BIRKHOFF'S PROBLEM WITH ALL PROOFS

QI ZHOU AND ZIXING WANG

Abstract. By giving the permutation representation on the finite dimensional space, we discovered a basis for generating all possible matrices. Then, we researched the space spanned by $n \times n$ permutation matrices and finally proved Birkhoff's Theorem.

1. Introduction

In this paper, we will make effort to prove **Birkhoff's Theorem** ([**?**]). Firstly, we introduce two concepts.

- An $n \times n$ matrix with nonnegative real entries is doubly stochastic if the sum of the entries along any of its rows or columns is equal to 1.
- A linear combination is called convex if the coefficients are nonnegative and their sum is equal to 1.

In 1946, Birkhoff showed the following result.

Theorem 1.1. *Every doubly-stochastic $n \times n$ matrix can be represented as a convex combination of at most $n^2 - 2n + 2$ permutation matrices. The number $n^2 - 2n + 2$ cannot be replaced by a smaller number.*

2. From Abstraction: The Permutation Representation on Finite-dim Space

The polyhedron Ω_n is obtained by taking the convex hull of the set of n-square permutation matrices, so how to describe permutation matrices is the first thing we need to do.

We cosider a linear space $V = \text{span}\{e_1, e_2, \cdots, e_n\} = \text{span}\{B\}$, then naturally an action(endomorphisms) $E(V)$ from S_n onto V can be defined, as well as the operation

$$(\phi + \chi)(x) = \phi(x) + \chi(x), \quad (\phi\chi)(x) = \phi\{\chi(x)\}, \quad \forall \phi, \chi \in E(V), x \in V \tag{2.1}$$

for which we call it permutation endomorphism. It is obvious that all permutation endomorphism form a multiplicative group, denoted by Π.

Here, we mainly consider those permutations with a single k-cycle $(e_{i_1} e_{i_2} \cdots e_{i_k}) = \gamma$, as well as the identity element θ. For different $\alpha \in \Pi$,

2010 *Mathematics Subject Classification.* Primary 05E15.
Key words and phrases. Birkhoff's Problem, doubly-stochastic, convex combination.

we denote those basis moving by α by $B(\alpha)$. For each pair $B(\alpha)$ and $B(\beta)$, one can verify that the following lemma holds.

Lemma 2.1. *Assume $\alpha, \beta \in \Pi$, with $B(\alpha) \cap B(\beta) = \emptyset$, then*

$$\alpha\beta = \alpha + \beta - \theta \tag{2.2}$$

Proof. We consider the following conditions for $x \in B$.

- $x \in B(\alpha)$. Note that $\alpha(x)$ still in its k-cycle, hence $\alpha(x) \in B(\alpha)$ and $\alpha(x)$ is invariant for permutation β. Then

$$\begin{aligned} (\alpha\beta)(x) &= \alpha(x) \\ (\alpha + \beta - \theta)(x) &= \alpha(x) + x - x = \alpha(x) \end{aligned} \tag{2.3}$$

- $x \in B(\beta)$. It is analogous with $x \in B(\alpha)$.
- $x \in (B(\alpha) \cup B(\beta))^c$. Then $\alpha(x) = \beta(x) = x$, which infers both sides are just x.

□

Now we can say a natural but not trivial conclusion, that is, the integral linear generator of all k-cycle.

Lemma 2.2. *Each k-cycle can be expressed as a linear combination of θ, 2-cycles and 3-cycles, with integral coefficients.*

Proof. For $k = 4$, WLOG, we set $\gamma = (e_1e_2e_3e_4)$, then the formula above makes the assertion correct.

