Lecture 8: Semidefinite programming

Rajat Mittal

IIT Kanpur

Semidefinite programming is a class of convex optimization where:

- Cost/optimization function is linear,
- Constraints are either linear equalities/inequalities or generalized inequalities with respect to the semidefinite cone.

Hence, it can be viewed as linear programming with the additional power of generalized inequalities for the positive semidefinite cone ($\mathcal{S}_n$). In this lecture note we will look at the standard form of a semidefinite program. The lecture will end with some situations which can be modelled as a semidefinite program.

1 Definition

A semidefinite program (SDP) in a standard form looks like,

$$\begin{aligned} \max \quad & C \bullet X \\ \text{s.t.} \quad & A_i \bullet X = b_i, \; i = 1, \cdots, m \\ & X \succeq 0. \end{aligned}$$

Note 1. Remember that $A \bullet B = tr(A^T B)$ is the hadamard product between two matrices. It can also be viewed as the standard inner product if A, B are viewed as n^2 dimensional vectors.

In the semidefinite program above, X is the *variable matrix* of dimension $n \times n$. The matrix C is called the cost or *objective matrix*. A_i's are the *constraint matrices*. C and A_i's have the same dimension as X ($n \times n$). b_i's are scalars and the vector b (with b_i) as components is known as *constraint vector*.

Many of the standard tricks used in linear programming to convert non-standard form into standard form can also be used here. For example, converting inequalities into equalities, changing minimum to maximum and change of variables.

Let us take a look at the following program,

$$\begin{aligned} \max \quad & Tr(X) \\ \text{s.t.} \quad & X = \begin{pmatrix} 1 & x \\ 1 & x \end{pmatrix} \succeq 0. \end{aligned}$$

Show that it is a semidefinite program. In other words, what are the constraint matrices and constraint vector?

Exercise 1. Find the value of this semidefinite program.

Generalizing the above example, a semidefinite program can be viewed as,

- variables $x_{i,j}$ arranged in a matrix X,
- linear constraints and objective over these variable $x_{i,j}$,
- and the positive semidefinite constraint $X \succeq 0$.

You will prove in the assignment that the positive semidefinite constraint can be viewed as infinite number of linear constraints in variables $x_{i,j}$. Look at another program,

$$\begin{aligned} \inf / \min \quad & x_1 \\ \text{s.t.} \quad & \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \succeq 0. \end{aligned}$$

Exercise 2. Find the value of this semidefinite program.

1.1 Equivalent definitions

In general, many totally different looking programs can be transformed into a semidefinite program. Let us take a look at few forms which arise often in practice.

Form with positive semidefinite constraints Another standard form for semidefinite programming is:

$$\begin{gathered}\min b^T y \\ \text{s.t. } \sum_{i=1}^m y_i A_i - C \succeq 0\end{gathered}$$

Let us take a look at why these two forms are equivalent. Denote the matrix $\sum_i y_i A_i - C$ by a new variable matrix Z.

Now the variables in the program are Z and the scalars y_i's. The linear constraints can be changed into

$$\forall\ i,j;\ z_{ij} = \sum_k y_k (A_k)_{ij}.$$

Here z_{ij} are the entries of matrix Z. So, the program changes to,

$$\begin{gathered}\min b^T y \\ \text{s.t. } \forall\ i,j;\ z_{ij} = \sum_k y_k (A_k)_{ij} \\ Z \succeq 0\end{gathered}$$

It almost looks like the standard form but variables y_i's do not occur in the semidefinite constraint. Remember the old trick of converting unrestricted variables to positive variables.

Replace y_i by two positive variables, i.e., $y_i = y'_i - y''_i$ and $y'_i, y''_i \geq 0$. Now, these positive variables y'_i, y''_i can be put in a separate block and included in the semidefinite constraint.

$$\begin{gathered}\min \sum_k b_k (y'_k - y''_k) \\ \text{s.t.} \quad \forall\ i,j;\ z_{ij} = \sum_k (y'_k - y''_k)(A_k)_{ij} \\ Z = \begin{pmatrix} Z & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & y'_1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \vdots & \ddots & \vdots & 0 & 0 & 0 \\ 0 & 0 & 0 & y'_m & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y''_1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & y''_m \end{pmatrix} \succeq 0\end{gathered}$$

You will show in the assignment that we don't need explicit constraints on the off-diagonal being zero.

