Semidefinite Programming	**Lecture 9**
OR 6327 Spring 2012	**Febrary 21, 2012**
Scribe: Sin–Shuen Cheung	

Today we talk about SDP formulations of the Lovász theta function.

1 SDP formulations of the Lovász theta function

Recall that in the last lecture, we introduced $\vartheta(G)$, which satisfies $\alpha(G) \leq \Theta(G) \leq \vartheta(G)$. Here we recall the notation

$$\begin{aligned} \vartheta := \vartheta(G) := \quad & \min \quad \max_i \frac{1}{(c^T u_i)^2} \\ & \text{subject to} \quad c^T c = 1; \\ & \qquad u_i^T u_i = 1, \quad \text{for } i = 1, \ldots, n; \\ & \qquad u_i^T u_j = 0, \quad \text{for } ij \notin E,\ i \neq j. \end{aligned}$$

Thus we derive the following SDP formulation.

$$\begin{aligned} \frac{1}{\sqrt{\vartheta}} = \quad & \max \quad t \\ & \text{subject to} \quad \mathrm{Diag}\,(\bar{X}) = \bar{e}; \\ & \qquad \bar{x}_{ij} = 0, \quad \text{for } ij \notin E,\ i \neq j; \\ & \qquad \bar{x}_{0i} \geq t, \quad \text{for } i = 1, \ldots, n; \\ & \qquad \bar{X} \succeq 0; \end{aligned}$$

here $\bar{X} \in \mathbb{M}^{1+n}$ corresponds to $[c, u_1, \ldots, u_n]^T [c, u_1, \ldots, u_n]$ and $\bar{e} \in \mathbb{R}^{1+n}$ is a vector of ones.

Remark: we can assume that c lies in the span of $\{u_1, \ldots, u_n\}$ in any such factorization of $\bar{X}$, since replacing c by a scaled version of its projection into this span can only increase the objective function. So we can take all these vectors in $\mathbb{R}^n$ rather than $\mathbb{R}^{1+n}$ without loss of generality.

Notice that we can write

$$\bar{X} = \bar{I} - \sum_{ij \in E} y_{ij} \bar{F}_{ij} - \sum_i z_i \bar{F}_{0i},$$

where $\bar{F}_{ij} := \bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T$ and $\bar{I}$ denotes the identity matrix of dimension $1 + n$; then the SDP can be reformulated as follows:

$$\begin{aligned} \frac{1}{\sqrt{\vartheta}} = \quad & \max_{y,z,t} \quad t \\ & \text{subject to} \quad \bar{I} - \sum_{ij \in E} y_{ij} \bar{F}_{ij} - \sum_i z_i \bar{F}_{0i} \succeq \mathbf{0}; \\ & \qquad -z_i - t \geq 0, \quad \text{for } i = 1, \ldots, n. \end{aligned}$$

By taking the dual, we get

$$
\begin{array}{rcl}
\frac{1}{\sqrt{\vartheta}} = & \min\limits_{\bar{V},w} & \bar{I} \bullet \bar{V} \\
& \text{subject to} & \bar{F}_{ij} \bullet \bar{V} = 0, \quad ij \in E; \\
& & \bar{F}_{0i} \bullet \bar{V} + w_i = 0, \quad i = 1, \ldots, n; \\
& & e^T w = 1; \\
& & \bar{V} \succeq \mathbf{0}; \\
& & w \geq 0, \quad w \in \mathbb{R}^n.
\end{array}
$$

The last two lines of the above formulation can also be written as:

$$
\begin{bmatrix} \bar{V} & 0 \\ 0 & \operatorname{Diag}(w) \end{bmatrix} \succeq \mathbf{0}.
$$

From the constraint

$$
\bar{F}_{0i} \bullet \bar{V} + w_i = 0, \quad i = 1, \ldots, n,
$$

we can let

$$
\bar{V} = \begin{bmatrix} v_{00} & -\frac{1}{2} w^T \\ -\frac{1}{2} w & V \end{bmatrix},
$$

where $V \in \mathbb{M}^n$. Thus, we can reformulate the SDP as:

$$
(*) \quad
\begin{array}{rcl}
\frac{1}{\sqrt{\vartheta}} = & \min & v_{00} + I \bullet V \\
& \text{subject to} & v_{ij} = 0, \quad \text{for } ij \in E; \\
& & e^T w = 1, \ w \geq 0; \\
& & \begin{bmatrix} v_{00} & -\frac{1}{2} w^T \\ -\frac{1}{2} w & V \end{bmatrix} \succeq \mathbf{0}.
\end{array}
$$

Remarks:
1. Note that we can find strictly feasible solutions to both the primal and the dual, so, as we will see soon, there is no duality gap.
2. In fact, we can remove the nonnegativity constraints on w. Suppose w is not nonnegative; then if it is replaced by $|w|$, and we correspondingly change the signs of some rows and columns of V, we can get another solution to $(*)$ which is feasible except that $e^T|w| > 1$. By scaling down v_{00}, $|w|$, and V, we obtain a feasible solution to $(*)$ with a smaller objective function. So, it is safe to remove the constraint $w \geq \mathbf{0}$.

