Hierarchical subspace models for contingency tables

Hisayuki Hara, Tomonari Sei and Akimichi Takemura

University of Tokyo

March 28, 2010

Special Session on Advances in Algebraic Statistics

AMS Sectional Meeting at University of Kentucky

Notations on contingency tables

- Notations for m-way contingency tables
  - $I := I_1 \times \cdots \times I_m$: number of cells of an m-way table
    - I_k : number of levels for k-th variable
  - $\mathcal{I} := [I_1] \times \cdots \times [I_m]$: set of cells
    - $[I_k] = \{1, \ldots, I_k\}$
  - $\boldsymbol{i} = (i_1 i_2 \cdots i_m) \in \mathcal{I}$: each cell
  - For $D \in [m]$,
    - $\boldsymbol{i}_D$: marginal cell
    - $\mathcal{I}_D$: the set of marginal cells for D
    - I_D : the number of marginal cells for D
  - $x(\cdot)$: frequencies
  - $p(\cdot)$: cell probabilities

Space of m-way tables

- $V = \mathbb{R}^I = \mathbb{R}^{I_1 \times \cdots \times I_m}$
  - the set of m-way tables with real entries
  - I-dimensional real vector space of functions $\psi : \mathcal{I} \mapsto \mathbb{R}$

- L_D for $D \subset [m]$:
  the set of functions depending only on D marginal cells

  $$L_D = \{\psi \in V \mid \psi(i_1, \ldots, i_m) = \psi(i'_1, \ldots, i'_m) \text{ if } i_v = i'_v, \forall v \in D\}$$

  - L_D is considered as $\mathbb{R}^{I_D}$, where $I_D = \prod_{v \in D} I_v$
  - If $D = [m]$, $L_D = L_{[m]} = V$

Log affine model

- L : a linear subspace of V s.t. $1 \in L$
- Log-affine model :

$$\log p(\cdot) := \{\log p(\boldsymbol{i}), \boldsymbol{i} \in \mathcal{I}\} \in L$$

- V and L_D specify saturated models
- If $L_D \subset L$, we say that D is saturated in L

Hierarchical model

- Δ : a simplicial complex
- $\mathrm{red}\Delta$: the set of maximal elements (facets) of Δ
- Hierarchical model L_Δ :

$$\log p(\cdot) \in L_\Delta := \sum_{D \in \mathrm{red}\Delta} L_D,$$

- If $\mathrm{red}\Delta = [m]$, L_Δ is m-way saturated model

Example 1. modeling for $I \times J$ tables

- Two-way saturated model : $\mathrm{red}\Delta := \{\{1,2\}\}$

$$V = L_{\{1,2\}} \;:\; \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_1 i_2}$$

- A log-affine model L :

$$L \;:\; \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma \phi_{i_1 i_2} \;\subset V$$

  - $\phi_{i_1 i_2}$: known functions

ex 1. uniform association model : $\phi_{i_1 i_2} = i_1 i_2$

ex 2. two-way change point model (Hirotsu(1997)) :

$$\phi_{i_1 i_2} = \begin{cases} 1, & \text{if } i_1 \le I_1' < I_1 \text{ and } i_2 \le I_2' < I_2, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Downarrow$$

- Modeling strategy for higher dimensional tables has not been fully discussed

Example 1. modeling for $I \times J$ tables

- Two-way saturated model : $\mathrm{red}\Delta := \{\{1,2\}\}$

$$V = L_{\{1,2\}} \ : \ \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_1 i_2}$$

- A log-affine model L :

$$L \ : \ \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma \phi_{i_1 i_2} \ \subset V$$

  - $\phi_{i_1 i_2}$: known functions

ex 1. uniform association model : $\phi_{i_1 i_2} = i_1 i_2$

ex 2. two-way change point model (Hirotsu(1997)) :

$$\phi_{i_1 i_2} = \begin{cases} 1, & \text{if } i_1 \leq I_1' < I_1 \text{ and } i_2 \leq I_2' < I_2, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Downarrow$$