Gram matrix formulation We know that any $n \times n$ positive semidefinite matrix X can be written as the gram matrix of vectors $u_1, u_2, \cdots, u_n$. Suppose, our variable matrix X can be expressed as the gram matrix of $u_1, \cdots, u_n \in \mathbb{R}^k$. Notice that $u_1, u_2, \cdots, u_n$ are vector variables now and we have not imposed any constraint on k.

Then the semidefinite program looks like,

$$\begin{gathered}\max \quad \sum_{ij} C_{ij} u_i^T u_j \\ \text{s.t. } \sum_{ij} A_{ij}^{(k)} u_i^T u_j = b_k,\ \forall k = 1, \cdots, m\end{gathered}$$

We have removed the $X \succeq 0$ constraint from the new form.

The new form can be understood as,

– $u_1, u_2, \cdots, u_n$ as vector variables,
– and linear constraints over the *inner product* of these vector variables.

This form is specially useful in combinatorial optimization. The reason is, many problems can be expressed as integer programs in combinatorial optimization. When you relax these variables to be vectors instead of integers, this form arises naturally. We will see such an example later in the course.

Note 2. This is not a linear program, since constraints are on the inner-products.

2 Examples

Let us take a look at various natural problems which can be turned into a semidefinite program.

2.1 Minimizing the maximum eigenvalue

Suppose, we are given a matrix $M(x)$, which depends affinely on the variables in x. That means every entry in $M(x)$ can be written as an affine function of variables in x ($M(x)_{ij} = a_1x_1 + \cdots + a_nx_n + b$). The problem is to minimize the maximum eigenvalue of $M(x)$ over all x, i.e.,

$$\min_x \max_i \lambda_i(M(x))$$

Clearly this is not in the standard form of SDP. The trick here is to introduce another variable η to change min max to only min. Suppose $\lambda_{\max}(M)$ represents the maximum eigenvalue of M, then

$$\begin{aligned} &\min \eta \\ \text{s.t. } \ &\eta \geq \lambda_{\max}(M(x)) \end{aligned}$$

Now use the fact that $\lambda_{\max} I - M(x) \succeq 0$. Hence,

$$\begin{aligned} &\min \eta \\ \text{s.t. } \ &\eta I - M(x) \succeq 0 \end{aligned}$$

This is one of the alternative form discussed in the last section (why?). Here the variables are (η, x).

Exercise 3. Write an SDP to find the maximum eigenvalue of a matrix M.

2.2 Fan's theorem

The sum of first k eigenvalues can also be written as a semidefinite program.

Theorem 1 (Fan). *Given a symmetric matrix M and its eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$,*

$$\begin{aligned} \lambda_1 + \lambda_2 + \cdots \lambda_k \ &= \max M \bullet X \\ &\quad s.t. \ \ tr(X) = k \\ &\quad\quad I \succeq X \succeq 0 \end{aligned}$$

.

Proof. Say, M has spectral decomposition

$$\lambda_1 x_1 x_1^T + \cdots + \lambda_n x_n x_n^T.$$

Taking $X = x_1x_1^T + \cdots + x_kx_k^T$, we will get the objective value $\lambda_1 + \cdots + \lambda_k$. Since, we have a maximization problem, $\sum_{i=1}^k \lambda_i$ is a lower bound on the optimal value of the SDP. So, the only thing needed to prove the theorem is: any feasible X gives value less than $\lambda_1 + \cdots + \lambda_k$.

Since M is a symmetric matrix, by spectral decomposition it has a complete basis of eigenvectors ($x_1, \cdots, x_n$ span the entire space).

Exercise 4. Show that M is diagonal in this basis with diagonal entries as eigenvalues.

Notice that the trace and positive semidefiniteness properties are preserved under change of basis.

Exercise 5. Show that $tr(AB) = tr(BA)$, hence $tr(U^T BU) = tr(B)$ for any orthogonal U. Also show that if $B \succeq 0$ then $U^T BU \succeq 0$ for any orthogonal U.

Lets look at the semidefinite program in the theorem in this eigenvector basis. Then $M \bullet X = \sum_i X_{ii}\lambda_i$. From the constraints, $tr(X) = k$ and $I \succeq X \succeq 0$ (the identity matrix and zero matrix remain same under any basis transformation). Since $I - X \succeq 0$, all the entries $X_{ii} \leq 1$. Since $X \succeq 0$, implies $X_{ii} \geq 0$. Consider another optimization program (it is a linear program),

$$\begin{aligned} &\max \textstyle\sum_i z_i\lambda_i \\ \text{s.t.} \quad &\textstyle\sum_i z_i = k \\ &1 \geq z_i \geq 0. \end{aligned}$$

By substituting $z_i = X_{ii}$ it is clear that the value of this program is at least the value of the semidefinite program in the theorem.

