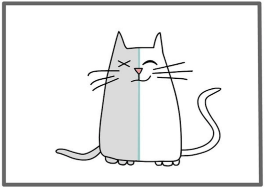
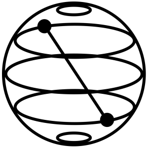
Aprendizaje automático cuántico

Mayra Alejandra Rivera Ruiz
mrivera@gdl.cinvestav.mx

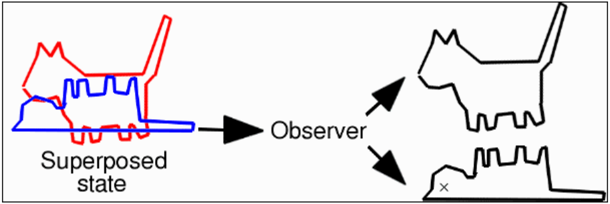


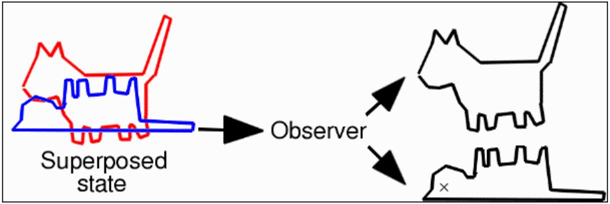


July 24, 2020

Contenido

Mecánica cuántica





Función de onda:

$$|\Psi\rangle = \frac{|\text{Gato muerto}\rangle + |\text{Gato vivo}\rangle}{\sqrt{2}}$$

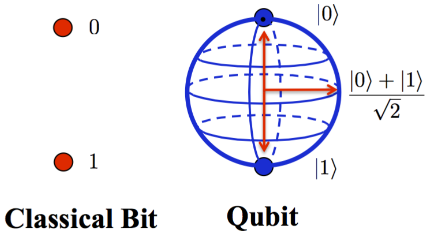
Cantidades observables representadas por operadores

- Los únicos valores que pueden resultar de una medida del observable físico A son los valores propios a de la ecuación $A\Psi_i = a_i\Psi_i$.
- Cualquier función de estado puede expresarse como combinación lineal de los estados propios de A.

$$\Psi = \Sigma c_i \Psi_i$$

donde $|c_i|^2 = |\langle \Psi_i | \Psi \rangle|^2$ es la probabilidad de medir Ψ_i.

El qubit



La base ortonormal $\{|0\rangle, |1\rangle\}$ se puede expresar como:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

La función de onda del qubit es:

$$|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

- $|\ \alpha\ |^2$: Nos dice la probabilidad de encontrar $|\Psi\rangle$ en el estado $|0\rangle$
- $|\ \beta\ |^2$: Nos dice la probabilidad de encontrar $|\Psi\rangle$ en el estado $|1\rangle$
- α y β satisfacen:

$$|\alpha|^2 + |\beta|^2 = 1$$

- Matemáticamente, para entender un sistema multipartícula, es necesario construir un espacio de Hilbert que es una composición de diferentes espacios de Hilbert independientes que están asociados a cada partícula.
- La maquinaria requerida para hacer esto es el producto tensorial.

Para el caso de dos qubits:

$$|00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Similarmente,

$$|01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$|11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Ejemplo

$$|\Psi\rangle = \frac{i}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$$

Encontramos que la probabilidad de encontrar el sistema en el estado $|0\rangle$

$$p_0 = \left|\frac{i}{2}\right|^2 = \left(\frac{i}{2}\right)^* \left(\frac{i}{2}\right) = \left(\frac{-i}{2}\right)\left(\frac{i}{2}\right) = \frac{1}{4}$$

Similarmente, la probabilidad de que el sistema se encuentre en el estado $|1\rangle$ es

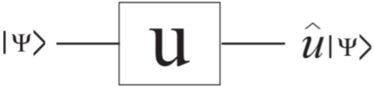
$$p_1 = \left|\frac{\sqrt{3}}{2}\right|^2 = \frac{3}{4}$$

Comprobando la suma de probabilidades:

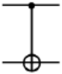
$$\sum p_i = p_0 + p_1 = \frac{1}{4} + \frac{3}{4} = 1$$

Compuertas cuánticas

En mecánica cuántica, podemos actuar sobre un estado $|\Psi\rangle$ con una transformación unitaria $\widehat{U}$. Esta transformación será el análogo a una compuerta lógica de un qubit.



Operaciones básicas sobre un qubit

Gate	Notation	Matrix
NOT (Pauli-X)		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli-Z		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Hadamard		$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
CNOT (Controlled NOT)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Compuerta NOT

$$X = U_{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$U_{NOT}\,|0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle$$

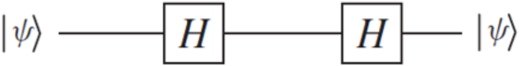
$$U_{NOT}\,|1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$$

Compuerta de Hadamard

Un importante paso en los algoritmos cuánticos es el uso de las compuertas Hadamard para crear estados superpuestos.

