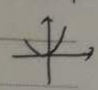
二、求解偏置b的公式推导

推导思路：由最小二乘法导出损失函数 $E(w,b)$

↓

证明损失函数 $E(w,b)$ 是关于 w 和 b 的 凸函数 → 形如 $f(x)=x^2$



↓

对损失函数 $E(w,b)$ 关于 b 求一阶偏导数

↓

令一阶偏导数等于0解出 b.

由最小二乘法导出损失函数 $E(w,b)$：

$$E(w,b)=\sum_{i=1}^{m}(y_i-f(x_i))^2$$

均方误差、平方损失。令 $f(x_i)=wx_i+b$；

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

$$=\sum_{i=1}^{m}(y_i-wx_i-b)^2$$

二元函数判断凹凸性：

设 $f(x,y)$ 在区域D上具有二阶连续偏导数，记 $A=f''_{xx}(x,y)$　$B=f''_{xy}(x,y)$，$C=f''_{yy}(x,y)$

则：(1) 在D上恒有 $A>0$，且 $AC-B^2\geq 0$ 时，$f(x,y)$ 在区域D上是凸函数；

(2) 在D上恒有 $A<0$，且 $AC-B^2\geq 0$ 时，$f(x,y)$ 在区域D上是凹函数。

二元凹凸函数求最值：

设 $f(x,y)$ 是在开区域D内具有连续偏导数的凸（或者凹）函数，$(x_0,y_0)\in D$ 且 $f'_x(x_0,y_0)=0$，$f'_y(x_0,y_0)=0$ 则 $f(x_0,y_0)$ 必为 $f(x,y)$ 在D内的最小值或最大值.

②

证明损失函数 $E(w,b)$ 是关于 w 和 b 的凸函数 —— 求 $A=f''_{xx}(x,y)$：

$$\frac{\partial E(w,b)}{\partial w}=\frac{\partial}{\partial w}\left[\sum_{i=1}^{m}(y_i-wx_i-b)^2\right]$$
$$=\sum_{i=1}^{m}\frac{\partial}{\partial w}(y_i-wx_i-b)^2$$
$$=\sum_{i=1}^{m}2(y_i-wx_i-b)\cdot(-x_i)$$
$$=2\left(w\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}(y_i-b)x_i\right)$$

$$\frac{\partial^2 E(w,b)}{\partial w^2}=\frac{\partial}{\partial w}\left(\frac{\partial E(w,b)}{\partial w}\right)$$
$$=\frac{\partial}{\partial w}\left[2\left(w\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}(y_i-b)x_i\right)\right]$$
$$=\frac{\partial}{\partial w}\left[2w\sum_{i=1}^{m}x_i^2\right]$$
$$=2\sum_{i=1}^{m}x_i^2$$ ，此即为 $A=f''_{xx}(x,y)$

证明损失函数 $E(w,b)$ 是关于 w 和 b 的凸函数 —— 求 $B=f''_{xy}(x,y)$

$$\frac{\partial^2 E(w,b)}{\partial w\partial b}=\frac{\partial}{\partial b}\left(\frac{\partial E(w,b)}{\partial w}\right)$$
$$=\frac{\partial}{\partial b}\left[2\left(w\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}(y_i-b)x_i\right)\right]$$
$$=\frac{\partial}{\partial b}\left[-2\sum_{i=1}^{m}(y_i-b)x_i\right]$$
$$=\frac{\partial}{\partial b}\left(-2\sum_{i=1}^{m}y_ix_i+2\sum_{i=1}^{m}bx_i\right)$$
$$=\frac{\partial}{\partial b}\left(2\sum_{i=1}^{m}bx_i\right)$$
$$=2\sum_{i=1}^{m}x_i$$ ，此即为 $B=f''_{xy}(x,y)$

