Semidefinite Programming	**Lecture 20**
OR 6327 Spring 2012	**April 5, 2012**
Scribe: Patrick Steele	

Notation

- The current iterate is $(X, y, S) \in \mathcal{N}_F(\beta)$, $0 < \beta < 1/\sqrt{2}$.
- The direction corresponding to $\nu = \sigma\mu$ (and P) is $(\Delta X, \Delta y, \Delta S)$.
- For P invertible, $(\hat{X}, \hat{y}, \hat{S}) \in \hat{\mathcal{N}}_F(\beta)$, with $\hat{\mu} := \hat{X} \bullet \hat{S}/n$, where
  - $\hat{X} := PXP^{\mathrm{T}}$, $\quad \widehat{\Delta X} := P\Delta X P^{\mathrm{T}}$,
  - $\hat{S} := P^{-\mathrm{T}}SP^{-1}$, $\quad \widehat{\Delta S} := P^{-\mathrm{T}}\Delta S P^{-1}$.
- $(X(\alpha), y(\alpha), S(\alpha)) = (X, y, S) + \alpha(\Delta X, \Delta y, \Delta S)$, with $(\hat{X}(\alpha), \hat{y}(\alpha), \hat{S}(\alpha))$ defined analogously.

Lemma (4). *With the notation above, for all $\alpha \in [0, 1]$,*

(a) $\hat{X}(\alpha) \bullet \hat{S}(\alpha) = (1 - \alpha + \alpha\sigma)\hat{X} \bullet \hat{S}$;

(b) $J_{\hat{X}^{-1/2}}(\hat{X}(\alpha)\hat{S}(\alpha) - \hat{\mu}(\alpha)I) = (1-\alpha)(\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}I) + \alpha J_{\hat{X}^{-1/2}}(R) + \alpha^2 J_{\hat{X}^{-1/2}}(\widehat{\Delta X}\widehat{\Delta S})$

where

$$\hat{\mu}(\alpha) := \frac{\hat{X}(\alpha) \bullet \hat{S}(\alpha)}{n} = (1 - \alpha + \alpha\sigma)\hat{\mu}$$

and

$$R := \hat{X}^{-1/2}(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S} + \hat{X}\hat{S} - \sigma\mu I)\hat{X}^{1/2}.$$

Proof. From the definitions,

$$\hat{X}(\alpha)\hat{S}(\alpha) = \hat{X}\hat{S} + \alpha(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S}) + \alpha^2\widehat{\Delta X}\widehat{\Delta S}.$$

For (a), note that $\widehat{\Delta X} \bullet \widehat{\Delta S} = 0$. Also, $\mathbf{tr}(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S}) = \mathbf{tr}(J(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S})) = \mathbf{tr}(J(\sigma\hat{\mu}I - \hat{X}\hat{S})) = \mathbf{tr}(\sigma\hat{\mu}I - \hat{X}\hat{S}) = \sigma\hat{\mu}n - \hat{\mu}n$. So

$$\mathbf{tr}(\hat{X}(\alpha)\hat{S}(\alpha)) = n\hat{\mu} + \alpha(\sigma\hat{\mu}n - \hat{\mu}n) = (1 - \alpha + \alpha\sigma)n\hat{\mu},$$

as required.

For (b),

$$\begin{aligned} J_{\hat{X}^{-1/2}}(\hat{X}(\alpha)\hat{S}(\alpha) - \hat{\mu}(\alpha)I) &= \hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}(\alpha)I + J_{\hat{X}^{-1/2}}(\alpha(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S})) + \alpha^2 J_{\hat{X}^{-1/2}}(\widehat{\Delta X}\widehat{\Delta S}) \\ &= (1-\alpha)(\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}I) + \alpha J_{\hat{X}^{-1/2}}(\hat{X}\widehat{\Delta S} + \widehat{\Delta X}\hat{S} + \hat{X}\hat{S} - \sigma\hat{\mu}I) \\ &\quad + \alpha^2 J_{\hat{X}^{-1/2}}(\widehat{\Delta X}\widehat{\Delta S}). \end{aligned}$$

□

Lemma (5). *If $(\hat{X}, \hat{y}, \hat{S}) \in \hat{\mathcal{N}}_F(\beta)$, then $\|\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \sigma\hat{\mu}I\|_F \le (\beta^2 + (1-\sigma)^2 n)^{1/2}\hat{\mu}$.*

Proof. Observe that $\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \sigma\hat{\mu}I = (\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \hat{\mu}I) + (1-\sigma)\hat{\mu}I$. Since the two matrices on the right-hand side are orthogonal symmetric matrices, we get the desired bounds. □

Lemma (6). *Assume $(\hat{X}, \hat{y}, \hat{S}) \in \mathcal{N}_F(\beta)$, $0 < \beta < 1/\sqrt{2}$. Then for $0 \le \alpha \le 1$,*

$$\|J_{\hat{X}^{-1/2}}(\hat{X}(\alpha)\hat{S}(\alpha) - \hat{\mu}(\alpha)I)\|_F \le (1-\alpha)\beta + \alpha\beta\theta + \alpha^2\theta^2\hat{\mu},$$

where $\theta := (\beta^2 + (1-\sigma)^2 n)^{1/2}/(1-\sqrt{2}\beta)$.