- Modeling strategy for higher dimensional tables has not been fully discussed

Example 1. modeling for $I \times J$ tables

- Two-way saturated model : $\mathrm{red}\Delta := \{\{1,2\}\}$

$$V = L_{\{1,2\}} \ : \ \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_1 i_2}$$

- A log-affine model L :

$$L \ : \ \log p_{i_1 i_2} = \alpha_{i_1} + \beta_{i_2} + \gamma \phi_{i_1 i_2} \ \subset V$$

  - $\phi_{i_1 i_2}$: known functions

ex 1. uniform association model : $\phi_{i_1 i_2} = i_1 i_2$

ex 2. two-way change point model (Hirotsu(1997)) :

$$\phi_{i_1 i_2} = \begin{cases} 1, & \text{if } i_1 \leq I_1' < I_1 \text{ and } i_2 \leq I_2' < I_2, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Downarrow$$

- Modeling strategy for higher dimensional tables has not been fully discussed

Example 2. 3-way Split model

- Split model (Højsgaard (2003)):
  The conditional independence structures are different for specific values of the conditioning variables

  ex. $L = L^{i_2=1}_{\{1\}} + L^{i_2=1}_{\{3\}} + L^{i_2=2}_{\{1,3\}}$

  - $i_2 = 1$ slice : $L^{i_2=1} = L_{\{1\}} + L_{\{3\}}$
  - $i_2 = 2$ slice : $L^{i_2=2} = L_{\{1,3\}}$
  - $L \subset L_{\{1,2,3\}}$

$$\Downarrow$$

- Sophisticated modeling of interaction terms is required for the analysis of contingency tables

Example 2. 3-way Split model

- Split model (Højsgaard (2003)):
  The conditional independence structures are different
  for specific values of the conditioning variables

  ex. $L = L^{i_2=1}_{\{1\}} + L^{i_2=1}_{\{3\}} + L^{i_2=2}_{\{1,3\}}$

  - $i_2 = 1$ slice : $L^{i_2=1} = L_{\{1\}} + L_{\{3\}}$
  - $i_2 = 2$ slice : $L^{i_2=2} = L_{\{1,3\}}$
  - $L \subset L_{\{1,2,3\}}$

$$\Downarrow$$

- Sophisticated modeling of interaction terms is required for the analysis of contingency tables

Main purpose

- We propose "hierarchical subspace model (HSM)" as a generalization of the hierarchical model
- The notion of HSM gives a modeling strategy of multiway tables and unifies various models of interaction effects
- In this talk we discuss HSM from a viewpoint of localization of the computation of MLE and Markov bases
- We also illustrate practical advantage of our modeling strategy with some data sets

Hierarchical subspace model

Conformality of Log-affine model

Definition (conformality)

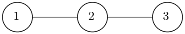
$W_1, \dots, W_K$: linear subspaces of V
$W := W_1 + \cdots + W_K$
L is *conformal* to $\{W_j\}_{j=1}^K$ if

$$L = L \cap W = (L \cap W_1) + \cdots + (L \cap W_K)$$

- $L \supset (L \cap W_1) + \cdots + (L \cap W_K)$ always holds
- The inclusion is strict in general

Ex. 3-way conditional independence model

- $W_1 := L_{\{1,2\}}$, $W_2 := L_{\{2,3\}}$
- $W = W_1 + W_2 = L_{\{1,2\}} + L_{\{2,3\}}$
- W : $\log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta_{i_1 i_2} + \delta'_{i_2 i_3}$

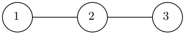


- L : $\log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta\phi_{i_1 i_2} + \delta'\psi_{i_2 i_3}$
  - $\phi_{i_1 i_2}$, $\psi_{i_2 i_3}$: known functions
  - δ, δ': free parameters
  - $L \cap W_1 = \{\alpha_{i_1} + \beta_{i_2} + \delta\phi_{i_1 i_2}\}$
  - $L \cap W_2 = \{\beta_{i_2} + \gamma_{i_3} + \delta'\psi_{i_2 i_3}\}$

$\rightarrow$ $\boxed{L = (L \cap W_1) + (L \cap W_2)}$ conformal!

