监督学习：（1 or 0）

1.回归（regression）：指我们的目标是预测一个连续的输出值。目标是预测

离散值输出。

2.分类（classfication）：

无监督学习：（归类）

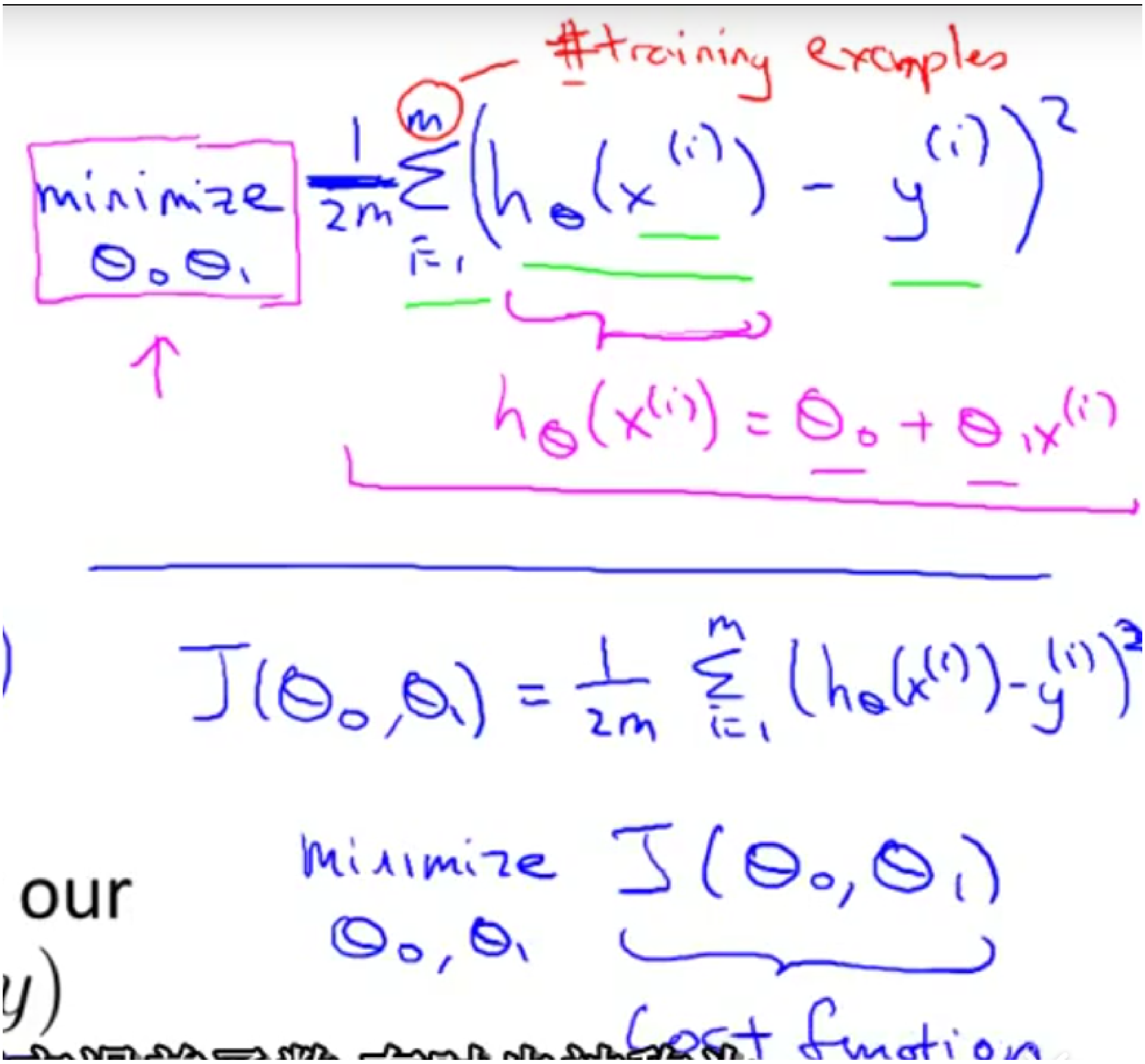
1.聚类算法：市场分割分类客户 社交群体分类组员

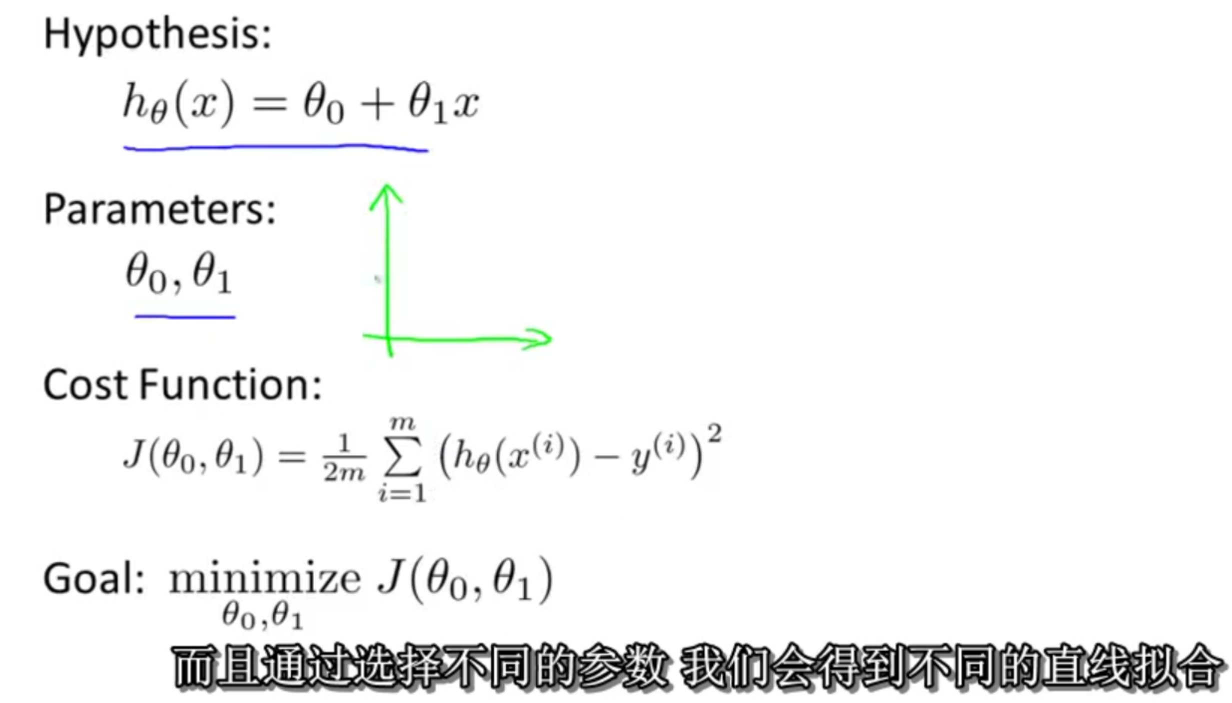
2.鸡尾酒算法：

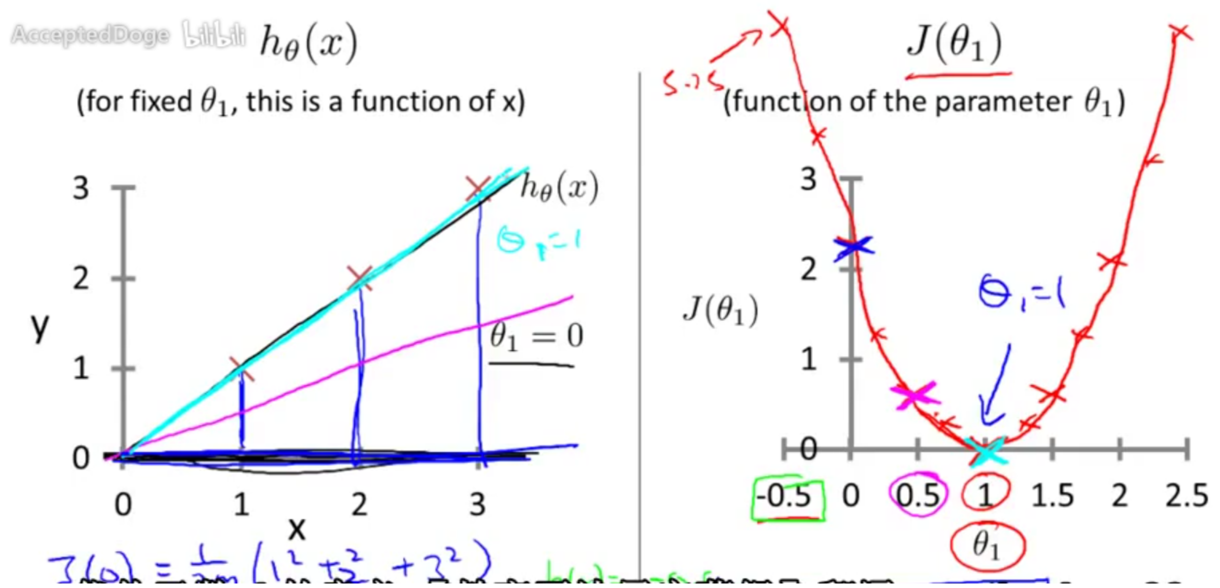
Octave软件

线性回归（linear regression）模型：

代价函数：平方误差函数




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: $\underset{\theta_0, \theta_1}{\text{minimize}} \; J(\theta_0, \theta_1)$
而且通过选择不同的参数 我们会得到不同的直线拟合


AcceptedDoge bilibili
$h_\theta(x)$
(for fixed θ_1, this is a function of x)
$h_\theta(x)$
$\theta_1 = 1$
$\theta_1 = 0$
y
x
$J(\theta_1)$
(function of the parameter θ_1)
5.25
$\theta_1 = 1$
$J(\theta_1)$
-0.5
0
0.5
1
1.5
2
2.5
θ_1

梯度下降：

Gradient descent algorithm

repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \qquad (\text{for } j = 0 \text{ and } j = 1)$$

}

Correct: Simultaneous update

$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$
$\text{temp1} := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$
$\theta_0 := \text{temp0}$
$\theta_1 := \text{temp1}$

让我们来看看 这个公式有很多细节问题

α-learning rate：以多大的幅度更新参数θj，j=0，1

导数项：下降斜率

α过小，下降过慢，α过大下降过快会错过最低点。