MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Markov Bases for Typical Block Effect Models of Two-way Contingency Tables

Mitsunori OGAWA and Akimichi TAKEMURA

METR 2011–40 December 2011

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t-u-tokyo.ac.jp/research/techrep/index.html

Markov Bases for Typical Block Effect Models of Two-way Contingency Tables

Mitsunori Ogawa* and Akimichi Takemura*†

December 2011

Abstract

Markov basis for statistical model of contingency tables gives a useful tool for performing the conditional test of the model via Markov chain Monte Carlo method. In this paper we derive explicit forms of Markov bases for change point models and block diagonal effect models, which are typical block-wise effect models of two-way contingency tables, and perform conditional tests with some real data sets.

Keywords and phrases: block diagonal effect model, change point model, Markov chain Monte Carlo, quadratic Gröbner basis, toric ideal.

1 Introduction

Goodness-of-fit tests for statistical models of contingency tables are usually performed by the large sample approximation to the null distribution of test statistics. However, as shown in Haberman [6], the large sample approximation may not be appropriate when the expected frequencies are not large enough. In such cases it is desirable to use a conditional testing procedure. In this paper we discuss a conditional testing via Markov chain Monte Carlo (MCMC) method with Markov bases.

Diaconis and Sturmfels [3] showed the equivalence of a Markov basis and a binomial generator for the toric ideal arising from a statistical model of discrete exponential families and developed an algebraic sampling method for conditional distributions. Thanks to their algorithm, once we have a Markov basis for a given statistical model, we can perform a conditional test for the model via MCMC method. However, the structure of Markov bases are complicated in general. Many researchers have studied the structures of Markov bases in algebraic statistics (e.g. Dobra and Sullivant [4], Aoki and Takemura [1], Rapallo [13], Hara et al. [7]).

In this paper we derive Markov bases for some statistical models of two-way contingency tables considering the effect of subtables. It is a well-known fact that the set of square-free moves of degree two (basic moves) forms the minimal Markov basis for complete independence model of two-way contingency tables. On the other hand, when a subtable effect is added to the model, the set of basic moves does not necessarily form a Markov basis. This problem is called

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo.

†JST CREST

two-way subtable sum problem (Hara et al. [8]). In the previous researches statistical models with one subtable effect are considered. We consider some statistical models including several subtable effects.

The organization of this paper is as follows. In Section 2 we summarize the notations and definitions on Markov bases and introduce two-way subtable sum problems. In Section 3 we derive the minimal Markov basis for the configuration arising from two-way change point model and discuss the algebraic properties of the toric ideal arising from the change point model. Section 4 and 5 give the explicit forms of Markov bases for the configurations arising from some block diagonal effect models. In Section 6 we apply the MCMC method with our Markov bases to some data sets and confirm that it works well in practice. We conclude the paper with some remarks in Section 7.

2 Two-way subtable sum problem

In this section we summarize notations and definitions on Markov bases and give a brief review of two-way subtable sum problems.

2.1 Preliminaries

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ and let $\boldsymbol{x} = \{x_{ij}\}, x_{ij} \in \mathbb{N}, i = 1, \ldots, R, j = 1, \ldots, C$ be an $R \times C$ two-way contingency table with nonnegative integer entries. Let $\mathcal{I} = \{(i,j) \mid 1 \leq i \leq R, 1 \leq j \leq C\}$ be the set of cells. We order the elements of contingency table $\boldsymbol{x}$ lexicographically and regard $\boldsymbol{x}$ as a column vector.

Let A be a $T \times |\mathcal{I}|$ zero-one matrix where T is a positive integer and $|\mathcal{I}| = RC$. We assume that the subspace of $\mathbb{R}^{|\mathcal{I}|}$ spanned by the rows of A contains the $|\mathcal{I}|$-dimensional row vector $(1, \ldots, 1)$. For a given $\boldsymbol{t} \in \mathbb{N}^T$ the set of contingency tables

$$\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x} \in \mathbb{N}^{|\mathcal{I}|} \mid A\boldsymbol{x} = \boldsymbol{t}\}$$

is called the $\boldsymbol{t}$-fiber. An integer table $\boldsymbol{z}$ with $A\boldsymbol{z} = 0$ is called a move for A. Define the degree of move $\boldsymbol{z}$ as $\|\boldsymbol{z}\|_1/2 = \sum_{(i,j)\in\mathcal{I}} |z_{ij}|/2$. Let $\mathcal{M}_A = \{\boldsymbol{z} \mid A\boldsymbol{z} = 0\}$ denote the set of moves for A. A subset $\mathcal{B} \subseteq \mathcal{M}_A$ is called sign-invariant if $\boldsymbol{z} \in \mathcal{B}$ implies $-\boldsymbol{z} \in \mathcal{B}$. Markov basis for A is defined as follows.

Definition 1. *A sign-invariant finite set of moves $\mathcal{B} \subseteq \mathcal{M}_A$ is a Markov basis for A, if for any $\boldsymbol{t}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}} (\boldsymbol{x} \neq \boldsymbol{y})$ there exist $U > 0, \boldsymbol{z}_{v_1}, \ldots, \boldsymbol{z}_{v_U} \in \mathcal{B}$ such that*

$$\boldsymbol{y} = \boldsymbol{x} + \sum_{s=1}^{U} \boldsymbol{z}_{v_s} \quad \textit{and} \quad \boldsymbol{x} + \sum_{s=1}^{u} \boldsymbol{z}_{v_s} \in \mathcal{F}_{\boldsymbol{t}} \textit{ for } 1 \leq u \leq U.$$

In this paper we only consider sign-invariant sets of moves as Markov bases. Since the row vector $(1, \ldots, 1)$ is contained in the subspace of $\mathbb{R}^{|\mathcal{I}|}$ spanned by the rows of A, $\sum_{(i,j)\in\mathcal{I}, z_{ij}>0} z_{ij} = -\sum_{(i,j)\in\mathcal{I}, z_{ij}<0} z_{ij}$ holds for every move $\boldsymbol{z} \in \mathcal{M}_A$. A move $\boldsymbol{z}$ can be written as $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ where $\boldsymbol{z}^+ = \{\max(z_{ij}, 0)\}$ and $\boldsymbol{z}^- = \{\max(-z_{ij}, 0)\}$. If there exists a fiber $\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{z}^+, \boldsymbol{z}^-\}$, we say that $\boldsymbol{z}$ is an indispensable move.

Suppose that $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the same fiber $\mathcal{F}_{\boldsymbol{t}}$ and the l_1-norm $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = \sum_{(i,j)\in\mathcal{I}} |x_{ij} - y_{ij}|$ is not equal to zero. We say that $\|\boldsymbol{x} - \boldsymbol{y}\|_1$ can be reduced by a subset $\mathcal{B} \subseteq \mathcal{M}_A$ if there exist

$\tau^+ \geq 0, \tau^- \geq 0$, $\tau^+ + \tau^- > 0$, and sequences of moves $B_s^+ \in \mathcal{B}$, $s = 1, \ldots, \tau^+$, and $B_s^- \in \mathcal{B}$, $s = 1, \ldots, \tau^-$, satisfying

$$\left\| \boldsymbol{x} - \boldsymbol{y} + \sum_{s=1}^{\tau^+} B_s^+ + \sum_{s=1}^{\tau^-} B_s^- \right\|_1 < \|\boldsymbol{x} - \boldsymbol{y}\|_1,$$

$$\boldsymbol{x} + \sum_{s=1}^{\tau'} B_s^+ \in \mathcal{F}_{\boldsymbol{t}}, \qquad \tau' = 1, \ldots, \tau^+,$$

$$\boldsymbol{y} - \sum_{s=1}^{\tau'} B_s^- \in \mathcal{F}_{\boldsymbol{t}}, \qquad \tau' = 1, \ldots, \tau^-.$$

It is easy to see that a subset $\mathcal{B} \subseteq \mathcal{M}_A$ is a Markov basis for $\mathcal{M}_A$ if $\|\boldsymbol{x} - \boldsymbol{y}\|_1$ can be reduced by $\mathcal{B}$ for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in every fiber $\mathcal{F}_{\boldsymbol{t}}$.

