On a perturbation method for determining group of invariance of hierarchical models

Tomonari SEI[1] Satoshi AOKI[2] Akimichi TAKEMURA[1]

[1]University of Tokyo

[2]Kagoshima University

Dec. 16, 2008

Table of contents

For details: Sei, Aoki & Takemura (2008) arXiv:0808.2725v1

Introduction

Introduction

- Consider a 2×3 contingency table:

p_{11}	p_{12}	p_{13}
p_{21}	p_{22}	p_{23}

- p_{ij} is probability and x_{ij} is observation.
- Consider the independence model $p_{ij} = a_i b_j$.
- The set of sufficient statistics is $\{x_{i+}\}$ and $\{x_{+j}\}$. ($x_{i+} = \sum_j x_{ij}$ etc.)
- A minimal Markov basis for this model is, for example,

$M_1 =$

1	−1	
−1	1	

, $M_2 =$

1		−1
−1		1

.

- M_2 is obtained by permutation of columns 2 and 3 of M_1.

Introduction

- Similarly, for $\mathbf{I} \times \mathbf{J}$ table, a Markov basis is obtained by

$$\mathbf{M_1} = \begin{array}{|r|r|r|c|}\hline 1 & -1 & & \cdots \\ \hline -1 & 1 & & \cdots \\ \hline & & & \cdots \\ \hline \vdots & \vdots & \vdots & \ddots \\ \hline \end{array}$$

  and its permutations of rows and columns. (i.e. generated by a single move if symmetry is considered.)
- However algorithms obtaining Gröbner bases and Markov bases do not utilize the symmetry at present.
- For a larger table, algorithms just take longer.
- If $\mathbf{I} = \mathbf{J}$, we can also consider a permutation of axes $\mathbf{x_{ij}} \leftrightarrow \mathbf{x_{ji}}$.

Introduction

- Example : three-way tables with fixed two-dimensional marginals (no-three-factor-interaction model)

Table: Number of elements in the unique minimal Markov basis and reduced Gröbner basis for $\mathbf{3} \times \mathbf{3} \times \mathbf{K}, \mathbf{K} \leq \mathbf{7}$.

K	**3**	**4**	**5**	**6**	**7**
# unique minimal MB	81	450	2670	10665	31815
# reduced GB	110	622	3240	12085	34790
# orbits in the MB	4	5	6	6	6

- The number of orbits remains 6 for all $\mathbf{K} \geq \mathbf{5}$ (Aoki and Takemura (2003)): "Markov complexity"

Introduction

- Our problem is... [Aoki & Takemura (2008) J.Symb.Comp.]
    - *Determine all permutations of cells that preserve a given configuration determining a toric ideal.*
    - *In particular, for the configuration associated with hierarchical models of contingency tables.*
- We call the set of allowable permutations the group of invariance.
- An interesting example is Sudoku invariance group (explained later)

Introduction

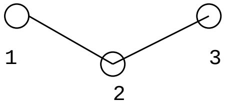
- Is our definition of the group of invariance useful?
- Why look at the largest symmetry?
- A configuration usually has an obvious symmetry. However it is difficult to prove that it is indeed the largest. (i.e. the is no more symmetry than the obvious symmetry.)
- Mathematically it makes sense to look at the largest symmetry.
- The result of this talk shows that our definition is useful, leading to a mathematically meaningful questions and results.

Group of invariance of hierarchical models

Definition of hierarchical models (1/2)

- $\mathbf{m}$: number of factors of contingency tables.
- $\mathcal{I}_\mathbf{j} = [\mathbf{I_j}] = \{\mathbf{1}, \ldots, \mathbf{I_j}\}$: the set of levels of the $\mathbf{j}$-th factor ($\mathbf{j} = \mathbf{1}, \ldots, \mathbf{m}$).
- $\mathcal{I} = \prod_{\mathbf{j}=\mathbf{1}}^{\mathbf{m}} \mathcal{I}_\mathbf{j}$: the set of cells.
- $\mathbf{\Delta}$: an abstract simplicial complex of $[\mathbf{m}] = \{\mathbf{1}, \ldots, \mathbf{m}\}$.
- $\mathbf{red\Delta}$: maximal simplices in $\mathbf{\Delta}$ w.r.t. inclusion order.

