

Evaluating a Learning Algorithm

Bias vs. Variance

- Diagnosing Bias vs. Variance 7 min
- Diagnosing Bias vs. Variance 3 min
- Regularization and Bias/Variance 11 min
- Regularization and Bias/Variance 3 min
- Learning Curves 11 min
- Learning Curves 3 min
- Deciding What to Do Next Revisited 6 min
- Deciding What to do Next Revisited 3 min

Review

Building a Spam Classifier

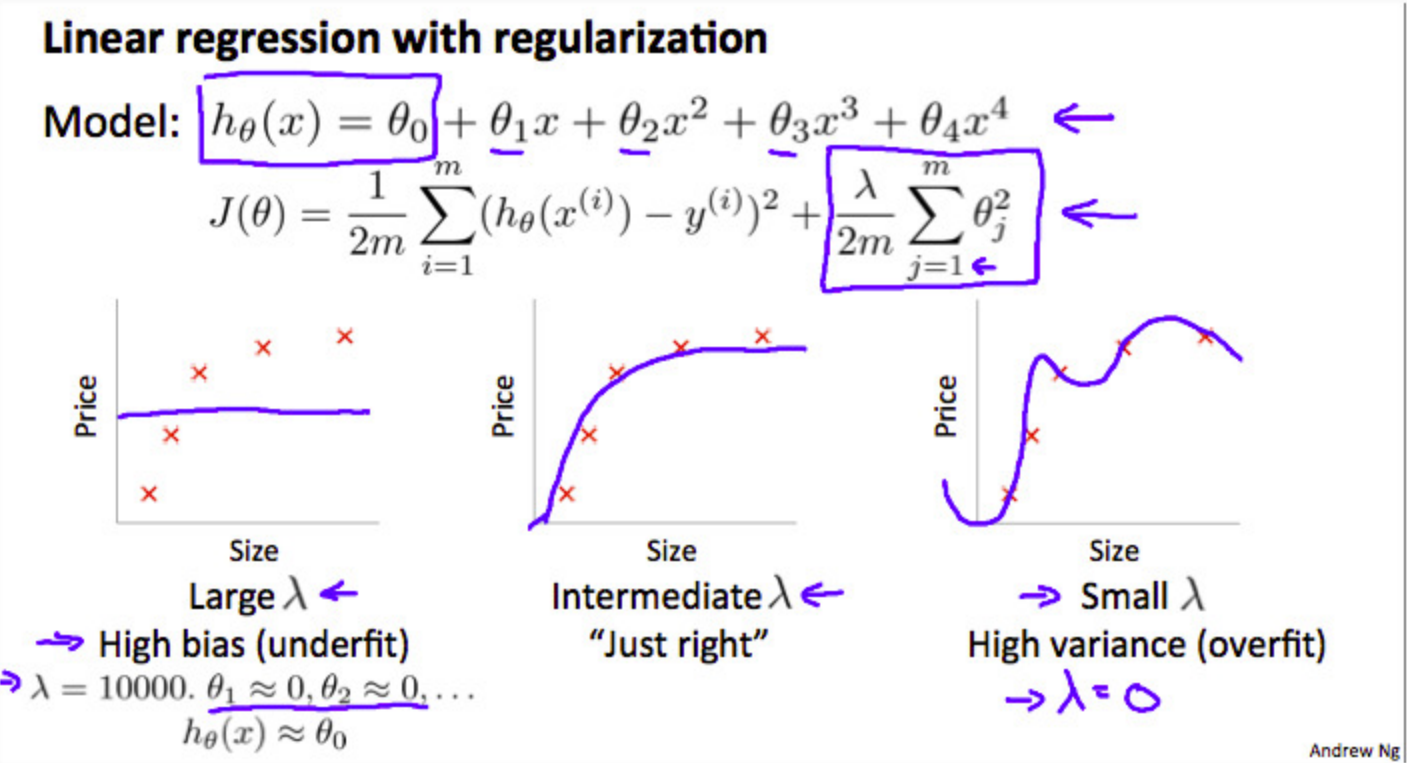
Handling Skewed Data

Using Large Data Sets

Review

Regularization and Bias/Variance

Note: [The regularization term below and through out the video should be $\frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$ and **NOT** $\frac{\lambda}{2m} \sum_{j=1}^m \theta_j^2$]



In the figure above, we see that as λ increases, our fit becomes more rigid. On the other hand, as λ approaches 0, we tend to over overfit the data. So how do we choose our parameter λ to get it 'just right' ? In order to choose the model and the regularization term λ , we need to:

- Create a list of lambdas (i.e. $\lambda \in \{0, 0.01, 0.02, 0.04, 0.08, 0.16, 0.32, 0.64, 1.28, 2.56, 5.12, 10.24\}$);
- Create a set of models with different degrees or any other variants.
- Iterate through the λ s and for each λ go through all the models to learn some Θ .
- Compute the cross validation error using the learned Θ (computed with λ) on the $J_{CV}(\Theta)$ **without** regularization or $\lambda = 0$.
- Select the best combo that produces the lowest error on the cross validation set.
- Using the best combo Θ and λ , apply it on $J_{test}(\Theta)$ to see if it has a good generalization of the problem.

Marcar como completo

