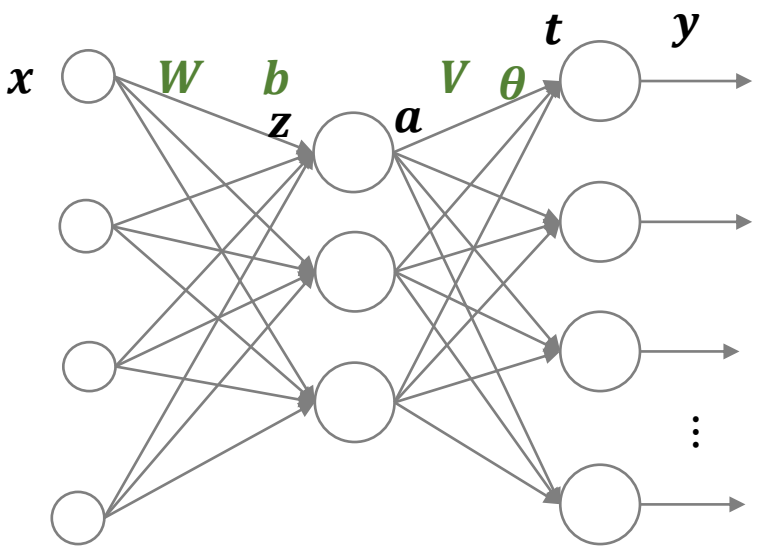
MNIST 手写数字识别数据集



前向传播

$$z = Wx + b$$

$$a = \sigma(z)$$

$$t = Va + \theta$$

$$y = \sigma(t)$$

$$L = \frac{1}{2}(y - h)^2$$

$x(784 \times 1)$

$W(12 \times 784)$

$b(12 \times 1)$

$z(12 \times 1)$

$a(12 \times 1)$

$V(10 \times 12)$

$\theta(10 \times 1)$

$t(10 \times 1)$

$y\ (10 \times 1)$

$h(10 \times 1)$

$$\frac{\partial L}{\partial \theta} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial t}\frac{\partial t}{\partial \theta} = (y - h) \circ y \circ (1 - y)$$

$$\frac{\partial L}{\partial V} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial t}\frac{\partial t}{\partial V} = (y - h) \circ y \circ (1 - y)a^{\mathrm{T}}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial t}\frac{\partial t}{\partial a}\frac{\partial a}{\partial z}\frac{\partial z}{\partial b} = V^T[(y - h) \circ y \circ (1 - y)] \circ a \circ (1 - a)$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial t}\frac{\partial t}{\partial a}\frac{\partial a}{\partial z}\frac{\partial z}{\partial W} = V^T[(y - h) \circ y \circ (1 - y)] \circ a \circ (1 - a)x^{\mathrm{T}}$$

$$\theta = \theta - \alpha\frac{\partial L}{\partial \theta}$$

$$V = V - \alpha\frac{\partial L}{\partial V}$$

$$b = b - \alpha\frac{\partial L}{\partial b}$$

$$W = W - \alpha\frac{\partial L}{\partial W}$$

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
# read data from files
train_data = pd.read_csv('experiment_05_training_set.csv')
test_data = pd.read_csv('experiment_05_testing_set.csv')
train_data = np.array(train_data)
test_data = np.array(test_data)
train_x = train_data[:,1:]/255
train_y = train_data[:,0]
test_x = test_data[:,1:]/255
test_y = test_data[:,0]
# one-hot encoding
n = train_x.shape[0]
one_hot_h = np.zeros((n,10))
one_hot_h[np.arange(n),train_y.reshape((n,))]  = 1
# initialization
theta = np.random.randn(10,1)
V = np.random.randn(10,12)
b = np.random.randn(12,1)
W = np.random.randn(12,784)
```