On the facial structure of cone-LP's and semi-definite programs

Gábor Pataki *

Graduate School of Industrial Administration

Carnegie Mellon University

Management Science Research Report #MSRR-595

February 3, 1994

Abstract

In this article we study the facial structure of the primal and dual pair of *cone-LP's*

$$\max \{ cx \ : \ x \in K, \ Ax = b \} \tag{1}$$

and

$$\min \{ yb \ : \ yA - c \in K^* \}. \tag{2}$$

We give particular consideration to *semi-definite programs* (SDP's) , where $K = K^*$ is the cone of symmetric positive semi-definite matrices. We derive a characterization of the faces of both the primal and dual feasible sets. As an application, we give bounds on the rank of optimal matrices in SDP's, and on the multiplicity of eigenvalues in eigenvalue-optimization.

1 Introduction

Consider the primal and dual pair of cone-LP's

$$\begin{array}{ll} max & cx \\ s.t. & x \in K \\ & Ax = b \end{array} \tag{P}$$

and

$$\begin{array}{ll} min & yb \\ s.t. & yA - c \in K^* \end{array} \tag{D}$$

When $K = \Re^n_+$ we get standard linear programming. Another interesting special case is when $K = K^* = \{X : X \in \Re^{n\times n}, X \text{ is positive semi-definite}\}$; the resulting optimization problems are called *semi-definite programs* (SDP's). For results and background on cone-LP's we refer to [BICK69],[BW81a], [BW81b], [BW81c], [Wol81]. For applications of SDP and a generalization of a polynomial-time interior-point LP-algorithm, see [Ali92]. For non-polynomial time, efficient algorithms, see [Ove92], [RVW93].

In this article we study the facial structure of cone-LP's with special consideration given to SDP's. In the remainder of the Introduction we review some preliminaries and fix notations. In Section 2. we present some results in convex analysis, which might be of interest on their own right. In Section 3. we use these results to characterize the faces of the primal and dual feasible regions in cone-LP's, and of SDP's in particular. For SDP we give a bound on the rank of the extreme matrices; we can view these results as a characterization of "basic solutions". In Section 4. we use these results to analytically explain an empirical observation: why do optimal eigenvalues in eigenvalue-optimization problems have a large multiplicity? Section 4. uses a "dual SDP" together with an expression for the optimal solution to determine the sum of the k largest eigenvalues of a symmetric matrix. This formulation is derived in Appendix A. In Appendix B we give a result, which could be of interest on its own right. It presents an orthogonal set of generators for a special subspace of the space of n by n symmetric matrices, thus also determining its dimension.

In a forthcoming sequel of our paper we apply our results to generalize primal and dual simplex as well as primal-dual algorithms for cone-LP's.

Notation

$\Re^{n\times n}$ denotes the set of n by n matrices, $\mathcal{S}^{n\times n}$ the set of n by n symmetric matrices, $\mathcal{S}^{n\times n}_+$ the set of n by n symmetric positive semi-definite matrices. For $A, B \in \mathcal{S}^{n\times n}$ $A \succeq B$ means that $A - B$ is positive semidefinite. If $A = (a_{ij})^n_{i,j=1}$ and $B = (b_{ij})^n_{i,j=1}$, the inner product of A and B is

$$A \circ B = \sum_{i,j=1}^{n} a_{ij} b_{ij} \tag{1.5}$$

For $X \in \mathcal{S}^{n\times n}$, $\lambda_i(X)$ denotes the i'th largest eigenvalue of X. Let L be a subset of $\Re^n$. We denote

$$LL^T = \{uu^T : u \in L\} \tag{1.6}$$

Let S be a closed convex set. The relative interior and the relative boundary of S are denoted by ri S and rb S, respectively. Conv, cone, aff and lin denote convex, conical, affine, and linear hull, respectively.

