从经典神经网络到量子神经网络

张鹏

2025.01.10

从经典神经网络到量子神经网络

问题一：如何找到一个能够超越经典算法、展现量子潜力的现实应用场景?

找到应用场景

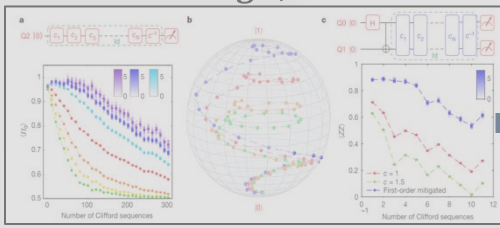
人工智能 + **量子计算**

突破算力瓶颈

问题二：如何突破人工智能的算力瓶颈?

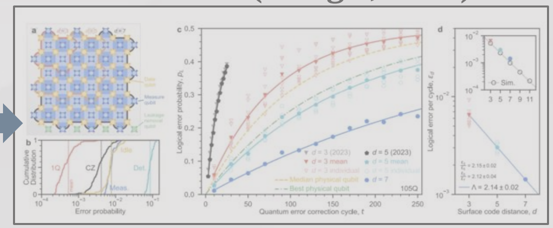
量子纠错

Google, 2019



误差率10e-2量级

Willow(Google, 2024)



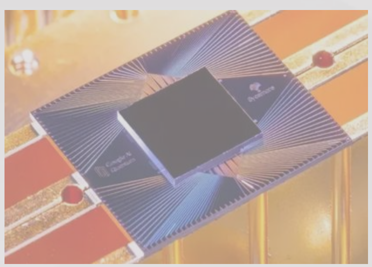
误差率10e-10量级

量子纠错率

除了加速优势，量子计算是否拥有能够短期内利用的其它优势?

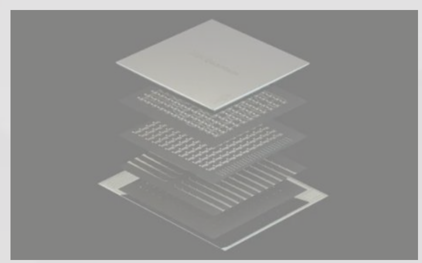
量子比特

(Google, 2019)



53量子比特

(IBM, 2023)



127量子比特

量子比特数

从经典神经网络到量子神经网络

AI for Science

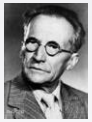
量子机器学习

量子语言模型

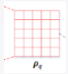
量子认知

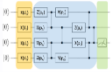
起点

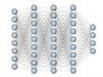
量子力学




建模工具	特点
非经典概率空间 	能够建用户**认知过程的非经典现象**[1-3]

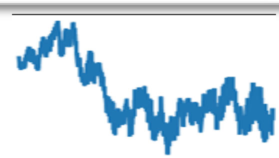
建模工具	特点
密度矩阵 	**语言的不确定**封装于密度矩阵[4-9]


建模工具	特点
量子变分线路 	利用变分量子线路的**指数级傅里叶序列拟合能力**[10-12]

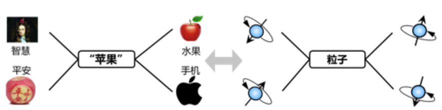

建模工具	特点
深度神经网络 	利用深度神经网络**预测蛋白质复杂结构**[13]

[1] AAAI 2010, [2] ECIR 2011, [3] IJCAI 2015, [4] CIKM 2018, [5] AAAI 2018, [6] SIGIR 2020, [7] ICLR 2020, [8] TOIS, [9] AAAI 2024, [10] ICASSP 2024, [11] NeurIPS 2019, 2022, [12] ICML 2024. [13] Nature

全世界最前沿的125个科学问题　量子人工智能可以模仿人脑吗?



√ 量子机器学习拥有
傅里叶序列拟合能力



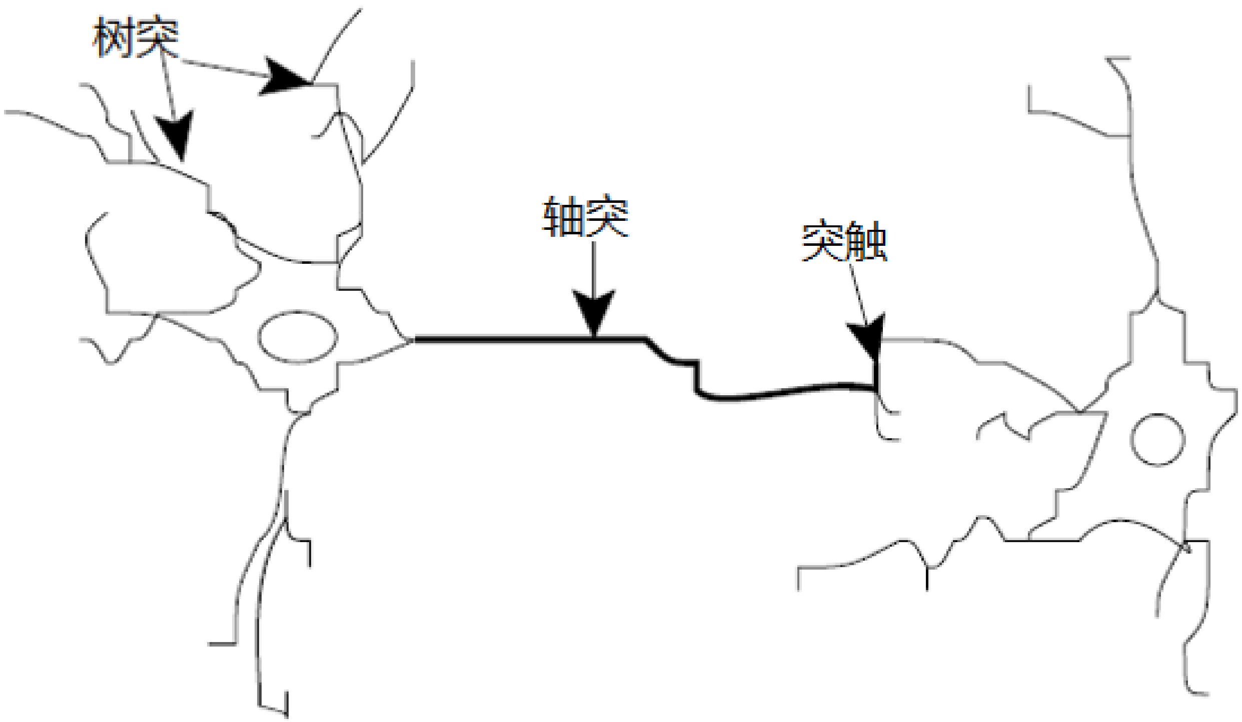
量子力学数学原理
具有建模自然语言的强大能力

量子概率能够建模
用户认知过程的非经典现象

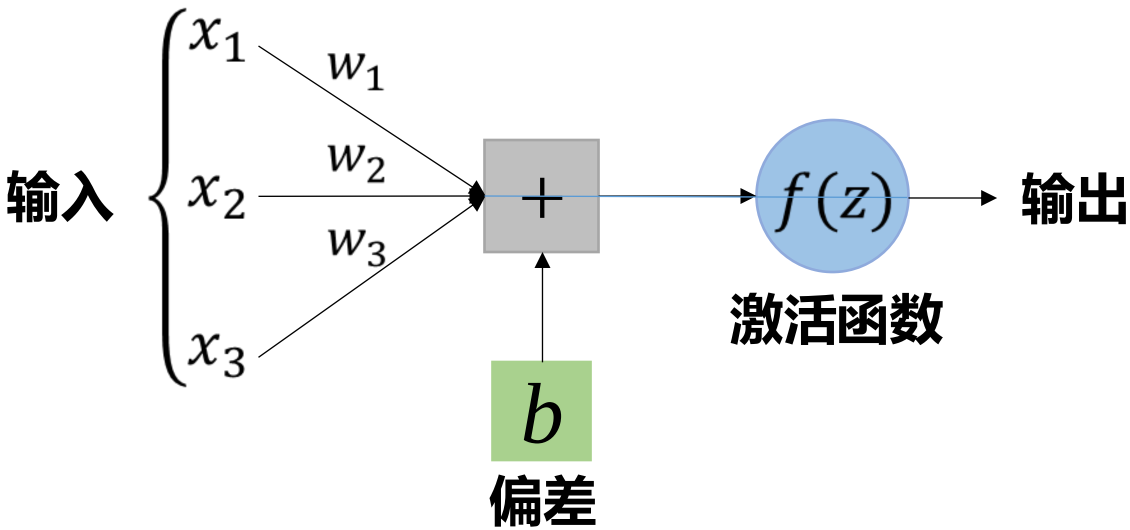
[1] Penrose R, Mermin N D. The emperor's new mind: Concerning computers, minds, and the laws of physics[J]. 1990.
[2] Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces[J]. Physical review letters, 2019, 122(4): 040504.
[3] Jerbi, Sofiene, et al. "Quantum machine learning beyond kernel methods.[J] Nature Communications 14.1 (2023): 1-8.
[4] Melko R G, Carrasquilla J. Language models for quantum simulation[J]. Nature Computational Science, 2024, 4(1): 11-18.

目录

神经元



神经元

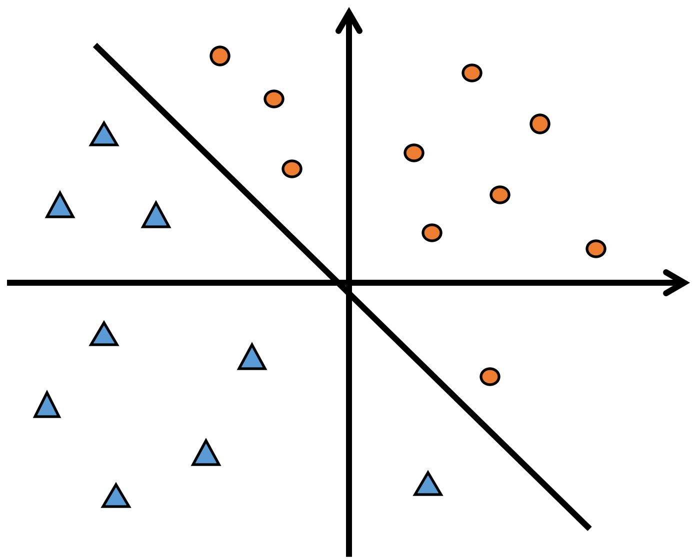


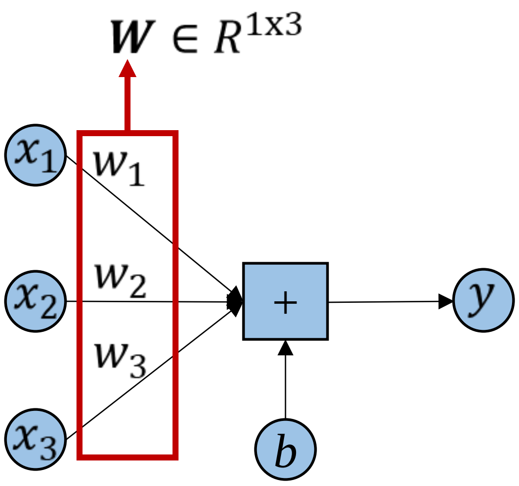
$$y = f(z)$$

$$z = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$$

单层感知器

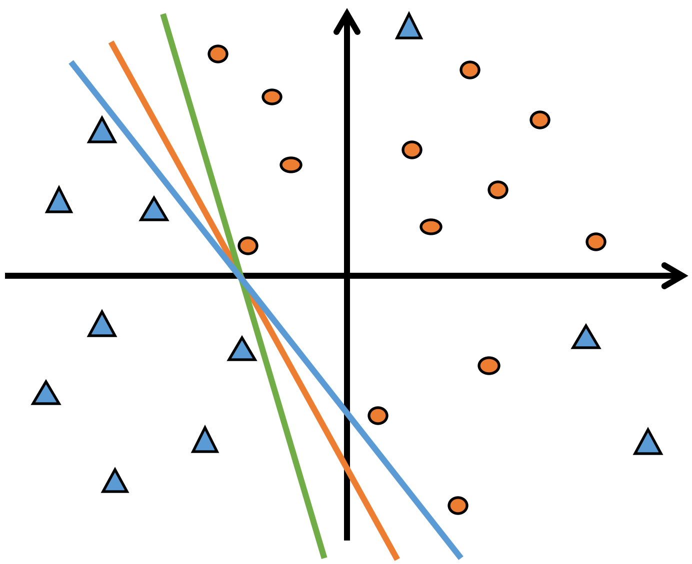
作用：解决线性可分的问题

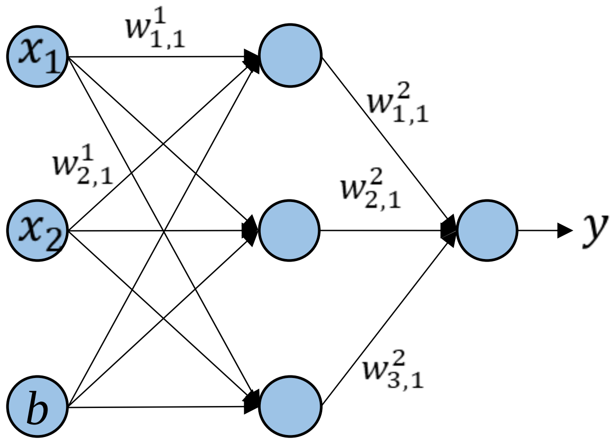




$$y = \boldsymbol{wx} + b$$
$$= w_1x_1 + w_2x_2 + w_3x_3 + b$$

多感知器

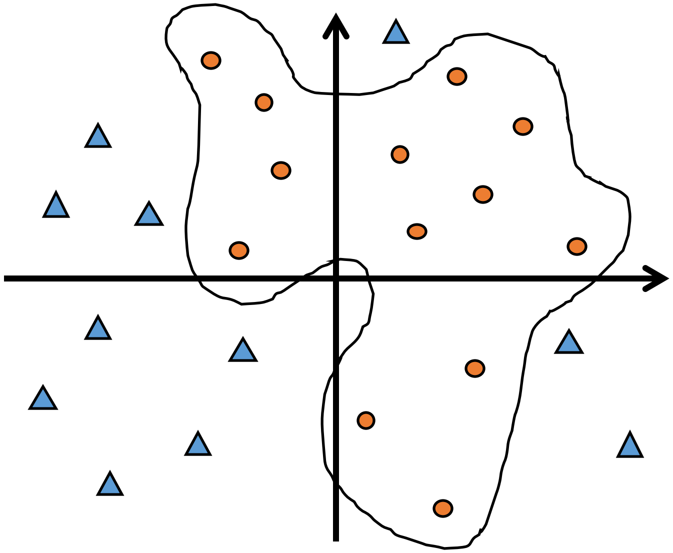




是否能够实现正确分类?

激活函数

作用：引入非线性因素，解决线性模型所不能解决的问题

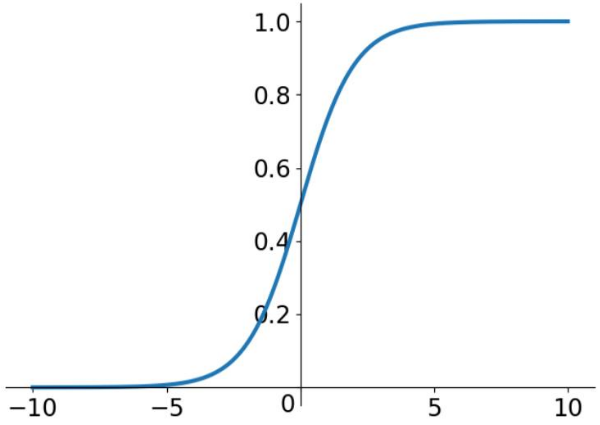


Sigmoid函数又称Logistic函数

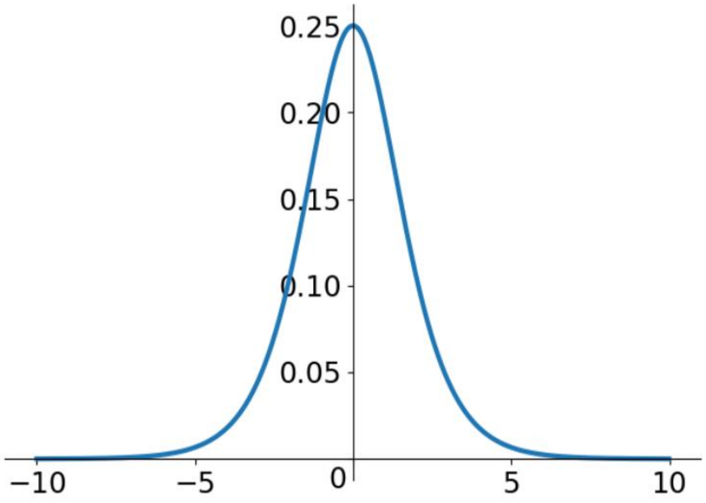
取值范围为(0,1)

公式表示为

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



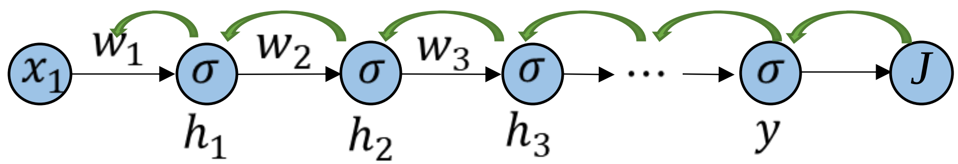
导数：$\frac{d\sigma(x)}{dx} = \frac{e^{-x}}{(1+e^{-x})^2} = \sigma(x)(1-\sigma(x))$



值域：(0, 0.25]

梯度消失

发生阶段：最小化目标函数J 时

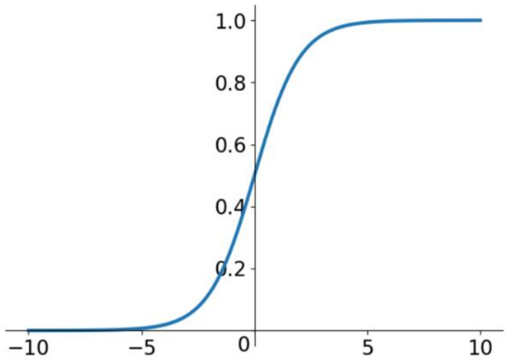


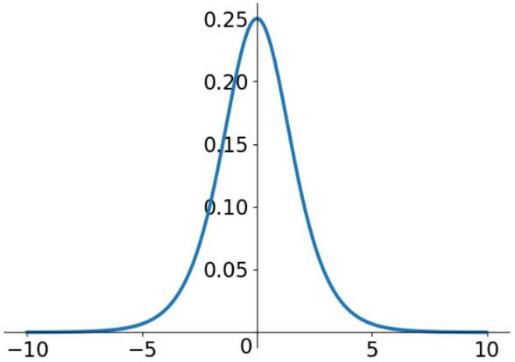
$$\frac{\partial J}{\partial w_1} = \frac{\partial J}{\partial y} \boxed{\frac{\partial y}{\partial h_n} \frac{\partial h_n}{\partial h_{n-1}} \cdots \frac{\partial h_1}{\partial w_1}}$$

