arXiv:2305.00966v1 [cs.DS] 1 May 2023

相对弗罗贝尼乌斯范数下的列表可解码协方差估计的谱算法

Ilias Diakonikolas[†]威斯康星大学麦迪逊分校 ilias@cs.wisc.edu

Daniel M. Kane[‡]加州大学圣迭戈分校 dakane@cs.ucsd.edu

Jasper C.H. Lee[§]威斯康星大学麦迪逊分校 jasper.lee@wisc.edu

Ankit Pensia[¶]威斯康星大学麦迪逊分校 ankitp@cs.wisc.edu

Thanasis Pittas[‖]威斯康星大学麦迪逊分校 pittas@wisc.edu

2023年5月2日

摘要

我们研究列表可解码高斯分布协方差估计问题。给定一个由 T 个 $\mathbb{R}^d$ 中的 n 点构成的多重集，其中未知比例 $\alpha < 1/2$ 的 T 中的点是来自未知高斯分布 $\mathcal{N}(\mu, \Sigma)$ 的独立同分布样本，目标是输出一个包含 $O(1/\alpha)$ 个假设的列表，其中至少一个与 Σ 的相对弗罗贝尼乌斯范数误差较小。我们的主要结果是一个用于解决该问题的算法，其样本和时间复杂度均为 poly$(d, 1/\alpha)$，并保证相对弗罗贝尼乌斯范数误差为 poly$(1/\alpha)$。重要的是，我们的算法完全依赖谱方法。作为推论，我们得到了一种高效的谱算法，用于高斯混合模型（GMM）的鲁棒部分聚类——这是近期 [BDJ+22] 鲁棒学习任意 GMM 的工作的关键组成部分。结合 [BDJ+22] 的其他组成部分，我们的新方法得到了首个无需平方和方法的鲁棒 GMM 学习算法。在技术层面上，我们针对列表可解码协方差估计提出了一种新颖的多重过滤方法，该方法或可用于其他场景。

*作者按字母顺序排列。

[†]由美国国家科学基金会中型项目奖 CCF-2107079、美国国家科学基金会 CCF-1652862 奖（CAREER）以及美国国防高级研究计划局“少标签学习”（LwLL）项目资助。

[‡]由美国国家科学基金会中型项目奖 CCF-2107547 和美国国家科学基金会奖 CCF-1553288（CAREER）资助。[§]由 Croucher 博士后研究奖学金、美国国家科学基金会奖 DMS-2023239、美国国家科学基金会中型项目奖 CCF-2107079 以及美国国家科学基金会 AiTF 奖 CCF-2006206 资助。

[¶]由美国国家科学基金会奖项 CCF-1652862（CAREER）以及美国国家科学基金会资助项目 CCF-1841190 和 CCF-2011255 支持。

[‖]由美国国家科学基金会中型项目奖 CCF-2107079 资助。

目录

1 引言 **1**

1.1 我们的结果.2 1.2 技术概述.4 1.3 相关工作.5 1.4 组织结构.6

2 预备知识 6

2.1 基本符号.6 2.2 其他事实.6 2.3 确定性条件.7 2.4 关于稳定性的事实.7

3 算法单次递归调用的分析 8

3.1 证书引理：有界四阶矩蕴含接近性.10 3.2 过滤：移除极端离群点.14 3.3 划分器：识别多个簇并递归处理.16

4 定理 1.2 的证明 174.1 引理 4.6 的证明.24

5 鲁棒学习高斯混合模型的应用 25

A 线性代数结果 **31**

A.1 相对弗罗贝尼乌斯范数到弗罗贝尼乌斯范数：引理 5.1 的证明.31 A.2 引理 3.7 的证明.34

B 第2节中的省略的证明 **35**

B.1 Fact 2.1 的证明.35 B.2 样本复杂度：引理 2.6 的证明.36 B.3 关于稳定性的事实：推论 2.7 和引理 2.8 的证明 38

C 归一化引理 3.2 的证明 **39**

1 引言

稳健统计学研究给定样本时底层分布的高效（参数）可学习性，其中一部分样本可能已被任意篡改。虽然这些问题的统计理论早已确立 [Hub64，ABH+72，HR09]，但直到最近，这类估计问题的算法理论才开始得到理解 [DKK+16，LRV16，DK23]。

鲁棒估计中的一个经典问题是多元鲁棒均值估计——也就是在存在少量离群点的情况下，估计未知分布的均值。该领域最初的一些研究成果之一表明：给定来自高斯分布 $\mathcal{N}(\mu, I)$ 的样本，其中含有 ϵ 比例的离群点（对于某个 $\epsilon < 1/2$），未知均值 μ 可以被高效估计，误差为 $O(\epsilon\sqrt{\log 1/\epsilon})$，这里使用 ℓ_2-范数 [DKK+16]。值得注意的是，此处使用 ℓ_2-范数相当自然，因为两个单位协方差高斯分布之间的全变差距离大致正比于其均值之间的 ℓ_2-距离；因此，该估计器能够以 $O(\epsilon\sqrt{\log 1/\epsilon})$ 的误差，在全变差距离意义下学习底层分布。这个界无法得到实质性改进，因为从信息论角度而言，不可能将全变差距离误差降至 $o(\epsilon)$。

虽然上述算法结果在 ϵ 较小时（小于 $1/2$）适用，但如果超过一半的样本受到污染，仅返回一个假设就无法实现学习——污染样本可能只是模拟其他高斯分布，因此任何算法都无法判断哪个高斯分布才是原始分布。我们可以采用列表可解码均值估计范式 [CSV17]，来规避这一问题。在这种范式下，算法可以输出一个较小的假设列表，并保证其中至少有一个与目标相对接近。特别是对于高斯均值估计，如果有 α比例的样本是干净的（即未受污染），对于某个 $\alpha < 1/2$，存在多项式时间算法能够返回一个包含 $O(1/\alpha)$个假设的列表，并以高概率保证其中至少有一个假设与 $\tilde{O}(\sqrt{1/\alpha})$ 的均值相距不超过 ℓ_2距离 [CSV17]。请注意，尽管这一ℓ_2范数距离界并不能保证真实分布与学习得到的假设之间的全变差距离较小，但它能保证二者的距离小于 1，从而确保存在非平凡的重叠。

稳健统计学中另一个重要且可以说更为复杂的问题（也是本文的研究重点）是稳健地估计多元高斯分布的协方差。已有研究表明，在 [DKK+16] 中，对于来自 ϵ-受污染样本的 $\mathcal{N}(0, \Sigma)$（对于 $\epsilon < 1/2$），存在一种多项式时间算法，可以将 Σ 估计到误差 $O(\epsilon\log 1/\epsilon)$，其中误差采用相对 Frobenius 范数衡量；也就是说，输出一个假设协方差 $\widetilde{\Sigma}$，满足 $\|\widetilde{\Sigma}^{-1/2}\Sigma\widetilde{\Sigma}^{-1/2} - I\|_{\mathrm{F}}$。这同样是相关的度量，因为当 $\|\widetilde{\Sigma}^{-1/2}\Sigma\widetilde{\Sigma}^{-1/2} - I\|_{\mathrm{F}}$ 小于某个较小的普适常数时，两个此类高斯分布之间的全变差距离与其协方差之间的相对 Frobenius 范数成正比。

一个自然且具有多种应用的目标，是将上述算法结果推广到列表可解码设定。也就是说，我们希望有一种多项式时间算法，给定来自 $\mathcal{N}(0, \Sigma)$ 的受污染样本（其中干净样本所占比例为 α，对某个 $\alpha < 1/2$而言），返回一个包含$O(1/\alpha)$ 个假设的列表，并保证其中至少一个假设与真实分布在全变差距离下存在一定的非平凡重叠。我们首先指出，列表可解码协方差估计的样本复杂度为 $\mathrm{poly}(d/\alpha)$，尽管目前这是通过指数时间算法实现的。目前已知的唯一一种列表可解码协方差估计算法（具有全变差误差保证）由 [IK22]提出。该算法本质上依赖平方和方法，其（样本和计算）复杂度为 $d^{\mathrm{poly}(1/\alpha)}$。耐人寻味的是，有力证据表明这一复杂度界无法改进。具体而言， [DKS17、DKP+21] 证明，任何用于此任务的统计查询（SQ）算法都需要复杂度 $d^{\mathrm{poly}(1/\alpha)}$。结合 [BBH+21] 的归约，这意味着低次多项式检验也存在类似的下界。这些下界表明，该问题存在内在的信息-计算差距。

上述 述的下界结果 [DKS17，DKP+21] 证明了，要区分标准的低阶）困难 -

阶）上很难区分标准多元高斯分布，$\mathcal{N}(0, I)$，与分布 P，后者在低阶矩方面表现得类似于 $\mathcal{N}(0, I)$，但 P包含比例为 α 的高斯分布样本 $\mathcal{N}(0, \Sigma)$，其中 Σ 沿某个隐藏方向 v非常薄。由于分布$\mathcal{N}(0, \Sigma)$ 沿 v 非常薄，也就是说方差极小，因此显然选择的 $\mathcal{N}(0, I)$ 与 $\mathcal{N}(0, \Sigma)$ 基本没有重叠——这使得这类强列表解码保证（全变差距离意义下的接近性）在计算上很可能难以实现。

有趣的是，两个零均值高斯分布可能通过两种方式彼此区分： [DHKK20, BDH+20]: (1) 在某个方向上，一个分布可能比另一个窄得多；或者 (2) 它们可能有许多方差不同的正交方向，这些差异累积起来会变得更加显著。尽管 [DKS17, DKP+21] 的下界似乎表明，无法在完全多项式时间内（即 poly(d/α)）检测出前一种类型的偏差，但这并不意味着无法高效检测后一种偏差。具体而言，我们可以希望在 poly(d/α) 时间内找到一个良好的假设 $\widetilde{\Sigma}$，使得 $\|\widetilde{\Sigma}^{-1/2}\Sigma\widetilde{\Sigma}^{-1/2} - I\|_{\mathrm{F}}$ 不太大。虽然这并不排除 $\widetilde{\ }$在少数相互独立方向上比 Σ 窄得多的可能性 Σ，但它确实排除了两者之间第二种类型的差异。本文的主要目标是提供一种此类的基础列表解码算法（仅依赖谱方法）。给定来自高斯分布 $\mathcal{N}(0, \Sigma)$ 的受污染样本，我们给出一个 poly(d/α) 时间算法，该算法返回一个较小的假设列表 $\widetilde{\Sigma}$，其中至少有一个假设满足 $\|\widetilde{\Sigma}^{-1/2}\Sigma\widetilde{\Sigma}^{-1/2} - I\|_{\mathrm{F}} <$ poly$(1/\alpha)$。

除了在完全多项式时间内提供我们所能期望的最佳定性保证之外，上述这类“弱”列表解码算法还对一个研究充分的问题——鲁棒学习高斯混合模型（GMM）——具有重要意义。 [BDJ+22] 提出了一种多项式时间算法，可鲁棒地学习任意高斯混合（含常数个组分）。他们方法的两个主要组成部分之一，是一个子程序，能够将点部分聚类到各个组分中，并满足恰好这种弱接近性保证。 [BDJ+22] 针对样本来自轻度受污染的高斯混合（其中离群点所占比例较小且为常数）的情形，通过关键性地使用平方和（SoS）松弛，构造了这样一个子程序。作为我们技术的推论，我们为这一部分聚类任务得到了一个初等谱算法（并且如前所述，我们的结果适用于更一般的列表解码情形）。由此，我们得到了首个无平方和算法来鲁棒学习高斯混合模型，从而解决了文献中的一个开放问题 [Vem22、Kot21]。

1.1 我们的结果

我们的主要结果是一种多项式时间算法，可在相对 Frobenius 范数下对高斯分布的协方差进行列表解码，即使在对抗性污染中，超过一半的样本都可能是离群点，也能实现这一目标。 定义 1.1 精确定义了污染模型，而 定理 1.2 陈述了我们主要算法的保证（算法 1）。

定义 1.1 (列表解码的污染模型)。 设参数为 $\epsilon, \alpha \in (0, 1/2)$，分布族为 $\mathcal{D}$。统计学家指定样本数 m。随后，从某个未知的 $D \in \mathcal{D}$中抽取一组 $n \geq \alpha m$ 个独立同分布点。我们称这些 n 样本为内点。检查过这些 n 内点后，一个（恶意且计算能力不受限制的）对手可以将任意数量 $\ell \leq \epsilon n$ 的内点替换为任意点，并进一步向数据集添加 $m - n$ 个任意点，然后将全部 m 个点组成的集合交给统计学家。参数 α 也预先为统计学家所知，但对手选择的数量 n 未知。我们将这组来自 D 的 m 个点称为 (α, ϵ)- 污染样本集。

在我们的语境中，列表解码的含义如下：我们的算法将返回一个大小为多项式的矩阵列表 H_i，其中至少有一个 H_i 是实际

协方差 Σ 的"近似平方根"，其与维度无关的误差有界 $\|H_i^{-1/2}\Sigma H_i^{-1/2} - I\|_{\mathrm{F}}$。如前所述，我们保证的界为 poly$(1/\alpha)$，通常大于 1，因此无法得到非平凡的全变差界，从而绕开了相关的统计查询下界。

我们的主要结果的正式表述如下：

定理 *1.2*（相对弗罗贝尼乌斯范数下的列表可解码协方差估计）。设环境维度为 $d \in \mathbb{N}$，C' 是一个足够大的绝对常数；参数 $\alpha \in \in (0, 1/2)$、$\epsilon \in \in (0, \epsilon_0)$，其中 ϵ_0 是一个足够小的正数；并且失败概率 $\delta \in \in (0, 1/2)$ 均为算法已知。设 $D = \mathcal{N}(\mu, \Sigma)$ 是一个均值为 $\mu \in \mathbb{R}^d$、协方差为满秩的分布$\Sigma \in \mathbb{R}^{d\times d}$。存在一个运行时间为 $\tilde{O}(m^2 d^2)$ 的算法（算法 *1*），对于输入 α, δ，以及来自 D（定义 *1.1*）的 m 个点组成的一个 (α, ϵ)-受污染集合，只要 $m > C'^{d^2} \log^5(d/\alpha\delta)\, \alpha^6$，该算法就能以至少 $1-\delta$ 的概率返回一个至多包含 $O(1/\alpha)$ 个集合的列表 T_i；这些集合是样本的互不相交子集，每个集合的大小至少为 $0.5\alpha m$，并且输出列表中存在一个 T_i，使得：

- 回顾污染模型（定义 *1.1*）中的记号，其中 n 是原始内点集的大小，S 而 ℓ 是 S 中被对手替换的点数——n 并且 ℓ 对算法而言是未知的，但 $n \geq \alpha m$ 和 $\ell \leq \epsilon n$除外。返回列表中的集合 T_i 满足$|T_i \cap S| \geq (1 - 0.01\alpha)(n - \ell)$。

- *Denote* $H_i := \mathbf{E}_{X\sim T_i}[XX^\top]$. *The matrix* H_i *satisfies* $\|H_i^{-1/2}\Sigma H_i^{-1/2} - I\|_{\mathrm{F}} \lesssim \frac{1}{\alpha^4}\log(\frac{1}{\alpha})$.

算法 1 是一种迭代谱算法，不涉及基于平方和层级的大型凸规划。我们在算法 1 中给出算法，并简要概述其满足定理 1.2 的证明，详见第 3 节。定理 1.2 的形式证明见定理 1.2 的第 4 节（其中我们也允许 Σ 为秩亏矩阵）。

作为主要结果的推论，该算法（尽管样本复杂度略高）同样能以相对弗罗贝尼乌斯范数对高斯混合模型各组分的协方差进行抗离群值列表解码。定义 1.3 和定理 1.4 给出了相应的污染模型以及关于算法 1的理论保证。

定义 1.3（高斯混合模型样本的污染模型）。设 $\epsilon \in (0, 1/2)$。考虑一个高斯混合模型 $\sum_p \alpha_p \mathcal{N}(\mu_p, \Sigma_p)$，其中参数 α_p、μ_p和 Σ_p未知，且满足 $\alpha_p \geq \alpha$，其中 α是某个已知参数。统计学家指定样本数 m，并从该高斯混合模型中抽取m 个独立同分布样本，称为内点。随后，具有恶意且计算能力不受限的对手检查这 m 个内点，并可将其中任意一个包含 $\epsilon\alpha m$ 个内点的子集替换为任意离群点，然后将修改后的数据集交给统计学家。我们称之为 ϵ-污染样本集。

定理 *1.4*（含离群点鲁棒的*GMM*聚类与协方差估计）。设环境维度为 $d \in \mathbb{N}$，且 C'是一个充分大的绝对常数。设 $\alpha \in (0, 1/2)$、$\epsilon \in (0, \epsilon_0)$ 以及失败概率 $\delta \in (0, 1/2)$ 是算法已知的，其中 ϵ_0是某个充分小的正常数。存在一个运行时间为 $\tilde{O}(m^2 d^2)$ 的算法，使得对于输入 α、δ以及来自某个未知的 k组分高斯混合模型 $\sum_{p=1}^k \alpha_p \mathcal{N}(\mu_p, \Sigma_p)$ 的 $m > C'^{d^2} \log^5(d/\alpha\delta)$ 个α^6 ϵ-污染样本（该模型如定义 *1.3* 所述），其中所有 Σ_p均为满秩，且所有 α_p 均满足 $\alpha_p \geq \alpha$，而 $k \leq \frac{1}{\alpha}$ 对算法未知时，该算法以至少 $1-\delta$ 的概率（对污染样本及算法的随机性取概率）返回一个由至多 k 个两两不相交的样本子集 $\{T_i\}$ 组成的列表，使得：

- 对于 pth 高斯分量，记集合 S_p 为从组分 p中抽取且属于内点集 S 的样本。令 n_p 为 S_p的大小，令 ℓ_p 为点的数量

在 S_p 中，对手替换了 n_p，而算法对 ℓ_p 一无所知，除了$\mathbf{E}[n_p] = \alpha_p m \geq \alpha m$ 对每个 p 和 $\sum_p \ell_p \leq \epsilon\alpha m$ 均成立。于是，对于混合分布中的每个高斯分量 p，返回的列表中都存在一个集合 T_{i_p}，使得 $|T_{i_p} \cap S_p| \geq (1 - 0.01\alpha)(n_p - \ell_p)$。

- 对于每个成分 p，返回的列表中都存在一个样本集合 T_{i_p}，使得定义$H_{i_p} = \mathbf{E}_{X \sim T_{i_p}}[XX^\top]$ 后，我们有 H_{i_p} 满足 $\|H_{i_p}^{-1/2} \Sigma_p H_{i_p}^{-1/2} - I\|_{\mathrm{F}} \lesssim (1/\alpha^4)\log(1/\alpha)$。
- 设 Σ 为高斯混合模型的（总体层面的）协方差矩阵。对于任意两个成分 $p \neq p'$，若 $\|\Sigma^{-1/2}(\Sigma_p - \Sigma_{p'})\Sigma^{-1/2}\|_{\mathrm{F}} > C(1/\alpha)^5 \log(1/\alpha)$，且常数 C足够大，则可保证前一个要点中的集合 T_{i_p} 和 $T_{i_{p'}}$ 不同。

该定理指出，算法 1 不仅能够对高斯分量的协方差进行列表解码，还能依据协方差在相对 Frobenius 范数下的分离程度对样本进行聚类。近期关于鲁棒学习高斯混合模型的研究 [BDJ+22] 也包含用于聚类的算法组件。他们的方法基于平方和层级（因此需要求解大型凸规划），而算法 1 则是纯粹的谱算法。

我们还要强调的是，对于一个 算法 1 返回的列表大小，在 k 组分高斯混合模型的情形下，至多为 k ——而不是诸如 $O(1/\alpha)$ 或者关于 $1/\alpha$ 多项式级／指数级大的较弱结果。这之所以可行，是因为 算法 1 会仔细追踪样本，并确保每个组分中被移除或被错误归入另一个组分的样本比例不超过 0.01α。我们在 定理 1.4 中证明了这一点，见 第 5 节。

1.2 技术概述

从整体上看，我们的方法将[DKK+16] 中的鲁棒协方差估计技术与 [DKS18]用于列表解码的多重滤波器相结合。对于集合 S，我们用 $SS^\top$ 表示集合 $\{xx^\top : x \in S\}$。 [DKK+16] 指出，要估计真实协方差 Σ，只需找到一个较大的点子集 S，使其与原始优质点集合有较大的重叠；如此一来，若 $\Sigma' = \mathrm{Cov}(S)$，则 $\mathrm{Cov}((\Sigma')^{-1/2} S S^\top (\Sigma')^{-1/2})$ 不会有较大的特征值。这将意味着$\|(\Sigma')^{-1/2}(\Sigma - \Sigma')(\Sigma')^{-1/2}\|_{\mathrm{F}}$ 不会太大（引理 3.5）。

与 [DKK+16] 一样，我们寻找此类集合 S 的基本方法是反复执行以下过程：首先令 S 为所有样本构成的集合，然后反复检查 $\mathrm{Cov}((\Sigma')^{-1/2} S S^\top (\Sigma')^{-1/2})$ 是否存在较大的特征值。若不存在，我们就完成了。若存在一个较大的特征值，对应于矩阵 A （归一化为弗罗贝尼乌斯范数为 1），我们便考察 $f(x) = \langle A, (\Sigma')^{-1/2} x x^\top (\Sigma')^{-1/2} \rangle$（其中 x 属于 S），并尝试利用它们找出离群点。遗憾的是，由于内点可能只占样本中的少数，我们从这个公式得到的值最终可能形成若干规模相当大的簇，其中任何一个都可能合理地包含真实样本；因此，我们无法有把握地断定任何特定点是离群点。我们采用 [DKS18] 的多重过滤方法来解决这个问题——我们要么 (i) 迭代地移除离群点，要么 (ii) 将数据划分为多个簇，并在每个簇上递归处理。特别地，我们注意到，如果在 A 方向上存在大方差，那么以下两种情况必有其一：(i) $f(x)$ 的绝大多数值落在同一个簇中，另有一些极端离群点。在这种情况下，我们可以确信这些极端离群点确实是错误，并将其移除（第 3.2 节）。(ii) $f(x)$ 的值至少形成两个彼此相距甚远的簇。在这种情况下，我们必须采取更复杂的方法。我们不只是移除明显的离群点，而是将 S 替换为两个子集 S_1 和 S_2，并保证其中至少一个 S_i 几乎包含所有内点。简单来说，可以找到一个位于两个簇之间的值 y，使得只有极少数样本的 $f(x)$ 接近 y，然后令 S_1 为满足 $f(x) < y$ 的点集，令 S_2 为满足 $f(x) > y$ 的点集（第 3.3 节）。无论哪种情况，我们都将

清理样本集合后，就可以在返回的 S 子集上分别递归处理。对返回的所有较小子集递归地反复应用此技术，可以确保始终至少有一个子集包含大多数内点；最终，当协方差不再过大时，我们将返回对 Σ 的适当近似。

我们想强调我们的方法与[DKS18]有显著差异的核心之处。要实现上述算法概述，需要预先对 $f(x)$ 在内点上的方差有较好的界。由于 $f(\cdot)$ 是一个二次多项式， f 在内点上的方差本身取决于协方差 Σ，而这正是我们要估计的量。这种循环依赖的难题在 [DKS18]中并不存在：他们的目标是估计单位协方差高斯分布的未知均值，因此使用线性多项式f（而非二次多项式）就足够了。重要的是，线性多项式的协方差不依赖于（未知的）均值（它只取决于协方差，而在他们的设定中协方差是已知的）。为克服这一难题，我们注意到，如果 S 包含了大多数内点，那么 S 的协方差就不可能比真实协方差 Σ小太多。这使我们能够求出 Σ的一个上界，进而得到 $f(x)$ 在优质样本上的方差上界（引理 3.3）。

