Semidefinite Programming
Lecture 4

scribe: Lorant Porkolab

February 20, 1995

Eigenvalues as Semidefinite Programs

In many cases semidefinite programs arise in the form of minimizing or maximizing an appropriate linear combination of eigenvalues of a symmetric matrix subject to linear constraints on the matrix. In this lecture we study problems of this form and the corresponding semidefinite programs.

1 Notations

Let X be an $n \times n$ symmetric matrix. Let $\lambda_i(X)$ and $\lambda^i(X)$ denote the i^{th} largest and the absolute-valu-wise i^{th} largest eigenvalue of X, respectively. Let $A, A_0, A_1, \ldots, A_m$ be $n \times n$ real, symmetric matrices. For any $\mathbf{x} \in R^m$, let $A(\mathbf{x})$ denote the symmetric matrix $A_0 + x_1A_1 + x_2A_2 + \ldots + x_mA_m$. Let $(\lambda_1 + \lambda_2 + \ldots + \lambda_k)(X)$ denote the sum of the first k largest eigenvalues of X.

Let $B = \{b_1, \ldots, b_m\}$, $B(x) = \{\mathbf{a}_1^T\mathbf{x} + b_1, \ldots, \mathbf{a}_m^T\mathbf{x} + b_m\}$, where all of the $\mathbf{a}_i$'s and b_i's are rational vectors in R^n and rational numbers. (So, B is just a finite set of numbers and $B(x)$ is a finite set of linear functions.) Let $f(B)$ be the largest element of B, and $f(\mathbf{x}) := f(B(\mathbf{x}))$ $\forall \mathbf{x} \in R^n$. Let $f_k(B)$ be the sum of the k largest elements of B, and $f_k(\mathbf{x}) := f_k(B(\mathbf{x}))$ $\forall \mathbf{x} \in R^n$. Finally, let $f_{|k|}(B)$ be the sum of the k absolute-value-wise largest elements of B, and $f_{|k|}(\mathbf{x}) := f_{|k|}(B(\mathbf{x}))$ $\forall \mathbf{x} \in R^n$.

2 Some Simple Problems and their LP Formulations

2.1 Problem 1: Find $f(B)$.

We can formulate this problem as a linear program:

$$
\begin{array}{lll}
(\mathbf{P_1}) & \min & z \\
 & \text{s.t.} & z \geq b_i, \quad i = 1, \ldots, m.
\end{array}
$$

The dual of $(\mathbf{P_1})$ is :

$$
\begin{array}{lll}
(\mathbf{D_1}) & \max & \mathbf{b}^T\mathbf{y} \\
 & \text{s.t.} & \mathbf{e}^T\mathbf{y} = 1 \\
 & & y_i \geq 0, \quad i = 1, \ldots, m.
\end{array}
$$

We can see that the set of the feasible solutions of $(\mathbf{D_1})$ is the standard simplex.

2.2 Problem 2: Find min $\{ f(\mathbf{x}) : \mathbf{x} \in R^n \}$.

This problem is a generalization of Problem 1, since we replace the constant function b_i by the linear function $\mathbf{a}_i^T\mathbf{x} + b_i$ for each $i = 1, \ldots, m$. Hence, we hope that we can get easily some LP formulations for it by simple modifications of those of Problem 1. However, we can see immediately that if we want to keep linearity in both the objective function and in the constraints, then we have to replace b_i by $\mathbf{a}_i^T\mathbf{x} + b_i$ in $(\mathbf{P_1})$. (The same replacement in $(\mathbf{D_1})$ would give us a quadratic objective function.)
So, we can formulate also this problem as a linear program:

$$
\begin{array}{lll}
(\mathbf{P_2}) & \min & z \\
 & \text{s.t.} & z - \mathbf{a}_i^T\mathbf{x} \geq b_i, \quad i = 1, \ldots, m.
\end{array}
$$

The dual of $(\mathbf{P_2})$ is :

$$
\begin{array}{lll}
(\mathbf{D_2}) & \max & \mathbf{b}^T\mathbf{y} \\
 & \text{s.t.} & \mathbf{e}^T\mathbf{y} = 1 \\
 & & A^T\mathbf{y} = 0 \\
 & & y_i \geq 0, \quad i = 1, \ldots, m,
\end{array}
$$

where A is the $m \times n$ matrix containing $\mathbf{a}_i^T$ as its i^{th} row.

