MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Connecting Tables with Zero-One Entries by a Subset of a Markov Basis

Hisayuki HARA and Akimichi TAKEMURA

METR 2009–35 August 2009

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

Connecting tables with zero-one entries by a subset of a Markov basis

Hisayuki Hara* and Akimichi Takemura†‡

August 2009

Abstract

We discuss connecting tables with zero-one entries by a subset of a Markov basis. In this paper, as a Markov basis we consider the Graver basis, which corresponds to the unique minimal Markov basis for the Lawrence lifting of the original configuration. Since the Graver basis tends to be large, it is of interest to clarify conditions such that a subset of the Graver basis, in particular a minimal Markov basis itself, connects tables with zero-one entries. We give some theoretical results on the connectivity of tables with zero-one entries. We also study some common models, where a minimal Markov basis for tables without the zero-one restriction does not connect tables with zero-one entries.

Key words: Graver basis, Latin squares, logistic regression, Rasch model

1 Introduction

Markov bases methodology initiated by Diaconis and Sturmfels [1998] for performing conditional tests of discrete exponential family models have been extensively studied in recent years. Since the size of a Markov basis tends to be large for large-scale problems, researchers are interested in a subset of Markov basis for a specific fiber. In most applications of Markov basis there are no restrictions on the cell counts. However in some problems, the counts are either zero or one. The most well-known case is the Rasch model used in educational statistics.

The Rasch model can be interpreted as a logistic regression (logit model), where the number of trials is just one for each combination of covariates. In this model, tables with zero-one entries (zero-one tables) are elements of a specific fiber, where the marginal frequencies corresponding to the response variable are all equal to one in the logistic

*Department of Technology Management for Innovation, University of Tokyo
†Graduate School of Information Science and Technology, University of Tokyo
‡JST, CREST

regression. In other cases, zero-one table appears as truncation or dichotomization of a variable, where for example only an occurrence or non-occurrence of certain large event is recorded. A convenient statistical model for zero-one tables is a log-linear model for contingency tables, where the support of the distribution is restricted to zero-one tables. Then we can use the Markov basis methodology for conditional tests of the fit of the model.

Another source of zero-one tables is the set of incidence matrices satisfying certain combinatorial restrictions. For example the set of Latin squares or sudoku can be considered as a set of zero-one tables with fixed marginals. From combinatorial viewpoint it is of interest to construct a connected Markov chain over the set of these tables.

Note that a minimal Markov basis without the zero-one restriction may not connect zero-one tables, because by applying a move from the Markov basis, some cells may contain frequencies greater than one. On the other hand the set of square-free moves of the Graver basis connects tables with zero-one entries. Therefore it is of interest to study when a minimal Markov basis connects zero-one tables, and if this is not the case, to find a subset of the Graver basis connecting zero-one tables.

In this paper we give some theoretical results on the connectivity of tables with zero-one entries. Unfortunately we found that our sufficient conditions for connectivity are satisfied only in a few examples. Therefore we investigate some common models, where a minimal Markov basis for tables without the zero-one restriction does not connect tables with zero-one entries.

The organization of the paper is as follows. For the rest of this section we summarize our notation and preliminary facts. In Section 2 we give some theoretical results on connectivity of zero-one tables with a minimal Markov basis and with some other subsets of the Graver basis. In Section 3 we study connectivity of zero-one tables in some common models for contingency tables, including the Rasch model, its multivariate version and the logistic regression. We also discuss Latin squares. We conclude the paper with some remarks in Section 4.

1.1 Notation and preliminary facts

Here we set up our notation and summarize preliminary facts on Markov bases and the Graver basis. We mostly follow the notation in Hara et al. [2007]. Let $\mathcal{I}$ denote the set of cells of a table. $i \in \mathcal{I}$ is usually a multi-index. Let $I = |\mathcal{I}|$ be the number of cells. A contingency table or a frequency vector is denoted by $\boldsymbol{x} = (x(i))_{\{i \in \mathcal{I}\}}$. For a given $\boldsymbol{x}$, $\mathrm{supp}(\boldsymbol{x}) = \{i \mid x(i) > 0\}$ denotes the set of positive cells of $\boldsymbol{x}$. Given a loglinear model (more precisely a "toric model"), the sufficient statistic $\boldsymbol{t}$ can be written as $\boldsymbol{t} = A\boldsymbol{x}$ for some integral matrix A. We call A a configuration of the model. I_A denotes the toric ideal of A. The set of contingency tables with the common sufficient statistic $\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x} \geq 0 \mid \boldsymbol{t} = A\boldsymbol{x}\}$ is a *fiber*.

An integer vector $\boldsymbol{z}$ is called a move if $A\boldsymbol{z} = \boldsymbol{0}$. Separating positive elements and negative elements of $\boldsymbol{z}$, we write $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$, where $\boldsymbol{z}^+$ is the positive part of $\boldsymbol{z}$ and $\boldsymbol{z}^-$ is the negative part of $\boldsymbol{z}$. The sample size of $\boldsymbol{z}^+$ ($\boldsymbol{z}^-$) is called degree of $\boldsymbol{z}$. We denote a square-free move $\boldsymbol{z}$ of degree m by $\boldsymbol{z} = (i_1)(i_2)\cdots(i_m) - (j_1)(j_2)\cdots(j_m)$, where $\mathrm{supp}(\boldsymbol{z}^+) = \{i_1,\ldots,i_m\}$ and $\mathrm{supp}(\boldsymbol{z}^-) = \{j_1,\ldots,j_m\}$. $|\boldsymbol{z}| = \sum_{i\in\mathcal{I}}|z(i)|$ denotes the L_1-norm of $\boldsymbol{z}$. For two moves $\boldsymbol{z}_1, \boldsymbol{z}_2$, the sum $\boldsymbol{z}_1 + \boldsymbol{z}_2$ is called *conformal* if there is no cancellation of signs in $\boldsymbol{z}_1 + \boldsymbol{z}_2$, i.e., $\emptyset = \mathrm{supp}(\boldsymbol{z}_1^+) \cap \mathrm{supp}(\boldsymbol{z}_2^-) = \mathrm{supp}(\boldsymbol{z}_1^-) \cap \mathrm{supp}(\boldsymbol{z}_2^+)$. A move $\boldsymbol{z}$ which can not be written as a conformal sum of two (non-zero) moves is called *primitive*. The set of primitive moves is finite and it is called the Graver basis of I_A. We denote the Graver basis as $\mathcal{B}_{\mathrm{GR}}$.

Let E_I denote the $I \times I$ identity matrix. The configuration

$$\Lambda(A) = \begin{pmatrix} A & 0 \\ E_I & E_I \end{pmatrix} \tag{1}$$

is called the Lawrence lifting of A. In statistical terms, the Lawrence lifting corresponds to the logistic regression, where the interaction effects of the covariates are specified by A. It is known ([Sturmfels, 1996, Theorem 7.1]) that the unique minimal Markov basis of $I_{\Lambda(A)}$ coincides with the Graver basis of I_A.

