一种基于量子线路的支持向量机训练方案

张毅军[1)3)] 慕晓冬[3)] 郭乐勐[2)] 张朋[4)3)] 赵导[4)3)] 白文华[2)†]

1) (国防科技大学, 试验训练基地, 西安 710106)

2) (国防科技大学, 信息通信学院, 武汉 430079)

3) (火箭军工程大学, 研究生院, 西安 710043)

4) (西安卫星测控中心, 西安 710043)

(2022 年 10 月 20 日收到; 2023 年 1 月 19 日收到修改稿)

本文针对支持向量机提出一种基于量子态内积的量子线路训练方案. 该方案以量子基础力学理论为基础, 通过量子化, 生成支持向量机训练样本元素对应的量子态; 以量子初始基态和对应的量子逻辑门为基础, 构建可以实现训练样本元素量子态的量子线路; 通过建立量子态内积与 SWAP 量子逻辑门之间的关系, 采用量子态振幅的交换演化操作来实现量子态内积. 验证结果表明, 该方案不但使得支持向量机完成了正确分类, 还针对该方案的量子部分实现了在真实量子计算机上运行, 与经典算法相比, 多项式程度上降低了算法的时间复杂度, 扩展了支持向量机的训练思路.

关键词: 量子线路, 内积, 量子态, 支持向量机

PACS: 03.67.Lx, 03.65.Aa **DOI:** 10.7498/aps.72.20222003

1 引 言

支持向量机由 Cortes 和 Vapnik[1] 于 1995 年提出, 属于机器学习中的监督学习模型, 其基本原理主要是基于统计学习理论和最优化理论[2]. 它主要用于求解分类和回归方面的问题[3], 由于其具有很强的适应性和泛化性, 以及全局优化性等特点, 受到了广泛关注, 取得了丰硕的研究成果, 在图像识别[4,5]、图像分类[6,7]、模式识别[8,9]、人脸检测[10,11]、蛋白质预测[12,13]、邮件分类[14]、数据挖掘[15,16] 等领域得到了广泛的应用.

但是, 随着训练样本数据规模的不断增大, 支持向量机在处理问题时, 计算复杂度成多项式增长, 明显表现出计算速度下降和预测准确率降低等弱点[17,18]. 针对这些弱点, 大量的专家学者对支持向量机进行了研究, 提出了许多改进的方法. Mangasarian 和 Wild[19] 提出利用两个彼此不平行的平面进行划分的概念. 该方法放宽了支持向量机对生成边界或者平面的要求, 通过将问题转化为广义特征值问题进行求解, 一定程度提升了算法的速率. Zhang 和 Suganthan[20] 提出了一种基于倾斜决策树生成决策超平面的方法, 根据训练数据的几何性质将数据分成两个超类, 提升算法的运行效率. Xu 等[21] 提出了一种基于一致遍历马尔可夫链样本采样的支持向量机算法, 通过马尔克夫方法对给定的数据集进行采样, 使训练样本具有泛化性, 使分类器具有稀疏性. Zou 等[22] 提出了一种基于 K 次马尔克夫抽样的支持向量机算法, 使算法的采样和训练时间减少, 分类器具有更好的稀疏性. Wu 等[23] 提出了一种基于二叉树支持向量机的教学评价模型, 通过对评价的数据结构和类型进行变换, 实现对教学质量评价的快速准确分类. 然而这些经典算法虽然提升了支持向量机的计算速率, 但

† 通信作者. E-mail: baiwenhua@nudt.edu.cn

© 2023 中国物理学会 Chinese Physical Society

基本都是通过分解训练数据样本或者减少训练数据样本的方式实现, 本质上并没有降低算法的时间复杂度. 近年来, 随着量子技术的不断发展, 在一些领域量子优势已经体现出来[24–28]. 由于量子本身所具有的相干和纠缠等物理特性, 使其天然具有并行计算的能力, 在一些问题的处理上, 相对于经典算法已经实现了指数级加速[29–34]. Lloyd 等[35]在 2013 年对机器学习中的聚类分配和聚类搜索问题进行分析研究, 提出了有监督和无监督量子机器学习算法, 相对经典算法具有指数加速. Rebentrost 等[36]又于 2014 年对支持向量机进行了研究分析, 利用非解析矩阵求幂技术, 提出了量子支持向量机算法, 相对经典支持向量机可以获得指数级加速. 在上述两个研究成果的启发下, 本工作通过对经典支持向量机中的核函数进行量子化, 设计出对应的量子线路并采用量子方法完成求解, 实现对支持向量机的量子加速.

本文通过构建基于量子态内积的量子线路训练方案, 实现对支持向量机训练样本的多项式级加速. 首先, 在对经典支持向量机的计算复杂度进行分析的基础上, 从影响算法计算复杂度的主要因素核函数入手, 以量子力学和量子计算理论为基础, 对核函数中的训练样本元素进行量子化, 生成对应的量子态; 其次, 根据训练样本元素的量子态推导计算出所需的量子逻辑门种类与数量, 设计出可以通过量子初始基态和量子逻辑门演化生成训练样本元素对应的量子态的量子线路; 接着, 根据量子态内积范式与 SWAP 量子逻辑门之间的联系, 设计出对应的量子线路, 完成对相应量子态振幅的交换操作; 最后, 通过对控制位量子态进行测量, 求出核函数对应的内积数值, 实现对支持向量机训练的加速效果.

2 基础理论及推导

对于一个支持向量机, 设训练集的样本数量为 M, 特征数量为 N, 引入松弛因子 ε_i 后, 对应的数学模型可以表示为

$$\min_{w,b,\varepsilon_i} \frac{1}{2}||w||^2 + C\sum_{i=1}^{M}\varepsilon_i, \tag{1}$$

$$\text{s.t. } y_i\left(w^{\mathrm{T}}x_i + b\right) \geqslant 1-\varepsilon_i,\ (i=1,2,\cdots,M), \tag{2}$$

$$\varepsilon_i \geqslant 0,\ (i=1,2,\cdots,M), \tag{3}$$

其中, C 为惩罚参数, C 值越大说明模型更加关注这个弹性范围带来的代价, 对误分类的样本点的惩罚就越大.

