Optimization with Birkhoff Polytopes

Utkarsh Sharma*

Electrical Engineering, IIT Bombay

April 29, 2017

Abstract

We study the characterization of families of parallel planes based on their intersections with the Birkhoff polytope. We use two methods to this end. The first is by reembedding the polytope in a Euclidean space of the same number of dimensions as the dimensionality of the polytope. The second is by an analysis of projections of the polytope onto an appropriate Euclidean space, such that the projection completely characterizes the polytope.

Contents

*utkarsh.sharma@iitb.ac.in

1 Introduction

Matrices of shape $n \times n$ over a field $\mathbb{F}$ can be looked at as elements of $\mathbb{F}^{n^2}$. Here we confine ourselves to matrices over the real field $\mathbb{R}$. It therefore suffices to study the space $\mathbb{R}^{n^2}$. Since we deal with orthogonality, our analysis also requires a metric, for which the natural choice is the euclidean metric. We therefore look at $n \times n$ matrices over $\mathbb{R}$ as points in $\mathbb{E}^{n^2}$, the n^2 dimensional euclidean space.

We deal here only with $n \times n$ doubly stochastic matrices, which are defined as follows:

$$a_{ij} \geq 0 \qquad \forall i, j \in \{1, 2 \ldots n\} \tag{1}$$

$$\sum_{i=1}^{n} a_{ij} = 1 \qquad \forall j \in \{1, 2 \ldots n\} \tag{2}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \qquad \forall i \in \{1, 2 \ldots n\} \tag{3}$$

where a_{ij} is the entry in row i and column j of the matrix. Doubly stochastic matrices arise in various optimization problems. The advantage of the geometric picture described above arises from the following theorem a proof of which is given in [1]. We reproduce the proof in appendix A

Theorem 1.1 (Birkhoff-von Neumann theorem). *The set of $n \times n$ doubly stochastic matrices forms a polytope (Birkhoff polytope) in n^2 dimensional Euclidean space, with $n \times n$ permutation matrices at its vertices*

As a result, solving linear programming problems concerning optimization of doubly stochastic matrices reduces to solving extremization problems on the corresponding Birkhoff polytope. Given a polytope and a family of parallel hyper planes (i.e. having the same direction cosines), we know that two, and only two members of the family intersect the polytope in a face of the polytope. We say that a face is the extreme face of a family of parallel planes if the family contains a plane that intersects the polytope in that face. Optimization problems require that each family be characterized according to the pair of its extreme faces. Note that with each face is associated a cone formed by direction cosines such that a family of planes corresponding to any of these direction cosines has the given face as its extreme face. Characterizing families of planes according to extreme face pairs also solves the problem of finding the cones associated with the faces of the polytope. We now review the geometric properties of Birkhoff polytopes which will help us solve the above problems. We shall denote the Birkhoff polytope associated with $n \times n$ doubly stochastic matrices as an $n \times n$ Birkhoff polytope.

Corollary 1.1. *The $n \times n$ Birkhoff polytope is $(n-1)^2$ dimensional*

Proof. There are a total of $2n$ equations that constrain the dimensionality of the polytope. Observe that only $2n-1$ of these are linearly independent. The polytope therefore has a dimensionality of $n^2 - (2n-1) = (n-1)^2$ □

The paper is organised as follows: in section 2 we attempt to re-embed this polytope into an $(n-1)^2$ dimensional Euclidean space. The characterization of families of planes becomes particularly simple in this formalism, although not much insight is obtained regarding the numerical values of the results (despite the formalism being very intuitive and natural), the reasons for which are explained towards the end of the section. In section 3 we present a projective formalism in which we consider the projection of the polytope onto an $(n-1)^2$ dimensional Euclidean space. The results obtained this way provide more insight into the problem.

2 Reembedding Formalism

We have established that the polytope is $(n-1)^2$ dimensional, and we know that it lies naturally in an n^2 dimensional Euclidean space (which we shall call the full space). It can therefore be embedded in an $(n-1)^2$ Euclidean space.