- For example, if $\mathbf{m} = \mathbf{3}$ and $\mathbf{\Delta} = \{\emptyset, \{\mathbf{1}\}, \{\mathbf{2}\}, \{\mathbf{3}\}, \{\mathbf{1}, \mathbf{2}\}, \{\mathbf{2}, \mathbf{3}\}\}$, then $\mathbf{red\Delta} = \{\{\mathbf{1}, \mathbf{2}\}, \{\mathbf{2}, \mathbf{3}\}\}$.



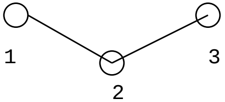
Definition of hierarchical models (2/2)

Definition

The hierarchical model specified by $\boldsymbol{\Delta}$ is the set of probability distributions $(\mathbf{p}_i)_{i\in\mathcal{I}}$ written in a log-linear form

$$\log \mathbf{p}_i = \sum_{\mathbf{D}\in \mathrm{red}\boldsymbol{\Delta}} \phi_{\mathbf{D}}(i_{\mathbf{D}}), \quad i_{\mathbf{D}} = (i_{\mathbf{j}})_{\mathbf{j}\in\mathbf{D}}.$$

- For example, let $\mathbf{m} = \mathbf{3}$ and $\mathrm{red}\boldsymbol{\Delta} = \{\{1,2\},\{2,3\}\}$.



- Then the hierarchical model is, by putting $(\mathbf{i},\mathbf{j},\mathbf{k}) = (\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3)$,

$$\log \mathbf{p}_{\mathbf{ijk}} = \alpha_{\mathbf{ij}} + \beta_{\mathbf{jk}}.$$

Definition of group of invariance (1/2)

- Let $\mathbf{S}_{\mathcal{I}}$ be the set of all permutations on the cells $\mathcal{I}$.
- Any $\mathbf{g} \in \mathbf{S}_{\mathcal{I}}$ acts on the linear space $\mathbb{Q}^{\mathcal{I}}$ (i.e. tables) by

$$(\mathbf{g}\theta)_i = \theta_{\mathbf{g}^{-1}(i)} \quad \text{for} \quad \theta \in \mathbb{Q}^{\mathcal{I}}.$$

- Now recall that the hierarchical model is $\log \mathbf{p}_i = \sum_{\mathbf{D} \in \mathbf{\Delta}} \phi_{\mathbf{D}}(i_{\mathbf{D}})$.
- Define a linear subspace of $\mathbb{Q}^{\mathcal{I}}$, the range of parameters, by

$$\mathbf{r}(\mathbf{\Delta}) = \left\{ \sum_{\mathbf{D} \in \mathrm{red}\mathbf{\Delta}} \phi_{\mathbf{D}}(i_{\mathbf{D}}) \;\middle|\; \phi_{\mathbf{D}} \in \mathbb{Q}^{\mathcal{I}_{\mathbf{D}}} \right\}.$$

(same as the row space of the configuration.)

Definition of group of invariance (2/2)

Definition

The group of invariance $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})}$ is the setwise stabilizer of $\mathbf{r}(\boldsymbol{\Delta})$:

$$\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})} := \{\mathbf{g} \in \mathbf{S}_{\mathcal{I}} \mid \mathbf{g}(\mathbf{r}(\boldsymbol{\Delta})) = \mathbf{r}(\boldsymbol{\Delta})\}.$$

- Remark that the subspace $\mathbf{r}(\boldsymbol{\Delta})$ is dual to the kernel of the sufficient statistics in $\mathbb{Q}^{\mathcal{I}}$, which contains Markov bases.
- The group of invariance of the kernel is the same as $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})}$.
- It is a linear-algebraic notion.
- It can be easily seen that $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})}$ also acts on the set of fibers.

Some known results

- For the completely independent model ($\mathbf{red}\boldsymbol{\Delta} = \{\{\mathbf{1}\}, \ldots, \{\mathbf{m}\}\}$), the group of invariance was derived by [Aoki & Takemura, JSC2008].
- In their paper, cases of $\mathbf{I} \times \mathbf{J} \times \mathbf{K}$ no three-factor model and the Hardy-Weinberg model (not hierarchical) were also solved.
- [Bailey et al. Proc. London Math. Soc. 1983] studied a related concept for design of experiments.

Example

- Consider the two-way independence model

$$\log p_{ij} \;=\; \alpha_i + \beta_j.$$

- If $|\mathcal{I}_1| \neq |\mathcal{I}_2|$, the the group of invariance is known to be $G_{r(\Delta)} = S_{\mathcal{I}_1} \times S_{\mathcal{I}_2}$, the direct product of the row and column permutations.

