

量子谱回归算法

潘世杰 高 飞 万林春 秦素娟 温巧燕

¹(网络与交换技术国家重点实验室(北京邮电大学) 北京 100876)

²(密码科学技术国家重点实验室 北京 100878)

(panshijie@bupt.edu.cn)

Quantum Algorithm for Spectral Regression

Pan Shijie, Gao Fei, Wan Linchun, Qin Sujuan, and Wen Qiaoyan

¹(State Key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications), Beijing 100876)

²(State Key Laboratory of Cryptology, Beijing 100878)

Abstract Subspace learning is an important research direction in the field of machine learning. In order to reduce the complexity of subspace learning, Cai et al. proposed a spectral regression dimensionality reduction framework, and proposed efficient spectral regression for subspace learning that constructed their graph by incorporating the label information. In recent years, the development of quantum computing has made it possible to further reduce the complexity of subspace learning algorithms. Meng et al. pioneered the quantum algorithm for spectral regression (MYXZ algorithm). In this article, the limitations of the MYXZ algorithm are pointed out. We point out that MYXZ algorithm uses sparse Hamiltonian simulation technology to process the matrix generated by the weight matrix, but this matrix is a dense matrix in some cases. In response to this situation, we propose an improved quantum spectral regression algorithm. Our improved algorithm uses quantum singular value estimation technology, which has a polynomial acceleration compared with MYXZ algorithm when dealing with dense matrices. In addition, we propose a new quantum algorithm to accelerate the classical efficient spectral regression. The problems that our new algorithm can handle cannot be handled by MYXZ algorithm. Our algorithm utilizes quantum ridge regression and quantum matrix vector multiplication technology. Under the same parameter conditions, our algorithm has polynomial acceleration effect compared with the classical algorithm.

Key words quantum algorithm; quantum machine learning; spectral regression; subspace learning; dense matrix

摘 要 子空间学习是机器学习领域的重要研究方向.为了降低子空间学习的复杂度,Cai 等人提出了谱回归降维框架,并针对结合标签构造对应图的子空间学习提出了高效谱回归.近年来,量子计算的发展使进一步降低子空间学习算法的复杂度成为了可能.Meng 等人率先提出了量子谱回归算法(MYXZ 算法).MYXZ 算法用了稀疏哈密顿量模拟技术来处理由权重矩阵生成的矩阵,但这个矩阵在较多的情况

收稿日期:2021-04-06;修回日期:2021-05-25

基金项目:中央高校基本科研业务费专项资金(2019XD-A01);国家自然科学基金项目(61976024,61972048)

This work was supported by the Fundamental Research Funds for the Central Universities (2019XD-A01) and the National Natural Science Foundation of China (61976024, 61972048).

通信作者:高飞(gaof@bupt.edu.cn)

下是稠密矩阵.针对这种情况,指出了 MYXZ 算法的局限性,提出了一个改进的量子谱回归算法.改进算法采用了量子奇异值估计技术,在处理稠密矩阵时相对 MYXZ 算法有多项式加速.另外,提出了一个新的量子算法,对经典的高效谱回归进行加速.新算法能处理的这类问题是 MYXZ 算法无法处理的.新算法利用了量子岭回归和量子矩阵向量乘技术,在相同的参数条件下相对经典算法具有多项式加速效果.

关键词 量子算法;量子机器学习;谱回归;子空间学习;稠密矩阵

中图法分类号 O413

量子计算展示了其相对经典计算的强大优势^[1-2].在处理特定的问题时,如大数质因数分解^[3]、无结构数据搜索^[4]、稀疏线性方程组求解^[5]和低秩矩阵主成分分析^[6],采用量子计算的方式具有比经典计算方式低得多的复杂度.近年来,一些研究人员将量子强大的计算能力用在处理经典计算机上具有较高复杂度的机器学习问题,取得了一系列成果,包括量子回归算法^[7]、量子支持向量机^[8]、量子生成对抗网络^[9-10]、量子神经网络和量子数据挖掘^[11]等.

子空间学习是机器学习的重要部分.现实世界的的数据大多数是以高维的形式出现,而决定计算任务——例如分类和回归——结果的有效维度往往很低.由于直接处理高维数据不仅需要消耗大量计算资源,还会导致维数灾难(the curse of dimensionality)问题,因此子空间学习被提出,用来把数据集从高维降到低维,同时保持数据潜在的结构.人脸识别的结果表明,通过子空间学习进行降维,可以显著地提高算法的性能^[12].尽管子空间学习可以部分解决高维数据的学习问题,但子空间学习本身通常也需要很高的复杂度,因此如何降低子空间学习的复杂度是一个很有实际意义的问题.一个自然而然的想法是将量子计算应用于子空间学习,以达到更低的算法复杂度.

常用的子空间学习算法包括主成分分析(principal components analysis, PCA)^[13]、线性判别分析(linear discriminant analysis, LDA)^[14]、局部保持投影(locality preserving projections, LPP)^[12]和邻域保持嵌入(neighborhood preserving embedding, NPE)^[15]等.子空间学习算法如 LDA, LPP 和 NPE,通常需要进行稠密矩阵的特征分解,这往往需要消耗较高的计算资源.为了降低计算资源,2007 年 Cai 等人^[16]提出了谱回归降维框架,该框架包含了很多现有的子空间学习算法,如 LDA, LPP 和 NPE.在同一篇文章中,针对结合标签信息来构造对应图的子空间学习, Cai 等人提出了高效谱回归. Cai 等人提

出可利用权重矩阵的特殊结构,结合格拉姆-施密特正交化来得到稠密矩阵的特征向量,而不需要对稠密矩阵进行特征分解,从而把复杂度从维数的立方降低为维数的一次方.另外,因为引入了线性回归,所以谱回归中也可以很自然地加入正则化参数,使得算法可用于数据矩阵是病态矩阵的情况.

