**Classification and representation**

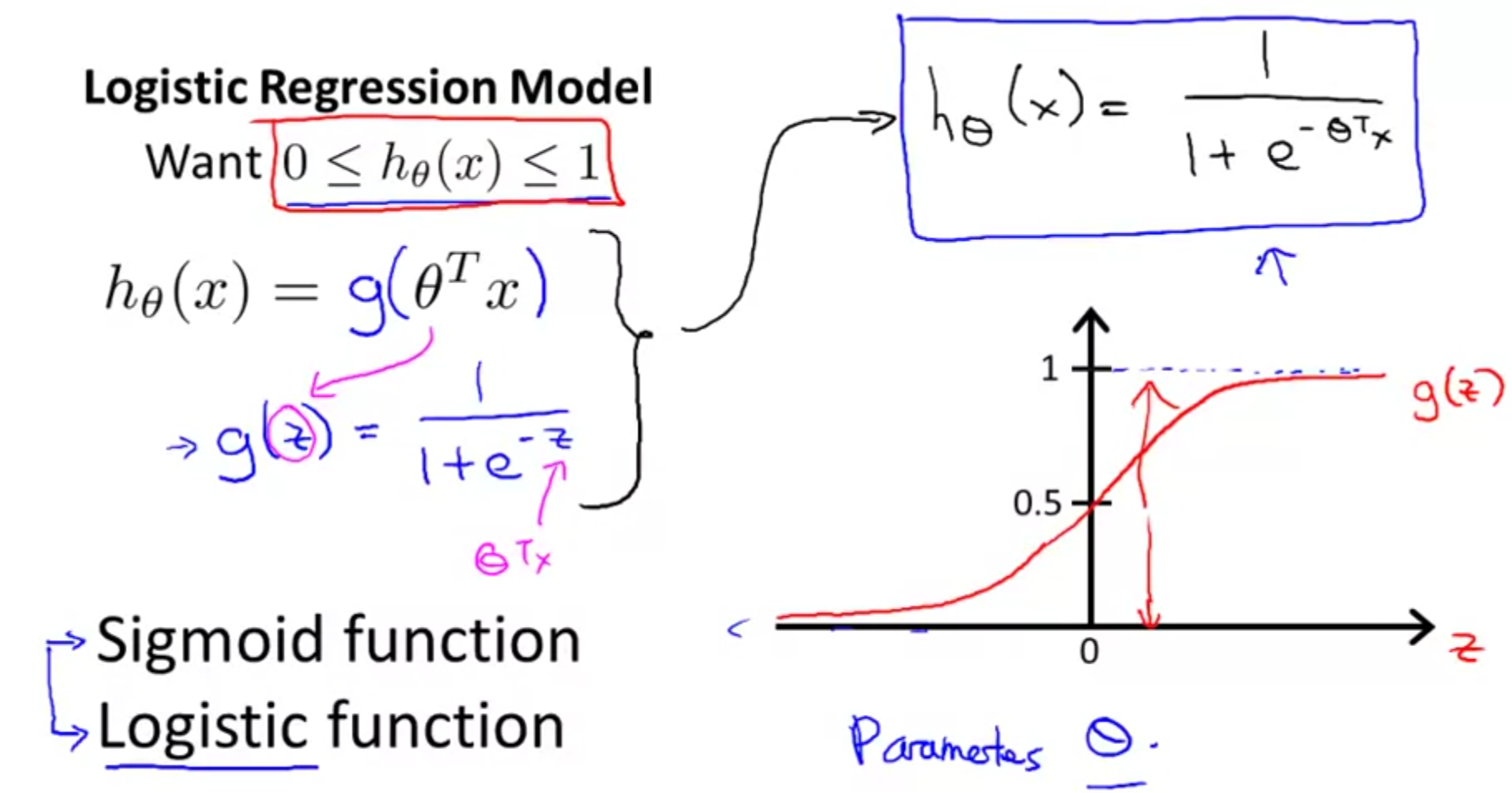
We classify the data into classes:

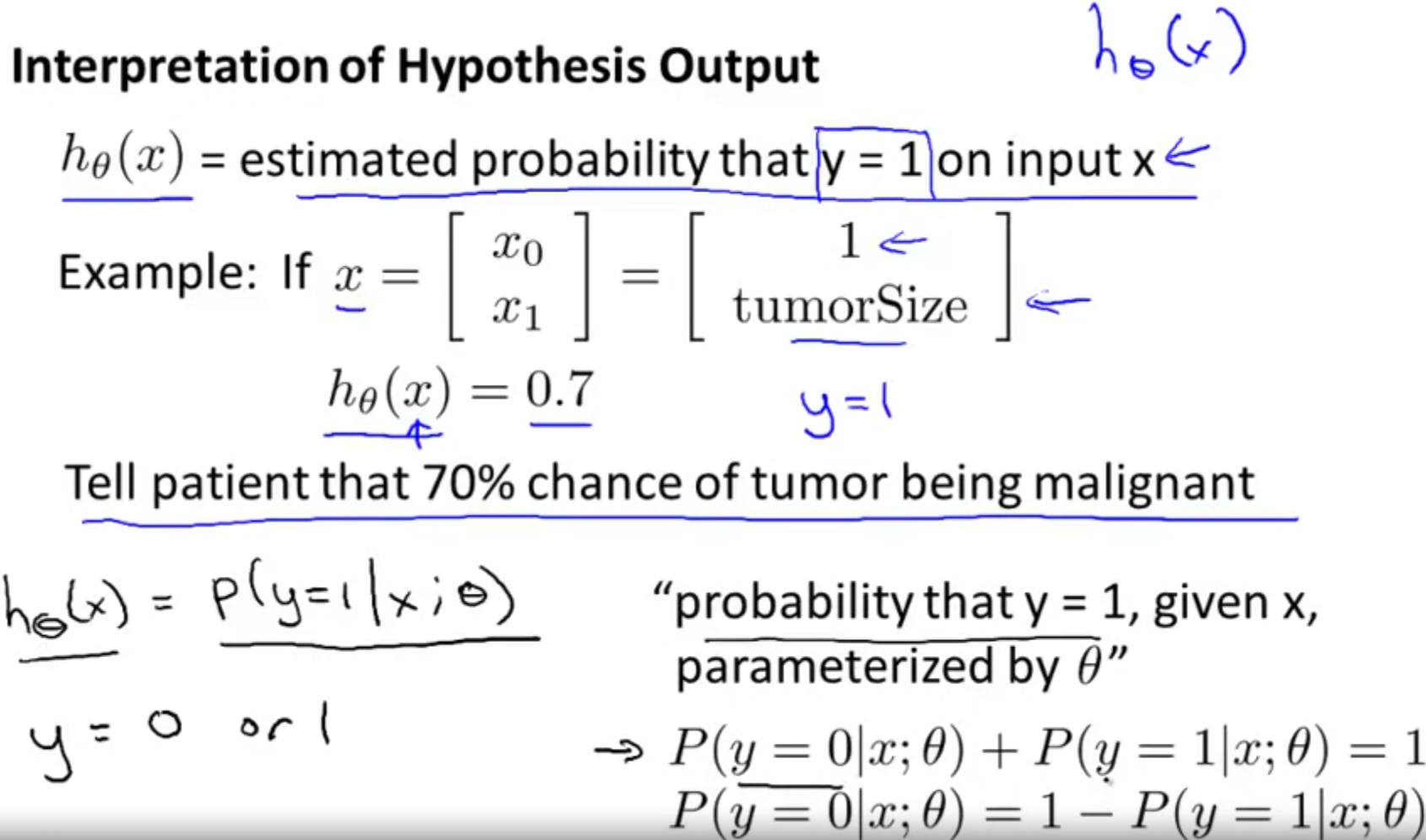
Where each number corresponds to a class. It’s for discrete data.

Threshold classifier: we set a limit so that if the data is over or below this value, it belongs to one class or another.