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$



En general, la aplicación de $H^{\otimes n}$ a un estado de productos de n copias de $|0\rangle$ es

$$H^{\otimes n}(|0\rangle^{\otimes n}) = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}} |x\rangle$$

Ejemplo:

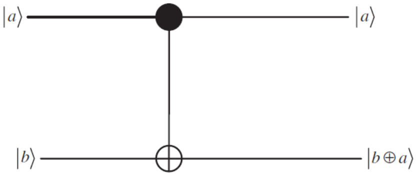
$$\begin{aligned} (H \otimes H)|0\rangle|0\rangle &= (H|0\rangle)(H|0\rangle) \\ &= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \\ &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \end{aligned}$$

Compuerta CNOT

$$CN = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$|00\rangle \to |00\rangle$$
$$|01\rangle \to |01\rangle$$
$$|10\rangle \to |11\rangle$$
$$|11\rangle \to |10\rangle$$

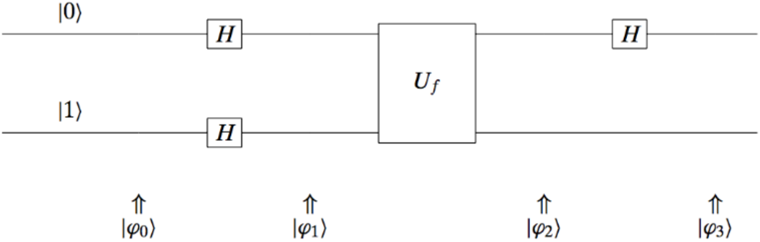
$$|a, b\rangle \to |a, b \oplus a\rangle$$



El algoritmo de Deutsch

$$U_f |x, y\rangle = |x, y \oplus f(x)\rangle$$

$$|\Psi_{out}\rangle = (H \otimes I)U_f(H \otimes H)|0\rangle |1\rangle$$



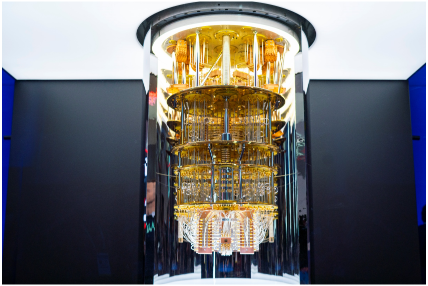
$$|\Psi_{out}\rangle = -\,|0\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \qquad (f(0) = f(1))$$

$$|\Psi_{out}\rangle = \pm\,|1\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \qquad (f(0) \neq f(1))$$

IBM Quantum Experience

Computadoras cuánticas disponibles:

- bmq _ 16 _ melbourne (15 qubits)
- bmq _essex (5 qubits)
- bmq _burlington (5 qubits)
- bmq _ london (5 qubits)
- bmq _vigo (5 qubits)
- bmq _ ourense (5 qubits)
- bmq _5 _yorktown-ibmqx2 (5 qubits)
- bmq _ armonk (1 qubit)

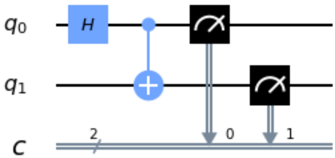


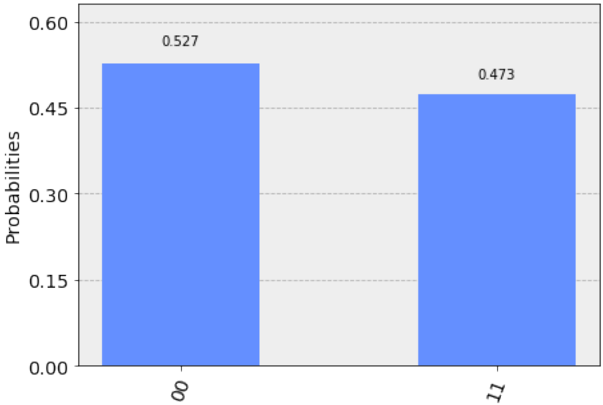
$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|\Psi_1\rangle = H|0\rangle \otimes |0\rangle = \frac{|00\rangle + |10\rangle}{\sqrt{2}}$$

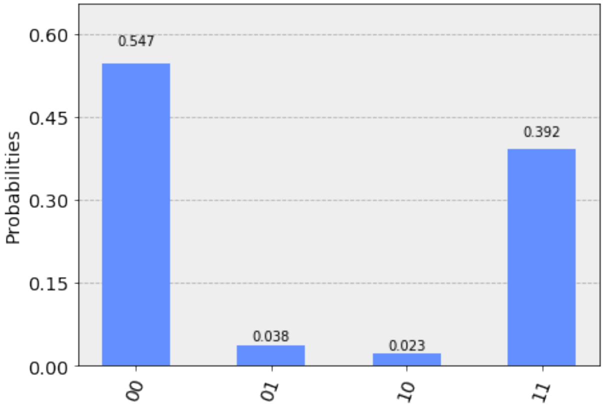
$$|\Psi\rangle = U_{CNOT}|\Psi_1\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

$$P_{00} = P_{11} = \left| \frac{1}{\sqrt{2}} \right|^2 = 0.5$$




0.60
0.45
0.30
0.15
0.00
Probabilities
0.527
0.473
00
11

Experimento



Algoritmos cuánticos

Algunos algoritmos cuánticos presentan una notable superioridad respecto a su contraparte clásica.