在量子领域, Lloyd 等人^[6]率先提出了量子主成分分析算法(quantum principal component analysis algorithm, QPCA),算法相对经典算法具有指数加速效果.需要指出的是,尽管该算法被称为 QPCA 算法,但实际上如果要维持量子算法的加速效果,该算法只能用来得到低秩(或者近似低秩)矩阵的较大的特征值,以及以量子态形式得到对应的特征向量;另外,如果矩阵有重特征值,那么我们无法通过量子主成分分析得到重特征值对应的一组特征向量,而只能得到特征向量的叠加态.随后, Yu 等人基于量子 PCA 算法,提出了量子 PCA 降维算法,当降维后数据的维数为训练数据维数的对数时,他们的量子算法相对经典算法实现了指数级的加速^[17]. Cong 等人实现了量子 LDA 算法,与经典算法相比有指数级的加速效果^[18]. Meng 等人基于谱回归降维框架^[16],首次提出了正则化子空间学习的量子谱回归算法(MYXZ 算法,以作者姓名首字母命名),并在降维数据的项数 m 和数据的特征数 n 满足 $m = O(\ln n)$ 时,算法具有多项式加速效果^[19].尽管如此, MYXZ 算法仍然存在有待改进的地方.首先, MYXZ 算法在算法的第一步,即求一般特征值问题时用了稀疏哈密顿量模拟,当处理的矩阵是稀疏矩阵时,算法具有良好的加速效果,但实际应用中存在很多矩阵不稀疏的例子,例如 LDA, 此时用稀疏哈密顿量模拟没有加速效果.其次,如果算法第一步的解中有相同的特征值,也会导致 MYXZ 算法无法得到相应的特征向量(只能得到特征向量的叠加态)而导致后续步骤出错.最后, MYXZ 算法的加速基于假设

$m=O(\ln n)$, 其中 m 为数据的项数, n 为数据的特征数. 因为谱回归的原始文献[16]的重点在于把算法复杂度中的因子 m^3 降至 m , 而对于 $m=O(\ln n)$ 的情况, 谱回归本身只有很小的加速效果, 因此我们认为 $m=O(n)$ 是更符合实际情况的假设.

本文首先对 MYXZ 算法进行了细致分析, 并指出了算法的局限性. 然后, 我们提出了一个改进算法, 该算法采用了量子奇异值估计技术 (quantum singular value estimation, QSVE)^[20], 而不是稀疏量子主成分分析来求解一般特征值问题, 在处理稠密矩阵特征值问题时, QSVE 具有更低的复杂度. 当 $m=O(n)$ 时, 我们的改进算法相对经典算法有多项式加速效果; 当算法的第 1 步涉及的矩阵是稠密矩阵时, 我们的改进算法相对 MYXZ 算法有多项式加速效果. 其次, 针对结合标签信息来构造对应图的子空间学习算法, 例如 LDA 和带标签的 LPP, NPE, 我们提出了相应的量子算法. 我们把文献[16]中提出的高效算法中复杂度较高的岭回归部分设计成量子算法, 并行得到多个岭回归结果; 而复杂度较低的部分 (即格拉姆施密特正交化) 仍然采用原算法实现. 这样做的好处在于既规避了稠密矩阵特征分解的问题, 也利用了量子的加速效果. 另外, 为了算法的完整性, 我们给出了对新输入数据进行降维的实现方式. 当 $m=O(n)$ 时, 我们的算法相对经典的高效谱回归算法具有多项式加速效果.

1 正则化子空间学习的谱回归

在本节中, 我们将对正则化子空间学习的谱回归进行回顾, 并介绍其复杂度.

1.1 谱回归

假设有 m 张人脸图像, 表示为 $\{\mathbf{x}_i\}_{i=1}^m \in \mathbb{R}^n$, 令 $\mathbf{X}=(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$. 我们要做的是对这些图像数据进行降维处理, 降维后数据矩阵表示为 $\mathbf{Y}=(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m)$. 为了对新输入数据进行降维, 以及对测试集的数据进行检验, 通常需要输出降维矩阵 \mathbf{A} , 对新输入数据 \mathbf{x} , 降维后数据记为 $\mathbf{y}_x=\mathbf{A}^\top \mathbf{x}$.

令 G 表示由 m 个顶点构成的图, 每个顶点对应一张人脸图像即 \mathbf{x}_i , \mathbf{W} 为一个 $m \times m$ 的权重矩阵, 其元素表示图 G 中顶点 i 和顶点 j 之间的权重. 谱回归是一个降维算法的框架, 通过 \mathbf{W} 矩阵的不同选取方式, 能得到不同的降维算法, 如线性判别分析、局部保持投影和邻域保持嵌入. 谱回归的提出是为了避免降维算法中复杂度较高的稠密矩阵特征分解.

谱回归主要包含 3 个步骤:

1) 求解最大特征向量问题

$$\mathbf{W}\mathbf{y}=\lambda\mathbf{D}\mathbf{y}, \quad (1)$$

其中, \mathbf{D} 为对角矩阵, $D_{ii}=\sum_{j=1}^m W_{ij}$. 得到最大的 $c-1$ 个特征值对应的特征向量 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(c-1)}$, c 为类别个数.

2) 求解 $c-1$ 个最小二乘问题

$$\mathbf{a}_i=\arg \min_{\mathbf{a}_i} \sum_{j=1}^m (\mathbf{a}_i^\top \mathbf{x}_j - y_j^{(i)})^2, \quad (2)$$

其中, $y_j^{(i)}$ 表示向量 $\mathbf{y}^{(i)}$ 的第 j 个元素, $i \in \{1, 2, \dots, c-1\}$.