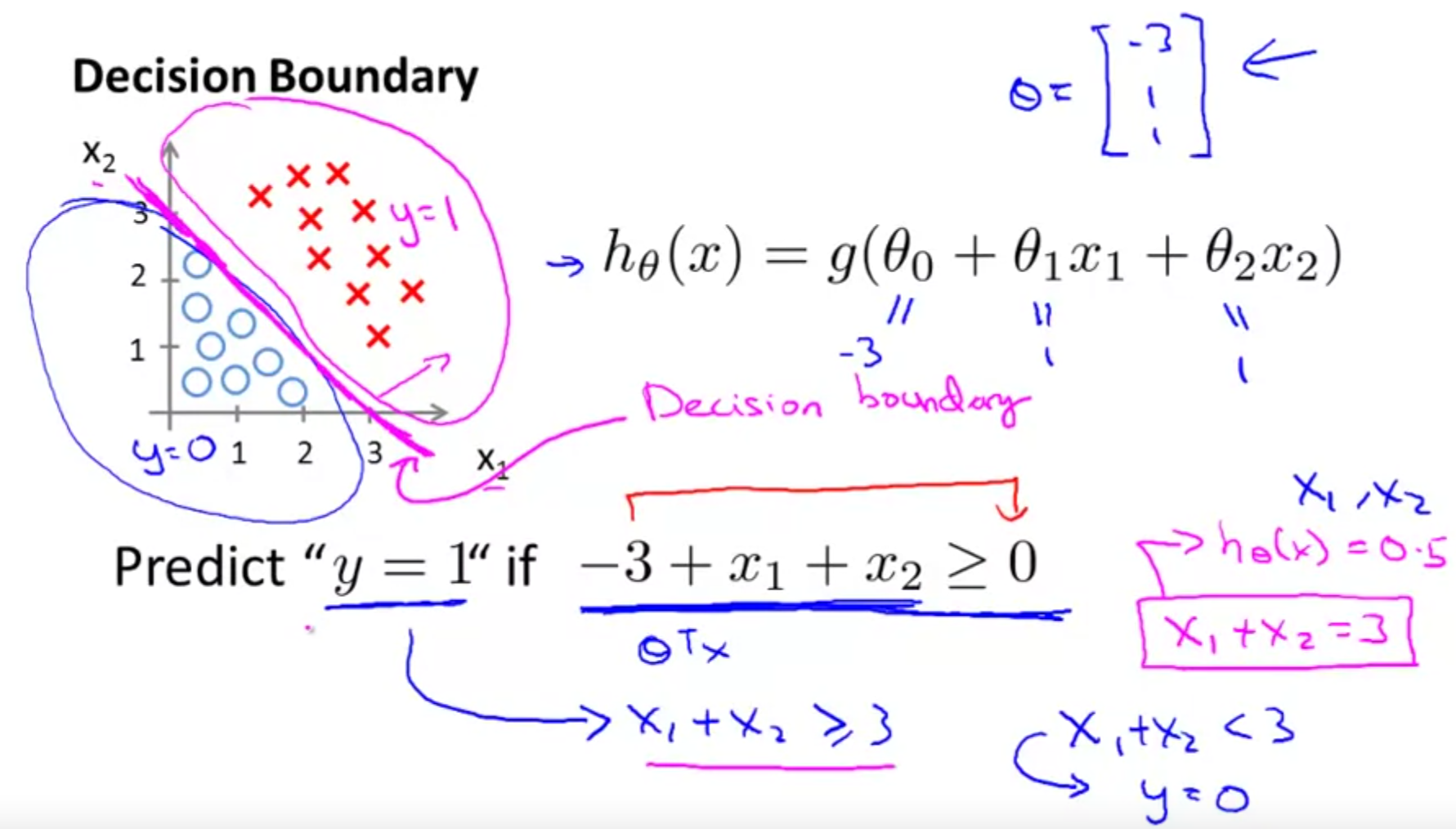
**Logistic regression**

We apply a transformation to the basic hØ(x), so we make sure our hypothesis function will always exist between 0 and 1.

