Czechoslovak Mathematical Journal, 49 (124) (1999), 249–254

TRIANGULAR STOCHASTIC MATRICES GENERATED BY INFINITESIMAL ELEMENTS

Inheung Chon and Hyesung Min, Seoul

(Received February 19, 1996)

Abstract. We show that each element in the semigroup S_n of all $n\times n$ non-singular upper (or lower) triangular stochastic matrices is generated by the infinitesimal elements of S_n, which form a cone consisting of all $n\times n$ upper (or lower) triangular intensity matrices.

MSC 2000: 22E99

1. Introduction

Let G be a Lie group, let $L(G)$ be its Lie algebra, and let $\exp\colon L(G)\to G$ denote the exponential mapping. Let $\mathrm{gl}(n,\mathbb{R})$ denote the set of all real $n\times n$ matrices and $\mathrm{GL}(n,\mathbb{R})$ the general linear group of degree n over $\mathbb{R}$. Here $\mathbb{R}$ denotes the set of all real numbers and hereafter we shall use this notation. For $G=\mathrm{GL}(n,\mathbb{R})$ and $L(G)=\mathrm{gl}(n,\mathbb{R})$, it is well known that the exponential map exp: $\mathrm{gl}(n,\mathbb{R})\to\mathrm{GL}(n,\mathbb{R})$ is defined by $\exp(tX)=I+tX+\frac{1}{2!}(tX)^2+\dots$ for $X\in\mathrm{gl}(n,\mathbb{R})$.

Let S_n be a subsemigroup of $\mathrm{GL}(n,\mathbb{R})$ and let $X(t)$ be a differentiable matrix function of the real parameter t in an interval $0\leqslant t\leqslant t_0$ such that $X(t)\in S_n$ for each t and $X(0)=I$. We call the matrix $(\frac{\mathrm{d}X(t)}{\mathrm{d}t})|_{t=0}$ an *infinitesimal element* of S_n and denote the totality of all infinitesimal elements of S_n by $\mathscr{D}(S_n)$. Let $A(t)$ be a sectionwise continuous function of t $(0\leqslant t\leqslant t_0)$ such that $A(t)\in\mathscr{D}(S_n)$ for each t. It is standard that the differential equation

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t}=A(t)X(t);\quad X(0)=I$$

This paper was supported by the Natural Science Research Institute of Seoul Women's University, 1994

has a unique continuous solution and $X(t_0) \in S_n$. This $X(t_0)$ in S_n is called *generated by the infinitesimal elements* $A(t)$ $(0 \leqslant t \leqslant t_0)$.

Loewner [3] showed that each element in the semigroup of all $n \times n$ non-singular totally positive matrices is generated by the infinitesimal elements of the semigroup, which form a set of all $n \times n$ Jacobi matrices with non-negative off-diagonal elements. In general, a semigroup is not completely recreated from its infinitesimal elements, even if the semigroup is connected, and it is quite difficult to compute a semigroup generated by its infinitesimal elements.

In Section 2, we show that the infinitesimal elements of the semigroup of all $n \times n$ non-singular upper (or lower) triangular stochastic matrices are $n \times n$ upper (or lower) triangular intensity matrices. Finally, in Section 3, we show that each element in the semigroup S_n of all $n \times n$ non-singular upper (or lower) triangular stochastic matrices is generated by the infinitesimal elements of S_n, which form a cone consisting of all $n \times n$ upper (or lower) triangular intensity matrices.

2. Infinitesimal elements of triangular stochastic matrices

Definition. A matrix $A = \|a_{ij}\|$ $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$ over $\mathbb{R}$ is called a *stochastic matrix* if $a_{ij} \geqslant 0$ and $\sum\limits_{j=1}^{n} a_{ij} = 1$ for $i = 1, 2, \ldots, m$. A matrix $B = \|b_{kl}\|$ $(k = 1, 2, \ldots, m;\ l = 1, 2, \ldots, n)$ over $\mathbb{R}$ such that $b_{kl} \geqslant 0$ for $k \neq l$ and $\sum\limits_{l=1}^{n} b_{kl} = 0$ for $k = 1, 2, \ldots, m$ is called an *intensity matrix*. An intensity matrix C is called an *extreme intensity matrix* if C has only one nonzero off-diagonal element which is equal to 1. An extreme intensity matrix $C = \|c_{kl}\|$ is denoted by E_{pq} $(p \neq q)$ if $c_{pp} = -1$ and $c_{pq} = 1$.