Ex. 3-way conditional independence model

- $W_1 := L_{\{1,2\}}$, $W_2 := L_{\{2,3\}}$
- $W = W_1 + W_2 = L_{\{1,2\}} + L_{\{2,3\}}$
- $L: \ \log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta_{i_1 i_2} + \delta'_{i_2 i_3}$



- $L': \ \log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta(\phi_{i_1 i_2} + \psi_{i_2 i_3})$
  - $L' \cap W_1 = \{\alpha_{i_1} + \beta_{i_2}\}$
  - $L' \cap W_2 = \{\beta_{i_2} + \gamma_{i_3}\}$

$$\rightarrow \quad (L' \cap W_1) + (L' \cap W_2) = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3}$$

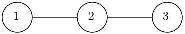
$$\rightarrow \quad \boxed{L' \supset (L' \cap W_1) + (L' \cap W_2)}$$

not conformal

  - Intuitively conformality represents decomposability of L

Ex. 3-way conditional independence model

- $W_1 := L_{\{1,2\}}$, $W_2 := L_{\{2,3\}}$
- $W = W_1 + W_2 = L_{\{1,2\}} + L_{\{2,3\}}$
- $L:\ \log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta_{i_1 i_2} + \delta'_{i_2 i_3}$



- $L':\ \log p_{i_1 i_2 i_3} = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta(\phi_{i_1 i_2} + \psi_{i_2 i_3})$
  - $L' \cap W_1 = \{\alpha_{i_1} + \beta_{i_2}\}$
  - $L' \cap W_2 = \{\beta_{i_2} + \gamma_{i_3}\}$

$$\rightarrow \quad (L' \cap W_1) + (L' \cap W_2) = \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3}$$

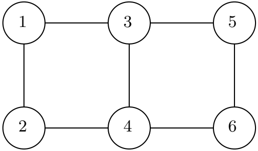
$$\rightarrow \quad \boxed{L' \supset (L' \cap W_1) + (L' \cap W_2)}$$ not conformal

  - Intuitively conformality represents decomposability of L

Terminologies on hypergraphs

- $\mathrm{red}\Delta$ is considered as a hypergraph
- A divider S of $\mathrm{red}\Delta$:
  $\exists u, v$ s.t. S is a minimal clique separator separating u and v
  - $\mathcal{S}$: the set of dividers of $\mathrm{red}\Delta$
- u and v are tightly connected:
  there is no divider separating u and v
- compact component C:
  any two vertices in C are tightly connected
  - $\mathcal{C}$: the set of maximal compact components of $\mathrm{red}\Delta$

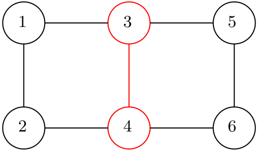
ex. A divider : $\{3, 4\}$,
Compact components : $\{1, 2, 3, 4\}$, $\{3, 4, 5, 6\}$



Terminologies on hypergraphs

- $\mathrm{red}\Delta$ is considered as a hypergraph
- A divider S of $\mathrm{red}\Delta$:
  $\exists u, v$ s.t. S is a minimal clique separator separating u and v
  - $\mathcal{S}$: the set of dividers of $\mathrm{red}\Delta$
- u and v are tightly connected:
  there is no divider separating u and v
- compact component C:
  any two vertices in C are tightly connected
  - $\mathcal{C}$: the set of maximal compact components of $\mathrm{red}\Delta$