将目标函数以及约束条件使用拉格朗日函数转化为无约束问题:

$$L(w,b,\varepsilon,\alpha,\mu) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\varepsilon_i - \sum_{i=1}^{M}\alpha_i\left[y_i\left(w^{\mathrm{T}}x_i+b\right)-1+\varepsilon_i\right] - \sum_{i=1}^{M}\mu_i\varepsilon_i, \tag{4}$$

其中 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_M)^{\mathrm{T}}$ 和 $\mu=(\mu_1,\mu_2,\cdots,\mu_M)^{\mathrm{T}}$ 均为拉格朗日乘子向量, 且 $\alpha_i\geqslant 0, \mu_i\geqslant 0$. 这样优化目标函数就变为

$$\min_{w,b,\varepsilon}\max_{\alpha_i\geqslant 0,\mu_i\geqslant 0} L(w,b,\varepsilon,\alpha,\mu). \tag{5}$$

根据 KKT 条件, (5) 式表示的优化目标函数等价于求解其对偶问题:

$$\max_{\alpha_i\geqslant 0,\,\mu_i\geqslant 0}\min_{w,b,\varepsilon} L(w,b,\varepsilon,\alpha,\mu). \tag{6}$$

接着, 通过目标函数对 w,b,ε 求偏导, 经过优化和化简后, 将目标函数转化为

$$\min_{\alpha}\frac{1}{2}\sum_{i=1,j=1}^{M}\alpha_i\alpha_j y_i y_j\langle x_i,x_j\rangle - \sum_{i=1}^{M}\alpha_i, \tag{7}$$

$$\text{s.t. } \sum_{i=1}^{M}\alpha_i y_i = 0, \tag{8}$$

$$0\leqslant\alpha_i\leqslant C. \tag{9}$$

2.1 支持向量机中拉格朗日向量乘子

在通常求解过程中, 对于支持向量机中求解拉格朗日向量乘子 α 问题一般都转化为二次规划问题进行求解. 选块算法[37]、分解算法[38]、序列最小优化 (sequential minimal optimization, SMO) 算法[39]、增量算法[40]都可以用来求解支持向量机中的二次规划问题, 其中 SMO 算法采用的是启发式策略, 在求解速率和质量方面相较于其他方法表现较为突出.

SMO 算法对每个拉格朗日乘子 $(\alpha_1,\alpha_2,\cdots,\alpha_M)$ 进行优化过程中, 它先将原始的 M 个参数二次规划问题分解成多个子二次规划问题进行求解, 再对每个子二次规划问题分别进行求解. 这样对于每个子问题只需要求解 2 个参数, 节省了大量的时间和内存需求.

对于一个子问题, 假设选取的两个需要优化的参数为 α_1,α_2, 剩下的参数 $\alpha_3,\alpha_4,\cdots,\alpha_M$ 则作为常数处理. 经过一系列的优化计算, 可以得到如下结果:

$$\alpha_1^{\text{new}}=\frac{-\sum_{i=3}^{M}\alpha_i y_i-y_2\alpha_2^{\text{old}}}{y_1}-\frac{E_1-E_2}{\eta y_1},\tag{10}$$

$$\alpha_2^{\text{new}}=\alpha_2^{\text{old}}+\frac{y_2\left(E_1-E_2\right)}{\eta},\tag{11}$$

$$\begin{aligned}b_1^{\text{new}}=&\ y_1-\sum_{i=3}^{M}\alpha_i y_i\left\langle x_i,x_1\right\rangle\\&-\alpha_1^{\text{new}}y_1\left\langle x_1,x_1\right\rangle-\alpha_2^{\text{new}}y_2\left\langle x_2,x_1\right\rangle,\end{aligned}\tag{12}$$

$$\begin{aligned}b_2^{\text{new}}=&\ -E_2-y_1\left\langle x_1,x_2\right\rangle\left(\alpha_1^{\text{new}}-\alpha_1^{\text{old}}\right)\\&-y_2\left\langle x_2,x_2\right\rangle\left(\alpha_2^{\text{new}}-\alpha_2^{\text{old}}\right)+b_{\text{old}},\end{aligned}\tag{13}$$

其中 α_1^{new} 和 α_2^{new} 分别表示 α_1 和 α_2 经过优化后的取值, α_1^{old} 和 α_2^{old} 分别表示 α_1 和 α_2 优化前的取值, b_1^{new} 和 b_2^{new} 分别表示通过 α_1^{new} 和 α_2^{new} 更新阈值 b 的取值. $E_1=\sum_{i=1}^{M}\alpha_i y_i\left\langle x_i,x_1\right\rangle+b-y_1$ 表示训练集取 (x_1,y_1) 时, 支持向量机预测值与真实值的误差. $E_2=\sum_{i=1}^{M}\alpha_i y_i\left\langle x_i,x_2\right\rangle+b-y_2$ 表示训练集取 (x_2,y_2) 时, 支持向量机预测值与真实值的误差. η 等于 $\left\langle x_1,x_1\right\rangle+\left\langle x_2,x_2\right\rangle-2\left\langle x_1,x_2\right\rangle$.

通过上述公式, 可以分析出在优化拉格朗日向量乘子 α 和阈值 b 的过程中, 都需进行训练样本集之间的内积计算. 当训练样本集中每个元素的特征数量较大时, 这种内积计算就非常消耗时间和内存资源, 从而增加算法的计算复杂度, 降低算法的执行效率.

2.2 向量内积与量子态内积

在希尔伯特空间中, 一个量子位对应的量子态 $|\psi\rangle$.可以表示为二维向量的形式:

$$|\psi\rangle=\beta|0\rangle+\gamma|1\rangle=\beta\begin{bmatrix}1\\0\end{bmatrix}+\gamma\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}\beta\\\gamma\end{bmatrix},\tag{14}$$

其中, β 和 γ 均为复数, 称为振幅, 且 $|\beta|^2+|\gamma|^2=1$. 同理, 根据张量积运算, 两个量子位对应的量子态就可以表示成一个四维向量形式, 以此类推, n 个量子位对应的量子态就可以表示成一个 2^n 维向量. 设任意两个量子态分别为 $|\varphi\rangle=[e_1,e_2,\cdots,e_K]^{\mathrm{T}}$, $|\phi\rangle=[f_1,f_2,\cdots,f_K]^{\mathrm{T}}$, 其中 $K=2^n$. 根据量子力学基础理论和量子态 $|\varphi\rangle$ 与 $|\phi\rangle$ 的向量表达形式, 通过计算推导, 量子态 $|\varphi\rangle$ 与 $|\phi\rangle$ 的内积可以表示为

$$\left\langle\varphi|\,\phi\right\rangle=[e_1^*,e_2^*,\cdots,e_K^*]\,[f_1,f_2,\cdots,f_K]^T=\sum_{i=1}^{K}e_i^*f_i.\tag{15}$$

