STT 465

Gibbs Sampler

Gibbs Sampler

$\Rightarrow$ So far we have discussed models where the joint posterior distribution have a closed form.

$\Rightarrow$ However in most models, the joint posterior distribution does not have a closed form.

$\Rightarrow$ In those cases we can use Monte Carlo Markov Chain algorithms to draw samples from the joint posterior without explicitly drawing from that distribution.

$\Rightarrow$ Distributions we have discussed so far

Prior $p(\theta_1, \theta_2)$

Joint posterior $p(\theta_1, \theta_2 \mid y)$

Fully conditional distribution $p(\theta_1 \mid y, \theta_2)$; $p(\theta_2 \mid y, \theta_1)$

$\Rightarrow$ In Gibbs sampler we draw samples iteratively from fully conditional distributions.

Gibbs Sampler

$\Rightarrow$ Target Posterior Density: $p\left(\theta_1, \theta_2, \ldots, \theta_p \mid y\right)$

$\Rightarrow$ Fully Conditionals $p\left(\theta_j \mid y, ELSE\right) = p\left(\theta_j \mid y, \theta_{-j}\right)$

$\Rightarrow$ Algorithm

- Initialize parameters (use values that have no-zero prior prob.)
- Iteratively sample from

$$\theta_1^s \sim p\left(\theta_1 \mid y, \theta_2^{s-1}, \ldots, \theta_p^{s-1}\right)$$

$$\theta_2^s \sim p\left(\theta_2 \mid y, \theta_1^s, \theta_3^{s-1}, \ldots, \theta_p^{s-1}\right)$$

$$\vdots$$

$$\theta_p^s \sim p\left(\theta_2 \mid y, \theta_1^s, \theta_2^s, \ldots, \theta_p^s\right)$$

Gibbs Sampler

$\Rightarrow$ The algorithm generates a dependence sequence of samples of

$$\theta^s = \left[\theta_1^s, \theta_2^s, \ldots, \theta_p^s\right]$$

$\Rightarrow$ After convergence these samples can be regarded as draws from the joint posterior density.

$\Rightarrow$ Given the way these samples were generated, the sequence is not an IID sequence.

$\Rightarrow$ Samples are not independent.

$\Rightarrow$ Markov Chain Property. However, the sequence has the property that

$$p\left(\theta^s, \theta^{s-1}, \theta^{s-2}, \ldots, \theta^0 \mid y\right) = p\left(\theta^s \mid \theta^{s-1}, y\right) \times p\left(\theta^{s-1} \mid \theta^{s-2}, y\right) \times \ldots \times p\left(\theta^1 \mid \theta^0, y\right)$$

$\Rightarrow$ Since θ^s can be regarded as a sample from the posterior distribution

$\Rightarrow$ $g\left(\theta^s\right)$ can be regarded as a sample from $p\left(g\left(\theta\right) \middle| y\right)$

Gibbs Sampler

$\Rightarrow$ Diagnostics

- $\Rightarrow$ Some of the original samples may need to be discarded (burn-in)
- $\Rightarrow$ Due to auto-correlation, we may need to thin (i.e., take x samples out of q)
- $\Rightarrow$ Due to auto-correlation, MC errors cannot be computed using the formulas we used so far.

$\Rightarrow$ Diagnostics:

- $\Rightarrow$ Trace plot
- $\Rightarrow$ Auto-correlation between samples
- $\Rightarrow$ Time-series MC error
- $\Rightarrow$ Effective number of samples