原因：多个偏导连乘

结果：$\frac{\partial J}{\partial w_1}$趋近于0

Sigmoid函数





优点：连续函数，便于求导

缺点：容易产生梯度消失问题

偏导为指数函数，复杂度高

tanh函数

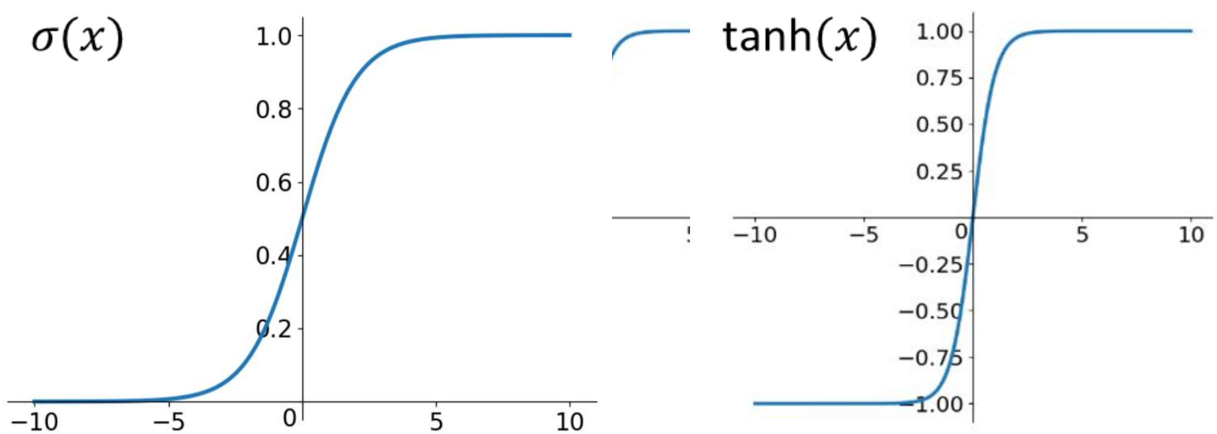
tanh函数也称双曲正切函数

取值范围为 (-1,1)

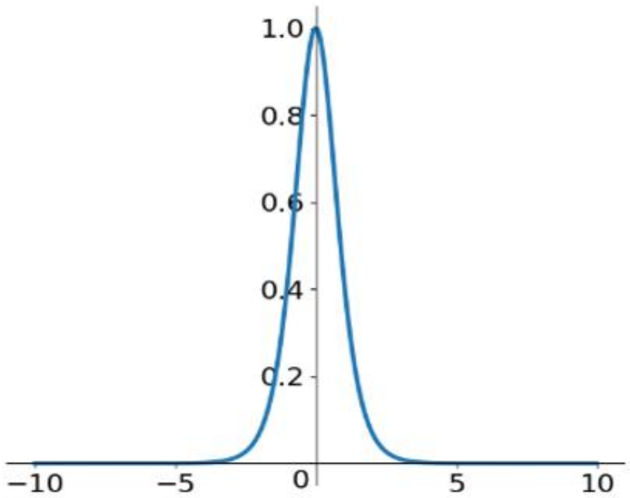
公式表示为

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\tanh(x) = 2\sigma(2x) - 1 = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

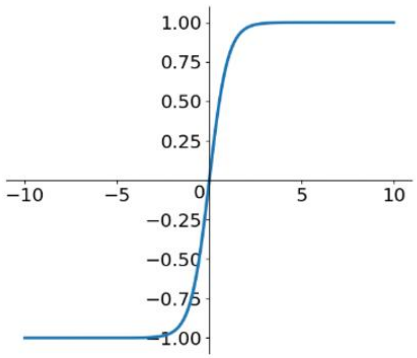


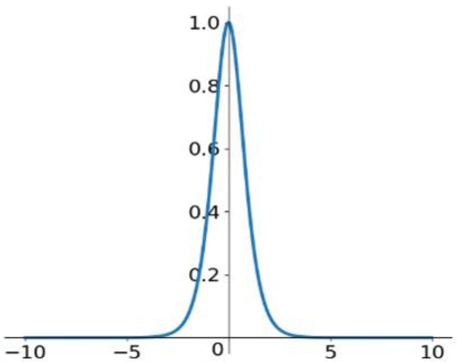
导数：$\frac{d tanh(x)}{dx} = 1 - (\tanh(x))^2$



值域：(0, 1]

tanh函数





与sigmoid函数相比

优点：函数的导数取值变大，有助于缓解梯度消失

缺点：仍存在梯度消失

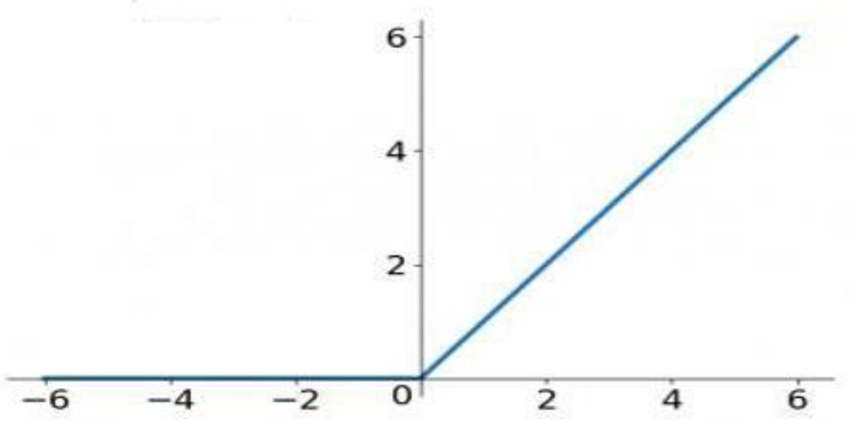
偏导为指数函数，复杂度高

ReLU函数

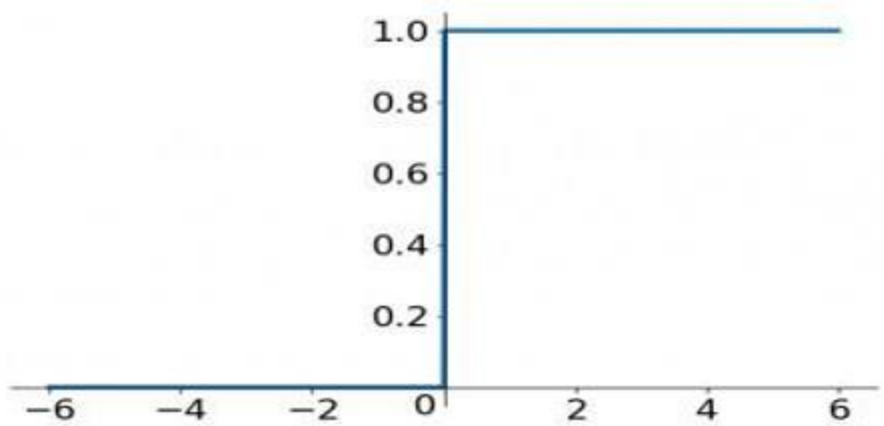
取值范围为 $[\mathbf{0}, +\infty)$

公式表示为

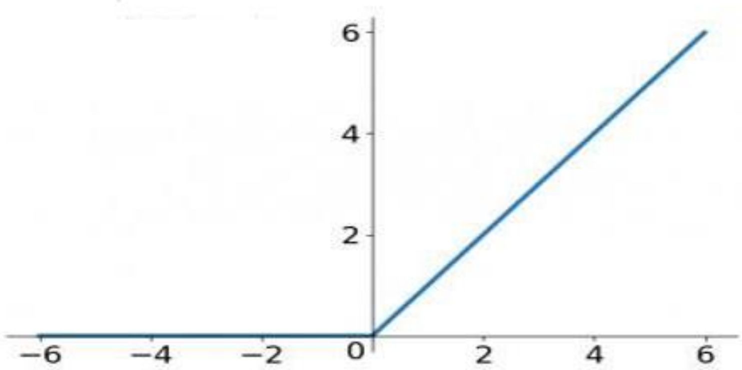
$$ReLU(x) = \max(0, x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$$

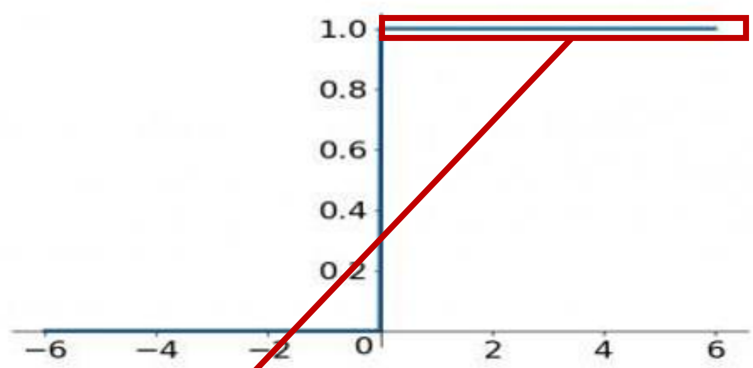


导数：$\frac{dReLU(x)}{dx}=\begin{cases}0 & x<0\\1 & x\geq 0\end{cases}$



ReLU函数





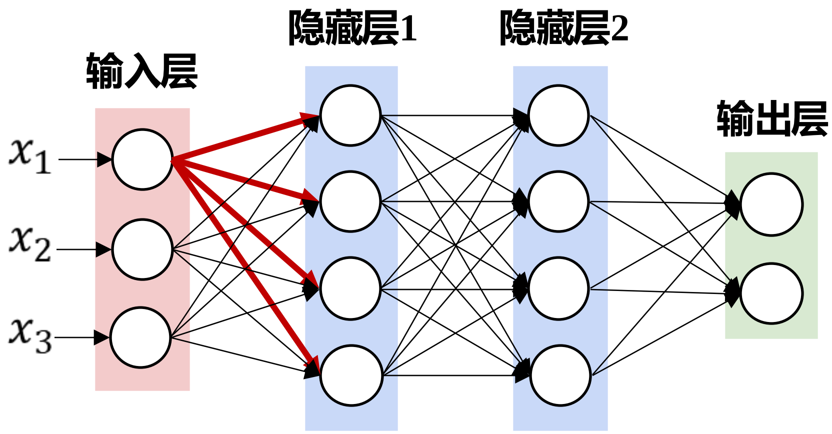
优点：有效缓和了梯度消失的问题

计算复杂度低

缺点：神经元”坏死”现象

x<0时，函数输出为0，函数求偏导为0，权重永远不会被更新

前馈神经网络



输入层：$\boldsymbol{x} \in R^{3x1}$

权重矩阵：$\boldsymbol{W}^{\mathbf{1}} \in R^{4x3}$

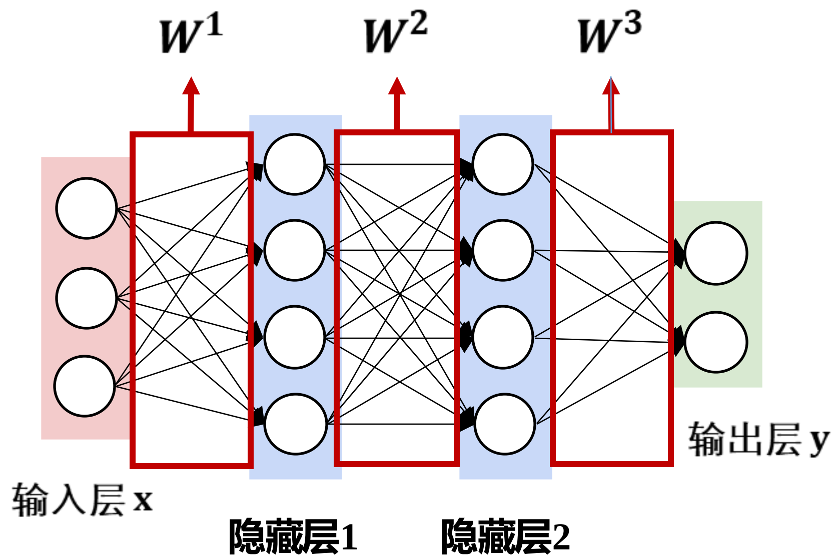
偏差：$\boldsymbol{b}^{\mathbf{1}} \in R^{4x1}$

隐藏层1：$\boldsymbol{h}^{\mathbf{1}} \in R^{4x1}$

$$\boldsymbol{h}^{\mathbf{1}} = f(\boldsymbol{W}^{\mathbf{1}}\boldsymbol{x} + \boldsymbol{b}^{1})$$

$$= f\left(\begin{bmatrix} \boxed{\begin{matrix} w_{1,1} \\ w_{1,2} \\ w_{1,3} \\ w_{1,4} \end{matrix}} & \begin{matrix} w_{2,1} \\ w_{2,2} \\ w_{2,3} \\ w_{2,4} \end{matrix} & \begin{matrix} w_{3,1} \\ w_{3,2} \\ w_{3,3} \\ w_{3,4} \end{matrix} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}\right)$$

$\boldsymbol{x}_{\mathbf{1}}$的所有权重

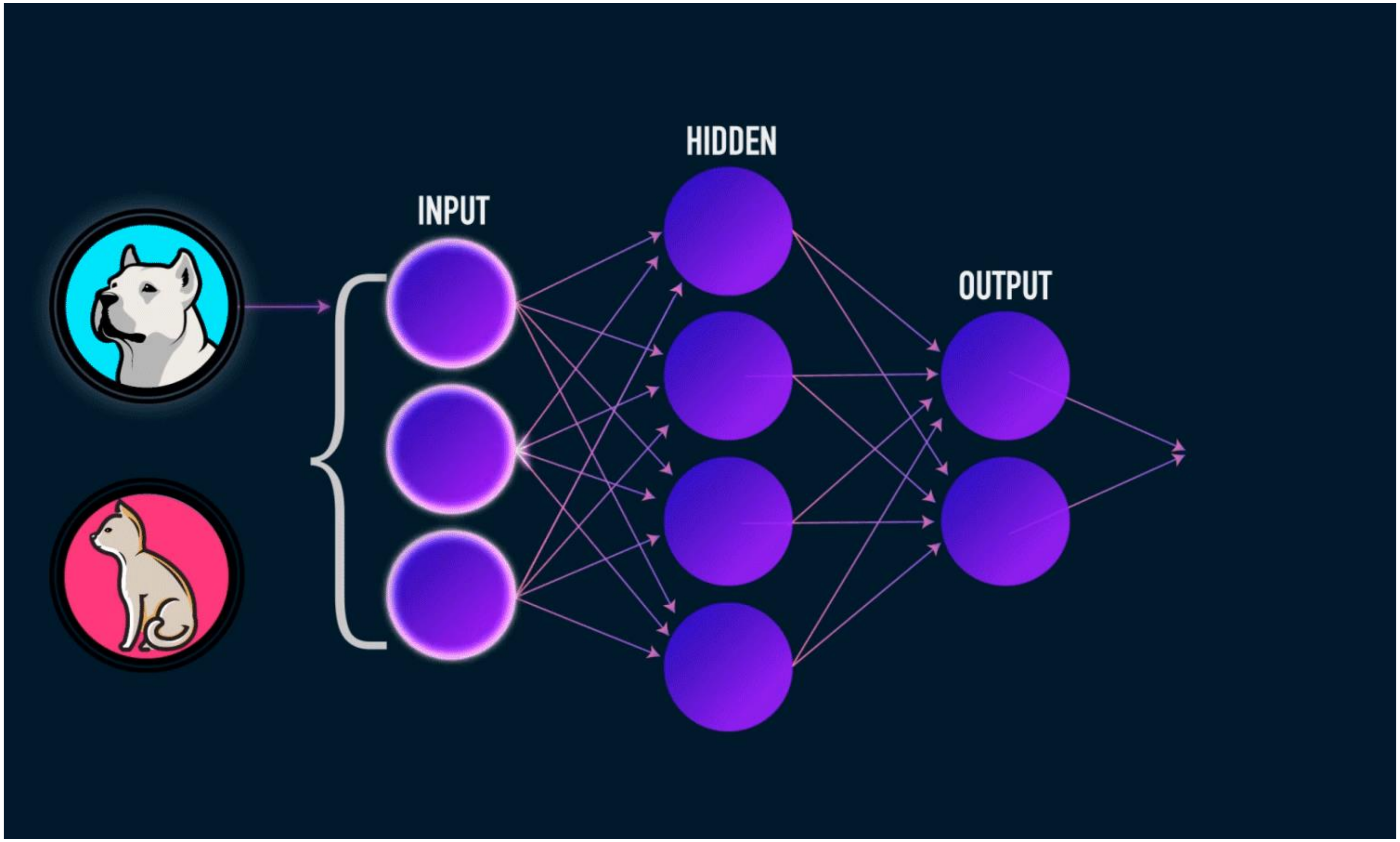


$$\boldsymbol{y} = \boxed{f^3}(\boldsymbol{W}^3\left(f^2\left(\boldsymbol{W}^2\left(f^1(\boldsymbol{W}^1\boldsymbol{x} + \boldsymbol{b}^1)\right) + \boldsymbol{b}^2\right)\right) + \boldsymbol{b}^3)$$