1.3 相关工作

有关算法鲁棒统计的概述，请参阅 [DK23] 。列表可解码学习模型最早由 [BBV08] 引入，并首次应用于高维估计领域，见 [CSV17]。特别是在均值估计方面，已有大量研究在对干净数据作出各种假设的条件下，提出了保证逐步提升的算法 [DKS18, KS17, RY20a, CMY20, DKK20, DKK+21, DKK+22b, DKK+22a, ZS22]。在子空间估计和线性回归这两个特殊情形下，研究者已开发出列表可解码协方差估计算法 [KKK19, BK21, RY20b, RY20a, DJKS22]。另一方面， [DKS17, DKP+21] 分别给出了学习高斯混合模型和列表可解码线性回归（因而也包括列表可解码协方差估计）的统计查询下界。

[IK22] 首次提出了针对一般列表可解码协方差估计的算法，利用强大的平方和层级，在全变差距离上实现了非平凡的界。该算法输出一个大小为 $\exp(\mathrm{poly}(1/\alpha))$ 的矩阵列表，其中包含一个 H_i ，它在两种度量下都接近真实的 Σ ：(i) 相对弗罗贝尼乌斯范数： $\|H_i^{-1/2}\Sigma H_i^{-1/2} - I\|_{\mathrm{F}} = \mathrm{poly}(1/\alpha)$；(ii) 乘法谱近似，即 $\mathrm{poly}(\alpha)\Sigma \preceq H_i \preceq \mathrm{poly}(1/\alpha)\Sigma$。该算法使用$d^{\mathrm{poly}(1/\alpha)}$ 样本；对于能够实现乘法谱近似的高效（统计查询）算法而言，这似乎是必要的 [DKS17, DKP+21]。相比之下，定理 1.2 仅使用 $\mathrm{poly}(d/\alpha)$ 个样本，返回一个包含 $O(1/\alpha)$ 个矩阵的列表，但只在相对弗罗贝尼乌斯范数下进行近似，即 $\|H_i^{-1/2}\Sigma H_i^{-1/2} - I\|_{\mathrm{F}} = \mathrm{poly}(1/\alpha)$。

高斯混合模型高斯混合模型在统计学和理论计算机科学中有着悠久的研究历史。在理论计算机科学领域， [Das99], 率先系统研究了高斯混合模型学习的高效算法，随后一系列工作对其进行了改进；例如，参见 [VW04、AK05、DS07、KK10]。这些算法对高斯混合模型的参数施加了额外假设，例如组分之间需要分离。后续工作[MV10、BS10] 提出了不对组分作任何假设的算法。上述所有算法都容易受到离群点的影响。近期的一系列工作 [DHKK20、BDH+20、Kan21、LM21、BDJ+22] 发展出了用于稳健地学习高斯混合模型的高效算法。

1.4 组织结构

在第2节中，我们列出必要的符号和背景知识，并给出算法将使用的一组确定性条件。第 3 节介绍主要的算法要素及其分析。随后，第 4 节将这些思路整合为最终的递归算法，并通过归纳法证明定理 1.2 。中间的技术性引理和证明放在附录中。

2 预备知识

2.1 基本记号

我们用 $\mathbb{Z}_+$ 表示正整数。对于 $n \in \mathbb{Z}_+$，我们记 $[n] \stackrel{\text{def}}{=} \{1, \ldots, n\}$，并用 $\mathcal{S}^{d-1}$ 表示 d维单位球面。对于向量 v，我们用 $\|v\|_2$ 表示其 ℓ_2范数。我们用 I_d 表示 $d \times d$ 阶单位矩阵。上下文明确时，我们省略下标。对于矩阵A，我们分别用 $\|A\|_{\mathrm{F}}$ 和 $\|A\|_{\mathrm{op}}$ 表示其弗罗贝尼乌斯范数和谱范数（或算子范数）。我们称对称的 $d \times d$ 矩阵 A 为半正定矩阵（PSD），并记作 $A \succeq 0$；若对所有向量 $x \in \mathbb{R}^d$，均有 $x^\top A x \geq 0$。我们记 $\lambda_{\max}(A) := \max_{x \in \mathcal{S}^{d-1}} x^\top A x$。当 $B - A$ 为半正定矩阵时，我们写作 $A \preceq B$。对于矩阵 $A \in \mathbb{R}^{d \times d}$，$\mathrm{tr}(A)$ 表示矩阵 A 的迹。我们用$\langle v, u \rangle$表示向量 u和 v之间的标准内积。对于矩阵 U、$V \in \mathbb{R}^{d \times d}$，我们将内积推广如下：$\langle U, V \rangle = \sum_{ij} U_{ij} V_{ij}$。对于矩阵 $A \in \mathbb{R}^{d \times d}$，我们用 $A^\flat$ 表示其在 $\mathbb{R}^{d^2}$中的展平向量；对于 $v \in \mathbb{R}^{d^2}$，我们用 $v^\sharp$ 表示唯一满足 $A^\flat = v$的矩阵 A。对于矩阵 A，我们用 $A^\dagger$ 表示其伪逆。对于半正定矩阵 A，我们也用记号 $A^{\dagger/2}$ 表示 $A^\dagger$ 的半正定平方根（其本身也为半正定矩阵）。我们用 $\otimes$ 表示克罗内克积。对于矩阵 A，我们用 $\ker(A)$表示 A 的零空间，并用 $\mathbb{P}_A$ 表示投影到 A的值域上的投影，即 $\mathbb{P}_A = AA^\dagger$ [Ber18, 命题 8.1.7 (xii)]。

我们使用记号 $X \sim D$ 来表示随机变量 X 服从分布 D。对于随机变量 X，我们使用 $\mathbf{E}[X]$ 表示其期望。我们使用 $\mathcal{N}(\mu, \Sigma)$ 表示均值为 μ、协方差矩阵为 Σ的高斯分布。对于集合 S，我们使用 $\mathcal{U}(S)$ 表示其上的均匀分布，并使用 $X \sim S$ 作为 $X \sim \mathcal{U}(S)$ 的简写。

我们用 $a \lesssim b$ 表示存在一个绝对常数 $C > 0$（它与 a 和 b 所依赖的变量或参数无关），使得 $a \leq Cb$。

2.2 其他事实

本节陈述一些将在分析中用到的标准事实。

事实 2.1。对于任意矩阵 $A \in \mathbb{R}^{d \times d}$，**方差**$_{X \sim \mathcal{N}(\mu, \Sigma)}[X^\top A X] \leq 4\|\Sigma^{1/2} A \Sigma^{1/2}\|_{\mathrm{F}}^2 + 8\|\Sigma^{1/2} A \mu\|_2^2$。

为完整起见，我们在附录 B.1中给出这一结论的证明。

事实 2.2（Azuma-Hoeffding不等式）。设 $\{X_n\}_{n \in \mathbb{N}}$ 是关于序列 $\{Y_n\}_{n \in \mathbb{N}}$的一个次鞅或超鞅。若对所有 $n = 1, 2, \ldots$ 均有 $|X_n - X_{n-1}| \leq c_n$ 几乎必然成立，则对任意 $\epsilon > 0$ 和 $n \in \mathbb{N}$

$$\Pr\left[X_n - X_t > \epsilon\right] \leq \exp\left(-\frac{2\epsilon^2}{\sum_{t=1}^{\infty} c_t}\right).$$

事实 2.3（例如，参见 [Ber18，命题 11.10。34]）。设 R 和 M 是两个方阵，则$\|RM\|_{\mathrm{F}} \leq \min\left((\|R\|_{\mathrm{F}}\|M\|_{\mathrm{op}}, \|R\|_{\mathrm{op}}\|M\|_{\mathrm{F}}\right)$。

事实 2.4 (弗罗贝尼乌斯范数与投影矩阵)。设 A 为一个对称矩阵。若 L_1 和 L_2是满足 $L_1 \preceq L_2$ 的两个半正定投影矩阵，则 $\|L_1 A L_1^\top\|_{\mathrm{F}} \leq \|L_2 A L_2^\top\|_{\mathrm{F}}$。

2.3 确定性条件

我们的算法将依赖未受污染的内点满足一组性质，这些性质类似于 [DKK+16、DK19]中的“稳定性条件”。直观而言，稳定性条件是针对样本集合的集中性质，但还额外要求样本的每个较大子集也满足这些性质。

定义 2.5 ((η,ϵ)-稳定集合)。设 D 是一个均值为 μ 、协方差为 Σ的分布。若对于任意满足 $|A'| \geq (1-\epsilon)|A|$的子集 $A' \subseteq A$，以下条件均成立，则称点集 $A \subset \mathbb{R}^d$ 相对于 D是(η,ϵ)-稳定的：

(L.1) 对每个 $v \in \mathbb{R}^d$, $\left|(1/|A'|)\sum_{x\in A'} v^\top (x-\mu)\right| \leq 0.1\sqrt{(v^\top \Sigma v)}$。

(L.2) For every symmetric $U \in \mathbb{R}^{d\times d}$, $\left|\left\langle \frac{1}{|A'|}\sum_{x\in A'}(x-\mu)(x-\mu)^\top - \Sigma, U\right\rangle\right| \leq 0.1\left\|\Sigma^{1/2}U\Sigma^{1/2}\right\|_{\mathrm{F}}$.

(L.3) For every even degree-2 polynomial p, $\mathop{\mathbf{Pr}}_{X\sim A'}\left[|p(X) - \mathop{\mathbf{E}}_{X\sim D}[p(X)]| > 10\sqrt{\mathop{\mathbf{Var}}_{Y\sim D}[p(Y)]}\ln\left(\frac{2}{\eta}\right)\right] \leq \eta$.

(L.4) 对于所有偶数次-2 多项式 p，我们有 $\mathbf{Var}_{X\sim A'}[p(X)] \leq 4\mathbf{Var}_{X\sim D}[p(X)]$.

(L.5) 二阶矩矩阵的 A' 的零空间包含于 Σ 的零空间，即 $\ker\left(\sum_{x\in A'} xx^\top\right) \subseteq \ker(\Sigma)$。

众所周知，高斯分布数据集以高概率具有“稳定”性 [DKK+16]；形式化地，我们有 引理 2.6，证明见 附录 B.2。正如我们在 附录 B.2 中的 附录 B.2所指出的，引理 2.6可以扩展到多种分布。

引理 2.6 （确定性条件以高概率成立）。对于充分小的正数 ϵ_0 和充分大的绝对常数 C，从 $\mathcal{N}(\mu,\Sigma)$ 中抽取的一个包含 $m > Cd^2\log^5(d/(\eta\delta))/\eta^2$ 个样本的集合，以概率 $1-\delta$ 是相对于 μ、Σ 对所有 $\epsilon \leq \epsilon_0$ 都成立的 (η,ϵ)-稳定集合。

2.4 关于稳定性的若干事实

以下推论将在后文中发挥作用。特别是，由于我们的算法将对数据集进行变换，我们需要推论 2.7 来论证这些集合的稳定性性质。此外，接下来的引理（引理 2.8）将用于在给出该定理的第一条要点后，推导出定理最终的弗罗贝尼乌斯范数保证。证明见附录 B.3。

推论 2.7 （变换集合的稳定性）。设 D 是定义在 $\mathbb{R}^d$ 上的一个分布，其均值为 μ，协方差为 Σ。设 A 是 $\mathbb{R}^d$ 中的一个点集，满足定义 2.5中的条件 *(L.1)* 和 关于(η,ϵ)-稳定性。则对于任意对称变换 D以及任意满足 $P \in \mathbb{R}^{d\times d}$的 $A' \subset A$，$|A'| \geq (1-\epsilon)|A|$，以下结论成立：

(i) $\left\|P\left(\frac{1}{|A'|}\sum_{x\in A'} x - \mu\right)\right\|_2 \leq 0.1\sqrt{\|P\Sigma P\|_{\mathrm{op}}}$.

(ii) $\left\|P\left(\frac{1}{|A'|}\sum_{x\in A'}(x-\mu)(x-\mu)^\top - \Sigma\right)P\right\|_{\mathrm{F}} \leq 0.1\|P\Sigma P\|_{\mathrm{op}}$.

下面的引理将用于证明我们的证书引理，从而将算法的最终保证从样本集的经验协方差转移到高斯分布的实际协方差。

引理 2.8. 设 A 是 $\mathbb{R}^d$ 中的一个点集，满足 条件 *(L.1)* 和 *(L.2)*，并且关于 (η,ϵ)-稳定（定义 *2.5*），且相对于 μ,Σ。设 $A' \subseteq A$ 满足 $|A'| \geq (1-\epsilon)|A|$。若 $P, L \in \mathbb{R}^{d\times d}$ 是具有如下性质的对称矩阵：$\|$**协方差**$_{X\sim A'}[PX] - L\|_{\mathrm{F}} \leq r$，则

$$\|P\Sigma P - L\|_{\mathrm{F}} = O(r + \|L\|_{\mathrm{op}}) .$$

3 算法单次递归调用的分析

我们的算法 算法 1会对样本进行过滤并递归地划分为多个集合，直到我们能够证明“当前数据集”的经验二阶矩矩阵适合纳入返回的协方差列表。作为参考，下面定义分析每次递归调用所需的符号和假设。不过，在进行正式分析之前，我们先非正式地概述算法的各个步骤及其背后的高层次思路。

假设 3.1.（算法单次递归调用的假设和符号）).

- $S = \{x_i\}_{i=1}^n$ 是一个 n 未受对手篡改的样本集合（在对手进行修改之前），假设该集合相对于内点分布 $(\eta, 2\epsilon_0)$-稳定，该内点分布的 D 均值和协方差为 μ,Σ（参见 定义 2.5）。假设 $\eta \leq 0.001$, $\epsilon_0 = 0.01$, 并且对所有对称的 **方差**$_{X\sim D}[X^\top AX] \leq C_1(\|\Sigma^{1/2}A\Sigma^{1/2}\|_F^2 + \|\Sigma^{1/2}A\mu\|_2^2)$ 矩阵 $d\times d$ A 以及常数 C_1。

- T 是当前算法递归调用的输入集（经过对抗性损坏后），满足 $|S\cap T| \geq (1-2\epsilon_0)|S|$ 和 $|T| \leq (1/\alpha)|S|$。这两者意味着$|T\cap S| \geq \alpha|T|(1-2\epsilon_0)$和 $|T| \geq |S|(1-2\epsilon_0)$。

- 我们记 $H = \mathbf{E}_{X\sim T}\left[XX^\top\right]$。

- 我们用 $\widetilde{S},\widetilde{T}$ 表示经过 $H^{\dagger/2}$归一化的 S 和 T 版本，即 $\widetilde{S} = \{H^{\dagger/2}x : x \in S\}$ 和 $\widetilde{T} = \{H^{\dagger/2}x : x \in T\}$。我们用记号 $\widetilde{x}$ 表示 $\widetilde{S}$ 和 $\widetilde{T}$中的元素，用 x 表示 S 和 T中的 $\widetilde{}$元素。类似地，我们用记号 X 表示支撑集$\widetilde{}$ $\widetilde{}$位于 S（或 T 或类似的集合“波浪号”版本）中的随机变量。

- 经过 $H^{\dagger/2}$ 变换后，内点分布 D 的均值和协方差分别记为 $\widetilde{\mu} := H^{\dagger/2}\mu$、$\widetilde{\Sigma} := H^{\dagger/2}\Sigma H^{\dagger/2}$。我们将 T中$\widehat{}$ $\widetilde{}$变换后的内点的经验均值和协方差记为$\widehat{\mu} := \mathbf{E}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[X]$ 和 $\Sigma := \mathbf{Cov}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[X]$。

下面的引理（引理 3.2）表明，假设输入数据集是稳定内点集的污染版本，那么用其经验二阶矩对数据进行归一化后，各种形式的均值和协方差的范数都有界。证明主要是常规计算，可参见 附录 C。

$\widetilde{}$ $\widehat{}$**引理 3.2**（归一化）。令 假设 *3.1* 成立，并回顾记号 $\widetilde{\mu}$，Σ，$\widehat{\mu}$，Σ，这些记号定义于 假设 *3.1*。我们有：

(i) $\|\hat{\Sigma}\|_{\mathrm{op}} \leq 2/\alpha.$

(ii) $\|\hat{\mu}\|_2 \leq \sqrt{2/\alpha}.$

(iii) $\|\widetilde{\Sigma}\|_{\mathrm{op}} \leq 3/\alpha$

(iv) $\|\widetilde{\mu}\|_2 \leq \sqrt{3/\alpha}$

~ ~
$\widetilde{X}\sim\widetilde{S}[X]\|_{\mathrm{op}} \le 4/\alpha$。 ~ ~ ~

6. 如果它们彼此接近，但方差仍然很大，我们就利用这一信息设计一个 得分函数 ，并进行过滤，以与得分成正比的概率随机移除一个样本。然后我们回到步骤 1。这样做之所以有效，是因为根据 引理 3.3 以及稳定性（定义 2.5），（未经过滤的）内点彼此之间的经验方差很小，这意味着较大的总经验方差主要是由离群点造成的。

理想情况下，样本的得分应与该样本和内点均值之间的平方距离成正比——此时，内点的总得分将等于内点方差。然而，由于我们不知道哪些点是内点，因此改用所有投影样本的中位数作为未知内点均值的代理。我们证明， αm个最小值与最大 $\tilde{y}$之间的距离界定了理想得分与代理得分之间的差异。

7. 否则， αm个最小值与 αm个最大值之间相距甚远。这种一维投影下的稳定性条件定义 2.5 (具体而言， 条件 (L.3)）意味着大多数内点在该一维投影下彼此接近。因此，较大的分位数范围必然意味着在该投影下存在一个阈值，可将样本分为两个集合，使得每个集合至少包含 αm 个点，并且大多数内点都留在同一个集合中。

步骤 6 中提到的得分函数为步骤 3 中检查的最大方差提供上界。为简单起见，在实际算法（算法 1）中，我们直接使用得分进行终止检查，而不检查协方差；不过从技术上说，使用哪个量并不重要。

备注 3.4（算法 1 的运行时间）。我们断言，算法 1 中的每个“循环”计算都需要 $\tilde{O}(md^2)$ 时间。运行“循环”的次数至多为 $O(m)$，因为每次循环要么以终止结束，要么从数据集中移除 1 个元素，要么拆分数据集，而这些情况各自最多发生 $O(m)$ 次。由此可得，算法的运行时间为 $\tilde{O}(m^2d^2)$。根据引理 2.6 以及定理 4.1 中参数 $\eta = \Theta(\alpha^3)$ 的选取，我们的算法的样本复杂度也可以明确计算为 $\tilde{O}(d^2/\alpha^6)$。

来看单次循环的运行时间：每次循环中最耗时的操作是计算 H_t、其伪逆，以及在第 7 行计算对称矩阵 A ，它是一个$d^2 \times d^2$ 矩阵的最大特征值对应的特征向量。计算 H_t 显然需要 $O(md^2)$ 的时间，得到一个 $d \times d$ 矩阵。其伪逆可在 $O(d^\omega)$ 的时间内计算，而这由 $O(md^2)$ 主导，因为 $m \gtrsim d$。最后，我们注意到，在分析中，无须真的计算第 7 行的最大特征值对应的特征向量来得到矩阵 A；只需计算一个矩阵 B，使其瑞利商 $(B^\flat)^\top \left(\mathrm{Cov}_{\widetilde{X}\sim\widetilde{T}_t}[\widetilde{X}^{\otimes 2}]\right) B^\flat/((B^\flat)^\top B^\flat)$ 至少为 $1/2$ 倍的 $(A^\flat)^\top \left(\textbf{协方差}_{\widetilde{X}\sim\widetilde{T}_t}[\widetilde{X}^{\otimes 2}]\right) A^\flat/((A^\flat)^\top A^\flat)$。我们可以通过 $O(\log d)$ 次幂迭代来实现这一点。由于

$$\mathbf{Cov}_{\widetilde{X}\sim\widetilde{T}_t}[\widetilde{X}^{\otimes 2}] = \frac{1}{|\widetilde{T}_t|}\sum_{\widetilde{X}\in\widetilde{T}_t} zz^\top - \left(\frac{1}{|\widetilde{T}_t|}\sum_{\widetilde{X}\in\widetilde{T}_t} z\right)^\top \left(\frac{1}{|\widetilde{T}_t|}\sum_{\widetilde{X}\in\widetilde{T}_t} z\right),$$

我们可以在 $O(|T_t|d^2) \le O(md^2)$ 时间内计算每次矩阵-向量乘积，因此幂迭代的运行时间为 $\tilde{O}(m^2d^2)$。

3.1 证书引理：有界四阶矩蕴含接近性

分析的第一个部分是我们的证书引理，它指出：如果当前数据集（归一化后）的（二阶矩矩阵展平后的）经验协方差有界，那么当前数据集的经验二阶矩矩阵 H 就是我们要估计的高斯分量的协方差的良好近似。

算法 1 相对弗罗贝尼乌斯范数下的列表可解码协方差估计

1：常数：$m, \alpha, \eta, R := C(1/\alpha^2)\log(1/(\epsilon_0\alpha))$，适用于 $C > 6000\sqrt{C_1}$, $C' > 720/\epsilon_0$（其中 C_1, ϵ_0 在假设 3.1 中定义）。2：函数 CovarianceListDecoding(T_0) 3：$t \leftarrow 0$。4：循环 5：计算 $H_t = \mathrm{E}_{X\sim T_t}[XX^\top]$。6：令 $\widetilde{T}_t = \{H_t^{\dagger/2}x : x \in T_t\}$ 为变换后的样本集合。7：令 A 为与 $\mathrm{Cov}_{\widetilde{X}\sim\widetilde{T}_t}[\widetilde{X}^{\otimes 2}]$ 的最大特征值对应的特征向量相对应的对称矩阵。8：对 A 归一化，使得 $\|A\|_{\mathrm{F}} = 1$。9：计算集合 $\widetilde{Y}_t = \{\widetilde{x}^\top A\widetilde{x} : \widetilde{x} \in \widetilde{T}_t\}$ 10：计算第 $\alpha m/9$ 小的元素 q_{left} 以及第 $\alpha m/9$ 大的元素 q_{right}, ~ 以及 Y_t 的中位数 $\widetilde{y}_{\mathrm{median}}$。11：定义函数 $f(\widetilde{x}) = (\widetilde{x}^\top A\widetilde{x} - \widetilde{y}_{\mathrm{median}})^2$。12 $\widetilde{X}\sim\widetilde{T}_t[f\ \widetilde{X}\ ] \le C'R^2/\alpha^3$ ▷ 35：13

如果 E ( )，则参见引理。：如果 $|T_t| \ge 0.5\alpha m$，则 14：返回 $\{T_t\}$。15：否则 16：返回空列表。17：否则，如果 $q_{\mathrm{right}} - q_{\mathrm{left}} \le R$，则▷ 随机化筛选（参见引理 3.8）18：令概率质量函数 $p(\widetilde{x}) := f(\widetilde{x})/\sum_{\widetilde{x}\in\widetilde{T}_t} f(\widetilde{x})$。19：根据 $p(H_t^{\dagger/2}x)$ 选取 $x_{\mathrm{removed}} \in T_t$。20：$T_{t+1} \leftarrow T_t \setminus \{x_{\mathrm{removed}}\}$。21：否则 ~ 22：$\tau \leftarrow$ FindDivider($Y_t, \alpha m/9$)。▷ 参见引理 3.11 23：$T' \leftarrow \{H_t^{1/2}\widetilde{x} : \widetilde{x} \in \widetilde{T}_t, \widetilde{x}^\top A\widetilde{x} \le \tau\}$, $T'' \leftarrow \{H_t^{1/2}\widetilde{x} : \widetilde{x} \in T_t, \widetilde{x}^\top A\widetilde{x} > \tau\}$。24：$L_1 \leftarrow$ CovarianceListDecoding(T')。25：$L_2 \leftarrow$ CovarianceListDecoding(T'') 26：返回 $L_1 \cup L_2$。27：$t \leftarrow t + 1$。

