DEMONSTRATIO MATHEMATICA
Vol. XXXII No 4 1999

Paweł Kopciuszewski

THE CLASS OF JOINT DENSITIES UNIQUELLY DETERMINED BY CONDITIONALS USED IN GIBBS SAMPLING

Abstract. This work is about uniqueness of joint densities, where the conditionals are given. We formulate a necessary and sufficient condition for uniqueness of joint densities, at the same time answering the question risen by Arnold and Press [1] as well as Chan [4]. The theorem is illustrated in a few examples and discussed in context of Gibbs sampling method.

1. Introduction

Lately, the most dynamically growing Monte-Carlo methods are the ones based on the Markov chains theory. The most popular of those are Gibbs sampling and Metropolis methods. They find many applications in statistical calculations, in evaluations of multidimensional integrals, for example in estimation of expectations of a-posteriori densities. Most of the Monte-Carlo methods usage leads to generating samples from joint densities. Gibbs sampling method find an application in generating samples from a-posteriori densities in Hierarchical Generalized Linear Mixed Models (Hobert and Casella [8]).

Gibbs sampling method allows us to generate samples from a density $\pi(x_1, \dots, x_n)$ by given conditional densities $\pi(x_i|x_j, i \neq j)$ $i = 1, \dots, n$. Of course, a density $\pi(x_1, \dots, x_n)$ must be uniquely determined by its conditional densities $\pi(x_i|x_j, i \neq j)$ $i = 1, \dots, n$, for the chain to be ergodic. Even though it is not a sufficient condition for ergodicity of the chain. Besag [3] found that a density $\pi(x_1, \dots, x_n)$ is unique by its conditional densities, if the support of the joint density is a product of some finite sets.

Hobert, Robert and Goutis [9] proposed convergence conditions for Gibbs sampling. We give conditions for uniqueness of a joint density $\pi(x)$ by its conditional densities $\pi(x_i|x_j, i \neq j)$, $i = 1, \dots, n$. Uniqueness of a joint density is not a sufficient condition for the ergodicity of the chain. In section

4 we analyse the example when a joint density π is uniquely determined by its conditional densities but Markov chain generated by Gibbs sampling method is not ergodic. Besides we proof when the joint density $\pi(x)$ is determined uniquely by conditional densities $\pi(w_i|w_j, i \neq j)$, where w_i are some subvectors of the vector x, $i = 1, \ldots, m < n$. Theorems are proved for σ-finite measures.

2. Gibbs sampling method

Let $(\mathcal{R}^n, \mathcal{B}^{(n)}, P)$ be a probability space with σ−field $\mathcal{B}^{(n)}$ and probability measure P. Let ν be a σ−finite measure defined on $\mathcal{B}^{(n)}$, such that P is absolutely continuous with respect to ν, and let π be the density of the measure P with respect to the measure ν. Let $N \in \mathcal{B}^{(n)}$ be the support of the density π. Let conditional densities $\pi(x_i|x_j, i \neq j)$, $i = 1, \ldots, n$ exist for all $(x_1, \ldots, x_n) \in N$. Let N_i be the support of the conditional density $\pi(x_i|x_j, i \neq j)$, $i = 1, \ldots, n$. We define the Markov chain $\{x^{(k)}\}$ with states from the set N as follows.

Let $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in N$ be a starting point of the chain.

The first step is defined as follows:

$x_1^{(1)}$ is the realization from $\pi(x_1|x_j^{(0)}, j > 1)$,

$x_2^{(1)}$ is the realization from $\pi(x_2|x_1^{(1)}, x_j^{(0)}, j > 2)$,

$\vdots$

$x_{n-1}^{(1)}$ is the realization from $\pi(x_{n-1}|x_1^{(1)}, \ldots, x_{n-2}^{(1)}, x_n^{(0)})$,

$x_n^{(1)}$ is the realization from $\pi(x_n|x_j^{(1)}, j < n)$.

The k-th step $x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$ with given $k - 1$-th step is defined as follows:

$x_1^{(k)}$ is the realization from $\pi(x_1|x_j^{(k-1)}, j > 1)$,

$x_2^{(k)}$ is the realization from $\pi(x_2|x_1^{(k)}, x_j^{(k-1)}, j > 2)$,

$\vdots$

$x_{n-1}^{(k)}$ is the realization from $\pi(x_{n-1}|x_1^{(k)}, \ldots, x_{n-2}^{(k)}, x_n^{(k-1)})$,

$x_n^{(k)}$ is the realization from $\pi(x_n|x_j^{(k)}, j < n)$.