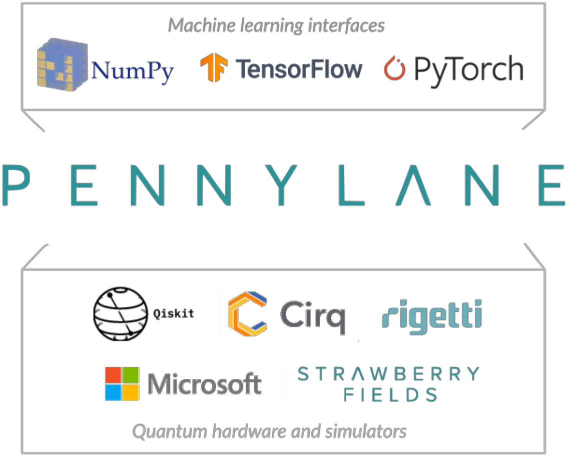
- La búsqueda en una base de datos desordenada (algoritmo de Grover)
- La factorización de números grandes (algoritmo de Shor)

Principios de la computación cuántica

Para realizar una computación con estados cuánticos necesitamos:

- Una colección finita de qubits, cuyo estado inicial corresponde al input.
- Un circuito de compuertas cuánticas, diseñado para ejecutar una trasformación unitaria sobre el estado inicial de los qubits.
- Finalmente, se debe realizar una medición sobre los qubits.

PennyLane


Machine learning interfaces
NumPy
TensorFlow
PyTorch
PENNYLANE
Qiskit
Cirq
rigetti
Microsoft
STRAWBERRY FIELDS
Quantum hardware and simulators

Planteamiento del problema

$$\min_{\theta \in \mathbb{R}^d} \mathcal{L}(\theta) = \min_{\theta \in \mathbb{R}^d} \frac{1}{2} \langle \psi_\theta, H\psi_\theta \rangle$$

donde $\psi_\theta = U_\theta |0\rangle$.

Expresión matemática

$$\theta_{t+1} = \theta_t - \mu g^{+}(\theta_t) \nabla L(\theta_t)$$

donde g es la parte real del Tensor Geométrico Cuántico, el cual, para este caso se expresa como:

$$G_{ij}^{(l)} = \langle \psi_l | K_i K_j | \psi_l \rangle - \langle \psi_l | K_i | \psi_l \rangle \langle \psi_l | K_j | \psi_l \rangle \quad .$$

Tensor geometrico cuántico

- Tensor Geométrico

$$G_{ij}(\theta) = \left\langle \frac{\partial\psi_\theta}{\partial\theta^i}, \frac{\partial\psi_\theta}{\partial\theta^j} \right\rangle - \left\langle \frac{\partial\psi_\theta}{\partial\theta^i}, \psi_\theta \right\rangle \left\langle \psi_\theta, \frac{\partial\psi_\theta}{\partial\theta^j} \right\rangle \quad = g_{\mu\nu} + iF_{\mu\nu}/2 \tag{1}$$

  - Parte real= Fubini-Study metric tensor

$$g_{ij} = Re\left[G_{ij}\right] \tag{2}$$

  - Parte imaginaria= Curvatura de Berry

$$F_{ij} = -2Im\left[G ij\right] \tag{3}$$

- El tensor geometrico cuántico describe la geometria del manifold del estado cuántico.

Notación:

Consideremos un operador unitario actuando sobre n qubits de la siguiente forma

$$U_L(\theta) := V_L(\boldsymbol{\theta}_L)W_L \cdots V_1(\boldsymbol{\theta}_1)W_1 \quad , \tag{4}$$

Introducimos la siguiente notación para representar un circuito entre las capas $l_1 \leq l_2$

$$U_{[l_1:l_2]} := V_{l_2}W_{l_2} \cdots V_{l_1}W_{l_1} \quad . \tag{5}$$

Ejemplo

$$U_L(\theta) = U_{(l:L]}V_lW_lU_{[1:l)} \quad , \tag{6}$$

Además, definiremos el siguiente estado para cada capa $l \in [L]$:

$$\psi_l := U_{[1:l]}|0\rangle \quad . \tag{7}$$

Generadores hermitianos

Para cada capa $l \in [L]$ existen los generadores hermitianos K_i y K_j tales que,

$$\partial_i V_l(\boldsymbol{\theta}_l) = -\mathrm{i} K_i V_l(\boldsymbol{\theta}_l) \ , \tag{8}$$

$$\partial_j V_l(\boldsymbol{\theta}_l) = -\mathrm{i} K_j V_l(\boldsymbol{\theta}_l) \ , \tag{9}$$

$[K_i, K_j] = 0$

La conmutatividad de los operadores de derivadas parciales combinado con la unitariedad de $V_l(\boldsymbol{\theta}_l)$ implica que $[K_i, K_j] = 0$.

