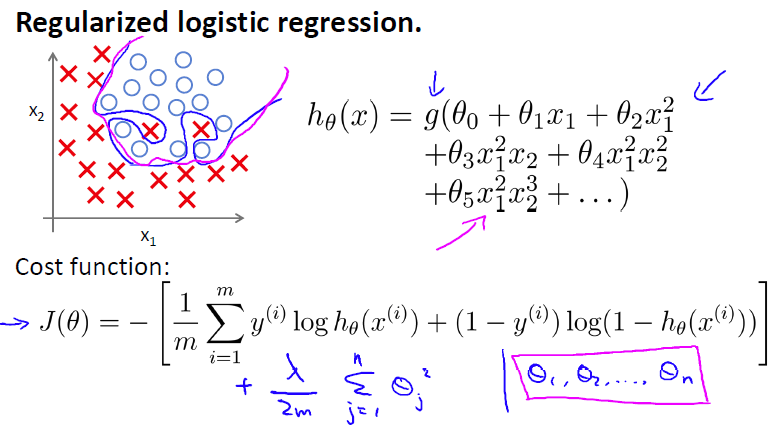


通常不惩罚theta zero，约定俗成，没有什么特别大的差别。

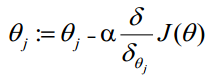


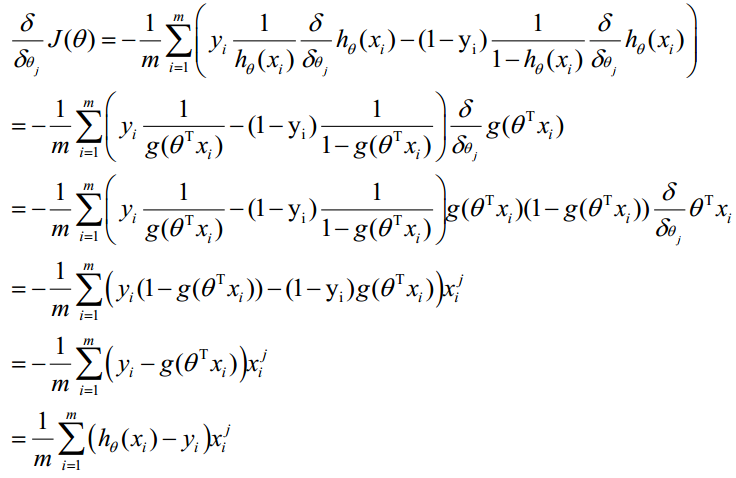
正则化参数可以帮助解决过拟合，使得原来的蓝色的拟合曲线变成玫红色相对光滑的曲线。

逻辑回归导数推导：

🡨sigmoid函数

http://img.my.csdn.net/uploads/201407/16/1405496926_1023.png http://img.my.csdn.net/uploads/201407/16/1405496967_1305.png



梯度推导过程：<https://blog.csdn.net/weixin_30014549/article/details/52850870>

**1. Which  
of the following statements about regularization are true? Check all that apply.【D】**

A.Because regularization causes J(θ) to no longer be convex, gradient descent may not always converge to the global minimum (when λ>0, and when using an appropriate learning rate α).

【解析】Regularized logistic regression and regularized linear regression are both convex, and thus gradient descent will still converge to the global minimum.

B.Using too large a value of λ can cause your hypothesis to overfit the data; this can be avoided by reducing λ.

【解析】Using a very large value of λ can lead to underfitting of the training set.

C.Because logistic regression outputs values 0≤hθ(x)≤1, it's range of output values can only be "shrunk" slightly by regularization anyway, so regularization is generally not helpful for it.

【解析】Regularization affects the parameters θ and is also helpful for logistic regression.

D.Consider a classification problem. Adding regularization may cause your classifier to incorrectly classify some training examples (which it had correctly classified when not using regularization, i.e. when λ=0).

【解析】Regularization penalizes complex models (with large values of θ).They can lead to a simpler models, which misclassifies more training examples.

**2. You are training a classification model with logistic regression. Which of the following statements are true? Check all that apply.【D】**

A. Introducing regularization to the model always results in equal or better performance on the training set.

    Introducing regularization to the model always results in equal or better performance on the training set.

 【解析】If we introduce too much regularization, we can underfit the training set and have worse performance on the training set.

           B.Adding many new features to the model helps prevent overfitting on the training set.

【解析】Adding many new features gives us more expressive models which are able to better fit our training set. If too many new features are added, this can lead to overfitting of the training set.

C. Adding a new feature to the model always results in equal or better performance on examples not in  
the training set.

【解析】Adding  more features might result in a model that overfits the training set, and thus can lead to worse performs for examples which are not in the training set.

D.Adding a new feature to the model always results in equal or better performance on the training set.