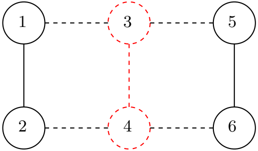
ex. A divider : $\{3, 4\}$,
Compact components : $\{1, 2, 3, 4\}$, $\{3, 4, 5, 6\}$



Terminologies on hypergraphs

- $\text{red}\Delta$ is considered as a hypergraph
- A divider S of $\text{red}\Delta$:
  $\exists u, v$ s.t. S is a minimal clique separator separating u and v
  - $\mathcal{S}$: the set of dividers of $\text{red}\Delta$
- u and v are tightly connected:
  there is no divider separating u and v
- compact component C:
  any two vertices in C are tightly connected
  - $\mathcal{C}$: the set of maximal compact components of $\text{red}\Delta$

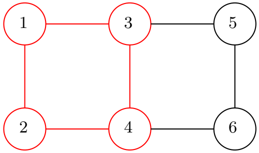
ex. A divider : $\{3, 4\}$,
Compact components : $\{1, 2, 3, 4\}$, $\{3, 4, 5, 6\}$



Terminologies on hypergraphs

- $\mathrm{red}\Delta$ is considered as a hypergraph
- A divider S of $\mathrm{red}\Delta$:
  $\exists u, v$ s.t. S is a minimal clique separator separating u and v
  - $\mathcal{S}$: the set of dividers of $\mathrm{red}\Delta$
- u and v are tightly connected:
  there is no divider separating u and v
- compact component C:
  any two vertices in C are tightly connected
  - $\mathcal{C}$: the set of maximal compact components of $\mathrm{red}\Delta$

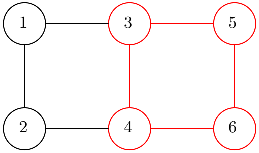
ex. A divider : $\{3, 4\}$,
Compact components : $\{1, 2, 3, 4\}$, $\{3, 4, 5, 6\}$



Terminologies on hypergraphs

- $\mathrm{red}\Delta$ is considered as a hypergraph
- A divider S of $\mathrm{red}\Delta$:
  $\exists u, v$ s.t. S is a minimal clique separator separating u and v
  - $\mathcal{S}$: the set of dividers of $\mathrm{red}\Delta$
- u and v are tightly connected:
  there is no divider separating u and v
- compact component C:
  any two vertices in C are tightly connected
  - $\mathcal{C}$: the set of maximal compact components of $\mathrm{red}\Delta$

ex. A divider : $\{3, 4\}$,
Compact components : $\{1, 2, 3, 4\}$, $\{3, 4, 5, 6\}$



Hierarchical subspace model (HSM)

Definition (hierarchical subspace model)

L is a hierarchical subspace model of a hierarchical model L_Δ if

1. $L_S \subset L$ for each $S \in \mathcal{S}$
2. L is conformal to $\{L_C, C \in \mathcal{C}\}$

- $L_S \subset L \Rightarrow \hat{p}(\boldsymbol{i}_S) = x(\boldsymbol{i}_S)/n$
  - $\hat{p}(\boldsymbol{i}_S)$: MLE for $L \cap L_S$, $x(\boldsymbol{i}_S)$: a marginal frequency for $\boldsymbol{i}_S$
- Conformality : the same decomposability as L_Δ
  - $L_\Delta = L_{\{1,2\}} + L_{\{2,3\}}$
  - L: $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta\phi_{ij} + \delta'\psi_{jk}$
    $\Rightarrow$ HSM of L_Δ
  - L': $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta(\phi_{ij} + \psi_{jk})$
    $\Rightarrow$ L' is not conformal to $\{L_{\{1,2\}}, L_{\{2,3\}}\}$
    $\Rightarrow$ L' is not an HSM of L_Δ
  - Every log-affine model is an HSM of the saturated model

Hierarchical subspace model (HSM)

Definition (hierarchical subspace model)