2 Results in convex analysis

Throughout this section let $S \subseteq \Re^n$ be a closed convex set, $A : \Re^n \to \Re^m$ an affine function. Let $C = \{x : A(x) \in S\}$. It is easy to see, that C is convex. Our aim is to characterize the faces of C. Our main result is the following

Theorem 2.1 *F is a face of C if and only if $F = \{x : A(x) \in G\}$ for some face G of S.*

First, we need some preliminary results.

Lemma 2.2 *Let G be a face of S. Let $x \in C$ be s.t. $A(x) \in \text{ri } G$. Let $y \in \Re^n$ s.t. $A(y) \in \text{aff } G$. Then there exists $\epsilon > 0$ s.t.*

$$\begin{aligned} z_1 &= (1-\epsilon)x + \epsilon y \\ z_2 &= (1+\epsilon)x - \epsilon y \end{aligned}$$

are both in C.

Proof Let z_1, z_2 be as above for some $\epsilon \geq 0$. Since A is an affine function,

$$\begin{aligned} A(z_1) &= (1-\epsilon)A(x) + \epsilon A(y) \\ A(z_2) &= (1+\epsilon)A(x) - \epsilon A(y) \end{aligned}$$

Since $A(x) \in \text{ri } G$ it follows that we may choose $\epsilon > 0$. □

Lemma 2.3 *Let S' be a convex set, $S' \subseteq G$, where G is a face of S. Then G is the smallest face of S containing S' if and only if $S' \cap \text{ri } G \neq \emptyset$.*

Proof S' is convex, thus by Theorem 18.2 in [Roc70] there is a unique face H of G, s.t. $S' \subseteq H$. Since $G = \text{ri } G \cup \text{rb } G$ the lemma follows. □

Lemma 2.4 *Let F be a face of C, G the smallest face of S containing $\bigcup\{A(x) : x \in F\}$. Then*

$$\text{aff } F = \{x : A(x) \in \text{aff } G\}.$$

Proof The inclusion $\subseteq$ is obvious.

$\supseteq$: Let y be s.t. $A(y) \in \text{aff } G$.

By Lemma 2.3 there exists $\bar{x} \in F$ s.t. $A(\bar{x}) \in \text{ri } G$. By Lemma 2.2 there exists $\epsilon > 0$ s.t.

$$\begin{aligned} z_1 &= (1-\epsilon)x + \epsilon y \\ z_2 &= (1+\epsilon)x - \epsilon y \end{aligned}$$

are both in C. Since $x = \frac{1}{2}(z_1 + z_2)$, and F is a face, we must have $z_1, z_2 \in F$. Thus $y \in \text{aff } \{z_1, z_2\} \subseteq \text{aff } F$, as required. $\square$

Proof of Theorem 2.1 The "if" part is immediate from the definition of a face.

To show "only if", let F be a face of C. By Lemma 2.4 there exists a face G of S s.t. $\text{aff } F = \{x : A(x) \in \text{aff } G\}$. Then

$$\begin{aligned} F &= C \cap \text{aff } F \\ &= \{x : A(x) \in S\} \cap \{x : A(x) \in \text{aff } G\} \\ &= \{x : A(x) \in S \cap \text{aff } G\} \\ &= \{x : A(x) \in G\} \end{aligned}$$

$\square$

Corollary 2.5 *Let C_1, C_2 be closed convex sets, $C = C_1 \cap C_2$. Then F is a face of C if and only if $F = F_1 \cap F_2$ for some faces F_1 of C_1 and F_2 of C_2.*

Proof Define $A(x) = (x, x)$, $S = C_1 \times C_2$. Thus $C = \{x : A(x) \in S\}$. By Theorem 2.1 F is a face of C if and only if $F = \{x : (x, x) \in G\}$ for some face G of $C_1 \times C_2$. From the definition of a face we have: G is a face of $C_1 \times C_2$ if and only if $G = F_1 \times F_2$ for faces F_1 of C_1 and F_2 of C_2. Thus our theorem follows.

