International Journal of Mathematics and
Computer Science, **12**(2017), no. 2, 107–115

On the Helmert Matrix and Application in Stochastic Processes

Reza Farhadian, Nader Asadian

Department of Mathematics and Statistics
Lorestan university
khorramabad, Iran

email: farhadian.reza@yahoo.com, asadian.n@lu.ac.ir

(Received April 2, 2017, Accepted May 1, 2017)

Abstract

In this work we prove that if $H_n = [h_{ij}]_{n\times n}$ is the well-known Helmert matrix, then $P_n = [h_{ij}^2]_{n\times n}$ is a regular doubly stochastic matrix. Also, we show that the Markov chain by stochastic matrix P_n is an ergodic chain. Afterwards, we prove that $\lim_{k\to\infty} P_n^k = \frac{1}{n}[1]_{n\times n}$, which is the stationary distribution for the stochastic matrix P_n.

1 Introduction

In linear algebra and matrix theory there are many special and important matrices. For example, the well-known *Helmert matrix* is one of them. A Helmert matrix of order n is a square matrix that was introduced by H. O. Lancaster in 1965 [7]. Usually, the Helmert matrix is used in mathematical statistics for analysis of variance (ANOVA) [2, 3, 13]. In this work, we will show that the Helmert matrix can be used in stochastic processes. In fact, we know that in modern probability theory and dynamical systems, the stochastic matrices are nonnegative real matrices that are used for showing the transition probabilities [10, 12]. In this article we will show that one can construct a stochastic matrix using a Helmert matrix. For the next sections, the following notation will be used:

Key words and phrases: Helmert matrix; Stochastic matrix; Stationary distribution.
AMS (MOS) Subject Classifications: 15B51, 15B10, 60J10, 60G10.
7 **ISSN** 1814-0432, 2017, http://ijmcs.future-in-tech.net

(a) I_n denotes an identity matrix of order n.
(b) J_n denotes an $n \times n$ matrix whose elements are all 1.
(c)$A_n > 0$ stands for a matrix A_n all of whose elements are positive.
(d) A_n^T denotes the transpose of a matrix A_n.

2 Definition and Particulars

2.1 Some definitions of linear-algebra and matrix theory
In this part, we present some important definitions and properties of linear algebra and matrix theory [5, 14]:

Definition 2.1. *The n-dimensional vector* $V = \begin{bmatrix} v_1 & v_2 & v_3 & \dots & v_n \end{bmatrix}$ *is called unit vector if*

$$V.V = VV^T = v_1^2 + v_2^2 + v_3^2 + ... + v_n^2 = 1. \tag{2.1}$$

Note that in the above product, $V.V$ is *the Standard inner or Dot product* and VV^T is *the Cayley or Matrix product*, which in this case are both equivalent. The generalized form of this definition for a square matrix of order n is defined as:

Definition 2.2. *The square matrix* $A_n = [aij]_{n\times n}$ *is called orthogonal, if*

$$A_n A_n^T = A_n^T A_n = I_n. \tag{2.2}$$

In other words, the rows and columns of A_n *are unit vectors.*

Definition 2.3. *A nonnegative square matrix* $A_n = [a_{ij}]_{n\times n}$ *is called a regular or semi-positive matrix, if there is a finite positive integer* m *such that* $A_n^m > 0$.

Definition 2.4. *A real* $n \times n$ *matrix such as* $P_n = [p_{ij}]_{n\times n}$ *is called a row stochastic or probability or transition matrix, if*
i) $p_{ij} \geq 0, \qquad 1 \leq i, j \leq n$
ii) $\sum_{j=1}^{n} p_{ij} = 1, \qquad \forall i \in \{1, 2, ..., n\}$

Definition 2.5. *A doubly stochastic matrix, is a square matrix of nonnegative real number with each row and column summing to 1 (in other words, a stochastic matrix is doubly, if its transpose is stochastic matrix).*

Definition 2.6. *A non-zero row vector* $T = \begin{bmatrix} t_1 & t_2 & t_3 & \dots & t_n \end{bmatrix}$, *is called a fixed point of an* $n \times n$ *matrix* A, *if*

$$TA = T. \tag{2.3}$$

Remark 2.7. *The fixed point of a stochastic matrix is usually called the stationary distribution and is denoted by* $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \dots & \pi_n \end{bmatrix}$.

