arXiv:2302.03734v1 [math.ST] 7 Feb 2023

度校正随机块模型的一致模型选择

Andressa Cerqueira、Sandro Gallo、Florencia Leonardi 和 Cristel Vera

2023年2月9日

摘要

度修正随机块模型（DCSBM）由Karrer 和 Newman（2011）提出，是随机块模型的一种推广。在该模型中，同一社群的顶点可以具有不同的度分布。在建模方面，这种变异性使 DCSBM 更适用于现实中的复杂网络。在统计方面，由于社区检测涉及大量参数，分析起来更具挑战性。本文证明，惩罚边际似然估计量能够一致地估计社区数量。我们考虑 $dense$ 或 $semi\text{-}sparse$随机网络，并且我们的估计量 $unbounded$，也就是说，所考虑的社区数量k 最大可达 n，即网络中的节点数量。

1 引言

许多现实世界现象都可以通过对象在网络中的交互来描述。例如，社交网络中个体之间的互动、一个国家内机场之间的连接、脑区之间的连接等。大多数此类网络都具有社区结构；也就是说，属于同一群体的对象（网络的节点）往往表现出相似的行为。因此，旨在描述真实网络的概率模型需要纳入这些社区结构。

由 Holland 等人 (1983年) 提出的随机块模型（SBM）是一种允许存在社区结构的随机网络模型。每一对顶点彼此独立地以泊松分布的边数相连，其速率仅取决于相关顶点所属的社区。因此，SBM 对同一社区中节点具有相同均值的网络进行建模。这一特性可能限制其在现实网络中的应用，因为现实网络中，同一社区内顶点的度分布往往会呈现异质性（枢纽）。考虑到这一点，Karrer 和 Newman (2011) 提出了度校正随机块模型（DCSBM），该模型考虑了社区内节点度的异质性。在度修正模型中，每个节点都关联一个非负实数参数，即权重，用于指定其与网络中其他节点相连的“能力”。每个社区内权重之和对应于该社区包含的节点数，这推广了均质的 SBM：在该模型中，可以将每个节点视为权重均为 1。在 SBM 和 DCSBM 中，研究两种不同情形是一种标准做法，即稠密和半稀疏情形。在前一种情形中，决定边数的泊松分布速率固定不变（不随网络中的节点数变化）

，因此每个节点的期望度线性增长。在后一种情形中，这些速率可以趋于零，但下降方式使节点的期望度增长速度远低于 n。

文献中已有多项研究探讨了随机区块模型和度修正随机块模型的社区检测问题，其目标是估计网络中节点的 k_0 潜在群组。对于随机区块模型，研究者提出了基于谱方法（Rohe 等人，2011；Lei 和 Rinaldo，2015；Sarkar 和 Bickel，2015）、模块度（Newman 和 Girvan，2004）、似然方法（Bickel 和 Chen，2009；Celisse 等人，2012；Amini 等人，2013）以及贝叶斯视角（Decelle 等人，2011；Latouche 等人，2012；van der Pas 等人，2017）的社区检测方法。对于度修正随机块模型，Zhao 等人（2012）研究了基于模块度和基于似然的方法的一致性；Qin 和 Rohe（2013）提出了一种正则化谱聚类算法；Jin（2015）则提出了一种基于邻接矩阵特征向量逐元素比值的方法。为从随机区块模型和度修正随机块模型中选出最适合数据的模型，Yan 等人（2014）提出了一种基于似然比检验的方法，并利用置信传播对其进行近似计算。所有这些方法都假设社区数量 k_0 已知，而在实践中这种情况几乎不会发生。估计社区数量可以视为一个模型选择问题。

关于估计 k_0 社区数量的文献较为新近，至少从理论角度来看并不十分广泛。对于标准随机块模型，一些方法包括序贯假设检验（Lei 等人，2016）、交叉验证（Chen 和 Lei，2018）、谱方法（Le 和 Levina，2022）、惩罚似然准则（Wang 和 Bickel，2017；Hu 等人，2020）以及惩罚边际似然估计量（Daudin 等人，2008；Biernacki 等人，2010；Latouche 等人，2012；Cerqueira 和 Leonardi，2020）。具体而言，对于具有 n 节点、权重未知且处于半稀疏区间的 DCSBM，当 $n^{1/2}\rho_n/\log\ n \to \infty$时，Wang 和 Bickel（2017）证明了惩罚似然估计量的一致性，其中惩罚函数的阶为$k^2n\log\ n$， k是候选模型的社区数量。最近，Ma 等人（2021）提出了用于估计社区数量的似然比检验，并证明了当 $n\rho_n/\log n$ 足够大时，该方法在半稀疏区间内具有一致性。他们的方法基于谱算法，因此假设了许多额外条件，以便正确检测各个群组。这两种方法都假设社区数量有一个已知的常数上界。

本文研究了度修正随机块模型（DCSBM）下社区数量的惩罚边际似然估计量，该估计量在权重未知的情况下由 Cerqueira 和 Leonardi （2020年针对标准随机区块模型提出。该估计量可视为一种 最小描述长度 原则，在信息论领域也称为 Krichevsky-Trofimov 估计量。我们证明，在更一般的度修正模型下，无需假设 k_0 的上界，当顶点数量 n 足够大时，该估计量以概率 1 渐近几乎必然等于正确的社区数量。也就是说，优化是在所有介于 1 和 k_0之间的可能社区数量上进行的。正如上文所指出的，度修正模型具有 n额外参数，这是由于每个节点都带有关联权重。为此，与 n Cerqueira 和 Leonardi （2020年）所使用的标准随机区块模型惩罚函数相比，我们需要在惩罚函数中增加一个阶为 $kn\log\ n$ 的附加项。我们考虑与 Ma 等人 （2021年）相同的半稀疏区间，其中当 $n\rho_n/\log\ n$ 足够大时，该速率即为社区精确恢复的相变（参见 Abbe （2018年））。

本文结构如下。我们在第2节定义度修正随机块模型（DCSBM）及其相应的似然函数（参数已知时）。在第3节中，我们介绍参数的先验分布，定义惩罚边际似然估计量，并陈述主要定理，即

一致性结果，以及一个将边际似然与最大似然联系起来的关键命题。最后，我们在第4 节给出主要结果的证明。技术性证明及其他辅助结果放在附录中。

2 度修正随机块模型：定义与似然

对于任意 $n \in \mathbb{N}$，令 $X = (X_{ij})_{i,\ j\in[n]}$ ($[n] := \{1, \ldots, n\}$) 表示一个具有 n 个顶点、 $X_{ij} \in \mathbb{N}$ 的随机网络的对称邻接矩阵。对于每一对 i、 $j \in [n]$，其中 $i \neq j$ 变量 X_{ij} 表示顶点 i 和 j 之间无向边的数量（或者连接强度）。为方便起见，我们将 X_{ii} 定义为顶点 i 上自环数量的两倍。

这些顶点被随机划分为 $k_0 \geq 1$ 个社区，这种社区归属由向量 $Z = (Z_1, \ldots, Z_n)$ 表示，该向量由 $[k_0]$ 值随机变量组成（独立同分布，边缘分布为 π）。我们常用记号 $i \in [a]$ ($i \in [n], a \in [k_0]$) 表示 $Z_i = a$。

在齐性随机区块模型中，顶点 i 和 j 之间的边数期望不取决于具体顶点，而只取决于所属社区。假设社区 a和 b 之间的边数服从参数为 λ_{ab} 的泊松分布，则有

$$\mathbb{E}(X_{ij} | Z_i = z_i, Z_j = z_j) = \lambda_{z_i z_j}$$

对于任意 $i \neq j$。

在每个社区内，所有顶点的行为都相同。这一特点使随机区块模型在对现实世界中的复杂网络建模时存在不足。为了允许同一社区内不同顶点表现出不同的行为，度修正随机区块模型（DCSBM）为每个顶点 $i \in [n]$ 引入权重 w_i，该权重会影响该顶点与其他顶点建立连接的能力。在这种情况下，顶点 $i \neq j$ 之间的边数期望为

$$\mathbb{E}(X_{ij} | Z_i = z_i, Z_j = z_j) = w_i w_j \lambda_{z_i z_j}$$

而且，同一社区内不同节点的这一期望可能不同。因此，这类网络有时也称为非齐性网络。

我们假设矩阵 $\lambda = (\lambda_{ab})_{a,b\in[k_0]}$ 是对称的，且所有元素均大于零。在稠密区制下，矩阵 λ 是固定的（不依赖于 n），且所有元素均有一个正的常数下界。在这种情况下，每个节点的期望度随 n线性增长，这使得网络连接过于密集。因此，考虑一种半稀疏区制很有意义，在这种区制下，允许 λ 随 n趋于零。本文采用这种方法~，并假设对于每个 n，由 n 个节点组成的网络的分布参数为 $\lambda = \rho_n \lambda$,~，其中 $\rho_n \to 0$ 为 $n \to \infty$，而 λ 是一个元素具有正的常数下界的对称矩阵。我们将在主要结果中给出关于 ρ_n 的条件。为了保证阶数 k_0（即模型的社区数量）的可辨识性，我们假设 λ 中没有任何一列与另一列成比例。文献中通常采用这一假设，例如 Ma 等人（2021）。

现在考虑具有 $k \in [n]$ 个社区的度修正随机块模型。为了计算 (Z, X) 的联合分布，我们需要定义以下计数器。对于任意 $z \in [k]^n$ 和 $a \in [k]$，令 $n_a(z)$ 表示网络中属于社群 a 的顶点数量，即

$$n_a(z) = \sum_{i=1}^{n} \mathbb{1}\{z_i = a\}.$$

根据 Karrer 和 Newman (2011)，我们假设每个社群内部的权重向量满足

$$\sum_{i:\, z_i=a} w_i \;=\; n_a(z)$$

对所有 $a \in [k]$ 均成立。这尤其意味着网络中的总权重为 n，即顶点数；令 $w_i \equiv 1$，即可恢复经典随机区块模型。

对于任意对称矩阵 $x \in \mathbb{N}^{n\times n}$（即网络的一种实现），还定义计数器 $o_{ab}(x,z)$，其值为社区 a 和 b 中节点之间的边数，即

$$o_{ab}(x,z) = \begin{cases} \sum_{1\le i,j\le n} x_{ij}\mathbb{1}\{z_i=a, z_j=b\} & \text{for } a\neq b\,; \\ \frac{1}{2}\sum_{1\le i,j\le n} x_{ij}\mathbb{1}\{z_i=a, z_j=a\} & \text{for } a=b \end{cases} \tag{2.1}$$

节点 i 的度定义为

$$d_i(x) = \sum_{1\le j\le n} x_{ij}. \tag{2.2}$$

社群 a 的总度数记为 $d_a^t(x,z)$，其定义为

$$d_a^t(x,z) = \sum_{i:\, z_i=a} d_i(x) = \sum_{1\le i\le n} d_i(x)\mathbb{1}\{z_i=a\} = \sum_{1\le i,j\le n} x_{ij}\mathbb{1}\{z_i=a\}\,.$$

注意，我们有

$$d_a^t(x,z) = \sum_{b\neq a} o_{ab}(x,z) + 2o_{aa}(x,z)$$

社群 a 和 b 中节点对的数量记为 $n_{ab}(z)$，其定义为

$$n_{ab}(z) = \begin{cases} n_a(z)n_b(z) & \text{for } a\neq b\,; \\ \frac{1}{2}n_a(z)^2 & \text{for } a=b. \end{cases} \tag{2.3}$$

现在，在给定其他所有参数的情况下，我们可以写出具有 k 个社区的模型中 (Z,X) 的联合分布

$$p(z,x|\pi,\lambda,w) = p(z|\pi)p(x|z,w,\lambda) \tag{2.4}$$

其中

$$p(z|\pi) = \prod_{1\le a\le k} \pi_a^{n_a(z)} \tag{2.5}$$

and

$$p(x|z,w,\lambda) = \frac{1}{c(x)}\left[\prod_{1\le i\le n} w_i^{d_i(x)}\right]\left[\prod_{1\le a\le b\le k} \lambda_{ab}^{o_{ab}(x,z)}\exp\{-n_{ab}(z)\lambda_{ab}\}\right] \tag{2.6}$$