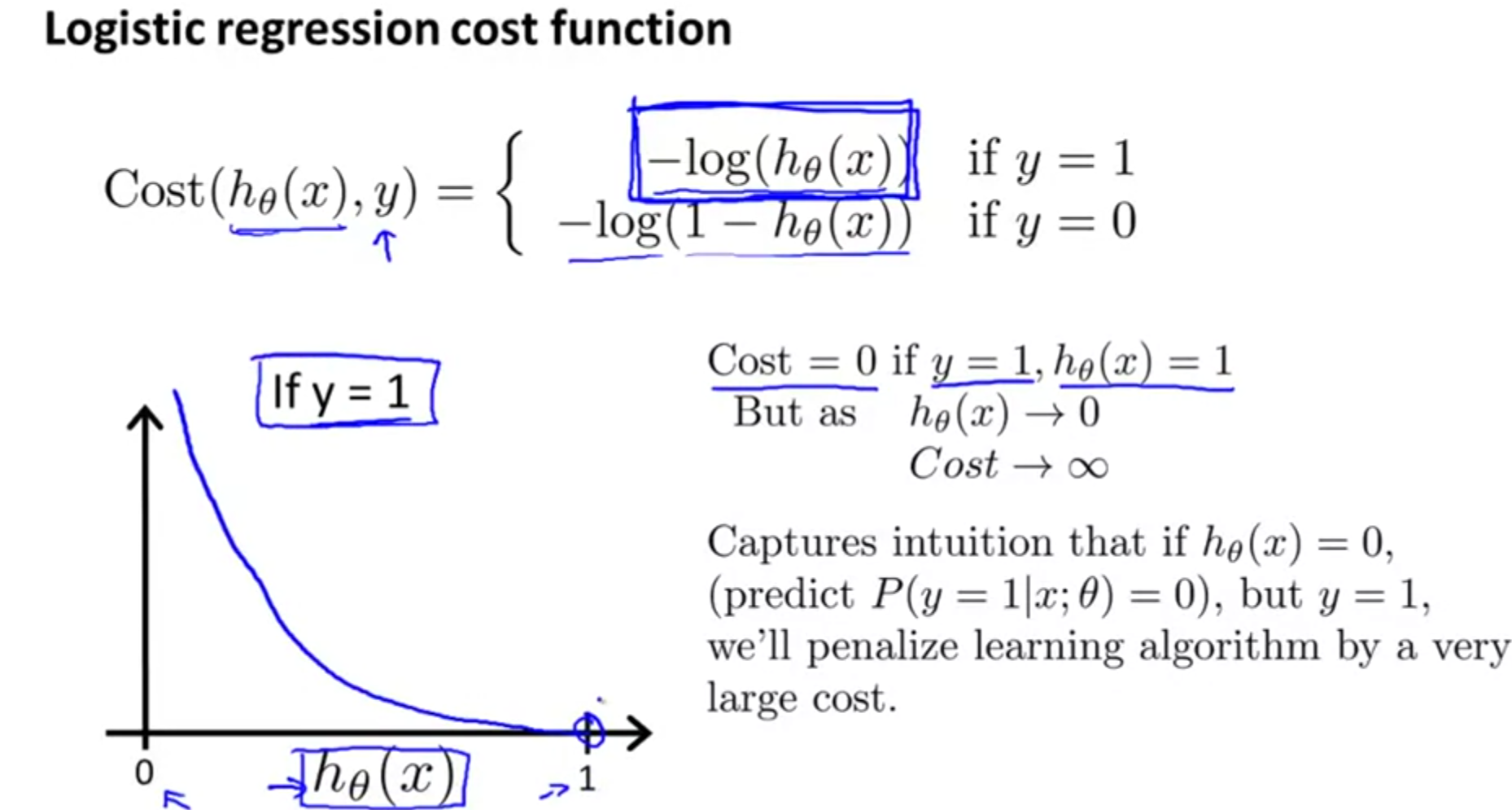


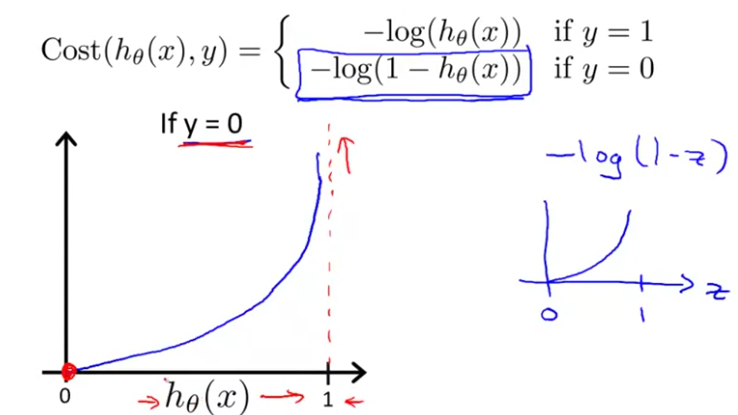


**Decision Boundary**