The chain $\{x^{(k)}\}$, $k = 1, 2, \ldots$, is called the Gibbs sampler. The transition probability $P(x, A)$ from the state x into any ν−measurable set A is expressed by the formula $P(x, A) = \int_A \prod_{i=1}^n \pi(y_i|x_j, y_k, i < j, i > k)\nu(d(y))$, where $y = (y_1, \ldots, y_n)$. The chain $\{x^{(k)}\}$, $k = 1, 2, \ldots$ has the limiting density π, if it is irreducible and aperiodic.

Then for all $\nu-$measurable sets $A \subset N$, the joint density π satisfies the following condition

$$\int_N P(x, A)\pi(x)\nu(d(x)) = P(A).$$

3. Two theorems about uniqueness of a joint density

The main result of the paper is a generalization of Besag's result [3] and Hobert, Robert and Goutis result [9] on probability measures P absolutely continuous with respect to some σ-finite measures ν. These authors consider the case of ν equals Lebesque measure without explicity pointing out to this special case. The basic definition [2.2, 9] is hard to understand but our theorems develop important ideas contained in this definition.

As we pointed out in the introduction, Besag [3] proved that, if $N = N_1 \times \ldots \times N_n$ where $N_i \subset \mathcal{R}$, $i = 1, \ldots, n$, are some finite sets, then every joint density π defined on the support N is unique by its conditional densities $\pi(x_i|x_j, i \neq j)$, $i = 1, \ldots, n$. In our work we are going to show that results of [3,9] can be generalized to a wider class of supports. When $n = 2$ Besag's proof shows as follows.

In this case $N = N_1 \times N_2$, where $N_1 = \{u_1, \ldots, u_n\}$, $N_2 = \{v_1, \ldots, v_m\}$.

From the definition of the conditional density it appears that for all $(u, v) \in N$ comes out the relation

$$\frac{\pi(u, v)}{\pi(u_1, v_1)} = \frac{\pi(u|v)\pi(v|u_1)}{\pi(u_1|v)\pi(v_1|u_1)}.$$

Let π^* be a density defined on the support N, such that $\pi^*(u|v) = \pi(u|v)$ and $\pi^*(v|u) = \pi(v|u)$ for all $(u, v) \in N$. Then it appears

$$\sum_{(u,v)\in N} \frac{\pi(u, v)}{\pi(u_1, v_1)} = \sum_{(u,v)\in N} \frac{\pi^*(u, v)}{\pi^*(u_1, v_1)},$$

from which we obtain $\pi(u_1, v_1) = \pi^*(u_1, v_1)$. Similarly, it is proved that $\pi(u_i, v_j) = \pi^*(u_i, v_j)$ for all $(u_i, v_j) \in N$. Hence $\pi \equiv \pi^*$.

To make results of [3,9] wider to any support $N \in \mathcal{B}^{(n)}$ let us introduce the following definition.

Let R_A be a relation defined on the set $A \subset N$ as follows.

DEFINITION 1. A point $x \in A$ is in relation R_A to a point $y \in A$ which we write as xR_Ay, iff there exists a finite sequence of vectors $c^{(k)} \in R^n$, $k = 1, \ldots, l$ which have only one coordinate different from zero, such that

$$z^{(1)} = x + c^{(1)} \in A,$$
$$z^{(k+1)} = z^{(k)} + c^{(k+1)} \in A, \text{ where } k = 1, \ldots, l-1,$$
$$y = z^{(l)} \in A.$$

The relation R_A is an equivalence relation.

If N is of the form $N = N_1 \times \ldots \times N_n$, where $N_i \subset \mathcal{R}$, $i = 1, \ldots, n$, then xR_Ny holds for all $x, y \in N$. It is possible to construct a support N of a joint density $\pi(x_1, \ldots, x_n)$, such that N does not have a product form, and such that xR_Ny for all $x, y \in N$. Notice that if the set N consists of eight points as shown in figure 1, then the relation $x_iR_Nx_j$ holds for all $i, j = 1, \ldots, 7$. However, the point x_8 is not in relation R_N to any other point in N.

LEMMA 1. *If N is a connected and open set, then the relation xR_Ny holds for all $x, y \in N$.*

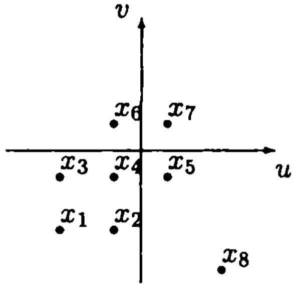

Figure 1: $x_iR_Nx_j$ and the point x_8 is not in relation R_N to the point x_i, $i, j = 1, \ldots, 7$

Proof. Because N is a connected and open set, there exists a path L in N (homeomorphic with interval $[0,1]$) connecting points x and y, for all $x, y \in N$. Moreover, the path L can be chosen so that, it consists of intervals parallel to the axes.

