arXiv:2303.06256v1 [cs.CC] 11 Mar 2023

量子异常值定理

Samuel Epstein
samepst@jptheorygroup.org

2023年3月14日

摘要

近期的研究结果已经证明，所有采样方法都会产生异常值。本文将这些结果推广到量子信息论。高秩投影算子的像中必定包含离群的纯量子态。否则，这些投影算子就是奇异的，因为它们与停机序列具有高度互信息。因此，使用投影的量子编码方案（例如 Schumacher 压缩）必须通过离群量子态进行通信。

1 引言

在算法信息论中，异常值的概念通过随机性缺陷来建模。该模型由自然数上的概率 p 定义，而数据点 x 是一个自然数。随机性缺陷的形式化定义为

$$\mathbf{d}(x|p) = \log \mathbf{t}_p(x),$$

其中 $\mathbf{t}_p$ 是一个通用下半可计算的 p-检验。它衡量数据点相对于某个模型的非典型程度。在 [GÓ1]中，引入了量子随机性缺陷的概念。量子随机性缺陷衡量纯态或混合量子态 ρ 相对于另一个量子混合态 σ的算法非典型性。其定义为

$$\mathbf{d}(\rho|\sigma) = \log \mathrm{Tr}\rho\mathbf{t}_\sigma,$$

其中 $\mathbf{t}_\sigma$ 是一个通用下半可计算的厄米矩阵，满足 $\mathrm{Tr}\sigma\mathbf{t}_\sigma \leq 1$。假设密度矩阵 σ 是可计算的。混合态用于对纯态 $\{p_i\}$ 的随机混合态 $\{|\psi_i\rangle\}$进行建模，因此量子随机性缺陷衡量量子态相对于某种混合态的非典型程度。关于量子随机性缺陷定义的更多动机，请参见第 2 节。在 [Eps19], 中，量子随机性缺陷被扩展到不可计算量子态。

量子随机性缺陷有哪些有趣的性质？在 [Eps19], 证明了量子随机性缺陷在偏迹和酉操作下的守恒。经过一些努力，可以证明其在量子操作下也守恒。关于经典随机性缺陷，最近的一项结果表明，采样方法会产生异常值 [Eps21]。这一结果有几种证明方法，其中一种源于这样一个事实：随机性缺陷较低的自然数大集合具有奇异性，因为它们与停机序列有很高的互信息。

在本文中，我们证明了这一结果的量子版本。秩较大的投影，其像中必定包含离群的纯量子态。否则，这些投影就是奇异的，

也就是说，它们与停机序列具有很高的互信息。因此，使用投影的量子编码方案（如 Schumacher压缩）必须通过离群量子态进行通信。经典定理和量子定理彼此类似，但它们的证明截然不同！

2 约定

对于正实函数 f、$<^+ f$、$>^+ f$ 和 $=^+ f$，分别用来表示 $< f + O(1)$、$> f + O(1)$ 和 $= f \pm O(1)$。对于非负实函数 f，术语 $<^{\log} f$、$>^{\log} f$ 和 $=^{\log} f$ 分别表示术语 $<f+O(\log(f+1))$、$>f-O(\log(f+1))$ 和 $=f\pm O(\log(f+1))$。

令 $\mathrm{K}(x)$ 表示前缀自由柯尔莫哥洛夫复杂度。我们用 $\mathrm{I}(x;\mathcal{H}) = \mathrm{K}(x) - \mathrm{K}(x|\mathcal{H})$ 表示停机序列 $\mathcal{H} \in \{0,1\}^\infty$ 关于 $x \in \{0,1\}^*$所拥有的信息量。如果概率的支撑集有限且其所有取值均为有理数，则称该概率为初等概率。字符串 x 相对于初等概率度量 Q 的随机性亏损为 $\mathrm{d}(x|Q) = \lfloor -\log Q(x) \rfloor - \mathrm{K}(x|\langle Q\rangle)$。字符串的随机性为 $\mathrm{Ks}(x) = \min_Q\{\mathrm{K}(Q) + 3\log \max\{\mathrm{d}(x|Q), 1\}\}$。