Prueba de $[K_i, K_j] = 0$

Tomando la derivada parcial de (8) respecto a j y usando (9):

$$\partial_j \partial_i V_I(\theta_I) = -iK_i \partial_j V_I(\theta_I) \tag{10}$$

$$\partial_j \partial_i V_I(\theta_I) = (-iK_i)(-iK_j V_I(\theta_I)) \tag{11}$$

Por consiguiente

$$\boldsymbol{\partial_j \partial_i V_I(\theta_I) = -K_i K_j V_I(\theta_I)} \tag{12}$$

Similarmente

$$\partial_i \partial_j V_l(\theta_l) = iK_j \partial_i V_l(\theta_l) \tag{13}$$

$$\partial_i \partial_j V_l(\theta_l) = (-iK_j)(-iK_i V_l(\theta_l)) \tag{14}$$

$$\partial_{\mathbf{i}} \partial_{\mathbf{j}} \mathbf{V_l}(\theta_{\mathbf{l}}) = -\mathbf{K_j K_i V_l}(\theta_{\mathbf{l}}) \tag{15}$$

De (12) y (15) tenemos que

$$(K_i K_j - K_j K_i) V_l(\theta_l) = 0 \tag{16}$$

Lo cual implica

$$[\mathbf{K_i}, \mathbf{K_j}] = \mathbf{0} \tag{17}$$

Usando (6), (8) y (9) calculamos

$$\partial_j U_L(\theta) = U_{(l:L]}\partial_j V_l(\boldsymbol{\theta}_l) W_l U_{[1:l)} \quad , \tag{18}$$
$$= U_{(l:L]}(-\mathrm{i}K_j) V_l(\boldsymbol{\theta}_l) W_l U_{[1:l)} \quad , \tag{19}$$
$$= U_{(l:L]}(-\mathrm{i}K_j) U_{[1:l]} \quad . \tag{20}$$

Similarmente, tenemos que

$$\partial_i U_L(\theta)^\dagger = U_{[1:l]}^\dagger (\mathrm{i}K_i^\dagger) U_{(l:L]}^\dagger \quad . \tag{21}$$

Utilizando 20 y 21 y recordando que

$$\psi_\theta = U_\theta \left|0\right\rangle \tag{22}$$

$$\begin{aligned}
\langle \partial_i \psi_\theta | \partial_j \psi_\theta \rangle &= \left\langle 0\right| U^\dagger_{[1:l]} (iK_i^\dagger) U^\dagger_{(l:L]} U_{(l:L]} (-iK_j) U_{[1:l]} \left|0\right\rangle && (23)\\
&= \left\langle 0\right| U^\dagger_{[1:l]} K_i^\dagger K_j U_{[1:l]} \left|0\right\rangle && (24)\\
&= \left\langle 0\right| U^\dagger_{[1:l]} K_i K_j U_{[1:l]} \left|0\right\rangle && (25)\\
&= \left\langle \psi_l\right| K_i K_j \left|\psi_l\right\rangle && (26)
\end{aligned}$$

Finalmente

$$\langle \partial_{\mathbf{i}} \psi_\theta | \partial_{\mathbf{j}} \psi_\theta \rangle = \langle \psi_{\mathbf{l}} | \mathbf{K_i K_j} | \psi_{\mathbf{l}} \rangle \quad . \tag{27}$$

Calculo de $\mathrm{i}\langle\psi_\theta|\partial_j\psi_\theta\rangle$

$$
\begin{aligned}
\langle \psi_\theta \mid \partial_j \psi_\theta \rangle &= \langle 0| \, U_L^\dagger(\theta) U_{(l:L]}(-iK_j) U_{[1:l]} \, |0\rangle \\
&= \langle 0| \, U_{[1:l)}^\dagger W_l^\dagger V_l^\dagger U_{(l:L]}^\dagger U_{(l:L]}(-iK_j) U_{[1:l]} \, |0\rangle \\
&= \langle 0| \, (V_l W_l U_{[1:l)})^\dagger (-iK_j) U_{[1:l]} \, |0\rangle \\
&= -i \, \langle \psi_l| \, K_j \, |\psi_l\rangle
\end{aligned}
\tag{28}
$$

Finalmente

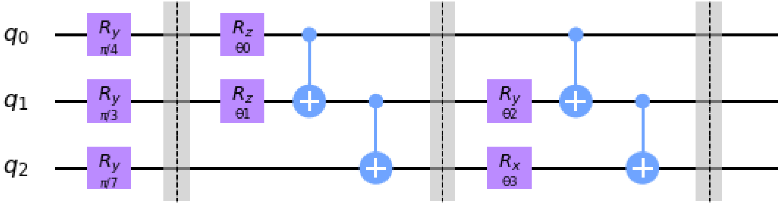
$$
\mathbf{i} \, \langle \psi_\theta \mid \partial_{\mathbf{j}} \psi_\theta \rangle = \langle \psi_{\mathbf{l}}| \, \mathbf{K_j} \, |\psi_{\mathbf{l}}\rangle \tag{29}
$$