**Logistic Regression Model**

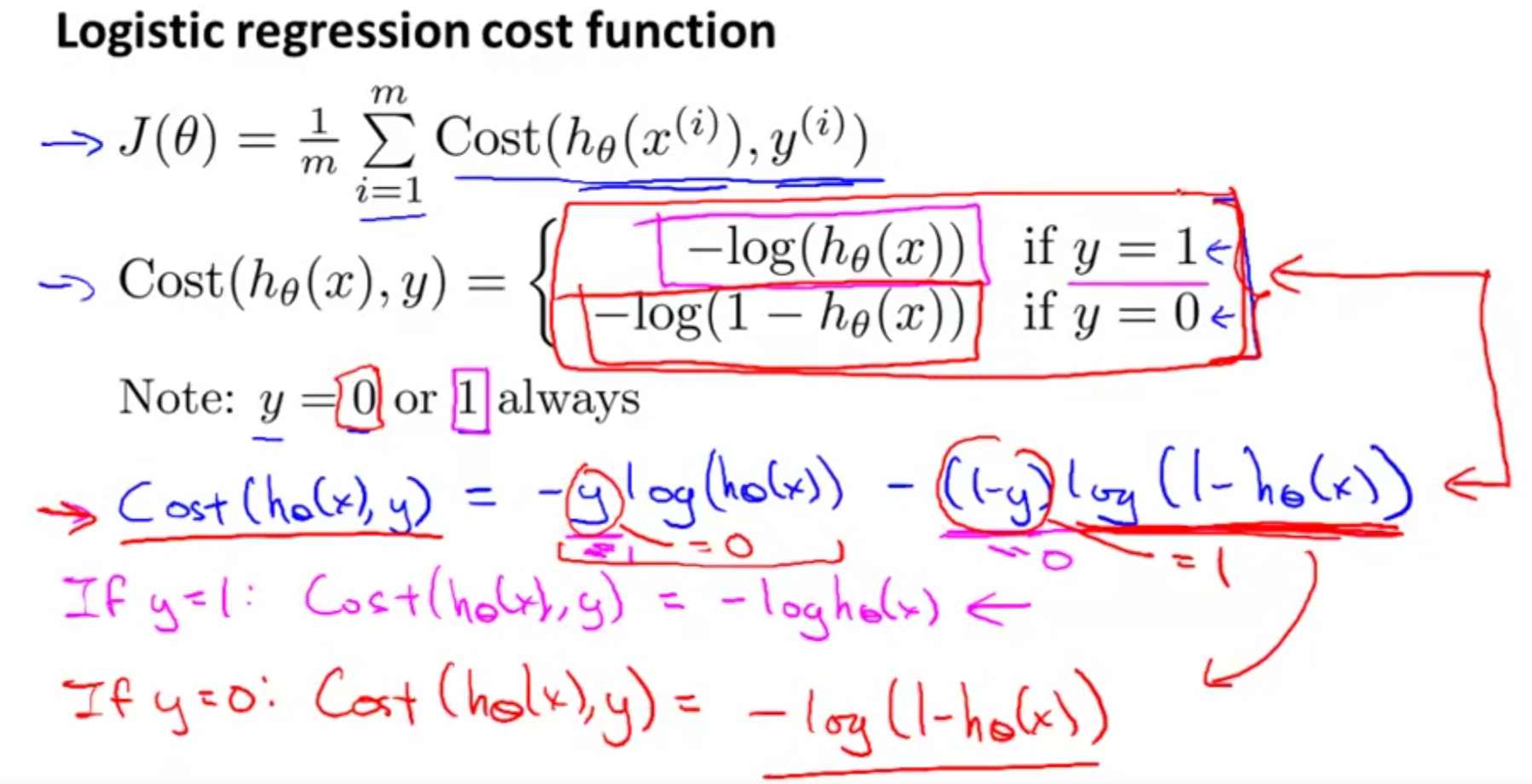


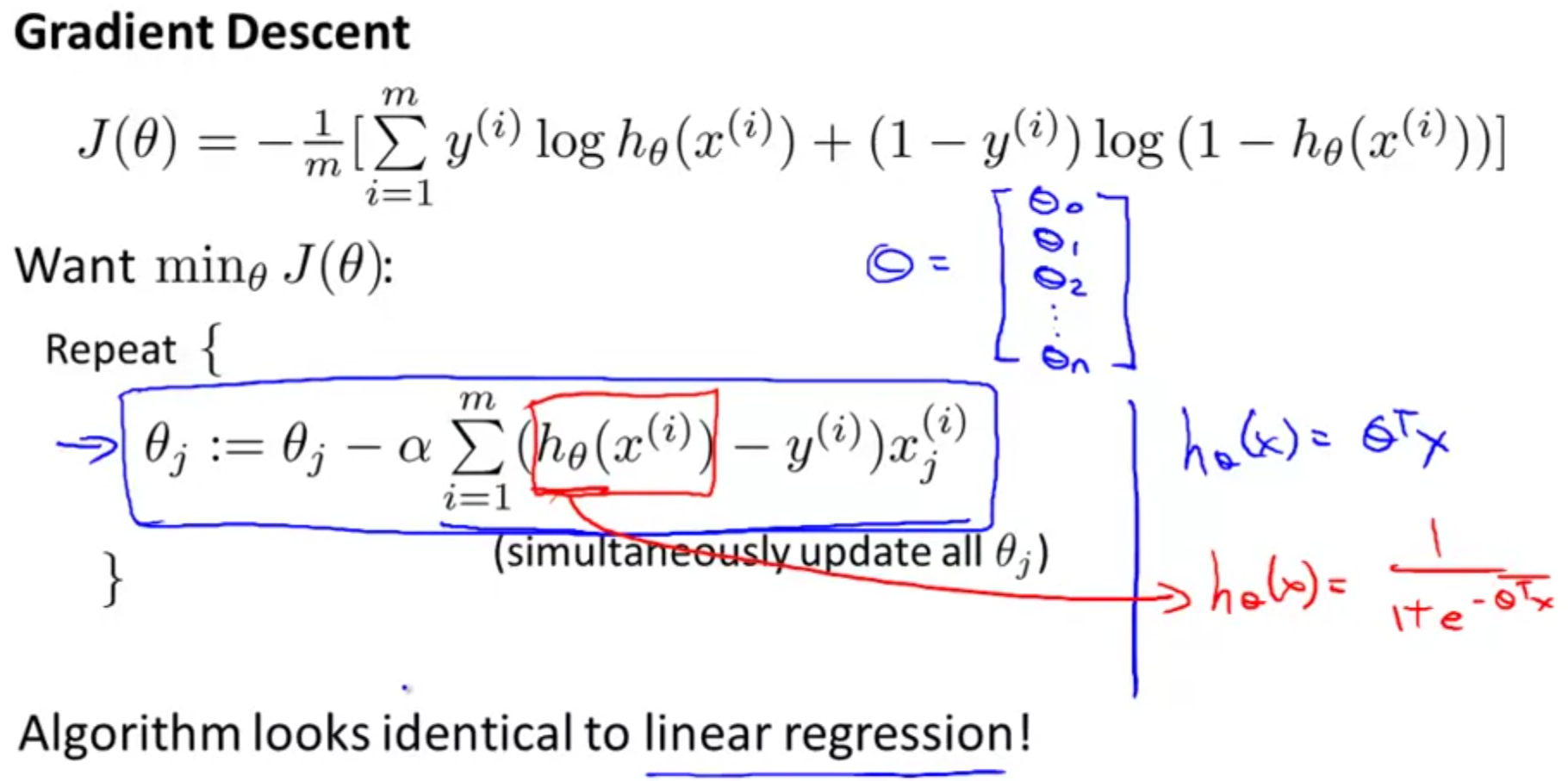


If hθ(x) = 0 and y=0, the cost function is 0.

As hθ(x) approaches 1 and y = 0, the cost function tends to infinite.

