Semidefinite and Second Order Cone Programming Seminar
Fall 2012
Lecture 8

Instructor: Farid Alizadeh
Scribe: Ai Kagawa

11/05/2012

1 Overview

This lecture covers transformations of Linear Programming (LP), Semi-definite Programming (SDP), and Second Order Cone Programming (SOCP), algebraic properties of SOCP, analysis of eigenvalue decomposition, and applications of eigenvalue optimization problem.

2 Transformations and Properties of SDP, LP, and SOCP

In our analysis of interior point methods for SDP in the previous lectures, there were several tools that were used effectively. Our goal here is to underscore the fact these features have analogs in SOCP. They can also be generalized to more other algebraic structures. Below we list some of these features for SDP and then construct their analogs for LP and SOCP.

scribe: Ai Kagawa
Lecture 8 **Date:** 01/17/2001

2.1 SDP

In SDP, we used the following transformations.

$$\begin{aligned}
X &\rightarrow \ln\det(X) = \sum \ln \lambda_i(\mathbf{x}) \\
Q_Y : X &\rightarrow YXY \\
X \circ S &\rightarrow \frac{XS + SX}{2} \quad \text{Note that this is commutative, but not associative} \\
\|X\|_2 &= \max_i |\lambda_i(X)| \\
\|X\|_F &= \sqrt{\sum \lambda_i^2(X)} = \sqrt{\sum_{i,j} X_{ij}^2}
\end{aligned}$$

2.2 LP

We can think of linear programming as a special case of semidefinite programming where matrices are 1×1. In this case the $\circ$ operation coincided with the usual real number multiplication: Thus it is both associative and commutative.

$$\begin{aligned}
\mathbf{x} &\rightarrow \sum \ln x_i \\
Q_{\mathbf{y}} : \mathbf{x} &\rightarrow \begin{pmatrix} y_1 x_1 y_1 \\ \vdots \\ y_n x_n y_n \end{pmatrix} = \begin{pmatrix} y_1^2 x_1 \\ \vdots \\ y_n^2 x_n \end{pmatrix} \\
\mathbf{x}_0 \circ \mathbf{x} &\rightarrow \mathrm{Diag}(\mathbf{x}_0)\mathbf{x} = \begin{pmatrix} (\mathbf{x}_0)_1 x_1 \\ \vdots \\ (\mathbf{x}_0)_n x_n \end{pmatrix} \\
\|\mathbf{x}\| &= \sqrt{x_1^2 + \ldots x_n^2} \\
\|\mathbf{x}\|_\infty &= \max_i |x_i|
\end{aligned}$$

2.3 Algebraic Properties of SOCP and the Lorentz cone

For the second order cone, we have already seen that the binary operation

$$\mathbf{x} \circ \mathbf{y} = \begin{pmatrix} x_0 y_0 + \cdots + x_n y_n \\ \vdots \\ x_0 y_i + x_i y_0 \\ \vdots \\ x_0 y_n + x_n y_0 \end{pmatrix} = \mathrm{Arw}(\mathbf{x})\mathbf{y}$$

where

$$\mathrm{Arw}(\mathbf{x}) = \begin{pmatrix} x_0 & x_1 & \cdots & x_n \\ x_1 & x_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \cdots & x_0 \end{pmatrix}$$

The algebra $(\mathbb{R}^{n+1}, \circ)$ is sometimes called the *spin factor algebra* due to connection to physics.

But how do we define the analog of the operation $X \to YXY$ in symmetric matrices for the spin factor algebra? The key is to try to write XYX purely in terms of $X \circ Y = \frac{XY+YX}{2}$. Here is how:

$$\begin{aligned} XYX &= X\frac{XY+YX}{2} + \frac{XY+YX}{2}X - \frac{X^2Y - YX^2}{2} \\ &= 2X \circ (X \circ Y) - X^2 \circ Y \end{aligned}$$

Note that we completely removed ordinary multiplication and replaced it with the "$\circ$" operation. We can now do the same for the spin factor algebra:

Definition 1

$$Q_{\mathbf{x}} : \mathbf{y} \to 2\mathbf{x} \circ (\mathbf{x} \circ \mathbf{y}) - \mathbf{x}^2 \circ \mathbf{y}$$