其中

$$c(x) := \left[\prod_{i<j} x_{ij}!\right]\left[\prod_i 2^{x_{ii}/2}(x_{ii}/2)!\right]. \tag{2.7}$$

3 模型选择：主要结果

为定义估计量，我们为参数（π、λ、w）引入方便的先验分布。于是，DCSBM 的层次模型分布为

$$\begin{aligned}
\pi &\sim \text{Dirichlet}(\overbrace{1/2,\ldots,1/2}^{k}) \\
\lambda_{ab} &\sim \text{Gamma}(1/2,1)\,, \quad \text{for } a,b\in[k],\, a\le b \\
z_i\,|\,\pi &\sim \pi\,, \quad i\in[n] \\
(w_i)_{i\in[a]}\,|\,z &\sim n_a(z)\,\text{Dirichlet}(\overbrace{1/2,\ldots,1/2}^{n_a(z)})\,, \quad a\in[k] \\
x_{ij}\,|\,z_i,z_j,w_i,w_j,\lambda &\sim (2-\mathbb{1}\{i\neq j\})\,\text{Poisson}(w_iw_j\lambda_{z_i,z_j})\,, \quad \text{for } i,j\in[n],\, i\le j\,.
\end{aligned}$$

令 $(\lambda_{ab})_{a,\ b\in[k]}$ 为由 $(\lambda_{ab})_{1\le a\le b\le k}$ 构造的对称矩阵。记 Θ_k 为超参数 (π,λ,w) 取值的空间，记 $\nu_k(\theta)$ 为 Θ_k 上的先验分布。

对于任意 $x\in\mathbb{N}^{n\times n}$，边际似然 $p_k(x)$ 由下列积分给出

$$p_k(x)=\sum_{z\in[k]^n}\int_{\Theta_k}p(x,z|\theta)\nu_k(\theta)d\theta$$

其中 ν_k 按照（5.7）（见附录）进行了显式分解。现在，我们可以将社区数量的估计量定义为

$$\widehat{k}_n(x):=\operatorname*{arg\,max}_{1\le k\le n}\left\{\,\log p_k(x)-(k^3+3kn)\log(n+1)\,\right\}\,. \tag{3.1}$$

ˆ注 3.1。 注意，估计量 k_n 在 (3.1) 比 $n\log n$ 多了一个阶为 Cerqueira 和 Leonardi (2020 年）所定义的估计量的惩罚项。模型中增加 n 参数 $w_i,\ i=1,\ldots,n$ 带来了这一额外项，而在齐性随机块模型中则不需要这一项。

现在陈述我们的主要定理。

定理 3.1。 对于具有 k_0 个社区且 $\rho_n\ge C\log_n^{\,n}$，$n\ge 1$，其中 C 是一个足够大的常数，式 (3.1) 中定义的估计量满足

$$\widehat{k}_n=k_0 \tag{3.2}$$

当 n 趋于无穷大时，最终几乎必然成立。

定理3.1 的证明基于一个关键命题（见下文），该命题将边际似然 $p_k(x)$ 与最大似然 $\sup_{\theta\in\Theta_k}p(x|\theta)=\sup_{\pi}\ p(z|\pi)\sup_{\lambda,\ w}\ p(x|z,w,\lambda)$ 联系起来。我们现在对其进行定义。对于任意 $x\in\mathbb{N}^{n\times n}$、$z\in[k]^n$、$i\in[n]$ 以及 a、$b\in[k]$，π、λ 和 w 的最大似然估计量为

$$\widehat{\pi}_a(z)=\frac{n_a(z)}{n}\,,\quad \widehat{\lambda}_{ab}(x,z)=\frac{o_{ab}(x,z)}{n_{ab}(z)}\,,\quad \widehat{w}_i(x,z)=\sum_{1\le a\le k}\mathbb{1}\{z_i=a\}\frac{n_a(z)d_i(x)}{d_a^t(x,z)}\,,$$

因此

$$\sup_{\pi}\ p(z|\pi) = \prod_{1\le a\le k}\left(\frac{n_a(z)}{n}\right)^{n_a(z)} \tag{3.3}$$

and

$$\begin{aligned}\sup_{\lambda,w}\ p(x|z,w,\lambda) = \frac{1}{c(x)}\prod_{1\le a\le b\le k}\left(\frac{o_{ab}(x,z)}{n_{ab}(z)}\right)^{o_{ab}(x,z)} e^{-o_{ab}(x,z)} \\ \times \prod_{a\in[k],i\in[n]:z_i=a}\left(\frac{n_a(z)d_i(x)}{d_a^t(x,z)}\right)^{d_i(x)}.\end{aligned} \tag{3.4}$$

现在我们陈述关键命题，该命题对一组“良好”的网络成立（我们将证明其以高概率成立），这组网络定义如下：

$$\Omega_n := \{x\colon x_{ij} \le \log n,\ \text{for all } i,j\in[n]\}\,. \tag{3.5}$$

命题 3.1. 对所有 $k\ge 1$、所有 $n \ge \max(k,3)$ 以及所有 $x\in\Omega_n$，我们有

$$0\ \le\ \log\left(\frac{\sup_{\theta\in\Theta_k} p(x|\theta)}{p_k(x)}\right)\ \le\ k(k+2)\log(n+1)+3n\log n\,.$$

该命题的证明见附录。

4 主要结果的证明

定理 3.1 的证明分为两部分：首先，我们证明社区数量不低估，然后证明不高估。附录中给出了其他结果和定义。

我们首先给出估计量的证明 $\widehat{k}_n$ 不会低估社区的真实数量 k_0。

命题 4.1。对于满足 $\rho_n \ge {}^{C\,\text{对数}_n}{}^{\,n}$ 的稀疏度校正随机块模型，当常数 $C>0$ 足够大时，$\widehat{}$估计量 k_n 按 (3.1) 定义，满足

$$\widehat{k}_n \ge k_0,$$

当 $n\to\infty$ 时最终几乎必然成立。

证明。我们基于模型下的观测图，为含有 k 个社区的情形定义剖面估计量如下：

$$\widehat{z}_k = \operatorname*{arg\,max}_{z\in\{1,2,\dots,k\}^n}\ \sup_{\theta\in\Theta_k}\ p(x,z|\theta)\,. \tag{4.1}$$

根据引理 5.6，我们可以（并且今后将）取 n 足够大，使得 $x\in\Omega_n$，即$\widehat{}$在（3.5）中定义的“良好”网络集合。为了证明 $k_n(x)\ge k_0$在 $n\to\infty$时几乎必然成立，只需证明对所有 $k<k_0$，

$$\log p_{k_0}(x) - (k_0^3+3k_0n)\log(n+1)\ >\ \log p_k(x) - (k^3+3kn)\log(n+1)\,, \tag{4.2}$$

当 $n \to \infty$ 时几乎必然成立。但如果我们证明

$$\liminf_{n\to\infty} \frac{1}{\rho_n n^2} \log \frac{p_{k_0}(x)}{p_k(x)} > 0\,, \tag{4.3}$$

并且由于

$$\frac{1}{\rho_n n^2}\left(k_0^3 + 3k_0 n - k^3 - 3kn\right)\log(n+1)$$

根据对 ρ_n 的假设，可以使其足够小，因此（4.3）蕴含（4.2）。注意，对于所有 $x \in \Omega_n$，根据命题 3.1 以及 $p_k(x) \leq \sup^{\theta\in\Theta_k} p(x|\theta)$ 这一事实，我们有

$$\frac{1}{\rho_n n^2}\log\frac{p_{k_0}(x)}{p_k(x)} = \frac{1}{\rho_n n^2}\log\frac{p_{k_0}(x)}{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)} + \frac{1}{\rho_n n^2}\log\frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)} + \frac{1}{\rho_n n^2}\log\frac{\sup_{\theta\in\Theta_k} p(x|\theta)}{p_k(x)}$$

$$\geq -\gamma(k_0)\frac{\log n}{\rho_n n} + \frac{1}{\rho_n n^2}\log\frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)}\,,$$

对于某个常数 $\gamma(k_0)$。接下来，为了证明（4.3），只需证明对于 $k < k_0$

$$\liminf_{n\to\infty} \frac{1}{\rho_n n^2}\log\frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)} > 0 \tag{4.4}$$

因为根据对 ρ_n 的假设，$-\gamma(k_0)\frac{\log n}{\rho_n n}$ 也足够小。首先注意，对于所有 $z \in [k_0]^n$，我们有

$$\log \sup_{\theta\in\Theta_{k_0}} p(x|\theta) \;\geq\; \log \sup_{\theta\in\Theta_{k_0}} p(x,z|\theta)\,. \tag{4.5}$$

Let

$$\tilde{o}_{ab}(x,z) := \sum_{1\leq i,j\leq n} x_{ij}\mathbb{1}\{z_i = a, z_j = b\}$$

对于所有数对 a，b 并注意到它与 o_{ab} 通过 $\tilde{o}_{ab}(x,\ z) = o_{ab}(x,\ z)$ 对所有 $a \neq b$ 以及 $\tilde{o}_{aa}(x,\ z) = 2o_{aa}(x,\ z)$ 对所有 a。由式（2.4）-（2.7）以及最大似然估计量的定义可得

$$\begin{aligned}\log \sup_{\theta\in\Theta_{k_0}} p(x,z|\theta) \;=\; & L(x) + n\sum_{a=1}^{k_0}\widehat{\pi}_a(z)\log\widehat{\pi}_a(z) + \frac{1}{2}\sum_{1\leq a,b\leq k_0}\tilde{o}_{ab}(x,z)\log\widehat{\lambda}_{ab}(x,z)\\ & + \sum_{1\leq a\leq k_0} d_a^t(x,z)\log\frac{n_a(z)}{d_a^t(x,z)}\end{aligned} \tag{4.6}$$

with

$$\begin{aligned} L(x) \;&=\; -\log c(x) - \sum_{i,j} x_{ij} + \sum_i d_i(x)\log d_i(x)\,;\\ \widehat{\pi}_a(z) \;&=\; \frac{n_a(z)}{n}, \qquad 1 \leq a \leq k_0;\\ \widehat{\lambda}_{ab}(x,z) \;&=\; \frac{\tilde{o}_{ab}(x,z)}{n_a(z)n_b(z)}, \qquad 1 \leq a,b \leq k_0;\\ d_a^t(x,z) \;&=\; \sum_{1\leq i,j\leq n} x_{ij}\mathbb{1}\{z_i = a\} \;=\; \sum_b \tilde{o}_{ab}(x,z), \qquad 1 \leq a \leq k_0\,. \end{aligned}$$

对于式（4.4）中的分母，我们利用

$$\begin{aligned}\log \sup_{\theta\in\Theta_k} p(x|\theta) &\le \log k^n \sup_{\theta\in\Theta_k} p(x,\widehat{z}_k|\theta)\\ &\le n\log k + \log \sup_{\theta\in\Theta_k} p(x,\widehat{z}_k|\theta)\,,\end{aligned} \tag{4.7}$$

其中 $\widehat{z}_k$ 定义为 (4.1)。类似于 (4.6)，我们有

$$\begin{aligned}\log \sup_{\theta\in\Theta_k} p(x,\widehat{z}_k|\theta) &= L(x) + n\sum_{1\le a\le k}\widehat{\pi}_a(\widehat{z}_k)\log\widehat{\pi}_a(\widehat{z}_k) + \frac{1}{2}\sum_{1\le a,b\le k}\tilde{o}_{ab}(\widehat{z}_k)\log\widehat{\lambda}_{ab}(x,\widehat{z}_k)\\ &\quad + \sum_{1\le a\le k} d_a^t(x,\widehat{z}_k)\log\frac{n_a(\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)}\,.\end{aligned} \tag{4.8}$$

