Logistic Regression

▷ $h_\theta(x) = \frac{1}{1+e^{-\theta x}}$ ⇒ Hypothesis which generates $\hat{y}$ (predictions)

▷ cost function:

$$J(\theta) = -\frac{1}{m}\sum^{m} y^{(i)}\log\left(h_\theta(x^{(i)})\right) + (1-y^{(i)})\left(\log(1-h_\theta(x^{(i)}))\right)$$

▷ Gradient update rule:

$$\theta_j = \theta - \frac{\alpha}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)x_j^{(i)}$$

▷ VECTORIZATION OF COST:

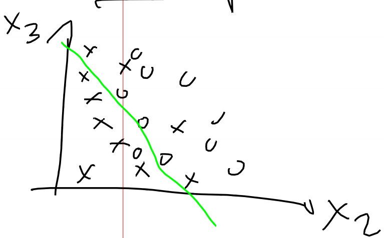
$$\frac{1}{m}\left[\underbrace{\vec{y}}_{m\times 1} \cdot \underbrace{\log(h_\theta(x))}_{m\times 1} - \underbrace{(1-y)}_{m\times 1}\left(\underbrace{\log(h_\theta(x))}_{m\times 1}\right)\right]$$

$\Rightarrow \vec{y}^T \in \mathbb{R}^{1\times m}$ $\quad (1-\vec{y})^T \in \mathbb{R}^{1\times m}$

▷ VECTORIZATION OF GRADIENT

$$\theta = \frac{1}{m}\underbrace{X'}_{n+1\times m} \cdot \underbrace{\text{sigmoid}(X\theta)}_{\in \mathbb{R}^{m\times 1}} \Rightarrow \theta \in \mathbb{R}^{n+1\times m}$$ ✓

▷ Plotting decision boundary



decision boundary defined as sigmoid(0)

⇒ Argument of sigmoid func must be 0!

⇔ $h_\theta(x) = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \overset{!}{=} 0 \Leftrightarrow x_3 = -\frac{\theta_1 x_1 + \theta_2 x_2}{\theta_3}$

Idea: get (min, max) of either x_2 (x-coord) or x_3 (y-coord) and solve above equation for the other var.

e.g.: plot_x = [min(x_2), max(x_2)] (x-coords of plot)

⇒ need y-coords now:

plot_y = $\left(-\frac{1}{\theta_3} \times [\theta_1 \overset{1}{x_1} + x_2\theta_2]\right)$

⇒ plot(plot_x, plot_y)

Code for creating higher order Polynomials

Transform X

$$\begin{bmatrix} x_1^1 & x_1^2 \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix} \to \begin{bmatrix} x_1^1 & x_1^2 & x_1^3 & x_1^4 & \dots \\ \vdots & & & & \vdots \end{bmatrix}$$

where $x^{(i)}$ for $i > 3$ represents another polynomial (eg $(x^1)^2, (x^2)^2, (x^1 x^2), \dots$)

```
degree = 6;
out = ones(size(X(:,1)))   -> [row-Dim]
for i = 1:degree
    for j = 0:i
        out(:, end+1) = (X1^(i-j) .* X^j)
    end
end
```

Exp: $\boxed{i = 3}$

$\Rightarrow$ for j 0:3

1. Step out $= (X1^{(3-0)} \cdot \cancel{X2^0}) \Rightarrow X1^3$
2. Step out $= (X1^{3-1} \cdot X2^1) \Rightarrow X1^2 X2$
3. Step out $= (X1^{3-2} \cdot X2^2) \Rightarrow X1 X2^2$
4th Step $X1^{3-3} \cdot X2^3 \Rightarrow X2^3$

$\boxed{i = 2}$

1. $X1^2$
2. $X1\,X2$
3. $X2^2$

$\boxed{i = 1}$

$X1$

$X2$

$$\left[\; X_1 \;\middle|\; X_2 \;\middle|\; X_1^2 \;\middle|\; X_1X_2 \;\middle|\; X_2^2 \;\middle|\; X_1^3 \;\middle|\; X_1^2X_2 \;\middle|\; X_1X_2^2 \;\middle|\; X_2^3 \;\middle|\; \dots \;\right]$$

REGULARIZED LOG REG

▷ cost func

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

Intercept — Remember to exclude θ_1 in the regularization term!

▷ Gradient

$$\theta = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)} + \frac{\lambda}{m} \theta_j \,, \quad j = 2$$