Wreath product

Wreath product

- Wreath product naturally arises in our problem.
- In general, the wreath product of two group actions $(\mathbf{G}, \mathbf{X})$ and $(\mathbf{H}, \mathbf{Y})$ is formally defined by

$$\mathbf{H}\mathrm{wr}\mathbf{G} \;=\; \mathbf{G} \times \mathbf{H}^{\mathbf{X}} \quad (\text{acts on } \mathbf{X} \times \mathbf{Y}),$$

  where $\mathbf{H}^{\mathbf{X}}$ denotes all functions from $\mathbf{X}$ to $\mathbf{H}$.
- Let us consider a table

A	B	C
D	E	F

- Let $\mathbf{S}_{\mathcal{I}_1}$ and $\mathbf{S}_{\mathcal{I}_2}$ be the permutation group of rows and columns, resp.
- Then, the wreath product $\mathbf{S}_{\mathcal{I}_2}\mathrm{wr}\mathbf{S}_{\mathcal{I}_1}$ is generated from
  - Permutation of rows, and
  - Permutation of columns in each row:

A	B	C
F	D	E

Why the wreath product arises?

- Now we explain why the wreath product arises in our problem.
- Consider the row-effect-only model: $\log p_{ij} = \alpha_i$.
- More visually, consider a table

α_1	α_1	α_1
α_2	α_2	α_2

.
- Then the wreath product $S_{\mathcal{I}_2} \mathrm{wr} S_{\mathcal{I}_1}$ preserves the range the row-effect-only model.
- Similarly $S_{\mathcal{I}_1} \mathrm{wr} S_{\mathcal{I}_2}$ preserves the range of the column-effect-only model.

- Since the range of the independence model is the vector sum of the above two models,

$$(\mathbf{S}_{\mathcal{I}_1} \mathrm{wr} \mathbf{S}_{\mathcal{I}_2}) \cap (\mathbf{S}_{\mathcal{I}_2} \mathrm{wr} \mathbf{S}_{\mathcal{I}_1})$$

  is a subgroup of of $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})}$ for the two-way independence model.
- Equality holds if $|\mathcal{I}_1| \neq |\mathcal{I}_2|$.
- Furthermore

$$(\mathbf{S}_{\mathcal{I}_1} \mathrm{wr} \mathbf{S}_{\mathcal{I}_2}) \cap (\mathbf{S}_{\mathcal{I}_2} \mathrm{wr} \mathbf{S}_{\mathcal{I}_1}) = \mathbf{S}_{\mathcal{I}_1} \times \mathbf{S}_{\mathcal{I}_2}.$$

  (even when $|\mathcal{I}_1| = |\mathcal{I}_2|$.)

Main theorem

Main Theorem 1

Theorem (Main Theorem 1)

Under a weak assumption on $\mathcal{I}$, the group of invariance is given by

$$\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})} = \bigcap_{\mathbf{D} \in \mathbf{red}(\boldsymbol{\Delta})} (\mathbf{S}_{\mathcal{I}_{\mathbf{D}^{\mathbf{C}}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathbf{D}}}),$$

where $\mathbf{S}_{\mathcal{I}_{\mathbf{D}}}$ is the set of all permutations on $\mathcal{I}_{\mathbf{D}}$.

Our assumption is

- $|\mathcal{I}_{\mathbf{D}}|$ are mutually distinct, and
- $\mathbf{I}_{\mathbf{j}} = |\mathcal{I}_{\{\mathbf{j}\}}| > \mathbf{2}$ except for at most one $\mathbf{j} \in \{\mathbf{1}, \ldots, \mathbf{m}\}$.

Even if the assumption fails, the inclusion $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})} \supset \bigcap_{\mathbf{D}} (\mathbf{S}_{\mathcal{I}_{\mathbf{D}^{\mathbf{C}}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathbf{D}}})$ is always true.

Generalized wreath product

- We want more explicit expression of the right-hand side.
- Wreath product is generalized for an indexed set of group actions $(\mathbf{G}_\rho, \mathbf{X}_\rho)_{\rho \in \mathcal{P}}$, where $\mathcal{P}$ is a poset.
- The generalized wreath product is defined by $\prod_{\rho \in \mathcal{P}} (\mathbf{G}_\rho)^{\mathbf{X}_{\mathbf{A}(\rho)}}$ [Wells1976], where $\mathbf{A}(\rho)$ denotes the ancestor set of ρ.
- The generalized version is useful for our problem, because it allows us to sample a random element of $\mathbf{G}_{\mathbf{r}(\Delta)}$.