2.2 Stochastic processes and Markov chins

Let $\ell = \{0, 1, 2, ..., d\}$ be the state space of a stochastic process and suppose that $i, j \in \ell$ are two different states of this process. Now, suppose that process starts from state i to state j. This transition shown by $i \longrightarrow j$, and p_{ij} denote its probability. Now, if X_t denotes the state at time t, then the transition $i \longrightarrow j$ at time t, is indicated by $X_t = i$ and $X_{t+1} = j$. Furthermore, a process is called a Markov chain if the transition probability p_{ij} is independent of time t for every states i and j of state space. Hence, the transition probability under the Markov property, is defined as follows:

$$p_{ij} = Prob\{X_{t+1} = j | X_t = i, X_{t-1} = i_{n-1}, ..., X_0 = i_0\} = Prob\{X_{t+1} = j | X_t = i\}. \tag{2.4}$$

For every $i_1, i_2, ..., i_{t-1}, i, j \in \ell$, we can construct a $(d+1) \times (d+1)$ matrix by the p_{ij}, $0 \leq i, j \leq d$ that is called the transition matrix. This transition matrix is a stochastic matrix.
Chapman and Kolmogorov independently showed that if $P_n = [p_{ij}]_{n \times n}$ is a 1-step stochastic matrix, then the k-step stochastic matrix, is equal to the kth power of the 1-step stochastic matrix [10, 11, 12]. Hence, if $p_{ij}^{(k)}$ denotes the k-step transition probability from state i to state j and ℓ is the state space of a Markov chain, then we have

$$p_{ij}^{(k)} = Prob\{X_{t+1} = j | X_t = i\} = \sum_{z \in \ell} p_{iz}^{(r)} p_{zj}^{(k-r)}, \qquad 0 < r < k. \tag{2.5}$$

For some Markov chains, it is possible that there is a finite step such as kth step such that $p_{ij}^{(k)} > 0$ for every i, j from state space. This property is defined as:

Definition 2.8. *[4] A Markov chain is called a regular chain if some power of the transition matrix has only positive elements (in other words, the transition matrix of chain be regular).*

Regularity is an important property for Markov chains since it has a strong relationship with another important *ergodicity* property [4]. We know that a Markov chain is ergodic if it is possible to go from every state to every state (not necessarily in one move). By [6, Theorem 1.8] we know if a Markov transition matrix P is regular, then it has exactly one ergodic class and in general this process is ergodic. Hence, the following proposition shows the relationship between the regularity property and ergodicity one:

Proposition 2.9. *[4, 6] Every regular Markov chain is ergodic.*

Ergodic Markov chains are very important because there is a unique stationary distribution vector for their states. This means that if P is the transition matrix of a ergodic Markov chain, then there is a fixed point vector such as $\pi = \begin{Bmatrix} \pi_1 & \pi_2 & \pi_3 & \ldots & \pi_n \end{Bmatrix}$ such that $\pi P = \pi$. To compute the stationary distribution of an ergodic Markov chain, we have:

Proposition 2.10. *[10] Let $p_{ij}^{(k)}$ be the k-step transition probability from state i to state j and $\lim_{k\to\infty} p_{ij}^{(k)}$ exists. If π_j denotes the stationary distribution and the chain is ergodic, then $\pi_j = \lim_{k\to\infty} p_{ij}^{(k)}$.*

Since every regular chain is ergodic, if chain be regular, the stationary distribution exists.

2.3 The Helmert matrix

Following [3, page 67], the Helmert matrix of order n is a square matrix that is defined as:

$$H_n = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-1)}{\sqrt{n(n-1)}} \end{bmatrix}_{n\times n} \tag{2.6}$$

Moreover, the first row of the Helmert matrix of order n, has the following form

$$[\underbrace{\frac{1}{\sqrt{n}} \quad \frac{1}{\sqrt{n}} \quad \frac{1}{\sqrt{n}} \quad \cdots \quad \frac{1}{\sqrt{n}}}_{n \quad items}] \tag{2.7}$$

And the other i-th rows $2 \leq i \leq n$ are formed by

$$[\underbrace{\frac{1}{\sqrt{i(i-1)}} \quad \frac{1}{\sqrt{i(i-1)}} \quad \cdots \quad \frac{1}{\sqrt{i(i-1)}}}_{i-1 \quad items} \quad \frac{-(i-1)}{\sqrt{i(i-1)}} \quad \underbrace{0 \quad \cdots \quad 0}_{n-i \quad items}] \tag{2.8}$$

Furthermore, we know that the Helmert matrix is orthogonal [2]:

$$H_n H_n^T = H_n^T H_n = I_n. \tag{2.9}$$

Usually, the Helmert matrix is used in mathematical statistics for the analysis of variance (ANOVA), see [2, 3, 13].

