

以非凸Riemannian梯度下降重建量子態

arXiv:2210.04717v1 (2022) | 兩頁中文導讀

研究問題與低秩假設

q 個量子位元的密度矩陣 ρ 大小為 $d \times d$ ，其中 $d=2^q$ ，必須Hermitian、半正定且跡為1。完整量子態重建希望由有限量測估計整個 ρ ；本文假設 $\text{rank}(\rho)=r$ 遠小於 d ，利用低秩結構降低計算負擔。這是實質假設，不是所有量子態的普遍性質。秩 r Hermitian矩陣的實參數數量為 $2dr-r^2$ ，跡1再少一個自由度，遠低於一般 d^2 尺度。本文採非凸Riemannian gradient descent (RGD)，並分析有限shot雜訊下的收斂與誤差。

取樣設定與量測次數是兩件事

隨機抽 m 個Pauli可觀測量 S_i ，定義 $A(X)_i = \sqrt{d/m} \text{Tr}(S_i X)$ ，伴隨 $A \dagger y = \sqrt{d/m} \sum y_i S_i$ 。這個尺度使平均意義下 $A \dagger A$ 接近恆等。每個設定實際量測 l 次，得到 ± 1 結果，其平均再乘 $\sqrt{d/m}$ 才是輸入 y_i 。總shots為 $m \cdot l$ ，不能只用 m 描述實驗成本。資料模型是 $y=A(\rho)+z$ ，噪音以矩陣spectral norm $\|A \dagger z\| \leq \lambda$ 控制。所有迭代重用同一批資料，不需要每步重新量測。

把量子態問題放寬成低秩最小平方

演算法最小化 $\frac{1}{2} \|y - A(X)\|_2^2$ ，限制 X 為Hermitian且 $\text{rank}(X) \leq r$ ，但放寬PSD與跡1。因此任意有限步輸出不應無條件稱為物理密度矩陣。初始化 $X_0 = H_r(A \dagger y)$ ，其中 H_r 保留最大 r 個奇異值，給最佳秩 r 近似。對目前 X 的行列支撐投影 P_U 、 P_V ，切空間投影為 $P_T(G) = P_U G + G P_V - P_U G P_V$ ；投影方向秩至多 $2r$ 。它保留與低秩曲面相容的局部方向，避免每步在全部矩陣自由度中搜尋。

每次更新的清楚流程

先算殘差方向 $G = A \dagger (y - A(X))$ ，再投影 $D = P_T(G)$ 。若 D 非零，取 $\alpha = \|D\|_F^2 / \|A(D)\|_2^2$ ，形成 $W = X + \alpha D$ ，最後 $X_{\text{new}} = H_r(W)$ 。 α 是截斷前沿直線的精確line search，不是沿截斷後曲線的精確最小化； $D=0$ 需停止或另處理，不能算 $0/0$ 。切向 W 至多秩 $2r$ ，因此截斷可透過小矩陣處理。此方法與FGD的 $X=UU \dagger$ 因子化不同，後者目標對 U 是四次式，線搜尋較複雜，但不能絕對說其數學上不可能精確求解。

一個手算例子與RIP意義

以兩量子位元 $\rho = \text{diag}(0.8, 0.2, 0, 0)$ 、手選目前 $X = \text{diag}(0.7, 0.3, 0, 0)$ 、 $r=2$ 為例，假設完整等距量測且無雜訊，便有 $D = \text{diag}(0.1, -0.1, 0, 0)$ 、 $\alpha=1$ ， $W=\rho$ 且 $H_2(W)=W$ 。這只是更新示範，完整量測下的規定初始化其實已是 ρ 。一般壓縮情形需要RIP： $(1-\delta_s) \|X\|_F^2 \leq \|A(X)\|_2^2 \leq (1+\delta_s) \|X\|_F^2$ 對秩 $\leq s$ 成立，表示量測不過度扭曲低秩幾何，而非對所有矩陣都等距。