Exercise 6. Show that the maximum for this new program will occur on z's which have exactly k non-zero entries equal to 1.

If we are allowed to choose k co-ordinates, s.t., $\sum_i \lambda_i z_i$ is maximum, then the best choice is the first k co-ordinates. Hence, the maximum value of this linear program is $\lambda_1 + \cdots + \lambda_k$. This implies that maximum value of the semidefinite program in Thm. 1 is $\lambda_1 + \cdots + \lambda_k$.

□

2.3 Linear programs as a special case

Linear programming is a special case of semidefinite programs. It is obtained by considering the diagonal matrices in the standard form of semidefinite programming. Suppose the linear program is,

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & a_i^T x = b_i, \ \forall i = 1, \cdots, m \\ & x \geq 0. \end{aligned}$$

in the standard form.

Look at the semidefinite program,

$$\begin{aligned} \max \quad & C \bullet X \\ \text{s.t.} \quad & A_i \bullet X = b_i, \ \forall i = 1, \cdots, m \\ & X \succeq 0; \end{aligned}$$

Here, C is the diagonal matrix with entries from c and A_i's are the diagonal matrices with diagonals a_i. Then the above mentioned two programs are actually equal.

Given a solution of the linear program, it can be converted into a solution for semidefinite program by taking X to be the diagonal matrix with diagonal x. Similarly, if X is any solution for the semidefinite program, then $x = diag(X)$ will be a solution for the linear program with the same objective value.

Hence, any linear program can be converted into a semidefinite program by taking corresponding diagonal matrices for the constraints as well as the objective matrix.

2.4 Sum of squares

If a polynomial $p(x)$ can be written as a sum of squares then it is clearly positive for all values of x (here $x = (x_1, \cdots, x_n)$). In general, this gives a sufficient condition for positivity of the polynomial. To check whether a polynomial can be written as a sum of squares is a semidefinite programming feasibility problem.

To see this, first notice that a polynomial (of degree d) can be written as the dot product between two vectors $p(x) = p^T x_d$. Here, x_d is the list of all degree $\leq d$ monomials. Then p is the vector of coefficients corresponding to those monomials.

Now, suppose $p(x) = \sum_i q_i(x)^2$, i.e., it can be written as the sum of squares. Say the degree of p is $2d$. Hence,

$$p(x) = \sum_i q_i(x)^2 = \sum_i x_d^T q_i^T q_i x_d = x_d^T (\sum_i q_i^T q_i) x_d$$

This shows that a polynomial is a sum of squares, iff, it can be written as $x_d^T Q x_d$ for some positive semidefinite Q. So, a polynomial is a sum of squares iff

$$\begin{aligned} p(x) &= x_d^T Q x_d \\ Q &\succeq 0. \end{aligned}$$

It might be confusing that why this is a semidefinite program. First, the constraint $p(x) = x_d^T Q x_d$ is a linear constraint on the elements of Q. Also the absence of max/min might be confusing. This kind of problem without objective function and only constraint is called a semidefinite programming feasibility problem. It is a special case of semidefinite programming (why ?).

This semidefinite program is important in giving bounds on the minimum value of a polynomial. Consider,

$$\begin{aligned} &\max \lambda \\ \text{s.t.}\ \ & p(x) - \lambda = x_d^T Q x_d \\ & Q \succeq 0. \end{aligned}$$

The value of this program (say s^*) satisfies $p(x) \geq s^*$ for all x. So, it gives a lower bound on the minimum value of the polynomial.

This might not be the biggest s, s.t. $p(x) \geq s$ for all x; since the representation as sum of squares is only a sufficient condition for positivity. Though if $n = 1$, this gives us the tight bound.

3 Assignment

Exercise 7. Show that semidefinite program can be viewed as an optimization problem with linear cost function and infinite linear constraints.

Exercise 8. Prove that we don't need to put the explicit constraint that the off-diagonal entries (blocks) are zero, when we converted form with positive semidefinite constraint into standard form.

Exercise 9. Prove that for a single variate polynomial p, it is positive iff it can be written as a sum of squares (Hint: Look at the factorization of p in complex domain).

Exercise 10. Show that the following problem can be converted into a SDP.

$$\begin{aligned} &\min \tfrac{(c^T x)^2}{d^T x} \\ \text{s.t.}\ \ & Ax + b \geq 0 \end{aligned}$$

Here, x is a vector and c, d are vectors of the same dimension as x.