Proof. Let $\delta_X = \hat{\mu}\|\hat{X}^{-1/2}\widehat{\Delta X}\hat{X}^{-1/2}\|_F$, $\delta_S = \|\hat{X}^{1/2}\widehat{\Delta S}\hat{X}^{1/2}\|_F$. From Lemma 3 with $L := \sigma\hat{\mu}I - \hat{X}\hat{S}$ and $\nu = \hat{\mu}$, we see that $K_{\hat{X}^{-1/2}}(L) = 0$, and get

$$\max\{\delta_X, \delta_S\} \le \frac{1}{1-\sqrt{2}\beta}\|\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \sigma\hat{\mu}I\|_F \le \theta\hat{\mu},$$

using Lemma 5, and

$$\|J(R)\|_F \le \beta\delta_X \le \beta\theta\hat{\mu},$$

with R as in Lemma 4. We also have

$$\alpha^2\|J_{\hat{X}^{-1/2}}(\widehat{\Delta X}\widehat{\Delta S})\|_F \le \alpha^2\|\hat{X}^{-1/2}\widehat{\Delta X}\hat{X}^{-1/2}\|_F\|\hat{X}^{1/2}\widehat{\Delta S}\hat{X}^{1/2}\|_F \le \alpha^2\frac{\theta\hat{\mu}}{\hat{\mu}} \bullet \theta\hat{\mu},$$

and then combining these bounds with Lemma 4 gives the desired result. □

Lemma (7). *Suppose $\hat{X} \succ 0$, $\hat{S} \succ 0$, and Q is invertible. Then for any $\nu > 0$,*

$$\|\hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \nu I\|_F \le \|J_Q(\hat{X}\hat{S} - \nu I)\|_F.$$

Proof. Using Lemma 1(b) with $P = \hat{X}^{-1/2}$ and $R = \hat{X}\hat{S} - \nu I$ and the fact that $J_P(\hat{X}\hat{S} - \nu I) = \hat{X}^{1/2}\hat{S}\hat{X}^{1/2} - \nu I$ and $K_P(\hat{X}\hat{S} - \nu I) = 0$ yields the desired result. □

Lemma (8). *Let $V \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n}$ invertible be given. If $\|J_Q(V) - I\| < 1$, then V is invertible.*

Proof. Let $W = QVQ^{-1}$, so $J_Q(V) = (W + W^{\mathrm{T}})/2$. So all the eigenvalues of $(W + W^{\mathrm{T}})/2)$ are positive, and so $(W + W^{\mathrm{T}})/2$ is positive definite. If $Wu = 0$, then $u^{\mathrm{T}}(W + W^{\mathrm{T}})u = 0$, and so $u = 0$; thus, W is invertible, and hence so is V. □

Theorem. *Suppose $(X, y, S) \in \mathcal{N}_F(\beta)$, $0 < \beta < 1/\sqrt{2}$, and $0 \le \delta < \sqrt{n}$ are such that*

$$\frac{2(\beta^2 + \delta^2)}{(1-\sqrt{2}\beta)^2}\frac{1}{1-\delta/\sqrt{n}} = \gamma < 1.$$

If $(\Delta X, \Delta y, \Delta S)$ is defined at (X, y, S) using any invertible P and $\sigma = 1 - \delta/\sqrt{n}$, then

(a) $(X(1), y(1), S(1)) \in \mathcal{N}_F(\gamma)$, and

(b) $X(1) \bullet S(1) = (1 - \frac{\delta}{\sqrt{n}})X \bullet S$.

Proof. Define the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ as usual. Part (b) is immediate from Lemma 4. We now show part (a). From Lemma 6, using $\beta \leq \theta$ and $\alpha \leq 1$,

$$\begin{aligned} \|J_{\hat{X}^{-1/2}}(\hat{X}(\alpha)\hat{S}(\alpha) - \hat{\mu}(\alpha)I)\|_F &\leq ((1-\alpha)\beta + 2\alpha\theta^2)\hat{\mu} \\ &= \left((1-\alpha)\beta + 2\alpha\frac{\beta^2+\delta^2}{(1-\sqrt{2}\beta)^2}\right)\hat{\mu} \\ &= ((1-\alpha)\beta + \alpha\sigma\gamma)\,\hat{\mu} \\ &\leq (\max\{\beta, \gamma\})\hat{\mu}(\alpha), \end{aligned}$$

for every $0 \leq \alpha \leq 1$. So, dividing by $\hat{\mu}(\alpha)$, we have

$$\|J_{\hat{X}^{-1/2}}\left(\frac{\hat{X}(\alpha)\hat{S}(\alpha)}{\hat{\mu}(\alpha)} - I\right)\|_F < 1.$$

So by Lemma 8, $\hat{X}(\alpha)\hat{S}(\alpha)$ is invertible. Then both $\hat{X}(\alpha)$ and $\hat{S}(\alpha)$ must be positive definite for $0 \leq \alpha \leq 1$. We also get

$$\|J_{\hat{X}^{-1/2}}(\hat{X}(1)\hat{S}(1) - \hat{\mu}(1)I)\|_F \leq \gamma\hat{\mu}(1).$$

Then Lemma 7 gives

$$\|\hat{X}^{1/2}(1)\hat{S}(1)\hat{X}^{1/2} - \hat{\mu}(1)I\|_F \leq \gamma\hat{\mu}(1),$$

so $(\hat{X}(1), \hat{y}(1), \hat{S}(1))$ is in the required neighborhood. (Showing that the linear constraints are satisfied is easy using the equations defining the directions.) Scaling back yields (a). □

If $\beta = 1/10$ and $\delta = 1/7$, then $\gamma < 1/10$. Hence the theorem above implies Theorem 2 in Lecture 18 by an easy induction.

N.B. The preceding results are due to Renato Monteiro.