引理 3.5（停止并返回的情况）。令 假设 *3.1*。令 $w \in \mathbb{R}^{d^2}$ 为 $\mathbf{Cov}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^{\otimes 2}]$ 的主特征向量，且 $\|w\|_2 = 1$，并令 $A \in \mathbb{R}^{d\times d}$ 为 $w^\sharp$。注意，$\|w\|_2 = 1$ ~ ~ ~ ~意味着 $\|A\|_{\mathrm{F}} = 1$。对于每个子集 $T' \subseteq T$，满足 $|T'| \ge (\alpha/2)|T|$，我们有

$$\left\|\mathop{\mathbf{Cov}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] - \mathbb{P}_H\right\|_{\mathrm{F}}^2 \lesssim \frac{1}{\alpha}\mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] + \frac{1}{\alpha^2}\,.$$

特别地，我们有：

$$\left\|H^{\dagger/2}\Sigma H^{\dagger/2} - \mathbb{P}_H\right\|_{\mathrm{F}}^2 \lesssim \frac{1}{\alpha}\mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] + \frac{1}{\alpha^2}$$

and

$$\left\|H^{\dagger/2}\Sigma H^{\dagger/2} - \mathbb{P}_\Sigma\right\|_{\mathrm{F}}^2 \lesssim \frac{1}{\alpha}\mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] + \frac{1}{\alpha^2}\,.$$

我们的终止检查第 12 行使用得分 $f(\widetilde{x}) = (\widetilde{x}^\top A\widetilde{x} - \widetilde{y}_{\text{median}})^2$ 其中 $\widetilde{y}_{\text{median}}$ 是 $\{\widetilde{x}^\top A\widetilde{x}$ 的中位数: $\widetilde{x} \in \widetilde{T}\}$。由于**方差**$_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] \le$ **期望**$_{\widetilde{X}\sim\widetilde{T}}[f(\widetilde{X})]$ 我们的检查确保**方差**$_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] \le$ 多项式$(1/\alpha)$ 然后再返回。

事实上，我们证明了一个稍微更一般的引理，引理 3.6，它指出，在任意线性变换矩阵 P之后，如果 $X^{\otimes 2}$ 在受污染点集上的协方差矩阵的算子范数有界，那么 T 的经验二阶矩矩阵就是内点协方差在弗罗贝尼乌斯范数下的良好估计。引理 3.5则基本上可通过选择 $P = H^{\dagger/2}$得出。

引理 3.6 （有界的四阶矩和二阶矩蕴含协方差的接近性）。设 T 是 $\mathbb{R}^d$中的 n 个点。设 P 是任意对称矩阵。设 $w \in \mathbb{R}^{d^2}$ 是$\mathbf{Cov}_{X\sim T}[(PX)^{\otimes 2}]$ 的主特征向量，其中 $\|w\|_2 = 1$；并设 $A \in \mathbb{R}^{d\times d}$ 为 $w^\sharp$。注意， $\|w\|_2 = 1$蕴含 $\|A\|_{\mathrm{F}} = 1$。对于满足 $|T'| \ge (\alpha/2)|T|$的任意子集 $T' \subseteq T$，我们有

$$\left\|\mathop{\mathbf{Cov}}_{X\sim T'}[PX] - \mathop{\mathbf{E}}_{X\sim T}[PXX^\top P]\right\|_{\mathrm{F}}^2 \le \frac{4}{\alpha}\mathop{\mathbf{Var}}_{X\sim T}[X^\top PAPX] + \frac{8}{\alpha^2}\left\|\mathop{\mathbf{E}}_{X\sim T}[PXX^\top P]\right\|_{\mathrm{op}}^2 .$$

特别地，假设存在一个集合 S 满足 条件 *(L.1)* 和 *(L.2)*，即满足关于 η， ϵ-稳定性（定义 *2.5*），其中涉及 μ 和 Σ，对于某个 η 和 ϵ；并假设 $|S\cap T| \ge \min(\alpha|T|/2, (1-\epsilon)|S|)$。那么，将上述结果应用于 $T' = S\cap T$，可得：

$$\left\|P\Sigma P - \mathop{\mathbf{E}}_{X\sim T}[PXX^\top P]\right\|_{\mathrm{F}} \lesssim \frac{1}{\alpha}\mathop{\mathbf{Var}}_{X\sim T}[X^\top PAPX] + \frac{1}{\alpha^2}\left\|\mathop{\mathbf{E}}_{X\sim T}[PXX^\top P]\right\|_{\mathrm{op}}^2 .$$

在给出正式计算之前，我们先简要概述引理 3.6 的证明。对于引理的第一部分，将协方差项分解后，左边可由以下两项之和界定：$\mathbf{E}_{X\sim T'}[PXX^\top P] - \mathbf{E}_{X\sim T}[PXX^\top P]$的弗罗贝尼乌斯范数的平方，以及 $\mathbf{E}_{X\sim T'}[PX]\mathbf{E}_{X\sim T'}[X^\top P]$的弗罗贝尼乌斯范数的平方。前一项是 T' 中的点对右侧方差项的贡献，至多相差一个因子 $|T'|/|T|$，根据假设，该项至多为 $O(1/\alpha)$。后一项是一个秩一矩阵的弗罗贝尼乌斯范数的平方，它等于其算子范数，并且可进一步由 $\|\mathbf{E}_{X\sim T'}[PXX^\top P]\|_{\mathrm{op}}$界定。为了界定 $\|\mathbf{E}_{X\sim T'}[PXX^\top P]\|_{\mathrm{op}}$，我们采用同样的论证：根据 $|T'|/|T|$ 的构造，该项至多为 $\|\mathbf{E}_{X\sim T}[PXX^\top P]\|_{\mathrm{op}}$的 T' 倍。

引理的第二部分由 $T' = S\cap T$ 的选取得出，并应用条件（L.2）对内点集 S 的稳定性。

引理 3.6 的形式证明如下。

~引理 *3.6* 的证明。令 T定义集合 $\{PX\colon\ X \in T\}$。用这种记号， w是$\mathbf{Cov}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^{\otimes 2}]$ 的主特征向量，其中 $\|w\|_2 = 1$，并且 $A = w^\sharp$。那么，对于任意 $\widetilde{T}' \subseteq \widetilde{T}$满足 $|\widetilde{T}'| \ge \alpha|\widetilde{T}'|/2$的，我们有如下结果

$$\begin{aligned}
\mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] &= \sup_{v\in\mathbb{R}^{d^2},\|v\|_2^2=1} \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}\left[\left\langle \widetilde{X}^{\otimes 2} - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}^{\otimes 2}], v\right\rangle^2\right] \\
&= \sup_{v\in\mathbb{R}^{d^2},\|v\|_2^2=1} \frac{|\widetilde{T}'|}{|\widetilde{T}|}\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}\left[\left\langle \widetilde{X}^{\otimes 2} - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}^{\otimes 2}], v\right\rangle^2\right] \\
&\qquad + \frac{|\widetilde{T}\setminus\widetilde{T}'|}{|\widetilde{T}|}\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}\setminus\widetilde{T}'}\left[\left\langle \widetilde{X}^{\otimes 2} - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}^{\otimes 2}], v\right\rangle^2\right] \\
&\ge \frac{\alpha}{2}\sup_{v\in\mathbb{R}^{d^2},\|v\|_2^2=1} \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}\left[\left\langle \widetilde{X}^{\otimes 2} - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}^{\otimes 2}], v\right\rangle^2\right] \qquad (\text{since } |\widetilde{T}'| \ge (\alpha/2)|\widetilde{T}|)
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{\alpha}{2} \sup_{v\in\mathbb{R}^{d^2},\|v\|_2^2=1} \left\langle \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^{\otimes 2}] - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}^{\otimes 2}], v \right\rangle^2 \qquad (\mathbf{E}[Y^2] \geq (\mathbf{E}[Y])^2 \text{ for any } Y) \\
&= \frac{\alpha}{2} \sup_{J\in\mathbb{R}^{d\times d},\|J\|_{\mathrm{F}}^2=1} \left[\left\langle \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{Y}\sim\widetilde{T}}[\widetilde{Y}\widetilde{Y}^\top], J \right\rangle^2 \right] \\
&= \frac{\alpha}{2} \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 ,
\end{aligned}
$$

上面的倒数第二步成立，是因为将上一行中展平后的矩阵内积提升为矩阵内积；最后一行则利用了弗罗贝尼乌斯范数的变分刻画。因此，我们得到如下结果：

$$
\left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 \leq \frac{2}{\alpha} \mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}]. \tag{2}
$$

现在我们来证明引理的第一部分。利用二阶矩矩阵与协方差之间的关系，我们有：

$$
\begin{aligned}
\left\| \mathop{\mathbf{Cov}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 &= \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 \\
&\leq 2 \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 + 2 \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top] \right\|_{\mathrm{F}}^2 \\
&\leq \frac{4}{\alpha} \mathop{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}] + 2 \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top] \right\|_{\mathrm{op}}^2 ,
\end{aligned}
$$

其中，最后一步用到了 (2) 以及 $\mathbf{E}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}]\mathbf{E}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top]$ 是秩一矩阵这一事实，因此其 Frobenius 范数与算子范数相等。由此，要证明所需的不等式，只需证明 $\mathbf{E}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}]\mathbf{E}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top] \preceq (2/\alpha)\cdot\mathbf{E}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top]$。确实，我们有

$$
\begin{aligned}
\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}^\top] \preceq \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{T}'}[\widetilde{X}\widetilde{X}^\top] &= \frac{1}{|\widetilde{T}'|} \sum_{\widetilde{X}\in\widetilde{T}'} \widetilde{X}\widetilde{X}^\top = \frac{|\widetilde{T}|}{|\widetilde{T}'|} \frac{1}{|\widetilde{T}|} \sum_{\widetilde{X}\in\widetilde{T}'} \widetilde{X}\widetilde{X}^\top \\
&\preceq \frac{2}{\alpha} \frac{1}{|\widetilde{T}|} \sum_{\widetilde{X}\in\widetilde{T}'} \widetilde{X}\widetilde{X}^\top \preceq \frac{2}{\alpha} \frac{1}{|\widetilde{T}|} \sum_{\widetilde{X}\in\widetilde{T}} \widetilde{X}\widetilde{X}^\top,
\end{aligned}
$$

其中最后一个表达式实际上等于 $(2/\alpha)\cdot\mathbf{E}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top]$。这就完成了第一个结论的证明。

现在考虑 T 包含一个较大的稳定子集的情形。令 $T' = T\cap S$，其满足 $|T'| \geq (\alpha/2)|T|$ 且 $|T'| \geq (1-\epsilon)|S|$。将 引理 2.8 应用于 $L = \mathbf{E}_{X\sim T}[PXX^\top P]$，并取陈述中的第二个不等式，其中 T'，即可得到所需结果。

□

现在通过选择 引理 3.6，利用 引理 3.5，我们即可证明 $P = H^{\dagger/2}$。

引理 3.5 的证明。 我们将应用主要证书引理，即 引理 3.6，并取 $P = H^{\dagger/2}$。注意到 $\mathbf{E}_{X\sim T}[PXX^\top P] = H^{\dagger/2}HH^{\dagger/2} = \mathbb{P}_H$，且 $\mathbb{P}_H$ 的算子范数至多为 1。

对于最后一个不等式，我们将使用附录 A中证明的以下结果，来给出 $H^{\dagger/2}\Sigma H^{\dagger/2} - \mathbb{P}_\Sigma$的弗罗贝尼乌斯范数的上界：

引理 3.7。设 A 和 B 是两个半正定矩阵，满足 $\mathbb{P}_B \preceq \mathbb{P}_A$，或等价地满足 $\ker(A) \subseteq \ker(B)$。则 $\|ABA - \mathbb{P}_B\|_{\mathrm{F}} \le 2\|ABA - \mathbb{P}_A\|_{\mathrm{F}}$。

由于 H 是 T的二阶矩矩阵，因此 $\ker(H) \subseteq \ker(\sum_{x\in S\cap T} xx^\top)$，而它又包含于 $\ker(\Sigma)$，根据条件 (L.5) 和 定义 2.5。我们应用 引理 3.7，取 $B = \Sigma$和 $A = H$，即可得到引理陈述中的最后一个不等式。 □

引理 3.7 的证明涉及较长的技术计算，因此推迟到附录 A；这里我们给出一个简要概述。根据三角不等式，只需证明 $\|\mathbb{P}_A - \mathbb{P}_B\|_{\mathrm{F}} \le \|ABA - \mathbb{P}_A\|_{\mathrm{F}}$。由于 $\ker(A) \subset \ker(B)$，$\mathbb{P}_A - \mathbb{P}_B$ 仍然是一个秩等于 $\mathrm{rank}(A) - \mathrm{rank}(B)$ 的投影矩阵，因此其弗罗贝尼乌斯范数等于 $\mathrm{rank}(A) - \mathrm{rank}(B)$ 的平方根。另一方面，矩阵 ABA 和 $\mathbb{P}_A$ 的秩之差至少为 $\mathrm{rank}(A) - \mathrm{rank}(B)$。结合 $\mathbb{P}_A$ 具有二元特征值这一事实，我们可以将 $\|ABA - \mathbb{P}_A\|_{\mathrm{F}}$ 的下界估计为 $\mathrm{rank}(A) - \mathrm{rank}(B)$ 的平方根。

3.2 过滤：移除极端离群值

如算法概述所述，如果终止检查未通过，即整个集合 $\widetilde{\ }$ 上的期望得分 T 很大，那么我们就继续过滤样本，或将样本二分。本小节说明过滤过程的保证（引理 3.8），该过程假设集合 αm 中第 $\{\tilde{x}^\top A\tilde{x} \text{ 小和最大的元素 } \tilde{x} \in \widetilde{T}\}$ 之间的距离至多为 R，其中$R = \mathrm{poly}(1/\alpha)$。

回顾引理 3.3，$\mathrm{Var}_{\widetilde{X}\sim H^{\dagger/2}D}[\widetilde{X}^\top A\widetilde{X}] = \mathrm{E}_{\widetilde{X}\sim H^{\dagger/2}D}[(\widetilde{X}^\top A\widetilde{X} - \mathrm{E}_{\widetilde{X}\sim H^{\dagger/2}D}[\widetilde{X}^\top A\widetilde{X}])^2]$的上界为 $O(1/\alpha^2)$。根据稳定性，$\mathrm{E}_{\widetilde{X}\sim \widetilde{S}\cap\widetilde{T}}[(\widetilde{X}^\top A\widetilde{X} - \mathrm{E}_{\widetilde{X}\sim \widetilde{S}\cap\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}])^2]$也同样有此上界。

注意，这看起来几乎就是我们的得分函数 $f(X)$，只不过在 $f(X)$ 中，我们使用 $\widetilde{y}_{\mathrm{median}}$进行中心化，而不是使用 $\mathrm{E}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[\widetilde{X}^\top A\widetilde{X}]$，因为后者按定义是算法未知的量。在引理 3.9 中，我们证明这两个量之间的距离上界为 $O(R)$，其中 R 是我们假设中的分位数距离；这进而意味着，$S\cap T$ 中的内点对 $\mathrm{E}_{\widetilde{X}\sim\widetilde{T}}[f(X)]$的贡献很小。鉴于 $\mathrm{E}_{\widetilde{X}\sim\widetilde{T}}[f(X)]$很大，也就是说终止检查未通过，我们可以据此得出结论：得分贡献大部分来自离群点。因此，我们可以放心地依据得分，从数据集中随机选择一个元素予以移除，选择概率与其得分成正比；而该元素极有可能是离群点，而非内点。下文的引理 3.8 给出了内点的总得分与整个数据集的总得分之比的精确保证。

引理 *3.8*（过滤）。作出假设 *3.1*。令 A 为任意一个满足 $\|A\|_{\mathrm{F}} = 1$ 的对称矩阵。对于 $C \ge 100\sqrt{C_1}$，令 $R = C(1/\alpha)\log(1/\eta)$。定义 $\widetilde{y}_{\mathrm{median}} = \mathit{Median}(\{\widetilde{x}^\top A\widetilde{x} \mid \widetilde{x} \in \widetilde{T}\})$。定义函数 $f(\widetilde{x}) := (\widetilde{x}^\top A\widetilde{x} - \widetilde{y}_{\mathrm{median}})^2$。令 m_1 为小于 $|S|/3$ 的数。用 q_i 表示 $\{\widetilde{x}^\top A\widetilde{x} \mid \widetilde{x} \in \widetilde{T}\}$ 中第 i 小的点。

如果 $q_{|T|-m_1} - q_{m_1} \le R$且 $\mathbf{E}_{X\sim T}[f(x)] > C'R^2/\alpha^3$对于 $C' \ge 720/\epsilon_0$，也就是说，在 第 *12* 行 中的检查失败时，函数 $f(\cdot)$ 满足

$$\sum_{\widetilde{x}\in\widetilde{T}} f(\widetilde{x}) > \frac{40}{\epsilon_0}\frac{1}{\alpha^3}\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}} f(\widetilde{x})\,.$$

得分比决定了（在期望意义下）被移除的离群点数与内点数之比。在 引理 3.8 中，该比值的量级为 $1/\alpha^3$——这将使我们能够保证，在 算法 1 的整个递归执行结束时，移除的内点至多占 0.01α 的比例。更多详情请参见备注 4.5 （见第 4 节）。

本小节余下部分将证明引理 3.8和3.9。我们先证明引理 3.9：若对受污染集合上的二次多项式 p 取中位数，则如果左分位数和右分位数接近，内点上的 p 样本均值与该中位数相差不大。

引理 3.9（归一化后二次型的分位数）。令 假设 *3.1*成立。设 A 为任意满足 $\|A\|_{\mathrm{F}}=1$的对称矩阵。定义 $\widetilde{y}_{\mathrm{median}}=$中位数 $\left(\left\{\widetilde{x}^{\top}A\widetilde{x}\mid\widetilde{x}\in\widetilde{T}\right\}\right)$。设 m_1 为任意小于 $|S|/3$的数。记 q_i 为 i的第 $\left\{\widetilde{x}^{\top}A\widetilde{x}\mid\widetilde{x}\in\widetilde{T}\right\}$小的点。

假设 $q_{|T|-m_1}-q_{m_1}\le R$对 R成立：$=C\,(1/\alpha)\log(1/\eta)$ 其中 $C>1\overline{00}\sqrt{C_1}$（回顾一下，绝对常数 C_1 定义于 假设 *3.1*中）。那么，$\left|\mathbf{E}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^{\top}A\widetilde{X}]-\widetilde{y}_{\mathrm{median}}\right|\le 2R$。

证明。令 $\mu'=\mathbf{E}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^{\top}A\widetilde{X}]$。（不要与 $\hat{}\mu$和 $\widetilde{\mu}$在 假设 3.1中的$\widetilde{}$混淆；它们是 X 在样本中和总体层面的期望，而不是 μ'定义中的二次型。）

由于 $\widetilde{y}_{\mathrm{median}}$（根据定义）位于区间 $[q_{m_1},q_{|T|-m_1}]$内，且该区间的长度至多为R，因此只需证明 μ' 也位于距该区间 R 的范围内。也就是说，$\mu'\ge q_{m_1}-R$ 且$\mu'\le q_{|T|-m_1}+R$。

令 $\sigma^2:=\mathrm{Var}_{X\sim D}[(H^{\dagger/2}X)^{\top}A(H^{\dagger/2}X)]$。假设 3.1 中 S的 $(\eta,2\epsilon_0)$-稳定性（见条件 (L.3)）意味着，除去 η比例的 $\{\widetilde{x}^{\top}A\widetilde{x}$外：$\widetilde{x}\in\widetilde{S}\cap\widetilde{T}\}$都位于区间 $\mathbf{E}_{X\sim D}[(H^{\dagger/2}X)^{\top}A(H^{\dagger/2}X)]\pm 10\sigma\log(1/\eta)$ 内，因为集合 $\widetilde{S}\cap\widetilde{T}\}$的大小至少为 $|S|(1-2\epsilon_0)$（根据假设 3.1）。

因此，对于 T 中至少 $(1-\eta)|S\cap T|\ge|S|(1-2\epsilon_0-\eta)|S|$ 个点，$\widetilde{x}^{\top}A\widetilde{x}$ 的值落在区间 $\mathbf{E}_{X\sim D}[(H^{\dagger/2}X)^{\top}A(H^{\dagger/2}X)]\pm 10\sigma\log(1/\eta)$ 内。由于 $\eta\le 0.01$ 和 $\epsilon_0<1/4$，此类点的数量至少为 $|S|/3$。

换言之，存在一个区间 $[y_{\mathrm{left}},y_{\mathrm{right}}]$，其中至少包含 $|S|/3$ 个点，$\widetilde{T}$，其中

$$y_{\mathrm{left}}=\mathop{\mathbf{E}}_{X\sim D}[(H^{\dagger/2}X)^{\top}A(H^{\dagger/2}X)]-10\sigma\log(1/\eta)\,,$$
$$y_{\mathrm{right}}=\mathop{\mathbf{E}}_{X\sim D}[(H^{\dagger/2}X)^{\top}A(H^{\dagger/2}X)]+10\sigma\log(1/\eta)\,.$$

$\widetilde{}$因此，T中至多有 $|T|-|S|/3$ 个点小于 y_{left}，这意味着$q_{|T|-|S|/3}\ge y_{\mathrm{left}}$。此外，由于根据假设我们有 $m_1\le|S|/3$，这意味着$q_{|T|-m_1}\ge y_{\mathrm{left}}$。通过对称论证，我们还可以得到 $q_{m_1}\le y_{\mathrm{right}}$.

