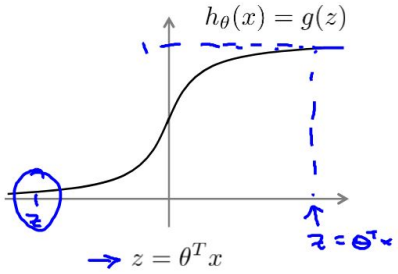
**Plan**

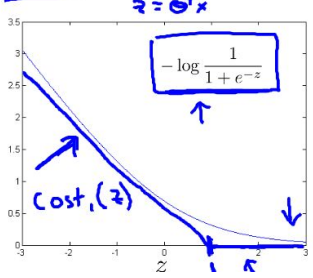
1. **Large margin classification - Take a look at support vector machine supervised algorithm**
   a. **Optimization objective**

$$\min_\theta \frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)}\left(-\log h_\theta(x^{(i)})\right) + (1-y^{(i)})\left((-\log(1-h_\theta(x^{(i)}))\right)\right] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2$$

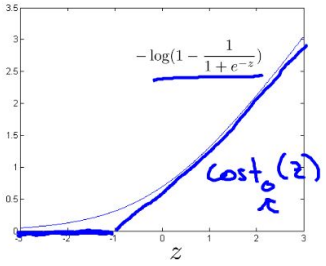
$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$




If $y = 1$ (want $\theta^T x \gg 0$):

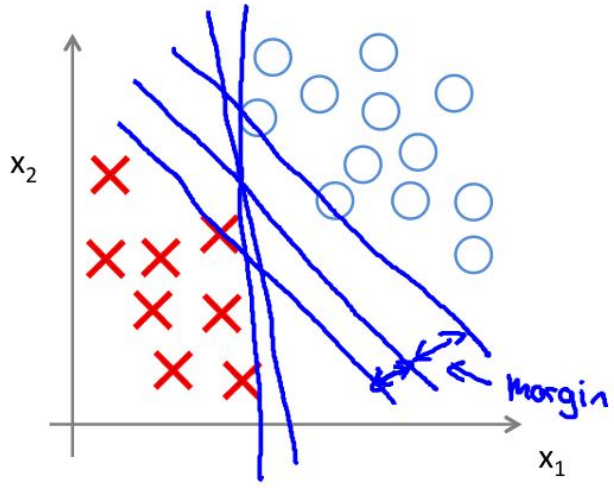



If $y = 0$ (want $\theta^T x \ll 0$):




$$\min_\theta C\sum_{i=1}^{m}\left[y^{(i)}cost_1(\theta^T x^{(i)}) + (1-y^{(i)})cost_0(\theta^T x^{(i)})\right] + \frac{1}{2}\sum_{i=1}^{n}\theta_j^2$$

   b. **Large margin intuition**



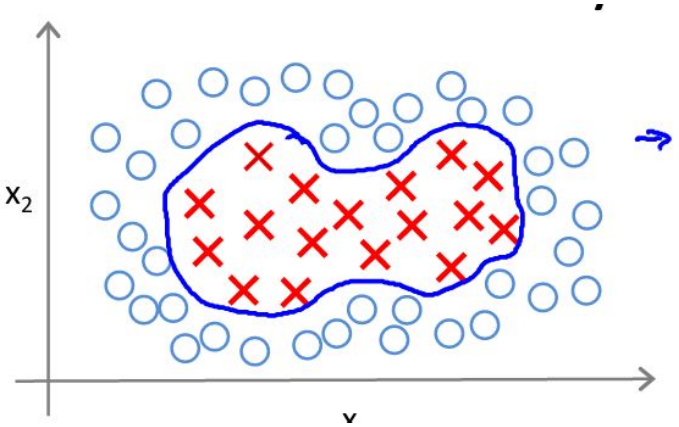

   c. **Math behind LMC**

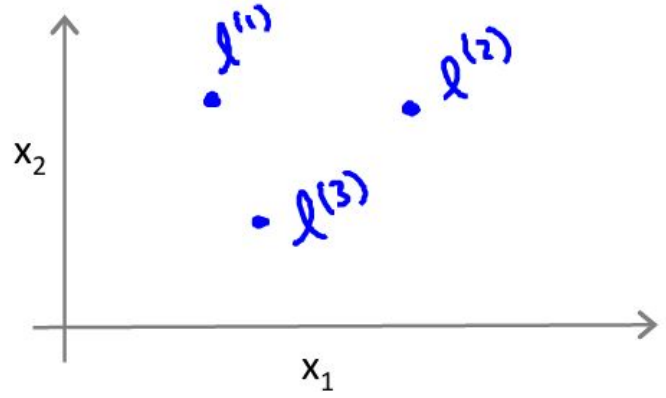
$$\min_\theta \frac{1}{2}\sum_{j=1}^{n}\theta_j^2 = \frac{1}{2}\|\theta\|^2$$