A finite set of moves $\mathcal{B}$ is *distance reducing* (Takemura and Aoki [2005]) if for all $\boldsymbol{t}$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ there exists an element $\boldsymbol{z} \in \mathcal{B}$ and $\epsilon = \pm 1$ such that

$$\boldsymbol{x} + \epsilon\boldsymbol{z} \in \mathcal{F}_{\boldsymbol{t}},\ |\boldsymbol{x} + \epsilon\boldsymbol{z} - \boldsymbol{y}| < |\boldsymbol{x} - \boldsymbol{y}| \quad \text{or} \quad \boldsymbol{y} + \epsilon\boldsymbol{z} \in \mathcal{F}_{\boldsymbol{t}},\ |\boldsymbol{x} - (\boldsymbol{y} + \epsilon\boldsymbol{z})| < |\boldsymbol{x} - \boldsymbol{y}|.$$

If $\mathcal{B}$ is distance reducing, it is obviously a Markov basis and we call $\mathcal{B}$ a distance reducing Markov basis. Furthermore $\mathcal{B}$ is *strongly distance reducing* if for all $\boldsymbol{t}$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}_{\boldsymbol{t}}$ there exist elements $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{B}$ and $\epsilon_1, \epsilon_2 = \pm 1$ such that $\boldsymbol{x} + \epsilon_1\boldsymbol{z}_1, \boldsymbol{y} + \epsilon_2\boldsymbol{z}_2 \in \mathcal{F}_{\boldsymbol{t}}$, $|\boldsymbol{x} + \epsilon_1\boldsymbol{z}_1 - \boldsymbol{y}| < |\boldsymbol{x} - \boldsymbol{y}|$ and $|\boldsymbol{x} - (\boldsymbol{y} + \epsilon_2\boldsymbol{z}_2)| < |\boldsymbol{x} - \boldsymbol{y}|$.

Since we are considering zero-one tables in this paper, let us denote

$$\tilde{\mathcal{F}}_{\boldsymbol{t}} = \{\boldsymbol{x} \mid \boldsymbol{t} = A\boldsymbol{x},\ x(i) = 0 \text{ or } 1\}. \tag{2}$$

As in the usual setting for Markov basis, we call a finite set $\mathcal{B}$ of moves a Markov basis for zero-one tables, if $\mathcal{B}$ connects all fibers $\tilde{\mathcal{F}}_{\boldsymbol{t}}$. If $\mathcal{B}$ is distance reducing for zero-one tables, then it is a distance reducing Markov basis for zero-one tables. Since there are $2^{|\mathcal{I}|}$ zero-one tables, there are only finite number of fibers and finite number of differences of two elements belonging to the same fiber. Therefore the set of these differences is the largest trivial Markov basis. However clearly this set is too large and we are interested in a much smaller set of moves connecting all fibers $\tilde{\mathcal{F}}_{\boldsymbol{t}}$.

2 Some theoretical results

The starting point of our investigation of connectivity of zero-one tables is the following basic fact on the Graver basis $\mathcal{B}_{\mathrm{GR}}$ for I_A.

Proposition 1. *Let $\mathcal{B}_0$ denote the set of square-free moves of the Graver basis $\mathcal{B}_{\mathrm{GR}}$ of I_A. Then $\mathcal{B}_0$ is strongly distance reducing for tables with zero-one entries.*

Proof. Let $\boldsymbol{x}$, $\boldsymbol{y}$ be two zero-one tables of the same fiber. They are connected by a conformal sum of primitive moves:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_K. \tag{3}$$

Since there is no cancellation of signs on the right-hand side, once an entry greater than equal to 2 appears in an intermediate sum of the right-hand side, it can not be canceled. Therefore it follows that that $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K \in \mathcal{B}_0$. Since there are no sign cancellations in (3), $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K$ can be added to $\boldsymbol{x}$ in any order and $-\boldsymbol{z}_1, \ldots, -\boldsymbol{z}_K$ can be added to $\boldsymbol{y}$ in any order. Therefore $\mathcal{B}_0$ is strongly distance reducing. $\square$

Since the Graver basis tends to be large, we are interested in conditions for connecting tables with zero-one entries with a subset of the Graver basis. We consider the following condition.

Condition 1. *(Existence of strong crossing pattern)*
Let $\mathbf{e}_i$ denote the frequency vector with just 1 frequency in the i-th cell and 0 otherwise. For every fiber and every $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$, in the same fiber, there exist distinct cells i_1, i_2, i_3, i_4 such that $x(i_1) > y(i_1), x(i_2) > y(i_2), x(i_3) < y(i_3), x(i_4) \leq y(i_4)$ or $y(i_1) > x(i_1), y(i_2) > x(i_2), y(i_3) < x(i_3), y(i_4) \leq x(i_4)$ and

$$\boldsymbol{z} = \boldsymbol{e}_{i_3} + \boldsymbol{e}_{i_4} - \boldsymbol{e}_{i_1} - \boldsymbol{e}_{i_2} \tag{4}$$

is a move.

Note the set $\mathcal{B}$ of the moves $\boldsymbol{z}$ in (4) forms a distance reducing Markov basis. Therefore Condition 1 is a sufficient condition for existence of a distance reducing Markov basis consisting of square-free moves of degree two. However existence of such a Markov basis does not imply Condition 1. We discuss this point at the end of this section. Under Condition 1 we have the following result.

Theorem 1. *Under Condition 1, the set $\mathcal{B}$ of moves (4) is distance reducing for tables with zero-one entries.*

Proof. Let $\boldsymbol{x}$, $\boldsymbol{y}$ be two zero-one tables in the same fiber. We first ignore the restriction that the entries of the tables are restricted to $\{0, 1\}$. By Condition 1, we can find distinct cells i_1, i_2, i_3, i_4 such that

$$x(i_1) \geq y(i_1) + 1,\ x(i_2) \geq y(i_2) + 1 \quad \Rightarrow \quad x(i_1) = 1,\ x(i_2) = 1,\ y(i_1) = 0, y(i_2) = 0$$

and $0 \leq x(i_3) < y(i_3) \leq 1 \Rightarrow x(i_3) = 0, y(i_3) = 1$. If $x(i_4) = 0$ then we can add $\boldsymbol{z} = \boldsymbol{e}_{i_3} + \boldsymbol{e}_{i_4} - \boldsymbol{e}_{i_1} - \boldsymbol{e}_{i_2}$ to $\boldsymbol{x}$ and reduce the L_1-distance by four. Furthermore $\boldsymbol{x} + \boldsymbol{z}$ is a table of zeros and ones.

It remains to consider the case $x(i_4) = 1$. Since $x(i_4) \leq y(i_4)$, we have $y(i_4) = 1$. Therefore $y(i_1) = 0, y(i_2) = 0, y(i_3) = 1, y(i_4) = 1$. Then we can subtract $\boldsymbol{z}$ from $\boldsymbol{y}$ and $\boldsymbol{y} - \boldsymbol{z}$ is a table of zeros and ones. Furthermore $|\boldsymbol{x} - (\boldsymbol{y} - \boldsymbol{z})| = |\boldsymbol{x} - \boldsymbol{y}| - 2$. Therefore under Condition 1 we can reduce the distance always by at least 2. Therefore $\mathcal{B}$ is distance reducing for fibers of zero-one tables. $\square$

Theorem 1 is simple and effective to prove that a particular Markov basis connects zero-one tables for some simple configurations. We now present several generalizations of Theorem 1. The following proposition is an obvious extension of Theorem 1 and we omit a proof.

Proposition 2. *Assume that there exists a positive integer M, such that for every fiber and every $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$, in the same fiber there exists a positive integer $m \leq M$ and distinct cells $i_1, \ldots, i_{2m}$ such that*

$$\boldsymbol{z} = \sum_{j=m+1}^{2m} \boldsymbol{e}_{i_j} - \sum_{j=1}^{m} \boldsymbol{e}_{i_j} \tag{5}$$

is a move such that at least one of the following conditions hold: i) $x(i_j) > y(i_j)$, $j = 1, \ldots, m$, $x(i_j) < y(i_j)$, $j = m+1, \ldots, 2m-1$, $x(i_{2m}) \leq y(i_{2m})$, or ii) $y(i_j) > x(i_j)$, $j = 1, \ldots, m$, $y(i_j) < x(i_j)$, $j = m+1, \ldots, 2m-1$, $y(i_{2m}) \leq x(i_{2m})$. Then the set $\mathcal{B}$ of moves $\boldsymbol{z}$ in (5) is distance reducing for tables with zero-one entries.

