Journal of Engineering and Applied Sciences 15 (7): 1609-1617, 2020
ISSN: 1816-949X

Toric Ideals for $(25n^3-66n^2+41n)\times3\times n$ Contingency Tables

Hussein S. Mohammed Hussein and Abdulrahman H. Majeed
Department of Mathematics, College of Science, University of Baghdad,
Baghdad, Iran

Abstract: In this study, we find Markov basis and toric ideals for $(25n^3-66n^2+41n)\times3\times n$ contingency tables with fixed two dimensional marginals.

Key words: Computational algebraic statistics, sufficient statistics, linear transformation, connected graph, bipartite graph, dimensional

INTRODUCTION

Let I be a finite set and $|I| = n$, a cell is the element of I and it denoted by $i\in I$. $i = i_1,..., i_m$, i is often multi-index. A non-negative integer $x_i\in N = \{1, 2,...\}$ denoted a frequency of the cell i. A contingency table is a set of frequencies and stated as $x = \{x_i\}_{i\in I}$ with an suitable arrangement of the cell, considered a contingency table $x = \{x_i\}_{i\in I}\in N$ as a n-dimansional column vector of non-negative integers. The contingency table can be treated as a function from I to N defined as $i\rightarrow x_i$. Denoted $\mathbb{Z}$ be the set of integer numbers, also denoted the $a_j\in\mathbb{Z}^n$, $j = 1, \ldots, v$ as fixed vectors consisting of integers. A v-dimantional column vector $t = (t_1, \ldots, t_v)'\varepsilon Z^v$ as $t_j = a'_j x$, $j = 1, ..., v$. Here, denotes a transpose of the matrix or vector. Also, define v×p matrix A with its j-row being a'_j given by:

$$A = \begin{bmatrix} a'_1 \\ M \\ a'_v \end{bmatrix}$$

And if $t = Ax$ is a v-dimensional column vector, we define the set, $T = \{\mathbf{t} : \mathbf{t} = A\mathbf{x}, \mathbf{x}\in\mathbb{N}^n\} = A\mathbb{N}^n \subset \mathbb{Z}^v$ where denoted $\mathbb{N}$ is a set of natural numbers. The set of x's for t, $A^{-1}[t] = \{x\in\mathbb{N}^n: Ax = t\}$ (t-fibers) is treat for result similar tests. A set of t-fibers deigns a taking apart of $\mathbb{N}^n$. An important noting is that t-fiber depend on given out of its kernel ker (A). In fact, defined $x_1\sim x_2 \leftrightarrow x_1\sim x_2 \in$ ker (A). With oneself kernel for different A's, the set of t-fibersare the same (Aoki and Takemura, 2003a, b).

Diaconis and Sturmfels publication in 1998 found a new path in the rapid-advancing field of computational algebraic statistics (Diaconis *et al*., 1998; Russell, 2001). In 2000, M. Dyer and C. Greenhill, found a polynomial-time compute and sampling of contingency tables (Diaconis and Sturmfels, 1998). Dobra 2003 showed that the only moves have to be inclusive in a Markov basis that connects all contingency tables with fixed marginals (Dobra, 2003). Dobra and Sullivant (2002), described an algorithm for generate a Markov basis of multi-way tables that links all tables with fixed marginal totals (Aoki and Takemura, 2003) proved that there exist a uniqe minimal basis for 3×3×K contingency tables consisting of four types of indispensable moves (Sullivant, 2005) and in the same year S. Aoki and A. Takemura presented a list of moves of 3×4×K and 4×4×K contingency tables with fixed dimensional marginals Aoki and Takemura (2003a, b), also Takemura and Aoki (2003) given some description of a minimal Markov basis for connected Markov chain and given a sufficient and necessary condition for uniqueness of a minimal Markov basis (Takemura and Aoki, 2004). Takemura and Aoki (2005) studied the Markov basis for sampling from discret sample space which is equipped with som convent metric and they started from two state in the sample space.

In this study, we give a new algorithm to find the Markov basis and toric ideals for $(25n^3-69n^2+44n)\times3\times n$ contingency tables with it have a fixed dimensional marginals.

Some basic concepts: In this section, we review some basic definitions and notations of contingency tables, moves, Markov basis and toric ideals that, we need in our work.

Definition 1: Let $A: Z^n \rightarrow Z^d$ be a linear transformation, the Toric ideal I_A is the ideal $<P^u-P^v: u, v\in\mathbb{N}^n, A(u) = A(v)> \subseteq K[P_1, ..., P_2]$ where $P^u = P_1^{u1}P_2^{u2} \ldots P_p^{up}$ (Sullivant *et al*., 2012)).

Definition 2: A move is a n-dimensional vector of integer numbers $z = \{z_i\}_{i\varepsilon I}\,\varepsilon\mathbb{Z}^n$ and if it is in the kernel of A (i.e., Az = 0) (Aoki and Takemura, 2003a, b).

Corresponding Author: Hussein S. Mohammed hussein, Department of Mathematics, College of Science, University of Baghdad, Baghdad, Iran

Definition 3: Let $B \subset M_A$ be the set of moves and let $x_1, x_2 \in A^{-1}$ [t]. Say that x_2 accessible from x_1 by B if there exists a moves $z_1, ..., z_s \in B$ and $\varepsilon_s \in \{1, -1\}$, s = 1, ..., S such that $x_2 = x_1 + \sum_{s=1}^{S} \varepsilon_s z_s$ (Takemura and Aoki, 2004).

$$x_1 + \sum_{s=1}^{S} \varepsilon_s z_s \in A^{-1}[\mathbf{t}] \quad \text{for } 1 \le s \le S$$

Now, we give some concepts about graph theory that we use later.

Definition 4: Let G = (V, E) be A graph then it consists of two sets:

- V(G), the vertices setof a graph G, predominantly denoted by V which is a nonempty set of elements and
- E(G), the edges set of a graph G, predominantly denoted by E such that every edge e in the edges set is assigned of vertices [u, v] (Clark and Holton, 1995)

Definition 5: A connected graph G = (V, E) is a graph where every pair of distinct vertices $u, v \in V$ (G) the graph G has a u, v-path. Otherwise, we say the graph is disconnected (Agnarsson and Greenlaw, 2007).

