arXiv:2302.00242v1 [stat.ML] 1 Feb 2023

充分分离的球形高斯混合的参数稳定性

Hanyu Zhang
统计系，华盛顿大学，
西雅图，WA 98115
hanuyz6@uw.edu

Marina Meil˘a
统计系，华盛顿大学，西雅
图，WA 98115
mmp@stat.washington.edu

摘要

我们量化了球形高斯混合模型（sGMM）在分布空间中受到小扰动时的参数稳定性。具体而言，我们推导出首个显式界，表明对于预先定义的模型类中的球形高斯 P混合分布（sGMM），该模型类中所有与 P 的全变差距离较小的其他 sGMM，其参数距离也较小。进一步说，这个上界仅取决于 P。这项工作的动机是为拟合高斯混合分布提供保证；基于这一目标，所有涉及的常数都有明确定义，并给出了拟合球形高斯分布混合模型的分布无关条件。我们的结果显著收紧了现有的可计算界，并在渐近意义上达到该问题已知的尖锐阈值。

1 引言

我们研究将球形高斯混合模型（sGMM）拟合到未知分布 Q的问题。在不了解目标 Q的情况下，对于估计得到的 sGMM，我们能给出何种保证？又在什么条件下才能给出保证?

先前的研究（例如，[6, 27, 1, 20], 等）在对数据源 Q （例如，它是由充分分离的高斯分布构成的混合分布）作出模型假设的前提下，确立了混合参数的估计保证。具体而言，就本文讨论的范围来说，当 Q 是由球面对称高斯分布构成的混合分布，并满足以下两个条件时：(i) 分量比例非零，且 (ii) 分量间距足够大，就存在多项式运行时间的估计算法。
我们的研究提出了这样一个问题：如果没有对 Q如此精确的了解，我们能得出什么结论？作为替代，我们假设对 Q有间接了解，即一个充分分离的混合模型 P已拟合到 Q，并且模型拟合良好。主要区别在于，现在不要求 Q属于该模型类，而只要求它与该模型类“接近”，具体定义将在下文给出（本文中，模型拟合程度由 P与 Q之间的全变差距离衡量， $TV(P,\ Q)$）。我们还假设 P是由充分分离的高斯分布构成的混合分布，且分量比例有正的下界。事实证明，知道Q 接近某个“良好”的模型类，几乎与知道 Q 属于该模型类一样有用。在这些关于 P和 Q的条件下，我们证明 P的参数在扰动范围内唯一，并给出了该范围的上界。总之，我们旨在证明如下形式的陈述。

定理1 (一般 (ϵ, δ)–稳定性)。 对于分布 Q 和一类球形高斯混合模型 $\mathcal{M}$，如果给定的模型$P \in \mathcal{M}$满足 $TV(Q, P) \leq \epsilon$，那么对于任意满足 $TV(Q, P') \leq \epsilon$的 $P' \in \mathcal{M}$，都有$d_{param}(P, P') \leq \delta$，其中 $d_{param}(\cdot, \cdot)$是参数空间中定义的一种散度。

此前， [23] 在基于损失的聚类语境下提出过这类论述，但迄今为止，它尚未在任何参数模型拟合问题中得到具体化。

预印本。审稿中。

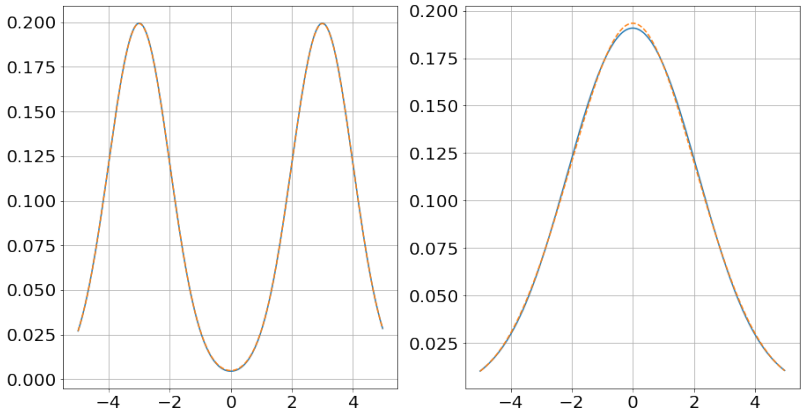


图 1：一维稳定与不稳定的球形高斯混合。左：由 $K = 2$ 个高斯分布组成、分离良好的混合分布 $P = 0.5N(-3, 1) + 0.5N(3, 1)$，其分离度为 $c = 3$（定义见第 2 节，A3），叠加在一个分布 Q 上，使得 $TV(P, Q) \leq 0.001$。第 2 节中的定理 4 保证，任何由 P' 个分量构成、最小权重为 0.45 、间隔常数为 3，且拟合 Q 的效果同样好的混合分布，其参数都必须接近 P；具体而言（参数定义见第 2 节），μ'_i 与 -3 和 3 的距离在 0.0200 以内，$\max\{1/\sigma_i'^2, \sigma_i'^2\} \leq 1.034$，且 $|0.5 - \pi'_i| \leq 0.004$。右：两个不稳定的球形高斯混合分布 P、P'；它们在全变差距离上很接近，但参数却大不相同；$P = 0.0625N(-3, \sigma^2) + 0.4375N(-1, \sigma^2) + 0.4375N(1, \sigma^2) + 0.0625N(3, \sigma^2)$，$P' = 0.0078125N\ -4\ \sigma^2 + 0.21875N\ -2\ \sigma^2 + 0.546875N\ 0\ \sigma^2 + 0.21875N\ 2\ \sigma^2 + (,)(,)(,)0.0078125N\ 4\ \sigma^2$ $\sigma^2 = 2.25$)(,)(,)，其中 。在这个例子中，混合成分的分离度较低，而且部分混合比例也较小。

满足定理1的分布 P 称为稳定。显然，这样的分布在 $\mathcal{M}$中通常并非唯一，也没有什么特别之处，因此稳定性应定义为 $\mathcal{M}$（的某个子集）的性质。将 $Q = P^* \in \mathcal{M}$代入也能得出这一结论。因此，我们将模型类的参数稳定性定义如下。

定义 2。设 $\mathcal{M}$ 为球面高斯混合模型类，$d_{param}(\cdot, \cdot)$ 为参数之间的散度；对于足够小的 $\epsilon > 0$，若任意满足 $d_{TV}(P, P') \leq 2\epsilon$ 的模型 $P' \in \mathcal{M}$ 均满足 $d_{param}(P, P') \leq \delta$，则称 P在 $\mathcal{M}$ 中是(ϵ, δ)–稳定的。

本文的技术贡献在于为 $\mathcal{M}$ 提出条件，并在总体情形下为任意 $P \in \mathcal{M}$建立上界$\delta(\epsilon, \mathcal{M})$。这些结果见定理 4。我们得到的 $\delta(\epsilon, \mathcal{M})$ 上界易于计算，其显式常数仅取决于模型类。作为一种统计程序，界 $\delta(\epsilon, \mathcal{M})$ 可在无需任何先验知识的情况下，为高斯混合模型的拟合提供定量的估计后评估或诊断分析。例如，对于定义在 $\mathbb{R}^d$ 上的任意总体 Q（密度为 q），设 $\hat{P} = \sum_{i=1}^{K} \hat{\pi}_i N_d(\hat{\mu}_k, \hat{\sigma}_k^2 I_d)$ 是一个学习得到的球形高斯混合模型。图 1 展示了稳定和不稳定分布的球形高斯混合模型示例；此外，对于稳定的球形高斯混合模型，图中还构造了良好拟合区域，使得所有与总体的全变差距离较小的球形高斯混合模型，其参数都必须位于这些区域内。这些区域可以像常规参数估计问题中的置信区域那样报告。

此外，定义 2 仅涉及模型类 $\mathcal{M}$，它是球形高斯混合的一个子集。因此，通过研究模型类内分布的稳定性（例如球形高斯混合模型中的分布），通常即可获得分布无关稳定性结果。虽然这一观察几乎显而易见，但令人意外的是，很少有研究关注这样一种可能性：将分布无关保证作为一致性或可识别性证明的副产品来获得。我们希望本文的贡献之一体现在概念层面，即让人们注意到这种可能性；这仍是本研究的主要动机。

在第 2 节中，我们定义感兴趣的模型类 $\mathcal{M}$，并具体设定 $d_{param}(\cdot, \cdot)$ 及所有其他参数，随后在定理 4 中陈述我们的主要稳定性结果。第3节通过一些数值示例说明该结果。在补充材料中，附录A给出了主要结果定理 4 的详细证明，此外，B 节还证明了若干引理。

2 问题表述与主要结果

我们首先对问题进行正式表述。然后，我们在定理 4 中陈述主要结果：如果两个充分分离的球面高斯混合分布在分布意义上足够接近，即全变差距离足够小，那么它们的参数也足够接近。

2.1 问题设定

在问题表述中，我们确定了三个关键要素：(i) 模型类 $\mathcal{M}$，(ii) 参数空间中模型之间的距离或散度 $d_{param}(P, P')$，以及 (iii) 拟合优度度量。

模型类 一个球形高斯混合模型 P 在 $\mathbb{R}^d$ 上可写成如下形式

$$P = \sum_{k=1}^{K} \pi_k N_d(\mu_k, \sigma_k^2 I_d), \quad \text{with } \sum_{k=1}^{K} \pi_k = 1, \quad \pi_k \geq 0. \tag{1}$$

上述内容采用了标准记法，其中 $N_d(\mu_k, \sigma_k^2 I_d)$ 称为混合成分，是均值为 $\mu_{1:K} \in \mathbb{R}^d$、协方差矩阵为对角协方差矩阵 $\sigma_{1:K}^2 I_d \in \mathbb{R}^{d\times d}$ 的正态分布；而 $\pi_{1:K}$ 称为混合比例。

接下来，我们假设分量数 K 固定，并对最小分量比例和分量间隔施加限制。因此，我们考虑的模型类记为 $\mathcal{M}(K, \pi_{\min}, \pi_{\max}, c)$。

定义 3. </style}一个球形高斯混合模型 属于模型类 (K, $\pi_{\min}$, $\pi_{\max}$, c) *</style}*当且仅当满足以下条件：

A1 $K \geq 2$
A2 $\min_{k\in[K]} \pi_k \geq \pi_{\min}, \max_{k\in[K]} \pi_k \leq \pi_{\max}$
A3 $||\mu_i - \mu_j|| > c(\sigma_i + \sigma_j)$ *for any* $i, j \in [K]$ *and* $i \neq j$

对于 $P \in \mathcal{M}(K, \pi_{\min}, \pi_{\max}, c)$，最大比例始终小于或等于 $1-(K-1)\pi_{\min}$。因此，当 $\pi_{\max} = 1-(K-1)\pi_{\min}$ 时，我们将模型类的记法简写为 $\mathcal{M}(K, \pi_{\min}, c)$。间隔常数 c 是必要的。 [26] 中的定理3.1指出，随着 K 增大，当间隔常数 c 为 $o(\sqrt{\log K})$ 时，可以找到两个球形高斯混合模型，它们的参数相差 $\Omega(1)$，但全变差距离在 K 中超多项式级地小。我们通过对间隔常数 c 作出假设来排除这一困难，从而规范模型类。

模型参数的散度介于任意两个模型 P, P' 含有 K 个分量之间

$$d_{param}(P, P') = \min_{\text{perm}\in\mathbb{S}_K} \max_{k\in[K]} \frac{|\pi_i - \pi'_{\text{perm}(i)}|}{\min\{\pi_i, \pi'_{\text{perm}(i)}\}} + \frac{||\mu_i - \mu'_{\text{perm}(i)}||}{\max\{\sigma_i, \sigma'_{\text{perm}(i)}\}} + \frac{|\sigma_i^2 - \sigma'^2_{\text{perm}(i)}|}{\min\{\sigma_i^2, \sigma'^2_{\text{perm}(i)}\}} \tag{2}$$

上述 perm $\in \mathbb{S}_K$ 是集合 $[K]$的一个置换； d_{param} 不一定是度量[1]。下面给出的定理 4 分别为式 (2) 中的三个项提供了上界。

考虑 $G = \sum_{k=1}^{K} \pi_k \delta_{(\mu_k, \sigma_k^2)}$，其中 δ 是点质量。也可以将每个球形高斯混合模型分布 $P_G = \sum_{k=1}^{K} \pi_k N(\mu_k, \sigma_k^2 I_d)$ 视为参数空间上离散概率的平滑分布。设 G, G' 是参数空间上的两个此类分布，则定义 Wasserstein-1距离如下：

$$W_1(G, G') = \inf_{\Pi} \int d((\mu_k, \sigma_k^2), (\mu'_k, \sigma'^2_k)) \Pi(G, G') \tag{3}$$

[1] d_{param} 在信息几何意义上也不是一种散度 。

本文考虑该距离。先前的几篇论文（例如 [16, 15, 31, 11]）表明，它是一种自然参数散度。我们的 d_{param} 无法捕捉参数上 W_1 距离所对应的相同拓扑。考虑两个序列

$$P_n = 0.5N\ -1\ \ n+1\quad +0.5N\ 1,\ n+1 \qquad Q_n = (,(\ )^2)\ (\ ) \\ 0.5N\ -1\ n^2\ +0.5N\ 1\ n^2 \qquad\qquad n\to\infty\ \ d_{param}\ P_n\ Q_n \qquad\qquad\qquad W_1\ n^2\ \ n+$$

$^2)$ 和 $(,)$ $(,)$。于是，当 $,$ $(,)$ 收敛至零时，$(,(\ 1)^2)$ 却发散。另一方面，也可以证明，对于固定的 G 及其对应的 P_G，当$W_1(G, G') \to 0$时，有

$$d_{param}(P_G, P_{G'}) \asymp W_1(G, G') \asymp \min_{\text{perm}\in\mathbb{S}^K} \sum_{i=1}^{K} |\pi_i - \pi_{\text{perm}(i)}| + ||\mu_i - \mu'_{\text{perm}(i)}|| + |\sigma_i^2 - \sigma'^2_{\text{perm}(i)}| \ . \tag{4}$$

因此，我们对 d_{param} 的定义与模型拟合准则更加一致，不需要参数空间紧致这一假设 [16]，而且由于我们分别给出了 d_{param} 中各项的上界，因此可以直接构造拟合优度区域或置信区域。此外，d_{param} 在重新缩放下保持不变，这在基于模型的聚类等场景中会很有用。

拟合优度可以用多种准则衡量（例如似然或 KL散度、 W_2距离），但本文为方便起见，采用全变差距离 [14]。对于定义在 $\mathbb{R}^d$ 上的两个概率分布 P，Q，其全变差距离定义为

$$TV(P, Q) = \sup_{A\in\mathcal{B}} |P(A) - Q(A)| \ , \tag{5}$$

其中，$\mathcal{B}$ 表示 $\mathbb{R}^d$ 上的所有博雷尔集。

2.2 主要结果的准备：定义常数

在针对球形高斯混合这一特定情形陈述定理1之前，我们先引入若干取决于模型类 $\mathcal{M}(K,$ $\pi_{\min}$，$c)$ 和 ϵ、但与所考虑的具体分布无关的常数。这些常数用于刻画我们即将给出的 d_{param} (2) 中各项的上界。下文将定义的常数 η_0，$\eta^* > 1$ 将界定 σ_i 中的扰动，目标是使其尽可能接近 1。常数 c_0 与 η_0 一起给出分离度 c 的最小充分值，而 c^* 则界定分量均值的扰动。

用 $\Phi(x)$ 表示一维标准正态分布的累积分布函数，用 F_d 表示 Gamma($\frac{d}{2}$，$\frac{d}{2}$) 分布的累积分布函数。–

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \ , \quad F_d(x) = \int_0^x \frac{(\frac{d}{2})^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} t^{\frac{d}{2}-1} \exp\left(-\frac{d}{2}t\right) dt \ . \tag{6}$$

由 $\pi_{\min}$，ϵ，$\pi_{\max}$（或 K）确定的常数 c_0，η_0 定义如下

$$c_0 = 2\Phi^{-1}(1 - \frac{\pi_{\min} - 2\epsilon}{2}) \ ; \tag{7}$$

η_0 满足 $\eta_0 \geq 1$ 且

$$1 - \frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} + \frac{2(1-\pi_{\max})}{\pi_{\max}} \Phi(-\frac{1}{2}\eta_0 c_0) = F_d(\frac{2\eta_0^2 \log \eta_0}{\eta_0^2 - 1}) - F_d(\frac{2\log\eta_0}{\eta_0^2 - 1}) \ . \tag{8}$$

引理19（补充材料）将证明，(8) 式右侧实际上为 $TV(N(0,\ I_d)\ N(0,\ \eta^2 I_d))$，且其下界为 $1 - (2\eta/(\eta^2+1))^{d/2}$。

利用数值 c_0，η_0，可以得到方程 (2) 中三个项的上界；特别地，$c_0 \max\{\sigma_i, \sigma'_{\text{perm}(i)}\}$ 给出了第一项的界，该项表示均值变化。附录 A.3 节对此进行了详细说明。我们将 c_0，η_0 初始界 称为初始界，因为实际上还可以通过我们在 A.4 节介绍的技术进一步缩小它们。由此，我们得到依赖于常数 c^*，η^* 的精细化参数距离上界，这些常数满足 $0 \leq c^* \leq c_0$、$1 \leq \eta^* \leq \eta_0$ 以及以下条件

$$1 - 2\Phi(-\frac{c^*}{2}) = \frac{2\epsilon}{\pi_{\min}} + \frac{2(1-\pi_{\min})}{\pi_{\min}} \Phi(-\frac{1}{2}[c(1+\frac{1}{\eta^*}) - c^*]) \ , \tag{9}$$

$$F_d(\frac{2(\eta^*)^2 \log \eta^*}{(\eta^*)^2 - 1}) - F_d(\frac{2\log\eta^*}{(\eta^*)^2-1}) = \frac{2\epsilon}{\pi_{\min}} + \frac{2(1-\pi_{\min})}{\pi_{\min}} \Phi(-\frac{1}{2}[c(1+\frac{1}{\eta^*}) - c^*]) \ . \tag{10}$$

注意，通过方程（9）和（10），c^*，η^* 也由 $\pi_{\min}$，ϵ、c 和 $\pi_{\max}$ 确定。下一节将证明四个常数 c_0，η_0，c^*，η^* 存在且唯一。