设支持向量机的特征数量为 N, 设任意两组训练样本数据分别为 $x_1=[x_{11},x_{12},\cdots,x_{1N}]^{\mathrm{T}}$, $x_2=[x_{21},x_{22},\cdots,x_{2N}]^{\mathrm{T}}$. 对 x_1 和 x_2 的向量表达式进行归一化, 设 $F=||x_2||=\sqrt{\sum_{i=1}^{N}x_{2i}^2}$, $F=\|x_2\|=\sqrt{\sum_{i=1}^{N}x_{2i}^2}$. 将 E 和 F 分别代入 x_1 和 x_2 表示式中, 可将其变换为

$$x_1=E\left[\frac{x_{11}}{E},\frac{x_{12}}{E},\cdots,\frac{x_{1N}}{E}\right]^{\mathrm{T}},\tag{16}$$

$$x_2=F\left[\frac{x_{21}}{F},\frac{x_{22}}{F},\cdots,\frac{x_{2N}}{F}\right]^{\mathrm{T}}.\tag{17}$$

根据 (15)—(17) 式及上述相关理论可知, 当支持向量机的特征数量为 N 时, 需要的量子位数为 $\lceil\log_2 N\rceil$. 当量子态对应的向量形式中出现多余的维度时, 将其对应的振幅数值设置为 0 即可. 因此,当 $K=N$ 时, 令 $e_1=\frac{x_{11}}{E},e_2=\frac{x_{12}}{E},\cdots,e_K=\frac{x_{1N}}{E}$, 令 $f_1=\frac{x_{21}}{F},f_2=\frac{x_{22}}{F},\cdots,f_K=\frac{x_{2N}}{F}$; 当 $K>N$ 时, 令 $e_1=\frac{x_{11}}{E},e_2=\frac{x_{12}}{E},\cdots,e_N=\frac{x_{1N}}{E},e_{N+1}=0,e_{N+2}=0,\cdots,e_K=0$, 令 $f_1=\frac{x_{21}}{F},f_2=\frac{x_{22}}{F},\cdots,f_N=\frac{x_{2N}}{F},f_{N+1}=0,f_{N+2}=0,\cdots,f_K=0$. 通过推导可得如下公式:

$$\left\langle\varphi|\,\phi\right\rangle=\left\langle\frac{x_1}{E},\frac{x_2}{F}\right\rangle=\frac{1}{EF}\left\langle x_1,x_2\right\rangle.\tag{18}$$

这样通过量子计算方法求解经典内积的理论基础就被建立起来, 使接下来构建设计量子线路进行内积求解成为可能.

3 量子线路

由于实验中量子计算机等硬件设备和可展示性等限制, 设置训练样本数据的维度为 N=4, 根据本文 2.2 节内容的分析结果, 可知需要具有两个量子比特的量子态进行量子态内积计算. 设两个量子态分别为 $|\varphi\rangle=[e_1,e_2,e_3,e_4]^{\mathrm{T}}$, $|\phi\rangle=[f_1,f_2,f_3,f_4]^{\mathrm{T}}$, 分别与支持向量机的两个训练样本数据为 $x_1=[x_{11},x_{12},x_{13},x_{14}]^{\mathrm{T}}$, $x_2=[x_{21},x_{22},x_{23},x_{24}]^{\mathrm{T}}$ 相对应.

3.1 制备量子态的量子线路

在目前量子线路系统中, 用于改变振幅的量子逻辑门主要有以 Pauli-X 为生成元的 $\mathrm{RX}(\theta)$ 门和以 Pauli-Y 为生成元的 $\mathrm{RY}(\theta)$ 门. $\mathrm{RX}(\theta)$ 门在对量子态进行演化过程中, 对振幅改变的同时也向振幅中引入了虚数. $\mathrm{RY}(\theta)$ 门在对量子态进行演化过程中, 只对振幅进行改变没有引入虚数. 根据本文的实验需求, 选用 $\mathrm{RY}(\theta)$ 门对量子态振幅进行操作. 在本文设计的量子态生成量子线路中, 主要通过 $\mathrm{RY}(\theta)$ 量子逻辑门和 X 量子逻辑门来具体实现. 其中 $\mathrm{RX}(\theta)$ 门的矩阵形式为

$$\mathrm{RX}(\theta)=\mathrm{e}^{-\mathrm{i}\theta X/2}=\begin{bmatrix}\cos\left(\dfrac{\theta}{2}\right) & -\mathrm{i}\sin\left(\dfrac{\theta}{2}\right)\\ -\mathrm{i}\sin\left(\dfrac{\theta}{2}\right) & \cos\left(\dfrac{\theta}{2}\right)\end{bmatrix}. \tag{19}$$

$\mathrm{RY}(\theta)$ 门的矩阵形式为

$$\mathrm{RY}(\theta)=\mathrm{e}^{-\mathrm{i}\theta Y/2}=\begin{bmatrix}\cos\left(\dfrac{\theta}{2}\right) & -\sin\left(\dfrac{\theta}{2}\right)\\ \sin\left(\dfrac{\theta}{2}\right) & \cos\left(\dfrac{\theta}{2}\right)\end{bmatrix}. \tag{20}$$

X 门的矩阵形式为

$$X=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}. \tag{21}$$

以生成量子态 $|\varphi\rangle$ 为例. 首先, 在初始量子态 $|00\rangle$ 的高位上作用一个 $\mathrm{RY}(\theta_1)$ 量子逻辑门, 其中 $\theta_1=2\cos^{-1}\left(\sqrt{e_1^2+e_2^2}\right)$, 并且 $0\leqslant\theta_1\leqslant\pi/2$; 其次, 在量子态的低位上作用一个控制 $\mathrm{RY}(\theta_2)$ 量子逻辑门, 控制位为量子态的高位, 其中 $\theta_2=2\cos^{-1}\left(e_3/\sqrt{e_3^2+e_4^2}\right)$, 并且 $0\leqslant\theta_2\leqslant\pi/2$; 接着, 在量子态的高位上作用一个 X 量子逻辑门; 随后, 在量子态的低位上作用一个控制 $\mathrm{RY}(\theta_3)$ 量子逻辑门, 控制位为量子态的高位, 其中 $\theta_3=2\cos^{-1}\left(e_1/\sqrt{e_1^2+e_2^2}\right)$, 并且 $0\leqslant\theta_3\leqslant\pi/2$; 最后, 在量子态的高位上作用一个 X 量子逻辑门. 制备量子态 $|\varphi\rangle$ 的量子线路简略图如图 1 所示.