3 Main results

To prove the main theorem, we need the following lemmas:

Lemma 3.1. *Let* $H_n = [h_{ij}]_{n\times n}$ *be the Helmert matrix of order* n*. Then* $P_n = [h_{ij}^2]_{n\times n}$*, is a doubly stochastic matrix.*

Proof. Using (2.6), we have

$$P_n = [h_{ij}^2]_{n\times n} = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} & 0 & \cdots & 0 & 0 \\ \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{3}{4} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{n(n-1)} & \frac{1}{n(n-1)} & \frac{1}{n(n-1)} & \frac{1}{n(n-1)} & \cdots & \frac{1}{n(n-1)} & \frac{(n-1)}{n} \end{bmatrix}_{n\times n} \tag{3.1}$$

Since the Helmert matrix is orthogonal (see [2]), by definition (2.2) all the rows and columns of it are unit vectors and, by definition (2.1), we have

$$h_{i1}^2 + h_{i2}^2 + ... + h_{in}^2 = 1, \qquad \forall i \in \{1, 2, ..., n\} \tag{3.2}$$

And

$$h_{1j}^2 + h_{2j}^2 + ... + h_{nj}^2 = 1, \qquad \forall j \in \{1, 2, ..., n\} \tag{3.3}$$

By (3.2) and (3.3), each row and column of (3.1) sum to 1 and $h_{ij}^2 \geq 0$ for every $1 \leq i, j \leq n$. Therefore, by definition (2.5), the above matrix P_n is a doubly stochastic matrix. □

With $\{X_t : t = 0, 1, 2, ..., n-1\}$ denoting the finite Markov chain by doubly stochastic matrix P_n in lemma (3.1), the probability of transition $i \longrightarrow j$ is equal to the (i, j)th element of the doubly stochastic matrix P_n and using (3.1) this transition probability is defined as follows:

$$p_{ij} = Prob\{X_{t+1} = j | X_t = i\} = \begin{cases} \frac{1}{n} & \text{if } i = 0, j = 0, 1, 2, ..., n-1 \\ \frac{1}{i(i-1)} & \text{if } 1 \leq i \leq n-1, i > j \\ \frac{i-1}{i} & \text{if } 1 \leq i \leq n-1, i = j \\ 0 & \text{if } 1 \leq i \leq n-1, i < j \end{cases} \tag{3.4}$$

Now, we want to show that the stochastic matrix P_n defined in lemma (3.1) is a regular and ergodic stochastic matrix. First, consider the following lemma:

Lemma 3.2. *Let $A_n = [a_{ij}]_{n\times n}$ be an $n \times n$ matrix of nonnegative numbers. If all elements in the first row and first column of matrix A_n are positive, then all elements in A_n^2 are positive ($A_n^2 > 0$).*

Proof. We have

$$A_n = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}_{n\times n}, \tag{3.5}$$

where $a_{ij} \geq 0$ and $a_{i1}, a_{1j} > 0$ for every $1 \leq i, j \leq n$. Hence, for A_n^2, we have

$$A_n^2 = A_n \times A_n = \begin{bmatrix} \sum_{k=1}^n a_{1k}a_{k1} & \sum_{k=1}^n a_{1k}a_{k2} & \dots & \sum_{k=1}^n a_{1k}a_{kn} \\ \sum_{k=1}^n a_{2k}a_{k1} & \sum_{k=1}^n a_{2k}a_{k2} & \dots & \sum_{k=1}^n a_{2k}a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{nk}a_{k1} & \sum_{k=1}^n a_{nk}a_{k2} & \dots & \sum_{k=1}^n a_{nk}a_{kn} \end{bmatrix}_{n\times n} \tag{3.6}$$