LEMMA 2. *If $\nu\{x \in N : \nu\{y \in N : x$ is not in relation R_N to $y\} > 0\} > 0$, then there exists a set $B^* \subset N$ with $\nu(B^*) > 0$ and $\nu(N \setminus B^*) > 0$ such that*

$$\text{(1)} \qquad \textit{if } x \in B^* \textit{ then } \nu\{y \in N \setminus B^* : xR_Ny\} = 0.$$

Proof. Let B' denote the set satisfying

$$B' = \{x \in N : \nu(B'_x) > 0\}.$$

where $B'_x = \{y \in N : x \text{ is not in relation } R_N \text{ to } y\}$.

Let B'' denote the set satisfying

$$B'' = \{x \in B' : \nu\{y \in N : xR_Ny\} > 0\}.$$

Suppose that $B'' \neq \emptyset$ and choose an arbitrary point $x \in B''$. Let the set B^* be defined as follows

$$B^* = \{y \in N : xR_Ny\}.$$

From the definition of the set B'' it follows that $\nu(B^*) > 0$. Notice that $B'_x \subset N \setminus B^*$ and $\nu(B'_x) > 0$. Hence $\nu(N \setminus B^*) > 0$. Finally (1) holds under the equivalence of the relation R_N.

Now suppose that $B'' = \emptyset$. Then B^* is an arbitrary set satisfying $B^* \subset B'$, $\nu(B^*) > 0$, $\nu(N \setminus B^*) > 0$. From this we conclude that condition (1) holds.

The theorem below gives a necessary and sufficient condition for the uniqueness of a joint density $\pi(x_1, \ldots, x_n)$ by conditional densities $\pi(x_i | x_j, i \neq j)$, $i = 1, \ldots, n$.

THEOREM 1. *A joint density $\pi(x_1, \ldots, x_n)$ is unique by its conditional densities $\pi(x_i | x_j, i \neq j)$, $i = 1, \ldots, n$, iff for all sets $A \subset N$ with $\nu(A) = 0$*

(2) $\nu\{x \in N \setminus A : \nu\{y \in N \setminus A : x \text{ is not in relation } R_{N \setminus A} \text{ to } y\} > 0\} = 0.$

Proof. Suppose that (2) holds for all sets $A \subset N$ with $\nu(A) = 0$. Hence for ν-almost all $x \in N \setminus A$ the following condition is satisfied

(3) $$\nu\{y \in N \setminus A : x \text{ is not in relation } R_{N \setminus A} \text{ to } y\} = 0.$$

Let π^* be a different density defined on the support N with respect to the measure ν such that

(4) $$\pi(x_i | x_j, i \neq j) = \pi^*(x_i | x_j, i \neq j),\ i = 1, \ldots, n,$$

for $\nu-$almost all $x = (x_1, \ldots, x_n) \in N$.

Let $A \subset N$ be the set defined as follows

$$A = \{x \in N : \ \pi(x_i | x_j, i \neq j) \neq \pi^*(x_i | x_j, i \neq j) \text{ for any } i = 1, \ldots, n\}.$$

For simplicity, N^* will be used to denote the set $N \setminus A$ and R will be used to denote the relation $R_{N \setminus A}$. Let $x \in N^*$ be a point satisfying (4) and such that

(5) $$\pi^*(x) = a\pi(x),$$

where $a > 0$.

Since xRy for $\nu-$almost all $y \in N^*$ there exists a finite sequence of vectors $c^{(k)} \in R^n$, $k = 1, \ldots, l$ which have only one coordinate different from zero such that

$$z^{(1)} = x + c^{(1)} \in N^*,$$
$$z^{(k+1)} = z^{(k)} + c^{(k+1)} \in N^*, \text{ where } k = 1, \ldots, l-1,$$
$$y = z^{(l)} \in N^*.$$

Let m_k be the number of non-zero coordinate of the vector $c^{(k)}$, $k = 1, \ldots, l$. Hence from (5)

$$\pi^*(x_{m_1} | x_i, m_1 \neq i)\pi^*(x_i, m_1 \neq i) = a\pi(x_{m_1} | x_i, m_1 \neq i)\pi(x_i, m_1 \neq i).$$

