**Lecture #8 Notes Summary**

Semidefinite Programming

# A short introduction to semidefinite programming (SDP)

Recall the standard form of a linear programming (LP) problem (in primal and dual form):

$$\begin{array}{llcll} \max & \boldsymbol{c}^T \boldsymbol{x} & \quad & \min & \boldsymbol{b}^T \boldsymbol{y} \\ s.t. & A\boldsymbol{x} = \boldsymbol{b} & & s.t. & A^T \boldsymbol{y} \geq \boldsymbol{c} \\ & \boldsymbol{x} \geq \boldsymbol{0} & & & \end{array}$$

If we denote the rows of $A$ by $\boldsymbol{a}_1^T, \ldots, \boldsymbol{a}_m^T$, and we use the notation $[m] := \{1, \ldots, m\}$ we can rewrite this pair of LPs as follows:

$$\begin{array}{llcll} \max & \langle \boldsymbol{c}, \boldsymbol{x} \rangle & \quad & \min & \sum_{i=1}^m b_i y_i \\ s.t. & \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = b_i \ (\forall i \in [m]) & & s.t. & \sum_{i=1}^m y_i \boldsymbol{a}_i \geq \boldsymbol{c} \\ & \boldsymbol{x} \geq \boldsymbol{0} & & & \end{array}$$

By analogy, we would like to define a pair of optimization problems by replacing the vector of variables $\boldsymbol{x} \in \mathbb{R}^n_+$ by a matrix $X \in \mathbb{S}^n_+$. To to this, we consider a list of matrices $C, A_1, \ldots, A_m \in \mathbb{S}_m$:

$$\begin{array}{llcll} \max\limits_X & \langle C, X \rangle & \quad & \min\limits_{\boldsymbol{y}} & \sum_{i=1}^m b_i y_i \\ s.t. & \langle A_i, X \rangle = b_i \ (\forall i \in [m]) & & s.t. & \sum_{i=1}^m y_i A_i \succeq C \\ & X \succeq 0 & & & \end{array} \qquad \text{(P)-(D)}$$

**Definition 1** (LMI)**.** A *linear matrix inequality* (LMI) is an equation of the form

$$\sum_{i=1}^m y_i A_i \succeq C \qquad (1)$$

for some symmetric matrices $C, A_1, \ldots, A_m$. The set of feasible solutions $S := \{\boldsymbol{y} \in \mathbb{R}^m : \ (1) \text{ holds}\}$ defines a convex region of $\mathbb{R}^m$ (called a *spectrahedron*).

**Definition 2.** A *semidefinite program* (SDP) is an optimization problem, where a linear function is optimized over a convex region that is defined by linear (in)equalities and LMIs.

**Proposition 1.** *Every SDP can be rewritten under the standard primal form* (*P*)*, and every SDP can also be rewritten under the standard dual form* (*D*)*.*

We do not prove this statement, but we demonstrate how this works by an example:

Consider the SDP

$$\begin{aligned} \max \quad & z - x \\ s.t. \quad & \begin{pmatrix} 1 & y & z \\ y & 1 & y \\ z & y & x \end{pmatrix} \succeq 0 \\ & 2x + y - z = 3 \end{aligned} \tag{2}$$

This problem can be put under the stadard dual form $(D)$ as follows:

$$\begin{aligned} \min \quad & -z + x \\ s.t. \quad & x \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + z \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \succeq \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ & x \begin{pmatrix} 2 & \\ & -2 \end{pmatrix} + y \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} + z \begin{pmatrix} -1 & \\ & 1 \end{pmatrix} \succeq \begin{pmatrix} 3 & \\ & -3 \end{pmatrix}, \end{aligned}$$ #1

and then the two LMIs can be combined in a big $5 \times 5$ LMI with a block structure. But the problem could also be cast as a standard SDP in the primal form $(P)$:

$$\begin{aligned} \max \quad & \left\langle \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 0 & -1 \end{pmatrix}, X \right\rangle \\ s.t. \quad & \left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 1, \qquad \left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 1 \\ & \left\langle \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, X \right\rangle = 0, \qquad \left\langle \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 2 \end{pmatrix}, X \right\rangle = 3 \\ & X \succeq 0 \end{aligned}$$

**Theorem 2.** *A matrix $X$ is positive semidefinite if and only if*

$$\forall U \succeq 0, \ \langle U, X \rangle \geq 0.$$

*Proof.* $\Longrightarrow$ was done in exercises last week.