Proposition 2 suggests a possibility to choose a subset of $\mathcal{B}_0$ of Proposition 1, which still guarantees the connectivity of tables with zero-one entries. Let $\mathcal{B}$ be a subset of $\mathcal{B}_0$ with the following property.

Condition 2. *(Generalized strong crossing pattern for the Graver basis)*
For every element $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^- \in \mathcal{B}_0 \setminus \mathcal{B}$, there exists a move $\boldsymbol{z}' = \sum_{j=m+1}^{2m} \boldsymbol{e}_{i_j} - \sum_{j=1}^{m} \boldsymbol{e}_{i_j} \in \mathcal{B}$ such that $i_1, \ldots, i_{2m}$ are distinct and at least one of the following conditions hold: i) $\boldsymbol{z}^+(i_j) > \boldsymbol{z}^-(i_j)$, $j = 1, \ldots, m$, $\boldsymbol{z}^+(i_j) < \boldsymbol{z}^-(i_j)$, $j = m+1, \ldots, 2m-1$, $\boldsymbol{z}^+(i_{2m}) \leq \boldsymbol{z}^-(i_{2m})$, or ii) $\boldsymbol{z}^-(i_j) > \boldsymbol{z}^+(i_j)$, $j = 1, \ldots, m$, $\boldsymbol{z}^-(i_j) < \boldsymbol{z}^+(i_j)$, $j = m+1, \ldots, 2m-1$, $\boldsymbol{z}^-(i_{2m}) \leq \boldsymbol{z}^+(i_{2m})$.

Combining Proposition 1 and Proposition 2 we have the following proposition.

Proposition 3. *If $\mathcal{B}$ satisfies Condition 2, then $\mathcal{B}$ is distance reducing for tables with zero-one entries.*

Proof. As in the proof of Proposition 1, consider (3), where $\boldsymbol{x}$, $\boldsymbol{y}$ are two zero-one tables in the same fiber. By induction on the number K of primitive moves, it suffices to prove the distance reduction for $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}_1$. By the same argument as in Lemma 2.4 of Takemura and Aoki [2004], it suffices to check the distance reduction by moves from $\mathcal{B}$ in moving from $\boldsymbol{z}_1^-$ to $\boldsymbol{z}_1^+$. If $\boldsymbol{z}_1 \in \mathcal{B}$, we can reduce the distance at once. If $\boldsymbol{z}_1 \in \mathcal{B}_0 \setminus \mathcal{B}$, we can find $\boldsymbol{z} \in \mathcal{B}$ which can be applied either to $\boldsymbol{z}_1^-$ or $\boldsymbol{z}_1^+$ such that $|\boldsymbol{z}_1|$ is reduced. The resulting move can now be decomposed into a conformal sum of primitive moves and we can recursively use the distance reduction argument. This proves the proposition. □

By Proposition 3, once $\mathcal{B}_0$ is given we can remove some elements from $\mathcal{B}_0$ and obtain a smaller set of moves $\mathcal{B}$ as follows. Find a pair $\boldsymbol{z}, \tilde{\boldsymbol{z}} \in \mathcal{B}_0$, $\boldsymbol{z} \neq \tilde{\boldsymbol{z}}$, such that $\boldsymbol{z} + \tilde{\boldsymbol{z}}$ has just one sign cancellation, i.e. there is only one cell i such that $z(i)\tilde{z}(i) < 0$. If $\boldsymbol{z} + \tilde{\boldsymbol{z}} \in \mathcal{B}_0$ then we can remove $\boldsymbol{z} + \tilde{\boldsymbol{z}}$ from $\mathcal{B}_0$ and still guarantee connectivity of zero-one tables.

As the last topic of this section we clarify the interpretation of Condition 1 by discussing a weaker condition which is equivalent to the existence of distance reducing Markov basis consisting of square-free moves of degree two. In our previous works (e.g. Aoki and Takemura [2005], Hara et al. [2007]) we have obtained such Markov bases and in these works we have used a similar argument as in the proof of Theorem 1. By omitting the requirement $x(i_4) \le y(i_4)$ in Condition 1 consider the following weaker condition:

Condition 3. *(Existence of weak crossing pattern)*
For every fiber and every $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$, *in the same fiber, there exist distinct cells* i_1, i_2, i_3, i_4 *such that* $x(i_1) > y(i_1), x(i_2) > y(i_2), x(i_3) < y(i_3)$ *or* $y(i_1) > x(i_1), y(i_2) > x(i_2), y(i_3) < x(i_3)$ *and* $\boldsymbol{z} = \boldsymbol{e}_{i_3} + \boldsymbol{e}_{i_4} - \boldsymbol{e}_{i_1} - \boldsymbol{e}_{i_2}$ *is a move.*

We now show that Condition 3 is equivalent to the existence of a distance reducing Markov basis consisting of square-free moves of degree two.

Proposition 4. *There exists a distance reducing Markov basis consisting of square-free moves of degree two if and only if Condition 3 holds.*

Proof. It is clear that under Condition 3 the set $\mathcal{B}$ of moves in (4) is distance reducing. Therefore it suffices to show the converse. Let $\mathcal{B}$ be a distance reducing Markov basis consisting of square-free moves of degree two. Let $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$, be in the same fiber. We can find $\pm \boldsymbol{z} \in \mathcal{B}$ such that $\boldsymbol{z}$ is applicable to $\boldsymbol{x}$ or $\boldsymbol{y}$ and $|(\boldsymbol{x}+\boldsymbol{z}) - \boldsymbol{y}| < |\boldsymbol{x} - \boldsymbol{y}|$ or $|\boldsymbol{x} - (\boldsymbol{y}+\boldsymbol{z})| < |\boldsymbol{x} - \boldsymbol{y}|$, respectively. For $\boldsymbol{x}, \boldsymbol{y}$ in the same fiber and for distinct indices i_1, i_2, i_3, i_4 let

$$\begin{aligned} g(i_1, i_2, i_3, i_4, \boldsymbol{x}, \boldsymbol{y}) = {} & I(x(i_1) > y(i_1)) + I(x(i_2) > y(i_2)) \\ & + I(x(i_3) < y(i_3)) + I(x(i_4) < y(i_4)) - 2, \end{aligned}$$

where $I(E)$ denotes the indicator function of the event E. When $\boldsymbol{z}$ can be added to $\boldsymbol{x}$, we have

$$|\boldsymbol{x} - \boldsymbol{y}| - |(\boldsymbol{x}+\boldsymbol{z}) - \boldsymbol{y}| = 2g(i_1, i_2, i_3, i_4, \boldsymbol{x}, \boldsymbol{y}).$$

Therefore

$$|\boldsymbol{x} - \boldsymbol{y}| - |(\boldsymbol{x}+\boldsymbol{z}) - \boldsymbol{y}| > 0 \quad \Leftrightarrow \quad g(i_1, i_2, i_3, i_4, \boldsymbol{x}, \boldsymbol{y}) > 0,$$

i.e., at least 3 inequalities among $x(i_1) > y(i_1), x(i_2) > y(i_2), x(i_3) < y(i_3), x(i_4) < y(i_4)$ hold. It is easy to see that then Condition 3 holds. Similarly if $\boldsymbol{z}$ can be added to $\boldsymbol{y}$ and $|\boldsymbol{x} - (\boldsymbol{y}+\boldsymbol{z})| < |\boldsymbol{x} - \boldsymbol{y}|$, then Condition 3 holds. □

3 Connectivity results for some models

In this section we investigate connectivity of zero-one tables for some common models for contingency tables.