3) 对新数据 \mathbf{x} 进行降维, 令 $\mathbf{A}=(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_c)$, 则降维后的坐标为

$$\mathbf{y}_x=\mathbf{A}^\top \mathbf{x}.$$

当样本数 m 小于特征数 n 时, 优化问题 (2) 是病态的, 一种常见的做法是采用岭回归的方式, 引入正则化参数 α , 即求解优化问题:

$$\mathbf{a}_i=\arg \min_{\mathbf{a}_i} \sum_{j=1}^m (\mathbf{a}_i^\top \mathbf{x}_j - y_j^{(i)})^2 + \alpha \mathbf{a}_i^2,$$

此时 \mathbf{a}_i 的解析解为

$$\mathbf{a}_i=(\mathbf{X}\mathbf{X}^\top + \alpha \mathbf{I})^{-1} \mathbf{X}\mathbf{y}^{(i)}.$$

引入正则化参数的谱回归被称为正则化子空间学习的谱回归.

在文献[16]中, 作者提出一种高效的谱回归求解方法. 对于结合标签信息来构造对应图的降维算法, 如 LDA 以及有监督的 LPP 和 NPE, 其最大特征向量问题 (即式 (1)) 的求解是平凡的. 不妨假设数据是按标签信息排列的, 则

$$\mathbf{W}=\begin{bmatrix} \mathbf{W}^{(1)} & 0 & \cdots & 0 \\ 0 & \mathbf{W}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{W}^{(c)} \end{bmatrix},$$

其中 $\mathbf{W}^{(t)}$ 是 $m_t \times m_t$ 的矩阵, m_t 为第 t 个类别的样本数, 此时可以直接给出式 (1) 的最大特征值 1 对应的 c 个特征向量, 即

$$\mathbf{y}^{(i)}=[\underbrace{0, \dots, 0}_{m_1}, \underbrace{1, \dots, 1}_{m_i}, \underbrace{0, \dots, 0}_{m_c}].$$

$$\int_{j=1}^{i-1} m_j \uparrow \quad m_i \uparrow \quad \int_{j=i+1}^c m_j \uparrow$$

另外, 显然全 1 向量 $\mathbf{1}=(1, 1, \dots, 1)^\top$ 也是式 (1) 中最大特征值 1 对应的特征向量. 因为对所有的 \mathbf{W} , $\mathbf{1}$ 都是式 (1) 中特征值 1 对应的特征向量, 因此 $\mathbf{1}$ 并不包含关于问题的任何信息. 为了去除和 $\mathbf{1}$ 相关的冗余信息, 文献[16]中提出的方法是取 $\mathbf{1}$ 为第 1 个正交向量, 其余 $c-1$ 个特征向量由 $\mathbf{y}^{(i)}$ 通过格拉姆-

施密特正交化得到.最后去掉 $\mathbf{1}$, 就得到了我们想要的 $c-1$ 个相互正交的特征向量, 也即式(1)的解.

1.2 谱回归的复杂度

根据文献[16]中的算法, 回归算法的第 1 步的复杂度主要来源于格拉姆-施密特正交化, 复杂度为 $mc^2 - \frac{1}{3}c^3$; 第 2 步做 $c-1$ 次岭回归的复杂度为 $s(c-1)(2mn+3m+5n)$, 其中 s 为迭代次数, 岭回归的收敛很快, 因此 s 通常取值较小; 第 3 步矩阵向量乘法的复杂度为 $O(cn)$. 因此总算法复杂度为 $O(c^2m + cmns)$.

2 正则化子空间学习的量子谱回归

本节将对正则化子空间学习的量子谱回归进行回顾, 并对算法进行分析, 给出算法复杂度.

2.1 量子谱回归

2019 年, Meng 等人提出了正则化子空间学习的量子谱回归算法^[19] (MYXZ 算法). 该算法没有利用文献[16]中的核心思想, 即对于结合标签信息来构造对应图的降维算法, 其最大特征向量问题的求解(式(1))是平凡的, 而是直接用量子主成分分析^[6]的方式来对矩阵进行特征分解.

MYXZ 算法主要包含 2 个步骤:

1) 用稀疏矩阵量子主成分分析和量子矩阵向量乘法来求式(1)的解 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(c-1)}$.

MYXZ 算法假设矩阵 $\bar{\mathbf{L}} = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{1/2} + \mathbf{I}_m$ 事先被存储在一个树状的量子存储器 (QRAM) 中, 其中 \mathbf{I}_m 为 $m \times m$ 的矩阵. 用量子算法 (即量子主成分分析) 求最大特征向量问题:

$$\bar{\mathbf{L}} \bar{\mathbf{y}} = (\lambda + 1) \bar{\mathbf{y}},$$

得到的 $c-1$ 个最大特征值对应的特征向量 (量子态形式) 表示为 $|\bar{\mathbf{y}}^{(1)}\rangle, |\bar{\mathbf{y}}^{(2)}\rangle, \dots, |\bar{\mathbf{y}}^{(c-1)}\rangle$, 则式(1)的解 $|\mathbf{y}^{(i)}\rangle = \mathbf{D}^{1/2} |\bar{\mathbf{y}}^{(i)}\rangle$ 可通过量子矩阵向量乘法得到^[6].

2) 用量子算法求

$$\mathbf{a}_i = (\mathbf{X} \mathbf{X}^T + \alpha \mathbf{I})^{-1} \mathbf{X} \mathbf{y}^{(i)}.$$

MYXZ 算法假设矩阵 \mathbf{X} 事先存储在一个树状的 QRAM 中, 基于量子稠密线性方程组求解技术^[16], 设计出量子稠密矩阵岭回归技术, 可直接用于求 \mathbf{a}_i .

