Semidefinite Programming	**Lecture 19**
OR 6327 Spring 2012	**April 3, 2012**
Scribe: Ilker Birbil	

We shall devote the next two lectures to proving Proposition 2 and Theorem 2 stated on March 29 (Lecture 18). This effort will require us to show a series of lemmas. We first need to relate different norms because the update of XS is much more straightforward than that of $X^{\frac{1}{2}}SX^{\frac{1}{2}}$. (The hard part of the proof is showing that the iterates remain in the neighborhood of the central path.)

Some notation: For any $R \in \mathbb{R}^{n\times n}$, we will denote its symmetric part $(R+R^T)/2$ by $J(R)$ and its skew-symmetric part $(R-R^T)/2$ by $K(R)$. Similarly,

$$J_P(R) := J(PRP^{-1}) \quad \text{and} \quad K_P(R) := K(PRP^{-1}).$$

Lemma 1 *For every $R \in \mathbb{R}^{n\times n}$, we have the following:*

(a) $\operatorname{trace}(R^2) = \|J(R)\|_F^2 - \|K(R)\|_F^2$ *and* $\|R\|_F^2 = \|J(R)\|_F^2 + \|K(R)\|_F^2$.

(b) For every pair P, Q of invertible matrices in $\mathbb{R}^{n\times n}$, we can write

$$\|J_P(R)\|_F^2 - \|K_P(R)\|_F^2 = \|J_Q(R)\|_F^2 - \|K_Q(R)\|_F^2.$$

Proof: (a) Let $J = J(R)$ and $K = K(R)$. Then, $R = J + K$ and $R^2 = J^2 + JK + KJ + K^2$. But

$$J \bullet K = \operatorname{trace}(JK) = \operatorname{trace}(KJ) = \operatorname{trace}(JK^T) = -\operatorname{trace}(JK) = 0$$

and so

$$\operatorname{trace}(R^2) = \operatorname{trace}(J^TJ) - \operatorname{trace}(K^TK) = \|J\|_F^2 - \|K\|_F^2.$$

Also,

$$\|R\|_F^2 = \operatorname{trace}(R^TR) = \operatorname{trace}((J-K)(J+K)) = \operatorname{trace}(J^2) - \operatorname{trace}(K^2) = \|J\|_F^2 + \|K\|_F^2.$$

(b) We have

$$\|J_P(R)\|_F^2 - \|K_P(R)\|_F^2 = \|J(PRP^{-1})\|_F^2 - \|K(PRP^{-1})\|_F^2.$$

Then, by using part a), we obtain

$$\|J_P(R)\|_F^2 - \|K_P(R)\|_F^2 = \operatorname{trace}((PRP^{-1})^2) = \operatorname{trace}(R^2),$$

which is independent of P. $\square$

Lemma 2 *Suppose $R \in \mathbb{R}^{m\times n}$ has $J_P(R) = 0$ for some invertible P. Then,*

(a) $\|J(R)\|_F \leq \frac{1}{\sqrt{2}}\|R\|_F \leq \|K(R)\|_F$.

(b) If $R = S + T$, $S \in \mathbb{M}^n$, then $\|S\|_F \leq \sqrt{2}\|T\|_F$.

Proof: (a) Take $Q = I$ in Lemma 1(b) to get $\|J(R)\|_F^2 - \|K(R)\|_F^2 \le 0$. So

$$2\|J(R)\|_F^2 \le \|J(R)\|_F^2 + \|K(R)\|_F^2 \le 2\|K(R)\|_F^2.$$

(b) Note that $J(R) = J(S) + J(T) = S + J(T)$ and $K(R) = K(S) + K(T) = K(T)$. By using the triangle inequality, we obtain

$$\begin{aligned}\|S\|_F &\le \|S + J(T)\|_F + \|J(T)\|_F \\ &= \|J(R)\|_F + \|J(T)\|_F \\ &\le \|K(R)\|_F + \|J(T)\|_F \\ &= \|K(T)\|_F + \|J(T)\|_F \\ &\le \sqrt{2}\left(\|J(T)\|_F^2 + \|K(T)\|_F^2\right)^{\frac{1}{2}}. \quad \square\end{aligned}$$

Lemma 3 *Let* $\widehat{X} \succ 0$ *and* $\widehat{S} \succ 0$ *with*

$$\|\widehat{X}^{\frac{1}{2}}\widehat{S}\widehat{X}^{\frac{1}{2}} - \nu I\| \le \beta\nu, \quad 0 \le \beta \le \frac{1}{\sqrt{2}}$$

for some $\nu > 0$. *Suppose* $L \in \mathbb{R}^{n\times n}$ *and* $U, V \in \mathbb{M}^n$ *satisfy*

$$J(U\widehat{S} + \widehat{X}V) = J(L), \quad U \bullet V \ge 0. \tag{1}$$

Then with

$$\delta_U := \nu\|\widehat{X}^{-\frac{1}{2}}U\widehat{X}^{-\frac{1}{2}}\|_F, \quad \delta_V := \|\widehat{X}^{\frac{1}{2}}V\widehat{X}^{\frac{1}{2}}\|_F,$$

and

$$E := \widehat{X}^{-\frac{1}{2}}(U\widehat{S} + \widehat{X}V - L)\widehat{X}^{\frac{1}{2}},$$

we have

$$\max\{\delta_U, \delta_V\} \le \frac{1}{1-\sqrt{2}\beta}\left(\sqrt{2}\|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F + \|J_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F\right)$$

and

$$\|J(E)\|_F \le \|K(E)\|_F \le \beta\delta_U + \|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F.$$