具体的计算推导如下:

$$\begin{aligned}
&|00\rangle\xrightarrow{\text{High position}\mathrm{RY}(\theta_1)}\left(\sqrt{e_1^2+e_2^2}\,|0\rangle+\sqrt{e_3^2+e_4^2}\,|1\rangle\right)|0\rangle\\
&\xrightarrow{\text{Low position}\mathrm{RY}(\theta_2)}\sqrt{e_1^2+e_2^2}\,|0\rangle\,|0\rangle\\
&+\sqrt{e_3^2+e_4^2}\,|1\rangle\left(e_3/\sqrt{e_3^2+e_4^2}\,|0\rangle+e_4/\sqrt{e_3^2+e_4^2}\,|1\rangle\right)\\
&\xrightarrow{\text{High position}X}\sqrt{e_1^2+e_2^2}\,|1\rangle\,|0\rangle+e_3\,|0\rangle\,|0\rangle+e_4\,|0\rangle\,|1\rangle\\
&\xrightarrow{\text{Low position}\mathrm{RY}(\theta_3)}\sqrt{e_1^2+e_2^2}\,|1\rangle\left(e_1/\sqrt{e_1^2+e_2^2}\,|0\rangle\right.\\
&\left.+e_2/\sqrt{e_1^2+e_2^2}|1\rangle\right)+e_3\,|0\rangle\,|0\rangle+e_4\,|0\rangle\,|1\rangle\\
&\xrightarrow{\text{High position}X}e_1\,|0\rangle\,|0\rangle+e_2|0\rangle|1\rangle+e_3|1\rangle|0\rangle\\
&+e_4|1\rangle|1\rangle.
\end{aligned} \tag{22}$$

同理, 也可以生成制备量子态 $|\phi\rangle$ 的量子线路, 只是将依次作用于量子线路的 $\mathrm{RY}(\theta_1)$ 量子逻辑门、控制 $\mathrm{RY}(\theta_2)$ 量子逻辑门、控制 $\mathrm{RY}(\theta_3)$ 量子逻辑门的旋转角度分别设置为 $\theta_1=2\cos^{-1}\left(\sqrt{f_1^2+f_2^2}\right)$, $\theta_2=2\cos^{-1}\left(f_3/\sqrt{f_3^2+f_4^2}\right)$, $\theta_3=2\cos^{-1}\left(f_1/\sqrt{f_1^2+f_2^2}\right)$, 并且 $0\leqslant\theta_1,\theta_2,\theta_3\leqslant\pi/2$.

根据制备量子态对应的量子线路, 对其进行编译并在本源量子云平台上进行实现. 本源量子云平台不仅可以提供真实量子计算, 还可以提供量子计算模拟器. 目前该平台提供的真实量子计算机主要是包含 6 个量子比特的超导量子计算机, 也称本源悟源; 提供的量子计算模拟器主要有全振幅量子模拟机、含噪声量子模拟机、部分振幅量子模拟机和单振幅量子模拟机. 在量子云平台上, 如图 2 所示为制备量子态 $|\varphi\rangle$ 和 $|\phi\rangle$ 对应的量子线路.

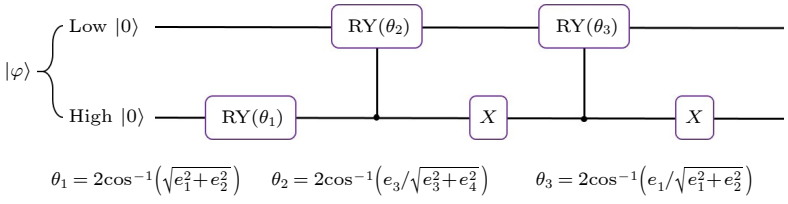


图 1 制备量子态 $|\varphi\rangle$ 的量子线路简略图

Fig. 1. The schematic diagram of the quantum circuit for preparing the quantum state $|\varphi\rangle$.

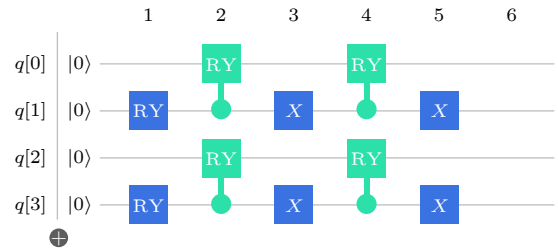


图 2 制备量子态 $|\varphi\rangle$ 和 $|\phi\rangle$ 的量子线路

Fig. 2. The quantum circuit for preparing the quantum states $|\varphi\rangle$ and $|\phi\rangle$.

3.2 量子态内积的量子线路

在完成$|\varphi\rangle$和$|\phi\rangle$两个量子态制备后, 接下来就是构建实现两个量子态内积的量子线路. 因为量子线路最后的测量结果只与对应的量子态振幅有关, 而要想让测量结果与两个量子态的内积有关, 可以在量子线路演化过程中, 通过增加控制位线路对两个量子态状态进行交换演化来实现. 紧接着可以利用量子态之间的纠缠关系, 通过对控制位量子态的测量来求出两个量子态的内积.

在生成量子线路过程中, 主要通过 H 量子逻辑门和控制 SWAP 量子逻辑门来实现的. 其中 H 量子逻辑门的矩阵形式为

$$H=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}. \tag{23}$$

SWAP 量子逻辑门的矩阵形式为

$$\mathrm{SWAP}=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}. \tag{24}$$

首先, 在初始量子态$|0\rangle$的控制位上作用一个 H 量子逻辑门; 其次, 在$|\varphi\rangle$和$|\phi\rangle$两个量子态上作用一个控制 SWAP 量子逻辑门; 接着, 在控制位上作用一个 H 量子逻辑门; 最后, 对控制位进行测量. 构建的量子线路简略图如图 3 所示.

