A note on nonnegative diagonally dominant matrices

Geir Dahl

Report 269, ISBN 82-7368-211-0

April 1999
A note on nonnegative diagonally dominant matrices

Geir Dahl *

April 1999

We make some observations concerning the set $C_n^*$ of real nonnegative, symmetric and diagonally dominant matrices of order $n$. This set is a convex cone and we determine its extreme rays. From this we derive different results, e.g., that the rank and the kernel of each matrix $A \in C_n^*$ is determined by a certain support graph of $A$, and may be found explicitly. Moreover, the set of doubly stochastic matrices in $C_n^*$ is studied.

Keywords: Diagonally dominant matrices, convex cones, graphs and matrices.

1 An observation

We recall that a real matrix $A$ of order $n$ is called diagonally dominant if $|a_{i,i}| \geq \sum_{j \neq i} |a_{i,j}|$ for $i = 1, \ldots, n$. If all these inequalities are strict, $A$ is strictly diagonally dominant. These matrices arise in many applications as e.g., discretization of partial differential equations ([14]) and cubic spline interpolation ([10]), and a typical problem is to solve a linear system $Ax = b$ where $A$ is (strictly) diagonally dominant, see also [13]. Strict diagonal dominance is a criterion (which is easy to check) for nonsingularity, and this is important for the estimation of eigenvalues (confer Gerschgorin disks, see e.g. [7]). For more about diagonally dominant matrices, see [7] or [13]. A matrix is called nonnegative (positive) if all its elements are nonnegative (positive).

Let $D_n \subset \mathbb{R}^{n,n}$ denote the set of all matrices of order $n$ that are nonnegative and diagonally dominant. The set of symmetric matrices in $D_n$ is denoted by $D_n^s$. Both these sets are pointed polyhedral (convex) cones in the vector space $\mathbb{R}^{n,n}$ of real matrices of order $n$ as we have

\begin{align*}
D_n &= \{ A \in \mathbb{R}^{n,n} : \quad a_{i,j} \geq 0 \quad \text{for} \quad 1 \leq i, j \leq n; \nonumber \\
&\quad \quad a_{i,i} \geq \sum_{j \neq i} a_{i,j} \quad \text{for} \quad i = 1, \ldots, n \}, \quad (1) \\
D_n^s &= \{ A \in D_n : \quad a_{i,j} = a_{j,i} \quad \text{for} \quad 1 \leq i, j \leq n \}.
\end{align*}

*University of Oslo, Dept. of Mathematics and Dept. of Informatics, P.O.Box 1080, Blindern, 0316 Oslo, Norway (Email: geird@i.uio.no)
Note that the set of diagonally dominant matrices in \( \mathbb{R}^{n,n} \) is a nonconvex cone. The interior of \( D_n \) consists of the positive and strictly diagonally dominant matrices. Similarly, the relative interior of \( D_n^* \) consists of the symmetric, positive and strictly diagonally dominant matrices.

We mention an interesting result from \([9]\) that is relevant to this note. It was shown that if \( A \in D_n^* \), then \( A \) is completely positive. This means that \( A \) can be factored as \( A = BB^T \) for some nonnegative \( n \times m \) matrix. We return to this result in connection with Theorem 3 below.

Let \( S = \{v_1, \ldots, v_k\} \) be a set of \( k \geq 1 \) vectors in a vector space \( V \) (over the reals). The finitely generated convex cone

\[
\text{cone}(S) = \{ \sum_{j=1}^{k} \lambda_j v_j : \lambda_1, \ldots, \lambda_k \geq 0 \}
\]

is said to be spanned by \( S \). If the vectors \( v_1, \ldots, v_k \) are linearly independent, \( \text{cone}(S) \) is called a simplex cone. We need a simple result on such cones.

**Lemma 1** Let \( \text{cone}(S) \) be the convex cone spanned by \( S = \{v_1, \ldots, v_k\} \subset V \). Then \( \text{cone}(S) \) is a simplex cone if and only if each point in \( \text{cone}(S) \) may be written uniquely as a conical (i.e., nonnegative linear) combination of the vectors \( v_1, \ldots, v_k \).

**Proof.** If \( v_1, \ldots, v_k \) are linearly independent, then the representation is clearly unique. Conversely, assume that the uniqueness of such representations hold and that \( \sum_{j=1}^{k} \mu_j v_j = 0 \). Choose nonnegative numbers \( \lambda_j \) and \( \lambda'_j \) for \( j = 1, \ldots, k \) such that \( \mu_j = \lambda_j - \lambda'_j \) for each \( j \). Then \( 0 = \sum_j \mu_j v_j = \sum_j \lambda_j v_j - \sum_j \lambda'_j v_j \). Therefore \( \sum_j \lambda_j v_j = \sum_j \lambda'_j v_j \) so by assumption \( \lambda_j = \lambda'_j \) and \( \mu_j = 0 \) for all \( j \). This shows that \( v_1, \ldots, v_k \) are linearly independent. \( \square \)

We call the unique representation of a point \( v \) in a given simplex cone the conical representation of \( v \).

