多层神经网路

Author： 中山大学 17数据科学与计算机学院 YSY

https://github.com/ysyisyourbrother

前馈计算

$$g_k(x)=z_k=f(\sum_{j=1}^{n_H}w_{kj}f(\sum_{i=1}^{d}w_{ji}x_i+w_{j0})+w_{k0})$$

反向传播

考虑隐藏层到输出层

$$\frac{\partial J}{\partial w_{kj}}=\frac{\partial J}{\partial net_k}\frac{\partial net_k}{\partial w_{kj}}$$

对于

$$J(w)=\frac{1}{2}\sum_{k=1}^{n}(t_k-z_k)^2$$

有

$$\frac{\partial J}{\partial net_k}=\frac{\partial J}{\partial z_k}\frac{\partial z_k}{\partial net_k}=(z_k-t_k)f'(net_k)$$

$$\therefore \frac{\partial J}{\partial w_{kj}}=(z_k-t_k)f'(net_k)y_j$$

考虑输入层到隐藏层

$$\frac{\partial J}{\partial w_{ji}}=\frac{\partial J}{\partial y_j}\frac{\partial y_j}{\partial net_j}\frac{\partial net_j}{\partial w_{ji}}$$

其中

$$\begin{aligned}\frac{\partial J}{\partial y_j}&=\sum_{k=1}^{c}\frac{\partial J}{\partial z_k}\frac{\partial z_k}{\partial y_j}\\&=\sum_{k=1}^{c}(z_k-t_k)\frac{\partial z_k}{\partial net_k}\frac{\partial net_k}{\partial y_j}\\&=\sum_{k=1}^{c}(z_k-t_k)f'(net_k)w_{kj}\end{aligned}$$

$$\begin{aligned}\therefore \frac{\partial J}{\partial w_{ji}}&=\sum_{k=1}^{c}(z_k-t_k)f'(net_k)w_{kj}\cdot f'(net_j)x_i\\&=\sum_{k=1}^{c}\delta_k w_{kj}\cdot f'(net_j)x_i\\&=\delta_j x_i\end{aligned}$$

贝叶斯理论

当采用均方差准则进行反向传播训练时，多层神经网络可产生一个相应于贝叶斯判别函数的最小二乘判别。