2.1 Construction of the Transformation

The problem is to find such an embedding. Looking at each $n \times n$ matrix as a point in the full space, we can write it as a $n^2 \times 1$ matrix, i.e. an length n^2 column vector. The precise prescription for the above expansion, as we shall see, is not significant. The inner product of two vectors is given by the sum of element wise products.

$$\langle v, w \rangle = v^T w = \sum_{i=1}^{n^2} v_i w_i \tag{4}$$

Where v and w are being regarded n^2 length column vectors. In the language of $n \times n$ matrices, this translates to:

$$\langle V, W \rangle = Tr(V^T W) \tag{5}$$

where we regard V and W as square matrices. To carry out the appropriate transformations, we need to find $2n-1$ $(= n^2 - (n-1)^2)$ hyperplanes in which the polytope lies. Observe now that matrices R_k's and C_k's defined below satisfy the required properties. An element r_{ij} of R_k for $i, j, k \in \{1, 2 \ldots n\}$ equals 1 iff it belongs to the k^{th} row, i.e. iff $i = k$ else it is zero. The matrices

C_k's are defined in a similar fashion with c_{ij} being one iff $j = k$, else it equals zero. With this construction, for any doubly stochastic matrix V,

$$\langle V, R_k \rangle = 1 \tag{6}$$
$$\langle V, C_k \rangle = 1 \tag{7}$$

All doubly stochastic matrices therefore lie in hyperplanes perpendicular to R_k's and C_k's. Observe however that only $2n-1$ of these are linearly independent since

$$\sum_{k=1}^{n} C_k = \sum_{k=1}^{n} R_k = K_{11} \tag{8}$$

We need to rotate the coordinate axes such that the resulting polytope lies in $2n-1$ of these planes (any $2n-1$ chosen from the above $2n$). This amounts to constructing an $n^2 \times n^2$ orthogonal matrix O, $2n-1$ rows of which are the normal vectors chosen from above. Take these to be the last $2n-1$ rows of the matrix. To complete the construction, we need $(n-1)^2$ more vectors to fill in the rows of the orthonormal matrix. To this end observe that for any bistochastic matrices V_1 and V_2,

$$\langle R_k, (V_1 - V_2) \rangle = 0 \tag{9}$$
$$\langle C_k, (V_1 - V_2) \rangle = 0 \tag{10}$$

In particular permutation matrices satisfy these equations. Therefore a vector corresponding to any matrix of the form $P - I$, where P is a permutation matrix and I is the identity matrix of appropriate size, is orthogonal to the last $2n-1$ rows of O. Since the number of permutation matrices is $n!$, this procedure can produce $n! - 1$ non-trivial vectors. Since $n! - 1 \geq (n-1)^2 \ \ \forall n \in \mathbb{N}$, there are enough vectors to construct O.

Hence we have obtained two classes of vectors, such that each vector in one class is orthogonal to each vector in the other. We can apply Gram-Schmidt orthonormalization to these classes separately. Hence resulting matrix Q is an orthonormal matrix. Note that multiplication of matrix O with any vector in the polytope results in a vector the bottom $2n-1$ entries of which are all 1. The application of Gram-Schmidt does not change this fact. Hence the matrix Q is an orthonormal matrix that reembeds the polytope into an $(n-1)^2$ dimensional Euclidean space (called the reduced space).

Note that the vector obtained after the transformation has n^2 entries, although the bottom $2n-1$ are ones. We call the full n^2 length representation of this reduced space vector the extended representation.

2.2 Computing Cones of Faces

We shall first carry out all computations in the reduced space and later extend our results to the full space.