2.2 Two-way subtable sum problem

In this subsection we introduce two-way subtable sum problems and give a brief review of previous researches.

For a contingency table $\boldsymbol{x}$ denote the row sums and column sums of $\boldsymbol{x}$ by

$$x_{i+} = \sum_{j=1}^{C} x_{ij}, \ i = 1, \ldots, R, \quad x_{+j} = \sum_{i=1}^{R} x_{ij}, \ j = 1, \ldots, C.$$

Let $S_1, \ldots, S_N$ be subsets of $\mathcal{I}$ and define the subtable sums $x_{S_n}, n = 1, \ldots, N$, by

$$x_{S_n} = \sum_{(i,j) \in S_n} x_{ij}.$$

We summarize the set of row sums, column sums and the subtable sums as a column vector

$$\boldsymbol{t} = (x_{1+}, \ldots, x_{R+}, x_{+1}, \ldots, x_{+C}, x_{S_1}, \ldots, x_{S_N})'.$$

Then, with an appropriate zero-one matrix A_S, the relation between $\boldsymbol{x}$ and $\boldsymbol{t}$ is written by

$$A_{S_1,\ldots,S_N} \boldsymbol{x} = \boldsymbol{t}.$$

The set of columns of $A_{S_1,\ldots,S_N}$ is a configuration defining a toric ideal $I_{A_{S_1,\ldots,S_N}}$. For simplicity we call $A_{S_1,\ldots,S_N}$ the configuration for $S_1, \ldots, S_N$ and abbreviate the set of moves $\mathcal{M}_{A_{S_1,\ldots,S_N}}$ for $A_{S_1,\ldots,S_N}$ as $\mathcal{M}_{S_1,\ldots,S_N}$.

Consider a square-free move of degree 2 (basic move) of the form

$$\begin{array}{ccc} & j & j' \\ i & +1 & -1 \\ i' & -1 & +1 \end{array}$$

for $i \neq i'$ and $j \neq j'$. For simplicity we denote this move by $(i,j)(i',j') - (i',j)(i,j')$. Similarly a move of degree d is denoted by $(i_1, j_1) \cdots (i_d, j_d) - (i_2, j_1) \cdots (i_d, j_{d-1})(i_1, j_d)$ for appropriate

$i_1, \ldots, i_d$ and $j_1, \ldots, j_d$. Hara et al. [8] and [9] discussed Markov bases for the configuration A_S for one subtable $S \subseteq \mathcal{I}$. In [8] it is shown that the set of basic moves in $\mathcal{M}_S$ is a Markov basis for A_S if and only if S is either 2×2 block diagonal or triangular. Ohsugi and Hibi [12] discussed the same problem from algebraic viewpoint.

In Sections 3–5 we use the notations on signs of subtables: Let $\boldsymbol{x}, \boldsymbol{y}$ be the two contingency tables in the same fiber $\mathcal{F}_{\boldsymbol{t}}$ and let $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$. Consider a subset $\hat{\mathcal{I}} \subseteq \mathcal{I}$. If $z_{ij} = 0$ for every $(i, j) \in \hat{\mathcal{I}}$, we denote it by $\boldsymbol{z}(\hat{\mathcal{I}}) = 0$. If $\boldsymbol{z}(\hat{\mathcal{I}}) \neq 0$ and $z_{ij} > 0$ for $\exists(i, j) \in \hat{\mathcal{I}}$, we denote it by $\boldsymbol{z}(\hat{\mathcal{I}}) > 0$. Other notations such as $\boldsymbol{z}(\hat{\mathcal{I}}) \geq 0, \boldsymbol{z}(\hat{\mathcal{I}}) \leq 0$ and $\boldsymbol{z}(\hat{\mathcal{I}}) < 0$ are defined in the same way.

3 Markov bases for the change point models

In this section we derive Markov bases for the configuration arising from the change point model of two-way contingency tables and discuss its algebraic properties.

3.1 The unique minimal Markov bases for the change point models

In this subsection we derive the unique minimal Markov bases for the change point models. Two-way change point models with an unknown change point is introduced in Hirotsu [10]. We consider the two-way change point models with several fixed change points. We call a subtable $S \subseteq \mathcal{I}$ a rectangle in $\mathcal{I}$, if S has a form

$$S = \{(i, j) \mid a_1 \leq i \leq a_2, b_1 \leq j \leq b_2\}$$

for $1 \leq a_1 < a_2 \leq R$ and $1 \leq b_1 < b_2 \leq C$. Let $S_n, n = 1, \ldots, N$, be rectangles of $\mathcal{I}$ satisfying $S_1 \subset S_2 \subset \cdots \subset S_N \subset \mathcal{I}$. Then the change point model is defined by

$$\log p_{ij} = \mu + \alpha_i + \beta_j + \sum_{n=1}^{N} \gamma_n I_{S_n}(i, j), \tag{1}$$

where $I_{S_n}(i, j) = 1$ if $(i, j) \in S_n$ and $I_{S_n}(i, j) = 0$ otherwise. The sufficient statistic for this model consists of the row sums, the column sums and the sums of frequencies in $S_n, n = 1, \ldots, N$.

The first main result of this paper is stated as follows.

Theorem 1. *Let $S_n, n = 1, \ldots, N$, be the rectangles of $\mathcal{I}$ with $S_1 \subset S_2 \subset \cdots \subset S_N \subset \mathcal{I}$. The set of basic moves in $\mathcal{M}_{S_1,\ldots,S_N}$ is the unique minimal Markov basis for the configuration $A_{S_1,\ldots,S_N}$.*

Proof. After an appropriate interchange of rows and columns, we may assume that $S_n, n = 1, \ldots, N$, are the subsets of $\mathcal{I}$ defined by

$$S_n = \{(i, j) \mid 1 \leq i \leq r_n, 1 \leq j \leq c_n\}, \qquad n = 1, \ldots, N,$$

where each $r_n, c_n, n = 1, \ldots, N$, is a positive integer with $1 < r_1 \leq \cdots \leq r_N \leq R$ and $1 < c_1 \leq \cdots \leq c_N \leq C$. Since $S_1 \subset S_2 \subset \cdots \subset S_N \subset \mathcal{I}$, at least one of $r_n < r_{n+1}$ or $c_n < c_{n+1}$ holds for each $n = 1, \ldots, N-1$ and at least one of $r_N < R$ or $r_N < C$ holds.