L is a hierarchical subspace model of a hierarchical model $L_\triangle$ if

1. $L_S \subset L$ for each $S \in \mathcal{S}$
2. L is conformal to $\{L_C, C \in \mathcal{C}\}$

- $L_S \subset L \Rightarrow \hat{p}(\boldsymbol{i}_S) = x(\boldsymbol{i}_S)/n$
  - $\hat{p}(\boldsymbol{i}_S)$: MLE for $L \cap L_S$, $\quad x(\boldsymbol{i}_S)$: a marginal frequency for $\boldsymbol{i}_S$
- Conformality : the same decomposability as $L_\triangle$
  - $L_\triangle = L_{\{1,2\}} + L_{\{2,3\}}$
  - L: $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta\phi_{ij} + \delta'\psi_{jk}$

    $\Rightarrow$ HSM of $L_\triangle$
  - L': $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta(\phi_{ij} + \psi_{jk})$

    $\Rightarrow$ L' is not conformal to $\{L_{\{1,2\}}, L_{\{2,3\}}\}$

    $\Rightarrow$ L' is not an HSM of $L_\triangle$
  - Every log-affine model is an HSM of the saturated model

Hierarchical subspace model (HSM)

Definition (hierarchical subspace model)

L is a hierarchical subspace model of a hierarchical model L_Δ if

1. $L_S \subset L$ for each $S \in \mathcal{S}$
2. L is conformal to $\{L_C, C \in \mathcal{C}\}$

- $L_S \subset L \Rightarrow \hat{p}(\boldsymbol{i}_S) = x(\boldsymbol{i}_S)/n$
  - $\hat{p}(\boldsymbol{i}_S)$: MLE for $L \cap L_S$, $x(\boldsymbol{i}_S)$: a marginal frequency for $\boldsymbol{i}_S$
- Conformality : the same decomposability as L_Δ
  - $L_\Delta = L_{\{1,2\}} + L_{\{2,3\}}$
  - L: $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta\phi_{ij} + \delta'\psi_{jk}$
    $\Rightarrow$ HSM of L_Δ
  - L': $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta(\phi_{ij} + \psi_{jk})$
    $\Rightarrow$ L' is not conformal to $\{L_{\{1,2\}}, L_{\{2,3\}}\}$
    $\Rightarrow$ L' is not an HSM of L_Δ
  - Every log-affine model is an HSM of the saturated model

MLE of HSM

- L : HSM of L_Δ
  - L has the same conditional independence structure as L_Δ
- $p(\boldsymbol{i}_C)$: cell prob for $L \cap L_C$

$$p(\boldsymbol{i}) = \frac{\prod_{C \in \mathcal{C}} p(\boldsymbol{i}_C)}{\prod_{S \in \mathcal{S}} p(\boldsymbol{i}_S)}$$

- $\hat{p}(\boldsymbol{i})$: MLE of $p(\boldsymbol{i})$
- $\hat{p}(\boldsymbol{i}_C)$: MLE for $L \cap L_C$

$$\hat{p}(\boldsymbol{i}) = \frac{\prod_{C \in \mathcal{C}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S \in \mathcal{S}} \hat{p}(\boldsymbol{i}_S)} = \frac{\prod_{C \in \mathcal{C}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S \in \mathcal{S}} x(\boldsymbol{i}_S)/n}$$

Decomposability of log-affine model

- So far we have discussed the definition of HSM
- For a given L_Δ, we can obtain an HSM with the same decomposability as L_Δ
- Next we discuss decomposability of a given log-affine model L
- Every log affine model has a hierarchical model for which L is an HSM
- From a viewpoint of localization of the inference, a natural question is to look for a small L_Δ for which L is an HSM
- We derive the smallest decomposable model for which L is an HSM