2.2.1 Normals to Facets

The transformation of the previous subsection has resulted in an $(n-1)^2$ $(= d)$ dimensional polytope lying in an $(n-1)^2$ Euclidean space. The facets of this polytope therefore form hyperplanes of co-dimension one, i.e. each facet has a unique normal associated with it. Finding the direction normal to a facet requires only to solve a set of homogeneous linear equations. For example if $A_0, A_1 \dots A_{d-1}$ form a facet, the normal to which is $n = (n_1, n_2 \dots n_d)$, then the set of equations $\langle (A_k - A_0), n \rangle = 0$ form a system of homogeneous equations the solution to which exists provided the determinant of the coefficient matrix of $n_1, n_2 \dots n_d$ vanishes, which is true because these points indeed lie in the hyperplane (facet). Lemma 2.1 in the next sub-subsection gives a neat way of finding a facet.

2.2.2 Cones to Other Faces

The directions of planes for which a given face is an extreme form the interior of the convex hull of the normals to the facets that contain the given face. These directions therefore form an open cone. To determining this cone, one needs to find out all facets that contain a given face. This is done easily using the formalism described below, given by Brualdi et al.[2].

We now state a few definitions and a theorem which will allow us to deal with faces of the polytope. We give the proof of the theorem in Appendix B

Definition 2.1. *For a zero-one matrix A, the face corresponding to it, denoted by $\mathcal{F}(A)$ is defined as the set of all doubly stochastic matrices X that satisfy $X \leq A$, where the $\leq$ operator implies an element wise comparison (i.e. $x_{ij} \leq a_{ij}, i, j \in \{1, 2 \dots n\}$ in this case).*

The $n \times n$ permutation matrices P such that $P \leq B$ are precisely the vertices of the faces $\mathcal{F}(B)$. Hence if $B^{'}$ is an $n \times n$ zero-one matrix such that for each permutation matrix P, $P < B$ if and only if $P < B^{'}$, then $\mathcal{F}(B) = \mathcal{F}(B^{'})$. As a consequence, if there exist $r, s \in 1, \dots, n$ with $b_{rs} = 1$ for which there is no permutation matrix $P = [p_{ij}]$ with $p_{rs} = 1$ and $P \leq B$, then $\mathcal{F}(B) = \mathcal{F}(B^{'})$, where $B^{'}$ is obtained from B by replacing b_{rs} by 0. Hence in determining the nonempty faces of the polytope, we need only consider those those $n \times n$ nonzero zero-one matrices that satisfy the property of total support defined next.

Definition 2.2 (Total Support). *An $n \times n$ nonzero zero-one matrix $B = [b_{ij}]$, with the property that $b_{rs} = 1$ implies there exists a permutation matrix $P = [p_{ij}]$ with $p_{rs} = 1$ and $P \leq B$ is said to have total support.*

Lemma 2.1. *The face $\mathcal{F}(A)$ corresponding to an $n \times n$ zero-one matrix A that has only one entry equal to zero, is a facet.*

Proof. One entry is equal to zero implies that one more linear equation (over and above the equations that define the polytope) is needed to define $\mathcal{F}(A)$. The dimensionality of $\mathcal{F}(A)$ is therefore one less than the dimensionality of the polytope. Hence $\mathcal{F}(A)$ is a facet. $\square$

Theorem 2.1. *Let $P_1, \ldots P_t$ be distinct $n \times n$ permutations matrices. Let $A = [a_{ij}]$ be the $n \times n$ zero-one matrix such that $a_{ii} = 1$ if and only if the (i, j)-entry of at least one of the P_k's is 1. Then A has total support and $\mathcal{F}(A)$ is the smallest face of the polytope, which contains the vertices $P_1, \ldots P_t$, . Moreover, $P_1, \ldots P_t$ are the vertices of a face of a face of the polytope if and only if the permanent of A equals t.*

Lemma 2.2. *For any two zero-one matrices A and B of total support, if $B \leq A$ then $\mathcal{F}(B) \subseteq \mathcal{F}(A)$ and vice versa.*

Proof. Any permutation matrix P such that $P \leq B$ satisfies $P \leq A$. Since the face defined by $\mathcal{F}(B)$ is the convex combination of all permutation matrices that satisfy $P \leq B$, this face is contained in $\mathcal{F}(A)$. For the converse, consider a vertex p of $\mathcal{F}(B)$. This point p corresponds to a permutation matrix P such that $P \leq B$. Since p also belongs to $\mathcal{F}(A)$, $P \leq A$ is satisfied i.e. all entries of P are less or equal to the corresponding entries of A. The property of total support implies that every non zero entry of B is a one in some permutation matrix $P \leq B$. Therefore $B \leq A$. $\square$

We can now find all facets that contain a given face. This is the key result of this subsection and is summarized in the following theorem.