Suppose $\boldsymbol{z}(S_1) \neq 0$. Then $\boldsymbol{z}$ contains both of positive cell and negative cell in S_1. Denote these cells by (i, j) and (i', j'). If $j = j'$, letting (i, j'') be a negative cell in the i-th row, $\|\boldsymbol{z}\|_1$ can be reduced by a basic move $(i, j)(i', j'') - (i', j)(i, j'') \in \mathcal{M}_{S_1,\ldots,S_N}$. Let us consider the case of $i \neq i'$ and $j \neq j'$. Let (i, j'') and (i'', j) be negative cells in the i-th row and in the j-th column, respectively. Then $\|\boldsymbol{z}\|_1$ can be reduced by a sequence of two basic moves $(i, j')(i', j'') - (i', j')(i, j'')$ and $(i, j)(i'', j') - (i'', j)(i, j')$. For the case of $r_N < R$ and $c_N < C$, it can be shown by the same argument that if $\boldsymbol{z}$ contains both of positive cell and negative cell in $\hat{\mathcal{I}} := \{(i, j) \mid r_N < i \leq R, c_N < j \leq C\}$, then $\|\boldsymbol{z}\|_1$ can be reduced by the set of basic moves in $\mathcal{M}_{S_1,\dots,S_N}$. If $r_N = R$ or $c_N = C$ holds, say $r_N = R$, $\boldsymbol{z}$ contains both of positive cell and negative cell in $\mathcal{I} \setminus S_N$. Then $\|\boldsymbol{z}\|_1$ can be reduced by the set of basic moves in $\mathcal{M}_{S_1,\dots,S_N}$.

Consider the case of $\boldsymbol{z}(S_1) = 0$. We claim that if $\boldsymbol{z}(S_{n-1}) = 0$ and $\boldsymbol{z}(S_n) \neq 0$ for $1 < \exists n \leq N$, then $\|\boldsymbol{z}\|_1$ can be reduced by the set of basic moves in $\mathcal{M}_{S_1,\dots,S_N}$. If either $r_{n-1} = r_n$ or $c_{n-1} = c_n$ holds, we see that $\|\boldsymbol{z}\|_1$ can be reduced by a basic move in $\mathcal{M}_{S_1,\dots,S_N}$. For the case of $r_{n-1} < r_n$ and $c_{n-1} < c_n$, let $S_n^{12} = \{(i, j) \mid 1 \leq i \leq r_{n-1}, c_{n-1} < j \leq c_n\}$, $S_n^{21} = \{(i, j) \mid r_{n-1} < i \leq r_n, 1 \leq j \leq c_{n-1}\}$ and $S_n^{22} = \{(i, j) \mid r_{n-1} < i \leq r_n, c_{n-1} < j \leq c_n\}$. If $\boldsymbol{z}$ contains both of positive cell and negative cell in one of S_n^{12}, S_n^{21} or S_n^{22}, it can be similarly shown that $\|\boldsymbol{z}\|_1$ can be reduced by the set of basic moves in $\mathcal{M}_{S_1,\dots,S_N}$. Then we only need to consider the case of $\boldsymbol{z}(S_n^{kl}) \geq 0$ or $\boldsymbol{z}(S_n^{kl}) \leq 0$ for each $(k, l) \in \{(1, 2), (2, 1), (2, 2)\}$. Without loss of generality we can assume $S_n^{12} > 0$. Let (i, j) be a positive cell in S_n^{12} and let (i', j) be a negative cell in the j-th column. If $(i', j) \in S_n^{22}$, using a negative cell (i, j') in the i-th row, $\|\boldsymbol{z}\|_1$ can be reduced by a basic move $(i, j)(i', j') - (i', j)(i, j') \in \mathcal{M}_{S_1,\dots,S_N}$. Suppose $(i', j) \notin S_n^{22}$. There exists a negative cell $(i'', j'') \in S_n^{21} \cup S_n^{22}$. Then $\|\boldsymbol{z}\|_1$ can be reduced by a sequence of two basic moves $(i', j)(i'', j'') - (i', j'')(i'', j)$ and $(i, j)(i', j') - (i', j)(i, j')$. Therefore the claim is proved.

The remaining part is the case that $\boldsymbol{z}(S_N) = 0$ and one of $\boldsymbol{z}(\hat{\mathcal{I}}) \geq 0$ or $\boldsymbol{z}(\hat{\mathcal{I}}) \leq 0$ holds. If $\boldsymbol{z}(\hat{\mathcal{I}}) = 0$, $\boldsymbol{z}$ contains a nonzero cell in $\{(i, j) \mid r_N < i \leq R, 1 \leq j \leq c_N\}$ or $\{(i, j) \mid 1 \leq i \leq r_N, c_N < j \leq C\}$. It is easy to see that $\|\boldsymbol{z}\|_1$ can be reduced by a basic move in $\mathcal{M}_{S_1,\dots,S_N}$. Suppose $\boldsymbol{z}(\hat{\mathcal{I}}) > 0$ and let (i, j) be a positive cell in $\hat{\mathcal{I}}$. There exist a negative cell (i, j') and a positive cell (i', j') with $i' \neq i$ in $\{(i, j) \mid r_N < i \leq R, 1 \leq j \leq c_N\}$. Then $\|\boldsymbol{z}\|_1$ can be reduced by a basic move $(i, j)(i', j') - (i', j)(i, j')$.

Since every basic move in $\mathcal{M}_{S_1,\dots,S_N}$ is indispensable, the set of basic moves in $\mathcal{M}_{S_1,\dots,S_N}$ is the unique minimal Markov basis for $A_{S_1,\dots,S_N}$ (see [15]). □

3.2 Algebraic properties of the configuration arising from the change point model

In this subsection we investigate the algebraic properties of the configuration $A_{S_1,\dots,S_N}$ arising from the change point model.

Let K be a field and let $K[\{u_i\}_{1 \leq i \leq R} \cup \{v_j\}_{1 \leq j \leq C} \cup \{w_n\}_{1 \leq n \leq N+1}]$ be a polynomial ring in $R + C + N$ variables over K. We associate each cell $(i, j) \in S_n \setminus S_{n-1}, 1 \leq n \leq N + 1$, to a monomial $u_i v_j w_n$ where $S_0 = \emptyset$ and $S_{N+1} = \mathcal{I}$. Define $R_{S_1,\dots,S_N}$ as a semigourp ring generated by those monomials. Let $K[\boldsymbol{x}] = K[\{x_{ij}\}_{(i,j) \in \mathcal{I}}]$ be a polynomial ring in RC variables over K. Define a surjective map $\pi : K[\boldsymbol{x}] \to R_{S_1,\dots,S_N}$ by $\pi(x_{ij}) = u_i v_j w_n$ for $1 \leq n \leq N + 1$. Define the toric ideal for the change point model as the kernel of π and denote it by $I_{S_1,\dots,S_N}$. See [14] and [2] for general facts on toric ideals and their Gröbner bases. From Theorem 1 we already know that the toric ideal $I_{S_1,\dots,S_N}$ is generated by the quadratic binomials corresponding to basic moves in $\mathcal{M}_{S_1,\dots,S_N}$. Furthermore we have the following Theorem 2. Although its proof is similar to [12], we need to use a lexicographic order different from the order used in [12]. In fact the toric ideal $I_{S_1,\dots,S_N}$ does not have a quadratic Gröbner basis with respect to the lexicographic order used in [12] if $N \geq 2$ and there exist m, n such that $2 \leq m < n \leq N + 1$ and

$r_{m-1} < r_m, c_{m-1} < c_m, r_{n-1} < r_n, c_{n-1} < c_n$ where $r_{N+1} = R$ and $c_{N+1} = C$.

