Semidefinite Programming	**Lecture 18**
OR 6327 Spring 2012	**March 29, 2012**
Scribe: Xiaoting Zhao	

Scale-invariance

We'll say a direction is *scale-invariant*, if whenever it gives $(\Delta X, \Delta y, \Delta S)$ at (X, y, S) for a given pair of problems, it gives $(U\Delta X U^T, \Delta y, U^{-T}\Delta S U^{-1})$ at $(UXU^T, y, U^{-T}SU^{-1})$ for the corresponding scaled problems.

Proposition 1 *The primal, dual, HKM and dual HKM directions are all scale-invariant, while the AHO direction is not.*

Proof: We'll just consider the HKM direction. Suppose

$$\Delta X + (X \odot S^{-1})\Delta S = \nu S^{-1} - X;\ \text{or}$$

$$\Delta X + \frac{1}{2}X\Delta S S^{-1} + \frac{1}{2}S^{-1}\Delta S X = \nu S^{-1} - X.$$

Pre- and post multiply by U and U^T to get

$$(U\Delta X U^T) + \frac{1}{2}(UXU^T)(U^{-T}\Delta S U^{-1})(US^{-1}U^T) + \frac{1}{2}(US^{-1}U^T)(U^{-T}\Delta S U^{-1})(UXU^T)$$
$$= \nu\, US^{-1}U^T - UXU^T,$$

and so $(U\Delta X U^T) + ((UXU^T) \odot (U^{-T}SU^{-1}))U^{-T}\Delta S U^{-1} = \nu(U^{-T}SU^{-1})^{-1} - (UXU^T)$,

which defines the HKM direction for the scaled problems. $\square$

Nesterov-Todd direction

In some sense, the *Nesterov-Todd direction* is "halfway between" primal and dual directions, and the HKM and dual HKM directions.

$$(\nu^{-\frac{1}{2}}X)^{-1}\Delta X(\nu^{-\frac{1}{2}}X)^{-1} + \Delta S = \nu X^{-1} - S \quad \leftarrow \text{primal}$$

$$(\Delta X + (\nu^{-\frac{1}{2}}S)^{-1}\Delta S(\nu^{-\frac{1}{2}}S)^{-1} = \nu S^{-1} - X \quad \leftarrow \text{dual}$$

Suppose we replace $\nu^{-\frac{1}{2}}X$ in the first equation by $W \succ 0$ and $\nu^{-\frac{1}{2}}S$ in the second equation by W^{-1}. We get

$$W^{-1}\Delta X W^{-1} + \Delta S = \nu X^{-1} - S$$

$$\Delta X + W\Delta S W = \nu S^{-1} - X$$

These give the same direction as long as $WX^{-1}W = S^{-1}$ and $WSW = X$.
Note that

$$WSW = X \iff S^{\frac{1}{2}}WSWS^{\frac{1}{2}} = S^{\frac{1}{2}}XS^{\frac{1}{2}} \text{ since } S \succ 0,\ \text{or}$$

$$(S^{\frac{1}{2}}WS^{\frac{1}{2}})^2 = S^{\frac{1}{2}}XS^{\frac{1}{2}}.$$

So $W := W(X,S) := S^{-\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{\frac{1}{2}}S^{-\frac{1}{2}}$ satisfies $WSW = X$, which implies $W^{-1}X = SW$ (and so $W^{-1}S^{-1}W^{-1} = X^{-1}$, and $WX^{-1}W = S^{-1}$).

This gives the Nesterov-Todd (NT) direction, corresponding to

$$\mathcal{E} = \mathcal{I},$$

$$\mathcal{F} = W \odot W,$$

$$R_{EF} = \nu S^{-1} - X.$$

(Alternatively, $\mathcal{E} = W^{-1} \odot W^{-1}, \mathcal{F} = \mathcal{I}, R_{EF} = \nu X^{-1} - S$, so the direction can be seen as in between the HKM and dual HKM directions.) Since $\mathcal{E}$ is invertible and $\mathcal{E}^{-1}\mathcal{F} = W \odot W$ is positive definite, the NT direction is well-defined for all $X \succ 0, S \succ 0$.

Exercise: Show that the NT direction is self-dual and scale-invariant.

Summary Table:

	Directions depends on primal and dual	Well-defined	Self-dual	Scaling-invariant
Primal	X	√	X	√
Dual	X	√	X	√
AHO	√	X	√	X
HKM	√	√	X	√
Dual HKM	√	√	X	√
NT	√	√	√	√

Remarks:
W is called the *metric geometric mean* of X and S^{-1}. Also, $W^{\frac{1}{2}}SW^{\frac{1}{2}} = W^{-\frac{1}{2}}XW^{-\frac{1}{2}}$, so $P = U = W^{-\frac{1}{2}}$ scales X and S to the *same* matrix, which obviously commutes with itself. The matrix $W^{\frac{1}{2}}SW^{\frac{1}{2}} = W^{-\frac{1}{2}}XW^{-\frac{1}{2}}$ is called the *spectral geometric mean* of X and S.

