**线性回归**

* 回归（regression）是能为一个或多个自变量与因变量之间关系建模的一类方法。
* 线性回归（linear regression）
  + 自变量和因变量之间的关系是线性的，即可以表示为中元素的加权和，通常允许包含观测值的一些噪声。
  + 任何噪声都比较正常，如噪声符合正态分布。
* 训练数据集（training data set）或训练集（training set）
* 样本（sample）或数据点（data point）或数据样本（data instance）
* 标签（label）或目标（target），例如房价预测中的房屋价格
* 特征（feature）或协变量（covariate），例如房价预测中的面积、房龄等

表示数据集中样本数，对索引为的样本，其特征表示为，其对应的标签是。

* 线性模型：
  + 是输入特征，。
  + 是权重，权重决定了每个特征对预测值的影响，。
  + 是偏置（bias）、偏移量（offset）或截距（intercept）。
  + 是输入特征的一个仿射变换（affine transformation），仿射变换的特点是通过加权和，对输入特征进行线性变换（linear transformation），并通过偏置项进行平移（translation）。