MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

The largest group of invariance for Markov bases and toric ideals

Satoshi AOKI and Akimichi TAKEMURA

METR 2005–14 June 2005

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.i.u-tokyo.ac.jp/mi/mi-e.htm

The largest group of invariance for Markov bases and toric ideals

Satoshi AOKI and Akimichi TAKEMURA

Graduate School of Information Science and Technology
University of Tokyo, Tokyo, Japan

SUMMARY

We define the largest group of invariance for a given toric ideal and the associated Markov bases. Reduction by invariance leads to a concise description of an invariant Markov basis and a sampling scheme in terms of the group and a list of representative elements from the orbits of the Markov basis. We also give explicit forms of the largest group of invariance for several standard statistical problems.

Key words and phrases: computational algebraic statistics, exact tests, group action, MCMC, sufficient statistic.

1 Introduction

Since the publication of Diaconis and Sturmfels (1998), the new field of computational algebraic statistics has been developing rapidly. See for example the papers in the upcoming special issue (2005) on computational algebraic statistics of *Journal of Symbolic Computation*. Diaconis and Sturmfels (1998) defined the notion of Markov basis for constructing a connected Markov chain for sampling from a conditional distribution over a discrete sample space and proved the fundamental fact that a Markov basis corresponds to a set of binomial generators of a toric ideal. See also Sturmfels (1995) and Dinwoodie (1998). This enables application of Gröbner basis technology for obtaining Markov bases in a general setting and Markov bases for various problems have been obtained in the form of reduced Gröbner bases.

However Gröbner basis computation depends on a particular term order and the symmetry inherent in the problems tends to be ignored. As demonstrated in Aoki and Takemura (2003a) and Aoki and Takemura (2003c), for some problems with enough symmetry, elementary arguments exploiting the symmetry as much as possible lead to explicit description of symmetric and minimal Markov bases without relying on Gröbner basis computation.

In the case of standard multiway contingency tables, the symmetry among the cells, or, equivalently, among the indeterminates of a polynomial ring, is formalized as the action of a group, which is a subgroup of a symmetric group permuting the cells. In Aoki and Takemura (2003b), we considered an action of a direct product of symmetric groups on each axis of multiway contingency table. If the categories of each axis do not have any order relations among them, it is natural to consider such a group action (e.g. Section 8C of Diaconis, 1988).

In this paper we consider a general finite sample space, not restricted to standard multiway contingency tables. For a given problem of obtaining a Markov basis, or for a given toric ideal, we define the largest group of invariance for the problem. In the following we simply call it *invariance group* for the problem. Given a toric ideal, we can often guess the form of the invariance group from the obvious symmetry in the problem. The hard part is actually proving that the candidate group is the largest group, which leaves the problem invariant.

The construction of this paper is as follows. For the rest of this introduction we discuss two motivating examples of contingency tables for considering the invariance group. In Section 2, we give some notations on contingency tables, toric ideals, symmetric group and its action. In Section 3, we define the invariance group for a given toric ideal and discuss its relations to our previous works. We give the structures and interpretations of the invariance group for some standard statistical problems of contingency tables in Section 4. In Section 5, we give lists of representative moves of minimal invariant Markov bases for two-way and three-way problems considered in Section 4. Finally in Section 6, we give some discussions.

1.1 Motivating examples

Here we discuss two simple examples of contingency tables from Aoki and Takemura (2003b) and Takemura and Aoki (2004), for motivating consideration of the general invariance group of this paper. Readers might skip them, although we believe that these examples are helpful for understanding the definition of the invariance group in Section 3.

Example 1 Consider $2 \times 2 \times 2$ contingency tables with fixed one-dimensional marginals. They are relevant for exact tests of the complete independence model, i.e., $p_{ijk} = \alpha_i \beta_j \gamma_k$, where p_{ijk} is the probability of the cell ijk. In considering a Markov basis for this problem, we encounter the set

$$\left\{ \boldsymbol{x} = \{x_{ijk}\}_{1 \le i,j,k \le 2} \;\middle|\; \sum_{i,j} x_{ijk} = \sum_{i,k} x_{ijk} = \sum_{j,k} x_{ijk} = 1,\ x_{ijk} \in \mathbb{N},\ 1 \le i,j,k \le 2 \right\}, \tag{1}$$

where $\mathbb{N} = \{0, 1, 2, \ldots\}$. x_{ijk} is the frequency of the cell ijk, and the set of frequencies $\boldsymbol{x} = \{x_{ijk}\}$ is a contingency table. (We give a precise definition in Section 2.) The diophantine equation (1) has four solutions. We write the solution by the cells of positive frequencies as

$$\{(111)(222),\ (112)(221),\ (121)(212),\ (122)(211)\}.$$

For example, (111)(222) means the solution $\boldsymbol{x} = \{x_{ijk}\}$, where

$$x_{ijk} = \begin{cases} 1, & ijk = 111 \text{ or } 222, \\ 0, & \text{otherwise.} \end{cases}$$

To construct a Markov basis for this problem, we have to connect the above four elements. There are many ways of connecting the four elements. The reduced Gröbner basis with respect to the graded reverse lexicographic order consists of the following three moves.

$$(121)(212) - (111)(222),\ (122)(211) - (111)(222),\ (112)(221) - (111)(222). \tag{2}$$

In Aoki and Takemura (2003b) we considered symmetry with respect to permuting the levels of each axis and showed that a union of each two of the three orbits

$$\begin{gathered} \{(111)(222) - (112)(221),\ (121)(212) - (122)(211)\}, \\ \{(111)(222) - (121)(212),\ (112)(221) - (122)(211)\}, \\ \{(111)(222) - (122)(211),\ (112)(221) - (121)(212)\} \end{gathered}$$

connects the four elements. Therefore the list of representative elements for two orbits such as

$$(111)(222) - (112)(221),\ (111)(222) - (121)(212) \tag{3}$$

is sufficient to describe a Markov basis. Note that (3) is more concise than (2). However in Aoki and Takemura (2003b) we did not consider permuting the axes. In the $2 \times 2 \times 2$ case, since the number of categories is common to the axis, we can permute the axes as well. If we consider invariance with respect to this larger group, then a single representative element such as

$$(111)(222) - (112)(221) \tag{4}$$

is sufficient to describe an invariant Markov basis.

Example 2 Consider the Hardy-Weinberg model for four alleles. Here we omit the background material on the model, but just consider the following diophantine equations for $x_{ij} \in \mathbb{N},\ 1 \leq i \leq j \leq 4$,

$$\begin{aligned} 2x_{11} + x_{12} + x_{13} + x_{14} &= 1, \\ 2x_{22} + x_{12} + x_{23} + x_{24} &= 1, \\ 2x_{33} + x_{13} + x_{23} + x_{34} &= 1, \\ 2x_{44} + x_{14} + x_{24} + x_{34} &= 1. \end{aligned} \tag{5}$$

As is shown in Takemura and Aoki (2004), there are three solutions for the equation (5) as

$$\{(12)(34),\ (13)(24),\ (14)(23)\},$$

and we have to connect these three elements to construct a Markov basis for performing exact tests of the Hardy-Weinberg proportions. This can be achieved for example by

$$(12)(34) - (13)(24),\ (12)(34) - (14)(23).$$

There are three ways to connect the three elements by two edges, each of which corresponds to a minimal Markov basis. The group considered in Aoki and Takemura (2003b) cannot be applied in this case, since the contingency table $\boldsymbol{x} = \{x_{ij}\}_{1 \leq i \leq j \leq 4}$ is of an upper triangular form. However it is clear that this problem has the symmetry with respect to a simultaneous permutation of the levels (equivalently, a permutation of alleles), which can be handled by the invariance group of this paper. (See Section 4.5.)

2 Preliminaries

In this section, we give some notations on contingency tables, toric ideals, symmetric group and its action.

2.1 Notations on contingency tables and toric ideals

Let $\mathcal{I}$ be a finite set with $p = |\mathcal{I}|$ elements. In this paper $\mathcal{I}$ is a general finite set. However with contingency tables in mind, we call an element of $\mathcal{I}$ a *cell* and denote it by $\boldsymbol{i} \in \mathcal{I}$. $\boldsymbol{i}$ is often a multi-index $\boldsymbol{i} = i_1 \cdots i_m$. A non-negative integer $x_{\boldsymbol{i}} \in \mathbb{N} = \{0, 1, 2, \ldots\}$ denotes the *frequency* of a cell $\boldsymbol{i}$. The set of frequencies is called a *contingency table* and denoted as $\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}}$. With an appropriate ordering of the cells, we treat a contingency table $\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}} \in \mathbb{N}^p$ as a p-dimensional column vector of non-negative integers. Note that a contingency table can also be considered as a function from $\mathcal{I}$ to $\mathbb{N}$ defined as $\boldsymbol{i} \mapsto x_{\boldsymbol{i}}$. The L_1-norm of $\boldsymbol{x} \in \mathbb{N}^p$ is called the *sample size* of $\boldsymbol{x}$ and denoted as $|\boldsymbol{x}| = \sum_{\boldsymbol{i}\in\mathcal{I}} x_{\boldsymbol{i}}$. Let $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ and let $\boldsymbol{a}_j \in \mathbb{Z}^p, j = 1,\ldots,\nu$, denote fixed column vectors consisting of integers. We define a ν-dimensional column vector $\boldsymbol{t} = (t_1,\ldots,t_\nu)' \in \mathbb{Z}^\nu$ as $t_j = \boldsymbol{a}_j'\boldsymbol{x}, j = 1,\ldots,\nu$. Here $'$ denotes the transpose of a vector or a matrix. We also define a $\nu \times p$ matrix A, with its j-th row being $\boldsymbol{a}_j'$, given by

$$A = \begin{bmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_\nu' \end{bmatrix}. \tag{6}$$

Then the ν-dimensional column vector $\boldsymbol{t}$ is written as $\boldsymbol{t} = A\boldsymbol{x}$. In the following the set of $\boldsymbol{t}$ is denoted as $\mathcal{T} = \{\boldsymbol{t} \mid \boldsymbol{t} = A\boldsymbol{x},\ \boldsymbol{x} \in \mathbb{N}^p\} = A\mathbb{N}^p \subset \mathbb{Z}^\nu$.

In typical situations of a statistical theory, $\boldsymbol{t}$ is the *sufficient statistic* for the nuisance parameter, and the set of $\boldsymbol{x}$'s for a given $\boldsymbol{t}$, $\mathcal{F}_{\boldsymbol{t}} = \{\boldsymbol{x} \in \mathbb{N}^p \mid A\boldsymbol{x} = \boldsymbol{t}\}$, is considered for performing *similar tests*. For the case of the independence model of the two-way contingency tables, for example, $\boldsymbol{t}$ is the row sums and column sums of $\boldsymbol{x}$, and $\mathcal{F}_{\boldsymbol{t}}$ is the set of $\boldsymbol{x}$'s with the same row sums and column sums to $\boldsymbol{t}$. Following Sturmfels (1995), we call $\mathcal{F}_{\boldsymbol{t}}$ a $\boldsymbol{t}$*-fiber* in this paper. As in Takemura and Aoki (2004), we assume that the toric ideal is homogeneous (Chapter 4 of Sturmfels, 1995, Section 4.1 of Hibi, 2003), i.e., the p-dimensional column vector $(1,\ldots,1)'$ is a rational linear combination of $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_\nu$. Under this assumption, all elements in $\mathcal{F}_{\boldsymbol{t}}$ have the same sample size. Therefore the sample size of $\boldsymbol{t}$ is well-defined as $|\boldsymbol{t}| = |\boldsymbol{x}|$ for $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{t}}$.

The set of $\boldsymbol{t}$-fibers gives a decomposition of $\mathbb{N}^p$. An important observation is that $\boldsymbol{t}$-fiber depends on a given A only through its kernel, $\ker(A)$. For different A's with the same kernel, the sets of $\boldsymbol{t}$-fibers are the same. In fact, if we define

$$\boldsymbol{x}_1 \sim \boldsymbol{x}_2 \ \Leftrightarrow\ \boldsymbol{x}_1 - \boldsymbol{x}_2 \in \ker(A),$$

this relation is an equivalence relation and $\mathbb{N}^p$ is partitioned into disjoint equivalence classes. The set of $\boldsymbol{t}$-fibers is simply the set of these equivalence classes. Furthermore, $\boldsymbol{t}$ may be considered as labels of these equivalence classes. In statistical theory, this non-uniqueness of the matrix A corresponds to the non-uniqueness of the sufficient statistic. For example, in the case of multiway contingency tables, it is often advantageous to keep some linearly dependent rows in the matrix A for the sake of symmetry. Note that linearly dependent rows does not alter $\ker(A)$.