FACT 2.1. $f(\mathbf{x})$ *is a convex, piecewise linear function.*

We will see and use also later the fact that at generalizations the corresponding replacement (e.g. replacement of a constant by an affine function or replacement of a constant matrix by some affine function of matrices) has to be done in those halves of the primal dual pairs in which the term(s) we want to replace appear in an isolated form. (E.g. b_i is isolated in $(\mathbf{P_1})$, but it is not isolated in $(\mathbf{D_1})$, so the replacement above had to be done in $(\mathbf{P_1})$.)

2.3 Problem 3: Find $f_k(B)$.

We can formulate this problem as an integer program:

$$
\begin{array}{lll}
(\mathbf{I_3}) & \max & \mathbf{b}^T\mathbf{y} \\
 & \text{s.t.} & \mathbf{e}^T\mathbf{y} = k \\
 & & y_i \in \{0,1\} \quad i = 1,\dots,m
\end{array}
$$

The LP relaxation of this 0-1 integer program is:

$$
\begin{array}{lll}
(\mathbf{D_3}) & \max & \mathbf{b}^T\mathbf{y} \\
 & \text{s.t.} & \mathbf{e}^T\mathbf{y} = k \\
 & & 0 \le y_i \le 1 \quad i = 1,\dots,m
\end{array}
$$

The following simple Lemma proves that this LP relaxation is exact, i.e. $(\mathbf{D_3})$ has the same optimum as $(\mathbf{I_3})$.

LEMMA 2.1. *The extreme points of*

$$S_k = \{\mathbf{x} \in R^n : \sum_{i=1}^{n} x_i = k, 0 \le x_i \le 1\}$$

are n-dimensional 0-1 vectors, i.e. S_k is an integer polytop.

Proof. It is easy to see that the constraint matrix of the obove LP is totally unimodular, hence all of the extreme points are integral. ■

The dual of $(\mathbf{D_3})$ is:

$$
\begin{array}{llll}
(\mathbf{P_3}) & \min & kz + \mathbf{e}^T\mathbf{v} & \\
 & \text{s.t.} & z + v_i \ge b_i, & i = 1,\dots,m, \\
 & & v_i \ge 0, & i = 1,\dots,m.
\end{array}
$$

2.4 Problem 4: Find min $\{\ f_k(\mathbf{x}) : \mathbf{x} \in R^n\ \}$.

We can formulate also this problem as a linear program:

$$
\begin{array}{llll}
(\mathbf{P_4}) & \min & kz + \mathbf{e}^T\mathbf{v} & \\
 & \text{s.t.} & z + v_i - \mathbf{a}_i^T\mathbf{x} \ge b_i, & i = 1,\dots,m, \\
 & & v_i \ge 0, & i = 1,\dots,m.
\end{array}
$$

The dual of $(\mathbf{P_4})$ is :

$$\begin{array}{lll} (\mathbf{D_4}) & \max & \mathbf{b}^T\mathbf{y} \\ & \text{s.t.} & \mathbf{e}^T\mathbf{y} = k, \\ & & A^T\mathbf{y} = 0, \\ & & 0 \leq y_i \leq 1, \quad i = 1, \ldots, m, \end{array}$$

where A is the $m \times n$ matrix containing $\mathbf{a}_i^T$ as its i^{th} row.

FACT 2.2. *$f_k(\mathbf{x})$ is a convex function.*

Later, we will see that also $(\lambda_1 + \ldots + \lambda_k)(A(\mathbf{x}))$ is a convex function and the problem of finding $\min\{(\lambda_1 + \ldots + \lambda_k)(A(\mathbf{x})) : \mathbf{x} \in R^n\}$ can be formulated as a convex program, in particular as a Semidefinite Program (SDP). To show this, we need the matrix analog of the above Lemma.

3 Some Eigenvalue Problems and their SDP Formulations

3.1 Problem 5: Find $(\lambda_1 + \ldots + \lambda_k)(A)$.

THEOREM 3.1. *(The matrix analog of Lemma 1.1) Let*

$$\begin{array}{l} S_1 = \{QQ^T : Q \in R^{n \times k}, Q^TQ = I_k\}, \\ S_2 = \{Y \in R^{n \times n} : Y = Y^T, tr(Y) = k, 0 \preceq Y \preceq I\}. \end{array}$$

Then: $Conv(S_1) = S_2$ and $Ext(S_2) = S_1$.

