What to do With Bias and Variance?

Diagnosis

- If your model cannot even fit the training examples, then you have a large bias.

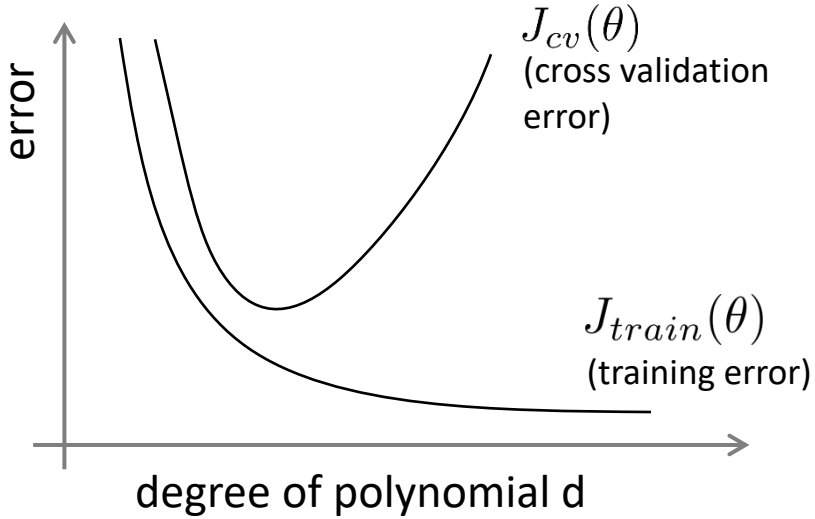
Underfitting

- If your model can fit the training data, but there is a large error on the testing data, then you probably have a large variance.

Overfitting

Diagnosis, Continued

Suppose your learning algorithm is performing less well than you were hoping. ($J_{cv}(\theta)$ *or* $J_{test}(\theta)$ is high.) Is it a bias problem or a variance problem?

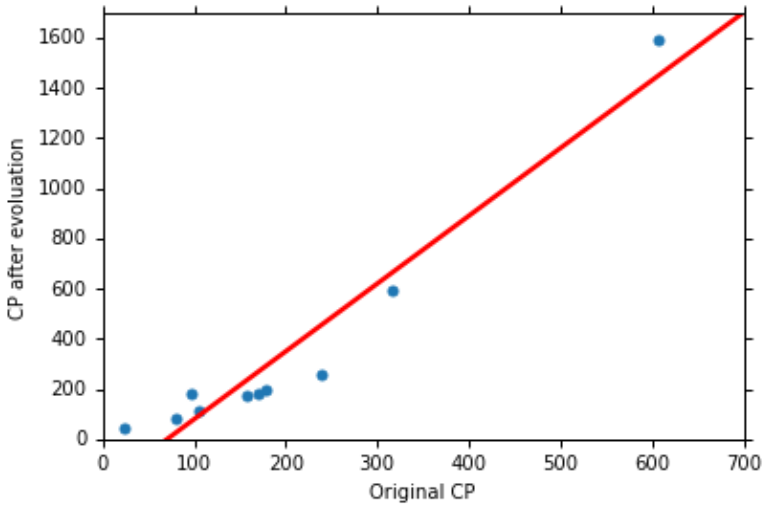


Bias (underfit): Both errors are high.

Variance (overfit): Training error is low while cross validation error is high.