From the above and from the equation (4) it follows that

$$\pi^*(x_i, m_1 \neq i) = a\pi(x_i, m_1 \neq i).$$

The coordinates of the vectors x and $z^{(1)}$ are identical except for the coordinate of the index m_1. Hence

$$\pi^*(z_i^{(1)}, m_1 \neq i) = a\pi(z_i^{(1)}, m_1 \neq i).$$

From the above and from the equation (4) it now follows that

$$\pi^*(z_{m_1}^{(1)}|z_i^{(1)}, m_1 \neq i)\pi^*(z_i^{(1)}, m_1 \neq i) = a\pi(z_{m_1}^{(1)}|z_i^{(1)}, m_1 \neq i)\pi(z_i^{(1)}, m_1 \neq i)$$

Hence

$$\pi^*(z^{(1)}) = a\pi(z^{(1)}).$$

By induction it can be proved that

$$\pi^*(z^{(k)}) = a\pi(z^{(k)}), \ \ k = 1, \dots, l.$$

From this

$$\pi^*(y) = a\pi(y) \text{ for } \nu\text{-almost all } y \in N^*.$$

By integrating the both sides of the above equation, we obtain $a = 1$. From the equation (5) it follows that $\pi(x) = \pi^*(x)$. The point x was arbitrary chosen, hence $\pi(x) = \pi^*(x)$ for ν-almost all $x \in N$.

The reverse proposition is proved by contradiction.

Suppose that there exists a set $A \subset N$ with $\nu(A) = 0$ satisfying

$$\nu\{x \in N \setminus A : \nu\{y \in N \setminus A : x \text{ is not in relation } R_{N\setminus A} \text{ to } y\} > 0\} > 0,$$

From Lemma 2 it follows that there exists a set $B^* \subset N \setminus A$ with $\nu(B^*) > 0$ and $\nu(N \setminus B^*) > 0$ such that

$$\text{if } x \in B^* \text{ then } \nu\{y \in (N \setminus A) \setminus B^* : xR_{N\setminus A}y\} = 0. \tag{6}$$

The support N_i of the density $\pi(x_i|x_j, i \neq j)$, $i = 1, \dots, n$, is included either in the set B^* or in the set $N \setminus B^*$ with precision of a set zero measure, because xRy for all $x, y \in N_i$.

Let π^* be a density defined on the support N with respect to measure ν satisfying

$$\pi^*(x) = \begin{cases} a\pi(x), & \text{for } \nu - \text{almost all } x \in B^*, \\ a^*\pi(x), & \text{for } \nu - \text{almost all } x \in N \setminus B^*, \end{cases}$$

where a is a positive constant such that $1 - aP(B^*) > 0$, and $a^* = \frac{1-aP(B^*)}{P(N\setminus B^*)}$. The conditional densities $\pi^*(x_i|x_j, i \neq j)$ are given by the formula

$$\begin{aligned} \pi^*(x_i|x_j, i \neq j) &= \frac{\pi^*(x)}{\int_{N_i} \pi^*(x_1, \dots, x_n)\nu(d(x_i))} \\ &= \frac{a\pi(x)}{a\int_{N_i} \pi(x_1, \dots, x_n)\nu(d(x_i))} = \pi(x_i|x_j, i \neq j), \end{aligned}$$

for $\nu-$almost all $x \in B^*$, and by the formula

$$\pi^*(x_i|x_j, i \neq j) = \frac{\pi^*(x)}{\int_{N_i} \pi^*(x_1, \dots, x_n)\nu(d(x_i))}$$
$$= \frac{a^*\pi(x)}{a^* \int_{N_i} \pi(x_1, \dots, x_n)\nu(d(x_i))} = \pi(x_i|x_j, i \neq j),$$

for $\nu-$almost all $x \in N \setminus B^*$. This implies that $\pi(x_i|x_j, i \neq j) = \pi^*(x_i|x_j, i \neq j)$, $i = 1, \dots, n$ for $\nu-$almost all $x \in N$, and consequently, that the conditional densities $\pi(x_i|x_j, i \neq j)$, $i = 1, \dots, n$, do not determine uniquely the joint density $\pi(x_1, \dots, x_n)$. Thus the theorem is proved.

COROLLARY 1. *A joint density $\pi(x_1, \dots, x_n)$ is unique by its conditional densities $\pi(x_i|x_j, i \neq j)$, $i = 1, \dots, n$, if its support is a connected and open set.*

The proof follows from Lemma 1 and Theorem 1.

