Office hour: today 15:00 - 16:00

Kavli building, room K02.

Second Example Sheet: On Moodle (WS2.pdf)

Example class will be a week on Friday.

Suppose we have MC $x_0, x_1, \dots, x_{n-1}$ chain length n (with burn-in removed)

has as stationary dist P.

$x_i \in \mathcal{X}$ sample space

if we have real-valued function $f(x)$, i.e. $f: X \longrightarrow \mathbb{R}$.

the expectation of f $\qquad E[f] = \int dx \; P(x) \, f(x)$

this can be estimated $\qquad S_n = \frac{1}{n} \sum_{i=0}^{n-1} f(x_i)$

this will converge $S_n \to E[f]$ in limit $n \to \infty$.

Variance in the estimator S_n,

$$\boxed{\begin{array}{l} \sigma^2 = \frac{1}{n} Var[f] \\ \text{if } x_i \overset{iid}{\sim} P \text{ then} \end{array}}$$

where $Var[f] = \int dx \; P(x) \left(f(x) - E[f]\right)^2$; this also has to be estimated from sample

$$\Rightarrow \quad \sigma^2 = \frac{1}{n(n-1)} \sum_{i=0}^{n-1} \left(f(x_i) - \cancel{E[f]}\, S_n\right)^2$$

Because nearby points in the MC are correlated, the actual variance in S_n will be higher.

$$\sigma^2 = \frac{\tau_f}{n} \mathrm{Var}[f]$$

Def of τ_f the Integrated Autocorrelation Time (IAT).

The effective number of independent samples in MC is

$$n_{eff} = n/\tau_f \qquad \tau_f \geq 1.$$

IAT depends on a choice of f. Typically consider $f(x) = x^k$ for $k = 0, 1, 2, \ldots d-1$. These "component functions" give us d different $\tau_k \equiv \tau_{f_k}$.

Use $\tau = \max_k (\tau_k)$. $\leftarrow$ "THE IAT"

Recall

$$S_n = \frac{1}{n} \sum_i f(x_i)$$
$$= \frac{1}{n} \sum_i f_i$$

$$\sigma^2 = \frac{1}{n^2} \sum_i \sum_j \mathrm{Cov}(f_i, f_j)$$
$$= \frac{1}{n} \sum_\varepsilon \underbrace{\sum_\delta \mathrm{Cov}(f_\varepsilon, f_{\varepsilon+\delta})}_{\text{doesn't depend on } \varepsilon}$$

$\varepsilon = i+j$
$\delta = i-j$

Assuming MC is stationary.

$$\approx \sum_\delta \mathrm{Cov}(f_\varepsilon, f_{\varepsilon+\delta})$$

estimate this from chain use sample covariance of all pairs of points δ apart.

There exist fast FFT-based methods for evaluating this.

a couple of other diagnostic tests.

- Gelman-Rubin statistic.

  compares multiple variance estimates from different chains

- Geweke statistic.

  compares parts of a single chain.

These are all only consistency checks.

Target dist $\qquad P(x,y) \propto \exp\left(\frac{-[x^2+y^2-2\beta xy]}{2[1-\beta^2]} \right)$

Gibbs needs conditional distributions

$$x, y \sim P$$

$$x|y \sim \mathcal{N}(\underbrace{\beta y}_{=\mu}, \underbrace{1-\beta^2}_{=\sigma^2})$$

think about $P(x,y)$ just as a function of x

$$\cancel{P(x)} \propto P(x|y) \propto \exp\left(\frac{-x^2+2\beta xy}{2(1-\beta^2)} \right)$$

$$\propto \exp\left(\frac{-(x-\beta y)^2 - \beta^2 y^2}{2(1-\beta^2)} \right)$$

$$\propto \exp\left(\frac{-(x-\beta y)^2}{2(1-\beta^2)} \right)$$

by symmetry $\quad y|x \sim \mathcal{N}(\beta x, 1-\beta^2)$

$$E(x,y) = -\log P(x,y)$$

$$= \frac{x^2+y^2-2\beta xy}{2(1-\beta^2)}$$

$$\frac{\partial E}{\partial x} = \frac{x-\cancel{2}\beta y}{1-\beta^2} \qquad \text{sim} \qquad \frac{\partial E}{\partial y} = \frac{y-\beta x}{1-\beta^2}.$$