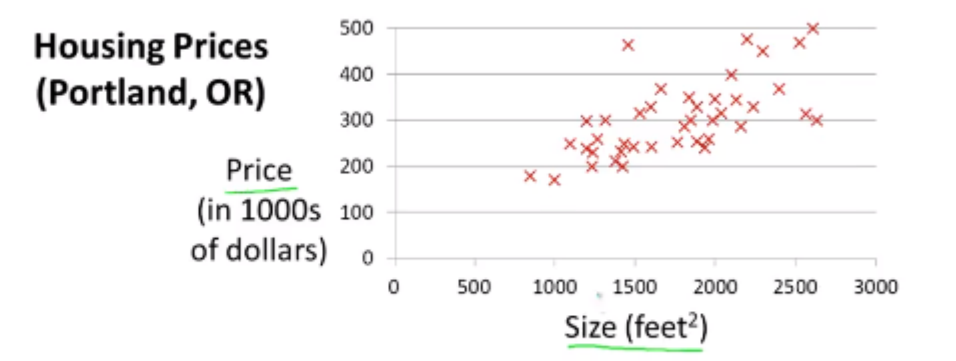
一、监督学习完整流程（房价预测为例）



给出正确答案，根据我们的数据来说房子实际的价格是多少（回归问题-根据之前的数据预测出一个准确的输出值）
另一种监督学习流程-分类问题（寻找癌症肿瘤问题）
当我们想要预测离散的输出值，例如寻找癌症肿瘤并想要确认肿瘤是良性的还是恶性的，这就是 0/1 离散输出问题。
在监督学习中我们有一个数据集（训练集，包含不同房屋价格），我们的任务就说从这个数据中学习预测房屋的价格。

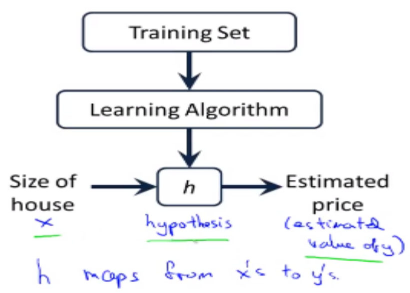
Training set of housing prices (Portland, OR)

Size in feet² (x)	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
…	…

常见的符号定义

Notation:
- **m** = Number of training examples
- **x's** = "input" variable / features
- **y's** = "output" variable / "target" variable

注意：m 代表训练样本的数量,(x，y)—x 作为训练样本，y 作为输出数据（结果）
(X(i)，y(i))i 上标，i 代表训练行
监督学习算法工作方式：



注：讲训练集里的房屋价格，放到学习算法中，然后输出一个函数，通常大小写成 h（h 代表 hypothesis(假设)h 表示一个函数）
输入的是房屋尺寸大小，就像想出售的房屋；h 根据输入的 x 值来得出 y 值