$\Downarrow$

Ambient decomposable model

Ambient decomposable model

Decomposition of log-affine model

- In the case of the hierarchical model
  interaction terms $\Leftrightarrow \Delta$
  decomposition of $L_\Delta \Leftrightarrow$ decomposition of $\mathrm{red}\Delta$ (hypergraph)
- L is not necessarily an HSM of a hierarchical model which has the same interaction as L
  ex) 3-way conditional independence model
  - $L_\Delta = L_{\{1,2\}} + L_{\{2,3\}}$
    $$\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta_{ij} + \delta'_{jk}$$
  - L: $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta\phi_{ij} + \delta'\psi_{jk}$
  - L': $\log p_{ijk} = \alpha_i + \beta_j + \gamma_k + \delta(\phi_{ij} + \psi_{jk})$
- Decomposition of L does not always correspond to simplicial complex induced by interaction terms
- We define the decomposition of L by using the conformality

Hypergraph induced by HSM

- $S \subset [m]$ is a partial edge separator (pes) of L if
  - $L_S \subset L$
  - For disjoint subsets $A_1 \cup A_2 \cup S = [m]$,
    L is conformal to $\{L_{A_1 \cup S}, L_{A_2 \cup S}\}$
- (A_1, A_2, S) is called a decomposition of L
- u and v are tightly connected in L if there is no pes of L s.t. $u \in A_1$ and $v \in A_2$
- Extended compact component (ECC)
  a set of vertices any two of which are tightly connected in L
- Hypergraph $\mathcal{H}$ induced by L
  the set of maximal ECCs

Ambient decomposable model

- $L_{\mathcal{H}}$: the hierarchical model induced by $\mathcal{H}$

Theorem

$L_{\mathcal{H}}$ is the ambient decomposable model of L

- MLE

$$\hat{p}(\boldsymbol{i}) = \frac{\prod_{C\in\mathcal{H}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S\in\mathcal{S}} \hat{p}(\boldsymbol{i}_S)} = \frac{\prod_{C\in\mathcal{H}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S\in\mathcal{S}} x(\boldsymbol{i}_S)/n}.$$

  - $\mathcal{S}$: the set of divider of $\mathcal{H}$
  - $\hat{p}(\boldsymbol{i}_C)$ depends only on the marginal table $x(\boldsymbol{i}_C)$

$\Downarrow$

The computation of the MLE is localized to ECCs

Ambient decomposable model

- $L_{\mathcal{H}}$: the hierarchical model induced by $\mathcal{H}$

Theorem

$L_{\mathcal{H}}$ is the ambient decomposable model of L

- MLE

$$\hat{p}(\boldsymbol{i}) = \frac{\prod_{C \in \mathcal{H}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S \in \mathcal{S}} \hat{p}(\boldsymbol{i}_S)} = \frac{\prod_{C \in \mathcal{H}} \hat{p}(\boldsymbol{i}_C)}{\prod_{S \in \mathcal{S}} x(\boldsymbol{i}_S)/n}.$$

  - $\mathcal{S}$: the set of divider of $\mathcal{H}$
  - $\hat{p}(\boldsymbol{i}_C)$ depends only on the marginal table $x(\boldsymbol{i}_C)$

$\Downarrow$

The computation of the MLE is localized to ECCs

Local computation of Markov bases

- Dobra and Sullivant (2004)
  A Markov basis of a hierarchical model L_Δ is computed recursively from Markov bases of marginal models L_C for all $C \in \mathcal{C}$
- A Markov basis of an HSM is also computed from Markov bases of $L \cap L_C$, $C \in \mathcal{H}$

Local computation of Markov bases

- (A_1, A_2, S) : decomposition of L
- $\boldsymbol{z}_1$: a move $L \cap L_{A_1 \cup S}$

$$\boldsymbol{z}_1 = [\{(\boldsymbol{i}_1, \boldsymbol{j}_1), \ldots, (\boldsymbol{i}_d, \boldsymbol{j}_d)\} || \{(\boldsymbol{i}'_1, \boldsymbol{j}_1), \ldots, (\boldsymbol{i}'_d, \boldsymbol{j}_d)\}],$$

$$\boldsymbol{i}_k, \boldsymbol{i}'_k \in \mathcal{I}_{A_1}, \quad \boldsymbol{j}_k \in \mathcal{I}_S$$