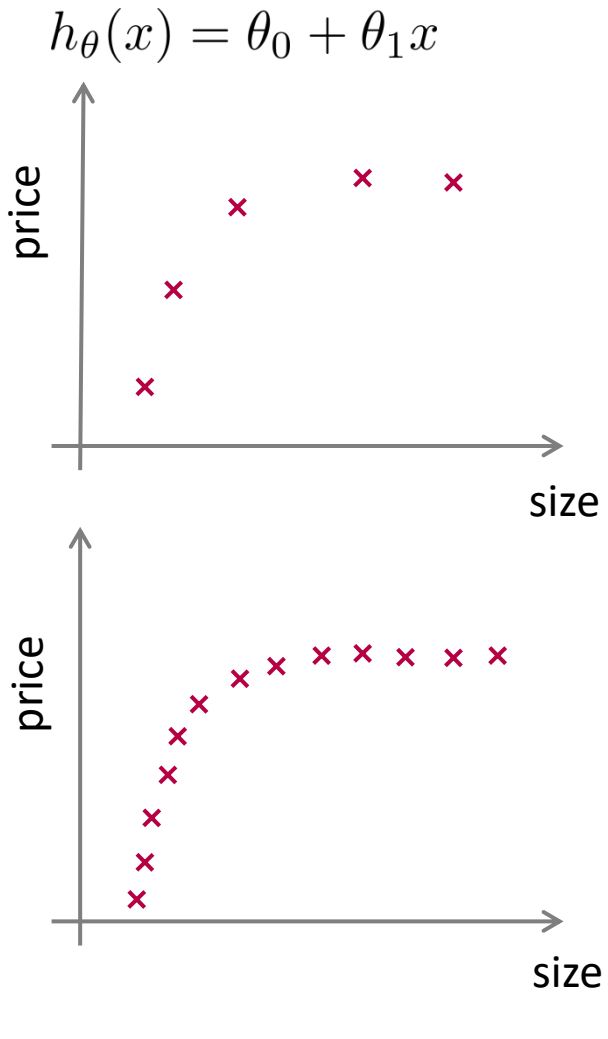
What to do With a Large Bias? Part I

- For bias, redesign your model:
  - Add more features (additional features and polynomial features) as input
  - Create a more complex model

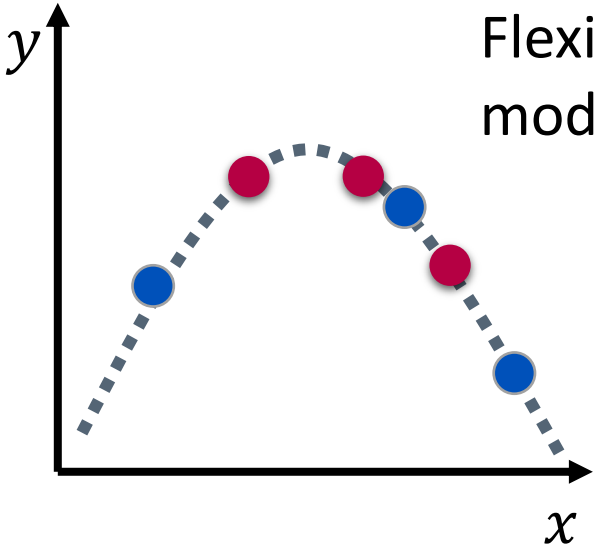


What to do With a Large Bias? Part II

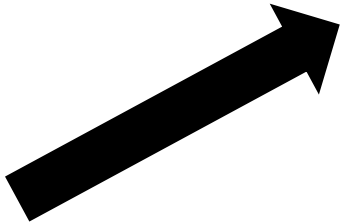
If a learning algorithm is suffering from high bias, getting more training data will not help much.