3.1 Rasch model

Rasch model (Rasch [1980]) has long received much attention in the item response theory. Suppose that I persons take a test with J dichotomous questions. Let $x_{ij} \in \{0,1\}$ be a response to the jth question of the ith person. Hence the table $\boldsymbol{x} = (x_{ij})_{1\le i\le I, 1\le j\le J}$ is considered as a two-way contingency table with zero-one entries. The Rasch model is expressed as

$$\Pr(x_{ij}=1) = \frac{\exp(\alpha_i - \beta_j)}{1+\exp(\alpha_i - \beta_j)}, \tag{6}$$

where α_i is an individual's latent ability parameter and β_j is an item's difficulty parameter. Suppose that each x_{ij} is independent. Then the set of row sums $x_{i+} = \sum_{j=1}^{J} x_{ij}$ and column sums $x_{+j} = \sum_{i=1}^{I} x_{ij}$ are the sufficient statistics.

The Rasch model has been extensively studied and practically used for evaluating educational and psychological tests. Many inference procedures have been developed (e.g. Glas and Verhelst [1995]) and most of them rely on asymptotic theory. However, as Rasch [1980] pointed out, a sufficiently large sample size is not necessarily expected in practice. In such cases the asymptotic inference may be inappropriate.

Rasch [1980] proposed to use an exact test procedure. The conditional distribution of a table given person scores and item totals is easily shown to be uniform. In order to implement exact test for Rasch model via Markov basis technique, we need a set of moves which connects every fiber of two-way zero-one tables with fixed row sums and column sums. Ryser [1957] first showed that the set of basic moves in two-way complete independence model of form

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

connects any fiber of zero-one tables with fixed row sums and column sums. Since then, many Monte Carlo procedures via Markov basis technique to compute distribution of test statistics under the null hypothesis of the Rasch model have been proposed (e.g. Besag and Clifford [1989], Ponocny [2001], Cobb and Chen [2003]). Chen and Small [2005] provided a computationally more efficient Monte Carlo procedure for implementing exact tests by using sequential importance sampling.

In the framework of the present paper, the connectivity by basic moves is a consequence of Theorem 1 and Proposition 1. The Rasch model can be regarded as the Lawrence lifting of the independence model for $I\times J$ two-way tables. From p.382 of Diaconis and Sturmfels [1998] we know that the set of loops of degree $2n$, $n \le \min(I,J)$, such as

$$\boldsymbol{z} = \begin{array}{|cccccc|} \hline 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \dots & \dots & 1 & -1 \\ -1 & 0 & \dots & \dots & 0 & 1 \\ \hline \end{array}$$

forms the Graver basis for the independence model of $I \times J$ contingency tables. Since the set of basic moves satisfies Condition 1, where $\boldsymbol{x}$ is the positive part and $\boldsymbol{y}$ is the negative part of these loops, it follows that the set of basic moves connects $I \times J$ zero-one tables with fixed row sums and columns sums.

3.2 Many-facet Rasch model

The many-facet Rasch model is an extension of the Rasch model to multiple items and polytomous responses (e.g. Linacre [1989], Linacre [1994]) and has also been extensively used in practice for evaluating essay exams and scoring systems of judged sports (e.g. Yamaguchi [1999], Zhu et al. [1998], Basturk [2008]).

Suppose that I_1 articles are rated by I_2 reviewers from I_3 aspects on the grade of I_4 scales. Let $i = (i_1, \dots, i_4)$. $x(i) = 1$ if the reviewer i_2 rates the article i_1 as the i_4th grade from the aspect i_3 and otherwise $x(i) = 0$. Then the three-facet Rasch model is expressed by

$$P(X(\boldsymbol{i})) = \frac{\exp\left[i_4(\beta_{i_1} - \beta_{i_2} - \beta_{i_3}) - \beta_{i_4}\right]}{\sum_{i_4=0}^{I_4-1} \exp\left[i_4(\beta_{i_1} - \beta_{i_2} - \beta_{i_3}) - \beta_{i_4}\right]}.$$

In general, the V-facet Rasch model is expressed as

$$P(X(\boldsymbol{i})) = \frac{\exp\left[i_{V+1}(\beta_{i_1} - \beta_{i_2} - \ldots - \beta_{i_V}) - \beta_{i_{V+1}}\right]}{\sum_{i_{V+1}=0}^{I_{V+1}-1} \exp\left[i_{V+1}(\beta_{i_1} - \beta_{i_2} - \ldots - \beta_{i_V}) - \beta_{i_{V+1}}\right]}.$$

When $V = 1$, the model coincides with the Rasch model. The sufficient statistic is the following set of marginal sums

$$\boldsymbol{t} = \{x(i_v i_{V+1}) \mid i_v \in \mathcal{I}_v, v = 1, \dots, V,\ i_{V+1} \in \mathcal{I}_{V+1}\}, \tag{7}$$

where $x(i_v i_{V+1}) = \sum_{v' \neq v, V+1} \sum_{i_{v'}=0}^{I_{v'}-1} x(\boldsymbol{i})$. Hence in order to implement exact tests for the many-facet Rasch model, we need a set of moves which connects any fiber $\tilde{\mathcal{F}}_{\boldsymbol{t}}$ of zero-one tables. In general, however, it is not easy to derive such a set of moves.

As seen in the previous section, in the case of the Rasch model, the set of basic moves for two-way complete independence model connects any fiber. For $V > 1$, however, the basic moves do not connect all fibers. Consider the case where $V = 3$ and $I_4 = 2$. In this case, the sufficient statistic $\boldsymbol{t}$ is the set of one dimensional marginals

$$\boldsymbol{t} = \{x(i_v) \mid i_v \in \mathcal{I}_v, v = 1, 2, 3\}.$$

From Proposition 1, the set of square-free moves of the Graver basis for three-way complete independence model connects any fiber $\tilde{\mathcal{F}}_{\boldsymbol{t}}$. Table 1 shows the number of square-free moves of the Graver basis for $I_1 \times I_2 \times I_3$ three-way complete independence model computed via 4ti2 (4ti2 team). We see that when the number of levels is larger than two, the sets include moves with degree larger than two. This fact does not necessarily imply that higher degree moves are required to connect every fiber for the three-way complete independence model. However we can give an example, which shows that the degree two moves do not connect all fibers of the three-way complete independence model.

Table 1: The number of square-free moves of the Graver basis for three-way complete independence model

	degree of moves				
$I_1 \times I_2 \times I_3$	2	3	4	5	6
$2 \times 2 \times 2$	12	0	0	0	0
$2 \times 2 \times 3$	33	48	0	0	0
$2 \times 2 \times 4$	64	192	96	0	0
$2 \times 2 \times 5$	105	480	480	0	0
$2 \times 3 \times 3$	90	480	396	0	0
$2 \times 3 \times 4$	174	1632	5436	1152	0
$2 \times 3 \times 5$	285	3840	23220	33120	720
$3 \times 3 \times 3$	243	3438	19008	12312	0

Example 1 (a fiber for $3 \times 3 \times 3$ three-way complete independence model)**.** *Consider the following two zero-one tables* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *in the same fiber of three-way complete independence model.*

$$
\boldsymbol{x} := \quad
\begin{array}{cc}
 & k \\
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
j \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{array}{|ccc|} \hline 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \hline \end{array} \\
 & i = 1
\end{array}
\qquad
\begin{array}{c}
\begin{array}{|ccc|} \hline 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ \hline \end{array} \\
i = 2
\end{array}
\qquad
\begin{array}{c}
\begin{array}{|ccc|} \hline 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ \hline \end{array} \\
i = 3
\end{array}
$$

$$
\boldsymbol{y} := \quad
\begin{array}{cc}
 & k \\
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
j \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{array}{|ccc|} \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ \hline \end{array} \\
 & i = 1
\end{array}
\qquad
\begin{array}{c}
\begin{array}{|ccc|} \hline 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ \hline \end{array} \\
i = 2
\end{array}
\qquad
\begin{array}{c}
\begin{array}{|ccc|} \hline 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ \hline \end{array} \\
i = 3
\end{array}.
$$