Theorem 2. *For the toric ideal $I_{S_1,\dots,S_N}$ the following statements hold:*

(i) $I_{S_1,\dots,S_N}$ *possesses a quadratic Gröbner basis;*

(ii) $I_{S_1,\dots,S_N}$ *possesses a square-free initial ideal;*

(iii) $R_{S_1,\dots,S_N}$ *is normal;*

(iv) $R_{S_1,\dots,S_N}$ *is Koszul.*

Proof. Generally, (i) $\Rightarrow$ (iv) and (ii) $\Rightarrow$ (iii) hold. Since $R_{S_1,\dots,S_N}$ is generated by the monomials of the same degree, (i) $\Rightarrow$ (ii) holds from the proof of Proposition 1.6 in Ohsugi and Hibi [11]. Therefore it suffices to show that the statement (i) holds.

By an appropriate interchange of rows and columns, we may assume that $S_n, n = 1, \dots, N$, share their upper-left corners. From Theorem 1 the toric ideal $I_{S_1,\dots,S_N}$ is generated by

$$\mathcal{G} = \{x_{ik}x_{jl} - x_{il}x_{jk} \mid 1 \le i < j \le R, 1 \le k < l \le C, \pi(x_{ik}x_{jl}) = \pi(x_{il}x_{jk})\}.$$

Fix a lexicographic order $\succ$ satisfying $x_{RC} \succ x_{RC-1} \succ \cdots \succ x_{R1} \succ x_{R-1C} \succ \cdots \succ x_{11}$. Then $x_{ik}x_{jl}$ is the initial monomial of $x_{ik}x_{jl} - x_{il}x_{jk}, 1 \le i < j \le R, 1 \le k < l \le C$. We prove that $\mathcal{G}$ is a Gröbner basis of $I_{S_1,\dots,S_N}$ with respect to $\succ$ using Buchberger's criterion.

Let f be the S-polynomial of $g_1, g_2 \in \mathcal{G}$. Suppose that f is not reduced to zero by $\mathcal{G}$. By Proposition 4 in Section 9 of Chapter 2 in [2], the initial monomials of g_1 and g_2 are not relatively prime. On the other hand, if the monomials of f share a common variable f is reduced to zero by $\mathcal{G}$. Then f is a cubic binomial and is represented as $f = x_{i_1l_1}x_{i_2l_2}x_{i_3l_3} - x_{i'_1l'_1}x_{i'_2l'_2}x_{i'_3l'_3}$ with the initial monomial $x_{i_1l_1}x_{i_2l_2}x_{i_3l_3}$. Since $f \in I_{S_1,\dots,S_N}$, we have $\{i_1, i_2, i_3\} = \{i'_1, i'_2, i'_3\}$ and $\{l_1, l_2, l_3\} = \{l'_1, l'_2, l'_3\}$. Since the monomials of f have no common variable, $|\{i_1, i_2, i_3\}| = |\{l_1, l_2, l_3\}| = 3$. We assume $1 \le i_1 = i'_1 < i_2 = i'_2 < i_3 = i'_3 \le R$ without loss of generality. By the definition of $\succ$, $l_3 > l'_3 \in \{l_1, l_2\}$. Then f is represented as one of the following forms:

(1) $x_{i_1j_1}x_{i_2j_2}x_{i_3j_3} - x_{i_1j_2}x_{i_2j_3}x_{i_3j_1}$,

(2) $x_{i_1j_1}x_{i_2j_2}x_{i_3j_3} - x_{i_1j_3}x_{i_2j_1}x_{i_3j_2}$,

(3) $x_{i_1j_1}x_{i_2j_3}x_{i_3j_2} - x_{i_1j_3}x_{i_2j_2}x_{i_3j_1}$,

(4) $x_{i_1j_2}x_{i_2j_1}x_{i_3j_3} - x_{i_1j_1}x_{i_2j_3}x_{i_3j_2}$,

(5) $x_{i_1j_2}x_{i_2j_1}x_{i_3j_3} - x_{i_1j_3}x_{i_2j_2}x_{i_3j_1}$,

(6) $x_{i_1j_3}x_{i_2j_1}x_{i_3j_2} - x_{i_1j_2}x_{i_2j_3}x_{i_3j_1}$,

where $1 \le i_1 < i_2 < i_3 \le R$ and $1 \le j_1 < j_2 < j_3 \le C$. The candidates (1)–(6) of the form of f are obtained as follows: Suppose $l_1 < l_2 < l_3$ and $l'_3 = l_1$. By $l_2 \ne l'_2$ we have $(l'_1, l'_2) = (l_2, l_3)$. This implies that f corresponds to the type (1). The forms (2)–(5) are obtained by the same argument.

For each form of (1)–(6), if there exists a quadratic binomial $\hat{f}$ such that $in_\prec(\hat{f})$ divides $in_\prec(f)$, then f can be reduced to a cubic binomial whose two monomials share a common variable. Hence such $\hat{f}$ does not belong to $\mathcal{G}$. We derive a contradiction for each form of f. Note that $x_{ik}x_{jl} - x_{il}x_{jk} \notin \mathcal{G}$, $1 \le i < j \le R, 1 \le k < l \le C$, is equivalent to the existence of $n, 1 \le n \le N$, such that $(i,k) \in S_n \setminus S_{n-1}$ and $(i,l), (j,k), (j,l) \notin S_n$. We refer to this equivalence by $(*)$.

(1) Consider a quadratic binomial $\hat{f} = x_{i_1j_1}x_{i_2j_2} - x_{i_1j_2}x_{i_2j_1}$ and let $S_n \setminus S_{n-1}$ be a subtable containing (i_1, j_1). Since $\hat{f} \notin \mathcal{G}$ and $(*)$, $(i_1, j_2), (i_2, j_3), (i_3, j_1) \notin S_n$. This contradicts $f \in I_{S_1,\dots,S_N}$.

(2) Consider a quadratic binomial $\hat{f} = x_{i_1j_1}x_{i_2j_2} - x_{i_1j_2}x_{i_2j_1}$ and let $S_n \setminus S_{n-1}$ be a subtable containing (i_1, j_1). Since $\hat{f} \notin \mathcal{G}$ and $(*)$, $(i_1, j_3), (i_2, j_1), (i_3, j_2) \notin S_n$. This contradicts $f \in I_{S_1,\dots,S_N}$.

(3) Consider two quadratic binomials $\hat{f}_1 = x_{i_1j_1}x_{i_2j_3} - x_{i_1j_3}x_{i_2j_1}$ and $\hat{f}_2 = x_{i_1j_1}x_{i_3j_2} - x_{i_1j_2}x_{i_3j_1}$. Let $S_n \setminus S_{n-1}$ be a subtable containing (i_1, j_1). Since $\hat{f}_1 \notin \mathcal{G}$ and $(*)$, $(i_2, j_1) \notin S_n$. Since $\hat{f}_2 \notin \mathcal{G}$ and $(*)$, $(i_1, j_2) \notin S_n$. Then $(i_1, j_3), (i_2, j_2), (i_3, j_1) \notin S_n$, which contradicts $f \in I_{S_1,\dots,S_N}$.

(4) Consider two quadratic binomials $\hat{f}_1 = x_{i_2j_1}x_{i_3j_3} - x_{i_2j_3}x_{i_3j_1}$ and $\hat{f}_2 = x_{i_1j_2}x_{i_3j_3} - x_{i_1j_3}x_{i_3j_2}$. Let $S_n \setminus S_{n-1}$ be a subtable containing (i_1, j_1). Since $\hat{f}_1 \notin \mathcal{G}$, $(*)$ and $f \in I_{S_1,\dots,S_N}$, $(i_2, j_1) \in S_n \setminus S_{n-1}$. Similarly $(i_1, j_2) \in S_n \setminus S_{n-1}$ follows from $\hat{f}_2 \notin \mathcal{G}$, $(*)$ and $f \in I_{S_1,\dots,S_N}$. These contradict $f \in I_{S_1,\dots,S_N}$.