Theorem 2.2. *For an $n \times n$ Birkhoff polytope with $n > 2$, any face $\mathcal{F}(B)$, where B is an $n \times n$ zero-one matrix of total support, a matrix A defines a facet $\mathcal{F}(A)$ containing $\mathcal{F}(B)$ if and only if the following two conditions are satisfied:*

1. *It contains only one zero*
2. *$B \leq A$*

Proof. Condition 1 is needed for $\mathcal{F}(A)$ to be a facet. If condition 2 is satisfied then according to lemma 2.2 $\mathcal{F}(B) \subseteq \mathcal{F}(A)$. Since A has total support when $n > 2$, lemma 2.2 also ensures the converse. The theorem is not significant for $n = 2$, since the polytope in that case is one dimensional. $\square$

One can find the normals corresponding to these facets and the interior of the cone formed by those normals is the required region

2.2.3 Extending to Full Space

We have found the required vectors in the reduced space. The extended representation of the column vector of a point in the reduced space has ones as the bottom $2n-1$ entries. We have found the $(n-1)^2$ component representation of the normal vectors. We need to extend them into the extended representation of the reduced space, so that we can act Q^{-1} on them to get normal vectors in the full space. The following lemma allows us to do that.

Lemma 2.3. *Let L be a plane in an n dimensional Euclidean space (N) such that the last $n-m$ components of the coordinates of any point in L are equal to a constant, i.e. L lies in an m dimensional subspace (M) of the full space. If L is regarded as a plane in M, and a vector $u \in M$ is orthogonal to it in M, then the vector v in N, the first m components of which are proportional to those of u, while the last $n-m$ components are arbitrary, is orthogonal to L in N.*

Proof. Consider any two points p_1 and p_2 lying in L. When p_1 and p_2 are viewed as points in M, only their first m coordinates are considered. Since u is orthogonal to L, $\langle u, (p_1 - p_2)\rangle = 0$. In the extended representation, we want $\langle v, (p_1 - p_2)\rangle = 0$. But since the last $n-m$ coordinates of p_1 and p_2 are equal to a constant (in the extended representation), the last $n-m$ coordinates of $(p_1 - p_2)$ are equal to 0. Therefore if the first m coordinates of v are the same as the corresponding ones in u, and the last $n-m$ coordinates of v take arbitrary values, the equation $\langle v, (p_1 - p_2)\rangle = 0$ is satisfied. □

We can therefore concatenate the $(n-1)^2$ direction cosines of the normal obtained above with any values in the bottom $2n-1$ entries to form a normal vector in the extended representation(one has to normalize the vector again though). The transformation from the extended representation of the reduced space to the full space is given by the matrix multiplication with the $n \times n$ matrix Q^{-1}. We thus obtain normals in the full space.

2.3 Results and Shortcomings

The authors have written codes following the above procedure. Numerical results have been obtained for $n = 3$. The problem with this approach is that there does not seem to be any obvious pattern in the values of the normals obtained. This is because the transformations for $(n-1)^2$ coordinates onto which the polytope is projected are chosen arbitrarily. There does not seem to be any obvious natural choice for the top $(n-1)^2$ rows of the transformation matrix.

3 Projective Formalism

We show in this section that there exists a way to project the polytope onto an $(n-1)^2$ dimensional Euclidean space such that a bijection exists between the polytope and its projection, i.e. the polytope is completely determined from its projection. The procedure outlined next constructs such a projection.