Let \( e_i \) denote the \( i \)th unit vector in \( \mathbb{R}^n \) and define the following matrices of order \( n \):

\[
\begin{align*}
(i) \quad \Delta^i &= e_i e_i^T \quad \text{for } i = 1, \ldots, n; \\
(ii) \quad \Delta^{i,j} &= (e_i + e_j)(e_i + e_j)^T \quad \text{for } 1 \leq i < j \leq n; \\
(iii) \quad \Delta^{i,j} &= e_i (e_i + e_j)^T \quad \text{for } i \neq j.
\end{align*}
\]

These are all \((0,1)\)-matrices. \( \Delta^i \) has a single one which is in position \((i,i)\). The four ones in the matrix \( \Delta^{i,j} \) are in positions \((i,i), (i,j), (j,i) \) and \((j,j) \). Finally, \( \Delta^{i,j} \) has two ones, in positions \((i,i) \) and \((i,j) \). Let \( S_n^* \) be the set of matrices in \((2)(i) \) and \((ii) \), and let \( S_n^* \) be the set of matrices in \((2)(i) \) and \((ii) \). Note that all these matrices are nonnegative, diagonally dominant and have rank one.

Moreover, the matrices in \( S_n^* \) are symmetric and positive semidefinite.

**Proposition 2** \( D_n = \text{cone}(S_n) \) and \( D_n^* = \text{cone}(S_n^*) \). Moreover, both \( D_n \) and \( D_n^* \) are simplex cones.
Proof. Let $A \in \text{cone}(S^n_+)$ so there are nonnegative numbers $\lambda_i$ for $i \leq n$ and $\lambda_{i,j}$ for $1 \leq i < j \leq n$ such that

$$(*) \quad A = \sum_{i=1}^{n} \lambda_i \Delta^i + \sum_{i<j} \lambda_{i,j} \Delta^{i,j}.$$  

From this it follows that $A$ is symmetric (as a linear combination of symmetric matrices) and nonnegative. Moreover, $(*)$ gives $a_{i,j} = \lambda_{i,j}$ for $1 \leq i < j \leq n$ as only the matrix $\Delta^{i,j}$ has a nonzero in position $(i, j)$. Moreover, due to the structure of the matrices $\Delta^{i,j}$ we also get from $(*)$ that $a_{i,i} = \lambda_i + \sum_{j<i} \lambda_{j,i} + \sum_{i<j} \lambda_{i,j}$, $\lambda_i \lambda_{j,i} = \lambda_i + \sum_{j\neq i} a_{i,j}$ so $\lambda_i = a_{i,i} - \sum_{j\neq i} a_{i,j}$. From this we conclude that $A$ is nonnegative and diagonally dominant and therefore $A \in D_n^+$. Conversely, each $A \in D_n^+$ may be written in the form $(*)$ so we conclude that $D_n^+ = \text{cone}(S^n_+)$. Moreover, we see that each $A \in D_n^+$ has a unique representation as a conical combination of the matrices $\Delta^i$ and $\Delta^{i,j}$. So, according to Lemma 1, $D_n^+$ is a simplex cone. The proof of the results for $D_n$ is similar.

Related results on generators for certain cones are found in [1]. They study different convex cones associated with diagonally dominant matrices, e.g., the complex (or real) matrices of order $n$ satisfying $a_{i,i} \geq \sum_{j \neq i} |a_{i,j}|$ (so the only nonnegativity requirements are on the diagonal elements). Note that the set of these matrices is a convex cone, but not a simplex cone.

We hereafter concentrate our study on the symmetric diagonally dominant matrices, i.e., the set $D_n^+$.

From the previous proof we see that the conical representation of a symmetric, nonnegative and diagonally dominant matrix $A$ is simply

$$A = \sum_{i=1}^{n} (a_{i,i} - \sum_{j \neq i} a_{i,j}) \Delta^i + \sum_{i<j} a_{i,j} \Delta^{i,j}. \quad (3)$$

We define the support graph of a matrix $A \in D_n^+$ as the graph $G_A = (V, E_A)$ with node set $V = \{v_1, \ldots, v_n\}$ and edges (i) $[v_i, v_i]$ (a loop) when $a_{i,i} > \sum_{j \neq i} a_{i,j}$ for $i = 1, \ldots, n$ and (ii) $[v_i, v_j]$ when $a_{i,j} > 0$ for $1 \leq i < j \leq n$. Thus, edges of $G_A$ correspond to the positive coefficients in the conical representation of $A$. This graph will be used below.