Main Theorem 2

Theorem (Main theorem 2; can be deduced from Bailey et al. 1983)

The intersection is rewritten as a generalized wreath product:

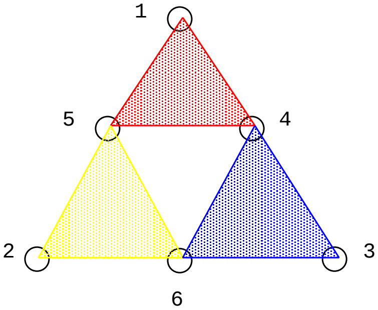
$$\bigcap_{\mathbf{D}\in\mathbf{red}(\mathbf{\Delta})} (\mathsf{S}_{\mathcal{I}_{\mathsf{D}^\mathsf{C}}} \mathbf{wr} \mathsf{S}_{\mathcal{I}_\mathsf{D}}) = \prod_{\rho\in\mathcal{P}} (\mathsf{S}_{\mathcal{I}_\rho})^{\mathcal{I}_{\mathsf{A}(\rho)}},$$

where the poset $\mathcal{P}$ is defined as follows.

- For each $\mathbf{i} \in [\mathbf{m}]$, define $(\mathbf{red}\mathbf{\Delta})(\mathbf{i}) := \{\mathbf{D} \in \mathbf{red}\mathbf{\Delta} \mid \mathbf{D} \ni \mathbf{i}\}$.
- We write $\mathbf{i} \sim \mathbf{j}$ if $(\mathbf{red}\mathbf{\Delta})(\mathbf{i}) = (\mathbf{red}\mathbf{\Delta})(\mathbf{j})$.
- $\mathcal{P} = [\mathbf{m}]/\sim$, the quotient space.
- Define a partial order $\rho \leq \rho'$ in $\mathcal{P}$ if $(\mathbf{red}\mathbf{\Delta})(\rho) \subset (\mathbf{red}\mathbf{\Delta})(\rho')$.

Example

- Let $\mathbf{m} = \mathbf{6}$ and $\mathbf{red\Delta} = \{\{\mathbf{1,4,5}\},\{\mathbf{2,5,6}\},\{\mathbf{3,4,6}\}\}$.



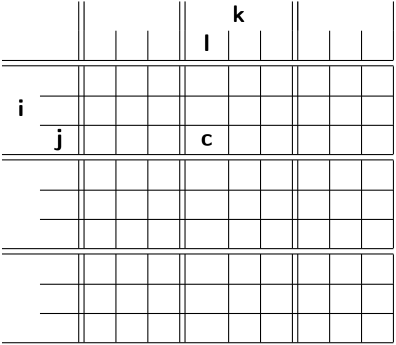
- From Theorem 1 and 2, the group of invariance is

$$\mathbf{G}_{\mathbf{r}(\mathbf{\Delta})} = (\mathbf{S}_{\mathcal{I}_1})^{\mathcal{I}_{\{4,5\}}} \times (\mathbf{S}_{\mathcal{I}_2})^{\mathcal{I}_{\{5,6\}}} \times (\mathbf{S}_{\mathcal{I}_3})^{\mathcal{I}_{\{4,6\}}} \times \mathbf{S}_{\mathcal{I}_4} \times \mathbf{S}_{\mathcal{I}_5} \times \mathbf{S}_{\mathcal{I}_6}$$

Sudoku contingency table (1/4)

- Sudoku is a puzzle on $\mathbf{9} \times \mathbf{9}$ contingency table.
- Each row, column and $\mathbf{3} \times \mathbf{3}$ block contains the $\mathbf{9}$ digits exactly once.
- For a sudoku solution, we define a $\mathbf{3^4} \times \mathbf{9}$ table $(\mathbf{x_{ijklc}})$ by

$\mathbf{x_{ijklc}} = \mathbf{1}$ if



, and $\mathbf{0}$ otherwise.

- **i**:band, **j**:row, **k**:stack, **l**:column, **c**:color.