引理 1 ([Eps21, Lev16]) $\mathbf{Ks}(x) <^{\log} \mathbf{I}(x;\mathcal{H})$。

我们用 $\mathcal{H}_n$ 表示一个维数为 n、由基 $|\beta_1\rangle, \ldots, |\beta_n\rangle$ 张成的希尔伯特空间。量子比特是希尔伯特空间 $\mathcal{H}_2$ 中的一个单位向量，该空间由向量 $|0\rangle, |1\rangle$ 张成。为了描述 n 个量子比特，我们使用 $\mathcal{H}_{2^n}$ 中的一个单位向量，该空间由基向量 $|x\rangle$ 张成，其中 x 是一个长度为 n 的字符串。

长度为 n 的纯量子态 $|\psi\rangle$ 是 $\mathcal{H}_{2^n}$ 中的一个单位向量。其对应的对偶空间元素记作 $\langle\phi|$。矩阵 A 的共轭转置为 A^*。Tr 用于表示矩阵的迹。投影矩阵是特征值位于 $\{0,1\}$ 中的厄米矩阵。

复矩阵 A 若其元素为具有有理系数的复数、可编码为 $\langle A\rangle$，且其柯尔莫哥洛夫复杂度为 $\mathrm{K}(A)$，则称为初等矩阵。

对于厄米矩阵，$\sigma \le \rho$ 当且仅当 $\rho - \sigma$ 是厄米矩阵。我们称程序 $q \in \{0,1\}^*$ 下计算厄米矩阵 σ，如果将其作为输入交给通用图灵机 U，该机器 U 读取 $\le \|q\|$位，并输出（无论是否停机）一个初等半密度矩阵序列 $\{\sigma_i\}$，使得 $\sigma_i \le \sigma_{i+1}$ 且 $\lim_{i\to\infty}\sigma_i = \sigma$。如果存在一个程序能够下计算矩阵 T，则称该矩阵是下可计算的。其复杂度为 $\mathrm{K}(T) = \min\{\mathrm{K}(q) : q \text{ 下计算 } T\}$。给定一个密度矩阵$\sigma$，$\sigma$-检验是一个下可计算的厄米矩阵 T，满足 $\mathrm{Tr}T\sigma = 1$。如果 σ 是可计算的，则存在一个通用的 σ 检验 t_σ，它相对于量子比特数 n下可计算，满足 $\mathrm{Tr}\sigma\mathrm{t}_\sigma \le 1$，且对于每个下可计算的 σ检验 T，都有 $O(1)\mathrm{t}_\sigma > 2^{-\mathbf{K}(T|\sigma)}T$。这个通用检验可以按照标准方式计算，与经典情形类似（参见 [Ǵ21]）。

定义 1（量子　　随机性缺陷）对于混合态 σ 和 ρ，$\mathrm{d}(\rho|\sigma) = \log \mathrm{Tr}\ \mathrm{t}_\sigma\rho$.

量子随机性亏损，除其他解释外，是衡量某个纯态相对于算法生成的量子源有多典型的分数。事实上，假设存在一个可计算概率 P，定义在基本正交纯态 $\{\langle|\psi_i\rangle\rangle\}$（正交纯态 $\{|\psi_i\rangle\}$）的编码上，其对应的密度矩阵为 $\sigma = \sum_i P(\langle|\psi_i\rangle\rangle)\ |\psi_i\rangle\langle\psi_i|$。那么就存在一个下半可计算的 σ-测试 $T = \sum_i 2^{\mathbf{d}(\langle|\psi_i\rangle\rangle|P)}|\psi_i\rangle\langle\psi_i|$，满足 $O(1)\mathrm{t}_\sigma > T$。因此 $\mathrm{d}(|\psi_i\rangle|\sigma) >^+ \mathrm{d}(\langle|\psi_i\rangle\rangle|P)$，从而为那些 $|\psi_i\rangle$ 不符合该量子源典型特征的纯态赋予较高分数。一般而言，任意 $|\phi\rangle$ 的 $\mathrm{d}(|\phi\rangle|\sigma)$ 分数将大于若干 $\mathrm{d}(\cdot|P)$ 分数的组合，其中 $\mathrm{d}(|\phi\rangle|\sigma) >^+ \log \sum 2^{\mathbf{d}(\langle|\psi_i\rangle\rangle|P)}|\langle\phi|\psi_i\rangle|^2$。