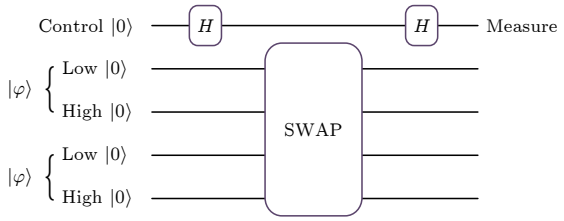


图 3 量子态内积对应的量子线路简略图

Fig. 3. The quantum circuit schematic diagram of the IPQS.

具体的计算推导如下:

$$|0\rangle_1|\varphi\rangle_2|\phi\rangle_3 \xrightarrow{1:H} \frac{1}{\sqrt{2}}\left(|0\rangle_1|\varphi\rangle_2|\phi\rangle_3+|1\rangle_1|\varphi\rangle_2|\phi\rangle_3\right) \xrightarrow{\text{The control qubit is 1, 2 and 3: SWAP}}$$

$$\frac{1}{\sqrt{2}}\left(|0\rangle_1|\varphi\rangle_2|\phi\rangle_3+|1\rangle_1|\phi\rangle_2|\varphi\rangle_3\right) \xrightarrow{1:H} \frac{1}{2}|0\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3+|\phi\rangle_2|\varphi\rangle_3\right)+\frac{1}{2}|1\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3-|\phi\rangle_2|\varphi\rangle_3\right). \tag{25}$$

用标准基态$|0\rangle$测量控制位得到 0 的概率为

$$\begin{aligned}&\left[\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2+\langle\varphi|_3\langle\phi|_2\right)\langle 0|_1+\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2-\langle\varphi|_3\langle\phi|_2\right)\langle 1|_1\right]|0\rangle_1\langle 0|_1\\&\left[\frac{1}{2}|0\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3+|\phi\rangle_2|\varphi\rangle_3\right)+\frac{1}{2}|1\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3-|\phi\rangle_2|\varphi\rangle_3\right)\right]\\=&\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2+\langle\varphi|_3\langle\phi|_2\right)\langle 0|_1|0\rangle_1\langle 0|_1\frac{1}{2}|0\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3+|\phi\rangle_2|\varphi\rangle_3\right)=\frac{1}{2}\left(1+\langle\varphi|\phi\rangle^2\right).\end{aligned} \tag{26}$$

用标准基态$|1\rangle$测量控制位得到 1 的概率为

$$\begin{aligned}&\left[\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2+\langle\varphi|_3\langle\phi|_2\right)\langle 0|_1+\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2-\langle\varphi|_3\langle\phi|_2\right)\langle 1|_1\right]|1\rangle_1\langle 1|_1\\&\left[\frac{1}{2}|0\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3+|\phi\rangle_2|\varphi\rangle_3\right)+\frac{1}{2}|1\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3-|\phi\rangle_2|\varphi\rangle_3\right)\right]\\=&\frac{1}{2}\left(\langle\phi|_3\langle\varphi|_2-\langle\varphi|_3\langle\phi|_2\right)\langle 1|_1|1\rangle_1\langle 1|_1\frac{1}{2}|1\rangle_1\left(|\varphi\rangle_2|\phi\rangle_3-|\phi\rangle_2|\varphi\rangle_3\right)=\frac{1}{2}\left(1-\langle\varphi|\phi\rangle^2\right).\end{aligned} \tag{27}$$

将生成量子态的量子线路与求解量子内积的量子线路相结合, 统一进行编译并在本源量子云平台上实现. 在量子云平台上运行的量子线路如图 4 所示.

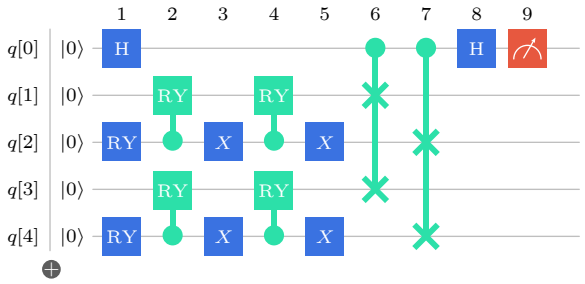


图 4 量子态内积对应的量子线路图

Fig. 4. The quantum circuit diagram of the IPQS.

4 仿真实验分析

为了能在真实的量子计算机上进行实验验证, 本文从 UCI 数据库中的 Haberman 数据集中随机提取了 100 组数据 (数据特征维度为 3), 在进行缩放处理后作为训练数据. 下文将对构建的量子线路和所提出的训练方案进行仿真实验验证.

4.1 验证量子线路

为了验证构建的量子线路精度, 从训练样本中随机选取了 5 组数据 (表 1), 按照如图 4 的量子线路结构, 在本源悟源 6 比特超导量子计算机上进行实验验证.

实验结果如图 5 所示, 为 5 组数据用标准基态 $|0\rangle$和$|1\rangle$分别测量 0 和 1 的结果. 根据测量结果, 利用 (18), (26) 和 (27) 式等分别计算出每组数据的内积, 再与通过经典方法计算出数据组的内积进行对比. 如表 1 所示, 采用基于量子态内积的量子线路求解方案计算内积数值与经典方法计算内积数值之间的误差最小 0.18%, 最大为 2.89%. 通过对误差数值进行分析, 可以得出量子线路的测量误差在可接受的范围内, 量子线路的测量精度比较高.

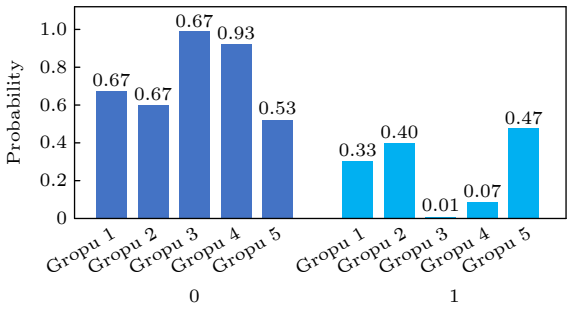


图 5 5 组数据用标准基态 $|0\rangle$和$|1\rangle$分别测量 0 和 1 的概率

Fig. 5. The probability of 0 and 1 separately measured by standard ground state $|0\rangle$ and $|1\rangle$ for five groups of data.