The difference of the two tables is

$$
\boldsymbol{z} = \begin{array}{|ccc|} \hline 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \\ \hline \end{array}
\;
\begin{array}{|ccc|} \hline 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ \hline \end{array}
\;
\begin{array}{|ccc|} \hline 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ \hline \end{array}
$$

and we can easily check that $\boldsymbol{z}$ *is a move for the three-way complete independence model.*

Let $\bar{\Delta}$ *be the set of degenerate variables defined in Hara et al. [2007]. Then degree two moves for three-way complete independence model are classified into the following four patterns.*

1. $\bar{\Delta} = \{1, 2, 3\}$

$$
\begin{array}{cc}
 & \begin{array}{cc} i_3 & i_3' \end{array} \\
\begin{array}{c} i_2 \\ i_2' \end{array} & \begin{array}{|cc|} \hline 1 & 0 \\ 0 & -1 \\ \hline \end{array} \\
 & i_1
\end{array},
\qquad
\begin{array}{cc}
 & \begin{array}{cc} i_3 & i_3' \end{array} \\
\begin{array}{c} i_2 \\ i_2' \end{array} & \begin{array}{|cc|} \hline -1 & 0 \\ 0 & 1 \\ \hline \end{array} \\
 & i_1'
\end{array}.
$$

2. $\bar{\Delta} = \{1, 2\}$

$$\underset{i_1}{\begin{array}{c|cc|} & i_3 & i'_3 \\ \hline i_2 & 1 & 0 \\ i'_2 & -1 & 0 \\ \hline \end{array}}, \qquad \underset{i'_1}{\begin{array}{c|cc|} & i_3 & i'_3 \\ \hline i_2 & 0 & -1 \\ i'_2 & 0 & 1 \\ \hline \end{array}}.$$

3. $\bar{\Delta} = \{1, 3\}$

$$\underset{i_1}{\begin{array}{c|cc|} & i_3 & i'_3 \\ \hline i_2 & 1 & -1 \\ i'_2 & 0 & 0 \\ \hline \end{array}}, \qquad \underset{i'_1}{\begin{array}{c|cc|} & i_3 & i'_3 \\ \hline i_2 & -1 & 1 \\ i'_2 & 0 & 0 \\ \hline \end{array}}.$$

4. $\bar{\Delta} = \{2, 3\}$

$$\underset{i_1}{\begin{array}{c|cc|} & i_3 & i'_3 \\ \hline i_2 & 1 & -1 \\ i'_2 & -1 & 1 \\ \hline \end{array}}.$$

However it is easy to check that if we apply any move in this class to $\boldsymbol{x}$ *or* $\boldsymbol{y}$*,* -1 *or* 2 *has to appear. Therefore we cannot apply any degree two moves to both* $\boldsymbol{x}$ *and* $\boldsymbol{y}$*. Hence a degree three move is required to connect this fiber.*

This example indicates that it may be difficult to obtain a set of moves which connects every fiber of the many-facet Rasch model theoretically. As seen in Table 1, the number of square-free moves in the Graver basis is too large even for tables of small sizes and the computational cost of computing the Graver basis for larger table is too high. Hence in practice the use of Graver basis may be inappropriate to implement exact tests for many-facet Rasch model. The implementation of exact tests for many-faceted Rasch model is left as a future task.

3.3 Univariate and bivariate logistic regression

In Hara et al. [2008] we have obtained a subset of Markov basis, which connects all fibers with positive marginals for the response variables in the univariate and the bivariate logistic regression models. We can consider a zero-one table as one slice (corresponding to the "success") of the logistic regression with the sample size equal to one for each combination of covariates. Therefore a connectivity result for logistic regression with positive response variable marginals guarantees the connectivity of zero-one tables.

A frequency vector $\boldsymbol{x}$ and a move $\boldsymbol{z}$ for the Lawrence lifting $\Lambda(A)$ in (1) can be indexed as $\boldsymbol{x}(u, i)$ and $\boldsymbol{z}(u, i)$, where $u = 0, 1$ and i indexes the columns of A. Then the slice $\tilde{\boldsymbol{z}} = \boldsymbol{z}(0, \cdot)$ (i.e. $\tilde{z}(i) = z(0, i)$, $i \in \mathcal{I}$) is a move for A. The response marginal of $\boldsymbol{x}$ is given as $\{t_2(i) = x(0, i) + x(1, i) \mid i \in \mathcal{I}\}$, which corresponds the second block $\boldsymbol{t}_2$ of the sufficient statistic for $\Lambda(A)$:

$$\begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{pmatrix} = \begin{pmatrix} A & 0 \\ E_I & E_I \end{pmatrix} \boldsymbol{x}.$$

Let $\mathcal{B}$ be a set of moves which connects all fibers with positive response variable marginals $\{\mathcal{F}_{\boldsymbol{t}} \mid \boldsymbol{t}_2 > 0\}$. Then by the argument above, the slices of the moves from $\mathcal{B}$ connects all zero-one tables for A, i.e., all fibers $\tilde{\mathcal{F}}_{\boldsymbol{t}}$ in (2). We state this fact in the following lemma.

Lemma 1. *Let $\mathcal{B}$ be a set of moves connecting all fibers with positive response variable marginals $\{\mathcal{F}_{\boldsymbol{t}} \mid \boldsymbol{t}_2 > 0\}$ for $\Lambda(A)$. Then the set of slices $\tilde{\mathcal{B}} = \{\tilde{\boldsymbol{z}} \mid \boldsymbol{z} \in \mathcal{B}\}$ connects all zero-one tables for A.*

Based on this lemma, we have the following results from Proposition 2 and Theorem 3 of Hara et al. [2008].

Proposition 5. *Consider the configuration*

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & I \end{pmatrix}.$$

Let $\boldsymbol{e}_i$ denote the I-dimensional frequency vector with 1 in the i-th cell and zero everywhere else. The set of moves

$$\mathcal{B} = \{\pm(\boldsymbol{e}_{i_1} + \boldsymbol{e}_{i_4} - \boldsymbol{e}_{i_2} - \boldsymbol{e}_{i_3}) \mid 1 \le i_1 < i_2 < i_3 < i_4 \le I, \quad i_2 - i_1 = i_4 - i_3\}$$

connects all zero-one tables for A.

Proposition 6. *Consider the configuration A consisting of $I \times J$ column vectors of the form*

$$\begin{pmatrix} 1 \\ i \\ j \end{pmatrix}, \quad 1 \le i \le I, \ 1 \le j \le J.$$

Let $\boldsymbol{e}_{jk}$ be denote the $I \times J$ frequency vector with 1 at the cell (ij) and and 0 everywhere else. The set of moves $\boldsymbol{z} = (z_{ij})$ satisfying the following conditions

1. *$\boldsymbol{z} = \boldsymbol{e}_{i_1j_1} - \boldsymbol{e}_{i_2j_2} - \boldsymbol{e}_{i_3j_3} + \boldsymbol{e}_{i_4j_4}$;*

2. *$(i_1, j_1) - (i_2, j_2) = (i_3, j_3) - (i_4, j_4)$,*

connects all zero-one tables for A.

3.4 Multivariate logistic regression with binary regressors

As discussed in Hara et al. [2008], it seems to be difficult to obtain a subset of Markov basis, which connects all fibers with positive marginals for the response variables in logistic regression model with more than two covariates. In this section we consider the logistic regression model with three dummy (binary) variables as covariates. Then we can extend the result in Hara et al. [2008].