When $E$ is some finite set, $S \subseteq E$ and $x \in \mathbb{R}^E$ we use the notation $x(S) := \sum_{e \in S} x_e$. We also let $x(\emptyset) := 0$.

2 Some consequences

We now look at some consequences of our proposition.

Dimension and faces. An immediate consequence of Proposition 2 is that $\dim(D_n) = n^2$ (so it is full-dimensional) and $\dim(D_n^+) = n(n-1)/2$.

The kernel. In order to study the kernel of matrices in $D_n^+$ we need some graph equations. Consider a matrix $A \in D_n^+$. Let $C_1, \ldots, C_k \subseteq V$ be the connected components of the support graph $G_A$. We may assume (after reordering)
that for $1 \leq j \leq p$ the subgraph $G_A[C_j]$ is bipartite and without any loop. Here $0 \leq p \leq t$ and $p = 0$ means that no component is bipartite and without loops. For each $j \leq p$, let $C_j^+$ and $C_j^-$ be the two color classes of $C_j$. Thus, $G_A[C_j]$ is a partition of $C_j$ and each edge of $G_A[C_j]$ joins a node in $C_j^+$ and a node in $C_j^-$. Let $z^{C_j} \in \mathbb{R}^n$ be a vector whose support is $C_j$, and $z^{C_j}_i = 1$ if $v_i \in C_j^+$ and $z^{C_j}_i = -1$ if $v_i \in C_j^-$. If the color classes change role, we obtain the negative of $z^{C_j}$, but this ambiguity will not matter below). Note that we allow the component $C_j$ to be trivial, i.e., with a single node $v_i$ (but no loop), and then $z^{C_j}_i = e_i$.

With this notation we have the following result on the kernel $\ker(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ of a matrix $A \in \mathcal{D}^*_n$.

**Theorem 3** Let $A \in \mathcal{D}^*_n$. Then $\text{rank}(A) = n - p$ and

$$\ker(A) = \text{span}(\{z^{C_1}, \ldots, z^{C_p}\}).$$

**Proof.** Consider the conical representation $A = \sum_{i=1}^n \lambda_i \Delta^i + \sum_{i<j} \lambda_{i,j} \Delta^{i,j}$ where $\lambda_i = a_{i,i} - \sum_{j \neq i} a_{i,j} \geq 0$ and $\lambda_{i,j} = a_{i,j} \geq 0$. Let $x \in \mathbb{R}^n$. From the simple structure of the matrices $\Delta^i$ and $\Delta^{i,j}$ we obtain the following identities

1. $Ax = \sum_{i=1}^n \lambda_i x_i e_i + \sum_{i<j} \lambda_{i,j} (x_i + x_j)(e_i + e_j)$;
2. $x^T Ax = \sum_{i=1}^n \lambda_i x_i^2 + \sum_{i<j} \lambda_{i,j} (x_i + x_j)^2$.

Moreover, in (4) it suffices to sum over those $i$ for which $[v_i, v_i] \in E_A$ (i.e., $\lambda_i > 0$) and those $i < j$ for which $[v_i, v_j] \in E_A$ (i.e., $\lambda_{i,j} > 0$). Note that $x^T Ax \geq 0$ (so A is positive semidefinite).

Let now $x \in \ker(A)$. Then $Ax = 0$ and therefore $x^T Ax = 0$. Thus, from (4) (ii) we see that

a. $x_i = 0$ whenever $[v_i, v_i] \in E_A$ and
b. $x_i = -x_j$ whenever $[v_i, v_j] \in E_A$ (and $i \neq j$).

From (a) and (b) it easily follows that $x_i = 0$ for each node $v_i$ that lies in a component $C_j$ (of $G_A$) which contains an odd cycle or a loop. Consider a component $C_j$ where $1 \leq j \leq p$ (so $G_A[C_j]$ is a bipartite component with no loop). Then it follows from (b) that, for some real number $\alpha$, $x_i = \alpha$ for each $v_i \in C_j^+$ and $x_i = -\alpha$ for each $v_i \in C_j^-$. Thus, the restriction of $x$ to the nodes in $C_j$ lies in $\text{span}(z^{C_j})$. This holds for every $j \leq p$ and we conclude that $x \in \text{span}(\{z^{C_1}, \ldots, z^{C_p}\})$.