Definition 6: A graph G is bipartite graph if there are X, $Y \subseteq V$ (G) meeting the following conditions: (Agnarsson and Greenlaw, 2007):

- V (G) = XUY
- $X \cap Y = \varphi$
- G [X] and G [Y]

Are both null graphs where G [X] and G [Y] are sub graph of the graph G induced by the set of vertices X, $Y \subseteq V$ (G), respectively.

Theorem 1: For a graph G is bipartite if and only if every cycle in the graph has an even length (Agnarsson and Greenlaw, 2007).

Definition 7: Let $A: \mathbb{Z}^n \to \mathbb{Z}^v$ a linear transformation $\mathbf{t} \in \mathbb{Z}^v$ and A^{-1} [t] be the set of t-fibers and let $B \subset \ker_z$ (A), then we define A^{-1} $[t]_B$ be the graph with vertex set A^{-1} [t] and u – –v an edge if and only if $u\text{-}v \in \pm B$ (Sullivant *et al*., 2012).

Definition 8: Let A^{-1} [t] = $\{x \in \mathbb{N}^n : Ax = t\}$. A set of a finite moves B is called a Markov basis if A^{-1} [t] constitutes one B equivalence class for all t (Takemura and Aoki, 2004).

Definition 9: If $B \subseteq \ker_z$ (A) is a nonempty set such that A^{-1} $[t]_B$ is a connected for all t, then B is Markov basis for A (Sullivant *et al*., 2012).

Definition 10: Let $B \subset M_A$ be the set of moves and let $x_1, x_2 \in A^{-1}$ [t]. Say that x_1 accessible from x_2 by B if there exists sequence of a moves $z_1, ..., z_2 \in B$ and $\varepsilon_m \in \{-1, 1\}$, m = 1, ..., D such that (Takemura and Aoki, 2004).

$$\mathbf{x}_1 = \mathbf{x}_2 + \sum_{m=1}^{D} \varepsilon_m z_m$$

$$\mathbf{x}_1 + \sum_{m=1}^{h} \varepsilon_m z_m \in A^{-1}[\mathbf{t}] \quad \text{for } 1 \le h \le D$$

Theorem 2: A collection of binomials $\{p^{z+}\text{-}p^{z-}: z \in B\} \subset I_A$ is generating set of toric ideal I_A if and only if $\pm B$ is a Markov basis for A (Diaconis and Sturmfels, 1998).

Genomics and phylogenetics (Aoki and Takemura, 2008): Deoxyribonucleic Acidulous (DNA) molecules are noticeiota encoding the heritable blueprint used in the further and working of all familiar living and many viruses. Forever with proteins and RNA, DNA is one of the three essential macromolecules that are fundamental for all known life forms. Genetic indication is cryptographic as a sequence of nucleotides (cytosine, guanine, thymine and adenine) using the letters C, G, T, A and most DNA molecules are double-cutspiral, consisting of two polymers of nucleotides, molecules with fundamental made of alternate sugars and phosphate groups with the nucleobases (C, G, T, A) linked to the sugars. DNA is suited for biological input storage (Fig. 1).

Main results: Let n be a natural number and let $xj \in A^{-1}[t]$, j = 1, ..., k be the representative elements of the set of 3×n contingency tables and B = $\{z_1, z_2, ..., z_k\}$ such that each z_m, m = 1, 2, ..., k is a matrix of dimension 3×n either has two non-zero columns and the other columns are zero denoted by $2z_m$ or it has treenon-zero columns and the other columns are zero denoted by $3z_m$, like:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 2 & 0 \\ 0 & 0 & 0 \\ 2 & -2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix}$$

Also, we write the elements of B as one dimensional column vectoras follows:

$z_m = (z_1, ..., z_{3n})'$ m = 1, ..., k and z_s = 0, 1-1, 2 or -2, s = 1, 2, ..., 3n such that. If s = 1, 2, ..., n then:

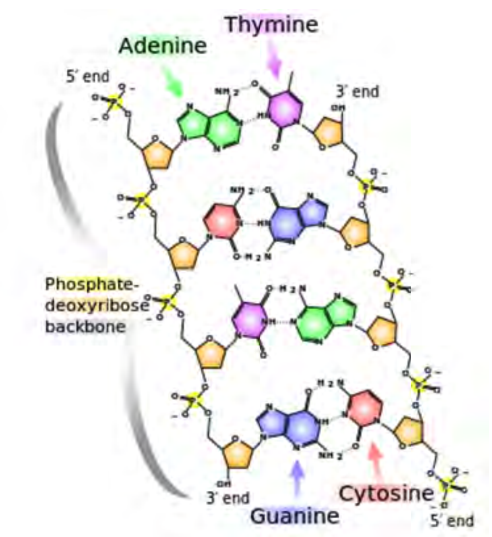

Fig. 1: Four DNA bases shown with complementary bases paired

$$z_s = \begin{cases} 1 & \text{if } z_{s+n} + z_{s+2n} = -1 \text{ and } \sum_{i=1, i \neq s}^{n} z_i = -1 \\ 2 & \text{if } z_{s+n} + z_{s+2n} = -2 \text{ and } \sum_{i=1, i \neq s}^{n} z_i = -2 \\ 0 & \text{if } z_{s+n} + z_{s+2n} = 0 \text{ and } \sum_{i=1, i \neq s}^{n} z_i = 0 \\ -1 & \text{if } z_{s+n} + z_{s+2n} = 1 \text{ and } \sum_{i=1, i \neq s}^{n} z_i = 1 \\ -2 & \text{if } z_{s+n} + z_{s+2n} = 2 \text{ and } \sum_{i=1, i \neq s}^{n} z_i = 2 \end{cases} \quad (1)$$

