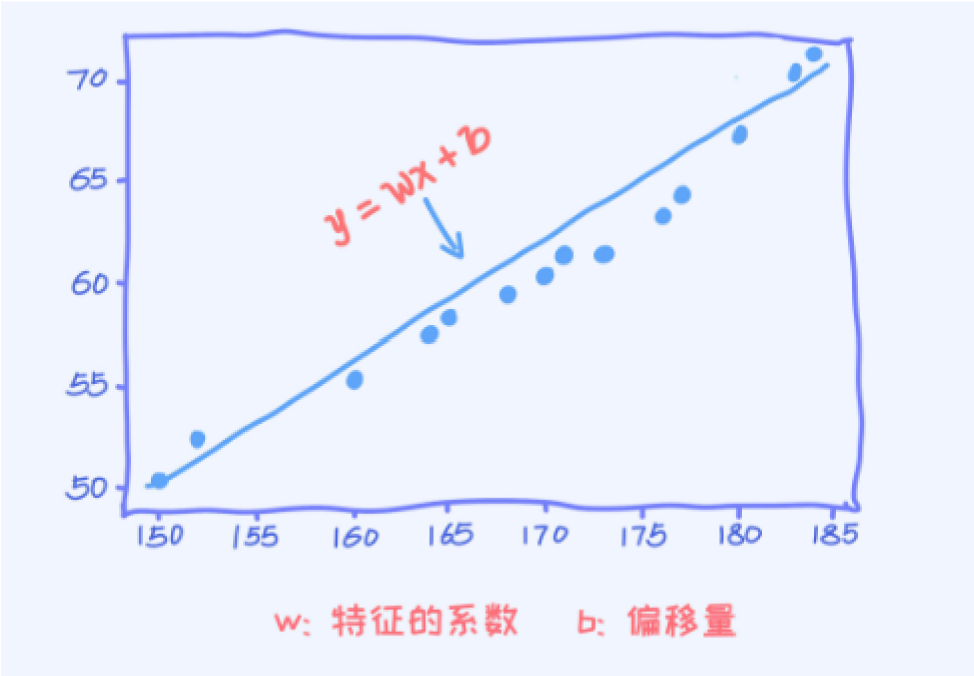
一、一元线性回归模型

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，训练速度块，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

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

$$l=\sum_{i=1}^{N}(\boxed{\overset{\hat{y}_i}{wx_i+b}}-y_i)^2$$

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

$$=2\left(\sum_{i=1}^{N}wx_i+b-y_i\right)=0$$

$$\Rightarrow w\sum_{i=1}^{N}x_i+\sum_{i=1}^{N}b-\sum_{i=1}^{N}y_i=0$$

$$\Rightarrow w\cdot N\cdot\bar{x}+bN-N\bar{y}=0$$

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

$$\frac{\partial l}{\partial w}=\sum_{i=1}^{N}2(wx_i+b-y_i)\cdot x_i=0$$

$$\Rightarrow\sum_{i=1}^{N}(wx_i+\bar{y}-w\bar{x}-y_i)\cdot x_i=0$$

$$\Rightarrow w\sum_{i=1}^{N}x_i^2+\sum_{i=1}^{N}\bar{y}\cdot x_i-\sum_{i=1}^{N}w\bar{x}\cdot x_i-\sum_{i=1}^{N}x_iy_i=0$$

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

$$\Rightarrow w(N\cdot\overline{x^2}-N\cdot\bar{x}^2)=N\cdot\overline{xy}-\bar{y}\cdot\bar{x}\cdot N$$

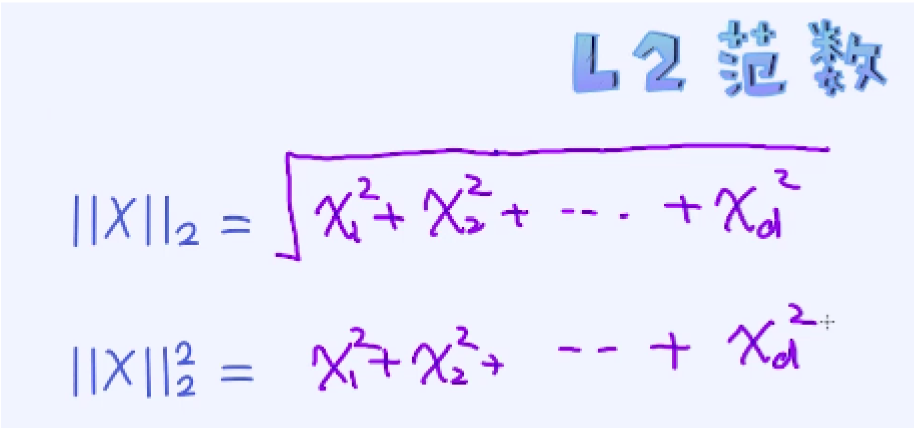
$$w=\frac{\overline{xy}-\bar{x}\cdot\bar{y}}{\overline{x^2}-\bar{x}^2}$$