于是，(4.4) 中的对数可以用 (4.5) 与 (4.7) 之差作为下界，并利用 (4.6) 和 (4.8) 中的表达式，可得

$$\begin{aligned}\log\frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)} &\ge n\sum_{1\le a\le k_0}\widehat{\pi}_a(z)\log\widehat{\pi}_a(z) + \frac{1}{2}\sum_{1\le a,b\le k_0}\tilde{o}_{ab}(x,z)\log\widehat{\lambda}_{ab}(x,z)\\ &\quad + \sum_{1\le a\le k_0} d_a^t(x,z)\log\frac{n_a(z)}{d_a^t(x,z)} - n\log k\\ &\quad - n\sum_{1\le a\le k}\widehat{\pi}_a(\widehat{z}_k)\log\widehat{\pi}_a(\widehat{z}_k) - \frac{1}{2}\sum_{1\le a,b\le k}\tilde{o}_{ab}(x,\widehat{z}_k)\log\widehat{\lambda}_{ab}(x,\widehat{z}_k)\\ &\quad - \sum_{1\le a\le k} d_a^t(x,\widehat{z}_k)\log\frac{n_a(\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)}\,.\end{aligned} \tag{4.9}$$

现在我们将重新整理右侧的六项。首先，令

$$\mathcal{A}(n) := n\sum_{1\le a\le k_0}\widehat{\pi}_a(z)\log\widehat{\pi}_a(z) - n\log k - n\sum_{1\le a\le k}\widehat{\pi}_a(\widehat{z}_k)\log\widehat{\pi}_a(\widehat{z}_k). \tag{4.10}$$

其次，我们可以写成

$$\begin{aligned}\sum_{1\le a\le k_0} d_a^t(x,z)\log\frac{n_a(z)}{d_a^t(x,z)} &= \frac{1}{2}\sum_{1\le a\le k_0} d_a^t(x,z)\log\frac{n_a(z)}{d_a^t(x,z)} + \frac{1}{2}\sum_{1\le b\le k_0} d_b^t(x,z)\log\frac{n_b(z)}{d_b^t(x,z)}\\ &= \frac{1}{2}\sum_{1\le a,b\le k_0}\tilde{o}_{ab}(x,z)\log\frac{n_a(z)n_b(z)}{d_a^t(x,z)d_b^t(x,z)}\end{aligned} \tag{4.11}$$

因此

$$\frac{1}{2}\sum_{1\le a,b\le k_0}\tilde{o}_{ab}(x,z)\log\widehat{\lambda}_{ab}(x,z) + \sum_{1\le a\le k_0} d_a^t(x,z)\log\frac{n_a(z)}{d_a^t(x,z)} = \frac{1}{2}\sum_{1\le a,b\le k_0}\tilde{o}_{ab}(x,z)\log\frac{\tilde{o}_{ab}(x,z)}{d_a^t(x,z)d_b^t(x,z)}. \tag{4.12}$$

使用（4.10）、（4.12）以及在 k 阶模型下，将 z 替换为 $\widehat{z}_k$ 后，（4.9）的右侧现在为

$$\log \frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)} \;\geq\; \mathcal{A}(n) + \frac{1}{2}\sum_{1\leq a,b\leq k_0} \tilde{o}_{ab}(x,z)\log\frac{\tilde{o}_{ab}(x,z)}{d_a^t(x,z)d_b^t(x,z)} - \frac{1}{2}\sum_{1\leq a,b\leq k} \tilde{o}_{ab}(x,\widehat{z}_k)\log\frac{\tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\,.$$

现在，将等式两边除以4.9，并在右侧加上以下项（该项等于 0） $\rho_n n^2$，然后求和

$$\frac{1}{\rho_n n^2}\left(\frac{1}{2}\sum_{1\leq a,b\leq k_0}\tilde{o}_{ab}(x,z)\log\rho_n n^2 - \frac{1}{2}\sum_{1\leq a,b\leq k}\tilde{o}_{ab}(x,\widehat{z}_k)\log\rho_n n^2\right)$$

我们最终得到

$$\begin{aligned}\frac{1}{\rho_n n^2}\log\frac{\sup_{\theta\in\Theta_{k_0}} p(x|\theta)}{\sup_{\theta\in\Theta_k} p(x|\theta)} \;\geq\; \frac{1}{2}\Bigg(&\sum_{1\leq a,b\leq k_0}\frac{\tilde{o}_{ab}(x,z)}{\rho_n n^2}\log\frac{\rho_n n^2\tilde{o}_{ab}(x,z)}{d_a^t(x,z)d_b^t(x,z)} \\ &- \sum_{1\leq a,b\leq k}\frac{\tilde{o}_{ab}(x,\widehat{z}_k)}{\rho_n n^2}\log\frac{\rho_n n^2\tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\Bigg) + \frac{\mathcal{A}(n)}{\rho_n n^2}.\end{aligned} \tag{4.13}$$

由于 $\frac{\mathcal{A}(n)}{\rho_n n^2}$ 几乎必然收敛于 0，因此证明（4.13）最终几乎必然被一个正常数从下方界定（当 $n\to\infty$时），等价于证明

$$\begin{aligned}\liminf_{n\to\infty}\;&\sum_{1\leq a,b\leq k_0}\frac{d_a^t(x,z)d_b^t(x,z)}{\rho_n^2 n^4}\varphi\Big(\frac{\rho_n n^2\tilde{o}_{ab}(x,z)}{d_a^t(x,z)d_b^t(x,z)}\Big) \\ &- \sum_{1\leq a,b\leq k}\frac{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4}\varphi\Big(\frac{\rho_n n^2\tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\Big) \;>\; 0\,,\end{aligned} \tag{4.14}$$

其中 $\varphi(u) = u$ 对数 $u_\circ$ 根据引理 5.7 可得

$$\begin{aligned}\frac{\rho_n n^2\tilde{o}_{ab}(x,z)}{d_a^t(x,z)d_b^t(x,z)} \;=\; \frac{\tilde{o}_{ab}(x,z)/\rho_n n^2}{d_a^t(x,z)d_b^t(x,z)/\rho_n^2 n^4} \;&\to\; \frac{[\mathrm{diag}(\pi)\tilde{\lambda}\mathrm{diag}(\pi)^T]_{ab}}{[\mathrm{diag}(\pi)\tilde{\lambda}\mathrm{diag}(\pi)^T\mathbf{1}_k]_a[\mathrm{diag}(\pi)\tilde{\lambda}\mathrm{diag}(\pi)^T\mathbf{1}_k]_b} \\ &=\; \frac{\pi_a\pi_b\tilde{\lambda}_{ab}}{\pi_a[\tilde{\lambda}\pi]_b\pi_b[\tilde{\lambda}\pi]_a} \\ &=\; \frac{\tilde{\lambda}_{ab}}{[\tilde{\lambda}\pi]_a[\tilde{\lambda}\pi]_b}\end{aligned} \tag{4.15}$$

˜其中回顾一下， λ 是满足 $\lambda = \rho_n \lambda$ 的矩阵。于是我们有

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{2} \sum_{1\le a,b\le k_0} \frac{d_a^t(x,z) d_b^t(x,z)}{\rho_n^2 n^4} \varphi\Big(\frac{\rho_n n^2 \tilde{o}_{ab}(x,z)}{d_a^t(x,z) d_b^t(x,z)}\Big) \\ = \frac{1}{2} \sum_{1\le a,b\le k_0} \pi_a \pi_b [\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b \varphi\Big(\frac{\tilde{\lambda}_{ab}}{[\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b}\Big).\end{aligned} \tag{4.16}$$

另一方面，根据引理 5.8 可知

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{2} \sum_{1\le a,b\le k} \frac{d_a^t(x,\widehat{z}_k) d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4} \varphi\Big(\frac{\rho_n n^2 \tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k) d_b^t(x,\widehat{z}_k)}\Big) \\ \le \frac{1}{2} \sum_{1\le a,b\le k} \pi_a^* \pi_b^* [\lambda^*\pi^*]_a [\lambda^*\pi^*]_b \varphi\Big(\frac{\lambda_{ab}^*}{[\lambda^*\pi^*]_a [\lambda^*\pi^*]_b}\Big)\end{aligned} \tag{4.17}$$

存在某个 $k \times k$ 正矩阵 λ^* 以及 k 维向量 π^* 由（5.35）定义。最后，根据引理 5.9 可知，（4.16）与（4.17）之差的下界为

$$\frac{1}{2}\left(\sum_{1\le a,b\le k_0} \pi_a \pi_b [\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b \varphi\Big(\frac{\tilde{\lambda}_{ab}}{[\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b}\Big) - \sum_{1\le a,b\le k} \pi_a^* \pi_b^* [\lambda^*\pi^*]_a [\lambda^*\pi^*]_b \varphi\Big(\frac{\lambda_{ab}^*}{[\lambda^*\pi^*]_a [\lambda^*\pi^*]_b}\Big)\right) > 0$$

˜除非 λ 有两列成比例，这与 k_0可辨识性的假设矛盾。由此完成命题4.1的证明。

□

ˆ我们通过证明 3.1 来完成定理的证明： $k_n(x)$ 不会高估 k_0，即真实社区数量。

ˆ命题 4.2。 对于度修正随机块模型，估计量 k_n 定义于 (3.1) 满足

$$\widehat{k}_n \le k_0$$

最终几乎必然成立，当 $n \to \infty$。

请注意，该命题对 ρ_n 没有任何假设。

证明。根据博雷尔-坎泰利引理，只需证明下列级数收敛

$$\begin{aligned}\sum_{n=1}^{\infty} \sum_{k=k_0+1}^{n} p(\widehat{k}_n = k) &= \sum_{n=1}^{\infty} \sum_{k=k_0+1}^{n} \sum_{x\in\Omega_n^c} p(x) \mathbb{1}\{\widehat{k}_n(x) = k\} \\ &\quad + \sum_{n=1}^{\infty} \sum_{k=k_0+1}^{n} \sum_{x\in\Omega_n} p(x) \mathbb{1}\{\widehat{k}_n(x) = k\}.\end{aligned} \tag{4.18}$$

我们先来看右侧的第一项，并注意到对于每个 n

$$\sum_{k=k_0+1}^{n} \sum_{x\in\Omega_n^c} p(x) \mathbb{1}\{\widehat{k}_n(x) = k\} = \sum_{x\in\Omega_n^c} p(x) \sum_{k=k_0}^{n} \mathbb{1}\{\widehat{k}_n(x) = k\} \le \sum_{x\in\Omega_n^c} p(x) = p(\Omega_n^c).$$

现在利用引理 5.6 我们得出，式（4.18）的第一项确实在 n 中可求和。因此，我们现在需要证明式（4.18）的第二项在 n 中也可求和。首先注意，对于某些固定的 $k = k_0 + 1$、.....、 n，有

$$\sum_{x\in\Omega_n} p(x)\mathbb{1}\{\widehat{k}_n(x) = k\} \;=\; \sum_{x\in\Omega_n} p(x)\mathbb{1}\{\arg\max_{\ell}(\log p_\ell(x) - (\ell^3 + 3\ell n)\log(n+1)) = k\}. \tag{4.19}$$

根据 $k_n(x)$ 的定义，当 $k > k_0$ 时，我们有

$$\begin{aligned}&\Big\{\arg\max_{\ell}[\log\,\log p_\ell(x) - (\ell^3 + 3\ell n)\log(n+1)] = k\Big\}\\ &\qquad\subset\; \Big\{\log p_k(x) - (k^3 + 3kn)\log(n+1) \geq \log p_{k_0}(x) - (k_0^3 + 3k_0 n)\log(n+1)\Big\}.\end{aligned}$$

So

$$\begin{aligned}&\sum_{x\in\Omega_n} p(x)\mathbb{1}\{\widehat{k}_n(x) = k\}\\ &\quad\leq \sum_{x\in\Omega_n} p(x)\mathbb{1}\big\{\log p_k(x) - (k^3 + 3kn)\log(n+1) \geq \log p_{k_0}(x) - (k_0^3 + 3k_0 n)\log(n+1)\big\}\\ &\quad= \sum_{x\in\Omega_n} p(x)\mathbb{1}\big\{p_{k_0}(x) \leq p_k(x)\exp[(-k^3 - 3kn + k_0^3 + 3k_0 n)\log(n+1)]\big\}.\end{aligned} \tag{4.20}$$

利用命题 3.1，对于 $x \in \Omega_n$，我们有

$$\begin{aligned}\log p(x) \;&\leq\; \log\sup_{\theta\in\Theta^{k_0}} p(x|\theta).\\ &\leq\; \log p_{k_0}(x) + k_0(k_0+2)\log(n+1) + 3n\log n,\end{aligned} \tag{4.21}$$