If s = n+1, n+2, ..., 2n then:

$$z_s = \begin{cases} 1 & \text{if } z_{s-n} + z_{s+n} = -1 \text{ and } \sum_{i=n+1, i \neq s}^{2n} z_i = -1 \\ 2 & \text{if } z_{s-n} + z_{s+n} = -2 \text{ and } \sum_{i=n+1, i \neq s}^{2n} z_i = -2 \\ 0 & \text{if } z_{s-n} + z_{s+n} = 0 \text{ and } \sum_{i=n+1, i \neq s}^{2n} z_i = 0 \\ -1 & \text{if } z_{s-n} + z_{s+n} = 1 \text{ and } \sum_{i=n+1, i \neq s}^{2n} z_i = 1 \\ -2 & \text{if } z_{s-n} + z_{s+n} = 2 \text{ and } \sum_{i=n+1, i \neq s}^{2n} z_i = 2 \end{cases} \quad (2)$$

If s = 2n+1, 2n+2, ..., 3n, then:

$$z_s = \begin{cases} 1 & \text{if } z_{s-n} + z_{s-2n} = -1 \text{ and } \sum_{i=2n+1, i \neq s}^{3n} z_i = -1 \\ 2 & \text{if } z_{s-n} + z_{s-2n} = -2 \text{ and } \sum_{i=2n+1, i \neq s}^{3n} z_i = 2 \\ 0 & \text{if } z_{s-n} + z_{s-2n} = 0 \text{ and } \sum_{i=2n+1, i \neq s}^{3n} z_i = 0 \\ -1 & \text{if } z_{s-n} + z_{s-2n} = 1 \text{ and } \sum_{i=2n+1, i \neq s}^{3n} z_i = 1 \\ -2 & \text{if } z_{s-n} + z_{s-2n} = 2 \text{ and } \sum_{i=2n+1, i \neq s}^{3n} z_i = 2 \end{cases} \quad (3)$$

Theorem 3: The number of elements in B equal to $25n^3-66n^2+41n$.

Proof: Since, there are three rows and n columns in $2z_m$, such that it has two columns (1, -1, 0)', (-1, 1, 0)' or (2, -2, 0)', (-2, 2, 0)' or (1, 0, -1)', (-1, 0, 1)' or (2, 0, -2)', (-2, 0, 2)' or (0, 1, -1)', (0, -1, 1)' or (0, 1, -1)', (0, -2, 2)' or (2, -1, -1)', (-2, 1, 1)' or (-1, 2, -1)', (1, -2, 1)' or (-1, -1, 2)', (1, 1, -2)') and the other columns are zero, then the number of elements $2z_m$ in B is $9 \times (n)!/(n-2)! = 9n(n-1) = 9n^2 - 9n$ but $3z_m$ has three columns (2, -2, 0)', (-1, 1, 0)', (-1 ,1, 0)' or (-2, 2, 0)', (1, -1, 0)', (1, -1, 0)' or (2, 0, -2)', (-1, 0, 1)', (-1, 0, 1)' or (-2, 0, 2)', (1, 0, -1)', (1, 0, -1)' or (0, 2, -2)', (0, -1, 1)', (0, -1, 1)' or (0, -2, 2)', (0, 1, -1)', (0, 1, -1)' or (2, -2, 0)', (0, 2, -2)', (2, 0, -2)' or (-2, 2, 0)', (0, - 2, 2)', (-2, 0, 2)' or (2, -1, -1)' and (-1, 2, -1)', (-1, -1, 2)' or (-1, 1, 0)', (-1, 0, 1)' or (-2, 2, 0)', (0, -1, 1)' or (-2, 0, 2)', (0, 1, -1)' or (-2, 1, 1)' and (1, -2, 1)', (1, 1, -2)' or (1, -1, 0)', (1, 0, -1)' or (2, 0, - 2)', (0, -1, 1)' or (2, 0, -2)', (0, -1, 1)' or (-1, 2, -1)' and (0, -1, 1)', (1, -1, 0)' or (2, -2, 0)', (-1, 0, 1)' or (0, -2, 2)', (-1, 0, 1)' or (1, -2, 1)' and (0, 1, -1)', (-1, 1, 0)' or (-2, 2, 0)', (1, 0, -1)' or (0, 2, -2)', (1, 0, -1)' or (-1, -1, 2)' and (0, 1, -1)', (1, 0, -1)' or (2, 0, -2)', (-1, 1, 0)' or (0, 2, -2)', (1, -1, 0)' or (1, 1, -2)' and (0, 1, -1)', (1, 0, -1)' or (0, -2, 2)', (-1, 1, 0)' or (-2, 0, 2)', (1, -1, 0)' and the other columns are zero, then the number of elements $3z_m$ in B is:

$$6 \times \frac{(n)!}{2(n-3)!} + 22 \times \frac{(n)!}{(n-3)!} = 3n(n-1)(n-2)$$
$$+22n(n-1)(n-2) = 25n^3 - 75n^2 + 50n$$

are an element in B for all m = 1, 2, …, k, since, each element in B is either $2z_m$ or $3z_m$ then the numbers of elements in B is:

$$9n^2-9n+25n^3-75n^2+50n = 25n^3-66n^2+41n \; \square$$

Now, we will show all elements in B are moves.

Remark 1: Given a contingency table, the entry of the matrix A in the column indexed by $x = \{x_1, x_2, ..., x_n\}$ and row:

$$\left(\sum_{i=1}^{n} x_i, \sum_{i=n+1}^{2n} x_i, \sum_{i=2n+1}^{3n} x_i x_1 + x_{n+1} + x_{2n+1}, x_2 + x_{n+2} + x_{2n+2}, \ldots, x_n + x_{2n} + x_{3n} \right)$$

will be equal to one if x_i a pears in the $\left(\sum_{i=1}^{n} x_i\right)$ and it will zero otherwise. Then:

$$A=\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n}$$

Theorem 4: $B = \{z_1, ..., z_{(25n^3-66n^2+41n)}\}$ is a set of moves.