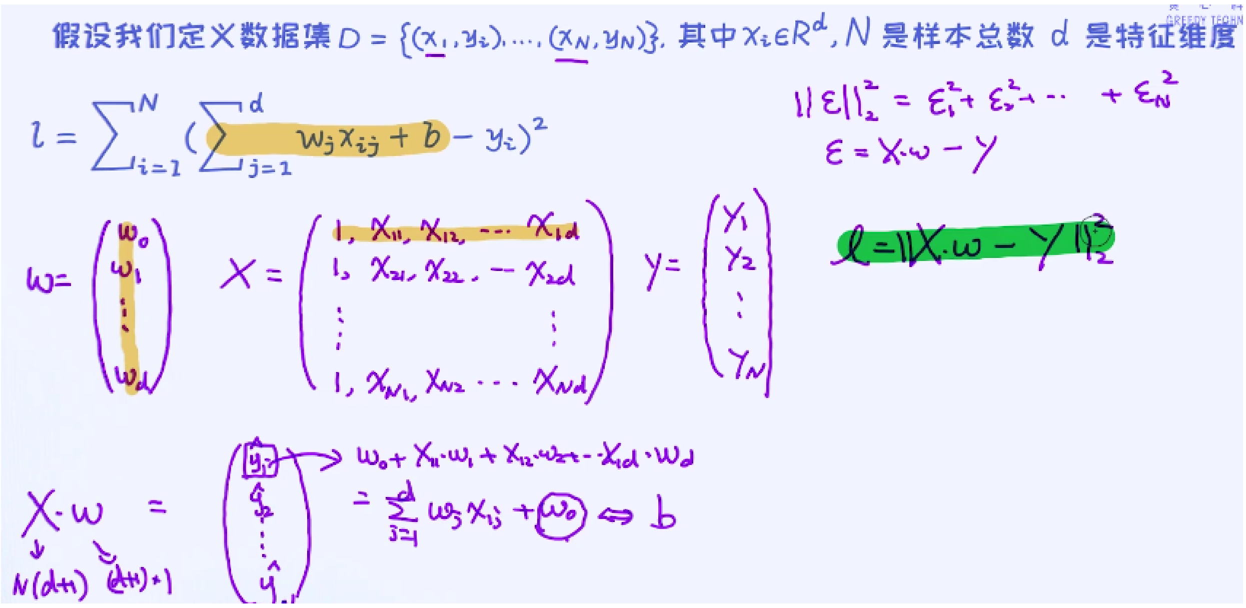
二、多元线性回归模型

涉及向量、矩阵、张量知识。

向量的范数(Vector Norm)：

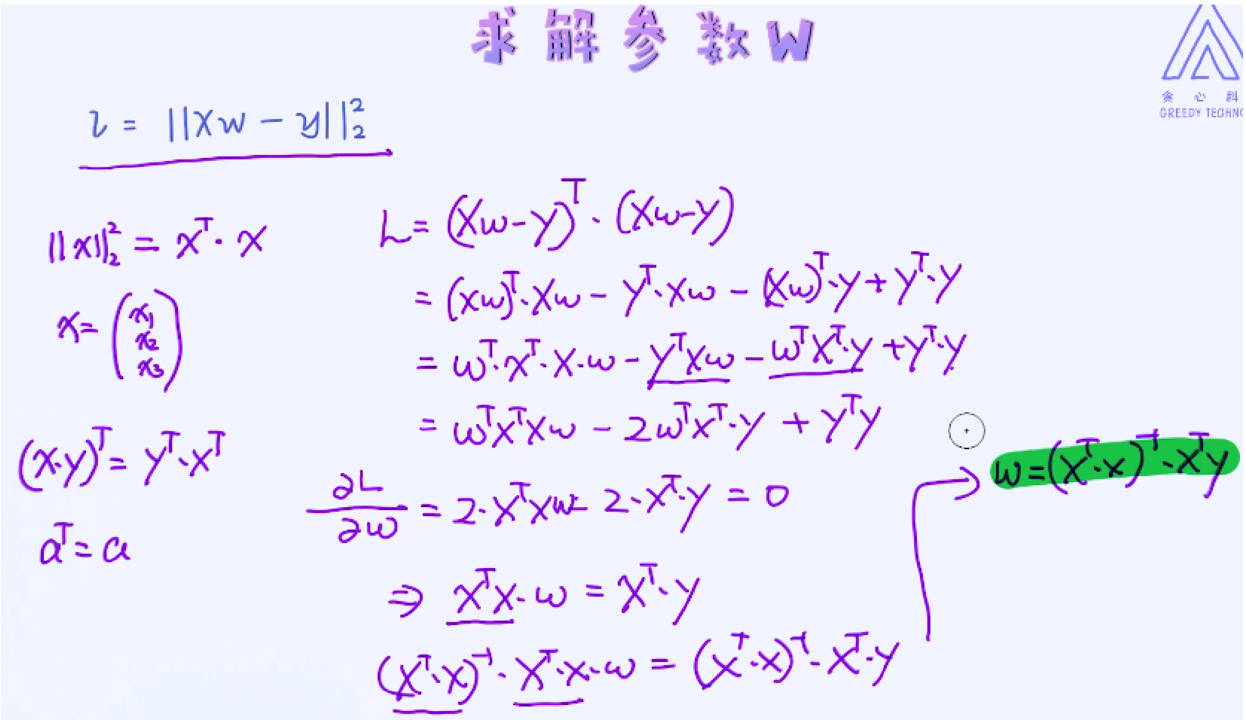
- L2范数





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



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XtX的基础上加一个单位矩阵：XtX + 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)。