证明损失函数$E(w,b)$是关于w和b的凸函数——求$C=f''_{yy}(x,y)$:

$$\frac{\partial E(w,b)}{\partial b}=\frac{\partial}{\partial b}\left[\sum_{i=1}^{m}(y_i-wx_i-b)^2\right]$$

$$=\sum_{i=1}^{m}\frac{\partial}{\partial b}(y_i-wx_i-b)^2$$

$$=\sum_{i=1}^{m}2(y_i-wx_i-b)\cdot(-1)$$

$$=2\left(mb-\sum_{i=1}^{m}(y_i-wx_i)\right)$$

$$\frac{\partial^2 E(w,b)}{\partial b^2}=\frac{\partial}{\partial b}\left(\frac{\partial E(w,b)}{\partial b}\right)$$

$$=\frac{\partial}{\partial b}\left[2\left(mb-\sum_{i=1}^{m}(y_i-wx_i)\right)\right]$$

$$=\frac{\partial}{\partial b}(2mb)$$

$$=2m$$ 此即为$C=f''_{yy}(x,y)$

$A=2\sum_{i=1}^{m}x_i^2>0$ $\quad B=2\sum_{i=1}^{m}x_i$ $\quad C=2m$ $\quad AC-B^2\geqslant 0$，$A>0$

$$AC-B^2=2m\cdot 2\sum_{i=1}^{m}x_i^2-\left(2\sum_{i=1}^{m}x_i\right)^2$$

$$\boxed{\frac{1}{m}\sum_{i=1}^{m}x_i=\bar{x}}$$

$$=4m\sum_{i=1}^{m}x_i^2-4\left(\sum_{i=1}^{m}x_i\right)^2$$

$$=4m\sum_{i=1}^{m}x_i^2-4m\cdot\frac{1}{m}\left(\sum_{i=1}^{m}x_i\right)^2$$

$$=4m\sum_{i=1}^{m}x_i^2-4m\cdot\bar{x}\cdot\sum_{i=1}^{m}x_i$$

$$=4m\left(\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}x_i\bar{x}\right)$$

$$=4m\sum_{i=1}^{m}(x_i^2-x_i\bar{x})$$

又$\because \sum_{i=1}^{m}x_i\bar{x}=\bar{x}\sum_{i=1}^{m}x_i=\bar{x}\cdot m\cdot\frac{1}{m}\cdot\sum_{i=1}^{m}x_i=m\bar{x}^2=\sum_{i=1}^{m}\bar{x}^2$

$$=4m\sum_{i=1}^{m}(x_i^2-x_i\bar{x}-x_i\bar{x}+x_i\bar{x})=4m\sum_{i=1}^{m}(x_i^2-x_i\bar{x}-x_i\bar{x}+\bar{x}^2)$$

$$=4m\sum_{i=1}^{m}(x_i-\bar{x})^2$$

④ 令一阶偏导数等于0解出b:

$$\frac{\partial E(w,b)}{\partial b}=2\left(mb-\sum_{i=1}^{m}(y_i-mx_i)\right)=0$$

$$mb-\sum_{i=1}^{m}(y_i-mx_i)=0$$

$$b=\frac{1}{m}\sum_{i=1}^{m}(y_i-wx_i)$$

$$b=\frac{1}{m}\sum_{i=1}^{m}y_i-w\cdot\frac{1}{m}\sum_{i=1}^{m}x_i$$

$$=\bar{y}-w\bar{x}$$

二. 求解权重w的公式推导

推导思路: 由最小二乘法导出损失函数$E(w,b)$

↓

证明损失函数$E(w,b)$是关于w和b的凸函数

↓

对损失函数$E(w,b)$关于w求一阶偏导数

↓

令一阶偏导数等于0解出w.

令一阶偏导数等于0解出w:

$$\frac{\partial E(w,b)}{\partial w}=2\left(w\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}(y_i-b)x_i\right)=0$$

$$w\sum_{i=1}^{m}x_i^2-\sum_{i=1}^{m}(y_i-b)x_i=0$$

$$w\sum_{i=1}^{m}x_i^2=\sum_{i=1}^{m}y_ix_i-\sum_{i=1}^{m}bx_i$$ 将$b=\bar{y}-w\bar{x}$代入

$$w\sum_{i=1}^{m}x_i^2=\sum_{i=1}^{m}y_ix_i-\sum_{i=1}^{m}(\bar{y}-w\bar{x})x_i$$

$$w\sum_{i=1}^{m}x_i^2=\sum_{i=1}^{m}y_ix_i-\bar{y}\sum_{i=1}^{m}x_i+w\bar{x}\sum_{i=1}^{m}x_i$$

$$w\left(\sum_{i=1}^{m}x_i^2-\bar{x}\sum_{i=1}^{m}x_i\right)=\sum_{i=1}^{m}y_ix_i-\bar{y}\sum_{i=1}^{m}x_i$$

