11 - Gibbs sampling - 01

Master in Foundations of Data Science
Bayesian Statistics and Probabilistic Programming
Fall 2018-2019

Josep Fortiana

Wednesday, December 05, 2018

Universitat de Barcelona

11 - Gibbs sampling - 01

Gibbs sampling (two-stage)

Slice sampling

Gibbs sampling (multistage)

Slice sampling for a product pdf

11 - Gibbs sampling - 01

General idea

Gibbs sampling is a *relaxation* algorithm.

Meaning:

Approximate a multivariate simulation by a sequence, where variables are updated one at a time.

Setting (two-stage)

Two r.v. X and Y, with joint pdf $f(x,y)$,

Marginal densities $f_X(x)$, $f_Y(y)$,

Conditional densities:

$$f_{Y|X=x} = \frac{f(x,y)}{f_X(x)}, \qquad f_{X|Y=y} = \frac{f(x,y)}{f_Y(y)},$$

Dual sequence

Initial: $x_0 \in Supp(X)$.

We generate a dual sequence:

$$(x, y) = \{(x_t, y_t); t \in \mathbb{Z}, t \geq 1\},$$

defined by:

$$\begin{cases} y_t & \sim & f_{Y|X=x_{t-1}} \\ x_t & \sim & f_{X|Y=y_t} \end{cases}$$

Properties of the sequence

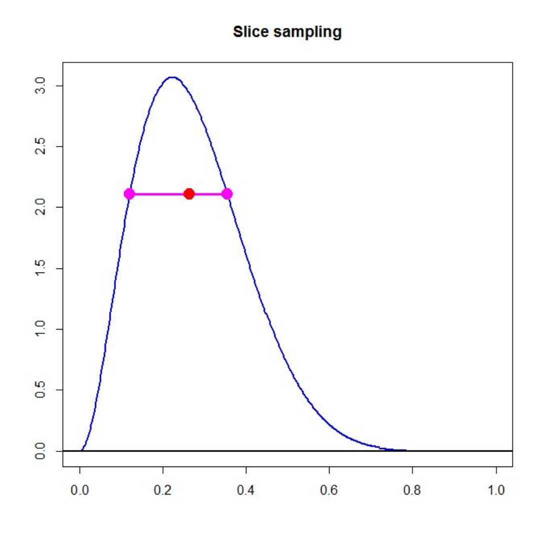
The sequence is a trajectory of a Markov chain.

The chain converges to the joint distribution

$$\sim f(x, y).$$

11 - Gibbs sampling - 01

Simple slice sampling



Rationale of the method

A common principle with the acceptance-rejection method:

For a pdf f, a uniform distribution on the region $\mathscr{A}(f) \subset \mathbb{R}^2$ comprised between the X axis and the curve $y = f(x)$, has X marginal pdf precisely f.

It is a particular case of Gibbs sampling.

Slice sampling as Gibbs sampling

Given an univariate pdf $f(x)$ we construct the following bivariate pdf:

$$f(x, y) = \mathbb{1}_{\mathscr{A}(f)}(x, y), \qquad (x, y) \in \mathbb{R}^2,$$

where $\mathscr{A}(f) \subset \mathbb{R}^2$ is the region of $\mathbb{R}^2$ comprised between the X axis and the curve $y = f(x)$.

11 - Gibbs sampling - 01

Gibbs sampling (two-stage)

Slice sampling

Gibbs sampling (multistage)

Slice sampling for a product pdf

Multivariate context

Suppose we want to simulate a random vector:

$$\boldsymbol{X} = (X_1, \ldots, X_k),$$

having joint pdf $f(\boldsymbol{x})$, either unknown or complicated, but for each i, the conditional pdf of:

$$X_i \,|(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$$

is known and easy to simulate.

Gibbs sampling: first cycle, first step

Set an initial value: $\boldsymbol{x}_0 = (x_{01}, \dots, x_{0k})$.