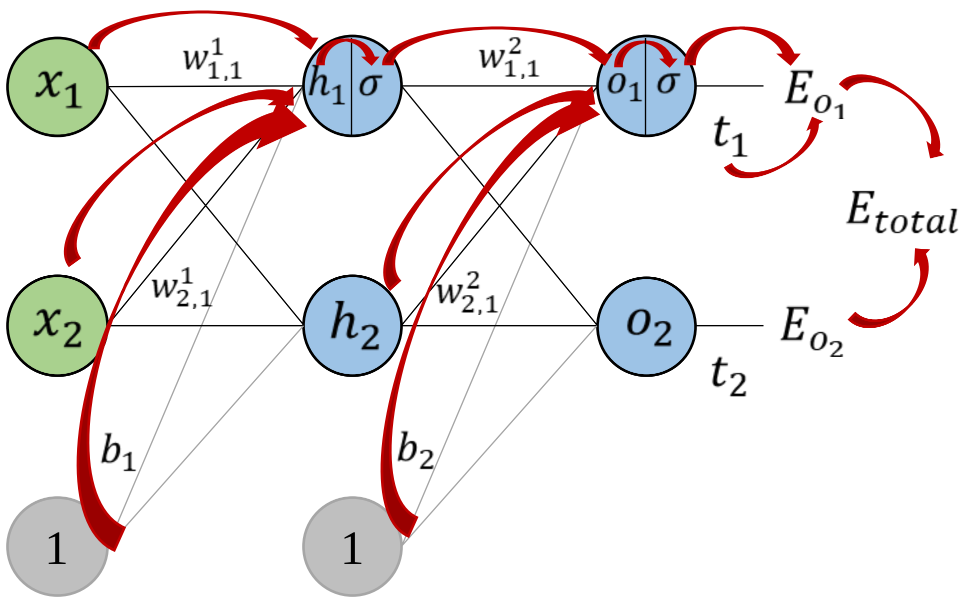
$$softmax(y_i) = \frac{e^{y_i}}{\sum_{j=1}^{N} e^{y_j}}, \quad y_i, y_j \in \mathbf{y}$$

$$\mathbf{x} \in R^{3x1}, \boldsymbol{W}^1 \in R^{4x3}, \boldsymbol{b}^1 \in R^{4x1}, \boldsymbol{W}^2 \in R^{4x4}$$
$$\boldsymbol{b}^2 \in R^{4x1}, \boldsymbol{W}^3 \in R^{2x4}, \boldsymbol{b}^2 \in R^{2x1}$$

前馈神经网络



前向传播



$$h_1 = w^1_{1,1}x_1 + w^1_{2,1}x_2 + b_1$$

$$\sigma(h_1) = \frac{1}{1+e^{-h_1}}$$

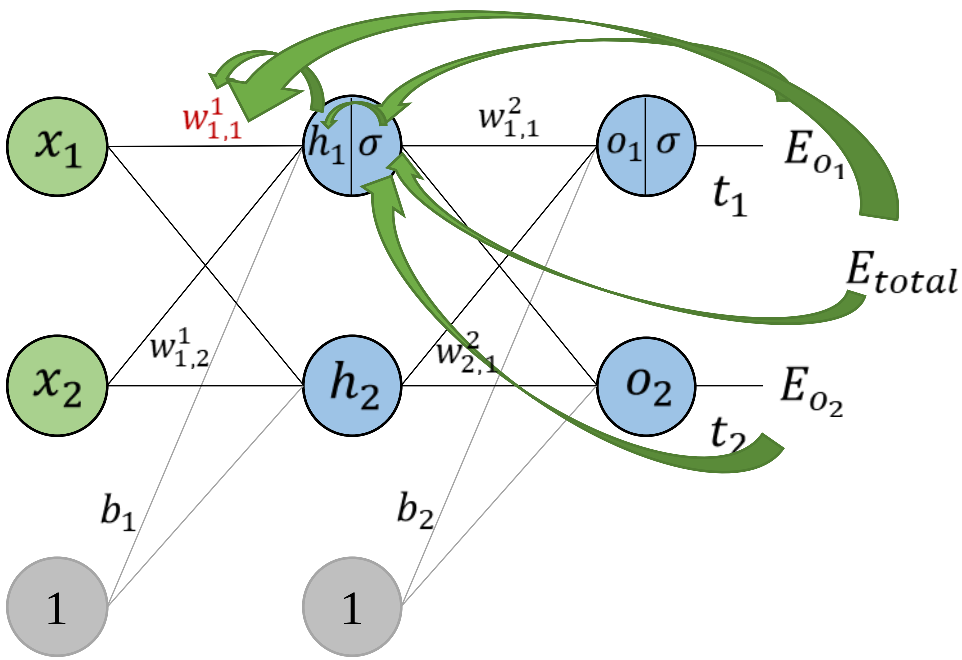
$$o_1 = w^2_{1,1}\sigma(h_1) + w^2_{2,1}\sigma(h_2) + b_2$$

$$\sigma(o_1) = \frac{1}{1+e^{-o_1}}$$

$$E_{o_1} = \frac{1}{2}(t_1 - \sigma(o_1))^2$$

$$E_{\text{total}} = E_{o_1} + E_{o_2}$$

反向传播



$$\frac{\partial E_{o_1}}{\partial \sigma(h_1)} = \frac{\partial E_{o_1}}{\partial \sigma(o_1)} \frac{\partial \sigma(o_1)}{\partial o_1} \frac{\partial o_1}{\partial \sigma(h_1)}$$

$$\frac{\partial E_{o_2}}{\partial \sigma(h_1)} = \frac{\partial E_{o_2}}{\partial \sigma(o_2)} \frac{\partial \sigma(o_2)}{\partial o_2} \frac{\partial o_2}{\partial \sigma(h_1)}$$

$$\frac{\partial E_{total}}{\partial \sigma(h_1)} = \frac{\partial E_{o_1}}{\partial \sigma(h_1)} + \frac{\partial E_{o_2}}{\partial \sigma(h_1)}$$

$$\frac{\partial \sigma(h_1)}{\partial h_1} = \sigma(h_1)(1 - \sigma(h_1))$$

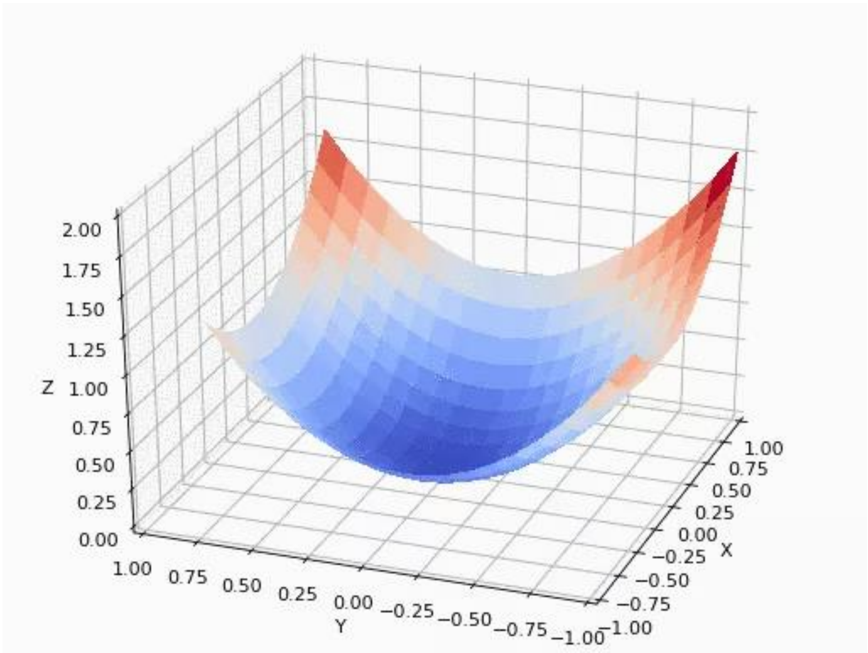
$$\frac{\partial h_1}{\partial w^1_{1,1}} = x_1$$

$$\boxed{\frac{\partial E_{total}}{\partial w^1_{1,1}} = \frac{\partial E_{total}}{\partial \sigma(h_1)} \frac{\partial \sigma(h_1)}{\partial h_1} \frac{\partial h_1}{\partial w^1_{1,1}}}$$

梯度下降

公式表示： $w_{1,1}^{1}{}^{+} = w_{1,1}^{1} - \eta \frac{\partial E_{total}}{\partial w_{1,1}^{1}}$

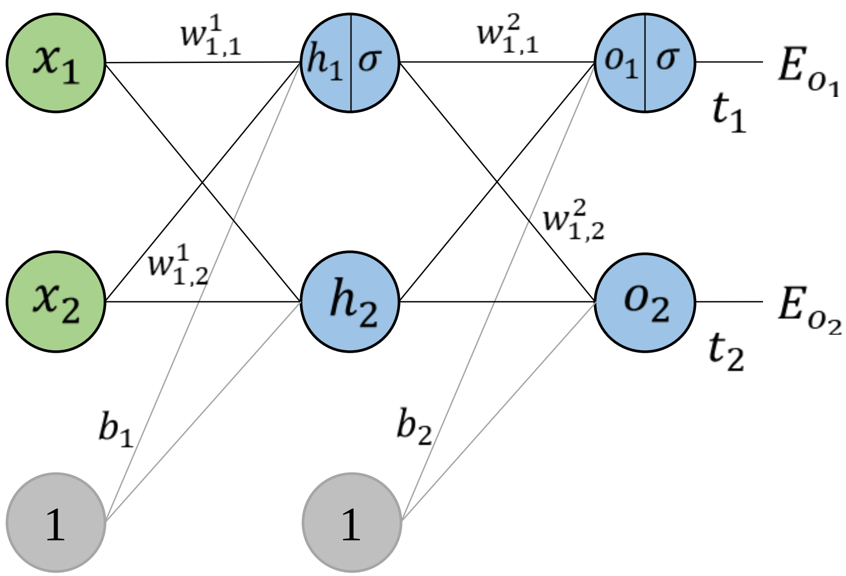
下降 **步长** **函数增长最快的方向**



权重更新

方法：梯度下降

目标：最小化目标函数



$$w^{1}_{1,1}{}^{+} = w^{1}_{1,1} - \eta \frac{\partial E_{total}}{\partial w^{1}_{1,1}}$$

$$w^{1}_{1,2}{}^{+} = w^{2}_{1,2} - \eta \frac{\partial E_{total}}{\partial w^{2}_{1,2}}$$

$$\vdots$$

小结

- 神经元：生物神经元、人工神经元
- 激活函数：三种激活函数
- 前馈神经网络：网络输出的计算方式
- 反向传播：权重更新过程

目录

狄拉克符号

- **左矢和右矢**

左矢：$\langle\psi|$ （行向量） → $\langle 1| = \ (0,1)$

右矢：$|\varphi\rangle$ （列向量） → $|0\rangle = (1,0)^T$

- **内积**

$\langle\psi|\varphi\rangle$ → $\langle 1|0\rangle = 0$

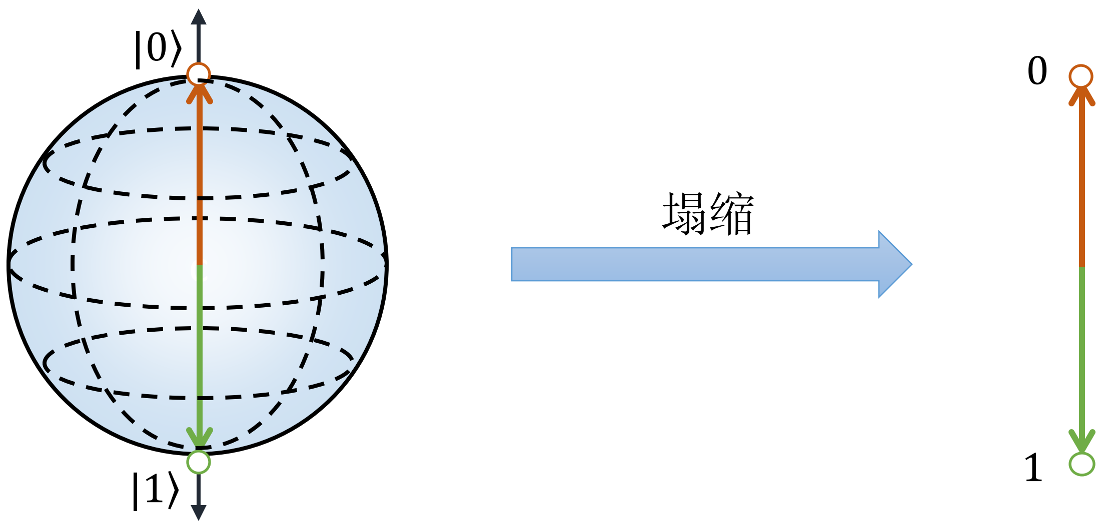
- **外积**

$|\varphi\rangle\langle\varphi|$ → $|0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

量子叠加态

一个物理系统可以由状态向量来描述，它是**系统状态空间里的一个单位向量，可以处于基向量的叠加**，比如一个2维空间：

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle，其中\ \alpha^2+\beta^2=1$$



量子比特（0和1）　　　　比特（0或1）

Nielsen, M. A., & Chuang, I. L. (2010). Quantum Computation and Quantum Information (10th Anniversary Edition). Cambridge University Press.

量子演化

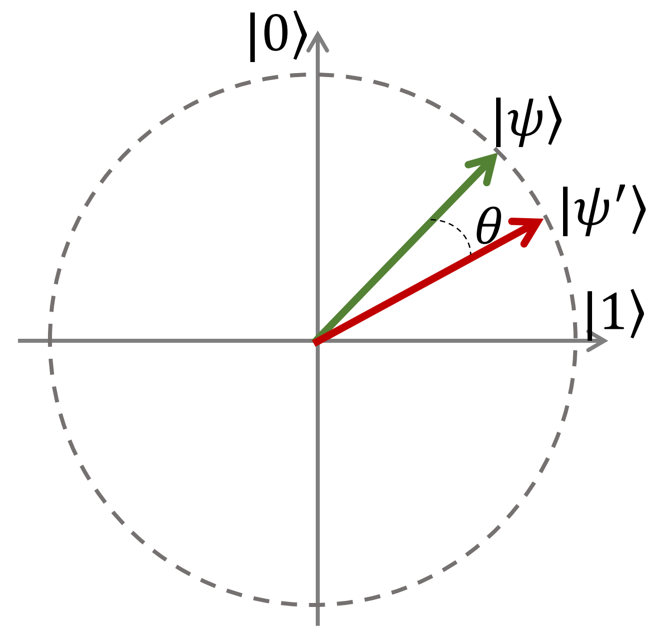
封闭量子系统的演化可用酉变换来描述，**系统在t_1处于状态ψ和在t_2变化为状态ψ'**可以通过一个仅与时间t有关的酉算子U联系起来的：

$$|\psi'\rangle = U|\psi\rangle$$

几何意义

在实数域的演化可以理解为一个向量旋转角度θ

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$



一个量子态在实数域上进行演化的几何意义

量子测量

假设当前时刻量子态演变到状态$|\psi'\rangle$，我们**将以一定概率观测到0和1**：

$$|\psi'\rangle = \alpha|0\rangle + \beta|1\rangle$$

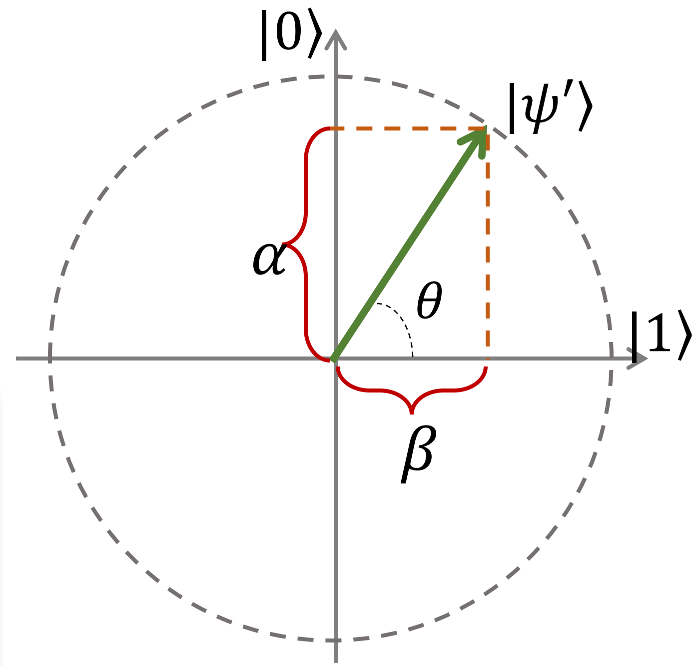
$$p(0|\psi') = |\alpha|^2 = \langle\psi'|0\rangle^2$$

$$p(1|\psi') = |\beta|^2 = \langle\psi'|1\rangle^2$$

几何意义

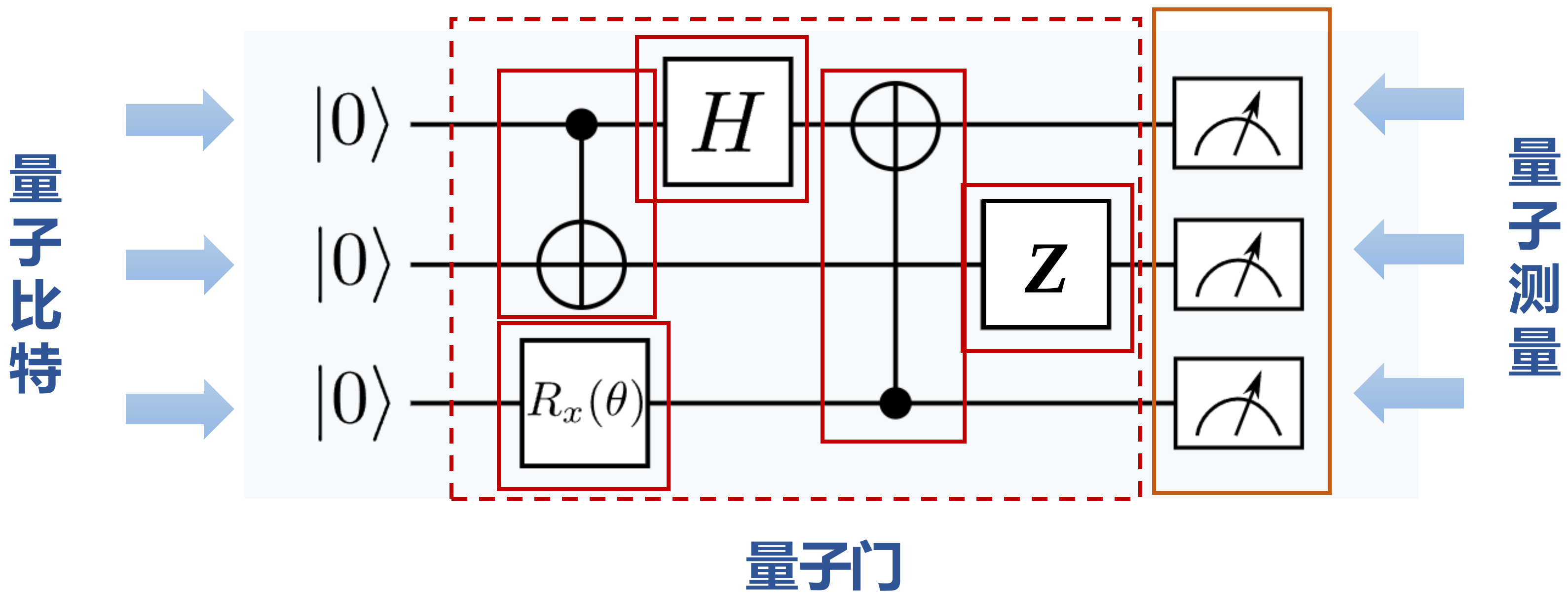
测量可以理解为量子态和基坐标的余弦值的平方

$$\beta = \cos\theta = \frac{\beta}{\sqrt{\alpha^2+\beta^2}}$$



一个量子态在实数域上进行测量的几何意义

量子线路



量子线路是一种直观地描述**量子计算过程中对量子比特执行的操作序列的方法**

量子比特

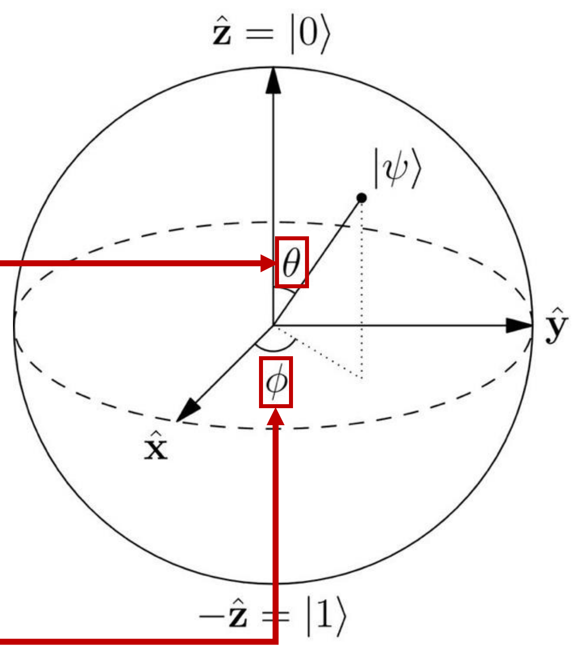
量子线路的基本单元，一个量子比特可以处于的0和1的任意叠加态

几何意义：

α 和 β是复数，可以表示为幅值与相位的形式：

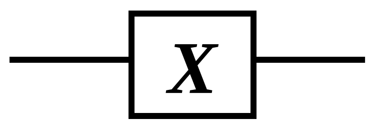
$$\alpha = |\alpha|e^{i\gamma}, \beta = |\beta|e^{i(\gamma+\varphi)}$$

$$|\psi\rangle = \cos\frac{\boxed{\theta}}{2}|0\rangle + e^{i\boxed{\varphi}}\sin\frac{\theta}{2}|1\rangle$$