Theorem 1 *The NT direction lies in the MZ family, corresponding to $P = W^{-\frac{1}{2}}, M = W^{-1}$.*

Proof: Let $(\Delta X, \Delta y, \Delta S)$ be the NT direction, and consider (using $W^{-1}X = SW$)

$$\begin{aligned}
(W^{-1} \odot S)\Delta X + (W^{-1}X \odot I)\Delta S &= \frac{1}{2}(W^{-1}\Delta X S + S\Delta X W^{-1} + SW\Delta S + \Delta S W S) \\
&= \frac{1}{2}(W^{-1}\Delta X S + \Delta S W S + S\Delta X W^{-1} + SW\Delta S) \\
&= \frac{1}{2}(W^{-1}(\Delta X + W\Delta S W)S + S(\Delta X + W\Delta S W)W^{-1}) \\
&= \frac{1}{2}(W^{-1}(\nu S^{-1} - X)S + S(\nu S^{-1} - X)W^{-1}) \\
&= \frac{1}{2}(\nu W^{-1} - W^{-1}XS + \nu W^{-1} - SXW^{-1}) \\
&= \nu W^{-1} - \frac{1}{2}W^{-1}XS - \frac{1}{2}SXW^{-1},
\end{aligned}$$

so we satisfy the last equation corresponding to

$$P = W^{-\frac{1}{2}}, \; M = W^{-1}.$$

To complete the proof, we show that this choice of P and M gives a unique direction. It suffices that $\mathcal{E}^{-1}$ exists and $\mathcal{E}^{-1}\mathcal{F}$ is positive definite, with

$$\mathcal{E} = W^{-1} \odot S, \quad \mathcal{F} = (W^{-1}X \odot I),$$

so $\mathcal{E}$ is self-adjoint and positive definite, so invertible.

Let $0 \neq V \in \mathbb{M}^n$. Let $U = (W^{-1} \odot S)^{-1}V \neq 0$. Then

$$\begin{aligned}
V \bullet \mathcal{E}^{-1}\mathcal{F}V &= (\mathcal{E}^{-1}V) \bullet \mathcal{F}V \\
&= U \bullet \mathcal{F}\mathcal{E}U \\
&= \frac{1}{2}U \bullet (\mathcal{F}(W^{-1}US + SUW^{-1})) \\
&= \frac{1}{4}U \bullet (W^{-1}XW^{-1}US + W^{-1}XSUW^{-1} + W^{-1}USXW^{-1} + SUW^{-1}XW^{-1}) \\
&= \frac{1}{4}(\mathrm{tr}(USUS) + \mathrm{tr}(USUS) + \mathrm{tr}(USWSUW^{-1}) + \mathrm{tr}(UW^{-1}USWS)) \\
&= \frac{1}{2}||S^{\frac{1}{2}}US^{\frac{1}{2}}||_F^2 + \frac{1}{2}||W^{-\frac{1}{2}}U(SWS)^{\frac{1}{2}}||_F^2 > 0. \qquad \square
\end{aligned}$$

Short-Step Algorithm

- Choose $0 < \beta < \frac{1}{\sqrt{2}}$, $0 < \delta < 1$, and $\epsilon > 0$. Suppose we have $(X_0, y_0, S_0) \in \mathcal{N}_F(\beta)$.
- Iteration k $(k = 0, 1, \ldots)$

- Let $(X,y,S) = (X_k, y_k, S_k) \in \mathcal{N}_F(\beta)$ and stop if $X_k \bullet S_k \le \epsilon(X_0 \bullet S_0)$. Otherwise, $\mu := \mu_k := \frac{X_k \bullet S_k}{n}$, $\nu := \sigma\mu := (1 - \frac{\delta}{\sqrt{n}})\mu$.

- Then compute the direction $(\Delta X, \Delta y, \Delta S)$ from the **MZ** family equations using some invertible $P := P_k$. Set $(X_{k+1}, y_{k+1}, S_{k+1}) = (X, y, S) + (\Delta X, \Delta y, \Delta S)$.

We first state the main results and then prove a set of lemmas to establish them:

Proposition 2 *: Let $(X,y,S) \in \mathcal{N}_\infty(\beta)$ for $0 < \beta < \frac{1}{\sqrt{2}}$. Let $0 \le \sigma \le 1$ and $\nu = \sigma\mu = \sigma \bullet \frac{X \bullet S}{n}$. Then, for any invertible P, the corresponding direction $(\Delta X, \Delta y, \Delta S)$ is well-defined.*

Theorem 2 *: Suppose $\beta = \frac{1}{10}, \delta = \frac{1}{7}$; then for the Short-Step Algorithm:*

(i). (X_k, y_k, S_k) is well-defined for all k;

(ii). $(X_k, y_k, S_k) \in \mathcal{N}_F(\beta)$ for all k;

(iii). $X_{k+1} \bullet S_{k+1} = (1 - \frac{\delta}{\sqrt{n}})(X_k \bullet S_k)$ for all k; and

(iv). the algorithm terminates within $O(\sqrt{n} \ln \frac{1}{\varepsilon})$ iterations.