$$\text{s.t.} \quad p^{(i)} \cdot \|\theta\| \geq 1 \quad \text{if } y^{(i)}$$

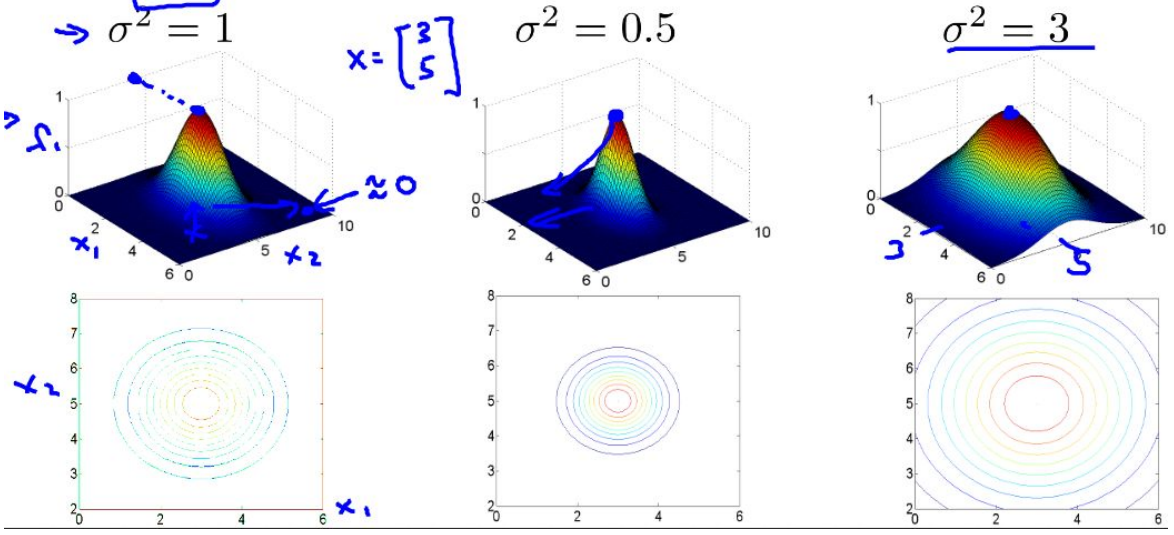
$$p^{(i)} \cdot \|\theta\| \leq -1 \quad \text{if } y^{(i)}$$

2. **Kernels - Defining algorithms with non-linear decision boundaries**
   a. **Gaussian kernel**

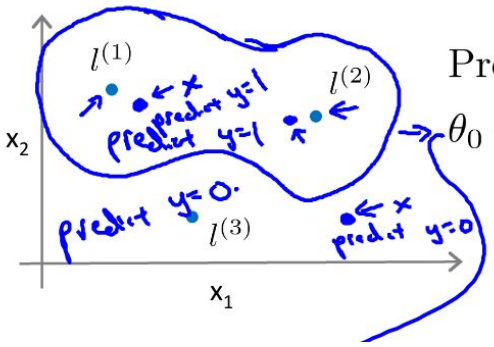



$l^{(1)} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ $\quad f_1 = \exp\left(-\frac{\|x - l^{(1)}\|^2}{2\sigma^2}\right)$




   b. **Creating landmarks**




Predict "1" when

$\theta_0 + \theta_1 f_1 + \theta_2 f_2 + \theta_3 f_3 \geq 0$

$\theta_0 = -0.5, \quad \theta_1 = 1, \quad \theta_2 = 1, \quad \theta_3 = 0$

3. **Using an SVM**

**Questions:**

1. **In which type of ML problems we can apply SVM?**
2. **An alternative name for SVM?**
3. **What kind of kernel do you know? (This question isn't about Linux kernels btw)**

**Glossary**

**Kernel -** point we use as a feature based on our learning datasets - *By default it's highly recommended to use Gaussian kernel*

**Support Vector Machine -** yet another learning algorithm used in classification problems - *Support Vector Machine aka Large Margin Classifier*

**Large Margin Classifier -** learning algorithm with an alternative approach on classification problem - *Support Vector Machine aka Large Margin Classifier*