Combinando estas expresiones obtenemos:

$$G^{I}_{ij}(\theta) = \left\langle \frac{\partial \psi_\theta}{\partial \theta^i}, \frac{\partial \psi_\theta}{\partial \theta^j} \right\rangle - \left\langle \frac{\partial \psi_\theta}{\partial \theta^i}, \psi_\theta \right\rangle \left\langle \psi_\theta, \frac{\partial \psi_\theta}{\partial \theta^j} \right\rangle \tag{30}$$

$$= \langle \psi_I | K_i K_j | \psi_I \rangle - [i \langle \psi_I | K_i | \psi_I \rangle] [-i \langle \psi_I | K_j | \psi_I \rangle] \tag{31}$$

$$= \langle \psi_I | K_i K_j | \psi_I \rangle - \langle \psi_I | K_i | \psi_I \rangle \langle \psi_I | K_j | \psi_I \rangle \tag{32}$$



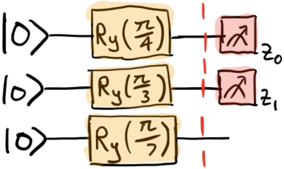
$$G = \begin{bmatrix} 0.125 & 0 & 0 & 0 \\ 0 & 0.185 & 0 & 0 \\ 0 & 0 & 0.24973433 & 0 \\ 0 & 0 & 0 & 0.20293623 \end{bmatrix}$$

Cálculo de G

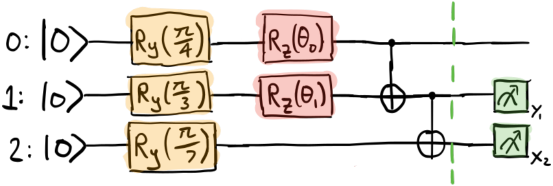
$$G = \begin{pmatrix} G^{(0)} & 0 \\ 0 & G^{(1)} \end{pmatrix} \tag{33}$$

- Una matriz de 2×2 para la capa 0
- Una matriz de 2×2 para la capa 1

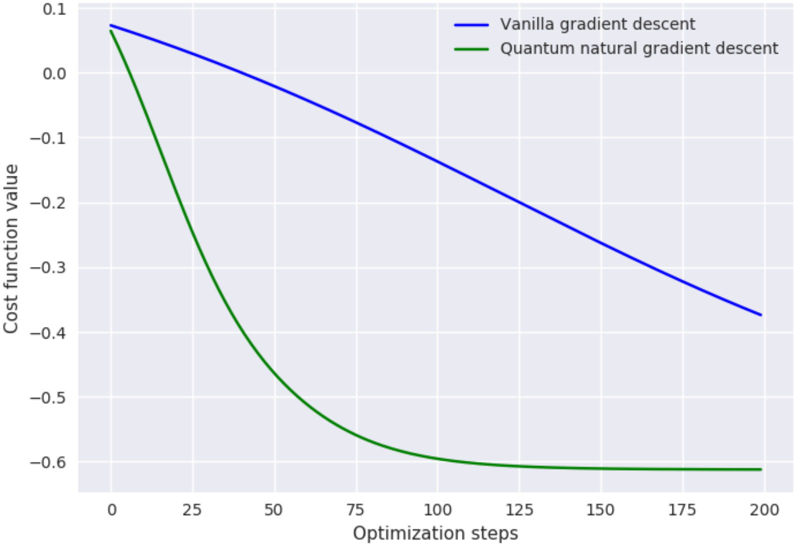
Capa 0



$$
\begin{array}{cc}
 & \begin{array}{cc} \qquad\qquad \theta_0 \qquad\qquad\qquad & \qquad\qquad\qquad \theta_1 \end{array} \\
G^{(0)} = \begin{array}{c} \theta_0 \\ \theta_1 \end{array} & \begin{pmatrix} < z_0^2 > - < z_0 >^2 & < z_0 z_1 > - < z_0 >< z_1 > \\ < z_0 z_1 > - < z_0 >< z_1 > & < z_1^2 > - < z_1 >^2 \end{pmatrix}
\end{array}
\tag{34}
$$



$$\begin{array}{cc} & \begin{array}{cc} \boldsymbol{\theta}_2 & \boldsymbol{\theta}_3 \end{array} \\ G^{(0)} = \begin{array}{c} \boldsymbol{\theta}_2 \\ \boldsymbol{\theta}_3 \end{array} & \begin{pmatrix} < y_1^2 > - < y_1 >^2 & < y_1 x_2 > - < y_1 >< x_2 > \\ < y_1 x_2 > - < y_1 >< x_2 > & < x_2^2 > - < x_2 >^2 \end{pmatrix} \end{array} \quad (35)$$