2.3 主定理：球形高斯混合模型的良好拟合只能接近一个良好拟合。

现在我们可以陈述主要结果了。

定理 *4.* 设 $P \in \mathcal{M}(K, \pi_{\min}, \pi_{\max}, c)$。假设 P' 是 $\mathcal{M}(K', \pi_{\min}, \pi_{\max}, c)$ 中的任意模型，满足 $TV(P, P') \leq 2\epsilon$，其中 $\max\{K, K'\} \leq 1/\pi_{\min}$，$\pi_{\max} \leq 1 - (\min\{K, K'\} - 1)\pi_{\min}$。令$c_0, \eta_0$ 如 *(7)* 和 *(8)* 中所定义。则，当 $c \geq c_0\eta_0$ 且 $\pi_{\min} > 2\epsilon$时，我们有 $K = K'$；进一步地，存在一个置换 $perm \in \mathbb{S}_K$ 以及满足 *(9)* 和 *(10)* 的常数 $c^* \in [0, c_0], \eta^* \in [1, \eta_0]$，使得对于每个 $i \in [K]$,

$$||\mu_i - \mu'_{\text{perm}(i)}|| \leq c^*\eta^*\sigma_i \tag{11}$$

$$\max\{\sigma_i/\sigma'_{\text{perm}(i)}, \sigma'_{\text{perm}(i)}\sigma_i\} \leq \eta^* \tag{12}$$

$$|\pi_i - \pi'_{\text{perm}(i)}| \leq 2\epsilon + (1 - \pi_{\min} + \pi_{\max})\Phi(-C(c, c^*, \eta^*)) \,, \tag{13}$$

其中 $C(c, c^*, \eta^*)$ 定义为

$$C(c, c^*, \eta^*) := \sqrt{\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2}} - \frac{c^*}{2} \,. \tag{14}$$

该定理拓展了通常的可辨识性结果。关于与以往工作的更多讨论，请参见第 4 节。此外，这个上界易于处理，因为所有常数都是明确的或可计算的，并且不要求关于参数的先验知识。可计算性使我们有可能将该结果应用于有限样本。

以下命题给出了渐近区域中常数 c_0，η_0，c^*，η^* 的估计，其中 $K \to \infty$。

命题 *5.* 给定 $\pi_{\min} > 2\epsilon > 0$，则方程 *(7)–(9)* 中定义的 c_0、η_0、c^*、η^* 均存在且唯一。进一步假设存在非负常数 α、β 以及足够大的正常数 γ，使得 $\beta \leq 1 + \alpha$、$\pi_{\min} = \Omega(K^{-(1+\alpha)})$、$\pi_{\max}/\pi_{\min} = O(K^{\beta})$ 且 $2\epsilon/\pi_{\min} \leq K^{-\gamma}$。当 $K \to \infty$时，我们有初始分离条件 $c_0 = O(\sqrt{\log K})$、$\eta_0 = O(K^{2\beta/d})$。对于任意分离度 $c > c_0\eta_0$，最终上界为 $c^* = O(K^{-\gamma} + K\left[\sqrt{\log K}/K\right]^4)$ 且 $\eta^* = O(K^{-\gamma} + K\left[\sqrt{\log K}/K\right]^4)$。

这里我们作几点说明。首先，具体而言，对于满足 $\pi_{\max}/\pi_{\min} = 1$ 的平衡模型类，我们得出最小分离条件为 $c > c_0\eta_0 = O(\sqrt{\log K})$。在已知标准差比值的先验知识下，我们的结果与 [26] 中为学习充分分离的高斯混合而确立的尖锐分离阈值 $O(\sqrt{\log K})$ 相符。其次，[26] 也表明，对于足够大的 C 和 $K > C^8$，模型类 $\mathcal{M}(K, 1/K, C^{-24}\sqrt{\log K})$ 中存在两个混合分布 P，P'，它们的每个分量都具有单位方差，且参数距离至少为 $C^{-24}\sqrt{\log K}$，但 $d_{TV}(P, P') \leq K^{-C}$。然而，当分离常数中的 $\sqrt{\log K}$ 乘数更大时，我们的结果进一步表明，参数距离不会像 $K \to \infty$ 那样发散。也就是说，在我们的分离条件下，无论簇的数量有多少，都能逐分量地识别这些混合分布。

第三，当 K 固定时，分离下界 $c_0\eta_0$ 随 d增加而降低。直观地说，在具有相同K $\pi_{\min}$、$\pi_{\max}$的情况下，在更高维度中更容易区分两个高斯混合分布。这是高斯混合模型 [2]中维数祝福的一个实例。

第四，对每个分量的最大和最小标准差进行约束，也是实现正则化的一种有效方式。如果 P、P'的分量满足关系 $c > \rho_\sigma c_0$，其中$\rho_\sigma = \max_{i,\,j}\{\sigma_i, \sigma'_j\}/\min_{i,\,j}\{\sigma_i, \sigma'_j\} \leq \eta_0$，则一一对应关系同样成立。最后，当定理 4适用时，两个混合分布必须具有相同数量的分量。这进一步导出以下推论：一个充分分离的高斯混合分布不可能在全变差距离意义下接近单个球面对称高斯分布。在附录A的证明中，我们将证明这一推论并明确给出常数。

推论 6. 对于任意 $\epsilon < \pi_{\min}$，存在一个正的常数 c_{single}，它取决于 $\pi_{\min}$ 和 ϵ；当 $c > c_{\text{single}}$时，任意球形高斯分布与$P \in \mathcal{M}(K, \pi_{\min}, c)$中的任意 P之间的全变差距离都不小于 2ϵ。

我们要指出，定理 4 中的上界在 $d_{TV}(P, P')$ 趋于 0 时并不会趋于 0。理想情况下，对于固定的混合分布 $P \in \mathcal{M}(K, \pi_{\min}, c)$，由于 [16] 中提出的局部泰勒展开分析，应有 $\liminf TV(P, P')/d_{param}(P, P') > 0$，当 $d_{param}(P, P') \to 0$ 时。另一方面，我们最终得到的界 c^* 和 η^* 包含两部分：第一部分来自两个高斯分布之间的全变差距离，第二部分来自分量间距不足。据作者所知，目前尚不清楚能否得到 P, P' 之间可计算的参数距离界，使其随$O(TV(P, P'))$ 缩放，其中全变差距离 $TV(P, P')$ 趋于零。

2.4 定理 4 证明概要

此处提及的所有定理和引理的详细证明见补充材料。证明分为初始阶段和细化阶段：在初始阶段，得到界 c_0，η_0；随后进入细化阶段。

初始阶段 首先，我们在定理 15 中证明，如果 $\pi_{\min}$ 足够大且 ϵ 足够小，那么对于 P'的每个分量 P'_j，都存在 P的一个“匹配”分量 P_i，使得 $||\mu'_j - \mu_i|| \leq c_0 \max\{\sigma'_j, \sigma_i\}$，其中 c_0 由式（7）给出。

接下来，我们在定理 16 中证明，如果分离度足够大（$c > \eta_0 c_0$），那么对于 P'的每个分量 P'_j，匹配的分量 P_i是唯一的。这意味着 $K = K'$。对 P'_j的分量重新编号后，我们得到初始界（c_0, η_0），使得对于所有 i，$||\mu_i - \mu'_i|| \leq c_0 \max\{\sigma'_i, \sigma_i\}$ 和 $\max\{\sigma_i/\sigma'_i, \sigma'_i/\sigma_i\} \leq \eta_0$ 均成立。

细化 我们证明，若对于某个 (c_b, η_b)，以及来自 $\mathcal{M}(K, \pi_{\min}, \epsilon)$ 的两个球形高斯混合模型中的所有对应对，$||\mu_i - \mu'_i|| \leq c_b \max\{\sigma'_i, \sigma_i\}$ 和 $\max\{\sigma_i/\sigma'_i, \sigma'_i/\sigma_i\} \leq \eta_b$ 均成立，则对于 $j \neq i$，μ_i, μ'_j 的分离下界（即 P' 中未与 P_i 匹配的分量）可以得到改进。由此，引理 12 推出匹配的高斯分量的每一对 $TV(P_i, P'_i)$ 之间的全变差距离有一个新的上界 $UB(c_b, \eta_b)$，其中

$$UB(c_b, \eta_b) := \frac{2\epsilon}{\pi_{\min}} + \frac{2(1 - \pi_{\min})}{\pi_{\min}} \Phi\left(-(c + \frac{c}{\eta_b} - c_b)/2\right) \tag{15}$$

对于所有 $c_b \in [0, c_0]$ 和 $\eta_b \in [1, \eta_0]$，均有定义。

因此，由 (c_0, η_0)，我们从引理 13 得到一对新的上界常数 (c_1, η_1)，其中$c_1 < c_0, \eta_1 < \eta_0$。利用它们，我们可以将全变差距离 $TV(P_i, P'_i)$ 上界为 $UB(c_1, b_1) < UB(c_0, b_0)$，从而得到更紧的常数 (c_2, b_2)，如此反复。

上界序列（c_t，η_t）收敛于 η^*，c^*，其中 $\{c_t\}$ 和 $\{\eta_t\}$ 均为正数，且随 t 增大而递减。这些极限见方程 (9)，证明见补充材料 A.4。结合比例的分量级比较引理 14，现在即可完成定理 4 的证明。

3 数值示例

定理 4 中的上界和条件均可计算，下面我们将在多种示例中对其进行评估。

最小分离度 回顾一下，要使定理 4 适用，两个高斯混合 P，P' 的分量必须充分分离，相对分离度为 $c \geq c_0 \eta_0$。这里我们计算极限情形 $\epsilon = 0$ 下的最小分离度 $c_0 \eta_0$。这将展示该定理的适用范围。

我们考虑在 $d = 5$,20,35 维欧几里得空间中的球面高斯混合分布，分量数量 K 从 2 到 40。对于每个 d 和 K，比率 $\eta_\pi = \pi_{\max}/\pi_{\min}$ 设为 1、2、4、8 和 16。根据 K 和 η_π，我们设定 $\pi_{\min}, \pi_{\max}$，使得 $\pi_{\max} = 1 - (K - 1)\pi_{\min}$。图 3 展示了应用定理 4 所需的最小分离度。我们观察到，混合比例的异质性 η_π 确实会对分离度要求产生很大影响；当 η_π 高达 16 时，即使只有 2 个簇，在 5 维空间中应用定理 4 也可能需要间隔常数 $c \geq 12$。另一方面，对于均衡混合，要求并不那么严格。对于 $\mathcal{M}(2, 1/2, c)$ 中的混合分布，即两个分量相等的混合分布，即使对于 $d = 1$，c 的下界也是

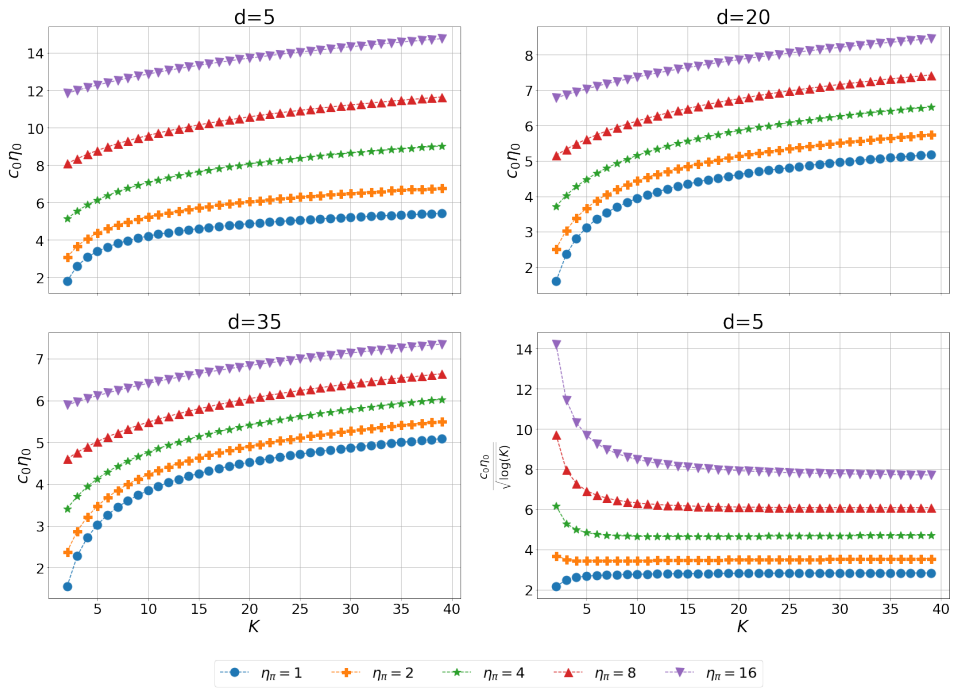


图 2：**定理 4 中不同设置下的充分最小分离度** $c_0\eta_0$ 。**左上、右上和左下**分别展示了在维度为$d = 5$、20、35时， $c_0\eta_0$ 对 K 和 $\eta_\pi = \pi_{\max}/\pi_{\min}$ 的依赖关系。**右下**展示了 $c_0\eta_0$ 对 K 的依赖关系渐近于$\sqrt{\log\ K}$。

2.29。正如预期的那样， [2], 当 d 增加时，这一要求会降低，最低可降至 $c = 1.55$（当 $d = 35$时）。当 K 增加时， $c_0\eta_0$ 以约 $\sqrt{\log K}$ 的速度增加。

参数散度的上界这里给出不同模型拟合水平 ϵ下，上界 c^*、 η^* 的数值，以及方程 (13) 中相对差 $|\pi_i - \pi'_{\text{perm}(i)}|/\pi_{\min}$ 的上界。我们在 $\mathbb{R}^{20}$上考虑 $P \in \mathcal{M}(K,\ 1/(K+1)\ c)$，并令 $K = 2$、5、10，以及 $c = 3$、4、5、6取不同值。图 3 展示了这一情形下三个上界的数值。例如，根据定理 4，在 $d = 20$ 维空间中，一个含有 $K = 2$ 个分量的混合分布要满足 $\epsilon \approx 0.01$ ，其分离度须达到 $c \approx 3$。一旦满足这一条件，我们就能获得良好的保证：对于每一对对应的分量，

$$||\mu_i - \mu'_i|| \leq 0.151 \max\{\sigma_i, \sigma'_i\}, \quad \frac{\max\{\sigma_i, \sigma'_i\}}{\min\{\sigma_i, \sigma'_i\}} \leq 1.035, \quad |\pi_i - \pi'_i| \leq 0.02 \approx 0.06\pi_{\min}\ . \quad (16)$$

从图 3 可以看出，当 ϵ 较小时，所有上界都由分离度 c主导。随着 ϵ增大，我们观察到，三个上界都由 ϵ主导。请注意，这些图与图 3 中橙色曲线 $\eta_\pi = 2$之间的关系。

4 讨论

与近期稳健可识别性结果的比较 据我们所知，定理 4 是首个为充分分离的球形高斯混合模型提供可计算的全局参数稳定性上界的结果。该结果也可视为针对 $\mathcal{M}(K,\ \pi_{\min},\ c)$的稳健可识别性结果。

当 $TV(P,\ P') \to 0$ 时，存在一些基于参数 W_1 距离的可辨识性结果： [16],[15],等表明，对于两组参数分布 $G,\ G'$，若 $W_1(G,\ G')$ 较小，则 $TV(P_G,\ P'_{G'})$ 渐近上大于 $W_1(G,\ G')$ 距离（无需假设分离度）。这一结果几乎是我们定理 4 的逆命题。

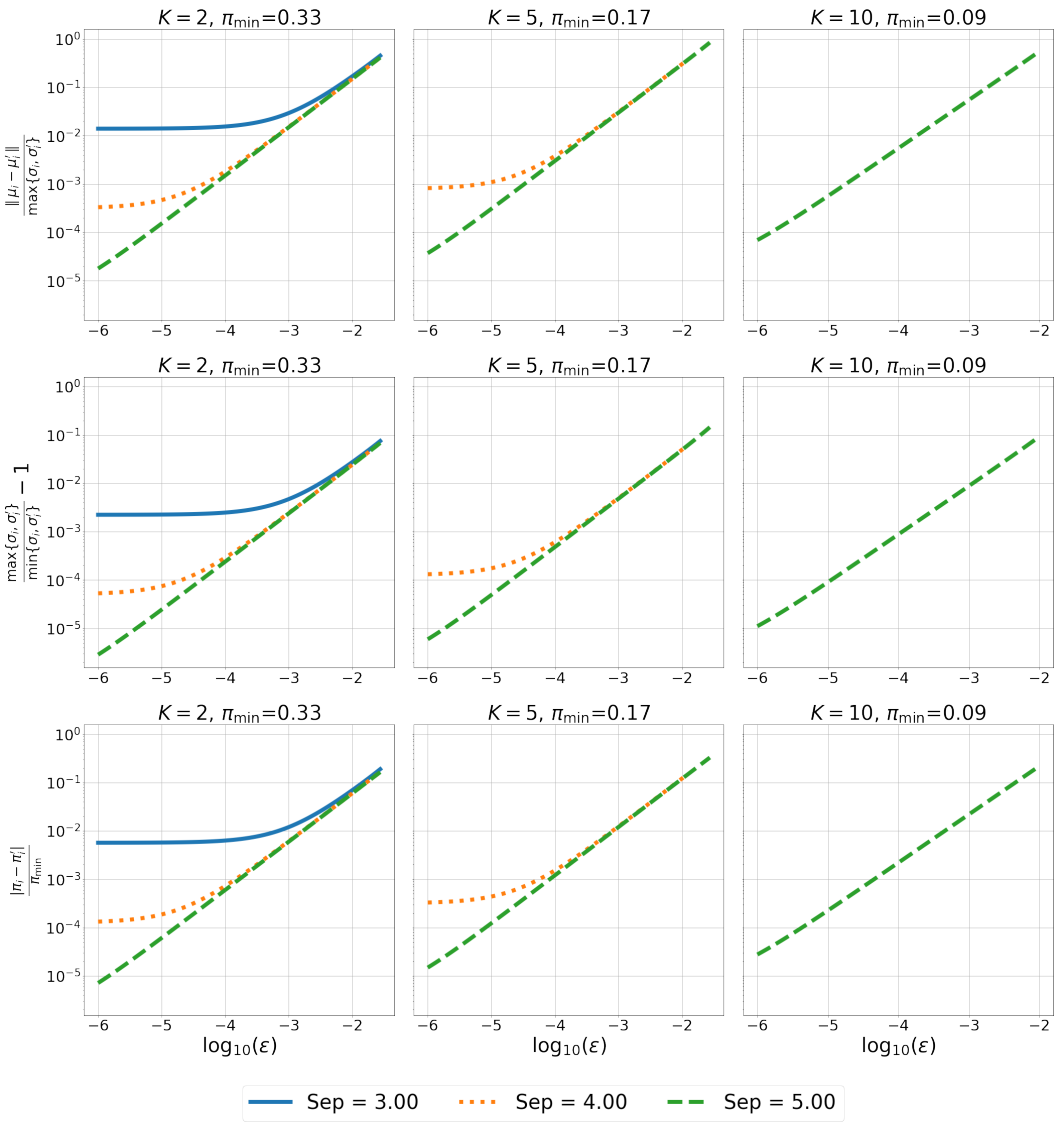


图 3：定理 4 给出的对应均值之间距离的**上界** c^*（上排）、标准差之比减一 η^*（中排）以及以 $\pi_{\min}$ 的倍数衡量的混合比例差异（下排），分别对应不同的 K 值和分离度 Sep $= c$，维度为 $d = 20$。部分曲线未显示，因为分离度不够大，无法得到上界；图 3 左上面板中的橙色曲线说明了这一点。

最相关的工作是高斯混合鲁棒学习中的参数可识别性定理，即 [10], 中的定理 B.1、 [22], 中的定理 8.1、 [3], 中的定理 9.1；这些结果基于新近发展的矩方法证明技巧。

