

数据相关

- 数据整体叫数据集 (data set)
- 每一行数据称为一个样本 (sample)
- 除最后一列，每一列表达样本的一个特征 (feature)
 - 特征可以很抽象
 - 对于图像，每一个像素点都是特征
 - 28×28 的（灰度图）图像有 $28 \times 28 = 784$ 个特征
 - 如果是彩色图像则特征更多
- 所有数据的特征用一个 X （矩阵）来表示
 - 第 i 个样本行写作 $X^{(i)}$ ，也可以说，这是第 i 个样本的特征向量。
 - 第 i 个样本第 j 个特征值 $X_j^{(i)}$
- 最后一列，称为标记 (label)，所有数据的标记，用一个 y 表示， y 是一个向量
 - 第 i 个样本的标记写作 $y^{(i)}$

- 数学上，我们通常把向量表示成列向量 $\begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}$ ，因此，整个数据的矩阵可以

表示为 $\begin{pmatrix} (X^{(1)})^T \\ (X^{(2)})^T \\ (X^{(3)})^T \\ \vdots \end{pmatrix}$ （它把每个特征向量看作是一个列向量，因此用 T 上标把它转置一下变为行向量）

- 特征空间 (feature space)
就是特征值分布的空间（二维空间，三维空间，四维空间.....）
- 分类任务本质就是在特征空间切分
- 在低维空间思考，再推断到高维空间