量子门 量子线路的基本操作，用于改变量子比特的状态

Pauli-X门（X门）：类似于经典计算中的**NOT门**，也称为量子非门，即将量子比特的状态从$|0\rangle$翻转为$|1\rangle$，或从$|1\rangle$翻转为$|0\rangle$。

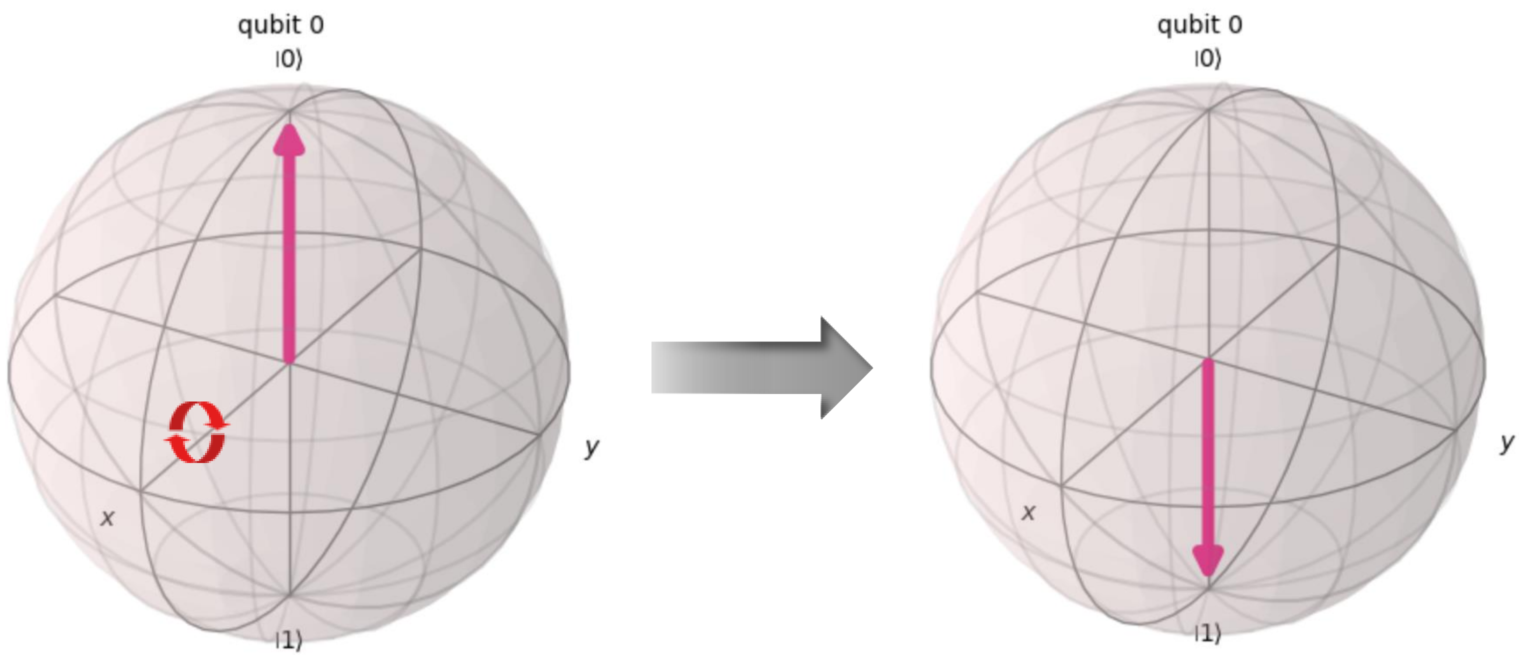


$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle \quad X=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \quad \alpha|1\rangle+\beta|0\rangle$$

例子:

$$X|\psi\rangle=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\alpha X|0\rangle+\beta X|1\rangle=\alpha|1\rangle+\beta|0\rangle=\begin{pmatrix}\beta\\ \alpha\end{pmatrix}$$

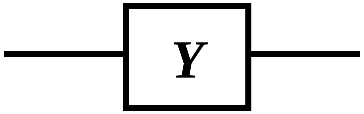
Pauli-X门表示量子态围绕布洛赫球**x轴**旋转180°



$|\psi\rangle = |0\rangle$

$|\psi'\rangle = |1\rangle$

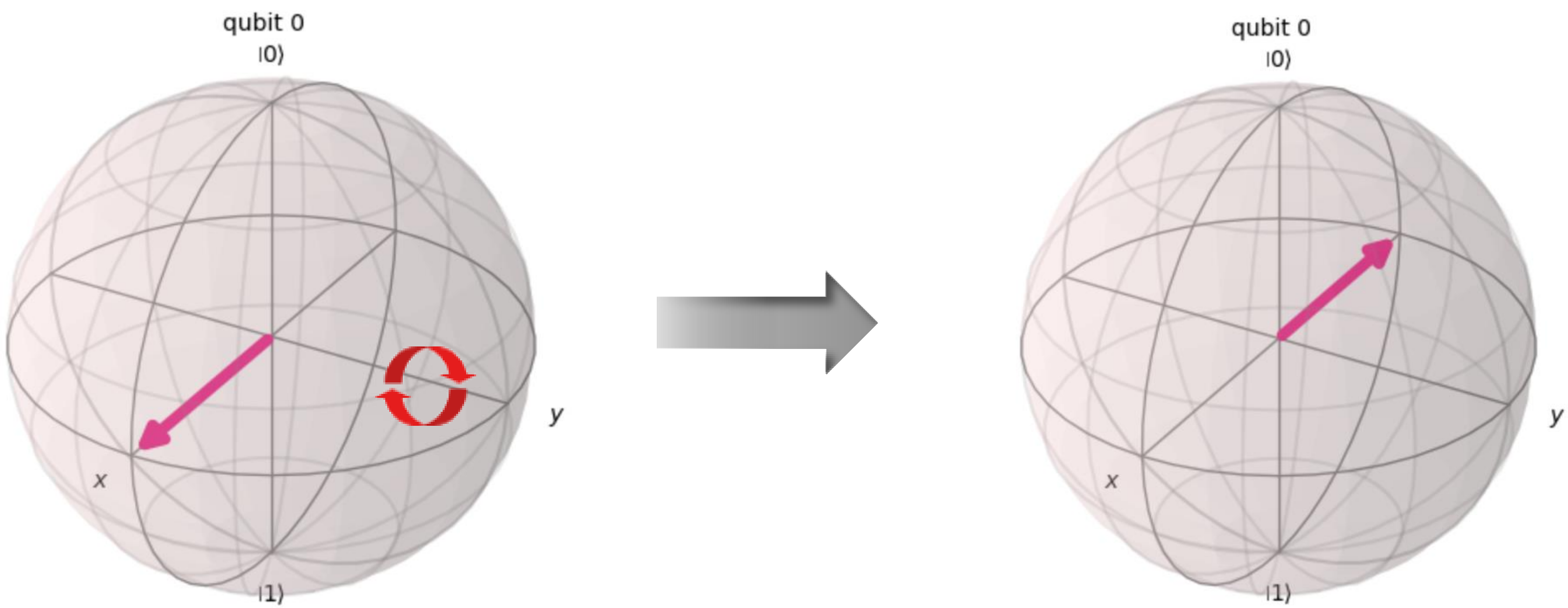
Pauli-Y门（Y门）：量子态围绕布洛赫球**y轴**旋转180°



$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle \quad \Rightarrow \quad Y=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \quad \Rightarrow \quad -i\alpha|1\rangle+i\beta|0\rangle$$

例子:

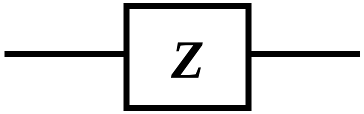
$$Y|\psi\rangle=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}-i\beta\\ i\alpha\end{pmatrix}$$



$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

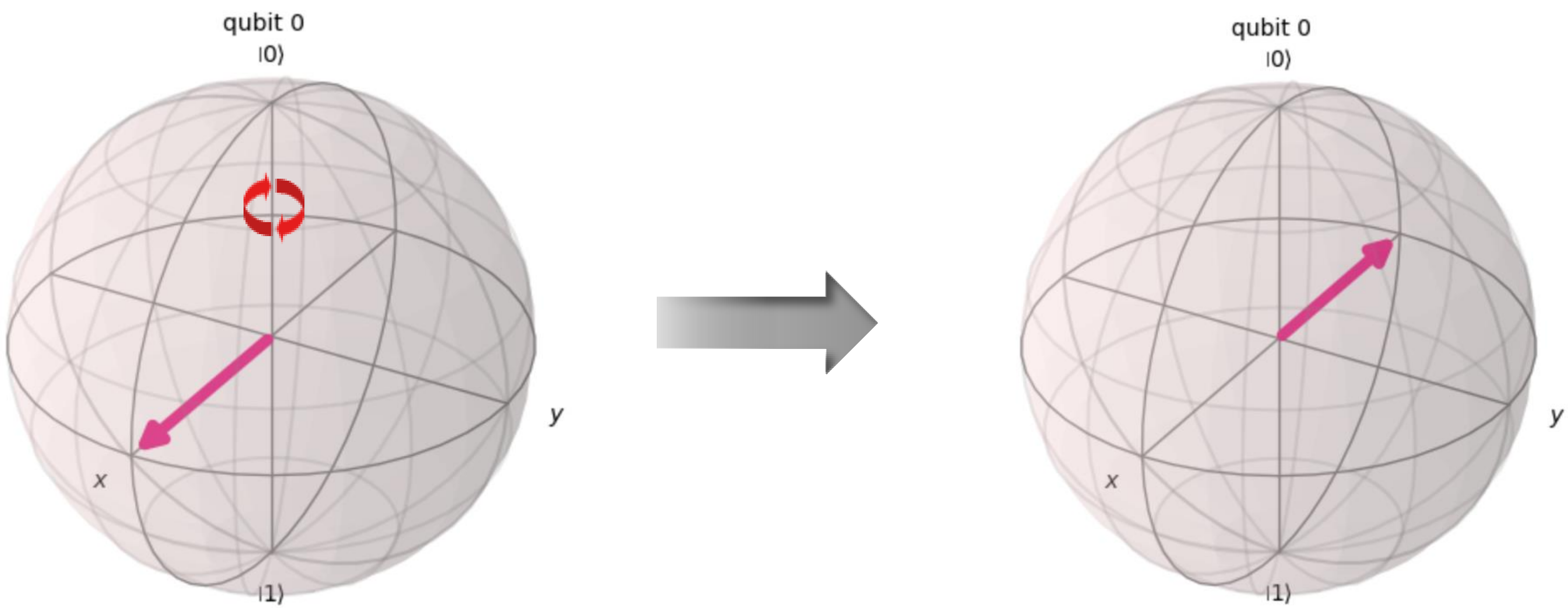
Pauli-Z门（Z门）：量子态围绕布洛赫球**z轴**旋转180°



$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle \quad Z=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \quad \alpha|0\rangle-\beta|1\rangle$$

例子:

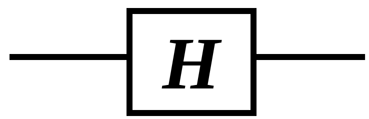
$$Z|\psi\rangle=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}\alpha\\ -\beta\end{pmatrix}$$



$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Hadamard门（H门）：创建量子态的叠加态。将一个经典基态$|0\rangle$或$|1\rangle$转换为两个基态的叠加态，使得量子比特以等概率处于$|0\rangle$和$|1\rangle$

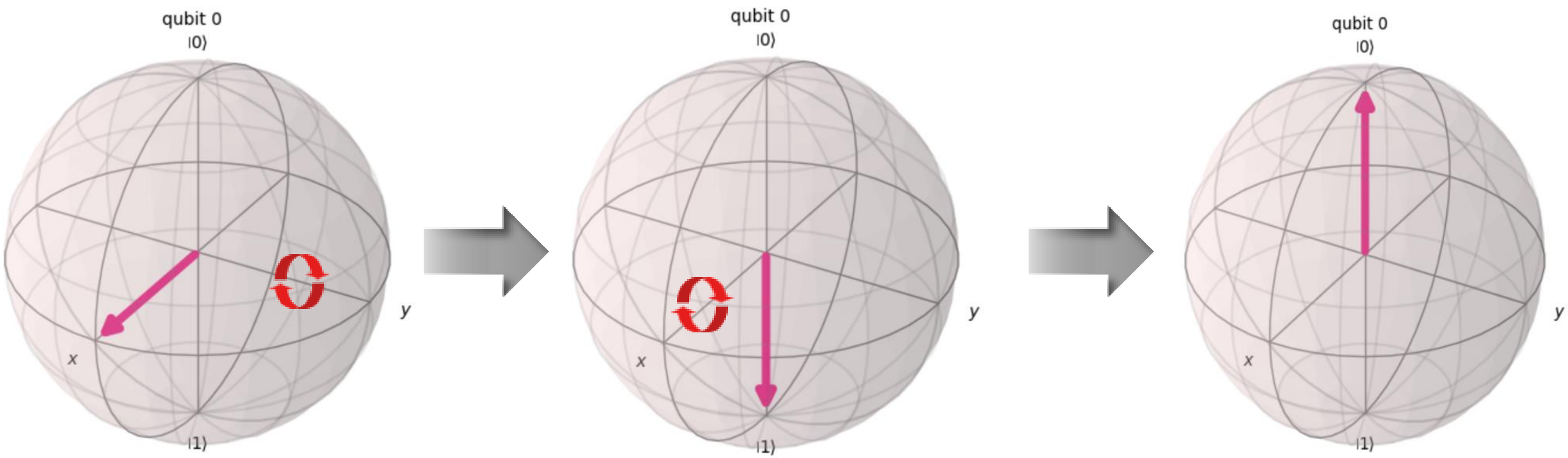


$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \Rightarrow H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \Rightarrow \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$$

例子:

$$H|\psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \frac{\alpha+\beta}{\sqrt{2}} \\ \frac{\alpha-\beta}{\sqrt{2}} \end{pmatrix}$$

在布洛赫球上，H门表示量子态先**绕y轴旋转90°**，再**绕x轴旋转180°**



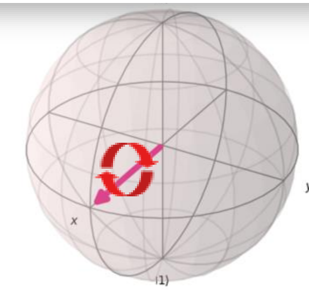
$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$|\psi'\rangle = |0\rangle$$

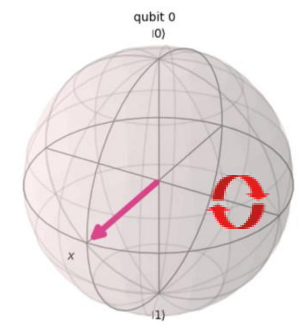
旋转门

旋转门是量子门的进一步泛化，能在x, y, z轴等方向实现旋转任意角度的旋转

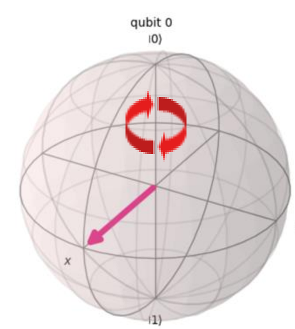
$R_x(\theta)$门表示绕x轴旋转θ角度：$R_x(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$



$R_y(\theta)$门表示绕y轴旋转θ角度：$R_y(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$



$R_z(\theta)$门表示绕z轴旋转θ角度：$R_z(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$



双量子比特门

量子纠缠使量子计算有望超越经典计算，在量子线路上可以使用双量子比特门来构建这种特性

双量子比特门的作用是将**两个量子比特的量子态**通过一个**联合操作**进行改变。

双量子比特的量子态表示为：

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$$

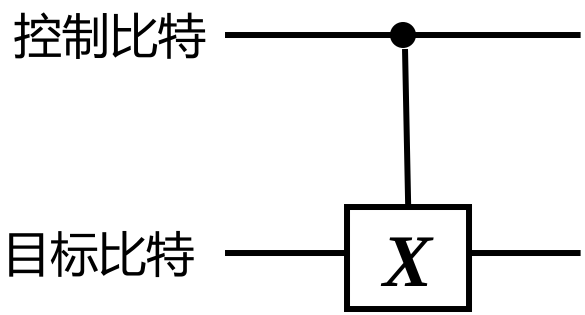
复系数c_{ij}满足归一化条件：

$$|c_{00}|^2 + |c_{01}|^2 + |c_{10}|^2 + |c_{11}|^2 = 1$$

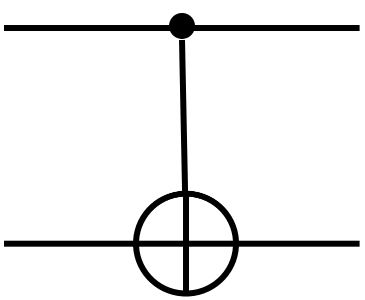
矩阵形式为：

$$|\psi\rangle = \begin{pmatrix} c_{00} \\ c_{01} \\ c_{10} \\ c_{11} \end{pmatrix}$$