考虑两个一般高斯分布的混合分布 P, P'，其分量的最大数量为K。所作假设为： $TV(P, P') \leq \epsilon$， P, P'的最小权重大于ϵ^{b_1}，且 P, P'的不同分量之间的全变差距离分离度大于 ϵ^{b_2}, 其中常数 $b_1, b_2 > 0$ 足够小，且仅依赖于 K。证明中使用的关键观察是：如果厄米矩多项式有界于 $poly(\epsilon)$，则 $P_k, P'_{\text{perm}(k)}$ 的参数距离有界于 $poly(\epsilon)$；而这由 P, P' 之间的全变差距离上界为 ϵ这一假设保证。他们的证明技巧可以直接修改，用于证明 P 是$(\epsilon, \tilde{\delta}(\epsilon))$ 稳定的，且具有$\tilde{\delta}(\epsilon) = O_k(\epsilon^{b_3^K})$，其中 b_3 是另一个足够小的正数常数。

当 P 固定时（这既是我们关注的情形，也是一般稳健可识别性所关注的情形），如果 $\epsilon \to 0$ ~，这些结果比我们的结果更紧，因为上界为 $\delta \to 0$。在这种情况下，对 P 的假设比我们的假设宽松。然而，如何从他们的证明技巧中确定 δ 中的未指定常数尚不明确，而且对 K 的依赖关系以及 ϵ 的指数都不是最优的（关于这一点，参见第 2 节中的讨论[26],）。

对于规模更大、更贴近现实的 ϵ，速率 $\tilde{\delta}(\epsilon) \sim \epsilon^{b_3^K}$ 远慢于我们的数值模拟所达到的水平，其表现为 $\delta(\epsilon, \mathcal{M}) \sim \epsilon$。在这种情况下，定理 4 中关于 $\pi_{\min}$ 和分离度的假设保持不变，而 [22, 3] 中的假设随着 ϵ增大而变得更加严格（甚至可能不再适用）。因此，我们的结果更适用于鲁棒恢复，因为这类问题希望将稳定性扩展到更大的 ϵ；而 [22, 3] 在极限 $\epsilon \to 0$ 下以及对于算法 稳定性（这些工作的主要目标）更有用。$\delta(\epsilon, \mathcal{M}) \sim \epsilon$ 的这种增长与 [16], 的尖锐阈值相符；该阈值是针对 $\epsilon \to 0$渐近得到的。这表明，对于更大的 ϵ， [16] 可能仍然成立，而且至少在某些条件下，我们的最坏情况界可能与之相符。

关于对 K的依赖， [26], 给出了球形高斯混合模型的尖锐阈值，并证明要在多项式样本量和多项式时间内恢复这类高斯混合模型，分离度至少需要达到 $\Omega(\sqrt{\log K})$，且这一条件也是充分的。我们的定理 4 与该阈值近似匹配。

其他相关问题的结果 已证明，在球形高斯混合模型中求解精确最大似然估计是 NP-困难的 [30]。目前拟合高斯混合模型的主流方法是期望最大化 （EM）算法 [8]；若没有额外假设，就无法证明 EM 会收敛到对数似然函数的全局最大值。事实上，即使对于充分分离的高斯分布，EM 也可能以很高的概率收敛到较差的局部极大值 [19], ；同一篇论文还证实，总体似然函数存在局部极大值。这一负面结果使得设置后处理验证阶段更加必要，以便能够排除较差的局部最优解。从总体角度看，我们论文中的结果可以发挥这一作用。

第二种能够提供一致性保证的高斯混合模型参数估计方法依赖于矩方法 (MOM) [24], , [17, 31]中便是一个例子。 [17] 使用三阶及以下的矩来学习球形高斯混合模型。该方法不要求分离条件，但要求不同分量的均值线性无关。因此，这种方法无法应用于低 d 和大 K 的情形。 [31]使用了基于半定规划的去噪过程，对最高阶数为 O的矩进行处理(K)。
当簇分量充分分离时，开创性论文 [6] 提出并在 [27, 1, 20],等工作中改进的另一条研究路线，以及 [7], , 能够以高概率准确估计高斯混合模型参数。所有这些结果都以数据采样自充分分离的高斯分布混合分布为前提。

此外，迄今为止，关于如何评估高斯混合拟合结果的研究成果也寥寥无几。在这些研究中，例如 [12], [18]，主要目标是模型选择，尤其是对高斯混合总体中的分量数量进行一致估计。

自助法稳定性 一系列论文 [21, 25, 5, 4, 28, 29] 探讨了 K均值和高斯混合的“自助法稳定性”特性，重点关注对数据进行重抽样或扰动聚类算法时聚类的稳定性。这些研究的目标是利用自助法稳定性来选择簇的数量 K。

总之，本文在总体设定下，首次获得了球形高斯混合可计算的稳健可识别性界。我们的界可以推广到具有完整协方差矩阵 Σ_k 和有界偏心率的高斯混合。该界 $\delta(\epsilon, \mathcal{M})$ 与 [26], 中关于 K 的已知尖锐阈值相匹配，并且对整个模型类 $\mathcal{M}(K, \pi_{\min}, c)$ 一致成立。

在证明中，我们引入了一种据我们所知颇具原创性的迭代方法，用于收紧这些界。我们的若干其他引理也可能具有独立的研究价值，例如，关于单个高斯分布参数变化的更紧界；当 TV 距离上的扰动较大时，该界仍比 [9] 的先前界更具参考价值。

致谢

作者感谢美国国家科学基金会数学科学部项目1810975号奖项的支持。

参考文献

[1] Dimitris Achlioptas 和 Frank McSherry。《分布混合的谱学习》。载于第*18*届年度学习理论会议论文集，COLT'05，第458–469页，柏林、海德堡，2005年。施普林格出版社。ISBN 3540265562。doi：10.1007/11503415_31。URL：https://doi.org/10.1007/11503415_31。[2] Joseph Anderson、Mikhail Belkin、Navin Goyal、Luis Rademacher 和 James R. Voss。《越多越好：维数祝福与大型高斯混合的学习》。载于*COLT*，2014年。[3] Ainesh Bakshi、Ilias Diakonikolas、He Jia、Daniel M. Kane、Pravesh K. Kothari 和 Santosh S. Vempala。《鲁棒学习k个任意高斯分布的混合》。*CoRR*，abs/2012.02119，2020年。URL：https://arxiv.org/abs/2012.02119。[4] S. Ben-David 和 U. von Luxburg。《将聚类稳定性与聚类边界的性质联系起来》。载于*COLT2008*，第379–390页，美国威斯康星州麦迪逊，2008年7月。马克斯·普朗克学会，Omnipress。[5] Shai Ben-David、Dávid Pál 和 Hans Ulrich Simon。《k均值聚类的稳定性》。载于 Nader H. Bshouty 和 Claudio Gentile（编），学习理论，第20–34页，柏林、海德堡，2007年。施普林格出版社。[6] Sanjoy Dasgupta。《高斯分布混合的学习》。载于第*40*届年度计算机科学基础研讨会论文集，FOCS '99，第634页，美国，1999年。IEEE计算机学会。ISBN 0769504094。[7] Sanjoy Dasgupta 和 Leonard Schulman。《分离球形高斯分布混合的EM算法的概率分析》。《机器学习研究杂志》，8(7)：203–226，2007年。URL：http://jmlr.org/papers/v8/dasgupta07a.html。[8]A. P. Dempster、N. M. Laird 和 D. B. Rubin。《通过EM算法从不完整数据中进行最大似然估计》。《英国皇家统计学会会刊：*B*辑（方法论）》，39(1)：1–22，1977年。doi：https://doi.org/10.1111/j.2517-6161.1977.tb01600.x。URL：https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1977.tb01600.x。[9] Luc Devroye、Abbas Mehrabian 和 Tommy Reddad。《高维高斯分布之间的全变差距离》，2020年。[10] Ilias Diakonikolas、Daniel M. Kane 和 Alistair Stewart。《高维高斯分布及高斯混合鲁棒估计的统计查询下界》。载于*2017*年*IEEE*第*58*届年度计算机科学基础研讨会（*FOCS*），第73–84页，2017年。doi：10.1109/FOCS.2017.16。[11] Natalie Doss、Yihong Wu、Pengkun Yang 和 Harrison H. Zhou。《高维高斯混合的最优估计》，2020年。[12] Mathias Drton 和 Martyn Plummer。《奇异模型的贝叶斯信息准则》。《英国皇家统计学会会刊：*B*辑（统计方法论）》，79(2)：323–380，2017年2月。doi：10.1111/rssb.12187。[13] Harley Flanders。《积分号下求导》。美国数学月刊，80(6)：615–627，1973年。ISSN 00029890，19300972。URL：http://www.jstor.org/ stable/2319163。[14] Alison L. Gibbs 和 Francis Edward Su。《概率度量的选择与界定》。国际统计评论 / 国际统计评论，70(3)：419–435，2002年。ISSN 03067734，17515823。URL：http://www.jstor.org/stable/1403865。[15] Philippe Heinrich 和 Jonas Kahn。《有限混合估计的强可识别性与最优极小极大速率》。《统计年鉴》，46(6A)：2844 – 2870，2018年。doi：10.1214/ 17-AOS1641。URL：https://doi.org/10.1214/17-AOS1641。

[16] Nhat Ho 和 XuanLong Nguyen。某些弱可辨识有限混合模型参数估计的收敛速率。《统计年鉴》，44(6):2726 – 2755，2016。doi: 10.1214/ 16-AOS1444。URL https://doi.org/10.1214/16-AOS1444.[17] Daniel Hsu 和 Sham M. Kakade。球面高斯混合分布的学习：矩方法与谱分解。载于第四届理论计算机科学创新会议论文集，ITCS '13，第11–20页，美国纽约州纽约，2013。计算机协会。ISBN 9781450318594。doi: 10.1145/2422436.2422439。URL https://doi.org/10.1145/2422436.2422439.[18] Tao Huang、Heng Peng 和 Kun Zhang。高斯混合模型的模型选择。《统计学报》，27(1):147–169，2017。ISSN 10170405, 19968507。URL http://www.jstor.org/stable/44114365.[19] Chi Jin、Yuchen Zhang、Sivaraman Balakrishnan、Martin J. Wainwright 和 Michael I. Jordan。高斯混合模型似然中的局部极大值：结构性结果与算法后果。载于第*30*届神经信息处理系统国际会议论文集，NIPS'16，第4123–4131页，美国纽约州雷德胡克，2016。Curran Associates Inc. ISBN 9781510838819.[20] Ravindran Kannan、Hadi Salmasian 和 Santosh Vempala。一般混合模型的谱方法。《*SIAM*计算杂志》，38(3):1141–1156，2008。doi: 10.1137/ S0097539704445925.[21] Tilman Lange、Volker Roth、Mikio L. Braun 和 Joachim M. Buhmann。聚类解的稳定性验证。《神经计算》，16(6):1299–1323，2004。ISSN 0899-7667。doi: http://dx.doi.org/10.1162/089976604773717621.[22] Allen Liu 和 Ankur Moitra。解决高斯混合的鲁棒可学习性问题。载于第*53*届*ACM SIGACT*年度理论计算研讨会论文集，STOC 2021，第518–531页，美国纽约州纽约，2021。计算机协会。ISBN 9781450380539。doi: 10.1145/3406325.3451084。URL https://doi.org/10.1145/3406325.3451084.[23] Marina Meilˇa。k均值的良好最优解的唯一性。载于第*23*届国际机器学习会议论文集，ICML '06，第625–632页，美国纽约州纽约，2006。计算机协会。ISBN 1595933832。doi: 10.1145/1143844. 1143923。URL https://doi.org/10.1145/1143844.1143923.[24] 卡尔·皮尔逊。III. 对进化数学理论的贡献。伦敦皇家学会哲学汇刊（*A*辑），185:71–110，1894.[25] Alexander Rakhlin 和 Andrea Caponnetto。k均值聚类的稳定性。载于第*19*届神经信息处理系统国际会议论文集，NIPS'06，第1121–1128页，美国马萨诸塞州剑桥，2006。MIT出版社。[26] Oded Regev 和 Aravindan Vijayaraghavan。关于充分分离高斯分布混合的学习。*CoRR*，abs/1710.11592，2017。URL http://arxiv.org/abs/1710.11592.[27] Arora Sanjeev 和 Ravi Kannan。任意高斯分布混合的学习。载于第*33*届*ACM*年度理论计算研讨会论文集，STOC '01，第247–257页，美国纽约州纽约，2001。计算机协会。ISBN 1581133499。doi: 10.1145/380752.380808。URL https://doi.org/10.1145/380752.380808.[28] Ohad Shamir 和 Naftali Tishby。大样本情形下聚类稳定性的可靠性。载于 D. Koller、D. Schuurmans、Y. Bengio 和 L. Bottou 编，《style id='39'>《神经信息处理系统进展21》》，第1465–1472页。Curran Associates, Inc.，2009.[29] Ohad Shamir 和 Naftali Tishby。k均值聚类中的稳定性与模型选择。《机器学习》，80(2):213–243，2010年9月。[30] Christopher Tosh 和 Sanjoy Dasgupta。球形高斯混合的最大似然估计是NP-困难。《机器学习研究杂志》，18(175):1–11，2018。URL http://jmlr.org/papers/v18/16-657.html。

[31] Yihong Wu 和 Pengkun Yang。通过去噪矩估计法对高斯混合进行最优估计。《统计年鉴》，48(4):1981 – 2007，2020。doi: 10.1214/19-AOS1873。URL https://doi.org/10.1214/19-AOS1873。

A 主要结果的证明

A.1 命题 5 的证明

命题 *5* 的证明。 由 (7) 可知，首先 c_0 存在且唯一。 c_0 的定义表明 t

$$\Phi(-c_0/2) = 1 - \Phi(c_0/2) = 1 - \Phi(\Phi^{-1}(1 - \frac{\pi_{\min} - 2\epsilon}{2})) = \frac{1}{2}(\pi_{\min} - 2\epsilon) \,. \tag{17}$$

将 $R(\eta)$ 视为由 (8) 右侧定义的一个趋于零的函数，即

$$R(\eta) = F_d\left(\frac{2\eta^2 \log \eta}{\eta^2 - 1}\right) - F_d\left(\frac{2 \log \eta}{\eta^2 - 1}\right) \tag{18}$$

并且当 $\eta \to 1^+$ 时， $R(\eta) \to 0$ 。根据引理 21， $R(\eta)$ 关于 η 单调递增，并且当 $\eta \to \infty$ 时趋于 1。记

$$L(\eta) = 1 - \frac{\pi_{\min} - 2\epsilon}{\pi} + \frac{1 - \pi}{\pi} 2\Phi(-\frac{\eta_0 c_0}{2}) \,. \tag{19}$$

那么，(8) 的左侧均可写为 $L(\eta)$。由于

$$L(1) = 1 - \frac{\pi_{\min} - 2\epsilon}{\pi} + \frac{1 - \pi}{\pi}(\pi_{\min} - 2\epsilon) \tag{20}$$

$$= 1 - (\pi_{\min} - 2\epsilon)\left(-\frac{1}{\pi} + \frac{1}{\pi} - 1\right) \tag{21}$$

$$= 1 - \pi_{\min} + 2\epsilon > 0 \tag{22}$$

并且当 $\eta \to \infty$, $L(\eta)$ 趋于 $1 - (\pi_{\min} - 2\epsilon)/\pi < 1$ 时。因此，由于 $L(\eta)$ 关于 η 单调递减，而 $R(\eta)$ 单调递增，存在唯一的 η_0 满足方程 (8)。

对于这个引理的第二部分，我们现在考虑以下情形：对于某个 $\alpha \geq 0$,$\pi_{\max}/\pi_{\min} = O(K^\beta)$，有 $\pi_{\min} = \Omega(K^{-(1+\alpha)})$；对于某个 $0 \leq \beta \leq 1 + \alpha$ ，有 $\pi_{\max}/\pi_{\min} = O(K^\beta)$；并且 ϵ 足够小，使得 $\pi_{\min} - 2\epsilon = \Omega(K^{-(1+\alpha)})$。由 (7) 可得

$$\Phi(-\frac{1}{2}c_0) = \Omega\left(\frac{1}{2K^{1+\alpha}}\right) \,. \tag{23}$$

注意，对于任意常数 $\zeta > 0$,

$$\Phi(-\frac{1}{2}2\sqrt{2\log(2K^{1+\alpha}/\zeta)}) \leq \exp\left(-\frac{1}{8}(2\sqrt{2\log(2K^{1+\alpha}/\zeta)})^2\right) = \frac{\zeta}{2K^{1+\alpha}} \lesssim \Phi(-\frac{1}{2}c_0) \tag{24}$$

其中 $a \lesssim b$ 表示存在一个常数 $A > 0$，使得 $a \leq A \cdot b$。因此，存在某个 $\zeta_0 > 0$ 使得

$$c_0 \leq 2\sqrt{2\log(\frac{2K^{1+\alpha}}{\zeta_0})} = O(\sqrt{\log K})$$

.

