A Minimal Completion of Doubly Stochastic Matrix

Selcuk Koyuncu, Lei Cao and Timmothy Parmer (speaker)

This paper is submitted to the Journal of Linear and Multilinear Algebra

University of North Georgia

2015

Table of contents

Basic Definitions

The term stochastic matrix goes back at least to Romanovsky (1931,p.267). In 1935, Romanosvsky provided a detail discussion of stochastic matrices such as the use of it in the theory of discrete markov chains. Doubly stochastics matrices are also called Schur transformations, or they are said to be bistochastic. The term doubly stochastic apperas to have been first introduced by Feller. Stochastics matrices arise in the study of Markov chains and in the variety of modeling problems such as economics and operation research.

Basic Definitions

Definition

An $n \times n$ matrix $A = (a_{ij})$ is called doubly stochastic if

$$a_{ij} \geq 0$$

for all i, j,

$$\sum_{i=1}^{n} a_{ij} = 1$$

for all j, and

$$\sum_{j=1}^{n} a_{ij} = 1$$

for all i.

Basic Definitions

Definition

A matrix $B = (b_{ij})$ is called doubly substochastic if

$$b_{ij} \geq 0$$

for all i, j,

$$\sum_{i=1}^{n} b_{ij} \leq 1$$

for all j, and

$$\sum_{j=1}^{n} b_{ij} \leq 1$$

for all i.

Basic Definitions

Definition

A matrix $P \in M_n(\mathbb{R})$ is said to be a partial permutation matrix if it has at most one nonzero entry in each row and column, and these nonzero entries are all one.

Birkhoff's Theorem

Theorem (Birkhoff's Theorem)

A matrix $A \in M_n(\mathbb{R})$ *is a doubly stochastic matrix if and only if for some* $N < \infty$ *there are permutation matrices* $P_1, \cdots, P_2 \in M_n(\mathbb{R})$ *and positive scalars* $\alpha_1, \cdots, \alpha_N \in \mathbb{R}$ *such that* $\alpha_1 + \cdots + \alpha_N = 1$ *and* $A = \alpha_1 P_1 + \cdots + \alpha_N P_N$.

Graphs and Matching

Definition

A **bipartite graph** $G = (U, V, E)$ is a graph whose vertices can be divided into two disjoint sets such that no edge connects two vertices of the same set.

Definition

A **perfect matching** is a subset of the edge set E such that every vertex has exactly one edge incident on it, i.e, $|U| = |V| = n$

Perfect Matching

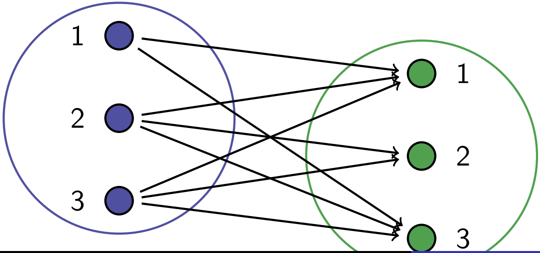
The proof of Birkhoff's theorem uses the following key Lemma.

Lemma

The associated graph of any doubly stochastic matrix has a perfect matching.

Perfect Matching

We represent each row and each column in our doubly stochastic matrix with a vertex and we connect the vertex representing row i with the vertex representing row j if the entry x_{ij} in the matrix is not zero. For instance, let $A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{4} & \frac{5}{12} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{12} \end{bmatrix}$ be a doubly stochastic matrix. Then the graph associated to matrix A is given in the picture below.



Proof of Birkhoff's Theorem

We proceed by induction on the number of nonzero entries in the matrix. Let M_0 be a doubly stochastic matrix. By the previous lemma, the associated graph has a perfect matching. For example, in the associated graph previously $(1,3),(2,1),(3,2)$ is a perfect matching and we underline x_{13}, x_{21} and x_{32}. Thus we underline exactly one element in each row and each column. Let α_0 be the minimum of the underlined entries. Let P_0 be the permutation matrix that has a 1 exactly at the position of the underlined elements. If $\alpha_0 = 1$ then all underlined entries are 1, and $M_0 = P_0$ is a permutation matrix.

Proof of Birkhoff's Theorem

If $\alpha_0 < 1$ then the matrix $M_0 - \alpha_0 P_0$ has non-negative entries, and the sum of the entries in any row or any column is $(1 - \alpha_0)$. Dividing each entry by $(1 - \alpha_0)$ in $M_0 - \alpha_0 P_0$ gives a doubly stochastic matrix M_1. Thus we may write $M_0 = \alpha_0 P_0 + (1 - \alpha_0) M_1$ where M_1 is not only doubly stochastic, but has less non-zero entries than M_0.

