线性回归的含义

线性

1. 线性关系：对于一元函数，因变量与自变量的关系能用一条直线表示，或导数为常数，则称二者为线性关系。
2. 线性运算：加减和数乘统称为线性运算。

回归

回归分析是研究因变量与自变量关系（定量关系）的一种技术，是训练一个回归函数来输出一个数值，常用于预测分析。回归也是最简单的监督学习任务之一，多用于解决“有多少”问题。

线性回归

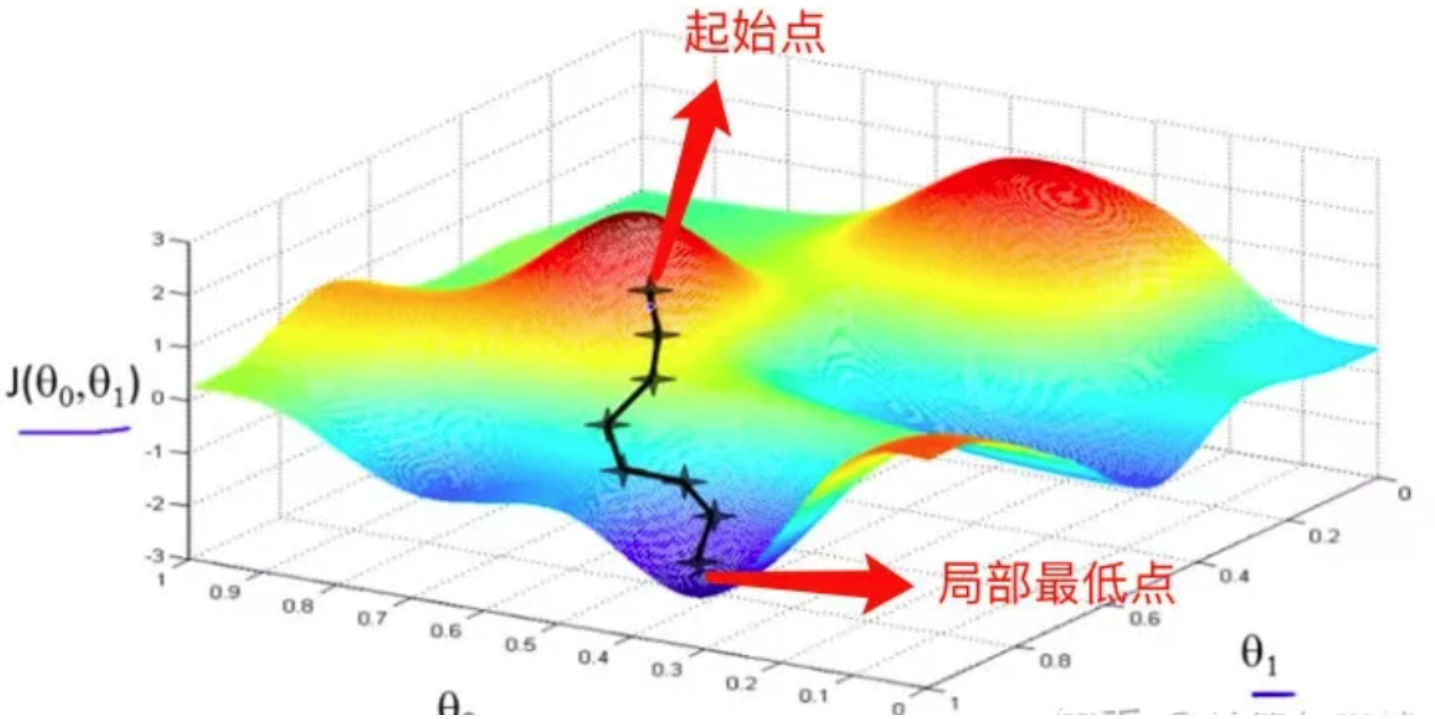
线性回归是研究具有线性关系的因变量与自变量的分析方法。
作用：

1. 通过一定的测试数据确定因变量与自变量定量关系后，可以通过自变量值预测因变量值。
2. 既然能确定定量关系，也就可以判断自变量与因变量的相关性强弱。
3. （个人的胡思幻想）线性回归本质是通过数据（自变量）确定数据（因变量），可以将因变量的变化量或者因变量大小作为判定条件执行代码，也就转化成了收集自变量数据->执行对应代码（操作），再加上非线性的回归分析，总的就可以实现观察->自动行动，人工智能的雏形?

梯度下降法

类比下山的过程：

假设现在有一个人在山上，现在他想要走下山，一定要沿着山高度下降的地方走。山高度下降的方向有很多，假定他选择最陡峭的方向，即山高度下降最快的方向。现在确定了方向，就要开始下山了。又有一个问题，在下山的过程中，最开始选定的方向并不总是高度下降最快的地方。所以选定一段距离，每走一段距离之后，就重新确定当前所在位置的高度下降最快的地方。这样，这个人每次下山的方向都可以近似看作是每个距离段内高度下降最快的地方。



下山与线性回归的对应关系

下山走的每一段路的距离-->**学习率**α
一次走一段距离-->**迭代**
山-->**损失函数J（θ)**
山底-->**损失函数最小的地方**
下山过程-->**求解参数矩阵使得损失函数最小**

算法过程：

1. 确定定参数的初始值，计算损失函数的偏导数
2. 将参数代入偏导数计算出梯度。若梯度为 0，结束；否则转到 3
3. 用步长乘以梯度，并对参数进行更新
4. 重复 2-3

拟合函数：$h_\theta(x)=\sum_{i=0}^{n}\theta_i x_i$
损失函数：$J(\theta)=\frac{1}{2m}\sum_{i=0}^{m}(y^{(i)}-h_\theta(x^{(i)}))^2$
损失函数的偏导数：$\frac{\partial J(\theta)}{\theta_i}=\frac{1}{m}\sum_{j=1}^{m}(h_{(\theta)}(x^{(j)})-y^{(i)})x_i^{(j)}$
每次更新参数的操作：$\theta_i=\theta_i-\frac{\alpha}{m}\sum_{j=1}^{m}(h_{(\theta)}(x^{(j)})-y^{(i)})x_i^{(j)}$

最小二乘法

假设所求关系式为 Y=$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon$。
ϵ称为残差，意为真实值与测量值之差。
将上式写成矩阵的形式：Y=$\beta X + \epsilon$。
线性回归就是要让ϵ越小越好，几乎为0则说明测量值与真实值非常接近。
那么，在上式中，也就是要找到一个β来使ϵ为0，上式的几何意义Y为超平面Xβ外一点，到超平面上某一点，使得ϵ(距离)最小。
⇒由于垂线段最短，当从Y点作垂线与平面相交时，ϵ最小。
⇒ϵ垂直于该超平面。
⇒ϵ垂直于该平面的所有基向量。
⇒$X^T \epsilon = X^T (Y - X\beta) = 0$。
解得$\beta = (X^T X)^{-1} \cdot X^T Y$。