现在我们给出 η_0 的上界。注意，对于任意 $t > 0$，令

$$g(t) = 1 - \Phi(t) - \frac{1}{\sqrt{2\pi}} \frac{t}{t^2 + 1} \exp(-\frac{1}{2}t^2). \tag{25}$$

由于 $g'(t) = -2e^{-t^2/2}/(t^2 + 1)^2$， $g(t)$ 严格递减。由于 $\lim_{t\to\infty} g(t) = 0$，我们有$g(t) \geq 0$，即，

$$1 - \Phi(t) \geq \frac{1}{\sqrt{2\pi}} \frac{t}{t^2 + 1} \exp(-\frac{1}{2}t^2) > \frac{1}{5\sqrt{2\pi}t} \exp(-\frac{1}{2}t^2) \tag{26}$$

对所有 $t \geq 1/2$ 均成立。因此，由于 $\Phi(-c_0/2) \leq 1/2K \leq 1/4$ 或 $c_0/2 \geq 0.67 > 0.5$

$$\frac{1}{2K} \geq \Phi(-\frac{1}{2}c_0) \geq \frac{1}{5\sqrt{2\pi}\frac{1}{2}c_0} \exp(-\frac{1}{8}c_0^2) \,, \tag{27}$$

我们得出 $\exp(-\frac{1}{8}c_0^2) = O(\sqrt{\log K/K})$。因此，

$$L(\eta) = 1 - \frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} + \frac{1 - \pi_{\max}}{\pi_{\max}} 2\Phi(-\frac{1}{2}\eta c_0) \tag{28}$$

$$\leq 1 - \frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} + 2K \exp\left(-\frac{1}{2}\eta^2 c_0^2\right) \tag{29}$$

$$= 1 - \frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} + O\left(K\left[\frac{\sqrt{\log K}}{K}\right]^{4\eta^2}\right) . \tag{30}$$

另一方面，(8) 的右侧有如下下界：

$$R(\eta) > 1 - \left(\frac{2\eta}{\eta^2 + 1}\right)^{\frac{d}{2}} > 1 - \left(\frac{2}{\eta}\right)^{\frac{d}{2}} . \tag{31}$$

因此，取 $\tilde{\eta} = 2(2\pi_{\max}/(\pi_{\min} - 2\epsilon))^{2/d} = O(K^{2\beta/d})$，可得

$$R(\tilde{\eta}) - L(\tilde{\eta}) > \frac{1}{2}\frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} - \Omega\left(K\left[\frac{\sqrt{\log K}}{K}\right]^{4\tilde{\eta}^2}\right) . \tag{32}$$

然而，根据 $\tilde{\eta}$ 的选取，对于充分大的 K， $R(\tilde{\eta}) - L(\tilde{\eta})$ 必须为正。因此$\eta_0 < \tilde{\eta}$，由此可得，对于充分大的 K $\eta_0 = O(K^{2\beta/d})$。
最后，我们将考虑 c^* 和 η^* 的估计。首先，(9) 的右侧始终有如下上界：

$$\frac{2\epsilon}{\pi_{\min}} + \frac{2(1 - \pi_{\min})}{\pi_{\min}}\Phi(-\frac{1}{2}[c(1 + \frac{1}{\eta^*}) - c^*]) \leq \gamma + 2K \exp\left(-\frac{1}{2}c_0^2\eta_0^2\right) \tag{33}$$

$$= \frac{1}{K^\gamma} + O\left(K\left[\frac{\sqrt{\log K}}{K}\right]^{4\eta_0^2}\right) . \tag{34}$$

(9) 和 (10) 的右侧为 $O(1)$，并且当 K 充分大时小于 1。注意，对于两个球形高斯分布 $P_1 = N_d(\mu_1, \sigma_1^2 I_d)$ $P_2 = N_d(\mu_2, \sigma_2^2 I_d)$，由引理 7 可得 c^* 和 η^* 的界。 □

A.2 证明定理 4 的技术工具

我们不直接着手证明定理 4，而是先讨论混合模型中的一些观察结果，以及球形高斯分布之间全变差距离的分析。值得注意的是，本小节中的引理并不要求两个混合模型 P， P' 具有相同数量的分量。

A.2.1 球形高斯分布之间的全变差距离

我们提出一个引理，根据球形高斯分布之间的全变差距离，给出其参数距离的上界。为此，可以使用海林格距离，它具有解析表达式，可作为自然的下界；Gibbs 和 Su [14], 给出了这一方法，但其常数通常不够紧，因此无法分别得到均值参数和方差参数的界。Devroye 等人[9] 从定义出发提出了另一个下界，但只有当全变差距离足够小时，他们的结果才有意义。

引理 7. 设 $P_1 = N_d(\mu_1, \sigma_1^2 I_d)$ 和 $P_2 = N_d(\mu_2, \sigma_2^2 I_d)$。令 $C_0(\rho) = 2\Phi^{-1}(1 - \frac{\rho}{2})$ 和 $\eta_0(\rho)$ 为以下方程的解

$$1 - \rho = F_d\left(\frac{2\eta^2 \log \eta}{\eta^2 - 1}\right) - F_d\left(\frac{2 \log \eta}{\eta^2 - 1}\right) . \tag{35}$$

则有如下结论：

- 若 $||\mu_1-\mu_2|| \geq C_0(\rho)(\sigma_1+\sigma_2)/2$，则 $TV(P_1,P_2) \geq 1-\rho$。当且仅当 $\sigma_1=\sigma_2$且 $||\mu_1-\mu_2||= C_0(\rho)(\sigma_1+\sigma_2)/2$时，等号成立。
- 如果 $\max\{\sigma_1/\sigma_2, \sigma_2/\sigma_1\} \geq \eta_0(\rho)$，那么 $TV(P_1,P_2) \geq 1-\rho$。取得等号当且仅当 $\mu_1=\mu_2$ 且 $\max\{\sigma_1/\sigma_2, \sigma_2/\sigma_1\} = \eta_0(\rho)$。

我们将高斯总变差下界分为三个不同的引理，并对其进行证明。

引理 8。 </style}设 ($\mu_1, \sigma_1^2 I_d$)和 $P_2 = N_d(\mu_2, \sigma_2^2 I_d)$。令 Φ 为标准正态分布的累积分布函数，并对每个 $0<\rho<1$定义，

$$C_0(\rho) = 2\Phi^{-1}(1-\frac{\rho}{2}) \tag{36}$$

如果 $C = ||\mu_1-\mu_2||/\max\{\sigma_1,\sigma_2\} \geq C_0(\rho)$，那么存在一个集合 A，使得 $P_1(A) - P_2(A) \geq 1-\rho$。取得等号当且仅当 $\sigma_1=\sigma_2$ 且 $C=C_0(\rho)$。

引理 8 的证明。令 p_1, p_2 为 P_1, P_2 的密度，且随机变量 X_1, X_2 分别服从 P_1, P_2。

不失一般性，假设 $\sigma_1 \leq \sigma_2$。若该结论在此情形下成立，那么当 $\sigma_1 > \sigma_2$ 时，可以选取 A，使得 $\overline{P_2(A) - P_1(A)} > 1-\rho$。接着取 $A = A$，则所需结果成立。

首先，我们考虑 $d=1$ 的情形，并假设 $\mu_2 - \mu_1 = C\sigma_2 = C\eta\sigma_1 > 0$。当 $\sigma_2 > \sigma_1$时，令 A 为集合 $\{x \in \mathbb{R}:\ x_{0,-} \leq x \leq x_{0,+}\}$，其中

$$x_{0,\pm} = \frac{\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} \pm \sqrt{\frac{(\mu_1-\mu_2)^2}{\sigma_1^2\sigma_2^2} + 2(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2})\log\frac{\sigma_2}{\sigma_1}}}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}} \tag{37}$$

$x_{0,\pm}$ 是两个正态密度相等时方程的两个根，即

$$\frac{1}{\sqrt{2\pi}\sigma_1}\exp\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\} = \frac{1}{\sqrt{2\pi}\sigma_2}\exp\{-\frac{(x-\mu_2)^2}{2\sigma_2^2}\} \tag{38}$$

记 $\eta = \sigma_2/\sigma_1 \geq 1$，当 $\eta > 1$ 时，有

$$x_{1,\pm} = \frac{x_{0,\pm}-\mu_1}{\sigma_1} = \frac{-C\eta \pm \eta\sqrt{C^2\eta^2 + 2(\eta^2-1)\log\eta}}{\eta^2-1} \tag{39}$$

$$x_{2,\pm} = \frac{x_{0,\pm}-\mu_2}{\sigma_2} = \frac{-C\eta^2 \pm \sqrt{C^2\eta^2 + 2(\eta^2-1)\log\eta}}{\eta^2-1} \tag{40}$$

令 φ 为标准正态分布的密度函数，则根据$x_{0,\pm}$的选取，$\varphi(x_{2,+}) = \eta\varphi(x_{1,+})$ $\varphi(x_{2,-}) = \eta\varphi(x_{1,-})$。由于所需表达式

$$P_1(A) - P_2(A) = Pr(x_{1,-} \leq \frac{x-\mu_1}{\sigma_1} \leq x_{1,+}) - Pr(x_{2,-} \leq \frac{x-\mu_2}{\sigma_2} \leq x_{2,+}) \tag{41}$$

$$= \Phi(x_{1,+}) - \Phi(x_{1,-}) - \Phi(x_{2,+}) + \Phi(x_{2,-}) \tag{42}$$

隐式地是 C，η 的函数，我们将其记为 $h_+(C,\ \eta)$。注意，由于

$$x_{0,\pm} = \frac{\left(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_2}{\sigma_2^2}\right)^2 - \frac{(\mu_1-\mu_2)^2}{\sigma_1^2\sigma_2^2} - 2(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2})\log\frac{\sigma_2}{\sigma_1}}{\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}\right)\left(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_2}{\sigma_2^2} \mp \sqrt{\frac{(\mu_1-\mu_2)^2}{\sigma_1^2\sigma_2^2} + 2(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2})\log\frac{\sigma_2}{\sigma_1}}\right)} = \frac{\frac{\mu_1^2}{\sigma_1^2}-\frac{\mu_2^2}{\sigma_2^2} - 2\log\frac{\sigma_2}{\sigma_1}}{\frac{\mu_1}{\sigma_1^2}-\frac{\mu_2}{\sigma_2^2} \mp \sqrt{\frac{(\mu_1-\mu_2)^2}{\sigma_1^2\sigma_2^2} + 2(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2})\log\frac{\sigma_2}{\sigma_1}}} \tag{43}$$

考虑极限 $\sigma_2 \to \sigma_1$，令 σ_1 固定，并回忆 $\mu_2 - \mu_1 = C\sigma_2 > 0$，可得 $\lim_{\sigma_2\to\sigma_1} x_{0,-} = -\infty$，且 $\lim_{\sigma_2\to\sigma_1} x_{0,+} = (\mu_1+\mu_2)/2$。此外

$$\lim_{\sigma_2\to\sigma_1^+} x_{1,+} = \frac{\mu_2-\mu-1}{2\sigma_1} = \frac{C}{2}, \quad \lim_{\sigma_2\to\sigma_1^+} x_{1,-} = -\infty \tag{44}$$

$$\lim_{\sigma_2\to\sigma_1^+} x_{2,+} = \frac{\mu_1-\mu_2}{2\sigma_2} = -\frac{C}{2}, \quad \lim_{\sigma_2\to\sigma_1^+} x_{2,-} = -\infty \tag{45}$$

由此可知，由于 Φ 连续，

$$\lim_{\eta\to 1^+} h_+(C,\eta) = 1 - 2\Phi(-\frac{C}{2}) \tag{46}$$

注意到 $x_{1,\pm} - \eta x_{2,\pm} = C\eta$ 或 $x_{2,\pm} = x_{1,\pm}/\eta - C$。于是

$$\frac{\partial x_{1,\pm}}{\partial \eta} - \eta\frac{\partial x_{2,\pm}}{\partial \eta} = \frac{\partial x_{1,\pm}}{\partial \eta} - \eta\left[-\frac{1}{\eta^2}x_{1,\pm} + \frac{1}{\eta}\frac{\partial x_{1,\pm}}{\partial \eta}\right] = \frac{1}{\eta}x_{1,\pm}\,. \tag{47}$$

(38) 表明 $\eta\varphi(x_{1,\pm}) = \varphi(x_{2,\pm})$。然后，对 η 求偏导数可得

$$\begin{aligned}
\frac{\partial h_+(C,\eta)}{\partial \eta} &= \varphi(x_{1,+})\frac{\partial x_{1,+}}{\partial \eta} - \varphi(x_{1,-})\frac{\partial x_{1,-}}{\partial \eta} - \varphi(x_{2,+})\frac{\partial x_{2,+}}{\partial \eta} + \varphi(x_{2,-})\frac{\partial x_{2,-}}{\partial \eta} && (48)\\
&= \varphi(x_{1,+})\left[\frac{\partial x_{1,+}}{\partial \eta} - \eta\frac{\partial x_{2,+}}{\partial \eta}\right] - \varphi(x_{1,-})\left[\frac{\partial x_{1,-}}{\partial \eta} - \eta\frac{\partial x_{2,-}}{\partial \eta}\right] && (49)\\
&= \frac{1}{\eta}(x_{1,+}\varphi(x_{1,+}) - x_{1,-}\varphi(x_{1,-})) && (50)\\
&> 0\,. && (51)
\end{aligned}$$

其中，在最后一个不等式中，我们利用了以下观察：$x_{1,-} < 0 < x_{1,+}$，$|x_{1,-}| > |x_{1,+}|$。因此，当 C 固定时，$h_+(C,\eta)$ 随 η 增大而增大。当 $C > C_0(\rho)$ 时，有 $h_+(C,\eta) > 1 - 2\Phi(-C_0(\rho)/2) = 1-\rho$。

当 $\mu_1 - \mu_2 = C\sigma_2 > 0$ 再次定义

$$x_{1,\pm} = \frac{x_{0,\pm} - \mu_1}{\sigma_1} = \frac{C\eta \pm \eta\sqrt{C^2\eta^2 + 2(\eta^2-1)\log\eta}}{\eta^2-1} \tag{52}$$

$$x_{2,\pm} = \frac{x_{0,\pm} - \mu_2}{\sigma_2} = \frac{C\eta^2 \pm \sqrt{C^2\eta^2 + 2(\eta^2-1)\log\eta}}{\eta^2-1}\,. \tag{53}$$

考虑 $h_-(C,\eta) = P_1(A) - P_2(A) = \Phi(x_{1,+}) - \Phi(x_{1,-}) - \Phi(x_{2,+}) + \Phi(x_{2,-})$。首先注意到 $\lim_{\eta\to 1^+} x_{1,+} = +\infty$ 以及 $\lim_{\eta\to 1^+} x_{1,-} = -C/2$。这表明

$$\lim_{\eta\to 1^+} h_-(C,\eta) = 1 - 2\Phi(-\frac{C}{2}) \tag{54}$$

仍然成立。注意，我们仍有 $x_{2,\pm} = x_{1,\pm}/\eta + C$。用同样的方法对 η 求导，可得

$$\frac{\partial h_-(C,\eta)}{\partial \eta} = \frac{1}{\eta}(x_{1,+}\varphi(x_{1,+}) - x_{1,-}\varphi(x_{1,-}))\,. \tag{55}$$

由于 $x_{1,-} < 0 < x_{1,+}$，偏导数仍为正。因此，用类似的论证可得该结果。结合这两种情况，我们完成了对情形 $d = 1$的证明。

最后，当 $d \geq 2$ 时，我们考虑投影到由$\mu_2 - \mu_1$张成的一维空间上的 x 集合。

当 $\sigma_2 = \sigma_1$ 时，集合 $A = \{x \in \mathbb{R}^d : p_1(x) \geq p_2(x)\}$ 是半空间

$$A = \left\{x \in \mathbb{R}^d : \langle x, \frac{\mu_2 - \mu_1}{||\mu_2 - \mu_1||}\rangle \leq \tfrac{\mu_1+\mu_2}{2}\right\}$$

因此，$P_1(A) - P_2(A) = 1 - 2\Phi(-\frac{C}{2}) \geq 1-\rho$，因为 $C \geq C_0(\rho) = 2\Phi^{-1}(1-\rho/2)$。当且仅当 $C = C_0(\rho)$ 时，等号成立。

当 $\sigma_2/\sigma_1 > 1$ 时，考虑满足 $\tilde{\mu}_1 = \langle \mu_1, \frac{\mu_2-\mu_1}{||\mu_2-\mu_1||}\rangle$ 的 $\tilde{P}_1 = N(\tilde{\mu}_1, \sigma_1^2)$ 和满足 $\tilde{\mu}_2 = \langle \mu_2, \frac{\mu_2-\mu_1}{||\mu_2-\mu_1||}\rangle$ 的 $\tilde{P}_2 = N(\tilde{\mu}_2, \sigma_2^2)$。利用 P_1, P_2，根据方程 (37) 计算 $\tilde{x}_{0,\pm}$，并令集合

$$A = \left\{x \in \mathbb{R}^d : \tilde{x}_{0,-} \leq \langle x, \frac{\mu_2 - \mu_1}{||\mu_2 - \mu_1||}\rangle \leq \tilde{x}_{0,+}\right\}\,, \quad \tilde{A} = \{x \in \mathbb{R} : \tilde{x}_{0,-} \leq x \leq \tilde{x}_{0,+}\}\,. \tag{56}$$

注意， $|\tilde{\mu}_2 - \tilde{\mu}_1| = ||\mu_2 - \mu_1|| = C\max\{\sigma_1, \sigma_2\}$ ，因此

$$P_1(A) - P_2(A) = \tilde{P}_1(\tilde{A}) - \tilde{P}_2(\tilde{A}) > 1 - \rho \,. \tag{57}$$

□

引理 9. 在引理 *8* 的相同设定下，若 $C = ||\mu_1 - \mu_2||/(\sigma_1 + \sigma_2) \geq C_0(\rho)/2$，则存在集合 A，使得 $P_1(A) - P_2(A) \geq 1 - \rho$。当且仅当 $\sigma_1 = \sigma_2$ 且 $C = C_0(\rho)/2$ 时，等号成立。

引理 *9* 的证明。与引理 8 相比，证明基本没有变化。我们仍然考虑 $d = 1$ 的情形，并首先假设 $\sigma_1 \leq \sigma_2$。一旦这一点得到证明，其余部分可通过与引理 8 的证明类似的论证得到。令$x_{0,\pm}$ 与引理 8 证明中的相同。

当 $\mu_2 - \mu_1 = C(\sigma_1 + \sigma_2) > 0$ 时，现在我们应定义

$$x_{1,\pm} = \frac{x_{0,\pm} - \mu_1}{\sigma_1} = \frac{-C(\eta+1) \pm \eta\sqrt{C^2(\eta+1)^2 + 2(\eta^2-1)\log\eta}}{\eta^2 - 1} \tag{58}$$

$$x_{2,\pm} = \frac{x_{0,\pm} - \mu_2}{\sigma_2} = \frac{-C\eta(\eta+1) \pm \sqrt{C^2(\eta+1)^2 + 2(\eta^2-1)\log\eta}}{\eta^2 - 1} \,, \tag{59}$$

与引理 8 的证明类似，可得

$$h_+(C, \eta) = \Phi(x_{1,+}) - \Phi(x_{1,-}) - \Phi(x_{2,+}) + \Phi(x_{2,-}) \,, \tag{60}$$

and

$$\lim_{\eta \to 1^+} h_+(C, \eta) = 1 - 2\Phi(-C) \,. \tag{61}$$

由于 $x_{1,\pm} - \eta x_{2,\pm} = C(\eta + 1)$ 或 $x_{2,\pm} = x_{1,\pm}/\eta - C(1 + 1/\eta)$，因此

$$\frac{\partial x_{1,\pm}}{\partial \eta} - \eta\frac{\partial x_{2,\pm}}{\partial \eta} = \frac{\partial x_{1,\pm}}{\partial \eta} - \eta\left[-\frac{1}{\eta^2}x_{1,\pm} + \frac{1}{\eta}\frac{\partial x_{1,\pm}}{\partial \eta} + \frac{C}{\eta^2}\right] = \frac{1}{\eta}x_{1,\pm} - \frac{C}{\eta} \,. \tag{62}$$

通过与方程（51）相同的计算，我们有

$$\frac{\partial h_+(C, \eta)}{\partial \eta} = \frac{1}{\eta}\left((x_{1,+} - C)\varphi(x_{1,+}) - (x_{1,-} - C)\varphi(x_{1,-})\right) \,. \tag{63}$$

注意

$$x_{1,+} - C = \frac{-C(\eta+1) - C(\eta^2-1) + \eta\sqrt{C^2(\eta+1)^2 + 2(\eta^2-1)\log\eta}}{\eta^2 - 1} \tag{64}$$

$$= \frac{-C\eta(\eta+1)}{\eta - 1} + \frac{\eta\sqrt{C^2(\eta+1)^2 + 2(\eta^2-1)\log\eta}}{\eta^2 - 1} > 0 \tag{65}$$

$$x_{1,-} - C = \frac{-C\eta(\eta+1)}{\eta - 1} - \frac{\eta\sqrt{C^2(\eta+1)^2 + 2(\eta^2-1)\log\eta}}{\eta^2 - 1} < 0 \tag{66}$$

