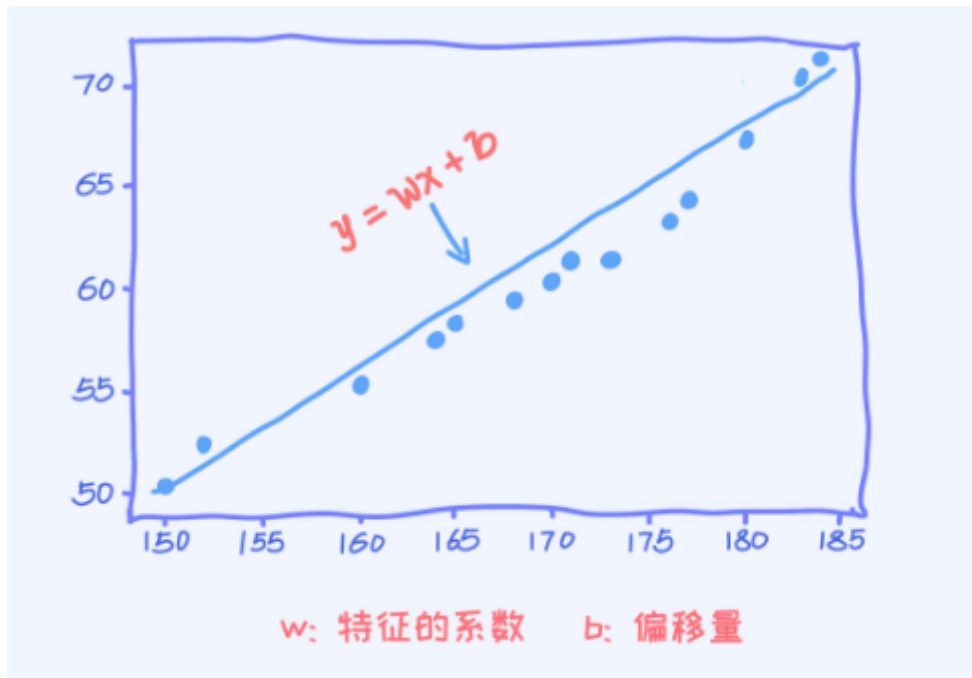


一、一元线性回归模型

线性回归特点：线性--具有线性的决策边界 回归--解决回归类的问题

- 模型简单且适用于大数据，训练速度快，资源消耗少
- 模型本身有非常好得可解释性--特征的有效、无效
- 变量x与y之间的线性关系



损失函数 (Loss Function) :

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

预测与测量值之间的差异成为误差，一般用最小二乘法来计算--距离的平方

线性回归的优化

假设我们定义数据集 $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$, N 是样本总数

$$l = \sum_{i=1}^N (wx_i + b - y_i)^2$$
$$\frac{\partial l}{\partial b} = \sum_{i=1}^N 2(wx_i + b - y_i) \cdot 1 = 2 \left(\sum_{i=1}^N wx_i + b - y_i \right) = 0$$
$$\Rightarrow w \sum_{i=1}^N x_i + \sum_{i=1}^N b - \sum_{i=1}^N y_i = 0$$
$$\Rightarrow w \cdot N \cdot \bar{x} + bN - N\bar{y} = 0$$
$$b = \bar{y} - w\bar{x}$$
$$\frac{\partial l}{\partial w} = \sum_{i=1}^N 2(wx_i + b - y_i) \cdot x_i = 0$$
$$\Rightarrow \sum_{i=1}^N (wx_i + \bar{y} - w\bar{x} - y_i) \cdot x_i = 0$$
$$\Rightarrow w \sum_{i=1}^N x_i^2 + \sum_{i=1}^N \bar{y} \cdot x_i - \sum_{i=1}^N w\bar{x} \cdot x_i - \sum_{i=1}^N x_i y_i = 0$$
$$\Rightarrow w \left(\sum_{i=1}^N x_i^2 - \bar{x} \sum_{i=1}^N x_i \right) = \sum_{i=1}^N x_i y_i - \bar{y} \sum_{i=1}^N x_i$$
$$\Rightarrow w(N\bar{x}^2 - N\bar{x}^2) = N\bar{x}\bar{y} - \bar{y} \cdot \bar{x} \cdot N$$
$$w = \frac{\bar{xy} - \bar{x} \cdot \bar{y}}{\bar{x}^2 - \bar{x}^2}$$

统计量:

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

二、多元线性回归模型

涉及向量、矩阵、张量知识。

向量的范数(Vector Norm):

- L2范数

L2范数

$$\|X\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$$

$$\|X\|_2^2 = x_1^2 + x_2^2 + \dots + x_d^2$$

假设我们定义数据集 $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$, 其中 $x_i \in \mathbb{R}^d$, N 是样本总数 d 是特征维度

$$l = \sum_{i=1}^N \left(\sum_{j=1}^d w_j x_{ij} + b - y_i \right)^2$$

$$\| \epsilon \|_2^2 = \epsilon_1^2 + \epsilon_2^2 + \dots + \epsilon_N^2$$

$$\epsilon = X \cdot w - y$$

$$w = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{pmatrix} \quad X = \begin{pmatrix} 1, x_{11}, x_{12}, \dots, x_{1d} \\ 1, x_{21}, x_{22}, \dots, x_{2d} \\ \vdots \\ 1, x_{N1}, x_{N2}, \dots, x_{Nd} \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

$$l = \|X \cdot w - y\|_2^2$$

$$X \cdot w = \begin{pmatrix} y_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{pmatrix} \rightarrow w_0 + x_{11} \cdot w_1 + x_{12} \cdot w_2 + \dots + x_{1d} \cdot w_d$$

$$= \sum_{j=1}^d w_j x_{ij} + (w_0) \Leftrightarrow b$$

\downarrow $N(d+1)$ \downarrow $(d+1)+1$

The Matrix Cookbook:多元线性回归损失函数求导法则参考书

主要第二章

求解参数W



$$L = \|Xw - y\|_2^2$$

$$\|x\|_2^2 = x^T \cdot x$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$(x \cdot y)^T = y^T \cdot x^T$$

$$a^T = a$$

$$L = (Xw - y)^T \cdot (Xw - y)$$

$$= (Xw)^T \cdot Xw - y^T \cdot Xw - (Xw)^T \cdot y + y^T \cdot y$$

$$= w^T \cdot X^T \cdot X \cdot w - \underline{y^T Xw} - \underline{w^T X^T y} + y^T \cdot y$$

$$= w^T X^T X w - 2w^T X^T y + y^T y$$

$$\frac{\partial L}{\partial w} = 2 \cdot X^T X w - 2 \cdot X^T y = 0$$

$$\Rightarrow \underline{X^T X} \cdot w = X^T \cdot y$$

$$(\underline{X^T X})^{-1} \cdot \underline{X^T X} \cdot w = (X^T X)^{-1} \cdot X^T \cdot y$$

$$w = (X^T X)^{-1} \cdot X^T y$$

在多元线性回归中求解参数时矩阵不一定可逆--行或列线性相关即非满秩矩阵，这时需要采取措施：

- 去掉线性相关的特征
- 在 $X^T X$ 的基础上加一个单位矩阵： $X^T X + E$

把导数设置为零的方式来求解参数，这种方法也叫做解析解(Analytic Solution)。