Proof: Let $z_m \in B$. To show z_m is a move.
By remark 3:

$$A=\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n}$$

We must show that:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ M \\ z_{3n} \end{bmatrix}_{3n\times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(n+3)\times 1}$$

From 1, 2 and 3 we get:

$$\text{If} \quad i=1 \Rightarrow \sum_{j=1}^{3n} a_{1j} z_j = \sum_{j=1}^{3n} z_j = 0$$

$$\text{If} \quad i=2 \Rightarrow \sum_{j=1}^{3n} a_{2j} z_j = \sum_{j=n+1}^{2n} z_j = 0$$

$$\text{If} \quad i=3 \Rightarrow \sum_{j=1}^{3n} a_{3j} z_j = \sum_{j=2n+1}^{3n} z_j = 0$$

$$\text{If} \quad i=4, ..., n+3 \Rightarrow \sum_{j=1}^{3n} a_{ij} z_j = z_j + z_{n+j} + z_{2n+j} = 0, \forall i, \forall j$$

Then Az_m is equal to:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ M \\ z_n \end{bmatrix}_{3n\times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(n+3)\times 1}$$

Implies that, $B \subset \ker A$. Then B is a set of moves (by definition 3).□

Remark 2: Now, we find the elements $x_i \in A^{-1}[t]$, $i = 1, ..., (25n^3-66n^2+41n)$ by using the elements of the set $= \{z_1, ..., z_{(25n^3-66n^2+41n)}\}$, Let z_m be an element of B such that $z_m = x_m - x_{m-1}$, $m = 1, 2, ..., (25n^3-66n^2+41n)-1$ and $z_{(25n^3-66n^2+41n)} = x_0 - x_{(25n^3-69n^2+44n)-1}$ where $x_i \in A^{-1}[t]$, $i = 0, 1, ..., (25n^3-66n^2+41n)-1$.

Corollary 1: The set B of a moves in theorem 4 is a Markov basis.

Proof: From definition 2, for all x_i, $x_j \in A^{-1}$, then $x_i \sim x_j \Leftrightarrow x_i \sim x_j \in \ker (A)$ Let $t = Ax$ for same $t \in N^n$, let x_i, $x_j \in A^{-1}[t]$. By Remark 4, we get:

$$x_i + z_{i+1} = x_{i+1}$$
$$x_{i+1} + z_{i+2} = x_{i+2}$$
$$x_{j-1} + z_j = x_j$$

Implies that:

$$\mathbf{x}_i + \sum_{k=i+1}^{j} \mathbf{z}_k = \mathbf{x}_j$$

$$\Rightarrow x_j - x_i = \sum_{k=i+1}^{j} z_k \quad \text{and}$$

$$z_k \in \text{Ker (A)}$$

For all $z_k \in B$

$$\Rightarrow \sum_{k=i+1}^{j} \mathbf{z}_k \in \text{Ker}(A) \Rightarrow \mathbf{x}_j - \mathbf{x}_i \in \text{Ker}(A) \text{ for all } \mathbf{x}_i, \mathbf{x}_j \in A^{-1}[\mathbf{t}]$$

$$\Rightarrow x_i \sim x_j \text{ for all } x_i, x_j \in A^{-1}[t]$$

$\Rightarrow A^{-1}[t]$ constitutes one B equivalence class for all t.

⇒By definition 8, the set B is a Markov basis. Now, we will find thetoric ideals that corresponding Markov basis for $(25n^3-66n^2+41n)\times3\times n$ contingency tables.

Corollary 2: Let B is a Markov basis for A, the toric ideal I_A for $(25n^3-66n^2+41n)\times3\times n$ contingency tables is: $I_A = \langle P_{i+l}P_{j+r}-P_{j+l}P_{i+r},\ P_{i+l}^2P_{j+r}P_{j+s}-P_{j+l}^2P_{i+r}P_{i+s},\ P_{i+l}^2P_{j+r}P_{k+r}-P_{i+r}^2P_{j+l}P_{k+l},\ P_{i+l}^2P_{j+r}^2P_{k+s}-P_{j+l}^2P_{i+r}P_{i+s}P_{k+r},\ P_{i+l}^2P_{j+r}^2P_{k+s}^2-P_{i+r}^2P_{j+s}^2P_{k+l}^2,\ P_{i+l}^2P_{j+r}^2P_{k+s}^2-P_{i+s}P_{i+r}P_{j+s}P_{j+l}P_{k+r}P_{k+l},\ P_{i+l}^2P_{j+r}P_{k+s}-P_{i+r}P_{i+s}P_{j+l}P_{k+l}$: i, j, k = 1, 2, …, n and l, s, r = 0, n, 2n, such that $i\neq j\neq k$ and $l\neq s\neq r\rangle\subset C[P_1, P_2, \ldots, P_{3n}]$.

