

Admin

Assignment 7 due Friday at 5pm

Project proposals due Thursday at 11:59pm

Grading update

Maximum likelihood estimation

Intuitive

Sets the probabilities so as to maximize the probability of the training data

Problems?

Overfitting!

- Amount of data
- particularly problematic for rare events
- Is our training data representative

Basic steps for probabilistic modeling

Step 1: pick a model

Step 2: figure out how to estimate the probabilities for the model

Step 3 (optional): deal with overfitting

Probabilistic models

Which model do we use, i.e. how do we calculate p(feature, label)?

How do train the model, i.e. how to we we estimate the probabilities for the model?

How do we deal with overfitting?

Priors

Coin1 data: 3 Heads and 1 Tail Coin2 data: 30 Heads and 10 tails Coin3 data: 2 Tails Coin4 data: 497 Heads and 503 tails

If someone asked you what the probability of heads was for each of these coins, what would you say?

Training revisited

From a probability standpoint, MLE training is selecting the Θ that maximizes:

 $p(\theta \mid data)$

i.e.

 $\operatorname{argmax}_{\theta} p(\theta \mid data)$

We pick the most likely model parameters given the data

Estimating revisited

We can incorporate a prior belief in what the probabilities might be!

To do this, we need to break down our probability

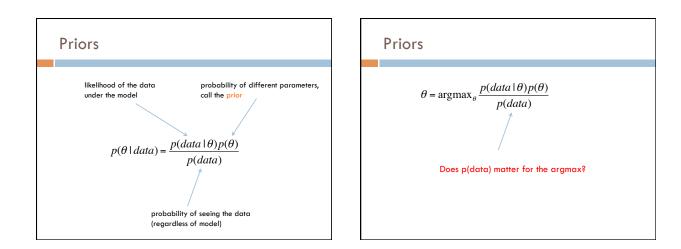
 $p(\theta \mid data) = ?$

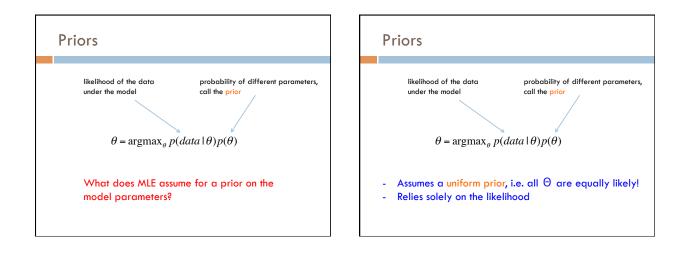
(Hint: Bayes rule)

Estimating revisited

What are each of these probabilities?

 $p(\theta \mid data) = \frac{p(data \mid \theta)p(\theta)}{p(data)}$





3

