*NOTES:*

**Week1:**

**Definition**

A computer program is said to learn form experience E with respect to some Task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

一个程序被认为能从经验E中学习, 解决任务T, 达到性能度量值P，当且仅当，有了经验E后, 经过P评判, 程序在处理 T 时的性能有所提升.

**Supervised Learning**

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

Supervised learning problems are categorized into "regression" and "classification" problems.

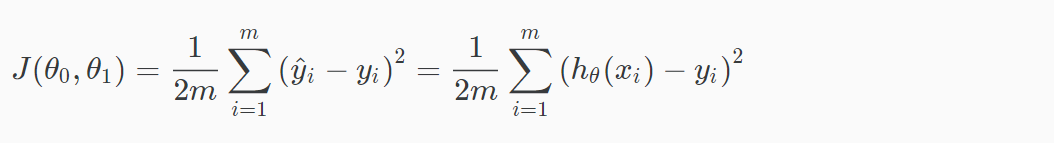
**Unsupervised Learning**

Unsupervised learning allows us to approach problems with little or no idea what our results should look like. We can derive (得到) structure from data where we don't necessarily know the effect of the variables. We can derive this structure by clustering the data based on relationships among the variables in the data.

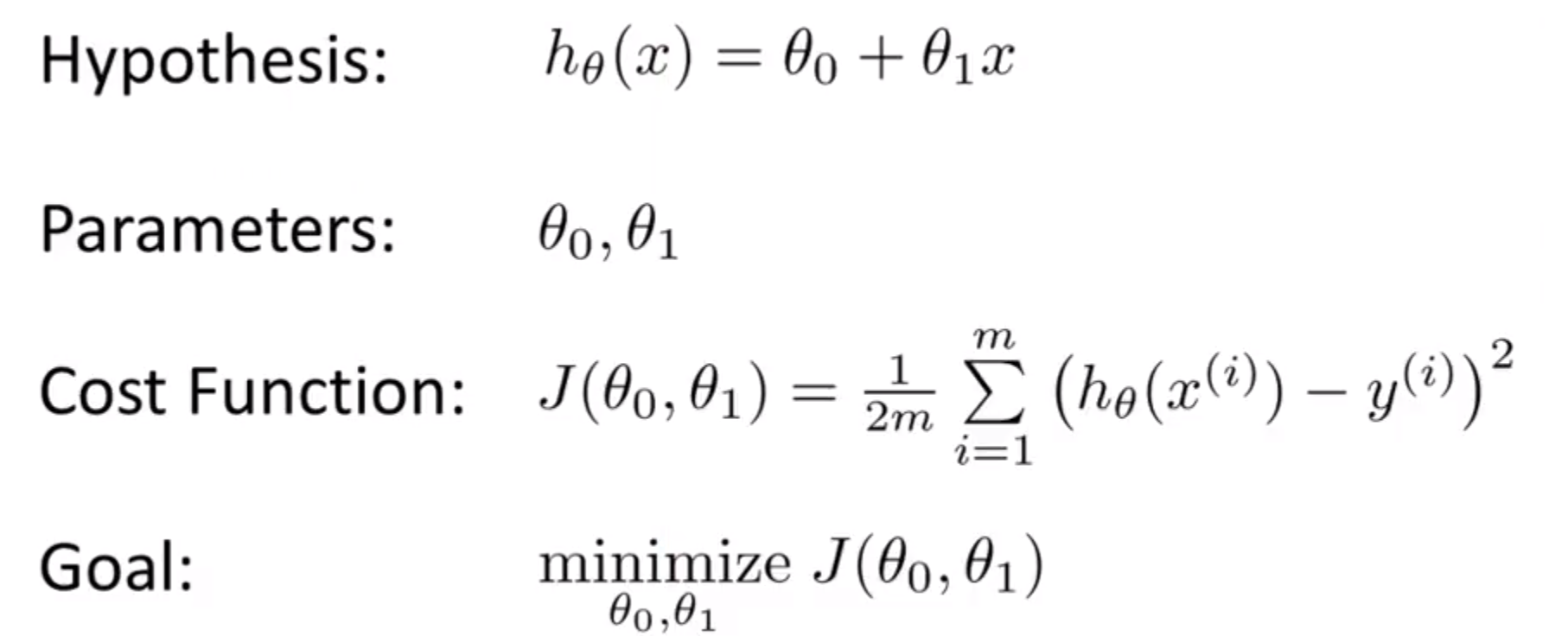
With unsupervised learning there is no feedback based on the prediction results.

**Learning Regression**

Cost function:



The 1/2 is as a convenience for the computation of the gradient descent, as the derivative term of the square function will cancel out the 1/2 term.



In order to minimize cost function, using gradient descent algorithm.

Hint: Difference between =(判断为真的声明) and :=(赋值)

**Gradient descent algorithm**

Definition of gradient (From neural computation course):