Vanilla gradient descent
Quantum natural gradient descent
Cost function value
Optimization steps
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
0
25
50
75
100
125
150
175
200

Relación con la información clásica de Fisher

$$\psi_\theta = \sum_{x\in[N]} \sqrt{p_\theta(x)}|x\rangle \quad . \tag{36}$$

Por regla de la cadena

$$\frac{\partial \psi_\theta}{\partial \theta^i} = \frac{1}{2}\sum_{x\in[N]} \frac{1}{\sqrt{p_\theta(x)}}\frac{\partial p_\theta(x)}{\partial \theta^i}|x\rangle \quad . \tag{37}$$

$$\left\langle \psi_\theta, \frac{\partial \psi_\theta}{\partial \theta^i} \right\rangle = \frac{1}{2} \sum_{x \in [N]} \sum_{x' \in [N]} \frac{\sqrt{p_\theta(x')}}{\sqrt{p_\theta(x)}} \frac{\partial p_\theta(x)}{\partial \theta^i} \langle x' | x \rangle \ , \tag{38}$$

$$= \frac{1}{2} \sum_{x \in [N]} \frac{\partial p_\theta(x)}{\partial \theta^i} \ , \tag{39}$$

$$= \frac{1}{2} \frac{\partial}{\partial \theta^i} \sum_{x \in [N]} p_\theta(x) \ , \tag{40}$$

$$= \frac{1}{2} \frac{\partial}{\partial \theta^i} 1 \ , \tag{41}$$

$$= 0 \ , \tag{42}$$

donde hemos utilizado $\langle x' | x \rangle = \delta_{xx'}$.

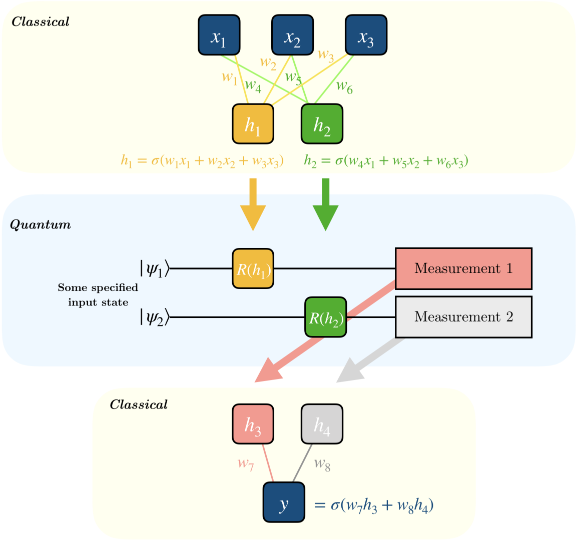
$$
\begin{aligned}
G_{ij}(\theta) &= \left\langle \frac{\partial \psi_\theta}{\partial \theta^i}, \frac{\partial \psi_\theta}{\partial \theta^j} \right\rangle \ , && (43)\\
&= \frac{1}{4} \sum_{x\in[N]} \sum_{x'\in[N]} \frac{1}{\sqrt{p_\theta(x)p_\theta(x')}} \frac{\partial p_\theta(x)}{\partial \theta^i} \frac{\partial p_\theta(x')}{\partial \theta^j} \langle x'|x\rangle \ , && (44)\\
&= \frac{1}{4} \sum_{x\in[N]} \frac{1}{p_\theta(x)} \frac{\partial p_\theta(x)}{\partial \theta^i} \frac{\partial p_\theta(x)}{\partial \theta^j} \ , && (45)\\
&= \frac{1}{4} \sum_{x\in[N]} p_\theta(x) \frac{\partial \log p_\theta(x)}{\partial \theta^i} \frac{\partial \log p_\theta(x)}{\partial \theta^j} \ , && (46)\\
&= \frac{1}{4} I_{ij}(\theta) \ . && (47)
\end{aligned}
$$

QNN

- QNN es una red neuronal artificial (ANN) incluida en un sistema con computación cuántica.
- Los investigadores intentan desarrollar algoritmos más eficientes en la clasificación de patrones o aprendizaje automático que lo que está disponible ahora en las capacidades de ANN.
- La computación cuántica promete acelerar el entrenamiento de las redes neuronales clásicas y la computación de aplicaciones de big data, generando resultados más rápidos.

Las ventajas de QNN incluyen lo siguiente:

- menor número de neuronas ocultas proporcionan un mayor rendimiento
- aprendizaje rápido
- problemas linealmente inseparables para soluciones de red de capa única



Classical
x_1
x_2
x_3
w_1 w_2 w_3 w_4 w_5 w_6
h_1
h_2
$h_1 = \sigma(w_1x_1 + w_2x_2 + w_3x_3)$
$h_2 = \sigma(w_4x_1 + w_5x_2 + w_6x_3)$
Quantum
$|\psi_1\rangle$
$R(h_1)$
Measurement 1
Some specified input state
$|\psi_2\rangle$
$R(h_2)$
Measurement 2
Classical
h_3
h_4
w_7
w_8
y
$= \sigma(w_7h_3 + w_8h_4)$