In the literature there are also considered some modifications of the above defined Gibbs sampling method. Let $w_i = (x_{n_{i-1}+1}, \dots, x_{n_i})$, where $i = 1, \dots, m$ and $0 = n_0 < n_1 < \dots < n_{m-1} < n_m = n$. Let W_i be a subspace of R^n generated by vectors such of the form $(0_{n_{i-1}}, w_i, 0_{n-n_i})^T$, where $0^T_{n_i}$ denotes n_i-dimentional vector of zeros, $i = 1, \dots, m$.

Define the k-th state $w^{(k)} = (w_1^{(k)}, \dots, w_m^{(k)})$ of Markov's chain $\{w^{(k)}\}$ as follows:

$w_1^{(k)}$ is the realization from $\pi(w_1|w_j^{(k-1)}, j > 1)$,
$w_2^{(k)}$ is the realization from $\pi(w_2|w_1^{(k)}, w_j^{(k-1)}, j > 2)$,
$\vdots$
$w_{m-1}^{(k)}$ is the realization from $\pi(w_{m-1}|w_1^{(k)}, \dots, w_{m-2}^{(k)}, w_m^{(k-1)})$,
$w_m^{(k)}$ is the realization from $\pi(w_m|w_j^{(k)}, j < m)$.

If the chain $\{w^{(k)}\}$, is irreducible and aperiodic, then its limiting density is π. Now we give a necessary and sufficient condition for the joint density $\pi(x_1, \dots, x_n)$ to be unique by its conditional densities $\pi(w_i|w_j, i \neq j)$, $i = 1, \dots, m$.

Let R^*_A be a relation defined on the set $A \subset N$ as follows.

DEFINITION 2. A point $x \in A$ is in relation R^*_A to a point $y \in A$ iff there exists a finite sequence of vectors $c^{(k)}$, $k = 1, \dots, l$, which belong to one of spaces W_i, $i = 1, \dots, m$, such that

$z^{(1)} = x + c^{(1)} \in A$,
$z^{(k+1)} = z^{(k)} + c^{(k+1)} \in A$, where $k = 1, \dots, l-1$,
$y = z^{(l)} \in A$.

The following theorem can be proved similarly as Theorem 1.

THEOREM 2. *A joint density $\pi(x_1, \ldots, x_n)$ is unique by its conditional densities $\pi(w_i | w_j, i \neq j)$, $i = 1, \ldots, m$, iff for all sets $A \subset N$ with $\nu(A) = 0$*

(7) $\nu\{x \in N \setminus A : \nu\{y \in N \setminus A : x$ is not in relation $R^*_{N\setminus A}$ to $y\} > 0\} = 0$.

This modification of the Gibbs sampling method can be used when the components of the vectors w_i, $i = 1, \ldots, m$, are strongly correlated. It can also find applications when Markov's chain generated by the ordinary Gibbs sampling method is not ergodic.

4. Examples

We illustrate Theorem 1 in six examples. In examples 1, 2 and 4 a joint density is not determined uniquely by its conditional densities. In example 6 a joint density is determined uniquely by its conditional densities. In example 3 and 5 we construct a set N and consider two different measures on this set N. In the first case a joint density π with support N is not determined uniquely by its conditional densities, in the second case a joint density is determined uniquely.

EXAMPLE 1. First we analyse the example presented by Athreya, Doss, and Sethuraman [2]. They constructed two different densities π and π^* with respect to Lebesque measure, with identical conditional densities.

The joint density $\pi(x_1, x_2)$ is defined as follows

$$\pi(x_1, x_2) = \begin{cases} \frac{1}{2}p(x_1) & \text{for } x_2 = x_1 + 1 \\ \frac{1}{2}p(x_1) & \text{for } x_2 = x_1 - 1 \end{cases},$$

where $p(x) = \frac{1}{2}exp(-|x|)$.

The support N of the density $\pi(x_1, x_2)$ is the set consists of two parallel lines $x_2 = x_1 + 1$ and $x_2 = x_1 - 1$.

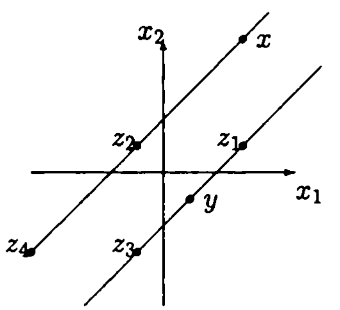

Figure 2: the point x is not in relation R_N to the point y

Conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$ are given by

$$\pi(x_1|x_2) = \begin{cases} \dfrac{p(x_2+1)}{p(x_2+1)+p(x_2-1)} & \text{for } x_1 = x_2+1 \\ \dfrac{p(x_2-1)}{p(x_2+1)+p(x_2-1)} & \text{for } x_1 = x_2-1 \end{cases},$$

$$\pi(x_2|x_1) = \begin{cases} \frac{1}{2} & \text{for } x_2 = x_1+1 \\ \frac{1}{2} & \text{for } x_2 = x_1-1 \end{cases}.$$