It is easy to see that the set of all non-singular $n \times n$ stochastic matrices forms a subsemigroup of $\mathrm{GL}(n, \mathbb{R})$.

Lemma 2.1. *Let S_n be the semigroup of all real $n \times n$ non-singular matrices with non-negative entries. Then $\mathscr{D}(S_n)$ coincides with the set of all real $n \times n$ matrices which are non-negative off the diagonal.*

P r o o f. Let $A = \|a_{ij}\| \in \mathscr{D}(S_n)$. Then $A = (\frac{\mathrm{d}X(t)}{\mathrm{d}t})|_{t=0}$ with $X(t) \in S_n$ for each t and $X(0) = I$. Since $X(t) \in S_n$, $x_{ij}(t) \geqslant 0$ for $i, j = 1, 2, \ldots n$. From $X(0) = I$, $x_{ij}(0) = 0$ for $i \neq j$. Thus $a_{ij} = (\frac{\mathrm{d}x_{ij}(t)}{\mathrm{d}t})|_{t=0} \geqslant 0$ for $i \neq j$.

Conversely let $E_{ij} (i \neq j)$ be an extreme intensity matrix as denoted in the above definition. Since ${E_{ij}}^2 = -E_{ij}$, $\exp(tE_{ij}) = I + tE_{ij} - \frac{t^2}{2!}E_{ij} + \frac{t^3}{3!}E_{ij} + \ldots = I + (1 - \mathrm{e}^{-t})E_{ij}$, and hence $\exp(tE_{ij}) \in S_n$ for $t \geqslant 0$. Since $E_{ij} = \frac{\mathrm{d}}{\mathrm{d}t}(\exp(tE_{ij}))|_{t=0}$,

$E_{ij} \in \mathscr{D}(S_n)$. Let E_k be the matrix whose elements are 0 except that the k-th diagonal element is equal to 1. Since ${E_k}^2 = E_k$, $\exp(tE_k) = I + tE_k + \frac{t^2}{2!}E_k + \frac{t^3}{3!}E_k + \ldots = I + (\mathrm{e}^t - 1)E_k$, and hence $\exp(tE_k) \in S_n$ for $t \geqslant 0$. Thus $E_k \in \mathscr{D}(S_n)$. Similarly we may show $-E_k \in \mathscr{D}(S_n)$. Since $\mathscr{D}(S_n)$ forms a convex cone in the matrix space $\mathrm{gl}(n, \mathbb{R})$, $\sum\limits_{1 \leqslant i \neq j \leqslant n} \alpha_{ij} E_{ij} + \sum\limits_{k=1}^{n} \beta_k E_k - \sum\limits_{k=1}^{n} \gamma_k E_k \in \mathscr{D}(S_n)$ for all α_{ij}, β_k, $\gamma_k \geqslant 0$. Thus every real $n \times n$ matrix which is non-negative off the diagonal is contained in $\mathscr{D}(S_n)$. □

Lemma 2.2. *Let T_n be the semigroup of all real non-singular $n \times n$ matrices with each row sum equal to* 1. *Then*

$$\mathscr{D}(T_n) = \left\{ \|c_{ij}\| \in \mathrm{gl}(n, \mathbb{R}) \colon \sum_{j=1}^{n} c_{ij} = 0 \ \textit{for}\ i = 1, 2, \ldots, n \right\}.$$

P r o o f. Let $\Omega = \|\omega_{ij}\| \in \mathscr{D}(T_n)$. Then there exists $U(t) \in T_n$ such that $\Omega = (\frac{\mathrm{d}U(t)}{\mathrm{d}t})|_{t=0}$, $\sum\limits_{j=1}^{n} u_{ij}(t) = 1$ for $i = 1, 2, \ldots, n$, and $U(0) = I$. Hence

$$\begin{aligned} \sum_{j=1}^{n} \omega_{ij} &= \sum_{j=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t}(u_{ij}(t))|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{j=1}^{n} u_{ij}(t)\right)\Bigg|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(1)\Big|_{t=0} = 0 \ \text{ for } i = 1, 2, \ldots, n. \end{aligned}$$