The parameter shift rule

Si en un circuito cuántico

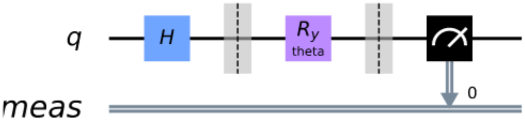
$$f(\theta) = \langle\Psi|\, U_G^\dagger(\theta) A U_G(\theta)\, |\Psi\rangle\,,$$

donde $U_G(\theta) = e^{-ia\theta G}$, a es un número real y G es un operador hermitiano con eigenvalores e_0 y e_1. Entonces

$$\frac{d}{d\theta} f(\theta) = r\left[f(\theta + \frac{\pi}{4r}) - f(\theta - \frac{\pi}{4r})\right], \tag{48}$$

con $r = {}^{a}/{}_{2}(e_1 - e_0)$.

Ejemplo



$$|\psi\rangle = H|0\rangle \tag{49}$$

$$R_y(\theta) = e^{-\frac{\theta Y}{2}} \tag{50}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{51}$$

$$e = \pm 1 \tag{52}$$

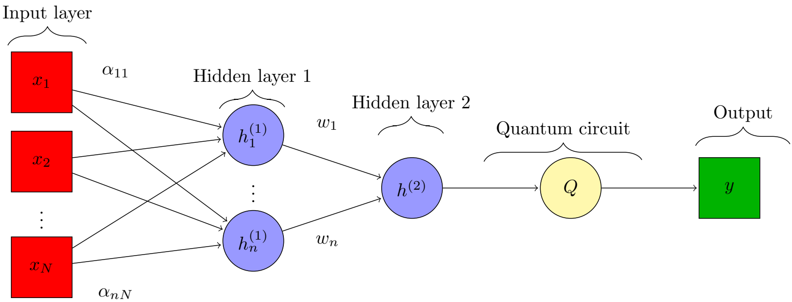
$$f(\theta) = \langle\psi| R_y(\theta)\sigma_z R_y(\theta) |\psi\rangle \tag{53}$$

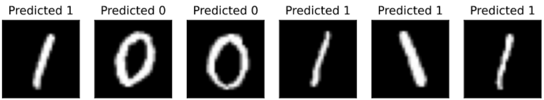
$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{54}$$

$$r = \frac{a}{2}(e_1 - e_0) = \frac{1}{4}(1 - (-1)) = \frac{1}{2} \tag{55}$$

$$shift = \frac{\pi}{4r} = \frac{\pi}{2} \tag{56}$$

$$\frac{df(\theta)}{d\theta} = \Delta Q = \frac{1}{2}\left[f(\theta + \frac{\pi}{2}) - f(\theta - \frac{\pi}{2})\right] \tag{57}$$


Input layer
x_1
x_2
x_N
α_{11}
α_{nN}
Hidden layer 1
$h_1^{(1)}$
$h_n^{(1)}$
w_1
w_n
Hidden layer 2
$h^{(2)}$
Quantum circuit
Q
Output
y


Predicted 1
Predicted 0
Predicted 0
Predicted 1
Predicted 1
Predicted 1

$$J = -t\log(y) - (1-t)\log(1-y) \tag{58}$$

$$y(\theta) = \langle\psi|\, R_y(\theta)\sigma_z R_y(\theta)\,|\psi\rangle \tag{59}$$

$$\frac{dy}{d\theta} = \Delta Q \tag{60}$$

$$\theta = h^{(2)} = \sum_j w_j h_j^{(1)} \tag{61}$$

$$h_j^{(1)} = \frac{1}{1+e^{-z_j}} \tag{62}$$

$$z_j = \sum_i \alpha_{ji} x_i \tag{63}$$

Hidden layer 2

$$\frac{\partial J}{\partial w_j} = \frac{\partial J}{\partial y}\frac{\partial y}{\partial \theta}\frac{\partial \theta}{\partial w_j} \tag{64}$$

$$\tag{65}$$

$$\begin{aligned}
\frac{\partial J}{\partial y} &= \frac{\partial}{\partial y}\left(-t log(y) - (1-t) log(1-y)\right) && (66)\\
&= -\frac{t}{y} + \frac{(1-t)}{(1-y)} && (67)\\
&= \frac{-t(1-y) + y(1-t)}{y(1-y)} && (68)\\
&= \frac{-t + ty + y - yt}{y(1-y)} && (69)\\
&= \frac{y-t}{y(1-y)} && (70)
\end{aligned}$$

$$\frac{\partial \theta}{\partial w_j} = \frac{\partial}{\partial w_j}\left(\sum_i w_i h_i^{(1)}\right) \tag{71}$$

$$= h_j^{(1)} \tag{72}$$

$$\Delta w_j = -\alpha \frac{\partial J}{\partial w_j} = -\alpha \left[\frac{y - t}{y(1 - y)}\right] \Delta Q h_j^{(1)} \tag{73}$$