CNOT（Controlled-NOT）门是一个控制门，作用于两个量子比特的组合态上，按照控制比特的状态执行操作：



或



控制比特：决定操作是否执行。

目标比特：在控制比特为$|1\rangle$时，目标比特会执行 NOT 操作。

CNOT门的**矩阵形式**是：

$$\mathrm{CNOT}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}$$

例子：

$$\mathrm{CNOT}|\psi\rangle=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\begin{pmatrix}c_{00}\\c_{01}\\\boxed{\begin{matrix}c_{10}\\c_{11}\end{matrix}}\end{pmatrix}=\begin{pmatrix}c_{00}\\c_{01}\\\boxed{\begin{matrix}c_{11}\\c_{10}\end{matrix}}\end{pmatrix}$$

量子测量

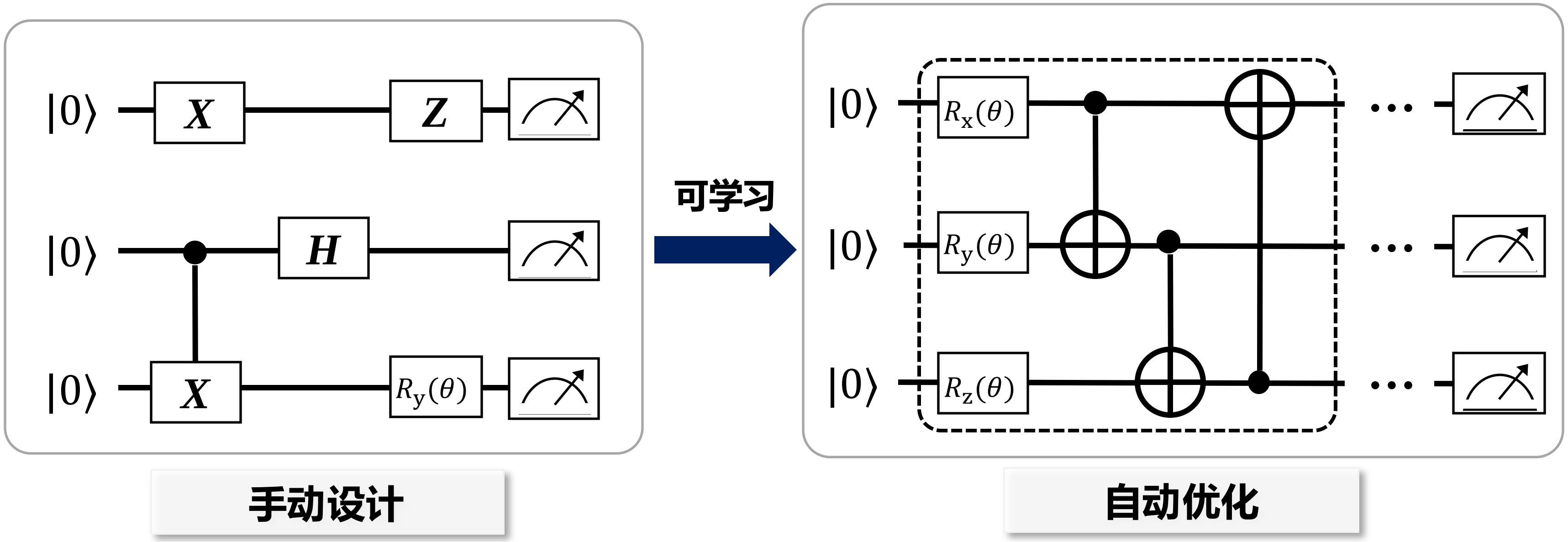
将量子比特的状态投射到经典比特，是量子线路得到最终结果的方式

在量子线路中，不确定观测结果无法得到完整信息，因此我们更关

注**量子线路测量的期望**：

$$E(M) = \sum_{m\in\{0,1\}} mp(m) = \sum_{m\in\{0,1\}} m\langle\psi|P_m|\psi\rangle$$
$$= \langle\psi|\left(\sum_m mP_m\right)|\psi\rangle$$
$$= \langle\psi|M|\psi\rangle$$

传统的量子线路需要**手动设计**，并且只能解决**特定的问题**



如何提高量子线路的灵活性和适应性
以解决更复杂的问题？

目录

探索能够真正实用的量子神经网络

输入x → 函数$f(x)$ → 输出$y*$

传统量子算法
关注加速

缺乏量子硬件
量子加速无法应用

范式的转变

量子神经网络
关注实用

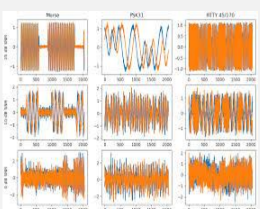
NISQ量子计算
量子模拟

**n, Tao, et al. Experimental quantum principal component analysis via parametrized quantum circuits. Physical Review Letters 126.11 (2021): 110502.

文本

像素

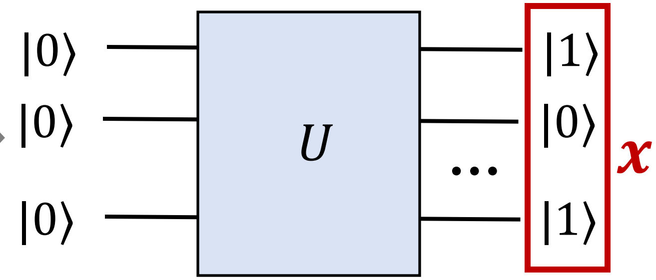
…



回归值

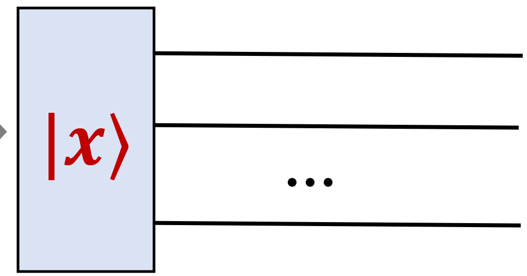
制备量子态

$$\mathbf{b} = (b_0, \ldots, b_{N-1}) \mapsto |b_0, \ldots, b_{N-1}\rangle$$



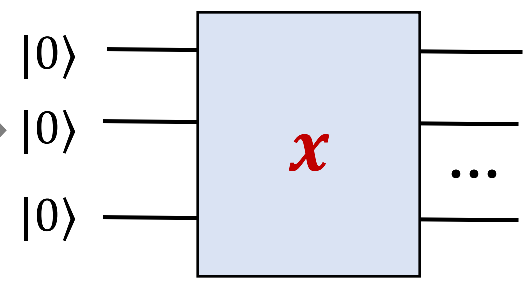
振幅编码

$$\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{2^N-1}) \mapsto \sum_{k=0}^{2^N-1} \alpha_k |k\rangle$$



角度编码

$$x \mapsto R_k(x)|0\rangle = e^{-\frac{ix\sigma_k}{2}}|0\rangle$$



制备量子态

$$\mathbf{b} = (b_0, \ldots, b_{N-1}) \mapsto |b_0, \ldots, b_{N-1}\rangle$$

容易在真实量子计算机上实现

只能编码N位的比特串

例子

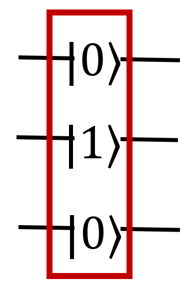
今天 天气 很好

0 1 2

由词表编号表达的文本

010

转为2进制



x

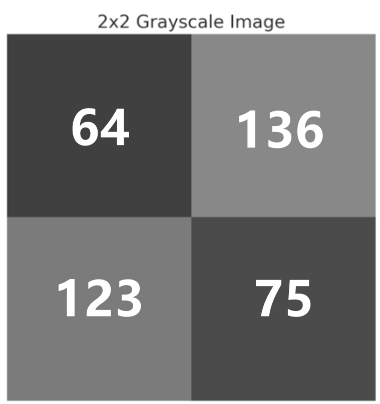
振幅编码

$$\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{2^N-1}) \mapsto \sum_{k=0}^{2^N-1} \alpha_k |k\rangle$$

可以编码2^N维的复数

难以在真实量子计算机上实现

例子



一张2×2的灰度图

归一化 $\begin{bmatrix} \alpha_0 = 0.3074 & \alpha_1 = 0.6532 \\ \alpha_2 = 0.5908 & \alpha_3 = 0.3602 \end{bmatrix}$

振幅编码

$$|\psi\rangle = \alpha_0|00\rangle + \alpha_1|01\rangle + \alpha_2|10\rangle + \alpha_3|11\rangle$$

N=2，表示有两个量子比特

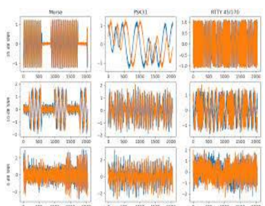
角度编码

$$x \mapsto R_k(x)|0\rangle$$

容易在真实量子计算机上实现

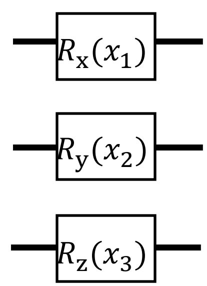
能编码N的多项式维的实数

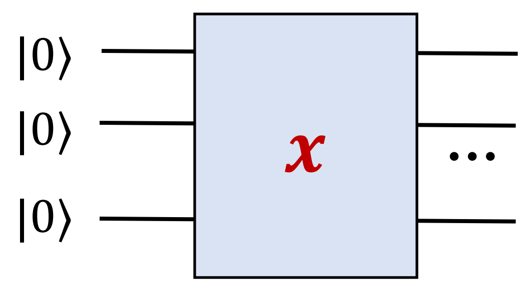
例子



$$x = [x_1, x_2, x_3]$$

3维特征值
如某3元函数





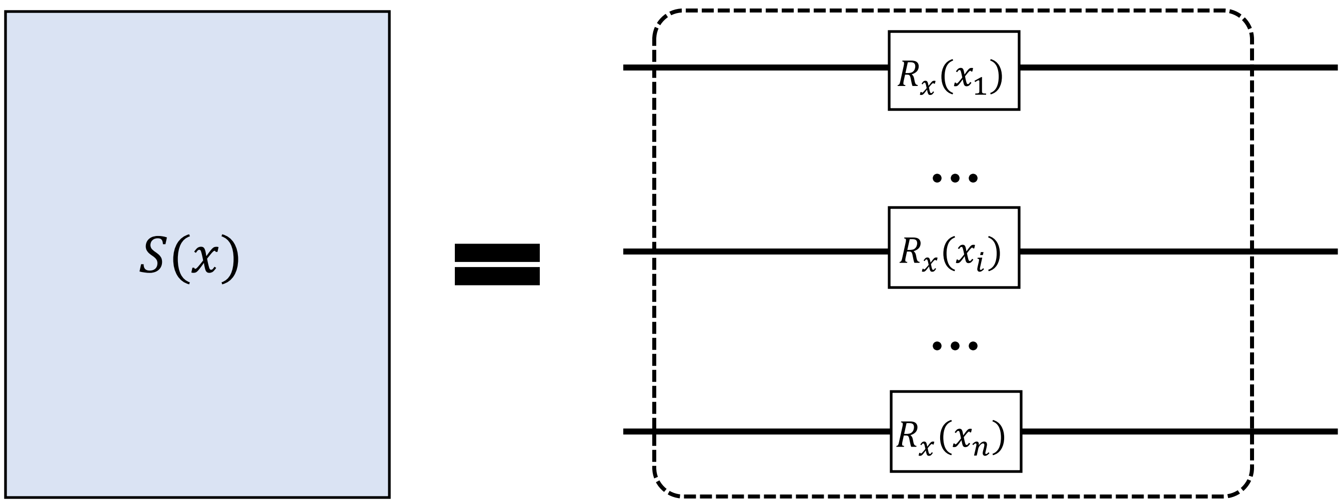
直接将特征值输入到旋转门

$$|x\rangle = R_{k_1}(x_1)|0\rangle \otimes R_{k_2}(x_2)|0\rangle \otimes R_{k_3}(x_3)|0\rangle$$

量子线路的参数化

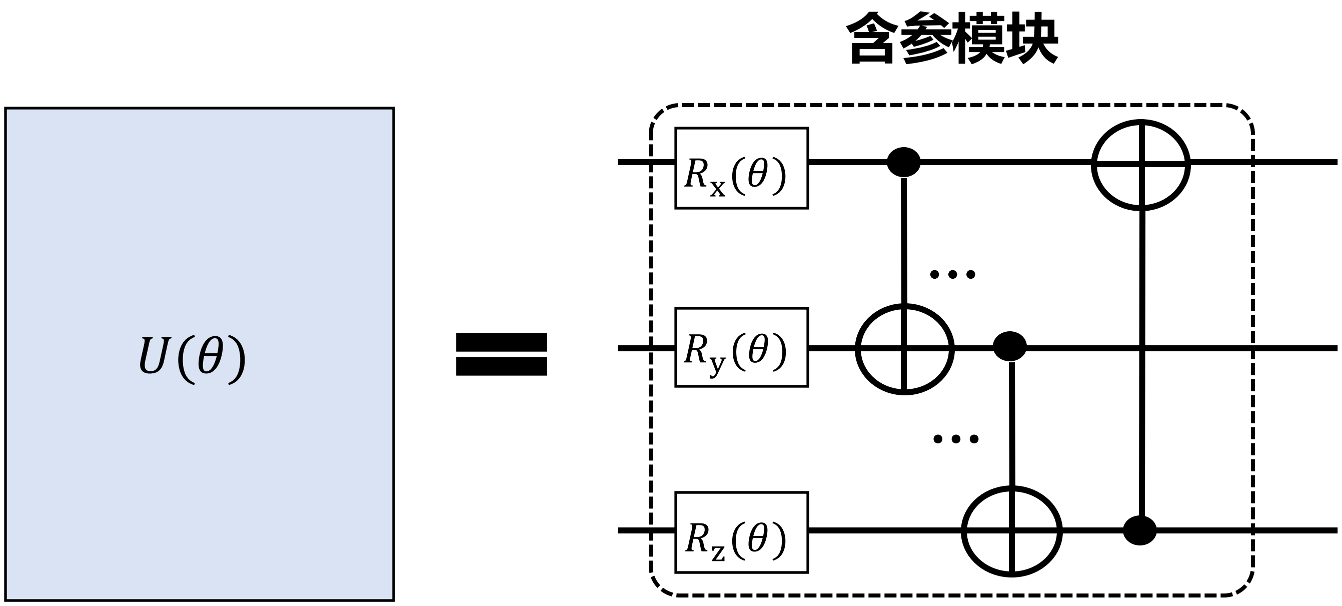
将**数据**输入旋转门，就构成了量子线路的输入

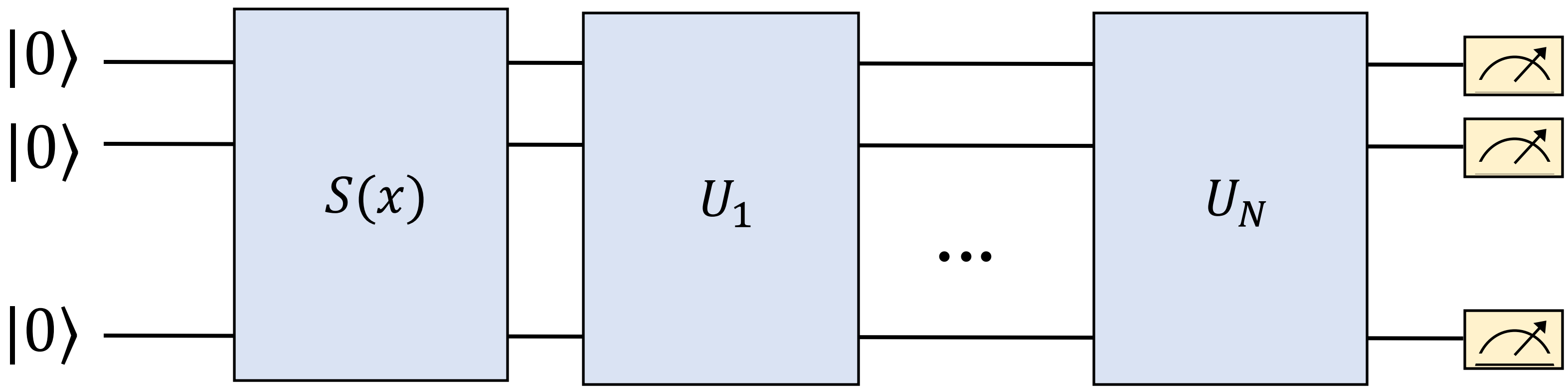
角度编码



$$S(x)|0\rangle = |x\rangle$$

将**参数**输入旋转门，就构成了量子线路的参数部分





一个完整量子线路上的操作可以表示为：

$$U(x;\theta) = U_N(\theta_N)U_{N-1}(\theta_{N-1})\cdots U_1(\theta_1)S(x)|0\rangle$$

其中，U_1, ……, U_N表示参数化的量子门。

量子线路表达了一个函数：

$$f(x;\theta_i) = \langle 0|S^\dagger(x)U_N^\dagger(\theta_N)U_1^\dagger(\theta_1)\widehat{M}U_N(\theta_N)\ldots U_1(\theta_1)S(x)|0\rangle$$

$$= \langle x|U^\dagger(\theta)\widehat{M}\underbrace{\boxed{U(\theta)}}_{\text{参数}}\underbrace{\boxed{|x\rangle}}_{\text{输入}}$$

参数位移算法

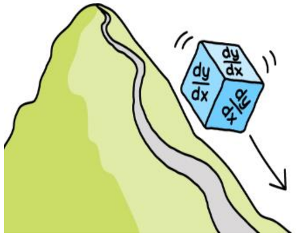
量子神经网络也可根据梯度更新参数**只需要测量一组特定的参数平移下的损失函数值**，就可以从解析上计算出偏导的值

经典数据 $\{X_i\}$ —输入→ **量子电路** $U(\theta_1,\theta_2,\dots,\theta_n)$ —期望 $\bar{x}$→ **回归、分类** $y_i = g(z_i)$ → **计算损失函数**