二、代价函数（最有可能的直线与我们的数据相拟合）

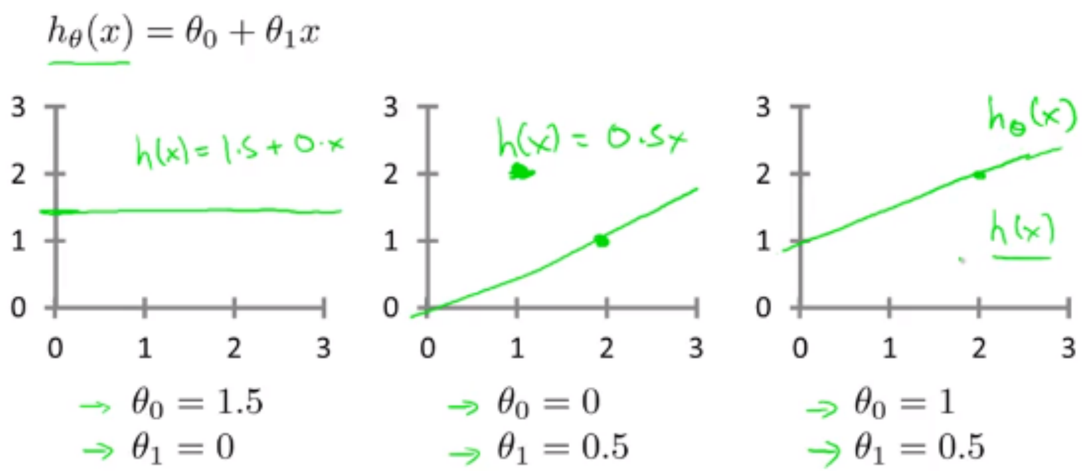
Training Set

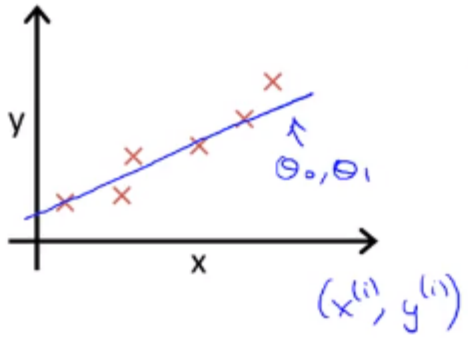
Size in feet² (x)	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
…	…

$m = 47$

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

通过 m（训练数据集），假设函数用来进行预测的函数就是这样是线性函数形式。引入术语 θ0，θ1（模型参数），主要讨论如何选择不同参数 θ0 和 θ1，我们会得到不同的假设函数，如下三组例子所示





$$\underset{\theta_0 \theta_1}{\text{minimize}} \ \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

\# training examples

Idea: Choose θ_0, θ_1 so that
$h_\theta(x)$ is close to y for our
training examples (x, y)

x , y

假设 θ0 和 θ1 是直线，得出这两个参数的值来假设函数表示的直线，尽量的与这些数据点很好的拟合。在我们训练集中我们会得到一定数量的样本，x 表示卖出哪所房子，并且知道这所房子的实际价格，所以我们尽量选择参数值，给出训练集中的 x 值，我们能合理的预测值，让我们给出标准的定义在线性回归中，我们要解决的是一个最小化问题。

$$\underset{\theta_0 \theta_1}{\text{minimize}} \ \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)$$

\# training examples

实现预测值与实际值之差平方值最小，尽量减少平均误差。
首先要定义一个代价函数

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

我们主要做的是关于 θ0 和 θ1 对函数(θ0,θ1)求最小值

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

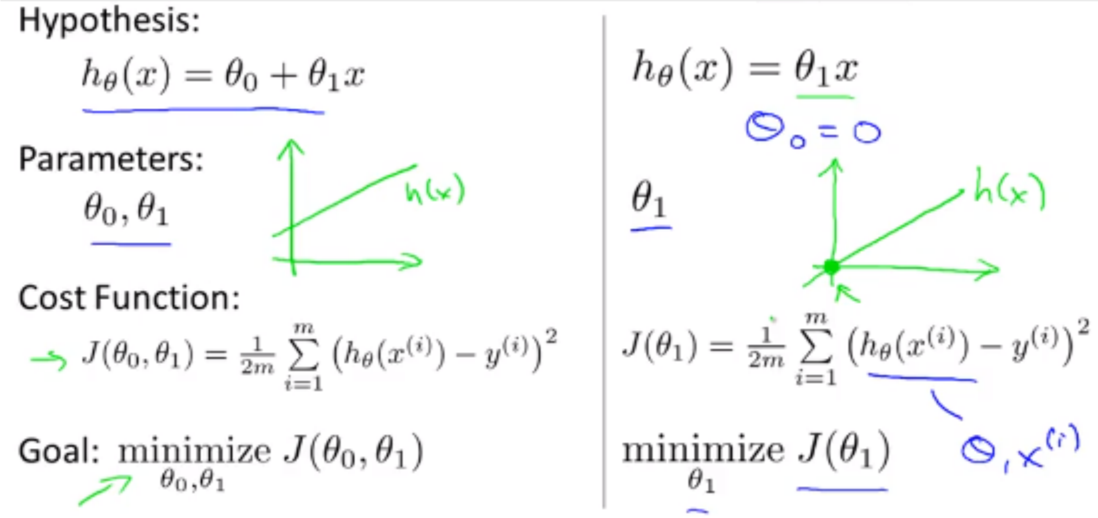
$$\underset{\theta_0, \theta_1}{\text{minimize}} \ \underbrace{J(\theta_0, \theta_1)}_{\text{Cost function}}$$