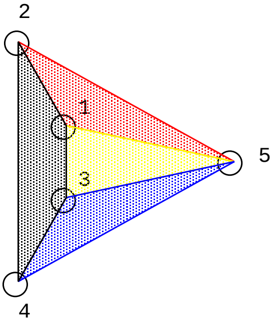
Sudoku contingency table (2/4)

- We add an additional factor (dimension) corresponding to the color.
- For example, if $\mathbf{c} = \mathbf{3}$ then we put a single frequency on the third level of the factor **c**.
- Then a sudoku solution has a one-to-one correspondence with $(\mathbf{x}_{\mathbf{ijklc}})$.
- Adding an additional dimension corresponds to "Lawrence lifting".
- Ordinary Lawrence lifting has just two levels in the additional dimension. Our case is a higher-order Lawrence lifting with 9 levels.
- The present treatment of Sudoku is different from the approach found in David A. Cox's tutorial (2007) on Gröbner basis approach to Sudoku. (He does not consider Lawrence lifting.)

Sudoku contingency table (3/4)

- A Sudoku solution satisfies $\mathbf{x_{ij++c}} = \mathbf{x_{i+k+c}} = \mathbf{x_{++klc}} = \mathbf{x_{ijkl+}} = \mathbf{1}$.
- This is a fiber of the model $\boldsymbol{\Delta}$, where

$$\mathbf{red}\boldsymbol{\Delta} = \{\{\mathbf{1,2,5}\},\{\mathbf{1,3,5}\},\{\mathbf{3,4,5}\},\{\mathbf{1,2,3,4}\}\}.$$



- We call it the Sudoku model.

Sudoku contingency table (4/4)

- Unfortunately, the Sudoku model does not satisfy the assumption in our main theorem because

$$|\mathcal{I}_{\{1,2,5\}}| = |\mathcal{I}_{\{1,3,5\}}| = |\mathcal{I}_{\{3,4,5\}}| = |\mathcal{I}_{\{1,2,3,4\}}| = 81.$$

- But, by the main theorem, we can deduce that $\mathbf{G}_{\mathbf{r}(\Delta)}$ contains

$$\mathbf{S}_{\mathcal{I}_1} \times (\mathbf{S}_{\mathcal{I}_2})^{\mathcal{I}_1} \times \mathbf{S}_{\mathcal{I}_3} \times (\mathbf{S}_{\mathcal{I}_4})^{\mathcal{I}_3} \times \mathbf{S}_{\mathcal{I}_5}.$$

- This group consists of
  - permutation of bands,
  - permutation of rows in each band,
  - permutation of stacks,
  - permutation of columns in each stack,
  - permutation of digits.
- But the transposition $(\mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}) \leftrightarrow (\mathbf{k}, \mathbf{l}, \mathbf{i}, \mathbf{j})$ is not included here.

How about classical Latin squares?

1	**2**	**3**
2	**3**	**1**
3	**1**	**2**

  (Just one block of Sudoku).
- By Lawrence lifting it corresponds to a $\mathbf{3} \times \mathbf{3} \times \mathbf{3}$ table.
- The model is the no-three-factor-interaction model.
- A Latin square is an element of the fiber with all marginals equal to 1.
- Researchers are interested in "non-isomorphic" Latin squares, i.e., in orbits of obvious group actions.

Perturbation method (for a proof of theorems)

Outline of the proof (simplest case)

- Consider $\mathbf{2} \times \mathbf{3}$ independence model again.
- The group of invariance is $\mathbf{S}_{\mathcal{I}_1} \times \mathbf{S}_{\mathcal{I}_2}$.
  i.e. permutation of rows and columns, resp.
- [Proof] Let $\theta = \begin{array}{|c|c|c|}\hline \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \end{array} \in \mathbf{r}(\mathbf{\Delta})$ and $\mathbf{g} \in \mathbf{G}_{\mathbf{r}(\mathbf{\Delta})}$.
- Since $\mathbf{g}\theta$ must be $\begin{array}{|c|c|c|}\hline \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \end{array}$ or $\begin{array}{|c|c|c|}\hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \end{array}$, we have $\mathbf{g} \in \mathbf{S}_{\mathcal{I}_2} \mathbf{wr} \mathbf{S}_{\mathcal{I}_1}$.
- Similarly, let $\theta = \begin{array}{|c|c|c|}\hline \mathbf{0} & \mathbf{1} & \mathbf{2} \\ \hline \mathbf{0} & \mathbf{1} & \mathbf{2} \\ \hline \end{array} \in \mathbf{r}(\mathbf{\Delta})$ and $\mathbf{g} \in \mathbf{G}_{\mathbf{r}(\mathbf{\Delta})}$.
- Then we can show that $\mathbf{g} \in \mathbf{S}_{\mathcal{I}_1} \mathbf{wr} \mathbf{S}_{\mathcal{I}_2}$.
- As already mentioned, $(\mathbf{S}_{\mathcal{I}_2} \mathbf{wr} \mathbf{S}_{\mathcal{I}_1}) \cap (\mathbf{S}_{\mathcal{I}_1} \mathbf{wr} \mathbf{S}_{\mathcal{I}_2}) = \mathbf{S}_{\mathcal{I}_1} \times \mathbf{S}_{\mathcal{I}_2}$.

