A SHORT PROOF OF THE BIRKHOFF-VON NEUMANN THEOREM

GLENN HURLBERT*

Abstract. The Birkhoff-von Neumann Theorem has been proved many times in the literature with a number of different methods, some inductive, some constructive, some existential. We offer a new proof that is a little more direct than most, though nonconstructive.

Key words. Doubly stochastic matrices, Convex combinations, Permutation matrices.

AMS subject classifications. 15A51, 52A20.

1. Introduction. A vector is *stochastic* if it is nonnegative and its components sum to 1. A matrix is *doubly stochastic* (DS) if each of its rows and columns is stochastic. A *permutation matrix* is a square $\{0, 1\}$-matrix with exactly one 1 per row and per column. The identity matrix is an example of a permutation matrix; indeed, every permutation matrix is a rearrangement of the columns (or rows) of an identity matrix. A permutation matrix is one type of doubly stochastic matrix; in fact, every integral doubly stochastic matrix is a permutation matrix. It is elementary that every convex combination of permutation matrices is DS. The converse is a 1936 theorem of Kőnig [7] (Chapter XIV, Section 3, in the context of generalizing the factorization of regular bipartite graphs), typically attributed instead to the 1946 and 1953 work of Birkhoff [2] and von Neumann [8], respectively.

Theorem 1.1. *Every DS matrix is a convex combination of permutation matrices.*

The traditional proof uses induction by removing an appropriate fraction of a permutation matrix P from the given DS matrix, and various methods have been found to find such a P, including von Neumann's iterated scheme (similar to our method below) as well as linear optimization (see [3, 4, 6]) — essentially an application of the integrality theorem for networks. Edmonds' proof (given on p. 331 of [3]) applies network theory more directly, instead of to the permutation matrix lemma. Another interesting proof is found in [9], reminscent of the Frobenius-Kőnig theorem (see [1], p.62) characterising 0-permanent matrices. The proof in [5] uses induction directly to prove Theorem 1.1. Our proof is also direct, avoiding the permutation matrix lemma; however it is consequently nonconstructive. The motivation for this

* Department of Mathematics and Statistics, Arizona State University, Tempe, Arizona 85287-1804, USA (hurlbert@asu.edu).

approach comes from presenting the material to undergraduates in the context of [6].

Proof. The key idea, as many have pointed out, is to think of an $n \times n$ matrix as a vector in $\mathbb{R}^{n^2}$. The strategy is to use the DS property to impose linear constraints on such vectors. If the extreme points of the polytope defined by the constraints correspond to permutation matrices (the bulk of the work in the proof) then the result follows by Straszewicz's theorem [10] that every polytope is the convex hull of its extreme points.

We let $\mathbf{X} = (x_{r,s})$ be an $n \times n$ DS matrix. The constraints of the system on $\{x_{r,s}\}$ that defines the DS property are as follows.

$$\begin{aligned} &\textstyle\sum_{r=1}^{n} x_{r,s} = 1 \\ &\textstyle\sum_{s=1}^{n} x_{r,s} = 1 \\ \\ &x_{r,s} \geq 0 \qquad\qquad (1 \leq r \leq n)\ (1 \leq s \leq n) \end{aligned}$$

The polyhedron P defined above is a polytope since the linear constraints imply that each $0 \leq x_{r,s} \leq 1$, and so P is bounded.

We now proceed to show that every extreme point of P is integral, by contrapositive. We will show that any nonintegral point of P is the center of some line segment residing inside P.

Suppose that $\mathbf{x} \in P$ is not integral, and let $0 < x_{r_1,s_1} < 1$. Because of the row constraint $\sum_{s=1}^{n} x_{r_1,s} = 1$, there must be some s_2 such that $0 < x_{r_1,s_2} < 1$. Likewise, because of the column constraint $\sum_{r=1}^{n} x_{r,s_2} = 1$, there must be some r_2 such that $0 < x_{r_2,s_2} < 1$. This process can be iterated, and we will stop when some index (r, s) is repeated. Moreover, we will assume that we chose the iterated process having the shortest such sequence of indices. Then we know that the final index is the first repeated index, namely (r_1, s_1).