代价函数如上图所示，代价函数也被称为平方误差函数，如下图所示：

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

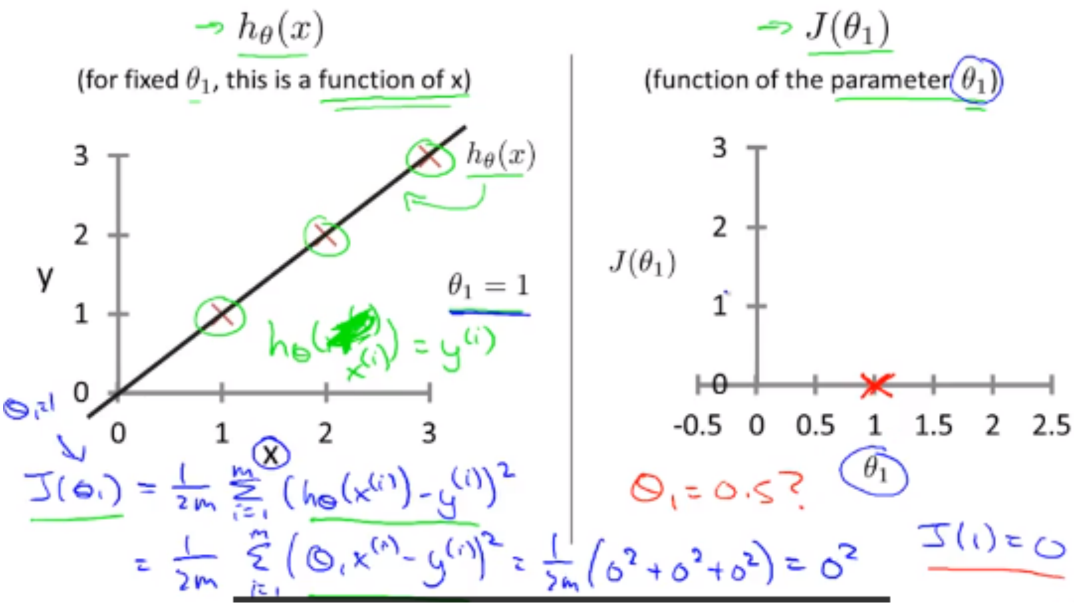
代价函数可视化

使用一个简化假设也就是 θ1*x 我们可以将这个函数看成是把 θ0 设为 0，所以只有一个参数就是 θ1 代价函数看起来与之前的很像唯一的区别是现在 h(x) 等于 θ1*x 只有一个参数 θ1 所以我的优化目标是将 J(θ1) 最小化 。用图形表示是过原点的直线。