MYXZ 算法没有考虑算法的第 3 步, 也即针对一个新输入的数据 \mathbf{x} , 输出其降维后坐标 $\mathbf{y}_x = \mathbf{A} \mathbf{x}$.

2.2 算法分析

本节将结合文献[19]中的假设, 对算法及其复杂度进行分析.

文献[19]认为矩阵 $\bar{\mathbf{L}}$ 是稀疏矩阵, 其稀疏度用 s_L 表示. 第 1 步的量子主成分分析中量子模拟部分采用的是稀疏哈密顿量模拟的方式. 稀疏哈密顿量模拟的复杂度为 $\tilde{O}[\text{lb}(m)ts_L^4]$ (\tilde{O} 表示一些增长慢的项被省略掉了)^[5]. 注意到, 矩阵 $\bar{\mathbf{L}}$ 通常不是低秩矩阵, 所以在量子主成分分析测量时, 测量一次得到最大的 c 个特征值中的一个的概率为 $\frac{c}{m}$ (需要指出的是, 文献[19]并没有考虑测量的概率, 导致算出来的复杂度比实际的复杂度低), 引入幅度放大, 量子主成分分析的复杂度为 $\tilde{O}\left[\frac{\sqrt{cm}s_L^4}{\epsilon} \text{poly}(\text{lb}(m))\right]$ (我们用 $\text{poly}(\text{lb}(m))$ 来表示 $\text{poly}(\text{lb } m)$, 以简化符号. $\text{poly}(\text{lb}(m))$ 来源于对量子态 $|\bar{\mathbf{L}}\rangle = \frac{1}{\|\bar{\mathbf{L}}\|_F} \sum_{ij} \bar{L}_{ij} |i\rangle |j\rangle$ 的制备), 算法的输出为 $|\bar{\mathbf{y}}^{(1)}\rangle, |\bar{\mathbf{y}}^{(2)}\rangle, \dots, |\bar{\mathbf{y}}^{(c-1)}\rangle$.

文献[19]没有明确给出 $\mathbf{D}^{1/2} |\bar{\mathbf{y}}^{(i)}\rangle$ 的量子算法及其复杂度, 尽管看起来这部分算法很直接, 且复杂度低于量子主成分分析部分, 为了完整性, 我们给出具体的算法并分析其复杂度. 类似于矩阵 $\bar{\mathbf{L}}$ 的存储假设, 我们假设 \mathbf{D} 的元素是存储在 QRAM 中的. 因为 \mathbf{D} 是对角矩阵, 且要得到 $\bar{\mathbf{L}}$, 需要先得到 \mathbf{D} , 因此这个假设比矩阵 $\bar{\mathbf{L}}$ 的假设弱. 此时输入 $|\bar{\mathbf{y}}^{(i)}\rangle = \sum_{j=1}^m \bar{y}_j^{(i)} |j\rangle$, 直接调用 QRAM 查询 (复杂度为 $O(\text{lb } m)$) 可得量子态 $\sum_{j=1}^m \bar{y}_j^{(i)} |j\rangle |D_{jj}\rangle$. 通过受控旋转^[5] (复杂度可忽略) 和测量 (平均测量次数为 $O(\kappa_D)$, 其中 κ_D 为矩阵 \mathbf{D} 的条件数, 即最大特征值和最小特征值的比值. 通过幅度放大, 调用 QRAM 查询的次数降低为 $O(\kappa_D^{1/2})$), 可得量子态 $\sum_{j=1}^m D_{jj}^{1/2} \bar{y}_j^{(i)} |j\rangle |D_{jj}\rangle$, 再次调用 QRAM 对第 2 量子寄存器进行逆运算, 可得 $\mathbf{D}^{1/2} |\bar{\mathbf{y}}^{(i)}\rangle = \sum_{j=1}^m D_{jj}^{1/2} \bar{y}_j^{(i)} |j\rangle$. 因此矩阵向量乘法的复杂度为 $O(\kappa_D^{1/2} \text{lb } m)$. 在实际问题中, κ_D 通常取值较小, 因此这个复杂度相对于量子主成分分析部分是可以忽略的.

综上, 算法的第 1 步总复杂度为 $O\left(\frac{\sqrt{cm}s_L^4}{\epsilon} \text{lb } m\right)$.

算法的第 2 步对稠密线性方程组技术进行了推广,应用于岭回归求解,这一步的算法的复杂度为 $O\left[\frac{c\sqrt{m}\kappa^2}{\epsilon}\text{polylb}(m+n)\right]$,其中 κ 为矩阵 \mathbf{X} 的条件数.因此算法的总复杂度为

$$O\left[\frac{\sqrt{cm}s_L^4}{\epsilon}\text{polylb}(m)+\frac{c\sqrt{m}\kappa^2}{\epsilon}\text{polylb}(m+n)\right].$$

作为对比,经典算法的复杂度为 $O(c^2m+cmns)$.

当 $m=O(\text{lb } n)$, $\frac{1}{\epsilon}=O(\text{lb } mn)$, $\kappa=O(\text{lb } mn)$, 因为 $c\ll m^{[16]}$, $s_L\leq m=O(\text{lb } n)$, 此时算法相对经典算法在 n 上有指数加速效果.当 $m=O(n)$, $\frac{1}{\epsilon}=O(\text{lb } mn)$, $\kappa=O(\text{lb } mn)$ 时,因为 $c\ll m^{[16]}$, 不妨令 $c=O(\text{lb } m)$, 若 $s_L=O(\text{lb } n)$, 那么算法相对经典算法有多项式加速效果;若 $s_L=\Omega\left(\frac{m}{c}\right)$ (这种情况常见于 LDA), 那么算法的复杂度比经典算法的复杂度更高.