Conversely, assume that $x \in \text{span}(\{z^{C_1}, \ldots, z^{C_p}\})$. Then $x_i = 0$ whenever $v_i$ lies in one of the components $C_{p+1}, \ldots, C_t$. Moreover, for each edge $[v_i, v_j] \in E_A$ that belongs to one of the components $C_1, \ldots, C_p$ we have $x_i = -x_j$. If $C_j = \{v_i\}$ is a trivial component (with no loop), then $\lambda_i = \lambda_{k,i} = \lambda_{i,j} = 0$ for $1 \leq k < i$ and $i < j \leq n$, so both the $i$th row and the $i$th column of $A$ are the zero vector. From these observations and (4)(i) it follows that $Ax = 0$ so $x \in \ker(A)$. This proves the description of the kernel. Finally, we note that all the vectors spanning the kernel are nonzero and have disjoint supports.
so they are linearly independent. Therefore, the kernel has dimension $p$ and $\text{rank}(A) = n - p$.

From this result we see the interesting fact that the kernel and the rank of a symmetric, nonnegative diagonally dominant matrix $A$ depends only on the support graph. In other words, the kernel and the rank are determined by which coefficients in the conical representation (3) that are positive; otherwise the magnitudes of these numbers are irrelevant. The reduced row-echelon form of $A$ also has this feature; it only depends on $G_A$. It is also interesting to note that the kernel has a basis consisting of orthogonal $(-1, 0, 1)$-vectors. Finally, we see that the calculation of $\text{rank}(A)$ and $\text{Ker}(A)$ is easily done by a breadth-first-search in the support graph $G_A$ (so no numerical calculation is required).

**Remark.** Theorem 3 and its proof is related to the already mentioned result of [9] saying that each matrix $A \in D^*_n$ is completely positive. In the proof of this result [9] considered the graph $G_A$ and defined its weighted incidence matrix $B$ as follows. $B$ has a row for each node of $G_A$ and a column for each edge in $E_A$, and $b_{v_i, [v_i, v_j]} = b_{u_i, [v_i, v_j]} = a_{i,j}$ when $[v_i, v_j] \in E_A$. For $b_{v_i, [v_i, v_i]} = (a_{i,i} - \sum_{j \neq i} a_{i,j})^{1/2}$ while all other entries are zero. Then one can check that $A = BB^T$. In connection with the proof of Theorem 3 we note that $\text{Ker}(B^T) = \text{Ker}(BB^T) = \text{Ker}(A)$, and that $B^Tx = 0$ is just conditions (a) and (b) in our proof.

**Range and linear systems.** Let $A \in D^*_n$. Since $A$ is symmetric, we have $\text{Ran}(A) = \text{Ker}(A)^\perp$ where $\text{Ran}(A) = \{Ax : x \in \mathbb{R}^n\}$ is the range of $A$. Thus, $\text{Ran}(A)$ consists of the vectors $x \in \mathbb{R}^n$ satisfying

$$x(C^+_j) = x(C^-_j) \quad \text{for } j = 1, \ldots, p.$$  \hspace{1cm} (5)

(If $C_j$ consists of a single node, the equation says that the corresponding variable $x_i$ is zero). Choose, for each $j \leq p$, an index $k(j)$ such that $v_{k(j)} \in C^-_j$ (if $C_j = \{v_i\}$, let $k(j) = i$). Then, a basis of $\text{Ran}(A)$ consists of the vectors $\mathbf{e}_i + \mathbf{e}_{k(j)}$, $i \in C^+_j$ and $\mathbf{e}_i - \mathbf{e}_{k(j)}$, $i \in C^-_j$ for $j = 1, \ldots, p$. Let now $b \in \mathbb{R}^n$ and consider the linear system of equations $Ax = b$. This system has a solution if and only if $b$ satisfies (5). Moreover, if this condition holds the solution set of $Ax = b$ is the affine set $x_0 + \text{span}\{\{z^{C_1}, \ldots, z^{C_p}\}\}$ where $x_0$ is some solution of $Ax = b$.

**Positive (semi)definite.** It is well-known that each symmetric diagonally dominant matrix is positive semidefinite. The fact that this is true for nonnegative matrices is an immediate consequence of Proposition 2 as we have noted that each matrix in $S^*_n$ is positive semidefinite (or it was observed in the proof of Theorem 3). The set of (symmetric) positive semidefinite matrices of order $n$ is a (nonpolyhedral) convex cone PSD$_n$ which contains $D^*_n$ as a subcone. See [5] for a discussion of many aspects of PSD$_n$, related cones and convex sets.

The positive definite matrices in $D^*_n$ may be characterized in terms of the support graph in the following way.