The density π^* with the support N is defined as follows

$$\pi^*(x_1,x_2) = \begin{cases} \dfrac{1}{2}\dfrac{p(x_1)f(x_1-[x_1])}{c_{x_1-[x_1]}} & \text{for } x_2 = x_1+1 \\ \dfrac{1}{2}\dfrac{p(x_1)f(x_1-[x_1])}{c_{x_1-[x_1]}} & \text{for } x_2 = x_1-1 \end{cases},$$

where $f(x)$ is a density function with the support $[0,1)$, $c_r = \sum_{-\infty<m<\infty} p(m+r)$, $r \in [0,1]$.

Hence $\pi^*(x_1|x_2) = \pi(x_1|x_2)$ and $\pi^*(x_2|x_1) = \pi(x_2|x_1)$ for almost all $(x_1,x_2) \in N$.

Let $x \in N$ be an arbitrary point in N. There exist only a countable number of points $z_i \in N$, such that xR_Nz_i, $i = 1,2,\ldots$. There are shown some points $z_1, z_2, z_3, z_4, y \in N$ in figure 2, such that xR_Nz_i, $i = 1,\ldots,4$, and the point x is not in relation R_N to the point y. It follows from Theorem 1 that the joint density $\pi(x_1,x_2)$ is not determined uniquely by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$.

EXAMPLE 2. Let N be the edge of the square, showed in figure 3. Let $4a$ be the length of edge of the square. As we see in figure 3, the point x (arbitrary chosen) is not in relation R_N to the point y. The only points related to the point x are the points z_1, z_2 and z_3. It follows from Theorem 1 that every joint density $\pi(x_1,x_2)$ defined on the support N, is not determined uniquely by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$.

We construct two joint densities defined on N with respect to Lebesque measure, having identical conditional densities. Let the density π be defined as follows $\pi(x_1,x_2) = \frac{1}{4}a$ for $\nu-$almost all $(x_1,x_2) \in N$. Its conditional densities are given by

$$\pi(x_1|x_2) = \frac{1}{2} \text{ for } \nu\text{-almost all } (x_1,x_2) \in N,$$

$$\pi(x_2|x_1) = \frac{1}{2} \text{ for } \nu\text{-almost all } (x_1,x_2) \in N.$$

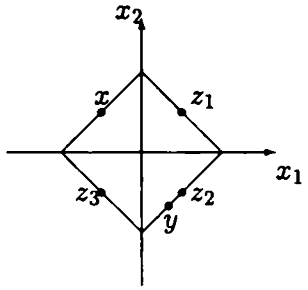

Figure 3: the point x is not in relation R_N to the point y

Let the density π^* be defined as follows

$$\pi^*(x_1,x_2) = \begin{cases} \frac{1}{2a^2}x_1 + \frac{1}{2a}, & \text{for } x_2 = x_1 + a, \quad x_1 \in [-a,0) \\ -\frac{1}{2a^2}x_1 + \frac{1}{2a}, & \text{for } x_2 = -x_1 + a, \quad x_1 \in [0,a) \\ -\frac{1}{2a^2}x_1 + \frac{1}{2a}, & \text{for } x_2 = x_1 - a, \quad x_1 \in [0,a) \\ \frac{1}{2a^2}x_1 + \frac{1}{2a}, & \text{for } x_2 = -x_1 - a, \quad x_1 \in [-a,0). \end{cases}$$

Hence $\pi(x_1|x_2) = \pi^*(x_1|x_2)$ and $\pi(x_2|x_1) = \pi^*(x_2|x_1)$ for ν−almost all $(x_1,x_2) \in N$.

EXAMPLE 3. Let us modify the support from the previous example. Let N be the edge of the square showed in figure 4. Let the set I consists of four points $A(0,0)$, $B(0,a)$, $C(a,a)$ and $D(a,0)$. Define Lebesque measure on the set N. Let us see that the set I has zero measure. The point x is not in relation $R_{N\setminus I}$ to the point y (fig. 4). Hence, every joint density $\pi(x_1,x_2)$ defined on the support N, is not determined uniquely by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$.