$\square$

3 Faces of cone-LP's and semi-definite programs

Let K be a closed, pointed, full-dimensional cone, K^* its polar, A a matrix, b a vector of appropriate dimensions (although the affine function in the previous chapter was also denoted A, the distinction should be clear from the context).

Theorem 3.1 *Let*

$$P = \{x : x \in K, Ax = b\}$$
$$D = \{y : yA - c \in K^*\}$$

Then

1. *F is a face of P if and only if* $F = \{x : x \in F', Ax = b\}$ *for some face F' of K.*
2. *G is a face of D if and only if* $G = \{y : yA - c \in G'\}$ *for some face G'of* K^*.

Proof Both 1. and 2. follow from Theorem 2.1, 1. by setting $A(x) = (x, Ax - b)$, $S = K \times \{0\}$, 2. by setting $A(y) = yA - c$, $S = K^*$. □

Now we consider the primal and dual programs (P) and (D) with $K = K^* = \mathcal{S}_+^{n \times n}$. Thus the rows of A are now symmetric matrices. We shall use the results of Section 2. to get bounds on the ranks of the optimal extreme matrices.

Let m denote the number of rows of A. In (D) let $S(y) = yA - c$, the "slack matrix". For $l \in \mathcal{N}$ let $d_l = \frac{1}{2}l(l+1)$. We prove in Appendix B that if L is a subspace of $\mathcal{R}^n$ s.t. $\dim L = l$ then $\dim LL^T = d_l$. It is known ([BC75]), that F is a face of $\mathcal{S}_+^{n \times n}$ if and only if $F = \text{cone } LL^T$ for some L subspace of $\mathcal{R}^n$.

Theorem 3.2 1. *Let X^* be an extreme point of the primal program. Then* rank $X^* \leq l$, *for some $l \in \mathcal{N}$ satisfying $d_l \leq m$.*

2. *Let y^* be an extreme point of the dual program. Then* rank $S(y^*) \leq j$, *for some $j \in \mathcal{N}$ satisfying $d_j \leq d_n - m$.*

Proof 1. By Lemma 2.4 and Theorem 2.1

$$x^* = \{x : x \in \text{cone } LL^T, Ax = b\}$$
$$\text{aff } x^* = \{x : x \in \text{lin } LL^T, Ax = b\}$$

for some subspace L of $\mathcal{R}^n$. Let $l = \dim L$. Since $\dim(\text{aff } x^*) = 0$ we must have $d_l \leq m$.

2. Again by Lemma 2.4 and Theorem 2.1

$$S(y^*) = \{y : S(y) \in \text{cone } JJ^T\}$$
$$\text{aff } S(y^*) = \{y : S(y) \in \text{lin } JJ^T\}$$

for some subspace J of $\mathcal{R}^n$. Let $j = \dim J$. Since $\dim JJ^T = d_j$, the condition $S(y^*) \in \text{lin } JJ^T$ imposes $d_n - d_j$ independent constraints on $S(y^*)$. Since A has m rows and y^* is an extreme point we must have $d_n - d_j \geq m$. □

4 Multiplicity of optimal eigenvalues

An important optimization problem that can be formulated as an SDP is the following: given matrices $A_0, A_1, ..., A_m$ find their affine combination with largest eigenvalue as small as possible. More generally, we may want to minimize the sum of the k largest eigenvalues for some fixed $k \in \mathcal{N}$. That is, we want to solve

$$\begin{array}{ll} min & \lambda_1(X) \\ s.t. & X = A_0 + y_1 A_1 + ... + y_m A_m \end{array} \tag{4.7}$$

or

$$\begin{array}{ll} min & \lambda_1(X) + ... + \lambda_k(X) \\ s.t. & X = A_0 + y_1 A_1 + ... + y_m A_m \end{array} \tag{4.8}$$

(Clearly, it makes no difference, whether X is given as an affine combination of matrices, or with linear constraints.)