Update $x_{01} \to x_{11}$, by simulation of:

$$X_1 \,|(X_2 = x_{02}, \dots, X_k = x_{0k}) \;,$$

This is called a *full conditional* pdf.

Then the current running value is:

$$\boldsymbol{x} = (x_{11}, x_{02}, \dots, x_{0k}).$$

Gibbs sampling: first cycle, second step

Update $x_{02} \to x_{12}$, by simulation of:

$$X_2 \,|(X_1 = x_{11}, X_3 = x_{03}, \ldots, X_k = x_{0k}) \;,$$

and continue in this fashion until updating $x_{0k} \to x_{1k}$, by simulation of:

$$X_k \,|(X_1 = x_{11}, X_2 = x_{12}, X_3 = x_{13}, \ldots, X_{k-1} = x_{1,k-1}) \;,$$

which ends the first cycle.

Gibbs sampling: second and following cycles

Similarly, in the second cycle $\boldsymbol{x}_1 = (x_{11}, \dots, x_{1k})$ is updated to $\boldsymbol{x}_2 = (x_{21}, \dots, x_{2k})$ and continue further by the same procedure.

The result is a sequence $\{\boldsymbol{X}_n\}_{n \in \mathbb{N} \cup \{0\}}$, which under fairly general conditions* can be proved to satisfy:

$$\boldsymbol{X}_n \xrightarrow[n \to \infty]{\mathscr{L}} \boldsymbol{X}.$$

*Not always!

11 - Gibbs sampling - 01

Gibbs sampling (two-stage)

Slice sampling

Gibbs sampling (multistage)

Slice sampling for a product pdf

Target: a product pdf

Assume the target pdf is a product of n functions:

$$f(x) = \prod_{i=1}^{n} g_i(x).$$

We construct an $(n+1)$-dimensional pdf, adding n auxiliary variables, $u_1, \ldots, u_n$, and define:

$$f(x, u_1, \ldots, u_n) = \prod_{i=1}^{n} \mathbb{1}_{\mathscr{A}_i(g_i)}(x, u_i),$$

Setting the slice sampling

$\mathscr{A}_i(g_i) \subset \mathbb{R}^2$ is the region of the (x, u_i) plane comprised between the X axis and the curve $u_i = g_i(x)$:

$$\mathbb{1}_{\mathscr{A}_i(g_i)}(x, u_i) = \begin{cases} 1, & \text{if } u_i < g_i(x), \\ 0, & \text{otherwise.} \end{cases}$$

That is,

$$u_i | x \sim \mathrm{Unif}(0, g_i(x)), \quad 1 \leq i \leq n.$$

Conditional pdf for x

Conditional to a value of $\boldsymbol{u} = (u_1, \ldots, u_n)$,

$$f(x|\boldsymbol{u}) \propto f(x, \boldsymbol{u}) = \prod_{i=1}^{n} \mathbb{1}_{\mathscr{A}_i(g_i)}(x, u_i),$$

which is equal to 1 when <u>all</u> the u_i belong to the $(0, g_i(x))$ interval.

Conditional pdf for x

That is, given $\boldsymbol{u} = (u_1, \ldots, u_n)$, the distribution of $x|\boldsymbol{u}$ is uniform on the subset of the real line defined by:

$$\mathscr{A}_x(\boldsymbol{u}) = \cap_{i=1}^{n} \mathrm{Proj}_x \left(\mathscr{A}_i(g_i) \cap [\text{The horizontal line at } u_i] \right),$$

the intersection of the individual slices (might be a reunion of disjoint intervals).

Implementation

Set an initial x_0. The j-th step is:

Generate independently:

$$u_{ji} \sim \mathrm{Unif}(0, g_i(x_{j-1})), \quad 1 \leq i \leq n,$$

Generate:

$$x_j \sim \mathrm{Unif}(\mathscr{A}_x(\boldsymbol{u}_j)).$$