Equations of Qubit Neural Network

Noriaki Kouda

1. Neuron model, network model

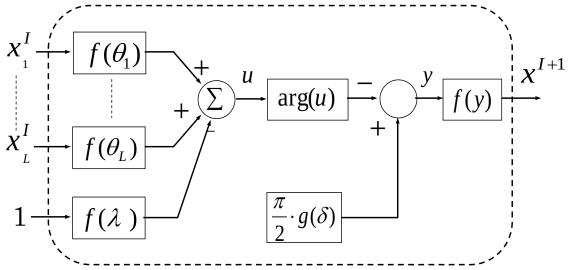

Fig.1 Qubit neuron model

$$x_l^I = e^{i\phi_l} \tag{1}$$

$$\begin{aligned} u &= \sum_l^L f(\theta_l)\cdot x_l^I - f(\lambda) \\ &= \sum_l^L e^{i(\theta_l+\phi_l)} - f(\lambda) \end{aligned} \tag{2}$$

$$y = \frac{\pi}{2}\cdot g(\delta) - \arg(u) \tag{3}$$

$$x^{I+1} = f(y) = e^{iy} \tag{4}$$

$g()$ is the sigmoid function.

Now, in implementation, Eq.(2), (3) are changed to below equations. Eq.(5) calculates Real part and Imaginary part respectively, and obtains output phase y.

$$y = \frac{\pi}{2}\cdot g(\delta) - \tan^{-1}(U(u)) = \frac{\pi}{2}\cdot g(\delta) - \tan^{-1}\left(\frac{\mathrm{Im}(u)}{\mathrm{Re}(u)}\right) \tag{5}$$

$$\mathrm{Im}(u) = \sum_l^L \sin(\theta_l+\phi_l) - \sin(\lambda)$$

$$\mathrm{Re}(u) = \sum_l^L \cos(\theta_l+\phi_l) - \cos(\lambda)$$

$$U(u) = \frac{\mathrm{Im}(u)}{\mathrm{Re}(u)}$$

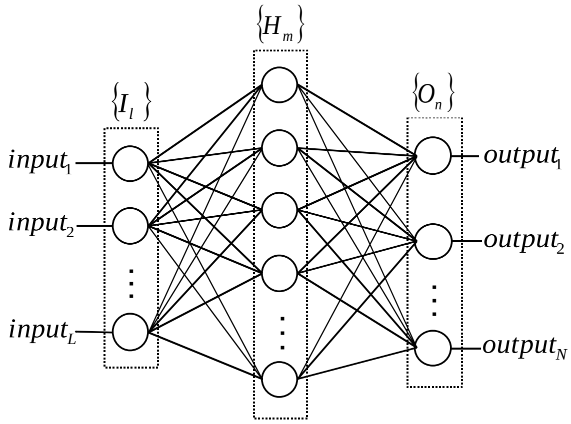

Fig.2 3-layerd network

In Fig.2, $\{I_l\}$ (l =1, 2･･･, L), $\{H_m\}$ (m = 1, 2･･･, M) and $\{O_n\}$ (n=1, 2･･･, N) mean network layers and l, m, n show position of neurons, respectively. $\{I_l\}$, $\{H_m\}$, $\{O_n\}$ are input, hidden and output layers.

Our qubit neural network transmits neuron signal between network layers by quantum state. From this feature, l-th network input $input_l$ = [0, 1] is converted to quantum phase ϕ_l^I = [0, π/2] and quantum state $x_l^I = e^{i\phi_l}$. x_l^I is put into input of l -th neuron on input layer.