Thus, as a matrix,

$$\begin{aligned} Q_{\mathbf{x}}(\mathbf{y}) &= 2\,\mathrm{Arw}(\mathbf{x})(\mathrm{Arw}(\mathbf{x})\mathbf{y}) - \mathrm{Arw}(\mathbf{x}^2)\mathbf{y} \\ &= \big(2\,\mathrm{Arw}^2(\mathbf{x}) - \mathrm{Arw}(\mathbf{x}^2)\big)\mathbf{y} \\ &\quad \text{Therefor: } Q_{\mathbf{x}} = 2\,\mathrm{Arw}^2(x) - \mathrm{Arw}(\mathbf{x}^2). \end{aligned}$$

Just as in the case of symmetric matrices, where the operation $Q_X : Y \to XYX$ maps the cone of positive semidefinite matrices to itself (that is, it is an *automorphism* of this cone), so is the operation $Q_{\mathbf{x}}$ an automorphism of the Lorentz cone.

In our analysis of potential reduction interior point methods, we used the transformation $Q_{X_0^{1/2}}(X) = X_0^{1/2} X X_0^{1/2}$. Remember that this transformation maps X_0 to the identity and the whole positive semidefinite cone to itself. To extend this to the Lorentz cone we must first define what $\mathbf{x}^{1/2}$ is in the spin factor algebra.

Before we do so we remark that it is possible to extend these notions to more abstract algebraic structure. By a *Euclidean Jordan Algebra*, we mean a vector space $\mathbb{J}$ (assume an n-dimensional vector space) with a binary operation $\circ$ satisfying the distributive law over addition, and

1. $\mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$
2. $\mathbf{x}^2 \circ (\mathbf{x} \circ \mathbf{y}) = \mathbf{x} \circ (\mathbf{x}^2 \circ \mathbf{y})$

3. $\sum x_i^2 = 0 \Rightarrow \mathbf{x}_i = 0$
4. $\exists$ an identity $\boldsymbol{e}$: $\mathbf{x} \circ \boldsymbol{e} = \boldsymbol{e} \circ \mathbf{x}$.

Next, consider the cone of squares with respect to $\circ$.

$$K = \{\mathbf{x}^2 : \mathbf{x} \in \mathbb{J}\}$$

The cone of squares of a Euclidean Jordan algebra are sometimes called *symmetric cones*. It is known that they satisfy the following properties.

1. Self-duality: $K = K^*$
2. Homogeneity: For any pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in \text{Int}(K)$ there is a linear transformation A such that $A(\mathbf{x}_1) = \mathbf{x}_2$, and $A(K) = K$

The cone of positive semidefinite matrices and the Lorentz cone are examples of symmetric cones. It turns out that there are only five classes of *basic* symmetric cones, as there are only five basic classes of Euclidean Algebras. Any other symmetric cone is a direct sum of a combination of these five basic classes:

1. The algebra associated with SOCP (sometimes called the *spin factor algebra*) and its cone of squares the Lorentz cone.
2. The set of real symmetric matrices under the product $\frac{XY+YX}{2}$ and its cone of squares, the set of real positive semidefinite matrices.
3. The set of complex Hermitian matrices under the product $\frac{XY+YX}{2}$ and its cone of squares, the set of complex positive semidefinite matrices.
4. The set of quaternionic Hermitian matrices under the product $\frac{XY+YX}{2}$ and its cone of squares, the set of quaternionic positive semidefinite matrices.
5. The set of 3×3 octonion Hermitian matrices under $\frac{XY+YX}{2}$ (known as the *Albert algebra* and its cone of squares, and exceptional 27 dimensional cone.

3 Spectral properties of second order cones

The characteristic polynomial of X is $p(\lambda) = p_0 + p_1 t + \ldots + p_n t^n = \det(\lambda I - X)$. By the Cayley-Hamiltonian theorem, substituting X for λ, $p(X) = \mathbf{0}$. Also recall that for any polynomial $q(\lambda)$ of degree less than $\deg(p)$, if $q(X) = 0$, then $q(\lambda)$ divides $p(\lambda)$. The smallest degree polynomial $q(\lambda)$ such that $q(X) = 0$ is the *minimum polynomial* of X. For a symmetric matrix, one can easily characterize the minimum and the characteristic polynomials from its eigenvalues. Let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of X with multiplicities $m_1, \ldots, m_k$. Then

$$\begin{aligned} p(\lambda) &= (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k} \\ q(\lambda) &= (\lambda - \lambda_1) \cdots (\lambda - \lambda_k) \end{aligned}$$

Here is another way to think of the minimum polynomials. For a symmetric matrix X consider the set $\{I, X, X^2, \cdots\}$. Since we are in a finite dimensional linear space, sooner or later this set becomes linearly dependent. Let k be the smallest integer such that $I, X, X^2, \ldots, X^k$ is linearly dependent. Then there are real numbers $p_0, p_1, \ldots, p_k$ dependent on X such that

$$p_0 + p_1 X + \cdots + p_k X^k = 0$$

The polynomial $p(\lambda)$ is the minimum polynomial.