**Corollary 4** Let $A \in D^*_n$. Then the following three statements are equivalent:

---

5
(i) \( A \) is positive definite.

(ii) \( A \) is nonsingular.

(iii) Each component of \( G_A \) contains a loop or an odd cycle.

**Proof.** The equivalence of (i) and (ii) follows from the fact that \( A \) is positive semidefinite. Moreover, \( A \) is nonsingular if and only if \( \text{rank}(A) = n \), which, by Theorem 3 means that \( p = 0 \), i.e., each component of \( G_A \) contains a loop or an odd cycle.

For instance, consider the matrices

\[
A = \begin{bmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 1 & 2
\end{bmatrix}, \quad B = \begin{bmatrix}
1 & 1 & 0 \\
1 & 2 & 1 \\
0 & 1 & 1
\end{bmatrix}.
\]

\( A \) is positive definite because \( G_A \) is an odd cycle (a triangle) while \( B \) is singular (\( G_B \) is a path). Note that these matrices are not strictly diagonally dominant, in fact, \( a_{i,i} = \sum_{j \neq i} a_{i,j} \) for each \( i \) (and similar equations hold for \( B \)).

Let \( A \in \mathcal{D}^+_n \) be tridiagonal, i.e., \( a_{i,j} = 0 \) when \( |i - j| > 1 \), and assume that \( a_{i,i+1} = a_{i+1,i} > 0 \) for \( i = 1, \ldots, n \). Then \( G_A \) contains the chain \( [v_i, v_{i+1}] \) for \( i = 1, \ldots, n-1 \), so \( G_A \) is connected. Thus, by Corollary 4, \( A \) is nonsingular (and positive definite) if and only if \( a_{i,i} > a_{i,i+1} + a_{i-1,i} \) for some \( i \). Such matrices are of interest in connection with cubic splines, see [10]. More generally, assume that \( A \in \mathcal{D}^+_n \) is not decomposable, i.e., there is no permutation matrix \( P \) such that \( P^T A P = A_1 \oplus A_2 \) where \( A_1 \) and \( A_2 \) are square, nonvacuous matrices.

This implies (in fact, it is equivalent to) that \( G_A \) is connected. Thus, Corollary 4 gives that \( A \) is nonsingular (and positive definite) if and only if \( G_A \) contains a loop (\( a_{i,i} > \sum_{j \neq i} a_{i,j} \)) or an odd cycle.

We refer to [6], [7] and [15] for other criteria for a diagonally dominant matrix to be nonsingular.

**Faces.** Recall that a (nontrivial) face of a convex set \( C \) is the intersection between \( C \) and one of its supporting hyperplanes. Consider \( A \in \mathcal{D}^+_n \) and let \( F(A) \) denote the smallest face of \( \mathcal{D}^+_n \) that contains \( A \). Then \( \dim(F(A)) = |E_A| \) and \( F(A) \) is the simplex cone spanned by the matrices \( \Delta^1 \) for \( [v_i, v_i] \in E_A \) and \( \Delta^{ij} \) with \( [v_i, v_j] \in E_A \). It follows from Theorem 3 that the maximum rank among the matrices in \( F(A) \) is \( n - p \) where \( p \) is the number of components of \( G_A \) that are bipartite and without any loop. This maximum rank is achieved for all matrices in the relative interior of \( F(A) \) (i.e., the matrices having conical representation with positive coefficient for each edge in \( E_A \)).

**Related polytopes.** Let \( \alpha > 0 \) and consider the set of matrices in \( \mathcal{D}^+_n \) with trace \( \alpha \), i.e., \( \mathcal{D}^+_n(\alpha) = \{ A \in \mathcal{D}^+_n : \sum_{i=1}^n a_{i,i} = \alpha \} \). Then \( \mathcal{D}^+_n(\alpha) \) is a simplex and its vertices are the zero matrix and the points in the intersection between the extreme rays of \( \mathcal{D}^+_n \) and the hyperplane \( \sum_{i=1}^n a_{i,i} = \alpha \). Thus the nonzero vertices of \( \mathcal{D}^+_n(\alpha) \) are the matrices \( \alpha \Delta^i \) for \( i = 1, \ldots, n \) and \( (\alpha/2) \Delta^{ij} \) for \( 1 \leq i < j \leq n \).