更新 $\theta_1,\theta_2,\dots,\theta_n$

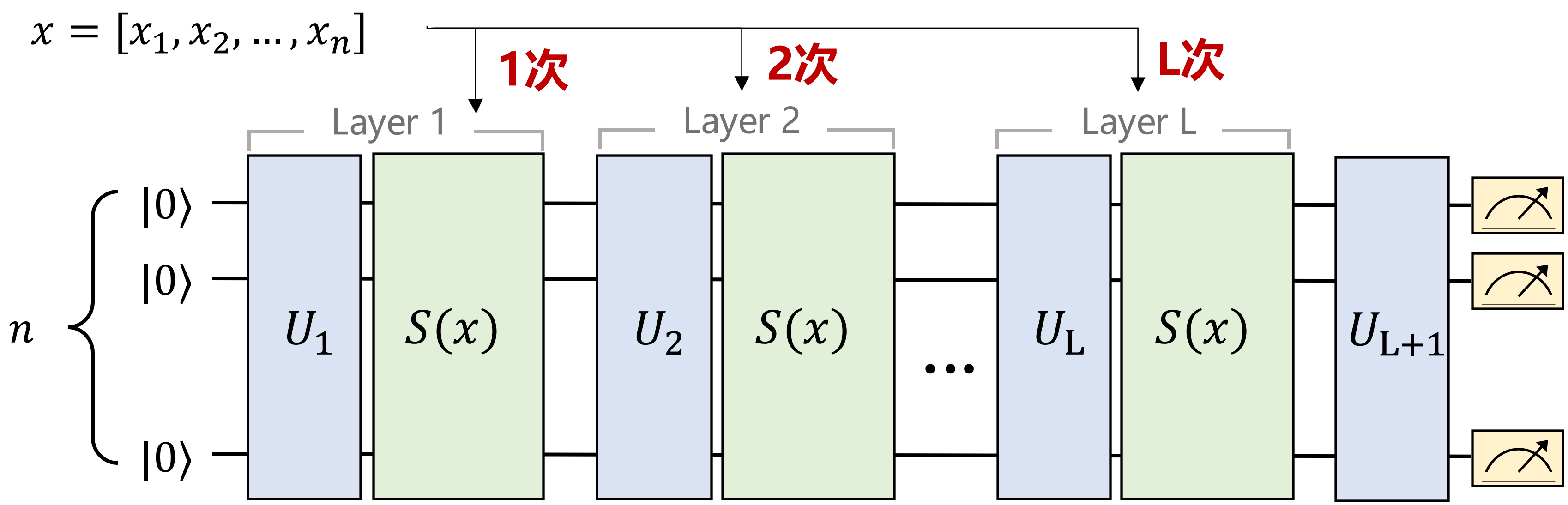
梯度更新:

$$\nabla_{\boldsymbol{\theta}_i}\langle\widehat{M}\rangle(\boldsymbol{\theta}) = \frac{1}{2}\left[\langle\widehat{M}\rangle\left(\boldsymbol{\theta}+\frac{\pi}{2}\hat{\mathbf{e}}_i\right) - \langle\widehat{M}\rangle\left(\boldsymbol{\theta}-\frac{\pi}{2}\hat{\mathbf{e}}_i\right)\right]$$



数据重上传电路

传统方法数据只在量子线路前端上传一次数据，而**重上传电路是将数据在参数层后面重复上传多次**



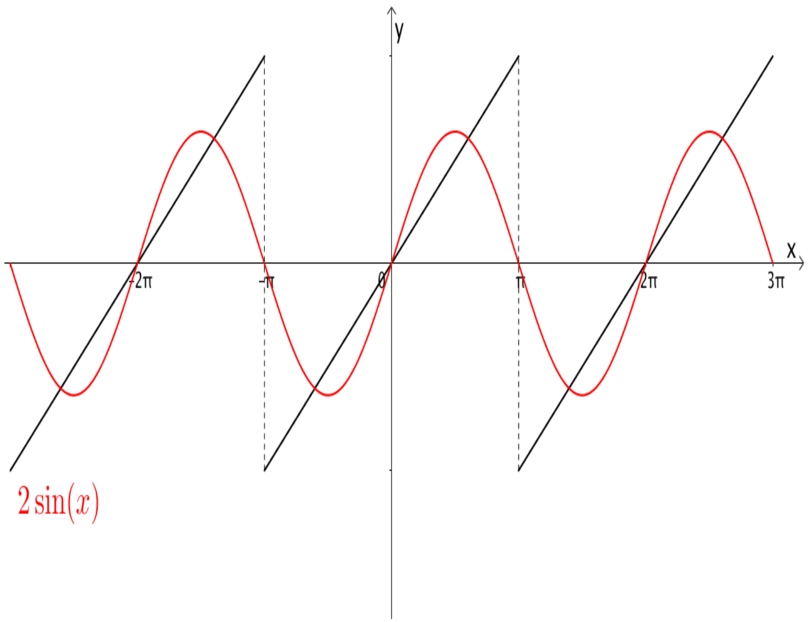
$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = U(x, \boldsymbol{\theta}) = U_{\mathrm{L}}(\boldsymbol{\theta}_{\mathrm{L}})S(x)\cdots U_2(\boldsymbol{\theta}_2)S(x)U_1(\boldsymbol{\theta}_1)|0\rangle^{\otimes n}$$

傅里叶序列

傅里叶序列能将一个函数表达成由正弦、余弦函数以系数a_n 和b_n为系数组合：

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)),$$

n是频率构成频谱$\boldsymbol{N} = \{\frac{\boldsymbol{n}}{\boldsymbol{T}}\}_{\boldsymbol{n}\in\mathbb{Z}}$



根据欧拉公式

$$e^{i\theta} = cos\theta + isin\theta$$

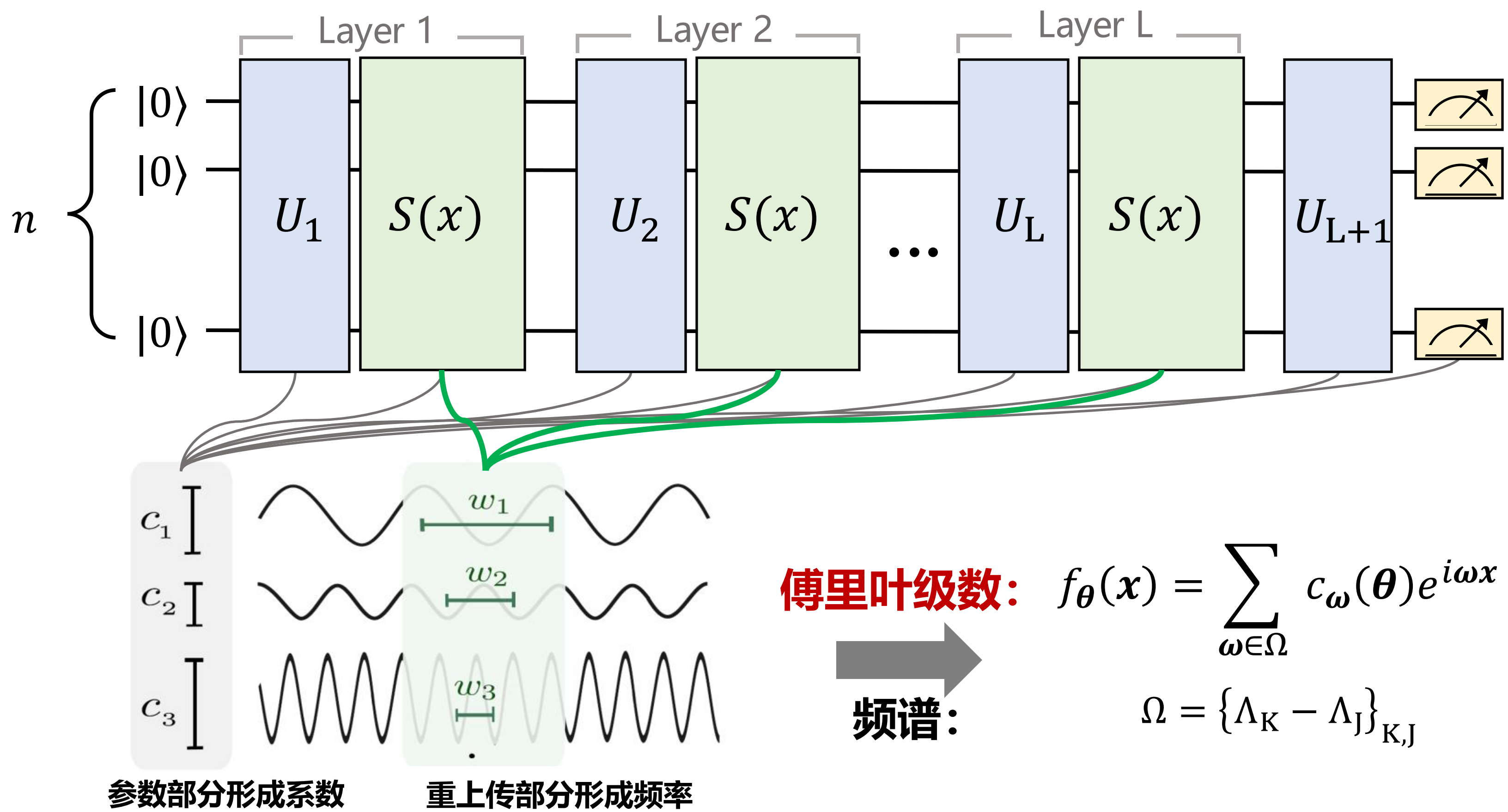
可得：

$$sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

带入可以进一步表达为指数函数的形式：

$$f(\mathbf{x}) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

其中 c_n 为系数



傅里叶级数： $f_\theta(x) = \sum_{\omega \in \Omega} c_\omega(\boldsymbol{\theta}) e^{i\omega x}$

频谱： $\Omega = \{\Lambda_K - \Lambda_J\}_{K,J}$

- 其中频谱$\{\Lambda_K - \Lambda_J\}_{K,J}$=$\{\boldsymbol{\lambda}_{\mathbf{k}^{(1)}} + \cdots + \boldsymbol{\lambda}_{\mathbf{k}^{(L)}} - (\boldsymbol{\lambda}_{\mathbf{j}^{(1)}} + \cdots + \boldsymbol{\lambda}_{\mathbf{j}^{(L)}})\}$= $\left\{\left(\lambda_{k_1^{(1)}} + \cdots + \lambda_{k_1^{(L)}}\right) - \left(\lambda_{j_1^{(1)}} + \cdots + \lambda_{j_1^{(L)}}\right)\right\}^{d_h}$
- λ是编码层中所表达矩阵的特征值，它们之间的**加减运算形成频率**， d_h是输入特征的维度

Jiaming Zhao, Wenbo Qiao，**Peng Zhang***, et al. Quantum Implicit Neural Representations. ICML, 2024 (CCF-A)
Schuld, M., Sweke, R., and Meyer, J. J. Effect of data encoding on the expressive power of variational quantum machine-learning models. Physical Review A, 2021.

Theorem (万能 (通用) 近似定理 (Universal Approximation Theorem))

令 $\phi(\cdot)$ 是一个非常数、有界、单调递增的连续函数，I_d 是一个 d 维的单位超立方体 $[0,1]^d$，$C(I_d)$ 是定义在 I_d 上的连续函数集合。对于任何一个函数 $f \in C(I_d)$，存在一个整数 m，和一组实数 $v_i, b_i \in \mathbb{R}$ 以及实数向量 $w_i \in \mathbb{R}^d, i = 1, \cdots, m$，以至于我们可以定义函数

$$F(x) = \sum_{i=1}^{m} v_i \phi(w_i^T x + b_i), \tag{1}$$

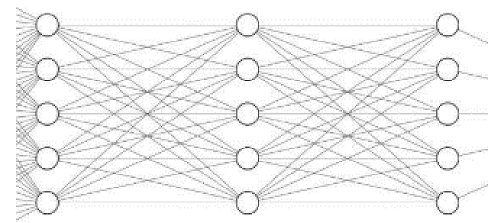
作为函数 f 的近似实现，即

$$|F(x) - f(x)| < \epsilon, \forall x \in I_d. \tag{2}$$

其中 $\epsilon > 0$ 是一个很小的正数。

神经网络

神经网络能够逼近任意复杂度的连续函数



Theorem (量子模型的万能 (通用) 近似定理 (Universal Approximation Theorem))

设 $\{H_m\}$ 是一个通用哈密顿量族，$\{f_m\}$ 是通过以下公式定义的关联量子模型族：

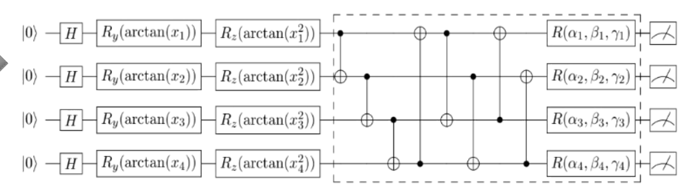
$$f_m(x) = \langle \Gamma | S_{H_m}^{\dagger}(x) M S_{H_m}(x) | \Gamma \rangle,$$

对于所有函数 $g \in L_2([0, 2\pi]^N)$，对于所有 $\epsilon > 0$，存在某个 $m' \in \mathbb{N}$，某个状态 $|\Gamma\rangle \in \mathbb{C}^{d^{m'}}$，以及某个可观测量 M，使得

$$\|f_{m'} - g\|_2 \leq \epsilon.$$

量子模型

量子模型可以拟合任意平方可积函数



Schuld, M., Sweke, R., and Meyer, J. J. Effect of data encoding on the expressive power of variational quantum machine-learning models. Physical Review A, 2021.

算法包与量子云平台

PENNYLANE　[M]s 昇思 MindSpore　MindQuantum

Qiskit　飞桨

本源量子 ORIGIN QUANTUM　TensorFlow

量子云平台支持

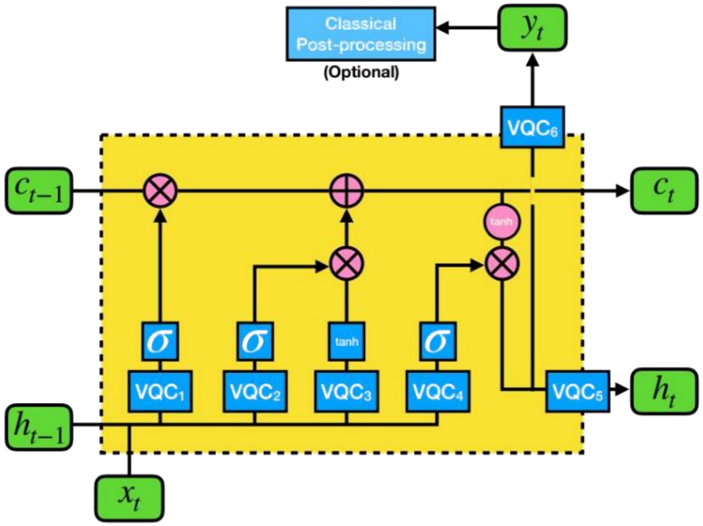
策略：模拟平台训练，量子平台推理

pennylane的量子神经网络

```python
import pennylane as qml
from pennylane import numpy as np

# 定义量子设备和量子节点
dev = qml.device("default.qubit", wires=2)
@qml.qnode(dev)
def quantum_circuit(inputs, weights):
    for i in range(2):
        qml.RX(inputs[i], wires=i)  # 数据编码
    qml.templates.StronglyEntanglingLayers(weights, wires=range(2))  # 可训练层
    return qml.expval(qml.PauliZ(0))  # 测量一个比特

# 初始化参数和数据
weights = np.random.uniform(size=qml.templates.StronglyEntanglingLayers.shape(2, 2))
X = np.array([[0.1, 0.2], [0.4, 0.3], [0.2, 0.9], [0.8, 0.5]])
y = np.array([0, 1, 0, 1])

# 定义优化器和损失函数
opt = qml.AdamOptimizer(stepsize=0.1)
def cost(weights):
    return np.mean([(quantum_circuit(X[i], weights) - y[i])**2 for i in range(len(X))])

# 训练量子神经网络
for epoch in range(50):
    weights = opt.step(cost, weights)
    if epoch % 10 == 0:
        print(f"Epoch {epoch}: Cost = {cost(weights):.4f}")

print("训练后的权重: ", weights)
```

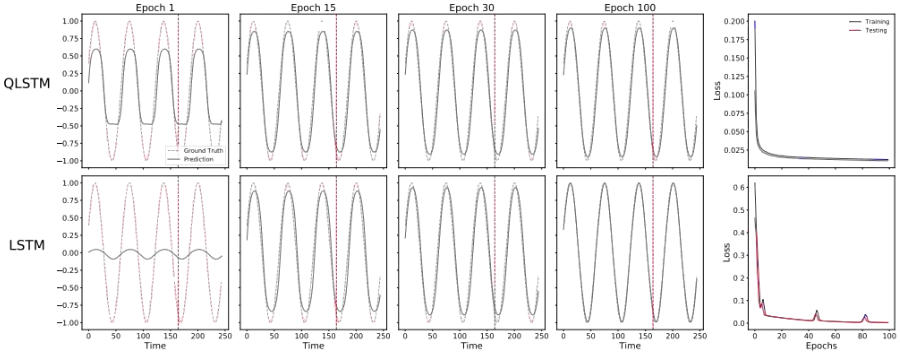
目录

简单的将经典机器学习的一些组件换成量子线路并不能超越经典神经网络，**量子神经网络缺乏能够证明并展现优势的应用场景**



$$
\begin{aligned}
f_t &= \sigma(\boxed{VQC_1}(v_t)) \\
i_t &= \sigma(\boxed{VQC_2}(v_t)) \\
\tilde{C}_t &= \tanh(\boxed{VQC_3}(v_t)) \\
c_t &= f_t * c_{t-1} + i_t * \tilde{C}_t \\
o_t &= \sigma(\boxed{VQC_4}(v_t)) \\
h_t &= \boxed{VQC_5}(o_t * \tanh(c_t)) \\
y_t &= \boxed{VQC_6}(o_t * \tanh(c_t))
\end{aligned}
$$

