Birkhoff's Theorem

Definition A square matrix is *doubly stochastic* if all its entries are non-negative and the sum of the entries in any of its rows or columns is 1.

Example The matrix

$$\begin{pmatrix} 7/12 & 0 & 5/12 \\ 1/6 & 1/2 & 1/3 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}$$

is doubly stochastic.

A special example of a doubly stochastic matrix is a *permutation matrix*.

Definition A permutation matrix is a square matrix whose entries are all either 0 or 1, and which contains exactly one 1 entry in each row and each column.

Example The matrix

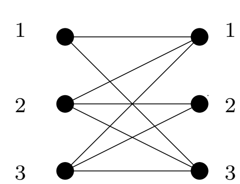
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

is a permutation matrix.

Recall that a *convex combination* of the vectors $v_1, \ldots, v_n$ is a linear combination $\alpha_1 v_1 + \cdots + \alpha_n v_n$ such that each α_i is non-negative and $\alpha_1 + \cdots + \alpha_n = 1$. (Necessarily, each α_i is at most 1.)

Theorem(Birkhoff) *Every doubly stochastic matrix is a convex combination of permutation matrices.*

The proof of Birkhoff's theorem uses Hall's marriage theorem. We associate to our doubly stochastic matrix a bipartite graph as follows. We represent each row and each column 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. The graph associated to our example is given in the picture below.



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

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

Proof: Assume, by way of contradiction that the graph has no perfect matching. Then, by Hall's theorem, there is a subset A of the vertices in one part such that the set $R(A)$ of all vertices connected to some vertex in A has strictly less than $|A|$ elements. Without loss of generality we may assume that A is a set of vertices representing rows, the set $R(A)$ consists then of vertices representing columns. Consider now the sum $\sum_{i \in A, j \in R(A)} x_{ij}$, i.e., the sum of all entries located in a row belonging to A and in a column in $R(A)$. In the rows belonging to A all nonzero entries are located in columns belonging to $R(A)$ (by the definition of the associated graph). Thus

$$\sum_{i \in A, j \in R(A)} x_{ij} = |A|$$

since the graph is doubly stochastic and the sum of elements located in any of given $|A|$ rows is $|A|$. On the other hand, the sum of all elements located in all columns belonging to $R(A)$ is at least $\sum_{i\in A, j\in R(A)} x_{ij}$ since the entries not belonging to a row in A are non-negative. Since the matrix is doubly stochastic, the the sum of all elements located in all columns belonging to $R(A)$ is also exactly $|R(A)|$. Thus we obtain

$$\sum_{i\in A, j\in R(A)} x_{ij} \leq |R(A)| < |A| = \sum_{i\in A, j\in R(A)} x_{ij},$$

a contradiction.

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 key lemma, the associated graph has a perfect matching. Underline the entries associated to the edges in the matching. For example in the associated graph above $(1,3),(2,1),(3,2)$ is a perfect matching so 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. 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. By our induction hypothesis M_1 may be written as $M_1 = \alpha_1 P_1 + \cdots + \alpha_n P_n$ 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 permutation matrices, and we have a convex combination since, $\alpha_0 \geq 0$, each $(1-\alpha_0)\alpha_i$ is non-negative and we have

$$\alpha_0 + (1-\alpha_0)\alpha_1 + \cdots (1-\alpha_0)\alpha_n = \alpha_0 + (1-\alpha_0)(\alpha_1 + \cdots \alpha_n) = \alpha_0 + (1-\alpha_0) = 1.$$

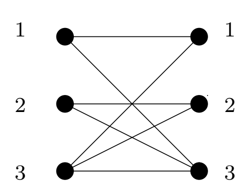
In our example

$$P_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and $\alpha_0 = 1/6$. Thus we get

$$M_1 = \frac{1}{1-1/6}\left(M_0 - \frac{1}{6}P_0\right) = \frac{6}{5}\begin{pmatrix} 7/12 & 0 & 1/4 \\ 0 & 1/2 & 1/3 \\ 1/4 & 1/3 & 1/4 \end{pmatrix} = \begin{pmatrix} 7/10 & 0 & 3/10 \\ 0 & 3/5 & 2/5 \\ 3/10 & 2/5 & 3/10 \end{pmatrix}.$$

The graph associated to M_1 is the following.



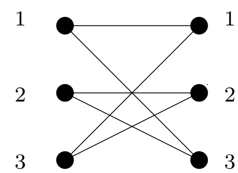
A perfect matching is $\{(1,1),(2,2),(3,3)\}$, the associated permutation matrix is

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and we have $\alpha_1 = 3/10$. Thus we get

$$M_2 = \frac{1}{1-3/10}\left(M_1 - \frac{3}{10}P_1\right) = \frac{10}{7}\begin{pmatrix} 4/10 & 0 & 3/10 \\ 0 & 3/10 & 2/5 \\ 3/10 & 2/5 & 0 \end{pmatrix} = \begin{pmatrix} 4/7 & 0 & 3/7 \\ 0 & 3/7 & 4/7 \\ 3/7 & 4/7 & 0 \end{pmatrix}$$

The graph associated to M_2 is the following.



A perfect matching in this graph is $\{(1,3),(2,2),(3,1)\}$, the associated permutation matrix is

$$P_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

and we have $\alpha_2 = 3/7$. Thus we get

$$M_3 = \frac{1}{1-3/7}\left(M_2 - \frac{3}{7}P_2\right) = \frac{7}{4}\begin{pmatrix} 4/7 & 0 & 0 \\ 0 & 0 & 4/7 \\ 0 & 4/7 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

We are done since $M_3 = P_3$ is a permutation matrix. Working our way backwards we get

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

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

and

$$M_0 = \alpha_0 P_0 + (1-\alpha_0)M_1 = \frac{1}{6}P_0 + \frac{5}{6}\left(\frac{3}{10}P_1 + \frac{3}{10}P_2 + \frac{4}{10}P_3\right) = \frac{1}{6}P_0 + \frac{1}{4}P_1 + \frac{1}{4}P_2 + \frac{1}{3}P_3.$$

We obtained that

$$\begin{pmatrix} 7/12 & 0 & 5/12 \\ 1/6 & 1/2 & 1/3 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$