接下来，回顾一下，根据第（vi）项和引理 3 2，$\mu'=|\ _{X\sim D}[\ H^{\dagger/2}X\ ^{\top}A\ H^{\dagger/2}X\ ]-\ _{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^{\top}A\widetilde{X}]|<1/\alpha$ $|\ _{X\sim D}[\ H^{\dagger/2}X\ ^{\top}A\ H^{\dagger/2}X\ ]-$，我们有$\mathbf{E}()()\mathbf{E}()()\mathbf{E}$。因此，我们有 $q_{|T|-m_1}\ge y_{\mathrm{left}}\ge$ $\mu'-10\sigma\log(1/\eta)-1/\alpha$。重新整理后，得到 $\mu'\le q_{|T|-m_1}+10\sigma\log(1/\eta)+1/\alpha$。对称地，我们得到 $\mu'\ge q_{m_1}-10\sigma\log(1/\eta)-1/\alpha$。

最后，我们论证 $10\sigma\log(1/\eta)+1/\alpha\le R$，从而证明 $\mu'\le q_{|T|-m_1}+R$和 $\mu'\ge q_{m_1}-R$；正如我们在证明开头所论证的，这足以证明引理陈述。为此，根据 引理 3.3，我们有 $\sigma\le 5\sqrt{C_1}/\alpha$。因此，只需像引理陈述中那样，取 $R=(C/\alpha)\log(1/\eta)$，其中 $C>0$ 足够大。

□

引理 3.9 意味着，如果我们计算以内点的受污染集合的经验中位数为中心的投影的“平方偏差”，那么其大小将与内点的经验方差同阶，相差一个 R^2 因子（根据三角不等式）。此外，内点的经验方差为 $O(1/\alpha^2)$ 同阶，依据 引理 3.3。因此，如果受污染集合的方差远

大于上界时，我们就可以为每个点赋予得分，函数 f 如算法 1所定义，使得离群点的贡献远大于内点。

现在我们可以利用引理 3.9来给出引理 3.8的证明。

引理 *3.8*的证明。 证明的核心论点是，在 $q_{|T|-m_1} - q_{m_1} \le R$的假设下，$(\widetilde{x})$ 可以得到上界，如下面的断言所示。

命题 3.10。 假设满足引理 *3.8*中的条件，则有 $\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}} f(\widetilde{x}) \le 9|S|R^2$。

证明。 令 $\mu' = \mathbf{E}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^\top A\widetilde{X}]$。我们有以下一系列不等式：

$$\begin{aligned}
\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}} f(\widetilde{x}) &\le \sum_{\widetilde{x}\in\widetilde{S}} f(\widetilde{x}) && \text{(since } f(x)\ge 0\text{)}\\
&= \sum_{\widetilde{x}\in\widetilde{S}} (\widetilde{x}^\top A\widetilde{x} - \widetilde{y}_{\text{median}})^2 \\
&\le 2\sum_{\widetilde{x}\in\widetilde{S}} \left(\left(x^\top Ax - \mu'\right)^2 + \left(\mu' - \widetilde{y}_{\text{median}}\right)^2 \right) && \text{(by } (a+b)^2 \le 2a^2+2b^2\text{)}\\
&= 2|S| \operatorname*{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^\top A\widetilde{X}] + 2\sum_{\widetilde{x}\in\widetilde{S}} (\mu' - \widetilde{y}_{\text{median}})^2 \\
&\le 2|S| \operatorname*{\mathbf{Var}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^\top A\widetilde{X}] + 8R^2|S| && \text{(by Lemma 3.9)}\\
&\le 2|S| \left(\frac{72C_1}{\alpha^2}\right) + 8R^2|S| && \text{(by Lemma 3.3 and Condition (L.4))}\\
&\le 9|S|R^2\,, && \text{(by definition of } R\text{)}
\end{aligned}$$

其中，最后一行利用了 $R \ge 12\sqrt{C_1}/\alpha$，该条件在 $C \ge 12\sqrt{C_1}$时成立。至此，命题 3.10的证明完成。

□

另一方面，我们有

$$\begin{aligned}
\sum_{\widetilde{x}\in\widetilde{T}} f(\widetilde{x}) &> C'R^2|T|/\alpha^3 && \text{(by assumption)}\\
&\ge 0.5C'R^2|S|/\alpha^3 && \text{(因为 } |T| \ge |S\cap T| \ge |S|(1-2\epsilon_0) \ge |S|/2\text{)}\\
&> (360/\epsilon_0)R^2|S|/\alpha^3 && \text{(since } C' \ge 720/\epsilon_0\text{)}\\
&\ge 40/(\epsilon_0\alpha^3) \sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}} f(\widetilde{x})\,. && \text{(using Claim 3.10)}
\end{aligned}$$

至此，引理 3.8的证明完成

□

3.3 划分器：识别多个簇并递归

前面的几个小节讨论了以下情形：（i）期望得分较小；或（ii） $\widetilde{T}$上的期望得分较大，且 $\{\widetilde{X}^\top A\widetilde{X}: \widetilde{x}\in\widetilde{T}\}$ 的 α 和 $1-\alpha$ 分位数彼此接近。剩下的情形是，期望得分较大，但分位数相距甚远。在这种情况下，我们无法通过上述论证利用过滤取得进展。这实际上符合直觉，因为离群点可能确实包含另一批 $\approx \alpha m$ 样本

，这些样本服从另一个协方差迥异的高斯分布——算法无法判断哪个高斯分布对应内点。我们将论证，当期望得分较大且分位数相距甚远时，实际上很容易将样本二分为 2 个簇，使大多数内点落在二分的一侧。这使我们能够在过滤之外取得进展；这种聚类机制也使我们能够处理高斯混合模型，并确保（大致上）分别处理各个组分。

关键直觉是，根据稳定性 条件 (L.3) 和 (L.4)，我们知道，一维投影 $\{\widetilde{X}^\top A\widetilde{X}:\ \ \widetilde{X}\in\widetilde{S}\cap\widetilde{T}\}$ 下的内点必定高度集中，事实上，它们落在长度为 ~的区间内，该区间长度为 $O(1/\alpha)$。分位数范围较宽这一事实意味着，根据平均论证，值域内必定存在某个点，使得其附近的样本 $\widetilde{X}^\top A\widetilde{X}$非常少。于是，我们可以将该点用作阈值，对样本进行二分。

具体算法及其保证如下所述，见 引理 3.11 和 算法 2。

引理 *3.11*（划分器算法）。设 $T=(y_1,\dots,y_{m'})$ 是实数轴上的一个包含 m' 个点的集合。设 $S_{\mathrm{proj}}\subset T$ 是一个包含 n' 个点的集合，并且 S_{proj} 的支撑集位于长度为 r 的区间 I 上。对于每个 $i\in[m']$，设集合 T 中第 i 小的点为 q_i。假设 $q_{|T|-m_1}-q_{m_1}\geq R$，使得 $R\geq 10(m'/n')r$。那么，给定 T 和 n'、m_1，存在一个算法（算法 2）返回一个点 t，使得若定义 $T_1=\{x\in T:\ x\leq t\}$ 且 $T_2=T\setminus T_1$，则：（*i*）$\min(|T_1|,|T_2|)\geq m_1$，且（*ii*）$S_{\mathrm{proj}}\subseteq T_1$ 或 $S_{\mathrm{proj}}\subseteq T_2$。

证明。我们在下面给出算法：

算法 2 用于列表解码的划分器

1: **函数** FindDivider(T,n',m_1)
2: 令第 m_1 小的点为 q_{m_1}，第 m_1 大的点为 $q_{|T|-m_1}$。
3: 将区间 $[q_{m_1},q_{|T|-m_1}]$ 划分为 $2m'/n'$ 个大小相等的子区间。
4: 查找一个最多包含 $n'/2$ 个点的子区间 I'，并返回其中点。这样的子区间由平均论证必然存在。

考虑返回的子区间 I'的中点 t，根据构造，它至少为 q_{m_1}且至多为 $q_{|T|-m_1}$。由于 T_1 包含所有不大于 q_{m_1}的点，而 T_2 包含所有不大于$q_{|T|-m_1}$的点，因此必有 $\min(|T_1|,\ |T_2|)\geq m_1$。最后，我们验证第二个要求；当 $t\notin I$时，该要求成立。反设 $t\in I$，那么由于 I 的长度为 r，且 I' 的长度至少为$R/(2m/n')\geq 5r$，便有 $I\subseteq I'$。然而，由于 $|I'\cap T|\leq n'/2$，我们知道 I 不可能严格包含于 I'，从而得到所需的矛盾。 □

4 证明定理1.2

本节将介绍本文的主要算法结果。如前所述，只要输入点集满足第 2.3 节中的确定性条件，且分布满足一个温和的要求，该定理对高斯分布以外的分布也成立。现在，我们将以这一更一般的形式重新陈述并证明该定理。

定理 4.1 (弱列表可解码协方差估计)。设环境维度为 $d\in\mathbb{Z}_+$，C 为充分大的绝对常数，参数 $\alpha\in(0,\ 1/2)$、$\epsilon\in(0,\ \epsilon_0)$（其中 ϵ_0 为充分小的正数）以及失败概率 $\delta\in(0,\ 1/2)$ 均为算法已知。设 D 为均值为 $\mu\in\mathbb{R}^d$、协方差为 $\Sigma\in\mathbb{R}^{d\times d}$ 的分布，并假设**方差**$_{X\sim D}[X^\top AX]=O(\|\Sigma^{1/2}A\Sigma^{1/2}\|_F^2+\|\Sigma^{1/2}A\mu\|_2^2)$。存在一个多项式时间算法，使得，

对于输入 α、δ，以及集合 S（定义 *1.1*）的 *(α, ϵ)*-污染，其中 $|S| > C/(\alpha^6\epsilon_0^2)$ *log*$(1/(\alpha\delta))$，且该集合关于 D（定义 *2.5*）是 *($(\epsilon_0/40)\alpha^3, 2\epsilon_0$)*-稳定的，算法以至少 $1-\delta$ 的概率返回一个至多包含 $O(1/\alpha)$ 个集合 T_i 的列表；这些集合是输入中样本的两两不相交子集，每个 T_i 的大小至少为 $0.5\alpha m$，并且输出列表中存在一个 T_i 使得：

- 回顾污染模型（定义 *1.1*）中的记号，其中 n 是原始内点集 S 的大小，而 ℓ 是对手替换的 S 中的点数——n 和 ℓ 对于算法而言是未知的，但 $n \geq \alpha m$ 且 $\ell \leq \epsilon n$。集合 T_i 满足 $|T_i \cap S| \geq (1-0.01\alpha)(n-\ell)$。

- 记 $H_i := \mathbf{E}_{X\sim T_i}[XX^\top]$。矩阵 H_i 满足

$$\max\left(\|(H_i^\dagger)^{1/2}\Sigma(H_i^\dagger)^{1/2} - \mathbb{P}_H\|_{\mathrm{F}} \ , \ \|(H_i^\dagger)^{1/2}\Sigma(H_i^\dagger)^{1/2} - \mathbb{P}_\Sigma\|_{\mathrm{F}}\right) \leq O((1/\alpha^4)\log(1/\alpha)).$$

高斯分布 $\mathcal{N}(\mu, \Sigma)$ 满足定理 4.1 的分布假设（参见 Fact 2.1），并且从 $\mathcal{N}(\mu, \Sigma)$ 中得到的大小为 $m = C'^{d^2}\log^5(d/\alpha\delta)\,\alpha^6$ 的集合以概率 $1-\delta$ 满足 $((\epsilon_0/40)\alpha^3, 2\epsilon_0)$-稳定性，其中 ϵ_0 为常数（参见引理 2.6）。因此，定理 4.1 涵盖了具有多项式样本复杂度的高斯情形$D = \mathcal{N}(\mu, \Sigma)$，并且直接推出了定理 1.2，即本工作的主要结果。

我们通过将上一节的过程结合到一个递归算法（定理 4.1 （算法 1）中，获得了假设 3.1 所述的保证，并在记号中添加下标以表示已执行的筛选步骤数。

We s概述定理 1.2 的整体思路，然后再给出完整证明。
回顾一下 算法 1 是一个递归算法：每次调用都会反复过滤掉样本，直到终止或将数据集拆分为两个部分并递归调用自身。算法 1 的执行算法 1 因此可以被视为一棵二叉树，其中每个节点都是一次递归调用。

证明 定理 1.2 的总体思路很直接，不过实现起来需要进行技术性计算。考虑从根递归调用开始生长的子树，直到并包括某个层 j。我们对该子树的高度 j 进行归纳法，并断言该子树中必定存在一个叶节点，使得大多数内点仍保留在该叶节点的输入数据集中。

具体而言，令 $\mathcal{T}_j$ 为从根节点生长出的高度为 j 的子树。我们断言，此子树中必定存在一个叶节点，其输入集 T 满足

$$\alpha^3(\epsilon_0/40)|T| + |S \setminus T| \leq (j+1)\alpha^3(\epsilon_0/20)m + \ell\,, \tag{3}$$

回顾一下，α 是内点所占的比例，ϵ_0 是对手移除内点的最大比例，ℓ 是实际（未知）移除的内点数量，而 m 是对手返回的原始数据集大小。左边的表达式同时反映了 a）我们意外移除的内点数量（通过 $|S \setminus T|$ 这一项），以及 b）我们移除的离群点与内点的相对比例，这可以通过比较左边的两项得出。

对于归纳步骤，我们基本上需要分析一次递归调用的执行。我们证明 (3) 蕴含 假设 3.1，因此归纳步骤可以遵循 第 3 节中概述的分类讨论——要么终止，要么决定对样本集进行过滤或二分。唯一缺少的部分是对重复过滤的分析：引理 3.8在第 3.2 节中证明，离群点的得分总和是内点得分的 $\Theta(1/\alpha^3)$ 倍，这意味着在期望意义下，我们移除的离群点比内点多 $\Theta(1/\alpha^3)$。为了将其转化为高概率陈述，我们使用标准的（次）鞅论证，见第 4.1 节。

下面给出定理 4.1的完整证明。

定理 *4.1* 的证明。 由于最终误差保证不依赖于 ϵ，不失一般性，在以下证明中我们使用最大污染层 $\epsilon = \epsilon_0 = 0.01$ 。我们还将用字母 η 表示表达式 $(\epsilon_0/40)\alpha^3$ ，如定理陈述中关于内点集的稳定性假设所示。我们将通过归纳法证明 定理 4.1，并关键地使用以下事实：“并集界在条件化下也成立”。

事实 4.2. 若事件 A 发生的概率为 $1-\tau_1$，且在事件 A发生的条件下，事件 B 发生的概率为 $1-\tau_2$，则事件 A 和 B 同时发生的概率至少为 $1-\tau_1-\tau_2$。

证明 定理 4.1的主要技术难点在于 引理 4.4，它是关于递归层级的归纳命题，指出对于任意一层，以高概率而言，要么 (i) 在前一层有一次递归调用，其输入包含了大部分内点，且根据定义该调用已终止；要么 (ii) 在当前层存在某次递归调用，其输入包含了大部分内点。我们将使用以下记号（定义 4.3）来表示算法递归执行过程中的各种对象。

定义 *4.3*（递归树的记号）。我们定义一棵有根二叉树 $\mathcal{T}$，它对应于递归算法（算法 *1*）的执行。第 *0* 层的根节点对应于 *CovarianceListDecoding*（算法 *1*）的顶层调用。对于每个非叶节点，其左子节点对应于第 *24* 行对 *CovarianceListDecoding* 的调用，右子节点对应于第 *25* 行的调用。因此，每个节点都由其所在的层以及它在该层中的位置唯一确定。我们用 $\mathcal{T}_{i,j}$ 表示第 j 层的 i^{th} 节点，其中节点编号也从 *0* 开始（例如，根节点为 $\mathcal{T}_{0,0}$）。对于节点 $\mathcal{T}_{i,j}$，我们用 $T^{(i,j)}$ *(0)* 表示相应递归调用的输入数据集。为了表示节点 $\mathcal{T}_{i,j}$ 执行主循环第 t^{th} 次迭代时的工作数据集，我们使用记号 $T^{(i,j)}$*(t)* 和 $\widetilde{T}$*(i,j)* (t) ，其用法与算法 *1* 中 $T_{(t)}$ 和 $T_{(t)}$ 的完全相同。最后， $\mathcal{T}_j$ （即使用单一下标）表示从根节点生长出的子树，其中包含截至并包括第 j^{th} 层的所有节点。

引理 4.4。 在 定理 *4.1* 的情形下，考虑 CovarianceListDe-coding *(定义 4.3)*。那么，对于任意 $j \in \{0,\dots,9/\alpha\}$，我们以至少$1-0.01\alpha j\delta$ 的概率有以下结论成立：

(i)（递归过程中规模减小）每个节点的输入$\mathcal{T}_{i,\,j}$ 层 j 满足 $|T^{(i,j)}{}_{(0)}| \le m - j\alpha m/9$

(ii)（不相交输入）考虑从根节点生长并在（包括）第 j 层截断的子树 $\mathcal{T}_j$。 $\mathcal{T}_j$ 中的所有叶节点都具有不相交的输入集。请注意，叶节点不一定都位于第 j 层，有些可能位于更早的层。

(iii)（移除内点数量的上界）考虑从根节点 $\mathcal{T}_j$ 生长并在第 j层截断（但包括该层）的子树。存在一个叶节点$\mathcal{T}_{i',\,j'}$ 位于 $\mathcal{T}_j$ 中（但叶节点 j' 的层不一定等于 j），使得其输入 $T^{(i',j')}{}_{(0)}$ 满足

$$\alpha^3(\epsilon_0/40)|T^{(i',j')}_{(0)}| + |S \setminus T^{(i',j')}_{(0)}| \le (j+1)\alpha^3(\epsilon_0/20)m + \ell \tag{4}$$

其中，与定义 *1.1*中一样， m 表示根节点 $\mathcal{T}_{0,0}$ 的原始输入集大小，而$\ell \le \epsilon_0 n$ 表示对手在 S 中替换的样本数。特别地，截至 $T^{(i',j')}$(0，算法移除的内点数至多为$(j+1)\alpha^3(\epsilon_0/20)m$。

备注 4.5。我们说明一下，为什么式 (4) 中使用了 α 的三次幂；这与引理 3.8 中出现的三次幂相同。引理 3.8 指出，所有离群点的总得分比所有内点的总得分高出 $1/\alpha^3$ 倍。α 的三个幂次源于以下原因：1）内点可能仅占所有样本的 α；2）递归树的深度可能达到 $O(1/\alpha)$，这意味着一个集合可能被过滤 $O(1/\alpha)$ 次；3）定理 4.1 保证，我们移除的内点至多占 $O(\alpha)$。我们还指出，通过调整算法中的参数，并提高协方差估计保证中 poly$(1/\alpha)$ 的幂次，可以使评分函数比（从而使式 (4) 中的 poly(α) 因子）达到我们期望的 α 的任意高次幂。这样即可保证移除的内点比例至多为更小的 poly(α)。

证明。我们对层数 j 进行归纳法，以证明 引理 4.4 。对于 $j = 0$ 的基本情形，条件 1 和 2 显然成立。至于条件 3，在 $j = 0$ 处，我们有 $|T^{(0,0)}{}_{(0)}| = m$ 和 $|S \setminus T^{(0,0)}{}_{(0)}| \le \ell$，根据污染模型的定义（定义 1.1），这直接推出$\alpha^3(\epsilon_0/40)|T^{(0,0)}{}_{(0)}| + |S \setminus T^{(0,0)}{}_{(0)}| \le \alpha^3(\epsilon_0/40)$ $m + \ell \le \alpha^3(\epsilon_0/20)m + \ell$，如所需。

为证明归纳情形，我们假设引理的陈述对某个 j 成立，并证明该陈述对情形 $j+1$ 也成立。

条件 1 和 2 显然成立，这是因为递归算法通过过滤移除点后，会使用 FindDivider划分输入集，并保证两个输出集合的大小都至少为 $m_1 = \alpha m/9$ (引理 3.11)。

条件 3 回顾 $T^{(i,j)}{}_{(t)}$ 这一记号，定义 4.3 中用它表示变量 T_t 在 t^{th} 层递归调用的第 i^{th} 次迭代后所包含的点集。j。

根据 事实 4.2，我们将条件化于条件 3 对 $\mathcal{T}_j$成立这一情形，其中该子树在第 j 层截断（包含该层），并证明条件 3 对 $\mathcal{T}_{j+1}$也成立，其中该子树在第 $j+1$层截断，除非发生概率为 $0.01\alpha\delta$的事件。具体而言，条件化意味着存在该子树 $\mathcal{T}_{i',j'}$ 中的一个叶节点 $\mathcal{T}_j$，其输入集 $T^{(i',j')}{}_{(0)}$ 满足条件 3 中关于 $j' \le j$的不等式。如果 $j' < j$，那么结论成立（因为它仍然是叶节点，而且公式（4）右侧所需的大小下界只会更大）。否则，在 $j' = j$的情况下，我们必须分析递归树 $\mathcal{T}_j$中与该叶节点关联的递归调用的执行过程。

首先进行分析，我们验证输入集 $T^{(i',j)}{}_{(0)}$ 满足 假设 3.1。我们首先检查 $|S \cap T^{(i',j)}{}_{(0)}| \ge n($ $1 - 2\epsilon_0)$：

$$
\begin{aligned}
|S \cap T^{(i',j)}_{(0)}| &= |S| - |S \setminus T^{(i',j)}_{(0)}| \\
&\ge n - (j+1)\alpha^3(\epsilon_0/20)m - \ell && \text{(by the inductive hypothesis)} \\
&\ge n - (9/\alpha + 1)\alpha^3(\epsilon_0/20)m - \ell && (j \le 9/\alpha) \\
&\ge n - \alpha^2\epsilon_0 m - \epsilon_0 n && (\ell \le \epsilon_0 n) \\
&\ge n - \alpha\epsilon_0 n - \epsilon_0 n && (n \ge \alpha m) \\
&\ge n - 1.5\epsilon_0 n\,, && (5)
\end{aligned}
$$

其中，最后一行用到了 $\alpha \le 1/2$。接下来只需验证 $|T^{(i',j)}{}_{(0)}| \le (1/\alpha)|S|$，而这显然成立，因为我们最初使用的输入集 $T^{(0,0)}{}_{(0)}$ 满足该条件：$|T^{(i',j)}{}_{(0)}| \le |T^{(0,0)}{}_{(0)}| = m \le (1/\alpha)n = (1/\alpha)|S|$。

在一次递归调用中，算法会迭代地移除点（通过 第 19 行），直到终止（第 12 行）或调用 FindDivider 并继续递归（第 24 行 和 25）。将终止或运行 FindDivider 之前的迭代记为 $t^*_{i'}$。我们需要论证，在第 $t^*_{i'}$ 次迭代之前对点进行的迭代过滤仍能大致保持条件 3。下面陈述并在 引理 4.6 中证明的 引理 4.6，概括了基于过滤的鲁棒算法中的标准鞅论证。直观地说，如果输入集 $T^{(i',j)}{}_{(0)}$ 满足 假设 3.1，那么条件 3 大致得以保持。

引理 4.6。 回顾 假设 *3.1* 中的记号和假设。考虑 CovarianceListDecoding*(*T_0*)* 的一次执行，其中 $T_0 = T$ *(*满足 T 中的假设，并且 假设 *3.1)*，且假设 $|T_0| \geq C'/(\alpha^6\epsilon_0^2)\log(1/(\alpha\delta))$。进一步假设 $|S \setminus T_0| \leq 1.5\epsilon_0|S|$*(* 这是对 假设 *3.1* 的加强 *)*。此外，令 T_t 表示 算法 *1* 循环迭代过程中经过的数据集。那么，以至少 $1 - 0.01\alpha\delta$的概率，有

$$\alpha^3(\epsilon_0/40)|T_t| + |S \setminus T_t| \leq \alpha^3(\epsilon_0/40)|T_0| + |S \setminus T_0| + \alpha^3(\epsilon_0/40)|T_0| \,,$$

对所有迭代 t 同时成立，直到 CovarianceListDecoding 的执行进入第 *14* 行或第 *22* 行。

具体而言，对输入集 $T^{(i',j)}{}_{(0)}$ 应用引理 4.6，并用失败概率 $0.01\alpha\delta$ 替代 δ。该引理适用，因为其第一个要求 $|S \setminus T^{(i',j)}{}_{(0)}| \leq 1.5\epsilon_0|S|$ 已在式 (5) 中验证；对于第二个要求，我们有 $|T^{(i',j)}{}_{(0)}| \geq |T^{(i',j)}{}_{(0)} \cap S| \geq (1 - 1.5\epsilon_0)|S| \geq C'/(\alpha^6\epsilon_0^2)\log(1/(\alpha\delta))$，其中最后一个不等式用到了式 (5)，而下一步用到了我们的假设 $|S| > C/(\alpha^6\epsilon_0^2)\log(1/(\alpha\delta))$。于是，该引理推出，以至少 $1 - 0.01\alpha\delta$ 的概率，$T^{(i',j)}{}_{(t^*_{i'})}$ 满足

$$\alpha^3(\epsilon_0/40)|T^{(i',j)}_{(t^*_{i'})}| + |S \setminus T^{(i',j)}_{(t^*_{i'})}| \leq (j+1)\alpha^3(\epsilon_0/20)m + \ell + (\epsilon_0/40)\alpha^3 m \,. \tag{6}$$

现在，递归调用要么在第 $t^*_{i'}$ 次迭代终止，要么进入第 22 行。前一种情况下，我们就完成了（因为此时它是一个叶节点）。否则，递归调用会使用 FindDivider 将 $T_{(i',j)}{}_{(t^*_{i'})}$ 划分为 T' 和 T''。我们需要证明 T' 和 T'' 中至少有一个满足条件 3（对于情形 $j+1$），这要用到引理 3.11。现在，我们来推导调用引理 3.11时应使用的参数和集合。

