Notes on machine learning

Daniel Miller

August 10, 2016

1 Supervised machine learning

1.1 Univariate linear regression

The basic idea is as follows. We have a set $\mathbf{x} = \{x^{(1)}, \ldots, x^{(m)}\}$ of "input variables," lying in some domain D, a set $\mathbf{y} = \{y^{(1)}, \ldots, y^{(m)}\}$ "output" or "target" variables in some range R, i.e. a map $[1, \ldots, m] \to D \times R$. Given this, we want to select a "hypothesis function" $h\colon D \to R$, such that $h(x) = y$ is a good fit for the data, i.e. $h(x^{(i)}) \approx y^{(i)}$ for some reasonable definition of $\approx$.

Univariate linear regression concerns $D = \mathbf{R}$, $R = \mathbf{R}$, and $h_\theta(x) = \theta_0 + \theta_1 x$. We try to find θ that minimizes the "cost function"

$$J_{\mathbf{x},\mathbf{y}}(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2.$$

The function $J_{\mathbf{x},\mathbf{y}}$ is quadratic in θ, so it should be easy to find the minimum point.