

# CSE 291, Neural Networks, Winter 2015,

## Project 1

Bobak Hashemi, William Fedus, Matthew Burns

April 14, 2015

Say we have  $m$  datapoints over  $n$  features (house prices). We lay out the data as a corresponding, tall,  $m$  by  $n$  matrix:  $X$ .  $X$  is augmented with a column of ones and  $\theta$  an extra variable,  $\theta_0$  (an intercept term).  $y$  is packaged as a column vector.

$$\theta = [\theta_0, \theta_1, \dots]^T \tag{1}$$

$$x_i = X e_i \tag{2}$$

Where  $e_i$  is the  $i$ th unit basis vector in  $\mathcal{R}^n$ . This gives us the  $i$ th column of  $X$ .

$$y^{(i)} = x_i^T \theta = \langle x_i, \theta \rangle \tag{3}$$

$$y = [y^{(1)}, y^{(2)}, \dots]^T \tag{4}$$

$$y = X \theta \tag{5}$$