当反映抽样变异与无应答，插补就能恢复一致性并支持有效推断。在实践中，这意味着插补模型至关重要：如果模型误设，所得估计可能存在偏差和/或效率低下。

人们已在广泛的插补模型下研究了插补估计量的性质。大量研究聚焦于参数插补，例如线性回归插补（如 Chauvet、Deville 和 Haziza （2011）），其中包括许多关于

方差估计的研究（Fay，1991；Shao 和 Steel，1999；Berger 和 Rao，2006；Kim 和 Rao，2009）。其他研究考察了非参数回归及相关方法，包括最近邻方法（Chen 和 Shao，2000，2001；Yang 和 Kim，2019）、得分法（Haziza 和 Beaumont，2007）、预测均值匹配（Yang 和 Kim，2017）以及随机森林（Dagdoug、Goga 和 Haziza，2025）。

在模型辅助估计的背景下，我们特别关注Opsomer 和 Miller(2005)的研究。该研究为调查抽样中的超参数选择提供了重要参考，尽管并未涉及插补情形。在常见的独立同分布情形下，变量选择已得到广泛研究，其主要目标有两个：为给定任务找出最佳预测模型（效率），或恢复真实的系数集合（一致性）（Yang，2005）。开创性研究包括Nishii (1984）、Rao 和 Wu(1989）、Shao (1993年）以及Shao(1997年），他们确立了众多模型选择准则的渐近性质，例如 AIC（Akaike，1970）、BIC（Schwarz，1978）以及交叉验证。

在有限总体抽样中，目标是估计有限总体参数，主要关注的是所得估计量的表现，尤其是其偏差和效率。因此，在选择插补模型时，调查统计学家的首要目标并非识别真实的数据生成模型，而是选择能够产生效率最高的插补估计量的模型，也就是均方误差最小的模型。尽管我们会借鉴独立同分布文献中的若干思路，但我们的目标有本质区别：我们旨在优化有限总体推断中的插补效率，而非恢复底层回归函数。然而，正如下文所示，对于线性模型而言，这两个目标密切相关。

本文的主要贡献如下。我们提出了一种预言损失函数，用于评估调查抽样下候选插补模型的效率。我们证明，在温和条件下，使该损失最小的模型渐近地就是真实模型，从而将插补的模型选择与模型识别联系起来。随后，我们研究了使用设定错误的模型所产生的影响。特别地，我们给出了一般条件，在这些条件下，所得插补估计量仍然具有一致性，并确定了过拟合（即添加不必要的协变量）会或不会增加渐近方差的情形。在标准正则条件下，我们证明，在独立同分布框架下一致的模型选择准则，在存在缺失数据的调查抽样下仍保持一致性。最后，当使用一致的选择准则时，所得插补估计量及其方差与真实模型下获得的结果渐近等价，从而实现预言效率。我们还证明，线性回归插补的标准方差估计量是一致的，并建立了所得点估计量的渐近正态性。我们提出了一套在模型选择后构造置信区间的完整方法论，并证明该方法渐近地达到名义覆盖率，且在候选模型类中具有渐近最小的区间宽度。在等概率和不等概率抽样设计下开展的模拟研究证实了所提方法的良好表现。

本文其余部分安排如下。第 2节介绍记号和一般框架。第 3 节定义预言损失函数及其性质。第4

节介绍插补模型选择的主要渐近结果。第 5 节报告模拟结果，第 6节总结全文并展望未来工作。所有证明均放在附录中。

2. 预备知识

2.1. 问题设定

考虑一个规模为 N、调查变量为 Y 的有限总体 $U:=\{1, 2, \ldots\ldots, N\}$。我们关注的是估计有限总体均值

$$\mu := \frac{1}{N}\sum_{k\in U} y_k,$$

其中，y_k 表示元素 $k \in U$ 的调查变量 Y 的测量值。根据抽样设计 $\mathcal{P}$，选取一个规模为 n 的样本 $S \subset U$。假设对所有 k，$l \in U$，由 $\pi_k := \mathbb{P}(k \in S)$ 和 $\pi_{kl} := \mathbb{P}(k, l \in S)$ 定义的一阶和二阶包含概率均严格为正。样本 S 也可用样本选择指示变量向量 $\mathbf{I}_U := [I_1, I_2, \ldots\ldots, I_N]^\top$ 表示，其中，若 $k \in S$，则 $I_k := 1$；否则，$I_k := 0$。

μ 的 Horvitz-Thompson 估计量定义为

$$\widehat{\mu}_\pi := \frac{1}{N}\sum_{k\in S} \frac{y_k}{\pi_k}. \tag{1}$$

实践中，当调查变量 Y 存在无应答时，式 (1) 中的估计量 $\widehat{\mu}_\pi$ 无法实现。我们将响应指示变量向量记为 $\mathbf{r}_U := [r_1, r_2, \ldots, r_N]^\top$，其中$r_k := 1$ 表示当 y_k 被观测到时，$r_k := 0$表示其他情况。应答者和无应答者集合分别记为 $S_r := \{k \in S : r_k = 1\}$ 和 $S_m := \{k \in S : r_k = 0\}$，其基数分别为$n_r$和 n_m。我们假设每个 $k \in S$都观测到了 p 个协变量 $X_1, X_2, \ldots, X_p$，其测量值为 $\mathbf{x}_k$。观测数据为

$$\mathcal{D}_{imp} := \{(\mathbf{x}_k, y_k) : k \in S_r\} \cup \{\mathbf{x}_k : k \in S_m\}.$$

向量 $[\mathbf{x}_k^\top, y_k, r_k]_{k\in U}^\top$被假定为独立同分布。假定抽样设计 P 是非信息性的，即对于每个 $k \in U$，都有 $\pi_k = \mathbb{P}(I_k = 1|\boldsymbol{X}_U, \boldsymbol{y}_U) = \mathbb{P}(I_k = 1|\boldsymbol{X}_U)$，其中 $\boldsymbol{X}_U$ 和 $\boldsymbol{y}_U$ 分别表示总体设计矩阵和调查变量的总体向量。有关详情，请参阅 Pfeffermann 和 Sverchkov（2009）。同样，假定缺失机制满足随机缺失（MAR，Rubin（1976））假设，即 $\mathbb{P}(r_k = 1|\mathbf{x}_k, y_k) = \mathbb{P}(r_k = 1|\mathbf{x}_k) := p(\mathbf{x}_k)$。最后，我们假设存在一个正常数 $\rho > 0$，使得$p(\mathbf{x}) \geq \rho$几乎必然成立。这就是通常所说的正值性假设。本文将研究范围限定于如下定义的常见同方差线性回归模型

$$y_k = \mathbf{x}_k^\top \boldsymbol{\beta} + \epsilon_k, \qquad k \in U, \tag{2}$$

其中 ϵ_k 满足 $\mathbb{E}[\epsilon_k|\mathbf{x}_k] = 0$ 且 $\mathbb{E}[\epsilon_k^2|\mathbf{x}_k] := \sigma^2$。

2.2. 补充记号

关于以下各项的概率、期望和方差：(i) 协变量分布$\mathbb{P}_{\mathbf{x}}$ 将以下标 x 表示；(ii) $\mathbb{P}_{y|\mathbf{x}}$ 将以下标 m表示；(iii) 抽样设计将以下标 p表示；(iv) 无应答机制将以下标 q表示。对于整数集 A，我们用 P(A) 表示其幂集，按惯例不包含空集。对于定义在相同定义域 D 上的两个函数 f 和 g，若 $\operatorname{argmin}_{x\in D} f(x) = \operatorname{argmin}_{x\in D} g(x)$，则记为 $f(x) \operatorname{arg\,min}\equiv g(x)$。像 $\mathrm{x} \in \mathbb{R}^p$ 这样的向量将以粗体表示，其第 j 个分量记为 $x(j)$。我们用 $\mathrm{x}, y, p(\mathrm{x}), r, \epsilon$ 表示 $(\mathrm{x}_k, y_k, p(\mathrm{x}_k), r_k, \epsilon_k)_{k\in U}$ 的一个独立副本。

3. 渐近最优插补

3.1. 线性回归插补与变量选择

线性回归插补使用基于应答者数据拟合的线性回归模型所作的预测值替换缺失值。设 $\boldsymbol{X}_r := (\mathbf{x}_k^\top)_{k\in S_r} \in \mathbb{R}^{n_r\times p}$ 和 $\boldsymbol{Y}_r \in \mathbb{R}^{n_r}$ 分别表示应答者的设计矩阵和测量值向量。Y 若矩阵 $\boldsymbol{X}_r^\top \boldsymbol{X}_r$ 为正定，则（2）中的普通最小二乘估计量为

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}_r^\top \boldsymbol{X}_r\right)^{-1} \boldsymbol{X}_r^\top \boldsymbol{Y}_r. \tag{3}$$

在调查抽样中，使用设计权重 $w_k = \pi_k^{-1}$对 $k \in S$进行加权最小二乘以拟合插补模型的情况并不少见。然而，由于假设抽样设计不具信息性，非加权最小二乘拟合是合理的。为简化代数运算，我们因此关注（3）中的非加权估计量。加权版本也会得出相同的结论。线性回归插补估计量 μ 定义如下

$$\widehat{\mu}_{lr} := \frac{1}{N}\left(\sum_{k\in S_r} \frac{y_k}{\pi_k} + \sum_{k\in S_m} \frac{\mathbf{x}_k^\top \widehat{\boldsymbol{\beta}}}{\pi_k}\right).$$

在实际应用中，$\boldsymbol{\beta}$的某些系数可能等于零。令 A $\subseteq$P ($\{1, 2, \ldots, p\}$) 表示一个候选模型族。也就是说，A 中的每个元素 $\alpha \in$都是 $\{1, 2, ..., p\}$ 的一个子集，我们将其视为纳入的协变量集合。模型 α 的补集表示为$\alpha^c := \{1, 2, \ldots, p\}\backslash\alpha$。例如，如果 $p = 3$，模型 $\alpha = \{1, 3\}$ 包含协变量X_1 和 X_3。其补集为 $\alpha^c = \{2\}$，对应于仅包含协变量 X_2 的模型。

对于任意 $\alpha \in \mathcal{A}$，令 $\mathbf{x}_{k,\alpha}$ 表示个体 k 的协变量子向量，对应于 α中的索引；并令 $\boldsymbol{X}_{r,\alpha} := (\mathbf{x}_{k,\alpha}^\top)_{k\in S_r}$ 表示相应的设计矩阵，该矩阵限于应答者。

若模型 $\alpha \in \mathcal{A}$ 满足 $\|\boldsymbol{\beta}_{\alpha^c}\|_2 = 0$，即 α 包含所有系数非零的协变量，则该模型是正确的。

正确模型的集合记为 C。真实模型 $\alpha^\star$ 定义为 $\boldsymbol{\beta}$的支撑集，也就是说，

$$\alpha^\star := \left\{ j \in \{1, 2, ..., p\} \; ; \; \beta_j \neq 0 \right\}.$$

真实模型 $\alpha^\star$ 是最小正确模型。在假设的线性模型下，它存在且唯一。若模型 α 不是正确模型，则称其为错误模型；也就是说，它至少遗漏了一个在 $\boldsymbol{\beta}$ 中系数非零的协变量。错误模型集合记为$\mathcal{W}$。

对于每个 $\alpha \in \mathcal{A}$，我们都可以对应一个普通最小二乘估计量

$$\widehat{\boldsymbol{\beta}}_\alpha = \left(\boldsymbol{X}_{r,\alpha}^\top \boldsymbol{X}_{r,\alpha} \right)^{-1} \boldsymbol{X}_{r,\alpha}^\top \boldsymbol{Y}_r, \tag{4}$$

从而得到相应的 μ插补估计量:

$$\widehat{\mu}_\alpha := \frac{1}{N} \left(\sum_{k \in S_r} \frac{y_k}{\pi_k} + \sum_{k \in S_m} \frac{\mathbf{x}_{k,\alpha}^\top \widehat{\boldsymbol{\beta}}_\alpha}{\pi_k} \right), \qquad \alpha \in \mathcal{A}.$$

本文旨在确定应使用

$$\Gamma_{\mathcal{A}} := \{ \widehat{\mu}_\alpha \; ; \; \alpha \in \mathcal{A} \}$$

中的哪一个估计量，并建立理论上可靠的方法论，以尽可能高效地对 μ 进行推断。

3.2. 插补的损失函数

对于任意模型 α，我们可以将 $\widehat{\mu}_\alpha$ 的总误差分解为

$$\widehat{\mu}_\alpha - \mu = \underbrace{(\widehat{\mu}_\alpha - \widehat{\mu}_\pi)}_{\text{imputation/nonresponse error}} + \underbrace{(\widehat{\mu}_\pi - \mu)}_{\text{sampling error}}.$$

第一项反映了通过插补程序由无应答引起的误差（因此取决于所选模型 α），而第二项是通常的抽样误差，与无应答无关。在插补阶段，目标是尽可能降低插补/无应答误差。为此，我们引入以下插补损失函数:

$$\mathcal{L}_n(\alpha) := \mathbb{E}_m \left[(\widehat{\mu}_\alpha - \widehat{\mu}_\pi)^2 \right], \qquad \alpha \in \mathcal{A}.$$

损失 $\mathcal{L}$ 是基于模型 α 的插补估计量 $\widehat{\mu}_\alpha$ 与完整数据 Horvitz-Thompson 估计量之间平方距离的度量。该值为正，并在我们倾向于用于插补的模型 α 下达到最小。更正式地说，我们将最优插补模型 α_{opt} 定义为 $\mathcal{L}_n$ 的任意极小化者，即

$$\alpha_{\text{opt}} \in \underset{\alpha \in \mathcal{A}}{\operatorname{argmin}} \, \mathcal{L}_n(\alpha).$$

原则上，人们希望使用插补估计量 $\widehat{\mu}_{\alpha_{\text{opt}}}$ 来估计 μ。然而，损失$\mathcal{L}_n$ 依赖于未观测到的量，无法根据现有数据加以评估，因此α_{opt}在实践中是未知的。在论文的其余部分，我们研究 $\widehat{\mu}_{\alpha_{\text{opt}}}$ 的性质，并开发实用的模型选择程序，以近似α_{opt} 用于推断。

命题 1. 固定 $\alpha \in \mathcal{A}$。则$\mathcal{L}_n(\alpha)$ 具有如下闭式表达式：

$$\begin{aligned}\mathcal{L}_n(\alpha) &\overset{\arg\min}{\equiv} \left\{\left(\sum_{k\in S_m}\frac{\mathbf{x}_k^\top}{\pi_k} - \sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^\top}{\pi_k}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\right)\boldsymbol{\beta}\right\}^2 \\ &\quad + \sigma^2\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^\top}{\pi_k}\right)\boldsymbol{A}_{r,\alpha}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}}{\pi_k}\right) \\ &:= \mathcal{L}_{1,n}(\alpha) + \mathcal{L}_{2,n}(\alpha).\end{aligned}$$

证明。 见附录B.1。 □

命题 1 给出了将损失分解为偏差项与方差项的有用分解。分量 $\mathrm{L1}_{,n}(\alpha)$ 对应于在模型 α 下进行插补所引起的模型偏差的平方，而 $\mathrm{L2}_{,n}(\alpha)$ 刻画方差贡献。对于任意正确模型 $\alpha \in \mathrm{C}$，我们有 $\mathrm{L1}_{,n}(\alpha) = 0$，因此正确模型仅通过 $\mathrm{L2}_{,n}(\alpha)$ 产生影响。引理 11 表明，$\mathrm{L2}_{,n}$ 是严格递增的集合函数，这意味着向模型中加入协变量只会增大 $\mathrm{L2}_{,n}$，尽管它可能减小 $\mathrm{L1}_{,n}$。因此，在 A 上最小化 $\mathrm{L}_n(\alpha)$ 相当于一种偏差—方差权衡。我们不假定极小化者的唯一性；任何达到 $\min_{\alpha\in\mathcal{A}}\mathrm{L}_n(\alpha)$ 的 α 均被视为最优。

命题 2。 若 $\alpha_{\mathrm{opt}} \in C$，则几乎必然

$$\mathbb{P}_{pq}\left(\alpha_{\mathrm{opt}} = \alpha^\star\right) = 1.$$

证明。 参见附录 B.1。 □

命题2 指出，对于任何有限样本量，若正确模型使损失达到极小（即 $\alpha_{\mathrm{opt}} \in C$），则最优模型必定是真实模型。特别地， $\alpha_{\mathrm{opt}} = \alpha^\star$。

3.3. 最优插补模型的渐近行为

一般而言，研究α_{opt} 在有限样本量下的性质颇具挑战。尽管如此，我们将在第4 节证明，渐近地，$\mathcal{L}_n$ 有唯一的极小化者；在温和条件下，该极小化者也是真实模型。

为此，我们采用Isaki andFuller (1982)的渐近框架。设$(U_v)_{v\in\mathbb{N}}$是一个有限总体递增序列，即 $U_1 \subset U_2 \subset \ldots$，其规模$(N_v)_{v\in\mathbb{N}}$也依次递增。在每个总体 U_v中，使用抽样设计 $\mathcal{P}_v$ 随机抽取样本 S_v，以估计有限总体均值 μ_v。所有后续步骤（即插补、方差估计等）均以类似方式在每个 $v \in \mathbb{N}$ 中进行。注意，协变量数量 p在 v 递增时保持不变。下文为简化记号，在不会引起混淆时，我们将省略索引 v 。

关于超总体模型，作出以下假设。

(S1) 协变量的分布 $\mathbb{P}_{\mathbf{x}}$ 相对于勒贝格测度绝对连续。

(S2) 协变量 $\mathbf{x}$ 具有有界支撑集，即存在 $C_0 > 0$ 使得几乎必然有 $\|\mathbf{x}\|_2 \leq C_0$。(S3) 存在一个常数 M_0 使得 $\mathbb{E}_m[\epsilon_1^4] \leq M_0$ 几乎必然成立。

假设 (S1) 确保矩阵 $\sum_{k\in U} \mathbf{x}_k \mathbf{x}_k^\top$ 几乎必然正定。条件(S2) 假设协变量几乎必然有界。严格来说，这一假设并非必需，但它能大大简化我们的证明。也可以对 $\|\mathbf{x}\|_2$的矩提出另一种假设，但代价是论证篇幅会大幅增加。因此，我们不再进一步探讨这种做法。假设(S3) 是我们在多个地方用到的矩条件。请注意，在大多数结果中，我们只要求有限二阶矩。一般而言，我们对超总体的条件可能并非最小条件。为简洁起见，我们直接假设存在有限四阶矩；后续进行方差估计时将需要这一条件。

注 *1.* 本文中，我们始终在事件 $E_{N_v} := \{\lambda_{min}(N_v^{-1} \sum_{k\in U_v} r_k x_k x_k^\top) \geqslant \gamma\}$ 上进行讨论，其中 $\gamma > 0$ 为某个值。我们在附录中（见引理 *2*）证明，对于某些通用常数C_1、 C_2，当 N_v 足够大时，有 $\mathbb{P}_q(E_{N_v}^c) \leq C_1 exp(-C_2 N_v)$ 几乎必然成立，因此 E_{N_v} 是一个高概率事件。通过考虑截断逆 $(N_v^{-1} \sum_{k\in U_v} r_k x_k x_k^\top)_{trunc}^{-1} = (N_v^{-1} \sum_{k\in U_v} r_k x_k x_k^\top)^{-1} \mathbb{1}_{\mathcal{E}_{N_v}}$，可以去掉这一条件。由于截断估计量与未截断估计量仅以指数级小的概率存在差异，两者之间的差异不会影响任何渐近结果。因此，不失一般性，我们继续以 E_{N_v} 为条件进行讨论。

同样，除了下文所述的抽样设计条件外，我们还假设 $\mathcal{P}$ 诱导出一个类似的高概率事件 $\{\boldsymbol{X}_r^\top \boldsymbol{X}_r \succ 0\}$，并以该事件为条件。这是一个自然的假设，在文献中也很常见，例如参见 *Chauvet和Goga (2022年)*。

我们将以下假设称为关于抽样设计序列 $(\mathcal{P}_v)_{v\in\mathbb{N}}$ 的假设。

(D1) 抽样比满足

$$\lim_{v\to\infty} \frac{n_v}{N_v} := f^\star > 0$$

几乎必然。

(D2) 存在正常数 $\lambda > 0$ 和 $\lambda^\star > 0$，使得对所有 $v \in \mathbb{N}$,

$$\min_{k\in U_v} \pi_{k,v} \geq \lambda, \qquad \text{and} \qquad \min_{k,l\in U_v} \pi_{kl,v} \geq \lambda^\star$$

几乎必然。

(D3) 抽样协方差 $\Delta_{kl} := \pi_{kl} - \pi_k \pi_l$，对于 k，$l \in U_v$，满足存在一个确定性常数 Δ，使得几乎必然，对所有 $v \in \mathbb{N}$,

$$n_v \max_{k,l\in U_v} |\Delta_{kl}| \leq \bar{\Delta}.$$

这些正则条件在文献中很常见。假设 (D1) 指出，样本量 n_v 以与总体规模 N_v 相同的速率增长，且 v 趋于无穷大。假设(D2) 要求一阶和二阶包含概率远离

零。最后，假设 (D3) 确保抽样协方差至少以 $O(n_v^{-1})$的速率趋于零。关于这些假设的更深入讨论，请参阅Breidt 和 Opsomer (2000)。

定理 1 . 设$(\alpha_{opt,v})_{v\in\mathbb{N}}$ 是$(\mathcal{L}_v)_{v\in\mathbb{N}}$的极小化解序列。对于 $\alpha \in \mathcal{A}$，*defin* *e*

$$\boldsymbol{S}_{U_v} := \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} - \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \left(\sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-1} \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v}. \quad (5)$$

假设*(S1)*-*(S2)* 以及*(D1)*-*(D3)*，并且对于 $\alpha \in \mathcal{W}$，

$$\liminf_{v\to\infty} \lambda_{min}\left(\boldsymbol{S}_{U_v}\right) > 0 \qquad a.s. \qquad (C_1)$$

则我们有

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\alpha_{opt,v} = \alpha^\star\right) = 1 \qquad a.s.$$

证明。 参见附录 B.1。 □

定理 1 证明，在 $\mathcal{A}$ 中的所有候选模型里，使（不可观测的）插补损失 α_{opt} 最小的模型 $\mathcal{L}_n$ 最终是真实模型 $\alpha^\star$。换言之，从渐近角度看，用于插补的最优协变量集与真实模型 $\alpha^\star$一致。当然，在实践中，真实模型未知，因此定理 1 主要具有理论意义。其价值在于指导制定实用的选择准则，以便渐近地识别真实模型。我们将在第 4 节重新讨论这一点。

假设（C_1）确保在考虑了纳入的协变量 $\mathbf{x}_{\alpha^c}$ 之后，遗漏协变量 $\mathbf{x}_\alpha$ 仍保留一定的变异。为说明这一点，考虑如下简单情形：$p(\mathbf{x}_k) = p$，$\pi_k = \pi$ 对于 $k \in U$，且截距包含在 α中。经过一些代数运算可得，式（$\boldsymbol{S}_{U_v}$ 5）中定义的矩阵与下列经验协方差矩阵成正比

$$\boldsymbol{S}_{U_v} \propto \frac{1}{N_v}\sum_{k\in U_v} \mathbf{r}_k\mathbf{r}_k^\top$$

其中 $\mathbf{r}_k$: $= \mathbf{x}_{k,\ \alpha^c} - \boldsymbol{B}_v^\top \mathbf{x}_{k,\ \alpha}$ 其中 $\boldsymbol{B}_v$ 求解矩阵最小二乘问题

$$\boldsymbol{B}_v = \underset{\boldsymbol{B}\in\mathbb{R}^{p_\alpha \times p_\alpha^c}}{\operatorname{argmin}} \frac{1}{N_v}\sum_{k\in U_v} \|\mathbf{x}_{k,\alpha^c} - \boldsymbol{B}^\top \mathbf{x}_{k,\alpha}\|_2^2 = \left(\sum_{k\in U_v} \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top\right)^{-1} \sum_{k\in U_v} \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top.$$

因此， $\boldsymbol{S}_{U_v}$ 表示将$\mathbf{x}_{\alpha^c}$ 对$\mathbf{x}_\alpha$回归所得残差的有限总体协方差矩阵。假设（C_1）要求每一个遗漏方向在去除纳入变量的线性效应后，都必须保留严格为正的变异性。直观而言，它排除了遗漏预测变量可由纳入的预测变量完全重构的情形，从而使某些不正确的模型表现得仿佛它们是

正确的。也就是说，缺失协变量实质上几乎不会产生缺失贡献。在我们的设定中，假设(S1) 与(S2) 意味着 $\mathbb{E}_{\mathbf{x}}[\mathbf{x}\mathbf{x}^\top]$ 是正定的，因此任何缺失协变量即使在计入所有其他协变量之后，仍会保留一定的变异性。下一条注记对此加以形式化。

注释 2. 对于等概率抽样设计，应用大数定律可得

$$\boldsymbol{S}_{U_\nu} \xrightarrow[\nu\to\infty]{a.s.} \boldsymbol{S}_* := \pi\left(\mathbb{E}_{\mathbf{x}}[\mathbf{x}_{\alpha^c}\mathbf{x}_{\alpha^c}^\top p(\mathbf{x})] - \mathbb{E}_{\mathbf{x}}[\mathbf{x}_{\alpha^c}\mathbf{x}_{\alpha}^\top p(\mathbf{x})](\mathbb{E}_{\mathbf{x}}[\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^\top p(\mathbf{x})])^{-1}\mathbb{E}_{\mathbf{x}}[\mathbf{x}_{\alpha}\mathbf{x}_{\alpha^c}^\top p(\mathbf{x})]\right).$$

现在，考虑分块矩阵

$$\boldsymbol{M} := \pi\,\mathbb{E}\begin{bmatrix} p\,\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^\top & p\,\mathbf{x}_{\alpha}\mathbf{x}_{\alpha^c}^\top \\ p\,\mathbf{x}_{\alpha^c}\mathbf{x}_{\alpha}^\top & p\,\mathbf{x}_{\alpha^c}\mathbf{x}_{\alpha^c}^\top \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^\top & \boldsymbol{C} \end{bmatrix},$$

并注意到 $\boldsymbol{S}_* = \boldsymbol{C} - \boldsymbol{B}^\top\boldsymbol{A}^{-1}\boldsymbol{B}$ 是 $\boldsymbol{A}$ 在 $\boldsymbol{M}$中的舒尔补。在 *(S1)* 和正性条件下，可以证明 $\boldsymbol{M}$ 是正定的。这进而意味着 $\boldsymbol{S}_*$ 也是正定的，参见例如（*Horn and Johnson, 2012*, 第 *495* 页）。因此，在这种情况下，假设(C_1) 成立。泊松抽样也可进行类似推理。更一般地说，假设 (C_1) 相当宽松，很可能适用于大多数抽样设计。从思想上讲，该假设与文献中的可识别性假设一致，参见例如*Shao (1993年)*。

回顾一下，若估计量序列 $(\widehat{\mu}_\nu)_{\nu\in\mathbb{N}}$ 满足 $\widehat{\mu}_\nu - \mu_\nu \xrightarrow{\mathbb{P}} 0_{\nu\to\infty}$，则称其对 μ_ν 一致。此外，若对所有 $\nu\in\mathbb{N}$， $\mathbb{V}(\widehat{\mu}_\nu)$均存在，且 $\mathbb{V}(\widehat{\mu}_\nu) \longrightarrow \tau^2$，则称 τ^2为 $\widehat{\mu}_\nu$的渐近方差$_{\nu\to\infty}$，并记作 $\mathbb{AV}(\widehat{\mu}_\nu)$：$=$极限$_{\nu\to\infty}\,\mathbb{V}(\widehat{\mu}_\nu)$。接下来的结果将进一步考察使用 $\widehat{\mu}_\nu$的插补估计量 $\widehat{\mu}_\alpha$ 且满足 $\alpha\neq\alpha^\star$时，在一致性和渐近方差方面产生的影响。

命题 3。 设 $\alpha\in\mathcal{A}$ 为任意模型，且 $(\widehat{\mu}_{\alpha,\nu})_{\nu\in\mathbb{N}}$ 为一列插补估计量。假设 *(S1)-(S2)*、*(D1)-(D3)*，并且还假设抽样设计产生相等的一阶包含概率（$\pi_k=\pi$ 对于 $k\in U_\nu$）以及相等的二阶包含概率（$\pi_{kl}=\pi^\star$ 对于 $k\neq l\in U_\nu$）。那么，我们有以下结果。

(i) 序列 $(\widehat{\mu}_{\alpha,\nu})_{\nu\in\mathbb{N}}$ 一致当且仅当

$$\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha^c}^\top\boldsymbol{\beta}_{\alpha^c}\right] = \mathbb{E}_{\mathbf{x}}\left[\mathbf{x}_{\alpha}^\top(1-p(\mathbf{x}))\right]\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^\top\right]^{-1}\mathbb{E}_{\mathbf{x}}\left[\mathbf{x}_{\alpha}\mathbf{x}_{\alpha^c}^\top\boldsymbol{\beta}_{\alpha^c}p(\mathbf{x})\right]. \tag{C_2}$$

(ii) (a) 设 $\alpha\in C$ 为正确模型。假设存在常数 K，使得以下极限存在 $\lim_{\nu\to\infty} N_\nu(\pi^\star-\pi^2) = K$。那么，存在一个常数 $C(\alpha^\star)$，其不依赖于 α，以及一个函数 $M: C\to\mathbb{R}_+^*$，使得

$$\mathbb{AV}\left(\sqrt{N_\nu}(\widehat{\mu}_{\alpha,\nu}-\mu_\nu)\right) = C(\alpha^\star) + \frac{\sigma^2}{\pi}M(\alpha),$$

其中

$$M:\alpha\mapsto \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha}^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha}\right].$$

(b) 函数 M 是一个非递减函数，也就是说，对于 $\alpha_1,\ \alpha_2 \in C$，

$$\alpha_1 \subset \alpha_2 \implies M(\alpha_1) \le M(\alpha_2),$$

等号成立当且仅当

$$\begin{aligned}&\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\right]\\&\quad=\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\right].\end{aligned}\tag{C3}$$

证明。 见附录 B.1。 □

对相等入样概率的限制本质上是技术性的；如果我们在（4）中考虑权重为 $(1/\pi_k)_{k\in S_r}$ 的加权最小二乘，则该限制可以去掉。第一个陈述（i）给出了使插补估计量 $\widehat{\mu}_\alpha$ 相合的条件（C_2）。第二个陈述（ii）推导了 $\alpha \in \mathsf{C}$ 时缩放误差 $\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v)$ 的渐近方差表达式。由此推广，术语“模型 α的渐近方差”将指 $\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v)$ 的渐近方差。量 $\mathbb{AV}(\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v))$ 分解为两项： $C(\alpha^\star)$ 与 $M(\alpha)$。第一项不依赖于 α。由于 $\alpha^\star$ 是最小正确模型，对每一个 $\alpha \in \mathsf{C}$，我们有 $M(\alpha^\star) \le M(\alpha)$。

因此，命题3 旨在强调使用（i）过少或（ii）过多协变量所带来的后果。尽管这些公式看起来可能有些抽象，我们将在下文进一步分析，以得出更易于理解的条件。

推论 1. 考虑命题3所述的设定。假设以下两种情形之一成立：

（*i*）模型正确，即 $\alpha \in C$。（*ii*）令 $\alpha_{\text{mis}} := \alpha^\star \cap \alpha^c$ 表示在 α中缺失的正确协变量集合。假设满足以下条件：

（*a*）正确的缺失协变量不会影响 p，即对于所有 $j \in \alpha_{\text{mis}}$，都有$\left[p(\mathbf{x}) \perp\!\!\!\perp \mathbf{x}_{\alpha^c,j}\right] | \mathbf{x}_\alpha$.（*b*）正确的缺失协变量与纳入的协变量存在线性关系，即对于所有 $j \in \alpha_{\text{mis}}$, $\mathbb{E}[x_{\alpha^c,j}|\mathbf{x}_\alpha] = \mathbf{x}_\alpha^\top \boldsymbol{\gamma}_j$，存在某个 。

于是，条件(C_2) 成立，因此 $\widehat{\mu}_{\alpha,v}-\mu_v \xrightarrow[v\to\infty]{\mathbb{P}} 0$。

证明。 参见附录 B.1。 □

因此，要获得一致估计量，只需满足以下条件之一：(i) 纳入所有与 Y相关的协变量；(ii) 纳入所有同时与 Y 和响应概率相关的协变量，但须满足条件 (b)。第二个条件较难理解，但相对宽松，却至关重要。它指出，与 Y 相关的遗漏协变量必须与已纳入的协变量呈线性关系。乍看之下，这或许像是为简化证明而引入的技术性假设，但事实上，要使 (C_2) 成立，就必须满足这一条件，下一例将对此进行说明。

例 1。*(i)*。回顾一下，如果 $r_k \perp\!\!\!\perp y_k$，则称缺失数据机制为完全随机缺失（*MCAR*）。此时，条件 *(a)* 显然成立，因为 r_k 必须与$\mathbf{x}$独立。如果协变量彼此独立，那么条件 *(b)* 就简化为要求在 α中纳入截距。如下例所示，这一要求是必要的。

考虑如下情形： $p = 2$ 满足$\mathbf{x}^\top = [x_1, x_2] = [1, x]$，真实模型为 $\alpha^\star = \{1,2\}$，候选模型为 $\alpha = \{2\}$，并假设 $p(\mathbf{x}) = \bar{p}, 0 < \bar{p} \leq 1$。那么， $\alpha_{\text{mis}} = \{1\}$。条件 *(a)* 得到满足，但条件 *(b)* 不满足，因为 $\mathbb{E}[1|\mathbf{x}_\alpha] = 1$，并且由于模型中不包含截距，不存在常数 γ 使得 γ 满足 $1 = x\gamma$ 几乎必然成立，除非 x 几乎必然为常数。条件 *(b)* 确实是必要的，因为可以证明

$$\widehat{\mu}_{\alpha,v} - \mu_v \xrightarrow[v\to\infty]{\mathbb{P}} -\beta_0 (1 - \bar{p}) \frac{\mathbb{V}_{\mathbf{x}}(x)}{\mathbb{E}_{\mathbf{x}}\left[x^2\right]},$$

其中， β_0 表示真实截距系数。只要$\beta_0 \neq 0$、 $\bar{p} \neq 1$ 且 $\mathbb{V}_{\mathbf{x}}(x) \neq 0$，这个渐近偏差就不为零；注意，如果 x 的方差为零，那么存在 γ 使得$1 = x\gamma$，并且一致性成立。

(ii). 若模型中包含截距，且$\mathbf{x}$服从椭圆对称分布（例如，多元高斯分布），或者其各分量相互独立，则条件 *(b)* 总是成立，而 *(a)* 足以确保一致性。

在下面的推论中，我们考察命题 3中陈述 (ii) 所得渐近方差的性质。

推论2。 考虑命题*3* 的设定，其中 $\alpha_1, \alpha_2 \in C$ 满足 $\alpha_1 \subset \alpha_2$。考虑以下假设。

(i) 假设：

(a) *For all* $j \in \alpha_2 - \alpha_1$, $\left[p(\mathbf{x}) \perp\!\!\!\perp x_{\alpha_2-\alpha_1,j}\right] | \mathbf{x}_{\alpha_1}$.

