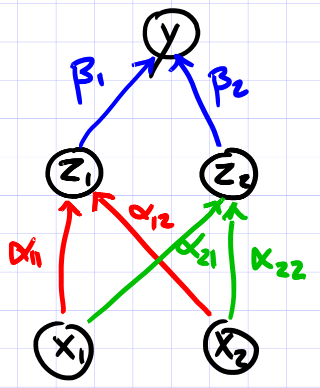
# Lecture 19: 3/29/17

## Neural Network: Models

**Ex#1:** NN w/ 1 Hidden Layer and 2 Hidden Units




(output)

(hidden layer)

(input)

Let $\sigma(\cdot)$ be the activation function

If $\sigma$ is sigmoid: $\sigma(a) = \frac{1}{1+\exp(-a)}$

$$z_1 = \sigma(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{10})$$

$$z_2 = \sigma(\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{20})$$

Each is just Logistic Regression

$$y = \sigma(\beta_1 z_1 + \beta_2 z_2 + \beta_0)$$

$$= \sigma(\beta_1 \sigma(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{10}) + \beta_2 \sigma(\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{20}) + \beta_0)$$

$Pr[Y=1 \mid \vec{x}, \alpha, \vec{\beta}]$

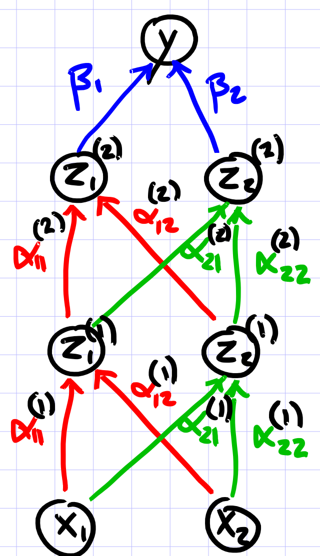
$x_i \in \mathbb{R}$

$z_i \in [0,1]$ if $\sigma$ is sigmoid

$z_i \in \mathbb{R}$ more generally

☆ Don't forget the bias term!

**Ex#2:** NN w/ 2 Hidden Layers and 2 Hidden Units each.




$$z_1^{(1)} = \sigma(\alpha_{11}^{(1)} x_1 + \alpha_{12}^{(1)} x_2 + \alpha_{10}^{(1)})$$

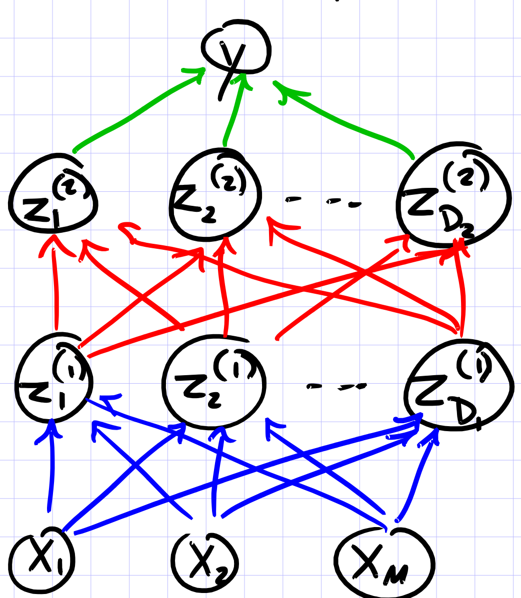
$$z_2^{(1)} = \ldots$$

$$z_1^{(2)} = \ldots$$

$$z_2^{(2)} = \sigma(\alpha_{21}^{(2)} z_1^{(1)} + \alpha_{22}^{(2)} z_2^{(1)} + \alpha_{20}^{(2)})$$

$$y = \sigma(\beta_1 z_1^{(2)} + \beta_2 z_2^{(2)} + \beta_0)$$

**Ex#3:** Arbitrary Feedforward NN (Matrix Form)




**Parameters**

$\vec{\beta} \in \mathbb{R}^{D_2 \times 1}$

$\alpha^{(2)} \in \mathbb{R}^{D_1 \times D_2}$

$\alpha^{(1)} \in \mathbb{R}^{M \times D_1}$

**Computation**

$$y = \sigma(\vec{\beta}^T \vec{z}^{(2)})$$

$$\vec{z}^{(2)} = \sigma((\alpha^{(2)})^T \vec{z}^{(1)})$$

applied element-wise

$$\vec{z}^{(1)} = \sigma((\alpha^{(1)})^T \vec{x})$$

**Bias Term?**

Assume $x_0 = 1$, $z_0^{(1)} = 1$, $z_0^{(2)} = 1$ and increment size of parameters.