The main object considered in this paper is the integer lattice in $\ker(A)$, i.e., $\mathbb{Z}^p \cap \ker(A)$. A p-dimensional column vector of integers $\boldsymbol{z} = \{z_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}} \in \mathbb{Z}^p$ is called a *move* if it is in the kernel of A, i.e., $A\boldsymbol{z} = \boldsymbol{0}$. For a move $\boldsymbol{z}$, the positive part $\boldsymbol{z}^+ = \{z_{\boldsymbol{i}}^+\}_{\boldsymbol{i}\in\mathcal{I}}$ and the negative part $\boldsymbol{z}^- = \{z_{\boldsymbol{i}}^-\}_{\boldsymbol{i}\in\mathcal{I}}$ are defined by $z_{\boldsymbol{i}}^+ = \max(z_{\boldsymbol{i}}, 0)$, $z_{\boldsymbol{i}}^- = \max(-z_{\boldsymbol{i}}, 0)$, respectively. Then $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$ and $\boldsymbol{z}^+, \boldsymbol{z}^- \in \mathbb{N}^p$. Moreover, $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ are in the same $\boldsymbol{t}$-fiber, i.e., $\boldsymbol{z}^+, \boldsymbol{z}^- \in \mathcal{F}_{\boldsymbol{t}}$ for $\boldsymbol{t} = A\boldsymbol{z}^+ = A\boldsymbol{z}^-$. We define the *degree* of $\boldsymbol{z}$ as the sample size of $\boldsymbol{z}^+$ (or $\boldsymbol{z}^-$) and denote it by $\deg(\boldsymbol{z}) = |\boldsymbol{z}^+| = |\boldsymbol{z}^-|$. In the following we denote the set of moves (for a given A) by

$$\mathcal{M} = \mathcal{M}_A = \mathbb{Z}^p \cap \ker(A).$$

Following Diaconis and Sturmfels (1998), the above notations on contingency tables can be translated to the objects on polynomial rings. Let $\boldsymbol{u} = \{u_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}}$ be the set of indeterminates and let $k[\boldsymbol{u}]$ denote the polynomial ring in the indeterminates $\boldsymbol{u}$ over a field k. Then a contingency table $\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathcal{I}}$ is specified as a monomial $\boldsymbol{u}^{\boldsymbol{x}} = \prod_{\boldsymbol{i}\in\mathcal{I}} u_{\boldsymbol{i}}^{x_{\boldsymbol{i}}} \in k[\boldsymbol{u}]$ and a move $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^- \in \mathcal{M}$ is specified as a binomial $\boldsymbol{u}^{\boldsymbol{z}^+} - \boldsymbol{u}^{\boldsymbol{z}^-} = \prod_{\boldsymbol{i}\in\mathcal{I}} u_{\boldsymbol{i}}^{z_{\boldsymbol{i}}^+} - \prod_{\boldsymbol{i}\in\mathcal{I}} u_{\boldsymbol{i}}^{z_{\boldsymbol{i}}^-} \in k[\boldsymbol{u}]$. The toric ideal I_A, which is the main object to be considered, is the ideal generated by all the homogeneous binomials corresponding to all the moves $\boldsymbol{z} \in \mathcal{M}$. Hereafter, we write $\mathrm{Mon}(\boldsymbol{u})$ or $\mathrm{Bin}(\boldsymbol{u})$ to denote the set of monomials $\boldsymbol{u}^{\boldsymbol{x}}$ and the set of homogeneous binomials $\boldsymbol{u}^{\boldsymbol{z}^+} - \boldsymbol{u}^{\boldsymbol{z}^-}$.

We also use a concise notation of contingency tables and moves (or monomials and binomials) of small sample sizes (or degrees) by specifying locations of their non-zero cells. For example, $\boldsymbol{x} \in \mathbb{N}^p$ is denoted as

$$\boldsymbol{x} = (\boldsymbol{i}_1)\cdots(\boldsymbol{i}_n),$$

where $n = |\boldsymbol{x}|$ and $\boldsymbol{i}_1, \ldots, \boldsymbol{i}_n$ are the cells of positive frequencies of $\boldsymbol{x}$. In the case of $x_{\boldsymbol{i}} > 1$, $\boldsymbol{i}$ is repeated $x_{\boldsymbol{i}}$ times. Similarly, we express a move $\boldsymbol{z}$ of degree n as

$$\boldsymbol{z} = (\boldsymbol{i}_1)\cdots(\boldsymbol{i}_n) - (\boldsymbol{j}_1)\cdots(\boldsymbol{j}_n)$$

where $\boldsymbol{i}_1, \ldots, \boldsymbol{i}_n$ are the cells of positive frequencies of $\boldsymbol{z}$ and $\boldsymbol{j}_1, \ldots, \boldsymbol{j}_n$ are the cells of negative frequencies of $\boldsymbol{z}$. For example, consider the independence model of 3×3 contingency tables with $\mathcal{I} = \{ij \mid 1 \le i, j \le 3\}$. The frequency vector $\boldsymbol{x}$ is defined by ordering cell frequencies lexicographically as

$$\boldsymbol{x} = (x_{11}, x_{12}, \ldots, x_{33})',$$

and a matrix A is written as

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix},$$

for example. Note that the ordering of columns of A is determined in accordance with the ordering of frequencies in $\boldsymbol{x}$. In our notation, a move

$$\boldsymbol{z} = \begin{array}{|ccc|} \hline 2 & -2 & 0 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \\ \hline \end{array}$$

is expressed as

$$\boldsymbol{z} = (2, -2, 0, -1, 0, 1, -1, 2, -1)',$$

$$\boldsymbol{z} = u_{11}^2 u_{23} u_{32}^2 - u_{12}^2 u_{21} u_{31} u_{33}$$

or

$$\boldsymbol{z} = (11)(11)(23)(32)(32) - (12)(12)(21)(31)(33).$$

2.2 Symmetric group and its action on contingency tables and binomials

First we give a brief list of definitions and notations of a group action. Let a group G act on a set $\mathcal{U}$. $G(u) = \{gu \mid g \in G\}$ is the orbit through u. For a subset $\mathcal{V}$ of $\mathcal{U}$, $G(\mathcal{V}) = \{gu \mid u \in \mathcal{V},\ g \in G\}$. $\mathcal{U}/G$ denotes the orbit space, i.e. the set of orbits. $G_u = \{g \mid gu = u\}$ denotes the stabilizer (isotropy subgroup) of u in G. If G acts on $\mathcal{U}$, the action of G on the set of functions f on $\mathcal{U}$ is induced by $(gf)(u) = f(g^{-1}u)$. Let $h : \mathcal{U} \to \mathcal{Y}$ be a surjection. If $h(u) = h(u') \Rightarrow h(gu') = h(gu), \forall g \in G$, then the action of G on $\mathcal{Y}$ is induced by defining $gy = h(gu)$, where $y = h(u)$.

Following Seress (2003), for a subset $\mathcal{V}$ of $\mathcal{U}$, $G_{(\mathcal{V})} = \{g \mid gu = u, \forall u \in \mathcal{V}\}$ denotes the pointwise stabilizer of $\mathcal{V}$. On the other hand $G_{\mathcal{V}} = \{g \mid g\mathcal{V} = \mathcal{V}\}$ denotes the setwise stabilizer of $\mathcal{V}$. For avoiding confusion, in the following we give a verbal description when pointwise or setwise stabilizers are defined.

In this paper we consider the action of the symmetric group S_p, $p = |\mathcal{I}|$, on the set of cells $\mathcal{I}$:

$$S_p \times \mathcal{I} \ni (g, \boldsymbol{i}) \mapsto g(\boldsymbol{i}) \in \mathcal{I}.$$

Each $g \in S_p$ can be identified with a $p \times p$ permutation matrix $P_g = \{p_{\boldsymbol{ij}}\} = \{\delta_{\boldsymbol{i},g(\boldsymbol{j})}\}$, where δ is the Kronecker's delta. Then $P_{g_1 \cdot g_2} = P_{g_1} P_{g_2}$ for $g_1, g_2 \in S_p$ and $P_{g^{-1}} = P'_g$. The identity matrix of order p is denoted by E_p. Therefore $P_e = E_p$ for the unit element $e \in S_p$. Hereafter, we occasionally write each element simply as $P \in S_p$, which means that $P = P_g$ for $g \in S_p$.

Since a contingency table $\boldsymbol{x} = \{x_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathcal{I}}$ can be considered as a function from $\mathcal{I}$ to $\mathbb{N}$: $\boldsymbol{i} \mapsto x_{\boldsymbol{i}}$, the action of S_p on $\mathbb{N}^p$, the set of contingency tables, is induced as

$$S_p \times \mathbb{N}^p \ni (g, \boldsymbol{x}) \mapsto g\boldsymbol{x} = \{x_{g^{-1}(\boldsymbol{i})}\} = P_g \boldsymbol{x} \in \mathbb{N}^p.$$

Similarly S_p acts on $\mathbb{Z}^p$, the set of integer arrays, by

$$S_p \times \mathbb{Z}^p \ni (g, \boldsymbol{z}) \mapsto g\boldsymbol{z} = g\boldsymbol{z}^+ - g\boldsymbol{z}^- = P_g \boldsymbol{z}^+ - P_g \boldsymbol{z}^- \in \mathbb{Z}^p.$$

Considering the correspondence between the contingency tables and the monomials, S_p acts on $\mathrm{Mon}(\boldsymbol{u}) \subset k[\boldsymbol{u}]$ by

$$S_p \times \mathrm{Mon}(\boldsymbol{u}) \ni (g, \boldsymbol{u}^{\boldsymbol{x}}) \mapsto \boldsymbol{u}^{g\boldsymbol{x}} \in \mathrm{Mon}(\boldsymbol{u}).$$

Then by linearity S_p also acts on the polynomial ring $k[\boldsymbol{u}]$. In particular S_p acts on $\mathrm{Bin}(\boldsymbol{u})$ by

$$S_p \times \mathrm{Bin}(\boldsymbol{u}) \ni (g, \boldsymbol{u}^{\boldsymbol{z}^+} - \boldsymbol{u}^{\boldsymbol{z}^-}) \mapsto \boldsymbol{u}^{g\boldsymbol{z}^+} - \boldsymbol{u}^{g\boldsymbol{z}^-} \in \mathrm{Bin}(\boldsymbol{u}). \tag{7}$$

3 Definition of the invariance group and some subgroups

In this section we define the invariance group. We also discuss some subgroups of the invariance group.

3.1 The invariance group

For the definition of the invariance group, we consider P_g, $g \in S_p$, acting on $\mathbb{Q}^p$, the p-dimensional vector space over the rationals. For a given subspace $L \subset \mathbb{Q}^p$, let

$$G_L = \{g \in S_p \mid P_g L = L\}$$

denote the setwise stabilizer of L in $G = S_p$. Now we give a definition of our invariance group.

Definition 1 *For a given $\nu \times p$ matrix A of integers, the invariance group is the setwise stabilizer $G_{\ker(A)}$ of $\ker(A)$ in S_p.*

By definition $G_{\ker(A)}$ is the largest subgroup of S_p, which acts on $\ker(A)$. Since the set of moves $\mathcal{M}_A$ spans $\ker(A)$ in $\mathbb{Q}^p$, we can also say that $G_{\ker(A)}$ is the largest subgroup acting on $\mathcal{M}_A$ and from the one-to-one correspondence between moves and binomials, $G_{\ker(A)}$ is the largest subgroup acting on the set of binomials in (7). Finally, since the toric ideal I_A is generated by the binomials, $G_{\ker(A)}$ is the largest subgroup acting on I_A.

It is also important to note that each subgroup $G \subset S_p$ which acts on $\ker(A)$ also acts on $\mathcal{T}$, the set of sufficient statistics, by

$$G \times \mathcal{T} \ni (g, \boldsymbol{t} = A\boldsymbol{x}) \mapsto g\boldsymbol{t} = gA\boldsymbol{x} = Ag\boldsymbol{x} = AP_g\boldsymbol{x} \in \mathcal{T}. \qquad (8)$$

This definition does not depend on the choice of $\boldsymbol{x}$ in $\boldsymbol{t} = A\boldsymbol{x}$. From this fact, all the results on structures of orbits and invariant Markov bases in Aoki and Takemura (2003b) hold with respect to the invariance group of Definition 1.

In Definition 1, it is desirable to clarify the relation of the definition of $G_{\ker(A)}$ and the freedom of choosing A with the same kernel. Since we are considering linear subspaces of $\mathbb{Q}^p$, we can consider A with rational elements. For the rest of this subsection we assume that elements of A are rational numbers. Suppose A_1 and A_2 are rational matrices of sizes $\nu_1 \times p, \nu_2 \times p$, respectively, such that $\ker(A_1) = \ker(A_2) = M$. M is a linear subspace in $\mathbb{Q}^p$ with $\dim M = d$, where $p - d = \mathrm{rank} A_1 = \mathrm{rank} A_2$. An important point here is that $M^\perp$, the orthogonal complement of M, coincides the row space of A_1 and A_2, i.e.,

$$M^\perp = r(A_1) = r(A_2), \quad \dim M^\perp = p - d,$$

where

$$r(A) = \{\boldsymbol{y} \in \mathbb{Q}^p \mid \boldsymbol{y} = A'\boldsymbol{c},\ \boldsymbol{c} \in \mathbb{Q}^\nu\}$$

is the row space of A. Furthermore, from the standard theory of a linear algebra, $(M^\perp)^\perp = M$ holds.

From the above relations, our invariance group is also interpreted as the setwise stabilizer of the row space of A. In this case, the action of $g \in S_p$ on the row space of A can be written as $gr(A) = r(AP_g')$ and the setwise stabilizer of the row space of A is given by

$$G_{r(A)} = \{g \in S_p \mid r(A) = r(AP_g')\}. \qquad (9)$$

From the fact that $\ker(A)$ and $r(A)$ are the orthogonal complements of each other, these two definitions are essentially equivalent. This relation is summarized as follows.

Proposition 1 *$G_M = G_{r(A)}$ for any subspace $M \subset \mathbb{Q}^p$ and any $\nu \times p$ rational matrix A such that $M = \ker(A)$.*

Proof. We first show that $G_M \subset G_{r(A)}$. Suppose $P \in G_M$. Let $\boldsymbol{y} \in r(AP'), \boldsymbol{z} \in M = \ker(A)$. $\boldsymbol{y}$ can be written as $\boldsymbol{y} = PA'\boldsymbol{c}$ for some $\boldsymbol{c} \in \mathbb{Q}^\nu$. Then

$$\boldsymbol{y}'\boldsymbol{z} = \boldsymbol{c}'AP'\boldsymbol{z} = 0$$

since $P' \in G_M$ and $P'\boldsymbol{z} \in M = \ker(A)$ from the definition of G_M. Therefore $r(AP') = M^\perp$ is shown, which means $G_M \subset G_{r(A)}$.