从而得到

$$p(x) \;\leq\; p_{k_0}(x)e^{k_0(k_0+2)\log(n+1)+3n\log n}. \tag{4.22}$$

那么，在集合 $\{p_{k_0}(x) \leq p_k(x)\exp[(k_0^3 + 3k_0 n - k^3 - 3kn)\log(n+1)]\}$ 上，和式 (4.20) 中的每一项都可以上界为

$$\begin{aligned}&p(x)\mathbb{1}\big\{p_{k_0}(x) \leq p_k(x)e^{(k_0^3+3k_0 n-k^3-3kn)\log(n+1)}\big\}\\ &\qquad\leq\; p_k(x)e^{k_0(k_0+2)\log(n+1)+3n\log n+(k_0^3+3k_0 n-k^3-3kn)\log(n+1)}.\end{aligned} \tag{4.23}$$

注意，当 $k \geq k_0 + 1$ 时，式 (4.23) 中的指数可以上界为

$$(-2k_0^2 - k_0 - 1)\log(n+1) \;\leq\; -4\log n\,.$$

现在将式 (4.23) 代入式 (4.20)，并对 $k = k_0 + 1, \ldots, n$ 求和，得到

$$\begin{aligned}\sum_{k=1}^{n}\sum_{x\in\Omega_n} p(x)\mathbb{1}\big\{\widehat{k}_n(x) = k\big\} \;&\leq\; n\,n^{-4}\sum_{x\in\Omega_n} p_k(x)\\ &\leq\; n^{-3}\end{aligned} \tag{4.24}$$

在 n 中可求和。这就完成了命题4.2的证明。 □

致谢

本研究是神经数学研究、创新与传播中心（FAPESP 资助项目 2013/07699-0）活动的一部分。此外，本研究还得到 FAPESP 项目（资助编号 2017/10555-0）“相互作用系统的随机建模”和 CNPq 普遍项目（资助编号 432310/2018-5）“统计学、随机过程与离散结构”的支持。FL 部分得到 CNPq 研究奖学金（资助编号 311763/2020-0）的支持。在开展本研究期间，CV 获得了 CAPES 博士奖学金。

参考文献

Emmanuel Abbe。社区检测与随机块模型：最新进展。机器学习研究杂志，18(177):1–86，2018年。网址：http://jmlr.org/papers/v18/16-480.html。Arash A. Amini、Aiyou Chen、Peter J. Bickel 和 Elizaveta Levina。大型稀疏网络中社区检测的伪似然方法。《统计年刊》，41(4):2097–2122，2013年。Peter J Bickel 和 Aiyou Chen。网络模型以及 Newman–Girvan 模块度和其他模块度的非参数视角。美国国家科学院院刊，106(50):21068–21073，2009年。Christophe Biernacki、G´erard Govaert。潜在类别模型积分似然的精确计算与蒙特卡罗计算。《统计规划与推断杂志》，140(11):2991–3002，2010年。St´ephane Boucheron、G´abor Lugosi 和 Pascal Massart。集中不等式：独立性的非渐近理论。牛津大学出版社，2013年。Alain Celisse、Jean-Jacques Daudin 和 Laurent Pierre。随机块模型中极大似然估计量和变分估计量的一致性。《电子统计学杂志》，6(none):1847–1899，2012年。Andressa Cerqueira 和 Florencia Leonardi。随机块模型中社区数量的估计。《IEEE 信息论汇刊》，66(10):6403–6412，2020年。doi：10.1109/TIT.2020.3016331。Kehui Chen 和 Jing Lei。用于确定网络数据中社区数量的网络交叉验证。《美国统计学会会刊》，113(521):241–251，2018年。J-J Daudin、Franck Picard 和 St´ephane Robin。随机图的混合模型。《统计与计算》，18(2):173–183，2008年。L Davisson、R McEliece、M Pursley 和 Mark Wallace。高效的通用无噪声源编码。《IEEE 信息论汇刊》，27(3):269–279，1981年。Aurelien Decelle、Florent Krzakala、Cristopher Moore 和 Lenka Zdeborov´a。模块化网络随机块模型的渐近分析及其算法应用。《物理评论 E》，84:066106，2011年12月。

Paul W Holland、Kathryn Blackmond Laskey 和 Samuel Leinhardt。随机区块模型：初探。《社会网络》，5(2):109–137，1983年。Jianwei Hu、Hong Qin、Ting Yan 和 Yunpeng Zhao。随机块模型的修正贝叶斯信息准则。《美国统计学会会刊》，115(532):1771–1783，2020年。Jiashun Jin。基于得分的快速社区检测。《统计年刊》，43(1):57–89，2015年。Brian Karrer 和 Mark EJ Newman。随机区块模型与网络中的社区结构。《物理评论 E》，83(1):016107，2011年。Pierre Latouche、Etienne Birmele 和 Christophe Ambroise。随机块模型的变分贝叶斯推断与复杂度控制。《统计建模》，12(1):93–115，2012年。Can M. Le 和 Elizaveta Levina。利用谱方法估计社区数量。《电子统计学杂志》，16(1):3315–3342，2022年。Jing Lei 和 Alessandro Rinaldo。随机块模型中谱聚类的一致性。《统计年刊》，43(1):215–237，2015年。Jing Lei 等人。随机块模型的拟合优度检验。《统计年刊》，44(1):401–424，2016年。Shujie Ma、Liangjun Su 和 Yichong Zhang。确定度校正随机块模型中的社区数量。机器学习研究杂志，22(69)，2021年。Mark E. J. Newman 和 Michelle Girvan。发现并评估网络中的社区结构。《物理评论 E：统计、非线性与软物质物理》，69 2 Pt 2:026113，2004年。Tai Qin 和 Karl Rohe。度校正随机块模型下的正则化谱聚类。载于 C.J. Burges、L. Bottou、M. Welling、Z. Ghahramani 和 K.Q. Weinberger 编，《神经信息处理系统进展》第26卷。Curran Associates, Inc.，2013年。Karl Rohe、Sourav Chatterjee 和 Bin Yu。谱聚类与高维随机区块模型。《统计年刊》，39(4):1878–1915，2011年。Purnamrita Sarkar 和 Peter J. Bickel。归一化在随机区块模型谱聚类中的作用。《统计年刊》，43(3):962–990，2015年。SL van der Pas、AW van der Vaart 等人。贝叶斯社区检测。《贝叶斯分析》，2017年。YX Rachel Wang 和 Peter J Bickel。随机块模型的基于似然的模型选择。《统计年刊》，45(2):500–528，2017年。Xiaoran Yan、Cosma Shalizi、Jacob E Jensen、Florent Krzakala、Cristopher Moore、Lenka Zde-borov´a、Pan Zhang 和 Yaojia Zhu。度修正区块模型的模型选择。《统计力学杂志：理论与实验》，2014(5):P05007，2014年。Yunpeng Zhao、Elizaveta Levina 和 Ji Zhu。度修正随机块模型下网络社区检测的一致性。《统计年刊》，40(4):2266–2292，2012年。

5 附录

5.1 基本结果

下面陈述引理 5.1 和 5.2 以供完整起见。证明收录于补充材料中。

引理 5.1. 对于整数 $m = m_1 + \cdots + m_J$，有

$$\frac{\prod_{j=1}^{J} \left(\frac{m_j}{m}\right)^{m_j}}{\prod_{j=1}^{J} \Gamma\left(m_j + \frac{1}{2}\right)} \leq \frac{1}{\Gamma\left(m + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right)^{J-1}}. \tag{5.1}$$

证明。对于整数 m，有

$$\Gamma\left(m + \frac{1}{2}\right) = (m+1)(m+2)\ldots(2m)\frac{\sqrt{\pi}}{2^m}.$$

因此，对于满足 $m = m_1 + \cdots + m_J$ 的整数 m_j、 $j = 1$、.....、 J，我们写作

$$\frac{\prod_{j=1}^{J} \Gamma\left(m_j + \frac{1}{2}\right)}{\Gamma\left(m + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right)^{J-1}} = \frac{\prod_{j=1}^{J} (m_j+1)(m_j+2)\cdots(2m_j)}{(m+1)(m+2)\cdots(2m)}. \tag{5.2}$$

定义，对于 $y \geq 0$ 和整数 $r \geq 1$

$$g_r(y) = \prod_{i=1}^{r} \left(y + \frac{i}{r}\right).$$

正如Davisson 等人（1981）附录中的引理所示，对于整数$m = m_1 + \cdots + m_J$和 $y \geq 0$，我们有

$$g_m(y) \leq \prod_{j=1}^{J} g_{m_j}(y). \tag{5.3}$$

利用这一结果，对于 $y = 1$ 和 $r = m$，我们有

$$\frac{\prod_{i=1}^{m} (m+i)}{m^m} \leq \prod_{j=1}^{J} \prod_{i=1}^{m_j} \frac{(m_j+i)}{m_j} = \frac{\prod_{j=1}^{J} (m_j+1)(m_j+2)\cdots(2m_j)}{\prod_{j=1}^{J} m_j^{m_j}}. \tag{5.4}$$

重新整理（5.4），并与（5.2）结合，可得

$$\prod_{j=1}^{J} \left(\frac{m_j}{m}\right)^{m_j} \leq \frac{\prod_{j=1}^{J} (m_j+1)(m_j+2)\cdots(2m_j)}{(m+1)(m+2)\cdots(2m)} = \frac{\prod_{j=1}^{J} \Gamma\left(m_j + \frac{1}{2}\right)}{\Gamma\left(m + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right)^{J-1}}. \qquad \square$$

Lemma 5.2。对于整数 $m = m_1 + \cdots + m_J$，其中 $J \geq 1$ and $m \geq \max(J, 3)$，我们有

$$\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m+\frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)\Gamma\left(m+\frac{1}{2}\right)} \;\leq\; m^J\,.$$

证明。 Γ 函数的斯特林公式指出，对于所有的 $y \geq 0$ ，我们有

$$y^{y-\frac{1}{2}}e^{-y}\sqrt{2\pi} \;\leq\; \Gamma(y) \;\leq\; y^{y-\frac{1}{2}}e^{-y}\sqrt{2\pi}e^{\frac{1}{12y}}\,.$$

Then

$$\begin{aligned}
\log\left(\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m+\frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)\Gamma\left(m+\frac{1}{2}\right)}\right) \;\leq\; & \left(m+\frac{J-1}{2}\right)\log\left(m+\frac{J}{2}\right) - \left(m+\frac{J}{2}\right) + \frac{1}{12\left(m+\frac{J}{2}\right)} \\
& - m\log\left(m+\frac{1}{2}\right) + \left(m+\frac{1}{2}\right) + \log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})} \\
\leq\; & \left(m+\frac{J-1}{2}\right)\log\left(m\left(1+\frac{J}{2m}\right)\right) + \frac{1}{12m} - m\log\left(m\left(1+\frac{1}{2m}\right)\right) \\
& - \frac{J-1}{2} + \log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})} \\
\leq\; & \left(\frac{J-1}{2}\right)\log m + \left(m+\frac{J-1}{2}\right)\log\left(1+\frac{J}{2m}\right) - m\log\left(1+\frac{1}{2m}\right) \\
& + \frac{1}{12m} - \frac{J-1}{2} + \log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})}\,.
\end{aligned} \tag{5.5}$$

利用 $1 - \frac{1}{y} \leq \log y \leq y - 1$，对于 $y > 0$，我们得到

$$\begin{aligned}
\log\left(\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m+\frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)\Gamma\left(m+\frac{1}{2}\right)}\right) &\leq \left(\frac{J-1}{2}\right)\log m + \left(m+\frac{J-1}{2}\right)\left(\frac{J}{2m}\right) + \frac{1}{12m} - \frac{J-1}{2} + \log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})} \\
&\leq \left(\frac{J-1}{2}\right)\log m + \frac{J(J-1)}{4m} + \frac{1}{12m} + \log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})}\,.
\end{aligned}$$