(b) *For all* $j \in \alpha_2 - \alpha_1$, $\mathbb{E}\left[x_{\alpha_2-\alpha_1,j}|\mathbf{x}_{\alpha_1}\right] = \mathbf{x}_{\alpha_1}^\top \boldsymbol{\gamma}_j$, *for some* $\boldsymbol{\gamma}_j \in \mathbb{R}^{p_{\alpha_1}}$.

那么，(C3) 成立，且基于 α_1 和 α_2 的模型具有相同的渐近方差。

(ii) 假设存在一个方向$\mathbf{c} \in \mathbb{R}^{\alpha_2-\alpha_1}$，使得

(a) $\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})\mathbf{c}^\top \mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^\top] = \mathbf{0}$.

(b) $\mathbb{E}_{\mathbf{x}}[(1 - p(\mathbf{x}))\mathbf{c}^\top \mathbf{x}_{\alpha_2-\alpha_1}] \neq 0$.

因此，(C3) 不成立，且 α_2 的方差严格大于 α_1。

证明。参见附录B.1。 □

在上述推论中，我们明确给出了一些充分条件，说明添加冗余协变量何时会“付出代价”，何时不会。具体而言，在第 (i) 部分，我们的条件意味着：(a) 与 α_1 相比，添加到 α_2 中的协变量在给定 α_1 中协变量的情况下，无法解释无应答机制。例如，以下情形便符合这一条件：

$p(\mathbf{x}) = p(\mathbf{x}_{\alpha_1})$。条件 (b) 意味着，额外的协变量与 α_1 中已有的协变量呈线性关系。这些条件共同描述了一种情形：纳入 α_1 中的协变量后，$\alpha_2 - \alpha_1$ 中的协变量对于无应答机制既“在均值上没有信息”，又“无关紧要”。例如，考虑真实模型 $\alpha^\star = \{1,2\}$，其中 $X_1 = 1$，并假设 X_2 与 X_3 相互独立。此时条件 (b) 成立；若 $p(\mathbf{x}) = p(X_1, X_2)$，那么使用 X_1、X_2 和 X_3，与仅使用 X_1 和 X_2 得到的渐近方差相同。在第 (ii) 部分，我们转向相反的情形，举例说明添加冗余协变量会增大方差。为了更好地理解这些条件，考虑一个新的随机变量 $z = \mathbf{c}^\top \mathbf{x}_{\alpha_2 - \alpha_1}$，它表示添加的协变量 $\alpha_2 - \alpha_1$ 所张成空间中的一个方向。第一个条件意味着，对于应答者而言，z 是一个新的方向，无法由 α_1 中的协变量捕捉；而第二个条件则要求沿这一方向存在非零的无应答信号。

4. 渐近最优变量选择的方法论

定理 1 表明，在损失函数 $\mathcal{L}$ 下，真实模型是插补的最优模型。然而在实践中，真实模型未知，相应的插补估计量 $\widehat{\mu}_{\alpha^\star}$ 仅作为神谕基准。本节将提出一种实用方法，使其达到与该神谕插补估计量相同的渐近效率，并给出渐近有效置信区间。

4.1. 与神谕的渐近等价

针对识别（即恢复线性回归模型中真实的非零系数）的模型选择已在独立同分布数据中得到广泛研究；教科书式的讨论可参见Shao （1993年、1997年）或Rao 等 （2001）。我们将任何用于选择模型 $\alpha \in \mathcal{A}$ 的（可能依赖于数据的）可测映射 $C_n : \mathcal{A} \to \mathbb{R}$ 称为模型选择准则，也就是说，选择满足以下条件的 $\widehat{\alpha}_{C_n} \in \mathcal{A}$

$$\widehat{\alpha}_{C_n} \in \operatorname*{argmin}_{\alpha \in \mathcal{A}} C_n(\alpha).$$

线性模型常用的模型选择准则包括赤池信息准则（AIC，Akaike （1970））、贝叶斯信息准则（BIC，Schwarz （1978）以及交叉验证等。在独立同分布（i.i.d.）数据上，如果模型选择准则 C_n 满足

$$\lim_{n \to \infty} \mathbb{P}_m\left(\widehat{\alpha}_{C_n} = \alpha^\star\right) = 1$$

几乎必然成立。也就是说，随着样本量 n 趋于无穷，模型选择最终只会选中真实模型。并非所有模型选择准则都具有一致性；例如，在适当条件下，BIC 准则具有一致性，而留一法交叉验证则不具有一致性。下一个引理将形式化地说明，在随机缺失和非信息性假设下，如果 C_n 在独立同分布数据上具有一致性，那么在调查数据上也具有一致性。

引理 1. 设 C^{U_v} 是在 $\{(\mathbf{x}_k, y_k) : k \in U_v\}$ 上拟合的模型选择准则，且 C^{S_r} 是在 $\{(\mathbf{x}_k, y_k) : k \in S_{r,v}\}$ 上拟合的相应算法。将 $\widetilde{\alpha}_{U_v}$ 和 $\widehat{\alpha}_{S_{r,v}}$ 称为它们的

相应的极小化者。假设 *(S1)-(S2),(D1)-(D3)*。
If

$$\lim_{\nu\to\infty} \mathbb{P}_m\left(\widetilde{\alpha}_{U_\nu} = \alpha^\star\right) = 1, \qquad a.s.$$

then

$$\lim_{\nu\to\infty} \mathbb{P}_{mpq}\left(\widehat{\alpha}_{S_{r,\nu}} = \alpha^\star\right) = 1, \qquad a.s.$$

证明。 参见附录B.1。 □

引理 1 形式化地阐述了这一点：只要抽样设计和无应答机制不改变条件分布，模型选择的一致性就会从总体传递到应答者样本中 $\mathbb{P}_m$。尽管这一结果相当自然，但它在后续论述中发挥着重要作用。引理 1的一个自然扩展，是考虑采用适当加权的模型选择准则的信息性抽样设计。例如，Lumley 和 Scott (2015年）；Wieczorek、Guerin 和 McMahon (2022年）；Iparragirre 等 (2023年）都曾提出过此类准则；然而，据我们所知，其一致性尚未得到正式证明。这一问题超出了本文的范围，将留待未来研究。

我们方法论的第一步是使用一致模型选择准则来选择一个模型 $\widehat{\alpha}$。然后，使用基于所选协变量集合 $\widehat{\alpha}$ 的插补估计量 $\widehat{\mu}_{\widehat{\alpha}}$。这将在渐近意义下，从所有可能的模型中得到一个最优估计量。

定理 2. 设 $(\widehat{\alpha}_\nu)_{\nu\in\mathbb{N}}$ 是一致模型选择程序选出的一列模型，且 $(\widehat{\mu}_{\widehat{\alpha},\nu})_{\nu\in\mathbb{N}}$ 是相应的一列插补估计量。假设*(S1)-(S2)*、*(D1)-(D3)*成立。那么，

$$\sqrt{n_\nu}\left(\widehat{\mu}_{\widehat{\alpha},\nu} - \mu_\nu\right) = \sqrt{n_\nu}\left(\widehat{\mu}_{\alpha^\star,\nu} - \mu_\nu\right) + o_{\mathbb{P}}(1).$$

证明。 见附录 B.1。 □

上述结果表明，一旦模型选择程序以概率 1 渐近地识别出真实模型 $\alpha^\star$，可行估计量 $\widehat{\mu}_{\widehat{\alpha}}$ 的渐近分布就与预言估计量相同。换言之，我们可以像事先已知真实模型那样进行推断。这使得在调查插补中使用标准模型选择工具（如贝叶斯信息准则）具有了依据，也表明使用数据驱动模型 $\widehat{\alpha}$ 所引入的偏差或额外变异性渐近消失。证明的关键思路是，在事件 $\widehat{\alpha}_\nu = \alpha^\star$上，我们有 $\widehat{\mu}_{\widehat{\alpha}} = \widehat{\mu}_{\alpha^\star}$。由于该事件的渐近概率为1，其补集上的情况无关紧要。

4.2. 一致方差估计

该方法论的第二步是基于由一致模型选择程序选出的模型 $\widehat{\alpha}$ 进行“经典”方差估计。所谓经典，是指使用与传统所用相同的方差估计量，但使用模型 $\widehat{\alpha}$ 而非所有可用的

协变量。在本文中，我们重点关注反向方法（Fay，1991；Shao 和 Steel，1999），尽管也可以采用类似的 Särndal 方法（Särndal，1992）。有关使用反向方法进行方差估计的更多细节，请读者参阅Kim 和 Rao（2009）和Haziza 和 Vallée （2020）。下一个定理表明，对于任意$\alpha \in C$，存在一个与 $\widehat{\mu}_\alpha$一阶渐近等价的、关于抽样指示变量为线性的估计量。更具体地说，令

$$\widetilde{\mu}_{\alpha,v} := \frac{1}{N_v} \sum_{k \in S_v} \frac{\eta_{k,\alpha}}{\pi_k},$$

表示　　线性化插补估计量，其形式为 $\eta_{k,\alpha} = \mathbf{x}_{k,\alpha}^\top \boldsymbol{\beta}_\alpha + r_k(1 + \pi_k \mathbf{c}_{\alpha,v}^\top \mathbf{x}_{k,\alpha})\epsilon_k$ 和

$$\mathbf{c}_{\alpha,v} = \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \sum_{k \in U_v} \frac{(1 - r_k)\mathbf{x}_{k,\alpha}}{N_v}. \tag{6}$$

下面的结果与 Kim 和 Rao (2009）的定理 1 非常相似；然而，他们的证明不能直接应用于我们的情形，因为我们使用的是非加权最小二乘估计量$\frown{\boldsymbol{\beta}}$，而他们依赖的是加权估计量。为完整起见，我们将他们的定理推广到非加权回归插补。

定理 3. 设 $\alpha \in C$ 是一个正确模型，并考虑序列 $(\widehat{\mu}_{\alpha,v})_{v \in \mathbb{N}}$ 和 $(\widetilde{\mu}_{\alpha,v})_{v \in \mathbb{N}}$。在假设*(S1)-(S3)* 和 *(D1)-(D3)* 下，有

$$\sqrt{n_v}\left(\widehat{\mu}_{\alpha,v} - \widetilde{\mu}_{\alpha,v}\right) = o_{\mathbb{P}}(1). \tag{7}$$

证明。 见附录B.1。 □

为进行方差估计，考虑如下对 $V_T(\alpha)$的分解，即 $\widetilde{\mu}_\alpha - \mu$的方差：

$$\begin{aligned} V_T(\alpha) = \mathbb{V}_{mpq}(\widetilde{\mu}_\alpha - \mu) &= \mathbb{E}_q\left[\mathbb{E}_m\left[\mathbb{V}_p(\widetilde{\mu}_\alpha - \mu)\right]\right] + \mathbb{E}_q\left[\mathbb{V}_m\left(\mathbb{E}_p[\widetilde{\mu}_\alpha - \mu]\right)\right] \\ &\quad + \mathbb{V}_q\left(\mathbb{E}_m\left[\mathbb{E}_p[\widetilde{\mu}_\alpha - \mu]\right]\right) \\ &:= V_1(\alpha) + V_2(\alpha) + V_3(\alpha). \end{aligned}$$

注意到对于 $\alpha \in C$,

$$\mathbb{E}_{mp}[\widetilde{\mu}_\alpha - \mu] = 0,$$

由此可得

$$\mathbb{V}_{mpq}(\widetilde{\mu}_\alpha - \mu) = V_1(\alpha) + V_2(\alpha).$$

^我们首先用 $\boldsymbol{\beta}_\alpha$ 估计 $\boldsymbol{\beta}_\alpha$，并用

$$\widehat{\mathbf{c}}_\alpha = \left(\sum_{k \in S} \frac{r_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N} \right)^{-1} \sum_{k \in S} \frac{(1 - r_k)\mathbf{x}_{k,\alpha}}{N\pi_k}. \tag{8}$$

^估计量 $V_1(\alpha)$ 的 $V_1(\alpha)$ 定义为

$$\widehat{V}_1(\alpha) = \frac{1}{N^2} \sum_{k \in S} \sum_{l \in S} \frac{\Delta_{kl}}{\pi_{kl}} \frac{\widehat{\eta}_k}{\pi_k} \frac{\widehat{\eta}_l}{\pi_l},$$

with

$$\widehat{\eta}_{k,\alpha} = \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha} + r_k\left(1 + \pi_k\widehat{\mathbf{c}}_{\alpha}^{\top}\mathbf{x}_{k,\alpha}\right)\left(y_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}\right), \qquad k \in S. \tag{9}$$

为了估计 $V_2(\alpha)$，我们首先使用以下方法估计 σ^2

$$\widehat{\sigma}_{\alpha}^2 = \sum_{k\in S}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}\right)^2}{n_r - p_{\alpha}}. \tag{10}$$

然后，我们通过 $V_2(\alpha)$ 估计

$$\widehat{V}_2(\alpha) = \widehat{\sigma}_{\alpha}^2\sum_{k\in S}\frac{1 - r_k + r_k\left(\pi_k\widehat{\mathbf{c}}_{\alpha}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{N^2\pi_k}.$$

最后，对于 $\alpha \in$ $\mathbb{V}_{mpq}\ \widehat{\mu}_{\alpha} - \mu$ $V_T\ \alpha\ = C$，总方差的一个估计量为 $V_1(\alpha) + V_2(\alpha)$。对于固定的 $\alpha \in C$，这是通过反向方法得到的常用方差估计量；例如，参见Kim 和 Rao (2009)；Haziza 和 Vallée (2020) 了解更多详情。尽管方差估计量 V_T 被广泛使用，但据我们所知，其一致性尚未得到严格证明。下一个结果将给出正式的论证。为此，我们对抽样设计作出以下附加假设：

(D4) 令 D_{4,N_v} 为 U_v中不同四元组构成的集合。抽样设计满足

$$\lim_{v\to\infty}\max_{(i,j,k,l)\in D_{4,N_v}}|(I_iI_j - \pi_{ij})(I_kI_l - \pi_{kl})| = 0,$$

几乎必然。

这一假设通常用于证明 Horvitz-Thompson 方差估计量的一致性，参见例如：Breidt 和 Opsomer (2000)。

定理 4. 令 $\alpha \in C$，并考虑方差估计量序列 $\{V_{T,v}(\alpha)\}_{v\in\mathbb{N}}$。假设*(S1)-(S3)*和*(D1)-(D4)*成立，且存在 $c_{1,\alpha}$ 和 $c_{2,\alpha}$，使得

$$N_v\mathbb{V}_p\left[(\tilde{\mu}_{\alpha,v} - \mu_v)\right] \xrightarrow[v\to\infty]{\mathbb{P}} c_{1,\alpha}, \qquad N_v\mathbb{V}_m\left(\mathbb{E}_p\left[\tilde{\mu}_{\alpha,v} - \mu_v\right]\right) \xrightarrow[v\to\infty]{\mathbb{P}} c_{2,\alpha}.$$

那么，我们有

$$n_v\left|\widehat{V}_{T,v}(\alpha) - V_{T,v}(\alpha)\right| = o_{\mathbb{P}}(1).$$

证明。 参见附录B.1。 □

我们提出，基于一致模型选择准则所选的模型 $\widehat{\alpha}$ 来估计方差。接下来的结果将证明这种方法的有效性。

定理 5. 设 $\{\widehat{\alpha}_v\}_{v\in\mathbb{N}}$ 是由一致的模型选择程序选出的一系列模型，并且 $\{V_T(\widehat{\alpha}_v)\}_{v\in\mathbb{N}}$ 是相应的一系列方差估计量。假设*(S1)-(S3)*、*(D1)-(D4)*成立，且存在常数 $c > 0$ 使得，几乎必然地，

$$\lim_{v\to\infty} n_vV_{T,v}(\alpha^{\star}) = c.$$

$\widehat{}$那么，估计量 $V_{T,\nu}(\widehat{\alpha}_\nu)$ 是一致的，即

$$\frac{\widehat{V}_{T,\nu}(\widehat{\alpha}_\nu)}{V_{T,\nu}(\alpha^\star)} \xrightarrow[\nu\to\infty]{\mathbb{P}} 1.$$

证明。 参见附录 B.1。 □

定理 5 表明，在选定模型下计算的方差估计量 $\widehat{\alpha}_\nu$ 渐近等价于在真实模型 $\alpha^\star$下将得到的方差。换言之，使用一致模型选择程序并不影响一阶方差估计：我们可以像已知真实模型那样估计方差。

4.3. 渐近有效且最优的置信区间

所提出方法论的最后一步是推导渐近有效置信区间。为此，我们首先建立基于一致模型选择程序的插补估计量的渐近分布。

定理 *6*。考虑一致模型序列 $(\widehat{\alpha}_\nu)_{\nu\in\mathbb{N}}$，并令 $(\widehat{\mu}_{\widehat{\alpha}})_{\nu\in\mathbb{N}}$ 与$\widehat{}$$(V_{T,\nu}(\alpha))_{\nu\in\mathbb{N}}$ 分别为相应的点估计量和方差估计量。令 $B_\nu := \sigma\left((r_k, x_k, \epsilon_k)_{k\in U_\nu}\right)$。假设 *(S1)-(S3)*、*(D1)-(D4)* 以及下列条件成立。

(i) 预言估计量 $\widetilde{\mu}_{\alpha^\star,\nu}$ 满足抽样设计中心极限定理，即

$$\mathbb{V}_p\left(\widetilde{\mu}_{\alpha^\star,\nu}\right)^{-1/2}\left(\frac{1}{N_\nu}\sum_{k\in U_\nu}\left(\frac{I_k}{\pi_k}-1\right)\eta_{k,\alpha^\star}\right)\Big|\,\mathcal{B}_\nu \xrightarrow[\nu\to\infty]{\mathcal{L}} \mathcal{N}(0,1),$$

(ii) 存在常数 $c_3>0$，使得

$$N_\nu\mathbb{V}_p\left(\widetilde{\mu}_{\alpha^\star,\nu}\right) \xrightarrow[\nu\to\infty]{\mathbb{P}} c_3.$$

(iii) 一阶包含概率 $(\pi_k)_{k\in U_\nu}$ 满足：存在正定矩阵 $\boldsymbol{C}_1$ 和 $\boldsymbol{C}_2$，使得

$$\frac{1}{N_\nu}\sum_{k\in U_\nu} r_k\pi_k\mathbf{x}_{k,\alpha^\star}\mathbf{x}_{k,\alpha^\star}^\top \xrightarrow[\nu\to\infty]{\mathbb{P}} \boldsymbol{C}_1, \qquad \frac{1}{N_\nu}\sum_{k\in U_\nu} r_k\pi_k^2\mathbf{x}_{k,\alpha^\star}\mathbf{x}_{k,\alpha^\star}^\top \xrightarrow[\nu\to\infty]{\mathbb{P}} \boldsymbol{C}_2.$$

那么，

$$\frac{\widehat{\mu}_{\widehat{\alpha},\nu}-\mu_\nu}{\sqrt{\widehat{V}_{T,\nu}(\widehat{\alpha}_\nu)}} \xrightarrow[\nu\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

证明。 参见附录B.1。 □

定理中的正则条件 (i) - (iii)6 是标准且相当宽松的。条件 (i) 意味着，取值为 $(\eta_{k,\alpha^\star})_{k\in U_\nu}$ 的 Horvitz-Thompson 估计量渐近正态。对于一些常用的抽样设计，Horvitz-Thompson 估计量的渐近正态性已得到证明： Hájek(1960) 针对不放回简单随机抽样，

Hájek (1964) 针对条件泊松抽样，Bickel 和 Freedman (1984) 针对分层抽样，以及 Krewski 和 Rao (1981) 针对有放回按规模概率比例抽样的整群抽样。例如，对于不放回简单随机抽样，只需证明（参见例如 Thompson (1997年)，第 59 页）几乎必然有，$\mathbb{E}_{mpq}[\eta_{k,\alpha}^{2+\delta}] < \infty$ 对于某个 $\delta > 0$。在我们的设定中，可以验证该条件成立。还可以证明，另外两个条件在简单随机抽样及其他常见抽样设计中也成立。

所提出的方法论总结如下，见算法 1。核心思想是，当使用一致模型选择准则选定插补模型时，模型不确定性的影响渐近可忽略。更确切地说，选定模型以趋近于 1 的概率与真实模型一致，因此，在给定选定模型的条件下应用标准推断程序，渐近而言仍然有效。因此，点估计和方差估计可以像预先已知选定模型那样进行。这大大简化了模型选择后的推断：无需进行选择后校正或采用更复杂的重抽样方案，即可获得渐近有效的置信区间，同时保持完全的效率。

结合此前的结果，可得出算法 1所得置信区间具有以下性质。

算法 1： 线性回归插补的变量选择与推断 n

输入： 观测数据 $\{(\mathbf{x}_k,\ y_k,\ \pi_k):\ k \in S_r\} \cup \{(\mathbf{x}_k,\ \pi_k):\ k \in S_m\}$，候选集合 A，以及一致模型选择准则 C_n。

步骤 1. 模型选择。
选择协变量集

$$\widehat{\alpha} = \arg\min_{\alpha \in \mathcal{A}} C_n(\alpha),$$

其中 C_n 是一个一致的选择准则。

步骤 2. 点估计。
使用步骤 1 中选定的模型 $\widehat{\alpha}$ 计算插补估计量 $\widehat{\mu}_{\widehat{\alpha}}$。

步骤 3. 方差估计。
使用与模型 $\widehat{\alpha}$ 相对应的方差估计量估计总方差 $V_T(\widehat{\alpha})$.

步骤 4。置信区间。
构造 $(1-2\alpha)$ 区间

$$\mathrm{CI}_{1-\alpha}(\widehat{\mu}_{\widehat{\alpha}}) = \left[\widehat{\mu}_{\widehat{\alpha}} - z_{1-\alpha/2}\sqrt{\widehat{V}_T(\widehat{\alpha})},\ \widehat{\mu}_{\widehat{\alpha}} + z_{1-\alpha/2}\sqrt{\widehat{V}_T(\widehat{\alpha})}\right],$$

其中 $z_{1-\alpha/2}$ 表示 $\mathrm{N}(0,1)$ 的 $1-\alpha/2$ 分位数。

输出： 有限总体均值 μ 的 $100(1-\alpha)\%$ 置信区间 $\mathrm{CI}_{1-\alpha}(\widehat{\mu}_{\widehat{\alpha}})$.

推论 3. 考虑置信区间序列 $(\text{置信区间}_{1-\alpha}(\widehat{\mu}_{\widehat{\alpha}}))_{\nu\in\mathbb{N}}$通过算法 *1*得到。假设满足定理 *6*的条件，则：

$$\lim_{\nu\to\infty} \mathbb{P}_{mpq}\left(\mu_\nu \in \mathrm{CI}_{1-\alpha}(\widehat{\mu}_{\widehat{\alpha}_\nu})\right) = 1-\alpha$$

几乎必然成立。此外，这些置信区间具有渐近最优性，达到了任何竞争模型下可实现的最小宽度。

根据定理可立即得到渐近覆盖率6。最优性源于 $\widehat{\mu}_{\widehat{\alpha}}$ 与 $\widehat{\mu}_{\alpha^\star}$具有相同的渐近方差；根据命题3 ，这是最低的，因为 M 是非递减的。

5. 模拟研究

本节介绍评估所提方法论效率的模拟研究结果。我们首先考察损失函数$\mathcal{L}_n(\alpha)$ 在不同模型设定下的表现，以及基于各种模型选择准则的点估计量的效率，随后评估其方差估计量的表现，包括其能否使通过算法1获得的置信区间达到预期的渐近覆盖率。

在模拟实验中，我们采用有限总体框架。具体而言，对于每种情形，我们都根据假定的超总体模型生成多个规模较大的有限总体 N 。从每个生成的总体中，按照指定的抽样设计选取一个大小为 n 的样本。随后，根据预先设定的响应机制，在每个选出的样本中生成项目无应答，然后应用所提出的插补程序。重复这一完整过程——总体生成、抽样、无应答生成和插补——20,000 次，以评估所提方法论在有限样本下的表现。

5.1. 模拟设置

我们生成了规模为 $N = (1000,\ 2000,\ 5000)$ 的有限总体，并从形状参数为 5、尺度参数为 2的伽马分布中生成了 $p = 20$ 个相互独立的协变量。调查变量的生成方式如下：

$$y_k = \mathbf{x}_k^\top \boldsymbol{\beta} + \epsilon_k, \qquad k \in U,$$

其中 $\boldsymbol{\beta} = [10, 9, 9, 8, 8, 7, \mathbf{0}_{14}^\top]^\top$，且当 $k \in U$ 时，$\epsilon_k \sim \mathcal{N}(0,\ 3600)$。因此，前六个协变量对应信号变量，其余十四个协变量则是对结果没有影响的噪声变量。

在以下设计下，我们考虑了三种样本量 $n = (100,\ 200,\ 500)$：

（i）不放回简单随机抽样。（ii）分层抽样：总体首先按 $-(3x_1 + 2x_2 + 4x_3 + 5x_4)$排序，然后被依次划分为 $H = 4$ 个层，这些层根据有序列表顺序定义，分别包含总体单位的50%、25%、20%和5% 。在每一层中，样本S_h 按样本量 n_h 、基于x_2最优分配，采用不放回简单随机抽样抽取。

这些样本量的选取与总体规模 $N = (1000, 2000, 5000)$相对应，从而得到恒定的抽样比10%。换言之，在各情景中，总体规模与样本量以相同速率增长，同时保持固定的抽样比。

我们选择了一个由以下 20 个模型组成的嵌入模型集$\mathcal{A}$

$$\underbrace{\{1\}}_{:=\alpha_1} \subseteq \underbrace{\{1,2\}}_{:=\alpha_2} \subseteq \cdots \subseteq \underbrace{\{1,2,\ldots,20\}}_{:=\alpha_{20}}.$$

尽管我们没有明确写出，但请注意，截距包含在每个模型中。

响应指示变量通过以下方式生成：

$$\operatorname{logit}(p(\mathbf{x}_k)) = 0.1 \times \left(-70 + \mathbf{x}_k^\top \boldsymbol{\zeta}\right), \qquad k \in U,$$

其中 $\boldsymbol{\zeta} = \left[1,1,1,1,0,0,1,1,1,\mathbf{0}_{11}^\top\right]^\top$。这使响应率约为50%。

5.2. 不同模型下$\mathcal{L}_n(\alpha)$ 的行为

我们首先研究了$\mathcal{L}_n(\alpha)$ 针对 $\alpha \in \mathcal{A}$的行为。此处，我们仅给出情形 $N = 5000$ 和 $n = 500$的结果。对于 N 和 n 的其他组合所得结果非常相似，因此为简洁起见此处不再报告。鉴于损失函数$\mathcal{L}_n(\alpha)$在实践中未知，我们通过 $B = 20,000$次迭代的蒙特卡罗近似对其进行估计。

作为点估计量偏差的度量，我们采用了如下定义的蒙特卡罗相对偏差（RB）

$$\mathrm{RB}(\widehat{\mu}) = 100\% \times \frac{1}{B}\sum_{b=1}^{B} \frac{\widehat{\mu}^{(b)} - \mu^{(b)}}{\mu^{(b)}},$$

其中 $\widehat{\mu}^{(b)}$ 表示任意估计量， $\mu^{(b)}$ 表示第 b次迭代的有限总体均值。作为效率的度量，我们计算了如下定义的蒙特卡罗相对效率（RE）

$$\mathrm{RE}(\widehat{\mu}) = 100\% \times \frac{\sum_{b=1}^{B}(\widehat{\mu}^{(b)} - \mu^{(b)})^2}{\sum_{b=1}^{B}(\widehat{\mu}_\pi^{(b)} - \mu^{(b)})^2},$$

其中 $\widehat{\mu}_\pi^{(b)}$表示第 b次重复中的HT估计量。模拟结果见表1。我们首先观察到，基于我们所提出的损失$\mathcal{L}_n$ 与基于相对效率（RE）的排序完全一致：在两种抽样设计下，这 20 个模型的排序完全相同。这证实损失$\mathcal{L}_n$ 的表现符合预期，并且使$\mathcal{L}_n$最小化的模型（我们称之为最优插补模型）确实是实践中的最佳选择。