Proof: Since, B is a Markov basis for A, by theorem 2 the set of binomials $\{p^{z+}-p^{z-}:z\in B\}$ is a generating of toric ideal I_A and since, $z_m\in B$, m = 1, 2, ..., k is a matrix of dimension 3×n and either it has two columns: (1, -1, 0)', (-1, 1, 0)' or (2, -2, 0)', (-2, 2, 0)' or (1, 0, -1)', (-1, 0, 1)' or (2, 0, -2)', (-2, 0, 2)' or (0, 1, -1)', (0, -1, 1)' or (0, 1, -1)', (0, -2, 2)' or (2, -1, -1)', (-2, 1, 1)' or (-1, 2, -1)', (1, -2, 1)' or (-1, -1, 2)', (1, 1, -2)' and the other columns are zero, or it has three columns: (2, -2, 0)', (-1, 1, 0)', (-1, 1, 0)' or (-2, 2, 0)', (1, -1, 0)', (1, -1, 0)' or (2, 0, -2)', (-1, 0, 1)', (-1, 0, 1)' or (-2, 0, 2)', (1, 0, -1)', (1, 0, -1)' or (0, 2, -2)', (0, -1, 1)', (0, -1, 1)' or (0, -2, 2)', (0, 1, -1)', (0, 1, -1)' or (2, -2, 0)', (0, 2, -2)', (2, 0, -2)' or (-2, 2, 0)', (0,- 2, 2)', (-2, 0, 2)' or (2, -1, -1)' and (-1, 2, -1)', (-1, -1, 2)' or (-1, 1, 0)', (-1, 0, 1)' or (-2, 2, 0)', (0, -1, 1)' or (-2, 0, 2)', (0, 1, -1)' or (-2, 1, 1)' and (1, -2, 1)', (1, 1, -2)' or (1, -1, 0)', (1, 0, -1)' or (2, 0, - 2)', (0, -1, 1)' or (2, 0, -2)', (0, -1, 1)' or (-1, 2, -1)' and (0, -1, 1)', (1, -1, 0)' or (2, -2, 0)', (-1, 0, 1)' or (0, -2, 2)', (-1, 0, 1)' or (1, -2, 1)' and (0, 1, -1)', (-1, 1, 0)' or (-2, 2, 0)', (1, 0, -1)' or (0, 2, -2)', (1, 0, -1)' or (-1, -1, 2)' and (0, 1, -1)', (1, 0, -1)' or (2, 0, -2)', (-1, 1, 0)' or (0, 2, -2)', (1, -1, 0)' or (1, 1, -2)' and (0, 1, -1)', (1, 0, -1)' or (0, -2, 2)', (-1, 1, 0)' or (-2, 0, 2)', (1, -1, 0)' and the other columns are zero. Implies that the toric ideal is the ideal: $I_A = \langle P_{i+l}P_{j+r}-P_{j+l}P_{i+r},\ P_{i+l}^2P_{j+r}P_{j+s}-P_{j+l}^2P_{i+r}P_{i+s},\ P_{i+l}^2P_{j+r}P_{k+r}-P_{i+r}^2P_{j+l}P_{k+l},\ P_{i+l}^2P_{j+r}^2P_{k+s}-P_{j+l}^2P_{i+r}P_{i+s}P_{k+r},\ P_{i+l}^2P_{j+r}^2P_{k+s}^2-P_{i+r}^2P_{j+s}^2P_{k+l}^2,\ P_{i+l}^2P_{j+r}^2P_{k+s}^2-P_{i+s}P_{i+r}P_{j+s}P_{j+l}P_{k+r}P_{k+l},\ P_{i+l}^2P_{j+r}P_{k+s}-P_{i+r}P_{i+s}P_{j+l}P_{k+l}$: i, j, k = 1, 2, …, n and l, s, r = 0, n, 2n, such that, $i\neq j\neq k$ and $l\neq s\neq r\rangle\subset C[P_1, P_2, \ldots, P_{3n}]$.

Example 1: For n = 2, there are 18 moves in a Markov basis according to theorem 3 for 3×2 contingency table, then:

$$\mathbf{B}=\left\{\begin{array}{l}\begin{bmatrix}1&-1\\-1&1\\0&0\end{bmatrix},\begin{bmatrix}1&-1\\0&0\\-1&1\end{bmatrix},\begin{bmatrix}0&0\\1&-1\\-1&1\end{bmatrix},\begin{bmatrix}-1&1\\1&-1\\0&0\end{bmatrix},\begin{bmatrix}-1&1\\0&0\\1&-1\end{bmatrix},\begin{bmatrix}0&0\\-1&1\\1&-1\end{bmatrix},\\ \begin{bmatrix}2&-2\\-2&2\\0&0\end{bmatrix},\begin{bmatrix}2&-2\\0&0\\-2&2\end{bmatrix},\begin{bmatrix}0&0\\2&-2\\-2&2\end{bmatrix},\begin{bmatrix}-2&2\\2&-2\\0&0\end{bmatrix},\begin{bmatrix}-2&2\\0&0\\2&-2\end{bmatrix},\begin{bmatrix}0&0\\-2&2\\2&-2\end{bmatrix},\\ \begin{bmatrix}2&-2\\-1&1\\-1&1\end{bmatrix},\begin{bmatrix}-1&1\\2&-2\\-1&1\end{bmatrix},\begin{bmatrix}-1&1\\-1&1\\2&-2\end{bmatrix},\begin{bmatrix}-2&2\\1&-1\\1&-1\end{bmatrix},\begin{bmatrix}1&-1\\-2&2\\1&-1\end{bmatrix},\begin{bmatrix}1&-1\\1&-1\\-2&2\end{bmatrix}\end{array}\right\}$$

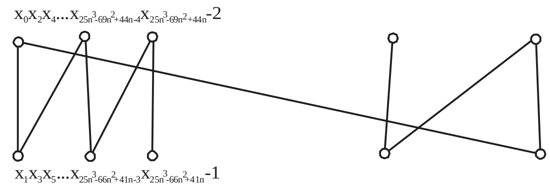

Fig. 2: Graph $G = (R, W, B) = A^{-1}[t]_B$ where the contingency tables explicated as vertices and connecting moves are explicated as edges of a graph, $R = \{x_0, x_2, \ldots, x_{(25n^3-66n^2+41n)-2}\}$ and $W = \{x_1, x_3, \ldots, x_{(25n^3-66n^2+41n-1)}\}$

By corollary 2 the toric ideal of 3×2 contingency table: $I_A = \langle P_{i+l}P_{j+r}-P_{j+l}P_{i+r},\ P_{i+l}^2P_{j+r}P_{j+s}-P_{j+l}^2P_{i+r}P_{i+s}$: i, j, = 1, 2, …, n and l, s, r = 0, n, 2n, such that, $i\neq j$ and $l\neq s\neq r\rangle\subset C[P_1, P_2, P_3, P_4, P_5, P_6]$.

Remark 3: We can find the toric idea by use corollary 2 for $(25n^3-66n^2+41n)\times3\times n$ contingency tables without find a Markov basis.

Remark 4: Now, we constructed a connected graph by use the elements of B. Let z_m be an element of B such that $z_m = x_m-x_{m-1}$, m = 1, 2, ..., $(25n^3-66n^2+41n)\times3\times n$ be an edge connected x_m and x_{m-1}, ... and $z_{(25n^3-66n^2+41n)} = x_0-x_{(25n^3-66n^2+41n)-1}$ be an edge connect x_0 and $x_{(25n^3-66n^2+41n)-1}$, where, $x_i\in A^{-1}[t]$, i = 0, 2, …, $(25n^3-66n^2+41n)-1$. Then we can connected all $(25n^3-66n^2+41n)\times3\times n$ contingency tables with fixed two dimensional marginals by used $(25n^3-66n^2+41n)$ edges by applying moves from B to x_0 one by one and go from x_0 to $x_{(25n^3-66n^2+41n)-1}$ without give rise to negative cell frequencies on the path, also from $x_{(25n3-66n2+41n)-1}$ to x_0 of this kind by forming undirected graph as shown in Fig. 2.