3 改进的量子谱回归

在实际应用中,往往样本数量 m 和提取的特征个数 n 的规模是相差不大的,另外,在谱回归的原始论文中^[16],谱回归算法的优势在于降低了复杂度中参数 m 的次数.如果 $m\ll n$,谱回归的加速效果并不明显.因此在改进的算法中,在其他参数假设不变的前提下,我们希望在 $m=O(n)$ 时,量子算法相对经典算法有显著加速效果.

在对 MYXZ 算法复杂度的分析中,我们发现,在较多的情况下,矩阵 $\bar{\mathbf{L}}$ 并不是一个稀疏矩阵(如在 LDA 中,有 $cs_L\geq m$,通常认为 c 取值较小^[16],不妨令 $c=O(\text{lb } m)$,此时有 $s_L\geq \frac{m}{\text{lb } m}$),在主成分分析步骤中采用稀疏矩阵模拟的方式并没有加速效果.

针对这种情况,我们的改进算法不采用稀疏矩阵模拟的方式来进行主成分分析,而是直接用量子奇异值估计^[20] (QSVE)来得到相应的特征向量. QSVE 算法在处理稠密矩阵时具有更低的复杂度.具体地,类似于 MYXZ 算法,假设 $\bar{\mathbf{L}}$ 的信息存储在树状的 QRAM 中.那么可以制备

$$|\bar{\mathbf{L}}\rangle=\frac{1}{\|\bar{\mathbf{L}}\|_{\text{F}}}\sum_{ij}\bar{L}_{ij}|i\rangle|j\rangle,$$

不妨令 $|\bar{\mathbf{L}}\rangle=\frac{1}{\|\bar{\mathbf{L}}\|_{\text{F}}}\sum_{j=1}^m(\lambda_j+1)|\bar{\mathbf{y}}^{(j)}\rangle|\bar{\mathbf{y}}^{(j)}\rangle$, 那么

根据 QSVE^[20] 可得:

$$\frac{1}{\|\bar{\mathbf{L}}\|_{\text{F}}}\sum_{j=1}^m(\lambda_j+1)|\bar{\mathbf{y}}^{(j)}\rangle|\bar{\mathbf{y}}^{(j)}\rangle|\widetilde{\lambda_j+1}\rangle,$$

算法复杂度为 $O\left[\frac{1}{\epsilon}\text{polylb}(m)\right]$, 其中 $\widetilde{\lambda_j+1}\in(\lambda_j+1)\pm\delta\|\bar{\mathbf{L}}\|_{\text{F}}$. 令 $\delta\|\bar{\mathbf{L}}\|_{\text{F}}=O(\delta\sqrt{m})<\epsilon$, 则有 $\delta=O\left(\frac{\epsilon}{\sqrt{m}}\right)$, 此时复杂度为 $O\left[\frac{\sqrt{m}}{\epsilon}\text{polylb}(m)\right]$.

测量第 4 寄存器,以 $\frac{c}{m}$ 的概率测得较大的 λ_i , 此时可得对应的量子态 $|\bar{\mathbf{y}}^{(j)}\rangle$. 可通过振幅放大,重复次数降低为 $O(\sqrt{m/c})$, 因此第 1 步的总复杂度为 $O\left(\frac{m\sqrt{c}}{\epsilon}\text{lb } m\right)$. 结合 MYXZ 算法的第 2 步,总算法复杂度为 $O\left[\frac{cm\kappa^2}{\epsilon}\text{lb}(m+n)\right]$.

对比经典算法复杂度 $O(csmn+c^2n)$, 当 $\kappa=O(\text{lb } mn)$, $\frac{1}{\epsilon}=O(\text{lb } mn)$, $m=O(n)$ 时,算法具有多项式加速效果.对比 MYXZ 算法,我们算法在 $m=O(n)$ 且矩阵 $\bar{\mathbf{L}}$ 不是稀疏矩阵时,具有更低的复杂度.

4 高效谱回归的量子算法

在文献[16]中,Cai 等人针对结合标签信息来构造对应图的降维算法,如有监督的局部保持投影和邻域保持嵌入,提出了一个高效的实现方式.对于这类降维算法,式(1)的最大特征值 1 是 c 重特征值,因而 MYXZ 算法无法进行处理.我们提出了这类子空间学习问题的量子谱回归算法.我们的算法中也用到了 MYXZ 中的部分过程.

在 MYXZ 算法中,假设矩阵 $\bar{\mathbf{L}}$ 的信息被存储在树状的 QRAM 中,而在实际情况中,通常我们并不会事先知道矩阵 $\bar{\mathbf{L}}$ 的信息.特别在一些问题中,矩阵 $\bar{\mathbf{L}}$ 的计算较为复杂,如带标签的 NPE 算法, \mathbf{W} 由几个矩阵的乘积及求和得到,而 $\bar{\mathbf{L}}=\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{1/2}+\mathbf{I}_m$, 此时如果直接进行存储假设,而不计算其复杂度,对与之作对比的经典算法不公平.我们的算法仅假设数据矩阵 \mathbf{X} 被按标签顺序存储在树状的 QRAM 中.

4.1 算法过程

考虑到量子主成分分析的局限性,我们直接采用经典的方式来求式(1)的解.我们的算法步骤为:

1) 令第 1 个特征向量为全 1 向量 $(1, 1, \dots, 1)^T$, 通过对

$$\mathbf{y}'^{(i)} = [\underbrace{0, \dots, 0}_{m_j \text{ 个}}, \underbrace{1, \dots, 1}_{m_i \text{ 个}}, \underbrace{0, \dots, 0}_{\sum_{j=i+1}^c m_j \text{ 个}}], i = 1, 2, \dots, c$$

进行格拉姆-施密特正交化, 得到 $c-1$ 正交的特征向量, 作为式(1)的解, 记为 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(c-1)}$.