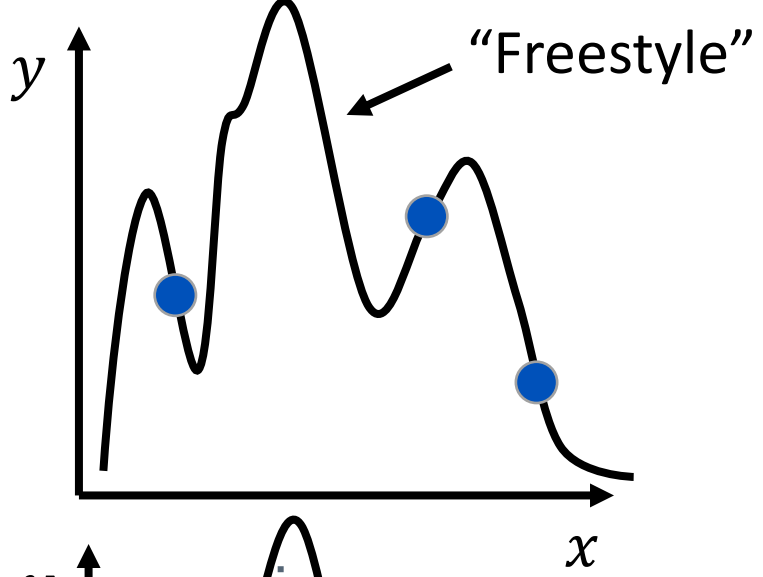


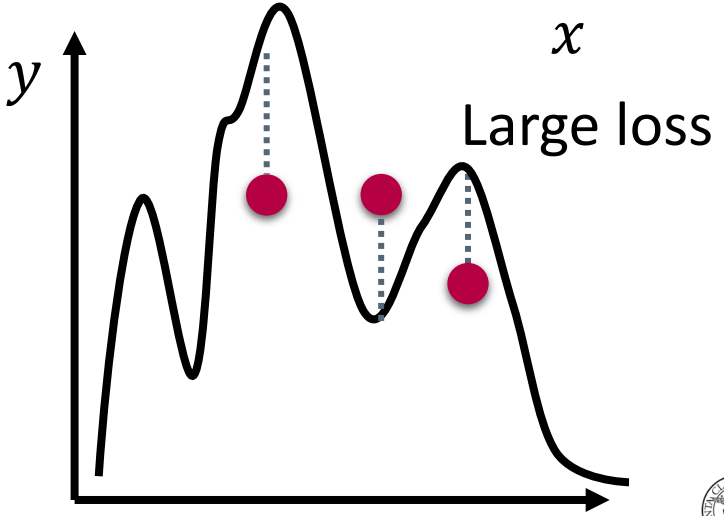
What to do With a Large Bias? Part III



Flexible model





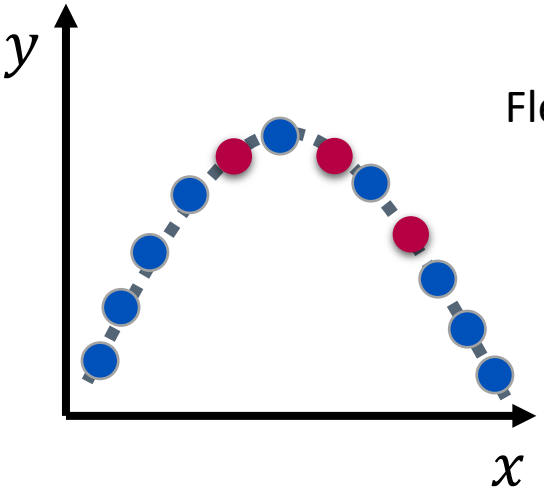


Real data distribution (not observable)

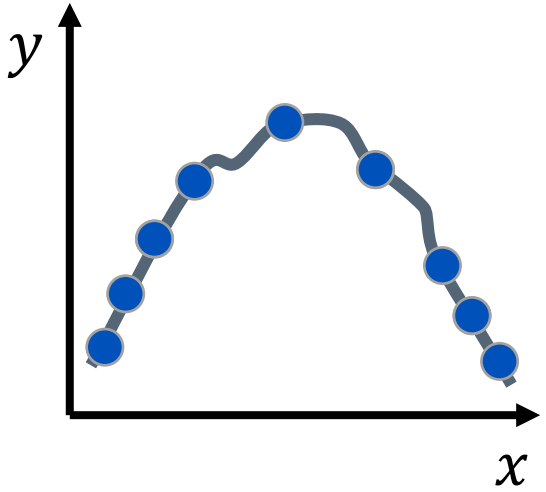
Training data

Testing data

What to do With a Large Variance? Part IV



Flexible model



More training data

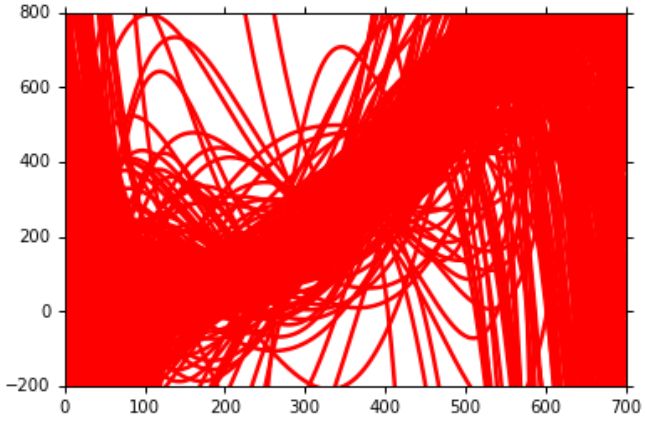
Real data distribution (not observable)