3 结果

定理 1 相对于一个 n 量子比特混合态 σ，对于基本 2^m 秩投影算子 P，$3m-2n <^{\log} \max_{|\phi\rangle\in \mathrm{Image}(P)} \mathrm{d}(|\phi\rangle\,|\sigma) + \mathrm{I}(\langle P\rangle; \mathcal{H})$。

证明。我们将通用图灵机相对化到 $\langle\sigma\rangle$ 和 $(3m-2n)$。因此，它实际上相对化到了 m、n和 σ。令初等概率测度 Q 和 $d\in\mathbb{N}$ 实现 $\mathrm{Ks}(P)$，其中$d=\max\{\mathrm{d}(P|Q)\ 1\}$。不失一般性，我们可以假设 Q 的支撑集是秩为 2^m的初等投影算子。共有 $d2^{n-m+2}$ 轮次。每一轮中，我们选择一个维数为 1 的 σ-测试 T，$\mathrm{Tr}\sigma T\le 1$，且对于某个 Q-概率的投影 B，$\mathrm{Tr}TB$ 很大。现在我们来描述这一选择过程。

选择一个随机测试 T，使其为 $2^{m-2}\,|\psi\rangle\,\langle\psi|$，其中 $|\psi\rangle$ 是从单位球面上均匀选取的 n 量子比特态，其分布为 Λ。

$$\mathbf{E}[\mathrm{Tr}T\sigma]=2^{m-2}\int \mathrm{Tr}\,\langle\psi|\,\sigma\,|\psi\rangle\,d\Lambda=2^{m-2}\mathrm{Tr}\sigma\int|\psi\rangle\,\langle\psi|\,d\Lambda=2^{m-n-2}\mathrm{Tr}\sigma=2^{m-n-2}.$$

因此，T 是一个 σ-测试的概率为 $\ge 1-2^{m-n-2}$。令 I_m 为一个 n量子比特单位矩阵，只有前 2^m 个对角元素非零。令 $K_m=I-I_m$。令 $p=2^{m-n}$ 和 $\hat{}T=T/2^{m-2}$。对于任意秩为 2^m的投影 B，

$$\begin{aligned}
&\Pr(\mathrm{Tr}B\hat{T}\le .5p)\\
=&\Pr(\mathrm{Tr}I_m\hat{T}\le .5p)\\
=&\Pr(\mathrm{Tr}K_m\hat{T}\ge 1-.5p)\\
\mathbf{E}[\mathrm{Tr}K_m\hat{T}]=&\,1-p\\
\Pr(\mathrm{Tr}K_m\hat{T}\ge 1-.5p)\le&\,(1-p)/(1-.5p)\\
\Pr(\mathrm{Tr}B\hat{T}\ge .5p)=&\,1-\Pr(\mathrm{Tr}K_m\hat{T}\ge 1-.5p)\\
\ge&\,1-(1-p)/(1-.5p)\\
=&\,.5p/(1-.5p)\ge .5p\\
\Pr(\mathrm{Tr}BT\ge 2^{2m-n-3})\ge&\,.5p.
\end{aligned}$$