□

Conversely suppose that $C = \|c_{ij}\|$ with $\sum\limits_{j=1}^{n} c_{ij} = 0$ for $i = 1, 2, \ldots, n$. Let

$$W = \left\{ \|b_{ij}\| \in \mathrm{gl}(n, \mathbb{R}) \colon \sum_{j=1}^{n} b_{ij} = 0 \text{ for } i = 1, 2, \ldots, n \right\}.$$

Then W is a cone in $\mathrm{gl}(n, \mathbb{R})$ and $C \in W$. Also

$$C = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tC}\Big|_{t=0} = \lim_{t \to 0^+} \frac{\mathrm{e}^{tC} - I}{t}.$$

Since $C \in W$ and W is a cone, $\exp(tC) \in I + tW = I + W$ for $t \geqslant 0$. Since $\exp(tC)$ is non-singular, $\exp(tC) \in \mathrm{GL}(n, \mathbb{R}) \cap (I + W) \subset T_n$. Thus $C \in \mathscr{D}(T_n)$.

Lemma 2.3. *Let S_n be the semigroup of all $n \times n$ non-singular stochastic matrices. Then $\Omega = \|\omega_{ij}\|$ is an element of $\mathscr{D}(S_n)$ iff Ω is an $n \times n$ intensity matrix.*

P r o o f. It is clear that if S_n and T_n are subsemigroups of $\mathrm{GL}(n, \mathbb{R})$, then $\mathscr{D}(S_n \cap T_n) = \mathscr{D}(S_n) \cap \mathscr{D}(T_n)$. Thus the lemma is proved from Lemma 2.1 and Lemma 2.2. □

Theorem 2.4. *Let S_n be the semigroup of all $n \times n$ non-singular upper (or lower) triangular stochastic matrices. Then A is an element of $\mathscr{D}(S_n)$ iff A is an $n \times n$ upper (or lower) triangular intensity matrix.*

P r o o f. It is obvious that if T_n is the semigroup of all real $n \times n$ non-singular upper (or lower) triangular matrices, A is an element of $\mathscr{D}(T_n)$ iff A is a real $n \times n$ upper (or lower) triangular matrix. Hence the theorem is proved from Lemma 2.3.

□

3. Infinitesimally generated triangular stochastic matrices

Lemma 3.1. *Let A be an $n \times n$ non-singular upper triangular stochastic matrix of the following form:*

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{pp} & a_{pp+1} & \dots & a_{pn} \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Then A can be represented as $A = \exp(t_{pp+1}E_{pp+1})\exp(t_{pp+2}E_{pp+2})\dots\exp(t_{pn}E_{pn})$, where E_{ij} is an extreme intensity matrix as denoted in the definition of Section 2.

P r o o f. Since A is stochastic, $a_{pp} + a_{pp+1} + \dots + a_{pn} = 1$. Since A is upper triangular and non-singular, determinant of $A = a_{pp} > 0$. Let

$$x_{p+i} = \frac{a_{pp} + a_{pp+i+1} + \dots + a_{pn}}{a_{pp} + a_{pp+i} + \dots + a_{pn}} \quad \text{for } i = 1, 2, \dots, n.$$

Then $0 < x_{p+i} \leqslant 1$ for $i = 1, 2, \dots, n$ since $a_{pp} > 0$. For $i = 1$, $x_{p+1} = a_{pp} + a_{pp+2} + \dots + a_{pn}$. Thus $a_{pp+1} = 1 - x_{p+1}$. Now,

$$x_{p+2} = \frac{a_{pp} + a_{pp+3} + \dots + a_{pn}}{a_{pp} + a_{pp+2} + \dots + a_{pn}} = \frac{a_{pp} + a_{pp+3} + \dots + a_{pn}}{x_{p+1}}.$$

Hence $a_{pp+2} = x_{p+1} - x_{p+1}x_{p+2} = x_{p+1}(1 - x_{p+2})$. Inductively,

$$x_{p+1}x_{p+2}\dots x_{p+k-1} = a_{pp} + a_{pp+k} + \dots + a_{pn}$$

for $k = 2, \dots, n-p$ and

$$x_{p+1}x_{p+2}\dots x_{p+k-1}x_{p+k} = a_{pp} + a_{pp+k+1} + \dots + a_{pn}.$$

Therefore

$$a_{pp+k} = x_{p+1}\dots x_{p+k-1}(1 - x_{p+k}) \quad \text{for } k = 2, \dots, n-p.$$