The model is expressed as

$$\mathrm{logit}(p_{1jkl}) = \log \frac{p_{1jkl}}{1 - p_{1jkl}} = \alpha_0 + \alpha_1 j + \alpha_2 k + \alpha_3 l, \tag{8}$$

where $j = 0, 1$, $k = 0, 1$, $l = 0, 1$. By following Hara et al. [2008], the sufficient statistic is

$$\boldsymbol{t} = \{x_{+jkl}, x_{0j++}, x_{0+k+}, x_{0++l}, j = 0, 1, k = 0, 1, l = 0, 1\}.$$

Hence a move $\boldsymbol{z}$ for (8) satisfies

$$z_{+jkl} = 0, \quad z_{0j++} = 0, \quad z_{0+k+} = 0, \quad z_{0++l} = 0. \tag{9}$$

We introduce the following class of moves.

Definition 1. *Let $\boldsymbol{e}_{jkl} = (e_{ijkl})$ be defined as an integer array with 1 at the cell $(0jkl)$, -1 at the cell $(1jkl)$ and 0 everywhere else. Define $\mathcal{B}$ as the set of moves $\boldsymbol{z} = (z_{ijkl})$ satisfying the following conditions,*

1. $\boldsymbol{z} = \boldsymbol{e}_{j_1k_1l_1} - \boldsymbol{e}_{j_2k_2l_2} - \boldsymbol{e}_{j_3k_3l_3} + \boldsymbol{e}_{j_4k_4l_4}$;
2. $(j_1, k_1, l_1) - (j_2, k_2, l_2) = (j_3, k_3, l_3) - (j_4, k_4, l_4)$.

Theorem 2. *$\mathcal{B}$ connects every fiber satisfying $x_{+jkl} > 0$, $\forall j, k, l$.*

By Lemma 1 we have the following corollary.

Corollary 1. *The set of $i = 0$ slices of $\boldsymbol{z}$ in $\mathcal{B}$ connects $2 \times 2 \times 2$ zero-one tables with fixed one-dimensional marginals.*

Proof of Theorem 2. We prove the theorem by the distance reducing argument in Takemura and Aoki [2005]. The distance in this proof refers to the L_1 distance. Let $\boldsymbol{x} \neq \boldsymbol{y}$ be two tables in the same fiber $\mathcal{F}_{\boldsymbol{t}}$. Define $\boldsymbol{z} := \boldsymbol{x} - \boldsymbol{y}$. Then $\boldsymbol{z}$ is a move.

Since $\boldsymbol{z}$ satisfies (9), the two-way marginal sum $\boldsymbol{z}_{0jk} := \{z_{0jk+}\}_{j=0,1,k=0,1}$ satisfies

$$\boldsymbol{z}_{1jk} = \begin{array}{|cc|} \hline z_{000+} & z_{001+} \\ z_{010+} & z_{011+} \\ \hline \end{array} = \begin{array}{|cc|} \hline 0 & 0 \\ 0 & 0 \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|rr|} \hline a & -a \\ -a & a \\ \hline \end{array}$$

for $a \neq 0$.

Case 1. Suppose that $\boldsymbol{z}_{0jk} = \begin{array}{|cc|} \hline 0 & 0 \\ 0 & 0 \\ \hline \end{array}$. Without loss of generality, we can assume that $z_{0000} > 0$. This implies that $z_{0001} < 0$, $z_{1001} > 0$ and $z_{1000} < 0$. Since $z_{0++0} = 0$, there exist j and k such that $(j, k) \neq (0, 0)$, $z_{0jk0} < 0$ and $z_{1jk0} > 0$. This implies that $z_{0jk1} > 0$ and $z_{1jk1} < 0$. Define

$$\boldsymbol{z}_0 = (0000)(0jk1)(1jk0)(1001) - (0001)(0jk0)(1jk1)(1000).$$

Then $\boldsymbol{x} - \boldsymbol{z}_0 \in \mathcal{F}_{\boldsymbol{t}}$ or $\boldsymbol{y} + \boldsymbol{z}_0 \in \mathcal{F}_{\boldsymbol{t}}$ and the distance is reduced by eight.

Case 2. The case that $\boldsymbol{z}_{1jk} = \begin{array}{|rr|} \hline a & -a \\ -a & a \\ \hline \end{array}$ for $a > 0$.

Without loss of generality we can assume that $z_{0000} > 0$ and $z_{1000} < 0$.

Case 2-1. The case that $z_{0010} < 0$, $z_{0100} < 0$ and $z_{0110} > 0$.
This assumption implies that $z_{1010} > 0$, $z_{1100} > 0$ and $z_{1110} < 0$,

$$\begin{array}{|cc|}\hline z_{0000} & z_{0010} \\ z_{0100} & z_{0110} \\ \hline z_{1000} & z_{101l} \\ z_{1100} & z_{1111} \\ \hline \end{array} = \begin{array}{|cc|}\hline + & - \\ - & + \\ \hline - & + \\ + & - \\ \hline \end{array}.$$

Define a move $\boldsymbol{z}_1$ by

$$\boldsymbol{z}_2 = (0010)(0100)(1000)(1110) - (0000)(0110)(1010)(1100)$$

Then $\boldsymbol{x} - \boldsymbol{z}_0 \in \mathcal{F}_{\boldsymbol{t}}$ or $\boldsymbol{y} + \boldsymbol{z}_0 \in \mathcal{F}_{\boldsymbol{t}}$ and the distance is reduced by eight.

Case 2-2. The case that $z_{0010} < 0$, $z_{0100} < 0$ and $z_{0110} = 0$.
This assumption implies that $z_{1010} > 0$, $z_{1100} > 0$ and $z_{1110} = 0$,

$$\begin{array}{|c|c|}\hline z_{0000} & z_{0010} \\ z_{0100} & z_{0110} \\ \hline z_{1000} & z_{101l} \\ z_{1100} & z_{1111} \\ \hline \end{array} = \begin{array}{|c|c|}\hline + & - \\ - & 0 \\ \hline - & + \\ + & 0 \\ \hline \end{array}.$$

Then the sign patterns of $l = 0$ slices of $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfy either

$$\begin{array}{|c|c|}\hline x_{0000} & x_{0010} \\ x_{0100} & x_{0110} \\ \hline x_{1000} & x_{1010} \\ x_{1100} & x_{1110} \\ \hline \end{array} - \begin{array}{|c|c|}\hline y_{0000} & y_{0010} \\ y_{0100} & y_{0110} \\ \hline y_{1000} & y_{1010} \\ y_{1100} & y_{1110} \\ \hline \end{array} = \begin{array}{|c|c|}\hline + & 0+ \\ 0+ & + \\ \hline 0+ & + \\ + & 0+ \\ \hline \end{array} - \begin{array}{|c|c|}\hline 0+ & + \\ + & + \\ \hline + & 0+ \\ 0+ & 0+ \\ \hline \end{array} \qquad (10)$$

or

$$\begin{array}{|c|c|}\hline x_{0000} & x_{0010} \\ x_{0100} & x_{0110} \\ \hline x_{1000} & x_{1010} \\ x_{1100} & x_{1110} \\ \hline \end{array} - \begin{array}{|c|c|}\hline y_{0000} & y_{0010} \\ y_{0100} & y_{0110} \\ \hline y_{1000} & y_{1010} \\ y_{1100} & y_{1110} \\ \hline \end{array} = \begin{array}{|c|c|}\hline + & 0+ \\ 0+ & 0+ \\ \hline 0+ & + \\ + & + \\ \hline \end{array} - \begin{array}{|c|c|}\hline 0+ & + \\ + & 0+ \\ \hline + & 0+ \\ 0+ & + \\ \hline \end{array}, \qquad (11)$$

where 0+ denotes that the cell count is nonnegative. In the case of (10), we can apply $\boldsymbol{z}_2$ to $\boldsymbol{x}$ and the distance is reduced by four. In the case of (11), we can apply $-\boldsymbol{z}_{2a}$ to $\boldsymbol{y}$ and the distance is reduced by four.