$$(e_1e_2e_3e_4) = (e_1e_2e_3) + (e_1e_3e_4) - (e_1e_3) \tag{2.4}$$

Assume for $\leq k - 1$ the assertion holds, then we consider a k-cycle $\gamma = (e_1 \cdots e_k)$, then γ can be written by

$$\gamma = (e_1e_k)(e_1e_2 \cdots e_{k-1}) \tag{2.5}$$

By induction assumption, $(e_1e_2 \cdots e_{k-1})$ can be written by an integral linear combination of θ, 2-cycle and 3-cycle. If a cycle is disjoint with (e_1e_k), then use Lemma **??** to divide $\alpha\beta$ into $\alpha + \beta - \theta$. And for other cases, we have

$$\begin{aligned} &(e_1e_k)(e_1e_j) = (e_1e_je_k) \\ &(e_1e_k)(e_1e_k) = \theta \\ &(e_1e_k)(e_ke_j) = (e_1e_j) \\ &(e_1e_k)(e_ke_je_i) = (e_1e_ke_je_i) = (e_1e_ke_j) + (e_1e_je_i) - (e_1e_j) \\ &(e_1e_k)(e_1e_ke_j) = (e_ke_j) \end{aligned} \tag{2.6}$$

hence the induction is correct, and the lemma is proved. □

Actually, there are also a batch of redundant generators in $\leq$ 3-cycles. For a much more precise estimation, we have the following lemma.

Lemma 2.3. *Fix e be any element in B, and denote $\Gamma(e)$ be the subset of Π consisting of θ, 2-cycles and 3-cycles that moves e. Then $\Gamma(e)$ can integrally and linearly generate Π.*

Proof. It suffices to show that for each 2-cycle and 3-cycle can be expressed by $\Gamma(e)$. In fact, the following relations hold

$$(e_1e_2e_3) = (ee_1e_2) + (ee_2e_3) - (ee_1e_3) - (ee_2) + (e_1e_3) \tag{2.7}$$

then $\Gamma(e)$ can be a generator set for Π. □

It can still be further simplified, hence becoming truely a set of basis!

Theorem 2.4. *For $\Gamma(e)$ defined on Lemma **??**, let Γ be a subset of $\Gamma(e)$ by deleting one of (ee_1e_2) and (ee_2e_1) for each distinct pair (e_1, e_2). Then every element of Γ can be expressed as a linear combination with integral coefficients of the elements of Γ uniquely.*

Proof. Note that

$$(ee_1e_2) + (ee_2e_1) = (ee_1) + (ee_2) + (e_1e_2) - \theta \tag{2.8}$$

we can deduce that Π can be integrally and linearly generated by Γ. It remains to show that elements in Γ are linearly independent.

Denote binary relation $i\Gamma j$ to indicate that $(eij) \in \Gamma$. So if

$$p\theta + \sum_{i\neq e} q_i(ei) + \sum_{j\Gamma i} r_{ij}(ij) + \sum_{k\Gamma l} s_{kl}(ekl) = 0 \tag{2.9}$$

Let $y\Gamma x$, then use both sides act on x, by linearly independence of B, we can derive that $r_{xy} = 0$ whenever $y\Gamma x$. Then let $x\Gamma y$, and the act is still on x, then by comparing the coeffients of y, we have $s_{xy} = 0$, which indicates

$$p\theta + \sum_{i\neq e} q_i(ei) = 0 \tag{2.10}$$

When acting both sides onto $x \neq e$, we obtain $q_x = 0$ by comparing the coefficients of e, therefore $p = 0$. They are indeed linearly independent, hence becoming a basis. □

Till now, for permutation matrices, we have discovered a basis for generating all possible matrices, then we can also describe its convex hull, as well as its dimension to approach our goal.

3. Ideal and Reality: Permutation Matrices with Doubly Stochastic Matrices

Here, with basis $\{e_1, \cdots, e_n\}$, we are to see what those special matrices reveals.

Lemma 3.1. *Every $n \times n$ permutation matrix can be expressed, in a unique manner, as a linear combination with integral coefficients of the $n^2 - 2n + 2$ matrices $P(e_i, e_j), 1 \leq i < j \leq n$ and $P(e_1, e_i, e_j), 1 < i < j \leq n$. In particular, the space spanned by $n \times n$ permutation matrices, over field $\mathbb{R}$, is of dimension $n^2 - 2n + 2$.*

This can be immediately derived by Theorem **??**.