(5) Consider two quadratic binomials $\hat{f}_1 = x_{i_2j_1}x_{i_3j_3} - x_{i_2j_3}x_{i_3j_1}$ and $\hat{f}_2 = x_{i_1j_2}x_{i_3j_3} - x_{i_1j_3}x_{i_3j_2}$. Let $S_n \setminus S_{n-1}$ be a subtable containing (i_2, j_1). Since $\hat{f}_1 \notin \mathcal{G}$ and $(*)$, $(i_2, j_3), (i_3, j_1) \notin S_n$ and $(i_1, j_3) \notin S_n$. Since $f \in \mathcal{G}$, $S_n \setminus S_{n-1}$ contains (i_2, j_2). Similarly $(i_1, j_2) \in S_n \setminus S_{n-1}$ follows from $\hat{f}_2 \notin \mathcal{G}$, $(*)$ and $f \in I_{S_1,\dots,S_N}$. This contradicts $f \in I_{S_1,\dots,S_N}$.

(6) Consider a quadratic binomial $\hat{f} = x_{i_2j_1}x_{i_3j_2} - x_{i_2j_2}x_{i_3j_1}$ and let $S_n \setminus S_{n-1}$ be a subtable containing (i_2, j_1). Since $\hat{f} \notin \mathcal{G}$ and $(*)$, $(i_1, j_2), (i_2, j_3), (i_3, j_1) \notin S_n$. This contradicts $f \in I_{S_1,\dots,S_N}$.

Therefore $\mathcal{G}$ is a Gröbner basis of $I_{S_1,\dots,S_N}$ with respect to $\succ$. $\square$

4 Markov bases for common block diagonal effect models

In this section we introduce the common diagonal effect model of two-way contingency tables and derive its Markov basis.

Let S denote the set of cells belonging to the diagonal blocks defined by

$$S = \{(i, j) \mid r_n \le i < r_{n+1}, c_n \le j < c_{n+1}, 1 \le \exists n \le N\},$$

where each $r_n, c_n, n = 1, \dots, N + 1$, is a non-negative integer with $1 = r_1 < r_2 < \cdots < r_{N+1} = R+1$ and $1 = c_1 < c_2 < \cdots < c_{N+1} = C+1$. S is an $N \times N$ block diagonal set in the contingency table. In the common block diagonal effect model, the cell probabilities $\{p_{ij}\}$ are defined by

$$\log p_{ij} = \mu + \alpha_i + \beta_j + \gamma_S I_S(i, j). \tag{2}$$

In the model (2), all cells in diagonal blocks have the same parameter γ_S. The sufficient statistic for (2) consists of the row sums, the column sums and the sum of frequencies in S. Note that the model (2) is a generalization of the common diagonal effect model whose Markov basis is discussed in Hara et al. [9].

Since Hara et al. [8] showed that for $N = 2$ the set of basic moves in $\mathcal{M}_S$ is a Markov basis for A_S, we assume $N \ge 3$ in this section. In order to describe a Markov basis for A_S we need

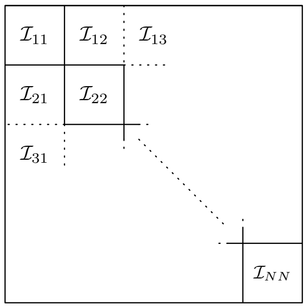


Figure 1: Block-wise indices.

some more notations. We index each block as in Figure 1, i.e., $\mathcal{I}_{kl} = \{(i,j) \mid r_k \leq i < r_{k+1}, c_l \leq j < c_{l+1}\}$ for $1 \leq k, l \leq N$. Note that S can be represented as $S = \mathcal{I}_{11} \cup \mathcal{I}_{22} \cup \cdots \cup \mathcal{I}_{NN}$.

Consider the following types of moves.

- Type I (square-free moves of degree 2):

$$\begin{array}{ccc} & j_1 & j_2 \\ i_1 & +1 & -1 \\ i_2 & -1 & +1 \end{array}$$

  where $i_1 \neq i_2$ and $j_1 \neq j_2$.

- Type II (square-free moves of degree 3):

$$\begin{array}{cccc} & j_1 & j_2 & j_3 \\ i_1 & 0 & +1 & -1 \\ i_2 & -1 & 0 & +1 \\ i_3 & +1 & -1 & 0 \end{array}$$

  where nonzero cells (i.e. ± 1) belong to distinct blocks in S^C.

- Type III (square-free moves of degree 3):

$$\begin{array}{cccc} & j_1 & j_2 & j_3 \\ i_1 & +1 & 0 & -1 \\ i_2 & 0 & -1 & +1 \\ i_3 & -1 & +1 & 0 \end{array}$$

  where (i_1, j_1) and (i_2, j_2) belong to distinct blocks in S and other nonzero cells belong to distinct blocks in S^C.

- Type IV (moves of degree 4):

$$\begin{array}{ccccc} & j_1 & j_2 & j_3 & j_4 \\ i_1 & +1 & 0 & -1 & 0 \\ i_2 & 0 & +1 & 0 & -1 \\ i_3 & 0 & -1 & +1 & 0 \\ i_4 & -1 & 0 & 0 & +1 \end{array}$$

where (i_1, j_1) and (i_3, j_2) belong to distinct blocks in S and other nonzero cells belong to (not necessarily distinct) blocks in S^C. The i_1-th and i_2-th rows belong to the same block of rows. Similarly the i_3-th and i_4-th rows belong to the same block of rows. There are both square-free and non-square-free moves of this type. Type IV includes the transpose of these moves.

We now give a Markov basis for A_S with its explicit form as follows.

Theorem 3. *The set of moves of Types I–VI in $\mathcal{M}_S$ forms a Markov basis for the configuration A_S.*

We establish Theorem 3 by the lemmas below. Suppose that $\boldsymbol{x}$ and $\boldsymbol{y}$ belong to the same fiber $\mathcal{F}_{\boldsymbol{t}}$ and let $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$. The first lemma is proved by the same argument as Lemma 2 of [8] and we omit its proof.

Lemma 1. *Suppose that $\boldsymbol{z}$ contains a block $\mathcal{I}_{kl}$ such that there exist two cells $(i, j), (i', j') \in \mathcal{I}_{kl}$ with $z_{ij} > 0$ and $z_{i'j'} < 0$. Then $\|\boldsymbol{z}\|_1$ can be reduced by moves of Type I in $\mathcal{M}_S$.*

By this lemma from now on we assume that every block $\mathcal{I}_{kl}$ in $\boldsymbol{z}$, $1 \le k, l \le N$, satisfies $\boldsymbol{z}(\mathcal{I}_{kl}) \ge 0$ or $\boldsymbol{z}(\mathcal{I}_{kl}) \le 0$. Let $K = \{k \mid 1 \le k \le N, \boldsymbol{z}(\mathcal{I}_{kk}) > 0\}$ denote the set of indices of positive diagonal blocks and let $L = \{l \mid 1 \le l \le N, \boldsymbol{z}(\mathcal{I}_{ll}) < 0\}$ denote the set of indices of negative diagonal blocks.