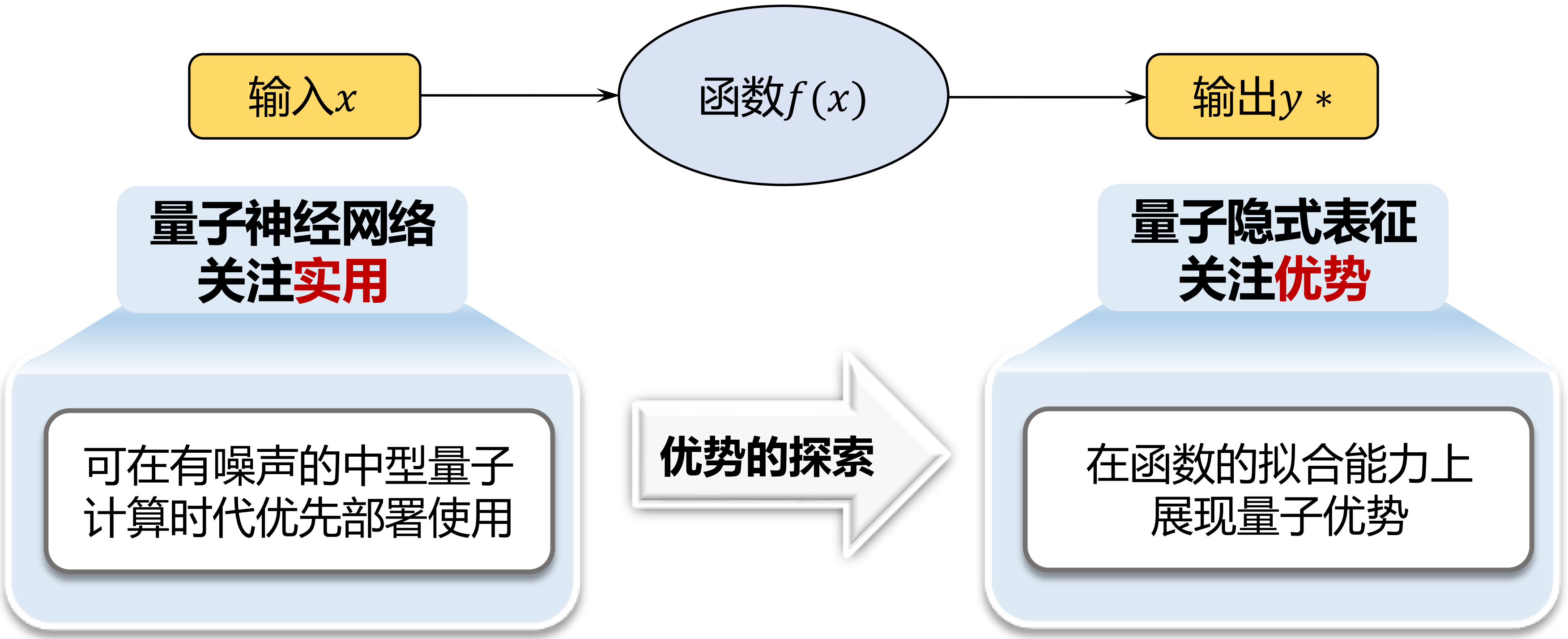
量子神经网络替换激活函数



没有发现显著优势

Chen S Y C, Yoo S, Fang Y L L. Quantum long short-term memory[C]//ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2022: 8622-8626.

探索能够体现“量子优势”的量子神经网络



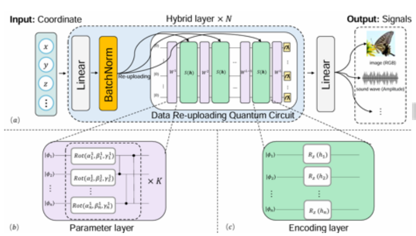
Wu, Shaojun, et al. Quantum reinforcement learning in continuous action space. arxiv preprint arxiv:2012.10711 (2020).

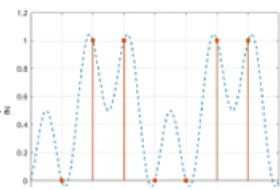
Singh, Gurmohan, et al. mplementation of quantum support vector machine algorithm using a benchmarking dataset. Indian Journal of Pure & Applied Physics (IJPAP) 60.5 (2022).

回顾之前的内容，**重上传的量子电路拥有拟合傅里叶序列的独特能力**

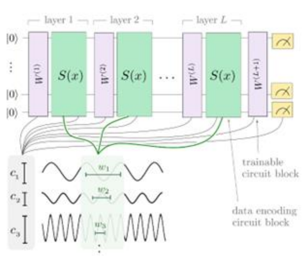
重上传量子电路

天生具有傅里叶序列的表征能力





$$f(\mathbf{h}) = \sum_{\mathbf{K},\mathbf{J}} a_{\mathbf{K},\mathbf{J}} e^{i(\Lambda_{\mathbf{K}}-\Lambda_{\mathbf{J}})\cdot\mathbf{h}},$$



- 少量的参数
- 拟合傅里叶序列的独特能力

联系

傅立叶级数

傅里叶级数在自然信号分析中具有普适性

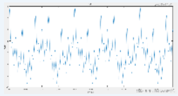
数据压缩 | 频域分析 | 图像滤波 | …

图像

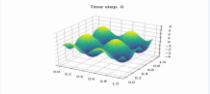
声学和音乐

$$f_\theta(x) = \sum_{\omega\in\Omega} c_\omega(\theta) e^{i\omega x}$$

电力系统



波动方程



Sitzmann, V. , Martel, J. N. P. , Bergman, A. W. , Lindell, D. B. , & Wetzstein, G. Implicit neural representations with periodic activation functions. NeurIPS , 2020.

傅立叶级数也是神经网络中的关键工具，**尤其常用于隐式神经表征任务**

傅立叶级数

傅里叶级数在自然信号分析中具有普适性

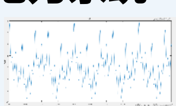
数据压缩 **频域分析** **图像滤波** …

图像

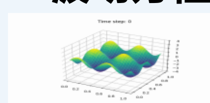
声学和音乐

$$f_\theta(x) = \sum_{\omega \in \Omega} c_\omega(\theta) e^{i\omega x}$$

电力系统



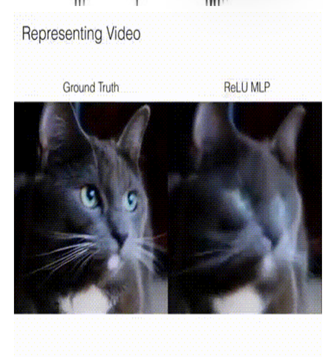
波动方程

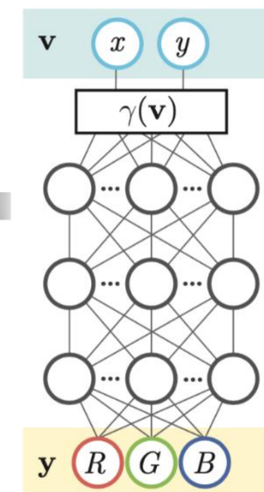


使用

隐式神经表征

压缩存储 **无限分辨率** **精细化表征**

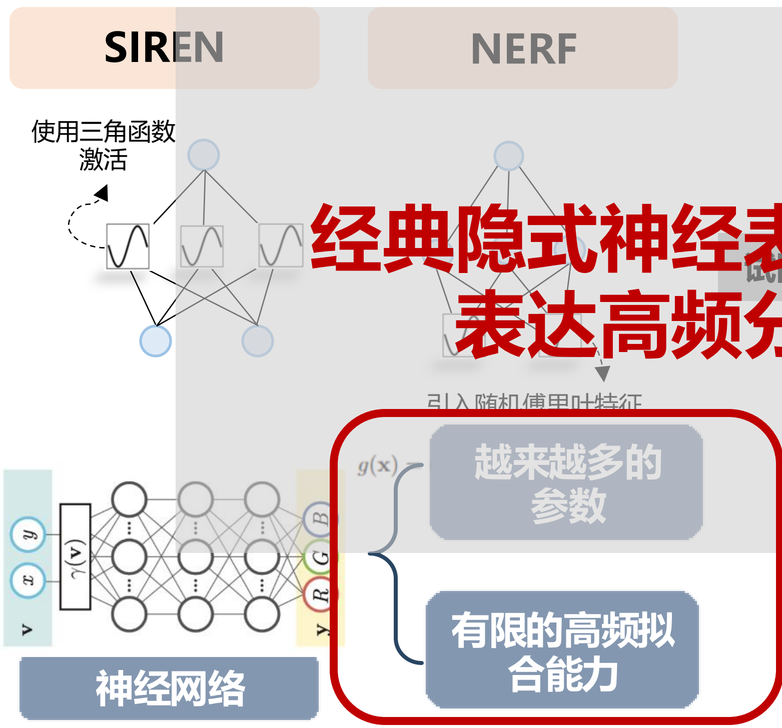




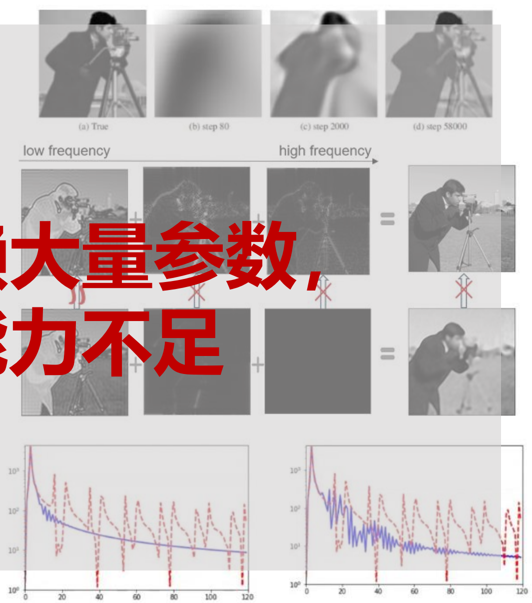
Sitzmann, V. , Martel, J. N. P. , Bergman, A. W. , Lindell, D. B. , & Wetzstein, G. Implicit neural representations with periodic activation functions. NeurIPS , 2020.

因为傅里叶序列的引入能解决隐式表征中存在的高频拟合难题

傅里叶特征



频率原理



低频易拟合，高频难拟合

经典隐式神经表征依赖大量参数，表达高频分量的能力不足

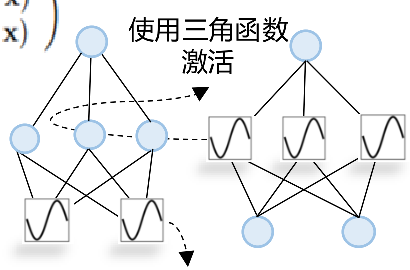
探索数据重上传量子线路的指数级增长的傅里叶序列拟合能力

经典隐式神经表示

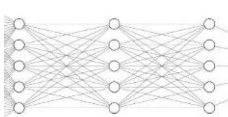
$$g(\mathbf{x}) = (a_{\mathbf{N}}, b_{\mathbf{N}}) \cdot \begin{pmatrix} \cos(2\pi \mathbf{N} \cdot \mathbf{x}) \\ \sin(2\pi \mathbf{N} \cdot \mathbf{x}) \end{pmatrix}$$

$$\mathbf{N} = \{\tfrac{\mathbf{n}}{\mathbf{T}}\}_{\mathbf{n} \in \mathbb{Z}^{d_{in}}}$$

线性长的频谱，为d



随神经网络参数**线性增长**的傅里叶序列表达能力



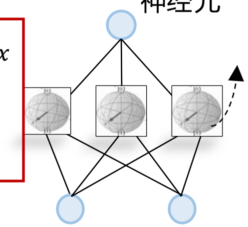
大量的参数和有限的表示精度

量子隐式神经表示

$$f(x) = \sum_{K,J} a_{K,J} e^{i(\Lambda_K - \Lambda_J)\cdot x}$$

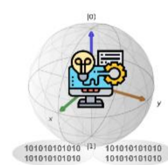
$$\Lambda_K - \Lambda_J = \{-dL, \ldots, dL\}^{d_x}$$

最优条件下指数长的频谱，为 $(2dL+1)^d$



量子优势

随量子比特数量**指数增长**的傅里叶序列表达能力



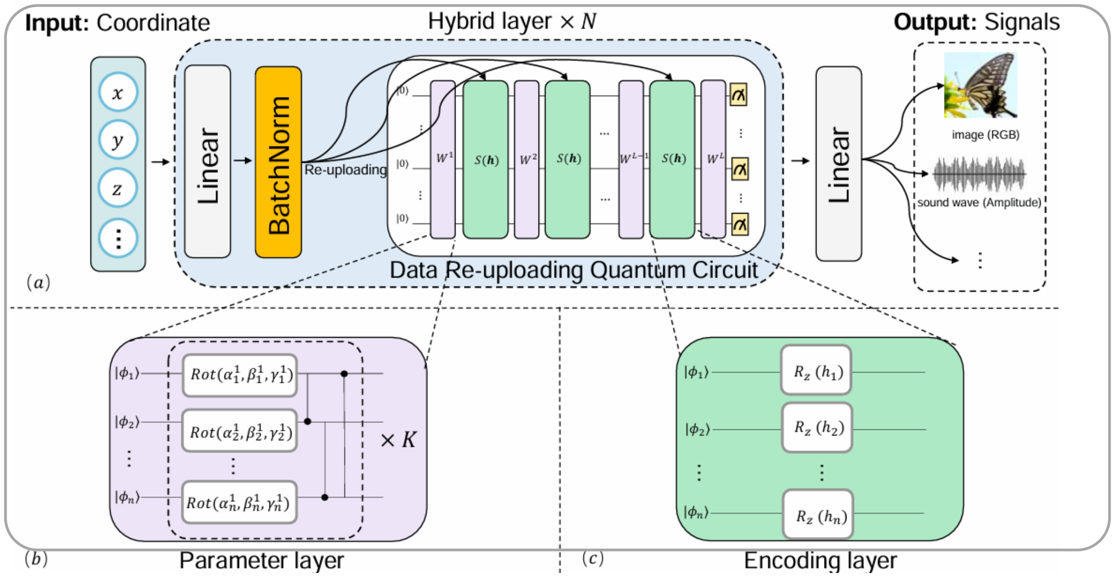
更少的参数和更精确的表示

Jiaming Zhao, Wenbo Qiao，**Peng Zhang***, et al. Quantum Implicit Neural Representations. **ICML, 2024 (CCF-A)**

成果

从理论上揭示了某种**量子线路具有指数级增长的傅里叶序列拟合能力**

量子机器学习从**理论到实践的一次跨越，为人工智能提供了量子视角的轻量化方案**



高频拟合能力

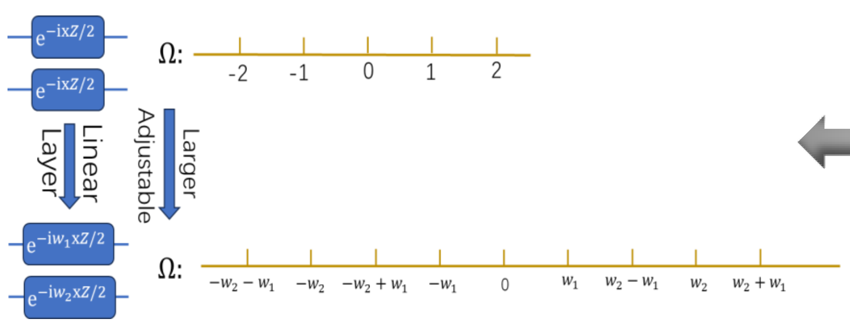
更少的训练参数

精细化表示

变分量子线路作为激活函数插入每层网络

Jiaming Zhao, Wenbo Qiao，**Peng Zhang***, et al. Quantum Implicit Neural Representations. **ICML, 2024 (CCF-A)**

在最佳条件下，数据重上传量子电路表示傅立叶级数的能力随着电路的大小呈指数增长



线性层扩展频谱

Step 1: $[U(x)|0\rangle]_i = \sum_{j_1\ldots j_L=1}^{d} e^{-i(\lambda_{j_1}+\cdots+\lambda_{j_L})x} \times W^{(L+1)}_{ij_L}\ldots W^{(2)}_{j_2j_1}W^{(1)}_{j_11}$

Step 2: $[U(x)|0\rangle]_i = \sum_{j\in[d]^L} e^{-i\Lambda_j x} W^{(L+1)}_{ij_L}\ldots W^{(2)}_{j_2j_1}W^{(1)}_{j_11}$

Step 3: $f(x) = \sum_{k,j\in[d]^L} e^{i(\Lambda_k-\Lambda_j)x} a_{k,j}$

$$a_{k,j} = \sum_{i,i'} (W^*)^{(1)}_{1k_1}(W^*)^{(2)}_{j_1j_2}\ldots(W^*)^{(L+1)}_{j_Li}M_{i,i'}\times W^{(L+1)}_{i'j_L}\ldots W^{(2)}_{j_2j_1}W^{(1)}_{j_11}.$$

推导

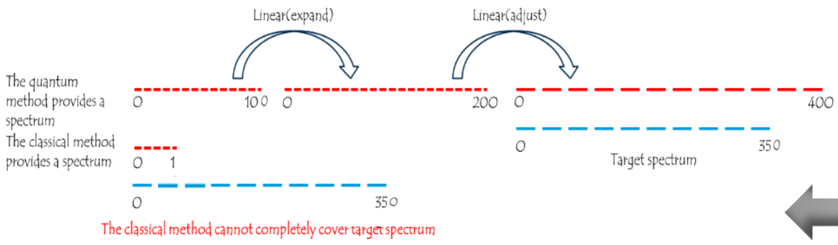
分析数据重上传电路的频谱

结论

1) 数据重上传量子线路的本质是傅里叶级数：$f(x) = \sum_{K,J} a_{K,J} e^{i(\Lambda_K-\Lambda_J)\cdot x}$

2) 对于一个大小为$d \times d_x$个qubits的数据重上传量子线路，**它能表征的傅里叶级数的的频谱大小为**: $\{\Lambda_K-\Lambda_J\}_{K,J} = \{-dL,\ldots,dL\}^{d_x}$