With this approach we can now define minimum polynomials for the elements of spin factor algebra (or any algebra that is *power associative*).

In fact, consider the polynomial

$$p(\lambda) = \lambda^2 - 2x_0\lambda + (x_0^2 - \|\bar{x}\|^2) = \big(\lambda - (x_0 - \|\bar{\mathbf{x}}\|)\big)\big(\lambda - (x_0 + \|\bar{\mathbf{x}}\|)\big).$$

Then it is very easy to verify that $p(\mathbf{x}) = \mathbf{0}$.

We now define concepts analogous to matrices:

Definition 2 *Given $\mathbf{x}$ in the spin factor algebra $(\mathbb{R}^{n+1}, \circ)$ and the quadratic polynomial $p(\lambda)$ as defined above,*

1. *the numbers $\lambda_1 = x_0 + \|\bar{\mathbf{x}}\|$ and $\lambda_2 = x_0 + \|\bar{\mathbf{x}}\|$ are the* eigenvalues *of $\mathbf{x}$.*
2. *The* trace*: $\operatorname{tr}(\mathbf{x}) = \lambda_1 + \lambda_2 = 2x_0$ and the determinant $\det(\mathbf{x}) = \lambda_1\lambda_2 = x_0^2 - \|\bar{\mathbf{x}}\|^2)$. Thus the barrier $-\ln(x_0^2 - \|\bar{\mathbf{x}}\|^2) = -\ln\det\mathbf{x}$.*
3. *The analog of max-norm is $\max\{\lambda_1, \lambda_2\} = \max\{x_0 + \|\bar{\mathbf{x}}\|, x_0 - \|\bar{\mathbf{x}}\|\}$.*

We may also define the following norms.

$$\begin{aligned} \|\mathbf{x}\|_F &= \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{2}\|\mathbf{x}\| \\ \|\mathbf{x}\|_2 &= \max\{x_0 + \|\bar{\mathbf{x}}\|, x_0 - \|\bar{\mathbf{x}}\|\} = x_0 + \|\bar{\mathbf{x}}\| \ \ \text{if } x \in Q \end{aligned}$$

For symmetric matrices there is a convenient way to expand the definition of a real valued continuous function $f(t)$ to a symmetric matrix: $f(X)$. Let the spectral decomposition of X be given by

$$X = Q^T \Lambda Q$$

where $\Lambda = \operatorname{Diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues of X, and Q is the orthogonal matrix whose i^{th} column is the eigenvector corresponding to λ_i. Then we can define $f(X)$ as follows:

$$f(X) \stackrel{\text{def}}{=} Q^T \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} Q$$

For instance

$$\sqrt{X} = Q^T \operatorname{Diag}\big(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\big) Q$$

Of course, the functions $f(X)$ is well-defined only if $f(\lambda_i)$ is well defined over the domain of f.

Note that the function $\sqrt{X}$ is defined only for positive semidefinite matrices and, itself is the unique positive semidefinite matrix whose square is X: $\sqrt{X}\sqrt{X} = X$, as expected.

3.1 Analysis of Eigenvalue Decomposition

To extend the notion of $\sqrt{\cdot}$, or in fact any function, we need to extend the notion of "eigenvalue decomposition" to the spin factor algebra. To do so we attempt to extract the essential properties of $Q\Lambda Q^\top$ decomposition for matrices.

$$\begin{aligned} X &= Q\Lambda Q^T \\ \text{where } Q &= [\mathbf{q}_1, \ldots \mathbf{q}_n] \\ &\quad \text{and each } \mathbf{q}_i \text{ is a column vector.} \end{aligned}$$

Since Q is an orthogonal matrix, $\|\mathbf{q}_i\| = 1$ and $\mathbf{q}_i^T \mathbf{q}_j = 0$ (each column vector is orthogonal to each other).

$$\begin{aligned} X &= [\mathbf{q}_1, \ldots \mathbf{q}_n] \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix} \\ &= \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \ldots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T \end{aligned}$$