令 Ω 为所有形如 $2^{m-2}\,|\phi\rangle\,\langle\phi|$ 的矩阵所构成的空间。令 R 为 Ω 上的均匀分布。令 $[A,B]$ 在 $\mathrm{Tr}AB > 2^{2m-n-3}$ 时取 1，否则取 0。根据上述方程，对于所有 $A\in\mathrm{Support}(Q)$，$\int_\Omega[A,B]dR(B)\ge .5p$。因此 $\sum_A\int_\Omega[A,B]Q(A)dR(B)\ge .5p$。对于厄米矩阵 A，$\{A\}$ 在 $\mathrm{Tr}A\sigma\le 1$ 时取 1，否则取 0。因此 $\int_\Omega\{A\}dR(A)\ge(1-p2^{-2})$。令 $f=\max_T\{T\}\sum Q(A)[T,A]$。

So

$$
\begin{aligned}
.5p &\leq \sum_A \int_\Omega [A,B]Q(A)dR(B) \\
&= \sum_A \int_\Omega \{B\}Q[A,B](A)dR(B) + \sum_A \int_\Omega (1-\{B\})[A,B]Q(A)dR(B) \\
&\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) + \int_\Omega (1-\{B\})dR(B) \\
&\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) + p2^{-2} \\
p/4 &\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) = \int_\Omega \left(\{B\}\sum_A [A,B]Q(A)\right) dR(B) \leq \int_\Omega f dR(B) \\
p/4 &\leq f.
\end{aligned}
$$

因此，对于每一轮次 i，关于 f 的下界证明存在一个一维矩阵$T_i = 2^{m-2}\left|\psi\right\rangle\left\langle\psi\right|$，使得 $\mathbf{Tr}\ T_i\sigma \leq 1$ 且 $\sum_R\{Q(R) : \mathbf{Tr}T_iR \geq 2^{2m-n-3}\} \geq p/4 = 2^{m-n-2}$。选取这样一个 T_i，并将 Q概率限定在满足$[T_i, B] = 0$的那些投影 B 上，然后开始下一轮次。假设共有 $d2^{n-m+2}$ 轮次，则不存在满足 $[T_i, B] = 1$ 的 T_i 的那些投影 B 的测度为

$$\leq (1-p/4)^{d2^{n-m+2}} \leq e^{-d}.$$

因此，存在一个 T_i 使得 $[T_i, P] = 1$；否则，可以构造一个 Q 测试 t，将 e^d 赋给所有满足不存在 T_i 使得 $[T_i, B] = 1$ 的投影 B，其他情况则赋值为 0。于是 $t(P) = e^d$，因此

$$1.44d < \log t(P) <^+ \mathbf{d}(P|Q) <^+ d.$$

这是一个矛盾，因为不失一般性，可以假设 d 很大。令 $T_i = 2^{m-2}\left|\psi\right\rangle\left\langle\psi\right|$ 满足 $[T_i, P] = 1$。令 $\left|\phi\right\rangle = P\left|\psi\right\rangle / \sqrt{\left\langle\psi\right| P \left|\psi\right\rangle}$。因此 $\left\langle\phi\right| T_i \left|\phi\right\rangle \geq 2^{2m-n-3}$ 且 $\left|\phi\right\rangle$ 属于 P的像。因此由引理 1，

$$
\begin{aligned}
2m-n &<^+ \log \left\langle\phi\right| T_i \left|\phi\right\rangle \\
2m-n &<^+ \log \max_{\left|\phi\right\rangle \in \mathrm{Image}(P)} \left\langle\phi\right| T_i \left|\phi\right\rangle \\
2m-n &<^+ \max_{\left|\phi\right\rangle \in \mathrm{Image}(P)} \mathbf{d}(P|\sigma) + \mathbf{K}(T_i) \\
2m-n &<^+ \max_{\left|\phi\right\rangle \in \mathrm{Image}(P)} \mathbf{d}(P|\sigma) + (n-m) + \log d + \mathbf{K}(d) + \mathbf{K}(Q) \\
2m-n &<^+ \max_{\left|\phi\right\rangle \in \mathrm{Image}(P)} \mathbf{d}(P|\sigma) + (n-m) + \mathbf{Ks}(P) \\
3m-2n &<^{\log} \max_{\left|\phi\right\rangle \in \mathrm{Image}(P)} \mathbf{d}(P|\sigma) + \mathbf{I}(P;\mathcal{H}).
\end{aligned}
$$