注意，对于 $J \geq 1$ 和 $m \geq \max(J, 3)$，我们有

$$\log\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{J}{2})} \;\leq\; \log(2)$$

and

$$\frac{J(J-1)}{4m} \;\leq\; \frac{J-1}{4} \;\leq\; \frac{J-1}{4}\log m\,.$$

Then

$$\log\left(\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m+\frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)\Gamma\left(m+\frac{1}{2}\right)}\right) \;\leq\; \left(\frac{J-1}{2}\right)\log m + \frac{J-1}{4}\log m + 1 \leq\; J\log m\,. \qquad \square$$

5.2 命题3.1 的证明

对于每个 $z \in [k]^n$，超参数空间 $\Theta_k = \Theta_k(z)$ 的形式为

$$\Theta_k = \Pi_k \times \Lambda_k \times \mathcal{W}(z)$$

其中

$$\Pi_k := \{(\pi_1, \ldots, \pi_k) \in (0,1]^k \colon \sum_i \pi_i = 1\}$$

是标准的 k维单纯形,

$$\Lambda_k := \{\lambda \in (\mathbb{R}^+)^{k \times k} \colon \lambda_{ab} = \lambda_{ba} \text{ for all } a, b \in [k]\}$$

是由正元素构成的 $k \times k$ 对称矩阵集合，并且

$$\mathcal{W}(z) = \mathcal{W}_1(z) \times \mathcal{W}_2(z) \times \cdots \times \mathcal{W}_k(z)$$

with

$$\mathcal{W}_a(z) := \{w \in (\mathbb{R}^+)^{n_a(z)} : \sum_i w_i = n_a(z)\} \tag{5.6}$$

这是社群 a 上可能的 w 集合。根据模型的定义，我们有， Θ_k 上的先验分布为

$$\nu_k(\theta) = \nu_k^{(1)}(\pi)\nu_k^{(2)}(\lambda)\nu^{(3)}(w|z) \tag{5.7}$$

其中

$$\nu_k^{(1)}(\pi) = \frac{\Gamma(k/2)}{\Gamma(1/2)^k} \prod_{1 \le a \le k} {\pi_a}^{-1/2},$$

$$\nu_k^{(2)}(\lambda) = \frac{1}{\Gamma(1/2)^{\frac{k(k+1)}{2}}} \prod_{1 \le a \le b \le k} \lambda_{ab}^{-1/2} e^{-\lambda_{ab}}$$

and

$$\nu^{(3)}(w|z) = \prod_{1 \le a \le k} \frac{\Gamma\left(\frac{n_a(z)}{2}\right)}{n_a(z)\Gamma\left(\frac{1}{2}\right)^{n_a(z)}} \prod_{i \colon z_i = a} \left(\frac{w_i}{n_a(z)}\right)^{-1/2}.$$

现在，我们可以将边际似然分解为

$$\begin{aligned} p_k(x) &= \sum_{z \in [k]^n} \int_{\Theta^k} p(x, z|\theta)\nu_k(\theta)d\theta \\ &= \sum_{z \in [k]^n} \int_{\mathcal{W}(z)} \int_{\Lambda^k} p(x|w, z, \lambda)\nu_k^{(2)}(\lambda)\nu^{(3)}(w|z)d\lambda dw \int_{\Delta^k} p(z|\pi)\nu_k^{(1)}(\pi)d\pi \\ &=: \sum_{z \in [k]^n} p_k(x|z)p_k(z) \end{aligned}$$

in 其中（条件）似然在（2.5）和（2.6）中给出。于是我们有

$$\begin{aligned}\int_{\Lambda^k} p(x|w,z,\lambda)\nu_k^{(2)}(\lambda)d\lambda \;&=\; \int_{\Lambda^k}\frac{1}{c(x)}\left[\prod_{1\le i\le n} {w_i}^{d_i(x)}\right]\left[\prod_{1\le a\le b\le k}\lambda_{ab}^{o_{ab}(x,z)}e^{-n_{ab}(z)\lambda_{ab}}\right]\nu_k^{(2)}(\lambda)d\lambda\\ &=\frac{1}{c(x)\Gamma(\frac{1}{2})^{\frac{k(k+1)}{2}}}\left[\prod_{1\le i\le n}{w_i}^{d_i(x)}\right]\int_{\Lambda^k}\prod_{1\le a\le b\le k}\lambda_{ab}^{o_{ab}(x,z)-\frac{1}{2}}e^{-(n_{ab}(z)+1)\lambda_{ab}}d\lambda\\ &=\frac{1}{c(x)}\underbrace{\frac{1}{\Gamma(\frac{1}{2})^{\frac{k(k+1)}{2}}}\left[\prod_{1\le a\le b\le k}\frac{\Gamma\left(o_{ab}(x,z)+\frac{1}{2}\right)}{[n_{ab}(z)+1]^{o_{ab}(x,z)+\frac{1}{2}}}\right]}_{A(x,z)}\left[\prod_{1\le i\le n}{w_i}^{d_i(x)}\right].\end{aligned} \tag{5.8}$$

因此

$$\begin{aligned}p_k(x|z)&=\frac{A(x,z)}{c(x)}\int_{\mathcal{W}}\prod_{1\le i\le n}{w_i}^{d_i(x)}\nu^{(3)}(w|z)dw\\ &=\frac{A(x,z)}{c(x)}\prod_{1\le a\le k}\frac{\Gamma\left(\frac{n_a(z)}{2}\right)}{n_a(z)\Gamma\left(\frac{1}{2}\right)^{n_a(z)}}\int_{\mathcal{W}_a(z)}\prod_{i:\ z_i=a}{w_i}^{d_i(x)}\Big(\frac{w_i}{n_a(z)}\Big)^{-1/2}dw_a\\ &=\frac{A(x,z)}{c(x)}\prod_{1\le a\le k}\frac{\Gamma\left(\frac{n_a(z)}{2}\right)n_a(z)^{d_a^t(x,z)-1}}{\Gamma\left(\frac{1}{2}\right)^{n_a(z)}}\int_{\mathcal{W}_a(z)}\prod_{i:\ z_i=a}\Big(\frac{w_i}{n_a(z)}\Big)^{d_i(x)-1/2}dw_a\\ &=\frac{A(x,z)}{c(x)}\prod_{1\le a\le k}\frac{\Gamma\left(\frac{n_a(z)}{2}\right)n_a(z)^{d_a^t(x,z)}}{\Gamma\left(\frac{1}{2}\right)^{n_a(z)}}\int_{\mathcal{Y}_a(z)}\prod_{i:\ z_i=a}{y_i}^{d_i(x)-1/2}dy\\ &=\frac{A(x,z)}{c(x)}\underbrace{\prod_{1\le a\le k}\frac{n_a(z)^{d_a^t(x,z)}\Gamma\left(\frac{n_a(z)}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^{n_a(z)}}\frac{\prod_{i:\ z_i=a}\Gamma(d_i(x)+\frac{1}{2})}{\Gamma(d_a^t(x,z)+\frac{n_a(z)}{2})}}_{B(x,z)}.\end{aligned} \tag{5.9}$$

我们还有

$$\begin{aligned}p_k(z)=&\int_{\Pi^k}\prod_{a=1}^k{\pi_a}^{n_a(z)}\nu_k^{(1)}(\pi)d\pi\\ =&\frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^k}\int_{\Pi^k}\prod_{a=1}^k{\pi_a}^{n_a(z)-\frac{1}{2}}d\pi\\ =&\underbrace{\frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^k}\frac{\prod_{a=1}^k\Gamma\left(n_a(z)+\frac{1}{2}\right)}{\Gamma\left(n+\frac{k}{2}\right)}}_{C(z)}.\end{aligned} \tag{5.10}$$

O另一方面，根据最大似然估计量的定义，我们有 t

$$\sup_{\lambda,w} p(x|z,w,\lambda) = \frac{1}{c(x)} \underbrace{\Big[\prod_{1\le a\le b\le k} \Big(\frac{o_{ab}(x,z)}{n_{ab}(z)} \Big)^{o_{ab}(x,z)} e^{-o_{ab}(x,z)} \Big]}_{\widehat{A}(x,z)} \times \underbrace{\Big[\prod_{a\in[k], i\in[n]: z_i=a} \Big(\frac{n_a(z) d_i(x)}{d_a^t(x,z)} \Big)^{d_i(x)} \Big]}_{\widehat{B}(x,z)} \tag{5.11}$$

and

$$\sup_{\pi} p(z|\pi) = \underbrace{\prod_{1\le a\le k_0} \Big(\frac{n_a(z)}{n} \Big)^{n_a(z)}}_{\widehat{C}(z)} . \tag{5.12}$$

现在，我们注意到，通过约去归一化常数 $c(x)$，可得

$$\frac{\sup_\theta p(x|\theta)}{p_k(x)} \;\le\; \frac{\sum_z \widehat{A}(x,z)\widehat{B}(x,z)\widehat{C}(z)}{\sum_z A(x,z)B(x,z)C(z)} . \tag{5.13}$$

现在，如果我们能够找到在 x 和 z 上一致的界 D_1、D_2 和 D_3，使得

$$\frac{\widehat{A}(x,z)}{A(x,z)} \;\le\; D_1 , \quad \frac{\widehat{B}(x,z)}{B(x,z)} \;\le\; D_2 \quad \text{and} \quad \frac{\widehat{C}(x,z)}{C(x,z)} \;\le\; D_3$$

那么我们便可自动得到

$$\frac{\sup_\theta p(x|\theta)}{p_k(x)} \;\le\; D_1\, D_2\, D_3 . \tag{5.14}$$

这些界　可由以下引理得到：5.3、5.4 和5.5 这些引理将在下文证明。利用这些引理，我们得到 at

$$\begin{aligned} \log \frac{\sup_\theta p(x|\theta)}{p_k(x)} \;&\le\; \log\Big[(n+1)^{k(k+1)} (n^2 \log n)^n n^k \Big] \\ &\le\; k(k+1)\log(n+1) + n(2\log n + \log\log n) + k\log n \\ &\le\; k(k+2)\log(n+1) + 3n\log n \end{aligned}$$

从而完成命题3.1的证明。

引理 5.3。对于 $(x,z) \in \Omega_n \times [k]^n$，我们有

$$\frac{\widehat{A}(x,z)}{A(x,z)} \;\le\; (n+1)^{k(k+1)} . \tag{5.15}$$

证明。 任取 $(x,z)\in\Omega_n\times[k]^n$。为简化记号，我们将 $o_{ab}=o_{ab}(x,z)$ 记作，并类似地将 $n_a(z)$ 和 $n_{ab}(z)$ 记作。由 (5.8) 和 (5.11) 可得

$$\begin{aligned}\frac{\widehat{A}(x,z)}{A(x,z)} &= \frac{\prod_{1\le a\le b\le k}\left(\frac{o_{ab}}{n_{ab}}\right)^{o_{ab}}e^{-o_{ab}}}{\frac{1}{\Gamma(\frac12)^{\frac{k(k+1)}{2}}}\prod_{1\le a\le b\le k}\frac{\Gamma\left(o_{ab}+\frac12\right)}{[n_{ab}+1]^{o_{ab}+\frac12}}}\\ &= \Gamma\left(\frac12\right)^{\frac{k(k+1)}{2}}\prod_{1\le a\le b\le k}\frac{\left(\frac{o_{ab}}{n_{ab}}\right)^{o_{ab}}e^{-o_{ab}}(n_{ab}+1)^{o_{ab}+\frac12}}{\Gamma\left(o_{ab}+\frac12\right)}.\end{aligned}\tag{5.16}$$

令

$$N:=\sum_{1\le a\le b\le k}o_{ab}$$

我们改写为

$$\frac{\widehat{A}(x,z)}{A(x,z)}=\Gamma\left(\frac12\right)^{\frac{k(k+1)}{2}}\prod_{1\le a\le b\le k}\frac{\left(\frac{o_{ab}}{N}\right)^{o_{ab}}}{\Gamma\left(o_{ab}+\frac12\right)}\left(\frac{N}{n_{ab}}\right)^{o_{ab}}(n_{ab}+1)^{o_{ab}+\frac12}e^{-o_{ab}}.\tag{5.17}$$