There are three properties for $\mathbf{q}_i \mathbf{q}_i^T$

i. $\mathbf{q}_i \mathbf{q}_i^\top$ is an idempotent: $Q_i^2 = Q_i$, since $\mathbf{q}_i^T \mathbf{q}_i = 1$, $(\mathbf{q}_i \mathbf{q}_i^T)^2 = \mathbf{q}_i \mathbf{q}_i^T \mathbf{q}_i \mathbf{q}_i^T = \mathbf{q}_i \mathbf{q}_i^T$

ii. Mutual Orthogonality:

$$Q_i \circ Q_j = \frac{\mathbf{q}_i \mathbf{q}_i^T \mathbf{q}_j \mathbf{q}_j^T + \mathbf{q}_i \mathbf{q}_i^T \mathbf{q}_j \mathbf{q}_j^T}{2} = 0$$

since $\mathbf{q}_i^T \mathbf{q}_j = 0$

iii. sum to identity:

$$\sum_i Q_i = QQ^T = I$$

Definition 3 *A set of symmetric matrices* $\{Q_1, \ldots Q_n\}$ *satisfying i, ii, and iii is called a Jordan Frame.*

It is now possible to extend this notion to Euclidean Jordan algebras, and in particular to the spin factor algebra.

scribe: Ai Kagawa
Lecture 8 **Date:** 01/17/2001

Definition 4 *A set of vectors* $\{\mathbf{c}_1, \ldots \mathbf{c}_m\}$ *is called a Jordan Frame if*

i. $\mathbf{c}_i^2 = \mathbf{c}_i$

ii. $\mathbf{c}_i \circ \mathbf{c}_j = 0$

iii. $\sum_i \mathbf{c}_i = \mathbf{e}$

In the case of the spin factor algebra, it turns out a Jordan frame contains only two elements. To see this Suppose

$$\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

and define

$$\mathbf{c}_1 = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} \quad \text{and} \quad \mathbf{c}_2 = \frac{1}{2}\begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}$$

Then it is easily verified that $\mathbf{c}_1, \mathbf{c}_2$ form a Jordan frame.

$$\mathbf{c}_i^2 = \frac{1}{4}\begin{pmatrix} 1 + \frac{\bar{x}^T\bar{x}}{\|\bar{x}\|}^2 \\ \frac{2x_i}{\|\bar{x}\|} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ \vdots \\ \frac{x_i}{\|\bar{x}\|} \end{pmatrix} = \mathbf{c}_i$$

Furthermore,

$$\mathbf{c}_1 \circ \mathbf{c}_2 = \frac{1}{4}\begin{pmatrix} 1 - \frac{\bar{x}^T\bar{x}}{\|\bar{x}\|} \\ \frac{1 \cdot x_i}{\|\bar{x}\|} - \frac{1 \cdot x_i}{\|\bar{x}\|} \end{pmatrix} = 0$$

$$\mathbf{c}_1 + \mathbf{c}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbf{e}$$

Finally, each vector $\mathbf{x}$ can be written as:

$$\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_0 + \|\bar{x}\|)\mathbf{c}_1 + (x_0 + \|\bar{x}\|)\mathbf{c}_2$$

The relation above is now defined as the *spectral decomposition* of $\mathbf{x}$. Based on this decomposition, we can now define

$$f(\mathbf{x}) = f(\lambda_1)\mathbf{c}_1 + f(\lambda_2)\mathbf{c}_2$$

In particular,

$$X^{\frac{1}{2}} = \sqrt{\lambda_1}\mathbf{c}_1 + \sqrt{\lambda_2}\mathbf{c}_2$$

Thus for the spin factor algebra we have all the ingredients necessary to extend, word-for-word, the analysis we gave for the potential reduction algorithm.

scribe: Ai Kagawa

4 Modeling problems as semidefinite and second order programs

4.1 Eigenvalue Optimization

Problem 1

For $A_0 + x_1A_1 + \ldots + x_nA_n$, consider the following problem where $\mathbf{x}$ is unknown.

$$\min_{x_1,\ldots,x_n} \lambda_{[1]}(A_0 + x_1A_1 + \ldots x_nA_n)$$

where in general, $\lambda_{[i]}(X)$ is the i^{th} largest eigenvalue of X. An alternative version is:

$$\begin{array}{ll} \min & \lambda_{[1]}(X) \\ \text{s.t.} & A_i \bullet X = b_i, \quad i = 1,\ldots,m \end{array}$$