Problems of this type arise both in combinatorial optimization, see [CDW75], [GLS88] as well as in engineering, see [Ove92]. For algorithms, we refer the reader to [CDW75], [Ove92].

A phenomenon observed by most authors studying eigenvalue-optimization is the following: at an optimal solution optimal eigenvalues tend to coalesce; when minimizing $\sum_{i=1}^{k} \lambda_i(X)$, $\lambda_k(X)$ usually has multiplicity larger than one. To our knowledge, so far no explanation has been given to the above phenomenon.

Several algorithms, e.g. [Ove92] work by maintaining and successively updating an estimate of the optimal eigenvalue-multiplicity. Thus having an a priori bound on the multiplicity can be of both theoretical and practical significance.

In the rest of this section we use the results of Section 2 and Section 3 to derive such bounds. Not surprisingly, they are in terms of m; the larger m, the larger the multiplicity.

To illustrate our technique we first consider (4.7). We write it as an SDP

$$\begin{array}{ll} min & z \\ s.t. & zI \succeq A_0 + y_1 A_1 + ... + y_m A_m \end{array} \tag{4.9}$$

Let us denote by $A(y) = A_0 + y_1 A_1 + ... + y_m A_m$, $S(z, y) = zI - A(y)$. Clearly, for y fixed, the optimal z is equal to $\lambda_1(A(y))$.

Theorem 4.1 *Let (z^*, y^*) be an extreme point optimal solution of (4.9). Then there exists $l \in \mathcal{N}$ s.t. $d_l \leq d_n - (m+1)$ and the multiplicity of $\lambda_1(A(y^*))$ is at least $n - l$.*

Proof Let (z^*, y^*) be an extreme point optimal solution of (4.9). Write $A(y^*) = Q\Lambda Q^T$ with Q an n by n orthogonal matrix, $\Lambda = \text{diag } \lambda,\ \lambda_1 \geq \ldots \geq \lambda_n$. Then $z^* = \lambda_1$ and $S(z^*, y^*) = Q(\lambda_1 I - \Lambda)Q^T$. By Theorem 3.2 there exists $l \in \mathcal{N}$ s.t. $d_l \leq d_n - (m+1)$, and rank $S(z^*, y^*) \leq l$. Thus the multiplicity of λ_1 must be at least $n - l$. □

Of course, for Theorem 4.1 to make sense, m must be large enough. We have the following simple

Corollary 4.2 *Assume $m \geq n$, and let (z^*, y^*) be an extreme point optimal solution of (4.9). Then the multiplicity of $\lambda_1(A(y^*))$ is at least 2.*

Proof Immediate, by observing that $m \geq n$ implies $d_{n-1} > d_n - (m+1)$, thus l in Theorem 4.1 can be at most $n-2$. □

Now consider our problem for arbitrary $k \geq 1$. (4.8) can then be formulated as the SDP

$$\begin{array}{ll} min & kz + I \circ V \\ s.t. & V \succeq 0 \\ & zI + V \succeq A_0 + y_1 A_1 + \ldots + y_m A_m \end{array} \tag{4.10}$$

This formulation is known; see e.g [Ali92]. For our purposes we also need the optimal (z, V) pair, once y is given. In Appendix A, we derive 4.10 from scratch; for fixed y the optimal (z, V) are given by Theorem A.3. Let us denote $A(y) = A_0 + y_1 A_1 + \ldots + y_m A_m$ (as previously) $S(z, y, V) = zI + V - A(y)$.