  - $(\boldsymbol{i}_1, \boldsymbol{j}_1), \ldots, (\boldsymbol{i}_d, \boldsymbol{j}_d)$:
    cells (with replication) of "+" elements of z_1
  - $(\boldsymbol{i}'_1, \boldsymbol{j}_1), \ldots, (\boldsymbol{i}'_d, \boldsymbol{j}_d)$:
    cells (with replication) of "$-$" elements of z_1
  - $L_S \subset L \Rightarrow z_1(\boldsymbol{i}_S) = 0$

Extension of moves

Definition $(\mathrm{Ext}(\mathcal{B}(A_1 \cup S) \to L))$

$V_1 := A_1 \cup S$
$\mathcal{B}(V_1)$: a Markov basis of $L \cap L_{V_1}$
$\boldsymbol{z}_1 \in \mathcal{B}(V_1)$
$\boldsymbol{k} := \{\boldsymbol{k}_1, \dots, \boldsymbol{k}_d\} \in \mathcal{I}_{A_2} \times \cdots \times \mathcal{I}_{A_2}$,
Define $\boldsymbol{z}_1^{\boldsymbol{k}}$ by

$$\boldsymbol{z}_1^{\boldsymbol{k}} := [\{(\boldsymbol{i}_1, \boldsymbol{j}_1, \boldsymbol{k}_1), \dots, (\boldsymbol{i}_d, \boldsymbol{j}_d, \boldsymbol{k}_d)\} || \{(\boldsymbol{i}'_1, \boldsymbol{j}_1, \boldsymbol{k}_1), \dots, (\boldsymbol{i}'_d, \boldsymbol{j}_d, \boldsymbol{k}_d)\}].$$

Then define $\mathrm{Ext}(\mathcal{B}(V_1) \to L)$ by

$$\mathrm{Ext}(\mathcal{B}(V_1) \to L) := \{\boldsymbol{z}_1^{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathcal{I}_{A_2} \times \cdots \times \mathcal{I}_{A_2},\ \boldsymbol{z}_1 \in \mathcal{B}(V_1)\}$$

$$\begin{array}{cccc} i_1j_1 & i'_1j_1 & i_2j_2 & i'_2j_2 \\ \hline 1 & -1 & 1 & -1 \end{array} \quad \Rightarrow \quad \begin{array}{c|cccc} & i_1j_1 & i'_1j_1 & i_2j_2 & i'_2j_2 \\ \hline k_1 & 1 & -1 & 0 & 0 \\ k_2 & 0 & 0 & 1 & -1 \\ \hline & 1 & -1 & 1 & -1 \end{array}$$

Local computation of Markov bases

Theorem

$V_1 := A_1 \cup S$, $V_2 := A_2 \cup S$
$\mathcal{B}(V_1)$, $\mathcal{B}(V_2)$: MB of $L \cup L_{V_1}$, $L \cup L_{V_2}$
$\mathcal{B}_{V_1,V_2}$: a MB of decomposable model with two cliques V_1, V_2

Then

$$\mathcal{B} := \mathrm{Ext}(\mathcal{B}(V_1) \to L) \cup \mathrm{Ext}(\mathcal{B}(V_2) \to L) \cup \mathcal{B}_{V_1,V_2}$$

is a Markov basis of L

Numerical example

Numerical example : WAM data

- Woman and Mathematics (WAM) data : 6-way contingency table
  a questionnaire from high school students in NJ
  (Fowlkes(1988),Højsgaard(2003))