Outline of the proof (not trivial case 1/2)

- Now we proceed to **2** $\times$ **4** independence table.
- Let $\theta = \begin{array}{|c|c|c|c|}\hline \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \end{array}$.
- Then $\mathbf{g}\theta$ is $\begin{array}{|c|c|c|c|}\hline \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \end{array}$ or $\begin{array}{|c|c|c|c|}\hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \end{array}$ ...?
- $\longrightarrow$ No! because $\mathbf{g}\theta$ can be $\begin{array}{|c|c|c|c|}\hline \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \hline \end{array}$ ($\in$ model)
- It needs a perturbation method. Now we explain it.

Outline of the proof (not trivial case 2/2)

- Now we prove the $\mathbf{2} \times \mathbf{4}$ case by using the perturbation method.
- Let $\theta =$

0	**1**	**2**	**4**
100	**101**	**102**	**104**

. Then $\theta \in \mathbf{r}(\boldsymbol{\Delta})$. "generic"
- Put $\phi := \mathbf{g}\theta$. Then $\phi \in \mathbf{r}(\boldsymbol{\Delta})$.
- We can write $\phi_{\mathbf{ij}} = \mathbf{a_{ij}} + \mathbf{b_{ij}}$, where $\mathbf{a_{ij}} \in \{\mathbf{0}, \mathbf{100}\}$, $\mathbf{b_{ij}} \in \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{4}\}$.
- Since $\phi_{11} + \phi_{22} - \phi_{12} - \phi_{21} = \mathbf{0}$, we have $\mathbf{a_{11}} + \mathbf{a_{22}} - \mathbf{a_{12}} - \mathbf{a_{21}} = \mathbf{0}$ and $\mathbf{b_{11}} + \mathbf{b_{22}} - \mathbf{b_{12}} - \mathbf{b_{21}} = \mathbf{0}$.
- By careful consideration, we obtain $\mathbf{a_{11}} = \mathbf{a_{12}} = \mathbf{a_{13}} = \mathbf{a_{14}} \in \{\mathbf{0}, \mathbf{100}\}$ and $\mathbf{a_{21}} = \mathbf{a_{22}} = \mathbf{a_{23}} = \mathbf{a_{24}} \in \{\mathbf{0}, \mathbf{100}\}$. Hence $\mathbf{g} \in \mathbf{S_4}\mathrm{wr}\mathbf{S_2}$.
- Similarly, from $\mathbf{b_{1j}} = \mathbf{b_{2j}}$, we have $\mathbf{g} \in \mathbf{S_2}\mathrm{wr}\mathbf{S_4}$ and so $\mathbf{g} \in \mathbf{S_2} \times \mathbf{S_4}$.
- These observations are the perturbation method!

Outline of the proof

Preliminary 1 (difference operator)

- Let ∂_j be **j**-th difference operator defined by

$$(\partial_j \theta)_i \;=\; \theta_{(i_1,\ldots,i_j,\ldots,i_m)} - \theta_{(i_1,\ldots,1,\ldots,i_m)}.$$

- For any $E \subset [m]$, define $\partial_E = \prod_{j \in E} \partial_j$.
- Fact: Let $\eta_F(i)$ depend only on i_F. Then $\partial_E \eta_F = 0$ if $E \not\subset F$.
- Fact: Let $\Delta' \subset \Delta$ be another simplicial complex. Then

$$r(\Delta') = r(\Delta) \cap \left(\bigcap_{E \in \Delta \setminus \Delta'} \ker(\partial_E) \right).$$

Outline of the proof

Preliminary 2 (perturbation method)

- Construct a generic table $\boldsymbol{\theta} = \{\theta_i\}$.

$$\theta_i = \sum_{\mathbf{D} \in \mathbf{red}\boldsymbol{\Delta}} \phi_{\mathbf{D}}(i), \quad \text{where } \phi_{\mathbf{D}} \text{ depends only on } i_{\mathbf{D}},$$

  in such a way that the following condition holds:

  Fix (sufficiently large) positive integer $\mathbf{b}$*.*
  If a quantity $\mathbf{z}$ *satisfies* $\mathbf{z} = \sum_{\mathbf{D}} \sum_i \mathbf{c}_{\mathbf{D},i} \phi_{\mathbf{D}}(i)$ *with* $\mathbf{c}_{\mathbf{D},i} \in \{-\mathbf{b}, \ldots, \mathbf{b}\}$*, then the coefficients* $\{\mathbf{c}_{\mathbf{D},i}\}$ *are uniquely determined from* $\mathbf{z}$*.*

- Perturbation lemma: There exists such a table $\boldsymbol{\theta}$. (See the next page.)