Proof of Birkhoff's Theorem

By our induction hypothesis M_1 may be written as $M_1 = \alpha_1 P_1 + \quad + \alpha_n P_n$ where $P_1, \cdots, P_n$ are permutation matrices, and $\alpha_1 P_1 + \cdots + \alpha_n P_n$ is a convex combination. But then we have

$$M_0 = \alpha_0 P_0 + (1 - \alpha_0)\alpha_1 P_1 + \cdots + (1 - \alpha_0)\alpha_n P_n$$

where $P_0, P_1, \cdots, P_n$ are permuation matrices,

Proof of Birkhoff's Theorem

Thus we have a convex combination since, $\alpha_0 \geq 0$, each $(1-\alpha_0)\alpha_i$ is nonnegative and we have $\alpha_0 + (1-\alpha_0)\alpha_1 + \cdots + (1-\alpha_0)\alpha_n = \alpha_0 + (1-\alpha_0)(\alpha_1 + \cdots + \alpha_n) = \alpha_0 + (1-\alpha_0) = 1$.

Example of Birkhoff's Theorem

Example

$A = \begin{bmatrix} \frac{7}{12} & 0 & \frac{5}{12} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$ is a doubly stochastic matrix. Let

$P_0 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $\alpha_0 = \frac{1}{6}$. Thus we get $M_1 = \frac{1}{1-\frac{1}{6}}(M_0 - \frac{1}{6}P_0)$.

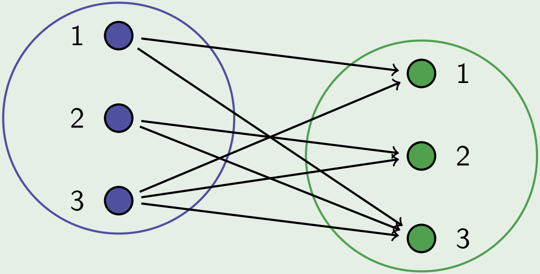
Then

$$M_1 = \begin{bmatrix} 7/10 & 0 & 3/10 \\ 0 & 3/5 & 2/5 \\ 3/10 & 2/5 & 3/10 \end{bmatrix}.$$

Example of Birkhoff's Theorem

Example

The graph associated to M_1 is the following.



Example of Birkhoff's Theorem

Example

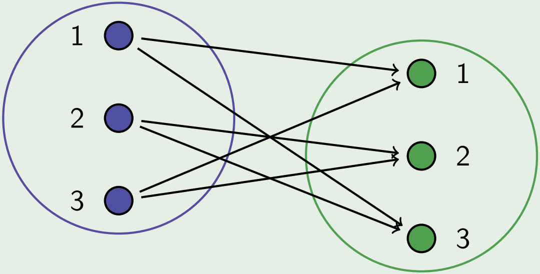
A perfect matching corresponding to the graph of M_1 is $\{(1,1),(2,2),(3,3)\}$, the associated matrix is $P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and we have $\alpha_1 = \frac{3}{10}$. Thus we get $M_2 = \frac{1}{1-\frac{3}{10}}(M_1 - \frac{3}{10}P_1)$. Thus

$$M_2 = \begin{bmatrix} 4/7 & 0 & 3/7 \\ 0 & 3/7 & 4/7 \\ 3/7 & 4/7 & 0 \end{bmatrix}.$$

Example of Birkhoff's Theorem

Example

The graph associated to M_2 is the following.



Example of Birkhoff's Theorem

Example

A perfect matching in this graph is $\{(1,3),(2,2),(3,1)\}$, the associated permutation matrix is $P_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, and we have $\alpha_2 = \frac{3}{7}$. Thus we get $M_3 = \frac{1}{1-\frac{3}{7}}(M_2 - \frac{3}{7}P_2)$. This gives us that

$$M_3 = P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Example of Birkhoff's Theorem

Example

Since $M_2 = \alpha_2 P_2 + (1-\alpha_2)M_3 = \frac{3}{7}P_2 + \frac{4}{7}P_3$ and

$M_1 = \alpha_1 P_1 + (1-\alpha_1)M_2 = \frac{3}{10}P_1 + \frac{7}{10}M_2 = \frac{3}{10}P_1 + \frac{3}{10}P_2 + \frac{4}{10}P_3$
then

$$M_0 = \alpha_0 P_0 + (1-\alpha_0)M_1 = \frac{1}{6}P_0 + \frac{1}{4}P_1 + \frac{1}{4}P_2 + \frac{1}{3}P_3.$$

thus by Birkhoff's Theorem a doubly stochastic matrix M_0 can be written as a convex combination of permutation matrices.