2) 把 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(c-1)}$ 存储到 QRAM 中, 制备量子态

$$|\psi\rangle = \frac{1}{\sqrt{c-1}} \sum_{i=0}^{c-2} |i\rangle |\mathbf{y}^{(i+1)}\rangle.$$

3) 对第 2 量子寄存器, 执行稠密矩阵岭回归量子算法, 得

$$|\phi\rangle = \frac{1}{\sqrt{c-1}} \sum_{i=0}^{c-2} |i\rangle |a_{i+1}\rangle.$$

4) 针对新输入的量子态 $|x\rangle$, 用量子低秩矩阵向量乘法, 求其降维后量子态 $|y\rangle$.

我们算法的第 4 步即对新输入的数据进行降维, 这正是子空间学习的目的. MYXZ 算法没有给出算法的降维过程, 为了算法的完整性, 我们将给出降维过程以及其复杂度分析.

4.2 算法分析

4.1 节所述的步骤 1 为经典步骤, 复杂度为 $mc^2 - \frac{1}{3}c^3$.

在 4.1 节所述的步骤 2 中, 我们先把 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(c-1)}$ 存储到 $c-1$ 个 QRAM 中, 因为元素个数为 $m(c-1)$, 因此 QRAM 存储的执行次数为 $O(mc)$. 通过调用存储 $\mathbf{y}^{(i)}$ 的 QRAM, 我们可以制备量子态 $\frac{1}{m} \sum_{j=0}^{m-1} |j\rangle |y_{j+1}^{(i)}\rangle$, 其中 $y_{j+1}^{(i)}$ 表示向量 $\mathbf{y}^{(i)}$ 的第 $j+1$ 个元素. 添加辅助比特进行受控旋转, 得

$$|\psi_1\rangle = \frac{1}{m} \sum_{j=0}^{m-1} |j\rangle |y_{j+1}^{(i)}\rangle \otimes \left[\frac{y_{j+1}^{(i)}}{C_z} |1\rangle + \sqrt{1 - \left(\frac{y_{j+1}^{(i)}}{C_z}\right)^2} |0\rangle \right],$$

其中 C_z 为绝对值不小于 $y_{j+1}^{(i)}$ 的绝对值的一个常数, 即 $\left|\frac{y_{j+1}^{(i)}}{C_z}\right| \leq 1$. 对第 2 寄存器做逆计算, 并测量辅助寄存器, 若测量结果为 1, 即得量子态 $|\mathbf{y}^{(i)}\rangle$. 可以看到, 对于一般的量子态, 测量结果为 1 的概率 $\frac{1}{m^2 C_z^2} \geq \frac{1}{m^2}$, 因此需要重复制备量子态 $|\psi_1\rangle$ 的次数为 $O(m^2)$. 在算法中引入幅度放大, 重复制备次数可以降低为 $O(m)$. 因此, 制备量子态 $|\mathbf{y}^{(i)}\rangle$ 的复杂度为

$O(mlbm)$. 进而, 制备 $|\psi\rangle$ 的复杂度为 $O(cmlbc \text{ lb} m)$. 其中因子 $\text{lb} c$ 来源于查找标签 i 所对应的存储位置, c 来源于读取 QRAM $c-1$ 次. 需要指出的是, QRAM 只需要存储一次, 就可以进行多次读取.

在 4.1 节所述的步骤 3 中, 我们采用 MYXZ 算法中稠密矩阵岭回归算法来求 $|\phi\rangle$. 具体地, 类似于 MYXZ 算法的假设, 假设矩阵 \mathbf{X} 事先存储在一个树状的 QRAM 中, 再在第 2 寄存器上直接执行量子稠密矩阵岭回归算法, 即得量子态 $|\phi\rangle$, 复杂度和 MYXZ 算法相应步骤一致, 为 $O\left[\frac{\sqrt{m}\kappa^2}{\epsilon} \text{polylb}(m+n)\right]$. 需要注意的是, 这个复杂度并没有考虑对 $|\psi\rangle$ 的制备, 因为需要进行 $O(\kappa)$ 次的调用, 因此这部分的复杂度为 $O(c\kappa m \text{ lb} c \text{ lb} m)$.

为了方便对 4.1 节所述的步骤 4 的分析, 我们定义对数据的量子访问^[21]:

定义 1^[21]. 如果存在以下变换可以在复杂度 T 内实现, 那么称我们可以通过量子访问矩阵 $\mathbf{A} \in \mathbb{R}^{n \times m}$:

$$|i\rangle |0\rangle \rightarrow |i\rangle |\mathbf{A}_i\rangle,$$

$$|0\rangle \rightarrow \frac{1}{\|\mathbf{A}\|_F} \sum_i \|\mathbf{A}_i\| |i\rangle,$$

其中, $i \in \{1, 2, \dots, n\}$, \mathbf{A}_i 表示矩阵 \mathbf{A} 的第 i 列构成的向量.