利用 (i) 原始内点集 S 的稳定性性质中的条件 (L.3)，以及 (ii) 对任意 H 和 A，$(H^{\dagger/2}X)^\top A(H^{\dagger/2}X) - \mathrm{E}_{X\sim\mathcal{N}(\mu,\Sigma)}[(H^{\dagger/2}X)^\top A(H^{\dagger/2}X)]$ 是 X 中的偶次二次式这一事实，必然有 $1-\eta$ 比例的点（回顾一下，η 是 $(\epsilon_0/40)\alpha^3$ 的简写）$x \in S$ 满足

$$\begin{aligned}\left|(H^{\dagger/2}x)^\top A(H^{\dagger/2}x) - \mathop{\mathbf{E}}_{X\sim\mathcal{N}(\mu,\Sigma)}[(H^{\dagger/2}X)^\top A(H^{\dagger/2}X)]\right| &\leq 10\log(2/\eta)\sqrt{\mathop{\mathbf{Var}}_{X\sim D}[(H^{\dagger/2}X)^\top A(H^{\dagger/2}X)]} \\ &\leq 50\log(2/\eta)\sqrt{\frac{C_1}{\alpha^2}} \,. \qquad \text{(Lemma 3.3)}\end{aligned}$$

将具有上述性质的点集合 $x \in S$ 记为 $S_{H,\,A} \subseteq S$，并令 $S^{(i',j)}{}_{\text{core}} = T^{(i',j)}{}_{(t^*_{i'})} \cap S_{H,\,A}$。因此，我们将对 $S_{\text{proj}} = \{(H^{\dagger/2}x)^\top A(H^{\dagger/2}x) : x \in S^{(i',j)}{}_{\text{core}}\}$, $T = \{(H^{\dagger/2}x)^\top A(H^{\dagger/2}x) : x \in T^{(i',j)}{}_{(t^*_{i'})}\}$ 应用引理 3.11，且直径为 $r = 100\log(2/\eta)\sqrt{C_1}/\alpha$。还需验证 —

对于 $q_{m_1} = q_{\text{left}}$ 和 $q_{|T^{(i',j)}(t^*_{i'})|-m_1} = q_{\text{right}}$ 属于 $T^{(i',j)}(t^*_{i'})$ 的情形，有 $q_{|T^{(i',j)}(t^*_{i'})|-m_1} - q_{m_1} \geq 10(|T^{(i',j)}(t^*_{i'})|/|S_{\text{proj}}|)r$，以确保应用引理 3.11 是有效的。

为说明这一点，回顾一下，每当调用 FindDivider 时，第 17 行都会失败，这意味着 $q_{|T^{(i',j)}(t^*_{i'})|-m_1} - q_{m_1} \geq R$。因此，我们需要验证 R 在 算法 1 中的定义满足 $R \geq 10(|T^{(i',j)}(t^*_{i'})|/|S_{\text{proj}}|)r$。关键步骤是给 $|S_{\text{proj}}| = |S^{(i',j)_{\text{core}}}|$ 一个下界。由于 $S^{(i',j)_{\text{core}}} \subset S$，我们可以将上述量下界估计为

$$
\begin{aligned}
|S^{(i',j)}_{\text{core}}| &\geq |S| - |S \setminus S_{H,A}| - |S \setminus T^{(i',j)}_{(t^*_{i'})}| && (S^{(i',j)}_{\text{core}} = T^{(i',j)}_{(t^*_{i'})} \cap S_{H,A}) \\
&\geq (1-\eta)|S| - |S \setminus T^{(i',j)}_{(t^*_{i'})}| && (\text{by the definition of } S_{H,A}) \\
&\geq (1-\eta)|S| - (j+1)\alpha^3(\epsilon_0/20)m - \ell - (\epsilon_0/40)\alpha^3 m && (\text{by (6)}) \\
&\geq (1-\eta)n - (j+2)\alpha^3(\epsilon_0/20)m - \ell && (|S| = n \text{ and basic inequality}) \\
&\geq (1-\eta)n - (j+2)\alpha^2(\epsilon_0/20)n - \ell && (n \geq \alpha m) \\
&\geq (1-\eta)n - \alpha\epsilon_0 n - \ell && (j \leq 9/\alpha \text{ and } \alpha < 1/2) \\
&\geq (1-\eta)n - \alpha\epsilon_0 n - \epsilon_0 n && (\ell \leq \epsilon_0 n) \\
&\geq (1-\eta-2\epsilon_0)n \geq n/2 && (\eta \leq 0.001 \text{ 和 } \epsilon_0 = 0.01)
\end{aligned}
$$

结合 $|T^{(i',j)}(t^*_{i'})| \leq m$ 这一事实，可得 $10(|T^{(i',j)}(t^*_{i'})|/|S_{\text{proj}}|)r \leq 20(m/n)r \leq (20/\alpha)r$。回顾 $r = 100\log(2/\eta)\sqrt{C_1/\alpha}$，$\eta = (\epsilon_0/40)\alpha^3$ 和 $\epsilon_0 = 0.001$，算法 1 中 R 的定义满足 $R \geq 6000\sqrt{C_1}(1/\alpha^2)\log(1/(\epsilon_0\alpha)) \geq (20/\alpha)r \geq 10(|T^{(i',j)}(t^*_{i'})|/|S_{\text{proj}}|)r$。

已知引理 3.11 的应用有效，该引理便保证T'或 T''包含全部 $S^{(i',j)_{\text{核心}}}$。不失一般性，我们假设这对 T'成立。现在我们针对 T'上的情形 $j+1$ 验证条件 3:

$$
\begin{aligned}
\alpha^3(\epsilon_0/40)|T'| + |S \setminus T'| &\leq \alpha^3(\epsilon_0/40)|T^{(i',j)}_{(t^*_{i'})}| + |S \setminus T'| && (\text{因为 } T' \subset T^{(i',j)}_{(t^*_{i'})}) \\
&\leq \alpha^3(\epsilon_0/40)|T^{(i',j)}_{(t^*_{i'})}| + |S \setminus S^{(i',j)}_{\text{core}}| && (\text{by Lemma 3.11}) \\
&\leq \alpha^3(\epsilon_0/40)|T^{(i',j)}_{(t^*_{i'})}| + |S \setminus T^{(i',j)}_{(t^*_{i'})}| + |S \setminus S_{H,A}| \\
& && (\text{根据 } S^{(i',j)_{\text{核心}}} \text{ 的定义以及并集界}) \\
&\leq \alpha^3(\epsilon_0/40)|T^{(i',j)}_{(t^*_{i'})}| + |S \setminus T^{(i',j)}_{(t^*_{i'})}| + \eta n && (\text{by the definition of } S_{H,A}) \\
&\leq (j+1)\alpha^3(\epsilon_0/20)m + \ell + (\epsilon_0/40)\alpha^3 m + \eta n && (\text{by Equation 6}) \\
&\leq (j+1)\alpha^3(\epsilon_0/20)m + \ell + (\epsilon_0/20)\alpha^3 m && (\text{since } \eta = (\epsilon_0/40)\alpha^3 \text{ and } n \leq m) \\
&= (j+2)\alpha^3(\epsilon_0/20)m + \ell\,. && (7)
\end{aligned}
$$

因此，情形 $j+1$下的 T' 满足条件 3，从而完成了该引理的归纳证明。

□

现在我们利用该引理完成定理 4.1的证明。

根据 引理 4.4，我们知道，以至少 $1 - 0.09\delta$ 的概率，生成到第 $9/\alpha$ 层的递归树，即 $\mathcal{T}_{9/\alpha}$，满足引理的保证。特别地，由于 (i) 根据 第 22 行 的算法 1，任何递归调用的输入规模都必须至少为 $m_1 = \alpha m/9$，且 (ii) 将 $j = m/m_1 = 9/\alpha$ 应用于引理 4.4 的条件 1 可得输入规模上界为 0，因此我们可以得出

算法的执行必须在层 $9/\alpha$之前完全终止。根据引理 4.4 可知，递归树中存在一个叶节点 i，其输入满足引理陈述第三部分中的条件 3。也就是说，令 j 表示该叶节点所在的层，则有

$$\alpha^3(\epsilon_0/40)|T_{(0)}^{(i,j)}| + |S \setminus T_{(0)}^{(i,j)}| \le (j+1)\alpha^3(\epsilon_0/20)m + \ell\,. \tag{8}$$

递归节点 $\mathcal{T}_{i,\,j}$ 从输入集 $T^{(i,j)}{}_{(0)}$ 开始，在终止之前，它可能会执行若干筛选步骤。与 引理 4.4 的证明一样，我们将使用 引理 4.6 来分析递归调用终止前的“工作集”。若用 t^* 表示终止前的筛选步骤数，引理 4.6 表明，除概率为 $0.01\alpha\delta$的情况外，我们有

$$\alpha^3(\epsilon_0/40)|T_{(t^*)}^{(i,j)}| + |S \setminus T_{(t^*)}^{(i,j)}| \le (j+1)\alpha^3(\epsilon_0/20)m + \ell + (\epsilon_0/40)\alpha^3 m\,. \tag{9}$$

注意，根据 事实 4.2，总失败概率的上界为 $0.09\delta + 0.01\alpha\delta$，小于 δ。

现在我们已准 备证明 定理 4.1的第一个要点。上述不等式意味着

$$\begin{aligned}
|T_{(t^*)}^{(i,j)} \cap S| &= |S| - |S \setminus T_{(t^*)}^{(i,j)}| \\
&\ge n - \ell - (\epsilon_0/20)(j+1)\alpha^3 m - (\epsilon_0/40)\alpha^3 m && \text{(by (9))} \\
&\ge n - \ell - (\epsilon_0/20)(j+2)\alpha^3 m \\
&\ge n - \ell - (\epsilon_0/20)(j+2)\alpha^2 n && (n \ge \alpha m) \\
&\ge n - \ell - (\epsilon_0/2)\alpha n && (j \le 9/\alpha \text{ and } \alpha < 1/2) \\
&\ge n - \ell - \epsilon_0(1-\epsilon_0)\alpha n && (1/2 \le 1-\epsilon_0) \\
&\ge n - \ell - \epsilon_0\alpha(n-\ell) && (\ell \le \epsilon_0 n) \\
&= (n-\ell)(1-\epsilon_0\alpha)\,. && (10)
\end{aligned}$$

(10) 进一步推出 $|T^{(i,j)}{}_{(t^*)} \cap S| \ge (n-\ell)(1-\epsilon_0\alpha) \ge n(1-\epsilon_0)(1-\epsilon_0\alpha) \ge (\alpha/2)m$ (因为 $n \ge \alpha m$、ϵ_0 和 α 都很小），这意味着集合 $T^{(i,j)}{}_{(t^*)}$ 确实会被返回。至此，定理 4.1第一个要点的证明完成。

现在转到第二个要点。定义 $H = \mathrm{E}_{X \sim T^{(i,j)}{}_{(t^*)}}[XX^\top]$ 后，我们需要证明 $\|H^{\dagger/2}\Sigma H^{\dagger/2} - HH^\dagger\|_{\mathrm{F}} \lesssim (1/\alpha^4)\log^2(1/\alpha)$。我们将应用引理 3.5 来证明第二个要点中的界，这需要检查假设 3.1 是否由 $T^{(i,j)}{}_{(t^*)}$ 满足。此前，在 (10) 正下方，我们已经检查过 $|T^{(i,j)}{}_{(t^*)} \cap S| \ge (\alpha/2)|T^{(i,j)}{}_{(t^*)}|$。唯一剩下需要检查的条件是 $|T^{(i,j)}{}_{(t^*)}| \le |S|/\alpha$，但这显然成立，因为 $T^{(i,j)}{}_{(t^*)} \subset T^{(0,0)}{}_{(0)}$（递归算法的原始输入集），且根据定义 1.1，$|T^{(0,0)}{}_{(0)}| \le |S|/\alpha$。因此，引理 3.5 得出

$$\begin{aligned}
\left\|H^{\dagger/2}\Sigma H^{\dagger/2} - \mathbb{P}_H\right\|_{\mathrm{F}}^2 &\lesssim \frac{1}{\alpha}\operatorname*{\mathbf{Var}}_{\widetilde{X} \sim T_{(t^*)}^{(i,j)}}[\widetilde{X}^\top A \widetilde{X}] + \frac{1}{\alpha^2} \\
&\le \frac{1}{\alpha}\operatorname*{\mathbf{E}}_{\widetilde{X} \sim T_{(t^*)}^{(i,j)}}[f(\widetilde{X})] + \frac{1}{\alpha^2} && (\mathbf{Var}(Y) \le \mathbf{E}[(Y-c)^2] \text{ for any } c \in \mathbb{R}) \\
&\lesssim \frac{R^2}{\alpha^4} && (\text{c.f. Line 12 of Algorithm 1}) \\
&\lesssim \frac{1}{\alpha^8}\log^2\left(\frac{1}{\alpha}\right)\,, && (R = \Theta((1/\alpha^2)\log(1/\alpha)), \text{ noting that } \epsilon_0 = 0.01)
\end{aligned}$$

其中倒数第二个不等式用到了如下事实：如果算法终止，那么必定有 $\mathbf{E}_{\widetilde{X}\sim T(i,\ j)\,(t^*)}[f(\widetilde{X})]\le C'R^2/\alpha^3$根据 第 12 行，在 算法 1中。对两边开平方可得

$$\left\|H^{\dagger/2}\Sigma H^{\dagger/2}-\mathbb{P}_H\right\|_{\mathrm{F}}\lesssim\frac{1}{\alpha^4}\log\left(\frac{1}{\alpha}\right)$$

对于 $\left\|H^{\dagger/2}\Sigma H^{\dagger/2}-\mathbb{P}_\Sigma\right\|_{\mathrm{F}}$，同样的保证也成立，这也可由 引理 3.5 得出。

最后，我们检查输出列表中只返回 $O(1/\alpha)$ 个集合。这由构造可知：(i) 第13行 仅允许输出大小至少为 $\Omega(\alpha m)$ 的集合；(ii) 起初只有 m 个点；(iii) 根据 引理 4.4，所有叶节点的输入集两两不相交。 □

4.1 引理 4.6 的证明

现在我们证明　　利用标准的鞅论证，过滤不会移除过多内点　　t.

引理 4.6。 回顾 假设 *3.1* 中的记号和假设。考虑一次 CovarianceListDecoding*(*T_0*)* 的执行，其中 $T_0=T$*(*其中 T 满足 假设 *3.1*中的假设*)*，并假设 $|T_0|\ge C'/(\alpha^6\epsilon_0^2)\log(1/(\alpha\delta))$。进一步假设 $|S\setminus T_0|\le 1.5\epsilon_0|S|$（这是对 假设 *3.1* 的加强）。此外，记 T_t 为算法循环迭代期间的数据集 算法 *1*。那么，以至少 $1-0.01\alpha\delta$ 的概率，有

$$\alpha^3(\epsilon_0/40)|T_t|+|S\setminus T_t|\le\alpha^3(\epsilon_0/40)|T_0|+|S\setminus T_0|+\alpha^3(\epsilon_0/40)|T_0|\,,$$

对所有迭代同时 t 直到执行CovarianceListDecoding 进入第 *14* 行或第 *22* 行。

证明。 我们将使用 假设 3.1 中针对 算法 1单次调用的记号。记 t 为迭代次数。另定义停止时间 t_{end} 为首次满足以下条件的迭代：$\alpha^3(\epsilon_0/40)|T_{t_{\mathrm{end}}}|+|S\setminus T_{t_{\mathrm{end}}}|>2\epsilon_0|S|$，或者迭代转入 第 12 行 或 第 22 行 而非 第 19 行 （也就是说，当 $\mathbf{E}_{X\sim T_{t_{\mathrm{end}}}}[f(H_t^{\dagger/2}X)]\le C'R^2/\alpha^3$，或者当 $q_{\mathrm{right}}-q_{\mathrm{left}}>R$ 对于迭代 t_{end}成立时）。现在，定义$\Delta_t=\alpha^3(\epsilon_0/40)|T_{\min\{t,t_{\mathrm{end}}\}}|+|S\setminus T_{\min\{t,t_{\mathrm{end}}\}}|$。

为了证明该引理，我们将证明，在第一个 t^* （如果存在）满足$\alpha^3(\epsilon_0/40)|T_{t^*}|+|S\setminus T_{t^*}|>\alpha^3(\epsilon_0/40)|T_0|+|S\setminus T_0|+\alpha^3(\epsilon_0/40)|T_0|$时，$\Delta_{t^*}>\alpha^3(\epsilon_0/40)|T_0|+|S\setminus T_0|+\alpha^3(\epsilon_0/40)|T_0|$ 也成立。之后，我们将证明 Δ_t 是一个次鞅，并利用次鞅尾界证明，序列 Δ_t 以高概率在$|T_0|$ 步的整个轨迹中保持较小。

第一步很容易证明，因为 $\alpha^3(\epsilon_0/40)|T_0|+|S\setminus T_0|+\alpha^3(\epsilon_0/40)|T_0|<\alpha^3(\epsilon_0/20)|T_0|+1.5\epsilon_0|S|\le\alpha^2(\epsilon_0/20)|S|+1.5\epsilon_0|S|\le 2\epsilon_0|S|$的阈值，其中第一个不等式由引理假设得出。因此，若 $t^*\le t_{\mathrm{end}}$ 存在，则它满足 $\Delta_{t^*}=\alpha^3(\epsilon_0/40)|T_{t^*}|+|S\setminus T_{t^*}|$。

现在我们需要证明 Δ_t 关于序列 $T_{\min\{t,t_{\mathrm{end}}\}}$是一个次鞅。如果$t\ge t_{\mathrm{end}}$，那么 Δ_{t+1} 根据定义等于 Δ_t，因此 $\mathbf{E}[\Delta_{t+1}\mid T_{\min\{t,t_{\mathrm{end}}\}}]=\Delta_t$。否则，我们有 $\mathbf{E}[\Delta_{t+1}\mid T_t]$ 等于 Δ_t 加上被移除的内点的期望数量，再减去 $\alpha^3(\epsilon_0/40)$乘以我们的过滤所移除的所有点的期望数量。由于停止条件不满足，我们有 $|S\setminus T_t|\le 2\epsilon_0|S|$，这意味着 假设 3.1 对于 T_t仍然成立，同时 引理 3.8 的其他条件也成立。因此，我们可以应用 引理 3.8，得到

$$\mathbf{E}[\Delta_{t+1}\mid T_t]=\Delta_t+\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}_t}f(\widetilde{x})-\alpha^3\frac{\epsilon_0}{40}\sum_{\widetilde{x}\in\widetilde{T}_t}f(\widetilde{x})\le\Delta_t\,.$$

总结而言，上述情况分析表明，$\{\Delta_t\}_{t\in\mathbb{N}}$ 是关于$T_{\min\{t,t_{\text{end}}\}}$的次鞅。

我们还注意到，对于每个 t，$|\Delta_t - \Delta_{t+1}| \le 1$ 其成立的概率为 1。因此，利用次鞅的标准 Azuma-Hoeffding不等式（事实 2.2），可得对于每个 $t \le |T_0|$,

$$\mathbf{Pr}\left[\Delta_t - \Delta_0 > (\epsilon_0/40)\alpha^3|T_0|\right] \le e^{-(2/40^2)\cdot\epsilon_0^2\alpha^6|T_0|}.$$

对所有 $t \in [|T_0|]$应用并集界，可得

$$\mathbf{Pr}\left[\exists t \in [m] : \Delta_t > \Delta_0 + \alpha^3(\epsilon_0/40)|T_0|\right] \le |T_0|e^{-(2/40^2)\cdot\epsilon_0^2\alpha^6|T_0|} \le 0.01\alpha\delta\,, \tag{11}$$

其中，最后一个不等式利用了：对于充分大的绝对常数 C'，$|T_0| > \frac{C'}{\alpha^6\epsilon_0^2}\log\left(\frac{1}{\alpha\delta}\right)$。

□

5 鲁棒学习高斯混合模型的应用

本节证明我们的结果在鲁棒学习高斯混合模型中的应用。

定理 *1.4*（含离群点鲁棒的*GMM*聚类与协方差估计）。设环境维度为 $d \in \mathbb{N}$，且 C' 为充分大的绝对常数。设 $\alpha \in \in (0, 1/2)$、$\epsilon \in \in (0, \epsilon_0)$（其中 ϵ_0为某个充分小的正数），并且算法已知失败概率 $\delta \in \in (0, 1/2)$。存在一个运行时间为 $\tilde{O}(m^2d^2)$ 的算法，使得对于输入 α、δ 以及来自某个未知的、含有 k 个组分的高斯混合模型 $\sum_{p=1}^{k}\alpha_p\mathcal{N}(\mu_p, \Sigma_p)$ 的 α^6个 ϵ-受污染样本，其中样本数量满足 $m > C'^{d^2}\log^5(d/\alpha\delta)$，且该模型如定义 *1.3* 所述，所有 Σ_p均为满秩，所有 α_p 均满足 $\alpha_p \ge \alpha$，而 $k \le \frac{1}{\alpha}$ 为算法所未知；该算法以至少 $1-\delta$ 的概率（对受污染样本及算法的随机性取概率），返回一个至多包含 k 个两两不相交的样本子集 $\{T_i\}$ 的列表，使得：

- 对于第 p 个高斯分布分量，记 S_p 为内点集 S 中从组分 p抽取的样本集合。令 n_p 为 S_p 的大小，令 ℓ_p 为对手替换的 S_p 中的点数——除了$\mathbf{E}[n_p] = \alpha_p m \ge \alpha m$ 对每个 p 和 $\sum_p \ell_p \le \epsilon\alpha m$成立外，算法并不知道 n_p 和 ℓ_p。那么，对于混合分布中的每个高斯分量 p，返回列表中都存在一个集合 T_{i_p}，使得 $|T_{i_p} \cap S_p| \ge (1-0.01\alpha)(n_p - \ell_p)$。

- 对于每个组分 p，返回列表中都存在一个样本集合 T_{i_p}，使得定义$H_{i_p} = \mathbf{E}_{X\sim T_{i_p}}[XX^\top]$ 后，有 H_{i_p}满足 $\|H_{i_p}^{-1/2}\Sigma_p H_{i_p}^{-1/2} - I\|_{\mathbb{F}} \lesssim (1/\alpha^4)\log(1/\alpha)$。

- 设 Σ 为高斯混合的（总体层面的）协方差矩阵。对于任意两个组分 $p \ne p'$，若 $\|\Sigma^{-1/2}(\Sigma_p - \Sigma_{p'})\Sigma^{-1/2}\|_{\mathbb{F}} > C(1/\alpha)^5\log(1/\alpha)$，且常数 C 足够大，则前一条要点中的集合 T_{i_p} 和 $T_{i_{p'}}$ 必定不同。

我们证明前两条要点可直接由定理 1.2推得，方法是将每个高斯分量 $D = \mathcal{N}(\mu_p, \Sigma_p)$ 作为内点分布。最后一条要点的证明需要一些更复杂的线性代数引理。这里我们只给出最后一条要点的高层次证明，并将这些引理各自的证明推迟到 附录 A。

证明。首先注意，若样本数为足够大的 $\text{poly}(d, \frac{1}{\alpha}, \log\frac{1}{\delta})$，则以至少$1-\delta/4$的概率，来自组分 p 的内点数 n_p 对所有组分 p 都至少为 $0.99\alpha m$。我们将在该事件成立的条件下进行讨论。

此外，注意到在定义 1.3中定义的污染模型里，最多总共有 $\epsilon\alpha m$ 个点可能被污染，这意味着 $\ell_p \le \epsilon\alpha m \le (\epsilon/0.99)n_p$（因为 $n_p \ge 0.99\alpha m$）。