接下来回顾一下，标记为 α_1–α_5 的模型存在模型误设，因为它们未纳入至少

表 1损失 $\mathcal{L}_n(\widehat{\mu}_\alpha)$，样本量为 $n = 500$ 时不同模型的相对偏差和相对效率。

模型	$\mathrm{L}_n(\alpha)$	RB	RE	$\mathcal{L}(\alpha)$	RB	RE
	无放回简单随机抽样			分层抽样		
α_1	491.3	4.3	2267.1	587.5	4.7	3940.9
α_2	235.5	2.9	1139.4	294.8	3.2	2023.6
α_3	63.6	1.3	381.3	89.2	1.7	677.3
α_4	15.5	0.1	166.6	16.3	0.0	202.5
α_5	12.1	0.0	152.2	12.7	0.0	180.1
α_6	9.7	0.0	142.2	10.0	0.0	162.6
α_7	10.3	0.0	144.5	10.5	0.0	165.9
α_8	11.0	0.0	147.7	11.3	0.0	170.7
α_9	11.7	0.0	150.7	11.9	0.0	174.9
α_{10}	11.8	0.0	151.0	12.0	0.0	175.3
α_{11}	11.9	0.0	151.1	12.0	0.0	175.5
α_{12}	11.9	0.0	151.3	12.1	0.0	175.9
α_{13}	11.9	0.0	151.5	12.1	0.0	176.2
α_{14}	12.0	0.0	151.7	12.2	0.0	176.6
α_{15}	12.0	0.0	151.9	12.2	0.0	177.0
α_{16}	12.1	0.0	152.3	12.3	0.0	177.4
α_{17}	12.1	0.0	152.5	12.4	0.0	177.7
α_{18}	12.2	0.0	152.6	12.4	0.0	178.1
α_{19}	12.3	0.0	153.0	12.4	0.0	178.5
α_{20}	12.3	0.0	153.2	12.5	0.0	178.8

一个重要预测变量： X_2、……、 X_6。变量 X_1 至 X_4 也与缺失机制相关。因此，由于这些变量与响应机制相关，将它们从插补模型中省略可能会导致渐近后仍不消失的偏差，从而产生不一致估计量，正如推论 1 所示。相反，一旦将 X_1、 X_2、 X_3 和 X_4 纳入插补模型，预期就不会出现渐近偏差。为说明这一点，以模型 α_4 为例，其中 $\alpha_{\text{mis}} = \{5, 6\}$。由于所有协变量彼此独立，推论 1 中的条件 (b) 得以满足：在给定$\mathbf{x}_{\alpha_4}$ 的条件下， X_5 和 X_6 无法解释$p(\mathbf{x})$。 此外，

$$\mathbb{E}[X_5 \mid X_1, X_2, X_3, X_4] = \mathbb{E}[X_5] \quad \text{and} \quad \mathbb{E}[X_6 \mid X_1, X_2, X_3, X_4] = \mathbb{E}[X_6]$$

是常数。由于模型中包含截距，推论 1的两个条件对于 α_4均成立。模拟清楚地反映了这一行为：模型 α_1 至 α_3 存在偏差，且效率较低；在不放回简单随机抽样中，RE 值介于 381.3 和 2267.1 之间。一旦将 X_4 加入 α_3，偏差便消失了。然而，模型 α_4 并非效率最高的模型。事实上，真实模型 α_6效率最高，从而说明了定理1。

模型 α_6 是真实模型。正如预期，它表现出可忽略的偏差、最小的 $\mathcal{L}_n(\alpha)$，以及最高的效率。这与定理1的结论一致：真实模型也应当是插补的最优模型，这一点在此处得到了验证。请注意，模型$\alpha_7 - \alpha_{20}$ 属于正确模型集合，因此它们的偏差也可忽略。此外，根据命题 2 的解释，对于 $j = 7, \ldots, 20$， $\mathcal{L}_n(\alpha_j) \geq \mathcal{L}_n(\alpha_6)$。虽然 $X_7 - X_9$解释了$p(\mathbf{x})$，但将这些变量纳入插补模型会导致明显增加

$\mathcal{L}_n(\alpha)$，并提高所得估计量的相对效率。进一步而言，命题 3表明，基于 $\alpha_7 - \alpha_9$ 的插补估计量的渐近方差大于基于 α_6 的插补估计量，因此这些模型的效率较低。相较之下， $X_{10} - X_{20}$ 在给定 $\mathbf{x}_{\alpha_9}$的条件下，不能解释 $p(\mathbf{x})$。此外，对于 $j = 10, \ldots, 20$, $x_{\alpha_9}] = \mathbb{E}_{\mathbf{x}}[X_j]$为常数。由于模型中包含截距，推论2 的两个条件对于 α_9均成立。因此，基于 $\alpha_{10} - \alpha_{20}$ 的插补估计量效率与 α_9 下得到的效率没有显著差异。

5.3. 使用模型选择准则进行点估计

我们现在研究基于常用模型选择准则（如 AIC、BIC 和 $k = 5$ 下的 K 折交叉验证）所选模型的插补估计量的表现。

为了进一步考察各项准则的模型选择能力，我们计算了如下定义的蒙特卡洛识别概率：

$$\mathbb{P}_{MC}\left(\widehat{\alpha} = \{\alpha^\star\}\right) = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left(\widehat{\alpha}^{(b)} = \alpha^\star\right),$$

其中 $\widehat{\alpha}$ 表示一个模型选择准则。结果见表 2。

表 2不同样本量下各模型选择准则的插补表现。注：我们使用记号 $C^- = C - \alpha\star$ 表示过拟合模型的集合，即包含所有真实协变量以及其他多余协变量的模型。

样本量	准则	RB	RE	$\mathcal{W}$	$\alpha^\star$	C^-	RB	RE	$\mathcal{W}$	$\alpha^\star$	C^-
		无放回简单随机抽样					分层抽样				
$n = 100$	AIC	0.0	164.1	1.6	49.9	48.5	0.0	193.8	1.5	52.2	46.3
	BIC	0.0	151.5	7.5	82.0	10.5	0.0	176.3	7.0	83.0	10.0
	交叉验证	0.0	157.7	3.9	37.5	58.6	0.0	185.7	3.5	37.5	59.0
	真实模型	0.0	147.9	-	-	-	0.0	170.2	-	-	-
$n = 200$	AIC	0.0	150.5	0.0	63.9	36.1	0.0	170.5	0.0	63.7	36.3
	BIC	0.0	145.0	0.5	94.5	5.0	0.0	163.6	0.3	94.9	4.8
	交叉验证	0.0	150.8	0.2	39.2	60.6	0.0	170.6	0.2	39.1	60.7
	真实模型	0.0	143.6	-	-	-	0.0	162.3	-	-	-
$n = 500$	AIC	0.0	146.8	0.0	69.2	30.7	0.0	168.9	0.0	68.5	31.5
	BIC	0.0	142.7	0.0	97.6	2.4	0.0	163.3	0.0	97.6	2.4
	交叉验证	0.0	146.9	0.0	39.3	60.7	0.0	170.1	0.0	38.5	60.5
	真实模型	0.0	142.2	-	-	-	0.0	162.6	-	-	-

在所有样本量下，基于赤池信息准则、贝叶斯信息准则和交叉验证准则的插补估计量在两种抽样设计下均表现出可忽略的偏差。这是因为已知这三种模型选择准则在总体层面满足 $\mathbb{P}(\widehat{\alpha}_v \in \mathcal{W}) \to 0$ as $v \to \infty$ （Zhang，1993年；Shao，1997年），因此根据引理 1，在样本层面也满足这一性质。我们的模拟证实了这一现象：对于这三种准则，选错模型的概率均收敛于0 。因此，基于这些模型的插补

估计量是一致的。然而，赤池信息准则和交叉验证并非一致模型选择程序，其过拟合概率分别为30.7%和60.7%；在我们的模拟中， $n = 500$ 。由于基于这些模型的插补估计量是一致的，效率较低意味着方差较大。对于 $n = 500$，基于赤池信息准则和交叉验证的插补估计量效率较低，其相对效率值分别为146.8% 和146.9%，而真实模型对应的值为142.2% 。正如命题3第 ii) 部分所述，真实模型的渐近方差最小，而基于过拟合模型的插补估计量的方差必然更大。相比之下，贝叶斯信息准则保持一致，选中真实模型的概率在97.6% 时达到$n = 500$。因此，基于贝叶斯信息准则的点估计量在所有情况下效率最高。此外，正如定理3所证明的，基于贝叶斯信息准则的插补估计量渐近等价于基于真实模型的预言式插补估计量。

5.4. 方差估计与置信区间

现在我们转而讨论方差估计和置信区间问题，具体按照算法 1 中所述的程序进行。更具体地说，我们关注的是 $V_T(\widehat{\alpha})$ 作为 $V_T(\widehat{\alpha})$ 的估计量时的相对偏差$\widehat{\ }$，其中 $\widehat{\alpha}$ 表示通过一致模型选择程序选出的模型；此处采用了 BIC 准则。我们计算了

$$\mathrm{RB}(\widehat{V}_T(\widehat{\alpha})) = 100\% \times \frac{1}{B}\sum_{b=1}^{B} \frac{\widehat{V}_T^{(b)}(\widehat{\alpha}) - \mathbb{V}_{MC}(\widehat{\mu}_{\widehat{\alpha}})}{\mathbb{V}_{MC}(\widehat{\mu}_{\widehat{\alpha}})},$$

with

$$\mathbb{V}_{MC}(\widehat{\mu}_{\widehat{\alpha}}) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\mu}_{\widehat{\alpha}}^{(b)} - \frac{1}{m}\sum_{m=1}^{m}\widehat{\mu}_{\widehat{\alpha}}^{(b)}\right)^2$$

表示 $\widehat{\mu}_{\widehat{\alpha}}$ 的蒙特卡洛方差。我们还希望通过实证验证 $\mathbb{P}(\mu_v \in$置信区间$_{1-\alpha}(\widehat{\mu}_{\widehat{\alpha}})) \approx 1-\alpha$，正如推论 3 所示，这是渐近成立的。我们计算了

$$\mathrm{CP}(\widehat{\mu}_{\widehat{\alpha}}) = 100\% \times \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left(\mu^{(b)} \in \mathrm{CI}_{1-\alpha}^{(b)}(\widehat{\mu}_{\widehat{\alpha}}^{(b)})\right),$$

其中 $\mathrm{CI}_{1-\alpha}^{(b)}(\widehat{\mu}_{\widehat{\alpha}}^{(b)})$表示第 1 个算法在第 b次迭代的输出。结果见表 3。

表 3随着 n 增大，不同抽样比下方差估计量的相对偏差和覆盖概率的变化。

	无放回简单随机抽样						分层抽样					
	5%		10%		20%		5%		10%		20%	
	RB	CP	RB	CP	RB	CP	RB	CP	RB	CP	RB	CP
$n = 200$	-2.9	94.5	-4.9	94.2	-8.0	93.8	-4.9	94.1	-5.5	94.1	-10.6	93.5
$n = 500$	-3.1	94.6	-3.0	94.6	-5.8	94.2	-4.0	94.5	-4.0	94.3	-8.2	93.8
$n = 1000$	-3.2	94.5	-1.9	94.8	-4.6	94.3	-1.8	94.7	-3.2	94.7	-4.5	94.3

对于较小的样本量，随着抽样比增加，相对偏差略有增大。因此，覆盖概率略低于标称的 95% 水平，尤其是在 $n = 200$ 时。然而，在抽样比固定的情况下，随着样本量和总体规模的增加，负偏差逐渐减小，最终变得可以忽略不计。这一规律在两种抽样设计中均一致。

这些结果证实了所提出的方差估计程序的渐近有效性：随着样本量增加，方差估计量的偏差逐渐消失，经验覆盖概率收敛于其名义水平。因此，算法 1在不放回简单随机抽样和分层抽样下，都能在大样本中产生可靠的置信区间。

6. 最后 remarks

本文针对调查抽样中的插补问题，构建了一个模型选择理论框架。我们引入了一个预言损失函数，用于量化插补模型的效率，并证明其极小化者在渐近意义下与真实模型一致。在标准正则条件和非信息性抽样设计下，在独立同分布情形中一致的模型选择程序，应用于调查数据时仍然保持一致性。特别是，BIC 准则能够在渐近意义下识别真实的插补模型。

基于这一框架，我们建立了使用一致准则选择插补模型时，插补估计量及其方差的渐近性质。所得估计量与真实模型下得到的估计量渐近等价，因此实现了预言效率。这些结果为在涉及调查数据的插补问题中使用标准模型选择工具提供了理论依据。

模拟结果支持理论发现。损失函数 $\mathcal{L}_n$ 能够有效区分模型，一旦纳入所有相关预测变量，偏差便会消失。BIC 准则能够稳定识别真实模型，而赤池信息准则和交叉验证往往倾向于选择过于复杂的设定。所提出的方差估计量在有限样本中表现良好，偏差可忽略，经验覆盖率接近名义水平。

所提出的框架在调查数据中的模型选择与插补之间建立了严格联系。该框架表明，在适当条件下，使用一致准则选择插补模型能够产生渐近有效的推断。将这些结果推广到更广泛的插补模型类别，或推广到具有信息性抽样设计的情形，将会很有意义。朝这一方向迈出的第一步，可以利用非参数回归中交叉验证良好的模型选择性质（Yang，2007）。另一个相关且颇具前景的方向，是研究使用聚合（Nemirovski，2000；Bunea、Tsybakov 和 Wegkamp，2007）来结合多个模型。Larbi 等人 对单位无应答的处理进行了实证检验，并取得了良好结果（2025），但目前仍缺乏对该主题的理论研究。

附录 A：补充记号

欧几里得向量范数记为 $\|\cdot\|_2$。矩阵的算子范数和弗罗贝尼乌斯范数分别记为 $\|\cdot\|_{op}$ 和 $\|\cdot\|_F$。对称矩阵 $\boldsymbol{S}$ 的最大和最小特征值分别记为 $\lambda_{min}(\boldsymbol{S})$ 和 $\lambda_{max}(\boldsymbol{S})$。我们用 $\mathrm{Tr}(\boldsymbol{S})$ 表示方阵 $\boldsymbol{S}$ 的迹。我们使用 $\boldsymbol{S} \succeq 0$（分别为 $\boldsymbol{S} \succ 0$）表示对称矩阵 $\boldsymbol{S}$ 是半正定的（分别为正定的），并写作 $\boldsymbol{S}_1 \succeq \boldsymbol{S}_2$ 表示 $\boldsymbol{S}_1 - \boldsymbol{S}_2 \succeq 0$。$\mathbb{R}^{n\times n}$ 的单位矩阵记为 $\boldsymbol{I}_n$。

这里，$z_n = o_{\mathbb{P}}(1)$ 表示对任意 $\epsilon > 0$，我们都有

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(|z_n| > \epsilon\right) = 0$$

几乎必然。在证明的某些中间步骤中，我们可能省略“几乎必然”或“a.s.”这一表述。为表示 $A \cup B$ 是不交并，我们使用 $A \uplus B$。我们用 $x_n \sim_p z_n$ 表示 $x_n - z_n = o_{\mathbb{P}}(1)$。若要表示两个随机序列 $(x_n)_{n\in\mathbb{N}}$ 和 $(z_n)_{n\in\mathbb{N}}$ 在概率意义下同阶，我们写作 $x_n \asymp_{\mathbb{P}} z_n$。

矩阵 $\boldsymbol{X}_r^\top \boldsymbol{X}_r$ 记作 $\boldsymbol{A}_r := \boldsymbol{X}_r^\top \boldsymbol{X}_r$。类似地，对于 $\alpha \in \mathcal{A}$，我们写作 $\boldsymbol{A}_{r,\,\alpha} := \boldsymbol{X}_{r,\,\alpha}^\top \boldsymbol{X}_{r,\,\alpha}$。

附录 B：主要结果的证明

B.1. 高概率下的可逆性

引理 2. 假设 *(S1)-(S2)*，*(D1)-(D2)*。令

$$\boldsymbol{A}_v^w := \frac{1}{N_v} \sum_{k\in U_v} w_k r_k \mathbf{x}_k \mathbf{x}_k^\top,$$

对于一组 $\sigma(\boldsymbol{X})$-可测权重 $(w_k)_{k\in U_v}$，满足存在 $\nu > 0$ 使得，对所有 $v \in \mathbb{N}$、$w_k \geqslant \nu$，对所有 $k \in U_v$。那么，存在常数 C_1，C_2，$C_3 \in \mathbb{R}_+^*$ 以及一个指标 $v_0 := v_0\left((\mathbf{x}_k)_{k\in U}\right) \in \mathbb{N}$，使得对所有 $v \geqslant v_0$，几乎必然，

$$\mathbb{P}_q\left(\lambda_{\min}(\boldsymbol{A}_v) < C_3\right) \leq C_1 \exp\left(-C_2 N_v\right). \tag{11}$$

证明。由于 $w_k \geqslant \nu$ 一致地成立，我们有

$$\boldsymbol{A}_v^w \succeq \nu \boldsymbol{A}_v, \quad \text{with} \quad \boldsymbol{A}_v := \frac{1}{N} \sum_{k\in U_v} r_k \mathbf{x}_k \mathbf{x}_k^\top.$$

特别是，这保证了

$$\lambda_{\min}(\boldsymbol{A}_v^w) \geq \nu \lambda_{\min}(\boldsymbol{A}_v),$$

因此，我们将改为证明不等式 $\boldsymbol{A}_v$。注意，

$$\mathbb{E}_q\left[\boldsymbol{A}_v\right] = \frac{1}{N} \sum_{k\in U_v} p(\mathbf{x}_k) \mathbf{x}_k \mathbf{x}_k^\top := \boldsymbol{A}_v^p \xrightarrow[v\to\infty]{a.s.} \mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}) \mathbf{x}\mathbf{x}^\top\right] := \boldsymbol{A}^{(\infty)}.$$

在我们的设定中， $\boldsymbol{A}^{(\infty)} \succ \mathbf{0}$，因此 $\sigma^{(\infty)} := \lambda_{\min}(\boldsymbol{A}^{(\infty)}) > 0$。由于收敛到 $\boldsymbol{A}^{(\infty)}$ 几乎必然成立，因此存在一个事件 E，使得 $\mathbb{P}_{\mathbf{x}}(E) = 1$；在该事件上，对于所有 $\omega \in E$，都存在$v_0(\omega)$，使得对于所有 $v \geq v_0$，

$$\lambda_{\min}(\boldsymbol{A}_v) \geq \frac{\sigma^{(\infty)}}{2}.$$

现在固定 $\omega \in E$和 $v \geq v_0$。令

$$\mathbf{Z}_k := N_v^{-1} r_k \mathbf{x}_k \mathbf{x}_k^{\top}, \qquad k \in U_v,$$

并注意到（$\mathbf{Z}_k$）$_{k\in U_v}$ 是（i）半正定的；（ii）算子范数一致有界$\|\mathbf{Z}_k\|_{op} \leq C_0^2/N$，其中 C_0满足 $\|\mathbf{x}_k\|_2 \leq C_0$。特别地，

$$\mu_{\min} := \lambda_{\min}\left(\mathbb{E}_q\left[\sum_{k\in U_v} \mathbf{Z}_k\right]\right) = \lambda_{\min}\left(\boldsymbol{A}_v^p\right) \geq \frac{\sigma^{(\infty)}}{2}.$$

应用矩阵 Chernoff 不等式（例如，参见Tropp 等人（2015）中的定理 5.1.1）可得

$$\mathbb{P}_q\left(\lambda_{\min}\left(\sum_{k\in U_v} \mathbf{Z}_k\right) \leq \frac{\mu_{\min}}{2}\right) \leq p\left(\sqrt{\frac{2}{e}}\right)^{\frac{\mu_{\min} N_v}{C_0^2}} \leq p\left(\sqrt{\frac{2}{e}}\right)^{\frac{\sigma^{(\infty)} N_v}{2C_0^2}}.$$

注意到$2/e < 1$，我们写作

$$p\left(\sqrt{\frac{2}{e}}\right)^{\frac{\sigma^{(\infty)} N_v}{2C_0^2}} = C_1 \exp\left(-C_2 N_v\right), \quad \text{with} \qquad C_1 := p, \quad C_2 := -\frac{\sigma^{(\infty)}}{2C_0^2} \ln\left(\sqrt{\frac{2}{e}}\right) > 0.$$

总之，由于

$$\left\{\lambda_{\min}\left(\sum_{k\in U_v} \mathbf{Z}_k\right) \leq \frac{\sigma^{(\infty)}}{4}\right\} \subset \left\{\lambda_{\min}\left(\sum_{k\in U_v} \mathbf{Z}_k\right) \leq \frac{\mu_{\min}}{2}\right\},$$

我们得到，对于所有 $v \geqslant v_0$，几乎必然成立，

$$\mathbb{P}_q\left(\lambda_{\min}\left(\sum_{k\in U_v} \mathbf{Z}_k\right) \leq \frac{\sigma^{(\infty)}}{4}\right) \leq C_1 \exp\left(-C_2 N_v\right)$$

这也可以转化为其加权版本。 □

命题 *1*的证明。

在证明的其余部分，我们固定一个任意的 $\alpha \in \mathcal{A}$。注意到

$$\begin{aligned}\mathcal{L}_n(\alpha) &= \mathbb{E}_m\left[\sum_{k\in S_m} \frac{\left(y_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}\right)^2}{\pi_k^2}\right] + \mathbb{E}_m\left[\sum_{k\in S_m}\sum_{\substack{l\in S_m \\ l\neq k}} \frac{y_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}}{\pi_k} \frac{y_l - \mathbf{x}_{l,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}}{\pi_l}\right] \\ &:= A_n(\alpha) + B_n(\alpha).\end{aligned}$$

我们分别计算每一项。

计算 $A_n(\alpha)$。 展开 $y_k = \mathbf{x}_{k,\ \alpha}^{\top}\boldsymbol{\beta} + \epsilon_k$ 对于 $k \in S_m$, 我们得到

$$
\begin{aligned}
A_n(\alpha) &= \mathbb{E}_m\left[\sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)^2}{\pi_k^2}\right] + \mathbb{E}_m\left[\sum_{k\in S_m}\frac{\epsilon_k^2}{\pi_k^2}\right] + 2\mathbb{E}_m\left[\sum_{k\in S_m}\frac{\epsilon_k}{\pi_k^2}\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)\right] \\
&= \mathbb{E}_m\left[\sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)^2}{\pi_k^2}\right] + \sum_{k\in S_m}\frac{\sigma^2}{\pi_k^2}.
\end{aligned}
$$

第一项可改写为

$$
\begin{aligned}
&\mathbb{E}_m\left[\sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)^2}{\pi_k^2}\right] \\
&= \mathbb{E}_m\left[\sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\boldsymbol{Y}_r\right)^2}{\pi_k^2}\right] \\
&= \mathbb{E}_m\left[\sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\boldsymbol{X}_r\boldsymbol{\beta}\right)^2}{\pi_k^2}\right] \\
&\quad + \sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\mathbb{E}_m\left[\boldsymbol{\epsilon}_r\boldsymbol{\epsilon}_r^{\top}\right]\boldsymbol{X}_{r,\alpha}\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{k,\alpha}}{\pi_k^2} \\
&= \sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\boldsymbol{X}_r\boldsymbol{\beta}\right)^2}{\pi_k^2} + \sigma^2\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{k,\alpha}}{\pi_k^2}.
\end{aligned}
$$

总的来说,

$$
A_n(\alpha) = \sum_{k\in S_m}\frac{\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\boldsymbol{X}_r\boldsymbol{\beta}\right)^2}{\pi_k^2} + \sigma^2\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{k,\alpha}}{\pi_k^2} + \sum_{k\in S_m}\frac{1}{\pi_k^2}\right). \quad (12)
$$

计算 $B_n(\alpha)$。 请注意

$$
\begin{aligned}
&\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\left(y_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)\left(y_l - \mathbf{x}_{l,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right) \\
&= \sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} + \epsilon_k - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)\left(\mathbf{x}_l^{\top}\boldsymbol{\beta} + \epsilon_l - \mathbf{x}_{l,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right) \\
&= \sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\epsilon_k\epsilon_l + \sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\epsilon_k\left(\mathbf{x}_l^{\top}\boldsymbol{\beta} - \mathbf{x}_{l,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right) + \sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\epsilon_l\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right) \\
&\quad + \sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}}\left(\mathbf{x}_k^{\top}\boldsymbol{\beta} - \mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right)\left(\mathbf{x}_l^{\top}\boldsymbol{\beta} - \mathbf{x}_{l,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_\alpha\right). \quad (13)
\end{aligned}
$$

由上述内容可知，式（13）的前三项均值为零。考虑对最后一项作如下分解：

$$
\begin{aligned}
&\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\left(\mathbf{x}_k^\top\boldsymbol{\beta}-\mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha\right)}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha\right)}{\pi_l}\\
&=\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\left(\mathbf{x}_k^\top\boldsymbol{\beta}-\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top(\boldsymbol{X}_r\boldsymbol{\beta}+\boldsymbol{\epsilon}_r)\right)}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top(\boldsymbol{X}_r\boldsymbol{\beta}+\boldsymbol{\epsilon}_r)\right)}{\pi_l}\\
&=\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\left(\mathbf{x}_k^\top\boldsymbol{\beta}-\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{r,\alpha}^\top\boldsymbol{X}_r\boldsymbol{\beta}\right)}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\boldsymbol{\beta}\right)}{\pi_l}\\
&\quad+2\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}\boldsymbol{\epsilon}_r}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\boldsymbol{\beta}\right)}{\pi_l}\\
&\quad+\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}\boldsymbol{\epsilon}_r}{\pi_k}\frac{\boldsymbol{\epsilon}_r^\top\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{l,\alpha}}{\pi_l}.
\end{aligned}
$$

由此可得

$$
\begin{aligned}
&\mathbb{E}_m\left[\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{y_k-\mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha}{\pi_k}\frac{y_l-\mathbf{x}_{l,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha}{\pi_l}\right]\\
&=\mathbb{E}_m\left[\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\left(\mathbf{x}_k^\top\boldsymbol{\beta}-\mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha\right)}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\widehat{\boldsymbol{\beta}}_\alpha\right)}{\pi_l}\right]\\
&=\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\left(\mathbf{x}_k^\top\boldsymbol{\beta}-\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\boldsymbol{\beta}\right)}{\pi_k}\frac{\left(\mathbf{x}_l^\top\boldsymbol{\beta}-\mathbf{x}_{l,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\boldsymbol{\beta}\right)}{\pi_l}\\
&\quad+\sigma^2\sum_{k\in S_m}\sum_{\substack{l\in S_m\\ l\neq k}} \frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\mathbf{x}_{l,\alpha}}{\pi_k\pi_l}.
\end{aligned}
\tag{14}
$$

Co结合以下闭式公式： $A_n(\alpha)$ 见（12）以及 $B_n(\alpha)$ 见（14），可得

$$
\begin{aligned}
\mathcal{L}_n(\alpha)=&\left\{\sum_{k\in S_m}\pi_k^{-1}\left(\mathbf{x}_k^\top-\mathbf{x}_{k,\alpha}^\top\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^\top\boldsymbol{X}_r\right)\boldsymbol{\beta}\right\}^2\\
&+\sigma^2\left\{\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^\top}{\pi_k}\right)\boldsymbol{A}_{r,\alpha}^{-1}\left(\sum_{l\in S_m}\frac{\mathbf{x}_{l,\alpha}}{\pi_l}\right)+\sum_{k\in S_m}\frac{1}{\pi_k^2}\right\}.
\end{aligned}
$$

命题的证明*2*.

设 α_1, $\alpha_2 \in C$ 使得 $\alpha_1 \subset \alpha_2$。回顾一下，对于任意 $\alpha \in C$, $\mathcal{L}_{1,n}(\alpha) = 0$。因此，

$$\mathcal{L}_n(\alpha_1) < \mathcal{L}_n(\alpha_2) \quad \Leftrightarrow \quad \mathcal{L}_{2,n}(\alpha_1) < \mathcal{L}_{2,n}(\alpha_2).$$

由于 $\alpha^\star$ 是最小正确模型，利用引理11即可得到结论。

定理的证明*1*.