Conversely, suppose $P \in G_{r(A)}$. From the definition of $G_{r(A)}$, $r(AP') = r(A) = M^\perp$. Therefore for any $\boldsymbol{y} \in r(AP')$ and $\boldsymbol{z} \in M$,

$$0 = \boldsymbol{y}'\boldsymbol{z} = \boldsymbol{c}'AP'\boldsymbol{z}$$

holds, which forces $AP'\boldsymbol{z} = \boldsymbol{0}$, i.e., $P'\boldsymbol{z} \in \ker(A) = M$ and $P\boldsymbol{z} \in M$ at the same time. Therefore $G_M \supset G_A$ is proved. Q.E.D.

3.2 Some subgroups

In the definitions of G_M and $G_{r(A)}$, we treated M and $r(A)$ as rational linear spaces. However we started with A with integral elements and we now go back to such an A. Also, by considering the relation (8), it is natural to view the group action as a permutation of the rows of A. To characterize a structure which is expressed as a permutation of the rows of A, we consider the lattice points in $r(A)$. Here we define two subgroups of $G_{r(A)}$, which are derived naturally from the viewpoint of a permutation of the rows of A.

The *additive group* of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\nu\}$ is defined as

$$r_{\mathbb{Z}}(A) = \{\boldsymbol{y} \in \mathbb{Z}^p \mid \boldsymbol{y} = A'\boldsymbol{c},\ \boldsymbol{c} \in \mathbb{Z}^\nu\},$$

where $\boldsymbol{a}_j'$ is the j-th row vector of A defined by (6). Similarly to the definition of $G_{r(A)}$, considering the group action of $P_g, g \in S_p$, on $\mathbb{Z}^p$, we define the setwise stabilizer of $r_{\mathbb{Z}}(A)$ for a given $\nu \times p$ matrix A as

$$G_{r_{\mathbb{Z}}(A)} = \{P \in S_p \mid r_{\mathbb{Z}}(A) = r_{\mathbb{Z}}(AP')\}.$$

$G_{r_{\mathbb{Z}}(A)}$ is a subgroup of $G_{r(A)}$ from the definition. The following fact is also obvious from the definition:

$$G_{r_{\mathbb{Z}}(A)} = G_{r(A)} \ \Leftrightarrow\ r_{\mathbb{Z}}(A) = r(A) \cap \mathbb{Z}^p.$$

In general, $G_{r_{\mathbb{Z}}(A)}$ is smaller than $G_{r(A)}$. A simple example is given as

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \text{ and } P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In this case, $P \in G_{r(A)}$ holds since

$$\begin{aligned} r(A) &= \left\{ c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 2 \end{pmatrix},\ c_1, c_2 \in \mathbb{Q} \right\} \\ &= \left\{ \frac{c_1}{2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \frac{c_1}{2}, 2c_2 \in \mathbb{Q} \right\} = r(AP'). \end{aligned}$$

On the other hand, though $(1,2)' \in r_{\mathbb{Z}}(A)$, the diophantine equation

$$c_1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ c_1, c_2 \in \mathbb{Z}$$

does not have a solution and therefore $(1,2)' \notin r_{\mathbb{Z}}(AP')$. This implies that $P \notin G_{r_{\mathbb{Z}}(A)}$. Conversely, for a given A, $r(A) \cap \mathbb{Z}^p$ is an additive group and we can choose a lattice basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\mu\}$ of $r(A) \cap \mathbb{Z}^p$. Therefore if we write $\tilde{A} = [\boldsymbol{b}_1' \ \cdots \ \boldsymbol{b}_\mu']'$, this $\tilde{A}$ satisfies $\ker(\tilde{A}) = \ker(A)$ and

$$G_{r(A)} = G_{r(\tilde{A})} = G_{r_{\mathbb{Z}}(\tilde{A})}.$$

Here we give the last definition. For a given $\nu \times p$ rational matrix A, we define a group $H(A)$ as

$$H(A) = \{P \in S_p \mid \exists Q \in S_\nu \text{ s.t. } A = QAP'\}.$$

By definition, a group $H(A)$ is the set of permutations of the columns of A, *which are also expressed as a permutation of the rows of* A. Compared with the definitions of $G_M, G_{r(A)}$ and $G_{r_{\mathbb{Z}}(A)}$, the meaning of $H(A)$ is easier to understand. It is seen that $H(A)$ is a subgroup of $G_{r_{\mathbb{Z}}(A)}$. This fact is obvious since $G_{r_{\mathbb{Z}}(A)}$ is also written as

$$G_{r_{\mathbb{Z}}(A)} = \{P \in S_p \mid \exists \text{unimodular matrix } Q \text{ s.t. } A = QAP'\}.$$

It should be noted that elementary row operations can be expressed as multiplying unimodular matrix from the left (see Section 4 of Schrijver, 1986). Since permutation matrices are also unimodular, $H(A) \subset G_{r_{\mathbb{Z}}(A)}$. To consider the situation that the equality holds, let $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\mu\}$ be any lattice basis of $r(A) \cap \mathbb{Z}^p$. By symmetrically enlarging the basis, define

$$C = \{P\boldsymbol{b}_j \mid P \in G_{r_{\mathbb{Z}}(A)}, \ j = 1, \ldots, \mu\}. \tag{10}$$

For any $\tilde{A}$ whose rows are the vectors of C, $\ker(\tilde{A}) = \ker(A)$ and the equality $G_{r(A)} = G_{r(\tilde{A})} = G_{r_{\mathbb{Z}}(\tilde{A})} = H(\tilde{A})$ holds.

As we have stated, an interpretation of $H(A)$ is clearer than the definition of $G_{r(A)}$. Moreover, if we choose A which is sufficiently symmetric in the sense of (10), permutations of the columns of A by elements of the invariance group can be canceled by permutations of the rows of A. As a relation to our previous work, Aoki and Takemura (2003b), we give the following remark.

Remark 1 We show that all the elements in the group of "permutation of levels for each axis" in Aoki and Takemura (2003b) can be written explicitly as the permutation of rows of A by defining A appropriately. As proved in Section 4.2, this group indeed coincides with the largest invariance group for the m-way contingency tables with fixed one-dimensional marginals if the numbers of levels for each axis are all distinct, and therefore this remark is an example of $H(A) = G_{r(A)}$.

Consider $I_1 \times \cdots \times I_m$ contingency tables with $\{D_1, \ldots, D_r\}$-marginal fixed, where $D_j \subset \{1, \ldots, m\}, \ j = 1, \ldots, r$. D_j-marginal total is given by $A_j\boldsymbol{x}$,

$$A_j = C_{j1} \otimes \cdots \otimes C_{jk}, \ C_{j\ell} = \begin{cases} E_{I_\ell}, & \text{if } \ell \in D_j, \\ \mathbf{1}_{I_\ell}', & \text{if } \ell \notin D_j, \end{cases} \tag{11}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{1}_n = (1, \ldots, 1)'$ denotes the n dimensional column vector of 1's. See Aoki and Takemura (2003b) and Dobra (2003) for notations on D-marginal totals. In this setting, the matrix A is written as

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_r \end{bmatrix}. \tag{12}$$

In Aoki and Takemura (2003b), the group G_0 of permutation of levels for each axis is considered, which is the direct product of symmetric groups

$$G_0 = S_{I_1} \times \cdots \times S_{I_k}. \tag{13}$$

Each element of G_0 is written as $P = \Sigma_1 \otimes \cdots \otimes \Sigma_k$, where $\Sigma_\ell \in S_{I_\ell}$ is a permutation matrix for the ℓ-th axis. In this case, if we define a permutation matrix $Q \in S_\nu$ as

$$Q = \begin{bmatrix} Q_{11} \otimes \cdots \otimes Q_{1k} & O & O & \cdots & O \\ O & Q_{21} \otimes \cdots \otimes Q_{2k} & O & \cdots & O \\ \vdots & & & \ddots & \vdots \\ O & \cdots & & O & Q_{r1} \otimes \cdots \otimes Q_{rk} \end{bmatrix},$$

where

$$Q_{j\ell} = \begin{cases} \Sigma'_\ell, & \text{if } \ell \in D_j, \\ 1, & \text{if } \ell \notin D_j, \end{cases}$$

it is shown that

$$QAP' = A.$$

As will be shown in Section 4.2, G_0 is indeed the largest invariance group if $r = m$, $\{D_1, \ldots, D_m\}$ $= \{\{1\}, \ldots, \{m\}\}$ and I_ℓ's are all distinct. Consequently, we have shown that the largest invariance group is $H(A)$ by defining A as (12) for this particular problem. See Section 6 for more discussion.

4 Structure of the invariance groups for some standard statistical models

From logical viewpoint, the definition of the invariance group in Definition 1 is simple enough. However it is a different matter whether it is simple to determine explicitly the invariance group for a given A. If the sizes and the elements of A are given numerically, we could use computational algebra softwares such as *GAP*, which is available at http://www.gap-system.org. Though we may obtain all the elements of G_M, $M = \ker(A)$, by some nearly linear-time algorithms in the order $|G|$ of G, the problem of computational feasibility arises since $p! = |S_p|$ increases exponentially with $p = |\mathcal{I}|$. In fact in our experience with *GAP*, the computation becomes infeasible quite rapidly when the number of cells $p = |\mathcal{I}|$ is increased. In this paper, we show that G has a simple structure for some standard statistical models of contingency tables.

Given a particular A, it is often easy to guess the form of the invariance group G_M and easy to check that the candidate group acts on $\ker(A)$. However we found that it is often difficult to prove that the candidate group is indeed the largest subgroup of S_p acting on $\ker(A)$. In our proofs below we employ simple investigations of possible patterns of each element in G_M to verify that a candidate group is indeed the largest.

4.1 Ingredients for proofs

For a given $\nu \times p$ matrix A, let G_M, $M = \ker(A)$, denote the invariance group. In this section, we write $G = G_M$ for simplicity. Let $\widetilde{G} \subset S_p$ denote a candidate group such that it is easy to verify $\widetilde{G} \subset G$. Our arguments in this section are all the same. First we define $\widetilde{G}$ for each problem, and we next show that indeed $\widetilde{G} = G$.

There are two fundamental items in our proofs. First, for any $g \in G$ and $\boldsymbol{z} \in \mathcal{M} = M \cap \mathbb{Z}^p$, it follows that $g\boldsymbol{z} \in \mathcal{M}$ and $\deg(\boldsymbol{z}) = \deg(g\boldsymbol{z})$ by definition. In particular, we consider moves $\boldsymbol{z}$ with the *minimum* degree for each problem.

Second, we look for some subset $\mathcal{J} \subset \mathcal{I}$ of the cells and consider the pointwise stabilizer $S_{(\mathcal{J})}$ of the cells in $\mathcal{J}$. In the following we write $G_{(\mathcal{J})} = G \cap S_{(\mathcal{J})}$, which is the subgroup of the invariance group fixing each cell $\boldsymbol{j} \in \mathcal{J}$. For an appropriate subset $\mathcal{J}$ of $\mathcal{I}$, we will show that for each $g \in G$ there exist $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g} \in S_{(\mathcal{J})}$. This implies that G can be written as $G = G_{(\mathcal{J})}\widetilde{G}$. On the other hand we will also show that $G_{(\mathcal{J})} \subset \widetilde{G}$, often by showing that $G_{(\mathcal{J})} = \{e\}$ is trivial. Then $G \subset \widetilde{G}\widetilde{G} \subset \widetilde{G}$, which completes the proof.

We give some notations on the symmetric group. Let $[\boldsymbol{i}_1, \boldsymbol{i}_2] \in S_p$ denote the transposition of $\boldsymbol{i}_1$ and $\boldsymbol{i}_2$. Note that we use $[\ \cdot\ ]$, not $(\ \cdot\ )$, to avoid confusing group elements with row vectors. Transpositions, $[\boldsymbol{i}_{1s}, \boldsymbol{i}_{2s}]$, $s = 1, \ldots, S$, are commutative if $\boldsymbol{i}_{1s}, \boldsymbol{i}_{2s}$, $s = 1, \ldots, S$, are all different cells. In such cases, we write the product of these group elements as $\prod_{s=1}^{S}[\boldsymbol{i}_{1s}, \boldsymbol{i}_{2s}]$. For $g_1, \ldots, g_s \in S_p$, we denote the subgroup generated by these elements as $\langle g_1, \ldots, g_s \rangle$. Also for subsets $G_1, \ldots, G_s$ of S_p, $\langle G_1, \ldots, G_s \rangle$ denotes the subgroup generated by $G_1, \ldots, G_s$ of S_p.

As in Remark 1, for m-way contingency tables with $\mathcal{I} = \{1, \ldots, I_1\} \times \cdots \times \{1, \ldots, I_m\}$, we will denote an element $g \in S_p$, which corresponds to permuting levels for each axis, by a direct product expression $g = g_1 \times \cdots \times g_m$, where $g_l \in S_{I_l}$. For avoiding triviality we also assume that $I_l \geq 2$ for all l.

For the rest of this section, we consider some standard statistical models for contingency tables. For ordinary m-way contingency tables, we consider the complete independence model in Section 4.2. The no three-factor interaction model for three-way contingency tables is considered in Section 4.3. As more specific models of contingency tables, in Section 4.4 we consider the quasi-symmetry model of square two-way contingency tables with their diagonal cells being structural zeros, and the Hardy-Weinberg model of upper triangular two-way tables in Section 4.5.

4.2 The invariance groups for m-way contingency tables with fixed one-dimensional marginals

First we consider the m-way contingency tables with fixed one-dimensional marginals. This setting is known as the *complete independence model for m-way contingency tables* in the statistical literature, and a matrix A for this problem is given by (11) and (12) with $r = m$, and $\{D_1, \ldots, D_r\} = \{\{1\}, \ldots, \{m\}\}$. As explained several times already, for this problem a natural group which acts on $\ker(A)$ is $S_{I_1} \times \cdots, S_{I_m}$, corresponding to permuting levels for each axis. Furthermore we can permute two axes, s and t, if $I_s = I_t$. Let $H_{st} = \{e, g_{st}\}$ denote a two-element group, where the action of g_{st}, $s < t$, on $\mathcal{I} = \{1, \dots, I_1\} \times \cdots \times \{1, \dots, I_m\}$ is given as

$$g_{st}(i_1 \cdots i_s \cdots i_t \cdots i_m) = i_1 \cdots \underset{\widehat{s}}{i_t} \cdots \underset{\widehat{t}}{i_s} \cdots i_m, \tag{14}$$

where $\underset{\widehat{s}}{i_t}$ means that i_t is in the s-th position. Then we have the following theorem.