Consider any feasible solution to $(*)$. Since w is non–zero, we must have that $v_{00} > 0$. Then by the Schur complement theorem,

$$
V - \frac{1}{4v_{00}} w w^T \succeq \mathbf{0}.
$$

So,

$$e^T V e \geq \frac{1}{4v_{00}}(e^T w)^2 = \frac{1}{4v_{00}} > 0.$$

Thus, $v_{00} \geq \frac{1}{4e^T V e}$. In fact, we can achieve $v_{00} = \frac{1}{4e^T V e}$ by setting $w = \frac{Ve}{e^T V e}$, since

$$\begin{bmatrix} \frac{1}{4e^T V e} & -\frac{(Ve)^T}{2e^T V e} \\ -\frac{(Ve)^T}{2e^T V e} & V \end{bmatrix} = \begin{bmatrix} \frac{e^T}{2e^T V e} \\ -I \end{bmatrix} V \begin{bmatrix} \frac{e}{2e^T V e} & -I \end{bmatrix} \succeq \mathbf{0}.$$

So, $(*)$ is equivalent to

$$\begin{array}{ll} \min & I \bullet V + \dfrac{1}{4e^T V e} \\ \text{subject to} & (e^T V e > 0;) \\ & v_{ij} = 0, \quad ij \in E; \\ & V \succeq \mathbf{0}. \end{array}$$

Let $V = \lambda X$, with $I \bullet X = 1$, $\lambda \geq 0$; then we can get

$$\begin{array}{ll} \min\limits_{X} & \min\limits_{\lambda \geq 0} \left(\lambda + \dfrac{1}{4\lambda e^T X e} \right) \\ \text{subject to} & I \bullet X = 1; \\ & (e^T X e > 0;) \\ & x_{ij} = 0, \quad ij \in E; \\ & X \succeq \mathbf{0}. \end{array}$$

The inner minimization gives $\lambda = \frac{1}{2\sqrt{e^T X e}}$. So, we have

$$\begin{array}{rll} \dfrac{1}{\sqrt{\vartheta}} = & \min\limits_{X} & \dfrac{1}{\sqrt{e^T X e}} \\ & \text{subject to} & I \bullet X = 1; \\ & & (e^T X e > 0;) \\ & & x_{ij} = 0, \quad ij \in E; \\ & & X \succeq \mathbf{0}. \end{array}$$

So,

$$(P) \quad \begin{array}{rll} \vartheta = & \max & ee^T \bullet X \\ & \text{subject to} & I \bullet X = 1; \\ & & x_{ij} = 0, \quad ij \in E; \\ & & X \succeq \mathbf{0}. \end{array}$$

Here the constraint $e^T X e > 0$ is dropped, because it is satisfied by any feasible solution at least as good as $X = \frac{1}{n} I$. The dual to (P) is

$$
\begin{array}{llll}
 & \min & \lambda & \\
(D) & \text{subject to} & ee^T - \lambda I - \sum_{ij \in E} y_{ij} F_{ij} \preceq \mathbf{0}; & \\
 & \text{(i.e.} & \lambda I \succeq ee^T - \sum_{ij \in E} y_{ij} F_{ij}), &
\end{array}
$$

or

$$
\begin{array}{ll}
\min_{F} & \lambda_{\max} \; \left(ee^T + F\right) \\
\text{subject to} & f_{ij} = 0, \quad \text{for } ij \notin E.
\end{array}
$$

Again there is no duality gap, because both problems have strictly feasible solutions.

Here is another derivation of (P) as an SDP relaxation of finding a maximum stable set:

$$
\begin{aligned}
\alpha(G) &= \max \left\{ e^T x : \; x_i x_j = 0, ij \in E; \; x \in \{0,1\}^n \right\} \\
&= \max \left\{ \frac{(e^T x)^2}{x^T x} : \; x_i x_j = 0, ij \in E; \; x \in \{0,1\}^n \right\} \\
&\leq \max \left\{ \frac{(e^T x)^2}{x^T x} : \; x_i x_j = 0, ij \in E; \; x \geq \mathbf{0}; \; x \neq \mathbf{0} \right\} \\
&= \max \left\{ (e^T x)^2 : \; x^T x = 1; \; x_i x_j = 0, ij \in E; \; x \geq \mathbf{0} \right\} \\
&\leq \max \left\{ ee^T \bullet xx^T : \; I \bullet xx^T = 1; \; (xx^T)_{ij} = 0, ij \in E; \; xx^T \geq 0 \right\} \\
&\quad \text{(where } xx^T \geq 0 \text{ means that all entries are nonnegative)} \\
&\leq \max \left\{ ee^T \bullet X : \; I \bullet X = 1; \; x_{ij} = 0, ij \in E; \; X \succeq \mathbf{0}; \; X \geq 0 \right\} =: \vartheta'(G) \\
&\quad \text{(where } X \geq 0 \text{ means that all entries are nonnegative)} \\
&\leq v(P) = \vartheta(G).
\end{aligned}
$$

(This derivation shows that $\vartheta'(G)$ might give a better bound on $\alpha(G)$ than $\vartheta(G)$; but if the graph is sparse, it has many more constraints in its SDP formulation than does $\vartheta(G)$.) In fact, $\vartheta(G)$ is sandwiched between the stability number $\alpha(G)$ and the clique covering number $\bar{\chi}(G)$ — the minimal number of cliques (sets of mutually adjacent nodes) required to cover all the nodes of G.

Lovász's sandwich theorem:

$$
\begin{array}{ccccc}
\alpha(G) & \leq & \vartheta(G) & \leq & \bar{\chi}(G). \\
\uparrow & & \uparrow & & \uparrow \\
\text{NP–hard} & & \text{SDP computable} & & \text{NP–hard}
\end{array}
$$