现在利用引理 5.1可得

$$\prod_{1\le a\le b\le k}\frac{\left(\frac{o_{ab}}{N}\right)^{o_{ab}}}{\Gamma\left(o_{ab}+\frac12\right)}\le\frac{1}{\Gamma\left(N+\frac12\right)\Gamma\left(\frac12\right)^{\frac{k(k+1)}{2}-1}}.\tag{5.18}$$

另一方面

$$\begin{aligned}\left(\frac{N}{n_{ab}}\right)^{o_{ab}}(n_{ab}+1)^{o_{ab}+\frac12}e^{-o_{ab}} &= N^{o_{ab}}\left(1+\frac{1}{n_{ab}}\right)^{o_{ab}}(n_{ab}+1)^{\frac12}e^{-o_{ab}}\\ &\le N^{o_{ab}}e^{\frac{o_{ab}}{n_{ab}}}(n_{ab}+1)^{\frac12}e^{-o_{ab}}.\end{aligned}\tag{5.19}$$

将 (5.18) 和 (5.19) 代入 (5.17，得到

$$\begin{aligned}\frac{\widehat{A}(x,z)}{A(x,z)} &\le \frac{\Gamma\left(\frac12\right)^{\frac{k(k+1)}{2}}}{\Gamma\left(N+\frac12\right)\Gamma\left(\frac12\right)^{\frac{k(k+1)}{2}-1}}\prod_{1\le a\le b\le k}N^{o_{ab}}e^{\frac{o_{ab}}{n_{ab}}}(n_{ab}+1)^{\frac12}e^{-o_{ab}}\\ &= \frac{\Gamma\left(\frac12\right)}{\Gamma\left(N+\frac12\right)}e^{-N}N^Ne^{\sum_{1\le a\le b\le k}\frac{o_{ab}}{n_{ab}}}\prod_{1\le a\le b\le k}(n_{ab}+1)^{\frac12}.\end{aligned}\tag{5.20}$$

对于任意实数 $r>0$, $\Gamma(r)\ge r^{r-1/2}e^{-r}\sqrt{2\pi}$，使得

$$\frac{e^{-N}N^N}{\Gamma\left(N+\frac12\right)}\le\left(1+\frac{1}{2N}\right)^{-N}e^{\frac12}\frac{1}{\sqrt{2\pi}}\le\frac{1}{\sqrt{2\pi}}.\tag{5.21}$$

此外，由于 $x\in\Omega_n$ ，我们有

$$\sum_{1\le a\le b\le k}\frac{o_{ab}}{n_{ab}}\le\sum_{1\le a\le b\le k}\frac{n_{ab}\log n}{n_{ab}}=\frac{k(k+1)}{2}\log n\tag{5.22}$$

and

$$\prod_{1\le a\le b\le k}(n_{ab}+1)^{\frac{1}{2}} \;\le\; \prod_{1\le a\le b\le k}\left(n^2+1\right)^{\frac{1}{2}} \;=\; \left(n^2+1\right)^{\frac{k(k+1)}{4}} \;\le\; (n+1)^{\frac{k(k+1)}{2}}. \tag{5.23}$$

将（5.21）、（5.22）和（5.23）代入（5.20）即可证明该引理。 □

引理 5.4. 对于 $(x,z)\in\Omega_n\times[k]^n$，我们有

$$\frac{\widehat{B}(x,z)}{B(x,z)} \;\le\; (n^2\log n)^n. \tag{5.24}$$

证明. 对于 $(x,z)\in\Omega_n\times[k]^n$，我们有

$$\begin{aligned}\frac{\widehat{B}(x,z)}{B(x,z)} &= \prod_{1\le a\le k}\frac{\prod_{i\in[n]:z_i=a}\left(\frac{n_a(z)d_i(x)}{d_a^t(x,z)}\right)^{d_i(x)}}{\frac{n_a(z)^{d_a^t(x,z)}\Gamma\left(\frac{n_a(z)}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^{n_a(z)}}\frac{\prod_{i:\ z_i=a}\Gamma(d_i(x)+\frac{1}{2})}{\Gamma(d_a^t(x,z)+\frac{n_a(z)}{2})}}\\ &= \prod_{1\le a\le k}\frac{\Gamma\left(\frac{1}{2}\right)^{n_a(z)}\Gamma(d_a^t(x,z)+\frac{n_a(z)}{2})}{\Gamma\left(\frac{n_a(z)}{2}\right)}\prod_{i:z_i=a}\frac{\left(\frac{d_i(x)}{d_a^t(x,z)}\right)^{d_i(x)}}{\Gamma(d_i(x)+\frac{1}{2})}\end{aligned}$$

但是，对于所有 $a\in[k]$，根据引理 5.1 可知

$$\prod_{i:z_i=a}\frac{\left(\frac{d_i(x)}{d_a^t(x,z)}\right)^{d_i(x)}}{\Gamma(d_i(x)+\frac{1}{2})} \;\le\; \frac{1}{\Gamma\left(d_a^t(x,z)+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)^{n_a(z)-1}}, \tag{5.25}$$

其中我们用到了等式 $d_a^t(x,z)=\sum_{i:\ z_i=a}d_i(x)$。综合以上所有界，我们得到

$$\frac{\widehat{B}(x,z)}{B(x,z)} \;\le\; \prod_{1\le a\le k}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma(d_a^t(x,z)+\frac{n_a(z)}{2})}{\Gamma\left(\frac{n_a(z)}{2}\right)\Gamma\left(d_a^t(x,z)+\frac{1}{2}\right)}.$$

最后，由引理 5.2 我们得出结论，因为 $d_a^t(x,\ z)\le n^2$ 对数 n 对于所有 $a\in[k]$ 以及 $\sum_a n_a(z)=n$ 可知

$$\frac{\widehat{B}(x,z)}{B(x,z)} \;\le\; \prod_{1\le a\le k}d_a^t(x,z)^{n_a(z)} \;\le\; (n^2\log n)^n\,.$$

□

引理 5.5。对于任意 $z\in[k]^n$，我们有

$$\frac{\widehat{C}(z)}{C(z)} \;\le\; n^k\,. \tag{5.26}$$

证明。根据定义，我们有

$$\frac{\widehat{C}(z)}{C(z)} = \frac{\prod_{a=1}^{k}\left(\frac{n_a(z)}{n}\right)^{n_a(z)}}{\frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^k}\frac{\prod_{a=1}^{k}\Gamma\left(n_a(z)+\frac{1}{2}\right)}{\Gamma\left(n+\frac{k}{2}\right)}}.$$

根据引理 5.1可得

$$\prod_{a=1}^{k}\frac{\left(\frac{n_a(z)}{n}\right)^{n_a(z)}}{\Gamma\left(n_a(z)+\frac{1}{2}\right)} \;\le\; \frac{1}{\Gamma\left(n+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)^{k-1}}. \tag{5.27}$$

根据引理 5.2 和（5.27），可得

$$\frac{\widehat{C}(z)}{C(z)} \;\le\; \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(n+\frac{k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(n+\frac{1}{2}\right)} \;\le\; n^k.$$

□

5.3 其他辅助引理的证明

引理 5.6。 设 $X=(X_{ij})_{i,\ j\in[n]}$ 由度修正随机块模型生成，并设 Ω_n 为 (3.5)中定义的集合。则 $X\in\Omega_n$ 随着 $n\to\infty$，最终几乎必然成立。

证明。我们有

$$\begin{aligned} p(\Omega_n^c) &= \sum_{z\in[k_0]^n} p(\Omega_n^c\cap\{Z=z\}) \\ &= \sum_{z\in[k_0]^n} p(z)p(\Omega_n^c|Z=z) \\ &\le \sum_{z\in[k_0]^n} p(z)\sum_{i,j=1}^{n} p(X_{ij}>\log n|Z=z). \end{aligned} \tag{5.28}$$

在给定 $Z=z$的条件下， X_{ij}服从参数为（$\lambda_{z_iz_j}$）的泊松分布。根据（Boucheron 等人，2013，第 2.2 节），若 $Y\sim$泊松(μ)，则 $r>0$

$$\begin{aligned} P(Y-\mu\ge r) &\le e^{-\mu\left(1+\frac{r}{\mu}\right)\log\left(1+\frac{r}{\mu}\right)+r} \\ &\le e^{-r\left(\log\left(1+\frac{r}{\mu}\right)-1\right)}, \end{aligned}$$

因此，对于满足 $\log n>\lambda_{z_iz_j}$的 n

$$\begin{aligned} p(X_{ij}\ge\log n|Z=z) &\le e^{-(\log n-\lambda_{z_iz_j})\left(\log\left(\frac{\log n}{\lambda_{z_iz_j}}\right)-1\right)} \\ &\le e^{-(\log n-\lambda_{\max})\left(\log\left(\frac{\log n}{\lambda_{\max}}\right)-1\right)} \end{aligned} \tag{5.29}$$

其中在最后一行我们使用了如下记号 $\lambda_{\max} = \max_{a,\, b} \lambda_{ab}$。然后，根据（5.28）和（5.29），我们有

$$p(\Omega_n^c) \le\ n^2 e^{-(\log n - \lambda_{\max})\left(\log\left(\frac{\log n}{\lambda_{\max}}\right)-1\right)} \sum_{z\in[k_0]^n} p(z) = n^{3+\log\lambda_{\max}-\log(\log n)}\,(\log n)^{\lambda_{\max}}\,(e\lambda_{\max})^{-\lambda_{\max}}$$

该式关于 n 可求和，因此根据 Borel–Cantelli引理，当 $n\to\infty$时，我们几乎必然有 $X \in \Omega_n$ 最终成立。□

为陈述下一个辅助结果，我们需要引入一些记号。对所有 $z \in \{1,\dots,k\}^n$ 和 $z^0 \in \{1,\dots,k_0\}^n$，定义矩阵 $k \times k_0$ $Q_n(z, z^0)$ 如下：

$$[Q_n(z,z^0)]_{aa'} \ =\ \frac{1}{n}\sum_{i=1}^n w_i\,\mathbb{1}\{z_i = a, z_i^0 = a'\}. \tag{5.30}$$

注意，对于 $a' \in [k_0]$，计数器 $n_a(z^0)$ 可写为

$$n_{a'}(z^0) = \sum_{i=1}^n w_i\,\mathbb{1}\{z_i^0 = a'\} = \sum_{i=1}^n w_i \sum_{a=1}^k \mathbb{1}\{z_i = a, z_i^0 = a'\}\,. \tag{5.31}$$

于是，$n_{a'}(z^0) = n(Q_n^T(z,z^0)\mathbf{1}_k)_{a'}$，其中 $\mathbf{1}_k$是一个维数为 k且所有元素均为 1 的列向量。此外，矩阵 $Q_n(z,z^0)$ 满足

$$\|Q_n(z,z^0)\|_1 \ =\ \sum_{a=1}^k\sum_{a'=1}^{k_0}[Q_n(z,z^0)]_{a,a'} = 1, \tag{5.32}$$

对所有 (z,z^0)，以及

$$\mathbb{E}(\tilde{o}_{ab}(x,z)\,|\,Z = z^0) \ =\ n^2 Q_n(z,z^0)\lambda Q_n(z,z^0)^T \ =\ \rho_n n^2 Q_n(z,z^0)\tilde{\lambda}Q_n(z,z^0)^T\,. \tag{5.33}$$

现在，我们证明 $\tilde{o}_{ab}(x,z)$ 在给定 $Z = z^0$的条件下的一个集中界；它是若干独立泊松随机变量之和。

引理 5.7. 对于任意 $\epsilon > 0$ 和 $a,\ b \in [k]$，我们有

$$\mathbb{P}\Big(\sup_{\bar{z}\in[k]^n}\Big|\frac{\tilde{o}_{ab}(X,\bar{z})}{\rho_n n^2} - [Q_n(\bar{z},Z)\tilde{\lambda}Q_n(\bar{z},Z)^T]_{ab}\Big| > \epsilon\Big) \ \le\ \exp\Big(-\frac{\rho_n n^2\epsilon^2}{\tilde{\lambda}_{\max}+\epsilon} + n\log k\Big)$$

and

$$\mathbb{P}\Big(\sup_{\bar{z}\in[k]^n}\Big|\frac{d_a^t(X,\bar{z})}{\rho_n n^2} - [Q_n(\bar{z},Z)\tilde{\lambda}Q_n(\bar{z},Z)^T\mathbf{1}_k]_a\Big| \ >\ \epsilon\Big) \ \le\ \exp\Big(-\frac{\rho_n n^2\epsilon^2}{\tilde{\lambda}_{\max}+\epsilon} + n\log k\Big)\,,$$