因此，（63）中的导数为正。通过相同的论证，有 $h_+(C, \eta) \geq 1 - 2\Phi(-C) > 1 - \rho$。

类似的论证可应用于 $\mu_1 > \mu_2$ 的情形，此处略去。对于等号成立的情形，当$\sigma_1 = \sigma_2$ 且 $C = C_0(\rho)/2$ 时，这与引理 8 中的证明相同。 □

引理 10。假设 $P_{1,\,2} = N_d(\mu_{1,\,2},\ \sigma^2_{1,\,2}I_d)\, 0 < \rho < 1$。令 F_d 为 ***Gamma***$(\frac{d}{2}, \frac{d}{2})$ 的累积分布函数，且 $\eta_0(\rho)$ 是下列方程的解

$$1 - \rho = F_d\left(\frac{2\eta^2\log\eta}{\eta^2 - 1}\right) - F_d\left(\frac{2\log\eta}{\eta^2 - 1}\right) \,. \tag{67}$$

如果 $\eta := \max\{\sigma_1/\sigma_2, \sigma_2/\sigma_1\} > \eta_0(\rho)$，那么存在一个集合 A，使得 $P_1(A) - P_2(A) > 1 - \rho$。

为证明引理 10，我们需要引用雷诺兹输运定理的一个推广版本。以下引理和所需定义可在 Flanders 的公式7.2中找到 [13]

引理 11。 设 Ω_t 是 $\mathbb{R}^d$ 中的一个 ℓ–维时变积分域，并且可表示为光滑映射 $(u,t) \to x(u,t)$ 的像，其中 u 遍历 $\mathbb{R}^\ell$ 中的一个固定域。设 ω 为一个外 ℓ–形式，它在局部坐标中可表示为

$$\omega = \sum_H a_H(x,t)dx^H, \quad dx^H = dx^{h_1} \wedge \cdots \wedge dx^{h_\ell} . \tag{68}$$

其中 H 是多重指标，且 $1 \le h_1 < h_2 < \cdots < h_\ell \le n$。于是

$$\frac{d}{dt}\int_{\Omega_t} \omega = \int_{\Omega_t} i_{\boldsymbol{v}}(d_x\omega) + \int_{\Omega_t} \dot{\omega} + \int_{\partial\Omega_t} i_v\omega \tag{69}$$

其中 $v = \partial x/\partial t$, i_v 表示与 v, $d_x\omega$ 的外微分，其中 ω 相对于 x ，而 $\dot{\omega} = \sum_H \frac{\partial a_H(x\ t}{\partial t}\ dx^{H,)}$

引理 *10* 的证明。 证明采用了与引理 8 的证明类似但更复杂的技巧。不失一般性，假设 $\sigma_2 = \eta\sigma_1 \ge \eta_0(\rho)\sigma_1 > \sigma_1$ 且

$$\mu_1 = (\mu_{11}, \cdots, \mu_{1d}), \quad \mu_2 = (C\sigma_2 + \mu_{21}, \cdots, \mu_{2d}). \tag{70}$$

根据对称性，我们还可以假设 $C \ge 0, \mu_{22} = \mu_{12}, \cdots, \mu_{2d} = \mu_{1d}$。

设 p_1, p_2 分别为 P_1, P_2 的概率密度。注意， p_1 恰好是标准d–维高斯分布的密度，我们将其记为 φ_d。于是，代数计算表明集合

$$A = \{x \in \mathbb{R}^d : p_1(x) \ge p_2(x)\} = \left\{x \in \mathbb{R}^d : ||x - \frac{\eta^2\mu_1 - \mu_2}{\eta^2 - 1}|| \le \frac{\sigma_2}{\eta^2 - 1}\sqrt{C^2\eta^2 + 2d(\eta^2 - 1)\log\eta}\right\} . \tag{71}$$

注意， A 是 d–维球。记

$$A_1(C,\eta) := \left\{x \in \mathbb{R}^d : ||x - (-\frac{C\eta}{\eta^2 - 1}, 0, \cdots, 0)|| \le \frac{\eta}{\eta^2 - 1}\sqrt{C^2\eta^2 + 2d(\eta^2 - 1)\log\eta}\right\} , \tag{72}$$

我们可以直接计算该概率:

$$\begin{aligned} P_1(A) &= \int_{x:||x - \frac{\eta^2\mu_1 - \mu_2}{\eta^2 - 1}|| \le \frac{\sigma_2}{\eta^2 - 1}\sqrt{C^2\eta^2 + 2d(\eta^2 - 1)\log\eta}} \frac{1}{(2\pi)^{\frac{d}{2}}\sigma_1^d} \exp\left(-\frac{||x - \mu_1||^2}{2\sigma_1^2}\right) dx \\ &= \int_{y:||y + \frac{C\eta}{\eta^2 - 1}|| \le \frac{\eta}{\eta^2 - 1}\sqrt{C^2\eta^2 + 2d(\eta^2 - 1)\log\eta}} \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{||y||^2}{2}\right) dy \qquad (73) \\ &= \int_{A_1(C,\eta)} \varphi_d(x)dx , \qquad (74) \end{aligned}$$

其中， φ_d 是 $N_d(0, I_d)$分布的密度。类似地，令 $A_2(C,\eta) = \{x \in \mathbb{R}^d : ||x - (-C\eta^2/(\eta^2 - 1), 0, \cdots, 0)|| \le \sqrt{C^2\eta^2 + 2d(\eta^2 - 1)\log\eta/(\eta^2 - 1)}\}$ ，我们有

$$P_2(A) = \int_{A_2(C,\eta)} \varphi_d(x)dx . \tag{75}$$

令 $h(C,\eta) = P_1(A) - P_2(A)$，因为后者仅由 C 和 η确定。现在我们证明，对于任意固定的 $\eta \ge \eta_0(\rho)$, $h(C,\eta)$ 关于 $C \ge 0$单调递增。

为证明这一点，我们考虑对 $h(C,\eta)$ 关于 C求偏导数。

$$\frac{\partial}{\partial C}h(C,\eta) = \frac{\partial}{\partial C}\int_{A_1(C,\eta)} \varphi_d(x)dx - \frac{\partial}{\partial C}\int_{A_2(C,\eta)} \varphi_d(x)dx , \tag{76}$$

我们首先处理 $\frac{\partial}{\partial C}\int_{A_1(C,\eta)} \varphi_d(x)dx$. $A_1(C,\eta)$ 是一个超球。设 n 为用极坐标表示的法向量，则可以进行如下变量代换

$$\begin{aligned} x : (n, r) &\to \mathbb{R}^d \qquad (77) \\ x_1 &= o_1 + r\cos\theta_1 \qquad (78) \\ x_2 &= r\sin\theta_1\cos\theta_2 \qquad (79) \\ &\cdots \qquad (80) \\ x_{d-2} &= r\sin\theta_1 \cdots \sin\theta_{d-2}\cos\theta_{d-1} \qquad (81) \\ x_d &= r\sin\theta_1 \cdots \sin\theta_{d-2}\sin\theta_{d-1} \qquad (82) \end{aligned}$$

其中 $o_1 = -C\eta/(\eta^2-1)$ 且 $0 \le r \le R_1$ 满足

$$R_1 = \frac{\eta}{\eta^2-1}\sqrt{C^2\eta^2 + 2d(\eta^2-1)\log\eta}\,. \tag{83}$$

对于每个 $C>0$，在 $C \in (C_a,\ C_b)$ 附近都存在一个小邻域，使得 $x = (x_1,\ \cdots,\ x_d)$ 是定义在 $[0,R_1]\times[0,\pi]^{d-2}\times[0,2\pi]\times(C_a,\ C_b)$ 上的连续可微映射 $\to \mathbb{R}^d$。此处便用到了引理 11。考虑 d–形式 $\omega = \varphi_d(x)dx_1\wedge\cdots dx_d$。由于 ω 是$\mathbb{R}^d$上的一个 d–形式，因此它关于 $x\in\mathbb{R}^d$ 的外微分为零。又由于 φ_d 与 C 无关，故 $\dot\omega = 0$。

因此，方程（69）可简化为

$$\frac{\partial}{\partial C}\int_{A_1(C,\eta)}\omega = \int_{\partial A_1(C,\eta)} i_{v_1}(\omega). \tag{84}$$

其中 $v_1 = \partial x/\partial C$。

现在我们来确定 $i_{v_1}(\omega)$。由于 $\partial A_1(C,\ \eta)$ 是一个维数为 $d-1$ 的超球面，因此它是可定向的。此外，考虑 $\partial A_1(C,\ \eta)$ 的切空间的以下基

$$e_i = \frac{\partial}{\partial\theta_i},\quad 1\le i\le d-1 \tag{85}$$

以及指向外侧的单位法向量 $n = \partial/\partial r$。那么对于任意向量场 $t = \sum_{1\le i\le d-1} t_i e_i + t_d n$，根据内积的定义以及 ω 是一个交替的 d–线性张量这一事实，有

$$\begin{aligned}
i_t(\omega)(e_1,\cdots,e_{d-1}) =& \omega(t,e_1,\cdots,e_{d-1}) && (86)\\
=& \sum_{i=1}^{d-1} t_i\omega(e_i,e_1,\cdots,e_{d-1}) + t_d\omega(n,e_1,\cdots,e_{d-1}) && (87)\\
=& \varphi_d(x)t_d(dx_1\wedge\cdots dx_d)(n,e_1,\cdots,e_{d-1}) && (88)\\
=& \varphi_d(x)\langle t,n\rangle d\Sigma(n)(e_1,\cdots,e_{d-1}) && (89)
\end{aligned}$$

其中，我们将 $d\Sigma(n) = i_n(dx_1\wedge\cdots\wedge dx_n)$ 记作面积元。换言之，$d\Sigma(n)$ 是满足以下条件的微分形式

$$\begin{aligned}
d\Sigma(n)(e_1,\cdots,e_{d-1}) &= (dx_1\wedge\cdots\wedge dx_d)(n,e_1,\cdots,e_{d-1}) && (90)\\
&= R_1^{d-1}\sin^{d-2}\theta_1\cdots\sin\theta_{d-2}(dr\wedge d\theta_1\wedge\cdots\wedge d\theta_{d-1})(n,e_1,\cdots,e_{d-1}) && (91)\\
&= R_1^{d-1}\sin^{d-2}\theta_1\cdots\sin\theta_{d-2}\,, && (92)
\end{aligned}$$

这表明 $d\Sigma(n) = R_1^{d-1}\sin^{d-2}\theta_1\cdots\sin\ \theta_{d-2}d\theta_1\wedge\cdots\wedge d\theta_{d-1}$。那么，由于 $i_{v_1(\omega)} = \varphi_d(x)\langle v_1,n\rangle d\Sigma(n)$，

$$\frac{\partial}{\partial C}\int_{A_1(C,\eta)}\omega = \int_{\mathbb{S}^{d-1}} R_1^{d-1}\varphi_d(x(n,R_1))\langle v_1,n\rangle d\Sigma(n)\,. \tag{93}$$

类似地，$A_2(C,\eta)$ 可表示为

$$\begin{aligned}
y:(n,r)&\to\mathbb{R}^d && (94)\\
y_1 &= o_2 + r\cos\theta_1 && (95)\\
y_2 &= r\sin\theta_1\cos\theta_2 && (96)\\
&\cdots && (97)\\
y_{d-1} &= r\sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1} && (98)\\
y_d &= r\sin\theta_1\cdots\sin\theta_{d-2}\sin\theta_{d-1} && (99)
\end{aligned}$$

其中 $o_2 = -C\eta^2/(\eta^2-1)$ 和 $0\le r\le R_2$，其中

$$R_2 = \frac{1}{\eta^2-1}\sqrt{C^2\eta^2 + 2d(\eta^2-1)\log\eta}\,. \tag{100}$$

同理，我们有

$$\frac{\partial}{\partial C}\int_{A_2(C,\eta)}\omega=\int_{\partial A_2(C,\eta)}\varphi_d(y(n,R_2))\langle v_2,n\rangle d\Sigma(n)=\int_{\mathbb{S}^{d-1}}R_2^{d-1}\varphi_d(y(n,R_2))\langle v_2,n\rangle d\Sigma(n)\,. \tag{101}$$

其中 $v_2=\partial y/\partial C$。

我们注意到两个恒等式。首先，对于相同的

$$n=(\cos\theta_1,\sin\theta_1\cos\theta_2,\cdots,\sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1},\sin\theta_1\cdots\sin\theta_{d-1})$$

成立

$$\begin{aligned}
&\frac{\varphi_d(x(n,R_1))}{\varphi_d(y(n,R_2))}\\
=&\frac{\exp\left(-\frac{1}{2}||x(n,R_1)||^2\right)}{\exp\left(-\frac{1}{2}||y(n,R_2)||^2\right)} &(102)\\
=&\frac{\exp\left(-\frac{1}{2}\left[o_1^2+R_1^2+2o_1R_1\cos\theta_1\right]\right)}{\exp\left(-\frac{1}{2}\left[o_2^2+R_2^2+2o_2R_2\cos\theta_1\right]\right)} &(103)\\
=&\frac{\exp\left(-\frac{1}{2}\left[\frac{C^2\eta^2}{(\eta^2-1)^2}+\frac{\eta^2}{(\eta^2-1)^2}(C^2\eta^2+2d(\eta^2-1)\log\eta)-2\frac{C\eta^2}{(\eta^2-1)^2}\sqrt{C^2\eta^2+2d(\eta^2-1)\log\eta}\cos\theta_1\right]\right)}{\exp\left(-\frac{1}{2}\left[\frac{C^2\eta^4}{(\eta^2-1)^2}+\frac{1}{(\eta^2-1)^2}(C^2\eta^2+2d(\eta^2-1)\log\eta)-2\frac{C\eta^2}{(\eta^2-1)^2}\sqrt{C^2\eta^2+2d(\eta^2-1)\log\eta}\cos\theta_1\right]\right)} &(104)\\
=&\exp\left(-\frac{1}{2}2d\log\eta\right) &(105)\\
=&\frac{1}{\eta^d} &(106)
\end{aligned}$$

其次，对于相同的 $\boldsymbol{n}$，有 $\boldsymbol{x}(\boldsymbol{n},\ R_1)-\eta\boldsymbol{y}(\boldsymbol{n},\ R_2)=(C\eta,\ 0,\ \cdots,\ 0)$。这表明 $\boldsymbol{v}_1-\eta\boldsymbol{v}_2=(\eta,\ 0,\ \cdots,\ 0)$

因此

$$\begin{aligned}
&\frac{\partial h(C,\ \eta)}{\partial C} &(107)\\
=&\int_{\mathbb{S}^{d-1}}\left(R_1^{d-1}\varphi_d(\boldsymbol{x}(\boldsymbol{n},\ R_1))\langle\boldsymbol{v}_1,\boldsymbol{n}\rangle-R_2^{d-1}\varphi_d(\boldsymbol{y}(\boldsymbol{n},\ R_1))\langle\boldsymbol{v}_2,\boldsymbol{n}\rangle\right)d\Sigma(\boldsymbol{n}) &(108)\\
=&\int_{\mathbb{S}^{d-1}}\left(R_1^{d-1}\varphi_d(\boldsymbol{x}(\boldsymbol{n},R_1))\langle\boldsymbol{v}_1,\boldsymbol{n}\rangle-\begin{pmatrix}R_1\\ \underline{\eta}\end{pmatrix}_{d-1}\eta^d\varphi_d(\boldsymbol{x}(\boldsymbol{n},\ R_1))\langle\boldsymbol{v}_2,\boldsymbol{n}\rangle\right)d\Sigma(\boldsymbol{n}) &(109)\\
=&\int_{\mathbb{S}^{d-1}}R_1^{d-1}\varphi_d(\boldsymbol{x}(\boldsymbol{n},\ R_1))\langle\boldsymbol{v}_1-\eta\boldsymbol{v}_2,\boldsymbol{n}\rangle d\Sigma(\boldsymbol{n}) &(110)\\
=&\int_{\mathbb{S}^{d-1}}R_1^{d-1}\eta\,\text{余弦}\,\theta_1\varphi_d(\boldsymbol{x}(\boldsymbol{n},R_1))d\Sigma(\boldsymbol{n}) &(111)\\
=&R_1^{d-1}\eta\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\cos\,\theta_1\frac{2\pi)^{\frac{d}{2}}\text{指数}\left(-\frac{1}{2}[o_1^2+R_1^2+2o_1R_1\,\text{余弦}\,\theta_1]\right)\text{正弦}^{d-2}\,\theta_1\cdots\text{正弦}}{(\,\theta_{d-\bar{2}}d\theta_1d\theta_2\cdots d\theta_{d-1}} &(112)\\
=&\Xi\int_0^{\pi}\cos\theta_1\exp(-o_1R_1\cos\theta_1)\sin^{d-2}\theta_1d\theta_1\,, &(113)
\end{aligned}$$

其中 Ξ 是一个正常数。最后，

$$\int_0^{\pi} \cos\theta_1 \exp(-o_1 R_1 \cos\theta_1) \sin^{d-2}\theta_1 d\theta_1 \tag{114}$$

$$= \int_0^{\frac{\pi}{2}} \cos\theta_1 \exp(-o_1 R_1 \cos\theta_1) \sin^{d-2}\theta_1 d\theta_1 + \int_{\frac{\pi}{2}}^{\pi} \cos\theta_1 \exp(-o_1 R_1 \cos\theta_1) \sin^{d-2}\theta_1 d\theta_1 \tag{115}$$

$$= \int_0^{\frac{\pi}{2}} \cos\theta_1 \exp(-o_1 R_1 \cos\theta_1) \sin^{d-2}\theta_1 d\theta_1 - \int_0^{\frac{\pi}{2}} \cos\theta_1 \exp(o_1 R_1 \cos\theta_1) \sin^{d-2}\theta_1 d\theta_1 \tag{116}$$

$$= \int_0^{\frac{\pi}{2}} \cos\theta_1 \sin^{d-2}\theta_1 \left[\exp(-o_1 R_1 \cos\theta_1) - \exp(o_1 R_1 \cos\theta_1)\right] d\theta_1 > 0 \tag{117}$$

其中，最后一个不等式是因为 $o_1 < 0$，且对于任意 $x > 0$，$e^x > 1 > e^{-x}$。这表明，当 $C > 0$ 时，$h(C, \eta)$ 关于 C 单调递增。因此，只要 $h(0, \eta) > 1 - \rho$，$P_1(A) - P_2(A) > 1 - \rho$。接下来只需证明，只要 $eta \geq \eta_0$，$h(0, \eta) \geq 1 - \rho$。当 $C = 0$ 时，这显然成立，

$$A_1 = \left\{ x \in \mathbb{R}^d : ||x|| \leq \sqrt{\frac{2d\eta^2 \log\eta}{\eta^2 - 1}} \right\} \tag{118}$$

$$A_2 = \left\{ x \in \mathbb{R}^d : ||x|| \leq \sqrt{\frac{2d \log\eta}{\eta^2 - 1}} \right\} \tag{119}$$