The "largest eigenvalue" is a convex, but in general nonsmooth, function of its argument matrix. The common way of proving this using the Rayleigh-Ritz quotient characterization of eigenvalues of a symmetric matrix. However, we can prove this, and simultaneously give an SDP representation by observing that if $\lambda_{[1]}(A)$ is the largest eigenvalue of A then $\lambda_{[1]}I - A \succcurlyeq 0$, and it is the smallest real number with this property. Thus, the optimization problems above can be formulated as:

$$\begin{array}{ll} \min & z \\ \text{s.t.} & zI - \sum_i x_iA_i \succcurlyeq A_0 \end{array} \quad \text{and} \quad \begin{array}{ll} \min & z \\ \text{s.t.} & zI - X \succcurlyeq 0 \\ & A_i \bullet X = b_i \text{ for } i = 1,\ldots,m \end{array}$$

Let's consider the following (admittedly trivial) problem: find the largest element of the set $\{a_1,\ldots,a_n\}$, using linear programming.

$$\begin{array}{ll} \min & x \\ \text{s.t.} & x \geq a_i, \quad i = 1,\ldots,n \end{array} \quad \text{with dual} \quad \begin{array}{ll} \max & a_1y_1 + \cdots + a_ny_n \\ \text{s.t.} & y_1 + \cdots + y_n = 1 \\ & y_i \geq 0 \text{ for } i = 1,\ldots,n \end{array}$$

Note that the LP formulation is quite similar to the largest eigenvalue problem. In fact, if instead of a set of numbers, we have a set of affine functions, and define:

$$f(\mathbf{x}) = \min_i\{\mathbf{a}_1^\top\mathbf{x} + b_1,\ldots,\mathbf{a}_n^\top\mathbf{x} + b_n\}$$

Then it is easy to see that the problem $\min_{\mathbf{x}} f(\mathbf{x})$ can be formulated by the following primal-dual linear programs:

$$\begin{array}{ll} \min & z \\ \text{s.t.} & z\mathbf{1} - A\mathbf{x} \geq \mathbf{b} \end{array} \quad \text{and} \quad \begin{array}{ll} \max & \mathbf{b}^\top\mathbf{y} \\ \text{s.t.} & \mathbf{1}^\top\mathbf{y} = 1 \\ & A^\top\mathbf{y} = 0 \\ & \mathbf{y} \geq 0 \end{array}$$

Everything above can be turned upside-down, and we can consider the smallest eigenvalue of an affine symmetric matrix valued functions. This function is concave, and its *maximization* can be formulated as an SDP.

4.2 optimization of the largest or smallest k eigenvalues

We now focus on the eigenvalue optimization problem for the largest k eigenvalues:

$$\min_{\lambda} \ (\lambda_{[1]} + \lambda_{[2]} + \ldots + \lambda_{[k]})(A_0 + x_1 A_1 + \ldots x_n A_n)$$

(Note that the "sum of k largest eigenvalues" includes multiplicities. So, for instance if the largest eigenvalue has multiplicity two, then the sum of the two largest eigenvalues is twice the largest eigenvalue.)

It turns out that this problem also can be expressed as a semidefinite program. To see how, let us first get back to the linear programming counterpart of this problem. Consider the again a set of real numbers $\{a_1, \ldots, a_n\}$ and let $\text{kmax}_k\{a_1, \ldots, a_n\}$ denote the *sum of the* k *largest elements in this set.* How can we formulate this as a linear program? The integer programming formulation is easy:

$$\begin{array}{ll} \max & x_1 a_1 + \ldots + x_m a_m \\ \text{s.t.} & \sum_i x_i = k \\ & x_i \in \{0, 1\} \end{array} \tag{1}$$

But notice that the linear programming relaxation

$$\begin{array}{ll} \max & x_1 a_1 + \ldots + x_m a_m \\ \text{s.t.} & \sum x_i = k \\ & 0 \le x_i \le 1 \end{array} \quad \text{and its dual} \quad \begin{array}{ll} \min & kz + \sum_i y_i \\ \text{s.t.} & z + y_i \ge a_i \text{ for } i = 1, \ldots, m \\ & y_i \ge 0 \text{ for } i = 1, \ldots, m \end{array} \tag{2}$$

is actually equivalent to the integer program. For instance, it is clear that this problem's matrix is totally unimodular, therefore, all sub-determinants are equal to 0, 1 or -1. As a result, the set of optimal solutions always contains integer solutions, which in this case means that it contains zero-one solutions.