因此，我们将定理 1.2 应用于所有 k 高斯分量，以 0.99α 替代 α、以 $\epsilon/0.99$ 替代 ϵ，并取失败概率 $\delta/(4k)$。由此可知，利用并集界，并使用引理 2.6 中的 $m = \mathsf{poly}(d, 1/\alpha, \log\, k/\delta) = \mathsf{poly}(d, 1/\alpha, \log\, 1/\delta)$ 个样本（因为 $k \lesssim 1/\alpha$，否则高斯混合的权重将大于 1），以至少 $1-\delta/4$ 的概率，定理中的两个要点对所有高斯分量都成立，这由定理 1.2 直接保证。

还需证明，该算法在其输出列表中返回的子集数量至多为 k。注意，对于算法针对各个组分输出的子集，其中已输出的点总数为$(1-0.01\alpha)\sum_{p=1}^{k}(n_p-\ell_p)$。回顾 $\sum_p n_p = m$ 以及 $\sum_p \ell_p \le \epsilon\alpha m$。因此，与这些组分对应的输出子集中，点的总数至少为 $m-(\epsilon+0.01)\alpha m \ge m-0.4\alpha m$（这里用到了 ϵ 有一个足够小的常数上界）。由于（i）输入中只有 m 个点，且（ii）所有输出子集都是输入的互不相交子集，这意味着，在任何其他与高斯分布组分无关的子集中，至多还能输出 $0.4\alpha m$个点。然而，根据定理 1.2，任一输出子集的大小至少为$0.5(0.99\alpha)m$，因此，输出列表中不可能存在任何与高斯分布组分无关的额外子集。这就证明了输出列表中的子集数量至多为 k。

我们来看定理的最后一条。我们将证明，算法输出中属于同一集合 T_i 的任意两个组分 p 和 p' 都满足 $\|\Sigma^{-1/2}\Sigma_p\Sigma^{-1/2} - \Sigma^{-1/2}\Sigma_{p'}\Sigma^{-1/2}\|_{\mathrm{F}} \lesssim \frac{1}{\alpha^5}\log(1/\alpha)$。取逆否命题即可得到定理的最后一条。

令 H 为该集合的二阶矩矩阵，即 $H = \mathbf{E}_{X\sim T_i}[XX^\top]$。那么 H 满足 $\|I - H^{\dagger/2}\Sigma_p H^{\dagger/2}\|_{\mathrm{F}} \le r$ 且 $\|I - H^{\dagger/2}\Sigma_{p'} H^{\dagger/2}\|_{\mathrm{F}} \le r$ 对于 $r \lesssim (1/\alpha^4)\log(1/\alpha)$，由定理的第二条可知。此外，$H$ 是满秩的，因为 H 是组分 p 的一个较大经验点子集的二阶矩矩阵，因此条件 (L.5) 意味着$\ker(H)$ 包含于 $\ker(\Sigma_p)$，而后者为空，因为根据定理假设，Σ_p 是满秩的。

现在我们将应用以下结果（引理 5.1），从而推断 Σ_p 和 $\Sigma_{p'}$ 必定彼此接近，具体含义见最终定理的最后一条要点。该引理的证明推迟到 附录 A。

引理 5.1。考虑两个任意的满秩正定矩阵 Σ_1 和 Σ_2，并假设存在一个满秩正定矩阵 H，使得 ***(i)*** $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}} \le \rho$ 且 ***(ii)***$\|I - H^{\dagger/2}\Sigma_2 H^{\dagger/2}\|_{\mathrm{F}} \le \rho$。那么，对于任意一个满秩正定矩阵 Σ，有

$$\|\Sigma^{-1/2}\Sigma_1\Sigma^{-1/2} - \Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}\|_{\mathrm{F}} \le 5\rho\max(\|\Sigma^{-1/2}\Sigma_1\Sigma^{-1/2}\|_{\mathrm{op}}, \|\Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}\|_{\mathrm{op}}).$$

现在我们将应用 引理 5.1 于 $\Sigma_1 = \Sigma_p$ 和 $\Sigma_2 = \Sigma_{p'}$，并对引理右侧进行上界估计，从而得到最终定理的最后一条要点。

我们断言，由于组分 p 的质量至少为 α，$\Sigma^{-1/2}\Sigma_p\Sigma^{-1/2}$ 的算子范数有上界 $1/\alpha$。事实上，我们有

$$\Sigma_p \preceq \mathop{\mathbf{E}}_{X\sim\mathcal{N}(\mu_p,\Sigma_p)}[(X-\mu)(X-\mu)^\top] \preceq \frac{1}{\alpha_p}\cdot \mathop{\mathbf{E}}_{X\sim\sum_i \alpha_i\mathcal{N}(\mu_i,\Sigma_i)}[(X-\mu)(X-\mu)^\top] = \frac{1}{\alpha_p}\cdot\Sigma\,.$$

这尤其意味着 $\Sigma^{-1/2}\Sigma_p\Sigma^{-1/2} \preceq \frac{1}{\alpha_p} I \preceq \frac{1}{\alpha} I$，因此$\|\Sigma^{-1/2}\Sigma_p\Sigma^{-1/2}\|_{\mathrm{op}} \le \frac{1}{\alpha}$。对于组分 p'，同样的结论也成立。回顾 $r \lesssim (1/\alpha^4)\log(1/\alpha)$，引理 5.1 右侧的上界可取为 $O((1/\alpha^5)\log(1/\alpha))$，由此得到定理中的最后一条，正如上文所述。

□

我们以简要讨论引理 5.1的证明作为本节的结尾。该引理指出，若存在一个矩阵 H，同时是 Σ_1 和 Σ_2的近似平方根，即

若 $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}}$ 和 $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}}$ 都有上界，则 Σ_1 和 Σ_2 不会相差太远，因为 $\Sigma_1 - \Sigma_2$ 的弗罗贝尼乌斯范数也必定有界。事实上，对于任意正定 Σ，都可以将其推广到归一化形式 $\Sigma^{-1/2}(\Sigma_1 - \Sigma_2)\Sigma^{-1/2}$。

为简化讨论，暂且假设 $\Sigma = I$。要证明引理 5.1，一个自然的初步想法是：如果 $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}}$ 有界，例如不超过某个量 ρ，那么 $\|HH^\top - \Sigma_1\|_{\mathrm{F}}$ 也应有界，其上界与 $O(\rho)$ 有关，再乘以经过适当缩放的 Σ_1 的算子范数。遗憾的是，这并不成立：即使对于 $\Sigma_1 = I$，我们也可以令 H 的第一个对角元素为无穷大，其余对角元素均为 1；此时 $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}}$ 仅为 1，但 $HH^\top - \Sigma_1$ 在第一个维度上却是无穷大。因此，关键的观察是，一个稍弱的结论确实成立：在一个仅取决于 H 而不取决于 Σ_1的标准正交基上，$HH^\top$ 在大多数维度上都近似于 Σ_1，具体来说，是在 $d - O(\rho^2)$ 个维度上。引理 A.3 对这一结论作了形式化表述。因此，在这些维度上，我们可以界定 Σ_1 与 Σ_2 之间的差异（按弗罗贝尼乌斯范数计）。对于其余维度，在这些维度上 $HH^\top$ 既不能很好地近似 Σ_1，也不能很好地近似 Σ_2，我们可以将这些维度对 $\Sigma_1 - \Sigma_2$的弗罗贝尼乌斯范数的贡献界定为 $\max(\|\Sigma_1\|_{\mathrm{op}}, \|\Sigma_2\|_{\mathrm{op}})$ 乘以这类维度的数量，而该数量很小。合并这些界，即可得到所需的引理 5.1。

参考文献

[ABH+72] D. F. Andrews、P. J. Bickel、F. R. Hampel、P. J. Huber、W. H. Rogers 和 J. W. Tukey。《位置的稳健估计：综述与进展》。普林斯顿大学出版社，美国新泽西州普林斯顿，1972 年。[AK05] S. Arora 和 R. Kannan。学习分离的非球形高斯混合。《应用概率年刊》，15(1A):69–92，2005 年。[BBH+21] M. Brennan、G. Bresler、S. Hopkins、J. Li 和 T. Schramm。统计查询算法与低次测试几乎等价。载于《学习理论会议》，2021 年。[BBV08] M. F. Balcan、A. Blum 和 S. Vempala。基于相似度函数的聚类判别框架。载于《*STOC*》，第 671–680 页，2008 年。[BDH+20] A. Bakshi、I. Diakonikolas、S. B. Hopkins、D. M. Kane、S. Karmalkar 和 P. K. Kothari。高斯及其他非球形混合分布的抗离群点聚类。载于《第*61*届*IEEE*计算机科学基础研讨会（*FOCS*）论文集》，2020 年。[BDJ+22] A. Bakshi、I. Diakonikolas、H. Jia、D. M. Kane、P. K. Kothari 和 S. S. Vempala。稳健学习任意 k 个高斯分布的混合。载于《第*54*届*ACM*计算理论年度研讨会（*STOC*）论文集》，2022 年。[Ber18] D. S. Bernstein。《标量、向量与矩阵数学：理论、事实与公式》。普林斯顿大学出版社，美国新泽西州普林斯顿，修订增补版，2018 年。[BK21] A. Bakshi 和 P. Kothari。列表可译码子空间恢复：多项式时间内实现维数无关误差。载于《第*32*届年度离散算法研讨会（*SODA*）论文集》，2021 年。[BS10] M. Belkin 和 K. Sinha。分布族的多项式时间学习。载于*FOCS*，第 103–112 页，2010 年。

[CMY20] Y. Cherapanamjeri、S. Mohanty 和 M. Yau。近线性时间的列表可解码均值估计。载于第61届IEEE计算机科学基础研讨会（FOCS）论文集，2020 年。[CSV17] M. Charikar、J. Steinhardt 和 G. Valiant。从不可信数据中学习。载于第49届ACM计算理论年度研讨会，第 47–60 页，2017 年。[Das99] S. Dasgupta。高斯混合模型学习。载于第40届计算机科学基础年度研讨会论文集，第 634–644 页，1999 年。[DHKK20] I. Diakonikolas、S. B. Hopkins、D. Kane 和 S. Karmalkar。稳健学习任意可聚类的高斯混合。CoRR，abs/2005.06417，2020 年。[DJKS22] A. Das、A. Jain、W. Kong 和 R. Sen。利用批次高效进行列表可解码回归。CoRR，abs/2211.12743，2022 年。[DK19] I. Diakonikolas 和 D. M. Kane。算法高维鲁棒统计的最新进展。arXiv 预印本 arXiv:1911.05911，2019 年。[DK23] I. Diakonikolas 和 D. M. Kane。《算法高维鲁棒统计》。剑桥大学出版社，2023 年。[DKK+16] I. Diakonikolas、G. Kamath、D. M. Kane、J. Li、A. Moitra 和 A. Stewart。无需计算难解性的高维鲁棒估计量。载于第57届 IEEE 计算机科学基础研讨会（FOCS）论文集，第 655–664 页，2016 年。[DKK20] I. Diakonikolas、D. M. Kane 和 D. Kongsgaard。通过迭代式多重过滤进行列表可解码均值估计。载于《神经信息处理系统进展》第33卷（NeurIPS），2020 年。[DKK+21] I. Diakonikolas、D. M. Kane、D. Kongsgaard、J. Li 和 K. Tian。近似主成分分析时间内的列表可解码均值估计。载于《神经信息处理系统进展》第34卷（NeurIPS），第 34 卷，2021 年。[DKK+22a] I. Diakonikolas、D. M. Kane、S. Karmalkar、A. Pensia 和 T. Pittas。通过差分对过滤进行列表可解码稀疏均值估计。载于《神经信息处理系统进展第35卷》（NeurIPS），2022 年。[DKK+22b] I. Diakonikolas、D. M. Kane、D. Kongsgaard、J. Li 和 K. Tian。通过列表可解码均值估计实现近线性时间的混合模型聚类。载于第54届ACM计算理论年度研讨会（STOC）论文集，2022 年。[DKP+21] I. Diakonikolas、D. M. Kane、A. Pensia、T. Pittas 和 A. Stewart。列表可解码线性回归的统计查询下界。载于《神经信息处理系统进展》第34卷（NeurIPS），2021 年。[DKS17] I. Diakonikolas、D. M. Kane 和 A. Stewart。高维高斯分布及高斯混合稳健估计的统计查询下界。载于第58届IEEE计算机科学基础研讨会（FOCS）论文集，第 73–84 页，2017 年。[DKS18] I. Diakonikolas、D. M. Kane 和 A. Stewart。列表可解码鲁棒均值估计与球面高斯混合模型学习。载于第50届ACM计算理论年度研讨会（STOC）论文集，2018 年。

[DS07] S. Dasgupta 和 L. J. Schulman。针对分离球形高斯分布混合的期望最大化算法的概率分析。机器学习研究期刊，8:203–226，2007。[HR09] P. J. Huber 和 E. M. Ronchetti。稳健统计。威利概率与统计丛书。John Wiley & Sons, Inc.，美国新泽西州霍博肯，2009。[Hub64] P. J. Huber。位置参数的稳健估计。数学统计年刊，35(1):73–101，1964年3月。[IK22] M. Ivkov 和 P. K. Kothari。列表可解码协方差估计。载于：第54届ACM计算理论年度研讨会（STOC）论文集，2022。[Kan21] D. M. Kane。高斯混合的稳健学习。载于：第32届年度离散算法研讨会（SODA）论文集，2021。[KK10] A. Kumar 和 R. Kannan。利用谱范数和k均值算法进行聚类。载于：2010年 IEEE 第51届年度计算机科学基础研讨会，页299–308。IEEE，2010。[KKK19] S. Karmalkar、A. Klivans 和 P. Kothari。列表可解码线性回归。载于：神经信息处理系统进展，页7423–7432，2019。[Kot21] P. Kothari。高斯混合的鲁棒学习。TCS+ Talk，2021。https://youtu.be/Yjk0XzvjyOw?t=2593.[KS17] P. K. Kothari 和 J. Steinhardt。通过松弛张量范数改进无先验聚类。CoRR，abs/1711.07465，2017。[Li18] J. Li。稳健机器学习及其拓展的原则性方法。博士论文，麻省理工学院，2018。[LM21] A. Liu 和 A. Moitra。确定高斯混合的稳健可学习性。载于：第53届年度 ACM 计算理论研讨会（STOC）论文集，2021。[LRV16] K. A. Lai、A. B. Rao 和 S. Vempala。均值与协方差的无先验估计。载于：第57届 IEEE 计算机科学基础研讨会（FOCS），页665–674，2016。[MV10] A. Moitra 和 G. Valiant。确定高斯混合的多项式可学习性。载于：FOCS，页93–102，2010。[RY20a] P. Raghavendra 和 M. Yau。通过平方和进行列表可解码学习。载于：第十四届年度 ACM-SIAM 离散算法研讨会论文集，页161–180。SIAM，2020。[RY20b] P. Raghavendra 和 M. Yau。列表可解码子空间恢复。载于：第33届年度学习理论会议（COLT）论文集，2020。[Vem22] S. Vempala。多项式时间内稳健地学习任意高斯分布的混合。班加罗尔理论研讨会系列，2022。https://youtu.be/eF9XTJRbFTg?t=3107.[VW04] S. Vempala 和 G. Wang。用于学习混合模型的谱算法。计算机与系统科学期刊，68(4):841–860，2004。

[ZS22] S. Zeng 和 J. Shen。列表可解码稀疏均值估计。载于神经信息处理系统进展第*35*卷（*NeurIPS*），2022 年。

附录

A 线性代数结果

<escape>在本节中，我们证明一些所需的线性代数结果。具体而言，我们给出引理 5.1和3.7 的证明，这些结果已在前文中使用。在此之前，我们先陈述以下将在后文中用到的事实：</escape>

事实 A.1 (矩阵的伪逆)。 令 G 为方阵，并定义 $R = G^\dagger$。令 G 的奇异值分解为 $U\Lambda V^\top$。则 $R = V\Lambda^\dagger U^\top$、$\mathbb{P}_G = U\Lambda\Lambda^\dagger U^\top$ 和 $R = R\mathbb{P}_G$。此外，令 Δ 为任意对角矩阵，使得若 $\Lambda_{i,\,i} = 0$，则 $\Delta_{i,\,i} = 0$，并定义 $L = V\Delta V^\top$。则 $L = LRG$。

证明。 第一条由 [Ber18，第 8], 二条由 [Ber18，命题 8.1.7 (xii)], ，第三条由前两条以及 $\Lambda^\dagger\Lambda\Lambda^\dagger = \Lambda$这一事实得出。最后的断言由以下一系列等式得出：

$$\begin{aligned} LRG &= V\Delta V^\top V\Lambda^\dagger U^\top U\Lambda V^\top \\ &= V\Delta V^\top V\Lambda^\dagger U^\top U\Lambda V^\top \\ &= V\Delta\Lambda^\dagger\Lambda V^\top \\ &= V\Delta V^\top = L, \end{aligned}$$

其中我们利用了 Δ 是对角矩阵，且对于每个满足 $\Delta_{i,\,i} \neq 0$, $\Lambda_i^\dagger{}_{,\,i}\Lambda_{i,\,i} = 1$的 i。

□

事实 A.2 ([Ber18，事实 8.4。16])。 设 A 和 B 是两个方阵，且假设 A 满秩。那么 $\mathbb{P}_{BA} = \mathbb{P}_B$。

A.1 相对弗罗贝尼乌斯范数到弗罗贝尼乌斯范数：引理 5.1 的证明

本节给出 引理 5.1 的证明。

引理 5.1. 考虑两个任意的满秩正定矩阵 Σ_1 和 Σ_2，并假设存在一个满秩正定矩阵 H，使得 ***(i)*** $\|I - H^{\dagger/2}\Sigma_1 H^{\dagger/2}\|_{\mathrm{F}} \leq \rho$ 且 ***(ii)***$\|I - H^{\dagger/2}\Sigma_2 H^{\dagger/2}\|_{\mathrm{F}} \leq \rho$。那么，对于任意满秩正定矩阵 Σ，我们有

$$\|\Sigma^{-1/2}\Sigma_1\Sigma^{-1/2} - \Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}\|_{\mathrm{F}} \leq 5\rho\max(\|\Sigma^{-1/2}\Sigma_1\Sigma^{-1/2}\|_{\mathrm{op}}, \|\Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}\|_{\mathrm{op}}).$$

证明引理 *5.1*。设 $G = H^{-1/2}\Sigma^{1/2}$。设 $\widetilde{\Sigma}_1 = \Sigma^{-1/2}\Sigma_1\Sigma^{-1/2}$ 和 $\widetilde{\Sigma}_2 = \Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}$均为正定矩阵。由于 Σ 是满秩的，我们得到

$$\begin{aligned} \Sigma_1 &= \Sigma^{1/2}\Sigma^{-1/2}\Sigma_1\Sigma^{-1/2}\Sigma^{1/2} = \Sigma^{1/2}\widetilde{\Sigma}_1\Sigma^{1/2} \\ \Sigma_2 &= \Sigma^{1/2}\Sigma^{-1/2}\Sigma_2\Sigma^{-1/2}\Sigma^{1/2} = \Sigma^{1/2}\widetilde{\Sigma}_2\Sigma^{1/2}. \end{aligned}$$

这直接推出 $G\widetilde{\Sigma}_1 G^\top = I - H\Sigma_1 H$ 和 $G\widetilde{\Sigma}_2 G^\top = I - H\Sigma_2 H$。总而言之，我们得到如下结果：

$$\|I - G\widetilde{\Sigma}_1 G^\top\|_{\mathrm{F}} \leq \rho \quad \text{and} \quad \|I - G\widetilde{\Sigma}_2 G^\top\|_{\mathrm{F}} \leq \rho.$$

我们首先陈述一个结果（将在本节末尾证明），它将相对 Frobenius 范数下的接近性转化为差矩阵的弗罗贝尼乌斯范数（在大多数方向上）。

引理 A.3。 设 G 是一个 $d \times d$ 矩阵， B 是一个维度相同的任意对称矩阵。回顾， $\mathbb{P}_G$ 表示到 G的值域上的投影矩阵。假设

$$\|\mathbb{P}_G - GBG^\top\|_{\mathrm{F}} \le \rho$$

对于某个 $\rho \ge 1$。设 G 的奇异值分解为 $G = U\Lambda V^\top$ ，其中 U 和 V 是正交归一矩阵， Λ 是对角矩阵。

那么存在一个 $d \times d$ 矩阵 L ，使得

1.（L 捕获 G 的方向） $L = V\Delta_B V^\top$ ，其中 $\Delta_B \preceq \Lambda\Lambda^\dagger$ 是一个二元对角矩阵。也就是说， Δ_B 的对角线元素均属于 $\{0,1\}$，并且 Δ_B 中所有值为 *1* 的对角线元素也是 Λ 中非零的对角线元素。

2.（ L 的秩） L 的秩满足 $\|\Lambda - \Delta_B\|_{\mathrm{F}} \le 2\rho$ ，并且

3. (closeness of $G^\dagger$ and B along L) $\left\|L\left(RR^\top - B\right)L^\top\right\|_{\mathrm{F}} \le 2\rho\|B\|_{\mathrm{op}}$, *where* $R = G^\dagger$.