Proof. 1.1: We will show that $Conv(S_1) \subseteq S_2$. Since S_2 is obviously convex, it suffices to prove that $S_1 \subseteq S_2$. Let QQ^T be an element of S_1. Then it is easy to see that QQ^T is symmetric, $0 \preceq QQ^T$ and $rank(QQ^T) = k$. On the other hand $tr(QQ^T) = tr(Q^TQ) = tr(I_k) = k$, where the first equation follows from the following:

$$tr(Q^TQ) = \sum_{j=1}^{k}(Q^TQ)_{jj} = \sum_{j=1}^{k}\sum_{i=1}^{n} Q_{ij}Q_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{k} Q_{ij}Q_{ij} = \sum_{i=1}^{n}(QQ^T)_{ii} = tr(QQ^T).$$

(Or it follows also from the fact that the set of nonzero eigenvalues of AB equal with set of nonzero eigenvalues of BA for for all A, B for which the multiplications are well defined.)

1.2: We will show that $S_2 \subseteq Conv(S_1)$. Let Y be a matrix in S_2. Then

$$\begin{array}{lll} Y & = QDiag(\mathbf{y})Q^T & \\ & = QDiag(\sum_{i=1}^{M} \alpha_i \mathbf{u}^i)Q^T & \text{(by Lemma 2.1)} \\ & = \sum_{i=1}^{M} \alpha_i QDiag(\mathbf{u}^i)Q^T & \text{(with } \alpha_i \geq 0, \sum_{i=1}^{M} \alpha_i = 1) \\ & = \sum_{i=1}^{M} \alpha_i Q_iQ_i^T \in Conv(S_1), & \end{array}$$

where $\mathbf{y} = (y_1, \ldots, y_m)$ is an $m-$dimensional vector containing the positive eigenvalues of Y, the $\mathbf{u}^i$'s are $m-$dimensional 0-1 vectors such that $\sum_{j=1}^m u_j^i = k$, and the last containment follows from the fact that $Q_i Q_i^T \in S_1$.

2: We will show that the extreme points of S_2 are exactly the matrices in S_1. To do this, first we can observe that no extreme point of S_2 can have rank larger than k. Otherwise Y could be written in the form of $QDiag(y_1, \ldots, y_m)Q^T$ with $m > k, y_i > 0, \sum_{i=1}^m y_i = k$. Hence, there would be y_i such that $0 < y_i < 1$. So $\mathbf{y}$ could be written as a nontrivial convex combination of some $\mathbf{u}^j$'s, which would imply that also Y can be written as some nontrivial convex combination of some $Q_j Q_j^T$'s, i.e. Y can not be an extreme point of S_2. On the other hand, it is obvious that no matrix in S_2 can have rank less than k.

Finally, if Y is a rank k matrix in S_2 then it is of the form QQ^T, where $Q \in R^{n \times k}, Q^T Q = I_k$, since Y has two eigenvalues, 1 with multiplicity k and 0 with multiplicity n-k, so it can be written in the above form. Thus, we have proved that all of the extreme points of S_2 are in S_1.

By using appropriate rotation matrices and referring to symmetry reasons, we can prove the converse too, i.e. $S_1 \subseteq Ext(S_2)$. ■

THEOREM 3.2.

$$(\mathbf{D_5}) \qquad (\lambda_1 + \ldots + \lambda_k)(A) = \max\{A \bullet Y : tr(Y) = k \text{ } and \text{ } 0 \preceq Y \preceq I\}.$$

Proof. We know from linear algebra that:

$$\begin{aligned}
\lambda_1(A) &= \max\{\mathbf{q}_1^T A \mathbf{q}_1 : \mathbf{q}_1^T \mathbf{q}_1 = 1\}, \\
\lambda_2(A) &= \max\{\mathbf{q}_2^T A \mathbf{q}_2 : \mathbf{q}_2^T \mathbf{q}_2 = 1, \mathbf{q}_2^T \mathbf{q}_1^* = 0\}, \\
&\vdots \\
\lambda_k(A) &= \max\{\mathbf{q}_k^T A \mathbf{q}_k : \mathbf{q}_k^T \mathbf{q}_k = 1, \mathbf{q}_k^T \mathbf{q}_1^* = 0, \ldots, \mathbf{q}_k^T \mathbf{q}_{k-1}^* = 0\},
\end{aligned}$$

where $\mathbf{q}_1^*, \mathbf{q}_2^*, \ldots, \mathbf{q}_k^*$ are optimal solutions for the $1^{st}, 2^{nd}, \ldots, k^{th}$ optimization problems.

Let $\lambda^* := \max\{A \bullet Y : tr(Y) = k \text{ and } 0 \preceq Y \preceq I\}$ and let S denote the set of feasible solutions for this optimization problem. Then

$$(\lambda_1 + \ldots + \lambda_k)(A) = \mathbf{q}_1^{*T} A \mathbf{q}_1^* + \ldots + \mathbf{q}_k^{*T} A \mathbf{q}_k^* = tr(Q^T A Q) = (QQ^T) \bullet A \leq \lambda^*,$$

where $Q = [\mathbf{q}_1^*, \ldots, \mathbf{q}_k^*]$ and the last inequality follows from the fact that $QQ^T \in S$, by the previous Theorem.