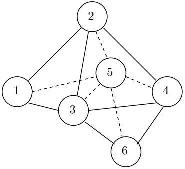
(1) Attendance in math lectures (attended=1, not=2)
(2) Sex (female=1, male=2)
(3) School type (suburban=1, urban=2)
(4) Agree in statement "I'll need mathematics in my future work" (agree=1, disagree=2)
(5) Subject preference (math-science=1, liberal arts=2)
(6) Future plans (college=1, job=2)

1. H_1 : decomposable model

$$L_1 \ = \ L_{\{1,2,3,5\}} + L_{\{2,3,4,5\}} + L_{\{3,4,5,6\}}$$

2. H_0 : split model

$$L_0 \ = \ L_{\{1,2,3,5\}} + L^{j_3=1}_{\{2,5\}} + L^{j_3=1}_{\{4,5\}} + L^{j_3=2}_{\{2,4,5\}} + L_{\{3,4,5,6\}}.$$



1.	attendance in math	4.	necessity of math in future
2.	sex	5.	subject preference
3.	school type	6.	future plan

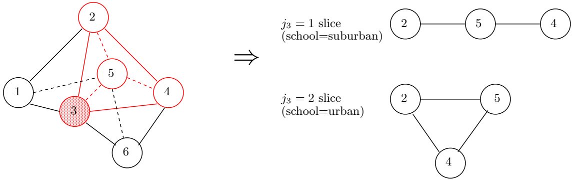
Numerical example : model

1. H_1 : decomposable model

$$L_1 \;=\; L_{\{1,2,3,5\}} + L_{\{2,3,4,5\}} + L_{\{3,4,5,6\}}$$

2. H_0 : split model

$$L_0 \;=\; L_{\{1,2,3,5\}} + L^{j_3=1}_{\{2,5\}} + L^{j_3=1}_{\{4,5\}} + L^{j_3=2}_{\{2,4,5\}} + L_{\{3,4,5,6\}}.$$



1.	attendance in math	4.	necessity of math in future
2.	sex	5.	subject preference
3.	school type	6.	future plan

Numerical example : Markov basis

- $\mathcal{B}_1$: Markov basis of L_1

$$\mathcal{B}_1 \;=\; \mathcal{B}_{\{1,2,3,5\},\{2,3,4,5,6\}} \cup \mathcal{B}_{\{1,2,3,4,5\},\{3,4,5,6\}}$$

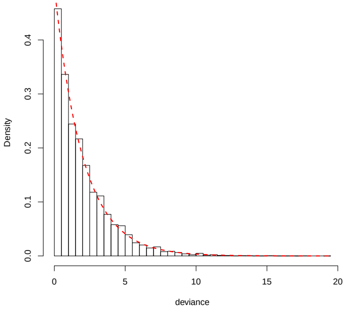
- $\mathcal{B}_0$: Markov basis of L_0

$$\mathcal{B}_0 \;=\; \mathcal{B}^{i_3=1}_{\{1,2,5\},\{4,5,6\}} \cup \mathcal{B}_1$$

  - $\mathcal{B}_{C_1,C_2}$: a Markov basis of a decomposable model with two cliques C_1 and C_2
  - $\mathcal{B}^{i_\delta}_{\mathcal{C}}$: $\mathcal{B}_{\mathcal{C}}$ on i_δ-slice

Numerical example : results

- We used LR statistic as a test statistic



	p-value	
Deviance	asymptotic χ^2_2	MCMC
1.851	0.396	0.399$\pm$0.012

- split model is accepted

- histgram : MCMC
- dotted line : asymptotic χ^2_2

Summary and future work

Summary

- We proposed a hierarchical subspace model by defining the notion of conformality of linear subspaces to a given hierarchical model
- The notion of an HSM gives a modeling strategy of multiway tables and unifies various models of interaction effects
- We illustrated practical advantage of our modeling strategy with some data sets

Future work

- Is it possible to treat nonlinear models such as the RC association model in the framework of HSM ?

References

Hara, H., Sei, T. and Takemura, A.(2009).
Hierarchical subspace models for contingency tables
arxiv 0909.4821, submitted