**Simplified cost function and gradient descent**

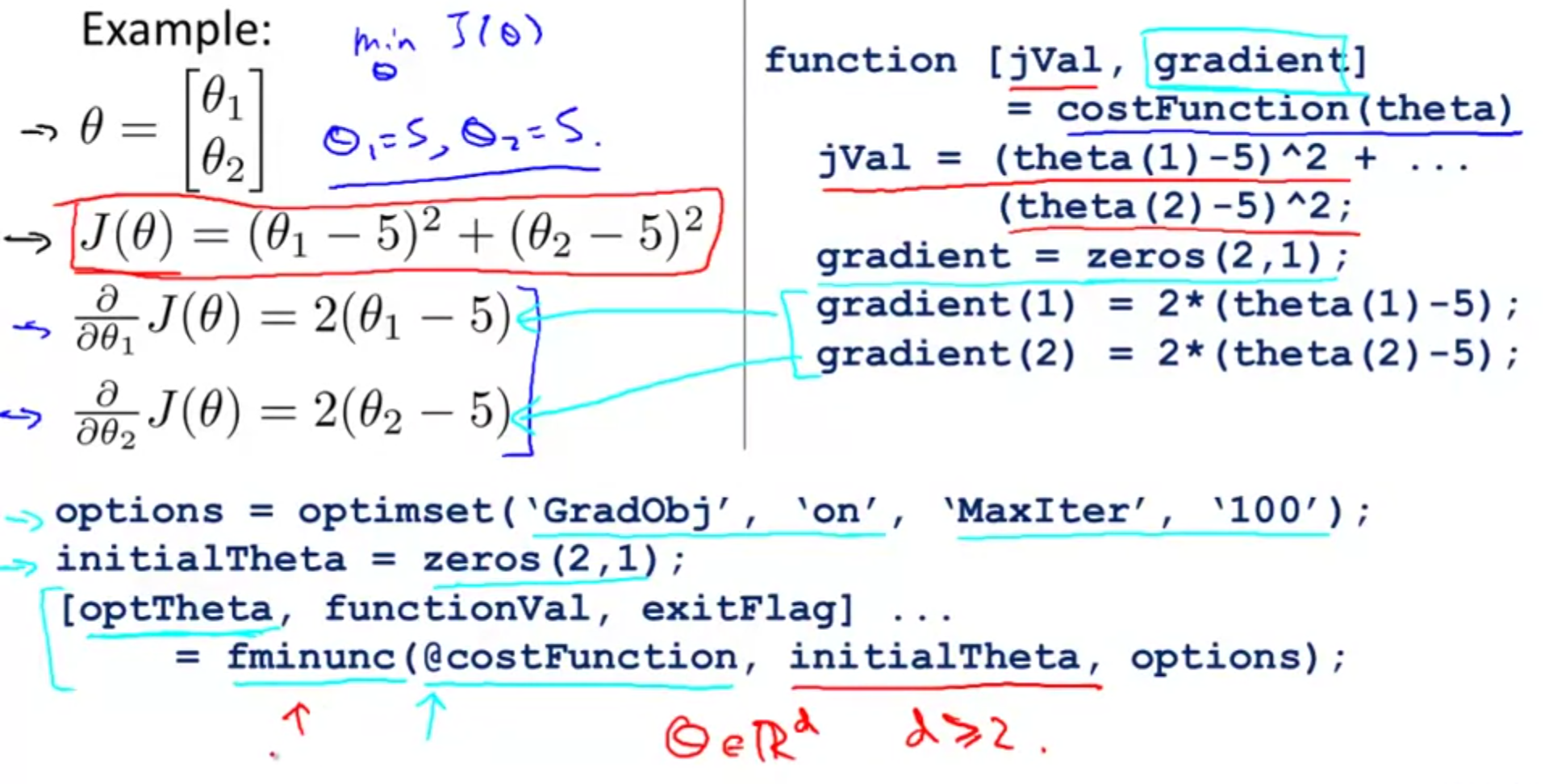




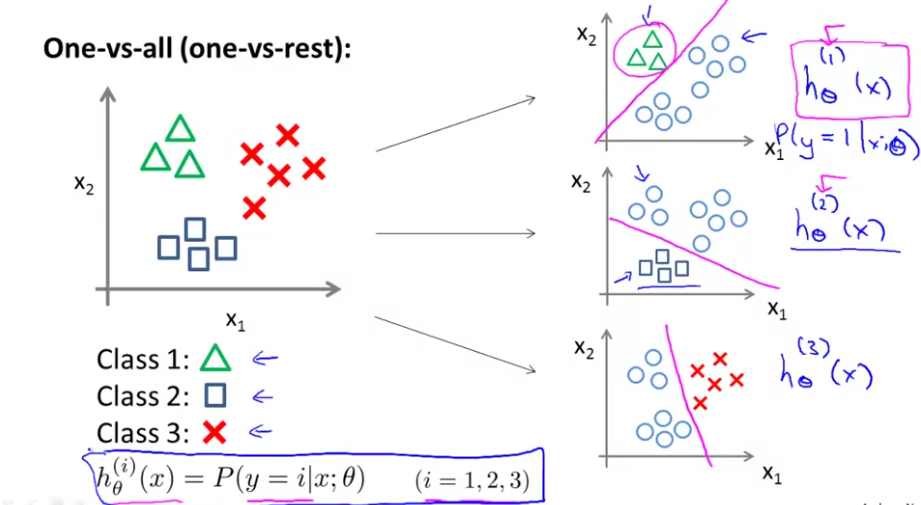
**Advanced optimization**



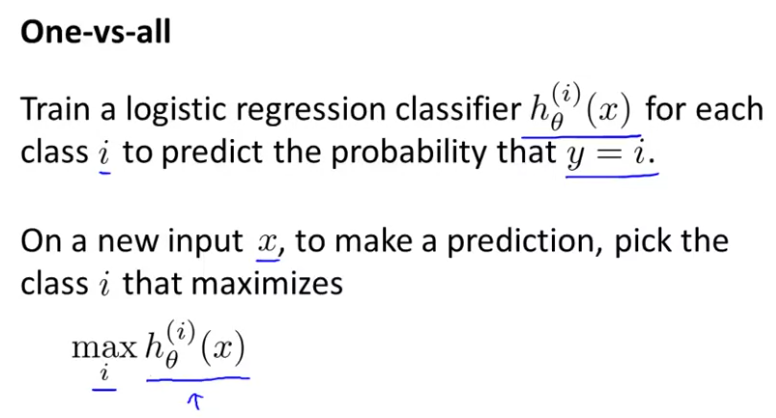
There are quite complex to implement by your own, but there are libraries for implementing them. An example of implementation on a library in octave:



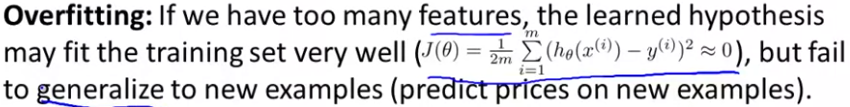
**Multiclass classification: one vs all**



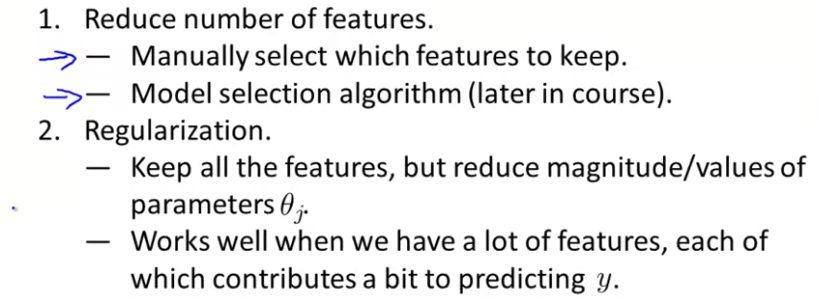
We create different classifiers where we discriminate one class vs the rest.