We claim that there is some k satisfying $(r_k, s_k) = (r_1, s_1)$; that is, the length of the sequence is even — otherwise a shorter sequence can be found. Suppose not, say $(r_k, s_{k+1}) = (r_1, s_1)$. Then, because (r_k, s_{k+1}), (r_1, s_1) and (r_1, s_2) are all in the same row, by deleting (r_1, s_2) and starting instead at (r_2, s_2) we obtain a valid sequence that is shorter, a contradiction.

Now let $\epsilon_0 = \min\{x_{r_j}, x_{1-r_j}, x_{s_j}, x_{1-s_j}\}_{j=1}^{k}$. Then for any $0 < \epsilon < \epsilon_0$ define $\mathbf{x}^+(\epsilon)$ (resp. $\mathbf{x}^-(\epsilon)$) by decreasing (resp. increasing) the value of each x_{r_j,s_j} by ϵ, while increasing (resp. decreasing) the value of each $x_{r_j,s_{j+1}}$ by ϵ. Note that $\mathbf{x}^+(\epsilon)$ (resp. $\mathbf{x}^-(\epsilon)) \in P$. Indeed, increasing x_{r_j,s_j} and decreasing x_{r_j,s_j} by the same amount ϵ maintains the sum of 1 in row r_j, while preventing both $x_{r_j,s_j} > 1$ and $x_{r_j,s_{j+1}} < 0$

because $\epsilon < \epsilon_0$. The same argument applies to column sum preservation. This shows that $\mathbf{x}^+(\epsilon) \in P$. The analogous argument shows that $\mathbf{x}^-(\epsilon) \in P$.

Thus we have shown that the line segment $\overline{\mathbf{x}^-(\epsilon)\mathbf{x}^+(\epsilon)}$ lies entirely in P and has $\mathbf{x}$ as its center. Therefore, $\mathbf{x}$ is not extreme. Hence, every extreme point of P is integral, and so corresponds to a permutation matrix. Thus every DS matrix is a convex combination of permutation matrices. □

REFERENCES

[1] R.B. Bapat and T.E.S. Raghavan, *Nonnegative Matrices and Applications*, Encyclopedia of Mathematics and its Applications 64, Cambridge University Press, Cambridge, 1997.

[2] G. Birkhoff, Three observations on linear algebra, *Univ. Nac. Tacumán Rev. Ser. A* 5:147–151, 1946.

[3] V. Chvátal, *Linear Programming*, W. H. Freeman, New York, 1983.

[4] G.B. Dantzig, Applications of the simplex method to a transportation problem, *Activity Analysis of Production and Allocation*, Cowles Commission Monograph 13:359–373, John Wiley & Sons, New York, 1951.

[5] A.J. Hoffman and H.W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.* 20:37–39, 1953.

[6] G. Hurlbert, *Linear Optimization: The Simplex Workbook*, Springer-Verlag, New York, in press.

[7] D. Kőnig, *Theorie der endlichen und unendlichen Graphen*, Akademische. Verlags Gesellschaft, Leipzig, 1936.

[8] J. von Neumann, A certain zero-sum two-person game equivalent to an optimal assignment problem, *Ann. Math. Studies* 28:5–12, 1953.

[9] J.V. Romanovsky, A simple proof of the Birkhoff-von Neumann theorem on bistochastic matrices, *A tribute to Ilya Bakelman*, Discourses Math. Appl. 3:51–53, Texas A & M Univ., College Station, TX, 1994.

[10] S. Straszewicz, Über exponierte Punkte abgeschlossener Punktmengen, *Fundam. Math.* 24:139–143, 1935.