Theorem 4.3 *Let (z^*, y^*, V^*) be an extreme point optimal solution of (4.10). Then there exists $l, j \in \mathcal{N}$ s.t. $l \leq k,\ j \leq n-k$, $d_l + d_j \leq d_n - (m+1)$ and the multiplicity of $\lambda_k(A(y^*))$ is at least $n - (l+j)$.*

Proof

Write $A(y^*) = Q\Lambda Q^T$ with Q an n by n orthogonal matrix, $\Lambda = \text{diag } \lambda,\ \lambda_1 \geq \ldots \geq \lambda_n$. Let (z^*, V^*) be determined by Theorem A.3 in Appendix A, i.e. $\lambda_k \leq z^* \leq \lambda_{k+1}$, $V = Q\Omega Q^T$, with $\Omega = \text{diag } \omega\ \omega_1 = \lambda_1 - z^*$, ..., $\omega_k = \lambda_k - z^*$, $\omega_{k+1} = \ldots = \omega_n = 0$. Thus $S(z^*, y^*, V^*) = Q(z^* I + \Omega - \Lambda)Q^T$, and the eigenvalues of $z^* I$, V^*, $A(y^*)$ and $S(z^*, y^*, V^*)$ are shown below:

	z^*I	V^*	$A(y^*)$	$S(z^*, y^*, V^*)$
1.	z^*	$\lambda_1 - z^*$	λ_1	0
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
k.	z^*	$\lambda_k - z^*$	λ_k	0
$(k+1)$.	z^*	0	λ_{k+1}	$z^* - \lambda_{k+1}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n.	z^*	0	λ_n	$z^* - \lambda_n$

On the other hand, the feasible set of (4.10) can be written as

$$\{(z, y, V) : V \succeq 0, S(z, y, V) \succeq 0\}$$

By Lemma 2.4 and Theorem 2.1

$$\begin{aligned} (z^*, y^*, V^*) &= \{(z, y, V) : V \in LL^T, S(z, y, V) \in JJ^T\} \\ \text{aff } (z^*, y^*, V^*) &= \{(z, y, V) : V \in \text{lin } LL^T, S(z, y, V) \in \text{lin } JJ^T\} \end{aligned}$$

for some subspaces L, J of $\mathcal{R}^n$. Let $l = \dim L$, $j = \dim J$. In the expression for aff (z^*, y^*, V^*) the constraint $V \in \text{lin } LL^T$ imposes $d_n - d_l$ independent constraints on V^*, while $S(z^*, y^*, V^*) \in \text{lin } JJ^T$ imposes $d_n - d_j$ independent constraints on $S(z^*, y^*, V^*)$. Since $\dim(\text{aff } (z^*, y^*, V^*)) = 0$, we must have $(d_n - d_l) + (d_n - d_j) \geq d_n + (m+1)$, i.e. $d_l + d_j \leq d_n - (m+1)$. By the definition of z^* and V^* $l \leq k$, $j \leq n - k$.

Now, rank $V^* \leq l$ implies that at least $k - l$ eigenvalues of $A(y^*)$ above λ_k are equal to z^*. Also, rank $S(z^*, y^*, V^*) \leq j$ implies that at least $n - k - j$ eigenvalues of $A(y^*)$ below λ_{k+1} are equal to z^*. Thus the multiplicity of $\lambda_k(A(y^*)$ is at least $n - (l + j)$, as required. □

Of course, to have the an optimal multiplicity at least 2, m must be large enough. The companion of Corollary 4.2 is

Corollary 4.4 *Assume $m \geq \max\{k(n-k+1), (k+1)(n-k)\}$, and let (z^*, y^*, V^*) be an extreme point optimal solution of* (4.10) *. Then the multiplicity of $\lambda_k(A(y^*))$ is at least 2.*

Proof Let l and j satisfy $l \leq k$, $j \leq n - k$, $d_l + d_j \leq d_n - (m+1)$. We show $l + j \leq n - 2$. It is easy to see that the maximum of $l + j$ subject to the above constraints is attained when both l and j are as close to their upper bounds as possible. It is straightforward to check that $d_{k-1} + d_{n-k} > d_n - (m+1)$ and $d_k + d_{n-k-1} > d_n - (m+1)$,thus our claim follows. $\square$

Remark 4.5 When we minimize the weighted and absolute value sum of eigenvalues, we can derive bounds on the multiplicities in a similar manner.

