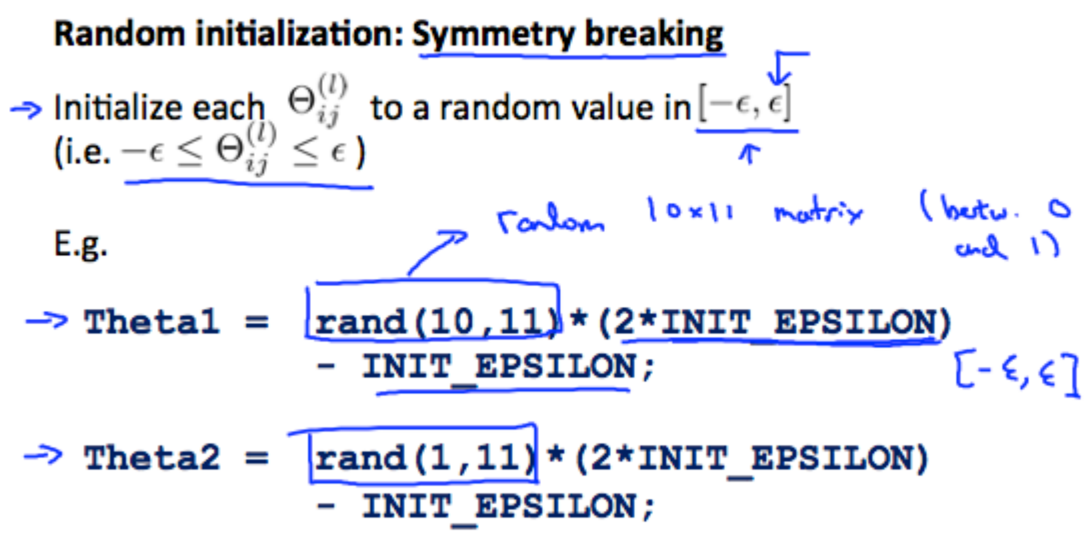
coursera | Stanford ONLINE

Random Initialization

Initializing all theta weights to zero does not work with neural networks. When we backpropagate, all r update to the same value repeatedly. Instead we can randomly initialize our weights for our Θ matrice following method:



Hence, we initialize each $\Theta_{ij}^{(l)}$ to a random value between $[-\epsilon, \epsilon]$. Using the above formula guarantee the desired bound. The same procedure applies to all the Θ's. Below is some working code you could experiment.

```
If the dimensions of Theta1 is 10x11, Theta2 is 10x11 and Theta3 is 1x11

Theta1 = rand(10,11) * (2 * INIT_EPSILON) - INIT_EPSILON;
Theta2 = rand(10,11) * (2 * INIT_EPSILON) - INIT_EPSILON;
Theta3 = rand(1,11) * (2 * INIT_EPSILON) - INIT_EPSILON;

```

rand(x,y) is just a function in octave that will initialize a matrix of random real numbers between 0 and

(Note: the epsilon used above is unrelated to the epsilon from Gradient Checking)

Go to next item ✓ Completed