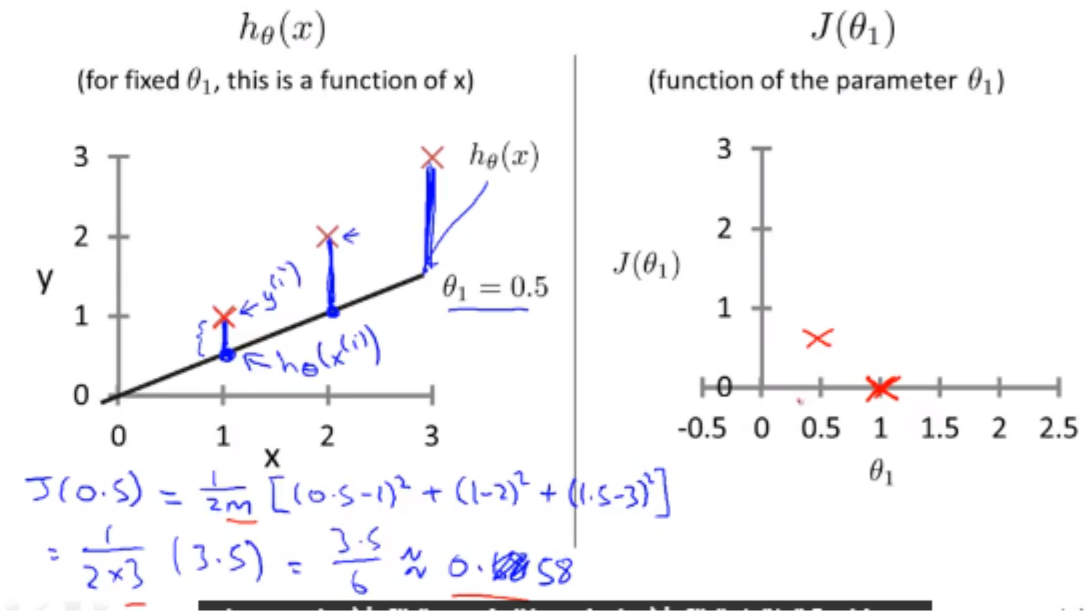


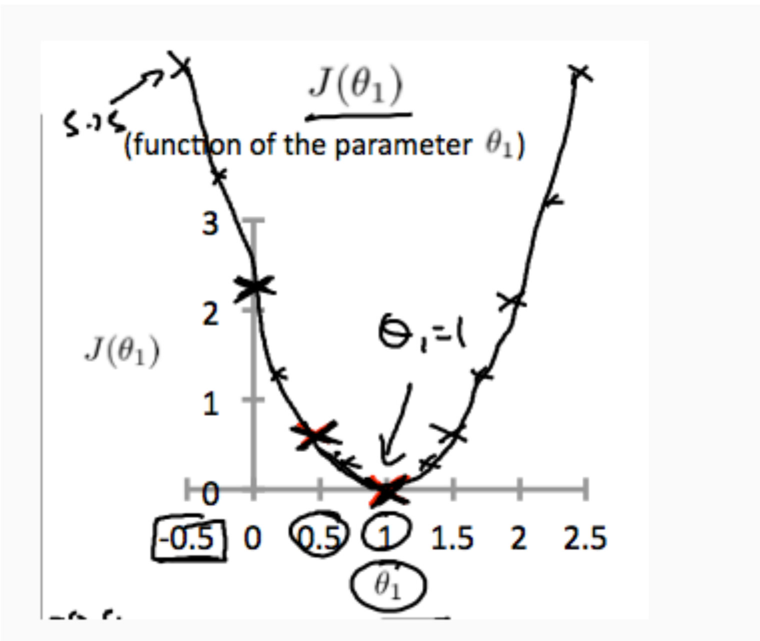
需要理解两个重要函数：第一个是假设函数。第二个是代价函数

假设函数：



代价函数：





最终目标：we should try to minimize the cost function. In this case, \theta_1 = 1ϑ1=1 is our global minimum