需要注意的是, 4.1 节算法步骤 2~3 可以实现 $|i\rangle |0\rangle \rightarrow |i\rangle |\mathbf{A}_i\rangle$, 但该过程不可逆(涉及到测量), 我们可以把这些过程转化为可逆的^[22], 而复杂度只有常数倍的增长. 在我们的算法中, $\|\mathbf{A}_i\| = 1$, 因此制备 $\frac{1}{\|\mathbf{A}\|_F} \sum_i \|\mathbf{A}_i\| |i\rangle$ 是平凡的. 因此我们有对数据矩阵的量子访问, 其中 $T = O\left[\frac{\sqrt{m}\kappa^2}{\epsilon} \text{polylb}(m+n)\right]$, 额外的复杂度为 $O(c\kappa m \text{ lb} c \text{ lb} m)$. 进而, 利用量子访问, 我们可以实现矩阵 $\begin{pmatrix} 0 & \mathbf{A} \\ \mathbf{A}^\dagger & 0 \end{pmatrix}$ 的 $(\|\mathbf{A}\|_F, \lceil \text{lb}(m+n) \rceil, 0)$ 块编码^[23], 复杂度为 $O\left[\frac{\sqrt{m}\kappa^2}{\epsilon} \text{polylb}(m+n)\right]$. 根据引理 1, 我们可得生成量子态

$$\begin{pmatrix} 0 & \mathbf{A} \\ \mathbf{A}^\dagger & 0 \end{pmatrix} |0\rangle |x\rangle = \begin{pmatrix} 0 & \mathbf{A} \\ \mathbf{A}^\dagger & 0 \end{pmatrix} \begin{pmatrix} |x\rangle \\ 0 \end{pmatrix} = |1\rangle |y\rangle$$

的复杂度(引理 1 中 γ 取常数, $\|\mathbf{A}\|_F = \sqrt{c}$)为 $\left[\frac{\sqrt{cm}\kappa^2}{\epsilon} \text{polylb}(m+n) \text{lb}\left(\frac{1}{\epsilon}\right)\right]$.

引理 1^[23]. 令 \mathbf{A} 为 $N \times N$ 的复矩阵, $|\mathbf{b}\rangle$ 为满足 $\|\mathbf{A}|\mathbf{b}\rangle\| \geq \gamma$ 量子态. 假制备 $|\mathbf{b}\rangle$ 的复杂度为 T_b ,

\mathbf{A} 的 $(\alpha, a, 0)$ 编码可以在 T_A 复杂度内实现,那么有量子算法可以以精度 ϵ 生成量子态 $\mathbf{A}|\mathbf{b}\rangle/\|\mathbf{A}|\mathbf{b}\rangle\|$, 复杂度为 $O\left[\frac{\alpha(T_b+T_A)}{\gamma}\text{lb}\left(\frac{1}{\epsilon}\right)\right]$.

综上,算法的总复杂度为

$$O\left[mc^2+\frac{\sqrt{cm}\kappa^2}{\epsilon}\text{polylb}(m+n)\text{lb}\left(\frac{1}{\epsilon}\right)+\right.\\ \left.c\kappa m\text{lb}c\text{lb}m\right]=O\left[\frac{c^2m\kappa^2}{\epsilon}\text{polylb}\left(\frac{m+n}{\epsilon}\right)\right],$$

相比经典算法的复杂度 $O(c^2m+cmns)$, 当 $\kappa, \frac{1}{\epsilon}=O(\text{lb}mn), m=O(n), c\ll m$ 时,我们提出的量子算法具有多项式加速效果.

相关算法的复杂度对比和适用条件如表 1 所示.需要注意的是,我们没有把降维过程(即 4.1 节中的算法步骤 4)的复杂度列在表中,因此表中量子高效谱回归的总复杂度比上面算出来的复杂度稍低,但并不影响对比结果.

Table 1 Comparison of Complexity and Applicable Conditions of Classical and Quantum Spectral Regression Algorithms

表 1 经典和量子谱回归算法的复杂度和适用条件对比

算法	最大特征问题	岭回归	适用条件	总复杂度
量子算法	$O\left[\frac{\sqrt{cm}s_L^4}{\epsilon}\text{polylb}(m)\right]$	$O\left[\frac{c\sqrt{m}\kappa^2}{\epsilon}\text{polylb}(m+n)\right]$	特征值不同, 稀疏	$O\left[\frac{\sqrt{cm}s_L^4}{\epsilon}\text{polylb}(m)+\frac{cm\kappa^2}{\epsilon}\text{polylb}(m+n)\right]$
经典算法	$mc^2-\frac{1}{3}c^3$	$O(csmn)$	整合标签信息的降维算法 (特征值相同)	$O(csmn+c^2m)$
改进算法 (第 3 节)	$O\left[\frac{m\sqrt{c}}{\epsilon}\text{polylb}(m)\right]$	$O\left[\frac{c\sqrt{m}\kappa^2}{\epsilon}\text{polylb}(m+n)\right]$	特征值不同	$O\left[\frac{cm\kappa^2}{\epsilon}\text{polylb}(m+n)\right]$
新算法 (第 4 节)	$mc^2-\frac{1}{3}c^3$	$O\left[\frac{\sqrt{m}\kappa^2}{\epsilon}\text{polylb}(m+n)+c\kappa m\text{lb}c\text{lb}m\right]$	整合标签信息的降维算法 (特征值相同)	$O\left[\frac{c^2\kappa^2m}{\epsilon}\text{polylb}(m+n)\right]$

5 结 论

在本文中,我们重新分析了 MYXZ 算法的细节,并纠正了其复杂度(并不影响 MYXZ 算法的结论).针对 MYXZ 算法的局限性,即在 $m=O(n)$ 时,只能在 \bar{L} 为稀疏矩阵的情形下才有量子加速,我们提出了一个改进算法.改进算法采用 QSVE 来得到 \bar{L} 的特征向量,因而复杂度不受 \bar{L} 是否稀疏的影响,算法在 $m=O(n)$ 且 \bar{L} 为稠密矩阵的情形下相对经典算法有多项式加速效果.另外,针对文献[16]中提出的高效谱回归算法,我们提出了其量子版本的算法.我们算法解决的问题并不能用 MYXZ 算法来解决,因为求解的问题有 $c-1$ 重特征值.实际上,如果特征值问题(1)的解中有重根,即有 2 个以上特征向量共享一个特征值,那么 MYXZ 算法将无法通过量子主成分分析得到相应的特征向量(只能得到特征向量的叠加态)而导致后续步骤出错.我们算法的优势在于把经典计算机上复杂度较高的部分用量子算法进行降维,而复杂度较低的部分保持原有经典算法.我们的算法设计思想不仅对子空间学习算

法,而且对其他类型的机器学习算法都具有一定的参考价值.