Training data

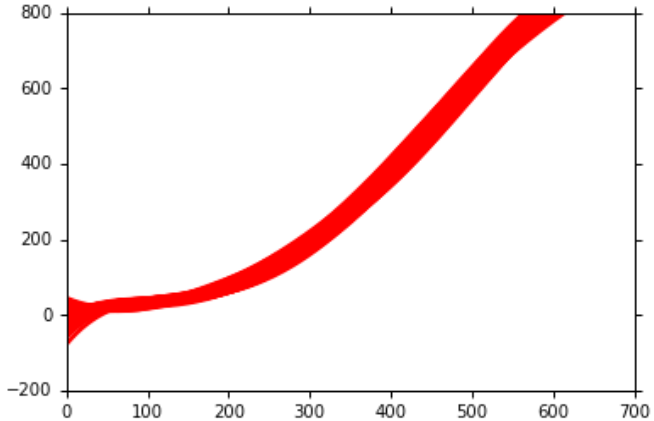
Testing data

What to do With a Large Variance? Part V

More data is very effective but not always practical.



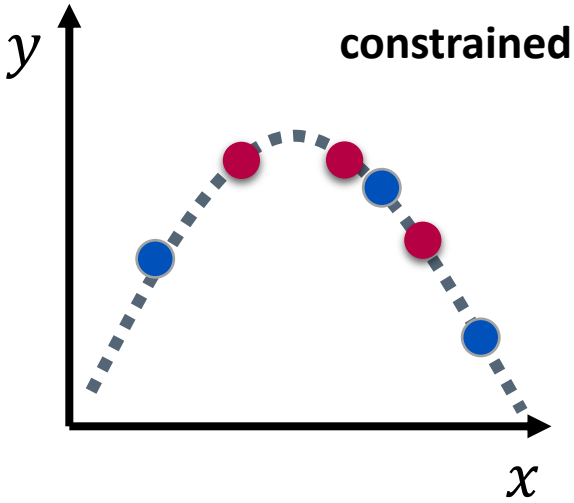
10 examples



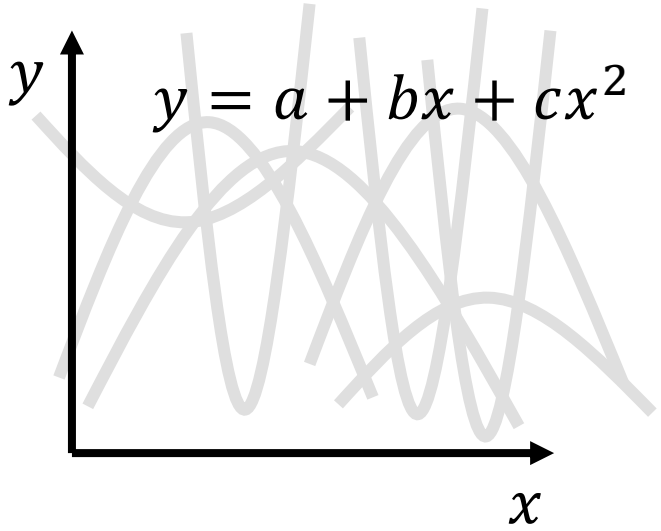
100 examples

Data augmentation

What to do With a Large Variance? Part VI



constrained model

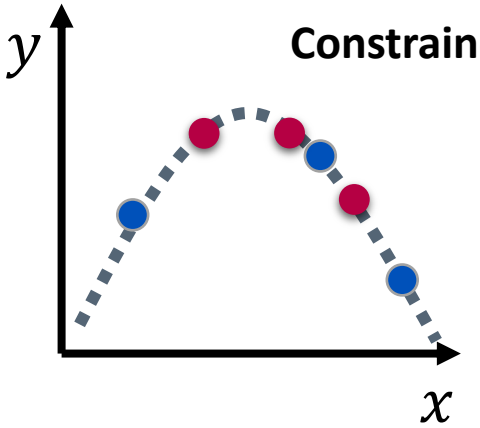


Real data distribution (not observable)