**Overfitting**



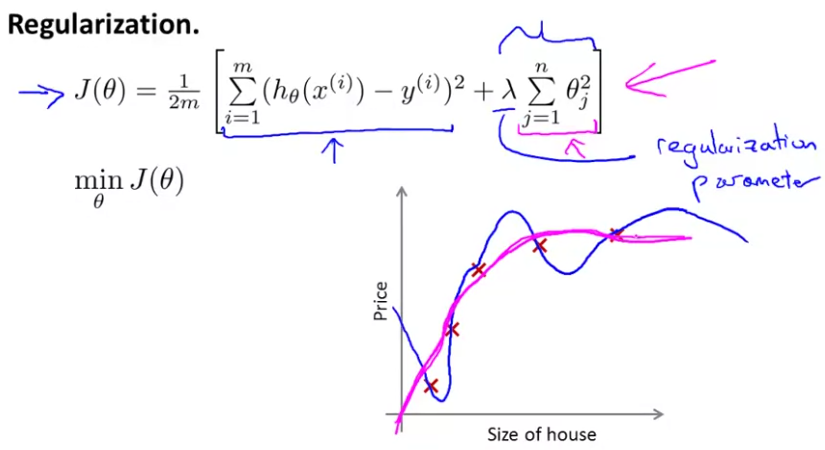
It appears when we have too many features in comparison to the training data. For solving the overfitting problems, we can:



**Regularization – Cost function**

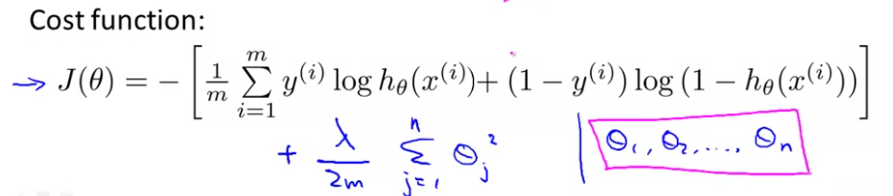
We try to make all the parameters smaller, which simplifies the hypothesis and solves the overfitting problem.

Θ0 usually is not regularized

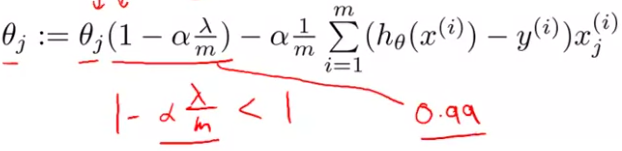


Note that there exists a regularization parameter λ, which controls the trade of between the goal of fitting the training set well and the goal of keeping the hypothesis relatively simple to avoid overfitting.

In the case of the logistic regression algorithm:

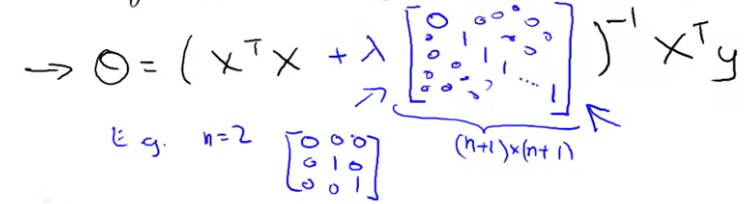


**Regularised linear regression**



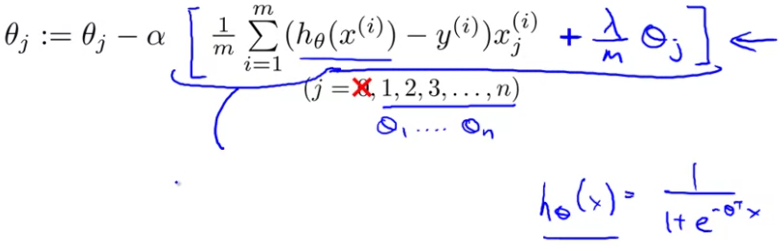
Regularised normal equation

Regularised gradient descent



When using the normal equation, the regularisation also makes sure that the matrix XTX will always be invertible.

**Regularised logistic regression**



Regularised gradient descent