Output of the network is defined as the probability that observes quantum state $|1\rangle$, therefore, n-th output of the network, $output_n$, is squared imaginary part of x_n^{output} that is the output quantum state of neuron on output layer. This definition is based on probability interpretation on quantum physics.

$$\phi_l^I = \frac{\pi}{2} \cdot input_l \tag{6}$$

$$output_n = \left|\mathrm{Im}(x_n^{output})\right|^2 = \left|\sin(y_n^O)\right|^2 \tag{7}$$

In training, the method of steepest descent is used. Our neural network uses follow squared error function.

$$E_{total} = \frac{1}{2} \cdot \sum_p^K E_p = \frac{1}{2} \cdot \sum_p^K \sum_n^N (t_{p,n} - output_{p,n})^2 \tag{8}$$

Here, K is the number of training input pattern and $output_{p,n}$, $t_{p,n}$ are $output_n$ and training target for the $output_n$ at p-th training pattern, respectively. Equations for parameter revision are given follow Eq.(9)～(11) and η is the learning rate.

$$\theta^{new} = \theta^{old} - \eta \frac{\partial E_{total}}{\partial \theta^{old}} \tag{9}$$

$$\lambda^{new} = \lambda^{old} - \eta \frac{\partial E_{total}}{\partial \lambda^{old}} \tag{10}$$

$$\delta^{new} = \delta^{old} - \eta \frac{\partial E_{total}}{\partial \delta^{old}} \tag{11}$$

2. Divergence in 2-layers network

2.1 Output Layer

・Network Output

$$\frac{\partial E_p}{\partial output_n} = output_{p,n} - t_{p,n}$$

$$\frac{\partial output_n}{\partial y_n^O} = 2\sin(y_n^O)\cdot\cos(y_n^O) = \sin(2y_n^O)$$

・ δ_n^O

$$\frac{\partial y_n^O}{\partial \delta_n^O} = \frac{\pi}{2}\cdot\frac{e^{-\delta_n^O}}{\left(1+e^{-\delta_n^O}\right)^2} = \frac{\pi}{2}\cdot\frac{1}{1+e^{-\delta_n^O}}\cdot\frac{e^{-\delta_n^O}}{1+e^{-\delta_n^O}} = \frac{\pi}{2}\cdot g(\delta_n^O)\cdot\left(1-g(\delta_n^O)\right)$$

$$\frac{\partial E_p}{\partial \delta_n^O} = \frac{\partial E_p}{\partial output_n}\cdot\frac{\partial output_n}{\partial y_n^O}\cdot\frac{\partial y_n^O}{\partial \delta_n^O}$$

・ $\theta_{mn}^O, \lambda_n^O$

$$\frac{\partial y_n^O}{\partial U_n^O} = -\frac{1}{1+\left(U_n^O\right)^2} = -\frac{\left(\mathrm{Re}(u_n^O)\right)^2}{\left(\mathrm{Re}(u_n^O)\right)^2+\left(\mathrm{Im}(u_n^O)\right)^2}$$

$$\frac{\partial U_n^O}{\partial \theta_{mn}^O} = \frac{\mathrm{Im}'(u_n^O)\,\mathrm{Re}(u_n^O)-\mathrm{Im}(u_n^O)\,\mathrm{Re}'(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2} = \frac{\cos(\theta_{mn}^O+y_m^H)\,\mathrm{Re}(u_n^O)+\sin(\theta_{mn}^O+y_m^H)\,\mathrm{Im}(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2}$$

$$\frac{\partial U_n^O}{\partial \lambda_n^O} = \frac{\mathrm{Im}'(u_n^O)\,\mathrm{Re}(u_n^O)-\mathrm{Im}(u_n^O)\,\mathrm{Re}'(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2} = \frac{-\cos(\lambda_n^O)\,\mathrm{Re}(u_n^O)-\sin(\lambda_n^O)\,\mathrm{Im}(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2}$$

$$\frac{\partial E_p}{\partial \theta_{mn}^O} = \frac{\partial E_p}{\partial output_n}\cdot\frac{\partial output_n}{\partial y_n^O}\cdot\frac{\partial y_n^O}{\partial U_n^O}\cdot\frac{\partial U_n^O}{\partial \theta_{mn}^O} \qquad \frac{\partial E_p}{\partial \lambda_n^O} = \frac{\partial E_p}{\partial output_n}\cdot\frac{\partial output_n}{\partial y_n^O}\cdot\frac{\partial y_n^O}{\partial U_n^O}\cdot\frac{\partial U_n^O}{\partial \lambda_n^O}$$