Outline of the proof

Lemma (Perturbation lemma)

1. *Let* $\mathbf{n}, \mathbf{b}$ *be positive integers. There exist* $\mathbf{n}$ *positive integers* $(\mathbf{Y_l})_{\mathbf{l=1}}^{\mathbf{n}}$ *such that a map*

$$\{-\mathbf{b}, \cdots, \mathbf{b}\}^{\mathbf{n}} \ni (\mathbf{c_l})_{\mathbf{l=1}}^{\mathbf{n}} \mapsto \sum_{\mathbf{l=1}}^{\mathbf{n}} \mathbf{c_l} \mathbf{Y_l} \in \mathbb{Z}$$

   is injective.

2. *Furthermore, we can choose* $\mathbf{n}$ *vectors* $\mathbf{Y^{(j)}} = (\mathbf{Y_l^{(j)}})$ *such that they span* $\mathbb{Q}^{\mathbf{n}}$ *and each of them satisfies the above condition.*

Proof.

Use van der Monde determinant. □

Outline of the proof

- Now we prove Theorem 1: $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})} = \bigcap_{\mathbf{D} \in \mathbf{red}\boldsymbol{\Delta}} (\mathbf{S}_{\mathcal{I}_{\mathbf{D}^{\mathbf{C}}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathbf{D}}})$.
- We employ induction on $\mathbf{K} = |\mathbf{red}\boldsymbol{\Delta}|$.
- For the case $\mathbf{K} = \mathbf{1}$, it is essentially the same as the row-factor-only model $\log \mathbf{p}_{\mathbf{ij}} = \alpha_{\mathbf{i}}$ for 2-way tables. And therefore we have $\mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})} = \mathbf{S}_{\mathcal{I}_{\mathbf{D}^{\mathbf{C}}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathbf{D}}}$.

Outline of the proof

- We next consider $\mathbf{K} \geq \mathbf{2}$ and choose $\mathbf{D} \in \mathbf{red}\boldsymbol{\Delta}$ such that $|\mathcal{I}_{\mathbf{D}}| = \min_{\mathbf{F} \in \mathbf{red}\boldsymbol{\Delta}} |\mathcal{I}_{\mathbf{F}}|$.
- Let $\boldsymbol{\theta}$ be a generic element in $\mathbf{r}(\boldsymbol{\Delta})$
- Fix $\mathbf{g} \in \mathbf{G}_{\mathbf{r}(\boldsymbol{\Delta})}$. Then $\mathbf{g}\boldsymbol{\theta} \in \mathbf{r}(\boldsymbol{\Delta})$.
- Define a simplicial complex $\boldsymbol{\Delta}'$ by $\mathbf{red}\boldsymbol{\Delta}' = (\mathbf{red}\boldsymbol{\Delta}) \setminus \mathbf{D}$.
- Let $\mathbf{E} \in \boldsymbol{\Delta} \setminus \boldsymbol{\Delta}'$. Then $(\partial_{\mathbf{E}}(\mathbf{g}\boldsymbol{\theta}))_{i}$ depends only on $i_{\mathbf{D}}$ because other terms vanish.

Outline of the proof

- However, the quantity $(\partial_{\mathsf{E}}(\mathbf{g}\theta))_i$ is a linear combination of $\{\phi_{\mathsf{F}}(j)\}$. Its coefficients are uniquely determined by $(\partial_{\mathsf{E}}(\mathbf{g}\theta))_i$ (perturbation lemma).
- The above two facts imply that $(\partial_{\mathsf{E}}(\mathbf{g}\theta))_i$ is a linear combination only of $\{\phi_{\mathsf{D}}(j)\}$.
- Now by the second part of the perturbation lemma, any table $\tilde{\theta}$ in $\mathbf{r}(\boldsymbol{\Delta}')$ is spanned by generic tables in $\mathbf{r}(\boldsymbol{\Delta})$.
- Therefore $\partial_{\mathsf{E}}(\mathbf{g}\tilde{\theta}) = \mathbf{0}$ and

$$\mathbf{g}\tilde{\theta} \in \mathbf{r}(\boldsymbol{\Delta}) \cap \left(\bigcap_{\mathsf{E} \in \boldsymbol{\Delta}'} \ker(\partial_{\mathsf{E}}) \right) = \mathbf{r}(\boldsymbol{\Delta}').$$