Lemma 2. *If $\boldsymbol{z}(S) = 0$, then $\boldsymbol{z}$ can be reduced by a move of Type I or II in $\mathcal{M}_S$.*

Proof. After an appropriate block-wise interchange of rows and columns, we assume $\boldsymbol{z}(\mathcal{I}_{12}) > 0$ and $\mathcal{I}_{11}, \mathcal{I}_{22} \subseteq S$ without loss of generality. Let (i, j) be a positive cell in $\mathcal{I}_{12}$ as shown in Figure 2. Then $\boldsymbol{z}$ contains negative cells in the i-th row and the j-th column. By an appropriate

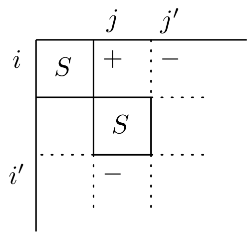


Figure 2: $\boldsymbol{z}$ with $\boldsymbol{z}(S) = 0$.

block-wise interchange of rows and columns, we assume that two of these cells belong to $\mathcal{I}_{13}$ and $\mathcal{I}_{32}$, respectively. Note that $\mathcal{I}_{33}$ may or may not be contained by S. Denote the two negative cells by (i, j') and (i', j). If $\mathcal{I}_{33} \not\subseteq S$, $\|\boldsymbol{z}\|_1$ can be reduced by $(i, j)(i', j') - (i', j)(i, j')$. Hence let us consider the case of $\mathcal{I}_{33} \subseteq S$. Since $z_{+j'} = 0$, the j'-th column contains a positive cell. By an appropriate block-wise interchange of rows, we assume that this positive cell belongs to $\mathcal{I}_{23}$ as in Figure 3. Here, $\mathcal{I}_{22}$ may or may not be contained by S. Denote the positive cell by (i'', j'). If $\mathcal{I}_{22} \not\subseteq S$, $\boldsymbol{z}$ can be reduced by $(i, j)(i'', j') - (i'', j)(i, j')$. If $\mathcal{I}_{22} \subseteq S$, there exists a negative cell $(i'', j'') \in \mathcal{I}_{2l}, l \ne 2, 3$ and $\|\boldsymbol{z}\|_1$ can be reduced by $(i, j)(i', j'')(i'', j') - (i', j)(i'', j'')(i, j')$. □

Lemma 3. *Suppose that $\boldsymbol{z}(\mathcal{I}_{kk}) > 0, \boldsymbol{z}(\mathcal{I}_{ll}) < 0$ and $\boldsymbol{z}$ contains a cell $(i, j) \in \mathcal{I}_{kk}$ with $z_{ij} > 0$ such that*

$$r_l \le \forall i' < r_{l+1} : z_{i'j} \ge 0 \quad \textit{and} \quad c_l \le \forall j' < c_{l+1} : z_{ij'} \ge 0.$$

Then $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type III in $\mathcal{M}_S$.

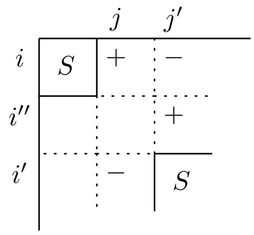


Figure 3: $\boldsymbol{z}$ with $\boldsymbol{z}(S) = 0$.

Proof. After an appropriate block-wise interchange of rows and columns, we assume that $k =$ and $l = 2$. Let (i', j') be a negative cell in $\mathcal{I}_{ll}$. Since $z_{ij} > 0$ and $z_{i+} = z_{+j} = 0$, there exist two positive cells $(i'', j), (i, j'')$ with $(i'', j) \notin \mathcal{I}_{kk}, \mathcal{I}_{lk}$ and $(i, j'') \notin \mathcal{I}_{kk}, \mathcal{I}_{kl}$ as in Figure 4. Hence, $\|\boldsymbol{z}\|_1$ can be reduced by $(i, j)(i'', j')(i', j'') - (i'', j)(i', j')(i, j'')$. $\square$

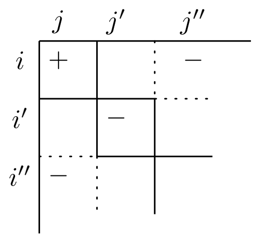


Figure 4: $\boldsymbol{z}$ with the condition of Lemma 3.

Lemma 4. *Suppose that $z_{ij} > 0, (i, j) \in \mathcal{I}_{kk}$ and there exists $l \in L$ satisfying*

$$r_l \leq \exists i' < r_{l+1} : z_{i'j} < 0 \quad \textit{and} \quad \boldsymbol{z}(\mathcal{I}_{kl}) > 0$$

or

$$c_l \leq \exists j' < c_{l+1} : z_{ij'} < 0 \quad \textit{and} \quad \boldsymbol{z}(\mathcal{I}_{lk}) > 0.$$

Then $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type III or IV in $\mathcal{M}_S$.

Proof. After an appropriate block-wise interchange of rows and columns, we assume that $k = 1, l = 2$, $r_2 \leq \exists i' < r_3 : z_{i'j} < 0$ and $\boldsymbol{z}(\mathcal{I}_{12}) > 0$. If there exists a pair of cells $(i_1, j') \in \mathcal{I}_{12}$ and $(i_2, j') \in \mathcal{I}_{22}$ with $z_{i_1 j'} > 0, z_{i_2 j'} < 0$ as in Figure 5, $\|\boldsymbol{z}\|_1$ can be reduced by $(i, j)(i', j_2)(i_1, j')(i_2, j_1) - (i', j)(i_1, j_2)(i_2, j')(i, j_1)$. If there exists no such pair, $\boldsymbol{z}$ satisfies $\boldsymbol{z}(\mathcal{I}_{11}) > 0, \boldsymbol{z}(\mathcal{I}_{22}) < 0$ and

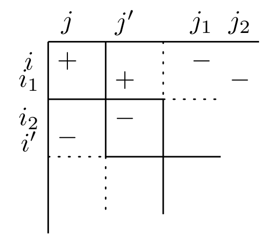


Figure 5: $\boldsymbol{z}$ with the condition of Lemma 4.

there exists a cell $(i', j') \in \mathcal{I}_{22}$ with $-z_{i'j'} > 0$ such that $r_1 \leq \forall i'' < r_2 : -z_{i''j'} \geq 0$ and $c_1 \leq \forall j'' < c_2 : -z_{i'j''} \geq 0$. Hence, by Lemma 3, $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type III. $\square$

Proof of Theorem 3. By Lemmas 1, 2 and 4, it is enough to show that $\|\boldsymbol{z}\|_1$ can be reduced by a move of types I–IV under the following conditions:

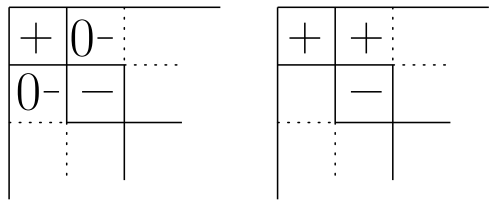
- $\boldsymbol{z}(\mathcal{I}_{kl}) \geq 0$ or $\boldsymbol{z}(\mathcal{I}_{kl}) \leq 0$ holds for all $1 \leq k, l \leq N$ and $\boldsymbol{z}(S) \neq 0$.
- For every $k \in K$, $l \in L$ and $(i, j) \in \mathcal{I}_{kk}$ with $z_{ij} > 0$,

$$r_l \leq \forall i' < r_{l+1} : z_{i'j} \geq 0 \quad \text{or} \quad \boldsymbol{z}(\mathcal{I}_{kl}) \leq 0$$

  and

$$c_l \leq \forall j' < c_{l+1} : z_{ij'} \geq 0 \quad \text{or} \quad \boldsymbol{z}(\mathcal{I}_{lk}) \leq 0.$$

