EMS702P Statistical Thinking and Applied Machine Learning

Week 9 – PBL

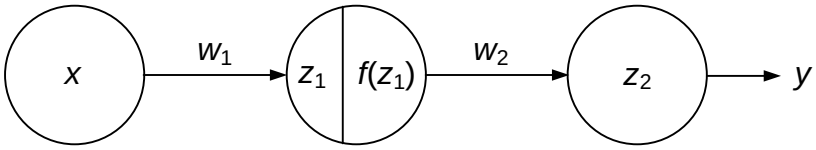
Yunpeng Zhu

Q1: (30 min)

Consider we have 3 sets of observed data

$$(x,y)=\left[(1.0,0.3);(1.5,0.6);(2.0,0.8)\right]$$

Use gradient decent method to determine the one-dimensional neural network model:



with $f(z)=\dfrac{1}{1+\exp(-z)}$ by solving the following problems:

(1) Calculate the feed-forward outputs and the cost function value with w_1 and w_2 being initialized as $\bar{w}_1=0.3$ and $\bar{w}_2=0.1$;

(2) Calculate the gradient of the cost function with respect to the weight w_2 for the first back-propagation step;

(3) Calculate the gradient of the cost function with respect to the weight w_1 for the first back-propagation step;

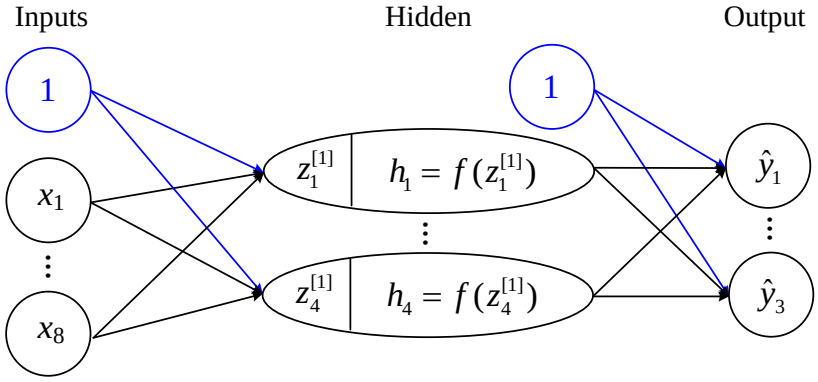
(4) Calculate the updated weights w_1 and w_2 for the first back-propagation step under the learning rate $\lambda=0.5$;

The required cost function is

$$J=\frac{1}{3}\sum_{t=1}^{3}\frac{1}{2}\left[\bar{y}(t)-y(t)\right]^2$$

with $\bar{y}(k)$ being the model prediction.

~~*Q2: (25 min)*~~



How to implement the BP training of the neural network (single data set)? Explain using symbolic calculations. The cost function is

$$J = \sum_{k=1}^{3} \frac{1}{2} \left[\hat{y}_k - y_k \right]^2$$