

## Environment Setup Instructions

### Multivariate Linear Regression

- ✓ **Video:** Multiple Features  
8 min
- ✓ **Reading:** Multiple Features  
3 min
- ✓ **Video:** Gradient Descent for  
Multiple Variables  
5 min
- ✓ **Reading:** Gradient Descent  
For Multiple Variables  
2 min
- ✓ **Video:** Gradient Descent in  
Practice I - Feature Scaling  
8 min
- ✓ **Reading:** Gradient Descent  
in Practice I - Feature Scaling  
3 min
- ✓ **Video:** Gradient Descent in  
Practice II - Learning Rate  
8 min
- ✓ **Reading:** Gradient Descent  
in Practice II - Learning Rate  
4 min
- ✓ **Video:** Features and  
Polynomial Regression  
7 min
- ✓ **Reading:** Features and  
Polynomial Regression  
3 min

### Computing Parameters Analytically

### Submitting Programming Assignments

### Review

### Octave/Matlab Tutorial

### Review



# 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:

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

In order to develop intuition about this function, we can think about  $\theta_0$  as the basic price of a house,  $\theta_1$  as the price per square meter,  $\theta_2$  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:

$$h_{\theta}(x) = [\theta_0 \quad \theta_1 \quad \dots \quad \theta_n] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \theta^T x$$

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

Remark: Note that for convenience reasons in this course we assume  $x_0^{(i)} = 1$  for  $(i \in 1, \dots, m)$ . 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:  $n+1$ ).]

✓ Complete

Go to next item