For such $\boldsymbol{z}$ fix $k \in K$ and $l \in L$ and consider the case that the above conditions are satisfied by $\boldsymbol{z}(\mathcal{I}_{kl}) \leq 0$ and $\boldsymbol{z}(\mathcal{I}_{lk}) \leq 0$ as in (a) of Figure 6. In Figure 6 we assume that $k = 1$ and $l = 2$ without loss of generality. In this case $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type III from



(a) $\mathbf{z}(\mathcal{I}_{kl}), \mathbf{z}(\mathcal{I}_{lk}) \leq 0$ (b) $\mathbf{z}(\mathcal{I}_{kl}) > 0$

Figure 6: $\boldsymbol{z}$ with the conditions in the proof of Theorem 3.

the sign-reverse case of Lemma 3. Finally, consider the case that at least one of $\boldsymbol{z}(\mathcal{I}_{kl}) \leq 0$ and $\boldsymbol{z}(\mathcal{I}_{lk}) \leq 0$ does not hold. Let $\boldsymbol{z}(\mathcal{I}_{kl}) > 0$ as in (b) of Figure 6. It is obvious from Lemma 3 that $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type III. □

5 Markov bases for general block diagonal effect models

In the common block diagonal effect model we assume that every diagonal block has the common parameter γ. In this section we discuss the case that each diagonal block S_n has its own parameter γ_n and more general cases of block diagonal effect.

We introduce the following model for block diagonal effect by the slight modification to (2). Let $S_n, n = 1, \ldots, N$, be the set of cells belonging to the n-th diagonal block defined as

$$S_n = \{(i, j) \mid r_n \leq i < r_{n+1}, c_n \leq j < c_{n+1}\}.$$

Then the block diagonal effect model with block-wise parameters $\gamma_n, n = 1, \ldots, N$, is defined by

$$\log p_{ij} = \mu + \alpha_i + \beta_j + \sum_{n=1}^{N} \gamma_n I_{S_n}(i, j). \tag{3}$$

Note that the model (3) contains the quasi-independence model considered in Hara et al. [9] as a special case.

The sufficient statistic for the model (3) consists of the row sums, column sums and the sums of frequencies in the block diagonal sets $S_n, n = 1, \ldots, N$, and is summarized as

$$\boldsymbol{t} = (x_{1+}, \ldots, x_{R+}, x_{+1}, \ldots, x_{+C}, x_{S_1}, \ldots, x_{S_N})'.$$

Then a Markov basis for $A_{S_1,\dots,S_N}$ is obtained by essentially the same arguments in the proof of Theorem 3.

Proposition 1. *If $N = 2$, the set of moves of Type I in $\mathcal{M}_{S_1,\dots,S_N}$ forms the unique minimal Markov basis for the configuration $A_{S_1,\dots,S_N}$. If $N \geq 3$, the set of moves of Types I and II in $\mathcal{M}_{S_1,\dots,S_N}$ forms the unique minimal Markov basis for the configuration $A_{S_1,\dots,S_N}$.*

Proof. When $N = 2$, it is easy to see that fixing the sums in $\boldsymbol{t}$ is equivalent to fixing the sums in $\boldsymbol{t}$ and the sums of $\mathcal{I}_{12}$ and $\mathcal{I}_{21}$. Then, if $\boldsymbol{z} \neq 0$ there exists a block containing both a positive cell and a negative cell. Hence by the same argument in the proof of Lemma 1 $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type I.

Let us consider the case of $N \geq 3$. If there exists a diagonal block $\boldsymbol{z}(S_n) \neq 0$ for $1 \leq n \leq N$, S_n contains both a positive cell and a negative cell. Then, by the same argument in the proof of Lemma 1, $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type I. In the case that $\boldsymbol{z}(S_n) = 0, n = 1, \dots, N$, by the same argument in the proof of Lemma 2, $\|\boldsymbol{z}\|_1$ can be reduced by a move of Type II.

The uniqueness of the minimal Markov basis follows, since the basic moves and the moves of Type II in $\mathcal{M}_{S_1,\dots,S_N}$ are indispensable. □

In Theorem 3 in Section 4 and Proposition 1, we assumed $r_{n+1} = R+1$ and $c_{n+1} = C+1$ in the definition of subtables. In fact, the set of moves of Types I – IV forms a Markov basis for the configurations arising from more general block diagonal effect models. Let $1 = r_1 < r_2 < \cdots < r_{N+1} \leq R+1$, $1 = c_1 < c_2 < \cdots < c_{N+1} \leq C+1$ and $S_n = \{(i,j) \mid r_n \leq i < r_{n+1}, c_n \leq j < c_{n+1}\}$. Denote $\mathcal{S} = \{S_1, \dots, S_N\}$. Let $T_q, q = 1, \dots, Q$, be the subtables of $\mathcal{I}$ of the form $T_q = \bigcup_{n \in \hat{\mathcal{N}}_q} S_n$ where $\hat{\mathcal{N}}_q$ is a subset of $\{1, 2, \dots, N\}$ with $\hat{\mathcal{N}}_q \cap \hat{\mathcal{N}}_{q'} = \emptyset$ for $1 \leq q < q' \leq Q$. Then the general block diagonal effect model is defined by

$$\log p_{ij} = \mu + \alpha_i + \beta_j + \sum_{q=1}^{Q} \gamma_q I_{T_q}(i,j). \tag{4}$$

The sufficient statistic of the model (4) is summarized as

$$(x_{1+}, \dots, x_{R+}, x_{+1}, \dots, x_{+C}, x_{T_1}, \dots, x_{T_Q})'.$$

By the same argument of the proofs of Theorem 3 and Proposition 1, we have the following corollary.

Corollary 1. *The set of moves of Types I–VI in $\mathcal{M}_{T_1,\dots,T_Q}$ forms a Markov basis for the configuration $A_{T_1,\dots,T_Q}$.*

6 Numerical experiments

In this section we apply the MCMC method with the Markov bases derived in the previous sections for performing conditional tests of some data sets.

The first example is Table 1 which shows the relationship between school and clothing for 1725 children. This data is from Gilby and Pearson [5]. Each row represents a primary school of the usual county-council type. The rows are arranged in ascending order of the wealth of the children. The children are also classified by their clothing and the columns are arranged

Table 1: Relationship between school and clothing.

	I	II	III	IV & V	Total
No. 1	86	49	10	1	146
No. 2	102	116	24	3	245
No. 3	25	19	2	0	46
No. 4	137	98	33	4	272
No. 5	209	222	73	16	520
No. 6	65	154	71	27	317
No. 7	9	33	1	1	44
No. 8	3	60	51	21	135
Total	636	751	265	73	1725

in ascending order of the neatness of their clothing. We set the model (1) with two subtables $S_1 = \{(i,1) \mid 1 \le i \le 3\}$, $S_2 = \{(i,j) \mid 1 \le i \le 5, 1 \le j \le 2\}$ as a null hypothesis. Starting from the observed data in Table 1 we run a Markov chain of 100,000 tables including 10,000 burn-in steps and compute the chi-square statistic for each sampled table. The histogram of chi-square statistics is shown in Figure 7. In the figure the black line shows the asymptotic distribution χ^2_{19}. Since the observed data is large enough, the estimated exact distribution is close to χ^2_{19}. For the observed data in Table 1 the value of chi-square statistic is 154 and the approximate p-value is essentially zero. Therefore the change point model (1) is rejected at the significant level of 5%.

