机器学习2

1 Multiple Features

Note:[$7 : 25 - \theta^T$is a 1 by (n+1) matrix and not an (n+1) by 1 matrix]

Linear regression with multiple variables is also known as "multivariate linear regression".

We now introduce notation for equations where we can have any number of input variables.

$x_j^{(i)}$=value of feature j in the i^{th} training example
$x^{(i)}$=the input (features) of the i^{th} training example
m=the number of training examples
n=the number of features

The multivariable form of the hypothesis function accommodating these multiple features is as follows:

$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \cdots + \theta_n x_n$

In order to develop intuition about this function, we can think about θ_0 as the basic price of a house, θ_1 as the price per square meter, θ_2 as the price per floor, etc. $x_1$will be the number of square meters in the house, x_2 the number of floors, etc.

Using the definition of matrix multiplication, our multivariable hypothesis function can be concisely represented as:

This is a vectorization of our hypothesis function for one training example; see the lessons on vectorization to learn more.

$$
h_\theta(x) = \begin{bmatrix} \theta_0 & \theta_1 & \cdots & \theta_n \end{bmatrix} * \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \theta^T x
$$

Remark: Note that for convenience reasons in this course we assume $x_0^{(i)}$ for (i $\in$ 1,...,m). This allows us to do matrix operations with theta and x. Hence making the two vectors $'\theta'$ and $x^{(i)}$match each other element-wise (that is, have the same number of elements: n+1).