# Approximation algorithm for generating $\mathrm{B}^m \times J$ contingency tables

01605000 University of Tokyo * Tomomi MATSUI
01605630 Tokai University Yasuko MATSUI
Science University of Tokyo Yoko ONO

## 1 Introduction

We propose a new Markov chain for sampling $\mathrm{B}^m \times J = \mathrm{B} \times \cdots \times \mathrm{B} \times J$ contingency tables where $\mathrm{B} = \{1,2\}$ and $J = \{1,2,\ldots,n\}$. This Markov chain is an extention of the Markov chain which is proposed by Dyer and Greenhill [3] for two rowed contingency tables. To show that our Markov chain is rapidly mixing, we use a path coupling method, which is proposed by Bubley and Dyer [1].

## 2 Contingency Tables

We denote the set of integers (non-negative integers, positive integers) by $\mathrm{Z}$ $(\mathrm{Z}_+, \mathrm{Z}_{++})$ respectively and consider a set of contingency tables indexed by $\mathrm{B}^m \times J$ where $\mathrm{B} = \{1,2\}$ and $J = \{1,2,\ldots,n\}$. Any index in $J$ is called a *column index.* For any vector $\boldsymbol{x} \in \mathrm{R}^{\mathrm{B}^m \times J}$, both $x(\boldsymbol{i};j)$ and $x(i_1,i_2,\ldots,i_m;j)$ denote the elements of $\boldsymbol{x}$ indexed by $\boldsymbol{i} = (i_1,i_2,\ldots,i_m) \in \mathrm{B}^m$ and $j \in J$. For any column index $j \in J$, $\boldsymbol{x}(j) \in \mathrm{R}^{\mathrm{B}^m}$ denotes the subvector of $\boldsymbol{x} \in \mathrm{R}^{\mathrm{B}^m \times J}$ consists of elements defined by indices in $\mathrm{B}^m \times \{j\}$. Given a vector of indices $\boldsymbol{i} \in \mathrm{B}^m$, $\boldsymbol{i}_{\bar{l}}$ denotes the vector $(i_1,\ldots,i_{l-1},i_{l+1},\ldots,i_m) \in \mathrm{B}^{m-1}$ and we also denote the vector $\boldsymbol{i}$ by $(\boldsymbol{i}_{\bar{l}}, i_l)$ by changing the order of elements. For any vector $\boldsymbol{x} \in \mathrm{R}^{\mathrm{B}^m \times J}$ and $l \in \{1,2,\ldots,m\}$, $x(\boldsymbol{i}_{\bar{l}}, i_l; j)$ denotes the element $x(\boldsymbol{i};j)$ by changing the order of indices.

Let $(\boldsymbol{r}^1, \boldsymbol{r}^2, \ldots, \boldsymbol{r}^m; \boldsymbol{c})$ be a sequence of non-negative integer vectors where $\boldsymbol{r}^l \in \mathrm{Z}_+^{\mathrm{B}^{m-1} \times J}$ for each $l \in \{1,2,\ldots,m\}$ and $\boldsymbol{c} \in \mathrm{Z}_+^{\mathrm{B}^m}$. The element of $\boldsymbol{r}^l$ indexed by $(\boldsymbol{i};j) \in \mathrm{B}^{m-1} \times J$ is denoted by $r^l(\boldsymbol{i};j)$. The set of contingency tables corresponding to $(\boldsymbol{r}^1, \boldsymbol{r}^2, \ldots, \boldsymbol{r}^m; \boldsymbol{c})$ is defined by

$$\mathcal{T} \stackrel{\text{def.}}{=} \left\{ \boldsymbol{x} \in \mathrm{Z}_+^{\mathrm{B}^m \times J} \;\middle|\; \begin{array}{ll} x(\boldsymbol{i}_{\bar{l}}, 1; j) + x(\boldsymbol{i}_{\bar{l}}, 2; j) = r^l(\boldsymbol{i}_{\bar{l}}; j) & (\forall l \in \{1,2,\ldots,m\},\ \forall \boldsymbol{i}_{\bar{l}} \in \mathrm{B}^{m-1},\ \forall j \in J), \\ \sum_{j \in J} x(\boldsymbol{i};j) = c(\boldsymbol{i}) & (\forall \boldsymbol{i} \in \mathrm{B}^m) \end{array} \right\}.$$

Each element in $\mathcal{T}$ is called a *table* for simplicity. In the following, $\sum_{\boldsymbol{i} \in \mathrm{B}^m} c(\boldsymbol{i})$ is denoted by $N$. Clearly, for any table $\boldsymbol{x} \in \mathcal{T}$, the sum total of elements of $\boldsymbol{x}$ is equivalent to $N$.

## 3 Markov Chain

Here, we propose a new Markov chain whose mixing time is bouded by a polynomial in $n$ and $\ln N$. We define the parity function $\mathrm{p} : \mathrm{Z} \to \{1,-1\}$ by

$$\mathrm{p}(x) = \begin{cases} 1 & (x \text{ is an even integer }), \\ -1 & (x \text{ is an odd integer }). \end{cases}$$

For any index $\boldsymbol{i} \in \mathrm{B}^m$, we denote $\mathrm{p}(i_1 + i_2 + \cdots + i_m)$ by $\mathrm{p}(\boldsymbol{i})$. The vector $\Delta \in \{1,-1\}^{\mathrm{B}^m}$ is defined by $\Delta(\boldsymbol{i}) = \mathrm{p}(\boldsymbol{i})$ for each vector of indices $\boldsymbol{i} \in \mathrm{B}^m$. Given a pair of distinct column indices $(j', j'')$, we define the vector $\Delta[j',j''] \in \mathrm{Z}^{\mathrm{B}^m \times J}$ by

$$\Delta[j',j''](j) \stackrel{\text{def.}}{=} \begin{cases} \boldsymbol{0} & (j \in J \setminus \{j',j''\}), \\ \Delta & (j = j'), \\ -\Delta & (j = j''). \end{cases}$$



© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．

For any table $x \in \mathcal{T}$ and any pair of distinct column indices $\{j', j''\}$, we define the following set of vectors;

$$\mathcal{N}(x;\{j',j''\}) \stackrel{\text{def.}}{=} \left\{ y \in \mathrm{Z}_+^{B^m \times \{j',j''\}} \;\middle|\; \begin{array}{l} x(i_{\bar{l}},1;j) + x(i_{\bar{l}},2;j) = y(i_{\bar{l}},1;j) + y(i_{\bar{l}},2;j) \\ \quad (\forall l \in \{1,2,\ldots,m\},\ \forall i_{\bar{l}} \in B^{m-1},\ \forall j \in \{j',j''\}), \\ x(i;j') + x(i;j'') = y(i,j') + y(i;j'') \quad (\forall i \in B^m) \end{array} \right\}$$
$$= \left\{ y \in \mathrm{Z}_+^{B^m \times \{j',j''\}} \;\middle|\; \exists \theta \in \mathrm{Z},\ (y(j'),y(j'')) = (x(j'),x(j'')) + \theta(\Delta,-\Delta) \geq \mathbf{0} \right\}.$$

By using the above set $\mathcal{N}(x;\{j',j''\})$, we propose our new Markov chain $\mathcal{M}^1$ with state space $\mathcal{T}$. For any table $x \in \mathcal{T}$ and any pair of distinct column indices $\{j', j''\}$, we define the following set of tables;

$$\mathrm{N}^1(x;\{j',j''\}) \stackrel{\text{def.}}{=} \{x' \in \mathcal{T} \mid x'(j) = x(j)\ (\forall j \in J \setminus \{j',j''\}),\ (x'(j'),x'(j'')) \in \mathcal{N}(x;\{j',j''\})\}$$
$$= \{x' \in \mathcal{T} \mid \exists \theta \in \mathrm{Z},\ x' = x + \theta\,\Delta[j',j''] \geq \mathbf{0}\}.$$

Let $\mathcal{M}^1$ denote the Markov chain with the state space $\mathcal{T}$ with the following transition procedure. If $X_t$ is the state of the chain $\mathcal{M}^1$ at time $t$ and the element of $X_t$ indexed by $(i;j)$ is denoted by $X_t(i;j)$. Then the state $X_{t+1}$ at time $t+1$ is determined as follows. First, choose a pair of distinct column indices $\{j',j''\}$ randomly. Next, choose a table $X_{t+1}$ from $\mathrm{N}^1(X_t;\{j',j''\})$ at random.

# 4 Mixing Time of New Markov Chain

The *mixing time* $\tau^1(\varepsilon)$ of $\mathcal{M}^1$ is defined by

$$\tau^1(\varepsilon) \stackrel{\text{def.}}{=} \max_{x \in \mathcal{T}} \min\{t \mid \forall t' \geq t,\ \forall \mathcal{T}' \subseteq \mathcal{T}, -\varepsilon \leq \pi(\mathcal{T}') - \Pr[X_0 = x \text{ and } X_{t'} \in \mathcal{T}'] \leq \varepsilon\},$$

where $\pi : \mathcal{T} \to [0,1]$ is a unique stationary distribution of $\mathcal{M}^1$. To prove that our Markov chain is rapidly mixing, we use the path coupling method. We define a special Markov process with respect to $\mathcal{M}^1$ called coupling. A *coupling* of $\mathcal{M}^1$ is a Markov chain $(X_t, Y_t)$ on $\mathcal{T} \times \mathcal{T}$ satisfying that each of $(X_t), (Y_t)$, considered marginally, is a faithful copy of the original Markov chain $\mathcal{M}^1$. More precisely, we require that

$$\Pr(X_{t+1} = x' | (X_t, Y_t) = (x, y)) = P_{\mathcal{M}^1}(x, x'),$$
$$\Pr(Y_{t+1} = y' | (X_t, Y_t) = (x, y)) = P_{\mathcal{M}^1}(y, y'),$$

for all $x, y, x', y' \in \mathcal{T}$ where $P_{\mathcal{M}^1}(x,x')$ and $P_{\mathcal{M}^1}(y,y')$ denote the transition probability from $x$ to $x'$ and from $y$ to $y'$ of the original Markov chain $\mathcal{M}^1$, respectively. The detail of our coupling is omitted. By using our coupling, it is known that we can analyse the mixing time of $\mathcal{M}^1$ (see [1]). Then the mixing time of our Markov chain $\mathcal{M}^1$ is as follows:

**Theorem 1** *The Markov chain $\mathcal{M}^1$ is rapidly mixing with mixing time $\tau^1(\varepsilon)$ satisfying* $\tau^1(\varepsilon) \leq n(n-1)\ln(\lceil N/2^m \rceil \varepsilon^{-1})/2$. □

# References


[1] R. BUBLEY AND M. DYER, Path coupling: A technique for proving rapid mixing in Markov chains, *38th Annual Symposium on Foundations of Computer Science*, IEEE, San Alimitos, 1997, pp. 223–231.

[2] P. DIACONIS AND L. SALOFF-COSTE, Random walk on contingency tables with fixed row and column sums, Technical Report, Department of Mathematics, Harvard University, 1995.

[3] M. DYER AND C. GREENHILL, Polynomial-time counting and sampling of two-rowed contingency tables, *Theoretical Computer Sciences*, 246(2000), pp. 265–278.




© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.