4.2 实现量子加速训练支持向量机

为了对该求解方案进行验证, 本文在 Python 框架下的量子软件开发环境中对该方案进行了编程实现, 具体的伪代码如表 2 所示. 通过表 2 可以看出, 所有涉及内积计算的部分都是通过量子线路来实现的. 为了便于对比分析, 也利用经典的 SMO 算法对训练支持向量机进行了编程实现, 这两种方法都利用同一个训练样本数据集对支持向量机进行训练, 训练样本数据集如图 6 所示.

两种求解方案在对同一训练样本进行相同次数优化后, 如图 7 所示取得了分别显示了基于量子态内积的量子线路训练方案与基于经典 SMO 算法训练支持向量机的训练结果. 根据图中显示的结果, 可以看出基于量子态内积的量子线路训练方案与基于经典 SMO 算法都实现了对训练样本的正确分类, 效果都挺好; 并且两种方案确定的支持向量点位也基本相同. 虽然两种方法的训练效果基本相同, 但是基于量子态内积的量子线路训练方案时

表 1 量子方案与经典方法求解内积结果对比

Table 1. Comparison of the inner product results between the quantum scheme and the classical method.

SN	Data group	Result of 0	Result of 1	Classical calculation	Error rate/%
		Inner product value			
1	(2.942485, 4.977398, 3.176513) (7.551510, 1.580030, 0.067732)	29.679279	29.679279	30.299795	2.05
2	(0.341367, 3.894998, 3.929515) (7.139979, 2.329896, 1.981083)	19.255815	19.255815	19.296989	0.21
3	(6.080573, 0.418886, 1.33507) (9.205805, 0.586480, 0.958476)	57.284265	57.284265	57.501870	0.38
4	(0.870296, 3.609952, 3.851484) (3.536555, 3.964960, 4.16744)	33.501898	33.501898	33.441993	0.18
5	(0.926310, 4.564359, 5.114204) (8.102154, 0.603875, 0.617218)	13.805552	13.805552	13.417987	2.89

间复杂度为 $\mathrm{O}(\mathrm{poly}(M)\log_2 N)$, 并且可以在真实的量子计算机上运行, 而经典的 SMO 算法时间复杂度为 $\mathrm{O}(\mathrm{poly}(MN))$.

表 2 基于量子态内积的量子线路训练方案的伪代码

Table 2. The pseudo code of the IPQS-based quantum circuit solution scheme.

算法1. 基于量子态内积的量子线路训练方案

输入: 训练样本集数据.

输出: 支持向量机相关参数.

初始化训练样本数据M和数据特征数量N, 初始化目标函数和约束条件, 初始化量子线路运行环境Q_r和量子比特数量$n \leftarrow \lceil \log_2 N \rceil$.

使用SMO算法求解拉格朗日乘子向量 $(\alpha_1, \alpha_2, \cdots, \alpha_M)$;// 2.1节

while 不满足停止条件时 **do**

  for $i, j = 1 : M$ **do**

    通过启发式策略选择一对 α_i and α_j;// 2.1节

    初始化量子态;

    生成量子态制备对应的量子线路;// (23)式

    生成量子态内积对应的量子线路;// (25)式

    通过量子线路计算对应训练样本之间的内积;// (19), (26) and (27) 式

    计算 α_i^{new};// (10)式

    计算 α_j^{new};// (11)式

    计算 b_i^{new};// (12)式

    计算 b_j^{new};// (13)式

    更新α_i, α_j, b的值;

  end for

end while

输出α与b

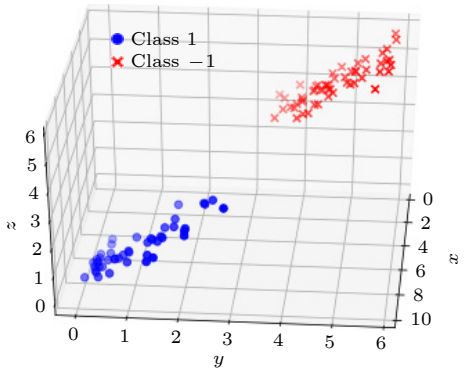


图 6 训练样本数据集

Fig. 6. The training sample data set.

4.3 时间复杂度分析

本文提出的基于量子态内积的量子线路训练方案属于量子经典混合算法, 因此时间复杂度主要由经典部分与量子部分组成, 其中经典部分的计算资源主要用于拉格朗日乘子向量的求解, 本文采用的是 SMO 算法, 每次只随机优化两个拉格朗日乘子, 根据分析这部分的时间复杂度为 $\mathrm{O}(\mathrm{poly}(M))$, 其中 M 为训练样本的数量. 量子部分的计算资源主要用于量子态的演化, 而根据文献 [36] 中的研究内容, 在本文设计的量子线路中, 通过 SWAP 量子逻辑门实现量子态演化是最消耗计算资源的部分. 在每次优化两个拉格朗日乘子过程中, 当数据的特征数量为 N时, 在量子线路中只需要$\lceil \log_2 N \rceil$个量子位就可以表示, 根据文献 [36], 分析得出这部分的时间复杂度为 $\mathrm{O}(\log_2 N)$, 其中 N 为数据的特征数量. 因此该方案总的时间复杂度为 $\mathrm{O}(\mathrm{poly}(M)\log_2 N)$. 而采用经典的 SMO 算法训练支持向量机时, 由于每次只随机优化两个拉格朗日乘子, 而在每次优化过程中, 都需要计算两个数据的特征之间的内积, 因此该算法的总的时间复杂度 $\mathrm{O}(\mathrm{poly}(MN))$, 其中 M 为训练数据的数量, N 为数据的特征数量. 经过对比分析, 可以得出本文提出的基于量子态内积的量子线路求解方案相对于经典的 SMO 算法不仅可以在真实计算机上运行, 还在时间复杂度上有