参 考 文 献

[1] Biamonte J, Wittek P, Pancotti N, et al. Quantum machine learning [J]. Nature, 2017, 549(7671): 195-202

[2] Wang Yongli, Xu Qiuliang. Principle and research progress of quantum computation and quantum cryptography [J]. Journal of Computer Research and Development, 2020, 57(10): 2015-2026
(王永利, 徐秋亮. 量子计算与量子密码的原理及研究进展综述[J]. 计算机研究与发展, 2020, 57(10): 2015-2026)

[3] Shor P. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer [J]. SIAM Journal on Computing, 1997, 26(5): 1484-150

[4] Grover L K. A fast quantum mechanical algorithm for database search [C] //Proc of the 28th Annual ACM Symp on Theory of Computing. New York: ACM, 1996: 212-219

[5] Harrow A W, Hassidim A, Lloyd S. Quantum algorithm for linear systems of equations [J]. Physical Review Letters, 2009, 103(15): No.150502

[6] Lloyd S, Mohseni M, Rebentrost P. Quantum principal component analysis [J]. Nature Physics, 2014, 10(9): 631-633

[7] Wiebe N, Braun D, Lloyd S. Quantum algorithm for data fitting [J]. Physical Review Letters, 2012, 109(5): No. 050505

[8] Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big data classification [J]. Physical Review Letters, 2014, 113(13): No.130503

[9] Lloyd S, Weedbrook C. Quantum generative adversarial learning [J]. Physical Review Letters, 2018, 121(4): No. 040502

[10] Situ Haozhen, He Zhimin, Wang Yuyi, et al. Quantum generative adversarial network for generating discrete distribution [J]. Information Sciences, 2020 (538): 193-208

[11] Yu Chaohua, Gao Fei, Wang Qingle, et al. Quantum algorithm for association rules mining [J]. Physical Review A, 2016, 94(4): No.042311

[12] He Xiaofei, Yan Shuicheng, Hu Yuxiao, et al. Face recognition using Laplacianfaces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(3): 328-340

[13] Hotelling H. Relations between two sets of variates [J]. Biometrika, 1936, 28(3/4): 321-377

[14] Fisher R A. The use of multiple measurements in taxonomic problems [J]. Annals of Eugenics, 1936, 7(2): 179-188

[15] He Xiaofei, Cai Deng, Yan Shuicheng, et al. Neighborhood preserving embedding [C] //Proc of the 10th IEEE Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2005: 1208-1213

[16] Cai Deng, He Xiaofei, Han Jiawei. Spectral regression for efficient regularized subspace learning [C] //Proc of IEEE 11th Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2007: 1-8

[17] Yu Chaohua, Gao Fei, Lin Song, et al. Quantum data compression by principal component analysis [J]. Quantum Information Processing, 2019, 18(8): No.249

[18] Cong I, Duan Luming. Quantum discriminant analysis for dimensionality reduction and classification [J]. New Journal of Physics, 2016, 18(7): No.073011

[19] Meng Fanxu, Yu Xutao, Xiang Ruiqing, et al. Quantum algorithm for spectral regression for regularized subspace learning [J]. IEEE Access, 2019, 7: 4825-4832

[20] Kerenidis I, Prakash A. Quantum recommendation systems [C] //Proc of the 8th Innovations in Theoretical Computer Science Conf. Dagstuhl: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017, 67: 49:1-49:21

[21] Kerenidis I, Landman J. Quantum spectral clustering [J]. Physical Review A, 2021, 103(4): No. 042415

[22] Kerenidis I, Landman J, Luongo A, et al. q-means: A quantum algorithm for unsupervised machine learning [C] //

Advances in Neural Information Processing Systems 32. Vancouver: Curran Associates, Inc., 2019: 1-11

[23] Chakraborty S, Gilyén A, Jeffery S. The power of block-encoded matrix powers: Improved regression techniques via faster Hamiltonian simulation [C] //Proc of the 46th Int Colloquium on Automata, Languages, and Programming. Dagstuhl: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019: 33:1-33:14



Pan Shijie, born in 1991. PhD candidate. His main research interests include quantum algorithms and quantum machine learning.
潘世杰,1991年生.博士研究生.主要研究方向为量子算法和量子机器学习.



Gao Fei, born in 1980. Professor, PhD supervisor. His main research interests include quantum algorithms, quantum cryptography and related quantum information problems.
高飞,1980年生.教授,博士生导师.主要研究量子算法、量子密码学及相关量子信息问题.



Wan Linchun, born in 1992. PhD candidate. His main research interests include quantum algorithm and privacy-preserving quantum machine learning.
万林春,1992年生.博士研究生.主要研究方向为量子算法和隐私保护的量子机器学习.



Qin Sujuan, born in 1979. Professor, PhD supervisor. Her main research interests include quantum cryptography, quantum algorithms and network security.
秦素娟,1979年生.教授,博士生导师.主要研究方向为量子密码学、量子算法和网络安全.



Wen Qiaoyan, born in 1959. Professor, PhD supervisor. Her main research interests include quantum cryptography, quantum algorithms, coding theory, information security, network security and applied mathematics.
温巧燕,1959年生.教授,博士生导师.主要研究方向为量子密码学、量子算法、编码理论、信息安全、网络安全和应用数学.