3.1 Constructing the Projection

Consider the top left $(n-1)\times(n-1)$ submatrix contained in an $n\times n$ bistochastic matrix. The elements of this submatrix. The entries of this submatrix obey the following constraints:

$$a_{ij} \geq 0 \qquad \forall i,j \in \{1,2\dots n-1\} \tag{11}$$

$$\sum_{i=1}^{n-1} a_{ij} \leq 1 \qquad \forall j \in \{1,2\dots n-1\} \tag{12}$$

$$\sum_{j=1}^{n-1} a_{ij} \leq 1 \qquad \forall i \in \{1,2\dots n-1\} \tag{13}$$

$$\sum_{i,j=1}^{n-1} a_{ij} \geq n-2 \tag{14}$$

The first of these is evident. The second and third arise by combining (1) with (2) and (3) respectively. The fourth arises since (1) implies that $a_{nn} \geq 0$. This proves that these conditions are necessary for any $(n-1)\times(n-1)$ matrix to be a submatrix of an $n\times n$ bistochastic matrix. It is not difficult to see that these are also sufficient. We can construct a unique $n\times n$ bistochastic matrix the top left $n-1\times n-1$ is the given matrix. To this end, choose

$$a_{in} = 1-\sum_{j=1}^{n-1} a_{ij} \forall i \in \{1,2\dots n-1\} \tag{15}$$

$$a_{nj} = 1-\sum_{i=1}^{n-1} a_{ij} \forall j \in \{1,2\dots n-1\} \tag{16}$$

$$a_{nn} = 1-\sum_{i=1}^{n-1} a_{in} = 1-\sum_{j=1}^{n-1} a_{nj} \tag{17}$$

Due to (11), (12), (13) and (14), the above satisfy (1). The construction imposes (2) and (3). No other choice of entries would have held (2) and (3). Uniqueness is therefore ensured. The result is that we can choose $(n-1)^2$ elements subject to certain inequalities and doing so yields a unique point lying in the polytope. This is also consistent with the fact that the polytope is $(n-1)^2$ dimensional.

3.2 Cones from the Projection

We again use the formalism outlined in [2] that we summarized in section 2.2.2. Consider the case of facets. The zero-one matrix A such that $\mathcal{F}(A)$ defines a given facet has one and only one entry zero. Since we look at only the top left $(n-1) \times (n-1)$ sub-matrix, there arise the following four cases

1. $\boldsymbol{a_{kn} = 0}$ **where** $\boldsymbol{k \in \{1, 2 \ldots n-1\}}$
   One of the inequalities that define the projection becomes an equation.

$$\sum_{j=1}^{n-1} a_{kj} = 1 \tag{18}$$

This equation defines the facet in the projection. Notice that this equation is also the equation of a hyperplane in the projected space. The family of planes containing this plane is defined by

$$\sum_{j=1}^{n-1} a_{kj} = r \quad \forall r \in \mathbb{R} \tag{19}$$

Where $r \in \mathbb{R}$ parametrizes the family. We therefore know the direction cosines of the normal to the facet in the projective space. This method also yields the other extreme face of the family. We prove next that this is obtained by choosing $r = 0$.

Theorem 3.1. *The other extreme face of the family of planes given by* (19) *is given by* $r = 0$

Proof. Since $r < 0$ is inconsistent with (11), any value of r that satisfies all constraints is non-negative. We only need to prove that there exists an $n \times n$ doubly stochastic matrix that satisfies (19) with $r = 0$. Take

$$a_{kn} = 1 \tag{20}$$

$$a_{in} = 0 \quad \forall i \neq k \tag{21}$$

Next consider the sub-matrix formed by removing the k^{th} row and the n^{th} column. Take this sub-matrix to be the identity matrix. Then the $n \times n$ matrix is indeed a doubly stochastic matrix. This proves that the intersection of the plane characterized by $r = 0$ with the polytope is nonempty. □

2. $\boldsymbol{a_{nk} = 0}$ **where** $\boldsymbol{k \in \{1, 2 \ldots n-1\}}$
   This case is similar to the one above. The equivalent of theorem 3.1 holds as well.