A Appendix: a "dual" SDP for the sum of the k largest eigenvalues of a symmetric matrix

Let A be an n by n matrix. In this section we consider the SDP

$$\begin{array}{ll} min & kz + I \circ V \\ s.t. & V \succeq 0 \\ & zI + V \succeq A \end{array} \tag{A.11}$$

It is known (see e.g. [Ali92], [Ove92]), that the optimal value of (A.11) is $\sum_{i=1}^{k} \lambda_i(A)$. Here we derive this result from scratch with the purpose of determining the optimal (z, V) for given A. Our treatment parallels the one if [OW93]; in Lemma A.1 we also use their notation.

We proceed as follows: first we determine the the optimal solution for $A = \Lambda$ diagonal with the restriction that $V = \text{diag } \omega$ for some $\omega \in \mathcal{R}^n$. Next we drop the diagonality restriction on V. Finally, we determine the optimal solution for arbitrary A and V.

Lemma A.1 *Assume $\lambda \in \mathcal{R}^n$, $\lambda_1 \geq ... \geq \lambda_n$.*

Consider the LP

$$\begin{array}{ll} min & kz + e\omega \\ s.t. & \omega \geq 0 \\ & ze + \omega \geq \lambda \end{array} \tag{A.12}$$

The optimal solution of (A.12) is given by choosing z between λ_k and λ_{k+1} and setting $\omega_1 = \lambda_1 - z, ..., \omega_k = \lambda_k - z, \omega_{k+1} = ... = \omega_n = 0$.

Proof Assume $\lambda_1 \geq ... \geq \lambda_r > \lambda_{r+1} = ... = \lambda_k = ... = \lambda_{r+t} > \lambda_{r+t+1} \geq ... \geq \lambda_n$. Consider the dual of (A.12)

$$\begin{array}{ll} max & \lambda u \\ s.t. & e \geq u \geq 0 \\ & eu = k \end{array} \tag{A.13}$$

It is easy to see, and it is also in Lemma 3.2 in [OW93], that the optimal solution for (A.13) is given by

$$\begin{array}{ll} u_i = 1 & i = 1, ..., r \\ 0 \le u_i \le 1 & i = r+1, ..., r+t \\ u_i = 0 & i = r+t+1, ..., n \\ \sum_{i=r+1}^{r+t} u_i = k - r & \end{array} \tag{A.14}$$

Any optimal (z, ω) must be feasible for (A.12) and complementary to the u given by (A.14). It is straightforward to check that the only (z, ω) pairs satisfying these conditions are the ones given above. □

Lemma A.2 *Let λ be as in Lemma A.1, $\Lambda = \mathrm{diag}\ \lambda$. Consider the SDP*

$$\begin{array}{ll} min & kz + I \circ V \\ s.t. & V \succeq 0 \\ & zI + V \succeq \Lambda \end{array} \tag{A.15}$$

The optimal value of (A.15) is $\sum_{i=1}^{k} \lambda_i$. (z, V) are optimal if and only if $V = \mathrm{diag}\ \omega$, and z and ω are chosen as in Lemma (A.1).

Proof Replacing the constraints $V \succeq 0$ and $zI + V \succeq \Lambda$ with $v_{ii} \ge 0$ $(i = 1, ..., n)$ and $z + v_{ii} \ge \lambda_i$ $(i = 1, ..., n)$ yields (A.12), a relaxation of (A.15). Thus A.15 has optimal value at least $\sum_{i=1}^{k} \lambda_i$. This value is attained for a diagonal V.