其中 $\tilde{\lambda}_{\max} = \max_{ab}\lambda_{ab}$。

证明。对于任意固定的 $z \in [k_0]^n$ 和 $\bar{z} \in [k]^n$，我们有

$$\begin{aligned}
&\tilde{o}_{ab}(X,\bar{z}) - n^2[Q_n(\bar{z},z)\lambda Q_n(\bar{z},z)^T]_{ab}\\
&\qquad = \tilde{o}_{ab}(X,\bar{z}) - \rho_n n^2[Q_n(\bar{z},z)\tilde{\lambda}Q_n(\bar{z},z)^T]_{ab}\\
&\qquad = \sum_{1\le i,j\le n}\ \sum_{1\le a',b'\le k_0}(X_{ij} - \rho_n w_i w_j\tilde{\lambda}_{a'b'})\mathbb{1}\{\bar{z}_i = a, z_i = a'\}\mathbb{1}\{\bar{z}_j = b, z_j = b'\}\,.
\end{aligned}$$

注意，给定 $Z = z$， $\tilde{o}_{ab}(X, \bar{z})$ 对应于以下 $n_a(\bar{z})n_b(\bar{z})$ 个独立泊松˜随机变量的和，它们由 $X_{ij}\mathbb{1}\{\bar{z}_i = a, \bar{z}_i = b\}$给出，期望值为 $\rho_n w_i w_j \lambda_{z_i z_j}$。因此，该和也服从泊松分布，其参数为相应参数之和。再次利用（Boucheron 等人，2013，第 2.2 节），对于 $Y \sim$ 泊松(μ) 和 $t > 0$

$$P(Y - \mu \geq r) \leq e^{-\mu\left(1+\frac{r}{\mu}\right)\log\left(1+\frac{r}{\mu}\right)+r}$$
$$P(Y - \mu \leq -r) \leq e^{-\mu\left(1-\frac{r}{\mu}\right)\log\left(1-\frac{r}{\mu}\right)-r}$$

经过一些代数运算后，得到

$$\mathbb{P}(|X - \mu| > r) \;\leq\; 2e^{-\frac{r^2}{2(\mu+r)}}.$$

因此，对于任意 $\delta > 0$

$$\mathbb{P}\Big(\Big|\tilde{o}_{ab}(X, \bar{z}) - \rho_n n^2 [Q_n(\bar{z}, z)\tilde{\lambda}Q_n(\bar{z}, z)^T]_{ab}\Big| \;>\; \delta \,|Z = z\Big)$$
$$\leq\; 2\exp\Big(-\frac{\delta^2}{2(\rho_n n^2 [Q_n(\bar{z}, z)\tilde{\lambda}Q_n(\bar{z}, z)^T]_{ab} + \delta)}\Big).$$

由于对于任意 $\bar{z}$ 和 z，我们有

$$\rho_n n^2 [Q_n(\bar{z}, z)\tilde{\lambda}Q_n(\bar{z}, z)^T]_{ab} \;\leq\; \rho_n n^2 \tilde{\lambda}_{\max}$$

˜ ˜满足 $\lambda_{\max} = \max_{ab} \lambda_{ab}$，因此，对于任意 z、$\bar{z}$ 和 $\epsilon > 0$

$$\mathbb{P}\Big(\Big|\frac{\tilde{o}_{ab}(X, \bar{z})}{\rho_n n^2} - [Q_n(\bar{z}, z)\tilde{\lambda}Q_n(\bar{z}, z)^T]_{ab}\Big| \;>\; \epsilon \,|\, Z = z\Big) \;\leq\; \exp\Big(-\frac{\rho_n^2 n^4 \epsilon^2}{\rho_n n^2 \tilde{\lambda}_{\max} + \epsilon\rho_n n^2}\Big)$$
$$\leq\; \exp\Big(-\frac{\rho_n n^2 \epsilon^2}{\tilde{\lambda}_{\max} + \epsilon}\Big).$$

现在，对所有 $\bar{z} \in [k]^n$ 应用并集界，并对 z 积分，可得

$$\mathbb{P}\Big(\sup_{\bar{z}\in[k]^n}\Big|\frac{\tilde{o}_{ab}(X, \bar{z})}{\rho_n n^2} - [Q_n(\bar{z}, Z)\tilde{\lambda}Q_n(\bar{z}, Z)^T]_{ab}\Big| \;>\; \epsilon\Big) \;\leq\; \exp\Big(-\frac{\rho_n n^2 \epsilon^2}{\tilde{\lambda}_{\max} + \epsilon} + n\log k\Big)$$

这就证明了引理的第一个不等式。现在，给定 $Z = z$，

$$d_a^t(X, \bar{z}) \;=\; \sum_{b\in[k]} \tilde{o}_{ab}(X, \bar{z})$$

也可以表示为若干相互独立的泊松分布随机变量之和，并且

$$\mathbb{E}(d_a^t(X, \bar{z}) \,|\, Z = z) \;=\; [\rho_n n^2 Q_n(\bar{z}, z)\tilde{\lambda}Q_n(\bar{z}, z)^T \mathbf{1}_k]_a,$$

因此，我们也得到

$$\mathbb{P}\Big(\sup_{\bar{z}\in[k]^n}\Big|\frac{d_a^t(X, \bar{z})}{\rho_n n^2} - [Q_n(\bar{z}, Z)\tilde{\lambda}Q_n(\bar{z}, Z)^T \mathbf{1}_k]_a\Big| \;>\; \epsilon\Big) \;\leq\; \exp\Big(-\frac{\rho_n n^2 \epsilon^2}{\tilde{\lambda}_{\max} + \epsilon} + n\log k\Big).$$

至此，引理的证明告一段落5.7 □

下文将陈述并证明命题证明中引用的引理4.1。

引理 5.8. 对于 $k < k_0$，存在一个 $k \times k$ 正矩阵 λ^* 和一个 k 维向量 π^*，使得

$$\begin{aligned}\limsup_{n\to\infty}\ \frac{1}{2}\sum_{1\le a,b\le k}\frac{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4}\,\varphi\Big(\frac{\rho_n n^2\tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\Big)\\ \le\ \frac{1}{2}\sum_{1\le a,b\le k}\pi_a^*\pi_b^*[\lambda^*\pi^*]_a[\lambda^*\pi^*]_b\,\varphi\Big(\frac{\lambda_{ab}^*}{[\lambda^*\pi^*]_a[\lambda^*\pi^*]_b}\Big)\,.\end{aligned}\tag{5.34}$$

此外，(π^*, λ^*) 由下式给出

$$\begin{aligned}\pi_a^* &= [R^*\mathbf{1}_{k_0}]_a, \qquad a\in\{1,\dots,k\}\\ \lambda_{ab}^* &= \frac{[R^*\lambda {R^*}^T]_{ab}}{[R^*\mathbf{1}_{k_0}\mathbf{1}_{k_0}^T {R^*}^T]_{ab}}\,, \qquad a,b\in\{1,\dots,k\}\,,\end{aligned}\tag{5.35}$$

对于满足 $\|R^*\|_1 = 1$ 的 $k \times k_0$ 实矩阵 R^*，且每一列恰有一个非零元素。

证明. 注意到

$$\begin{aligned}&\sum_{1\le a,b\le k}\frac{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4}\varphi\Big(\frac{\rho_n n^2\tilde{o}_{ab}(x,\widehat{z}_k)}{\rho_n\, d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\Big)\\ &\qquad=\sum_{1\le a,b\le k}\frac{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4}\varphi\left(\frac{\tilde{o}_{ab}(x,\widehat{z}_k)/\rho_n n^2}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)/\rho_n^2 n^4}\right),\end{aligned}\tag{5.36}$$

其中 $\varphi(u) = u\log u$。于是由引理 5.7，取 $\epsilon_n = \rho_n = \log_n{}^n$，可得

$$\Big|\frac{\tilde{o}_{ab}(x,\widehat{z}_k)}{\rho_n n^2} - [Q_n(\widehat{z}_k,z)\tilde{\lambda}Q_n^T(\widehat{z}_k,z)]_{ab}\Big|\ \le\ \epsilon_n$$

类似地

$$\Big|\frac{d_a^t(x,\widehat{z}_k)}{\rho_n n^2} - [Q_n(\widehat{z}_k,z)\tilde{\lambda}Q_n(\widehat{z}_k,z)^T\mathbf{1}_k]_a\Big|\ \le\ \epsilon_n$$

当 $n\to\infty$ 时最终几乎必然成立。然后，由于 φ 连续，将 (5.36) 右侧中的 $\tilde{o}_{ab}(x,\widehat{z}_k)/\rho_n n^2$ 替换为 $[Q_n(\widehat{z}_k,z)\tilde{\lambda}Q_n^T(\widehat{z}_k,z)]_{ab}$，并将 $d_a^t(x,\widehat{z}_k)/\rho_n n^2$ 替换为 $[Q_n(\widehat{z}_k,z)\tilde{\lambda}Q_n(\widehat{z}_k,z)^T\mathbf{1}_k]_a$，我们得到

$$\begin{aligned}&\sum_{1\le a,b\le k}\frac{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4}\,\varphi\Big(\frac{\rho_n n^2\tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k)d_b^t(x,\widehat{z}_k)}\Big)\\ &\qquad\le\ \sup_{\substack{Q_n:\,\|Q_n\|_1=1\\ Q_n^T\mathbf{1}_k=n(z)/n}}\ \sum_{1\le a,b\le k}[Q_n\tilde{\lambda}Q_n^T\mathbf{1}_k]_a[Q_n\tilde{\lambda}Q_n^T\mathbf{1}_k]_b\,\varphi\Big(\frac{[Q_n\tilde{\lambda}Q_n^T]_{ab}}{[Q_n\tilde{\lambda}Q_n^T\mathbf{1}_k]_a[Q_n\tilde{\lambda}Q_n^T\mathbf{1}_k]_b}\Big)+\eta_n\,.\end{aligned}\tag{5.37}$$

对于 序列 $\eta_n \to 0$ 当 $n \to \infty$ 时。然后，对两边取上极限，我们必有

$$\limsup_{n\to\infty} \sum_{1\le a,b\le k} \frac{d_a^t(x,\widehat{z}_k) d_b^t(x,\widehat{z}_k)}{\rho_n^2 n^4} \varphi\Big(\frac{\rho_n n^2 \tilde{o}_{ab}(x,\widehat{z}_k)}{d_a^t(x,\widehat{z}_k) d_b^t(x,\widehat{z}_k)}\Big)$$
$$\le \sup_{\substack{R:\ \|R\|_1=1 \\ R^T\mathbf{1}_k=\pi}} \frac{1}{2} \sum_{1\le a,b\le k} [R\tilde{\lambda}R^T\mathbf{1}_k]_a [R\tilde{\lambda}R^T\mathbf{1}_k]_b \, \varphi\Big(\frac{[R\tilde{\lambda}R^T]_{ab}}{[R\tilde{\lambda}R^T\mathbf{1}_k]_a [R\tilde{\lambda}R^T\mathbf{1}_k]_b}\Big) \tag{5.38}$$

几乎必然成立。因此，右侧（5.38）中的上确界是凸多面体上的凸函数的最大值，该凸多面体由 $\{R\colon \ \|R\|_1 = 1, \ R^T\mathbf{1}_k = \pi\}$定义。于是，该最大值必定在多面体的某个顶点处取得；也就是说，在满足以下条件的矩阵上取得：每一列恰有一个元素大于零，且 $R \ \ \pi_a > 0$ 对所有 $a \in \{1,\dots,k_0\}$均成立。我们将这些最大值（若不止一个）之一记为 R^*，并令

$$\begin{aligned} \pi_a^* &= [R^*\mathbf{1}_{k_0}]_a \qquad a \in \{1,\dots,k\} \\ \lambda_{ab}^* &= \frac{[R^*\tilde{\lambda}R^{*^T}]_{ab}}{[R^*\mathbf{1}_{k_0}\mathbf{1}_{k_0}^T R^{*^T}]_{ab}}, \qquad a,b \in \{1,\dots,k\}. \end{aligned} \tag{5.39}$$