Theorem 1 *The invariance group for $I_1 \times \cdots \times I_m$ contingency tables with fixed one-dimensional marginals is*

$$\langle S_{I_1} \times \cdots \times S_{I_m},\ \{H_{st} \mid s < t,\ I_s = I_t\} \rangle . \tag{15}$$

It is easy to check that the group (15) indeed acts on $\ker(A)$ and its meaning is clear. However, unfortunately it is not easy to prove that it is indeed the largest invariance group in S_p.

To show the theorem, we give some definitions and lemmas on m-way contingency tables with fixed one-dimensional marginals. In this section, we assume that $2 \le I_1 \le \cdots \le I_m$ without loss of generality, and write $q = I_1$. We denote *diagonal cells* as $\boldsymbol{i}_a = (a \cdots a)$ where $1 \le a \le q$. We partition $\mathcal{I}$ as

$$\mathcal{I} = \mathcal{I}_0 \cup \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_m,$$

where

$$\mathcal{I}_\ell = \{i_1 \cdots i_m \in \mathcal{I} \mid \text{there are } \ell \text{ 1's in } i_1, \dots, i_m\}. \tag{16}$$

For example $\mathcal{I}_m = \{\boldsymbol{i}_1\}$ and $\mathcal{I}_0 = \{2, \dots, I_1\} \times \cdots \times \{2, \dots, I_m\}$.

First we give a characterization of moves for A of degree 2, which has $+1$ entry at $\boldsymbol{i}_1 = 1 \cdots 1$. The proof of the following lemma is easy and omitted.

Lemma 1 *Let $\boldsymbol{z} = \{z(\boldsymbol{i})\}_{\boldsymbol{i} \in \mathcal{I}} \in \mathcal{M}$ be a move of degree 2 for the m-way contingency tables with fixed one-dimensional marginals and $z(\boldsymbol{i}_1) = +1$. Then the following relations hold.*

(a) $z(\boldsymbol{i}) \neq -1$ for all $\boldsymbol{i} \in \mathcal{I}_0$,

(b) $z(\boldsymbol{i}) \neq +1$ for all $\boldsymbol{i} \in \mathcal{I}_{m-1}$,

(c) Write $\boldsymbol{z} = (\boldsymbol{i}_1)(\boldsymbol{i}') - (\boldsymbol{i}'')(\boldsymbol{i}''')$, where $\boldsymbol{i}' \in \mathcal{I}_s$, $\boldsymbol{i}'' \in \mathcal{I}_t$ and $\boldsymbol{i}''' \in \mathcal{I}_u$. Then $m + s = t + u$.

To specify the structure of G, we consider the stabilizer $G_{\boldsymbol{i}_1} = \{g \in G \mid g(\boldsymbol{i}_1) = \boldsymbol{i}_1\}$ of $\boldsymbol{i}_1$ in G in the following three lemmas.

Lemma 2 *$G_{\boldsymbol{i}_1}$ acts on $\mathcal{I}_0$ transitively.*

Proof Since G is the largest invariance group in S_p, $S_{I_1-1} \times \cdots \times S_{I_m-1}$, the group of permuting levels for each axis in $\mathcal{I}_0$, is a subgroup in $G_{\boldsymbol{i}_1}$. If $G_{\boldsymbol{i}_1}$ acts on $\mathcal{I}_0$ then the action is clearly transitive since we have $g(\boldsymbol{i}) = \boldsymbol{i}'$ for any $\boldsymbol{i} = i_1 \cdots i_m \in \mathcal{I}_0, \boldsymbol{i}' = i'_1 \cdots i'_m \in \mathcal{I}_0$ by choosing

$$g = [i_1, i'_1] \times \cdots \times [i_m, i'_m] \in S_{I_1-1} \times \cdots \times S_{I_m-1}.$$

Therefore all we need is to show that $\mathcal{I}_0$ is $G_{\boldsymbol{i}_1}$-invariant.

Suppose $\tilde{g}(\boldsymbol{i}^*) \in \mathcal{I}_0$ for some $\boldsymbol{i}^* \notin \mathcal{I}_0$ and $\tilde{g} \in G_{\boldsymbol{i}_1}$. Then $\boldsymbol{i}^* \in \mathcal{I}_\ell$ for some $1 \le \ell \le m-1$. In this case, we can choose $\boldsymbol{i}' \in \mathcal{I}_0$ and $\boldsymbol{i}'' \in \mathcal{I}_{m-\ell}$ so that $\boldsymbol{z} = (\boldsymbol{i}_1)(\boldsymbol{i}') - (\boldsymbol{i}^*)(\boldsymbol{i}'')$ is a move. Applying $\tilde{g}$ on $\boldsymbol{z}$ yields a move

$$\tilde{g}\boldsymbol{z} = (\tilde{g}(\boldsymbol{i}_1))(\tilde{g}(\boldsymbol{i}')) - (\tilde{g}(\boldsymbol{i}^*))(\tilde{g}(\boldsymbol{i}'')).$$

However, since $\tilde{g}(\boldsymbol{i}_1) = \boldsymbol{i}_1$ and $\tilde{g}(\boldsymbol{i}^*) \in \mathcal{I}_0$, $\tilde{g}\boldsymbol{z}$ cannot be a move from Lemma 1(a), which is a contradiction. Q.E.D.

Lemma 3 *$\mathcal{I}_{m-1}$ is $G_{\boldsymbol{i}_1}$-invariant.*

Proof. We argue by contradiction. We suppose $\tilde{g}(\boldsymbol{i}^*) \in \mathcal{I}_{m-1}$ for some $\boldsymbol{i}^* \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_{m-2}$ and $\tilde{g} \in G_{\boldsymbol{i}_1}$ since $\mathcal{I}_0$ is $G_{\boldsymbol{i}_1}$-invariant by Lemma 2. We write $\boldsymbol{i}^* \in \mathcal{I}_\ell$ for some $1 \le \ell \le m-2$. In this case, for any $\ell + 1 \le s \le m-1$, we can choose $\boldsymbol{i}' \in \mathcal{I}_s$ and $\boldsymbol{i}'' \in \mathcal{I}_{m+\ell-s}$ so that $\boldsymbol{z} = (\boldsymbol{i}_1)(\boldsymbol{i}^*) - (\boldsymbol{i}')(\boldsymbol{i}'')$ is a move. Applying $\tilde{g}$ on $\boldsymbol{z}$ yields a move

$$\tilde{g}\boldsymbol{z} = (\tilde{g}(\boldsymbol{i}_1))(\tilde{g}(\boldsymbol{i}^*)) - (\tilde{g}(\boldsymbol{i}'))(\tilde{g}(\boldsymbol{i}'')).$$

However, since $\tilde{g}(\boldsymbol{i}_1) = \boldsymbol{i}_1$ and $\tilde{g}(\boldsymbol{i}^*) \in \mathcal{I}_{m-1}$, $\tilde{g}\boldsymbol{z}$ cannot be a move from Lemma 1(b). Q.E.D.

Lemma 4 *Each of $\mathcal{I}_0, \mathcal{I}_1, \ldots, \mathcal{I}_m$ is $G_{\boldsymbol{i}_1}$-invariant.*

Proof. $\mathcal{I}_m = \{\boldsymbol{i}_1\}$ is clearly $G_{\boldsymbol{i}_1}$-invariant. We have already shown that $\mathcal{I}_0$ and $\mathcal{I}_{m-1}$ are also $G_{\boldsymbol{i}_1}$-invariant. To show the lemma by induction, we assume that each of $\mathcal{I}_0, \mathcal{I}_m, \mathcal{I}_{m-1}, \ldots, \mathcal{I}_{m-\ell}$ is $G_{\boldsymbol{i}_1}$-invariant and show that $\mathcal{I}_{m-\ell-1}$ is $G_{\boldsymbol{i}_1}$-invariant.

We argue by contradiction. Suppose $\mathcal{I}_{m-\ell-1}$ is not $G_{\boldsymbol{i}_1}$-invariant, i.e., there exists some $1 \le j \le m-\ell-2$ and $\tilde{g} \in G_{\boldsymbol{i}_1}$ satisfying $\boldsymbol{i}^* \in \mathcal{I}_j$, $\tilde{g}(\boldsymbol{i}^*) \in \mathcal{I}_{m-\ell-1}$. In this case, we can choose $\boldsymbol{i}', \boldsymbol{i}''$ so that $\boldsymbol{z} = (\boldsymbol{i}_1)(\boldsymbol{i}^*) - (\boldsymbol{i}')(\boldsymbol{i}'')$ is a move. Applying $\tilde{g}$ on $\boldsymbol{z}$ yields a move

$$\tilde{g}\boldsymbol{z} = (\tilde{g}(\boldsymbol{i}_1))(\tilde{g}(\boldsymbol{i}^*)) - (\tilde{g}(\boldsymbol{i}'))(\tilde{g}(\boldsymbol{i}'')).$$

Here it follows that $g(\boldsymbol{i}_1) = \boldsymbol{i}_1$, $g(\boldsymbol{i}^*) \in \mathcal{I}_{m-\ell-1}$, from the assumption of induction. To ensure that $\tilde{g}\boldsymbol{z}$ is a move, the levels of the axes where the levels of $g(\boldsymbol{i}^*)$ are 1 must be 1 in both $g(\boldsymbol{i}')$ and $g(\boldsymbol{i}'')$. Therefore by Lemma 1(c), there exists some $1 \le s \le \ell$ satisfying $g(\boldsymbol{i}') \in \mathcal{I}_{m-\ell-1+s}$ and $g(\boldsymbol{i}'') \in \mathcal{I}_{m-s}$. In addition, because $m - \ell \le m - \ell - 1 + s$ and $m - s \le m - 1$, $\mathcal{I}_{m-\ell-1+s}$ and $\mathcal{I}_{m-s}$ are $G_{\boldsymbol{i}_1}$-invariant by the assumption of induction. It follows that $\boldsymbol{i}' \in \mathcal{I}_{m-\ell-1+s}$ and $\boldsymbol{i}'' \in \mathcal{I}_{m-s}$. Therefore Lemma 1(c) for $\boldsymbol{z}$ implies that $\boldsymbol{i}^* \in \mathcal{I}_{m-\ell-1}$, which contradicts the assumption that $\boldsymbol{i}^* \in \mathcal{I}_j$ for some $1 \le j \le m-\ell-2$, and the lemma is proved. Q.E.D.

So far we considered the stabilizer $G_{\boldsymbol{i}_1}$. We can perform the same procedure for each diagonal cell $\boldsymbol{i}_2, \dots, \boldsymbol{i}_q$ and take the product of the corresponding partitions. Define

$$\mathcal{I}_{k_1 k_2 \cdots k_q k_{q+1}} = \{i_1 \cdots i_m \in \mathcal{I} \mid k_j \text{ entries in } i_1, \dots, i_m \text{ are } j,\ j = 1, \dots, q \text{ and } \\ k_{q+1} \text{ entries in } i_1, \dots, i_m \text{ are larger than } q\}, \tag{17}$$

where $k_1 + \cdots + k_{q+1} = m$ and $k_1, \dots, k_{q+1} \geq 0$. Then these sets form a partition of $\mathcal{I}$. Furthermore let $\mathcal{I}_d = \{\boldsymbol{i}_1, \dots, \boldsymbol{i}_q\}$ denote the set of diagonal cells.

Then the above lemmas imply the following useful corollary

Corollary 1 *Each $\mathcal{I}_{k_1 k_2 \cdots k_q k_{q+1}}$ is $G_{(\mathcal{I}_d)}$-invariant, where $G_{(\mathcal{I}_d)}$ is the pointwise stabilizer of the set of diagonal cells.*

Using this corollary, we prove Theorem 1.

Proof of Theorem 1. As our candidate group, take

$$\widetilde{G} = \langle S_{I_1} \times \cdots \times S_{I_m},\ \{H_{st} \mid s < t,\ I_s = I_t\} \rangle \,.$$

Clearly $\widetilde{G} \subset G$. To prove $\widetilde{G} \supset G$, it suffices to show that for each $g \in G$ we can choose $\tilde{g} \in \widetilde{G}$ for any $g \in G$ so that $g\tilde{g} = e$. As we have stated, we assume $2 \leq q = I_1 \leq I_2 \leq \cdots \leq I_m$ without loss of generality. We specify the image of each $\boldsymbol{i} = i_1 \cdots i_m$ one by one by applying $g \in G$.

First we show that we can assume $g \in G_{\boldsymbol{i}_1}$ without loss of generality. To see this, suppose $g(\boldsymbol{i}) = \boldsymbol{i}_1$ for some $\boldsymbol{i} = i_1 \cdots i_m \neq \boldsymbol{i}_1$. In this case, we take $\tilde{g} = [1, i_1] \times \cdots \times [1, i_m] \in \widetilde{G}$ and see that

$$(g\tilde{g})(\boldsymbol{i}_1) = g(\tilde{g}(\boldsymbol{i}_1)) = g(\boldsymbol{i}) = \boldsymbol{i}_1.$$

This implies that we only need to consider $g \in G_{\boldsymbol{i}_1}$ and prove that for each $g \in G_{\boldsymbol{i}_1}$, there exists $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g} = e$. We use this reasoning repeatedly from now on.

