# 人工智能基础作业 5

2.

证明：

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i+\hat{y}_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2+2\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\overline{y})$$

将 $\hat{y}_i=wx_i+b$ 代入上式等式右边第三项，可得

$$2\sum_{i=1}^{n}(y_i-wx_i-b)(wx_i+b-\overline{y})$$

$$=2\sum_{i=1}^{n}(y_i-wx_i-\overline{y}+w\overline{x})(wx_i+\overline{y}-w\overline{x}-\overline{y})$$

$$=2w\sum_{i=1}^{n}(y_i-\overline{y}-w(x_i-\overline{x}))(x_i-\overline{x})$$

$$=2w\sum_{i=1}^{n}(y_i-\overline{y})(x_i-\overline{x})-2w^2\sum_{i=1}^{n}(x_i-\overline{x})(x_i-\overline{x})$$

$$=\frac{S_{xy}}{S_{xx}}S_{xy}-\left(\frac{S_{xy}}{S_{xx}}\right)^2S_{xx}=0$$

证毕.

3. Softmax 模型的推导

解：因为该式描述一个概率分布取到某个具体值的概率，

所以对于所有的 K 个分类，有 $\sum_{i=1}^{K}P(Y=i)=1$

将 $\log P(Y=k)=\beta_k x-\log Z\Rightarrow P(Y=k)=\frac{e^{\beta_k x}}{z}$ 代入上式

可得 $\sum_{j=1}^{K}\frac{e^{\beta_j x}}{z}=1$ 即 $z=\sum_{j=1}^{K}e^{\beta_j x}$，代回原式有 $P(Y=k)=\frac{e^{\beta_k x}}{\sum_{j=1}^{K}e^{\beta_j x}}$，证毕.