Now, we must also show that *any* optimal V is diagonal. Let (z, V) be an optimal solution of (A.15). Therefore $(z, v_{11}, ..., v_{nn})$ must be an optimal solution of (A.12). Hence by Lemma A.1 $\lambda_{k+1} \le z \le \lambda_k$, $v_{11} = \lambda_1 - z, ..., v_{kk} = \lambda_k - z, v_{k+1,k+1} = ... = v_{nn} = 0$. Let $S = zI + V - \Lambda$, and partition the matrices V and S as

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{bmatrix}$$

and

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix}$$

where the blocks are k by k and $n-k$ by $n-k$, respectively. Since the diagonal elements of V_{22} are 0, and $V \succeq 0$, we must have $V_{22} = 0$ and $V_{12} = 0$. Since the diagonal elements of S_{11} are 0, and $S_{11} \succeq 0$, V_{11} must be diagonal. □

Theorem A.3 *Consider the SDP (A.11). Write $A = Q\Lambda Q^T$ with Q an n by n orthogonal matrix, $\Lambda = \text{diag } \lambda, \ \ \lambda_1 \geq ... \geq \lambda_n$. Then the optimal z and V are given by choosing z between λ_k and λ_{k+1}, and setting $V = Q\Omega Q^T$, with $\Omega = \text{diag } \omega$, $\omega_1 = \lambda_1 - z, ..., \omega_k = \lambda_k - z$, $\omega_{k+1} = ... = \omega_n = 0$. Thus the optimal solution of (A.11) is unique up to the choice of the eigenvector-basis of A.*

Proof Given the above decomposition of A we can rescale (A.11) to get (A.15) without changing its objective value. The correspondence between the solutions of the rescaled and the original problem is: (z, V) is an optimal solution to the rescaled problem if and only if (z, QVQ^T) is an optimal solution to (A.15). Our theorem then follows from Lemma A.2. □

Remark A.4 In [OW93] Overton and Womersley use the characterization of optimal solutions for the dual of (A.11) to get a description for the generalized gradient of the function $f_k(x) = \sum_{i=1}^{k} \lambda_i(A(x))$, where $A(x)$ is a smooth matrix-valued function. They utilize this result then to develop an algorithm to minimize $f_k(x)$. One might develop a dual version of their algorithm, based on the results of the Appendix. The same applies for minimizing weighted and absolute value sums of eigenvalues.

B Appendix: dimension and orthogonal generators of lin LL^T

In this section we present Theorem B.1 which derives the dimension of LL^T by exhibiting a set of orthogonal generators of lin LL^T. Proving the dimension of LL^T could be done more easily; we believe however, that our stronger result could be of independent interest. We use the following notation: for $u, v \in \mathcal{R}^n$ (not necessarily distinct) vectors, we denote by $M(u, v) = uv^T + vu^T$.

Theorem B.1 *Let L and J be complementary linear subspaces of $\mathcal{R}^n$, $L = \text{lin } \{u_1, ..., u_l\}$, $J = \text{lin } \{u_{l+1}, ..., u_n\}$ where $u_1, ..., u_n$ are pairwise orthogonal vectors. Then*

$$\mathcal{S}^{n\times n} = \text{lin } LL^T \oplus \text{lin } JJ^T \oplus R$$

where

1. lin LL^T *is generated by $M(u_i, u_j)$ for $i, j = 1, ..., l$*
2. lin JJ^T *is generated by $M(u_i, u_j)$ for $i, j = l + 1, ..., n$*
3. *R is generated by $M(u_i, u_j)$ for $i = 1, ..., l; j = l + 1, ..., n$*

thus

$$\dim LL^T = d_l,\ \dim JJ^T = d_{n-l}\ and \dim R = l(n-l)$$

In order to prove our Theorem we need several simple claims.