我们将证明 $\lim_{v\to\infty} \mathbb{P}_{pq}(\alpha_{opt,v} \neq \alpha^\star) = 0$。我们可以将事件 $\{\alpha_{opt,v} \neq \alpha^\star\}$ 写成一个不交并

$$\{\alpha_{opt,v} \neq \alpha^\star\} = \{\alpha_{opt,v} \in C, \alpha_{opt,v} \neq \alpha^\star\} \biguplus \{\alpha_{opt,v} \in \mathcal{W}\},$$

使得

$$\mathbb{P}_{pq}(\alpha_{opt,v} \neq \alpha^\star) = \mathbb{P}_{pq}(\alpha_{opt,v} \in C, \alpha_{opt,v} \neq \alpha^\star) + \mathbb{P}_{pq}(\alpha_{opt,v} \in \mathcal{W}) := p_{1,v} + p_{2,v}.$$

根据命题2，我们可直接得到 $p_{1,v} = 0$ 对所有 $v \in \mathbb{N}$均成立。

令 $\alpha_F := \{1, \ldots, p\} \in C$ 为所有协变量组成的集合。对于 $p_{2,v}$，考虑以下分解

$$\begin{aligned}
p_{2,v} &= \mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_{opt,v}) < \mathcal{L}_n(\alpha^\star)\right) \\
&= \mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_{opt,v}) < \mathcal{L}_n(\alpha^\star), \mathcal{L}_n(\alpha^\star) \leq \mathcal{L}_n(\alpha_F)\right) \\
&\quad + \mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_{opt,v}) < \mathcal{L}_n(\alpha^\star), \mathcal{L}_n(\alpha^\star) > \mathcal{L}_n(\alpha_F)\right) \\
&\leq \mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_{opt,v}) < \mathcal{L}(\alpha_F)\right) + \mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_F) < \mathcal{L}_n(\alpha^\star)\right) := q_{1,v} + q_{2,v}.
\end{aligned}$$

收敛于 $q_{1,v}$。一方面，对于 $\alpha_{opt,v} \in \mathcal{W}$，损失 $\mathcal{L}_{1,n}(\alpha_{opt,v})$ 满足以下恒等式

$$\begin{aligned}
\mathcal{L}_{1,n}(\alpha_{opt,v}) &= \left\{ \sum_{k \in S_{m,v}} \pi_k^{-1} \left(\mathbf{x}_k^\top - \mathbf{x}_{k,\alpha}^\top \boldsymbol{A}_{r,\alpha}^{-1} \boldsymbol{X}_{r,\alpha}^\top \boldsymbol{X}_r \right) \boldsymbol{\beta} \right\}^2 \\
&= \left\{ \sum_{k \in S_{m,v}} \pi_k^{-1} \Big(\mathbf{x}_{k,\alpha_{opt,v}}^\top \boldsymbol{\beta}_{\alpha_{opt,v}} + \mathbf{x}_{k,\alpha_{opt,v}^c}^\top \boldsymbol{\beta}_{\alpha_{opt,v}^c} \right. \\
&\quad \underbrace{- \mathbf{x}_{k,\alpha_{opt,v}}^\top \boldsymbol{A}_{r,\alpha_{opt,v}}^{-1} \boldsymbol{X}_{r,\alpha_{opt,v}}^\top \boldsymbol{X}_{r,\alpha_{opt,v}} \boldsymbol{\beta}_{\alpha_{opt,v}}}_{= \mathbf{x}_{k,\alpha_{opt,v}}^\top \boldsymbol{\beta}_{\alpha_{opt,v}}} \\
&\quad \left. - \mathbf{x}_{k,\alpha_{opt,v}}^\top \boldsymbol{A}_{r,\alpha_{opt,v}}^{-1} \boldsymbol{X}_{r,\alpha_{opt,v}}^\top \boldsymbol{X}_{r,\alpha_{opt,v}^c} \boldsymbol{\beta}_{\alpha_{opt,v}^c} \Big) \right\}^2 \qquad (15) \\
&= \left\{ \sum_{k \in S_m} \pi_k^{-1} \left(\mathbf{x}_{k,\alpha_{opt,v}^c}^\top - \mathbf{x}_{k,\alpha_{opt,v}}^\top \boldsymbol{A}_{r,\alpha_{opt,v}}^{-1} \boldsymbol{X}_{r,\alpha_{opt,v}}^\top \boldsymbol{X}_{r,\alpha_{opt,v}^c} \right) \boldsymbol{\beta}_{\alpha_{opt,v}^c} \right\}^2 . \qquad (16)
\end{aligned}$$

另一方面，由于 $\alpha_F \in C$，我们有 $\mathcal{L}_{1,\ n}(\alpha_F)=0$。因此，$\mathcal{L}_n\left(\alpha_{opt,\ v}\right)<\mathcal{L}_n(\alpha_F)$等价于

$$\begin{aligned}\mathcal{L}_n\left(\alpha_{opt,v}\right)<\mathcal{L}_n\left(\alpha_F\right) \quad &\Leftrightarrow \quad \mathcal{L}_{1,n}\left(\alpha_{opt,v}\right)+\mathcal{L}_{2,n}\left(\alpha_{opt,v}\right)<\mathcal{L}_{1,n}\left(\alpha_F\right)+\mathcal{L}_{2,n}\left(\alpha_F\right) \\ &\Leftrightarrow \quad \mathcal{L}_{1,n}\left(\alpha_{opt,v}\right)<\mathcal{L}_{2,n}\left(\alpha_F\right)-\mathcal{L}_{2,n}\left(\alpha_{opt,v}\right). \end{aligned} \tag{17}$$

结合（15）和（17）可得

$$\begin{aligned}&\mathcal{L}_{1,n}\left(\alpha_{opt,v}\right)<\mathcal{L}_{2,n}\left(\alpha_F\right)-\mathcal{L}_{2,n}\left(\alpha_{opt,v}\right) \\ &\Leftrightarrow\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha^c_{opt,v}}^{\top}}{N_v\pi_k}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\mathbf{x}_{r,\alpha^c_{opt,v}}\right) \\ &\quad\times\boldsymbol{\beta}_{\alpha^c_{opt,v}}\boldsymbol{\beta}^{\top}_{\alpha^c_{opt,v}}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha^c_{opt,v}}^{\top}}{N_v\pi_k}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\mathbf{x}_{r,\alpha^c_{opt,v}}\right)^{\top} \\ &\quad<\sigma^2\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_F}^{\top}}{N_v\pi_k}\right)\boldsymbol{A}^{-1}_{r,\alpha_F}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_F}}{N_v\pi_k}\right) \\ &\quad-\sigma^2\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\right)\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}}{N_v\pi_k}\right).\end{aligned}$$

Usi等式（42），令 $\alpha_1:=\alpha_{opt,\ v}, \alpha_2:=\alpha_F, \alpha^c_{opt,\ v}:=\alpha^c_{opt,\ v}$，我们得到 in

$$\begin{aligned}&\sigma^2\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_F}^{\top}}{N_v\pi_k}\right)\boldsymbol{A}^{-1}_{r,\alpha_F}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_F}}{N_v\pi_k}\right)-\sigma^2\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\right)\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}}{N_v\pi_k}\right) \\ &=\sigma^2\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha^c_{opt,v}}^{\top}}{N_v\pi_k}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\mathbf{x}_{r,\alpha^c_{opt,v}}\right) \\ &\quad\times\boldsymbol{S}_v^{-1}\left(\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha^c_{opt,v}}^{\top}}{N_v\pi_k}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\mathbf{x}_{r,\alpha^c_{opt,v}}\right)^{\top},\end{aligned}$$

其中

$$\boldsymbol{S}_v=\mathbf{x}^{\top}_{r,\alpha^c_{opt,v}}\left(\boldsymbol{I}_{n_{r,v}}-\mathbf{x}_{r,\alpha_{opt,v}}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\right)\mathbf{x}_{r,\alpha^c_{opt,v}}$$

是一个正定矩阵。令

$$\mathbf{z}_v^{\top}=\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha^c_{opt,v}}^{\top}}{N_v\pi_k}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha_{opt,v}}^{\top}}{N_v\pi_k}\boldsymbol{A}^{-1}_{r,\alpha_{opt,v}}\mathbf{x}^{\top}_{r,\alpha_{opt,v}}\mathbf{x}_{r,\alpha^c_{opt,v}}.$$

则 $\mathcal{L}_{1,\ n}\left(\alpha_{opt,\ v}\right)<\mathcal{L}_{2,\ n}\left(\alpha_F\right)-\mathcal{L}_{2,\ n}\left(\alpha_{opt,\ v}\right)$化简为

$$\mathcal{L}_{1,n}\left(\alpha_{opt,v}\right)<\mathcal{L}_{2,n}\left(\alpha_F\right)-\mathcal{L}_{2,n}\left(\alpha_{opt,v}\right) \quad \Leftrightarrow \quad \mathbf{z}_v^{\top}\boldsymbol{\beta}_{\alpha^c_{opt,v}}\boldsymbol{\beta}^{\top}_{\alpha^c_{opt,v}}\mathbf{z}_v<\sigma^2\mathbf{z}_v^{\top}\boldsymbol{S}_v^{-1}\mathbf{z}_v.$$

考虑以下$q_{1,\ v}$的分解，

$$\lim_{v\to\infty}q_{1,v}=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\mathbf{z}_v^{\top}\boldsymbol{\beta}_{\alpha^c_{opt,v}}\boldsymbol{\beta}^{\top}_{\alpha^c_{opt,v}}\mathbf{z}_v<\sigma^2\mathbf{z}_v^{\top}\boldsymbol{S}_v^{-1}\mathbf{z}_v,\mathbf{z}_v\neq\mathbf{0}\right)$$

$$+\lim_{v\to\infty}\mathbb{P}_{pq}\left(\mathbf{z}_v^\top\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top\mathbf{z}_v<\sigma^2\mathbf{z}_v^\top\boldsymbol{S}_v^{-1}\mathbf{z}_v,\mathbf{z}_v=\mathbf{0}\right)$$

如果$\mathbf{z}_v=\mathbf{0}$，那么事件$\{\mathbf{z}_v^\top\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top\mathbf{z}_v<\sigma^2\mathbf{z}_v^\top\boldsymbol{S}_v^{-1}\mathbf{z}_v\}$可简化为$\{0<0\}=\emptyset$。因此，

$$\mathbb{P}_{pq}\left(\mathbf{z}_v^\top\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top\mathbf{z}_v<\sigma^2\mathbf{z}_v^\top\boldsymbol{S}_v^{-1}\mathbf{z}_v,\mathbf{z}_v=\mathbf{0}\right)=0,\quad\text{for all }v\in\mathbb{N}.$$

回顾一下，对于两个二次型$\mathrm{z}^\top\boldsymbol{A}\mathrm{z}$和$\mathrm{z}^\top\boldsymbol{B}\mathrm{z}$，若对于$\mathrm{z}\neq\mathbf{0}$，有$\mathrm{z}^\top\boldsymbol{A}\mathrm{z}<\mathrm{z}^\top\boldsymbol{B}\mathrm{z}$，则$\lambda_{max}(\boldsymbol{A}-\boldsymbol{B})<0$。此外，根据柯朗 - 费舍尔定理，对于实值对称矩阵$\boldsymbol{A}$和$\boldsymbol{B}$，我们有$\lambda_{max}(\boldsymbol{A})+\lambda_{min}(\boldsymbol{B})\leq\lambda_{max}(\boldsymbol{A}+\boldsymbol{B})\leq\lambda_{max}(\boldsymbol{A})+\lambda_{max}(\boldsymbol{B})$。因此，结合上述两个结论，我们得到

$$\begin{aligned}
&\lim_{v\to\infty}\mathbb{P}_{pq}\left(\mathbf{z}_v^\top\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top\mathbf{z}_v<\sigma^2\mathbf{z}_v^\top\boldsymbol{S}_v^{-1}\mathbf{z}_v,\mathbf{z}_v\neq\mathbf{0}\right)\\
&=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{max}(\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top-\sigma^2\boldsymbol{S}_v^{-1})<0\right)\\
&\leq\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{max}(\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top)+\sigma^2(\lambda_{min}(-\boldsymbol{S}_v^{-1}))<0\right)\\
&=\lim_{v\to\infty}\left(\lambda_{max}(\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top)-\sigma^2(\lambda_{max}(\boldsymbol{S}_v^{-1}))<0\right)\\
&=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{max}(\boldsymbol{\beta}_{\alpha_{opt,v}^c}\boldsymbol{\beta}_{\alpha_{opt,v}^c}^\top)-\sigma^2(\lambda_{min}(\boldsymbol{S}_v))^{-1}<0\right)\\
&=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2-N_v^{-1}\sigma^2(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v))^{-1}<0\right)\\
&=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<\frac{\sigma^2}{N_v\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2}\right).
\end{aligned}$$

根据引理 4，存在一个常数$K_0>0$使得

$$\begin{aligned}
&\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<\frac{\sigma^2}{N_v\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2}\right)\\
&=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<\frac{\sigma^2}{N_v\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2},\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)\geq K_0\right)\\
&\quad+\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<\frac{\sigma^2}{N_v\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2},\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<K_0\right)\\
&\leq\lim_{v\to\infty}\mathbb{P}_{pq}\left(K_0\leq\frac{\sigma^2}{N_v\|\boldsymbol{\beta}_{\alpha_{opt,v}^c}\|_2^2}\right)+\lim_{v\to\infty}\mathbb{P}_{pq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v)<K_0\right)=0.
\end{aligned}$$

因此可得，$\lim_{v\to\infty}q_{1,v}=0$几乎必然。

的收敛$q_{2,v}$。接下来要证明，几乎必然，

$$\lim_{v\to\infty}q_{2,v}=\lim_{v\to\infty}\mathbb{P}_{pq}\left(\mathcal{L}_n(\alpha_F)<\mathcal{L}_n(\alpha^\star)\right)=0.$$

然而，由于$\alpha^\star\subset\alpha_F\in C$，事件$\mathcal{L}_n(\alpha_F)<\mathcal{L}_n(\alpha^\star)$意味着$\mathcal{L}_{2,n}(\alpha_F)<\mathcal{L}_{2,n}(\alpha^\star)$，根据命题 2，这是一个可忽略事件。因此$q_{2,v}=0$对所有$v\in\mathbb{N}$都成立。因此，我们得出结论$\lim_{v\to\infty}p_{2,v}=0$，这进而意味着

$$\lim_{v\to\infty}\mathbb{P}_{pq}\left(\alpha_{opt,v}\neq\alpha^\star\right)=0$$

几乎必然。最后，应用勒贝格收敛定理可得

$$\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\alpha_{opt,v}=\alpha^{\star}\right)=1$$

几乎必然。

命题*3*的证明。

陈述（i）的证明。

对于 $\alpha\in\mathcal{A}$，考虑 $\widehat{\mu}_{\alpha}$ 的分解，$_v-\mu_v$

$$\widehat{\mu}_{\alpha,v}-\mu_v=\underbrace{\widehat{\mu}_{\alpha,v}-\widehat{\mu}_{\pi,v}}_{(a)}+\underbrace{\widehat{\mu}_{\pi,v}-\mu_v}_{(b)}.$$

由于对于所有 $k\in U_v$，$\pi_k=\pi$，我们有

$$\widehat{\boldsymbol{\beta}}_{\alpha,v}=\boldsymbol{A}_{r,\alpha}^{-1}\boldsymbol{X}_{r,\alpha}^{\top}\boldsymbol{Y}_r=\boldsymbol{A}_{r,\alpha}^{-1}\sum_{k\in S_{r,v}}\mathbf{x}_{k,\alpha}y_k=\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v\pi}\right)^{-1}\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v\pi}.$$

展开项 (a)，可得

$$\begin{aligned}\widehat{\mu}_{\alpha,v}-\widehat{\mu}_{\pi,v}&=\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}\right)+\epsilon_k}{N_v\pi}\\&=\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}}{N_v\pi}\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v\pi}\right)^{-1}\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v\pi}-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}}{N_v\pi}+\sum_{k\in S_{r,v}}\frac{\epsilon_k}{N_v\pi}\\&=\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}}{N_v\pi}\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v\pi}\right)^{-1}\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}\mathbf{x}_k^{\top}\boldsymbol{\beta}+\mathbf{x}_{k,\alpha}\epsilon_k}{N_v\pi}\\&\quad-\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}}{N_v\pi}+\sum_{k\in S_{r,v}}\frac{\epsilon_k}{N_v\pi}.\end{aligned}$$

由大数定律可得

$$\begin{aligned}&\sum_{k\in S_{m,v}}\frac{\epsilon_k}{N_v\pi}\xrightarrow[v\to\infty]{\mathbb{P}}0,\qquad\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}}{N_v\pi}\xrightarrow[v\to\infty]{\mathbb{P}}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha}^{\top}\boldsymbol{\beta}_{\alpha}\right],\\&\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v\pi}\right)^{-1}\xrightarrow[v\to\infty]{\mathbb{P}}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1},\\&\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}\mathbf{x}_k^{\top}\boldsymbol{\beta}+\mathbf{x}_{k,\alpha}\epsilon_k}{N_v\pi}\xrightarrow[v\to\infty]{\mathbb{P}}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}^{\top}\boldsymbol{\beta}\right].\end{aligned}$$

利用连续映射定理，我们得到

$$\begin{aligned}\widehat{\mu}_{\alpha,v}-\mu_v\xrightarrow[v\to\infty]{\mathbb{P}}&\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\,\mathbf{x}_{\alpha}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}^{\top}\boldsymbol{\beta}\right]\\&-\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\,\mathbf{x}^{\top}\boldsymbol{\beta}\right]\end{aligned}$$

回想一下，若 $\widehat{\mu}_{\alpha,v}-\mu_v \xrightarrow[v\to\infty]{\mathbb{P}} 0$，则 $\widehat{\mu}_{\alpha,v}-\mu_v$ 是相合的。等价地，

$$\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\,\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}^\top\boldsymbol{\beta}\right]-\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\,\mathbf{x}^\top\boldsymbol{\beta}\right]=0.$$

要得到（C_2），可对上述等式进行分解$\mathbf{x}^\top\boldsymbol{\beta}=\mathbf{x}_\alpha^\top\boldsymbol{\beta}_\alpha+\mathbf{x}_{\alpha^c}^\top\boldsymbol{\beta}_\alpha^c$，经过一些代数运算即可得到结果。

命题（ii）（a）的证明。

根据定理3，对于 $\alpha\in C$，我们有

$$\widehat{\mu}_{\alpha,v}-\mu_v=\widehat{\mu}_{\alpha,v}-\widetilde{\mu}_{\alpha,v}+o_{\mathbb{P}}\left(n_v^{-1/2}\right).$$

因此，$\mathbb{AV}\left(\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)\right)=\mathbb{AV}\left(\sqrt{N_v}\left(\widehat{\mu}_{\alpha,v}-\mu_v\right)\right)$。为得到渐近方差，我们定义

$$\begin{aligned}\mathbb{AV}\left(\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v)\right)=\lim_{v\to\infty}\Bigg(&\mathbb{E}_{mq}\left[\mathbb{V}_p\left(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)\right)\right]\\&+\mathbb{E}_q\left[\mathbb{V}_m\left(\mathbb{E}_p[\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)]\right)\right]\\&+\mathbb{V}_q\left(\mathbb{E}_{mp}[\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)]\right)\Bigg).\end{aligned}$$

回顾 $\mathbb{V}_q\left(\mathbb{E}_{mp}[\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)]\right)=0$，只剩下前两项，因此我们写作

$$\begin{aligned}\mathbb{AV}\left(\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v)\right)=\lim_{v\to\infty}\Bigg(&\mathbb{E}_{mq}\left[\mathbb{V}_p\left(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)\right)\right]\\&+\mathbb{E}_q\left[\mathbb{V}_m\left(\mathbb{E}_p[\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)]\right)\right]\Bigg)\\:=\ &\mathrm{AV}_1+\mathrm{AV}_2,\end{aligned}$$

AV的渐近行为1。主要思想。如果存在$c_{1,\alpha}$使得

$$\mathbb{V}_p\left(\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)\right)\xrightarrow[v\to\infty]{\mathbb{P}}c_{1,\alpha},$$

并且如果 $\mathbb{V}_p(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v))$一致可积，则几乎必然，我们将有

$$\lim_{v\to\infty}\mathbb{E}_{mq}\left[\left|\mathbb{V}_p\left(\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)\right)-c_{1,\alpha}\right|\right]=0.$$

此外，詹森不等式将意味着$\lim_{v\to\infty}\left|\mathbb{E}_{mq}[\mathbb{V}_p(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v))-c_{1,\alpha}]\right|=0$几乎必然。我们将遵循上述架构。

一致可积性。根据引理23可知，$\mathbb{V}_p(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v))$具有一致可积性；该引理证明了 $\mathbb{E}_{mq}[(\mathbb{V}_p(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)))^2]$有界。

极限确定。现在我们继续求出 $c_{1,\alpha}$，使得 $\mathbb{V}_p(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)) \xrightarrow[v\to\infty]{\mathbb{P}} c_{1,\alpha}$.为此，令

$$\begin{aligned}
\mathbb{V}_p\left(\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)\right) &= \frac{1}{N_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{\eta_{k,\alpha}}{\pi}\frac{\eta_{l,\alpha}}{\pi}\Delta \\
&= \frac{1}{N_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha}{\pi}\frac{\mathbf{x}_{l,\alpha}^{\top}\boldsymbol{\beta}_\alpha}{\pi}\Delta \\
&\quad+\frac{2}{N_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha}{\pi}\frac{r_l(1+\pi\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{l,\alpha})\epsilon_l}{\pi}\Delta \\
&\quad+\frac{1}{N_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{r_k(1+\pi\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha})\epsilon_k}{\pi}\frac{r_l(1+\pi\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{l,\alpha})\epsilon_l}{\pi}\Delta \\
&:= A_v+B_v+C_v.
\end{aligned}$$

接下来，我们分别研究 A_v、 B_v和 C_v 的渐近行为。

的行为 A_v。由于 $\alpha\in\mathsf{C}$，我们有$\mathbf{x}_{k,\ \alpha}^{\top}\boldsymbol{\beta}_\alpha=\mathbf{x}_{k\ ,\ \alpha^\star}^{\top}\boldsymbol{\beta}_{\alpha^\star}$ 因此 A_v 不依赖于α。因此，存在一个常数 C_1（$\alpha^\star$），使得 A_v-C_1（$\alpha^\star$）$=o_{\mathbb{P}}(1)$。

B_v**的行为**。记

$$\begin{aligned}
B_v &= 2\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha r_k\left(1+\pi\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v} \\
&\quad+\frac{2}{N_v}\frac{1-\pi}{\pi}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha}{\pi}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{l,\alpha}\right)\epsilon_l}{\pi}\Delta \\
&:= B_{1,v}+B_{2,v}.
\end{aligned}$$

通过代入 $\pi\mathbf{c}_{\alpha,\ v}$, $B_{1,\ v}$ 可得

$$\begin{aligned}
B_{1,v} &= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha r_k\epsilon_k}{N_v} \\
&\quad+\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha r_k}{N_v}\mathbf{x}_{k,\alpha}^{\top}\epsilon_k\left(\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\,\mathbf{x}_{k,\alpha}}{N_v}.
\end{aligned}$$

根据大数定律，我们得到

$$\frac{1}{N_v}\sum_{k\in U_v}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha r_k\epsilon_k\xrightarrow[v\to\infty]{\mathbb{P}}\mathbb{E}_{\mathbf{x}pq}\left[\mathbf{x}_\alpha^{\top}\boldsymbol{\beta}_\alpha p(\mathbf{x})\epsilon\right]=0,$$

$$\frac{1}{N_v}\sum_{k\in U_v}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha r_k\mathbf{x}_{k,\alpha}^{\top}\epsilon_k\xrightarrow[v\to\infty]{\mathbb{P}}\mathbb{E}_{\mathbf{x}pq}\left[\mathbf{x}_\alpha^{\top}\boldsymbol{\beta}_\alpha p(\mathbf{x})\mathbf{x}_\alpha\epsilon\right]=\mathbf{0}_\alpha^{\top},$$

and

$$\left(\frac{1}{N_v}\sum_{k\in U_v} r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top\right)^{-1}\frac{1}{N_v}\sum_{k\in U_v}(1-r_k)\mathbf{x}_{k,\alpha}\xrightarrow[v\to\infty]{\mathbb{P}}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\,\mathbf{x}_\alpha\right].$$

因此，$B_{1,\ v}\xrightarrow[v\to\infty]{\mathbb{P}}0$。 </style}对于 ，$_v$，回想一下， $K:=\lim_{v\to\infty}N_v\Delta$，我们有

$$\begin{aligned}
B_{2,v} &= \frac{2}{N_v}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{\pi}\Delta\\
&= \frac{2}{\pi^2}N_v\Delta\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{N_v}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{N_v}\\
&= \frac{2}{\pi^2}(K+o(1))\Bigg(\sum_{k\in U_v}\sum_{l\in U_v}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{N_v}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{N_v}\\
&\quad-\frac{1}{N_v}\sum_{k\in U_v}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v}\Bigg)\\
&= \frac{2}{\pi^2}(K+o(1))\left(\left(\sum_{k\in U_v}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{N_v}\right)\left(\sum_{k\in U_v}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v}\right)+o_{\mathbb{P}}\left(N_v^{-1}\right)\right)\\
&= \frac{2}{\pi^2}(K+o(1))\left((\mathbb{E}_{\mathbf{x}}[\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha]+o_{\mathbb{P}}(1))\left(0+o_{\mathbb{P}}(1))+o_{\mathbb{P}}(N_v^{-1}\right)\right)=o_{\mathbb{P}}(1),
\end{aligned}$$

因为

$$\begin{aligned}
\sum_{k\in U_v}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v} &= \sum_{k\in U_v}\frac{r_k\epsilon_k}{N_v}\\
&\quad+\sum_{k\in U_v}\frac{r_k\epsilon_k\mathbf{x}_{k,\alpha}^\top}{N_v}\left(\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\,\mathbf{x}_{k,\alpha}}{N_v}\\
&\xrightarrow[v\to\infty]{\mathbb{P}}0+\mathbf{0}_\alpha^\top\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}[(1-p(\mathbf{x}))\mathbf{x}_\alpha]=0. \quad (18)
\end{aligned}$$

最后，我们得到 $B_v\xrightarrow{\mathbb{P}}0.v\to\infty$

C_v的表现。写

$$\begin{aligned}
C_v &= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)^2\epsilon_k^2}{N_v}\\
&\quad+\frac{1}{N_v}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{\pi}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{\pi}\Delta\\
&:= C_{1,v}+C_{2,v}.
\end{aligned}$$

展开$C_{1,v}$ 得到

$$
\begin{aligned}
C_{1,v} &= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)^2\epsilon_k^2}{N_v}\\
&= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2}{N_v}+2\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}}{N_v}+\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2\left(\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)^2}{N_v}\\
&= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2}{N_v}+2\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2\pi\mathbf{x}_{k,\alpha}^\top\mathbf{c}_{\alpha,v}}{N_v}\\
&\quad+\frac{1-\pi}{\pi}\left(\pi\mathbf{c}_{\alpha,v}^\top\right)\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\left(\pi\mathbf{c}_{\alpha,v}\right)\\
&= \frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2}{N_v}+2\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2\mathbf{x}_{k,\alpha}^\top}{N_v}\left(\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v}\\
&\quad+\frac{1-\pi}{\pi}\sum_{k\in U_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}^\top}{N_v}\left(\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\left(\sum_{k\in U_v}\frac{r_k^2\epsilon_k^2}{N_v}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top\right)\\
&\quad\times\left(\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v}.
\end{aligned}
$$

因此，利用连续映射定理，$C_{1,\ v}$ 收敛到

$$
\begin{aligned}
&\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\right]-2\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha\right]\\
&\quad+\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\\
&\quad\times\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha\right]\\
&=\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]-2\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\right]\\
&\quad+\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha\right]\\
&\overset{(*)}{=}\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]-2\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})]+\sigma^2\frac{1-\pi}{\pi}M(\alpha)\\
&=\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}[2-p(\mathbf{x})]+\sigma^2\frac{1-\pi}{\pi}M(\alpha).
\end{aligned}
$$

等式 $(*)$ 由引理 12得出。对于$C_{2,\ v}$，利用（18），可写为

$$
\begin{aligned}
C_{2,v} &= \frac{1}{N_v}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{\pi}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{\pi}\Delta\\
&= \frac{N_v\Delta}{\pi^2}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{N_v}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{K+o(1)}{\pi^2}\left(\sum_{k\in U_v}\sum_{l\in U_v}\frac{r_k\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k}{N_v}\frac{r_l\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{l,\alpha}\right)\epsilon_l}{N_v}\right.\\
&\quad\left.-\frac{1}{N_v}\sum_{k\in U_v}\frac{r_k^2\left(1+\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\epsilon_k^2}{N_v}\right)\\
&= \frac{K+o(1)}{\pi^2}\left(\left(\sum_{k\in U_v}\frac{r_k(1+\pi\mathbf{c}_{\alpha,v}\mathbf{x}_{k,\alpha})\epsilon_k}{N_v}\right)^2+o_{\mathbb{P}}\left(n_v^{-1}\right)\right)\\
&= \frac{K+o(1)}{\pi^2}\left((0+o_{\mathbb{P}}(1))^2+o_{\mathbb{P}}\left(n_v^{-1}\right)\right)=o_{\mathbb{P}}(1).
\end{aligned}
$$

因此，$C_{2,\ v}\xrightarrow[v\to\infty]{\mathbb{P}}0$。最后，我们得出结论

$$
C_v\xrightarrow[v\to\infty]{\mathbb{P}}\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[2-p(\mathbf{x})\right]+\sigma^2\frac{1-\pi}{\pi}M(\alpha).
$$

总体而言，

$$
c_{1,\alpha}=C_1(\alpha^\star)+\sigma^2\frac{1-\pi}{\pi}\mathbb{E}_{\mathbf{x}}\left[2-p(\mathbf{x})\right]+\sigma^2\frac{1-\pi}{\pi}M(\alpha).
$$

AV 的渐近行为$_2$。我们采用与之前类似的方法。

一致可积性。在引理 23 中，我们证明 $\mathbb{E}_{mq}[(\mathbb{V}_m(\mathbb{E}_p[\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)]))^2]$ 几乎必然有界，由此可推出一致可积性。

极限确定。记

$$
\begin{aligned}
&\mathbb{V}_m\left(\mathbb{E}_p\left[\sqrt{N_v}\left(\widetilde{\mu}_{\alpha,v}-\mu_v\right)\right]\right)\\
&\quad=\sigma^2\sum_{k\in U_v}\frac{1-r_k+r_k(\pi\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha})^2}{N_v}\\
&\quad=\sigma^2\sum_{k\in U_v}\frac{1-r_k}{N_v}+\sigma^2\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{k,\alpha}^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\\
&\qquad\times\sum_{k\in U_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{k,\alpha}\right]\\
&\quad\xrightarrow[v\to\infty]{\mathbb{P}}\sigma^2\mathbb{E}_{\mathbf{x}}\left[1-p(\mathbf{x})\right]+\sigma^2\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_\alpha^\top\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_\alpha\mathbf{x}_\alpha^\top\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{k,\alpha}\right]\\
&\quad=\sigma^2\mathbb{E}_{\mathbf{x}}\left[1-p(\mathbf{x})\right]+\sigma^2M(\alpha).
\end{aligned}
$$

因此，

$$
c_{2,\alpha}=\sigma^2\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})]+\sigma^2M(\alpha).
$$

综合以上各项。将不依赖于 α 的项吸收到称为 $C(\alpha^\star)$ 的量中，我们得出

$$
\mathbb{AV}\left(\sqrt{N_v}(\widehat{\mu}_{\alpha,v}-\mu_v)\right)=\lim_{v\to\infty}\left(\mathbb{E}_{mq}\left[\mathbb{V}_p\left(\sqrt{N_v}(\widetilde{\mu}_{\alpha,v}-\mu_v)\right)\right]\right.
$$

$$
\begin{aligned}
&+ \mathbb{E}_q \left[\mathbb{V}_m \left(\mathbb{E}_p \left[\sqrt{N_v} (\widetilde{\mu}_{\alpha,v} - \mu_v) \right] \right) \right] \Bigg) \\
&= C_1(\alpha^\star) + \sigma^2 \frac{1-\pi}{\pi} \mathbb{E}_{\mathbf{x}}[2 - p(\mathbf{x})] + \sigma^2 \frac{1-\pi}{\pi} M(\alpha) \\
&\quad + \sigma^2 \mathbb{E}_{\mathbf{x}}[1 - p(\mathbf{x})] + \sigma^2 M(\alpha) \\
&= C(\alpha^\star) + \frac{\sigma^2}{\pi} M(\alpha).
\end{aligned}
$$

推论 *1*的证明。

陈述（i）显然成立，因为 $\|\boldsymbol{\beta}_{\alpha^c}\|_2 = 0$，这意味着 $\boldsymbol{\beta}_{\alpha^c} = \mathbf{0}_{\alpha^c}$。对于陈述（ii），我们需要证明在我们的假设下，

$$
\mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) \mathbf{x}_{\alpha^c}^\top \boldsymbol{\beta}_{\alpha^c} \right] = \mathbb{E}_{\mathbf{x}} \left[\mathbf{x}_\alpha^\top (1 - p(\mathbf{x})) \right] \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha \mathbf{x}_\alpha^\top \right] \right)^{-1} \mathbb{E}_{\mathbf{x}} \left[\mathbf{x}_\alpha \mathbf{x}_{\alpha^c}^\top \boldsymbol{\beta}_{\alpha^c} p(\mathbf{x}) \right]
$$

为简化记号，我们省略了下标。这等价于证明

$$
\mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) x_{\alpha^c,j} \right] = \mathbb{E}_{\mathbf{x}} \left[\mathbf{x}_\alpha^\top (1 - p(\mathbf{x})) \right] \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha \mathbf{x}_\alpha^\top \right] \right)^{-1} \mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha x_{\alpha^c,j} \right], \quad (19)
$$

$j \in \alpha_{\text{mis}}$。由于 $j \in \alpha^c \backslash \alpha_{\text{mis}}$，$\beta_{\alpha^c,\, j} = 0$，我们可以将（19）的左侧表示为

$$
L_j := \mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) x_{\alpha^c,j} \right] = \mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) \mathbf{x}_\alpha^\top \gamma_j \right],
$$

同时利用条件独立性（a）和线性连接（b）。同样，采用相同的方法，我们可以写出

$$
\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha x_{\alpha^c,j} \right] = \mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha \mathbf{x}_\alpha^\top \right] \boldsymbol{\gamma}_j.
$$

因此，（19）的右侧等于

$$
\begin{aligned}
R_j &:= \mathbb{E}_{\mathbf{x}} \left[\mathbf{x}_\alpha^\top (1 - p(\mathbf{x})) \right] \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha \mathbf{x}_\alpha^\top \right] \right)^{-1} \mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_\alpha x_{\alpha^c,j} \right] \\
&= \mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) \mathbf{x}_\alpha^\top \right] \boldsymbol{\gamma}_j = L_j,
\end{aligned}
$$

$j \in \alpha_{\text{mis}}$。结论由此得出。

推论的证明 *2*。

（*i*）的证明。

根据命题 3，当 $j \in \alpha_2 - \alpha_1$ 时等式成立，我们需要证明

$$
\mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) \mathbf{x}_{\alpha_1}^\top \right] \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_{\alpha_1} \mathbf{x}_{\alpha_1}^\top \right] \right)^{-1} \mathbb{E}_{\mathbf{x}} [p(\mathbf{x}) \mathbf{x}_{\alpha_1} x_{\alpha_2 - \alpha_1, j}] = \mathbb{E}_{\mathbf{x}} [(1 - p(\mathbf{x})) x_{\alpha_2 - \alpha_1, j}]. \quad (20)
$$