In our combinatorial mathematics course, we have already derived the following results:

Theorem 3.2. *A matrix lies in the convex hull of the set of permutation matrices if and only if it is doubly-stochastic.*

Let $X \subset \mathbb{R}^n$, we have the following notations for proof:

- $C(X)$: The liear variety spanned by X. More precisely,
$$C(X) = \{\lambda_1 x_1 + \cdots + \lambda_k x_k | \sum \lambda_i = 1, x_i \in X, k \in \mathbb{Z}+\} \tag{3.1}$$
- $D(X)$: The convex hull of X.
- $L(X)$: The vector space spanned by X.

By definition, one can deduce that

Lemma 3.3. *If* $0 \notin C(X)$*, then* $\dim(B(X)) = 1 + \dim(C(X))$*.*

We also have a significant covering lemma, which is the key to illustrate why the maximum $n^2 - 2n + 2$ can be reached.

Lemma 3.4. *Let X be a convex set in $\mathbb{R}^n$ with dimension* $\dim(C(X)) = m$*. Then X cannot be convered by a finite number of linear varieties of dimension less than m.*

Proof. Assume not, then we can write
$$X \subset C_1 \cup \cdots \cup C_k \tag{3.2}$$
where C_i are linear varieties of dimension less than m. Set r be the least integer that $X \subset C_1 \cup \cdots \cup C_r$, then $r > 1$, and
$$C_r \cap X \not\subset C_1 \cup \cdots \cup C_{r-1} \tag{3.3}$$
Hence $\exists u \in C_r \cap X$, and $u \notin C_1 \cup \cdots \cup C_{r-1}$. And also there exists $v \in X$ and $v \notin C_r$.

Consider the closed segment $L = uv$, then l is not contained in any C_i, which shows that $L \cap C_i$ has at most one point. Therefore
$$X \cap L \subset L \cap (\cup C_i) = \cup (C_i \cap L) \tag{3.4}$$
is a finite set. However, $u, v \in X$, and X is convex, there must have infinitely points in $X \cap L$, which makes a contradiction! □

Also, a traditional conclusion can be used for the final proof of Brinkhoff's theorem, one can find a brief proof in [**?**].

Lemma 3.5. *Suppose $X \subset \mathbb{R}^n$, and* $\dim(C(X)) = m$*, then each point in $D(X)$ belongs to the convex hull of $m+1$ suitable points in X. Furthermore, if X is finite, then there must be some points in $D(X)$ such that they cannot be represented by any m points in X.*

So here, we are enough to proof the Birkhoff's Theorem.

Proof. Suppose X_n be the set of $n \times n$ permutation matrices, then we consider

$$L_n = L(X_n), \quad C_n = C(X_n), \quad D_n = D(X_n) \tag{3.5}$$

From Lemma **??**, we can derive that

$$\dim(L_n) = n^2 - 2n + 2$$

Note that, the zero matrix does not belong to C_n (Note that the sum of each line and column must be 1), hence by Lemma **??** we have

$$\dim(C_n) = \dim(L_n) - 1 = (n-1)^2$$

Hence, by Theorem **??**, D_n is just the set of all $n \times n$ doubly stochastic matrices. By Lemma **??**, each doubly stochastic $n \times n$ matrix is in the convex hull of at most $n^2 - 2n + 2$ permutation matrices, and there must be some matrices that cannot be any convex conbination of every $n^2 - 2n + 1$ permutation matrices, thus we have already proved Birkhoff's theorem! □

References

[1] H. G. Eggleston, Convexity, Cambridge, 1998.
[2] H. K. Farahat, L. Mirsky , Permutation endomorphisms and refinement of a theorem of Birkhoff, Cambridge, 1959.

School of Mathematical Sciences, Shanghai Jiao Tong University, 800 Dongchuan Road, 200240 Shanghai, China
Email address: zhouqi1729@sjtu.edu.cn, nbwzx@126.com