Now, since $G_{\boldsymbol{i}_1}$ acts on $\mathcal{I}_0$ transitively by Lemma 2, we can also restrict our attention to $g \in G_{\boldsymbol{i}_1} \cap G_{\boldsymbol{i}_2}$. Moreover, applying Lemma 2 to $G_{\boldsymbol{i}_2}$ and combining it with Lemma 4, we see that $G_{\boldsymbol{i}_1} \cap G_{\boldsymbol{i}_2}$ acts on $\{3, \dots, I_1\} \times \cdots \times \{3, \cdots, I_m\}$ transitively. By the similar arguments, it is shown that we can assume $g \in G_{(\mathcal{I}_d)}$ without loss of generality. In other words, we have shown that $G = G_{(\mathcal{I}_d)}\widetilde{G}$.

Now we are going to show that, for any $g \in G_{(\mathcal{I}_d)}$, there exists a $\tilde{g} \in \widetilde{G}$ so that $g\tilde{g} = e$, implying that $G_{(\mathcal{I}_d)} \subset \widetilde{G}$. We separate our proof into two steps. First we consider all the cells in $\mathcal{I}^* = \{1, \dots, q\}^m \subset \mathcal{I}$ and next consider all the cells in $\mathcal{I} \setminus \mathcal{I}^*$.

(Step 1.) Partition $\mathcal{I}^*$ as

$$\mathcal{I}^* = \mathcal{I}^* \cap \mathcal{I} = \mathcal{I}^* \cap (\mathcal{I}_0 \cup \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_m) = \mathcal{I}^*_1 \cup \cdots \cup \mathcal{I}^*_m.$$

$\mathcal{I}^*_\ell$ is the set of cells in $\mathcal{I}^*$ which has the level 1 at ℓ axes.

We argue by induction. For $\mathcal{I}_m^* = \{\boldsymbol{i}_1\}$, it is clear that $g(\boldsymbol{i}_1) = \boldsymbol{i}_1$ for any $g \in G_{(\mathcal{I}_d)}$. We will show that, for any $g \in G_{(\mathcal{I}_d)}$, there exists $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g}(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}_{m-1}^*$. From the definition of $\mathcal{I}_{m-1}^*$, we can write

$$\mathcal{I}_{m-1}^* = \{\boldsymbol{i}_{\ell,a} \mid \ell = 1, \ldots, m, a = 2, \ldots, q\},$$

where $\boldsymbol{i}_{\ell,a} = i_1 \cdots i_m$ is specified as $i_\ell = a$ and $i_k = 1$ for $k \neq \ell$:

$$\boldsymbol{i}_{\ell,a} = 1 \cdots 1\underset{\hat{\ell}}{a}1 \cdots 1.$$

First we consider the cells with $a = 2$. Define m moves, $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ as

$$\boldsymbol{z}_\ell = (\boldsymbol{i}_1)(\boldsymbol{i}_2) - (\boldsymbol{i}_{\ell,2})(\overline{\boldsymbol{i}_{\ell,2}}),$$

where $\overline{\boldsymbol{i}_{\ell,a}} = a \cdots a\underset{\hat{\ell}}{1}a \cdots a$. Then, it is checked that any $g \in G_{(\mathcal{I}_d)}$ is a bijection from $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$ to itself. Here, if $g\boldsymbol{z}_t = \boldsymbol{z}_s$ for some $t \neq s$ with $I_t = I_s$, we can choose $g_{st} \in \widetilde{G}$ (or $g_{ts} \in \widetilde{G}$ if $t < s$) in (14), such that $(gg_{st})\boldsymbol{z}_\ell = \boldsymbol{z}_\ell$ for $\ell = s, t$. If $I_1 = \cdots = I_m$ we are done, otherwise by changing axes of the same sizes, there exist $\tilde{g}$ and $I_{\ell_1}, \ldots, I_{\ell_k}$, all distinct, such that

$$\begin{aligned} &(g\tilde{g})\boldsymbol{z}_{\ell_t} = \boldsymbol{z}_{\ell_{t+1}}, \quad t = 1, \ldots, k-1, \\ &(g\tilde{g})\boldsymbol{z}_{\ell_k} = \boldsymbol{z}_{\ell_1}. \end{aligned} \tag{18}$$

Here, note that there is at least one pair (ℓ, ℓ') in $\ell_1, \ldots, \ell_k$ satisfying

$$(g\tilde{g})\boldsymbol{z}_\ell = \boldsymbol{z}_{\ell'} \text{ and } I_\ell > I_{\ell'}. \tag{19}$$

However, by considering the action of $g\tilde{g}$ to the move

$$\boldsymbol{z} = (\boldsymbol{i}_1)(2 \cdots 2\underset{\hat{\ell}}{I_\ell}2 \cdots 2) - (1 \cdots 1\underset{\hat{\ell}}{I_\ell}1 \cdots 1)(2 \cdots 2\underset{\hat{\ell}}{1}2 \cdots 2),$$

we have that the $+1$ cell other than $\boldsymbol{i}_1$ of $(g\tilde{g})\boldsymbol{z}$ must be

$$(g\tilde{g})(2 \cdots 2\underset{\hat{\ell}}{I_\ell}2 \cdots 2) = 2 \cdots 2\underset{\hat{\ell}'}{I_\ell}2 \cdots 2.$$

This contradicts $I_\ell > I_{\ell'}$ in (19). From these considerations, it is shown that we can choose $\tilde{g} \in \widetilde{G}$ so that

$$(g\tilde{g})\boldsymbol{z}_\ell = \boldsymbol{z}_\ell, \ \ell = 1, \ldots, m \tag{20}$$

in (18) for any configuration of strict inequalities in $I_1 \leq \cdots \leq I_m$. Moreover, by this $\tilde{g} \in \widetilde{G}$, it also follows that

$$\begin{aligned} &(g\tilde{g})(\boldsymbol{i}_{\ell,2}) = \boldsymbol{i}_{\ell,2}, \quad \ell = 1, \ldots, m, \\ &(g\tilde{g})(\overline{\boldsymbol{i}_{\ell,2}}) = \overline{\boldsymbol{i}_{\ell,2}}, \quad \ell = 1, \ldots, m. \end{aligned} \tag{21}$$

Next we consider the cell $\boldsymbol{i}_{\ell,a} \in \mathcal{I}_{m-1}^*$ where $3 \leq a \leq q$. For each $\boldsymbol{i}_{\ell,a}, a \geq 3$, it suffices to consider the move

$$(\boldsymbol{i}_1)(2 \cdots 2\underset{\hat{l}}{a}2 \cdots 2) - (\boldsymbol{i}_{\ell,a})(\overline{\boldsymbol{i}_{\ell,2}})$$

and its image by $g\tilde{g}$, where $\tilde{g} \in \widetilde{G}$ is defined by the relation (20). Considering (21), it is easily checked that $(g\tilde{g})(\boldsymbol{i}_{\ell,a}) = \boldsymbol{i}_{\ell,a}$ for each $\ell = 1, \ldots, m$ and $a = 3, \ldots, q$. Therefore we have shown that for any $g \in G_{(\mathcal{I}_d)}$, there exists $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g}(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}^*_{m-1}$.

Let $\mathcal{J} = \mathcal{I}_d \cup \mathcal{I}^*_m \cup \mathcal{I}^*_{m-1}$. By now we have shown that for each $g \in G$, there exists $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g} \in G_{(\mathcal{J})}$. In the following we will prove that in fact $G_{(\mathcal{J})} = \{e\}$, completing the proof of the theorem.

For induction, we suppose $g(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}^*_{m-\ell}$ and show that $g(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}^*_{m-\ell-1}$. Each cell in $\mathcal{I}^*_{m-\ell-1}$ is written as $\boldsymbol{i}_{t_1,\ldots,t_{\ell+1};a_1,\ldots,a_{\ell+1}} = i_1 \cdots i_m \in \mathcal{I}^*_{m-\ell-1}$, where $i_{t_k} = a_k \in \{2, \ldots, q\}$ for $k = 1, \ldots, \ell+1$ and $i_s = 1$ for $s \notin \{t_1, \ldots, t_{\ell+1}\}$. For each $\boldsymbol{i} \in \mathcal{I}^*_{m-\ell-1}$ specified above, it suffices to consider the following two moves,

$$
\begin{array}{l}
(\boldsymbol{i}_{t_1,\ldots,t_{\ell+1};a_1,\ldots,a_\ell,1})(\boldsymbol{i}_{a_{\ell+1}}) - (\boldsymbol{i}_{t_1,\ldots,t_{\ell+1};a_1,\ldots,a_{\ell+1}})(\overline{\boldsymbol{i}_{t_{\ell+1},a_{\ell+1}}}),\\
(\boldsymbol{i}_{t_1,\ldots,t_{\ell+1};a_1,\ldots,a_{\ell-1},1,a_{\ell+1}})(\boldsymbol{i}_{a_\ell}) - (\boldsymbol{i}_{t_1,\ldots,t_{\ell+1};a_1,\ldots,a_{\ell+1}})(\overline{\boldsymbol{i}_{t_\ell,a_\ell}}),
\end{array}
$$

and the images of these moves by g. Considering the assumption of induction, we have $g(\boldsymbol{i}) = \boldsymbol{i}$ and Step 1 is completed.

(Step 2.) To complete the proof, we need to show that $g(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I} \setminus \mathcal{I}^*$. Using the notation of (17), $\mathcal{I} \setminus \mathcal{I}^*$ can be written as disjoint union,

$$
\mathcal{I} \setminus \mathcal{I}^* = \bigcup_{\substack{k_1+\cdots+k_{q+1}=m\\ k_1,\ldots,k_q \geq 0, k_{q+1} \geq 1}} \mathcal{I}_{k_1\cdots k_q k_{q+1}}.
$$

We argue by induction for k_{q+1}.

First consider the cells in $\mathcal{I}_{k_1\cdots k_q 1}$, i.e., the case of $k_{q+1} = 1$. For each $\boldsymbol{i} = i_1 \cdots i_m \in \mathcal{I}_{k_1\cdots k_q 1}$, suppose $i_\ell \geq q+1$ and $1 \leq i_j \leq q$ for $j \neq \ell$. Corresponding to this $\boldsymbol{i}$, we choose $\boldsymbol{i}' = i'_1 \cdots i'_m \in \mathcal{I}^*$ so that $i'_j \neq i_j$ and $1 \leq i'_j \leq q$ for $j = 1, \ldots, m$. Then it suffices to consider the move

$$
(i_1 \cdots i_m)(i'_1 \cdots i'_m) - (i_1 \cdots i_{\ell-1} i'_\ell i_{\ell+1} \cdots i_m)(i'_1 \cdots i'_{\ell-1} i_\ell i'_{\ell+1} \cdots i'_m)
$$

and its image by g, to check that $g(\boldsymbol{i}) = \boldsymbol{i}$.

Next we assume that $g(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}_{k_1\cdots k_q k_{q+1}}$ where $k_{q+1} = \ell (\geq 2)$ and show it for $\boldsymbol{i} \in \mathcal{I}_{k_1\cdots k_q k_{q+1}}$ where $k_{q+1} = \ell + 1$. It is again easily checked by considering the move

$$
(\boldsymbol{i}_1)(i_1 \cdots i_m) - (\boldsymbol{i}_{k,i_k})(i_1 \cdots i_{k-1} 1 i_{k+1} \cdots i_m)
$$

and its image by g, where $i_1 \ldots i_m \in \mathcal{I}_{k_1\cdots k_q(\ell+1)}$ and $i_k \geq q+1$. Q.E.D.

4.3 The invariance groups for three-way contingency tables with fixed two dimensional marginals

For ordinary m-way contingency tables, there are many hierarchical models other than the complete independence model considered in Section 4.2. There are some models such as the marginal independence models for three-way contingency tables expressed as $\{D_1, D_2\} = \{\{1,2\},\{3\}\}$, or the conditional independence models for three-way contingency tables expressed as $\{D_1, D_2\} =$

$\{\{1,2\},\{1,3\}\}$, which are essentially equivalent to the independence model for two-way contingency tables. For these models, the structures of the invariance groups can be obtained from those for the independence model of two-way tables. For example, the invariance group for the marginal independence model, $\{D_1, D_2\} = \{\{1,2\},\{3\}\}$, of $I_1 \times I_2 \times I_3$ tables is generated by, in addition to permutations of levels in each axis, permutations of the third axis and the combination of the first and the second axes if $I_1 I_2 = I_3$, by treating the combination of levels in the first and the second axes as a new single axis. Similarly, the invariance group for the conditional independence model, $\{D_1, D_2\} = \{\{1,2\},\{1,3\}\}$, of $I_1 \times I_2 \times I_3$ tables is generated by, in addition to permutations of levels in each axis, permutation of the second axis and the third axis if $I_2 = I_3$, since these two axes are conditionally independent by fixing the levels in the first axis.

An essentially different hierarchical model is the model of *no three-factor interactions*, where we need to consider m-way contingency tables with fixed *two-dimensional marginals*. Similarly to Section 4.2, it would be very desirable if we could specify the structure of invariance groups for general m-way tables with fixed two-dimensional marginals. However we found that for m-way tables it is quite complicated to perform the similar arguments as Section 4.2 for this problem. One reason for this difficulty is that the lattice basis of $\ker(A)$ consists of moves of degree 4, as shown in the following example.

Example 3 *Consider the $2 \times 2 \times 2$ contingency tables with fixed two-dimensional marginals. In this case, a single move $\boldsymbol{z} = (111)(122)(212)(221) - (112)(121)(211)(222)$ constitutes the lattice basis of $M = \ker(A)$, since $\dim(M) = 1$. Therefore any element $g \in S_8$ satisfying $g\boldsymbol{z} = \boldsymbol{z}$ or $g\boldsymbol{z} = -\boldsymbol{z}$ is a member of G. The order of G is calculated as $|G| = 4! \cdot 4! \cdot 2 = 1152$.*

In this section, we restrict our attention to the three-way contingency tables with fixed two-dimensional marginals, which is one of the most important models of three-way tables in applications. The result is very similar to Theorem 1.