以非凸Riemannian梯度下降重建量子態

arXiv:2210.04717v1 (2022) | 兩頁中文導讀

主要定理與不能省略的條件

設 $\kappa=\sigma_1/\sigma_r$ 為 ρ 的非零特徵值條件數。當 $\lambda\leq C_1\sigma_r/\sqrt{r}$ 、 $m\geq C_2\kappa^2r^2d\log^6d$ ，配合指定初始化，原文證明 $e_k=\|X_k-\rho\|_F\leq e_0\bar{\gamma}^k+2\sqrt{2r}\lambda/((1-\delta_3r)(1-\bar{\gamma}))$ ， $\bar{\gamma}<1$ 可取通用常數。第一項隨迭代下降，第二項是雜訊底限。收斂因子不依 κ ，不表示總成本不依 κ ：取樣充分條件有 κ^2 ，雜訊門檻也依 σ_r 。不能把 $\log(1/(\kappa\epsilon))$ 的表述解釋為條件數越差就無代價地越容易重建。

證明的骨架與誤差尺度

Eckart-Young與三角不等式先給 $e_{\text{new}}\leq 2\|\rho-W\|_F$ 。接著將誤差分成RIP造成的線性項、切空間曲率造成的 $2e_k^2/\sigma_r$ ，以及 $O(\sqrt{r}\lambda)$ 的加法雜訊項。初始化滿足 $e_0\leq 2\delta_2r\|\rho\|_F+2\sqrt{2r}\lambda$ ，足夠近時能控制二次項。原文一組充分條件為 $\lambda\leq\sigma_r/(20\sqrt{2r})$ 、 $\delta_3r\leq 1/(80\kappa\sqrt{r})$ ，得到 $\bar{\gamma}<0.45$ 。差矩陣的秩至多 $2r$ ，所以核範數誤差 $\leq\sqrt{2r}$ 倍Frobenius誤差，進而得到 $O(r\lambda)$ ；比較兩種誤差必須保留秩因子。

迭代次數與每步成本

無雜訊時要達絕對誤差 ϵ ，只需取 k 不小於 $\log(e_0/\epsilon)/\log(1/\bar{\gamma})$ 的非負整數上整數；有雜訊時只需把初始化項壓到雜訊尺度，繼續迭代不會提高資料本身精度。每步可用兩次 $d\times 2r$ 的QR與一次 $2r\times 2r$ 的SVD，但切空間投影仍有 $O(d^2r)$ 矩陣乘法。原文Remark5給 $O(d^2r+dr^2+r^3)$ ，另需計入 A 、 $A^\dagger$ 的應用、初始化及資料取得成本。不能只引用小矩陣SVD成本便宣稱全部每步為 $O(dr^2)$ 。

數值實驗與正確讀圖方式

原文用Qiskit與IBM量子模擬器產生6、8量子位元的 $|+\rangle$ 張量態與GHZ態資料，均為純態 $r=1$ 、 $\kappa=1$ 。四個設定的 m 依序為819、1638、13107、26214，各設定8192shots；最大案例總計214745088shots。比較MIFGD的固定步長0.01與五種動量設定，RGD在圖中較快收斂。橫軸是迭代步數而非秒數，縱軸是平方Frobenius誤差。ExactRGD使用精確期望值，與有限shot版本不同；這些純態實驗也未掃描不同 κ 來獨立展示條件數優勢。

物理性、來源與研究意義

秩限制由截斷維持，但PSD與跡1並未每步強制。若Hermitian秩 r 估計與 ρ 的算子範數距離小於 σ_r ，可由Weyl推得 r 個正特徵值；跡仍不必恰為1，必要時需另行說明後處理。本文價值在低秩計算與誤差保證的結合，範圍是足夠取樣、低雜訊與合適初始化。來源：Ming-Chien Hsu、En-Jui Kuo、Wei-Hsuan Yu、Jian-Feng Cai、Min-Hsiu Hsieh，Quantum state tomography via non-convex Riemannian gradient descent，arXiv:2210.04717v1 (2022)。本導讀未獨立重跑全部實驗，也不宣稱2022比較代表2026最新紀錄。