More generally, if $\boldsymbol{z}$ has either of the following patterns of signs,

$$\begin{array}{|c|c|}\hline z_{i0kl} & z_{i0k'l'} \\ z_{i1kl} & z_{i1k'l'} \\ \hline z_{i'0kl} & z_{i'0k'l'} \\ z_{i'1kl} & z_{i'1k'l'} \\ \hline \end{array} = \begin{array}{|c|c|}\hline + & - \\ - & 0 \\ \hline - & + \\ + & 0 \\ \hline \end{array}, \quad \begin{array}{|c|c|}\hline + & 0 \\ - & + \\ \hline - & 0 \\ + & - \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|c|c|}\hline + & - \\ 0 & + \\ \hline - & + \\ 0 & - \\ \hline \end{array}$$

for $i \neq i'$ and $(k,l) \neq (k',l')$, we can show that the distance is reduced by a move in $\mathcal{B}$ in the similar way.

Case 2-3. In the case where

$$\begin{array}{|cc|}\hline z_{ij0l} & z_{ij1l} \\ \hline z_{ij'0l'} & z_{ij'1l'} \\ \hline z_{i'j0l} & z_{i'j1l} \\ \hline z_{i'j'0l'} & z_{i'j'1l'} \\ \hline \end{array} = \begin{array}{|cc|}\hline + & - \\ \hline - & 0 \\ \hline - & + \\ \hline + & 0 \\ \hline \end{array}, \quad \begin{array}{|cc|}\hline + & 0 \\ \hline - & + \\ \hline - & 0 \\ \hline + & - \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|cc|}\hline + & - \\ \hline 0 & + \\ \hline - & + \\ \hline 0 & - \\ \hline \end{array}$$

for $i \neq i'$ and $(j, l) \neq (j', l')$, we can show that distance is reduced by a move in $\mathcal{B}$ in the same way as Case 2-2.

Case 2-4. In the case where $z_{0010} < 0$, $z_{0100} \geq 0$ and $z_{0110} \geq 0$.
Since $z_{010+} < 0$, we have $z_{0101} < 0$. If $z_{0001} > 0$, we can reduce the distance by a move in $\mathcal{B}$ from Case 2-2. In the same way, if $z_{0111} > 0$, we can reduce the distance from Case 2-1. We assume that $z_{0001} \leq 0$ and $z_{0111} \leq 0$. Since $z_{0++1} = 0$, we have $z_{0011} > 0$. Then we can reduce the distance from Case 2-2.

In the case where $z_{0010} \geq 0$, $z_{0100} < 0$ and $z_{0110} \geq 0$, the proof is similar.

Case 2-5. In the case where $z_{0010} < 0$, $z_{0100} \geq 0$ and $z_{0110} < 0$.
Since $z_{010+} < 0$ and $z_{011+} > 0$, we have $z_{0101} < 0$ and $z_{0101} > 0$. Hence we can reduce the distance from Case 2-3.

In the case where $z_{0010} \geq 0$ and $z_{0110} < 0$, the proof is similar.

Case 2-6. In the case where $z_{0010} < 0$, $z_{0100} < 0$ and $z_{0110} < 0$.
Since $z_{011+} > 0$, we have $z_{0101} > 0$. If $z_{0011} \leq 0$ or $z_{0101} \leq 0$, we can reduce the distance from Case 2-2 and 2-3.

Assume that $z_{0011} > 0$ and $z_{0101} > 0$. Since $z_{0++1} = 0$, we have $z_{0001} < 0$. Then we can reduce the distance from Case 2-2.

$\square$

3.5 Latin squares and zero-one tables for no-three-factor-interaction models

Zero-one tables also appear quite often in the form of incidence matrices for combinatorial problems. Here as an example we consider Latin squares. When the symbols of an $n \times n$ Latin square is considered as coordinates of the third axis (sometimes called the orthogonal array representation of a Latin square), it is a particular element of a fiber for the $n \times n \times n$ no-three-factor-interaction model with all two-dimensional marginals (line sums) equal to 1. One of the reasons to consider a Markov basis for Latin squares is to generate a Latin square randomly. Fisher and Yates [1934] advocated to choose a Latin square randomly from the set of Latin squares. Jacobson and Matthews [1996] gave a Markov basis for the set of $n \times n$ Latin squares.

Because the set of Latin squares is just a particular fiber, it may be the case that a minimal set of moves connecting all Latin squares is smaller than the set of moves connecting all zero-one tables. This is indeed the case as we show for the simple case of $n = 3$. We first present a connectivity result for $3 \times 3 \times 3$ zero-one tables with all line sums fixed.

Let $\boldsymbol{z} = \{z_{ijk}\}_{i,j,k=1,2,3}$ be a move for $3 \times 3 \times 3$ no-thee-factor-interaction model. From Diaconis and Sturmfels [1998] and Aoki and Takemura [2003] the minimal Markov basis consists of basic moves such as

$$\boldsymbol{z} = \begin{array}{|ccc|}\hline 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline\end{array} \tag{12}$$

and degree 6 moves such as

$$\boldsymbol{z} = \begin{array}{|ccc|}\hline 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline\end{array}\ . \tag{13}$$

However these moves do not connect zero-one tables of the $3 \times 3 \times 3$ no-three-factor-interaction model. We need the following type of degree 9 move, which corresponds to the difference of two Latin squares.

$$\boldsymbol{z} = \begin{array}{|ccc|}\hline 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ \hline\end{array} \quad \begin{array}{|ccc|}\hline -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ \hline\end{array}\ . \tag{14}$$

Proposition 7. *The set of basic moves (12), degree 6 moves (13) and degree 9 moves (14) forms a Markov basis for* $3 \times 3 \times 3$ *zero-one tables for the no-three-factor-interaction model.*

Proof. Consider any line sum, such as $0 = z_{+11} = z_{111} + z_{211} + z_{311}$ of a move $\boldsymbol{z}$. If $(z_{111}, z_{211}, z_{311}) \neq (0, 0, 0)$, then we easily see that $\{z_{111}, z_{211}, z_{311}\} = \{-1, 0, 1\}$. By a similar consideration as in Aoki and Takemura [2003], each i- or j- or k-slice is either a loop of degree two or loop of degree three, such as

$$\begin{array}{|ccc|}\hline 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ \hline\end{array} \quad \text{or} \quad \begin{array}{|ccc|}\hline 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ \hline\end{array}\ . \tag{15}$$

Now we consider two cases: 1) there exists a slice with a loop of degree two, or 2) all slices are loops of degree three.

Case 1. Without loss of generality, we can assume that the $i = 1$ slice of $\boldsymbol{z}$ is the loop of degree two in (15). Then we can further assume that $z_{211} = -1$ and $z_{311} = 0$. Now suppose that $z_{222} = -1$. If $z_{212} = 1$ or $z_{221} = 1$, then this constitutes a strong crossing pattern of Condition 1 and we can reduce $|\boldsymbol{z}|$ by a basic move. This implies $z_{212} = z_{221} = 0$. But then $z_{213} = z_{223} = 1$ and this contradicts the pattern of $\{z_{213}, z_{223}, z_{233}\} = \{-1, 0, 1\}$.