$\Longleftarrow$: in particular, the inequality must be true for all matrices $U = \boldsymbol{u}\boldsymbol{u}^T$ of rank one. So $\forall \boldsymbol{u} \in \mathbb{R}^n$, $0 \leq \langle U, X \rangle = \langle \boldsymbol{u}\boldsymbol{u}^T, X \rangle = \operatorname{trace} \boldsymbol{u}\boldsymbol{u}^T X = \boldsymbol{u}^T X \boldsymbol{u}$, which proves $X \succeq 0$. □

**Corollary 3.** *Let $U \in \mathbb{S}^m$. Then,*

$$\max_{X \succeq 0} \langle X, U \rangle = \begin{cases} 0 & \text{if } U \preceq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

*Proof.* cf. Exercises □

**Proposition 4.** *Problem $(D)$ is the Lagrangian dual of $(P)$*

*Proof.* The Lagrangian of Problem $(P)$ is

$$\mathcal{L}(X, \boldsymbol{y}) = \langle C, X \rangle + \sum_{i=1}^{m} y_i (b_i - \langle A_i, X \rangle).$$

Note that Problem $(P)$ can be written as a saddle point problem:

$$\max_{X\succeq 0}\ \min_{\boldsymbol{y}\in\mathbb{R}^m}\ \mathcal{L}(X,\boldsymbol{y}).$$

By definition, the Lagrangian dual of an optimization problem is obtained by swapping min and max:

$$\min_{\boldsymbol{y}\in\mathbb{R}^m}\ \max_{X\succeq 0}\ \mathcal{L}(X,\boldsymbol{y}).$$

In the dual problem, the function to minimize is thus

$$\begin{aligned}\max_{X\succeq 0}\ \langle C,X\rangle+\sum_{i=1}^m y_i(b_i-\langle A_i,X\rangle) &= \sum_{i=1}^m y_i b_i+\max_{X\succeq 0}\ \langle X, C-\sum_{i=1}^m y_i A_i\rangle\\ &=\begin{cases}\sum_{i=1}^m y_i b_i & \text{if } \sum_{i=1}^m y_i A_i\succeq C\\ +\infty & \text{otherwise}\end{cases}\end{aligned}$$

where the last equality follows from Corollary 3.

Hence, we obtain problem $(D)$ when we minimize the expression $(\max_{X\succeq 0}\ \mathcal{L}(X,\boldsymbol{y}))$ over $\boldsymbol{y}\in\mathbb{R}^m$. □

From now on, and for more generality, we consider an alternative standard form with inequalities in the primal (as for LPs, there are many concurrent *standard forms*):

$$\begin{array}{lll} \max\limits_X & \langle C,X\rangle \\ s.t. & \langle A_i,X\rangle\leq b_i\ (\forall i\in[m]) \\ & X\succeq 0 \end{array} \qquad \begin{array}{lll} \min\limits_{\boldsymbol{y}} & \sum\limits_{i=1}^m b_i y_i \\ s.t. & \sum\limits_{i=1}^m y_i A_i\succeq C \\ & \boldsymbol{y}\geq \mathbf{0}. \end{array} \qquad \text{(P')-(D')}$$

**Theorem 5** (Weak Duality)**.** *Let $X\succeq 0$ be feasible for $(P')$ and $\boldsymbol{y}\geq\mathbf{0}$ be feasible for $(D')$. Then, $\langle C,X\rangle\leq \boldsymbol{b}^T\boldsymbol{y}$. In particular, the optimal value of $(P')$ is $\leq$ than the optimal value of $(D')$.*

*Proof.* $\boldsymbol{b}^T\boldsymbol{y}=\sum_i y_i b_i\geq \sum_i y_i\langle A_i,X\rangle=\langle\underbrace{\sum_i y_i A_i}_{\succeq C},X\rangle\geq\langle C,X\rangle.$ □

Contrarily to what happens with LP, strong duality does not always hold. However, a sufficient condition is the *Slater's condition*, aka strict feasibility.