We know that within the matrix A_n^2, the sum $\sum_{k=1}^n a_{ik}a_{kj}$ is positive for every $1 \leq i, j \leq n$. Since all a_{ij} are nonnegative and every sum like $\sum_{k=1}^n a_{ik}a_{kj}$ is contained in the statement of $a_{i1} \times a_{1j}$ for every $1 \leq i, j \leq n$, then, according to assume of lemma, all a_{i1} and a_{1j} are positive. □

Corollary 3.3. *The doubly stochastic matrix P_n defined in lemma (3.1), is a semi-positive matrix.*

Proof. By (3.1), we know that all elements of the first row and first column of P_n are positive. So, by lemma(3.2), $P_n^2 > 0$, and by definition(2.3), P_n is a semi-positive or a regular matrix. □

Corollary 3.4. *The Markov chain in the doubly stochastic matrix P_n, defined in lemma (3.1), is a ergodic chain.*

Proof. By proposition (2.9), every regular chain is ergodic. On the other hand, by corollary(3.3), the stochastic matrix P_n is regular and hence its chain is regular. So, the chain is ergodic. □

In 2003, S. Baik and K. Bang proved that when a stochastic matrix is doubly and regular, the fixed point vector (or stationary distribution) always is equal to a unique vector [1].

Lemma 3.5. *[1, Theorem 2.1] If A_n be a semi-positive (regular) doubly stochastic matrix of order n, then $\lim_{m\to\infty} A_n^m = \frac{1}{n}J_n$.*

Theorem 3.6. *Given the doubly stochastic matrix P_n from lemma (3.1), let $\{X_t : t = 0, 1, 2, ..., n-1\}$ be the Markov chain by P_n. If π_j denotes the stationary distribution (or fixed point value) for jth state ($0 \le j \le n-1$), then $\pi_j = \frac{1}{n}$ for every $j = \{0, 1, 2, ..., n-1\}$.*

Proof. Consider the doubly stochastic matrix P_n in equation (3.1). Clearly, all elements in the first row and first column of P_n are positive. Hence, by lemma (3.2), $P_n^2 > 0$ and, by definition (2.3), P_n is semi-positive (regular). Also, by proposition (2.9), the Markov chain by stochastic matrix P_n must be ergodic. Therefore, since that P_n is a semi-positive and doubly stochastic matrix, by lemma (3.5), we have

$$\lim_{k\to\infty} P_n^k = \frac{1}{n}J_n = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}_{n\times n} \tag{3.7}$$

By proposition (2.10), if chain be ergodic and $\lim_{k\to\infty} p_{ij}^{(k)}$ exist, then $\pi_j = \lim_{k\to\infty} p_{ij}^{(k)}$. According to (3.7), we have $\pi_j = \frac{1}{n}$. □

References

[1] S. Baik, K. Bang, Limit theorem of the doubly stochastic matrices, Kangweon-Kyungki Math. J., 11 (2003), no. 2, 155-160.

[2] B. R. Clarke, Linear Models: The Theory and Application of Analysis of variance, Wiley, 2008.

[3] J. E. Gentie, Numerical Linear Algebra for Application in Statistics, Springer, 1998.

[4] Ch. M. Grinstead, J. L. Snell, Introduction to Probability, American Mathematical Society, 1997.

[5] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press, 2nd ed, 2013.

[6] A. Kehagias, Approximation of stochastic processes by hidden Markov models, U.M.I, 1992.

[7] H. O. Lancaster, The Helmert Matrices, The American Mathematical Monthly, 72(1965), no. 1, 4-12.

[8] E. Parzen, Stochastic Processes, Holden-Day, Inc., 1962.

[9] P. Perkins, A theorem on regular matrices, Pacific J. Math., 11 (1961), no. 4, 1529-1533.

[10] S. M. Roos, A first cours in probability, 6th ed, Pearson Prentice Hall, 2002.

[11] S. M. Ross, Introduction to Probability Models, 7th ed., Academic Press, Inc., 2000.

[12] S. M. Ross, Stochastic Processes, 2nd ed. John Wiley and Sons, Inc.,1996.

[13] G. A. F. Seber, A Matrix Handbook for Statistician, Wiley, 2007.

[14] G. Strang, Introduction to Linear Algebra, 5th ed, Wellesly-cambridge Press, 2016.