令 $X \sim N_d(0, I_d)$，则 $||X||^2/d \sim Gamma(\frac{d}{2}, \frac{d}{2})$。于是

$$h(0,\eta) = Pr\left(||X||^2 \leq \frac{2d\eta^2 \log\eta}{\eta^2 - 1} \right) - Pr\left(||X||^2 \leq \frac{2d \log\eta}{\eta^2 - 1} \right) = F_d(\frac{2\eta^2 \log\eta}{\eta^2 - 1}) - F_d(\frac{2\log\eta}{\eta^2 - 1}) . \tag{120}$$

根据引理 21，$h(0, \eta)$ 关于 η 单调递增。因此，$h(0, \eta) \geq h(0, \eta_0(\rho)) = 1 - \rho$，证明完毕。

□

A.2.2 分量逐一比较定理

证明我们的结果所需的第二个关键要素，是一组针对全变差距离较小的两个混合分布的逐分量比较引理。在以下引理中，考虑两个混合分布 $P = \sum_{i=1}^{K} \pi_i P_i$ 和 $P' = \sum_{j=1}^{K'} \pi'_j P_j$，并记 $\pi_{\min} = \min_{i,j}\{\pi_i, \pi'_j\}$ 和 $\pi_{\max} = \max_{i,j}\{\pi_i, \pi'_j\}$。证明见附录 A；这些引理主要表明：(1) 若两个混合分布彼此接近，则分量全变差差异的加权和有上界；(2) 当两个充分分离的混合分布具有相同数量的分量时，在特定的对应条件下，它们的混合比例彼此接近。

引理 12. 设 $TV(P, P') \leq 2\epsilon$。对于任意 $j \in [K']$，若存在一组集合$\{A_{ij}\}_{i\in[K]}$，使得对所有 $i \in [K]$均有 $\rho_{ij} = 1 - P_i(A_{ij}) + P'_j(A_{ij}) \in [0,1]$，则

$$\sum_{i=1}^{K} \max\{\pi_i, \pi'_j\}\rho_{ij} \geq \pi'_j - 2\epsilon , \quad \sum_{i=1}^{K} \pi_i \rho_{ij} \geq \pi_{\min} - 2\epsilon . \tag{121}$$

类似地，若存在一组集合 $\{B_{ij}\}_{i\in[K]}$，使得对所有 $\in [0,1]$，$\rho_{ij} = 1 - P'_j(B_{ij}) + P_i(B_{ij})$，方程也成立。

该引理源于这样一个事实：不存在某个集合，它在 P_i 中的 P 分量下具有较大的概率质量，却在 P' 的每个分量 P'_j 下都只有较小的概率质量。A_{ij} 是在 P_i 下具有较大概率质量、但在 P'_j 下概率质量较小的集合，而 B_{ij} 则恰好相反。交换 P 和 P' 的角色，可以得到一个对称的结果。

当某个分量 P'_j与所有分量 $P_i, i \neq j$ 之间的全变差距离都很大时，引理 12 尤其有用；此时，根据引理 12，可以对 P'_j与 P中剩余分量 P_j之间的全变差距离给出上界。

引理 13。设 $TV(P,P') \le 2\epsilon$，并记 $\pi_{\min} = \min_{i,j}\{\pi_i, \pi'_j\}$。考虑两个分量 P_{j_0}, P'_j，其中 $\pi_{j_0} \le \pi'_j$。若对所有 $i \in [K], i \ne j_0$，都存在一个集合 A_{ij}，使得 $P_i(A_{ij}) - P'_j(A_{ij}) \ge 1-\rho$，其中某个常数 ρ 满足 $0 \le \rho \le (\pi_{\min} - 2\epsilon)/(1-\pi_{\min})$，则有

$$TV(P_{j_0}, P'_j) \le \frac{2\epsilon}{\pi_{\min}} + \frac{1-\pi_{\min}}{\pi_{\min}}\rho\,. \tag{122}$$

现在，我们得到了一个关于对应分量之间混合比例差异的逐分量比较结果。

引理 *14.* 假设 $TV(P,P') \le 2\epsilon$。记 $\pi_{\min} = \min_{i,j}\{\pi_i, \pi'_j\}$ 和 $\pi_{\max} = \max_{i,j}\{\pi_i, \pi'_j\}$。如果存在一对分量 P_i, P'_i 以及一个集合 A，使得对于某个 $\rho \in (0,1)$，$\min\{P_i(A)\ P'_i(A)\} \ge 1-\rho$ 且 $\max_{j\ne i}\{P_j(A)\ P'_j(A)\} \le \rho$，那么

$$|\pi_i - \pi'_i| \le 2\epsilon + (1 + \pi_{\max} - \pi_{\min})\rho \tag{123}$$

引理 *12* 的证明。令 $Q = (P+P')/2$。则 $TV(P,\ Q) \le \epsilon, TV(P',\ Q) \le \epsilon$。令 P'_j 为 P' 的任意一个分量。我们首先通过反证法证明 (121) 的前半部分成立。假设 $\sum_{i=1}^K \max\{\pi_i, \pi'_j\}\rho_{ij} < \pi'_j - 2\epsilon$。现在考虑集合 $A = \cap_{i=1}^K A_{ij}$。一方面，

$$Q(A) \ge \pi'_j P'_j((\cup_{i=1}^K A_{ij})^{\complement}) - \epsilon \ge \pi'_j(1 - \sum_{i=1}^K P'_j(A_{ij})) - \epsilon = \pi'_j - \pi'_j \sum_{i=1}^K P'_j(A_{ij}) - \epsilon := LB \tag{124}$$

同时，

$$Q(A) \le \sum_{i=1}^K \pi_i P_i(A_i^{\complement}) + \epsilon = 1 - \sum_{i=1}^K \pi_i P_i(A_i) + \epsilon := UB \tag{125}$$

由于对于每个 $i \in [K]$，都有 $P_i(A_{ij}) - P'_j(A_{ij}) = 1-\rho_{ij}$，因此根据引理 20，$\pi_i P_i(A_{ij}) - \pi'_j P'_j(A_{ij}) \ge \pi_i - \max\{\pi_i, \pi'_j\}\rho_{ij}$。因此，将引理 20 应用于所有分量 $i \in [K]$，可得

$$LB - UB = \pi'_j - 1 - 2\epsilon + \sum_{i=1}^K \left(\pi_i P_i(A_{ij}) - \pi'_j P'_j(A_{ij})\right) \tag{126}$$

$$\ge \pi'_j - 1 - 2\epsilon + \sum_{i=1}^K \left(\pi_i - \max\{\pi_i, \pi'_j\}\rho_{ij}\right) \tag{127}$$

$$= \pi'_j - 1 - 2\epsilon + 1 - \sum_{i=1}^K \max\{\pi_i, \pi'_j\}\rho_{ij} > 0\,. \tag{128}$$

这与 $Q(A) \ge LB > UB \ge Q(A)$ 矛盾。因此，(121) 的前半部分对 P 成立 □

对于 (121) 的后半部分，若 $\sum_{i=1}^K \pi_i\rho_{ij} < \pi_{\min} - 2\epsilon$，则考虑 $Q(A)$ 的另一种下界

$$Q(A) \ge \pi'_j P'_j((\cup_{i=1}^K A_{ij})^{\complement}) - \epsilon \ge \pi_{\min} P'_j((\cup_{i=1}^K A_{ij})^{\complement}) - \epsilon \ge \pi_{\min} - \pi_{\min}\sum_{i=1}^K P'_j(A_{ij}) - \epsilon := LB';, \tag{129}$$

然后注意到

$$LB' - UB \ge \pi_{\min} - 1 - 2\epsilon + \sum_{i=1}^K \pi_i P_i(A_{ij}) - \pi_{\min} P'_j(A_{ij}) \tag{130}$$

$$\ge \pi_{\min} - 1 - 2\epsilon + \sum_{i=1}^K \pi_i(1-\rho_{ij}) \tag{131}$$

$$= \pi_{\min} - 1 - 2\epsilon + 1 - \sum_{i=1}^K \pi_i\rho_{ij} > 0\,, \tag{132}$$

这就构成了矛盾，并完成了 (121) 后半部分的证明。 □

引理 *13* 的证明。定义

$$\rho_1 = \frac{\pi_{\min} - 2\epsilon}{\pi_{\min}} - \frac{1-\pi_{\min}}{\pi_{\min}}\rho \tag{133}$$

注意，$0 \le \rho \le (\pi_{\min} - 2\epsilon)/(1-\pi_{\min})$ 意味着 $0 \le \rho_1 \le 1$。我们将通过反证法证明，全变差距离 $TV(P_{j_0}, P'_j) \le 1-\rho_1$。否则，若 $TV(P_{j_0}, P'_j) > 1-\rho_1$，则可以选取一个集合 $A_{j_0 j}$，使得 $P_{j_0}(A_{j_0 j}) - P'_j(A_{j_0 j}) > 1-\rho_1$。定义函数

$$f(x) = x(\rho_1 - 1) + \sum_{\substack{i\in[K]\\ i\neq j_0}} \max\{\pi_i, x\}\rho\,, \tag{134}$$

于是，$f(x)$ 是一个分段线性函数，其斜率的上界为

$$\rho_1 - 1 + \sum_{i:i\neq j}\rho = \rho_1 + (K-1)\rho - 1 \le \rho_1 + \left(\frac{1-\pi_{\min}}{\pi_{\min}}\right)\rho - 1 = \frac{\pi_{\min}-2\epsilon}{\pi_{\min}} - 1 < 0\,. \tag{135}$$

因此，$f(x)$ 关于 x 单调递减。由于 $\pi_{j_0} \le \pi'_j$，根据引理 12 之后的注释，我们有

$$f(\pi'_j) = \pi'_j\rho_1 + \sum_{\substack{i\in[K]\\ i\neq j_0}} \max\{\pi_i, \pi'_j\}\rho - \pi'_j \tag{136}$$

$$> \sum_{i=1}^{K} \max\{\pi_i, \pi'_j\}(1 - P_i(A_{ij}) + P'_j(A_{ij})) - \pi'_j \tag{137}$$

$$\ge -2\epsilon\,. \tag{138}$$

于是我们得出 $f(\pi_{\min}) \ge f(\pi'_j) > -2\epsilon$。但我们可以明确计算出

$$f(\pi_{\min}) = \pi_{\min}\left(\frac{\pi_{\min}-2\epsilon}{\pi_{\min}} - \frac{1-\pi_{\min}}{\pi_{\min}}\rho - 1\right) + \sum_{\substack{i\in[K]\\ i\neq j_0}} \max\{\pi_i, \pi_{\min}\}\rho \tag{139}$$

$$= -2\epsilon - (1-\pi_{\min})\rho + \rho\sum_{\substack{i\in[K]\\ i\neq j_0}} \pi_i \tag{140}$$

$$\le -2\epsilon - (1-\pi_{\min})\rho + (1-\pi_{\min})\rho \tag{141}$$

$$= -2\epsilon\,. \tag{142}$$

这一矛盾完成了对 (122) 的证明。 □

证明。引理 14] 的证明。注意，$\pi_i(A)^{\complement} \le \rho$ 且 $\pi'_i(A)^{\complement} \le \rho$

$$|\pi_i - \pi'_i| = |\pi_i P_i(A) - \pi'_i P'_i(A) + \pi_i P_i(A^{\complement}) - \pi'_i P'_i(A^{\complement})| \tag{143}$$

$$\le |P(A) - \sum_{j\neq i}\pi_j P_j(A) - P'(A) + \sum_{j\neq i}\pi'_j P'_j(A)| + |\pi_i P_i(A^{\complement}) - \pi'_i P'_i(A^{\complement})| \tag{144}$$

$$\le 2\epsilon + \max\{\sum_{j\neq i}\pi_j P_j(A), \sum_{j\neq i}\pi'_j P'_j(A)\} + \max\{\pi_i P_i(A^{\complement}), \pi'_i P'_i(A^{\complement})\} \tag{145}$$

$$\le 2\epsilon + (1 - \pi_{\min} + \pi_{\max})\rho\,. \tag{146}$$

□

A.3 初始化

令 c_0 如式（7）所定义，则以下定理保证：对于 P'的每个分量，都存在 P的一个分量，使其在总变差距离上与之接近。从这里开始，在附录A的其余部分中，我们将作如下假设：

B1 $P \in \mathcal{M}(K,\ \pi_{\min},\ \pi_{\max},\ c)$ 和 $P' \in \mathcal{M}(K',\ \pi_{\min},\ \pi_{\max},\ c)$。
B2 存在一个分布 Q，使得 $TV(P,\ Q \le \epsilon$ 且 $TV(P',\ Q) \le \epsilon$。B3 $\max\{K, K'\} \le 1/\pi_{\min}, \pi_{\max} \le 1 - (\min\{K, K'\} - 1)\pi_{\min}$。

定理 15. 假设 *B1-B3* 成立。对于任意分量 P'_j，必定存在一个分量 P_i，使得$TV(P_i, P'_j) \le 1 - \pi_{\min} + 2\epsilon$

定理*15*的证明. 不失一般性，我们证明关于 P'_1 的结论。假设定理的结论不成立，即对于任意分量 P_i，都有 $TV(P_i, P'_1) > 1 - \pi_{\min} + 2\epsilon$，这意味着存在一个集合 A_i，使得 $P_i(A_i) - P'_1(A_i) > 1 - \pi_{\min} + 2\epsilon$，或者 $\pi_{\min} - 2\epsilon > 1 - P_i(A_i) + P'_1(A_i)$。接着，对集合族 $\{A_i\}_{i\in[K]}$ 应用引理 12，可得

$$\pi_{\min} - 2\epsilon > \sum_{i=1}^{K} \pi_i (1 - P_i(A_i) + P_1(A'_i)) \ge \pi_{\min} - 2\epsilon \ . \tag{147}$$

因此，所需结果成立。 □

第二个定理指出，若 P 和 P' 均充分分离且分量近乎均衡，那么定理 15 所建立的匹配是一一对应的。具体而言，对于 P'中的任意一个分量，在 P中都存在且仅存在一个分量，使得它们的中心彼此接近。

定理 16. 假设 *B1-B3* 成立，若 $c > \eta_0 c_0$，其中 c_0、η_0 分别由(7) 和(8) 定义，则 $K = K'$，且存在一个置换 $\theta\colon [K] \to [K]$，使得 $||\mu_i - \mu'_{\theta(i)}|| \le c_0 \max\{\sigma_i, \sigma'_{\theta(i)}\}$且$\max\{\sigma_i/\sigma'_{\theta(i)}, \sigma'_{\theta(i)}/\sigma_i\} \le \eta_0$。

定理*16*的证明. 根据定理 15 和引理 8，每个 P'_j 都必须与至少一个分量 P_i 匹配，即它们的中心距离 $||\mu_i - \mu'_j|| \le c_0 \max\{\sigma_i, \sigma'_j\}$。注意，定理 15 对 P和 P'是对称的。对于每个 P_i，必定存在一个 P'_j，使得它们的中心彼此接近。

现在考虑以下移除过程。对于分量 P_j，如果存在唯一一个与 P_j匹配的 P'_j，且 P中只有分量 P_j 与 P'_j匹配，那么移除 P_j和 P'_j。对于被移除的一对 P_j，P'_j，可以将它们的标准差之比上界为 $\max\{\sigma_j, \sigma'_j\} \le \eta_0$。为说明这一点，不失一般性，假设 $\sigma_j \le \sigma'_j$。那么对于任意 $i \ne j$，设 P'_i为一个与 P_i匹配的分量。中心之间的距离为

$$||\mu'_i - \mu_j|| \ge ||\mu'_i - \mu'_j|| - ||\mu'_j - \mu_j|| \ge c(\sigma'_i + \sigma'_j) - c_0\sigma'_j = c\sigma'_i + (c - c_0)\sigma'_j \ . \tag{148}$$

- 如果 $\sigma'_i \ge \sigma_j$，则有 $||\mu'_i - \mu_j|| \ge c\max\{\sigma'_i, \sigma_j\}$。根据引理 8，可以选取集合 A_{ij}，使得 $P'_i(A_{ij}) - P_j(A_{ij}) > 1 - 2\Phi(-c/2) \ge 1 - 2\Phi(-c_0\eta_0/2)$
- 若 $\sigma'_i < \sigma_j$，则由于 P'_i未与 P_j匹配

$$||\mu'_i - \mu_j|| \ge c_0 \max\{\sigma'_i, \sigma_j\} = c_0\sigma_j \ . \tag{149}$$

将 (148) 和 (149) 相加，得到

$$||\mu'_i - \mu_j|| \ge \frac{1}{2}\left(c\sigma'_i + (c - c_0)\sigma'_j + c_0\sigma_j\right) \tag{150}$$

$$\ge \frac{1}{2}c(\sigma'_i + \sigma_j) \ . \tag{151}$$

根据引理 9，可以选取一个集合 A_{ij}，使得 $P'_i(A_{ij}) - P_j(A_{ij}) \ge 1 - 2\Phi(-c_0\eta_0/2)$。

因此，在这两种情况下，对于任意 $i \ne j$，都存在 $P'_i(A_{ij}) - P_j(A_{ij}) \ge 1 - \rho_2$，满足 $\rho_2 = 2\Phi(-c_0\eta_0/2) \in (0,\ 1)$。

现在假设 $\max\{\sigma_j, \sigma'_j\}/\min\{\sigma_j, \sigma'_j\} > \eta_0$，根据引理 10，可以选取一个集合 A_{jj}，使得$P'_j(A_{ij}) - P_j(A_{ij}) > 1 - \rho_1$，其中 ρ_1 定义为

$$\rho_1 := \frac{\pi_{\min} - 2\epsilon}{\pi_{\max}} - \frac{2(1 - \pi_{\max})}{\pi_{\max}}\Phi(-\frac{1}{2}\eta_0 c_0) \ . \tag{152}$$

注意， $1-\rho_1$ 是方程 (8) 的左侧。此外， $\rho_1 < \pi_{\min}/\pi_{\max} < 1$ 和

$$\rho_1 > (\pi_{\min} - 2\epsilon)\left(\frac{1}{\pi_{\max}} - \frac{1-\pi_{\max}}{\pi_{\max}}\right) = \pi_{\min} - 2\epsilon > \rho_2 \tag{153}$$

根据引理 12,

$$\pi'_j\rho_1 + \sum_{i:i\neq j}\pi'_i\rho_2 > \sum_{i=1}^{K}\pi'_i(1 - P'_i(A_{ij}) + P_j(A_{ij})) \geq \pi_{\min} - 2\epsilon;. \tag{154}$$

然而，式 (154) 左侧满足

$$\pi'_j\rho_1 + \sum_{i:i\neq j}\pi'_i\rho_2 = \pi'_j(\rho_1 - \rho_2) + \rho_2 \leq \pi_{\max}\rho_1 + (1-\pi_{\max})\rho_2 = \pi_{\min} - 2\epsilon \tag{155}$$

矛盾表明 $\max\{\sigma_j/\sigma'_j, \sigma'_j/\sigma_j\} \leq \eta_0$。

现在假设我们完成了这个移除过程，但仍有剩余分量。不失一般性，我们可以假设 P_1 是 P 和 P' 中所有剩余分量里标准差最小的分量。由于它没有被移除，因此存在一个 P'_1 和一个 P_2，使得 P_1 和 P_2 都与 P'_1 匹配。由于我们假设 $\sigma_1 \leq \sigma_2$，根据三角不等式

$$c(\sigma_1+\sigma_2) \leq ||\mu_1 - \mu_2|| \leq ||\mu'_1 - \mu_1|| + ||\mu'_1 - \mu_2|| \leq c_0\max\{\sigma_1, \sigma'_1\} + c_0\max\{\sigma_2, \sigma'_1\}. \tag{156}$$

注意，根据定理的假设， $c > c_0\eta_0 > c_0$，我们考虑以下三种情况:

- $\sigma'_1 \leq \sigma_1 \leq \sigma_2$. Then $c_0(\sigma_1+\sigma_2) < c(\sigma_1+\sigma_2) \leq c_0(\sigma_1+\sigma_2)$ and this is impossible.
- $\sigma_1 < \sigma'_1 \leq \sigma_2$. Then $c\sigma_1 + c_0\sigma_2 < c(\sigma_1+\sigma_2) \leq c_0\sigma'_1 + c_0\sigma_2$, implying that $\sigma'_1/\sigma_1 \geq c/c_0 > \eta_0$.
- $\sigma_1 \leq \sigma_2 < \sigma'_1$. Then $2c\sigma_1 \leq c(\sigma_1+\sigma_2) \leq 2c_0\sigma'_1$, also implying that $\sigma'_1/\sigma_1 \geq c/c_0 > \eta_0$.

