10/25/19

Physics 8805: Learning from Data: Lecture 18

Notebooks to use today:
Ⓐ . PyMC3_intro.ipynb
Ⓑ . PyMC3_docs_getting_started.ipynb

Follow-up to parallel tempering:
We wrote the Metropolis condition was that we accepted exchanges with probability

$$r = \min\left\{1, \frac{p_\beta(\vec{\theta}_{t,i+1}|D,\beta_i,I)\, p_\beta(\vec{\theta}_{t,i}|D,\beta_{i+1},I)}{p_\beta(\vec{\theta}_{t,i}|D,\beta_i,I)\, p_\beta(\vec{\theta}_{t,i+1}|D,\beta_{i+1},I)}\right.$$

This form is more intuitive if we connect back to the physics analogy.
$p_\beta(\vec{\theta}|D,I) \propto p(D|\vec{\theta},I)^\beta p(\vec{\theta}|I) \propto e^{-\beta(-\log p(D|\vec{\theta},I))} p(\vec{\theta}|I)$
and $-\log p(D|\vec{\theta},I)$ is the potential energy $U(\vec{r})$. Then the priors drop out of the ratio in r (the prior is independent of β) leaving:

$$r = \min\left\{1, e^{-(\beta_i-\beta_{i+1})(U(\vec{r}_{i+1})-U(\vec{r}_i))}\right\}$$

so the four terms in r arise from the Boltzman factors of the energy difference at the two temperatures.

Preview to PyMC3 notebooks:
Ⓐ Starts with sampling to find the posterior for μ, the mean of a distribution, given data that is generated according to a normal distribution with mean zero.
- Try changing the true μ, sampling sigma as well.
- The standard deviation is initially fixed at 1.

In the code, one specifies priors and then the likelihood. This is sufficient to specify the posterior for μ by Bayes' theorem, up to a normalization:

$$p(\mu|D,\sigma) \propto p(D|\mu,\sigma)\, p(\mu|\mu_\mu^0,\sigma_\mu^0)$$

cores = # of parallel chains

pdf to sample — observed — likelihood, specify last $\sim N(\mu,\sigma)$ — hyperparameters $\sim N(0,1)$ [specify first]