Proof: Let $E = S + T$, where

$$S = \widehat{X}^{\frac{1}{2}}V\widehat{X}^{\frac{1}{2}} + \nu\widehat{X}^{-\frac{1}{2}}U\widehat{X}^{-\frac{1}{2}} - J_{\widehat{X}^{-\frac{1}{2}}}(L)$$

and

$$T = \left(\widehat{X}^{-\frac{1}{2}}U\widehat{X}^{-\frac{1}{2}}\right)\left(\widehat{X}^{\frac{1}{2}}\widehat{S}\widehat{X}^{\frac{1}{2}} - \nu I\right) - K_{\widehat{X}^{-\frac{1}{2}}}(L).$$

But $J_{\widehat{X}^{\frac{1}{2}}}(E) = 0$; so by Lemma 2(b) we have $\|S\|_F \le \sqrt{2}\|T\|_F$. Also, using $\|AB\|_F \le \|A\|_F\|B\|$,

$$\|T\|_F \le \left(\frac{\delta_U}{\nu}\right)(\nu\beta) + \|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F. \tag{2}$$

Thus,

$$\begin{aligned}
\max\{\delta_U, \delta_V\} &\leq (\delta_U^2 + \delta_V^2)^{\frac{1}{2}} \\
&\leq \|\widehat{X}^{\frac{1}{2}} V \widehat{X}^{\frac{1}{2}} + \nu \widehat{X}^{-\frac{1}{2}} U \widehat{X}^{-\frac{1}{2}}\|_F \\
&= \|S + J_{\widehat{X}^{-\frac{1}{2}}}(L)\|\|_F \\
&\leq \sqrt{2}\|T\|_F + \|J_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F \\
&\leq \sqrt{2}\left(\beta\delta_U + \|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F\right) + \|J_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F \\
&\leq \sqrt{2}\beta \max\{\delta_U, \delta_V\} + \sqrt{2}\|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F + \|J_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F,
\end{aligned}$$

where the second inequality follows from $U \bullet V \geq 0$ and the last but one from (2). This gives the bound on $\max\{\delta_U, \delta_V\}$. Also, Lemma 2(a) shows that $\|J(E)\|_F \leq \|K(E)\|_F$ and

$$\|K(E)\|_F = \|K(T)\|_F \leq \|T\|_F \leq \beta\delta_U + \|K_{\widehat{X}^{-\frac{1}{2}}}(L)\|_F$$

following again by relation (2). $\square$

Proof of Proposition 2 (March 29, Lecture 18): Suppose X and S are as in the proposition and let $\widehat{X} = PXP^T$ and $\widehat{S} = P^{-T}SP^{-1}$ for the invertible P used in defining the directions. If λ is the vector of eigenvalues of $X^{\frac{1}{2}}SX^{\frac{1}{2}}$, then $\|\lambda - \nu e\|_\infty \leq \beta\nu$. However, $\widehat{X}^{\frac{1}{2}}\widehat{S}\widehat{X}^{\frac{1}{2}}$ is similar to $\widehat{X}\widehat{S}$, which is similar to XS, and which is in turn similar to $X^{\frac{1}{2}}SX^{\frac{1}{2}}$ (i.e., they have the same eigenvalues). Thus,

$$\|\widehat{X}^{\frac{1}{2}}S\widehat{X}^{\frac{1}{2}} - \nu I\| \leq \beta\nu,$$

which shows that we can use Lemma 3. We want to show that there exists a unique solution to

$$\begin{aligned}
\widehat{\mathcal{A}}^*\widehat{\Delta y} + \widehat{\Delta S} &= 0, \\
\widehat{\mathcal{A}}\widehat{\Delta X} &= 0, \\
J(\widehat{\Delta X}\widehat{S} + \widehat{X}\widehat{\Delta S}) &= J(\nu I - \widehat{X}\widehat{S}).
\end{aligned}$$

Suppose U, w, V is solution to the corresponding homogeneous equations. Then, U and V satisfy relation (1) with $L = 0$. Lemma 3 then shows that $\max\{\delta_U, \delta_V\} \leq 0$, so $U = V = 0$. Then, $w = 0$ since the A_i's are linearly independent. Therefore, there is a unique solution to the direction equations. $\square$