Theorem 2 *The invariance group for $I_1 \times I_2 \times I_3$ contingency tables with fixed two-dimensional marginals is*

$$\langle S_{I_1} \times S_{I_2} \times S_{I_3},\ \{H_{st} \mid s < t,\ I_s = I_t\}\rangle$$

if $\min(I_1, I_2, I_3) \geq 3$.

We give a proof of this theorem in Appendix.

4.4 Quasi-symmetry model of square two-way tables

Next model we consider is the quasi-symmetry model of square two-way contingency tables with their diagonal cells being structural zeros. It is shown in Aoki and Takemura (2005) that there is a unique minimal Markov basis for this model. For the frequency vector $\boldsymbol{x} = \{x_{ij}\}_{1 \leq i \neq j \leq I}$, the elements of the fixed marginals, i.e., sufficient statistic, $\boldsymbol{t}$ are given as

$$\left\{\sum_{j=1}^{I} x_{ij}\right\}_{1 \leq i \leq I},\ \left\{\sum_{i=1}^{I} x_{ij}\right\}_{1 \leq j \leq I},\ \{x_{ij} + x_{ji}\}_{1 \leq i < j \leq I}.$$

For this model, an intuitively natural subgroup is the group generated by permuting levels "for both axes simultaneously" and by permuting axes.

Theorem 3 *The invariance group for the quasi-symmetry model of $I \times I$ two-way contingency table with its diagonal cells being structural zero cells is $\langle S_I, H_{12}\rangle$ for $I \neq 3$, where the symmetric group S_I of order I acts on the set of cells $\{ij \mid 1 \leq i \neq j \leq I\}$, as $g : ij \mapsto g(i)g(j)$.*

For $I = 3$, there is only 1 degree of freedom for $\ker(A)$. In this case, the single move $\boldsymbol{z} = (12)(23)(31) - (13)(32)(21)$ constitutes the lattice basis of $M = \ker(A)$ since $\dim(M) = 1$. Therefore any element $g \in S_6$ satisfying $g\boldsymbol{z} = \boldsymbol{z}$ or $g\boldsymbol{z} = -\boldsymbol{z}$ is a member of G. Therefore in the following proof we assume $I \geq 4$, since the case $I = 2$ is trivial.

Proof. Write $\widetilde{G} = \langle S_I, H_{12}\rangle$ as our candidate group. We show that for any $g \in G$ there exists some $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g} = e$. Similarly to the proofs of Theorem 1 and 2, we can assume $g(12) = 12$ without loss of generality, since otherwise there exists some $i'j'$ satisfying $g(i'j') = 12$, and we can choose $\tilde{g} \in \widetilde{G}$ satisfying $(g\tilde{g})(12) = 12$.

First consider the action of $g \in G_{12}$ to a move

$$\boldsymbol{z}_1 = (12)(23)(31) - (13)(32)(21).$$

Since $g(12) = 12$, we can assume $g\boldsymbol{z}_1 = \boldsymbol{z}_1$ without loss of generality, since otherwise there exists some $i' \geq 3$ satisfying

$$g\boldsymbol{z}_1 = (12)(2i')(i'1) - (1i')(i'2)(21),$$

and we can choose $\tilde{g} \in \widetilde{G}$ as $\tilde{g} = [3, i'] \in S_I$ such that $(g\tilde{g})\boldsymbol{z}_1 = \boldsymbol{z}_1$. Therefore we have

$$\{g(23), g(31)\} = \{23, 31\},\ \{g(13), g(32), g(21)\} = \{13, 32, 21\}.$$

However, by considering the action of g to a move

$$\boldsymbol{z}_2 = (12)(24)(41) - (14)(42)(21),$$

we have $g(21) = 21$. Moreover, similarly to $\boldsymbol{z}_1$, we can assume $g\boldsymbol{z}_2 = \boldsymbol{z}_2$ without loss of generality. Therefore we have

$$\begin{aligned} &g(12) = 12,\ g(21) = 21,\\ &\{g(23), g(31)\} = \{23, 31\},\ \{g(13), g(32)\} = \{13, 32\},\\ &\{g(24), g(41)\} = \{24, 41\},\ \{g(14), g(42)\} = \{14, 42\}. \end{aligned} \tag{22}$$

Now consider the action of g to a move

$$\boldsymbol{z}_3 = (23)(34)(42) - (24)(43)(32).$$

By (22), there are two possible cases,

$$\begin{aligned} &(g(23), g(31)) = (23, 31),\ (g(13), g(32)) = (13, 32),\\ &(g(24), g(41)) = (24, 41),\ (g(14), g(42)) = (14, 42) \end{aligned} \tag{23}$$

or

$$\begin{aligned} &(g(23), g(31)) = (31, 23),\ (g(13), g(32)) = (32, 13),\\ &(g(24), g(41)) = (41, 24),\ (g(14), g(42)) = (42, 14). \end{aligned} \tag{24}$$

For the former case (23), we also have $g(34) = 34$ and $g(43) = 43$, i.e., we have $g(ij) = ij$ for all $1 \le i \ne j \le 4$. Now we show

$$\begin{gathered} g(i5) = i5,\ g(5i) = 5i,\ i = 1, \dots, 4, \\ g(i6) = i6,\ g(6i) = 6i,\ i = 1, \dots, 5, \\ \vdots \\ g(iI) = iI,\ g(Ii) = Ii,\ i = 1, \dots, I-1, \end{gathered}$$

inductively. Suppose we have $g(ij) = ij$ for all $1 \le i \ne j < J$ for some $4 < J < I$. Then by considering the actions of g to the moves

$$\begin{gathered} (12)(2J)(J1) - (1J)(J2)(21), \\ (23)(3J)(J2) - (2J)(J3)(32), \\ \vdots \\ (J''J')(J'J)(JJ'') - (J''J)(JJ')(J'J''), \end{gathered} \tag{25}$$

where $J' = J - 1$ and $J'' = J - 2$, it follows that

$$g(iJ) = iJ,\ g(Ji) = Ji,\ i = 1, \dots, J-1.$$

Consequently, we have shown that $g\tilde{g} = e$ for some $\tilde{g} \in \widetilde{G}$.

For the latter case (24), we have $g(34) = 43$ and $g(43) = 34$, which implies that g is a group element of interchanging the first and the second levels for both axes simultaneously, and interchanging axes for the upper left 4×4 subtable. Again we see inductively that g has the same structure for all the elements, i.e.,

$$\begin{gathered} g(15) = 52,\ g(25) = 51,\ g(i5) = 5i,\ i = 3, \dots, 4, \\ g(51) = 25,\ g(52) = 15,\ g(5i) = i5,\ i = 3, \dots, 4, \\ g(16) = 62,\ g(26) = 61,\ g(i6) = 6i,\ i = 3, \dots, 5, \\ g(61) = 26,\ g(62) = 16,\ g(6i) = i6,\ i = 3, \dots, 5, \\ \vdots \\ g(1I) = I2,\ g(2I) = I1,\ g(iI) = Ii,\ i = 3, \dots, I-1, \\ g(I1) = 2I,\ g(I2) = 1I,\ g(Ii) = iI,\ i = 3, \dots, I-1, \end{gathered}$$

by considering the moves (25) for $J = 5, \dots, I-1$. Consequently, we have shown that g is a group element of interchanging the first and the second levels for both axes simultaneously, and interchanging axes, which is also canceled by some $\tilde{g} \in \widetilde{G}$. Q.E.D.

4.5 Hardy-Weinberg model

Another model where G has a simple structure is the Hardy-Weinberg model, which we have considered in Section 1. We assume that there are I distinct alleles. $\boldsymbol{x} = \{x_{ij}\}_{1 \le i \le j \le I}$ is the allele frequency vector and A is written as

$$A = (A_I\ A_{I-1}\ \cdots\ A_1), \quad A_k = \left(O_{k \times (I-k)}\ B_k'\right)',$$

where B_k is the following $k \times k$ square matrix

$$B_k = \begin{bmatrix} 2 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

In this case, an intuitively natural group element is characterized as"interchanging alleles", which is formalized as follows. The symmetric group S_I acts on the set of cells $\{ij \mid 1 \leq i \leq j \leq I\}$ as

$$g : ij \mapsto \begin{cases} g(i)g(j), & \text{if } g(i) \leq g(j), \\ g(j)g(i), & \text{if } g(i) > g(j). \end{cases} \tag{26}$$

Theorem 4 *For the Hardy-Weinberg model of I-alleles the invariance group is S_I, where the action of S_I on the set of cells is defined in (26).*

Proof. Write $\widetilde{G} = S_I$ as a candidate group and let $g \in G$. First we note that a homozygote (i.e., diagonal) cell, $11, 22, \ldots, II$, is mapped to a homozygote cell by g, and a heterozygote (i.e., off-diagonal) cell, $12, 13, \ldots, (I-1)I$, is mapped to a heterozygote cell by g, respectively. To show this, consider a move $\boldsymbol{z}_{ij} = (ii)(jj) - (ij)(ij)$ for some $1 \leq i < j \leq I$. Since $\boldsymbol{z}_{ij}$ has one -2 heterozygote cell and two $+1$ homozygote cells, $g\boldsymbol{z}_{ij}$ also has one -2 cell and two $+1$ cells by definition. However, to ensure that $g\boldsymbol{z}_{ij}$ is a move, it follows that -2 cell must again be a heterozygote cell, and two $+1$ cells again must be homozygote cells. Therefore we can write $g\boldsymbol{z}_{ij} = (i'i')(j'j') - (i'j')(i'j')$ for some $1 \leq i' < j' \leq I$.

From the above considerations, it is easy to show that we can assume $g(ii) = ii$ for $i = 1, \ldots, I$ without loss of generality, because otherwise we can choose some $\tilde{g} \in \widetilde{G}$ so that $(g\tilde{g})(ii) = ii$ for $i = 1, \ldots, I$. Moreover, by considering the actions of g to $\boldsymbol{z}_{ij}$ again, it follows that $g(ij) = ij$ for all $1 \leq i < j \leq I$. This completes the proof. Q.E.D.

5 Application to invariant Markov bases

In this section, we apply the notion of the invariance group to the theory of invariant Markov basis discussed in Aoki and Takemura (2003b). There we considered a subgroup G_0 in (13) generated only by the group elements of "permuting levels for each axis". The arguments in Aoki and Takemura (2003b) is focused on a concise expression of Markov basis by the list of representative element for each orbit for the action of G_0 on $\mathcal{M}$. Since the invariance group of this paper is the largest, it yields the maximum reduction.

In this section, first we give a definition of invariant Markov basis and summarize the structure of minimal invariant Markov basis from Aoki and Takemura (2003b) in Section 5.1. Next, in Section 5.2, we give lists of representative elements of minimal invariant Markov basis for the invariance groups of this paper.

5.1 Definition of invariant Markov basis and the structure of minimal invariant Markov basis

Let $\mathcal{B} \subset \mathcal{M}_A$ be a set of moves and let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{F}_{\boldsymbol{t}}$. We say that $\boldsymbol{x}_2$ is *accessible* from $\boldsymbol{x}_1$ by $\mathcal{B}$ if there exists a sequence of moves $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_S \in \mathcal{B}$ and $\varepsilon_s \in \{-1, 1\}$, $s = 1, \ldots, S$, such that

$$\begin{aligned} \boldsymbol{x}_2 &= \boldsymbol{x}_1 + \sum_{s=1}^{S} \varepsilon_s \boldsymbol{z}_s, \\ \boldsymbol{x}_1 + \sum_{s=1}^{r} \varepsilon_s \boldsymbol{z}_s &\in \mathcal{F}_{\boldsymbol{t}} \text{ for } 1 \le r \le S \ , \end{aligned} \tag{27}$$

i.e., we can apply moves from $\mathcal{B}$ to $\boldsymbol{x}_1$ one by one and go from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$, without causing negative cell frequencies on the way. Since the notion of accessibility is symmetric and transitive, the accessibility by $\mathcal{B}$ is an equivalence relation and each $\mathcal{F}_{\boldsymbol{t}}$ is partitioned into disjoint equivalence classes by $\mathcal{B}$. We call these equivalence classes $\mathcal{B}$-*equivalence classes* of $\mathcal{F}_{\boldsymbol{t}}$. If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are elements from two different $\mathcal{B}$-equivalence classes of the same $\boldsymbol{t}$-fiber, we say that a move $\boldsymbol{z} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ *connects* these two equivalence classes. See Section 2 of Takemura and Aoki (2004a).

Here we give a definition of a Markov basis. A set of finite moves $\mathcal{B} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_L\} \subset \mathcal{M}_A$ is a *Markov basis* if for all $\boldsymbol{t} \in \mathbb{N}^\nu$, $\mathcal{F}_{\boldsymbol{t}}$ itself constitutes one $\mathcal{B}$-equivalence class. As is stated in Aoki and Takemura (2003b), there is a freedom of the signs of the elements of a Markov basis. In fact, if $\mathcal{B}$ is a Markov basis and $\boldsymbol{z}, -\boldsymbol{z} \in \mathcal{B}$, then $\mathcal{B} \setminus \{\boldsymbol{z}\}$ and $\mathcal{B} \setminus \{-\boldsymbol{z}\}$ are also Markov bases, respectively, and if we replace any element $\boldsymbol{z}$ of a Markov basis with $-\boldsymbol{z}$, the resulting set is again a Markov basis. Similarly to Aoki and Takemura (2003b), we identify an element $\boldsymbol{z}$ of a Markov basis with its sign change $-\boldsymbol{z}$ for convenience in this paper. It should be noted again that, similarly to the $\boldsymbol{t}$-fibers, the concept of a move, and therefore Markov basis, depends only on $\ker(A)$.

A Markov basis $\mathcal{B}$ is *minimal* if no proper subset of $\mathcal{B}$ is a Markov basis. A minimal Markov basis always exists, because from any Markov basis, we can remove redundant elements one by one, until none of the remaining elements can be removed any further. However, a minimal Markov basis is not always unique. Takemura and Aoki (2004a) gave some characterizations of a minimal Markov basis.