Now, suppose that there exists a distribution F satisfying

$$F = pF_s + qF_{ac}, \; 0 < p,q < 1$$

where F_s is a singular distribution with respect to Lebesque measure, F_{ac} is an absolutely continuous distribution with respect to Lebesque measure. If the distribution F_s is concentrated on the set I and the distribution F_{ac} is uniform on the set $N \setminus I$ then $xR_{N\setminus A}y$ for every set $A \subset N$, $\nu(A) = 0$ and all $x,y \in N \setminus A$. In this case the joint density $\pi(x)$ with distribution F is uniquely determined by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$.

As we can see the uniqueness of a joint density π does not depend as strongly on the shape of the support, as on its placement in the coordinate system.

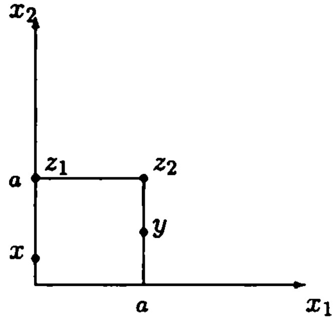

Figure 4: If $F = pF_s + qF_{ac}$, $0 < p, q < 1$ then $xR_{N\setminus A}y$ for every set $A \subset N$, $\nu(A) = 0$ and all $x, y \in N \setminus A$.

EXAMPLE 4. Let $N \subset \mathcal{R}^2$ be the set consists of two separate squares $K1$ and $K2$ (fig. 5)

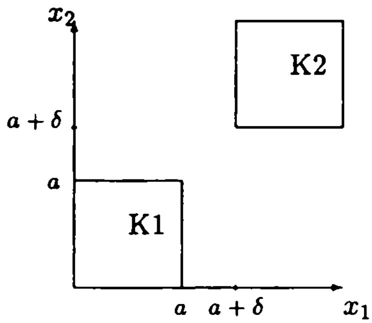

Figure 5: a point x is not in relation R_N to a point y for all $x \in K1$ and $y \in K2$

This way defined the set N disappoints the assumption of Theorem 1 because a point x is not in relation R_N to a point y for all points $x \in K1$ and $y \in K2$. That is why every joint density $\pi(x_1, x_2)$ defined on the support N is not determined uniquely by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$.

Define two different joint densities π and π^* with respect to Lebesque measure. Let the density π be defined as follows $\pi(x_1, x_2) = \frac{1}{2a^2}$ for ν−almost all $(x_1, x_2) \in N$. The conditional densities are given by

$$\pi(x_1|x_2) = \tfrac{1}{a} \text{ for } \nu\text{−almost all } (x_1, x_2) \in N,$$
$$\pi(x_2|x_1) = \tfrac{1}{a} \text{ for } \nu\text{−almost all } (x_1, x_2) \in N.$$

We define the joint density π^* as follows

$$\pi^*(x_1, x_2) = \tfrac{1}{4a^2} \text{ for } \nu\text{−almost all } (x_1, x_2) \in K1,$$
$$\pi^*(x_1, x_2) = \tfrac{3}{4a^2} \text{ for } \nu\text{−almost all } (x_1, x_2) \in K2.$$

Hence $\pi(x_1|x_2) = \pi^*(x_1|x_2)$ and $\pi(x_2|x_1) = \pi^*(x_2|x_1)$ for ν-almost all $x \in N$.

EXAMPLE 5. Now we analyse the example 2.1 [9]. Let N be the set consisting of two squares, $K1$ and $K2$ (fig. 6). Let $A(0,a)$ and $B(2a,a)$ be two points from the set N. Let $I = \overline{AB} \in N$ be the interval from the point A to the point B. Notice that if the set N is measurable with respect to two-dimensional Lebesque measure then the set I has zero measure and a point $x \in K_1$ is not in relation $R_{N\setminus I}$ to a point $y \in K_2$ (fig. 6). In this case, a joint density $\pi(x)$ with support N is not determined uniquelly by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$ (see example 2.1 [9]).

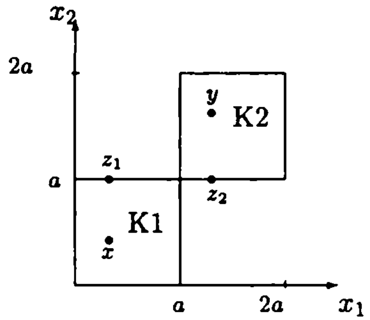

Figure 6: If $F = pF_s + qF_{ac},\ 0 < p, q < 1$ then $xR_{N\setminus A}y$ for every set $A \subset N$, $\nu(A) = 0$ and all $x, y \in N \setminus A$, but Gibbs chain is not ergodic.