Theorem 5: The graph G = (R, W, B) is a connected bipartite graph (up to graph isomorphism).

Proof: To prove G = (R, W, B) is a connected graph. Let:

$$x_i, x_j\in A^{-1}[t]$$

if, $0\leq i\leq j\leq 25n^3-66n^2+41n-1$, $i\neq j$, by Remark 6, there exists a path $\langle x_i, z_{i+1}, x_{i+1}, z_{i+2}, \ldots, x_{j-1}, z_j, x_j\rangle$ and if $0\leq j\leq i\leq 25n^3-66n^2+41n-1$, $i\neq j$ by Remark 6, there exists a path $\langle x_j, z_{j+1}, x_{j+1}, z_{j+2}, \ldots, x_{i-1}, z_i, x_i\rangle$, in addition, that implies there exists a path between every pair of distinct vertices $x_i, x_j\in A^{-1}[t]$ of the graph by (definition 5), G is a connected graph.

Now, we prove the graph G = (R, W, B) is a bipartite graph. Let $x_i, x_{i+1}, \ldots, x_{j-1}, x_j, x_{j+1} = x_i$ be a cycle in G.

Suppose $x_i\in R$. Then $x_{i+1}\in W$, since, the edge $z_{i+1} = x_{i+1}-x_i\in B$, implies that $x_{i+2}\in R$, since, the edge $z_{i+2} = x_{i+2}-x_{i+1}\in B$. Continuing in this way, we see that if k is odd, then $x_k \in W$ and if k is even then $x_k \in R$. Since, $x_{j+1} = x_i \in R$, it implies that j+1 is even and thus the cycle is of even length. By theorem 1, then the graph G = (R, W, B) is a bipartite graph.

Corollary 3: The set B of moves in theorem 4 is a Markov basis.

Proof: Let $B \subseteq \ker_Z(A)$ be a finite set of moves. From theorem 5 the graph $A^{-1}[t]_B$ is a connected graph. By definition 9 B is a Markov basis for A.

A new module of genetic algorithm of nucleotides in DNA Sequences: In this section, we use an algorithm to make a new model of genetic algorithm of the pieces of Nucleotides in these quences of aligned DNA. Now, we will describe the model in a following steps:

Step 1: Suppose, we have DNA sequences of m taxons, each taxon of length L like:

- **Taxon1:** ATCGA ACGGTA TGT⋯
- **Taxon2:** AGATCAGAAC CGAT⋯

  ⋮⋮⋮⋮⋮⋮⋮⋮⋮⋮⋮

- **Taxonm:** GCGTAGCGTGGCAC

Then, defined a pattern $i = i_1, i_2, \ldots, i_m$ to be the sequence of characters. We will look at a single site (column) of our sequence data. In the sequences above, we look at the first site in the sequences and we will see the pattern "AC ... G". A pattern frequency x_i is that i appears in our set of sequence data and we refer to the number of frequencies by 3n where n = 2> = 2.

Step 2: We can input the patterns frequency x_i of above sequences in 3×n:

x_1	x_2	⋯	x_n	$\sum_{i=1}^{n} x_i$
x_{n+1}	x_{n+2}	⋯	x_{2n}	$\sum_{i=n+1}^{2n} x_i$
x_{2n+1}	x_{2n+2}	⋯	x_{3n}	$\sum_{i=2n+1}^{3n} x_i$
$x_1+x_{n+1}+x_{2n+1}$	$x_2+x_{n+2}+x_{2n+2}$	⋯	$x_n+x_{2n}+x_{3n}$	$\lvert x \rvert = \sum_{i \in I} x_i$

contingency tables as follows:

where $|x| = \sum_{i \in I} x_i = L$ is the length of sequences (the sample size) and:

x_1 : Frequency of the first pattern
x_2 : Frequency of the second pattern
x_n : Frequency of the n pattern
x_{n+1} : Frequency of the n+1 pattern
x_{n+2} : Frequency of the n+2 pattern
⋮
x_{3n} : Frequency of the 3n pattern

Step 3: From remark 3, A is (n+3)×3n matrix and:

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n}$$

where, Ax = t is written as:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{(n+3)\times 3n} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{3n} \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ \vdots \\ t_{n+3} \end{bmatrix}$$

where the columns of the matrix A index by the elements of the column vector x.

Step 4: Find the Markov basis by use Corollary 1, therefore, the set is written as:

z_1	z_2	...	z_n	0
z_{n+1}	z_{n+2}	...	z_{2n}	0
z_{2n+1}	z_{2n+2}	...	z_{3n}	0
0	0	0	0	0

where z_s = 0, 1, 2, -1 or -2, s = 1, 2, ..., 3n as in Eq. 1, 2, and 3.

Step 5: Find the toric ideal by use Corollary 2.

Step 6: Find $(25n^3-66n^2+41n)\times 3\times n$ contingency tables where $A^{-1}[t]_B$ connected for all t by use the Markov basis.

Step 7: Using the contingency tables to describe apermutation of the pieces of nucleotides in aligned DNA sequences to $(25n^3-66n^2+41n)\times 3\times n$ contingency tables.

Example 2: Suppose we have three sequences of analigned DNA as followed:

- **Taxon 1:** CGATGCCCGATTTGGGC
- **Taxon 2:** AC TCGTAAC TCCCGGGT
- **Taxon 3:** ACA GACAACA GGGAAAC

Step 1: There are three taxons for above DNA sequences with $|x| = \sum_{i \in I} x_i = L = 16$ and six patterns CA A, G CC, ATA, TC G, GGA, CTC, CA A, CA A, G CC, TTA, TCG, TC G, TC G, GGA, GGA and GGA with frequencies 3, 2, 2, 4, 4 and 1, respectively where, 3n = 6.