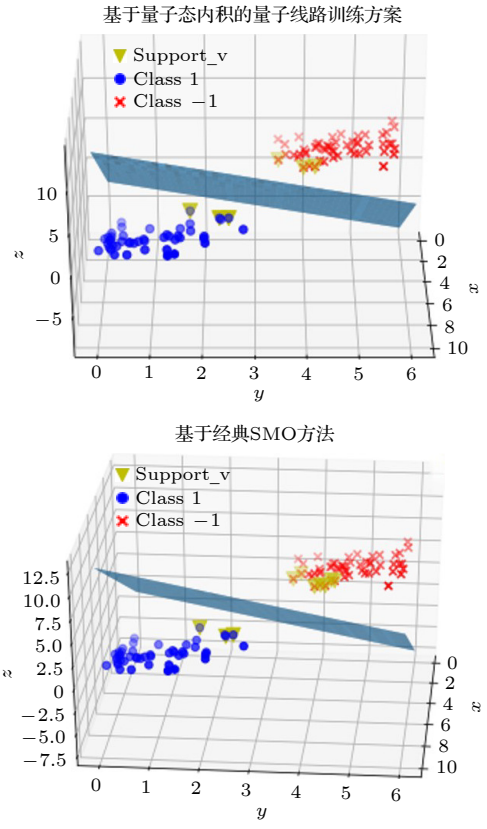


图 7 量子方案与经典方法训练效果对比

Fig. 7. Comparison of the training results between the IPQS-based quantum circuit solution scheme and the SMO-based method.

多项式级提高, 随着数据规模的提高该方案的优势相对经典的方法会更加明显.

5 结 论

本文针对训练支持向量机, 提出了一种基于量子态内积的量子线路训练方案. 该方案将问题转化为量子态表示, 通过构建量子线路, 利用量子逻辑门实现量子态的演化, 完成对问题的求解, 与经典求解方法相比, 实现了在真实量子计算机上运行效果, 进一步扩展了量子算法的应用范围. 实验结果表明, 该方案可以有效实现对支持向量机的正确训练. 与经典方法相比, 实现了多项式级加速, 大大缩短了模型的训练时间, 提升了算法的效率, 具有一定的可行性、有效性和创新性, 为训练支持向量机提供了一种新的求解思路.

本文提出的方案也有两个局限性: 1) 在该方案中, 只对涉及到经典内积部分进行了量子化转化, 量子化的范围还不够深入; 2) 该方案量子线路部分中的量子逻辑门数量和位置尚未进行优化处理.

我们未来的研究方向将集中在上述两个问题上. 主要从两个方面进行研究: 1) 加强对量子计算和经典机器学习算法的研究分析, 进一步加深量子计算与经典机器学习算法的结合深度和广度, 提高算法量子化程度的同时降低算法的计算复杂度. 2) 加强对量子线路和量子逻辑门的研究, 进一步提升对量子逻辑门的优化组合, 在降低量子逻辑门数量的同时提高量子线路的运行效率.

参考文献

[1] Cortes C, Vapnik V 1995 *Mach. Learn.* **20** 273
[2] Sebastian S, Matthias S, Christian S 2021 *ACM J. Exp. Algor.* **26** 1
[3] Vapnik V N 2000 *The Nature of Statistical Learning Theory* (New York: Springer-Verlag Press) pp25–314
[4] Jiang F, Lu Y, Chen Y, Cai D, Li G F 2020 *Comput. Electron. Agr.* **179** 105824
[5] Zhang Q, Liu J M, Tian Y 2022 *Optik* **260** 168807
[6] Okwuashi O, Ndehedehe C E, Olayinka D N, Eyoh A, Attai H 2021 *Int. J. Remote Sens.* **42** 6502
[7] Kalaiarasi G, Maheswari S 2021 *Neural Comput. Appl.* **33** 13391
[8] Liu L M, Li P, Chu M X, Cai H B 2021 *Int. J. Mach. Learn. Cyb.* **12** 2237
[9] Reshma R, Pooja S, Suresh C 2018 *Knowl. Based Syst.* **139** 64
[10] Riyazuddin Y M, Basha S M, Reddy K K 2020 *Int. J. Eng. Adv. Tech.* **9** 1336
[11] Ma W Z, Cao Y, Bao W Z, Yang B, Chen Y H 2020 *Sci. Programming-Neth* **20** 1
[12] Cheng Y W, Fu L Y, Luo P, Ye Q L, Liu F, Zhu W 2020 *Knowl. -Based Syst.* **210** 106488
[13] Deen A J, Gyanchandani M 2020 *Int. J. Adv. Comput. Sci. Appl.* **11** 187
[14] Kumaresan T, Palanisamy C 2017 *Int. J. Bio-Inspire. Com.* **9** 142
[15] Chen Y 2020 *Comput. Commun.* **154** 278
[16] Lu Y L, Li J T, Yang Z H, Ou X F, Xie W W 2021 *J. Comput. Methods Sci.* **21** 435
[17] Tukan M, Baykal C, Feldman D, Rus D 2021 *Theor. Comput. Sci.* **890** 171
[18] Zhao J 2021 *J. Phys. Conf. Ser.* **1748** 052006
[19] Mangasarian O L, Wild E W 2006 *IEEE T. Pattern Anal.* **28** 69
[20] Zhang L, Suganthan P N 2015 *IEEE T. Cybernetics* **45** 2165
[21] Xu J, Tang Y Y, Zou B, Xu Z B, Li L Q, Lu Y, Zhang B C 2015 *IEEE T. Cybernetics* **45** 1169
[22] Zou B, Xu C, Lu Y, Tang Y Y, Xu J, You X G 2018 *IEEE T. Neur. Net. Lear.* **29** 1328
[23] Wu X H, Shi Z F, Xing H H, Xue Y S 2022 *MATEC Web Conf.* **355** 03059
[24] Arute F, Arya K, Babbush R, et al. 2019 *Nature* **574** 505
[25] Zhong H S, Wang H, Deng Y H, Chen M C, Peng L C, Luo Y H, Qin J, Wu D, Ding X, Hu Y, Hu P, Yang Y X, Zhang W J, Li H, Li Y X, Jiang X, Gan L, Yang G W, You L X, Wang Z, Li L, Liu N L, Lu C Y, Pan J W 2020 *Science* **370** 1460
[26] Zhua Q L, Cao S R, Chen F S, et al. 2022 *Sci. Bull.* **67** 240
[27] Deshpande A, Mehta A, Vincent T, Quesada N, Hinsche M, Ioannou M, Madsen L, Lavoie J, Qi H Y, Eisert J, Hangleiter D, Fefferman B, Dhand I 2022 *Sci. Adv.* **8** eabi7894
[28] Yuan X 2020 *Science* **369** 1054
[29] Zhang Y J, Mu X D, Liu X W, Wang X Y, Zhang X, Li K, Wu T Y, Zhao D, Dong C 2022 *Appl. Soft Comput.* **118** 108554
[30] Fan J N, Wu S X, Yu C S 2021 *Quantum Inf. Process.* **20** 9
[31] Huang S, Yin H L, Chen Z B, Wu S J 2022 arXiv: 2203.12884 v1
[32] Booth K E C, O'Gorman B, Marshall J, Hadfield S, Rieffel E 2021 *Quantum* **5** 550
[33] Rujuta V, Nagraj D, Rajesh K, Akash S 2021 *Knowl. Based Syst.* **219** 106859
[34] Chen J W, Qi X M, Chen L F, Chen F L, Cheng G H 2020 *Knowl. Based Syst.* **203** 106167
[35] Lloyd S, Mohseni M, Rebentrost P 2013 arXiv: 1307.0411 v1
[36] Rebentrost P, Mohseni M, Lloyd S 2014 *Phys. Rev. Lett.* **113** 130503
[37] Boser B E, Guyou I M, Vapnik V N 1992 *Proceedings of the 5th Annual Workshop on Computational Learning Theory* Pittsburgh Pennsylvania, USA, July 1–2, 1992 p144
[38] Osuna E, Frenud R, Girosi F 1997 *Proceedings of the 1997 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing VII* Amelia Island, FL, USA, September 24–26, 1997 p276
[39] Platt J C 1998 *Advances in Kernel Methods-Support Vecotor Learning* (Cambridge MA: MIT Press) pp41–65
[40] Ralaivola L, Alché-Buc F D' 2001 *Proceedings of International Conference on Artificial Neural Networks* Vienna, Austria, August 21–25, 2001 p322