回顾一下，如果 Y和 Z在给定 X的条件下相互独立，那么 $\mathbb{E}[YZ|X] = \mathbb{E}[Y|X]\mathbb{E}[Z|X]$。接下来，我们需要证明（20）的左边和右边相等。（20）的左边为

$$
L_j = \mathbb{E}_{\mathbf{x}} \left[(1 - p(\mathbf{x})) \mathbf{x}_{\alpha_1}^\top \right] \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}) \mathbf{x}_{\alpha_1} \mathbf{x}_{\alpha_1}^\top \right] \right)^{-1} \mathbb{E}_{\mathbf{x}} [p(\mathbf{x}) \mathbf{x}_{\alpha_1} x_{\alpha_2 - \alpha_1, j}]
$$

$$
\begin{aligned}
&= \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}[\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})\mathbf{x}_{\alpha_1}x_{\alpha_2-\alpha_1,j}|\mathbf{x}_{\alpha_1}]] \\
&= \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}[\mathbf{x}_{\alpha_1}\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbb{E}_{\mathbf{x}}[x_{\alpha_2-\alpha_1,j}|\mathbf{x}_{\alpha_1}]] \\
&= \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\boldsymbol{\gamma}_j\right] \\
&= \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbf{x}_{\alpha_1}^{\top}\right]\boldsymbol{\gamma}_j \\
&= \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbf{x}_{\alpha_1}^{\top}\boldsymbol{\gamma}_j\right]
\end{aligned}
$$

类似地，（20）的右边为

$$
\begin{aligned}
R_j &= \mathbb{E}_{\mathbf{x}}[(1-p(\mathbf{x}))x_{\alpha_2-\alpha_1,j}] = \mathbb{E}_{\mathbf{x}}[\mathbb{E}_{\mathbf{x}}[(1-p(\mathbf{x}))x_{\alpha_2-\alpha_1,j}|\mathbf{x}_{\alpha}]] \\
&= \mathbb{E}_{\mathbf{x}}[\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbb{E}_{\mathbf{x}}[x_{\alpha_2-\alpha_1,j}|\mathbf{x}_{\alpha_1}]] \\
&= \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})|\mathbf{x}_{\alpha_1}]\mathbf{x}_{\alpha_1}^{\top}\boldsymbol{\gamma}_j\right].
\end{aligned}
$$

因此，在我们的假设下，两边相等。

（*ii*）的证明。

设$\mathbf{c}$为满足 (a) 和 (b) 的一个方向；我们需要证明

$$
\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\right] \neq \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\right].
$$

我们将上式与$\mathbf{c}$作欧几里得内积。左边得到

$$
L = \mathbf{c}^{\top}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\right] = \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{c}^{\top}\mathbf{x}_{\alpha_2-\alpha_1}\right] \neq 0
$$

根据 (b)。右边，

$$
\begin{aligned}
R_j &= \mathbf{c}^{\top}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\right] \\
&= \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{c}^{\top}\mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha_1}\right] \\
&= 0
\end{aligned}
$$

因为根据 (a),

$$
\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{c}^{\top}\mathbf{x}_{\alpha_2-\alpha_1}\mathbf{x}_{\alpha_1}^{\top}\right] = \mathbf{0}.
$$

引理 ***1***的证明。

假设

$$
\lim_{v\to\infty}\mathbb{P}_m\left(\widetilde{\alpha}_{U_v} = \alpha^{\star}\right) = 1 \tag{21}
$$

几乎必然成立，我们希望证明

$$
\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\widehat{\alpha}_{S_{r,v}} \neq \alpha^{\star}\right) = \lim_{v\to\infty}\mathbb{P}_{mpq}\left(C_{S_{r,v}}\left(\widehat{\alpha}_{S_{r,v}}\right) < C_{S_{r,v}}\left(\alpha^{\star}\right)\right) = 0
$$

几乎必然。稍稍滥用记号，我们用 $\mathbb{P}_{\cdot|\cdot}$ 表示一个一般的条件分布，其精确设定可能有所变化，但可从上下文中明确。利用随机缺失假设以及抽样设计的非信息性，我们得到

$$
\mathbb{P}_m\left(C_{S_{r,v}}\left(\widehat{\alpha}_{S_{r,v}}\right) < C_{S_{r,v}}\left(\alpha^{\star}\right)\right) = \int_{\mathbb{R}^{n_{r,v}}}\mathbb{1}\left(C_{S_{r,v}}\left(\widehat{\alpha}_{S_{r,v}}\right) < C_{S_{r,v}}\left(\alpha^{\star}\right)\right)\prod_{k\in S_{r,v}}\mathbb{P}_{y_k|\mathbf{x}_k}\left(\mathrm{d}y_k\right)
$$

$$\xrightarrow{v\to\infty} 0,$$

其中该极限由（21）和引理13得出。由于

$$\mathbb{P}_{mpq}\left(C_{S_{r,v}}\left(\widehat{\alpha}_{S_{r,v}}\right) < C_{S_{r,v}}\left(\alpha^{\star}\right)\right) = \mathbb{E}_{pq}\left[\mathbb{P}_{m}\left(C_{S_{r,v}}\left(\widehat{\alpha}_{S_{r,v}}\right) < C_{S_{r,v}}\left(\alpha^{\star}\right)\right)\right],$$

应用勒贝格控制收敛定理即可得出结果。

定理 *2* 的证明。

记

$$\sqrt{n_v}\left(\widehat{\mu}_{\widehat{\alpha},v} - \mu_v\right) = \sqrt{n_v}\left(\widehat{\mu}_{\alpha^{\star},v} - \mu_v\right) + \sqrt{n_v}\left(\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}\right).$$

我们将证明第二项依概率趋于零。注意到

$$\begin{aligned}\left\{\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon\right\} = &\left(\left\{\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon\right\} \cap \{\widehat{\alpha} = \alpha^{\star}\}\right) \\ &\biguplus \left(\left\{\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon\right\} \cap \{\widehat{\alpha} \neq \alpha^{\star}\}\right).\end{aligned}$$

此外，在事件 $\{\widehat{\alpha} = \alpha\}$ 上，我们有 $\widehat{\mu}_{\widehat{\alpha},\ v} = \widehat{\mu}_{\alpha^{\star},\ v}$ 因此

$$\left\{\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon\right\} \cap \{\widehat{\alpha} = \alpha^{\star}\} = \emptyset.$$

因此，

$$\begin{aligned}\mathbb{P}_{mpq}\left(\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon\right) &= \mathbb{P}_{mpq}\left(\sqrt{n_v}|\widehat{\mu}_{\widehat{\alpha},v} - \widehat{\mu}_{\alpha^{\star},v}| > \epsilon | \ \widehat{\alpha} \neq \alpha^{\star}\right)\mathbb{P}_{mpq}\left(\widehat{\alpha} \neq \alpha^{\star}\right) \\ &\leq \mathbb{P}_{mpq}\left(\widehat{\alpha} \neq \alpha^{\star}\right)\end{aligned}$$

根据假设，收敛到0 。

定理*3*的证明

我们将 Kim 和 Rao (2009) 的证明改用于无权情形。令

$$\widehat{U}\left(\boldsymbol{\beta}_{\alpha}\right) = \sum_{k\in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\left(y_k - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}\right)}{N_v}. \tag{22}$$

注意，$\boldsymbol{\beta}_{\alpha,\ v}$ 由（4）给出，是 $U(\boldsymbol{\beta}_{\alpha}) = \mathbf{0}_{\alpha}$的解。

令

$$\widehat{\mu}_v(\boldsymbol{\beta}_{\alpha}) = \frac{1}{N_v}\left(\sum_{k\in S_{r,v}} \frac{y_k}{\pi_k} + \sum_{k\in S_{m,v}} \frac{\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}}{\pi_k}\right),$$

它是 $\boldsymbol{\beta}_{\alpha}$的函数。我们进一步定义

$$\widetilde{\mu}_v\left(\boldsymbol{\gamma}, \boldsymbol{\beta}_{\alpha}\right) = \widehat{\mu}_v\left(\boldsymbol{\beta}_{\alpha}\right) + \boldsymbol{\gamma}^{\top}\widehat{U}\left(\boldsymbol{\beta}_{\alpha}\right),$$

它被视为 $\boldsymbol{\beta}_{\alpha}$ 和 $\boldsymbol{\gamma}$ 的函数。我们的目标是找到 $\boldsymbol{\gamma}$ 的一个特定取值，称为 $\boldsymbol{\gamma}^{\star}$，使得 $\widetilde{\mu}_v(\boldsymbol{\gamma}^{\star}, \widehat{\boldsymbol{\beta}}_{\alpha,v}) - \widetilde{\mu}_v(\boldsymbol{\gamma}^{\star}, \boldsymbol{\beta}_{\alpha,v}) = o_{\mathbb{P}}(n_v^{-1/2})$。若如此，通过选择 $\boldsymbol{\gamma} = \boldsymbol{\gamma}^{\star}$，即可忽略估计 $\boldsymbol{\beta}_{\alpha}$ 的影响 。注意到对于每个 p_{α} 维向量 $\boldsymbol{\gamma}$，$\widetilde{\mu}_v(\boldsymbol{\gamma}, \boldsymbol{\beta}_{\alpha,v}) = \widehat{\mu}_{\alpha,v}$，我们有 $\widehat{\mu}_{\alpha,v} - \widetilde{\mu}_v(\boldsymbol{\gamma}^{\star}, \boldsymbol{\beta}_{\alpha,v}) = o_{\mathbb{P}}(n_v^{-1/2})$。为求出 $\boldsymbol{\gamma}^{\star}$，我们使用 Randles（1982）的理论，

已证明 $\sqrt{N_v}(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_\alpha)=O_{\mathbb{P}}(1)$引理 3，则当满足以下条件时，$\widetilde{\mu}_v(\boldsymbol{\gamma}^\star,\widehat{\boldsymbol{\beta}}_{\alpha,v})-\widetilde{\mu}_v(\boldsymbol{\gamma}^\star,\boldsymbol{\beta}_\alpha)=o_{\mathbb{P}}(n_v^{-1/2})$成立

$$\mathbb{E}_{mp}\left[\frac{\partial\widetilde{\mu}_v(\boldsymbol{\gamma},\boldsymbol{\beta}_\alpha)}{\partial\boldsymbol{\beta}_\alpha}\bigg|_{\boldsymbol{\beta}_\alpha=\boldsymbol{\beta}_\alpha}\right]=\mathbb{E}_{mp}\left[\frac{\partial\widehat{\mu}_v(\boldsymbol{\beta}_\alpha)}{\partial\boldsymbol{\beta}_\alpha}\bigg|_{\boldsymbol{\beta}_\alpha=\boldsymbol{\beta}_\alpha}\right]-\boldsymbol{\gamma}^\top\mathbb{E}_{mp}\left[\frac{\partial\widehat{U}(\boldsymbol{\beta}_\alpha)}{\partial\boldsymbol{\beta}_\alpha}\bigg|_{\boldsymbol{\beta}_\alpha=\boldsymbol{\beta}_\alpha}\right]=0.$$

$\boldsymbol{\gamma}^\star$ 的解为

$$\begin{aligned}\boldsymbol{\gamma}^\star&=-\left(\mathbb{E}_{mp}\left[\frac{\partial\widehat{U}(\boldsymbol{\beta}_\alpha)^\top}{\partial\boldsymbol{\beta}_\alpha}\bigg|_{\boldsymbol{\beta}_\alpha=\boldsymbol{\beta}_\alpha}\right]\right)^{-1}\mathbb{E}_{mp}\left[\frac{\partial\widehat{\mu}_v(\boldsymbol{\beta}_\alpha)}{\partial\boldsymbol{\beta}_\alpha}\bigg|_{\boldsymbol{\beta}_\alpha=\boldsymbol{\beta}_\alpha}\right]\\&=\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v}=\mathbf{c}_{\alpha,v}.\end{aligned}$$

接下来我们将证明 $\widetilde{\mu}_v(\boldsymbol{\gamma}^\star,\boldsymbol{\beta}_\alpha)=\widetilde{\mu}_{\alpha,v}$。将 $\widetilde{\mu}_v(\boldsymbol{\gamma},\boldsymbol{\beta}_\alpha)$写作

$$\begin{aligned}\widetilde{\mu}_v(\boldsymbol{\gamma},\boldsymbol{\beta}_\alpha)&=\widehat{\mu}_v(\boldsymbol{\beta}_\alpha)+\boldsymbol{\gamma}^\top\widehat{U}(\boldsymbol{\beta}_\alpha)\\&=\frac{1}{N_v}\left(\sum_{k\in S_{r,v}}\frac{y_k}{\pi_k}+\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}+\boldsymbol{\gamma}^\top\sum_{k\in S_{r,v}}\frac{\pi_k\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)\mathbf{x}_{k,\alpha}}{\pi_k}\right)\\&=\frac{1}{N_v}\left(\sum_{k\in S_{r,v}}\frac{y_k+\pi_k\boldsymbol{\gamma}^\top\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)\mathbf{x}_{k,\alpha}}{\pi_k}+\sum_{k\in S_{m,v}}\frac{\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}\right)\\&=\frac{1}{N_v}\sum_{k\in S_v}\frac{r_ky_k+r_k\pi_k\boldsymbol{\gamma}^\top\mathbf{x}_{k,\alpha}\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)+(1-r_k)\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}\\&=\frac{1}{N_v}\sum_{k\in S_v}\frac{r_k\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)+r_k\pi_k\boldsymbol{\gamma}^\top\mathbf{x}_{k,\alpha}\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)+\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}\\&=\frac{1}{N_v}\sum_{k\in S_v}\frac{r_k\left(1+\pi_k\boldsymbol{\gamma}^\top\mathbf{x}_{k,\alpha}\right)\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)+\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}.\end{aligned}$$

回顾一下，

$$\widetilde{\mu}_{\alpha,v}=\frac{1}{N_v}\sum_{k\in S_v}\frac{r_k\left(1+\pi_k\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)\left(y_k-\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)+\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha}{\pi_k}.$$

取 $\boldsymbol{\gamma}^\star=\mathbf{c}_{\alpha,v}$，得到 $\widetilde{\mu}_v(\mathbf{c}_{\alpha,v},\boldsymbol{\beta}_\alpha)=\widetilde{\mu}_{\alpha,v}$，最终得出 $\widehat{\mu}_{\alpha,v}-\widetilde{\mu}_{\alpha,v}=o_{\mathbb{P}}(n_v^{-1/2})$。

定理*4*的证明。

写作

$$n_v\left|\widehat{V}_{T,v}(\alpha)-V_{T,v}(\alpha)\right|\le n_v\left|\widehat{V}_{1,v}(\alpha)-V_{1,v}(\alpha)\right|+n_v\left|\widehat{V}_{2,v}(\alpha)-V_{2,v}(\alpha)\right|:=(a)+(b).$$

$\widehat{}$ **（a）的一致性**。我们进一步将 $n_v|V_1(\alpha)-V_1(\alpha)|$分解为

$$\begin{aligned} n_v\left|\widehat{V}_{1,v}(\alpha)-V_{1,v}(\alpha)\right| &\le n_v\left|\widehat{V}_{1,v}(\alpha)-\widetilde{V}_{1,v}(\alpha)\right|+n_v\left|\widetilde{V}_{1,v}(\alpha)-\bar{V}_{1,v}(\alpha)\right| \\ &\quad +n_v\left|\bar{V}_{1,v}(\alpha)-V_{1,v}(\alpha)\right| \\ &:= A_v+B_v+C_v, \end{aligned}$$

其中

$$\widetilde{V}_{1,v}(\alpha)=\frac{1}{N_v^2}\sum_{k\in S_v}\sum_{l\in S_v}\frac{\eta_{k,\alpha}}{\pi_k}\frac{\eta_{l,\alpha}}{\pi_l}\frac{\Delta_{kl}}{\pi_{kl}},$$

$$\bar{V}_{1,v}(\alpha)=\mathbb{V}_p\left(\tilde{\mu}_{\alpha,v}-\mu_v\right)=\frac{1}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\frac{\eta_{k,\alpha}}{\pi_k}\frac{\eta_{l,\alpha}}{\pi_l}\Delta_{kl}, \tag{23}$$

A_v**的一致性**。令

$$c_{kl}=\frac{\Delta_{kl}I_kI_l}{\pi_k\pi_l\pi_{kl}},\qquad k,l\in U_v.$$

利用该等式，对于任意的 $k,\ l\in U_v$,

$$\widehat{\eta}_{k,\alpha}\widehat{\eta}_{l,\alpha}-\eta_{k,\alpha}\eta_{l,\alpha}=\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)\left(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha}\right)+\eta_{k,\alpha}\left(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha}\right)+\eta_{l,\alpha}\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right),$$

A_v 可分解为

$$\begin{aligned} A_v &\le \frac{n_v}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\left|c_{kl}\left(\widehat{\eta}_{k,\alpha}\widehat{\eta}_{l,\alpha}-\eta_{k,\alpha}\eta_{l,\alpha}\right)\right| \\ &\le \frac{n_v}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\left|c_{kl}\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)\left(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha}\right)\right| \\ &\quad +\frac{n_v}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\left|c_{kl}\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)\eta_{l,\alpha}\right|+\frac{n_v}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\left|c_{kl}\eta_{k,\alpha}\left(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha}\right)\right| \\ &:= A_{1,v}+A_{2,v}+A_{3,v}. \end{aligned}$$

接下来，我们将分别处理 $A_{1、v}$ 和 $A_{2、v}$以及 $A_{3、v}$。对于 $A_{1,\ v}$，由于引理8 表明

$$\sum_{k\in U_v}\frac{\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)^2}{N_v}=o_{\mathbb{P}}(1),$$

我们有

$$\begin{aligned} A_{1,v} &\le \frac{n_v}{N_v\lambda^2}\sum_{k\in U_v}\frac{\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)^2}{N_v} \\ &\quad +\frac{n_v}{N_v\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ k\neq l}}\frac{|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha})|}{N_v} \\ &\le \frac{n_v}{N_v\lambda^2}\sum_{k\in U_v}\frac{(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})^2}{N_v} \end{aligned}$$

$$+\frac{n_v}{N_v\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{l\in U_v}\frac{|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})(\widehat{\eta}_{l,\alpha}-\eta_{l,\alpha})|}{N_v}$$
$$\leq\left(\frac{n_v}{N_v\lambda^2}+\frac{n_v}{\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\right)\sum_{k\in U_v}\frac{(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})^2}{N_v}=o_{\mathbb{P}}(1).$$

根据对称性，对于 $A_{2,\ v}$，利用 $\sum_{k\in U_v}\sum_{l\in U_v}a_kb_l\leq N_v\left(\sum_{k\in U_v}a_k^2\right)^{1/2}\left(\sum_{l\in U_v}b_l\right)^{1/2}$，我们有

$$A_{2,v}\leq\frac{n_v}{N_v\lambda^2}\sum_{k\in U_v}\frac{|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})\eta_{k,\alpha}|}{N_v}+\frac{n_v}{N_v^2\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{\substack{l\in U_v\\k\neq l}}|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})\eta_{l,\alpha}|$$
$$\leq\frac{n_v}{N_v\lambda^2}\sum_{k\in U_v}\frac{|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})\eta_{k,\alpha}|}{N_v}+\frac{n_v}{N_v^2\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{l\in U_v}|(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})\eta_{l,\alpha}|$$
$$\leq\left(\frac{n_v}{N_v\lambda^2}+\frac{n_v}{\lambda^2\lambda^\star}\max_{k\neq l\in U_v}|\Delta_{kl}|\right)\sqrt{\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}}\sqrt{\sum_{k\in U_v}\frac{(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha})^2}{N_v}}=o_{\mathbb{P}}(1).$$

因此，我们有 $\widehat{A}_v=o_{\mathbb{P}}(1)$。此外，我们有 $n_v|V_{1,\ v}(\alpha)-V_{1,\ v}(\alpha)|=o_{\mathbb{P}}(1)$。

B_v的一致性。写出

$$\mathbb{E}_p\left[\left(n_v(\widetilde{V}_{1,v}(\alpha)-\bar{V}_{1,v}(\alpha))\right)^2\right]$$
$$\leq\frac{2n_v^2}{N_v^4}\sum_{k\in U_v}\sum_{l\in U_v}\frac{1-\pi_k}{\pi_k}\frac{1-\pi_l}{\pi_l}\frac{\eta_{k,\alpha}^2}{\pi_k}\frac{\eta_{l,\alpha}^2}{\pi_l}\Delta_{kl}$$
$$+\frac{2n_v^2}{N_v^4}\sum_{i\in U_v}\sum_{\substack{j\in U_v\\j\neq i}}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\l\neq k}}\frac{\eta_{i,\alpha}\eta_{j,\alpha}\eta_{k,\alpha}\eta_{l,\alpha}}{\pi_i\pi_j\pi_k\pi_l}\mathbb{E}_p\left[\frac{(I_iI_j-\pi_{ij})(I_kI_l-\pi_{kl})}{\pi_{ij}\pi_{kl}}\right]\Delta_{ij}\Delta_{kl}$$
$$:=B_{1,v}+B_{2,v}.$$

我们将采用Breidt 和 Opsomer （2000）中定理 3 的相同论证。对于 $B_{1,\ v}$，我们有

$$B_{1,v}\leq\left(\frac{2}{N_v\lambda^3}+\frac{2n_v\max_{k\neq l\in U_v}|\Delta_{kl}|}{N_v\lambda^4}\right)\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}.$$

回顾引理20 给出

$$\limsup_{v\to\infty}\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}<\infty,$$

几乎必然。因此，$B_{1,\ v}$ 几乎必然收敛于 0。对于 $B_{2,\ v}$，我们有

$$B_{2,v}\leq L_v+\frac{2\left(n_v\max_{k\neq l\in U_v}|\Delta_{kl}|\right)^2}{\lambda^4\lambda^{\star2}}$$
$$\times\max_{(i,j,k,l)\in D_{4,N_v}}\left|\mathbb{E}_p\left[\frac{(I_iI_j-\pi_{ij})(I_kI_l-\pi_{kl})}{\pi_{ij}\pi_{kl}}\right]\right|\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}.$$

这里， L_v 几乎必然收敛于 0。因此， $B_{2,\ v}$ 几乎必然收敛于 0。由此，对于任意 $\epsilon > 0$, 应用切比雪夫不等式可得

$$\lim_{v\to\infty} \mathbb{P}_p\left(\left|n_v(\widetilde{V}_{1,v}(\alpha) - \bar{V}_{1,v}(\alpha))\right| > \epsilon\right) = 0,$$

几乎必然。因此可得

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\left|n_v(\widetilde{V}_{1,v}(\alpha) - \bar{V}_{1,v}(\alpha))\right| > \epsilon\right) = 0,$$

几乎必然。由此得证 $B_v = o_{\mathbb{P}}(1)$。

一致性 C_v。 C_v 可分解为

$$C_v = n_v\left|\bar{V}_{1,v}(\alpha) - V_{1,v}(\alpha)\right| \le \left|n_v\bar{V}_{1,v}(\alpha) - c^{\star}_{1,\alpha}\right| + \left|c^{\star}_{1,\alpha} - n_vV_{1,v}(\alpha)\right|,$$

其中 $c^{\star}_{1,\ \alpha} = f^{\star}c_{1,\ \alpha}$ 且 $f_{\star} = \lim_{v\to\infty} n_v/N_v$。一方面，我们有

$$\left|n_v\bar{V}_{1,v}(\alpha) - c^{\star}_{1,\alpha}\right| = o_{\mathbb{P}}(1),$$

根据假设。剩下要证明的是

$$\lim_{v\to\infty}\left|c^{\star}_{1,\alpha} - n_vV_{1,v}(\alpha)\right| = \lim_{v\to\infty}\left|\mathbb{E}_{mq}[n_v\bar{V}_{1,v}(\alpha) - c^{\star}_{1,\alpha}]\right| = 0$$

几乎必然。我们的目标是应用维塔利收敛定理，即证明 $|n_v\bar{V}_{1、\ v}(\alpha) - c^{\star}_{1、\ \alpha}| = o_{\mathbb{P}}(1)$，并证明 $n_v\bar{V}_{1、\ v}(\alpha)$ 具有一致可积性。由此可得 $\lim_{v\to\infty}\mathbb{E}_{mq}[|n_v\bar{V}_{1,v}(\alpha) - c^{\star}_{1、\ \alpha}|] = 0$ 几乎必然。为了证明 $n_v\bar{V}_{1、\ v}(\alpha)$ 的一致可积性，只需证明 $\mathbb{E}_{mq}[(n_v\bar{V}_{1、\ v}(\alpha))^2]$ 几乎必然有界，这由引理 23 得出。

（b）的一致性。 项 $n_v|\hat{V}_{2、\ v}(\alpha) - V_{2、\ v}(\alpha)|$ 可以分解为 s

$$n_v\left|\hat{V}_{2,v}(\alpha) - V_{2,v}(\alpha)\right| \le n_v\left|\hat{V}_{2,v}(\alpha) - \bar{V}_{2,v}(\alpha)\right| + n_v\left|\bar{V}_{2,v}(\alpha) - V_{2,v}(\alpha)\right| := D_v + E_v,$$

其中

$$\bar{V}_{2,v}(\alpha) = \mathbb{V}_m\left(\mathbb{E}_p\left[\tilde{\mu}_{\alpha,v} - \mu_v\right]\right) = \sigma^2\sum_{k\in U_v}\frac{1 - r_k + r_k\left(\pi_k\mathbf{c}^{\top}_{\alpha,v}\mathbf{x}_{k,\alpha}\right)^2}{N_v^2} \tag{24}$$

的一致性。 结合引理 $|\hat{\sigma}^2_{\alpha、\ v} - \sigma^2| = o_{\mathbb{P}}(1)$ 中的结果 5 以及

$$\left|\sum_{k\in S_v}\frac{1 - r_k + r_k\left(\pi_k\widehat{\mathbf{c}}^{\top}_{\alpha,v}\mathbf{x}_{k,\alpha}\right)^2}{\pi_kN_v} - \sum_{k\in U_v}\frac{1 - r_k + r_k\left(\pi_k\mathbf{c}^{\top}_{\alpha,v}\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right| = o_{\mathbb{P}}(1),$$

由（40），可得

$$\left|\sum_{k\in U_v}\frac{1 - r_k + r_k\left(\pi_k\mathbf{c}^{\top}_{\alpha,v}\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right| \le 1 + \|\mathbf{c}_{\alpha,v}\|_2^2\sum_{k\in U_v}\frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v} = O_{\mathbb{P}}(1).$$

由此可得

$$n_v\left|\widehat{V}_{2,v}(\alpha) - \bar{V}_{2,v}(\alpha)\right| \le \frac{n_v}{N_v}\left|\left(\widehat{\sigma}^2_{\alpha,v} - \sigma^2\right)\left(\sum_{k\in S_v}\frac{1 - r_k + r_k(\pi_k\widehat{\mathbf{c}}^{\top}_{\alpha,v}\mathbf{x}_{k,\alpha})^2}{\pi_kN_v}\right.\right.$$

$$
\begin{aligned}
&- \sum_{k\in U_v} \frac{1-r_k+r_k(\pi_k \mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha})^2}{N_v}\Bigg)\Bigg| \\
&+ \frac{n_v}{N_v}\sigma^2\left|\sum_{k\in S_v}\frac{1-r_k+r_k\left(\pi_k \widehat{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{\pi_k N_v}\right. \\
&\left.- \sum_{k\in U_v}\frac{1-r_k+r_k\left(\pi_k \mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right| \\
&+ \frac{n_v}{N_v}\left|\sum_{k\in U_v}\frac{1-r_k+r_k\left(\pi_k \mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right|\left|\widehat{\sigma}_{\alpha,v}^2-\sigma^2\right| = o_{\mathbb{P}}(1).
\end{aligned}
$$

一致性 E_v。术语 E_v 可以分解为

$$
E_v = n_v\left|\bar{V}_{2,v}(\alpha)-V_{2,v}(\alpha)\right| \le \left|n_v\bar{V}_{2,v}(\alpha)-c_{2,\alpha}^{\star}\right| + \left|c_{2,\alpha}^{\star}-n_v V_{2,v}(\alpha)\right|,
$$

其中 $c_{2,\ \alpha}^{\star} = f^{\star}c_{2,\ \alpha}$ 且 $f_{\star}=$ 极限$_{v\to\infty}\, n_v/N_v$。一方面，我们有

$$
\left|n_v\bar{V}_{2,v}(\alpha)-c_{2,\alpha}^{\star}\right| = o_{\mathbb{P}}(1),
$$

根据假设。剩余部分需要证明

$$
\lim_{v\to\infty}\left|c_{2,\alpha}^{\star}-n_v V_{2,v}(\alpha)\right| = \lim_{v\to\infty}\left|\mathbb{E}_q\left[n_v\bar{V}_{2,v}(\alpha)-c_{2,\alpha}^{\star}\right]\right| = 0
$$

几乎必然成立。我们再次使用 Vitali 收敛定理。由引理 23 可知，， $_v(\alpha)$ 具有一致可积性。

定理5的证明。

记

$$
\frac{\widehat{V}_{T,v}(\widehat{\alpha})}{V_{T,v}(\alpha^{\star})} = 1 + \frac{n_v\left(\widehat{V}_{T,v}(\widehat{\alpha})-\widehat{V}_{T,v}(\alpha^{\star})\right)}{n_v V_{T,v}(\alpha^{\star})} + \frac{n_v\left(\widehat{V}_{T,v}(\alpha^{\star})-V_{T,v}(\alpha^{\star})\right)}{n_v V_{T,v}(\alpha^{\star})} := 1 + A_v + B_v. \tag{25}
$$

因此，只需证明 $A_v = o_{\mathbb{P}}(1)$和 $B_v = o_{\mathbb{P}}(1)$。注意，由假设可知，满足 $n_v \lim_{v\to\infty} V_{T,v}(\alpha^{\star}) > 0$，因此只需证明 A_v和B_v 的分子依概率收敛到0。对于 A_v，令 $\epsilon>0$，并写作

$$
\begin{aligned}
\mathbb{P}_{mpq}\left(n_v\left|\widehat{V}_{T,v}(\widehat{\alpha})-\widehat{V}_{T,v}(\alpha^{\star})\right|>\epsilon\right) &= \mathbb{P}_{mpq}\left(n_v\left|\widehat{V}_{T,v}(\widehat{\alpha})-\widehat{V}_{T,v}(\alpha^{\star})\right|>\epsilon,\ \widehat{\alpha}=\alpha^{\star}\right) \\
&+ \mathbb{P}_{mpq}\left(n_v\left|\widehat{V}_{T,v}(\widehat{\alpha})-\widehat{V}_{T,v}(\alpha^{\star})\right|>\epsilon,\ \widehat{\alpha}\neq\alpha^{\star}\right) \\
&\le \mathbb{P}_{mpq}\left(\widehat{\alpha}\neq\alpha^{\star}\right) \xrightarrow[v\to\infty]{} 0
\end{aligned}
$$

由于模型选择准则是一致的，几乎必然成立。

对于 B_v，鉴于 $\alpha^{\star}\in C$，由定理 $B_v = o_{\mathbb{P}}(1)$ 4可直接得出。

定理*6*的证明。

为证明 $\widehat{\mu}_{\alpha^\star}$ 的渐近正态性，$_v - \mu_v$，我们验证Chen 和 Rao (2007) 的定理 2 的条件。我们将其分解为

$$\widehat{\mu}_{\alpha^\star,v} - \mu_v = \underbrace{\frac{1}{N_v}\sum_{k\in U_v} v_k \eta_{k,\alpha^\star}}_{:=U_v} + \underbrace{\frac{1}{N_v}\sum_{k\in U_v}\left(\eta_{k,\alpha} - y_k\right)}_{:=V_v}, \tag{26}$$