现在我们将沿用引理 A.3 陈述中的记号，两次应用引理 A.3。特别地， $G = U\Delta V^\top$ 和 $R = G^\dagger$ 的奇异值分解。引理 A.3 表明，存在两个二元对角矩阵 Δ_1 和 Δ_2，其中 Δ_1 对应 Σ_1，而 Δ_2 对应 Σ_2，使得：(i) $\|I - \Delta_1\|_{\mathrm{F}}$ 和 $\|I - \Delta_1\|_{\mathrm{F}}$ 均为 2ρ；(ii)
$V\Delta_1 V^\top$ $RR^\top - \widetilde{\Sigma}_1$ $V\Delta_1 V^\top$ $V\Delta_2 V^\top$ $RR^\top - \widetilde{\Sigma}_2$ $V\Delta_2 V^\top$ ( ) 和 ( ) 的弗罗贝尼乌斯范数分别至多为 $2\rho\|\Sigma_1\|_{\mathrm{op}}$ 和 $2\rho\|\Sigma_2\|_{\mathrm{op}}$。令 $\Delta = \Delta_1 \times \Delta_2$，并定义 $L = V\Delta V^\top$。由 (i) 可知，
$\|I - L\|_{\mathrm{F}} = \|I - \Delta\|_{\mathrm{F}} \le 4\rho$。由于 $L \preceq V\Delta_1 V^\top$ 且 L 是投影矩阵，结合上述 (ii) 和事实 2.4 可得

$$\|L(RR^\top - \widetilde{\Sigma}_1)L^\top\|_{\mathrm{F}} \le 2\rho\|\widetilde{\Sigma}_1\|_{\mathrm{op}}.$$

对 Σ_2也同理。由三角不等式可得

$$\|L(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2)L^\top\|_{\mathrm{F}} \le 4\rho \max\left(\|\widetilde{\Sigma}_1\|_{\mathrm{op}}, \|\widetilde{\Sigma}_2\|_{\mathrm{op}}\right).$$

由于 L 是投影矩阵，三角不等式意味着

$$\begin{aligned}\left\|\left(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2\right)\right\|_{\mathrm{F}} &\le \left\|L\left(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2\right)L^\top\right\|_{\mathrm{F}} + \left\|(I-L)\left(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2\right)(I-L)^\top\right\|_{\mathrm{F}} \\ &\lesssim \left\|L\left(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2\right)L^\top\right\|_{\mathrm{F}} + \|I-L\|_{\mathrm{F}} \cdot \left\|\left(\widetilde{\Sigma}_1 - \widetilde{\Sigma}_2\right)\right\|_{\mathrm{op}} \\ &\lesssim \rho \cdot \max\left(\|\widetilde{\Sigma}_1\|_{\mathrm{op}}, \|\widetilde{\Sigma}_2\|_{\mathrm{op}}\right).\end{aligned}$$

证明至此完成。

□

现在我们给出引理 A.3的证明。

引理 A.3。 设 G 是一个 $d \times d$ 矩阵， B 是一个维数相同的任意对称矩阵。回顾一下， $\mathbb{P}_G$ 表示到 G 的值域上的投影矩阵。假设

$$\|\mathbb{P}_G - GBG^\top\|_{\mathrm{F}} \le \rho$$

对于某个 $\rho \ge 1$。设 G 的奇异值分解为 $G = U\Lambda V^\top$ ，其中 U和 V是正交归一矩阵， Λ是对角矩阵。

那么存在一个 $d \times d$ 矩阵 L ，使得

1.（L 捕获 G的方向） $L = V\Delta_B V^\top$，其中 $\Delta_B \preceq \Lambda\Lambda^\dagger$是某个二元对角矩阵。也就是说， Δ_B 的对角元素均属于 $\{0,1\}$，此外，Δ_B 的所有值为 *1* 的对角元素也都是 Λ中的非零对角元素。

2.（ L的秩） L 的秩满足 $\|\Lambda - \Delta_B\|_{\mathrm{F}} \le 2\rho$，并且
3.（沿 L的 $G^\dagger$ 与 B 的接近性） $\left\|L\left(RR^\top - B\right)L^\top\right\|_{\mathrm{F}} \le 2\rho\|B\|_{\mathrm{op}}$，其中 $R = G^\dagger$。

证明。令 u_1， u_2，……， u_d为 U的列，并令 v_1， v_2，……， v_d为 V的列。令 Λ 的元素为 λ_1， λ_2，……， λ_d，根据奇异值分解的性质，这些元素均为非负数。令 $\mathcal{J}$为 $[d]$中满足 $\lambda_i > 0$的 i的集合。

定义矩阵 $F\colon = U^\top GBG^\top U$和 $B'\colon = V^\top BV$。由于 $G^\top = V\Lambda U^\top$，我们有

$$F = U^\top GBG^\top U = U^\top U\Lambda V^\top BV\Lambda U^\top U = \Lambda V^\top BV\Lambda = \Lambda B'\Lambda.$$

因此， F 的第（i,j）个元素为$B'_{i,j}\lambda_i\lambda_j$。

现在定义矩阵 $M\colon = U^\top\mathbb{P}_G U$。根据 事实 A.1，我们有 $\mathbb{P}_G = U\Lambda\Lambda^\dagger U^\top$。因此，$M = \Lambda\Lambda^\dagger$，它是一个对角矩阵，其对角元素为 $M_{ii} = 1$，对每个 $i \in \mathcal{J}$，而 $M_{ii} = 0$ 在其他情况下。

利用弗罗贝尼乌斯范数在正交变换下的不变性，我们得到以下一系列等式：

$$\begin{aligned}
&\|\mathbb{P}_G - GBG^\top\|_{\mathrm{F}}^2 \\
&\quad= \left\|U^\top\left(\mathbb{P}_G - GBG^\top\right)U\right\|_{\mathrm{F}}^2 \\
&\quad= \|M - F\|_{\mathrm{F}}^2 \\
&\quad\ge \sum_{i\in\mathcal{J}}\left(M_{i,i} - F_{i,i}\right)^2 \\
&\quad= \sum_{i\in\mathcal{J}}\left(1 - B'_{i,i}\cdot\lambda_i^2\right)^2
\end{aligned} \tag{12}$$

令 $\mathcal{I} \subseteq \mathcal{J}$为 $\mathcal{J}$中满足$\lambda_i^2 B'_{i,i} < (1/2)$的 i的集合。于是 $\sum_{i\in\mathcal{I}}\left(1 - B'_{i,i}\cdot\lambda_i^2\right)^2 > |\mathcal{I}|/4$。因此，若 $|\mathcal{I}|$ 大于 $4\rho^2$，则弗罗贝尼乌斯范数的平方大于 r^2，这就产生了矛盾。

如果 $i \in \mathcal{J}\setminus\mathcal{I}$，那么 $\lambda_i^2 \ge 1/(2B'_{i,i})$，而不仅仅是 $\lambda_i > 0$。

定义矩阵 L为 $L\colon = \sum_{i:\ i\in\mathcal{J}\setminus\mathcal{I}} v_i v_i^\top$，它等于 $V\Delta_B V^\top$，其中 Δ_B是一个对角矩阵，其第 i 个对角元素在 $i \in \mathcal{J}\setminus\mathcal{I}$时等于 1，否则等于 0。根据定义， L的秩为$|\mathcal{J}| - |\mathcal{I}|$；此外， Δ_B 中值为 1 的对角元素始终是 Λ中的非零对角元素。

令 $R = G^\dagger$，根据 事实 A.1，可得 $R = V\Lambda^\dagger U^\top$。因此 $LR = V\Delta_B V^\top V\Lambda^\dagger U^\top = V\Delta_B\Lambda^\dagger U^\top$，从而 LR 的奇异值至多为

$$\frac{1}{\min_{i\in\mathcal{J}\setminus\mathcal{I}}\lambda_i} \le \max_i\sqrt{2B'_{i,i}} \le \sqrt{2\|B'\|_{\mathrm{op}}} = \sqrt{2\|V^\top BV\|_{\mathrm{op}}} = \sqrt{2\|B\|_{\mathrm{op}}}.$$

我们首先还需要两个观察结果： $L = LRG$ 和 $LRR^\top L^\top = LR\mathbb{P}_G R^\top L^\top$，这两者均可由事实 A.1得出。最后，我们证明 $RR^\top$ 和 B 沿着 L中的方向彼此接近。

$$\begin{aligned}
&\|L(RR^\top - B)L^\top\|_{\mathrm{F}} \\
&\quad= \|LRR^\top L^\top - LBL^\top\|_{\mathrm{F}} \\
&\quad= \|LR\mathbb{P}_G R^\top L^\top - LRGBG^\top R^\top L^\top\|_{\mathrm{F}}
\end{aligned}$$

$$
\begin{aligned}
&= \|LR(\mathbb{P}_G - GBG^\top)R^\top L^\top\|_{\mathrm{F}} \\
&\le \|LR\|_{\mathrm{op}}\|\mathbb{P}_G - GBG^\top\|_{\mathrm{F}}\|R^\top L^\top\|_{\mathrm{op}} \\
&= \|LR\|_{\mathrm{op}}^2\|\mathbb{P}_G - GBG^\top\|_{\mathrm{F}} \\
&\le 2\|B\|_{\mathrm{op}} \cdot \rho. \qquad (13)
\end{aligned}
$$

证明至此完成。 □

A.2 引理 3.7的证明

现在我们重述并证明引理 3.7。

引理 3.7。 设 A 和 B 是两个半正定矩阵，且满足 $\mathbb{P}_B \preceq \mathbb{P}_A$，或等价地，$\ker(A) \subseteq \ker(B)$。则 $\|ABA - \mathbb{P}_B\|_{\mathrm{F}} \le 2\|ABA - \mathbb{P}_A\|_{\mathrm{F}}$。

证明。 应用三角不等式可得

$$
\|ABA - \mathbb{P}_B\|_{\mathrm{F}} \le \|ABA - \mathbb{P}_A\|_{\mathrm{F}} + \|\mathbb{P}_A - \mathbb{P}_B\|_{\mathrm{F}}.
$$

因此，只需证明第二项 $\|\mathbb{P}_B - \mathbb{P}_A\|_{\mathrm{F}}$ 小于 $\|ABA - \mathbb{P}_A\|_{\mathrm{F}}$。设 d_A 为 A 的秩， d_B 为 B的秩。我们将证明以下两个结果，由此可得所需的界：

$$
\|\mathbb{P}_B - \mathbb{P}_A\|_{\mathrm{F}}^2 = |d_B - d_A| \le \|ABA - \mathbb{P}_B\|_{\mathrm{F}}^2 \qquad (14)
$$

在证明的其余部分，我们将给出上述两个不等式的证明。

我们先来看（14）中的第二个不等式。根据引理假设， A 的零空间是 B的零空间的子空间。这等价于说， B 的列空间是 A的列空间的子空间。特别地，这也意味着 $d_B \le d_A$。容易看出，$\mathrm{rank}(ABA) \le \mathrm{rank}(B) \le \mathrm{rank}(A) = \mathrm{rank}(\mathbb{P}_A)$，且 $\mathbb{P}_A$ 的特征值为 0 或1。现在我们给出一个简单的结果，对这类矩阵的弗罗贝尼乌斯范数进行下界估计，下界为秩之差：

断言 A.4 （秩差界）。 设 A' 是一个对称矩阵，且A' 的特征值属于 $\{0, 1, -1\}$。设 B' 是任意矩阵。则 $\|B' - A'\|_{\mathrm{F}}^2 \ge \mathrm{rank}(A') - \mathrm{rank}(B')$。

证明。 令 $u_1, \dots u_d$为 A'的正交特征向量，其对应的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_d$，满足对于 $i \le$ ，rank(A') 为 $|\lambda_i| = 1$ ，否则为 $\lambda_i = 0$ 。不失一般性，假设 rank(B') ≤rank(A')，否则结论显然成立。令 $\mathcal{I}$ 为满足 rank(A') 且 $B'u_i = 0$的 i的集合。

$$
\begin{aligned}
\|B' - A'\|_{\mathrm{F}}^2 = \sum_{i=1}^{d} \|B'u_i - A'u_i\|_2^2 \ge \sum_{i=1}^{\mathrm{rank}(A')} \|B'u_i - A'u_i\|_2^2 \\
= \sum_{i=1}^{\mathrm{rank}(A')} \|B'u_i - \lambda_i u_i\|_2^2. \ge \sum_{i\in\mathcal{I}} \|\lambda_i u_i\|_2^2 = |\mathcal{I}|.
\end{aligned}
$$

由于 B' 的秩小于 A'的秩，因此在 $\{u_i$ 中，至多有 rank(B') 个 u_i 使得 $B'u_i$ 非零： $i \le$ rank(A')}。因此， $|\mathcal{I}|$ 至少为 rank(A') − rank(B')。 □

根据 断言 A.4，我们有 $\operatorname{rank}(\mathbb{P}_B) - \operatorname{rank}(ABA) \leq \|\mathbb{P}_B - ABA\|_{\mathrm{F}}^2$。因此，这意味着所需的不等式成立：$|d_B - d_A| \leq \|ABA - \mathbb{P}_A\|_{\mathrm{F}}^2$。

现在我们来证明（14）中的第一个不等式。注意， $\mathbb{P}_A$ 和 $\mathbb{P}_B$ 都是半正定矩阵，其特征值只能是1 或 0。下面的结果给出了 $\mathbb{P}_A - \mathbb{P}_B$ 的弗罗贝尼乌斯范数上界，并以它们秩的差为依据。

断言 A.5. 若 A 和 B 是两个半正定矩阵，且 $\mathbb{P}_B \preceq \mathbb{P}_A$，则 $\|\mathbb{P}_A - \mathbb{P}_B\|_{\mathrm{F}}^2 = \text{秩}(A) - \text{秩}(B)$。

证明。 设 d_A 为 A 的秩， d_B 为 B的秩。我们有 $d_B = \text{秩}(\mathbb{P}_B) \leq \text{秩}(\mathbb{P}_A) = d_A$。任取 $\mathbb{R}^d$中的一组标准正交基 $v_1, \ldots, v_d$，使得 $v_1, \ldots, v_{d_A}$ 张成 $\mathbb{P}_A$的列空间，而 $v_1, \ldots, v_{d_B}$ 张成 B的列空间。由于 $\mathbb{P}_B$ 的列空间是 $\mathbb{P}_B$ 的列空间的子空间，这样的标准正交基总是存在的。

此外，由于 $\mathbb{P}_A$ 的特征值要么为 0，要么为 1，因此对于 $i \leq d_A$，有 $\mathbb{P}_A v_i = v_i$，其他情况下为 0。类似地，对于 $i \leq d_B$，有 $\mathbb{P}_B v_i = v_i$，其他情况下为 0。

我们还将使用以下基本事实：对于任意方阵 J 以及任意一组标准正交基向量 u_1、……、u_d，都有 $\|J\|_{\mathrm{F}}^2 = \sum_{i=1}^d \|Ju_i\|_2^2$。将这一事实应用于 $\mathbb{P}_A - \mathbb{P}_B$ 和上述标准正交基 v_1、……、v_d，可得：

$$\begin{aligned}\|\mathbb{P}_A - \mathbb{P}_B\|_{\mathrm{F}}^2 &= \sum_{i=1}^{d} \|\mathbb{P}_A v_i - \mathbb{P}_B v_i\|_2^2 = \sum_{i=1}^{d_B} \|\mathbb{P}_A v_i - \mathbb{P}_B v_i\|_2^2 + \sum_{d_B+1}^{d_A} \|\mathbb{P}_A v_i - \mathbb{P}_B v_i\|_2^2 = \sum_{d_B+1}^{d_A} \|v_i\|_2^2 \\ &= d_A - d_B,\end{aligned}$$

其中，第二个等式利用了对于 $i \geq d_A$ 有 $\mathbb{P}_A v_i = \mathbb{P}_B v_i = 0$，第三个等式利用了对于 $i \leq d_B$ 和 $\mathbb{P}_B v_i = 0$ 有 $\mathbb{P}_A v_i = \mathbb{P}_B v_i = v_i$，以及对于 $i \in \{d_B + 1, \ldots, d_A\}$ 有 $\mathbb{P}_A v_i = v_i$。 □

根据 断言 A.5，我们有 $\|\mathbb{P}_A - \mathbb{P}_B\|_{\mathrm{F}}^2 = |\text{秩}(A) - \text{秩}(B)|$。

□

B 第第2节

本节将给出第2节中省略结果的证明。

B.1 证明 事实2.1

Fact 2.1. 福 对于任意矩阵 $A \in \mathbb{R}^{d \times d}$，**方差**$_{X \sim \mathcal{N}(\mu, \Sigma)}[X^\top A X] \leq 4\|\Sigma^{1/2} A \Sigma^{1/2}\|_{\mathrm{F}}^2 + 8\|\Sigma^{1/2} A \mu\|_2^2$.

证明。 我们将 $X \sim \mathcal{N}(\mu, \Sigma)$ 写作 $X = \Sigma^{1/2} Z + \mu$，其中 $Z \sim \mathcal{N}(0, I)$。我们有 t

$$\begin{aligned}\operatorname*{\mathbf{Var}}_{X \sim \mathcal{N}(\mu, \Sigma)}[X^\top A X] &= \operatorname*{\mathbf{Var}}_{Z \sim \mathcal{N}(0, I)}[(\Sigma^{1/2} Z + \mu)^\top A (\Sigma^{1/2} Z + \mu)] \\ &= \operatorname*{\mathbf{Var}}_{Z \sim \mathcal{N}(0, I)}[Z^\top \Sigma^{1/2} A \Sigma^{1/2} Z + 2\mu^\top A \Sigma^{1/2} Z + \mu^\top A \mu] \\ &\leq 2\left(\operatorname*{\mathbf{Var}}_{Z \sim \mathcal{N}(0, I)}[Z^\top \Sigma^{1/2} A \Sigma^{1/2} Z] + 4 \operatorname*{\mathbf{Var}}_{Z \sim \mathcal{N}(0, I)}[\mu^\top A \Sigma^{1/2} Z]\right) \\ &\leq 4\|\Sigma^{1/2} A \Sigma^{1/2}\|_{\mathrm{F}}^2 + 8\|\Sigma^{1/2} A \mu\|_2^2 \,, \end{aligned} \tag{15}$$

其中，第一行不等式利用了 **方差**$[A + B] =$ **方差**$[A] +$ **方差**$[B] + 2$**协方差**$[A, B] \leq$ **方差**$[A] +$ **方差**$[B] + 2\sqrt{\text{**方差**}[A]\text{**方差**}[B]} = (\sqrt{\text{**方差**}[A]} + \sqrt{\text{**方差**}[B]})^2 \leq 2(\text{**方差**}[A] + \text{**方差**}[B])$这一点，而最后一个不等式

利用以下事实：第一项利用了如下事实：对于对称矩阵 B 和向量 v，**方差**$_{Z\sim\mathcal{N}(0,I)}[Z^\top BZ] = 2\|B\|_{\mathrm{F}}^2$和 **方差**$_{Z\sim\mathcal{N}(0,I)}[v^\top Z] = \|v\|_2^2$，其中 $B := \Sigma^{1/2}A\Sigma^{1/2}$和 $v := \Sigma^{1/2}A\mu$。

□

B.2 样本复杂度：证明 引理 2.6

本节中，我们将证明对于高斯分布，稳定性条件成立所需的样本复杂度。在此之前，我们指出，稳定性条件也应适用于更广泛的分布类别。

备注 B.1. 我们指出，定义 2.5 中的确定性条件并非高斯分布所特有；对于更广泛的分布类别，只要其四次多项式具有轻尾性质，粗略而言，这些条件都能以多项式样本复杂度 $\mathrm{poly}(d/\eta)$ 成立。

- 条件 (L.1) 对于二阶矩有界的分布，可在多项式样本复杂度下成立。
- 条件 (L.3) 在总体层面成立，直接源于二次多项式的超收缩性。当前条件中 $\log(1/\eta)$ 这一因子是针对高斯分布设定的，但对算法并非关键，可以放宽为 $1/\eta$的某个多项式；这样主定理的最终误差保证会增加一个 $\mathrm{poly}(1/\eta)$ 因子。在总体层面建立该条件后，可以用多项式样本复杂度将其推广到样本层面。
- 关于 条件（L.2），弗罗贝尼乌斯范数的每个元素都是一个二次多项式，因此，对于这些多项式，具有轻尾分布的分布满足该性质，例如，具有超收缩性的二次分布。
- 条件 (L.4) 也可以类似地通过 4 次多项式的超收缩性推导得出。
- 条件 (L.5) 对于连续分布几乎必然成立。

引理 2.6（确定性条件以高概率成立）。对于足够小的正数常数 ϵ_0 和足够大的绝对常数 C，从 $\mathcal{N}(\mu,\Sigma)$ 中抽取的 $m > Cd^2\log^5(d/(\eta\delta))/\eta^2$ 个样本，以概率 $1-\delta$ 构成相对于 μ,Σ 的(η,ϵ)-稳定集合，对所有 $\epsilon \le \epsilon_0$ 均成立。

证明。设 A 是来自 $\mathcal{N}(\mu,\ \Sigma)$的独立同分布样本集合。我们分别检查定义 2.5中的四个条件。

我们从条件（L.2）开始。令 $X \sim \mathcal{N}(\mu,\Sigma)$，则 $X-\mu \sim \mathcal{N}(0,\Sigma)$。等价地，$X-\mu = \Sigma^{1/2}Z$满足 $Z\sim\mathcal{N}(0,I)$。令 A为来自 $\mathcal{N}(\mu,\Sigma)$ 的 m个独立同分布样本，并且与之前一样，对于每个 $x\in A$，记 $x-\mu=\Sigma^{1/2}z$，其中 z是来自 $\mathcal{N}(0,I)$ 的相应独立同分布样本。令子集 $A'\subset A$ 满足 $|A'|\ge|A|(1-\epsilon)$。于是，我们有如下

$$\begin{aligned}\mathrm{tr}\left(\left(\frac{1}{n}\sum_{x\in A'}(x-\mu)(x-\mu)^\top-\Sigma\right)U\right) &= \mathrm{tr}\left(\Sigma^{1/2}\left(\frac{1}{n}\sum_z zz^\top - I\right)\Sigma^{1/2}U\right)\\ &= \mathrm{tr}\left(\left(\frac{1}{n}\sum_z zz^\top - I\right)\Sigma^{1/2}U\Sigma^{1/2}\right)\\ &\le C\epsilon\log(1/\epsilon)\|\Sigma^{1/2}U\Sigma^{1/2}\|_{\mathrm{F}}\,,\end{aligned}$$

其中第一行使用了我们之前引入的重写，第二行使用了迹算子的循环性质，最后一行使用了 [Li18，推论 2.1。13]。如该推论所述，只要样本复杂度是 $1-\delta/4$ 的足够大倍数，该不等式就以概率 $1-\delta/4$ 成立 $1-\delta/4$，其中 $(d^2/\epsilon^2)\log(4/\delta)$。固定 $\epsilon=\epsilon_0$ 一个足够小的正绝对常数，可以使上述不等式右侧 $0.1\|\Sigma^{1/2}U\Sigma^{1/2}\|_{\mathrm{F}}$。

条件 (L.1) 可以使用 [Li18，推论 2.1，以相同的方式证明。9]。

我们现在将注意力转向 条件 (L.3)。 [DKK+16，引理 5。17] 表明，使用所述数量的样本，以高概率有以下结论成立

$$\Pr_{X\sim A}\left[|p(x)|>\sqrt{\operatorname*{Var}_{Y\sim\mathcal{N}(\mu,\Sigma)}[p(Y)]}10\ln\left(\frac{2}{\eta}\right)\right]\le\frac{\eta}{2}.$$

由于 $|A'|\ge|A|(1-\epsilon_0)\ge|A|/2$，条件 (L.3) 成立。最后，注意到 $\mathrm{Var}_{X\sim A'}[p(X)]\le 2\mathrm{Var}_{X\sim A}[p(X)]$，并再次结合 [DKK+16、引理 5.17] 以及 $\epsilon=\epsilon_0$，即可推出条件 (L.4)。为了使该引理的结论以概率 $1-\delta/4$ 成立，高斯分量的样本数必须是 $d^2\log^5(4d/(\eta\delta))/\eta^2$ 的一个足够大的倍数。对对应于四个条件的事件应用并集界，即可完成证明。

最后，我们证明条件 (L.5)。实际上，我们将证明一个更强的结论：假设Σ的秩为 d_Σ。那么，如果我们从 $\mathcal{N}(\mu,\Sigma)$中取 $m\ge d_\Sigma+1$独立同分布样本 $\{\mu+z_1,\mu+z_2,\dots,\mu+z_m\}$，则必然以概率 1，对所有满足 $A'\subseteq A$ 的 $|A'|\ge d_\Sigma+1$，都有 $\sum_{x\in A'}(v^\top x)^2=0$ (等价地， v 与所有向量正交 $x\in A'$)，从而 $v^\top\Sigma v=0$ 对所有向量 $v\in\mathbb{R}^d$均成立。

我们将用到以下简单的断言：若 $A'=\{\mu+z'_0,\dots,\mu+z'_{d_\Sigma}\}$（大小为 $d_\Sigma+1$）满足：1）$\{z'_1-z'_0,\dots,z'_{d_\Sigma}-z'_0\}$线性无关；2） z'_i均位于 Σ的列空间内；那么对于任意向量 $v\in\mathbb{R}^d$，若 v与 A' 中所有向量正交，则 $v^\top\Sigma v=0$。为理解这一点，注意到 $\{z'_1-z'_0,\dots,z'_{d_\Sigma}-z'_0\}$是线性无关且均位于 Σ的列空间内的向量。根据维数计数，它们必然张成 Σ 的整个列空间。因此，若 v与 A'中所有点正交，则 v与 $\{z'_1-z'_0,\dots,z'_{d_\Sigma}-z'_0\}$中所有向量正交，这意味着 $v^\top\Sigma v=0$。

现在我们可以针对 A的大小证明以下归纳命题；结合上述断言，这将推出强化版的 条件 (L.5)。具体而言，我们通过归纳证明：对于 A的抽样，以概率 1 对任意大小至多为 $d_\Sigma+1$的子集 $A'\subset A$，断言条件对 A'均成立。也就是说，如果 $A'=\{\mu+z'_0,\dots,\mu+z'_j\}$ 对于某个 $j\le d_\Sigma$ 成立，并且 A'中的元素按任意顺序排列，那么 1) 集合 $\{z'_1-z'_0,\dots,z'_j-z'_0\}$线性无关，且 2)$z'_i$均属于 Σ的列空间。

$|A|=1$ 的基本情形显然成立。现在假设上述命题对 $|A_\ell|=\ell$成立，并考虑采样一个新点 $\mu+z_\ell$，将其加入 A_ℓ以形成 $A_{\ell+1}$。任取 $A_{\ell+1}$的一个子集 A'，其大小至多为 $d_\Sigma+1$。如果 $A'\subset A_\ell$，则根据归纳假设，我们已经完成证明。否则，$A'=\{\mu+z_\ell\}\cup A'_\ell$，其中 $A'_\ell=\{\mu+z'_0,\dots,\mu+z'_j\}$的大小至多为 d_Σ，且根据归纳假设满足归纳中的条件。满足条件 $z-z'_0$且使 $\mu+z$落在 $\{z'_1-z'_0,\dots,z'_j-z'_0\}$张成的空间中的点，在 $\mathcal{N}(\mu,\Sigma)$下构成零测度集合；这是因为 $\{z'_1-z'_0,\dots,z'_j-z'_0\}$中线性无关向量的数量严格少于 d_Σ，并且这些向量全都位于 Σ的列空间中。此外，这样的 A'_ℓ只有有限多个。因此，以概率 1，新点 $\mu+z_\ell$会被采样为：对所有这样的 A'_ℓ，$\{\mu+z_\ell\}\cup A'_\ell$仍然线性无关。此外，以概率 1， z_ℓ也会位于 Σ的列空间中。至此，归纳命题得证；正如我们所示，这意味着条件 (L.5) 成立。

□

B.3 关于稳定性的事实：以下内容的证明：推论 2.7 和引理 2.8

本节给出推论 2.7 和引理 2.8 的证明。

推论 2.7 （变换集合的稳定性）。设 D 是定义在 $\mathbb{R}^d$ 上的一个分布，其均值为 μ，协方差为 Σ。设 A 是 $\mathbb{R}^d$ 中的一个点集，满足定义 2.5中的条件 *(L.1)* 和 *(L.2)*，关于（，）η-ϵ稳定性，相对于。那么，对于任意对称变换 $P \in \mathbb{R}^{d\times d}$以及每个满足 $A' \subset A$ 的 $|A'| \geq (1-\epsilon)|A|$，下述结论成立：

(i) $\left\| P\left(\frac{1}{|A'|}\sum_{x\in A'} x - \mu\right)\right\|_2 \leq 0.1\sqrt{\|P\Sigma P\|_{\mathrm{op}}}$.