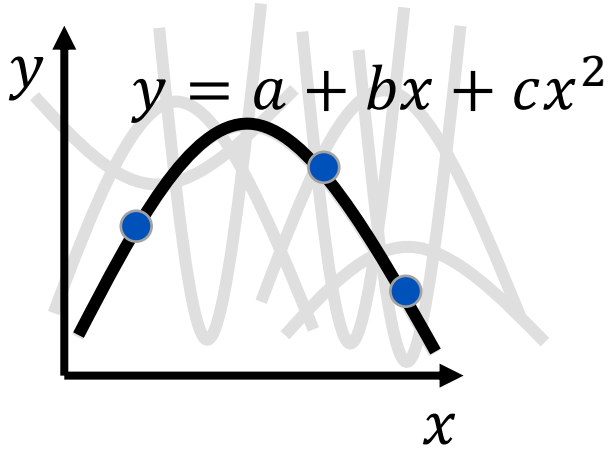
Training data

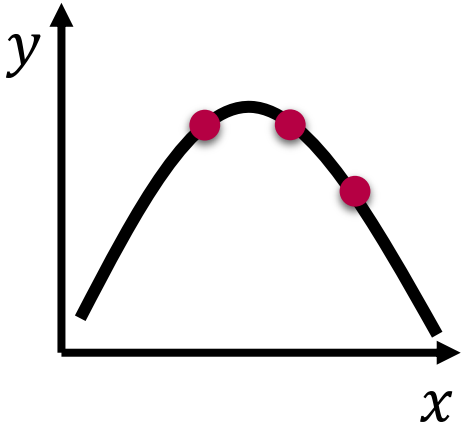
Testing data

What to do With a Large Variance? Part VII



Constrained model



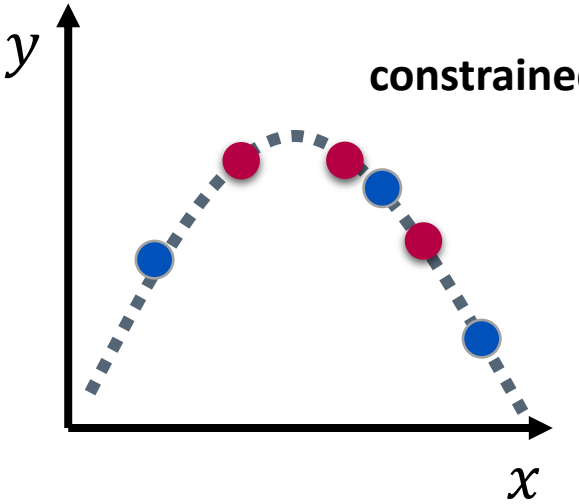


Real data distribution (not observable)