To show that $\lambda^* \leq (\lambda_1 + \ldots + \lambda_k)(A)$ let assume that Y^* is an optimal solution with optimum value $\lambda^* = A \bullet Y^*$. We can assume also that Y^* is an extreme point of S. Hence, by the above Lemma, $Y^* = QQ^T$ and $Q^T Q = I_k$. Thus,

$$\lambda^* = (QQ^T) \bullet A = tr(Q^T A Q) = \sum_{i=1}^k q_i^T A q_i \leq (\lambda_1 + \ldots + \lambda_k)(A),$$

where $q_1, \ldots, q_k$ are the pairwise orthogonal column vectors of Q and the last inequality follows easily from the representations of $\lambda_1, \ldots, \lambda_k$ mentioned at the beginning of this proof. ■

The dual of $(\mathbf{D_5})$ is:

$$(\mathbf{P_5}) \qquad (\lambda_1 + \ldots + \lambda_k)(A) = \begin{array}{ll} \min & kz + tr(V) \\ \text{s.t.} & zI + V \succeq A, \\ & \quad V \succeq 0. \end{array}$$

Let us now consider a more general eigenvalue problem.

3.2 Problem 6: Find $\min\{(\lambda_1 + \ldots + \lambda_k)(A(\mathbf{x})) : \mathbf{x} \in R^n\}$.

Since we want to replace A by $A(\mathbf{x})$, we are looking for a formulation of Problem 5 in which A has only isolated appearance. It is easy to see that $(\mathbf{D_5})$ is not good for our purpuses, since $A \bullet Y$ is the objective function, however $(\mathbf{P_5})$ is good. Hence, by using $(\mathbf{P_5})$, we get easily the following formulation:

$$\min\{(\lambda_1 + \ldots + \lambda_k)(A(\mathbf{x})) : \mathbf{x} \in R^n\} =$$

$$(\mathbf{P_6}) \qquad = \begin{array}{ll} \min & kz + tr(V) \\ \text{s.t.} & zI + V - x_1 A_1 - \ldots - x_m A_m \succeq A_0, \\ & \quad V \succeq 0. \end{array}$$

$$(\mathbf{D_6}) \qquad = \begin{array}{ll} \max & A_0 \bullet Y \\ \text{s.t.} & tr(Y) = k, \\ & A_i \bullet Y = 0, \quad i = 1, \ldots, m, \\ & 0 \preceq Y \preceq I. \end{array}$$

The complementary slackness result for primal feasible $\mathbf{x}, z, V$ and dual feasible Y states that these are optimal if and only if

$$Y(zI + V - A(\mathbf{x})) = 0 \text{ and } (I - Y)V = 0.$$

3.3 Problem 7: Find $\min\{(\lambda_1 + \ldots + \lambda_k)(X) : A_i \bullet X = b_i, i = 1, \ldots, m\}$.

Similarly as above, Problem 6 can be expressed by the following primal dual pair:

$$\min\{(\lambda_1 + \ldots + \lambda_k)(X) : A_i \bullet X, i = 1, \ldots, m\} =$$

$$(\mathbf{P_7}) \qquad = \begin{array}{ll} \min & kz + tr(V) \\ \text{s.t.} & A_i \bullet X = b_i, i = 1, \ldots, m, \\ & zI + V - X \succeq 0, \\ & \quad V \succeq 0. \end{array}$$

$$(\mathbf{D_7}) \qquad = \begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} & Y - x_1 A_- \ldots - x_m A_m = A_0, \\ & tr(Y) = k, \\ & A_i \bullet Y = 0, \quad i = 1, \ldots, m, \\ & 0 \preceq Y \preceq I. \end{array}$$

The complementary slackness result for primal feasible z, X, V and dual feasible Y states that these are optimal if and only if

$$Y(zI + V - X) = 0 \text{ and } (I - Y)V = 0.$$

Exercises:

1. Show that $f_{|k|}$ is convex.
2. Express $\min\{f_{|k|}(\mathbf{x}) : \mathbf{x} \in R^m\}$ as a small LP.
3. Find an SDP characterization of $|\lambda^1(B)| + \ldots + |\lambda^k(B)|$.
4. Find an SDP characterization of

$$\min\{|\lambda^1(B)| + \ldots + |\lambda^k(B)| : A_i \bullet X = b_i, i = 1, \ldots, m, X \succeq 0\}.$$