By the above consideration we have $z_{222} = 0$ and $z_{322} = -1$. By a similar consideration for the cells z_{i12} and z_{i21}, $i = 1, 2, 3$, we easily see that $\boldsymbol{z}$ is of the form

$$\begin{array}{|ccc|}\hline 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \\ \hline \end{array},$$

which is a degree 6 move.

Case 2. It is easily seen that the only case where degree 6 moves can not be applied is of the form of the move of degree 9 in (14). This proves that connectivity is guaranteed if we add degree 9 moves.

We also want to show that degree 9 moves are needed for connectivity. Consider

$$\boldsymbol{x} = \begin{array}{|ccc|}\hline 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ \hline \end{array} \quad \begin{array}{|ccc|}\hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline \end{array}.$$

By a simple program it is easily checked that if we apply any basic move or any move of degree 6 to $\boldsymbol{x}$, -1 or 2 has to appear. Hence degree 9 moves are required to connect zero-one tables.

□

Now consider 3×3 Latin squares. It is well-known that there is only one isotopy class of 3×3 Latin squares (Chapter III of Colbourn and Dinitz [2007]), i.e., all 3×3 Latin squares are connected by the action of the direct product $S_3 \times S_3 \times S_3$ of the symmetric group S_3 , which corresponds to permutation of levels of a factor of three-way contingency tables. Note that the symmetric group is generated by transpositions and a transposition corresponds to a move of degree 6 in (13). Therefore we have the following fact: *the set of 3×3 Latin squares in the orthogonal array representation is connected by the set of moves of degree 6 in (13)*. In view of Proposition 7, we see that we do not need basic moves nor degree 9 moves for connecting 3×3 Latin squares.

There are two isotopy classes for 4×4 Latin squares (1.18 of III.1.3 of Colbourn and Dinitz [2007]) and representative elements of these two classes are connected by a basic move. Transposition of two levels for a factor corresponds to a degree 8 move of the following form.

$$\boldsymbol{z} = \begin{array}{|cccc|}\hline 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ \hline \end{array} \quad \begin{array}{|cccc|}\hline -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ \hline \end{array} \quad \begin{array}{|cccc|}\hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \end{array} \quad \begin{array}{|cccc|}\hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \end{array}.$$

Therefore the set of 4×4 Latin squares is connected by the set of basic moves and moves of degree 8 of the above form. We can apply a similar consideration to the celebrated result of 22 isotopy classes of 6×6 Latin squares derived by Fisher and Yates [1934].

4 Concluding remarks

In this paper we discussed Markov bases for tables with zero-one entries. We derived several general results, where a particular subset of the Graver basis connects zero-one tables. However, in general, we found that a Markov basis for zero-one tables is difficult and we need separate arguments for each model. We obtained Markov bases for zero-one tables for some common models of contingency tables.

Rapallo and Yoshida [2009] gave some results for contingency tables with bounded entries, in particular for the case of two-way tables with structural zeros. Zero-one tables is a particular case of contingency tables with bounded entries. If the bound is large enough, compared to the sample size of a particular fiber, it seems that the bound is not binding. In this sense the bound of 1 in our case seems to be most stringent. On the other hand, our proof of Proposition 7 suggests that Markov basis for zero-one tables may have a simple structure. It is an interesting problem how Markov bases behave as we vary the upper bound for the cells.

In Section 3.5 we considered Latin squares. It is of interest to consider other combinatorial designs, such as sudoku. Markov basis for sudoku is considered in Fontana and Rogantin [2009]. Sei et al. [2009] discuss the invariance structure of sudoku. It is a challenging problem to derive a Markov basis for the ordinary $3 \times 3 \times 3 \times 3$ sudoku.

References

4ti2 team. 4ti2—a software package for algebraic, geometric and combinatorial problems on linear spaces. Available at www.4ti2.de.

Satoshi Aoki and Akimichi Takemura. Minimal basis for a connected Markov chain over $3 \times 3 \times K$ contingency tables with fixed two-dimensional marginals. *Aust. N. Z. J. Stat.*, 45(2):229–249, 2003. ISSN 1369-1473.

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.

R. Basturk. Applying the many-facet Rasch model to evaluate powerpoint presentation performance in higher education. *Meas. Phys. Educ. Exerc. Sci.*, 33(4):431–444, 2008.

G. Besag and P. Clifford. Generalized Monte Carlo significance tests. *Biometrika*, 76: 633–642, 1989.

Yuguo Chen and Dylan Small. Exact tests for the Rasch model via sequential importance sampling. *Psychometrika*, 70(1):11–30, 2005. ISSN 0033-3123.

G. W. Cobb and Y.-P. Chen. An application of markov chain monte carlo to community ecology. *American Mathematical Monthly*, 110:265–288, 2003.

Charles J. Colbourn and Jeffrey H. Dinitz, editors. *Handbook of combinatorial designs*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, second edition, 2007. ISBN 978-1-58488-506-1; 1-58488-506-8.

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

R.A. Fisher and F. Yates. The 6×6 Latin squares. *Proceedings of the Cambridge Philosophical Society*, 30:492–507, 1934.

Roberto Fontana and Maria-Piera Rogantin. Indicator function and sudoku designs, 2009. To appear in *Algebraic and Geometric Methods in Statistics* dedicated to Professor Giovanni Pistone, Cambridge University Press (P. Gibilisco, E. Riccomagno, M.P. Rogantin and H.P. Wynn, eds.).

C. A. W. Glas and N. D. Verhelst. Testing the rasch model. In *Rasch Models: Their Foundations, Recent Developments and Applications*, pages 69–95. Springer-Verlag, New York, 1995.

Hisayuki Hara, Satoshi Aoki, and Akimichi Takemura. Minimal and minimal invariant markov bases of decomposable models for contingency tables, 2007. arXiv:math/0701429v2. To appear in *Bernoulli*.

Hisayuki Hara, Akimichi Takemura, and Ruriko Yoshida. On connectivity of fibers with positive marginals in multiple logistic regression, 2008. arXiv:0810.1793v1.

Mark T. Jacobson and Peter Matthews. Generating uniformly distributed random Latin squares. *J. Combin. Des.*, 4(6):405–437, 1996. ISSN 1063-8539.

J. M. Linacre. *Many-facet Rasch Measurement*. MESA Press, Chicago, 1989.

J. M. Linacre. Constructing measurement with a many-facet rasch model. In M. Wilson, editor, *Objective Measurement: Theory into Practice*, pages 129–144. Northwood, NJ: Ablex Pub. Corp., 1994.

I. Ponocny. Nonparametric goodness-of-fit tests for the rasch model. *Psychometrika*, 66: 437–460, 2001.

Fabio Rapallo and Ruriko Yoshida. Markov bases and subbases for bounded contingency tables, 2009. arXiv:0905.4841v2.

G. Rasch. *Probabilistic Models for Some Intelligence and Attainment Tests*. University of Chicago Press, Chicago, 1980.

H. J. Ryser. Combinatorial properties of matrices of zeros and ones. *Canad. J. Math.*, 9: 371–377, 1957.

Tomonari Sei, Satoshi Aoki, and Akimichi Takemura. Perturbation method for determining group of invariance of hierarchical models. *Advances in Applied Mathematics*, 2009. doi:10.1016/j.aam.2009.02.005.

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

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.

Akimichi Takemura and Satoshi Aoki. Distance reducing Markov bases for sampling from a discrete sample space. *Bernoulli*, 11(5):793–813, 2005.

K. Yamaguchi. Comparing many-facet rasch model and anova model: analysis of ratings of essays. *Japanese Journal of Educational Psychology*, 47:383–392, 1999.

Weimo Zhu, Catherine D. Ennis, and Ang Chen. Many-faceted Rach modeling expert judgment in test development. *Meas. Phys. Educ. Exerc. Sci.*, 2(1):21–39, 1998.