Claim B.2 *Let $u, v, r, s \in \mathcal{R}^n$. Then*

$$M(u,v) \circ M(r,s) = 2[(s^T u)(v^T r) + (s^T v)(u^T r)]$$

Proof Straightforward by the definition of $M(u,v)$. □

Claim B.3 *Let $u, v \in \mathcal{R}^n$. Then the eigenvectors of $M(u,v)$ are $u+v$ and $u-v$.*

Proof Since $M(u,v)$ has rank at most 2, it can have at most 2 nonzero eigenvectors. The fact that $u+v$ and $u-v$ are indeed eigenvectors is straightforward to check. □

Proof of Theorem B.1 By Claim B.3

$$\begin{aligned} M(u_i,u_j) &\in \text{lin } LL^T \text{ if } i \le l,\ j \le l \\ M(u_i,u_j) &\in \text{lin } JJ^T \text{ if } i > l,\ j > l \\ M(u_i,u_j) &\in R \text{ if } i \le l,\ j > l \end{aligned}$$

By Claim B.2 it follows that the matrices $M(u_i,u_j)$ are pairwise orthogonal, thus independent, hence they span $\mathcal{S}^{n\times n}$. Since the above 3 subspaces of $\mathcal{S}^{n\times n}$ are orthogonal, we must have

$$\begin{aligned} \text{lin } \{M(u_i,u_j) : i \le l,\ j \le l\} &= \text{lin } LL^T \\ \text{lin } \{M(u_i,u_j) : i > l,\ j > l\} &= \text{lin } JJ^T \\ \text{lin } \{M(u_i,u_j) : i \le l,\ j > l\} &= R \end{aligned}$$

and the expressions for the dimensions follow by simply counting the generators. □

References

[Ali92] F. Alizadeh: Combinatorial Optimization with Semi-Definite Matrices. In E. Balas, G. Cornuejols and R. Kannan, editors: *Second Conference on Integer Programming and Combinatorial Optimization*, Carnegie-Mellon University Press, 1992.

[BC75] G. P. Barker and D. Carlson: Cones of Diagonally Dominant Matrices, *Pacific Journal of Mathematics*, 57:1, 15-32, 1975

[BICK69] A. Ben-Israel, A. Charnes and K. Kortanek: Duality and Asymptotic Solvability over Cones, *Bulletin of American Mathematical Society*, 75(2), 318-324, 1969

[BW81a] J. Borwein and H. Wolkowicz: Characterization of optimality for the abstract convex program with finite dimensional range, *J. Austral. Math. Soc. series A*, 30, 390-411, 1981

[BW81b] J. Borwein and H. Wolkowicz: Facial reduction for a cone-convex programming problem, *J. Austral. Math. Soc. series A*, 30:369-380, 1981

[BW81c] J. Borwein and H. Wolkowicz: Regularizing the Abstract Convex Program, *Journal of Mathematical Analysis and Applications*, Vol. 83, No. 2, 1981

[CDW75] J. Cullum, W.E. Donath and P. Wolfe: The minimization of certain nondifferentiable sums of eigenvalue problems, *Mathematical Programming Study*, 3:35-55, 1975

[Fle85] R. Fletcher: Semi-Definite Matrix Constraints in Optimization, *SIAM J. Control and Optimization*, 23(4):493-513

[GLS88] M. Grötschel, L. Lovász and A. Schrijver: Geometric Algorithms and Combinatorial Optimization, Springer-Verlag, 1988

[Ove92] M. L. Overton: Large-scale Optimization of Eigenvalues, *SIAM J. Optimization*, 2(1):88-120

[OW93] M. L. Overton and R. S. Womersley: Optimality Conditions and Duality Theory for Minimizing Sums of the Largest Eigenvalues of Symmetric Matrices, *Mathematical Programming, Series B*, Vol. 62 (1993), No. 2., 321-357

[Roc70] Rockafellar, R. T.: Convex Analysis, Princeton University Press, Princeton, N.J., 1970

[RVW93] F. Rendl, R. J. Vanderbei and H. Wolkowicz: Max-Min Eigenvalue Problems, Primal-Dual Interior Point Algorithms and Trust Region Subproblems, Technical Report CORR 93-30, University of Waterloo

[Wol81] H. Wolkowicz: Some applications of optimization in matrix theory, *Linear Algebra and its Applications*, 40:101-118, 1981