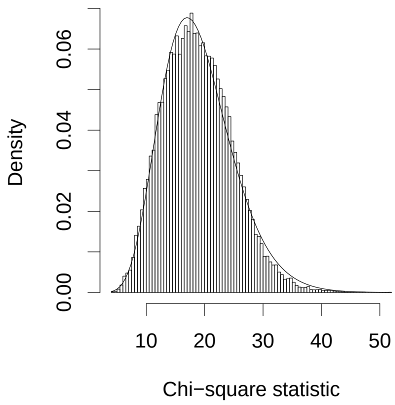


Figure 7: A histogram of chi-square statistic.

The second example is Table 2 which shows the relationship between birthday and deathday for 82 descendants of Queen Victoria. This data is from Diaconis and Sturmfels [3]. Let $r_i = c_i = 1 + 3(i-1)$ for $i = 1, \dots, 5$, permitting the replication modulo 12. We test the common block diagonal effect model (2) against the block diagonal model (3) with several parameters. Starting from the observed data in Table 2 we run a Markov chain of 1,000,000 tables including

Table 2: Relationship between birthday and deathday.

	Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec
Jan	1	0	0	0	1	2	0	0	1	0	1	0
Feb	1	0	0	1	0	0	0	0	0	1	0	2
March	1	0	0	0	2	1	0	0	0	0	0	1
April	3	0	2	0	0	0	1	0	1	3	1	1
May	2	1	1	1	1	1	1	1	1	1	1	0
June	2	0	0	0	1	0	0	0	0	0	0	0
July	2	0	2	1	0	0	0	0	1	1	1	2
Aug	0	0	0	3	0	0	1	0	0	1	0	2
Sep	0	0	0	1	1	0	0	0	0	0	1	0
Oct	1	1	0	2	0	0	1	0	0	1	1	0
Nov	0	1	1	1	2	0	0	2	0	1	1	0
Dec	0	1	1	0	0	0	1	0	0	0	0	0

100,000 burn-in steps and compute (2×) the log-likelihood ratio statistic

$$2 \sum_{(i,j)\in\mathcal{I}} x_{ij} \log \frac{\hat{m}_{ij}^2}{\hat{m}_{ij}^1}$$

for each sampled table $\boldsymbol{x} = \{x_{ij}\}$ where $\hat{m}_{ij}^1$ and $\hat{m}_{ij}^2$ denote the expected cell frequencies under the model (2) and (3), respectively. The histogram of log-likelihood statistics is shown in Figure 8. In the figure the black line shows the asymptotic distribution χ_3^2. From the sparsity of Table 2 the estimated exact distribution is different from the asymptotic distribution χ_3^2. For the observed data in Table 2 the value of log-likelihood ratio statistic is 3.07 and the approximate p-value is 0.43. Therefore the common diagonal block effect model (2) is accepted at the significant level of 5%.

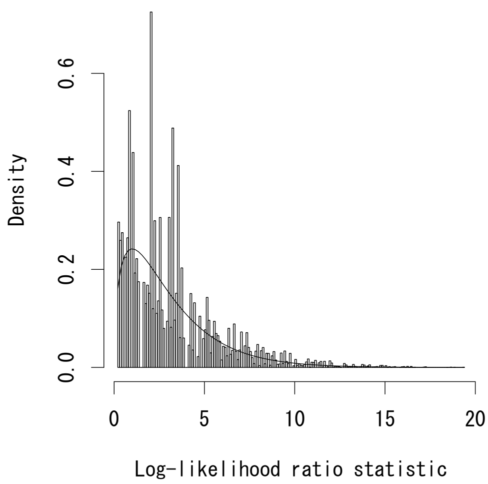


Figure 8: A histogram of log-likelihood ratio statistic.

7 Concluding remarks

In this paper we derive the explicit forms of the Markov bases for some statistical models with block-wise subtable effects and perform the conditional testing with some real data sets. For the change point model we also discussed the algebraic properties of the configuration arising from the model. Hara et al. [8] gave the necessary and sufficient condition on the subtable so that the set of basic moves forms a Markov basis. It is of interest to consider a necessary and sufficient condition on subtables $S_1, \ldots, S_N$ so that the set of basic moves forms a Markov basis.

Acknowledgment

We are very grateful to Hisayuki Hara for valuable discussions.

References

[1] Satoshi Aoki and Akimichi Takemura. Markov chain Monte Carlo exact tests for incomplete two-way contingency table. *Journal of Statistical Computation and Simulation*, 75(10):787–812, 2005.

[2] David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms.* Undergraduate Texts in Mathematics. Springer, New York, third edition, 2007. ISBN 978-0-387-35650-1; 0-387-35650-9. An introduction to computational algebraic geometry and commutative algebra.

[3] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363–397, 1998. ISSN 0090-5364.

[4] Adrian Dobra and Seth Sullivant. A divide-and-conquer algorithm for generating Markov bases of multi-way tables. *Comput. Statist.*, 19(3):347–366, 2004. ISSN 0943-4062.

[5] Walter H. Gilby and Karl Pearson. On the significance of the teacher's appreciation of general intelligence. *Biometrika*, 8(1):94–108, 1911.

[6] Shelby J. Haberman. A warning on the use of chi-squared statistics with frequency tables with small expected cell counts. *J. Amer. Statist. Assoc.*, 83(402):555–560, 1988. ISSN 0162-1459.

[7] Hisayuki Hara, Satoshi Aoki, and Akimichi Takemura. Minimal and minimal invariant Markov bases of decomposable models for contingency tables. *Bernoulli*, 16(1):208–233, 2010. ISSN 1350-7265.

[8] Hisayuki Hara, Akimichi Takemura, and Ruriko Yoshida. Markov bases for two-way subtable sum problems. *J. Pure Appl. Algebra*, 213(8):1507–1521, 2009. ISSN 0022-4049.

[9] Hisayuki Hara, Akimichi Takemura, and Ruriko Yoshida. A Markov basis for conditional test of common diagonal effect in quasi-independence model for square contingency tables. *Comput. Statist. Data Anal.*, 53(4):1006–1014, 2009. ISSN 0167-9473.

[10] Chihiro Hirotsu. Two-way change-point model and its application. *Austral. J. Statist.*, 39 (2):205–218, 1997. ISSN 0004-9581.

[11] Hidefumi Ohsugi and Takayuki Hibi. Toric ideals generated by quadratic binomials. *J. Algebra*, 218(2):509–527, 1999. ISSN 0021-8693.

[12] Hidefumi Ohsugi and Takayuki Hibi. Two way subtable sum problems and quadratic Gröbner bases. *Proc. Amer. Math. Soc.*, 137(5):1539–1542, 2009. ISSN 0002-9939.

[13] Fabio Rapallo. Markov bases and structural zeros. *Journal of Symbolic Computation*, 41: 164–172, 2006.

[14] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996. ISBN 0-8218-0487-1.

[15] Akimichi Takemura and Satoshi Aoki. Some characterizations of minimal Markov basis for sampling from discrete conditional distributions. *Ann. Inst. Statist. Math.*, 56(1):1–17, 2004. ISSN 0020-3157.