Step 2: Now, input the patterns frequency x_i of above sequences in 3×2 contingency table as follows:

3	2	5
2	4	6
4	1	5
8	7	16

Step 3: A is 5×6 matrix and:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}_{5\times6} \quad \text{and}$$

Ax = t, i.e:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}_{5\times6} \begin{bmatrix} 3 \\ 2 \\ 2 \\ 4 \\ 4 \\ 1 \end{bmatrix}_{6\times1}$$

$$= \begin{bmatrix} 5 \\ 6 \\ 5 \\ 8 \\ 7 \end{bmatrix}_{5\times1} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}_{5\times1}$$

Where, the columns of the matrix A index by the elements of the column vector x and:

$$t_1 = x_1+x_2 = 5,\ t_2 = x_3+x_4 = 4,\ t_3 = x_5+x_6 = 6$$
$$t_4 = x_1+x_3+x_5 = 7,\ t_5 = x_2+x_4+x_6 = 8$$

Step 4: We find the Markov basis from Eq. 1, 2 and 3. Then the number of moves is $25n^3-66n^2+41n = 18$ elements in the set:

z_1=

2	-2
-2	2
0	0

, z_2=

-1	1
1	-1
0	0

, z_3=

0	0
1	-1
-1	1

z_4=

1	-1
-1	1
0	0

, z_5=

0	0
2	-2
-2	2

, z_6=

-1	1
2	-2
-1	1

z_7=

-1	1
0	0
2	-2

, z_8=

-2	2
2	-2
0	0

, z_9=

0	0
-1	1
1	-1

z_{10}=

-1	1
0	0
1	-1

, z_{11}=

1	-1
-2	2
1	-1

, z_{12}=

1	-1
1	-1
-2	2

z_{13}=

2	-2
0	0
-2	2

, z_{14}=

-2	2
1	-1
1	-1

, z_{15}=

1	-1
0	0
-1	1

z_{16}=

2	-2
-1	1
-1	1

, z_{17}=

0	0
-2	2
2	-2

, z_{18}=

-2	2
0	0
2	-2

Step 5: We find the toric ideal $I_A = \langle P_{i+l}P_{j+r}- P_{j+l}P_{i+r}, P_{i+l}^2 P_{j+r}P_{j+s}P_{j+l}^2P_{i+r}P_{i+s}$: I, j, = 1, 2, ..., n and l, s, r = 0, n, 2n, such that $i \neq j$ and $l \neq s \neq r\rangle = \langle P_1P_4-P_2P_3, P_1P_6-P_2P_5, P_3P_6-P_4P_5, P_1^2P_4P_6-P_2^2P_3P_5, P_1P_4^2P_5-P_2P_3^2P_6, P_1P_3P_6^2-P_2P_4P_5^2\rangle \subset C [P_1, P_2, P_3, P_4, P_5, P_6]$ and by using corollary 2.

Step 6: The connected graph $A^{-1}[t]_B = G = (R, W, B)$ with $\frac{n^2-3n}{3} = 18$ (t-fibres) 3×2 contingency tables as vertices of it (Fig. 3).

Where:

x_0=

3	2	5
2	4	6
4	1	5
9	7	16

, x_1=

5	0	5
0	6	6
4	1	5
9	7	16

, x_2=

4	1	5
1	5	6
4	1	5
9	7	16

x_3=

4	1	5
2	4	6
3	2	5
9	7	16

, x_4=

5	0	5
1	5	6
3	2	5
9	7	16

, x_5=

5	0	5
3	3	6
1	4	5
9	7	16

x_6=

4	1	5
5	1	6
0	5	5
9	7	16

, x_7=

3	2	5
4	2	6
2	3	5
9	7	16

, x_8=

1	4	5
6	0	6
2	3	5
9	7	16

x_9=

1	4	5
5	1	6
3	2	5
9	7	16

, x_{10}=

0	5	5
5	1	6
4	1	5
9	7	16

, x_{11}=

1	4	5
3	3	6
5	0	5
9	7	16

x_{12}=

2	3	5
4	2	6
3	2	5
9	7	16

, x_{13}=

4	1	5
4	2	6
1	4	5
9	7	16

, x_{14}=

2	3	5
5	1	6
2	3	5
9	7	16

x_{15}=

3	2	5
5	1	6
1	4	5
9	7	16

, x_{16}=

5	0	5
4	2	6
0	5	5
9	7	16

, x_{17}=

5	0	5
2	4	6
2	3	5
9	7	16

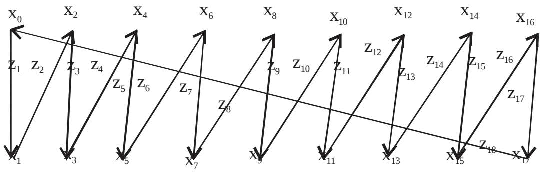

Fig. 3: Graph $G = (R, W, B) = A^{-1}[t]_B$ where the contingency tables explicated as vertices and connecting Markov basis are explicatedas edges of a graph, $R = \{x_0, x_2, x_4, x_6, x_8, x_{10}, x_{12}, x_{14}, x_{16}\}$ and $W = \{x_1, x_3, x_5, x_7, x_9, x_{11}, x_{13}, x_{15}, x_{17}\}$

Step 7: The change in the kind of DNA sequences under the Markov basis. Be as Fig. 2, where:

$x_0 =$ ③②②④④①
Taxon1: CGATGCC CGATTTGGG
Taxon2: AC TCGTAAC TCCCGGG
Taxon3: ACA GACAACA GGGAAA

$x_1 =$ ⑤⓿⓿⑥④①
Taxon1: CTGCC CC CTTTTTGGG
Taxon2: ACGT AAAACCCCCGGG
Taxon3: AGAC AA AAGGGGGAAA

$x_2 =$ ④①①⑤④①
Taxon1: CGATGCC CCTTTTGGG
Taxon2: AC TCGTAAACCCCGGG
Taxon3: ACA GACAAAGGGGAAA

$x_3 =$ ④①②④③②
Taxon1: CGATGCC CCATTTGGC
Taxon2: AC TCGTAAA TCCCGGT
Taxon3: ACA GACAAAA GGGAAC

$x_4 =$ ⑤⓿①⑤③②
Taxon1: CATGCCC CCTTTTGGC
Taxon2: A TCGTAAAACCCCGGT
Taxon3: AA GACAAAAGGGGAAC

