# Chapter 1 Optimization Objective

## 1.1 Optimization Objective

We will introduce supported vector machine, which is supervised.

For a sigmoid function, if the y = 1, we want the $h_\theta(x) \approx 1$ and $\theta^T x \gg 0$.

Support vector machine simplifies the cost function of logistic regression into a straight line.

$$J(\theta) = \frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)}\text{cost}_1(\theta^T x^{(i)}) + (1 - y^{(i)})\text{cost}_0(\theta^T x^{(i)})\right] + \frac{\lambda}{2m}\sum_{i=0}^{n}\theta_i^2$$

As the SVM is two order, the co-efficient before the expression could be dismissed.

The hypothesis will be this:

$$h_\theta(x) = 1 \text{ if } \theta^T x \leq 0; 0, \text{ else}$$

## 1.2 Large Margin Intuition

SVM is called large margin intuition sometimes. The threshold 1 and -1 make the SVM "safer".

SVM would give a decision boundary that seperate the region with a margin. A large co-efficient will turn the SVM to a sensitive one.

The SVM decision boundary is decided by inner product.

## 1.3 Kernels

If we need a non-linear decision boundary, we can apply some landmarks and compute new features depending on proximity to landmarks.

The similarity could be expressed as:

$$\text{similarity}(x, l^{(i)}) = \exp(-\frac{||x - l^{(i)}||^2}{2\sigma^2})$$