深度学习代价函数的作用

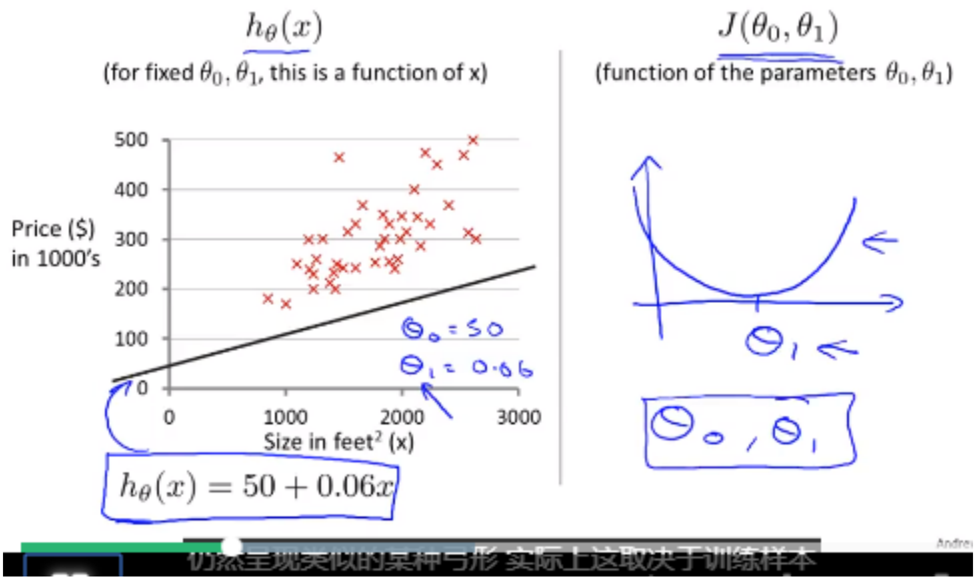
几个重要公式包括了假设 h、参数 θ、代价函数 J 以及优化目标

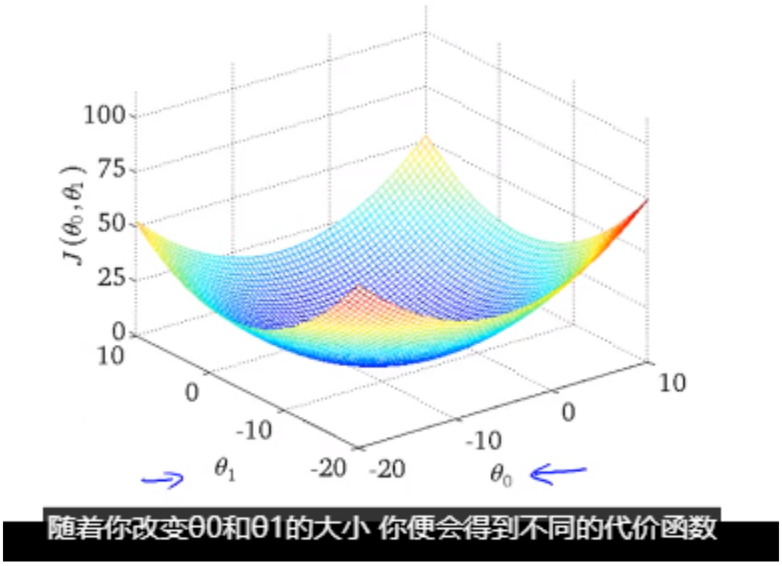
Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

Parameters: θ_0, θ_1

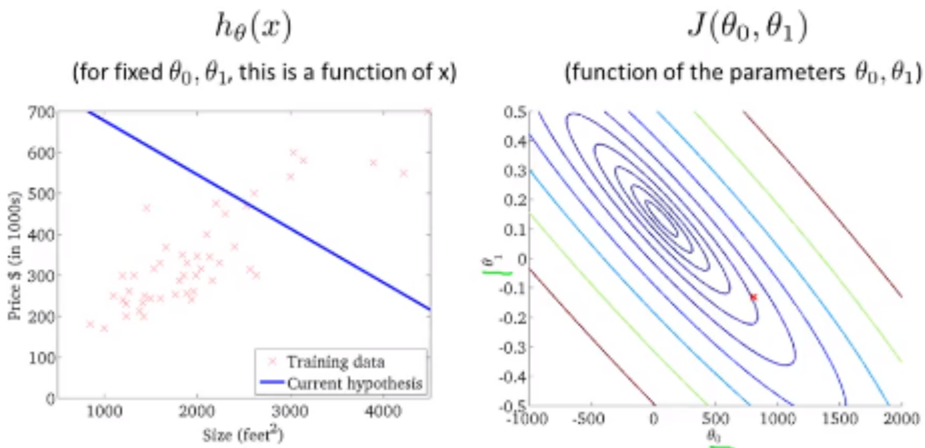
Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

Goal: $\underset{\theta_0,\theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$





竖直方向 的高度就表示代价函数 J(θ0,θ1) 的值不难发现这是一个弓形曲面 我们来看看三维图这是这个曲面的三维图 水平轴是 θ0、θ1 竖直方向表示 J(θ0,θ1) 旋转一下这个图 就更能理解这个弓形曲面所表示的代价函数了。



误差平方通过这些图形，更接近代价函数这些值。