**Convex optimization.** Let \( f \) be a real-valued convex function defined on \( \mathbb{R}^{n,n} \), i.e., \( f((1 - \lambda)A + \lambda B) \leq (1 - \lambda)f(A) + \lambda f(B) \) for each \( A, B \in \mathbb{R}^{n,n} \) and \( 0 \leq \lambda \leq 1 \). An example is \( f(A) = \|A\| \) where \( \| \cdot \| \) is an arbitrary matrix norm. Another example is the linear function \( f(A) = \langle C, A \rangle = \sum_{i,j} c_{i,j} a_{i,j} \).
From convexity we know that a convex function defined on a polytope achieves its maximum in one of the vertices, so \( \max \{ f(A) : A \in D^*_n(\alpha) \} \) equals the maximum of the numbers \( f(0), f(\alpha \Delta^i) \) for \( i = 1, \ldots, n \) and \( f((\alpha/2) \Delta^{i,j}) \) for \( 1 \leq i < j \leq n \). (When \( f \) is positively homogeneous, \( \alpha \) may be moved out of the maximization.) As an example, let \( f \) be the spectral norm so \( f(A) = \|A\|_2 \) is the largest eigenvalue of \( A \) (as \( A \) is symmetric). The characteristic polynomial of \( \Delta^i \) is \( (\lambda - 1)\lambda^{n-1} \) and the characteristic polynomial of \( \Delta^{i,j} \) is \( (\lambda - 2)\lambda^{n-1} \). This gives \( f(\alpha \Delta^i) = \alpha \cdot \lambda_{\text{max}}(\Delta^i) = \alpha \) and \( f((\alpha/2) \Delta^{i,j}) = (\alpha/2) \cdot \lambda_{\text{max}}(\Delta^{i,j}) = \alpha \). Therefore
\[
\max \{ \|A\|_2 : A \in D^*_n(\alpha) \} = \alpha
\]
and the maximum is attained for all the matrices \( \Delta^i \) and \( \Delta^{i,j} \).

**Matrices with nonpositive off-diagonal elements.** In the discretization of certain partial differential equations one is interested in symmetric diagonally dominant matrices with nonnegative diagonal elements, but nonpositive off-diagonal elements. Let \( M^*_n \) denote the set of such matrices of order \( n \). Using similar proof techniques as above one may show the following results. \( M^*_n \) is a simplex cone defined by the matrices \( \Delta^i \) (as before) and \( (e_i - e_j)(e_i - e_j)^T \) for \( 1 \leq i < j \leq n \). Define the support graph \( G_A = (V, E_A) \) nearly as before: \( V = \{v_1, \ldots, v_n\} \) and edges (i) \([v_i, v_j]\) (a loop) when \( a_{i,i} > \sum_{j \neq i} |a_{i,j}| \) for \( i = 1, \ldots, n \) and (ii) \([v_i, v_j]\) when \( a_{i,j} < 0 \) for \( 1 \leq i < j \leq n \). Again, the edges of \( G_A \) correspond to the positive coefficients in the conical representation of \( A \). We then have for \( A \in M^*_n \) that
\[
\ker(A) = \text{span}(\{\chi^{C_1}, \ldots, \chi^{C_q}\})
\]
where \( C_1, \ldots, C_q \) are the components of \( G_A \) without a loop and \( \chi^{C_j} \) is the \((0,1)\)-incidence vector of \( C_j \) (i.e., \( \chi_{C_j} \) is 1 if \( v_i \in C_j \) and 0 otherwise). Moreover, \( \text{rank}(A) = n - q \). Note that, in contrast to the case of nonnegative matrices, whether the components of \( G_A \) are bipartite or not plays no role for the kernel or the rank. But again we have the interesting fact that the kernel and the rank depends only on the support graph.

We also see that \( A \in M^*_n \) is nonsingular if and only if each component of \( G_A \) contains a loop. To recognize this condition, we see that after simultaneous permutations of rows and columns of \( A \) it may be written as the direct sum of smaller matrices, say \( A_1, \ldots, A_r \), each corresponding to a component of \( G_A \). Clearly, each \( A_i \) lies in \( M^*_k \) for some \( k \). Now, \( A_i \) is irreducible as it corresponds to a component of \( G_A \) (and \( A_i \) is symmetric). Moreover, the statement that this component has a loop just means that \( a_{i,i} > \sum_{j \neq i} |a_{i,j}| \) for some \( i \) (where node \( v_i \) lies in that component). Thus, \( A_i \) is irreducibly diagonally dominant, a known criterion for \( A_i \) to be nonsingular (see [7]).

Note that if \( A \in M^*_n \) is nonsingular and \( a_{i,i} > 0 \) for each \( i \) (meaning that \( A \) has no zero row), then \( A \) is a Stieltjes matrix, i.e., a symmetric \( M \)-matrix. Thus \( A^{-1} \geq 0 \).