Looking at the pair of problems (4.2) and considering the complementary slackness, we see that at the optimum, we have

$$(1 - x_i) y_i = 0 \tag{3}$$

$$x_i (z + y_i - a_i) = 0 \tag{4}$$

which implies that if $x_i = 1$ and $y_i = a_i - z$ for a_i one of the largest k elements, and $x_i = y_i = 0$ for the rest. Also z can be any real number larger than or equal the $a_{[k+1]}$.

Now consider the function:

$$f_k(\mathbf{x}) = \max_k \{\mathbf{a}_1^\top \mathbf{x} + b_1, \ldots, \mathbf{a}_m^\top \mathbf{x} + b_m\}$$

How can we formulate this as a linear program? We cannot simply replace a_i with $\mathbf{a}_i^\top \mathbf{x} + b_i$ in the maximization problem above, since the resulting optimization problem would not be linear any more. However, we can do so in the dual, that is the minimization problem:

$$\begin{array}{ll} \min & kz + \sum_i y_i \\ \text{s.t.} & z + y_i - \mathbf{a}_i^\top \mathbf{x} \geq b_i \text{ for } i = 1, \ldots, m \\ & y_i \geq 0 \end{array}$$

and taking dual again:

$$\begin{array}{ll} \max & \sum b_j u_j \\ \text{s.t.} & \sum_i u_i = k \\ & \sum_j a_{ij} u_j = 0 \text{ for } i = 1, \ldots, m \\ & 0 \leq u_i \leq 1 \end{array}$$

Question: How do we extend this approach to calculating $(\lambda_{[1]} + \cdots + \lambda_{[k]})(A)$?

Considering the eigenvalue version of the problems (4.2), we have,

Lemma 5 *For a symmetric* $n \times n$ *matrix* A *we have*

$$\text{k}\max\{\lambda_1(A), \ldots, \lambda_k(A)\} = \begin{array}{ll} \max & A \bullet X \\ s.t. & \text{trace}\, X = k \\ & 0 \preccurlyeq X \preccurlyeq I \end{array} \quad \textit{and dual} \quad \begin{array}{ll} \min & kz + \text{trace}\, Y \\ s.t. & zI + Y \succcurlyeq A \\ & Y \succcurlyeq 0 \end{array} \tag{5}$$

Proof: To see this suppose the optimal solution of the maximization problem is X, and assume that X and A commute, that is $AX = XA$, and therefore, they share a common system of eigenvectors, say the columns of matrix Q. In other words, $QQ^\top = I$ and $A = Q\Lambda Q^\top$ and $X = Q\Omega A^\top$ with Λ and Ω diagonal matrices containing eigenvalues of A and X, respectively. Then in this case it is easily seen that the primal and dual SDP's in (5) are the same as (4.2), with a_i replaced by $\lambda_i(A)$ and x_i replaced by $\lambda_i(X)$.

So all that is left is to show is that at the optimum, the solution X of (5) indeed commutes with A. To see this note that by complementary slackness, at the optimum, we must have

$$(I - X)Y = 0$$
$$X(zI + Y - A) = 0$$

The first implies that Y and X commute, and the second implies that–since X and $zI - Y$ already commute–X and A commute[1]. ■

Now we can deal with the SDP formulation of the problem

$$\min_x \big(\lambda_{[1]} + \cdots + \lambda_{[k]}\big)(A_0 + x_1 A_1 + \cdots + x_m A_m).$$

$$\begin{array}{lll} \min & kz + I \bullet Y & \\ \text{s.t.} & zI + Y - \sum X_i A_i \succcurlyeq A_0 & \leftrightarrow X \\ & Y \succcurlyeq 0 & \end{array}$$

[1]In general if the product of two symmetric matrices A and B is symmetric, then A and B commute.

and its dual:

$$\begin{array}{ll} \max & A_0 \bullet X \\ \text{s.t.} & I \bullet X = k \\ & 0 \preccurlyeq X \preccurlyeq I \\ & A_i \bullet X = 0 \text{ for } i = 1, \ldots, \end{array}$$

References

[1] Alizadeh, F., Goldfarb, D., *Second-Order Cone Programming*, Mathematical Programming manuscript, 2010.