Main Result

Theorem

Let $B = (b_{ij})_{i,j=1}^{n}$ *be an* $n \times n$ *doubly substochastic matrix. Denote the sum of all elements of* B *by* s, *i.e,*

$$s = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}.$$

Then the **sub-defect** *of* B *is* $k = \lceil n - s \rceil$, *where* $\lceil x \rceil$ *is the ceiling of* x, *that is the smallest integer greater than or equal to* x.

Main Result

Then the following is true:

(a) there exits an $(n+k) \times (n+k)$ doubly stochastic matrix containing B as a submatrix;

(b) there does not exist an $(n+l) \times (n+l)$ doubly stochastic matrix containing B as a submatrix with $l < k$.

Proof of the main result

Proof.

Let r_i and c_i be the sum of the ith row and ith column respectively, namely

$$r_i = \sum_{j=1}^{n} b_{ij} \ , \qquad c_i = \sum_{j=1}^{n} b_{ji}$$

for $i = 1, 2, \ldots, n$. Thus since B is substochastic, then we have

$$0 \leq r_i \leq 1 \ , \quad 0 \leq c_i \leq 1. \tag{1}$$

□

Proof of the main result

Proof.

Induction on sub-defect k.

Base case: Let

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

be an $n \times n$ doubly substochastic matrix with sub-defect $k = 1$. □

Proof of the main result

Proof.

Then

$$\begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} & 1-r_1 \\ b_{21} & b_{22} & \ldots & b_{2n} & 1-r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} & 1-r_n \\ 1-c_1 & 1-c_2 & \ldots & 1-c_n & 1-(n-s) \end{pmatrix}$$

is an $(n+1) \times (n+1)$ doubly stochastic matrix containing B as the $n \times n$ principal submatrix.

□

Proof of the main result

Proof.

On one hand, according to (1)

$$0 \leq 1 - r_i \leq 1, \quad 0 \leq 1 - c_i \leq 1$$

for $i = 1, 2, \ldots, n$, and $0 \leq n - s \leq 1$ due to the assumption $k = \lceil n - s \rceil = 1$. Hence all elements are between 0 and 1. □

Proof of the main result

Proof.

On the other hand,

$$\sum_{i=1}^{n}(1-r_i) = n - \sum_{i=1}^{n} r_i = n - s$$

and

$$\sum_{i=1}^{n}(1-c_i) = n - \sum_{i=1}^{n} c_i = n - s.$$

Hence, each row and column has sum 1. □

Proof of the main result

Proof.

Induction assumption: Suppose $\forall n$ and $k = \lceil n - s \rceil = t \geq 2$, there exists an $(n+t) \times (n+t)$ doubly stochastic matrix such that B is the $n \times n$ principal submatrix. Inductive step: We need to show

$\forall n$, if the sub-defect $k = t + 1$, then there exists an $(n+t+1) \times (n+t+1)$ doubly stochastic matrix such that B is the $n \times n$ principal submatrix. □

Proof of the main result

Proof.

First build a matrix $\tilde{B}$ as the following

$$\tilde{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} & x_1 \\ b_{21} & b_{22} & \dots & b_{2n} & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} & x_n \\ y_1 & y_2 & \dots & y_n & 0 \end{pmatrix}$$

□

Proof of the main result

Proof.

*Note x_i, y_i are

$$0 \leq x_i \leq 1 - r_i, \quad 0 \leq y_i \leq 1 - c_i$$

and

$$\sum_{i=1}^{n} x_i = 1, \sum_{i=1}^{n} y_i = 1$$

for $i = 1, 2, \ldots, n$. □

Proof of the main result

Proof.

The existence of such x_i and y_i is implied by $k - 1 < \sum_{i=1}^{n}(1 - r_i) = \sum_{i=1}^{n}(1 - c_i) = n - s \leq k$. One such choice is

$$x_1 = 1 - r_1, \quad y_1 = 1 - c_1$$

and

$$x_i = \min\{1 - r_i, 1 - \sum_{j=1}^{i-1} x_j\}, \quad y_i = \min\{1 - c_i, 1 - \sum_{j=1}^{i-1} y_j\}$$

for $i = 2, 3, \ldots, n$. □

Proof of the main result

Proof.

Note that $\tilde{B}$ is an $(n+1) \times (n+1)$ doubly substochastic matrix with the sum of all elements $s+2$ and hence the sub-defect $k = \lceil (n+1) - (s+2) \rceil = \lceil n - s - 1 \rceil = t$. Due to the assumption, there exists an $[(n+1)+t] \times [(n+1)+t)]$ doubly stochastic matrix containing $\tilde{B}$ as the $(n+1) \times (n+1)$ principal submatrix meaning that B is the $n \times n$ principle submatrix. □

Future Research

Lemma

An $n \times n$ doubly substochastic matrix B can be dilated to a $2n \times 2n$ doubly stochastic matrix.