**Eigenvalues in a restricted case.** Consider a matrix \( A \in D^*_n \) such that \( a_{i,j} \in \{0,1\} \) when \( i \neq j \) and \( a_{i,i} = \sum_{j \neq i} a_{i,j} \) for each \( i \). Thus, the support graph \( G_A \) has no loop and it determines \( A \) uniquely. We note that by changing sign on
all off-diagonal elements of $A$ we obtain the Laplacian matrix of $G_A$. Assume that $G_A$ is bipartite, say with color classes $I$ and $J$. Thus, $A$ is singular (by Corollary 4) and 0 is the smallest eigenvalue of $A$. Let $\mu_A$ denote the second smallest eigenvalue of $A$. Then $\mu_A$ is related to connectivity properties of the support graph $G_A$. To clarify this, note first that

$$\nu_A = \min\left\{ \sum_{[i,j] \in E_A} (x_i + x_j)^2 : \|x\| = 1, \ x(I) = x(J) \right\}$$

as $z = \chi^I - \chi^J$ is an eigenvector corresponding to the eigenvalue 0 ($z \in \text{Ker}(A)$). Consider the Courant-Fischer minmax theorem, see [7]. Introducing the change of variables $y_i = x_i$ for $i \in I$ and $y_i = -x_i$ for $i \in J$ we see from (7) that

$$\mu_A = \min\left\{ \sum_{[i,j] \in E_A} (y_i - y_j)^2 : y \in U \right\}$$

where $U$ consists of the vectors $y \in \mathbb{R}^n$ with $\|y\| = 1$ and $e^T y = \sum_{i=1}^n y_i = 1$. This means that $\mu_A$ is equal to the so-called algebraic connectivity of the graph $G_A$. We refer to [2] for a discussion of algebraic connectivity (and the Laplacian matrix of a graph). Several properties of $\mu_A$ are known, but here we just mention that (i) $\mu_A$ is positive if and only if $G_A$ is connected, and (ii) $\mu_A$ is no greater than the node connectivity of $G_A$.

## 3 Doubly stochastic diagonally dominant matrices

A matrix is doubly stochastic if it is nonnegative and each row and column sum is 1. We let $B_n$ denote the set of doubly stochastic matrices of order $n$. The set $B_n$ is a convex polytope in $\mathbb{R}^{n,n}$, often called the Birkhoff polytope. The classical Birkhoff-von Neumann theorem states that $B_n$ is the convex hull of all permutation matrices of order $n$. For more information about this theorem and doubly stochastic matrices, we refer to [2] and [7]. We are here concerned with the set $DB^*_n$ of symmetric, diagonally dominant and doubly stochastic matrices (of order $n$), i.e.,

$$DB^*_n = D^*_n \cap B_n.$$

Note that $DB^*_n$ is a (convex) polytope and that the only integral matrix in $DB^*_n$ is the identity matrix. We shall give different representations of $DB^*_n$. Let $B^*_n$ be the set of symmetric matrices in $B_n$. Let $\delta(v)$ denote the set of edges incident to a node $v$ in a graph (including, possibly, the loop $[v,v]$). We need the following lemma concerning the “fractional perfect matching polytope” in graphs with loops. It may be proved using techniques explained in [12] (see also [3]).

**Lemma 5** Let $G = (V,E)$ be a connected graph, possibly with loops and define the polytope $FM(G) = \{x \in \mathbb{R}^E : x \geq 0, \ x(\delta(v)) = 1 \text{ for all } v \in V \}$. 

8
Then \( x \in \mathbb{R}^E \) is a vertex of \( FM(G) \) if and only if \( x_e \in \{0, 1/2, 1\} \) for each \( e \in E \) and the edges \( e \) with \( x_e = 1/2 \) form node disjoint odd cycles.

This result may be reformulated in terms of matrices. A symmetric matrix \( A \in \mathbb{R}^{n,n} \) may be represented by a weighted graph \( G = (V, E) \) with nodes \( v_1, \ldots, v_n \) and edges \( \{v_i, v_j\} \) with associated weight \( x_{i,j} := a_{i,j} = a_{j,i} \) for \( 1 \leq i \leq j \leq n \) (when \( i = j \) we have a loop \( [v_i, v_i] \)). We see that \( A \) is symmetric and doubly stochastic if \( x \in \mathbb{R}^E \) is nonnegative and \( x(\delta(v_i)) = 1 \) for each \( i \leq n \). Thus, \( B_n^* \) and \( FM(G) \) are affinely isomorphic. Let \( A \) be a vertex of \( B_n^* \). Consider the corresponding vertex \( x \) of \( FM(G) \) and choose an ordering of the vertices so that (i) the vertices of each fractional cycle (having edges with \( x_e = 1/2 \) occur consecutively, and (ii) the endpoints of each edge with \( x_e = 1 \) occur consecutively. The node ordering corresponds to simultaneous line permutations in \( A \) and we see from Lemma 5 that the resulting matrix \( Q^T A Q \) may be written as the direct sum of the matrices