我们得出结论， $\sigma'_1/\sigma_1 \geq c/c_0$。根据引理 10 以及式 (8) 中 η_0 的定义， P_1, P'_1 的全变差距离可下界为

$$TV(P_1, P'_1) > F_d(\frac{2\eta_0^2\log\eta_0}{\eta_0^2-1}) - F_d(\frac{2\log\eta_0}{\eta_0^2-1}) = 1-\rho_1\ , \tag{157}$$

其中， ρ_1 与（152）中的相同。

另一方面，注意到 P_1 无法与 P' 中除 P'_1以外的任何分量匹配。这是因为，如果存在一个未被移除且与 P_1匹配的 P'_2，则根据上述相同的论证，可以证明$\sigma'_2 < \sigma_1 \leq \sigma'_1$。这与 σ_1的最小方差选择相矛盾。

因此，对于任意满足 $j \geq 2$且未被移除的 P'_j，令 P_j 为 P 中与 P'_j匹配的另一个分量（不是 P_1），则

$$\begin{aligned} ||\mu'_j - \mu_1|| &\geq ||\mu'_j - \mu'_1|| - c_0\max\{\sigma_1, \sigma'_1\} && (158)\\ &\geq c(\sigma'_j + \sigma'_1) - c_0\sigma'_1 && (159)\\ &\geq c\sigma'_j + (c-c_0)\sigma'_1 && (160)\\ &> c\sigma'_j && (161)\end{aligned}$$

同样，根据 σ_1 $||\mu'_j - \mu_1|| > c\sigma'_j \geq c\sigma_1$的选择可得。因此， $||\mu'_j - \mu_1|| > c\max\{\sigma'_j, \sigma_1\}$。

对于任何已被移除的 P'_j，仍将与 P'_j匹配的唯一分量记为 P_j，则（161）仍然成立。现在我们证明 $||\mu'_j - \mu_1|| > c\max\{\sigma_1, \sigma'_j\}$。若 $\sigma_1 \leq \sigma'_j$，则显然成立。当$\sigma'_j < \sigma_1$时，若 $||\mu'_j - \mu_1|| \leq c\sigma_1$不成立，则

$$c(\sigma_j + \sigma_1) \leq ||\mu_j - \mu_1|| \leq ||\mu_j - \mu'_j|| + ||\mu'_j - \mu_1|| \leq c_0\max\{\sigma_j, \sigma'_j\} + c\sigma_1\ , \tag{162}$$

这意味着$\max\{\sigma_j, \sigma_j'\}/\sigma_j \geq c/c_0 > \eta_0$。这是不可能的，因此对于任何已被移除的 P_j'，$||\mu_1 - \mu_j'|| \geq c\max\{\sigma_1, \sigma_j'\}$。

综合这两种情况，我们再次可以找到集合 A_i，使得

- For $i \neq 1$, whether P_i' is removed or not, $P_i'(A_i) - P_1(A_i) > 1 - \rho_2$.
- For $i = 1$, $P_1'(A_1) - P_1(A_1) > 1 - \rho_1$.

(154) 和 (155) 中的相同论证会导出矛盾，进一步表明所有分量都应已被移除，也就是说，当 $c > c_0\eta_0$ 时，匹配是一一对应的，因此 $K = K'$。关于中心距离和标准差比值的上界此前已经得到。□

A.4 均值和标准差的迭代细化

定理 16 给出了关于分离度、最小比例和最大比例的条件，在这些条件下，可以建立 P 与 P' 的分量之间的一一对应。本节将证明，一旦完成这一步 $K = K'$，我们就能进一步迭代改进这些界。因此，我们可以完成定理 4 的证明。我们假设定理 16 中的所有假设均成立。

定理 *4* 的证明：均值和标准差的上界。根据 $TV(P, P') \leq 2\epsilon$，取$Q = (P + P')/2$，我们有 $TV(P, Q) \leq \epsilon$，$TV(P', Q) \leq \epsilon$，因此可以应用定理 16，从而确认存在一一对应，且对应的中心彼此接近。为简化记号，我们重新排列 μ_j' 的分量，使得 P_i，P_i' 相互对应。具体而言，对于任意一对P_i，P_i'，根据假设 $||\mu_i - \mu_i'|| \leq c_0 \max\{\sigma_i, \sigma_i'\}$ 和 $\max\{\sigma_i/\sigma_i', \sigma_i'/\sigma_i\} \leq \eta_0$。

现在考虑一对 P_i，P_i'。不失一般性，假设 $\pi_i \leq \pi_i'$，否则可以交换 P和P'的角色。对于所有 $j \neq i$，

$$||\mu_j - \mu_i'|| \geq c(\sigma_j' + \sigma_i') - c_0 \max\{\sigma_j, \sigma_j'\} \geq c\sigma_i' + \left(\frac{c}{\eta_0} - c_0\right) \max\{\sigma_j', \sigma_j\} \geq c\sigma_i' + \left(\frac{c}{\eta_0} - c_0\right) \sigma_j \; ; \tag{163}$$

$$||\mu_j - \mu_i'|| \geq c(\sigma_j + \sigma_i) - c_0 \max\{\sigma_i, \sigma_i'\} \geq c\sigma_j + \left(\frac{c}{\eta_0} - c_0\right) \max\{\sigma_i, \sigma_i'\} \geq c\sigma_j + \left(\frac{c}{\eta_0} - c_0\right) \sigma_i' \; . \tag{164}$$

因此，将这两个方程相加可得

$$||\mu_j - \mu_i'|| \geq \frac{1}{2}\left(c + \frac{c}{\eta_0} - c_0\right)(\sigma_j + \sigma_i') \; . \tag{165}$$

注意到 $c > \eta_0 c_0$，因此 $||\mu_j - \mu_i'|| \geq \frac{1}{2} c_0 \eta_0 (\sigma_j + \sigma_i')$。根据引理 9，存在一个集合 A_j，使得 $P_i(A_j) - P_j'(A_j) > 1 - \rho$，其中

$$\rho = 2\Phi\left(-\frac{1}{2}(c + \frac{c}{\eta_0} - c_0)\right) < 2\Phi(-\eta_0 c_0/2) < 2\Phi(-c_0/2) = \pi_{\min} - 2\epsilon$$

。然后根据引理 13，P_i 和 P_i' 之间的全变差距离的上界为

$$TV(P_i, P_i') \leq UB(c_0, \eta_0) := \frac{2\epsilon}{\pi_{\min}} + \frac{2(1 - \pi_{\min})}{\pi_{\min}} \Phi\left(-\frac{1}{2}(c + \frac{c}{\eta_0} - c_0)\right) < 1 - \pi_{\min} + 2\epsilon \; . \tag{166}$$

由此进一步可得，根据引理 8，$||\mu_i - \mu_i'|| \leq c_1 \max\{\sigma_i, \sigma_i'\}$，$c_1 = 2\Phi^{-1}(1 - \frac{1 - UB(c_0, \eta_0)}{2})$。注意，$UB(c_0, \eta_0) < 1 - \pi_{\min} + 2\epsilon$，因此 $c_1 < c_0$。另外，根据引理 10，$\max\{\sigma_i/\sigma_i', \sigma_i'/\sigma_i\} \leq \eta_1$，其中 η_1 满足方程

$$F_d(\frac{2\eta_1^2 \log \eta_1}{\eta_1^2 - 1}) - F_d(\frac{2 \log \eta_1}{\eta_1^2 - 1}) = UB(c_0, \eta_0) \tag{167}$$

由于 $1-(K-1)\pi_{\min} \geq \pi_{\min}$,

$$F_d(\frac{2\eta_0^2 \log \eta_0}{\eta_0^2-1}) - F_d(\frac{2\log \eta_0}{\eta_0^2-1}) \tag{168}$$

$$= 1 - \frac{\pi_{\min}-2\epsilon}{1-(K-1)\pi_{\min}} + \frac{2(K-1)\pi_{\min}}{1-(K-1)\pi_{\min}} \Phi(-\frac{1}{2}\eta_0 c_0) \tag{169}$$

$$> 1 - \frac{\pi_{\min}-2\epsilon}{1-(K-1)\pi_{\min}} + \frac{2(K-1)\pi_{\min}}{1-(K-1)\pi_{\min}} \Phi\left(-\frac{1}{2}(c+\frac{c}{\eta_0}-c_0)\right) \tag{170}$$

$$= 1 - \frac{\pi_{\min}-2\epsilon-[1-(1-(K-1)\pi_{\min})]\,\rho}{1-(K-1)\pi_{\min}} \qquad (\text{recall } \rho = 2\Phi(-(c+c/\eta_0-c_0)/2)) \tag{171}$$

$$= 1 - \frac{\pi_{\min}-2\epsilon-\rho}{1-(K-1)\pi_{\min}} - \rho \tag{172}$$

$$\geq 1 - \frac{\pi_{\min}-2\epsilon-\rho}{\pi_{\min}} - \rho \tag{173}$$

$$= 1 - \frac{\pi_{\min}-2\epsilon}{\pi_{\min}} + \frac{1-\pi_{\min}}{\pi_{\min}}\rho \tag{174}$$

$$= F_d(\frac{2\eta_1^2 \log \eta_1}{\eta_1^2-1}) - F_d(\frac{2\log \eta_1}{\eta_1^2-1})\,, \tag{175}$$

成立的是 $\eta_1 < \eta_0$。总之，目前从 c_0，η_0 开始，我们能够给出更精细的上界 c_1，η_1。注意，利用与 (166) 相同的论证，可以用 c_1，η_1 将上界进一步细化为

$$TV(P_i, P_i') \leq UB(c_1,\eta_1) := \frac{2\epsilon}{\pi_{\min}} + \frac{2(1-\pi_{\min})}{\pi_{\min}} \Phi\left(-\frac{1}{2}(c+\frac{c}{\eta_1}-c_1)\right)\,. \tag{176}$$

类似地，$||\mu_i - \mu_i'|| \leq c_2 \max\{\sigma_i, \sigma_i'\}$，$c_2 = 2\Phi^{-1}(1 - \frac{1-UB(c_1,\eta_1)}{2}) \cdot \max\{\sigma_i/\sigma_i', \sigma_i'/\sigma_i\} \leq \eta_2$，其中 η_2 满足

$$F_d(\frac{2\eta_2^2 \log \eta_2}{\eta_2^2-1}) - F_d(\frac{2\log \eta_2}{\eta_2^2-1}) = UB(c_1,\eta_1)\,. \tag{177}$$

由于 $UB(c_1, \eta_1) < UB(c_0, \eta_0)$，我们有 $c_2 < c_1$，$\eta_2 < \eta_1$。这个过程可以重复。令

$$UB(c_t,\eta_t) := \frac{2\epsilon}{\pi_{\min}} + \frac{2(1-\pi_{\min})}{\pi_{\min}} \Phi\left(-\frac{1}{2}(c+\frac{c}{\eta_t}-c_t)\right) \tag{178}$$

于是可以找到

$$c_{t+1} = 2\Phi^{-1}(1 - \frac{1-UB(c_t,\eta_t)}{2}) = 2\Phi^{-1}(\frac{1}{2} + \frac{1}{2}UB(c_t,\eta_t)) \tag{179}$$

且 η_{t+1} 满足

$$F_d(\frac{2\eta_{t+1}^2 \log \eta_{t+1}}{\eta_{t+1}^2-1}) - F_d(\frac{2\log \eta_{t+1}}{\eta_{t+1}^2-1}) = UB(c_t,\eta_t) \tag{180}$$

使得

$$||\mu_i - \mu_i'|| \leq c_{t+1} \max\{\sigma_i, \sigma_i'\}, \quad \max\{\sigma_i'/\sigma_i, \sigma_i/\sigma_i'\} \leq \eta_{t+1}\,. \tag{181}$$

(180) 的左侧是关于 η_{t+1} 的严格递增函数，且存在连续反函数。因此我们得出，存在一个连续映射 $g\colon [0,1] \to [0,c_0] \times [1,\eta_0]$，使得

$$(c_{t+1}, \eta_{t+1}) = g(c_t, \eta_t) \tag{182}$$

通过归纳可得，对于 $t \geq 1$，$UB(c_t,\eta_t) < UB(c_{t-1},\eta_{t-1})$，因此，$c_{t+1} < c_t, \eta_{t+1} < \eta_t$。于是，$\{c_t\}, \{\eta_t\}$ 是两个递减且有下界的数列。因此，它们的极限存在。将极限记为 (c^*, η^*)，则它必定是 g 的不动点，即它们满足 (9)。此外，$||\mu_i - \mu_i'|| \leq c^* \max\{\sigma_i, \sigma_i'\}$,$\max\{\sigma_i'/\sigma_i, \sigma_i/\sigma_i'\} \leq \eta^*$。

比例的上界 $|\pi_i - \pi_i'|$ 已通过分量 c^*，η^* 将在下一节中证明；届时我们将建立其他技术，以完成定理 4 的证明。 □

在展示比例差异的证明之前，我们先给出推论 6 的证明。 .

推论 *6* 的证明。假设 P' 是一个球形高斯分布，使得 $TV(P,\ P') \leq 2\epsilon$。考虑两个新的球面高斯混合分布

$$\tilde{P} = \frac{1}{2}P + \frac{1}{2}Q \tag{183}$$

$$\tilde{P}' = \frac{1}{2}P' + \frac{1}{2}Q \tag{184}$$

其中 Q 是一个距离足够远的高斯分量。具体而言，对于 P 或 P' 中的任意分量 $N_d(\mu_i,\ \sigma_i^2 I_d)$，选择 $Q = N_d(\mu,\ \sigma^2 I_d)$，使得 $||\mu - \mu_i|| \geq c\max\{\sigma, \sigma_i\}$。因此 $P \in \mathcal{M}(K+1,\ \pi_{\min}/2,\ c)$ 和 $P' \in \mathcal{M}(2,\ \pi_{\min}/2,\ c)$。然而，由于 $TV(P,\ P') \geq 2\epsilon/2 = \epsilon$，~ ~，且 $P,\ P'$ 的最大比例有上界 $1/2$。令 η_0' 定义为下式的解

$$1 - \pi_{\min} + \epsilon + 2\Phi(-\frac{1}{2}\eta_0' c_0') = F_d(\frac{2\eta_0'^2 \log \eta_0'}{\eta_0'^2 - 1}) - F_d(\frac{2\log \eta_0'}{\eta_0'^2 - 1}) \tag{185}$$

其中 $c_0' = 2\Phi^{-1}(1 - \frac{\pi_{\min} - \epsilon}{4})$，因此如果 $c > c_0'\eta_0'$，根据定理 16，$K+1 = 2$，这意味着 $K = 1$，而这构成了矛盾。 □

A.5 比例差

最后，鉴于 $\eta^*,\ c^*$ 已在前文建立，我们将通过给比例差设定上界来完成定理 4 的证明。对于相应的一对 $\pi_i,\ \pi_i'$，其中 $||\mu_i - \mu_i'|| \leq c^* \max\{\sigma_i, \sigma_i'\}$且 $\max\{\sigma_i', \sigma_i\}/\min\{\sigma_i', \sigma_i\} \leq \eta^*$，我们通过构造集合 A 来为 $|\pi_i - \pi_i'|$ 设定上界，从而可以应用引理 14。

首先，我们介绍一个几何引理。

引理 17. 设 x_1, x_2, y_1, y_2 是 $\mathbb{R}^d$中的不同点，其中

$$\tilde{x} = \alpha x_1 + (1-\alpha)x_2\ , \tag{186}$$

$$\tilde{y} = \beta y_1 + (1-\beta)y_2\ . \tag{187}$$

Then

$$||\tilde{x} - \tilde{y}||^2 = \alpha\beta||x_1 - y_1||^2 + (1-\alpha)(1-\beta)||x_2 - y_2||^2 + (1-\alpha)\beta||x_2 - y_1||^2 + \alpha(1-\beta)||x_1 - y_2||^2 - \alpha(1-\alpha)||x_1 - x_2||^2 - \beta(1-\beta)||y_1 - y_2||^2$$

该引理的证明推迟到 B 节。接着考虑两对分量 P_i, P_i'和 P_j, P_j'，可以找到一个远离所有四个中心的超平面。

引理 18. 设 $P_i = N_d(\mu_i, \sigma_i^2 I_d)\ i = 1、2$ 和 $P_i' = N_d(\mu_i', \sigma_i'^2 I_d)$ 是两对球面高斯分布，且存在三个常数 $c^* \geq 0, \eta^* \geq 1, c \geq c^*\eta^*$，使得

$$\max\{\frac{\sigma_1}{\sigma_1'}, \frac{\sigma_1'}{\sigma_1}\} \leq \eta^*, \qquad \max\{\frac{\sigma_2}{\sigma_2'}, \frac{\sigma_2'}{\sigma_2}\} \leq \eta^* \tag{188}$$

$$||\mu_1 - \mu_1'|| \leq \frac{c^*}{2}(\sigma_1 + \sigma_1'), \quad ||\mu_2 - \mu_1|| \geq c(\sigma_1 + \sigma_2) \tag{189}$$

$$||\mu_2 - \mu_2'|| \leq \frac{c^*}{2}(\sigma_2 + \sigma_2'), \quad ||\mu_1' - \mu_2'|| \geq c(\sigma_1' + \sigma_2') \tag{190}$$

存在一个超平面 H，使得距离

$$dist(\mu_i, H) \geq C(c^*, \eta^*, c)\sigma_i, \quad dist(\mu_i', H) \geq C(c^*, \eta^*, c)\sigma_i', \quad i = 1, 2 \tag{191}$$

其中

$$C(c^*, \eta^*, c) = \sqrt{\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2}} - \frac{c^*}{2} \tag{192}$$