A support vector machine training scheme based on quantum circuits

Zhang Yi-Jun [1)3)] Mu Xiao-Dong [3)] Guo Le-Meng [2)] Zhang Peng [4)3)]
Zhao Dao [4)3)] Bai Wen-Hua [2)†]

1) (*Experimental Training Base, National University of Defense Technology, Xi'an* 710106, *China*)

2) (*Institute of Information and Communication, National University of Defense Technology, Wuhan* 430079, *China*)

3) (*Graduate School, Rocket Force University of Engineering, Xi'an* 710043, *China*)

4) (*Xi'an Satellite Control Center, Xi'an* 710043, *China*)

(Received 20 October 2022; revised manuscript received 19 January 2023)

Abstract

In order to improve the training efficiency of the support vector machine, a quantum circuit training scheme based on the inner product of the quantum state for the support vector machine is proposed in this work. Firstly, on the basis of the full analysis of the computational complexity of the classical support vector machine, the kernel function which is the main factor affecting the computational complexity of the algorithm is primarily analyzed. Based on quantum mechanics and quantum computing theory, the training sample elements in the kernel function are quantized to generate the corresponding quantum states. Secondly, according to the quantum states of the training sample elements, the types and quantities of the required quantum logic gates are derived and calculated, and the quantum circuit that can generate the corresponding quantum states of the training sample elements through the evolution of the quantum initial ground states and the quantum logic gates is designed. Then, in the light of the relationship between the inner product of the quantum state and the quantum logic gate SWAP, the quantum circuit is designed to complete the exchange operation of the corresponding quantum state amplitude. The inner product of the quantum state is realized by exchanging and evolving the amplitude of the quantum state in the quantum circuit. Finally, by measuring the quantum state of the controlling qubit, the inner product solution of the kernel function is obtained, and the acceleration effect of training support vector machine is realized. The verification results show that the scheme enables the support vector machine not only to complete the correct classification, but also to operate the quantum part of the scheme on the real quantum computer . Compared with the classical algorithm, the scheme reduces the time complexity of the algorithm for the polynomial degree, greatly shortens the training time of the model, and improves the efficiency of the algorithm. The scheme has certain feasibility, effectiveness and innovation, and expands the training idea of the support vector machine.

Keywords: quantum circuit, inner product, quantum state, support vector machine

PACS: 03.67.Lx, 03.65.Aa **DOI:** 10.7498/aps.72.20222003

† Corresponding author. E-mail: baiwenhua@nudt.edu.cn

物理学报 Acta Physica Sinica
Chinese Physical Society
Institute of Physics, CAS

一种基于量子线路的支持向量机训练方案

张毅军　慕晓冬　郭乐勐　张朋　赵导　白文华

A support vector machine training scheme based on quantum circuits

Zhang Yi-Jun　Mu Xiao-Dong　Guo Le-Meng　Zhang Peng　Zhao Dao　Bai Wen-Hua

引用信息 Citation: Acta Physica Sinica, 72, 070302 (2023)　DOI: 10.7498/aps.72.20222003
在线阅读 View online: https://doi.org/10.7498/aps.72.20222003
当期内容 View table of contents: http://wulixb.iphy.ac.cn

您可能感兴趣的其他文章
Articles you may be interested in

量子近似优化算法在指挥控制组织任务规划中的应用
Application of quantum approximate optimization algorithm to mission planning of command and control organization
物理学报. 2021, 70(23): 230304　https://doi.org/10.7498/aps.70.20211028

基于软件定义量子通信的自由空间量子通信信道参数自适应调整策略
Parameters adaptive adjustment strategy of quantum communication channel in free-space based on software-defined quantum communication
物理学报. 2019, 68(14): 140302　https://doi.org/10.7498/aps.68.20190462

浅光晶格中量子隧穿现象的实验观测
Experimental observation of quantum tunneling in shallow optical lattice
物理学报. 2022, 71(7): 073701　https://doi.org/10.7498/aps.71.20212038

激光诱导击穿光谱技术结合神经网络和支持向量机算法的人参产地快速识别研究
Rapid identification of ginseng origin by laser induced breakdown spectroscopy combined with neural network and support vector machine algorithm
物理学报. 2021, 70(4): 040201　https://doi.org/10.7498/aps.70.20201520

表面单分子量子态的探测和调控研究进展
Advances in detection and regulation of surface-supported molecular quantum states
物理学报. 2022, 71(6): 060701　https://doi.org/10.7498/aps.71.20212324

一维超导传输线腔晶格中的拓扑相变和拓扑量子态的调制
Modulation of topological phase transitions and topological quantum states in one-dimensional superconducting transmission line cavities lattice
物理学报. 2022, 71(19): 194203　https://doi.org/10.7498/aps.71.20220675