3. $\boldsymbol{a_{nn} = 0}$
   This constraint translates to

$$\sum_{i,j=0}^{n-1} a_{ij} = n - 2 \tag{22}$$

   This now becomes the equation of the plane containing the facet. Along the lines of the analysis done in case 1, we see that

$$\sum_{i,j=0}^{n-1} a_{ij} = r \qquad r \in \mathbb{R} \tag{23}$$

   The other extreme face of the family is given by $r = n - 1$. The proof for this is similar to that of theorem 3.1. Equations (12) and (13) imply that $r \leq n - 1$. Therefore we need only construct an $n \times n$ doubly stochastic matrix that satisfies the above constraints. It is not hard to see that the identity matrix is one such matrix. Hence the intersection of the plane with the polytope is nonempty.

4. $\boldsymbol{a_{pq} = 0}$ **where** $\boldsymbol{p, q \in \{1, 2 \ldots n-1\}}$
   In this case the equation of the plane itself is

$$a_{pq} = 0 \tag{24}$$

   The family of planes is

$$a_{pq} = r \qquad r \in \mathbb{R} \tag{25}$$

   Therefore the other extreme is clearly $r = 1$ (Since $0 \leq a_{pq} \leq 1$ implies that $r \leq 1$. Besides, there exists at least one permutation matrix that has $a_{pq} \neq 0$. Therefore the intersection of this plane with the polytope is nonempty)

The extension to all other faces is done as in section 2.2.2. The directions of planes for which a given face is an extreme form the interior of the convex hull of the normals to the facets that contain the given face. Lemma 2.2 and the discussion that follows it allow us to construct cones corresponding to arbitrary faces.

4 Acknowledgements

The authors would like thank Umang Mehta (senior undergraduate, department of physics, IIT Bombay) for proof reading the document.

A Proof of Birkhoff-von Neumann Theorem

The key idea is to think of an $n \times n$ matrix as a vector in $\mathbb{R}^{n^2}$. The strategy is to use the Doubly stochastic property (i.e. equations (1), (2) and (3)) 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 [3] that every polytope is the convex hull of its extreme points.

Let $X = [x_{rs}]$ be an $n \times n$ doubly stochastic matrix. The polyhedron P defined above is a polytope since the linear constraints (i.e. equations (1), (2) and (3)) imply that each $0 \leq x_{rs} \leq 1$, 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 $x \in P$ 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 $x^+(\epsilon)$ (resp. $x^-(\epsilon)$) 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 $x^+(\epsilon)$ (resp. $x^-(\epsilon)$) $\in P$. 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 $x^+(\epsilon) \in P$. The analogous argument shows that $x^-(\epsilon) \in P$.

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

B Proof of Theorem 2.1

From the definition of A, A has total support (the definition of total support is given in [2]). Clearly, $\mathcal{F}(A)$ is a face containing $P_1, \ldots, P_t$. Now suppose that B is a zero-one matrix with total support such that $\mathcal{F}(B)$ contains $P_1, \ldots, P_t$. Then $P_i \leq B (i = 1, \ldots, t)$, so that $A \leq B$. Hence $\mathcal{F}(A) \subseteq \mathcal{F}(A)$. Therefore $\mathcal{F}(A)$ is the smallest face containing $P_1, \ldots, P_t$. This, in turn, implies that $P_1, \ldots, P_t$ are the vertices of a face if and only if they are the vertices of $\mathcal{F}(A)$. Since permanent of A (denoted $per(A)$) equals the number of vertices of $\mathcal{F}(A)$, $P_1, \ldots, P_t$ are the vertices of $\mathcal{F}(A)$ if and only if $per(A) = t$.

References

[1] Glenn Hurlbert *A Short Proof Of The Birkhoff-Von Neumann Theorem* `http://people.vcu.edu/ ghurlbert/papers/SPBVNT.pdf`

[2] R. A. Brualdi, P. M. Gibson *Convex polyhedra of doubly stochastic matrices. I. Applications of the permanent function* Journal of Combinatorial Theory, Series A (1977) doi:10.1016/0097-3165(77)90051-6

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