2.2 Hidden Layer

$$\frac{\partial U_n^O}{\partial y_m^H}=\frac{\mathrm{Im}'(u_n^O)\,\mathrm{Re}(u_n^O)-\mathrm{Im}(u_n^O)\,\mathrm{Re}'(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2}=\frac{\cos(\theta_{mn}^O+y_m^H)\,\mathrm{Re}(u_n^O)+\sin(\theta_{mn}^O+y_m^H)\,\mathrm{Im}(u_n^O)}{\left(\mathrm{Re}(u_n^O)\right)^2}=\frac{\partial U_n^O}{\partial \theta_{mn}^O}$$

$$\frac{\partial E_p}{\partial y_m^H}=\sum_n^N\left(\frac{\partial E_p}{\partial output_n}\cdot\frac{\partial output_n}{\partial y_n^O}\cdot\frac{\partial y_n^O}{\partial U_n^O}\cdot\frac{\partial U_n^O}{\partial \theta_{mn}^H}\right)$$

· δ_m^H

$$\frac{\partial y_m^H}{\partial \delta_m^H}=\frac{\pi}{2}\cdot\frac{e^{-\delta_m^H}}{\left(1+e^{-\delta_m^H}\right)^2}=\frac{\pi}{2}\cdot\frac{1}{1+e^{-\delta_m^H}}\cdot\frac{e^{-\delta_m^H}}{1+e^{-\delta_m^H}}=\frac{\pi}{2}\cdot g(\delta_m^H)\cdot\left(1-g(\delta_m^H)\right)$$

$$\frac{\partial E_p}{\partial \delta_m^H}=\frac{\partial E_p}{\partial y_m^H}\cdot\frac{\partial y_m^H}{\partial \delta_m^H}$$

· $\theta_{lm}^H,\lambda_m^H$

$$\frac{\partial y_m^H}{\partial U_m^H}==-\frac{1}{1+\left(U_m^H\right)^2}=-\frac{\left(\mathrm{Re}(u_m^H)\right)^2}{\left(\mathrm{Re}(u_m^H)\right)^2+\left(\mathrm{Im}(u_m^H)\right)^2}$$

$$\frac{\partial U_m^H}{\partial \theta_{lm}^H}=\frac{\mathrm{Im}'(u_m^H)\,\mathrm{Re}(u_m^H)-\mathrm{Im}(u_m^H)\,\mathrm{Re}'(u_m^H)}{\left(\mathrm{Re}(u_m^H)\right)^2}=\frac{\cos(\theta_{lm}^H+\phi_l^I)\,\mathrm{Re}(u_m^H)+\sin(\theta_{lm}^H+\phi_l^I)\,\mathrm{Im}(u_m^H)}{\left(\mathrm{Re}(u_m^H)\right)^2}$$

$$\frac{\partial U_n^O}{\partial \lambda_n^O}=\frac{\mathrm{Im}'(u_m^H)\,\mathrm{Re}(u_m^H)-\mathrm{Im}(u_n^O)\,\mathrm{Re}'(u_n^O)}{\left(\mathrm{Re}(u_m^H)\right)^2}=\frac{-\cos(\lambda_m^H)\,\mathrm{Re}(u_m^H)-\sin(\lambda_m^H)\,\mathrm{Im}(u_m^H)}{\left(\mathrm{Re}(u_m^H)\right)^2}$$

$$\frac{\partial E_p}{\partial \theta_{lm}^H}=\frac{\partial E_p}{\partial y_m^H}\cdot\frac{\partial y_m^H}{\partial U_m^H}\cdot\frac{\partial U_m^H}{\partial \theta_{lm}^H}\qquad\frac{\partial E_p}{\partial \lambda_m^H}=\frac{\partial E_p}{\partial y_m^H}\cdot\frac{\partial y_m^H}{\partial U_m^H}\cdot\frac{\partial U_m^H}{\partial \lambda_m^H}$$