引理 *18* 的证明。考虑两个点

$$\tilde{\mu}_i = \frac{\sigma_i}{\sigma_i + \sigma_i'}\mu_i + \frac{\sigma_i'}{\sigma_i + \sigma_i'}\mu_i', \quad i = 1, 2\ , \tag{193}$$

注意，对于 $i = 1$，2，

$$||\tilde{\mu}_i - \mu_i|| = \frac{\sigma_i'}{\sigma_i + \sigma_i'}||\mu_i' - \mu_i|| \leq \frac{c^* \max\{\sigma_i, \sigma_i'\}}{2}, \quad ||\tilde{\mu}_i - \mu_i'|| = \frac{\sigma_i}{\sigma_i + \sigma_i'}||\mu_i' - \mu_i|| \leq \frac{c^* \max\{\sigma_i, \sigma_i'\}}{2} \tag{194}$$

也就是说，μ_i，μ_i'两者都位于一个以 $\tilde{\mu}_i$为中心、半径为 $c^* \max\{\sigma_i, \sigma_i'\}/2$ 的球内。我们考虑的超平面为

$$H : \left\{ x \in \mathbb{R}^d : \langle x - \tilde{\mu_1}, \tilde{\mu_2} - \tilde{\mu_1} \rangle = \frac{\sigma_1 + \sigma_1'}{\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2'} ||\tilde{\mu_2} - \tilde{\mu_1}|| \right\} \tag{195}$$

根据引理 17，距离

$$\begin{aligned} ||\tilde{\mu}_1 - \tilde{\mu}_2||^2 = & \underbrace{\frac{\sigma_1\sigma_2}{(\sigma_1 + \sigma_1')(\sigma_2 + \sigma_2')}||\mu_1 - \mu_2||^2 + \frac{\sigma_1'\sigma_2'}{(\sigma_1 + \sigma_1')(\sigma_2 + \sigma_2')}||\mu_1' - \mu_2'||^2}_{LB_1} \\ & + \underbrace{\frac{\sigma_1\sigma_2'}{(\sigma_1 + \sigma_1')(\sigma_2 + \sigma_2')}||\mu_1 - \mu_2'||^2 + \frac{\sigma_1'\sigma_2}{(\sigma_1 + \sigma_1')(\sigma_2 + \sigma_2')}||\mu_1' - \mu_2||^2}_{LB_2} \\ & - \underbrace{\frac{\sigma_1\sigma_1'}{(\sigma_1 + \sigma_1')^2}||\mu_1 - \mu_1'||^2 - \frac{\sigma_2\sigma_2'}{(\sigma_2 + \sigma_2')^2}||\mu_2 - \mu_2'||^2}_{UB} \end{aligned}$$

我们将分别对 LB_1，LB_2和 UB给出下界，从而为 $||\tilde{\mu}_1 - \tilde{\mu}_2||$ 建立一个下界。

首先，根据分离条件

$$LB_1 \geq c^2 \frac{\sigma_1\sigma_2(\sigma_1 + \sigma_2)^2 + \sigma_1'\sigma_2'(\sigma_1' + \sigma_2')^2}{(\sigma_1 + \sigma_1')(\sigma_2 + \sigma_2')(\sigma_1 + \sigma_2 + \sigma_1' + \sigma_2')^2}(\sigma_1 + \sigma_2 + \sigma_1' + \sigma_2')^2 \tag{196}$$

注意到 $\max\{\sigma_i/\sigma_i', \sigma_i'/\sigma_i\} \leq \eta^*$，$i = 1$，2，我们记

$$\alpha = \frac{\sigma_1}{\sigma_1 + \sigma_1'}\ , \quad \beta = \frac{\sigma_2'}{\sigma_2 + \sigma_2'}\ , \quad \gamma = \frac{\sigma_1 + \sigma_2}{\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2'}\ , \tag{197}$$

因此，由于 $1/(1 + \eta^*) \leq \alpha$，$\beta$，$\gamma \leq \eta^*/(1 + \eta^*)$，可得

$$\frac{LB_1}{c^2(\sigma_1 + \sigma_2 + \sigma_1' + \sigma_2')^2} = \alpha\beta\gamma^2 + (1 - \alpha)(1 - \beta)(1 - \gamma)^2 \tag{198}$$

$$\geq \left(\frac{1}{1 + \eta^*}\right)^2 [\gamma^2 + (1 - \gamma)^2] \tag{199}$$

$$\geq \frac{1}{2}\left(\frac{1}{1 + \eta^*}\right)^2 . \tag{200}$$

即

$$LB_1 \geq \frac{c^2}{2}\left(\frac{1}{1 + \eta^*}\right)^2 (\sigma_1 + \sigma_2 + \sigma_1' + \sigma_2')^2 \tag{201}$$

对于第二项，首先注意到

$$||\mu_1 - \mu_2'|| \geq c(\sigma_1 + \sigma_2) - \frac{c^*}{2}(\sigma_2 + \sigma_2'), \quad ||\mu_1 - \mu_2'|| \geq c(\sigma_1' + \sigma_2') - \frac{c^*}{2}(\sigma_1 + \sigma_1') \tag{202}$$

将这两个方程相加，得到

$$||\mu_1 - \mu_2'|| \geq \frac{1}{2}(c - \frac{c^*}{2})(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2')\ . \tag{203}$$

类似地， $||\mu_1' - \mu_2|| \geq (c - \frac{c^*}{2})(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2')/2$。因此

$$\frac{LB_2}{(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2')^2} \geq \frac{1}{4}(c - \frac{c^*}{2})^2[\alpha(1-\beta) + \beta(1-\alpha)] \,. \tag{204}$$

令 $g(\alpha, \beta) = \alpha(1-\beta) + \beta(1-\alpha) = \beta(1-2\alpha) + \alpha$。

- 若 $\alpha = 1/2$，则 $g(\alpha, \beta) = 1/2$。
- 若 $1/(1+\eta^*) \leq \alpha < 1/2$，则 $g(\alpha, \beta) \geq g(\alpha, 1/(1+\eta^*))$，因为它是关于$\beta$的仿射函数，且线性项为正。当 $\beta = 1/(1+\eta^*) < 1/2$时，采用相同的论证可知，$g(\alpha, \beta)$ 是关于 α 的仿射函数，且线性项为正。因此，最小值为

$$g(\frac{1}{1+\eta^*}, \frac{1}{1+\eta^*}) = \frac{2\eta^*}{(1+\eta^*)^2} \,. \tag{205}$$

- 若 $1/2 < \alpha \leq \eta^*/(1+\eta^*)$，相同的论证表明最小值为

$$g(\frac{\eta^*}{1+\eta^*}, \frac{\eta^*}{1+\eta^*}) = \frac{2\eta^*}{(1+\eta^*)^2} \,. \tag{206}$$

综上，我们得到

$$LB_2 \geq \frac{1}{4}\frac{2\eta^*}{(1+\eta^*)^2}(c - \frac{c^*}{2})^2(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2')^2 \tag{207}$$

最后，我们为 UB 建立一个上界，

$$UB = \frac{\sigma_1\sigma_1'}{(\sigma_1 + \sigma_1')^2}||\mu_1 - \mu_1'||^2 - \frac{\sigma_2\sigma_2'}{(\sigma_2 + \sigma_2')^2}||\mu_2 - \mu_2'||^2 \tag{208}$$

$$\leq \frac{(c^*)^2}{16}(\sigma_1 + \sigma_1')^2 + \frac{(c^*)^2}{16}(\sigma_2 + \sigma_2')^2 \tag{209}$$

$$\leq \frac{(c^*)^2}{16}(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2')^2 \tag{210}$$

因此

$$||\tilde{\mu}_1 - \tilde{\mu}_2|| \geq \sqrt{\frac{c^2}{2}\left(\frac{1}{1+\eta^*}\right)^2 + \frac{1}{4}\frac{2\eta^*}{(1+\eta^*)^2}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2}{16}}(\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2') \tag{211}$$

μ_1， μ_1' 到 H 的距离下界为

$$dist(\mu_1, H) \geq ||\tilde{\mu}_1 - \tilde{\mu}_2||\frac{\sigma_1 + \sigma_1'}{\sigma_1 + \sigma_1' + \sigma_2 + \sigma_2'} - \frac{c^*}{2}\max\{\sigma_1, \sigma_1'\} \tag{212}$$

$$\geq \sqrt{\frac{c^2}{2}\left(\frac{1}{1+\eta^*}\right)^2 + \frac{1}{4}\frac{2\eta^*}{(1+\eta^*)^2}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2}{16}}(\sigma_1 + \sigma_1') - \frac{c^*}{2}\max\{\sigma_1, \sigma_1'\} \tag{213}$$

$$\geq \sqrt{\frac{c^2}{2}\left(\frac{1}{1+\eta^*}\right)^2 + \frac{1}{4}\frac{2\eta^*}{(1+\eta^*)^2}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2}{16}}(\frac{1}{\eta^*} + 1)\max\{\sigma_1, \sigma_1'\} - \frac{c^*}{2}\max\{\sigma_1, \sigma_1'\} \tag{214}$$

$$\geq \left(\sqrt{\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2}} - \frac{c^*}{2}\right)\max\{\sigma_1, \sigma_1'\} \tag{215}$$

$$:= C(c^*, \eta^*, c)\max\{\sigma_1, \sigma_1'\} \tag{216}$$

注意，由于 $c > c^*\eta^*$ 和 $\eta^* \geq 1$,

$$\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2} - \frac{(c^*)^2}{4} \tag{217}$$

$$> \frac{(c^*)^2}{2} + \frac{(c^*)^2(\eta^* - \frac{1}{2})^2}{2\eta^*} - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2} - \frac{(c^*)^2}{4} \tag{218}$$

$$= (c^*)^2 \left[\frac{4(\eta^*)^2 + 8\eta^*(\eta^* - \frac{1}{2})^2 - (1+\eta^*)^2}{16(\eta^*)^2}\right] \tag{219}$$

$$= \left(\frac{c^*}{4\eta^*}\right)^2 \left[8(\eta^*)^3 - 5(\eta^*)^2 - 1\right] > 0 \tag{220}$$

式 (216) 中的系数 $C(c^*, \eta^*, c)$ 为正，因此

$$dist(\mu_1, H) \geq \left(\sqrt{\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2}} - \frac{c^*}{2}\right)\sigma_1 \tag{221}$$

$$dist(\mu_1', H) \geq \left(\sqrt{\frac{c^2}{2(\eta^*)^2} + \frac{1}{2\eta^*}(c - \frac{c^*}{2})^2 - \frac{(c^*)^2(1+\eta^*)^2}{16(\eta^*)^2}} - \frac{c^*}{2}\right)\sigma_1' \tag{222}$$

这也表明，μ_1，μ_1' 与 $\tilde{\mu}_1$ 位于 H 的同一侧。类似地，可以证明 μ_2，μ_2' 与 $\tilde{\mu}_2$ 位于 H 的同一侧，这与 μ_1，μ_1' 不同；另外两个距离 $dist(\mu_2, H)$ $dist(\mu_2', H)$ 也可以用类似的方法给出下界。□

定理 4 的证明：比例差异。现在我们可以完成定理 4 的证明。设 P_i，P_i' 是两个对应对，使得

$$\max\{\frac{\sigma_i}{\sigma_i'}, \frac{\sigma_i'}{\sigma_i}\} \leq \eta^*, \quad ||\mu_i - \mu_i'|| \leq \frac{c^*}{2}(\sigma_i + \sigma_i') \,. \tag{223}$$

这里中心之间的距离略有不同。事实上，对于这些对应对，有 t

$$TV(P_i, P_i') \leq \frac{2\epsilon}{\pi_{\min}} + \frac{2(1-\pi_{\min})}{\pi_{\min}}\Phi(-\frac{1}{2}\left(c + \frac{c}{\eta^*} - c^*\right)) = 1 - 2\Phi(-\frac{c^*}{2}) \,. \tag{224}$$

根据引理 9，还有 $||\mu_i - \mu_i'|| \leq c^*(\sigma_i + \sigma_i')/2$。考虑任意另一个对应对 P_j，P_j'，则 P_i，P_i'，P_j，P_j' 满足引理 18 的条件，因此我们可以选取一个超平面 H，使得 P_i，P_i' 和 P_j，P_j' 位于 H 的不同侧，且它们之间的距离有下界。令 A_j 为由 H 确定且包含 P_j，P_j' 的半空间，则

$$\max\{P_i(A_j^{\complement}), P_i'(A_j^{\complement}), P_j(A_j), P_j'(A_j)\} \leq \Phi(-C(c, c^*, \eta^*)) \tag{225}$$

其中 $C(c, c^*, \eta^*)$ 由式 (192) 确定。现在取 $A = \cap_{j\neq i} A_j$，则

$$P_i(A) = P_i((\cup_{j\neq i} A_j^{\complement})^{\complement}) \geq 1 - \sum_{j\neq i} P_i(A_j^{\complement}) \geq 1 - (K-1)\Phi(-C(c, c^*, \eta^*)) \,, \tag{226}$$

$$P_i'(A) = P_i'((\cup_{j\neq i} A_j^{\complement})^{\complement}) \geq 1 - \sum_{j\neq i} P_i'(A_j^{\complement}) \geq 1 - (K-1)\Phi(-C(c, c^*, \eta^*)) \,; \tag{227}$$

$$P_j(A) \leq P_j(A_j) \leq \Phi(-C(c, c^*, \eta^*)) \tag{228}$$

$$P_j'(A) \leq P_j'(A_j) \leq \Phi(-C(c, c^*, \eta^*)) \tag{229}$$

结果由引理 14 给出。□

B 辅助引理

B.1 第 2 节中的引理

引理19. 设 F_d 为 $Gamma(\frac{d}{2}, \frac{d}{2})$ 分布的累积分布函数，则对于任意 $\eta \geq 1$,

$$R.H.S.\ of\ (8) = F_d(\frac{2\eta^2 \log \eta}{\eta^2 - 1}) - F_d(\frac{2\log \eta}{\eta^2 - 1}) \geq 1 - \left(\frac{2\eta}{\eta^2 + 1}\right)^{\frac{d}{2}} \,, \tag{230}$$

引理*19*的证明. 考虑两个分布 $P_1 = N_d(\mathbf{0}, I_d)$ 和 $P_2 = N_d(\mathbf{0}, \eta^2 I_d)$，其中 $\eta \geq 1$，以及集合

$$A = \{x \in \mathbb{R}^d : p_1(x) \geq p_2(x)\} = \left\{x \in \mathbb{R}^d : ||x|| \leq \frac{\eta}{\eta^2 - 1}\sqrt{2d(\eta^2 - 1)\log \eta}\right\} . \tag{231}$$

其中，p_i，$i = 1$，2 是 P_i 的密度。因此，$P_1(A) - P_2(A)$ 是式 (8) 右侧的表达式。此外，它是 P_1 与 P_2 之间的全变差距离。众所周知，全变差距离的下界为海林格距离 [14], ，因此有

$$\text{R.H.S. of (8)} \geq \frac{1}{2}\int(\sqrt{p_1} - \sqrt{p_2})^2 = 1 - \left(\frac{2\eta}{\eta^2 + 1}\right)^{\frac{d}{2}} . \tag{232}$$

□

B.2 附录A中的引理

引理 20. 若存在集合 A，使得 $P_1(A) - P_2(A) \geq 1 - \rho \geq 0$，则对于任意权重 $0 \leq \pi_1$、$\pi_2 \leq 1$:

$$\pi_1 - \max\{\pi_1, \pi_2\}\rho \geq \pi_1 P_1(A) - \pi_2 P_2(A) \leq \pi_1 \tag{233}$$

引理 *20* 的证明。注意到

$$\pi_1 P_1(A) - \pi_2 P_2(A) \geq \pi_1(P_2(A) + 1 - \rho) - \pi_2 P_2(A) = \pi_1 + (\pi_1 - \pi_2)P_2(A) - \pi_1\rho .$$

将其视为关于 $P_2(A)$ 的仿射函数。由于 $0 \leq P_2(A) \leq \rho$，我们有

$$\pi_1 P_1(A) - \pi_2 P_2(A) \geq \min\{\pi_1 + (\pi_1 - \pi_2)\cdot 0 - \pi_1\rho, \pi_1 + (\pi_1 - \pi_2)\rho - \pi_1\rho\} = \pi_1 - \max\{\pi_1, \pi_2\}\rho . \tag{234}$$

□

引理 21. 当 $x > 1$, $f(x) = x$ 对数 $x/(x-1)$ 关于 x 单调递增，且 $g(x) =$ 对数 $x/(x-1)$ 关于 x 单调递减时。

引理 *21* 的证明。求导

$$f'(x) = \frac{(x-1)(1 + \log x) - x\log x}{(x-1)^2} = \frac{x - 1 - \log x}{(x-1)^2} \geq 0 \tag{235}$$

$$g'(x) = \frac{(x-1)/x - \log x}{(x-1)^2} = \frac{1}{(x-1)^2}[1 - \frac{1}{x} + \log\frac{1}{x}] \leq 0 \tag{236}$$

□

引理 *17* 的证明。 该引理可通过直接计算证明。首先，注意到

$$||y_1 - y_2||^2 = ||y_1 - x_1||^2 + ||y_2 - x_1||^2 - 2\langle y_1 - x_1, y_2 - x_1\rangle \tag{237}$$

Then

$$\begin{aligned}
&||x_1 - \tilde{y}||^2 \\
&= ||\beta(x_1 - y_1) + (1-\beta)(x_1 - y_2)||^2 && (238)\\
&= \beta^2||x_1 - y_1||^2 + (1-\beta)^2||x_1 - y_2||^2 + 2\beta(1-\beta)\langle x_1 - y_1, x_1 - y_2\rangle && (239)\\
&= \beta^2||x_1 - y_1||^2 + (1-\beta)^2||x_1 - y_2||^2 + 2\beta(1-\beta)\frac{||x_1 - y_1||^2 + ||x_1 - y_2||^2 - ||y_1 - y_2||^2}{2} && (240)\\
&= \beta||x_1 - y_1||^2 + (1-\beta)||x_1 - y_2||^2 - \beta(1-\beta)||y_1 - y_2||^2 . && (241)
\end{aligned}$$

同理，我们有

$$||x_2 - \tilde{y}||^2 = \beta||x_2 - y_1||^2 + (1-\beta)||x_2 - y_1||^2 - \beta(1-\beta)||y_1 - y_2||^2 . \tag{242}$$

综合以上所有结果，可得

$$
\begin{aligned}
&||\tilde{x}-\tilde{y}||^2 \\
&=\alpha||\tilde{y}-x_1||^2+(1-\alpha)||\tilde{y}-x_2||^2-\alpha(1-\alpha)||x_1-x_2||^2 \qquad (243)\\
&=\alpha\beta||x_1-y_1||^2+(1-\alpha)(1-\beta)||x_2-y_2||^2+(1-\alpha)\beta||x_2-y_1||^2+\alpha(1-\beta)||x_1-y_2||^2 \\
&\quad-\alpha(1-\alpha)||x_1-x_2||^2-\beta(1-\beta)||y_1-y_2||^2 \qquad (244)
\end{aligned}
$$

□