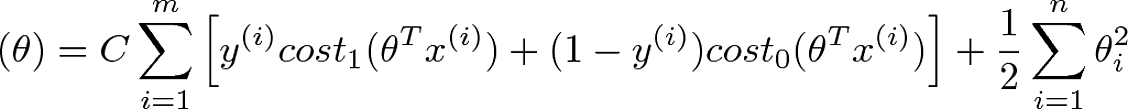
**SVM**

A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model sets of labelled training data for each category, they’re able to categorize new text.

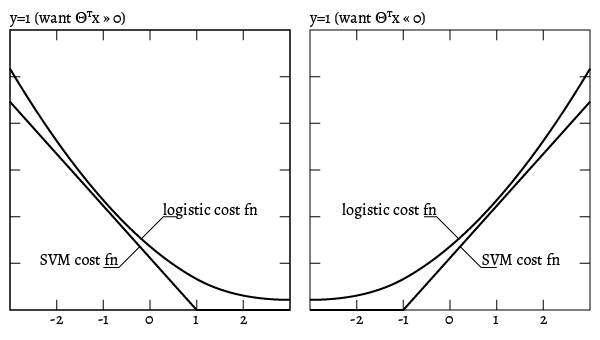
**The Cost function:**



Here C=1/lambda

The Cost Function is used to train the SVM. By minimizing the value of J(theta), we can ensure that the SVM is as accurate as possible. In the equation, the functions cost1 and cost0 refer to the cost for an example where y=1 and the cost for an example where y=0. For SVMs, cost is determined by kernel (similarity) functions.

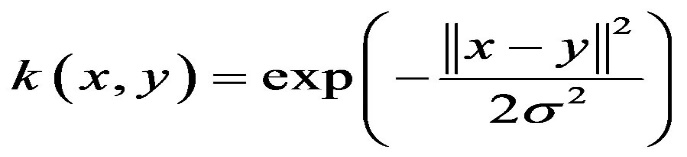
**Graphs:**



### ****Kernels****

Polynomial features tend to be computationally expensive, and may increase runtime with large datasets. Instead of adding more polynomial features, it's better to add landmarks to test the proximity of other datapoints against.  Each member of the training set can be considered a landmark, and a kernel is the similarity function that measures how close an input is to said landmarks.

In this model we have used Gaussian kernel, we may use linear and poly as well but the accuracy reduces a bit, so Gaussian kernel is most suitable based on our observation.

 (for gaussian kernel)

**Logistic Regression**

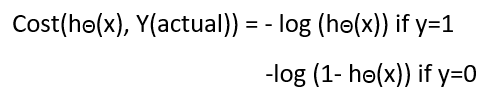
Logistic Regression is used when the dependent variable(target) is categorical. For example, to predict whether an email is spam (1) or (0), Whether the tumor is malignant (1) or not (0).

For logistic regression we use sigmoid function, so the cost function is

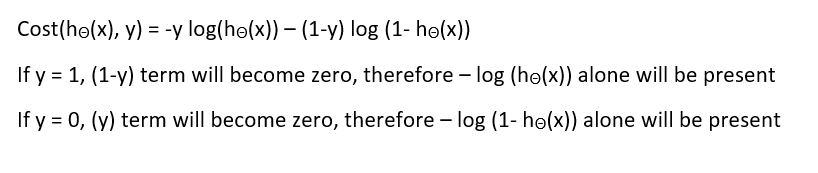
**Sigmoid Function**



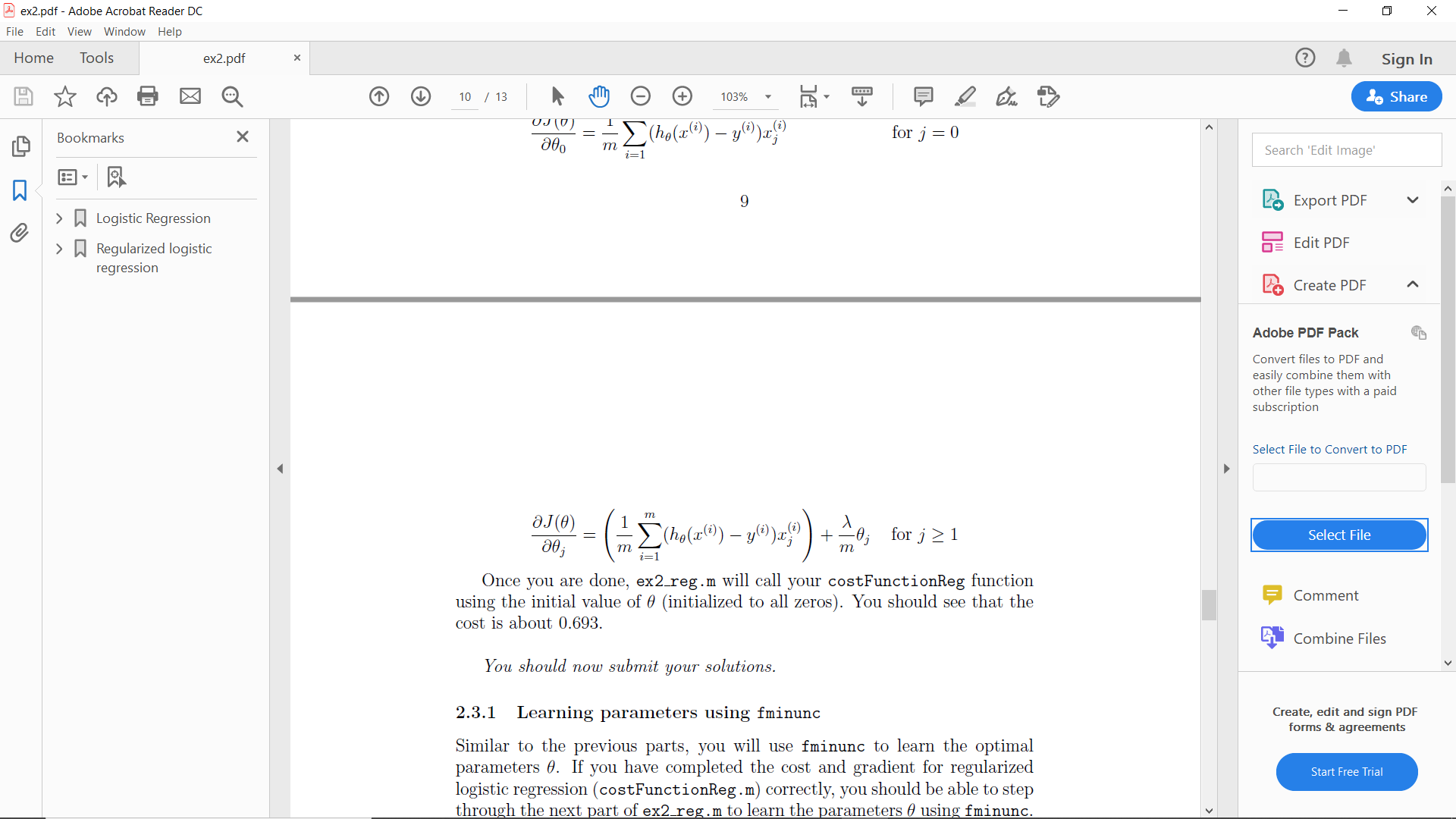
**Cost Function**



**Simplified cost function**



**Gradient descent**



Note that while this gradient looks identical to the linear regression gradient, the formula is actually different because linear and logistic regression have different definitions of h(ө(x)).

We may use regularization term (lambda) as well with the cost function.