Lecture 12

- Gibbs Sampling

  * Draw from parameter conditional posteriors one at a time.

e.g.

$x^0 \sim p(x)$

for $i = 1, 2, \ldots N$ do

$x_1^{(i)} \sim p(X_1 = x_1 \mid X_2 = x_2^{(i-1)}, \ldots)$

$x_2^{(i)} \sim p(X_2 = x_2 \mid X_1 = x_1^{(i)}, \ldots)$

$\vdots \qquad \vdots$

$x_{ndim}^{(i)} \sim p(X_{ndim} = x_{ndim} \mid X_1 = x_1^{(i)} \ldots)$

$i = i + 1$

end for

Pros $\Rightarrow$ (i) If conditionals are analytic, no rejection
(ii) very short auto-correlation

Cons $\Rightarrow$ (i) requires conjugate priors for analytic conditionals.
(ii) re-writing the data model in such conditionals is arduous.

Practical Evidence Evaluation & Model Selection

① Savage-Dickey Density Ratio

* good for nested models, i.e. M_1 = noise
  M_2 = noise + signal.
* this means $p_1(d|n) = p_2(d|A=0, n)$

$$\therefore \mathcal{B} = \frac{Z_2}{Z_1} = \frac{p(A=0)}{p(A=0|d)}$$

$$= \frac{\text{prior}(A=0)}{\text{posterior}(A=0)}$$

② Product-space Sampling

* search over ALL MODELS at the same time.
* a model indexing variable is also searched over.
* e.g. if n_{model} < threshold :
  activate model 1.
  else :
  activate model 2.
* $\mathcal{B}$ = RATIO of MCMC samples in M2 versus M1 range.

3 Thermodynamic Integration

* uses parallel tempering, where many chains launch in parallel.

$$p(d|\theta)_\beta = p(d|\theta)^\beta$$
$$= p(d|\theta)^{1/T}$$

TEMPERATURE

* temperature communication improves mixing.

$$\ln Z = \int_0^1 \langle \ln p(d|\theta) \rangle_\beta \, d\beta$$

4 Nested Sampling

* very different.

* start with many points spread throughout prior volume!
  - ↳ rank points by likelihood
  - ↳ remove lowest-ranked point
  - ↳ add newer, higher-likelihood point.
  - ↳ cluster moves up likelihood surface.
  - ↳ evidence integral done along way.

$$dX = p(\theta)\, d^n\theta \quad ; \quad X(\lambda) = \int_{p(d|\theta)>\lambda} p(\theta)\, d^n\theta$$

iso-likelihood contour.

$$Z = \int p(d|\theta)\, p(\theta)\, d^n\theta$$
$$\Rightarrow \int_0^1 \mathcal{L}(X)\, dX$$