Gibbs Sampling

Suppose that x is drawn from a bivariate Normal distribution with mean, μ, and variance, Σ. Let

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \; \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}.$$

The distribution of x_1 conditional on x_2 is Normal:

$$x_1|x_2 \sim \mathcal{N}\left(\mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}}(x_2 - \mu_2), \frac{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}{\Sigma_{22}}\right).$$

Similarly, the distribution of x_2 conditional on x_1 is:

$$x_2|x_1 \sim \mathcal{N}\left(\mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}}(x_1 - \mu_1), \frac{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}{\Sigma_{11}}\right).$$

Suppose that it is easy for us to compute the conditional distribution, but not so easy to compute the joint distribution. The Gibbs sampler provides a strategy for sampling from the joint distribution. It is a recursive strategy for computing a sequence of values of x, $x^{(1)}, x^{(2)}$,... . We could imagine collecting a sequence of length N in the matrix, M :

$$M = \left[x^{(1)}, x^{(2)}, ..., x^{(N)}\right].$$

The first column of M is the starting point of the Gibbs sampler. How this is chosen is not very important, but for definiteness suppose that $x^{(1)} = \mu$. The idea is that the way in which the Gibbs sampler works has the implication that

$$\lim_{N\to\infty} frequency\left[\text{`}x^{(j)} \text{ close to } \bar{x}\text{'}\right] \to p(\bar{x}),$$

where $p(x)$ is the joint distribution of x (in this case, the bivariate Normal). For example, for N large enough the histogram of the first row of M coincides with the actual marginal density of x_1, in this case $\mathcal{N}(\mu_1, \Sigma_{11})$.

The Gibbs sampler works like this. Suppose that the columns up to $s-1$ have been computed and we now want to compute the elements of the s^{th} column of M. We proceed as follows. The element in the first row of the s^{th} column, $x_1^{(s)}$, is a random draw from $x_1|x_2^{(s-1)}$, where $x_2^{(s-1)}$ is the 2×1 element of $x^{(s-1)}$. The second element of the s^{th} column of M is a random draw from $x_2|x_1^{(s)}$.