$x_5 =$ ⑤⓿③③①④
Taxon1: CATGCCC CCAATTCCC
Taxon2: A TCGTAAAA TTCC TT T
Taxon3: AA GACAAAAAAGGCCC

$x_6 =$ ④①⑤①⓿⑤
Taxon1: CGATCC CCA AAACCCC
Taxon2: AC TCTAAA TTTT TTTT
Taxon3: ACA GCAAAA AAACCCC

$x_7 =$ ③②④②②③
Taxon1: CGATGCC CGA AATGCC
Taxon2: AC TCGTAAC TTTCGTT
Taxon3: ACA GACAACA AAGACC

$x_8 =$ ①④⑥⓿②③
Taxon1: CGAGCGGGA AA AAGCC
Taxon2: AC TGTCCC TTT TTGTT
Taxon3: ACA ACCCCA AAA AACC

$x_9 =$ ①④⑤①③②
Taxon1: CGATGCGGGA A AAGGC
Taxon2: AC TCGTCCC TTT TGGT
Taxon3: ACA GACCCCA AAAAAC

$x_{10} =$ ⓿⑤⑤①④①
Taxon1: GATGCGGGGTTTTGGG
Taxon2: C TCGTCCCC TTT TGGG
Taxon3: CA GACCCCCA AAAAAA

$x_{11} =$ ①④③③⑤⓿
Taxon1: CGATGGGGAATTGGGG
Taxon2: AC TCGCCC TTCCGGGG
Taxon3: ACA GACCCA AGGAAAA

$x_{12} =$ ②③④②③②
Taxon1: CGATGCCGGA AATGGC
Taxon2: AC TCGTACC TTTCGGT
Taxon3: ACA GACACCA AAGAAC

$x_{13} =$ ④①④②①④
Taxon1: CGATGCC C CA AATCCC
Taxon2: AC TCGTAAA TTTC TTT
Taxon3: ACA GACAAAA AAGCCC

$x_{14} =$ ②③⑤①②③
Taxon1: CGATGCCGGA AAAGCC
Taxon2: AC TCGTACC TTT TGTT
Taxon3: ACA GACACCA AAAACC

$x_{15} =$ ③②⑤①①④
Taxon1: CGATGCCCGA AAACCC
Taxon2: AC TCGTAAC TTT T TTT
Taxon3: ACA GACAACA AAACCC

$x_{16} =$ ⑤⓿④②⓿⑤
Taxon1: CATCCC CCA AATCCCC
Taxon2: A TCTAAAA TTTC TTTT
Taxon3: AA GCAAAAAAAGCCCC

$x_{17} =$ ⑤⓿④②④①
Taxon1: CATGCCC CCA AATGGG
Taxon2: A TCGTAAAA TT TCGGG
Taxon3: AA GACAAAAAAAGAAA

The permutation of thepieces of nucleotides inaligned DNA sequences under the set of Markov basis in example 3.

Remark 5:

- We refer to ①, ②, ③, ④, ⑤ and ⑥ in Step 7 to the frequencies of the patterns in DNA sequences
- We refer to in the same figure to the hidden in the pattern frequency of DNA sequences

CONCLUSION

Dissection: In this study, we introduce a new model to change the pieces of nucleotides in aligned DNA sequences by using Markov basis and we shows that for Given A and t, there exists finite $B \subseteq \ker_Z(A)$ such that $A^{-1}[t]_B$ is constitutes one B equivalence class.

REFERENCES

Agnarsson, G. and R. Greenlaw, 2007. Graph Theory, Modeling, Applications and Algorithms. Pearson/Prentice Hall, Upper Saddle River, New Jersey, USA., ISBN:9780131423848, Pages: 446.

Aoki, S. and A. Takemura, 2003a. Minimal basis for a connected Markov chain over 3×3×K contingency tables with fixed two dimensional marginals. Aust. N. Z. J. Stat., 45: 229-249.

Aoki, S. and A. Takemura, 2003b. The list of indispensable moves of the unique minimal markov basis of 3×4×K and 4×4×K contingency tables with fixed two-dimensional marginals. Master Thesis, Department of Mathematical Engineering and Information Physics, The University of Tokyo, Tokyo, Japan.

Aoki, S. and A. Takemura, 2008. The largest group of invariance for Markov bases and toric ideals. J. Symbolic Comput., 43: 342-358.

Clark, J. and D.A. Holton, 1995. A First Look at Graph Theory. World Scientific Publishing, Singapore, ISBN:81-7023-463-8, Pages: 325.

Diaconis, P. and B. Sturmfels, 1998. Algebraic algorithms for sampling from conditional distributions. Ann. Stat., 26: 363-397.

Diaconis, P., D. Eisenbud and B. Sturmfels, 1998. Lattice Walks and Primary Decomposition. In: Mathematical Essays in Honor of Gian-Carlo Rota, Sagan, B. and R. Stanley (Eds.). Birkhauser Basel, Basel, Switzerland, ISBN:978-1-4612-8656-1, pp: 173-193.

Dobra, A. and S. Sullivant, 2002. A divide-and-conquer algorithm for generating markov bases of multi-way tables. Proceedings of the Workshop on Multidimensional Tables Statistics, Combinatorial Optimization and Groebner Bases, November 17-20, 2002, University of California Davis, Davis, California, pp: 1-20.

Dobra, A., 2003. Markov bases for decomposable graphical models. Bernoulli, 9: 1093-1108.

Russell, P., 2001. Genetics. Benjamin Cummings, New York, USA.,.

Sullivant, S., J. Rauh and T. Kahle, 2012. Postive margins and primary decomposition. Master Thesis, North Carolina State University, Raleigh, North Carolina.

Sullivant, S.M., 2005. Toric ideals in algebraic statistics. Ph.D Thesis, University of California, Berkeley, California.

Takemura, A. and S. Aoki, 2004. Some characterizations of minimal Markov basis for sampling from discrete conditional distributions. Annal. Inst. Stat. Math., 56: 1-17.

Takemura, A. and S. Aoki, 2005. Distance-reducing Markov bases for sampling from a discrete sample space. Bernoulli, 11: 793-813.