Then

$$\sup_{\substack{R:\ \|R\|_1=1 \\ R^T\mathbf{1}_k=\pi}} \frac{1}{2} \sum_{1\le a,b\le k} [R\tilde{\lambda}R^T\mathbf{1}_k]_a [R\tilde{\lambda}R^T\mathbf{1}_k]_b \, \varphi\Big(\frac{[R\tilde{\lambda}R^T]_{ab}}{[R\tilde{\lambda}R^T\mathbf{1}_k]_a [R\tilde{\lambda}R^T\mathbf{1}_k]_b}\Big)$$
$$= \frac{1}{2} \sum_{1\le a,b\le k} \pi_a^*\pi_b^* [\lambda^*\pi^*]_a [\lambda^*\pi^*]_b \, \varphi\Big(\frac{\lambda_{ab}^*}{[\lambda^*\pi^*]_a [\lambda^*\pi^*]_b}\Big). \tag{5.40}$$

引理 5.8 的证明至此完成。 □

˜引理 5.9。 假设 λ 没有两列成比例。那么对所有 $k < k_0$ 以及 $(\pi^*, \ \lambda^*)$ 如引理*5.8* 中所述，我们有

$$\sum_{1\le a,b\le k_0} \pi_a\pi_b [\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b \, \varphi\Big(\frac{\tilde{\lambda}_{ab}}{[\tilde{\lambda}\pi]_a [\tilde{\lambda}\pi]_b}\Big)$$
$$- \sum_{1\le a,b\le k} \pi_a^*\pi_b^* [\lambda^*\pi^*]_a [\lambda^*\pi^*]_b \, \varphi\Big(\frac{\lambda_{ab}^*}{[\lambda^*\pi^*]_a [\lambda^*\pi^*]_b}\Big) \ > \ 0. \tag{5.41}$$

证明。首先考虑 $k = k_0 - 1$ 的情况。由于 R^* 的每一列中有且仅有一个非零元素，因此存在一个满射函数 $h\colon [k_0] \to [k]$，将 $[k_0]$ 中的每个社区（ R^* 的列）与 $[k]$ 中对应的社区（含有非零元素的行）连接起来。于是，对于 $k = k_0 - 1$， $\{1,\dots,k_0\}$ 中有 $k-1$ 个社区映射到 $\{1,\dots,k\}$ 中的 $k-1$ 个社区，而 $\{1,\dots,k_0\}$ 中的两个社区映射到 $\{1,\dots,k\}$ 中的同一个社区。不失一般性地，假设社区 $k_0 - 1$ 和 k_0 满足 $h\,(k_0 - 1) \ = h\,(k_0) \ = k = k_0 - 1$。此外，由于 $R^{*T}\mathbf{1}_k = \pi$，我们必须有非零元素为

$$\begin{aligned} R_{aa}^* &= \pi_a, \qquad 1 \le a \le k_0 - 1 \\ R_{(k_0-1)k_0}^* &= \pi_{k_0}. \end{aligned} \tag{5.42}$$

然后，参数 π^* 和 λ^* 在（5.39）中定义为

$$\begin{aligned} \pi^*_a &= \pi_a\,, \qquad 1 \le a \le k_0 - 1 \\ \pi^*_{k_0-1} &= \pi_{k_0-1} + \pi_{k_0} \end{aligned}$$

and

$$\begin{aligned} \lambda^*_{ab} &= \tilde\lambda_{ab}\,, \qquad 1 \le a, b \le k_0 - 2 \\ \lambda^*_{a(k_0-1)} &= \frac{\pi_{k_0-1}\tilde\lambda_{a(k_0-1)} + \pi_{k_0}\tilde\lambda_{ak_0}}{\pi_{k_0-1} + \pi_{k_0}}\,, \qquad 1 \le l \le k_0 - 2, \\ \lambda^*_{(k_0-1)(k_0-1)} &= \frac{\pi^2_{k_0-1}\tilde\lambda_{(k_0-1)(k_0-1)} + 2\pi_{k_0-1}\pi_{k_0}\tilde\lambda_{(k_0-1)k_0} + \pi^2_{k_0}\tilde\lambda_{k_0k_0}}{\pi^2_{k_0-1} + 2\pi_{k_0-1}\pi_{k_0} + \pi^2_{k_0}}\,. \end{aligned}$$

˜注意，对于所有 $1 \le a \le k_0 - 1$，我们有 $[\lambda^*\pi^*]_a = [\lambda\pi]_a$，因此对于所有 $1 \le a,\ b \le k_0 - 2$

$$\pi^*_a\pi^*_b[\lambda^*\pi^*]_a[\lambda^*\pi^*]_b\ \varphi\!\left(\frac{\lambda^*_{ab}}{[\lambda^*\pi^*]_a[\lambda^*\pi^*]_b}\right) \;=\; \pi_a\pi_b[\tilde\lambda\pi]_a[\tilde\lambda\pi]_b\ \varphi\!\left(\frac{\tilde\lambda_{ab}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_b}\right).$$

另一方面，我们有

$$[\lambda^*\pi^*]_{k_0-1} = \frac{\pi_{k_0-1}[\tilde\lambda\pi]_{k_0-1} + \pi_{k_0}[\tilde\lambda\pi]_{k_0}}{\pi_{k_0-1} + \pi_{k_0}}\,.$$

那么对于 $1 \le a \le k_0 - 2$，由对数和不等式可知，

$$\begin{aligned} &\pi^*_a\pi^*_{k_0-1}[\lambda^*\pi^*]_a[\lambda^*\pi^*]_{k_0-1}\varphi\!\left(\frac{\lambda^*_{a(k_0-1)}}{[\lambda^*\pi^*]_a[\lambda^*\pi^*]_{k_0-1}}\right) \\ &\quad= \pi_a[\tilde\lambda\pi]_a(\pi_{k_0-1}[\tilde\lambda\pi]_{k_0-1} + \pi_{k_0}[\tilde\lambda\pi]_{k_0})\varphi\left(\frac{\pi_{k_0-1}\tilde\lambda_{a(k_0-1)} + \pi_{k_0}\tilde\lambda_{ak_0}}{[\tilde\lambda\pi]_a(\pi_{k_0-1}[\tilde\lambda\pi]_{k_0-1} + \pi_{k_0}[\tilde\lambda\pi]_{k_0})}\right) \\ &\quad= \pi_a(\pi_{k_0-1}\tilde\lambda_{a(k_0-1)} + \pi_{k_0}\tilde\lambda_{ak_0})\log\left(\frac{\pi_a(\pi_{k_0-1}\tilde\lambda_{a(k_0-1)} + \pi_{k_0}\tilde\lambda_{ak_0})}{\pi_a[\tilde\lambda\pi]_a(\pi_{k_0-1}[\tilde\lambda\pi]_{k_0-1} + \pi_{k_0}[\tilde\lambda\pi]_{k_0})}\right) \qquad (5.43) \\ &\quad\le \pi_a\pi_{k_0-1}\tilde\lambda_{a(k_0-1)}\log\!\left(\frac{\tilde\lambda_{a(k_0-1)}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0-1}}\right) + \pi_a\pi_{k_0}\tilde\lambda_{ak_0}\log\!\left(\frac{\tilde\lambda_{ak_0}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0}}\right) \\ &\quad= \pi_a\pi_{k_0-1}[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0-1}\varphi\!\left(\frac{\tilde\lambda_{a(k_0-1)}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0-1}}\right) + \pi_a\pi_{k_0}[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0}\varphi\!\left(\frac{\tilde\lambda_{ak_0}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0}}\right). \end{aligned}$$

此外，除非.....，否则该不等式必定是严格的

$$\frac{\tilde\lambda_{a(k_0-1)}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0-1}} \;=\; \frac{\tilde\lambda_{ak_0}}{[\tilde\lambda\pi]_a[\tilde\lambda\pi]_{k_0}} \qquad \text{for all } a \le k_0 - 2\,. \qquad (5.44)$$

在.....上 另一方面，对于 $a = k_0 - 1$ 和 $b = k_0 - 1$，同样根据对数和不等式，我们有 t

$$
\begin{aligned}
&\pi^*_{k_0-1}\pi^*_{k_0-1}[\lambda^*\pi^*]_{k_0-1}[\lambda^*\pi^*]_{k_0-1}\varphi\Big(\frac{\lambda^*_{(k_0-1)(k_0-1)}}{[\lambda^*\pi^*]_{k_0-1}[\lambda^*\pi^*]_{k_0-1}}\Big)\\
&= (\pi^2_{k_0-1}\tilde\lambda_{(k_0-1)(k_0-1)} + 2\pi_{k_0-1}\pi_{k_0}\tilde\lambda_{(k_0-1)k_0} + \pi^2_{k_0}\tilde\lambda_{k_0k_0})\\
&\qquad\qquad \times \log\Big(\frac{\pi^2_{k_0-1}\tilde\lambda_{(k_0-1)(k_0-1)} + 2\pi_{k_0-1}\pi_{k_0}\tilde\lambda_{(k_0-1)k_0} + \pi^2_{k_0}\tilde\lambda_{k_0k_0}}{\pi^2_{k_0-1}[\tilde\lambda\pi]^2_{k_0-1} + 2\pi_{k_0-1}\pi_{k_0}[\tilde\lambda\pi]_{k_0-1}[\tilde\lambda\pi]_{k_0} + \pi^2_{k_0}[\tilde\lambda\pi]^2_{k_0}}\Big)\\
&\le\ \pi^2_{k_0-1}\tilde\lambda_{(k_0-1)(k_0-1)}\log\Big(\frac{\tilde\lambda_{(k_0-1)(k_0-1)}}{[\tilde\lambda\pi]^2_{k_0-1}}\Big) + 2\pi_{k_0-1}\pi_{k_0}\tilde\lambda_{(k_0-1)k_0}\log\Big(\frac{\tilde\lambda_{(k_0-1)k_0}}{[\tilde\lambda\pi]_{k_0-1}[\tilde\lambda\pi]_{k_0}}\Big)\\
&\qquad + \pi^2_{k_0}\tilde\lambda_{k_0k_0}\log\Big(\frac{\tilde\lambda_{k_0k_0}}{[\tilde\lambda\pi]^2_{k_0}}\Big)\\
&= \pi^2_{k_0-1}[\tilde\lambda\pi]^2_{k_0-1}\varphi\Big(\frac{\tilde\lambda_{(k_0-1)(k_0-1)}}{[\tilde\lambda\pi]^2_{k_0-1}}\Big) + 2\pi_{k_0-1}\pi_{k_0}[\tilde\lambda\pi]_{k_0-1}[\tilde\lambda\pi]_{k_0}\varphi\Big(\frac{\tilde\lambda_{(k_0-1)k_0}}{[\tilde\lambda\pi]_{k_0-1}[\tilde\lambda\pi]_{k_0}}\Big)\\
&\qquad + \pi^2_{k_0}[\tilde\lambda\pi]^2_{k_0}\varphi\Big(\frac{\tilde\lambda_{k_0k_0}}{[\tilde\lambda\pi]^2_{k_0}}\Big),
\end{aligned}
\tag{5.45}
$$

当且仅当.....时取等号

$$
\frac{\tilde\lambda_{(k_0-1)(k_0-1)}}{[\tilde\lambda\pi]^2_{k_0-1}} \;=\; \frac{\tilde\lambda_{(k_0-1)k_0}}{[\tilde\lambda\pi]_{k_0-1}[\tilde\lambda\pi]_{k_0}} \;=\; \frac{\tilde\lambda_{k_0k_0}}{[\tilde\lambda\pi]^2_{k_0}}\,. \tag{5.46}
$$

由（5.44）和（5.46）可知，不等式（5.41）必为严格不等式，除非

$$
\tilde\lambda_{a(k_0-1)} \;=\; \frac{[\tilde\lambda\pi]_{k_0-1}}{[\tilde\lambda\pi]_{k_0}}\tilde\lambda_{ak_0} \qquad \text{for all } a \le k_0 \tag{5.47}
$$

这与 k_0 可辨识性的假设矛盾。 □