We have

$$\begin{aligned} 1 &= a_{pp} + a_{pp+1} + a_{pp+2} + \dots + a_{pn} \\ &= a_{pp} + (1 - x_{p+1}) + x_{p+1}(1 - x_{p+2}) + \dots + x_{p+1}\dots x_{n-1}(1 - x_n) \\ &= a_{pp} + 1 - x_{p+1}\dots x_n. \end{aligned}$$

Hence $a_{pp} = x_{p+1}x_{p+2}\dots x_n$. Let $A_{x_{p+j}}$ $(j = 1, 2, \dots, n-p)$ be an $n \times n$ upper triangular stochastic matrix of the following form:

$$A_{x_{p+j}} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & x_{p+j} & 0 & \dots & 1 - x_{p+j} & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix},$$

where x_{p+j} is in the pth low and pth column and $1 - x_{p+j}$ is in the pth low and $p+j$th column. Then $A = A_{x_{p+1}}A_{x_{p+2}}\dots A_{x_n}$. Since $0 < x_{p+j} \leqslant 1$, $A_{x_{p+j}} = \exp(t_{pp+j}E_{pp+j})$ for some $t_{pp+j} \geqslant 0$. Thus $A = \exp(t_{pp+1}E_{pp+1})\exp(t_{pp+2}E_{pp+2})\dots \exp(t_{pn}E_{pn})$. □

Lemma 3.2. *If U is an $n \times n$ non-singular upper triangular stochastic matrix, then it can be represented as $U = C_{n-1}C_{n-2}\dots C_1$, where $C_p = \exp(t_{pp+1}E_{pp+1})\dots \exp(t_{pn}E_{pn})$ for $p = 1, 2, \dots, n-1$ and $t_{ij} \geqslant 0$.*

Analogously, if L is an $n \times n$ non-singular lower triangular stochastic matrix, then it can be represented as $L = H_2H_3\dots H_n$, where $H_p = \exp(s_{p1}E_{p1})\exp(s_{p2}E_{p2})\dots \exp(s_{pp-1}E_{pp-1})$ for $p = 2, \dots, n$ and $s_{ij} \geqslant 0$.

P r o o f. Let $U_1, \ldots, U_n$ be the rows of U such that $U = (U_1, \ldots, U_n)^t$ and I_j be the jth row of $n \times n$ identity matrix. Then $U = C_{n-1}C_{n-2}\ldots C_1$, where C_p is an $n \times n$ matrix such that $C_p = (I_1, I_2, \ldots, I_{p-1}, U_p, I_{p+1}, \ldots, I_n)^t$ for $p = 1, 2, \ldots, n-1$. According to the Lemma 3.1, $C_p = \exp(t_{pp+1}E_{pp+1})\ldots\exp(t_{pn}E_{pn})$.

The proof for the lower triangular case is similar to that for the upper triangular case. □

Theorem 3.3. *Each element in the semigroup S_n of all $n \times n$ non-singular upper (or lower) triangular stochastic matrices is generated from the infinitesimal elements of S_n, which form a cone consisting of all $n \times n$ upper (or lower) triangular intensity matrices.*

P r o o f. Immediate from Theorem 2.4 and Lemma 3.2. □

References

[1] *I. Chon*: Lie group and control theory. Ph.D. thesis at Louisiana state university, 1988.
[2] *F. R. Gantmacher*: The Theory of Matrices vol. 1 and vol. 2. Chelsea Publ. Comp., New York, 1960.
[3] *C. Loewner*: On totally positive matrices. Math. Zeitschr. *63* (1955), 338–340.
[4] *C. Loewner*: A theorem on the partial order derived from a certain transformation semigroup. Math. Zeitschr. *72* (1959), 53–60.
[5] *H. Min*: One parameter semigroups in Lie groups. Master's thesis at Seoul women's university, 1995.
[6] *V. S. Varadarajan*: Lie Groups, Lie Algebras, and Their Representations. Springer-Verlag, New York, 1984.

Authors' address: Department of Mathematics, Seoul Women's University, Kongnung 2-Dong, Nowon-Ku, Seoul, 139-774, Korea.