$$w=\frac{\sum_{i=1}^{m}y_ix_i-\boxed{\bar{y}\sum_{i=1}^{m}x_i}}{\sum_{i=1}^{m}x_i^2-\boxed{\bar{x}\sum_{i=1}^{m}x_i}}$$

$$\bar{y}\sum_{i=1}^{m}x_i=\frac{1}{m}\sum_{i=1}^{m}y_i\sum_{i=1}^{m}x_i=\bar{x}\sum_{i=1}^{m}y_i$$

$$\bar{x}\sum_{i=1}^{m}x_i=\frac{1}{m}\sum_{i=1}^{m}x_i\sum_{i=1}^{m}x_i=\frac{1}{m}\left(\sum_{i=1}^{m}x_i\right)^2$$

$$w=\frac{\sum_{i=1}^{m} y_i x_i-\bar{x}\sum_{i=1}^{m} y_i}{\sum_{i=1}^{m} x_i^2-\frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2}=\frac{\sum_{i=1}^{m} y_i(x_i-\bar{x})}{\sum_{i=1}^{m} x_i^2-\frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2}$$

将w向量化. Vectorize w

$$w=\frac{\sum_{i=1}^{m} y_i(x_i-\bar{x})}{\sum_{i=1}^{m} x_i^2-\frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2}$$

将 $\frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2=\bar{x}\sum_{i=1}^{m} x_i=\sum_{i=1}^{m}\bar{x}x_i$ 代入分母:

$$w=\frac{\sum_{i=1}^{m} y_i(x_i-\bar{x})}{\sum_{i=1}^{m} x_i^2-\sum_{i=1}^{m} x_i\bar{x}}$$

$$=\frac{\sum_{i=1}^{m}(y_i x_i-y_i\bar{x})}{\sum_{i=1}^{m}(x_i^2-x_i\bar{x})}$$

由于

$$\begin{cases}\sum_{i=1}^{m} y_i\bar{x}=\bar{x}\sum_{i=1}^{m} y_i=\frac{1}{m}\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i=\sum_{i=1}^{m} x_i\cdot\frac{1}{m}\sum_{i=1}^{m} y_i=\sum_{i=1}^{m} x_i\bar{y}\\ \sum_{i=1}^{m} y_i\bar{x}=\bar{x}\sum_{i=1}^{m} y_i=\bar{x}\cdot m\cdot\frac{1}{m}\sum_{i=1}^{m} y_i=m\bar{x}\bar{y}=\sum_{i=1}^{m}\bar{x}\bar{y}\\ \sum_{i=1}^{m} x_i\bar{x}=\bar{x}\sum_{i=1}^{m} x_i=\bar{x}\cdot m\cdot\frac{1}{m}\sum_{i=1}^{m} x_i=m\bar{x}^2=\sum_{i=1}^{m}\bar{x}^2\end{cases}$$

$$w=\frac{\sum_{i=1}^{m}(y_i x_i-y_i\bar{x})}{\sum_{i=1}^{m}(x_i^2-x_i\bar{x})}=\frac{\sum_{i=1}^{m}(y_i x_i-y_i\bar{x}-y_i\bar{x}+y_i\bar{x})}{\sum_{i=1}^{m}(x_i^2-x_i\bar{x}-x_i\bar{x}+x_i\bar{x})}$$

$$=\frac{\sum_{i=1}^{m}(y_i x_i-y_i\bar{x}-x_i\bar{y}+\bar{x}\bar{y})}{\sum_{i=1}^{m}(x_i^2-x_i\bar{x}-x_i\bar{x}+\bar{x}^2)}$$

$$=\frac{\sum_{i=1}^{m}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{m}(x_i-\bar{x})^2}$$

⑥

将w向量化

$$x = (x_1, x_2, \cdots x_m)^T$$

$$y = (y_1, y_2, \cdots y_m)^T$$

$$x_d = (x_1 - \bar{x}, x_2 - \bar{x}, \cdots x_m - \bar{x})^T$$

$$y_d = (y_1 - \bar{y}, y_2 - \bar{y}, \cdots y_m - \bar{y})^T$$

$$w = \frac{\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{m}(x_i - \bar{x})^2} = \frac{x_d^T y_d}{x_d^T x_d}$$