

Lecture 10, Tuesday w6, 2014-10-21

Things we covered:

- Predicting outcomes with a Beta-Binomial model and a Dirichlet-Multinomial model.
- Effects of the “pseudo-counts” in these models.
- Alternative to adaptive model: fit parameters to whole file first, and encode in header. You’ll see the trade-offs for yourself in the assignment.
- Can have separate set of counts for each ‘context’ we predict in. For example, each possible setting of a small window of pixels.

Next time we’ll carry on talking about making predictions in a context, combining the predictions from contexts of different sizes.

Check your progress

Do you think the Dirichlet parameters for something like characters or words from language should be large or small? Why?

Explain how setting pseudo-counts to zero in the Beta-Binomial and Dirichlet-Multinomial models would break an arithmetic coding scheme.

Recommended reading

We’ve now done the ‘week 5’ slides except PPM which I’ll cover next time. Mark anything that’s unclear or that needs expanding on NB.

Extra reading

If keen, you could read Section 28.3, pp351–353 of MacKay. This section discusses ‘two-part codes’, which send parameters then an encoding in more detail. The ‘bits back’ method is an ingenious way of getting around the inefficiency of ‘sending the parameters twice’. Any extra details in these pages (not mentioned in lectures) are all non-examinable.