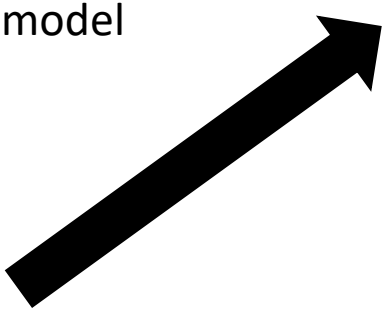
Training data

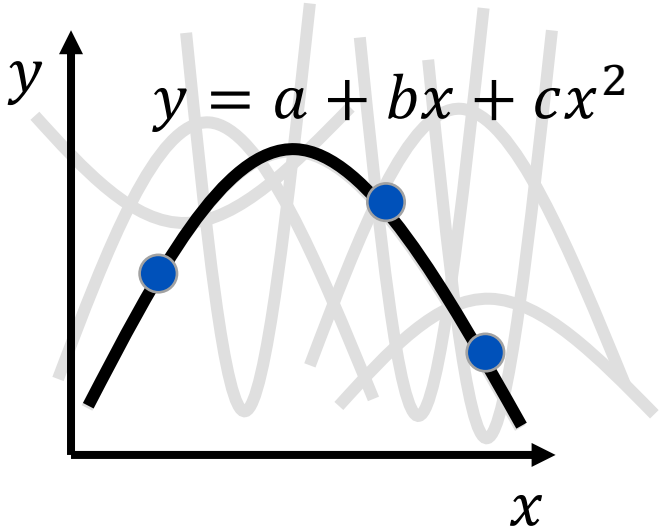
Testing data

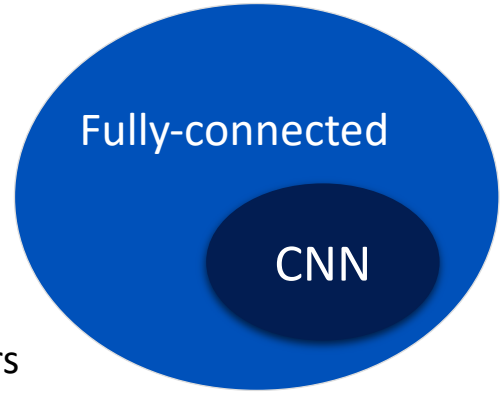
What to do With a Large Variance? Part VIII



constrained model

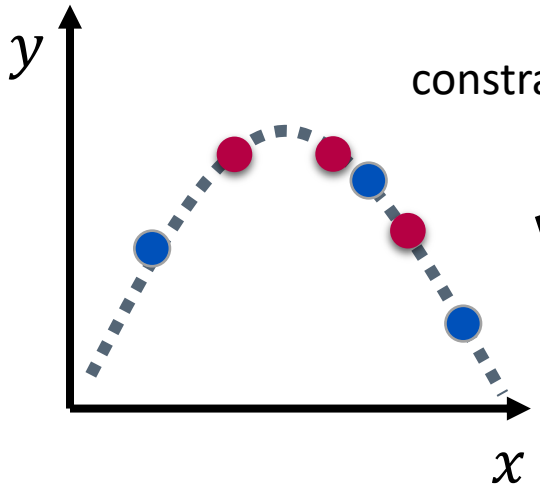




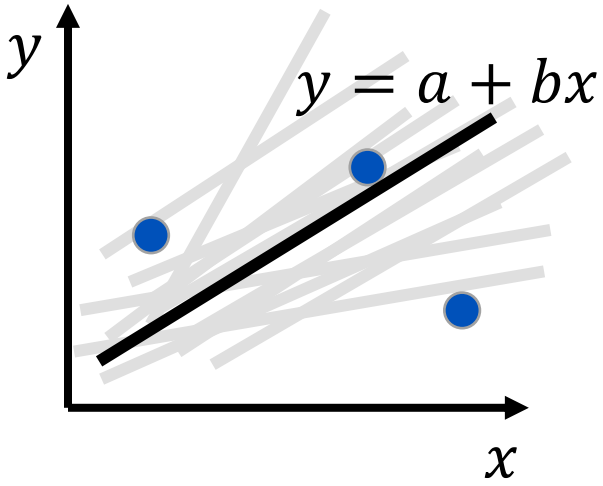


-Less features - regularization - less parameters, sharing parameters
-Early stopping - dropout

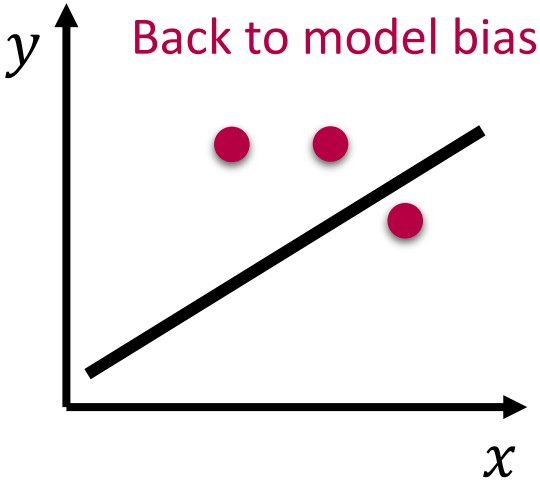
Overfitting



constrain too much



Back to model bias



Real data distribution (not observable)

Training data

Testing data