其中 $v_k := I_k/\pi_k - 1$ 对于 $k \in U_v$。定义 $\mathcal{B}_v := \sigma\left(\left(\mathbf{x}_k,\ r_k,\ y_k\right)_{k\in U_v}\right)$。我们需要验证Chen 和 Rao（2007）定理 2 的三个条件，分别记为（i）、（ii）和（iii）。

（i）的验证。我们希望证明 V_v 渐近服从正态分布且是$\mathcal{B}_v$-可测的。通过注意到对于任意 $k \in U_v$，可立即得到可测性，

$$\eta_{k,\alpha} - y_k = \epsilon_k(r_k + r_k\pi_k\mathbf{c}_{\alpha^\star,v}^\top\mathbf{x}_{k,\alpha^\star} - 1) := \epsilon_k w_k$$

其中$\mathbf{c}_{\alpha^\star,\ v}$ 仅取决于协变量和响应指示变量。此外，注意到$\mathbb{E}_m[\eta_{k,\alpha} - y_k] = 0$。我们将首先通过条件 Lyapunov 条件建立条件中心极限定理。更确切地说，我们需要证明存在 $\delta > 0$ 使得

$$L_v := \frac{1}{\left[\sum_{k\in U_v}\mathbb{V}_m\left(\frac{\eta_k - y_k}{N_v}\right)\right]^{1+\delta/2}}\sum_{k\in U_v}\mathbb{E}_m\left[\left(\frac{\eta_k - y_k}{N_v}\right)^{2+\delta}\right] \xrightarrow[v\to\infty]{\mathbb{P}} 0.$$

Give在我们的假设下，证明 $\delta = 2.$ 更为方便。我们首先注意到

$$\sum_{k\in U_v}\mathbb{V}_m\left(\frac{\eta_k - y_k}{N_v}\right) = \frac{\sigma^2}{N_v^2}\sum_{k\in U_v}\left((1 - r_k) + r_k(\pi_k\mathbf{c}_{\alpha^\star,v}^\top\mathbf{x}_{k,\alpha^\star})^2\right) \geq \frac{\sigma^2}{N_v^2}N_{m,v} \asymp_{\mathbb{P}} \frac{1}{N_v}.$$

此外，注意到 $w_k^4 = 1$ 若 $r_k = 0$ 且 $w_k^4 = (\pi_k\mathbf{c}_{\alpha^\star,v}^\top\mathbf{x}_{k,\alpha^\star})^4$ 否则。因此，利用$(a + b)^4 \leq 8(a^4 + b^4)$、柯西-施瓦茨不等式以及 (S3)，可得

$$\sum_{k\in U_v}\mathbb{E}_m\left[\left(\frac{\eta_k - y_k}{N_v}\right)^4\right] \leqslant \frac{8M_0}{N_v^4}\sum_{k\in U_v}\left(1+\|\mathbf{c}_{\alpha^\star,v}\|_2^4\|\mathbf{x}_{k,\alpha^\star}\|_2^4\right) = O_{\mathbb{P}}\left(\frac{1}{N_v^3}\right),$$

利用 (40) 和 (S2) 可得。这表明 $L_v = O_{\mathbb{P}}\left(N_v^{-1}\right)$，因此 Lyapunov 条件成立。因此，(i) 与

$$\sigma_{1v}^2 := \mathbb{V}_m\left(\sum_{k\in U_v}\frac{\eta_{k,\alpha^\star} - y_k}{N_v}\right)$$

成立。由此得到条件高斯性。此外，由于渐近分布不依赖于条件，因此可以使用控制收敛论证将其推广为无条件中心极限定理。

（ii）的验证。我们有

$$\mathbb{E}\left[U_v|\mathcal{B}_v\right] = \frac{1}{N_v}\sum_{k\in U_v}\mathbb{E}\left[\nu_k|\mathcal{B}_v\right]\eta_{k,\alpha^\star} = 0,$$

$$\mathbb{V}\left(U_v|\mathcal{B}_v\right) = \frac{1}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\Delta_{kl}\frac{\eta_{k,\alpha^\star}}{\pi_k}\frac{\eta_{l,\alpha^\star}}{\pi_l} := \sigma_{2v}^2,$$

因为对于所有 $k \in U_v, \eta_k$，都是 $\mathcal{B}_v$-可测的。根据假设，$\sigma_{2v}^{-1}U_v|\mathcal{B}_v \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$，这也可以等价地表述为

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\sigma_{2v}^{-1}U_v \leqslant t|\mathcal{B}_n\right) - \Phi(t)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

其中 Φ 表示标准正态随机变量的累积分布函数。假设(ii)得到验证。

验证(iii)。我们有

$$N_v\mathbb{V}_m\left(\sum_{k\in U_v}\frac{\eta_{k,\alpha^\star}-y_k}{N_v}\right) = \frac{\sigma^2}{N_v}\sum_{k\in U_v}(1-r_k) + \sigma^2\mathbf{c}_{\alpha^\star,v}^\top\frac{1}{N_v}\sum_{k\in U_v}r_k\pi_k^2\mathbf{x}_{k,\alpha^\star}\mathbf{x}_{k,\alpha^\star}^\top\mathbf{c}_{\alpha^\star,v}$$

$$\xrightarrow[v\to\infty]{\mathbb{P}} \sigma^2\left\{\mathbb{E}\left[1-p(\mathbf{x}_{\alpha^\star})\right] + \bar{\mathbf{c}}_{\alpha^\star}{}^\top\boldsymbol{C}_2\bar{\mathbf{c}}_{\alpha^\star}\right\}$$

其中

$$\bar{\mathbf{c}}_{\alpha^\star} = \boldsymbol{C}_1^{-1}\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha^\star}\right].$$

因此，我们有

$$\frac{\sigma_{1v}^2}{\sigma_{2v}^2} \xrightarrow[v\to\infty]{\mathbb{P}} \frac{\sigma^2\left\{\mathbb{E}\left[1-p(\mathbf{x}_{\alpha^\star})\right] + \bar{\mathbf{c}}_{\alpha^\star}{}^\top\boldsymbol{C}_2\bar{\mathbf{c}}_{\alpha^\star}\right\}}{c_3}.$$

因此，(iii) 得证。

综合以上各部分。由于 Chen 和 Rao (2007) 的三个条件均已满足，因此 $\widehat{\mu}_{\alpha^\star,v} - \mu_v$ 渐近正态。此外，根据定理 3，$\widehat{\mu}_{\widehat{\alpha},v} - \mu_v$ 的渐近分布与 $\widehat{\mu}_{\alpha^\star,v} - \mu_v$相同，因此也渐近正态。应用定理5 可证明其方差为单位方差。

B.2. 估计量的一致性

B.2.1. 单个估计量的一致性

引理 3. 设 $\alpha \in C$与 $\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)_{v\in\mathbb{N}}$为最小二乘估计量序列，由 (4)给出。假设 *(S1)-(S3)* 与 *(D1)-(D3)*。则，

$$\left\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

证明。设

$$\widetilde{\boldsymbol{\beta}}_{\alpha,v}=\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_k y_k}{N_v}.$$

我们将 $\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}$ 分解为

$$\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}=\underbrace{\widehat{\boldsymbol{\beta}}_{\alpha,v}-\widetilde{\boldsymbol{\beta}}_{\alpha,v}}_{(a)}+\underbrace{\widetilde{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}}_{(b)}.$$

我们需要证明，(a) 和 (b) 两项均以速率 $O_{\mathbb{P}}(n_v^{-1/2})$ 收敛至 0。

(a) 项的收敛。由三角不等式可得

$$\begin{aligned}\left\|\widehat{\boldsymbol{\beta}}_{\alpha,v}-\widetilde{\boldsymbol{\beta}}_{\alpha,v}\right\|_2=&\left\|\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v}\right. \\ &\left.-\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right\|_2 \qquad (27)\\ =&\left\|\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}\left(\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right)\right. \\ &\left.+\left(\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}-\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\right)\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right\|_2 \\ \le&\left\|\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}\left(\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right)\right\|_2 \\ &+\left\|\left(\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}-\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\right)\right. \qquad (28)\\ &\left.\times\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right\|_2 \\ :=&A_{1,v}+A_{2,v}\end{aligned}$$

接下来，我们分别考察 $A_{1,v}$ 和 $A_{2,v}$。

对 $A_{1,v}$ 的处理。对于 $A_{1,v}$，利用施瓦茨矩阵不等式，可得

$$A_{1,v}\le\left\|\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}\right\|_{op}\left\|\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right\|_2$$

我们将应用引理9 来得到

$$\left\|\sum_{k\in S_{r,v}}\frac{\mathbf{x}_{k,\alpha}y_k}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}y_k}{N_v}\right\|_2=O_{\mathbb{P}}\left(\frac{1}{\sqrt{N_v}}\right). \qquad (29)$$

令 $w_k := r_k, \boldsymbol{T}_k := \mathbf{x}_{k,\alpha} y_k$，由于 $\{r_k\}_{k\in U_v}$ 是独立同分布随机变量，且 $\mathbb{E}_q[r_k^2] = p_k \le 1$ 几乎必然，因此仍需验证

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\mathbb{E}_m\left[\|\mathbf{x}_{k,\alpha} y_k\|_2^2\right]}{N_v} < \infty$$

几乎必然。回顾(S2) 可知，协变量具有有界支撑集，由此得到

$$\sum_{k\in U_v} \frac{\mathbb{E}_m\left[\|\mathbf{x}_{k,\alpha} y_k\|_2^2\right]}{N_v} = \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2 \mathbb{E}_m\left[y_k^2\right]}{N_v} = \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2 \mathbb{E}_m\left[\left(\mathbf{x}_{k,\alpha}^\top \boldsymbol{\beta}_\alpha + \epsilon_k\right)^2\right]}{N_v}$$

$$\le C_0^2 \sum_{k\in U_v} \frac{2\sigma^2 + 2\left(\mathbf{x}_{k,\alpha}^\top \boldsymbol{\beta}_\alpha\right)^2}{N_v} \le 2C_0^2\sigma^2 + 2C_0^4 \|\boldsymbol{\beta}_\alpha\|_2^2 < \infty.$$

由此可得

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha} y_k\|_2^2}{N_v} < \infty,$$

几乎必然，由此可得（29）。另一方面，引理18表明

$$\left\| \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1).$$

因此，我们得到 $A_{1,v} = \mathrm{O}_{\mathbb{P}}(n_v^{-1/2})$。

对 $A_{2,v}$ 的处理。项 $A_{2,v}$ 可以分解为

$$A_{2,v} \le \left\| \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1} \right\|_{op} \left\| \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha} y_k}{N_v} \right\|_2.$$

另一方面，利用引理9 以及 $w_k := r_k$ 和 $\boldsymbol{T}_k := \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top$。由于

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\mathbb{E}_m\left[\|\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top\|_F^2\right]}{N_v} = \limsup_{v\to\infty} \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v} \le C_0^4 \tag{30}$$

几乎必然，因此可得

$$\left\| \frac{\boldsymbol{A}_{r,\alpha}}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right). \tag{31}$$

另一方面，回顾引理15表明

$$\left\| \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1).$$

引理的一个应用 14 给出

$$\left\| \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

与此同时，

$$\left\| \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha} y_k}{N_v} \right\|_2 \le \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2 |y_k|}{N_v} \le C_0 \sum_{k\in U_v} \frac{|y_k|}{N_v},$$

这意味着

$$\limsup_{v\to\infty} \left\| \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha} y_k}{N_v} \right\|_2 < \infty.$$

由此可得 $A_{2,\ v} = O_{\mathbb{P}}(n_v^{-1/2})$ 因此，

$$\left\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \widetilde{\boldsymbol{\beta}}_{\alpha,v}\right\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

(b) 的收敛性。根据引理16，我们有

$$\limsup_{v\to\infty} \left\| \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-2} \right\|_{op} < \infty$$

几乎必　然。应用切比雪夫不等式可知，对于任意 $\epsilon > 0$，我们有

$$\begin{aligned}
&\mathbb{P}_{mpq}\left(\left\|\widetilde{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right\|_2^2 > \epsilon\right) \qquad (32)\\
&\le \frac{1}{\epsilon^2}\mathbb{E}_{mpq}\left[\left(\widetilde{\boldsymbol{\beta}}_\alpha - \boldsymbol{\beta}_\alpha\right)^\top \left(\widetilde{\boldsymbol{\beta}}_\alpha - \boldsymbol{\beta}_\alpha\right)\right]\\
&= \frac{1}{\epsilon^2}\mathbb{E}_m\left[\sum_{k\in U_v} \frac{\epsilon_k \mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \left(\sum_{k\in U_v} \frac{\pi_k p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-2} \sum_{k\in U_v} \frac{\epsilon_k \mathbf{x}_{k,\alpha} p(\mathbf{x}_k)\pi_k}{N_v}\right]\\
&\overset{(*)}{=} \frac{\sigma^2}{\epsilon^2} \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \left(\sum_{k\in U_v} \frac{\pi_k p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-2} \frac{\mathbf{x}_{k,\alpha} p(\mathbf{x}_k)\pi_k}{N_v}\\
&\overset{(**)}{\le} \frac{\sigma^2}{N_v\epsilon^2} \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v} \left\| \left(\sum_{k\in U_v} \frac{\pi_k p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-2} \right\|_{op}\\
&\le \frac{\sigma^2 C_0^2}{N_v\epsilon^2} \left\| \left(\sum_{k\in U_v} \frac{\pi_k p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-2} \right\|_{op}. \qquad (33)
\end{aligned}$$

这里，**(*)** 是由于以下事实： $\mathbb{E}_m[\epsilon_k\epsilon_l] = 0$ 对于 $k \ne l \in U_v$ 而 **(**)** 则来自不等式$\mathbf{z}^\top \boldsymbol{A}\mathbf{z} \le \|\mathbf{z}\|_2^2\|\boldsymbol{A}\|_{op}$。式（32）的最后一行几乎必然收敛到 0。综合以上各点，我们得到

$$\left\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

□

引理 4. 假设 *(S1)-(S2)* 且 *(D1)-(D3)*。若 (C_1) 成立，则存在常数 K_0使得

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\lambda_{min}(N_v^{-1}\boldsymbol{S}_v) \ge K_0\right) = 1$$

几乎必然。

证明。 **步骤 1**。证明

$$\|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op} = O_{\mathbb{P}}(n_v^{-1/2}).$$

写作

$$\begin{aligned}
&\|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op} \\
&\quad= \Bigg\| \frac{1}{N_v}\mathbf{A}_{r,\alpha^c} \\
&\qquad - \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \left(\frac{\mathbf{A}_{r,\alpha}}{N_v}\right)^{-1} \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} - \Bigg(\sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} \\
&\qquad\quad - \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} \left(\sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \right)^{-1} \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} \Bigg) \Bigg\|_{op} \\
&\quad\le \left\| \frac{1}{N_v}\mathbf{A}_{r,\alpha^c} - \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} \right\|_{op} \\
&\qquad + \Bigg\| \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \left(\frac{\mathbf{A}_{r,\alpha}}{N_v}\right)^{-1} \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \\
&\qquad\quad - \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \left(\sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \right)^{-1} \sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top p(\mathbf{x}_k)\pi_k}{N_v} \Bigg\|_{op} \\
&\quad:= A_v + B_v。
\end{aligned}$$

回想一下，(31) 表明 $A_v = O_{\mathbb{P}}(n_v^{-1/2})$。对于 B_v，我们将其分解为

$$\begin{aligned}
B_v \le &\Bigg\| \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \Bigg(\left(\frac{\mathbf{A}_{r,\alpha}}{N_v}\right)^{-1} \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \\
&\quad - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \Bigg) \Bigg\|_{op} \\
&+ \Bigg\| \left(\sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) \\
&\quad \times \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \Bigg\|_{op} \\
:= &\, B_{1,v} + B_{2,v}.
\end{aligned}$$

我们需要证明 $B_{1,v}$ 和 $B_{2,v}$ 收敛到 0，且为 $O_{\mathbb{P}}(n_v^{-1/2})$。

收敛于 $B_{1,\ v}$。我们有

$$
\begin{aligned}
B_{1,v} \leq & \left\| \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right\|_{op} \\
& \times \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right. \\
& \left. - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_2 .
\end{aligned}
$$

使用 (S2)，我们得到

$$
\left\| \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right\|_{op} \leq \sum_{k \in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2 \|\mathbf{x}_{k,\alpha^c}\|_2}{N_v} \leq C_0^2,
$$

几乎必然。接下来，利用施瓦茨矩阵不等式，我们分解

$$
\begin{aligned}
& \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_{op} \\
& \leq \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} \left(\sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right) \right\|_{op} \\
& \quad + \left\| \left(\left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right) \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_{op} \\
& \leq \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_{op} \\
& \quad + \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_{op} . \quad (34)
\end{aligned}
$$

接下来，我们评估（34）中的每个分量。引理15和引理18 表明

$$
\left\| \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1), \qquad \left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1),
$$

分别地。我们继续应用引理9，其中 $w_k := r_k$ 和 $\boldsymbol{T}_k = \mathbf{x}_{k,\ \alpha} \mathbf{x}_{k,\ \alpha^c}^{\top}$。由于

$$
\limsup_{v \to \infty} \sum_{k \in U_v} \frac{\mathbb{E}_m \left[\|\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}\|_F^2 \right]}{N_v} = \limsup_{v \to \infty} \sum_{k \in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2 \|\mathbf{x}_{k,\alpha^c}\|_2^2}{N_v} \leq C_0^4
$$

几乎必然，由此可得

$$
\left\| \sum_{k \in S_{r,v}} \frac{\mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha^c}^{\top}}{N_v} \right\|_{op} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right).
$$

再次由（31）可知

$$\left\| \frac{\mathbf{A}_{r,\alpha}}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

因此，利用引理14 可得

$$\left\| \left(\frac{\mathbf{A}_{r,\alpha}}{N_v}\right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

最后，注意到

$$\left\| \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \right\|_{op} \le \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \|\mathbf{x}_{k,\alpha}\|_2 \|\mathbf{x}_{k,\alpha^c}\|}{N_v} \le C_0^2$$

几乎必然，因此我们得出 (34) 为 $\mathrm{O}_{\mathbb{P}}(n_v^{-1/2})$。

对 $B_{2、v}$ 的处理。 利用 施瓦茨矩阵不等式，我们有

$$\begin{aligned} B_{2,v} = & \left\| \left(\sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) \right. \\ & \left. \times \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \right\|_{op} \\ \le & \left\| \sum_{k\in S_{r,v}} \frac{\mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha^c}\mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} \\ & \times \left\| \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha^c}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right). \end{aligned}$$

综合以上各项，我们得到

$$\|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

步骤 2。一方面，步骤 1 对任意 $\epsilon > 0$ 给出，

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op} \le \epsilon \right) = 1$$

几乎必然。另一方面，根据韦尔不等式，

$$\left|\lambda_{min}(N_v^{-1}\mathbf{S}_v) - \lambda_{min}(\mathbf{S}_{U_v})\right| \le \|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op},$$

由此可得

$$\lim_{v\to\infty} \mathbb{P}_{pq}\left(\left|\lambda_{min}(N_v^{-1}\mathbf{S}_v) - \lambda_{min}(\mathbf{S}_{U_v})\right| \le \epsilon \right) = 1.$$

因此，令 $K_0 = C_0 - \epsilon$，可得

$$\begin{aligned}
&\lim_{v\to\infty} \mathbb{P}_{pq}\left(\left|\lambda_{min}(N_v^{-1}\mathbf{S}_v) - \lambda_{min}(\mathbf{S}_{U_v})\right| \le \epsilon\right) \\
&= \lim_{v\to\infty} \mathbb{P}_{pq}\left(-\epsilon \le \lambda_{min}(N_v^{-1}\mathbf{S}_v) - \lambda_{min}(\mathbf{S}_{U_v}) \le \epsilon\right) \\
&\le \lim_{v\to\infty} \mathbb{P}_{pq}\left(\lambda_{min}(\mathbf{S}_{U_v}) - \epsilon \le \lambda_{min}(N_v^{-1}\mathbf{S}_v)\right) \\
&\le \lim_{v\to\infty} \mathbb{P}_{pq}\left(C_0 - \epsilon \le \lambda_{min}(N_v^{-1}\mathbf{S}_v)\right) \\
&= \lim_{v\to\infty} \mathbb{P}_{pq}\left(K_0 \le \lambda_{min}(N_v^{-1}\mathbf{S}_v)\right) = 1
\end{aligned}$$

几乎必然。此外，利用勒贝格收敛定理，我们得到 n

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\|N_v^{-1}\mathbf{S}_v - \mathbf{S}_{U_v}\|_{op} \le \epsilon\right) = 1$$

几乎必然。

□

引理 5. 令 $\alpha \in C$ 和 $(\widehat{\sigma}^2_{\alpha,\ v})_{v\in\mathbb{N}}$ 为在(10)中定义的估计量序列。假设*(S1)-(S3)*和 *(D1)-(D3)*。我们有

$$|\widehat{\sigma}^2_{\alpha,v} - \sigma^2| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

证明。 令

$$\widehat{S}^2_v = \frac{1}{n_{r,v}} \sum_{k\in S_{r,v}} \left(y_k - \mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha\right)^2.$$

那么，$|\sigma^2_{\alpha,\ v} - \sigma^2|$ 可以分解为

$$\left|\widehat{\sigma}^2_{\alpha,v} - \sigma^2\right| \le \frac{n_{r,v}}{n_{r,v} - p_\alpha}\left|\widehat{S}^2_v - \sigma^2\right| + \frac{p_\alpha}{n_{r,v} - p_\alpha}\sigma^2.$$

由于 $\sigma^2 p_\alpha/(n_{r,v} - p_\alpha) = o_{\mathbb{P}}(n_v^{-1/2})$，剩下的部分是证明 $|\widehat{S}^2_v - \sigma| = O_{\mathbb{P}}(n_v^{-1/2})$。注意到 $|\widehat{S}^2_v - \sigma^2|$ 还可以进一步分解为

$$|\widehat{S}^2_v - \sigma^2| \le \underbrace{|\widehat{S}^2_v - \widehat{S}^2_r|}_{(a)} + \underbrace{|\widehat{S}^2_r - \widetilde{S}^2_v|}_{(b)} + \underbrace{|\widetilde{S}^2_v - \sigma^2|}_{(c)},$$

其中

$$\widehat{S}^2_r = \sum_{k\in S_v} \frac{r_k \epsilon_k^2}{N_v}\left(\frac{n_{r,v}}{N_v}\right)^{-1}$$

and

$$\widetilde{S}^2_v = \sum_{k\in S_v} \frac{p(\mathbf{x}_k)\pi_k \epsilon_k^2}{N_v}\left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-1}.$$

我们需要证明项 (a)、(b) 和 (c) 以速率 $O_{\mathbb{P}}(n_v^{-1/2})$ 收敛到 0。

(a) 的收敛。 $|\widehat{S}_v^2 - \widehat{S}_r^2|$ 可以表示为

$$\begin{aligned}\left|\widehat{S}_v^2 - \widehat{S}_r^2\right| &= \left|\frac{N_v}{n_{r,v}}\sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2}{N_v} - \frac{N_v}{n_{r,v}}\sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2}{N_v}\right| \\ &= \left|\left(\frac{n_{r,v}}{N_v}\right)^{-1}\right|\underbrace{\left|\sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2}{N_v} - \sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2}{N_v}\right|}_{:=A_v}.\end{aligned}$$

由引理 13可知

$$\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\frac{n_{r,v}}{n_v}\geq\xi\right) = 1$$

我们推得 $(n_{r,\ v}/N_v)^{-1} = O_{\mathbb{P}}(1)$。 利用

$$\left(\mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2 - \left(\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2 = \left(\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 + 2\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right),$$

我们将 A_v 分解为

$$\begin{aligned}&\left|\sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2}{N_v} - \sum_{k\in S_v}\frac{r_k\left(y_k - \mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2}{N_v}\right| \\ &\leq \sum_{k\in U_v}\frac{\left|\left(y_k - \mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2 - \left(y_k - \mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2\right|}{N_v} \\ &= \sum_{k\in U_v}\frac{\left|-2\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right) + \left(\mathbf{x}_{k,\alpha}^\top\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)^2 - \left(\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\right)^2\right|}{N_v} \\ &= \sum_{k\in U_v}\frac{\left|-2\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right) + \left(\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 + 2\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right|}{N_v} \\ &\leq \frac{2}{N_v}\sum_{k\in U_v}|y_k|\left|\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right| \\ &\quad + \frac{1}{N_v}\sum_{k\in U_v}\left(\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 + \frac{2}{N_v}\sum_{k\in U_v}\left|\mathbf{x}_{k,\alpha}^\top\boldsymbol{\beta}_\alpha\mathbf{x}_{k,\alpha}^\top\left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right| \\ &:= A_{1,v} + A_{2,v} + A_{3,v}.\end{aligned}$$

根据大数定律，我们有

$$\sum_{k\in U_v}\frac{y_k^2}{N_v}\xrightarrow[v\to\infty]{a.s.}\mathbb{E}_{\mathbf{x}m}[y_1^2] = \mathbb{E}_{\mathbf{x}m}\left[\left(\mathbf{x}_{1,\alpha}^\top\boldsymbol{\beta}_\alpha + \epsilon_1\right)^2\right]$$

$$\leq \mathbb{E}_{\mathbf{x}m}\left[2\left(\mathbf{x}_{1,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}\right)+\epsilon_1^2\right]=2\sigma^2+2C_0^2\|\boldsymbol{\beta}_{\alpha}\|_2^2.$$

对于 $A_{1,\ v}$，利用柯西-施瓦茨不等式，我们有

$$\begin{aligned}A_{1,v}&\leq 2\sqrt{\sum_{k\in U_v}\frac{y_k^2}{N_v}}\sqrt{\sum_{k\in U_v}\frac{\left(\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right)\right)^2}{N_v}}\\&\leq 2\|\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\|_2\sqrt{\sum_{k\in U_v}\frac{y_k^2}{N_v}}\sqrt{\sum_{k\in U_v}\frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v}}\\&\leq 2C_0\|\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\|_2\sqrt{\sum_{k\in U_v}\frac{y_k^2}{N_v}}=O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right),\end{aligned}$$

这也意味着 $A_{2,\ v}=O_{\mathbb{P}}(n_v^{-1/2})$。对于 $A_{3,\ v}$，利用柯西-施瓦茨不等式，可得

$$\begin{aligned}A_{3,v}&\leq 2\sqrt{\sum_{k\in U_v}\frac{\left(\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}\right)^2}{N_v}}\sqrt{\sum_{k\in U_v}\frac{\left(\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right)\right)^2}{N_v}}\\&\leq 2C_0^2\|\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\|_2\|\boldsymbol{\beta}_{\alpha}\|_2=O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).\end{aligned}$$

最后，我们得出 $|\widehat{S}_v-\widehat{S}_r|=O_{\mathbb{P}}(n_v^{-1/2})$。

对项 (b) 的处理。项 (b) 可分解为

$$\begin{aligned}\left|\widehat{S}_r^2-\widetilde{S}_v^2\right|&\leq\left|\left(\frac{n_{r,v}}{N_v}\right)^{-1}\left(\sum_{k\in S_v}\frac{r_k\epsilon_k^2}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\epsilon_k^2}{N_v}\right)\right|\\&\quad+\left|\left(\left(\frac{n_{r,v}}{N_v}\right)^{-1}-\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-1}\right)\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\epsilon_k^2}{N_v}\right|\\&:=B_{1,v}+B_{2,v}.\end{aligned}$$

对于 $B_{1,\ v}$，我们应用引理 9，其中 $\boldsymbol{T}_k:=\epsilon_k^2$ 以及 $w_k:=r_k$。由于

$$\limsup_{v\to\infty}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\epsilon_k^4\right]}{N_v}\leq M_0$$

几乎必然，因此可得

$$\left|\sum_{k\in S_v}\frac{r_k\epsilon_k^2}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\epsilon_k^2}{N_v}\right|=O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

根据引理 13 可知 $\left(n_{r,v}/N_v\right)^{-1}=O_{\mathbb{P}}(1)$，因此 $B_{1,v}=O_{\mathbb{P}}(n_v^{-1/2})$。对于 $B_{2,v}$，我们有

$$\left|\left(\frac{n_{r,v}}{N_v}\right)^{-1}-\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-1}\right|=\left|\left(\frac{n_{r,v}}{N_v}\right)^{-1}\right|\left|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-1}\right|\left|\frac{n_{r,v}}{N_v}-\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right|.$$

应用引理 9，令 $\boldsymbol{T}_k := 1$ 且 $w_k := r_k$，对于 $k \in U_v$，我们有

$$\left| \frac{n_{r,v}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k}{N_v} \right| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

根据(D2) 和正性，

$$\left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k}{N_v} \right)^{-1} \leq (\rho\lambda)^{-1}$$

因此 $B_{2,v} = O_{\mathbb{P}}(n_v^{-1/2})$ 且 $|\widehat{S}_r^2 - \widetilde{S}_v^2| = O_{\mathbb{P}}(n_v^{-1/2})$。

项 (c) 的处理。根据切比雪夫不等式

$$\begin{aligned}
\mathbb{P}_{mpq}\left(\left|\widetilde{S}_v^2 - \sigma^2\right| > \epsilon\right) &\leq \frac{1}{\epsilon^2}\mathbb{E}_{mpq}\left[\left(\widetilde{S}_v^2 - \sigma^2\right)^2\right] \\
&\leq \frac{1}{N_v}\sum_{k \in U_v} \frac{p(\mathbf{x}_k)^2\pi_k^2\mathbb{E}_m[\epsilon_k^4]}{N_v}\left(\sum_{k \in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-2} \\
&\leq \frac{1}{N_v}\sum_{k \in U_v}\frac{M_0}{N_v}\left(\sum_{k \in U_v}\frac{p(\mathbf{x}_k)\pi_k}{N_v}\right)^{-2} \\
&\leq \frac{M_0}{N_v\rho^2\lambda^2}. \qquad (35)
\end{aligned}$$

结合项 (a)、(b) 和 (c)，我们得出 $|\widehat{S}_v^2 - \sigma^2| = O_{\mathbb{P}}(n_v^{-1/2})$，因此 $|\widehat{\sigma}_{\alpha,v}^2 - \sigma^2| = O_{\mathbb{P}}(n_v^{-1/2})$。

□

引理 6。设 $\alpha \in C$ 和 $(\widehat{\boldsymbol{c}}_{\alpha,v})_{v \in \mathbb{N}}$ 是由 (8) 给出的一列估计量。假设*(S1)-(S3)*和*(D1)-(D3)*。则，

$$\|\widehat{\boldsymbol{c}}_{\alpha,v} - \boldsymbol{c}_{\alpha,v}\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

证明。设

$$\widetilde{\mathbf{c}}_{\alpha,v} = \left(\sum_{k \in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k \in U_v}\frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v}. \qquad (36)$$

记

$$\|\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2 = \underbrace{\|\widehat{\mathbf{c}}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2}_{(a)} + \underbrace{\|\widetilde{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2}_{(b)}.$$

（a）的处理。我们将其分解

$$\begin{aligned}
&\|\widehat{\mathbf{c}}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2 \\
&= \left\| \left(\sum_{k \in S_v}\frac{r_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k \in S_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}}{\pi_k N_v}\right.
\end{aligned}$$

$$
\begin{aligned}
&- \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \Bigg\|_2 \\
&= \Bigg\| \left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \left(\sum_{k \in S_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{\pi_k N_v} - \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right) \\
&\quad + \left(\left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right) \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \Bigg\|_2 \\
&\leq \Bigg\| \left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \left(\sum_{k \in S_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{\pi_k N_v} - \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right) \Bigg\|_2 \\
&\quad + \Bigg\| \left(\left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right) \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \Bigg\|_2 \\
&:= A_{1,v} + A_{2,v}.
\end{aligned}
$$

对于 $A_{1,\ v}$，我们有

$$
A_{1,v} \leq \left\| \left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k \in S_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{\pi_k N_v} - \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 .
$$