(ii) $\left\| P\left(\frac{1}{|A'|}\sum_{x\in A'}(x-\mu)(x-\mu)^\top - \Sigma\right)P\right\|_{\mathrm{F}} \leq 0.1\|P\Sigma P\|_{\mathrm{op}}$.

证明。我们分别说明每个条件。首先说明第一个条件：对于任意单位向量 u，属于 $\mathbb{R}^d$，有

$$\left|u^\top P\left(\frac{1}{|A'|}\sum_{x\in A'} x - \mu\right)\right| \leq 0.1\sqrt{u^\top P\Sigma P u} \leq \sqrt{\|P\Sigma P\|_{\mathrm{op}}}\,,$$

其中，第一个不等式用到了定义 2.5中的条件 (L.1) 对 $v = Pu$。

现在我们关注第二个条件。为简洁起见，记 $\hat{\Sigma} := \frac{1}{|A'|}\sum_{x\in A'}(x-\mu)(x-\mu)^\top$。设 Q 是任意满足 $\|Q\|_{\mathrm{F}} \leq 1$的对称矩阵 $Q \in \mathbb{R}^{d\times d}$。利用迹的循环性质，可得

$$\mathrm{tr}\left(P\left(\hat{\Sigma}-\Sigma\right)PQ\right) = \mathrm{tr}\left(\left(\hat{\Sigma}-\Sigma\right)PQP\right) \leq 0.1\left\|\Sigma^{1/2}PQP\Sigma^{1/2}\right\|_{\mathrm{F}}\,,$$

其中最后一步应用了条件（L.2），即定义 2.5中的 $V = PQP$。最后，利用迹算子的循环性质，

$$\begin{aligned}\left\|\Sigma^{1/2}PQP\Sigma^{1/2}\right\|_{\mathrm{F}}^2 &= \mathrm{tr}\left(\Sigma^{1/2}PQP\Sigma PQP\Sigma^{1/2}\right) = \mathrm{tr}(QP\Sigma PQP\Sigma P)\\ &= \|QP\Sigma P\|_{\mathrm{F}}^2 \leq \|Q\|_{\mathrm{F}}^2\|P\Sigma P\|_{\mathrm{op}}^2 \leq \|P\Sigma P\|_{\mathrm{op}}^2,\end{aligned}$$

其中第一个不等式用到了事实 2.3，第二个不等式用到了 $\|Q\|_{\mathrm{F}} \leq 1$。因此，我们有如下结论：

$$\begin{aligned}\left\|P\left(\hat{\Sigma}-\Sigma\right)P\right\|_{\mathrm{F}} &= \sup_{Q:\|Q\|_{\mathrm{F}}\leq 1}\mathrm{tr}\left(P\left(\hat{\Sigma}-\Sigma\right)PQ\right)\\ &= \sup_{\text{symmetric } Q:\|Q\|_{\mathrm{F}}\leq 1}\mathrm{tr}\left(P\left(\hat{\Sigma}-\Sigma\right)PQ\right)\\ &\leq 0.1\|P\Sigma P\|_{\mathrm{op}}.\end{aligned}$$

其中第二个等式成立，是因为 $P\left(\hat{\Sigma}-\Sigma\right)P$ 本身是对称的。

□

引理 2.8。设 A 是 $\mathbb{R}^d$ 中的一个点集，满足相对于 μ、Σ的 条件 *(L.1)* 和 *(L.2)*，即 (η,ϵ)-稳定性（定义 2.5）。设 $A' \subseteq A$ 满足 $|A'| \geq (1-\epsilon)|A|$。若 P、$L \in \mathbb{R}^{d\times d}$ 是满足以下性质的对称矩阵：$\|$ **协方差**$_{X\sim A'}[PX] - L\|_{\mathrm{F}} \leq r$，则

$$\|P\Sigma P - L\|_{\mathrm{F}} = O(r + \|L\|_{\mathrm{op}})\,.$$

证明。为简化记号，设 $\widehat{\Sigma} := \mathbf{协方差}_{X\sim A'}[X]$，并设 $\overline{\Sigma} := \mathbf{期望}_{X\sim A'}[(X-\mu)(X-\mu)^\top]$。注意到 $\overline{\Sigma} = \widehat{\Sigma} + (\mu - \hat{\mu})(\mu - \hat{\mu})^\top$。我们应用三角不等式和 推论 2.7，得到如下结果：

$$\begin{aligned}\|P(\widehat{\Sigma}-\Sigma)P\|_{\mathrm{F}} &\le \|P(\overline{\Sigma}-\Sigma)P\|_{\mathrm{F}} + \|P(\mu-\hat{\mu})(\mu-\hat{\mu})^\top P\|_{\mathrm{F}} \\ &= \|P(\overline{\Sigma}-\Sigma)P\|_{\mathrm{F}} + \|P(\mu-\hat{\mu})\|_2^2 \\ &\le 0.1\|P\Sigma P\|_{\mathrm{op}} + 0.1\|P\Sigma P\|_{\mathrm{op}} \\ &\le 0.2\|P\Sigma P\|_{\mathrm{op}}\,, \end{aligned} \tag{16}$$

其中，第一个不等式由三角不等式得出，第二个不等式使用了推论 2.7。

现在，我们使用三角不等式和引理假设来分析所需的表达式，如下所示：

$$\|P\Sigma P - L\|_{\mathrm{F}} \le \|P\widehat{\Sigma}P - L\|_{\mathrm{F}} + \|P(\widehat{\Sigma}-\Sigma)P\|_{\mathrm{F}} \le r + 0.2\|P\Sigma P\|_{\mathrm{op}},$$

其中，第一项根据假设由 r 界定，第二项使用了界（16）。因此，只需证明 $\|P\Sigma P\|_{\mathrm{op}} = O(r + \|L\|_{\mathrm{op}})$。为此，我们再次使用三角不等式，如下所示：

$$\begin{aligned}\|P\Sigma P\|_{\mathrm{op}} &\le \|P\widehat{\Sigma}P - L\|_{\mathrm{op}} + \|L\|_{\mathrm{op}} + \|P(\widehat{\Sigma}-\Sigma)P\|_{\mathrm{op}} \\ &\le r + \|L\|_{\mathrm{op}} + 0.2\|P\Sigma P\|_{\mathrm{op}}, \end{aligned} \tag{17}$$

其中，第二个不等式使用了界（16）。重新整理上面的展示不等式，可得 $\|P\Sigma P\|_{\mathrm{op}} = O(r + \|L\|_{\mathrm{op}})$。 □

C 归一化引理 3.2 的证明

本节证明引理 3.2，该引理指出，使用 $H^{\dagger/2}$ 进行归一化后，内点的经验均值和协方差以及总体均值和协方差都有界。这一点符合直觉，因为内点占总体样本的$\Omega(\alpha)$，且归一化后，完整集合的二阶矩被单位矩阵所界定。

我们重述假设 3.1，该假设来自正文，其中包含引理3.2所需的相关记号和假设。

假设 3.1。（算法的一次递归调用所用的假设和记号）。

- $S = \{x_i\}_{i=1}^n$ 是由 n 个未污染样本组成的集合（在对手进行修改之前）；假设该集合相对于具有均值和协方差$(\eta,\ 2\epsilon_0)$的内点分布 D 是稳定的 μ，Σ（参见定义 2.5）。我们假设 $\eta \le 0.001$，$\epsilon_0 = 0.01$，且对于所有对称的$\mathbf{方差}_{X\sim D}[X^\top AX] \le C_1(\|\Sigma^{1/2}A\Sigma^{1/2}\|_F^2 + \|\Sigma^{1/2}A\mu\|_2^2)$ 矩阵 $d\times d$ A 以及一个常数 C_1。

- T 是算法当前递归调用的输入集（经过对抗性破坏后），满足 $|S\cap T| \ge (1-2\epsilon_0)|S|$ 和 $|T| \le (1/\alpha)|S|$。这两者意味着$|T\cap S| \ge \alpha|T|(1-2\epsilon_0)$ 以及 $|T| \ge |S|(1-2\epsilon_0)$。

- 我们将 $H = \mathbf{E}_{X\sim T}\left[XX^\top\right]$记为。

- 我们将 $\widetilde{S},\widetilde{T}$表示由 $H^{\dagger/2}$归一化后的 S和 T的版本，即 $\widetilde{S} = \{H^{\dagger/2}x :\ x\in S\}$ 以及 $\widetilde{T} = \{H^{\dagger/2}x :\ x\in T\}$。对于 $\widetilde{S}$和 $\widetilde{T}$中的元素，我们使用记号 $\widetilde{x}$；对于

$\widetilde{S}$ 和 T中的元素，我们使用记号 X。类似地，对于支撑集$\widetilde{\quad}$ $\widetilde{\quad}$位于 S（或 T，或集合的类似“波浪号”版本）中的随机变量，我们使用该记号。

- 经过 $H^{\dagger/2}$ 变换后的内点分布 D 的均值和协方差分别记为 $\widetilde{\mu} := H^{\dagger/2}\mu$、$\widetilde{\Sigma} := H^{\dagger/2}\Sigma H^{\dagger/2}$。我们将 $\widehat{\widetilde{\ }}$ 中变换后的内点在 T 的经验均值和协方差分别记为 $\widehat{\mu} := \mathbf{E}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[X]$ 和 $\Sigma := \mathbf{Cov}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[X]$。

$\widetilde{\quad}$ $\widehat{\quad}$**引理 3.2** (归一化)。假设 假设 *3.1* 并回顾以下记号：$\widetilde{\mu}, \Sigma$、$\widehat{\mu}, \Sigma$，它们定义于 假设 *3.1*。我们有：

(i) $\|\Sigma\|_{\mathrm{op}} \le 2/\alpha$。*(ii)* $\|\widehat{\mu}\|_2 \le \sqrt{2/\alpha}$。*(iii)* $\|\Sigma\|_{\mathrm{op}} \le 3/\alpha$ *(iv)* $\|\widetilde{\mu}\|_2 \le \sqrt{3/\alpha}$ *(v)* $\|\mathbf{E}_{\widetilde{X}\sim\widetilde{S}}[X]\|_2 \le 2/\sqrt{\alpha}$ 和 $\|\mathbf{Cov}_{\widetilde{X}\sim\widetilde{S}}[X]\|_{\mathrm{op}} \le 4/\alpha$。

(vi) 对于每个满足 $\|A\|_{\mathrm{F}} \le 1$、$|\mathbf{E}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^\top A\widetilde{X}] - \mathbf{E}_{X\sim D}[(H^{\dagger/2}X)^\top A(H^{\dagger/2}X)]| < 1/\alpha$ 的 A 矩阵。

证明。我们将在下面证明各个部分：

建立 条目 (i) 变换 H 满足 $\mathbf{E}_{\widetilde{X}\sim\widetilde{T}}[\widetilde{X}\widetilde{X}^\top] = \mathbf{E}_{X\sim T}[H^{\dagger/2}XX^\top H^{\dagger/2}] = H^{\dagger/2}HH^{\dagger/2} = \mathbb{P}_H$，其算子范数至多为 1。根据假设，我们有 $\left|\widetilde{S}\cap\widetilde{T}\right| \ge (1-2\epsilon_0)\alpha|T| \ge 0.5\alpha|T|$。因此，我们得到以下不等式：

$$\frac{1}{\left|\widetilde{S}\cap\widetilde{T}\right|}\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}}\widetilde{x}\widetilde{x}^\top \preceq \frac{1}{\left|\widetilde{S}\cap\widetilde{T}\right|}\sum_{\widetilde{x}\in\widetilde{T}}\widetilde{x}\widetilde{x}^\top = \frac{|\widetilde{T}|}{\left|\widetilde{S}\cap\widetilde{T}\right|}\mathbb{P}_H \preceq \frac{2}{\alpha}\mathbb{P}_H\,. \tag{18}$$

通过应用（18），我们得到

$$\hat{\Sigma} = \frac{1}{\left|\widetilde{S}\cap\widetilde{T}\right|}\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}}(\widetilde{x}-\hat{\mu})(\widetilde{x}-\hat{\mu})^\top \preceq \frac{1}{\left|\widetilde{S}\cap\widetilde{T}\right|}\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}}\widetilde{x}\widetilde{x}^\top \preceq \frac{2}{\alpha}\mathbb{P}_H,$$

$\widehat{\quad}$ 这意味着 $\|\Sigma\|_{\mathrm{op}} \le 2/\alpha$，因为 $\mathbb{P}_H$ 只是一个具有 $0/1$ 个特征值的投影矩阵。

证明条目（ii）由于对于任意随机向量 X，我们有$\mathbf{E}[X]\mathbf{E}[X]^\top \preceq \mathbf{E}[XX^\top]$（这等价于$\mathbf{Cov}[X] \succeq 0$），我们得到

$$\hat{\mu}\hat{\mu}^\top \preceq \frac{1}{\left|\widetilde{S}\cap\widetilde{T}\right|}\sum_{\widetilde{x}\in\widetilde{S}\cap\widetilde{T}}\widetilde{x}\widetilde{x}^\top \preceq \frac{2}{\alpha}\mathbb{P}_H,$$

其中最后一个不等式用到了（18）。这意味着 $\|\widehat{\mu}\|_2^2 \le 2/\alpha$。

证明 命题 (iii) 目标是对总体协方差 $\|\Sigma\|_{\mathrm{op}}$进行界定。为此，我们将使用 命题 (i) 和 (ii) 中的界，这些界对其经验版本进行了界定，并通过确定性条件将经验版本与总体版本联系起来。

考虑任意子集 $\widetilde{S}_1$，其满足 $\widetilde{S}$条件 $\left|\widetilde{S}_1\right| \geq \left|\widetilde{S}\right|$ $(1-2\epsilon_0)$。首先注意到，根据推论 2.7，我们有

$$\left\|\widetilde{\Sigma} - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}_1}[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top]\right\|_{\mathrm{F}} \leq 0.1\left\|\widetilde{\Sigma}\right\|_{\mathrm{op}} \quad \text{and} \quad \left\|\widetilde{\mu} - \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}_1}[\widetilde{X}]\right\|_2 \leq 0.1\sqrt{\left\|\widetilde{\Sigma}\right\|_{\mathrm{op}}}. \tag{19}$$

现在定义 $\widetilde{\Sigma}_1 := \mathbf{E}_{\widetilde{X}\sim\widetilde{S}\cap\widetilde{T}}[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top]$，即 $\widetilde{S}\cap\widetilde{T}$ 的中心化二阶矩矩阵，其满足 $\widetilde{\Sigma}_1 = \hat{\Sigma} + (\widetilde{\mu}-\hat{\mu})(\widetilde{\mu}-\hat{\mu})^\top$。我们有

$$\begin{aligned}
\|\widetilde{\Sigma}\|_{\mathrm{op}} &\leq \|\widetilde{\Sigma}-\hat{\Sigma}\|_{\mathrm{F}} + \|\hat{\Sigma}\|_{\mathrm{op}} && \text{(triangle inequality)}\\
&\leq \|\widetilde{\Sigma}-\widetilde{\Sigma}_1\|_{\mathrm{F}} + \|\widetilde{\mu}-\hat{\mu}\|_2^2 + \|\hat{\Sigma}\|_{\mathrm{op}} && \text{(triangle inequality and } \widetilde{\Sigma}_1 = \hat{\Sigma} + (\mu-\hat{\mu})(\mu-\hat{\mu})^\top)\\
&\leq 0.2\|\widetilde{\Sigma}\|_{\mathrm{op}} + \|\hat{\Sigma}\|_{\mathrm{op}},
\end{aligned}$$

其中最后一行使用了（19）中的 $\widetilde{S}_1 = \widetilde{S}\cap\widetilde{T}$以及 $\left|\widetilde{S}\cap\widetilde{T}\right| \geq (1-2\epsilon_0)\,|\widetilde{S}|$这一事实。重新整理可得，$\|\widetilde{\Sigma}\|_{\mathrm{op}} < 1.25\|\hat{\Sigma}\|_{\mathrm{op}}$ 我们有 对此可以使用 条目（i）将其进一步上界为 $3/\alpha$。

证明 条目（iv） 我们使用与 条目（iii）类似的论证：

$$\|\widetilde{\mu}\|_2 \leq \|\hat{\mu}\|_2 + \|\widetilde{\mu}-\hat{\mu}\|_2 \leq \|\hat{\mu}\|_2 + 0.1\sqrt{\|\widetilde{\Sigma}\|_{\mathrm{op}}} \leq \sqrt{3/\alpha}, \tag{20}$$

其中第一步使用三角不等式，第二个不等式对 19）中的 $S_1 = S\cap T$使用了公式，而最后一个不等式使用了 条目（ii）和 （iii）。

建立 条目 (v) 对于协方差条件，我们有如下结论：

$$\begin{aligned}
\left\|\mathop{\mathbf{Cov}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]\right\|_{\mathrm{op}} &\leq \left\|\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}\left[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top\right]\right\|_{\mathrm{op}} && \text{(using } \mathbf{Cov}(\widetilde{X}) \preceq \mathbf{E}[\widetilde{X}\widetilde{X}^\top])\\
&\leq \left\|\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}\left[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top\right] - \widetilde{\Sigma}\right\|_{\mathrm{op}} + \|\widetilde{\Sigma}\|_{\mathrm{op}} && \text{(triangle inequality)}\\
&\leq 1.1\|\widetilde{\Sigma}\|_{\mathrm{op}},
\end{aligned}$$

其中，最后一步利用等式（19）对第一项取上界，适用于 $S_1 = S$，显然 满足 $|S| \geq |S|(1-2\epsilon_0)$。根据 $(1.1\times 3)/\alpha$，整个表达式的上界为 第 (iii) 项。

均值条件的证明也类似：

$$\left\|\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]\right\|_2 \leq \left\|\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\mu}\right\|_2 + \|\widetilde{\mu}\|_2 \leq 0.1\sqrt{\|\widetilde{\Sigma}\|_{\mathrm{op}}} + \sqrt{3/\alpha} \leq 2/\sqrt{\alpha},$$

其中我们使用公式（19）和（20）。

建立第（vi）项考虑任意方阵 A 其中 $\|A\|_{\mathrm{F}} \le 1$。我们有

$$
\begin{aligned}
&\left| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}^\top A\widetilde{X}] - \mathop{\mathbf{E}}_{X\sim D}[(H^{\dagger/2}X)^\top A(H^{\dagger/2}X)] \right| \\
&= \left| \left\langle A, \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}\widetilde{X}^\top] - \mathop{\mathbf{E}}_{X\sim D}[(H^{\dagger/2}X)(H^{\dagger/2}X)^\top] \right\rangle \right| \\
&= \left| \left\langle A, \left(\mathop{\mathbf{Cov}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] + \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]^\top \right) - \left(\widetilde{\Sigma} + \widetilde{\mu}\widetilde{\mu}^\top \right) \right\rangle \right| \\
&\le \left| \left\langle A, \mathop{\mathbf{Cov}}_{X\sim\widetilde{S}}[X] - \widetilde{\Sigma} \right\rangle \right| + \left| \left\langle A, \mathop{\mathbf{E}}_{X\sim\widetilde{S}}[X] \mathop{\mathbf{E}}_{X\sim\widetilde{S}}[X]^\top - \widetilde{\mu}\widetilde{\mu}^\top \right\rangle \right| && \text{(triangle inequality)} \\
&\le \left\| \mathop{\mathbf{Cov}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\Sigma} \right\|_{\mathrm{F}} + \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]^\top - \widetilde{\mu}\widetilde{\mu}^\top \right\|_{\mathrm{F}} && \text{(variational definition of Frobenius norm)} \\
&= \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top] - \widetilde{\Sigma} + \left(\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\mu} \right) \left(\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\mu} \right)^\top \right\|_{\mathrm{F}} \\
&\qquad + \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]^\top - \widetilde{\mu}\widetilde{\mu}^\top \right\|_{\mathrm{F}} \\
&\le \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[(\widetilde{X}-\widetilde{\mu})(\widetilde{X}-\widetilde{\mu})^\top] - \widetilde{\Sigma} \right\|_{\mathrm{F}} + \left\| \left(\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\mu} \right) \left(\mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] - \widetilde{\mu} \right)^\top \right\|_{\mathrm{F}} \\
&\qquad + \left\| \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}] \mathop{\mathbf{E}}_{\widetilde{X}\sim\widetilde{S}}[\widetilde{X}]^\top - \widetilde{\mu}\widetilde{\mu}^\top \right\|_{\mathrm{F}} \\
&\overset{(a)}{\le} \left\| \mathop{\mathbf{E}}_{X\sim\widetilde{S}}[(X-\widetilde{\mu})(X-\widetilde{\mu})^\top] - \widetilde{\Sigma} \right\|_{\mathrm{F}} + 3 \left\| \mathop{\mathbf{E}}_{X\sim\widetilde{S}}[X] - \widetilde{\mu} \right\|_2^2 \\
&\le 0.1\|\widetilde{\Sigma}\|_{\mathrm{op}} + 0.03\|\widetilde{\Sigma}\|_{\mathrm{op}} && \text{(using (19))} \\
&< 1/\alpha\,, && \text{(using Item (iii))}
\end{aligned}
$$

其中，不等式 (a) 用到了这样一个事实：$\|uv^\top - wz^\top\|_{\mathrm{F}}^2 \le \|u-w\|_2^2 + \|v-z\|_2^2$ 适用于最后一项。

□