3)**在线性层的帮助下，频谱可以进一步扩展**，从$(2dL+1)^{d_x}$扩展到$((3^d-1)L+1)^{d_x}$



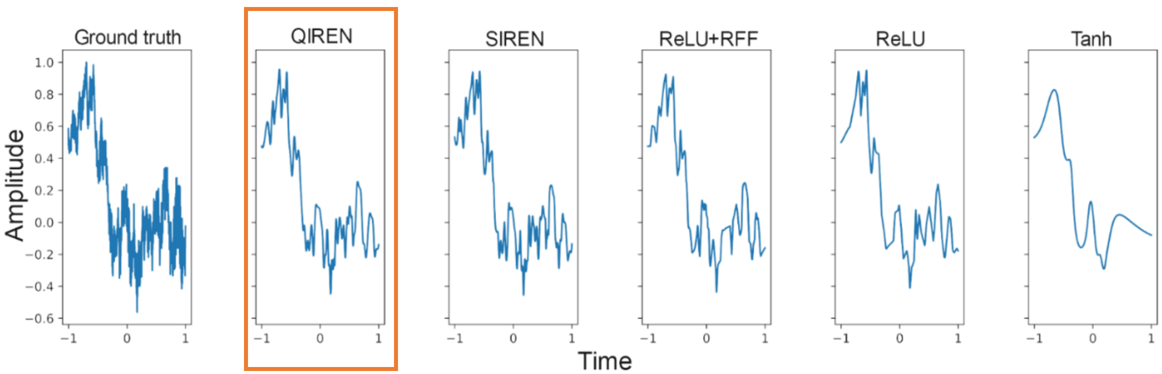
线性层扩展频谱和调整频率

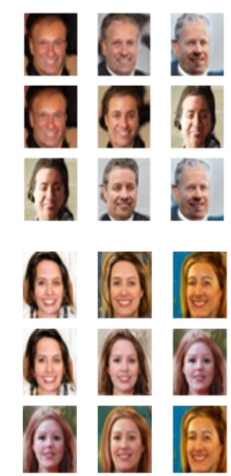
Jiaming Zhao, Wenbo Qiao, **Peng Zhang***, et al. Quantum Implicit Neural Representations. **ICML, 2024 (CCF-A)**

在信号表征、超分辨率和图像生成等众多任务中展现出精度和参数优势

Method	Sound Representation			Image Representation					Image Superresolution		
	Cello	#params	#mem(%)	Astronaut	Camera	Coffee	#params	#mem(%)	Astronaut	Camera	Coffee
Nearest	-	-	-	-	-	-	-	-	26.6	10.4	13.6
Bilinear	-	-	-	-	-	-	-	-	25.2	9.2	12.2
ReLU	6.8	831	16.9	9.9	2.7	4.2	841	17.9	33.8	11.3	13.5
Tanh	14.0	831	16.9	20.7	5.8	14.8	841	17.9	47.8	15.2	26.7
ReLU+RFF*	6.0	791	20.9	5.1	1.9	4.9	791	22.8	39.9	13.3	23.9
SIREN*	**5.5**	691	30.9	9.0	1.5	2.3	701	31.5	77.0	26.3	15.7
QIREN (ours)	**5.5**	**649** ↓	**35.1** ↑	**4.0** ↓	**1.1** ↓	**1.5** ↓	**657** ↓	**35.8** ↑	**24.3**	**7.9** ↓	**9.4** ↑

Method	FFHQ	CelebA-HQ	#params
Tanh	26.98	25.17	1.16M
ReLU	84.94	110.81	1.16M
ReLU+RFF*	15.01	13.91	1.14M
SIREN*	22.31	20.97	1.16M
QIREN (ours)	**11.53** ↓	**11.78** ↓	**1.13M** ↓



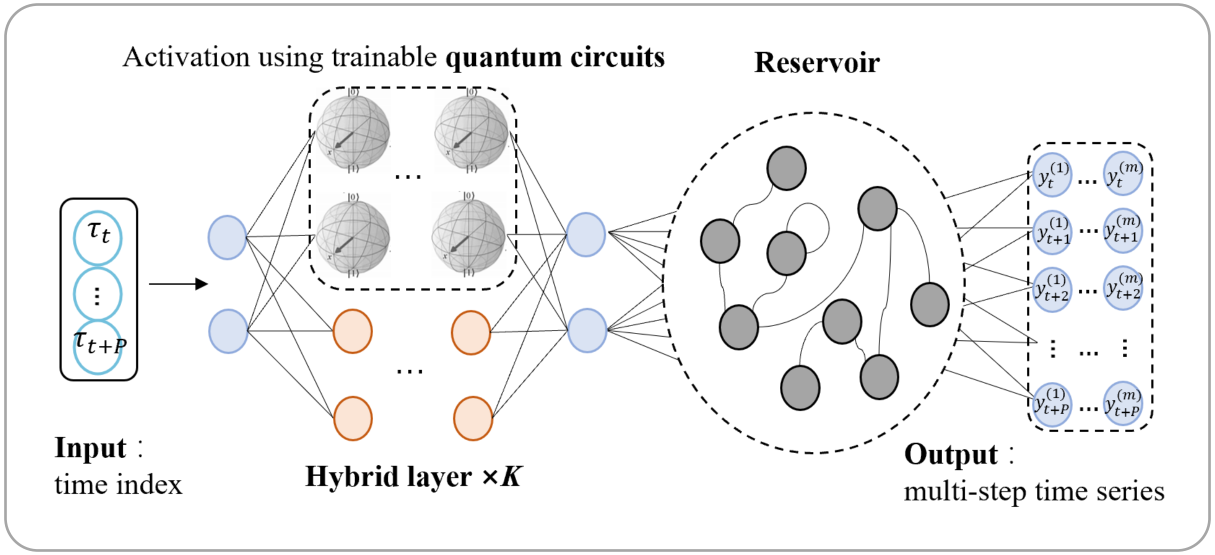


Jiaming Zhao, Wenbo Qiao，**Peng Zhang***, et al. Quantum Implicit Neural Representations. **ICML, 2024 (CCF-A)**

成果

使得量子机器学习**实现长序列时间序列预测（LSTF）问题成为可能**

揭示了**量子机器学习在电力、交通、汇率和非线性系统等领域的应用潜力**



量子时间指标模型

- 高频拟合能力
- 更少的训练参数
- 外推的能力

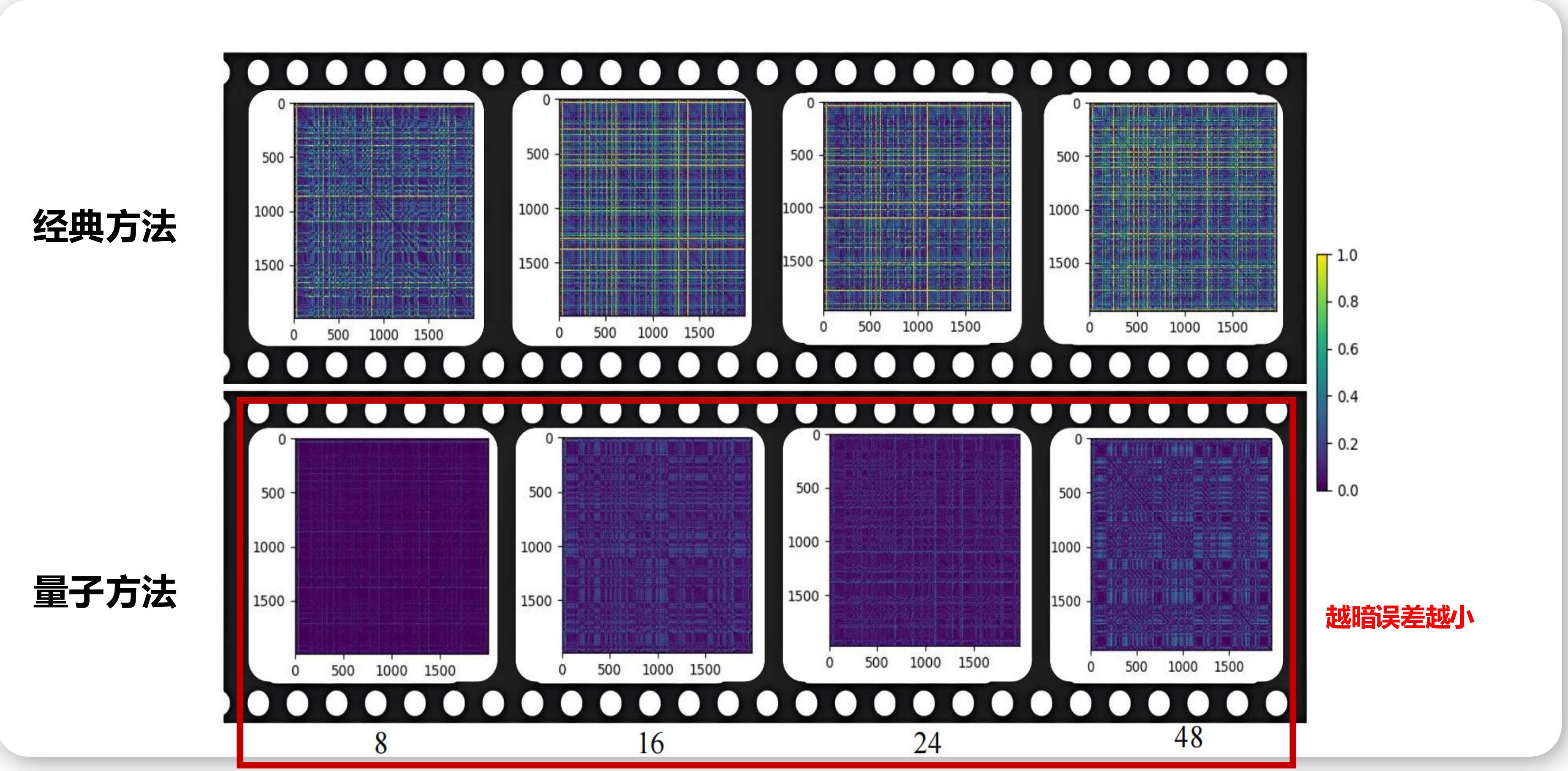
Wenbo Qiao , Jiaming Zhao，**Peng Zhang***, et al. Quantum Time-Index Models with Reservoir for Time Series Forecasting. **KDD 2025 (CCF-A)**

在电力、金融、交通等领域进行验证，最多可以节省SOTA方法98%的可训练参数

Methods		Informer		LogTrans		GP		NS Trans.		N-HiTS		ETSformer		FEDformer		Autoformer		DeepTime		QuantumTime (ours)		
Metrics		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	PR(%)
ECL	96	0.274	0.368	0.258	0.357	0.503	0.538	0.169	0.273	0.147	0.249	0.187	0.304	0.183	0.297	0.201	0.317	0.137	0.238	**0.135**	**0.232**	**97.66%**
	192	0.296	0.386	0.266	0.368	0.505	0.543	0.182	0.286	0.167	0.269	0.199	0.315	0.195	0.308	0.222	0.334	0.152	0.252	**0.148**	**0.244**	**97.51%**
	336	0.300	0.394	0.280	0.380	0.612	0.614	0.200	0.304	0.186	0.290	0.212	0.329	0.212	0.313	0.231	0.338	0.166	0.268	**0.165**	**0.261**	**97.37%**
	720	0.373	0.439	0.283	0.376	0.652	0.635	0.222	0.321	0.243	0.340	0.233	0.345	0.231	0.343	0.254	0.361	0.201	0.302	**0.198**	**0.297**	**97.29%**
Traffic	96	0.719	0.391	0.684	0.384	1.112	0.665	0.612	0.338	0.402	0.282	0.607	0.392	0.562	0.349	0.613	0.388	0.390	0.275	**0.386**	**0.272**	**97.66%**
	192	0.696	0.379	0.685	0.390	1.133	0.671	0.613	0.340	0.402	0.297	0.621	0.399	0.562	0.346	0.616	0.382	0.402	**0.278**	**0.400**	**0.278**	**97.51%**
	336	0.777	0.420	0.733	0.408	1.274	0.723	0.618	0.328	0.448	0.313	0.622	0.396	0.570	0.323	0.622	0.337	0.415	0.288	**0.412**	**0.284**	**96.86%**
	720	0.864	0.472	0.717	0.396	1.280	0.719	0.653	0.355	0.539	0.353	0.632	0.396	0.596	0.368	0.660	0.408	0.449	0.307	**0.447**	**0.302**	**96.20%**
Exchange	96	0.847	0.752	0.968	0.812	0.136	0.267	0.111	0.237	0.092	0.211	0.085	0.204	0.139	0.276	0.197	0.323	**0.081**	0.205	**0.081**	**0.198**	**98.24%**
	192	1.204	0.895	1.040	0.851	0.229	0.348	0.219	0.335	0.208	0.322	0.182	0.303	0.256	0.369	0.300	0.369	**0.151**	**0.284**	0.172	0.295	**97.51%**
	336	1.672	1.036	1.659	1.081	0.372	0.447	0.421	0.476	0.371	0.443	0.348	0.428	0.426	0.464	0.509	0.524	0.314	0.412	**0.292**	**0.400**	**96.34%**
	720	2.478	1.310	1.941	1.127	1.135	0.810	1.092	0.769	0.888	0.723	1.025	0.774	1.090	0.800	1.447	0.941	0.856	0.663	**0.595**	**0.583**	**96.20%**
ETTm2	96	0.365	0.453	0.768	0.642	0.442	0.422	0.192	0.274	0.176	0.255	0.189	0.280	0.203	0.287	0.255	0.339	0.166	0.257	**0.165**	**0.254**	**97.81%**
	192	0.533	0.563	0.989	0.757	0.605	0.505	0.280	0.339	0.245	0.305	0.253	0.319	0.269	0.328	0.281	0.340	0.225	0.302	**0.224**	**0.299**	**97.51%**
	336	1.363	0.887	1.334	0.872	0.669	0.569	0.334	0.361	0.295	0.346	0.314	0.357	0.325	0.366	0.339	0.372	0.277	0.336	**0.275**	**0.334**	**97.37%**
	720	3.379	1.388	3.048	1.328	0.959	0.669	0.417	0.413	0.401	0.426	0.414	0.413	0.421	0.415	0.422	0.419	**0.383**	**0.409**	0.397	0.422	**97.29%**

Wenbo Qiao , Jiaming Zhao , **Peng Zhang***, et al. Quantum Time-Index Models with Reservoir for Time Series Forecasting. **KDD 2025 (CCF-A)**

在非线性系统中也展现出更好的预测能力



Wenbo Qiao , Jiaming Zhao，**Peng Zhang***, et al. Quantum Time-Index Models with Reservoir for Time Series Forecasting. **KDD 2025 (CCF-A)**

量子隐式神经表征有望在各项任务上展现出应用潜力

大语言模型长序列建模

将序列视为一个连续时间函数，通过量子隐式神经表示学习该函数的连续映射

多媒体流数据压缩

高效压缩数据，同时保持高质量和支持实时处理

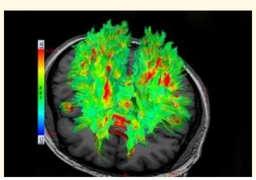
应用

量子隐式神经表征

应用

生物医学高维数据建模

通过连续函数映射高效处理和压缩复杂数据，提高分析精度和医学影像隐私安全



数字新媒体智能

实现高精度和高效的虚拟人类模拟，从而提升交互体验和多领域应用的智能化水平

量子隐式神经表征存在的挑战

模型尺度

✓ 随着任务规模扩大，所需的量子比特和门操作指数增长，导致模型难以扩展

训练效率

量子隐式神经表征仍需进一步探索

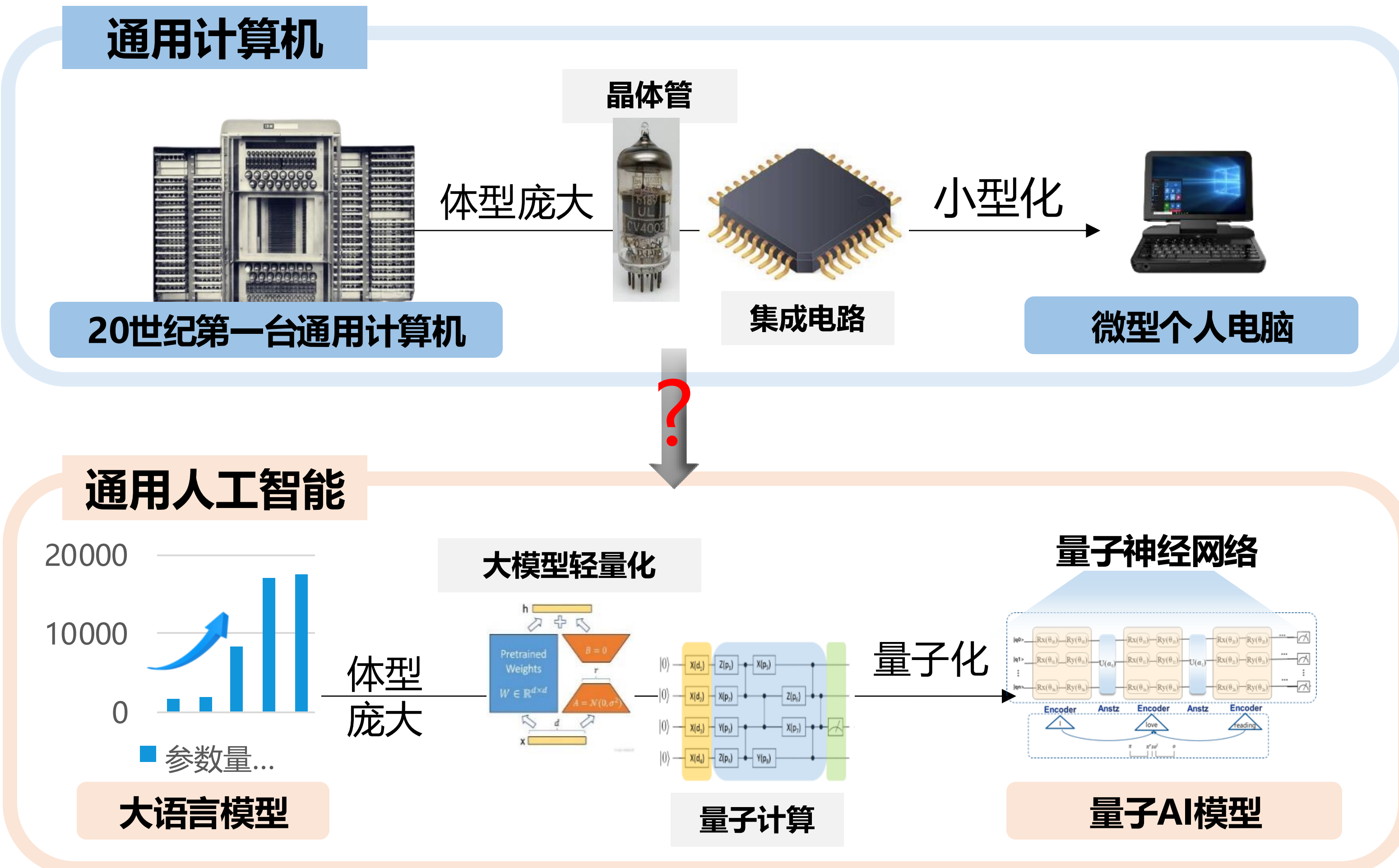
……理时间都较为耗时

硬件限制

✓ 当前量子计算硬件的比特数、退相干时间和门精度限制了模型在真机上的表现

未来展望

人工智能也许也会像通用计算机的发展历程一样不断被轻量化，其中量子神经网络有望扮演重要角色



请批评指正！

pzhang@tju.edu.cn

2025.01.10