**Definition 3** (strict feasibility)**.** An SDP is called strictly feasible if there is a feasible solution such that all linear (in)equalities are satisfied, and all LMIs are *strictly* satisfied. More precisely, an SDP of the form $(P')$ is called streactly feasible if

$$\exists X\succ 0:\forall i\in[m],\ \langle X,A_i\rangle\leq b_i.$$

Similarly, an SDP of the form $(D')$ is called strictly feasible if

$$\exists\boldsymbol{y}\geq 0:\sum_{i=1}^m y_i A_i\succ 0.$$

**Theorem 6** (Strong Duality)**.** *Consider a pair of primal-dual SDPs in the form* (P')-(D').

- *If* either $(P')$ or $(D')$ *is strictly feasible, then $OPT(P')=OPT(D')$ (but this value might be $\pm\infty$).*

- *If $(P')$ is strictly feasible and $OPT(P') < \infty$, then the infimum is attained in the dual problem (i.e., there exists a feasible $\boldsymbol{y}$ for $(D')$ such that $\boldsymbol{b}^T\boldsymbol{y} = OPT(P') = OPT(D')$.*
- *If $(D')$ is strictly feasible and $OPT(D') > -\infty$, then the supremum is attained in the primal problem (i.e., there exists a feasible $X$ for $(P')$ such that $\langle C, X\rangle = OPT(P') = OPT(D')$.*
- *If* both $(P')$ and $(D')$ *are feasible, then $OPT(P') = OPT(D')$ is finite, and the optimum is attained in both problems.*

*Proof.* Omitted. □

We can summarize the duality theorems as follows:

- $OPT(P) \leq OPT(D)$
- primal (*resp.* dual) feasible $\Longrightarrow$ dual (*resp.* primal) bounded
- primal or dual strictly feasible $\Longrightarrow OPT(P) = OPT(D)$
- primal (*resp.* dual) strictly feasible and bounded $\Longrightarrow$ dual (*resp.* primal) attainment

**Proposition 7** (Complementary Slackness)**.** *Consider a pair of primal/dual SDPs of the form* (P')-(D') *such that strong duality holds, and the optimal values are attained: there exists some feasible $X^* \in \mathbb{S}_+^n$ and $\boldsymbol{y^*} \in \mathbb{R}_+^m$ such that $\langle C, X^*\rangle = \boldsymbol{b}^T\boldsymbol{y^*}$ (and by weak duality, $X^*$ and $\boldsymbol{y^*}$ are optimal for* (P')-(D')*). Then, we have:*

$$\langle X^*, \sum_i y_i^* A_i - C\rangle = 0 \qquad \text{and} \qquad \forall i \in [m],\ y_i(b_i - \langle A_i, X^*\rangle) = 0.$$

*Proof.* These are simply the conditions for equality in the proof of the weak duality theorem. □

**Theorem 8.** *Every SDP can be solved to the desired precision in polynomial time. More precisely, given a SDP $(P)$ a target precision $\epsilon > 0$, the interior point method returns an $\epsilon-$approximation of the optimum in time polynomial with respect to $n$, $m$, and $\log \frac{1}{\epsilon}$.*

# Exercises

1. Proof of Corollary 3
2. Consider the SDP

$$\begin{array}{ll} \min & x_2 \\ s.t. & \begin{pmatrix} 1+x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{pmatrix} \succeq 0 \end{array}$$

   - What is the optimal value of this SDP?
   - What is the dual problem?
   - What is the optimal value of the dual?
   - Why isn't it a violation of the strong duality theorem?

3. Let $G = (V, E)$ be a simple undirected graph with $n$ vertices. We recall that the stable number $\alpha(G)$ of $G$ denotes the maximal cardinality of a *stable set* of $G$ (a set of vertices $S \subseteq V$ is stable iff $(i, j \in S \times S \Longrightarrow (i, j) \notin E)$). Consider the SDP:

$$\begin{array}{ll} \max\limits_{Y} & \langle J, Y \rangle \\ s.t. & \operatorname{trace} Y = 1 \\ & Y_{i,j} = 0, \quad \forall (i,j) \in E \\ & Y \succeq 0, \end{array}$$

   where $J$ is the $n \times n$ matrix of all ones. The optimal value of this SDP is denoted by $\vartheta(G)$, and is called the *Lovász theta number of G*.

   - Let $S$ be a subset of $V$, and denote by $\mathbf{1}_S$ the $\{0,1\}-$vector that indicates the vertices in $S$ ($(\mathbf{1}_S)_i = 1 \Longleftrightarrow i \in S$). Show that the matrix $Y = \frac{\mathbf{1}_S \mathbf{1}_S^T}{\mathbf{1}_S^T \mathbf{1}_S}$ is feasible for the SDP above if and only if $S$ is a stable set of $G$.
   - Deduce that $\alpha(G) \leq \vartheta(G)$.