根据引理17，

$$
\left\| \left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1),
$$

此外，应用引理9，令 $\boldsymbol{T}_k :=\pi_k^{-1}\mathbf{x}_{k,\ \alpha}, w_k :=1-r_k$，由于

$$
\limsup_{v \to \infty} \sum_{k \in U_v} \frac{\mathbb{E}_m \left[\| \pi_k^{-1} \mathbf{x}_{k,\alpha} \|_F^2 \right]}{N_v} \leq \frac{C_0^2}{\lambda^2}
$$

几乎必然，因此可得

$$
\left\| \sum_{k \in U_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{\pi_k N_v} - \sum_{k \in U_v} \frac{(1-p_k)\mathbf{x}_{k,\alpha}}{N_v} \right\|_2^2 = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right).
$$

因此，$A_{1,\ v} = o_{\mathbb{P}}$（1）。对于$A_{2,\ v}$，我们得到

$$
\begin{aligned}
A_{2,v} \leq & \left\| \left(\sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} - \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right)^{-1} \right\|_{op} \\
& \times \left\| \sum_{k \in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 .
\end{aligned}
$$

期望式（31）给出

$$
\left\| \sum_{k \in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v} \right\|_{op} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right).
$$

引理 15 和引理18 指出

$$\left\| \left(\sum_{k\in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)\pi_k}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1), \qquad \left\| \left(\sum_{k\in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1)$$

分别。利用引理14，我们得到

$$\left\| \left(\sum_{k\in S_v} \frac{r_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}} \right).$$

最后，注意到

$$\left\| \sum_{k\in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 \le \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2}{N_v} \le C_0$$

几乎必然，我们得出 $A_{2,\ v} = \mathrm{O}_{\mathbb{P}}(n_v^{-1/2})$。

对 (b) 的处理。(b) 项可由以下式子界定

$$\begin{aligned}
&\|\widetilde{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2 \\
&\le \left\| \left(\sum_{k\in U_v} \frac{r_k\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k\in U_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v} - \sum_{k\in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 \\
&\quad + \left\| \left(\sum_{k\in U_v} \frac{r_k\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} \\
&\quad \times \left\| \sum_{k\in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 \\
&:= B_{1,v} + B_{2,v}.
\end{aligned}$$

对于 $B_{1,v}$，应用引理 10，令 $w_k :=1-r_k, \boldsymbol{T}_k := \mathbf{x}_{k,\alpha}$ 并回忆

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\mathbb{E}_m\left[\|\mathbf{x}_{k,\alpha}\|_2^2\right]}{N_v} \le C_0^2 < \infty,$$

我们有

$$\left\| \sum_{k\in U_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v} - \sum_{k\in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}} \right).$$

此外，利用引理17，我们得到$B_{1,\ v} = O_{\mathbb{P}}(n_v^{-1/2})$。对于$B_{2,\ v}$，应用引理10，令$w_k := r_k$，$\boldsymbol{T}_k := \pi_k \mathbf{x}_{k,\ \alpha}\mathbf{x}_{k,\ \alpha}^\top$，（30）意味着

$$\left\| \sum_{k\in U_v} \frac{r_k\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} - \sum_{k\in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}} \right). \tag{37}$$

应用引理14 可得

$$\left\| \left(\sum_{k\in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} - \left(\sum_{k\in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}} \right).$$

因此，$B_{2,\,v} = O_{\mathbb{P}}(n_v^{-1/2})$。综合以上结果，我们得到 $A_v = O_{\mathbb{P}}(n_v^{-1/2})$和 $B_v = O_{\mathbb{P}}(n_v^{-1/2})$。由此得出 $\|\widehat{\mathbf{c}}_{\alpha,\,v} - \mathbf{c}_{\alpha,\,v}\|_2 = O_{\mathbb{P}}(n_v^{-1/2})$。□

B.2.2. 插件估计量的一致性

引理7。 设 $\alpha \in C$ 和 $(\widehat{\mathbf{c}}_{\alpha,\,v})_{v\in\mathbb{N}}$ 是由(8)定义的一个序列。假设 *(S1)-(S3)* 以及*(D1)-(D3)*。我们有

$$\left| \sum_{k\in S_v} \frac{1 - r_k + r_k \left(\pi_k \widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} - \sum_{k\in U_v} \frac{1 - r_k + r_k \left(\pi_k \mathbf{c}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v} \right| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}} \right).$$

证明。 考虑以下分解,

$$\begin{aligned}
&\left| \sum_{k\in S_v} \frac{1 - r_k + r_k \left(\pi_k \widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} - \sum_{k\in U_v} \frac{1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v} \right| \\
&\quad \le \left| \sum_{k\in S_v} \frac{1 - r_k + r_k \left(\pi_k \widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} - \sum_{k\in S_v} \frac{1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} \right| \\
&\quad + \left| \sum_{k\in S_v} \frac{1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} - \sum_{k\in U_v} \frac{1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v} \right| \\
&\quad + \left| \sum_{k\in U_v} \frac{1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v} - \sum_{k\in U_v} \frac{1 - r_k + r_k \left(\pi_k \mathbf{c}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v} \right| \\
&\quad := A_v + B_v + C_v,
\end{aligned}$$

其中，$\widetilde{\mathbf{c}}_{\alpha,\,v}$ 由（36）给出。接下来，我们将证明 A_v、B_v 和 C_v 以 $O_{\mathbb{P}}(n_v^{-1/2})$的速率收敛到 0。

对 A_v 的处理。使用

$$\left(\widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2 - \left(\widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2 = \left(\left(\widehat{\mathbf{c}}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha}\right)^2 + 2 \left(\widehat{\mathbf{c}}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha} \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha},$$

我们得到

$$\begin{aligned}
A_v &= \left| \sum_{k\in S_v} r_k \frac{\left(\pi_k \widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2 - \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2}{N_v \pi_k} \right| \\
&\le \sum_{k\in U_v} \frac{\left| \left(\widehat{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2 - \left(\widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\right)^2 \right|}{N_v \lambda} \\
&\le \sum_{k\in U_v} \frac{\left(\left(\widehat{\mathbf{c}}_{\alpha,v}^\top - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha}\right)^2 + 2 \left| \widetilde{\mathbf{c}}_{\alpha,v}^\top \mathbf{x}_{k,\alpha} \left(\widehat{\mathbf{c}}_{\alpha,v}^\top - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha} \right|}{N_v \lambda}
\end{aligned}$$

$$:= A_{1,v} + A_{2,v}. \tag{38}$$

对于$A_{1,v}$，已证明 $\|\widehat{\mathbf{c}}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2 = O_{\mathbb{P}}(n_v^{-1/2})$ 引理6中，我们有

$$\begin{aligned} A_{1,v} &\le \frac{1}{\lambda} \sum_{k \in U_v} \frac{\left(\left(\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha}\right)^2}{N_v} \\ &\le \frac{1}{\lambda} \|\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2^2 \sum_{k \in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v} \\ &\le \frac{C_0^2}{\lambda} \|\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2^2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right). \end{aligned}$$

对于$A_{2,v}$根据引理19，我们有

$$\limsup_{v \to \infty} \|\widetilde{\mathbf{c}}_{\alpha,v}\|_2 < \infty$$

几乎必然。因此

$$\begin{aligned} A_{2,v} &\le \frac{2}{\lambda} \sqrt{\sum_{k \in U_v} \frac{\left(\widetilde{\mathbf{c}}_{\alpha,v}^{\top} \mathbf{x}_{k,\alpha}\right)^2}{N_v}} \sqrt{\sum_{k \in U_v} \frac{\left(\left(\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha}\right)^2}{N_v}} \\ &\le \frac{2}{\lambda} \|\widetilde{\mathbf{c}}_{\alpha,v}\|_2 \|\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2 \sum_{k \in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v} \\ &\le \frac{2C_0^2}{\lambda} \|\widetilde{\mathbf{c}}_{\alpha,v}\|_2 \|\widehat{\mathbf{c}}_{\alpha,v}^{\top} - \widetilde{\mathbf{c}}_{\alpha,v}\|_2 = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right). \end{aligned}$$

因此，我们有 $A_v = \mathrm{O}_{\mathbb{P}}(n_v^{-1/2})$。

的处理 B_v。令 $a_k = 1 - r_k + r_k \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v} \mathbf{x}_{k,\alpha}\right)^2$ 对于 $k \in U_v$。应用引理 9，其中$\boldsymbol{T}_k$：$= a_k$ 且 w_k：$= 1$ 对于 $k \in U_v$，接下来只需证明

$$\limsup_{v \to \infty} \sum_{k \in U_v} \frac{\mathbb{E}_m\left[a_k^2\right]}{N_v} < \infty$$

几乎必然。为此，记

$$\begin{aligned} \limsup_{v \to \infty} \sum_{k \in U_v} \frac{\mathbb{E}_m\left[a_k^2\right]}{N_v} &\le \limsup_{v \to \infty} \frac{1}{N_v} \sum_{k \in U_v} \left(2(1 - r_k)^2 + 2r_k^2 \left(\pi_k \widetilde{\mathbf{c}}_{\alpha,v}^{\top} \mathbf{x}_{k,\alpha}\right)^2\right) \\ &\le \limsup_{v \to \infty} \frac{1}{N_v} \sum_{k \in U_v} \left(2 + 2\|\widetilde{\mathbf{c}}_{\alpha,v}\|_2^2 \|\mathbf{x}_{k,\alpha}\|_2^2\right) \\ &\le \limsup_{v \to \infty} 2\left(1 + C_0^2 \|\widetilde{\mathbf{c}}_{\alpha,v}\|_2^2\right) < \infty \end{aligned}$$

几乎必然。这就完成了 $B_v = \mathrm{O}_{\mathbb{P}}(n_v^{-1/2})$。

对 C_v的处理。请注意

$$\left(\mathbf{c}_{\alpha,v}^{\top} \mathbf{x}_{k,\alpha}\right)^2 - \left(\widetilde{\mathbf{c}}_{\alpha,v}^{\top} \mathbf{x}_{k,\alpha}\right)^2 = \left(\left(\mathbf{c}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha}\right)^2 + 2\left(\mathbf{c}_{\alpha,v} - \widetilde{\mathbf{c}}_{\alpha,v}\right) \mathbf{x}_{k,\alpha} \widetilde{\mathbf{c}}_{\alpha,v}^{\top} \mathbf{x}_{k,\alpha}.$$

利用引理 $\|\widetilde{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\|_2=O_{\mathbb{P}}(n_v^{-1/2})$中的结果，并6遵循式（38）中的相同分解，我们得到

$$\begin{aligned}
&\left|\frac{1}{N_v}\sum_{k\in U_v}1-r_k+r_k\left(\pi_k\widetilde{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2-\frac{1}{N_v}\sum_{k\in U_v}1-r_k+r_k\left(\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2\right| \\
&\quad\le\frac{1}{N_v}\sum_{k\in U_v}\left|\left(\widetilde{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2-\left(\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2\right| \\
&\quad\le\frac{1}{N_v}\sum_{k\in U_v}\left(\left(\widetilde{\mathbf{c}}_{\alpha,v}^{\top}-\mathbf{c}_{\alpha,v}\right)\mathbf{x}_{k,\alpha}\right)^2+2\left|\widetilde{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\left(\widetilde{\mathbf{c}}_{\alpha,v}^{\top}-\mathbf{c}_{\alpha,v}\right)\mathbf{x}_{k,\alpha}\right| \\
&\quad\le\|\widetilde{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\|_2\frac{1}{N_v}\sum_{k\in U_v}\|\mathbf{x}_{k,\alpha}\|_2^2\left(2\|\widetilde{\mathbf{c}}_{\alpha,v}\|_2+\|\widetilde{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\|_2\right) \\
&\quad\le C_0^2\|\widetilde{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\|_2\left(2\|\widetilde{\mathbf{c}}_{\alpha,v}\|_2+\|\widetilde{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\|_2\right)=O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right). \qquad (39)
\end{aligned}$$

总体

$$\left|\sum_{k\in S_v}\frac{1-r_k+r_k\left(\pi_k\widehat{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{N_v\pi_k}-\sum_{k\in U_v}\frac{1-r_k+r_k\left(\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right|=O_{\mathbb{P}}\left(\frac{1}{\sqrt{n_v}}\right).$$

□

引理8。 设 $\alpha\in\mathcal{A}$ 并假设*(D1)-(D3)*和 *(S1)-(S3)*。我们有

$$\frac{1}{N_v}\sum_{k\in U_v}\left(\widehat{\eta}_{k,\alpha}-\eta_{k,\alpha}\right)^2=o_{\mathbb{P}}(1),$$

其中 $\widehat{\eta}_{k,\ \alpha}$ 由下式给出 (9)。

证明。 对于任意 $k\in U_v$，我们分解

$$\begin{aligned}
\eta_{k,\alpha}-\widehat{\eta}_{k,\alpha}&=\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}+r_k\left(1+\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)\left(y_k-\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha,v}\right) \\
&\quad-\left(\mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha,v}+r_k\left(1+\pi_k\widehat{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)\left(y_k-\mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)\right) \\
&\quad+r_k\pi_k\widehat{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}-r_k\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta} \\
&=(1-r_k)\mathbf{x}_{k,\alpha}^{\top}\left(\boldsymbol{\beta}_{\alpha}-\widehat{\boldsymbol{\beta}}_{\alpha,v}\right)-r_k\pi_k\left(\mathbf{c}_{\alpha,v}-\widehat{\mathbf{c}}_{\alpha,v}\right)^{\top}\mathbf{x}_{k,\alpha}y_k \\
&\quad+r_k\pi_k(\widehat{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v})^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right) \\
&\quad+r_k\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right)+r_k\pi_k(\widehat{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v})^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha},
\end{aligned}$$

因为

$$\begin{aligned}
&r_k\pi_k\widehat{\mathbf{c}}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\widehat{\boldsymbol{\beta}}_{\alpha}-r_k\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta} \\
&\quad=r_k\pi_k\left(\widehat{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\right)^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right) \\
&\qquad+r_k\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\left(\widehat{\boldsymbol{\beta}}_{\alpha,v}-\boldsymbol{\beta}_{\alpha}\right)+r_k\pi_k\left(\widehat{\mathbf{c}}_{\alpha,v}-\mathbf{c}_{\alpha,v}\right)^{\top}\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_{\alpha}.
\end{aligned}$$

另一方面，根据引理17，我们得到：

$$\|\mathbf{c}_{\alpha,v}\|_2 \leq \left\| \left(\sum_{k\in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} \left\| \sum_{k\in U_v} \frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v} \right\|_2 = O_{\mathbb{P}}(1). \tag{40}$$

结合如下结果$\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_{\alpha,v}\| = o_{\mathbb{P}}(1)$在引理3中以及$\|\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2 = o_{\mathbb{P}}(1)$在引理6中，应用柯西-施瓦茨不等式可得

$$\begin{aligned}
&\sum_{k\in U_v} \frac{(\widehat{\eta}_{k,\alpha} - \eta_{k,\alpha})^2}{N_v} \\
&= \sum_{k\in U_v} \frac{(\eta_{k,\alpha} - \widehat{\eta}_{k,\alpha})^2}{N_v} \\
&= \frac{1}{N_v} \sum_{k\in U_v} \Bigg((1-r_k)\mathbf{x}_{k,\alpha}^\top \left(\boldsymbol{\beta}_\alpha - \widehat{\boldsymbol{\beta}}_{\alpha,v}\right) - r_k\pi_k \left(\mathbf{c}_{\alpha,v} - \widehat{\mathbf{c}}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha} y_k \\
&\quad + r_k\pi_k \left(\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right) \\
&\qquad + r_k\pi_k \mathbf{c}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right) + r_k\pi_k \left(\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \boldsymbol{\beta}_\alpha \Bigg)^2 \\
&\leq \frac{5}{N_v} \sum_{k\in U_v} \Bigg(\left(\mathbf{x}_{k,\alpha}^\top \left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 + \left(\left(\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha} y_k\right)^2 \\
&\quad + \left(\left(\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 + \left(\mathbf{c}_{\alpha,v}^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \left(\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\right)\right)^2 \\
&\quad + \left(\left(\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\right)^\top \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top \boldsymbol{\beta}_\alpha\right)^2 \Bigg) \\
&\leq 5\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\|_2^2 \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v} + 5\|\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2^2 \sqrt{\sum_{k\in U_v} \frac{y_k^4}{N_v}} \sqrt{\sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v}} \\
&\quad + 5\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\|_2^2 \|\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2^2 \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v} \\
&\quad + 5\|\widehat{\boldsymbol{\beta}}_{\alpha,v} - \boldsymbol{\beta}_\alpha\|_2^2 \|\mathbf{c}_{\alpha,v}\|_2^2 \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v} \\
&\quad + 5\|\boldsymbol{\beta}_\alpha\|_2^2 \|\widehat{\mathbf{c}}_{\alpha,v} - \mathbf{c}_{\alpha,v}\|_2^2 \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v} = o_{\mathbb{P}}(1).
\end{aligned}$$

最后，我们得出结论

$$\sum_{k\in U_v} \frac{(\widehat{\eta}_{k,\alpha} - \eta_{k,\alpha})^2}{N_v} = o_{\mathbb{P}}(1).$$

□

B.3. 补充引理

B.3.1. 技术引理

引理 9。 设 $(w_k)^{k\in U_v}$ 为满足 $\mathbb{E}_q[w_k^2] \le b$ 的独立同分布随机变量，其中某个 $b \ge 0$ 几乎必然成立。设 $(\widehat{\boldsymbol{Z}}_{\pi,v})_{v\in\mathbb{N}}$和 $(\boldsymbol{Z}_v)_{v\in\mathbb{N}}$为 $\mathbb{R}^{d_1\times d_2}$中的矩阵序列，其定义为

$$\widehat{\boldsymbol{Z}}_{\pi,v} = \sum_{k\in S_v} \frac{w_k \boldsymbol{T}_k}{N_v},$$

and

$$\boldsymbol{Z}_v = \sum_{k\in U_v} \frac{\mathbb{E}_q[w_k]\boldsymbol{T}_k}{N_v},$$

分别。假设 *(D1)-(D3)*。若 $\{\boldsymbol{T}_k\}^{k\in U_v}$ 满足

$$\limsup_{v\to\infty} \frac{1}{N_v} \sum_{k\in U_v} \mathbb{E}_m \left[\|\boldsymbol{T}_k\|_F^2\right] < \infty$$

几乎必然，则有

$$\left\|\widehat{\boldsymbol{Z}}_{\pi,v} - \boldsymbol{Z}_v\right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{N_v}}\right).$$

证明。 回顾一下，对于任意矩阵 $\boldsymbol{B}$，有 $\|\boldsymbol{B}\|_{op} \le \|\boldsymbol{B}\|_F$。 对于任意 $\epsilon > 0$，应用切比雪夫不等式可得

$$\begin{aligned}\mathbb{P}_{mpq}\left(\|\widehat{\boldsymbol{Z}}_{\pi,v} - \boldsymbol{Z}_v\|_{op} > \epsilon\right) &\le \frac{1}{\epsilon^2}\mathbb{E}_{mpq}\left[\|\widehat{\boldsymbol{Z}}_{\pi,v} - \boldsymbol{Z}_v\|_{op}^2\right] \\ &\le \frac{1}{\epsilon^2}\mathbb{E}_{mpq}\left[\|\widehat{\boldsymbol{Z}}_{\pi,v} - \boldsymbol{Z}_v\|_F^2\right].\end{aligned}$$

我们接下来评估

$$\begin{aligned}\mathbb{E}_{mpq}\left[\|\widehat{\boldsymbol{T}}_{\pi,v} - \boldsymbol{T}_v\|_F^2\right] &= \mathbb{E}_{mpq}\left[\left\|\sum_{k\in S_v}\frac{w_k\boldsymbol{T}_k}{N_v} - \sum_{k\in U_v}\frac{\mathbb{E}_q[w_k]\boldsymbol{T}_k}{N_v}\right\|_F^2\right] \\ &= \mathbb{E}_{mpq}\left[\mathrm{Tr}\left(\sum_{k\in U_v}\sum_{l\in U_v}\frac{(I_k w_k - \pi_k\mathbb{E}_q[w_k])(I_l w_l - \pi_l\mathbb{E}_q[w_l])\boldsymbol{T}_k^\top\boldsymbol{T}_l}{N_v^2}\right)\right] \\ &= \mathbb{E}_{mpq}\left[\sum_{k\in U_v}\sum_{l\in U_v}\frac{(I_k w_k - \pi_k\mathbb{E}_q[w_k])(I_l w_l - \pi_l\mathbb{E}_q[w_l])\mathrm{Tr}\left(\boldsymbol{T}_k^\top\boldsymbol{T}_l\right)}{N_v^2}\right].\end{aligned} \tag{41}$$

当 $k = l$，我们有 $\mathbb{E}_{pq}[(I_k w_k - \pi_k\mathbb{E}_q[w_k])(I_l w_l - \pi_l\mathbb{E}_q[w_l])] = \pi_k(1-\pi_k)\mathbb{V}_q[w_k]$。当$k \ne l \in U_v$，我们有 $\mathbb{E}_{pq}[(I_k w_k - \pi_k\mathbb{E}_q[w_k])(I_l w_l - \pi_l\mathbb{E}_q[w_l])] = \Delta_{kl}\mathbb{E}_q[w_k]\mathbb{E}_q[w_l]$。因此，（41）简化为

$$\mathbb{E}_{mpq}\left[\sum_{k\in U_v}\sum_{l\in U_v}\frac{(I_k w_k - \pi_k\mathbb{E}_q[w_k])(I_l w_l - \pi_l\mathbb{E}_q[w_l])\mathrm{Tr}\left(\boldsymbol{T}_k^\top\boldsymbol{T}_l\right)}{N_v^2}\right]$$

$$
\begin{aligned}
&= \sum_{k\in U_v} \frac{\pi_k(1-\pi_k)\mathbb{V}_q\left(w_k\right)}{N_v^2}\mathbb{E}_m\left[\mathrm{Tr}\left(\boldsymbol{T}_k^\top \boldsymbol{T}_k\right)\right] \\
&\quad + \sum_{k\in U_v}\sum_{\substack{l\in U_v \\ l\neq k}} \frac{\Delta_{kl}\mathbb{E}_q\left[w_k\right]\mathbb{E}_q\left[w_k\right]}{N_v^2}\mathbb{E}_m\left[\mathrm{Tr}\left(\boldsymbol{T}_k^\top \boldsymbol{T}_l\right)\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} + \max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{l\in U_v}\frac{|\mathbb{E}_q\left[w_k\right]\mathbb{E}_q\left[w_k\right]|}{N_v^2}\mathbb{E}_m\left[\left|\mathrm{Tr}\left(\boldsymbol{T}_k^\top \boldsymbol{T}_l\right)\right|\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} \\
&\quad + \frac{n_v \max_{k\neq l\in U_v}|\Delta_{kl}|}{n_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{|\mathbb{E}_q\left[w_k\right]\mathbb{E}_q\left[w_k\right]|}{N_v^2}\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F\|\boldsymbol{T}_l\|_F\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} + \frac{\bar{\Delta}}{n_v}\sum_{k\in U_v}\sum_{l\in U_v}\frac{|\mathbb{E}_q\left[w_k\right]\mathbb{E}_q\left[w_k\right]|}{N_v^2}\mathbb{E}_m^{1/2}\left[\|\boldsymbol{T}_k\|_F^2\right]\mathbb{E}_m^{1/2}\left[\|\boldsymbol{T}_l\|_F^2\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} + \frac{\bar{\Delta}}{n_v}\sum_{k\in U_v}\frac{\mathbb{E}_q^2\left[w_k\right]}{N_v}\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} + \frac{\bar{\Delta}}{n_v}\sum_{k\in U_v}\frac{\mathbb{E}_q\left[w_k^2\right]}{N_v}\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right] \\
&\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} + \frac{b\bar{\Delta}}{n_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v}.
\end{aligned}
$$

结论得证。 □

引理10。 设$(w_k)_{k\in U_v}$是满足 $\mathbb{E}_q[w_k^2]\leq b$ 的独立同分布随机变量，且该条件对某个 $b\geq 0$ 几乎必然成立。设$(\widehat{\boldsymbol{Z}}_{p,v})_{v\in\mathbb{N}}$和$(\boldsymbol{Z}_v)_{v\in\mathbb{N}}$是 $\mathbb{R}^{d_1\times d_2}$ 中定义的矩阵序列

$$
\widehat{\boldsymbol{Z}}_{p,v} = \sum_{k\in U_v}\frac{w_k\boldsymbol{T}_k}{N_v},
$$

and

$$
\boldsymbol{Z}_v = \sum_{k\in U_v}\frac{\mathbb{E}_q[w_k]\boldsymbol{T}_k}{N_v},
$$

分别。如果$(\boldsymbol{T}_k)_{k\in U_v}$ 满足

$$
\limsup_{v\to\infty}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v} < \infty
$$

几乎必然成立，因此我们有

$$
\left\|\widehat{\boldsymbol{Z}}_{p,v} - \boldsymbol{Z}_v\right\|_{op} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{N_v}}\right).
$$

证明。 通过计算 $\mathbb{E}_{mpq}[\|\widehat{\boldsymbol{Z}}_{p,v}-\boldsymbol{Z}_v\|_{op}^2]$，我们有

$$\begin{aligned}\mathbb{E}_{mpq}\left[\|\widehat{\boldsymbol{Z}}_{p,v}-\boldsymbol{Z}_v\|_{op}^2\right] &\leq \mathbb{E}_{mq}\left[\left\|\sum_{k\in U_v}\frac{(w_k-\mathbb{E}_m[w_k])\,\boldsymbol{T}_k}{N_v}\right\|_F^2\right]\\ &= \mathbb{E}_{mq}\left[\sum_{k\in U_v}\sum_{l\in U_v}\frac{(w_k-\mathbb{E}_m[w_k])(w_l-\mathbb{E}_m[w_l])\,\mathrm{Tr}\left(\boldsymbol{T}_k^\top\boldsymbol{T}_l\right)}{N_v^2}\right]\\ &= \sum_{k\in U_v}\frac{\mathbb{V}_q(w_k)\mathbb{E}_m\left[\mathrm{Tr}\left(\boldsymbol{T}_k^\top\boldsymbol{T}_k\right)\right]}{N_v^2}\\ &\leq \sum_{k\in U_v}\frac{\mathbb{E}_q(w_k^2)\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v^2}\\ &\leq \frac{b}{N_v}\sum_{k\in U_v}\frac{\mathbb{E}_m\left[\|\boldsymbol{T}_k\|_F^2\right]}{N_v},\end{aligned}$$

由此即可得出结论。 □

引理 11. 对于任意 $n\in\mathbb{N}$，映射 $\mathcal{L}_{2,\,n}:\mathcal{P}(\{1,2,...,\,p\})\longrightarrow\mathbb{R}_+$ 定义如下：

$$\alpha\;\longmapsto\;\sigma^2\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}^\top}{\pi_k}\right)\boldsymbol{A}_{r,\alpha}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha}}{\pi_k}\right)$$

几乎处处 是严格递增的集合函数，即 $\alpha_1\subset\alpha_2$ 蕴含 $\mathrm{L}_2(\alpha_1)\;<\mathrm{L}_2(\alpha_2$).