Hidden layer 1

$$\frac{\partial J}{\partial \alpha_{ji}} = \frac{\partial J}{\partial h_j^{(1)}} \frac{\partial h_j^{(1)}}{\partial z_j} \frac{\partial z_j}{\partial \alpha_{ji}} \tag{74}$$

$$\frac{\partial z_j}{\partial \alpha_{ji}} = \frac{\partial}{\partial \alpha_{ji}} \left[\sum_k \alpha_{jk} x_k \right] = x_i \tag{75}$$

$$\frac{\partial h_j^{(1)}}{\partial z_j} = \frac{\partial}{\partial z_j}\left(\frac{1}{1+e^{-z_j}}\right) \tag{76}$$

$$= -\left(1+e^{-z_j}\right)^{-2} e^{-z_j}(-1) \tag{77}$$

$$= (1+e^{-z_j})^{-1}(1+e^{-z_j})^{-1}e^{-z_j} \tag{78}$$

$$= h_j^{(1)}\left(\frac{e^{-z_j}}{1+e^{-z_j}}\right) \tag{79}$$

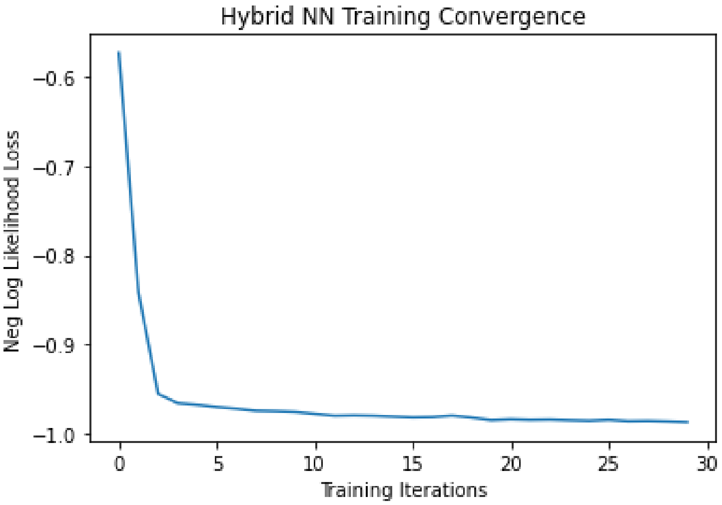
$$= h_j^{(1)}\left(\frac{e^{-z_j}+1-1}{1+e^{-z_j}}\right) \tag{80}$$

$$= h_j^{(1)}\left(\frac{e^{-z_j}+1}{1+e^{-z_j}} - \frac{1}{1+e^{-z_j}}\right) \tag{81}$$

$$= h_j^{(1)}(1-h_j^{(1)}) \tag{82}$$

$$
\begin{aligned}
\frac{\partial J}{\partial \alpha_{ji}} &= \frac{\partial J}{\partial h_j^{(1)}} \frac{\partial h_j^{(1)}}{\partial z_j} \frac{\partial z_j}{\partial \alpha_{ji}} && (83) \\
&= \frac{\partial J}{\partial h_j^{(1)}} h_j^{(1)} (1 - h_j^{(1)}) x_i && (84) \\
&= \delta_j x_i && (85)
\end{aligned}
$$

$$
\begin{aligned}
\Delta \alpha_{ji} &= -\alpha \frac{\partial J}{\partial \alpha_{ji}} && (86) \\
&= -\alpha \delta_j x_i && (87)
\end{aligned}
$$


Hybrid NN Training Convergence
Neg Log Likelihood Loss
−0.6
−0.7
−0.8
−0.9
−1.0
0
5
10
15
20
25
30
Training Iterations

Berry Phases in Electronic Structure Theory

David Vanderbilt

Rutgers University

2018

Experimental Measurement of the Quantum Metric Tensor and Related Topological Phase Transition with a Superconducting Qubit

"Tan, Xinsheng and Zhang, Dan-Wei and Yang, Zhen and Chu, Ji and Zhu, Yan-Qing and Li, Danyu and Yang, Xiaopei and Song, Shuqing and Han, Zhikun and Li, Zhiyuan and others"

"Physical review letters"

volume=122, number=21, pages=210401, year=2019, publisher=APS

Adiabatic approximation

B. Zwiebach

june 2, 2017

S. Amari and S. C. Douglas

"Why natural gradient?

" Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)

Seattle, WA, USA, 1998, pp. 1213-1216 vol.2.

doi: 10.1109/ICASSP.1998.675489

Bibliografía (cont.)

Quantum Natural Gradient

James Stokes and Josh A. Izaac and Nathan Killoran and Giuseppe Carleo

journal=ArXiv,

year=2019,

volume=abs/1909.02108

Ville Bergholm et al

PennyLane: Automatic differentiation of hybrid quantum- classical computations

arXiv: 1811.04968v3 [quant-ph]

14 Feb 2020

Gavin E. Crooks

Gradients of parameterized quantum gates using the parameter-shift rule and gatedecomposition

arXiv: 1905.13311v1 [quant-ph]

30 May 2019