Outline of the proof

- Hence $\mathbf{g}$ maps $\mathbf{r}(\boldsymbol{\Delta}')$ into itself.
- From the assumption of induction, we have

$$\mathbf{g} \in \bigcap_{\mathsf{F} \in \mathbf{red}\boldsymbol{\Delta}'} (\mathbf{S}_{\mathcal{I}_{\mathsf{FC}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathsf{F}}}) = \bigcap_{\mathsf{F} \in \mathbf{red}\boldsymbol{\Delta}, \mathsf{F} \neq \mathsf{D}} (\mathbf{S}_{\mathcal{I}_{\mathsf{FC}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathsf{F}}}).$$

- It remains to prove $\mathbf{g} \in \mathbf{S}_{\mathcal{I}_{\mathsf{DC}}} \mathbf{wr} \mathbf{S}_{\mathcal{I}_{\mathsf{D}}}$, which is not too difficult.

Summary and future works

We have proved that

- The group of invariance is determined (Theorem 1).
- The group is rewritten by generalized wreath product (Theorem 2).
- It enables us to draw a random sample from the group of invariance.
- (A subset of) Sudoku invariance group is automatically obtained.

Our future works are

- Remove the assumption of Theorem 1.
- In particular, give a natural conjecture for this problem.

Summary and future works

Suggestions from sudoku example

- Sudoku suggests a new connection between Markov bases (Diaconis-Stumbles school) and experimental design (Pistone-Wynn school) of algebraic statistics
- In experimental design, Gröbner bases have been used to clarify confounding and identification problems given a particular (typically non-regular) design.
- A large part of literature on experimental design looks at finding a good design for a given set of constraints (number of runs, number of treatments etc.)
- It is very similar to finding a good "sudoku puzzle".

Summary and future works

- But we saw that a sudoku puzzle is an element of a particular fiber (with all marginals $= \mathbf{1}$) of a hierarchical model.
- If we know the MB for the fiber, we can construct a Markov chain over the set of sudoku solutions (a paper by Fontana and Rogantine in Pistone volume.)
- For $\mathbf{2} \times \mathbf{2} \times \mathbf{2} \times \mathbf{2} (\times \mathbf{4})$ sudoku, Hisayuki Hara just obtained the MB.
- We can now walk around $\mathbf{2} \times \mathbf{2} \times \mathbf{2} \times \mathbf{2}$ Sudoku solutions!
- Standard $\mathbf{3} \times \mathbf{3} \times \mathbf{3} \times \mathbf{3}$ sudoku is a big challenge.

Bibliography

- S. Aoki & A. Takemura (2003). Minimal basis for a connected Markov chain over $\mathbf{3} \times \mathbf{3} \times \mathbf{K}$ contingency tables with fixed two-dimensional marginals. *Aust. N. Z. J. Stardust.*, **45** (2), 229–249.
- S. Aoki & A. Takemura (2008). The largest group of invariance for Markov bases and toric ideals. *J. Symbolic Com-put.*, **43** (5), 342–358.
- R. A. Bailey, C. E. Praeger, C. A. Rowley & T. P. Speed (1983). Generalized wreath products of permutation groups. *Proc. London Math. Soc.*, **47** (3), 69–82.
- David A. Cox. (2007). Gröbner basis tutorial. Part II. A sampler from recent developments. (available from Cox's homepage).
- R. Fontana & M.-P. Rogantin (2008). Indicator function and sudoku designs, in press.
- S. L. Lauritzen (1996). *Graphical Models*, Oxford University Press, Oxford.
- H. Monod & R. A. Bailey (1992). Pseudofactors: normal use to improve design and facilitate analysis, *Appl. Statist.*, **41** (2), 317–336.
- E. Russel & F. Jarvis (2006), Mathematics of Sudoku II, Preprint.
- C. Wells (1976). Some applications of the wreath product construction, *Amer. Math. Monthly*, **83** (5), 317–338.

Thank you for your attention!