Now we can define an invariant set of moves. Let G is a subgroup of S_p which acts on $\mathcal{M}_A$. Then $\mathcal{B} \subset \mathcal{M}_A$ is G-*invariant* if $G(\mathcal{B}) = \mathcal{B}$. Note that here we are identifying a move $\boldsymbol{z} \in \mathcal{B}$ with its sign change $-\boldsymbol{z}$. Therefore $\mathcal{B}$ is G-invariant if and only if

$$\forall g \in G, \forall \boldsymbol{z} \in \mathcal{B} \implies g\boldsymbol{z} \in \mathcal{B} \text{ or } -g\boldsymbol{z} \in \mathcal{B} \ .$$

A finite set $\mathcal{B} \subset \mathcal{M}_A$ is an *invariant Markov basis* for A if it is a Markov basis and it is G-invariant. An invariant Markov basis is *minimal* if no proper G-invariant subset of $\mathcal{B}$ is a Markov basis. A minimal invariant Markov basis always exists, because from any invariant Markov basis, we can remove orbits one by one, until none of the remaining orbits can be removed any further.

Following the arguments in Takemura and Aoki (2004a), two special sets of moves play important roles to characterize minimality of a Markov basis. One is the set of moves $\boldsymbol{z} =$

$\boldsymbol{z}^+ - \boldsymbol{z}^-$ with $A\boldsymbol{z}^+ = A\boldsymbol{z}^-$:

$$\mathcal{M}_{\boldsymbol{t}} = \{\boldsymbol{z} \in \mathcal{M}_A \mid \boldsymbol{t} = A\boldsymbol{z}^+ = A\boldsymbol{z}^-\}.$$

The other is a set of moves with degree less than or equal to n:

$$\mathcal{M}_n = \{\boldsymbol{z} \in \mathcal{M}_A \mid \deg(\boldsymbol{z}) \le n\}.$$

In addition, following Aoki and Takemura (2003b), we define another important set of moves. Consider a group action of G on the set of $\boldsymbol{t}$'s, $\mathcal{T} = A\mathbb{N}^p$. Let $G(\boldsymbol{t}) \in \mathcal{T}/G$ denote the orbit through a particular $\boldsymbol{t} \in \mathcal{T}$. Then we write

$$\mathcal{M}_{G(\boldsymbol{t})} = \bigcup_{\boldsymbol{t}' \in G(\boldsymbol{t})} \mathcal{M}_{\boldsymbol{t}'}$$

denotes the union of sets of moves $\mathcal{M}_{\boldsymbol{t}'}$ over the orbit $G(\boldsymbol{t})$ through $\boldsymbol{t}$.

Let $\mathcal{B} \subset \mathcal{M}_A$ be a finite set of moves. An important observation is that $\mathcal{B}$ is partitioned as

$$\mathcal{B} = \bigcup_{\alpha \in \mathcal{T}/G} \mathcal{B}_\alpha, \tag{28}$$

where $\mathcal{B}_\alpha = \mathcal{B} \cap \mathcal{M}_\alpha$, $\alpha \in \mathcal{T}/G$. Since $\mathcal{B}$ is invariant if and only if it is a union of orbits $G(\boldsymbol{z})$, $\mathcal{B}$ is invariant if and only if $\mathcal{B}_\alpha$ is invariant for each $\alpha \in \mathcal{T}/G$ (Lemma 2 of Aoki and Takemura, 2003b). In characterizing a Markov basis and its minimality, it is essential to consider $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$. In Aoki and Takemura (2003b), it is shown that the relation between the action of the stabilizer $G_{\boldsymbol{t}}$ and $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ is important. We write the set of $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ as $\mathcal{H}_{\boldsymbol{t}} = \mathcal{F}_{\boldsymbol{t}}/\mathcal{M}_{|\boldsymbol{t}|-1}$.

Now we give the main theorem of Aoki and Takemura (2003b) without proof.

Theorem 5 (Theorem 1 of Aoki and Takemura, 2003b) *Let $\mathcal{B}$ be a minimal G-invariant Markov basis and Let $\mathcal{B} = \bigcup_{\alpha \in \mathcal{T}/G} \mathcal{B}_\alpha$ be the partition in (28). Then each $\mathcal{B}_\alpha$, $\alpha \in \mathcal{T}/G$, is a minimal invariant set of moves, where $\mathcal{B}_\alpha \cap \mathcal{M}_{\boldsymbol{t}}$, $\boldsymbol{t} \in \alpha$, connects $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$ and*

$$\mathcal{B}_\alpha = G(\mathcal{B}_\alpha \cap \mathcal{M}_{\boldsymbol{t}}) \tag{29}$$

for any $\boldsymbol{t} \in \alpha$.

Conversely, from each $\alpha \in \mathcal{T}/G$ with $|\mathcal{H}_{\boldsymbol{t}}| \ge 2$, where $\boldsymbol{t} \in \alpha$ is a representative sufficient statistic, choose a minimal $G_{\boldsymbol{t}}$-invariant set of moves $\mathcal{B}_{\boldsymbol{t}} \subset \mathcal{M}_{\boldsymbol{t}}$ connecting $\mathcal{M}_{|\boldsymbol{t}|-1}$-equivalence classes of $\mathcal{F}_{\boldsymbol{t}}$, where $G_{\boldsymbol{t}} \subset G$ is the stabilizer of $\boldsymbol{t}$. Then

$$\mathcal{B} = \bigcup_{\substack{\alpha \in \mathcal{T}/G \\ |\mathcal{H}_{\boldsymbol{t}}| \ge 2, \boldsymbol{t} \in \alpha}} G(\mathcal{B}_{\boldsymbol{t}})$$

is a minimal G-invariant Markov basis.

5.2 Minimal invariant Markov bases for two-way and three-way models

As is stated in Aoki and Takemura (2003b), the main reason for considering invariant Markov bases is a concise expression of Markov bases by the list of representative moves for each orbit. Often the output of reduced Gröbner basis computation includes tens of thousands of elements, even when the number of orbits is only a few. We give the lists of representative moves of minimal invariant Markov basis for each problem considered in Section 4. It should be noted that the following lists are very concise.

Independence model for $I \times J$ contingency table
A representative move is $(11)(22)-(12)(21)$, i.e., minimal invariant Markov basis is written as

$$G((11)(22)-(12)(21)).$$

In this case, the unique minimal Markov basis exists, and it equals the unique minimal invariant Markov basis given above.

Quasi-symmetric model for $I \times I$ contingency table
The minimal invariant Markov basis consists of $I-2$ orbits, with the representative moves given as

$$\begin{gathered}(12)(23)(31)-(13)(32)(21),\\ (12)(23)(34)(41)-(14)(43)(32)(21),\\ \vdots\\ (12)(23)\cdots(I'I)(I1)-(1I)(II')\cdots(32)(21),\end{gathered}$$

where $I'=I-1$. In this case, the unique minimal Markov basis exists, and it equals the unique minimal invariant Markov basis given above.

Hardy-Weinberg model for I alleles
The minimal invariant Markov basis consists of 3 orbits, with the representative moves given as

$$\begin{gathered}(11)(22)-(12)(12),\\ (11)(23)-(12)(13),\\ (12)(34)-(13)(24).\end{gathered}$$

In this case, the unique minimal Markov basis does not exist as we have seen in Example 2 of Section 1. However, the minimal invariant Markov given above is the unique minimal invariant Markov basis.

Complete independence model $\{\{1\},\{2\},\{3\}\}$ for $I \times J \times K$ contingency table
Since the structure of G varies with the size of the table, the number of orbits also varies with the size of the table. Representative moves of minimal invariant Markov basis are given as follows.

- Case of $I \neq J \neq K \neq I$:

$$\begin{gathered}(111)(122) - (112)(121),\\ (111)(212) - (112)(211),\\ (111)(221) - (121)(211),\\ (111)(222) - (112)(221).\end{gathered}$$

- Case of $I \neq J = K$:

$$\begin{gathered}(111)(122) - (112)(121),\\ (111)(212) - (112)(211),\\ (111)(222) - (112)(221).\end{gathered}$$

- Case of $I = J = K$:

$$\begin{gathered}(111)(122) - (112)(121),\\ (111)(222) - (112)(221).\end{gathered}$$

In this case, the unique minimal Markov basis does not exist as we have seen in Example 1 of Section 1. However, the minimal invariant Markov given above is the unique minimal invariant Markov basis.

Conditional independence model $\{\{1,2\},\{1,3\}\}$ for $I \times J \times K$ contingency table
Similarly to the previous problem, the representative move of minimal invariant Markov basis is only $(111)(122) - (112)(121)$. The unique minimal Markov basis also exists, and it equals the unique minimal invariant Markov basis.

No three-factor interaction model $\{\{1,2\},\{1,3\},\{2,3\}\}$ for $I \times J \times K$ contingency table
For this problem, we do not know the general expressions of minimal Markov basis for general I, J, K. Aoki and Takemura (2003a) and Aoki and Takemura (2003c) gave a closed form expressions of minimal Markov bases for the problems of small sizes, and show that they are the unique minimal Markov bases. Of course, they are also the unique minimal invariant Markov basis. Here we give orbit expressions by representative moves for the unique minimal invariant Markov bases for some small size problems.

- For $(I, J, K) = (2, 3, 3), (2, 3, 4)$ or $(3, 3, 3)$, representative moves are

$$\begin{gathered}(111)(122)(212)(221) - (112)(121)(211)(222),\\ (111)(122)(133)(212)(223)(231) - (112)(123)(131)(211)(222)(233).\end{gathered}$$

- For $(I, J, K) = (2, 4, 4)$, representative moves are

$$\begin{gathered}(111)(122)(212)(221) - (112)(121)(211)(222),\\ (111)(122)(133)(212)(223)(231) - (112)(123)(131)(211)(222)(233),\\ (111)(122)(133)(144)(212)(223)(234)(241) - (112)(123)(134)(141)(211)(222)(233)(244).\end{gathered}$$

- For $(I, J, K) = (3, 3, 4)$, representative moves are

$$
\begin{gathered}
(111)(122)(212)(221) - (112)(121)(211)(222), \\
(111)(122)(133)(212)(223)(231) - (112)(123)(131)(211)(222)(233), \\
(111)(122)(221)(232)(312)(331) - (112)(121)(222)(231)(311)(332), \\
(111)(122)(213)(221)(234)(312)(324)(333) - (112)(121)(211)(224)(233)(313)(322)(334).
\end{gathered}
$$

- For $(I, J, K) = (3, 3, 5)$, representative moves are

$$
\begin{gathered}
(111)(122)(212)(221) - (112)(121)(211)(222), \\
(111)(122)(133)(212)(223)(231) - (112)(123)(131)(211)(222)(233), \\
(111)(122)(221)(232)(312)(331) - (112)(121)(222)(231)(311)(332), \\
(111)(122)(213)(221)(234)(312)(324)(333) - (112)(121)(211)(224)(233)(313)(322)(334), \\
(111)(122)(125)(134)(213)(221)(235)(312)(324)(333) - (112)(121)(124)(135)(211)(225)(233)(313)(322)(334).
\end{gathered}
$$

5.3 More symmetry of subspaces

Consider the complete independence model for $I \times J \times K$ contingency tables in the previous subsection. We see that we have minimum number of orbits for the case $I = J = K$, because in this case there is more symmetry than other cases and the invariance group contains permutations of the three axes. Assume $I \leq J \leq K$ without loss of generality. As a subtable of $I \times J \times K$ table, we have the table of size $I \times I \times I$. Even if $I < J < K$, we can permute the axes of moves totally contained in the subtable of size $I \times I \times I$ and can further reduce the number of orbits by exploiting the additional symmetry in subtables. We briefly discuss this additional reduction, although in this paper we do not pursue this topic in detail.

Let $\mathcal{J} \subset \mathcal{I}$ be a subset of cells. Define the $|\mathcal{J}|$-dimensional subspace $\mathbb{Q}^{\mathcal{J}}$ of $\mathbb{Q}^p$ by

$$
\mathbb{Q}^{\mathcal{J}} = \{\boldsymbol{x} \in \mathbb{Q}^p \mid x_{\boldsymbol{i}} = 0, \boldsymbol{i} \notin \mathcal{J}\}.
$$

$S_{(\mathcal{J}^C)}$, the set of permutations fixing all $\boldsymbol{i} \notin \mathcal{J}$, acts on $\mathbb{Q}^{\mathcal{J}}$. Furthermore let $M_{\mathcal{J}} = M \cap \mathbb{Q}^{\mathcal{J}}$. $M_{\mathcal{J}}$ corresponds to the kernel of a submatrix $A_{\mathcal{J}}$ of A, obtained by selecting columns of A corresponding to cells $\boldsymbol{i} \in \mathcal{J}$. Define

$$
G_{M_{\mathcal{J}}} = \{P \in S_{(\mathcal{J}^C)} \mid M_{\mathcal{J}} = P M_{\mathcal{J}}\}.
$$

Then in general we have

$$
G_{(\mathcal{J}^C)} = G_M \cap S_{(\mathcal{J}^C)} \subset G_{M_{\mathcal{J}}}, \tag{30}
$$

where G_M is the invariance group for A. The point is that the inclusion in (30) may be strict as shown in the example above and then we could employ this additional symmetry for further reduction of the list of representative elements.

More specifically we may proceed as follows. For a move $\boldsymbol{z}$, define its support $\mathrm{supp}(\boldsymbol{z})$ by $\mathrm{supp}(\boldsymbol{z}) = \{\boldsymbol{i} \mid \boldsymbol{z}(\boldsymbol{i}) \neq 0\}$. To each representative move $\boldsymbol{z}$, we specify $\mathcal{J} \supset \mathrm{supp}(\boldsymbol{z})$ and the invariance group $G_{M_{\mathcal{J}}}$ for the subtable. This may result in further reduction of the list of representative elements, although we need to specify the invariance group $G_{M_{\mathcal{J}}}$ for each $\boldsymbol{z}$.