证明。 设 $\alpha_1,\alpha_2\in\mathcal{P}\{1,2,...,\,p\})$ 使得 $\alpha_1\subset\alpha_2$。 则，

$$\begin{aligned}&\mathcal{L}_{2,n}(\alpha_1)<\mathcal{L}_{2,n}(\alpha_2)\\ &\Leftrightarrow\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^\top}{N\pi_k}\right)\boldsymbol{A}_{r,\alpha_1}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}}{N\pi_k}\right)<\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2}^\top}{N\pi_k}\right)\boldsymbol{A}_{r,\alpha_2}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2}}{N\pi_k}\right).\end{aligned}$$

这相当于证明一个二次型几乎必然小于另一个。我们先用 $\boldsymbol{A}_{r}^{-1}{}_{,\alpha_1}$ 表示 $\boldsymbol{A}_{r}^{-1}{}_{,\alpha_2}$。令 $\boldsymbol{A}_{r,\alpha_1,\alpha_2-\alpha_1}:=\mathbf{x}_{r,\alpha_1}^\top\mathbf{x}_{r,\alpha_2-\alpha_1}$，其中 $\alpha_1\subset\alpha_2$ 任意。注意到

$$\boldsymbol{A}_{r,\alpha_2}=\begin{pmatrix}\boldsymbol{A}_{r,\alpha_1} & \boldsymbol{A}_{r,\alpha_1,\alpha_2-\alpha_1}\\ \boldsymbol{A}_{r,\alpha_2-\alpha_1,\alpha_1} & \boldsymbol{A}_{r,\alpha_2-\alpha_1}\end{pmatrix}.$$

利用分块矩阵求逆公式（例如参见Horn and Johnson （2012），第 25 页），可以证明

$$\boldsymbol{A}_{r,\alpha_2}^{-1}=\begin{pmatrix}\boldsymbol{A}_{r,\alpha_1}^{-1}+\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_1,\alpha_2-\alpha_1}\boldsymbol{S}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1,\alpha_1}\boldsymbol{A}_{r,\alpha_1}^{-1} & -\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_1,\alpha_2-\alpha_1}\boldsymbol{S}^{-1}\\ -\boldsymbol{S}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1,\alpha_1}\boldsymbol{A}_{r,\alpha_1}^{-1} & \boldsymbol{S}^{-1}\end{pmatrix}.$$

因此，写出

$$\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2}^\top}{N\pi_k}\right)=\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^\top}{N\pi_k},\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^\top}{N\pi_k}\right),$$

经过一些代数运算可得

$$\begin{aligned}
&\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}}{N\pi_k},\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^{\top}}{N\pi_k}\right)\boldsymbol{A}_{r,\alpha_2}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}}{N\pi_k},\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}}{N\pi_k}\right)\\
&=\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}}{N\pi_k}\right)\boldsymbol{A}_{r,\alpha_1}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}}{N\pi_k}\right)\\
&\quad+\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^{\top}}{N\pi_k}-\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1}}{N\pi_k}\right)\\
&\quad\times\boldsymbol{S}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^{\top}}{N\pi_k}-\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1}}{N\pi_k}\right)^{\top}.
\end{aligned}\tag{42}$$

因此，证明 $\mathcal{L}_{2,\,n}(\alpha_1)<\mathcal{L}_{2,\,n}(\alpha_2)$等价于证明

$$\begin{aligned}
&\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^{\top}}{N\pi_k}-\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1}}{N\pi_k}\right)\\
&\times\boldsymbol{S}^{-1}\left(\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_2-\alpha_1}^{\top}}{N\pi_k}-\sum_{k\in S_m}\frac{\mathbf{x}_{k,\alpha_1}^{\top}\boldsymbol{A}_{r,\alpha_1}^{-1}\boldsymbol{A}_{r,\alpha_2-\alpha_1}}{N\pi_k}\right)^{\top}>0
\end{aligned}\tag{43}$$

几乎必然。为此，注意到 $\boldsymbol{S}$ 可以写成

$$\boldsymbol{S}=\mathbf{x}_{r,\alpha_2-\alpha_1}^{\top}\left(\boldsymbol{I}_{n_r}-\mathbf{x}_{r,\alpha_1}\boldsymbol{A}_{r,\alpha_1}^{-1}\mathbf{x}_{r,\alpha_1}^{\top}\right)\mathbf{x}_{r,\alpha_2-\alpha_1}.$$

矩阵 $\boldsymbol{S}$ 是线性无关向量的Gram矩阵，因此是正定的，这说明（43），从而 $\mathcal{L}_{2,\,n}(\alpha_1)<\mathcal{L}_{2,\,n}(\alpha_2)$几乎必然成立。□

引理12。 设 $\alpha\in\mathcal{A}$，并假设 α中包含截距项。那么，

$$\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\right]=\mathbb{E}_{\mathbf{x}}\left[1-p(\mathbf{x})\right].$$

证明。 令 $\mathbf{x}_{\alpha}^{\top}=\left(1,\ \mathbf{x}_{\alpha-1}^{\top}\right)$，其中 $\mathbf{x}_{\alpha-1}^{\top}$ 是不含截距项的协变量。因此，

$$\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]=\begin{pmatrix}\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})] & \mathbb{E}_{\mathbf{x}}[p(\mathbf{x})\mathbf{x}_{\alpha-1}^{\top}]\\ \mathbb{E}_{\mathbf{x}}[p(\mathbf{x})\mathbf{x}_{\alpha-1}] & \mathbb{E}_{\mathbf{x}}[p(\mathbf{x})\mathbf{x}_{\alpha-1}\mathbf{x}_{\alpha-1}^{\top}]\end{pmatrix}.$$

利用分块矩阵求逆公式，可得

$$\begin{aligned}
&\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1}\\
&=\begin{pmatrix}\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}+\frac{1}{\mathbb{E}_{\mathbf{x}}^{2}[p(\mathbf{x})]}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}^{\top}\right]\boldsymbol{M}_{\alpha-1}^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}\right] & -\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}\boldsymbol{B}_{\alpha-1}\\ -\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}\boldsymbol{B}_{\alpha-1}^{\top} & \boldsymbol{M}_{\alpha-1}^{-1}\end{pmatrix},
\end{aligned}$$

其中 $\boldsymbol{M}_{\alpha-1}=\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}\mathbf{x}_{\alpha-1}^{\top}\right]-\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}\right](\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})])^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}^{\top}\right]$是一个正定矩阵，而 $\boldsymbol{B}_{\alpha-1}=\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}^{\top}\right]\boldsymbol{M}_{\alpha-1}^{-1}$。因此，

$$\begin{aligned}
&\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\right]\\
&=\begin{pmatrix}\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}+\frac{1}{\mathbb{E}_{\mathbf{x}}^{2}[p(\mathbf{x})]}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}^{\top}\right]\boldsymbol{M}_{\alpha-1}^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}\right] & -\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}\boldsymbol{B}_{\alpha-1}\\ -\frac{1}{\mathbb{E}_{\mathbf{x}}[p(\mathbf{x})]}\boldsymbol{B}_{\alpha-1}^{\top} & \boldsymbol{M}_{\alpha-1}^{-1}\end{pmatrix}
\end{aligned}$$

$$\times \begin{pmatrix} \mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\right] \\ \mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha-1}\right] \end{pmatrix}$$
$$= \begin{pmatrix} 1 \\ \mathbf{0}_{\alpha-1} \end{pmatrix}.$$

这就得出

$$\mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha}^{\top}\right]\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\mathbf{x}_{\alpha}^{\top}\right]\right)^{-1}\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x})\mathbf{x}_{\alpha}\right]$$
$$= \left(\mathbb{E}_{\mathbf{x}}\left[1-p(\mathbf{x})\right], \mathbb{E}_{\mathbf{x}}\left[(1-p(\mathbf{x}))\mathbf{x}_{\alpha-1}^{\top}\right]\right)\begin{pmatrix} 1 \\ \mathbf{0}_{\alpha-1} \end{pmatrix} = \mathbb{E}_{\mathbf{x}}[1-p(\mathbf{x})].$$

□

引理13. 假设 *(S1)-(S2)*。还假设存在 $\rho > 0$，使得 $p(\mathbf{x}) > \rho$ 几乎必然成立。那么，存在 $\xi > 0$，使得

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\frac{n_{r,v}}{N_v} \geq \xi\right) = 1$$

几乎必然。

证明。 注意到

$$\mathbb{E}_q\left[\frac{n_{r,v}}{N_v}\right] = \frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right), \qquad \text{and,} \qquad \mathbb{V}_q\left(\frac{n_{r,v}}{n_v}\right) = \frac{1}{n_v^2}\sum_{k\in S_v} p\left(\mathbf{x}_k\right)\left\{1-p\left(\mathbf{x}_k\right)\right\} \leq \frac{1}{4n_v}.$$

因此，切比雪夫不等式表明，对所有 $\epsilon > 0$,

$$\mathbb{P}_q\left(\left|\frac{n_{r,v}}{n_v} - \frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right)\right| > \epsilon\right) \leq \frac{1}{4\epsilon^2 n_v} \xrightarrow[v\to\infty]{} 0$$

几乎必然成立。另一方面,

$$\frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right) \geq \min_{k\in S_v} p\left(\mathbf{x}_k\right) \geq \rho$$

几乎必然。另一方面,

$$\mathbb{P}_q\left(\left|\frac{n_{r,v}}{n_v} - \frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right)\right| \leq \epsilon\right) = \mathbb{P}_q\left(\frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right) - \epsilon \leq \frac{n_{r,v}}{n_v} \leq \frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right) + \epsilon\right)$$
$$\leq \mathbb{P}_q\left(\rho - \epsilon \leq \frac{n_{r,v}}{n_v}\right).$$

利用 最小值$_{k\in U_v}\, p(\mathbf{x}_k) \geq \rho$，我们设定 $\xi := \rho - \epsilon$，由此可得

$$\lim_{v\to\infty} \mathbb{P}_q\left(\xi \leq \frac{n_{r,v}}{n_v}\right) \geq \lim_{v\to\infty} \mathbb{P}_q\left(\left|\frac{n_{r,v}}{n_v} - \frac{1}{n_v}\sum_{k\in S_v} p\left(\mathbf{x}_k\right)\right| > \epsilon\right) = 1$$

几乎必然。此外,

$$\lim_{v\to\infty} \mathbb{P}_{mpq}\left(\frac{n_{r,v}}{N_v} \geq \xi\right) = 1$$

几乎必然。 □

引理14。 设$\{\boldsymbol{A}_v\}_{v\in\mathbb{N}}$与$\{\boldsymbol{B}_v\}_{\mathbb{N}}$为$p\times p$可逆矩阵，使得$\|\boldsymbol{A}_v^{-1}\|_{op}=O_{\mathbb{P}}(1)$且$\|\boldsymbol{B}_v^{-1}\|_{op}=O_{\mathbb{P}}(1)$。若存在一个序列$\{u_v\}_{v\in\mathbb{N}}$使得

$$\|\boldsymbol{A}_v-\boldsymbol{B}_v\|_{op}=O_{\mathbb{P}}\left(u_v^{-1}\right).$$

于是，我们有

$$\|\boldsymbol{A}_v^{-1}-\boldsymbol{B}_v^{-1}\|_{op}=O_{\mathbb{P}}\left(u_v^{-1}\right).$$

证明。 回顾施瓦茨矩阵不等式（参见Hansen (2022年)，第 981 页中的 B.15），对于任意方阵 $\boldsymbol{A}$和 $\boldsymbol{B}$，关于谱范数，有 $\|\boldsymbol{AB}\|_{op}\leq\|\boldsymbol{A}\|_{op}\|\boldsymbol{B}\|_{op}$。因此，

$$\begin{aligned}\|\boldsymbol{A}_v^{-1}-\boldsymbol{B}_v^{-1}\|_{op}&=\|\boldsymbol{A}_v^{-1}\left(\mathbf{I}_p-\boldsymbol{A}_v\boldsymbol{B}_v^{-1}\right)\|_{op}\leq\|\boldsymbol{A}_v^{-1}\|_{op}\|\boldsymbol{I}_p-\boldsymbol{A}_v\boldsymbol{B}_v^{-1}\|_{op}\\&=\|\boldsymbol{A}_v^{-1}\|_{op}\|\left(\boldsymbol{B}_v-\boldsymbol{A}_v\right)\boldsymbol{B}_v^{-1}\|_{op}\\&\leq\|\boldsymbol{A}_v^{-1}\|_{op}\|\boldsymbol{B}_v^{-1}\|_{op}\|\mathbf{B}_v-\boldsymbol{A}_v\|_{op}=O_{\mathbb{P}}\left(u_v^{-1}\right).\end{aligned}$$

由于对所有 $v\in\mathbb{N}$ 均有 $\|\boldsymbol{B}_v-\boldsymbol{A}_v\|_{op}=\|\boldsymbol{A}_v-\boldsymbol{B}_v\|_{op}$，证明完毕。 □

B.3.2. 若干关注统计量的紧性

引理 15。 假设*(S1)*-*(S2)*和*(D1)*-*(D3)*成立。对于 $\alpha\in\mathcal{A}$，我们有

$$\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\right\|_{op}=O_{\mathbb{P}}(1).$$

证明。 利用 (D2)可得

$$\left\|\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right\|_{op}\geq\lambda\left\|\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right\|_{op},$$

另一方面，利用大数定律和连续映射定理可得

$$\lambda_{min}\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\xrightarrow[v\to\infty]{a.s.}\lambda_{min}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha}\right]\right),$$

这是一个正定矩阵。因此，对于任意 $\epsilon>0$，我们有

$$\begin{aligned}&\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\left|\lambda_{min}\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)-\lambda_{min}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha}\right]\right)\right|\leq\epsilon\right)\\&\leq\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\lambda_{min}\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\geq\lambda_{min}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha}\right]\right)-\epsilon\right)=1.\end{aligned}$$

此外，

$$\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\lambda_{min}\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\geq\lambda\lambda_{min}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha}\right]\right)-\epsilon\right)=1$$

几乎必然。这就结束了

$$\left\| \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = \lambda_{min}^{-1} \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) = O_{\mathbb{P}}(1).$$

□

引理 16. 假设*(S1)-(S2)*和*(D1)-(D3)*。对于 $\alpha \in \mathcal{A}$，有

$$\limsup_{v \to \infty} \left\| \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-2} \right\|_{op} < \infty.$$

证明。我们有

$$\left\| \sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} \leq \lambda \left\| \sum_{k \in U_v} \frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}^\top \mathbf{x}_{k,\alpha}}{N_v} \right\|_{op},$$

此外,

$$\lambda_{min}^{-2} \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) \xrightarrow[v \to \infty]{a.s.} \lambda_{min}^{-2} \left(\mathbb{E}_{\mathbf{x}} \left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha} \right] \right),$$

这是一个正定矩阵。因此,

$$\limsup_{v \to \infty} \lambda_{min}^{-2} \left(\sum_{k \in U_v} \frac{\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top p(\mathbf{x}_k)}{N_v} \right) < \infty$$

几乎必然。由此可得

$$\limsup_{v \to \infty} \left\| \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-2} \right\|_{op} = \limsup_{v \to \infty} \lambda_{min}^{-2} \left(\sum_{k \in U_v} \frac{p(\mathbf{x}_k)\pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) < \infty$$

几乎必然。

□

引理 17. 假设*(S1)-(S2)*和*(D1)-(D3)*。对于 $\alpha \in \mathcal{A}$，我们有

$$\left\| \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1).$$

证明。 由引理 2可知，存在通用的 C_1， C_2，以及 $\gamma > 0$ 使得 t

$$\mathbb{P}_q \left(\lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) \geq \gamma \right) \geq 1 - C_1 e^{-C_2 N_v}$$

当 v 较大时几乎必然成立。因此，我们得到

$$\lim_{v \to \infty} \mathbb{P}_{mpq} \left(\lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v} \right) \geq \gamma \right) = 1.$$

由此可得

$$\left\| \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right)^{-1} \right\|_{op} = \lambda_{min}^{-1} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right) = O_{\mathbb{P}}(1).$$

证明至此结束。 □

引理 18. 假设*(S1)-(S2)* 和 *(D1)-(D3)*成立。对于 $\alpha \in \mathcal{A}$，我们有

$$\left\| \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v} \right)^{-1} \right\|_{op} = O_{\mathbb{P}}(1).$$

证明. 由（31）和（37），可得

$$\left\| \frac{\boldsymbol{A}_{r,\alpha}}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right)$$

and

$$\left\| \sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} - \sum_{k \in U_v} \frac{p(\mathbf{x}_k) \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right),$$

分别成立。应用算子范数的三角不等式可得

$$\left\| \frac{\boldsymbol{A}_{r,\alpha}}{N_v} - \sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right).$$

由此可得

$$\begin{aligned} \left| \lambda_{min} \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v} \right) - \lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right) \right| &\leq \left\| \frac{\boldsymbol{A}_{r,\alpha}}{N_v} - \sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right\|_{op} \\ &= O_{\mathbb{P}} \left(\frac{1}{\sqrt{n_v}} \right). \end{aligned}$$

因此，对于任意 $\epsilon > 0$

$$\begin{aligned} &\lim_{v \to \infty} \mathbb{P}_{mpq} \left(\left| \lambda_{min} \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v} \right) - \lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right) \right| \leq \epsilon \right) \\ &\leq \lim_{v \to \infty} \mathbb{P}_{mpq} \left(\lambda_{min} \left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v} \right) \geq \lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right) - \epsilon \right) = 1. \end{aligned}$$

回顾引理17，存在 $\gamma > 0$ 使得

$$\lim_{v \to \infty} \mathbb{P}_{mpq} \left(\lambda_{min} \left(\sum_{k \in U_v} \frac{r_k \pi_k \mathbf{x}_{k,\alpha} \mathbf{x}_{k,\alpha}^\top}{N_v} \right) \geq \gamma \right) = 1.$$

请注意，对于任意随机变量 X，Y 以及任意 γ，$\epsilon > 0$，我们有

$$\mathbb{P}_{mpq}(Y \geq \gamma) \leq \mathbb{P}_{mpq}(X \geq \gamma - \epsilon) + \mathbb{P}(|Y - X| > \epsilon).$$

Thus

$$\begin{aligned}
&\mathbb{P}_{mpq}\left(\lambda_{min}\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\geq\gamma\right)\\
&\leq\mathbb{P}_{mpq}\left(\lambda_{min}\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)\geq\gamma-\epsilon\right)\\
&\quad+\mathbb{P}_{mpq}\left(\left|\lambda_{min}\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)-\lambda_{min}\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\right|>\epsilon\right).
\end{aligned}$$

由此可得

$$\lim_{v\to\infty}\mathbb{P}_{mpq}\left(\lambda_{min}\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)\geq\gamma-\epsilon\right)=1.$$

这足以得出

$$\left\|\left(\frac{\boldsymbol{A}_{r,\alpha}}{N_v}\right)^{-1}\right\|_{op}=O_{\mathbb{P}}(1).$$

□

引理19。 设 $(\widetilde{\boldsymbol{c}}_{\alpha,\,v})_{v\in\mathbb{N}}$ 是由 (36)给出的一列估计量。假设 *(S1)-(S3)*和*(D1)-(D3)*成立。我们有

$$\limsup_{v\to\infty}\|\widetilde{\boldsymbol{c}}_{\alpha,v}\|_2^4<\infty.$$

证明。 回顾可知

$$\begin{aligned}
\|\widetilde{\mathbf{c}}_{\alpha,v}\|_2&=\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v}\right\|_2^4\\
&\leq\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\right\|_{op}^4\left\|\sum_{k\in U_v}\frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v}\right\|_2^4\\
&\leq\lambda^4\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\right\|_{op}^4\left\|\sum_{k\in U_v}\frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v}\right\|_2^4.
\end{aligned}$$

一方面，根据大数定律和连续映射定理，我们得到

$$\begin{aligned}
\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\right\|_{op}^4&=\lambda_{min}^{-4}\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)\\
&\xrightarrow[v\to\infty]{a.s.}\lambda_{min}^{-4}\left(\mathbb{E}_{\mathbf{x}}\left[p(\mathbf{x}_1)\mathbf{x}_{1,\alpha}\mathbf{x}_{1,\alpha}^\top\right]\right).
\end{aligned}$$

这意味着

$$\limsup_{v\to\infty}\left\|\left(\sum_{k\in U_v}\frac{p(\mathbf{x}_k)\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^\top}{N_v}\right)^{-1}\right\|_{op}^4<\infty.$$

另一方面，

$$\left\| \sum_{k\in U_v} \frac{(1-p(\mathbf{x}_k))\mathbf{x}_{k,\alpha}}{N_v} \right\|_2^4 \le \sum_{k\in U_v} \frac{\|\mathbf{x}_{k,\alpha}\|_2^4}{N_v} \le C_0^4$$

几乎必然。因此，我们得出

$$\limsup_{v\to\infty} \|\widetilde{\mathbf{c}}_{\alpha,v}\|_2^4 < \infty$$

几乎必然。 □

引理 20。 设 $\alpha \in C$。假设 *(S1)-(S3)* 以及 *(D1)-(D3)*。则，

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\eta_{k,\alpha}^4}{N_v} < \infty$$

几乎必然。

证明。 利用（46）的结果，我们得到

$$\sum_{k\in U_v} \frac{\eta_{k,\alpha}^4}{N_v} \le \sum_{k\in U_v} \frac{8\|\mathbf{x}_{k,\alpha}\|_2^4\|\boldsymbol{\beta}_\alpha\|_2^4 + 8(8+8\pi_k^2\|\mathbf{c}_{\alpha,v}\|_2^4\|\mathbf{x}_{k,\alpha}\|_2^4)\epsilon_k^4}{N_v}$$

$$\le 8C_0^4\|\boldsymbol{\beta}_\alpha\|_2^4 + 64(1+C_0^4\|\mathbf{c}_{\alpha,v}\|_2^4) \sum_{k\in U_v} \frac{\epsilon_k^4}{N_v}.$$

由于

$$\sum_{k\in U_v} \frac{\epsilon_k^4}{N_v} \xrightarrow[v\to\infty]{a.s.} \mathbb{E}_{\mathbf{x}m}\left[\epsilon_1^4\right] \le M_0,$$

我们有

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\epsilon_k^4}{N_v} < \infty$$

几乎必然。结合引理 19，我们得到

$$\limsup_{v\to\infty} \sum_{k\in U_v} \frac{\eta_{k,\alpha}^4}{N_v} < \infty$$

几乎必然。 □

B.3.3. 感兴趣的各种统计量的有限矩

引理21。 设 $\alpha \in \mathcal{A}$ 和 $\{c_{\alpha,v}\}_{v\in\mathbb{N}}$ 为由 (6)给出的统计量序列。假设*(S1)-(S3)* 和 *(D1)-(D3)* 成立。对于 $\alpha \in C$，存在一个常数 K 使得

$$\mathbb{E}_q\left[\|\boldsymbol{c}_{\alpha,v}\|_2^4\right] \le K$$

几乎必然。

证明。写出

$$\mathbf{c}_{\alpha,v}=\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\sum_{k\in U_v}\frac{(1-r_k)\mathbf{x}_{k,\alpha}}{N_v}:=\boldsymbol{A}_{\alpha,v}^{-1}\boldsymbol{b}_{\alpha,v}.$$

我们有

$$\|\boldsymbol{b}_{\alpha,v}\|_2=\left\|\frac{1}{N_v}\sum_{k\in U_v}(1-r_k)\mathbf{x}_{k,\alpha}\right\|_2\leq\frac{1}{N_v}\sum_{k\in U_v}(1-r_k)\|\mathbf{x}_{k,\alpha}\|_2\leq C_0. \tag{44}$$

此外，回顾一下，我们是在给定高概率事件 $\mathcal{E}_N$ （见备注 1）的条件下进行讨论的，因此有

$$\|\boldsymbol{A}_{\alpha,v}^{-1}\|_{op}=\left\|\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)^{-1}\right\|_{op}=\lambda_{min}^{-1}\left(\sum_{k\in U_v}\frac{r_k\pi_k\mathbf{x}_{k,\alpha}\mathbf{x}_{k,\alpha}^{\top}}{N_v}\right)\leq\frac{1}{\gamma}, \tag{45}$$

对于某个与 N_v 无关的常数 $\gamma>0$。最后，结合（44）和（45），我们逐点有

$$\|\mathbf{c}_{\alpha,v}\|_2^4\leq\|\boldsymbol{A}_{\alpha,v}^{-1}\|_{op}^4\|\boldsymbol{b}_{\alpha,v}\|_2^4\leq\frac{C_0^4}{\gamma^4}.$$

对两边进行积分即可得到该结果。

□

引理22。 假设*(S1)-(S3)* 和*(D1)-(D3)*成立。对于 $\alpha\in C$，存在一个常数 M 使得 t

$$\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}\right]\leq M$$

几乎必然。

证明。 使用$(a+b)^4\leq 8a^4+8b^4$，我们得到

$$\begin{aligned}
\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}\right]&=\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{(\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha+r_k(1+\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha})\epsilon_k)^4}{N_v}\right]\\
&\leq\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{8(\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha)^4+8(r_k(1+\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha})\epsilon_k)^4}{N_v}\right]\\
&\leq\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{8(\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha)^4+8(1+\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha})^4\epsilon_k^4}{N_v}\right]\\
&\leq\sum_{k\in U_v}\frac{8\left(\mathbf{x}_{k,\alpha}^{\top}\boldsymbol{\beta}_\alpha\right)^4+8\mathbb{E}_m\left[\epsilon_k^4\right]\mathbb{E}_q\left[\left(1+\pi_k\mathbf{c}_{\alpha,v}^{\top}\mathbf{x}_{k,\alpha}\right)^4\right]}{N_v}\\
&\leq\sum_{k\in U_v}\frac{8\|\mathbf{x}_{k,\alpha}\|_2^4\|\boldsymbol{\beta}_\alpha\|_2^4+8\mathbb{E}_m\left[\epsilon_k^4\right](8+8\pi_k^4\mathbb{E}_q\left[\|\mathbf{c}_{\alpha,v}\|_2^4\right]\|\mathbf{x}_{k,\alpha}\|_2^4)}{N_v}\\
&\leq 8\|\boldsymbol{\beta}_\alpha\|_2^4C_0^4+64M_0+64M_0C_0^4\mathbb{E}_q\left[\|\mathbf{c}_{\alpha,v}\|_2^4\right].
\end{aligned} \tag{46}$$

利用引理 21，我们得出

$$\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}\right]\le M,$$

几乎必然。 □

引理 23. 假设满足 *(S1)-(S3)*和*(D1)-(D3)*。对于 $\alpha\in C$，存在常数 K_1 和 K_2使得

$$\mathbb{E}_{mq}\left[\left(n_v\bar{V}_{1,v}(\alpha)\right)^2\right]\le K_1,\qquad \mathbb{E}_{q}\left[\left(n_v\bar{V}_{2,v}(\alpha)\right)^2\right]\le K_2$$

几乎必然。

证明。记

$$\begin{aligned}
n_v\left|\bar{V}_{1,v}(\alpha)\right| &= \frac{n_v}{N_v^2}\left|\sum_{k\in U_v}\sum_{l\in U_v}\frac{\eta_{k,\alpha}}{\pi_k}\frac{\eta_{l,\alpha}}{\pi_l}\Delta_{kl}\right|\\
&\le \frac{n_v}{N_v^2}\sum_{k\in U_v}\sum_{l\in U_v}\frac{|\eta_{k,\alpha}|}{\pi_k}\frac{|\eta_{l,\alpha}|}{\pi_l}|\Delta_{kl}|\\
&\le \frac{n_v}{N_v}\frac{1-\lambda}{\lambda}\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}+\frac{n_v}{N_v}\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{|\eta_{k,\alpha}||\eta_{l,\alpha}|}{N_v\pi_k\pi_l}|\Delta_{kl}|\\
&\le \frac{n_v}{N_v}\frac{1-\lambda}{\lambda}\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}+\frac{n_v}{N_v\lambda^2}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\sum_{\substack{l\in U_v\\ l\neq k}}\frac{|\eta_{k,\alpha}||\eta_{l,\alpha}|}{N_v}\\
&\le \frac{n_v}{N_v}\frac{1-\lambda}{\lambda}\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}+\frac{n_v}{\lambda^2}\max_{k\neq l\in U_v}|\Delta_{kl}|\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}.
\end{aligned}\tag{47}$$

因此，使用引理22 并对 (47) 两边积分可得

$$\begin{aligned}
\mathbb{E}_{mq}\left[\left(n_v\bar{V}_{1,v}(\alpha)\right)^2\right] &\le \left(\frac{n_v}{N_v}\frac{1-\lambda}{\lambda}+\frac{n_v}{\lambda^2}\max_{k\neq l\in U_v}|\Delta_{kl}|\right)^2\mathbb{E}_{mq}\left[\left(\sum_{k\in U_v}\frac{\eta_{k,\alpha}^2}{N_v}\right)^2\right]\\
&\le \left(\frac{n_v}{N_v}\frac{1-\lambda}{\lambda}+\frac{\bar{\Delta}}{\lambda^2}\right)^2\mathbb{E}_{mq}\left[\sum_{k\in U_v}\frac{\eta_{k,\alpha}^4}{N_v}\right]\le K_1
\end{aligned}$$

几乎必然。接下来，写出

$$\begin{aligned}
n_v\left|\bar{V}_{2,v}(\alpha)\right| &= \frac{n_v\sigma^2}{N_v}\left|\sum_{k\in U_v}\frac{1-r_k+r_k\left(\pi_k\mathbf{c}_{\alpha,v}^\top\mathbf{x}_{k,\alpha}\right)^2}{N_v}\right|\\
&\le \frac{n_v\sigma^2}{N_v}+\frac{n_v\sigma^2}{N_v}\|\mathbf{c}_{\alpha,v}\|_2^2\sum_{k\in U_v}\frac{\|\mathbf{x}_{k,\alpha}\|_2^2}{N_v}\\
&= \frac{n_v\sigma^2}{N_v}\left(1+C_0^2\|\mathbf{c}_{\alpha,v}\|_2^2\right)
\end{aligned}\tag{48}$$

对式（48）两边积分，得到

$$\mathbb{E}_q\left[\left(n_v\bar{V}_{2,v}(\alpha)\right)\right] \leq \frac{n_v^2\sigma^4}{N_v^2}\left(2+2C_0^4\mathbb{E}_q\left[\|\mathbf{c}_{\alpha,v}\|_2^4\right]\right) \leq K_2$$

几乎必然使用引理21。

□

参考文献

AKAIKE, H. (1970). 统计预测器识别。《统计数学研究所年刊》22 203–217。BERGER, Y. G. and RAO, J. (2006). 不等概率不放回抽样下用于插补的调整后的刀切法。《皇家统计学会会刊B辑：统计方法学》68 531–547。BICKEL, P. J. and FREEDMAN, D. A. (1984). 分层抽样中的渐近正态性与自助法。统计年刊 470–482。BREIDT, F. J. and OPSOMER, J. D. (2000). 调查抽样中的局部多项式回归估计量。统计年刊 1026–1053。BUNEA, F., TSYBAKOV, A. B. and WEGKAMP, M. H. (2007). 高斯回归聚合。统计年刊 1674–1697。CHAUVET, G., DEVILLE, J.-C. and HAZIZA, D. (2011). 关于调查中的平衡随机插补。Biometrika 98 459–471。CHAUVET, G. and GOGA, C. (2022). 高维数据情形下校准估计量的渐近效率。《统计规划与推断杂志》217 177–187。CHEN, J. and RAO, J. (2007). 两阶段抽样设计下的渐近正态性。Statistica Sinica 1047–1064。CHEN, J. and SHAO, J. (2000). 调查数据的最近邻插补。Journal of official statistics 16 113。CHEN, J. and SHAO, J. (2001). 最近邻插补的刀切法方差估计。《美国统计学会杂志》96 260–269。DAGDOUG, M., GOGA, C. and HAZIZA, D. (2025). 通过回归树和随机森林对含插补调查数据进行统计推断。《斯堪的纳维亚统计学杂志》52 960–998。FAY, R. E. (1991). 关于缺失数据方差的基于设计的视角。美国人口普查局[custodian]。HÁJEK, J. (1960). 从有限总体进行简单随机抽样的极限分布。Publications oftheMathematicalInstituteoftheHungarian Academy of Sciences 5 361–374。HÁJEK, J. (1964). 有限总体中变化概率拒绝抽样的渐近理论。《数学统计学年刊》35 1491–1523。HANSEN, B. (2022). 计量经济学。普林斯顿大学出版社。HAZIZA, D. and BEAUMONT, J.-F. (2007). 关于调查中插补类别的构建。《国际统计评论》75 25–43。HAZIZA, D. and VALLÉE, A.-A. (2020). 单次插补调查数据情形下的方差估计程序：批判性综述。《日本统计与数据科学杂志》3 583–623。

HORN, R. A. and JOHNSON, C. R. (2012). 矩阵分析。剑桥大学出版社。
IPARRAGIRRE, A., LUMLEY, T., BARRIO, I. and AROSTEGUI, I. (2023). 复杂抽样调查数据的LASSO回归变量选择。Stat 12 e578。ISAKI, C. T. and FULLER, W. A. (1982). 回归超总体模型下的调查设计。《美国统计学会杂志》77 89–96。KIM, J. K. and RAO, J. (2009). 对项目无应答插补后调查数据的线性化方差估计采用统一方法。Biometrika 96 917–932。KREWSKI, D. and RAO, J. N. (1981). 分层样本的推断：线性化法、刀切法和平衡重复复制法的性质。统计年刊 1010–1019。LARBI, K., TSANG, J., HAZIZA, D. and DAGDOUG, M. (2025). 关于机器学习方法在调查中处理单位无应答的应用。Survey Methodology 51 275–303。LUMLEY, T. and SCOTT, A. (2015). 复杂抽样调查数据建模的赤池信息准则和贝叶斯信息准则。调查统计与方法学期刊 3 1–18。NEMIROVSKI, A. (2000). 非参数统计专题，第1738卷，圣弗卢尔概率暑期学校 1998。数学讲义。施普林格，纽约。MR1775640。
NISHII, R. (1984). 多元回归中变量选择准则的渐近性质。统计年刊 758–765。
OPSOMER, J. and MILLER, C. (2005). 复杂调查中的非参数回归估计平滑程度选择。非参数统计 17 593–611。PFEFFERMANN, D. and SVERCHKOV, M. (2009). 信息性抽样下的推断。载于《统计学手册》29 455–487。爱思唯尔。RANDLES, R. H. (1982). 关于含估计参数统计量的渐近正态性。统计年刊 462–474。RAO, R. and WU, Y. (1989). 回归问题中一种强一致的模型选择程序。Biometrika 76 369–374。
RAO, C. R., WU, Y., KONISHI, S. and MUKERJEE, R. (2001). 关于模型选择。讲义专著系列 1–64。RUBIN, D. B. (1976). 推断与缺失数据。Biometrika 63 581–592。
SÄRNDAL, C.-E. (1992). 使用插补时估计调查估计精度的方法。Survey Methodology 18 241–252。SCHWARZ, G. (1978). 估计模型维数。统计年刊 461–464。SHAO, J. (1993). 通过交叉验证进行线性模型选择。《美国统计学会杂志》88 486–494。SHAO, J. (1997). 线性模型选择的渐近理论。Statistica Sinica 221–242。
SHAO, J. and STEEL, P. (1999). 复合插补和不可忽略抽样比下调查数据的方差估计。《美国统计学会杂志》94 254–265。THOMPSON, M. (1997). 抽样调查理论 74。CRC出版社。TROPP, J. A. et al. (2015). 矩阵集中不等式导论。机器学习基础与趋势 8 1–230。WIECZOREK, J., GUERIN, C. and MCMAHON, T. (2022). 复杂抽样调查的K折交叉验证。Stat 11 e454。YANG, Y. (2005). 能否兼取赤池信息准则和贝叶斯信息准则的优势？模型识别与回归估计之间的冲突。Biometrika 92 937–950。

Yang，Y.（2007）。用于比较回归程序的交叉验证一致性。统计年刊**35** 2450–2473。Yang，S. 和 Kim，J. K.（2017）。调查抽样中的预测均值匹配插补。*arXiv* 预印本 *arXiv:1703.10256*。Yang，S. 和 Kim，J. K.（2019）。调查抽样中用于一般参数估计的最近邻插补。在复杂调查数据计量经济学：理论与应用209–234。翡翠出版有限公司。Zhang，P.（1993）。通过多重交叉验证进行模型选择。统计年刊**21**299–313。