Future Research

Proof.

Let $B \in M_n(\mathbb{R})$ be doubly substochastic, let $e = [1, \cdots, 1]^T \in \mathbb{R}^n$, and let $D_r = \mathrm{diag}(Be)$ and $D_c = \mathrm{diag}(B^T e)$ be diagonal matrices containing the row and the column sums of B. The doubly stochastic $2n \times 2n$ matrix

$$\begin{bmatrix} B & I - D_r \\ I - D_c & B^T \end{bmatrix}$$

is a dilation of B. Conversely, it is evident that any square submatrix of a doubly stochastic matrix is doubly substochastic. □

Future Research

Let B be a given $n \times n$ doubly substochastic matrix.

1. Complete B to A by using pervious lemma.
2. Use Birkhoff's Theorem and write $A = c_1 P_1 + c_2 P_2 + \ldots + c_m P_m$ as convex combination of permutation matrices, where A and $P_1, \ldots, P_m$ are $2n \times 2n$ matrices, $\sum_{i=1}^{N} c_i = 1$ and $c_i \geq 0, N < \infty$. Next, write

   $$B = c_1 Q_1 + \ldots + c_m Q_m$$

   where Q_i is the $n \times n$ principal submatrix of P_i for $i = 1, \ldots, m$, and Q_i are partial permutation matrices.

Future Research

1. Complete each partial permutation matrix Q_i to a permutation matrix $\tilde{Q}_i$.
2. The minimum size of a completion of B is the difference between the size of given doubly substochastic matrix and the number of minimum $1's$ that Q_i. contains.

Future Research

Example

Let

$$B = \begin{pmatrix} 0.2010 & 0.1360 & 0.0030 \\ 0.2210 & 0.5420 & 0.2340 \\ 0.3440 & 0.3190 & 0.2010 \end{pmatrix}$$

be the 3×3 principal submatrix of the doubly stochastic matrix.
In order to find minimal completion of B, we follow the procedure given previously.
Use previous lemma to complete B to A as it is shown on the next slide.

Future Research

Example

$$A = \begin{pmatrix} 0.2010 & 0.1360 & 0.0030 & 0.6600 & 0 & 0 \\ 0.2210 & 0.5420 & 0.2340 & 0 & 0.0030 & 0 \\ 0.3440 & 0.3190 & 0.2010 & 0 & 0 & 0.1360 \\ 0.2340 & 0 & 0 & 0.2010 & 0.2210 & 0.3440 \\ 0 & 0.0030 & 0 & 0.1360 & 0.5420 & 0.3190 \\ 0 & 0 & 0.5620 & 0.0030 & 0.2340 & 0.2010 \end{pmatrix}$$

Using Birkoff's Theorem, we write A as a convex combination of partial permutation matrices.

Future Research

Example

$$A =$$

$$0.2210 \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} + 0.2010 \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} +$$

$$0.1360 \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} + 0.0030 \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Future Research

Example

$$+0.3410\begin{pmatrix}0&0&0&1&0&0\\0&1&0&0&0&0\\1&0&0&0&0&0\\0&0&0&0&0&1\\0&0&0&0&1&0\\0&0&1&0&0&0\end{pmatrix}+0.0980\begin{pmatrix}0&0&0&1&0&0\\0&0&1&0&0&0\\0&1&0&0&0&0\\1&0&0&0&0&0\\0&0&0&0&0&1\\0&0&0&0&1&0\end{pmatrix}$$

Future Research

Next, we extract B, the 3x3 principle submatrix

Example

$$B =$$
$$0.2210 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + 0.2010 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 0.1360 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} +$$
$$0.0030 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + 0.3410 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} + 0.0980 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

We then complete each 3x3 partial permutation matrix by adding one row and one column

Future Research

Example

$$\tilde{B} = 0.2210 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} + 0.2010 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} +$$

$$0.1360 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} + 0.0030 \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} +$$

$$0.3410 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} + 0.0980 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Future Research

Thus we have $\tilde{B}$, which is the doubly stochastic completion of B, with B as 3x3 principle submatrix.

Example

$$\tilde{B} = \begin{pmatrix} 0.2010 & 0.1360 & 0.0030 & 0.6600 \\ 0.2210 & 0.5420 & 0.2340 & 0.0030 \\ 0.3440 & 0.3190 & 0.2010 & 0.1360 \\ 0.2340 & 0.0030 & 0.5620 & 0.2010 \end{pmatrix}$$

Thank you