\[
[1], \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad C^{(p)}
\]

where \( C^{(p)} = [c_{i,j}] \in \mathbb{R}^{p \times p} \) is defined by \( c_{i,i} = c_{i-1,i} = 1/2 \) for \( 2 \leq i \leq p-1, c_{1,2} = c_{i,p} = c_{p,1} = c_{p,p-1} = 1/2 \) and \( c_{i,j} = 0 \) otherwise. Here the first and the second matrix corresponds to a loop and an edge with \( x_e = 1 \), respectively, while \( C^{(p)} \) corresponds to a fractional cycle of length \( p \) where \( p \) is odd. This also shows the following result due to [8] (see also [4]).

**Proposition 6** The set \( B_n^* \) of symmetric doubly stochastic matrices is the convex hull of matrices of the form \((1/2)(P + P^T)\) where \( P \) is a permutation matrix of order \( n \).

Observe that a matrix \( A \) with \( \sum_j a_{i,j} = 1 \) satisfies \( a_{i,1} \geq \sum_{j \neq i} a_{i,j} \) if and only if it satisfies \( a_{i,i} \geq 1/2 \). It follows that \( DB_n^* \) consists of the matrices \( A \) satisfying the following linear system

\[
\begin{align*}
\sum_{j=1}^{n} a_{i,j} &= 1 & & \text{for } i = 1, \ldots, n; \\
a_{i,j} &= a_{j,i} & & \text{for } 1 \leq i < j \leq n; \\
a_{i,i} &\geq 1/2 & & \text{for } i = 1, \ldots, n; \\
a_{i,j} &\geq 0 & & \text{for } 1 \leq i < j \leq n. \\
\end{align*}
\]

(6)

Other descriptions of \( DB_n^* \) are contained in the following proposition. We let \( I_n \) denote the identity matrix of order \( n \).

**Corollary 7** (i) \( DB_n^* = (1/2) \cdot I_n + (1/2) \cdot B_n^* \).

(ii) \( DB_n^* \) is the convex hull of the matrices \((1/2) \cdot I_n + (1/4) \cdot (P + P^T)\) where \( P \) is a permutation matrix of order \( n \).
Proof. Statement (i) is a direct consequence of (6). Next, from (i) we see that the vertices of $DB^*_n$ are of the form $(1/2) \cdot I_n + (1/2) \cdot Z$ where $Z$ is a vertex of $B^*_n$. Thus, statement (ii) is a consequence of Proposition 6.

Thus, the polytopes $DB^*_n$ and $B^*_n$ are affinely isomorphic. We note that the dimension of $DB^*_n$ is $n(n-1)/2$.

We get similar relations for $DB_n := D_n \cap B_n$: the diagonally dominant doubly stochastic matrices. So $DB_n = (1/2) \cdot I_n + (1/2) \cdot B_n$ and $DB_n$ is also equal to the convex hull of the matrices $(1/2) \cdot I_n + (1/2) \cdot P$ where $P$ is a permutation matrix of order $n$. The polytope $DB_n$ and the Birkhoff polytope $B_n$ are affinely isomorphic.

Finally, we point out that the set $DB_n$ may be of interest in connection with majorization (see [11]). If $x, y \in \mathbb{R}^n$, one says that $y$ is majorized by $x$, denoted by $y \prec x$, provided that $\sum_{j=1}^k y[j] \leq \sum_{j=1}^k x[j]$ for $k = 1, \ldots, n - 1$ and $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$. (Here $x[j]$ denotes the $j$th largest number among the components of $x$.) A well-known theorem of Hardy-Littlewood and Pólya (see [11]) says that $y \prec x$ if and only if $Bx = y$ for some $B \in B_n$. Consider now the stronger property that

$$Ax = y \quad \text{for some } A \in DB_n \tag{7}$$

so $A$ is not just doubly stochastic, but also diagonally dominant. From the description of $DB_n$ given above we see that (7) holds if and only if $y = (1/2)x + (1/2)z$ for some $z \prec x$. The geometrical interpretation is that $y$ is the midpoint of the line segment between $x$ and a point $z$ in the convex hull of all permutations of $x$. Or, equivalently, $y$ is a convex combination of points of the form $(1/2)x + (1/2)Pz$ where $P$ is a permutation matrix. A similar characterization may be given when (7) holds for a matrix in $DB^*_n$ (using Proposition 6 and Corollary 7).

References