6 Some discussions

In this paper, we give a definition of a subgroup of a symmetric group, which is characterized as the largest subgroup acting on $\ker(A)$. This definition provides an extension of the subgroup we considered in our previous work, Aoki and Takemura (2003b).

Concerning various notions of Markov bases, some relevant Markov bases, such as the Graver bases, the universal Gröbner basis, the minimal fiber Markov basis (Takemura and Aoki, 2005), are invariant with respect to the invariance group, since they are defined independent of any specific term order. Similarly, the set of indispensable moves, the set of circuits and the set of fundamental moves (Ohsugi and Hibi, 2005) are also invariant.

It is an interesting problem to express the invariance group as the permutation of rows of A as we show in Section 3.2. For the complete independence models of m-way contingency tables with mutually distinct levels of axes, we give the form of A explicitly so that the invariance group can be expressed as the permutation of rows. For general hierarchical models, on the other hand, it is still an open problem to verify the largest invariance group as $H(A)$.

For the two-way and three-way problems in Section 4, it is easy to describe a generating system of the invariance group. Using these representations, we can easily treat the invariance groups by computational group algebra softwares such as *GAP*. One example of utilization of *GAP* for Markov basis application is that we can obtain a group element randomly by the command `Random`. Random samples from our group enables us to perform a Markov chain Monte Carlo procedure using the list of representative elements of Markov basis in Section 5. In addition, it is of interest to characterize the invariance groups from viewpoint of the theory of finite groups. For example, by using *GAP*, we obtain two generators of the invariance group for the 3×3 independence model

$$
\begin{aligned}
&G = \langle g, h \rangle\,, \\
&g = [12, 21][13, 31][23, 32],\ \ h = [11, 23, 31, 13, 21, 33][12, 22, 32],
\end{aligned}
$$

where $[\boldsymbol{i}_1, \ldots, \boldsymbol{i}_m] \in S_p$ denotes the cyclic permutation $\boldsymbol{i}_1 \mapsto \boldsymbol{i}_2 \mapsto \cdots \mapsto \boldsymbol{i}_m \mapsto \boldsymbol{i}_1$. The fundamental relation for G is given as

$$
g^2 = h^6 = (gh)^4 = (h^5 ghg)^2 = e.
$$

Another important topic for investigation is to consider our invariance group in view of the existing large literature on computational invariant theory (e.g. Sturmfels, 1993, and Derksen and Kemper, 2002), although computational invariant theory is mainly concerned with the set of invariant polynomials with respect the action of a given subgroup of general linear group, whereas our invariance group is concerned as the setwise stabilizer for the action of symmetric group on a given set of homogeneous binomials.

Appendix: Proof of Theorem 2

Similarly to Theorem 1, take $\widetilde{G} = \langle S_{I_1} \times S_{I_2} \times S_{I_3},\ \{H_{st} \mid s < t,\ I_s = I_t\}\rangle$ as our candidate group. Hereafter, we write $I = I_1, J = I_2, K = I_3$ for simplicity and assume $I \leq J \leq K$ and $K \geq 3$ without loss of generality. To show $\widetilde{G} \supset G$, we show that for any $g \in G$ we can choose $\tilde{g} \in \widetilde{G}$ so that $g\tilde{g} = e$. Similarly to Theorem 1, we specify all the images of $\boldsymbol{i} \in \mathcal{I} = \{ijk \mid 1 \le i \le I, 1 \le j \le J, 1 \le k \le K\}$ one by one by applying $g \in G$.

Similarly to the proof of Theorem 1, we decompose $\mathcal{I}$ as

$$\mathcal{I} = \mathcal{I}_0 \cup \mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3,$$

where

$$\mathcal{I}_\ell = \{ijk \mid \ell \text{ of } \{i,j,k\} \text{ is/are 1 or 2}\}.$$

Therefore $\mathcal{I}_3 = \{1,2\} \times \{1,2\} \times \{1,2\}$, $\mathcal{I}_0 = \{3,\dots,I\} \times \{3,\dots,J\} \times \{3,\dots,K\}$ and so on. As in the proof of Theorem 1 we consider the pointwise stabilizer $G_{(\mathcal{I}_3)}$. We first show that for any $g \in G$ there exists some $\widetilde{g} \in \widetilde{G}$ such that $g\widetilde{g} \in G_{(\mathcal{I}_3)}$, and next show that for any $g \in G_{(\mathcal{I}_3)}$ there exists some $\widetilde{g} \in \widetilde{G}$ such that $g\widetilde{g} = e$.

It is obvious that for any $g \in G$ there exists $\widetilde{g} \in \widetilde{G}$ such that $g\widetilde{g} \in G_{111}$, because if $g(111) \neq 111$, there is $(i,j,k) \neq (1,1,1)$ satisfying $g(ijk) = 111$ and we see that $(g\tilde{g})(111) = 111$ by choosing $\tilde{g} = [1,i] \times [1,j] \times [1,k] \in \widetilde{G}$. Now consider the action of $g \in G_{111}$ on the move $\boldsymbol{z} = (111)(122)(212)(221) - (112)(121)(211)(222)$. It is seen that we can choose $\tilde{g} \in \widetilde{G}$ as the permutation of levels for each axis such that $g\tilde{g}$ is in the setwise stabilizer of the positive and negative cells of $\boldsymbol{z}$, i.e,

$$g\tilde{g} \in G_{\{122,212,221\}} \cap G_{\{112,121,211,222\}}.$$

As in other proofs of this paper, we can replace g by $g\tilde{g}$ and we have

$$g \in G_{111} \cap G_{\{122,212,221\}} \tag{31}$$

and

$$g \in G_{\{112,121,211,222\}}. \tag{32}$$

We consider all the possible cases of (31) and (32), one by one. Consider the constraint (31). There are $3! = 6$ possibilities, which we consider in the following.

- Case 1: $(g(122), g(212), g(221)) = (122, 212, 221)$.
  We shall show that other cases are reduced to this case.

- Case 2: $(g(122), g(212), g(221)) = (122, 221, 212)$.
  Considering the actions of g on moves

  $$\boldsymbol{z}_k = (111)(12k)(221)(21k) - (121)(11k)(211)(22k),\ k = 3,\dots,K,$$

  we have $\{g(121), g(211)\} = \{112, 211\}$ and for each $k = 3,\dots,K$, $\{g(12k), g(21k)\} = \{1j2, 2j1\}$ and $\{g(11k), g(22k)\} = \{1j1, 2j2\}$ for some $3 \le j \le J$. However, since the group action is bijective and $J \le K$, it follows that $J = K$. Consequently, since $\widetilde{G}$ includes permuting the j-th axis and the k-th axis, this case is reduced to Case 1 by replacing g by gg_{23}.

- Case 3: $(g(122), g(212), g(221)) = (212, 122, 221)$.
  If $J = I = 2$, we do not have to consider this case, since it reduces to Case 1 by permuting

the i-th axis and the j-th axis. If $J \geq 3$, similarly to Case 2, considering the actions of g on moves

$$(111)(1j2)(212)(2j1) - (112)(1j1)(211)(2j2),\ j = 3, \ldots, J,$$

we have again $I = J$. Therefore this case reduces to Case 1 by permuting the i-th axis and the j-th axis.

- Case 4: $(g(122), g(212), g(221)) = (221, 212, 122)$.
  Similarly to Case 2 again, considering the actions of g on the moves

  $$(111)(12k)(221)(21k) - (121)(11k)(211)(22k),\ k = 3, \ldots, K,$$

  we have $I = K$. Therefore this case reduces to Case 1 by permuting the i-th axis and the k-th axis.

- Case 5: $(g(122), g(212), g(221)) = (212, 221, 122)$.
  Considering the actions of g on the moves

  $$(111)(12k)(221)(21k) - (121)(11k)(211)(22k),\ k = 3, \ldots, K,$$

  we have $I = K$. Therefore by permuting the i-th axis and the k-th axis, this case reduces to Case 3, and we need not consider this case.

- Case 6: $(g(122), g(212), g(221)) = (221, 122, 212)$.
  Similarly to Case 5, we have $J = K$ by considering the actions of g on the same moves to Case 5, and we see that this case is also reduced to Case 3 by permuting the j-th axis and the k-th axis.

From the above considerations, we have $(g(122), g(212), g(221)) = (122, 212, 221)$, by choosing $\tilde{g} \in \widetilde{G}$ as the permutation of axes and replacing g by $g\tilde{g}$ if necessary. As for the constraint (32), applying g on two moves, $(111)(132)(212)(231)-(112)(131)(211)(232)$ and $(111)(123)(213)(221)-(121)(113)(211)(223)$ yields $(g(112), g(121), g(211)) = (112, 121, 211)$, and therefore $g(222) = (222)$. Then we have shown that for each $g \in G$ there exists some $\tilde{g} \in \widetilde{G}$ such that $g\tilde{g} \in G_{(\mathcal{I}_3)}$, i.e., $g\tilde{g}(\boldsymbol{i}) = \boldsymbol{i}$ for all $\boldsymbol{i} \in \mathcal{I}_3$.

Next we have to consider the cells in $\mathcal{I}_2, \mathcal{I}_1$ and $\mathcal{I}_0$. Since the arguments are similar, we give only an outline. It should be noted that hereafter we only have to consider $\tilde{g} \in \widetilde{G}$ as permutation of levels for each axis, i.e., permutations of axes have been fully considered above for any $g \in G_{(\mathcal{I}_3)}$. To show that $(g\tilde{g})(\boldsymbol{i}) = \boldsymbol{i}$ for $\boldsymbol{i} \in \mathcal{I}_2 \cup \mathcal{I}_1$, it suffices to consider the actions of g on the following moves.

$$\begin{array}{ll}
(111)(12k)(21k)(221) - (121)(11k)(211)(22k), & k = 3, \ldots, K, \\
(111)(1j2)(212)(2j1) - (112)(1j1)(211)(2j2), & j = 3, \ldots, J, \\
(111)(122)(i12)(i21) - (112)(121)(i11)(i22), & i = 3, \ldots, I, \\
(111)(1jk)(21k)(2j1) - (11k)(1j1)(211)(2jk), & j = 3, \ldots, J, k = 3, \ldots, K, \\
(111)(12k)(i1k)(i21) - (11k)(121)(i11)(i2k), & i = 3, \ldots, I, k = 3, \ldots, K, \\
(111)(1j2)(i12)(ij1) - (112)(1j1)(i11)(ij2), & i = 3, \ldots, I, j = 3, \ldots, J.
\end{array}$$

Moreover, to show that $(g\tilde{g})(\boldsymbol{i}) = \boldsymbol{i}$ for $\boldsymbol{i} \in \mathcal{I}_0$, it suffices to consider actions of g on the following moves.

$$(111)(1jk)(i1k)(ij1) - (11k)(1j1)(i11)(ijk),\ \ i = 3, \ldots, I,\ j = 3, \ldots, J,\ k = 3, \ldots, K.$$

Q.E.D.

References

[1] Aoki, S. and Takemura, A. (2003a) Minimal basis for connected Markov chain over $3\times3\times K$ contingency tables with fixed two-dimensional marginals. *Australian and New Zealand Journal of Statistics.* **45**, 229–249.

[2] Aoki, S. and Takemura, A. (2003b) Invariant minimal Markov basis for sampling contingency tables with fixed marginals. Technical Report METR 03-25, Department of Mathematical Engineering and Information Physics, The University of Tokyo. Submitted for publication.

[3] Aoki, S. and Takemura, A. (2003c) The list of indispensable moves of the unique minimal Markov basis for $3 \times 4 \times K$ and $4 \times 4 \times 4$ contingency tables with fixed two-dimensional marginals. Technical Report METR 03-38, Department of Mathematical Engineering and Information Physics, The University of Tokyo. Submitted for publication.

[4] Aoki, S. and Takemura, A. (2005) Markov chain Monte Carlo exact tests for incomplete two-way contingency tables. To appear in *Journal of Statistical Computation and Simulation.*

[5] Derksen, H. and Kemper, G. (2002) *Computational Invariant Theory.* Springer, Berlin.

[6] Diaconis, P. (1988) *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, Hayward, California.

[7] Diaconis, P. and Sturmfels, B. (1998) Algebraic algorithms for sampling from conditional distributions. *The Annals of Statistics*, **26**, 363–397.

[8] Dinwoodie, J. H. (1998) The Diaconis-Sturmfels algorithm and rules of succession. *Bernoulli*, **4**, 401–410.

[9] Dobra, A. (2003) Markov bases for decomposable graphical models. *Bernoulli*, **9**, 1093–1108.

[10] Hibi, T. (2003) *Gröbner Bases.* Asakura Shoten, Tokyo. (in Japanese).

[11] *Journal of Symbolic Computation.* (2005) A special issue on computational algebraic statistics, to appear.

[12] Ohsugi, H. and Hibi, T. (2005) Indispensable binomials of finite graphs. *Journal of Algebra and Its Applications*, to appear.

[13] Schrijver, A. (1986) *Theory of Linear and Integer Programming.* Wiley, New York.

[14] Seress, A. (2003) *Permutation Group Algorithms*, Cambridge University Press, Cambridge.

[15] Sturmfels, B. (1993) *Algorithms in Invariant Theory.*, Springer, Wien.

[16] Sturmfels, B. (1995) *Gröbner Bases and Convex Polytopes.* American Mathematical Society, Providence, RI.

[17] Takemura, A. and Aoki, S. (2004) Some characterizations of minimal Markov basis for sampling from discrete conditional distributions, *Annals of the Institute of Statistical Mathematics*, **56**, 1–17.

[18] Takemura, A. and Aoki, S. (2005) Distance reducing Markov bases for sampling from discrete space. *Bernoulli*, to appear.