Now, suppose that there exists a distribution F satisfying

$$F = pF_s + qF_{ac},\ 0 < p, q < 1$$

where F_s is a singular distribution with respect to Lebesque measure, F_{ac} is an absolutely continuous distribution with respect to Lebesque measure. If the distribution F_s is uniform on the support I and the distribution F_{ac} is uniform on the support $N \setminus I$ then $xR_{N\setminus A}y$ for every set $A \subset N$ with zero measure and all $x, y \in N \setminus A$. In this case the joint density $\pi(x)$ with distribution F is uniquely determined by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$. Suppose now that the chain in the k-th step stays at the point $x^{(k)} = (x_1^{(k)}, x_2^{(k)}) \in K1 \setminus I$. In the $k+1$-th step we generate value $x_2^{(k+1)}$ from the density $\pi(x_2|x_1^{(k)})$. The value $x_2^{(k+1)} = a$ is generated with zero probability. Hence transition probability $P(x^{(k)}, I)$ from the state $x^{(k)}$ to the set I is zero. We conclude that $P(x^{(k+1)} \in K2|x^{(k)}) = 0$ for ν-almost all $x^{(k)} \in K_1$ (see example 2.1 [9]). That is why Markov chain generated by Gibbs sampling method is not ergodic, in both cases. The example shows that uniqueness of a joint density is not a sufficient condition for the ergodicity of the chain.

EXAMPLE 6. Let N be the set consisting of two squares, $K1$ and $K2$ (fig. 7)

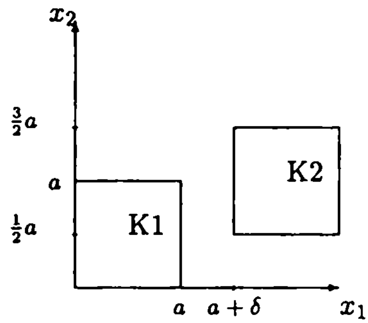

Figure 7: xR_Ny for all $x, y \in N$

We can see that xR_Ny for all $x, y \in N$. Hence every joint density π defined on the support N, is determined uniquely by its conditional densities $\pi(x_1|x_2)$ and $\pi(x_2|x_1)$. In this case, the Markov chain generated by Gibbs sampling algorithm is ergodic.

We have shown a few examples, which let us understand the idea of theorems and will help with such a planning of the experiment that Markov's chain generated by Gibbs sampling method, would be ergodic.

Acknowledgement. The author is grateful to prof. W. Klonecki for suggestion the problem. The author thank the Editor for many constructive comments that lead to substantial improvements.

References

[1] B. S. Arnold and J. Press, *Compatible conditional distributions*, J. Amer. Statist. 84 (1989), 152–156.

[2] K. B. Athreya, H. Doss and J. Sethuraman, *On the convergence of the Markov chain simulation method*, Ann. Statist. 24 (1989), 69–100.

[3] J. Besag, *Spatial interaction and the statistical analisys of lattice systems (with discussion)*, J. Royal Statist. Soc. Ser.B, 36 (1974), 192–236.

[4] K. S. Chan, *Asymptotic behaviour of the Gibbs sampler*, J. Amer. Statist. 88 (1993), 320–326.

[5] M. Evans and T. Swartz, *Methods for aproximating integrals in statistics with special emphasis on bayesian integration problems*, Statist. Sci. 10 (1995), 254–272.

[6] A. E. Gelfand and F. M. Smith, *Sampling-based approaches to calculating marginal densities*, J. Amer. Satist. Ser.B, 85 (1990), 398–409.

[7] A. Gelman and T. P. Speed, *Characterizing a joint probability distribution by conditionals*, J. Royal Statist. Soc. Ser. B, 55 (1993), 185–188.

[8] J. Hobert and G. Casella, *The effect of improper priors on Gibbs sampling in hierarchical linear mixed models*, J. Amer. Statist. 9 (1996), 1461–1473.

[9] J. P. Hobert, C. P. Robert and C. Goutis, *Connectedness conditions for the convergence of the Gibbs sampler*, Statist. Probab. Lett. 33 (1997), 235–240.
[10] C. P. Robert, *Convergence control methods for Markov chain Monte Carlo algorithms*, Statist. Sci. 10 (1995), 231–253.
[11] L. Tierney, *Markov chain for exploring posterior distributions*, Ann. Statist. 22 (1994), 1701–1762.

INSTIUTE OF MATHEMATICS AND COMPUTER SCIENCE
TECHNICAL UNIVERSITY OF CZĘSTOCHOWA
ul. Dąbrowskiego 73
42-200 CZĘSTOCHOWA, POLAND
e-mail pkop@matinf.pcz.czest.pl

Received September 15, 1997; revised version December 30, 1998.