请注意，由于等式左侧是 $(3m-2n)$ 且具有对数精度，因此可以据此将通用图灵机条件化为 $(3m-2n)$。
□

3.1 可计算投影

定理 1 是针对初等描述的投影而言的，并且可以推广到任意可计算投影。对于矩阵 M，令 $\|M\| = \max_{i,j} |M_{i,j}|$ 为最大范数。若程序 $p \in \{0,1\}^*$ 输出一列秩为 ℓ 的投影 $\{P_i\}_{i=1}^{\infty}$，且满足 $\|P - P_i\| \leq 2^{-i}$，则称其计算出了一个秩为 ℓ 的投影 P。对于可计算投影算子 P，$\mathrm{I}(P; \mathcal{H}) = \min\{\mathrm{K}(p) - \mathrm{K}(p|\mathcal{H}) : p$ 是一个计算 $P\}$的程序。

引理 2 ([Eps22a]) 对于部分可计算的 f，$\mathbf{I}(f(a); \mathcal{H}) <^{+} \mathbf{I}(a; \mathcal{H}) + \mathbf{K}(f)$.

推论 1 相对于 n 量子比特混合态 σ，对于可计算的 2^m 秩投影算子 P, $3m - 2n <^{\log} \max_{|\phi\rangle \in \mathrm{Image}(P)} \mathrm{d}(|\phi\rangle \langle\sigma|) + \mathrm{I}(\langle P\rangle; \mathcal{H})$。

证明。设 p 是一个计算 P的程序。存在一个定义简单的算法 A，给定 p和 σ时，输出 P_n，使得 $\max_{|\psi\rangle \in \mathrm{Image}(P)} \mathrm{d}(|\psi\rangle \langle\sigma|) =^{+} \max_{|\psi\rangle \in \mathrm{Image}(P_n)} \mathrm{d}(|\psi\rangle \langle\sigma|)$。因此，根据引理 2，可得 $\mathrm{I}(P_n; \mathcal{H}) <^{+} \mathrm{I}(P; \mathcal{H})$。推论由定理 1 得出。□

4 讨论

在先前的工作中 [Eps21，Eps22a，Eps22b，Eps23], 研究表明，异常值会在动力学过程中涌现，并出现在概率算法的输出中。这一结论是通过使用随机性缺陷函数对异常值进行建模而得出的。该函数已在字符串、无限序列或可计算度量空间中的点上定义。

本研究表明，异常值普遍存在于物理领域，具体而言，也存在于量子力学中。涉及高秩量子投影算子的量子源量子通信方案，必须与相对于这些量子源而言非典型的量子态进行通信。未来的研究可以探究其他物理领域中算法异常是否确实存在，例如量子场论或黑洞。

参考文献

[Eps19] S. Epstein. 算法不可克隆定理. IEEE 信息论汇刊, 65(9):5925–5930, 2019.[Eps21] Samuel Epstein. 所有采样方法都会产生异常值. IEEE 信息论汇刊, 67(11):7568–7578, 2021.[Eps22a] S. Epstein. 离群点定理再探. CoRR, abs/2203.08733, 2022.[Eps22b] S. Epstein. 离群点、动力学与独立性假设. CoRR, abs/2207.03955, 2022.[Eps23] S. Epstein. 一致检验与算法热力学熵, 2023. http://www.jptheorygroup.org/doc/Oscillation.pdf.[Gó1] P. G´acs. 量子算法熵. 《物理学杂志 A：数学与一般》, 34(35), 2001.[Gá1] Peter G´acs. 描述复杂度与随机性讲义. CoRR, abs/2105.04704, 2021.

[Lev16] L. A. Levin。《元素最低复杂度的奥卡姆界》。《纯粹与应用逻辑年刊》，167(10):897–900, 2016年。另见：S. Epstein 和 L.A. Levin，《集合具有简单成员》，arXiv 预印本 arXiv:1107.1458, 2011年。