Semidefinite Programming	**Lecture 17**
OR 6327 Spring 2012	**March 27, 2012**
Scribe: Tia Sondjaja	

So far, we have:

- the primal direction, given by:
$$\mathcal{E} = \nu X^{-1} \odot X^{-1},\ \mathcal{F} = \mathcal{I},\ R_{EF} = \nu X^{-1} - S,$$
- the dual direction, given by:
$$\mathcal{E} = \mathcal{I},\ \mathcal{F} = \nu S^{-1} \odot S^{-1},\ R_{EF} = \nu S^{-1} - X,$$
- the HRVW-KSH direction, given by:
$$\mathcal{E} = \mathcal{I},\ \mathcal{F} = X \odot S^{-1},\ R_{EF} = \nu S^{-1} - X,$$
- and the AHO direction, given by:
$$\mathcal{E} = S \odot I,\ \mathcal{F} = X \odot I,\ R_{EF} = \nu I - \frac{1}{2}(XS + SX).$$

Note that every direction/method has a dual obtained by re-writing (P) in dual form, rewriting (D) in primal form, applying the original direction/method to the new problems at the iterate (S, X), and obtaining $(\Delta S, \Delta X)$.

The dual of the primal (dual) direction is the dual (primal) direction, and the AHO direction is self-dual. The dual of the HRVW-KSH direction is called the "dual HRVW-KSH direction" and is defined by:
$$\mathcal{E} = S \odot X^{-1},\ \mathcal{F} = \mathcal{I},\ R_{EF} = \nu X^{-1} - S.$$

Before putting these directions (except for the first two) into a common framework, we revisit the question of whether the directions are well-defined.

Proposition 1 *Suppose that $G \succ 0,\ H \succ 0$. Then, $G \odot H$, viewed as an operator on $\mathbb{M}^n$, is self-adjoint and positive definite. Moreover, $(G \odot G)^{-1} = G^{-1} \odot G^{-1}$.*

Proof: Choose $U, V \in \mathbb{M}^n$. Then,

$$\begin{aligned}
U \bullet (G \odot H)(V) &= \text{trace}\left(U\left(\frac{1}{2}GVH + \frac{1}{2}HVG\right)\right) \\
&= \frac{1}{2}\text{trace}\,(UGVH) + \frac{1}{2}\text{trace}\,(UHVG) \\
&= \frac{1}{2}\text{trace}\left(\left(H^{1/2}UG^{1/2}\right)\left(G^{1/2}VH^{1/2}\right)\right) + \frac{1}{2}\text{trace}\,(GVHU) \\
&= \frac{1}{2}\text{trace}\left(\left(H^{1/2}UG^{1/2}\right)\left(G^{1/2}VH^{1/2}\right)\right) + \\
&\quad \frac{1}{2}\text{trace}\left(\left(H^{1/2}UG^{1/2}\right)\left(G^{1/2}VH^{1/2}\right)\right) \\
&= \left(G^{1/2}UH^{1/2}\right) \bullet \left(G^{1/2}VH^{1/2}\right) \\
&= \left(G^{1/2}VH^{1/2}\right) \bullet \left(G^{1/2}UH^{1/2}\right) \\
&= V \bullet (G \odot H)(U).
\end{aligned}$$

In particular if $V = U$, we get

$$U \bullet (G \odot H)(U) = \|G^{1/2} U H^{1/2}\|_F^2 > 0,$$

for $U \neq 0$. $\square$

Hence, in all the examples above, $\mathcal{E}$ is self-adjoint and positive definite, and so $\mathcal{E}^{-1}\mathcal{F}$ is positive definite for the primal, dual, HRVW-KSH, and dual HRVW-KSH directions.

This omits the AHO direction, which may not be well-defined, as we will see in the following example.

Example 1

$$m = 1,\ A_1 = \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix},$$

$$X = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix},\ S = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix}.$$

Then, $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is a 1×1 matrix. We start by computing $\mathcal{F}\mathcal{A}^*$ as follows.

$$\begin{aligned} \mathcal{F}\mathcal{A}^* = \mathcal{F}A_1 &= \frac{1}{2}(XA_1 + A_1X) \\ &= \frac{1}{2}\left(\begin{pmatrix} 1 & \sqrt{2} \\ 2\sqrt{2} & 2 \end{pmatrix} + \begin{pmatrix} 1 & 2\sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}\right) \\ &= \begin{pmatrix} 1 & 3/\sqrt{2} \\ 3/\sqrt{2} & 2 \end{pmatrix}. \end{aligned}$$

Next, we compute $\mathcal{E}^{-1}\mathcal{F}A_1$. Recall that $\mathcal{E} = S \odot I$. So,

$$\begin{aligned} (S \odot I)\begin{pmatrix} a & b \\ b & c \end{pmatrix} &= \frac{1}{2}\left(S\begin{pmatrix} a & b \\ b & c \end{pmatrix} + \begin{pmatrix} a & b \\ b & c \end{pmatrix}S\right) \\ &= \frac{1}{2}\left(\begin{pmatrix} a & b \\ 11b & 11c \end{pmatrix} + \begin{pmatrix} a & 11b \\ b & 11c \end{pmatrix}\right) \\ &= \begin{pmatrix} a & 6b \\ 6b & 11c \end{pmatrix}, \end{aligned}$$

which means that

$$(S \odot I)^{-1}\mathcal{F}A_1 = \begin{pmatrix} 1 & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{2}{11} \end{pmatrix},$$

obtained by reversing the action of $(S \odot I)$, by dividing the off-diagonal entries by 6 and the $(2, 2)$ entry by 11.

Hence,

$$A_1 \bullet \mathcal{E}^{-1}\mathcal{F}A_1 = \text{trace}\left(\begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{2}{11} \end{pmatrix}\right) = 0.$$

Moreover, the corresponding right-hand side is $-\mathcal{A}\mathcal{E}^{-1}R_{EF}$, which turns out to be nonzero for any positive ν. Therefore, there is no solution to the system.

[In fact, the AHO direction is well-defined if $(X, y, S) \in \mathcal{N}_\infty(\beta)$ for $\beta < \frac{1}{\sqrt{2}}$, as we shall see.]

We now see how all of our directions can be viewed in a unified framework, as follows:

- First, apply a similarity transformation to the last equation, to get:

$$P\tilde{X}\tilde{S}P^{-1} = \nu I,$$

  where $P \in \mathbb{R}^{n\times n}$ is invertible.

- Then, symmetrize the left-hand side, to get

$$\frac{1}{2}P\tilde{X}\tilde{S}P^{-1} + \frac{1}{2}P^{-T}\tilde{S}\tilde{X}P^{T} = \nu I.$$

  Note that this also defines the central path, since $P\tilde{X}\tilde{S}P^{-1}$ is similar to $\tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2}$, which is symmetric. So, it has all real eigenvalues, and we can use the same argument we used for the AHO direciton.

- Now, linearize at the current iterates to get

$$\frac{1}{2}P\Delta XSP^{-1} + \frac{1}{2}P^{-T}S\Delta XP^{T} + \frac{1}{2}PX\Delta SP^{-1} + \frac{1}{2}P^{-T}\Delta SXP^{T}$$
$$= \nu I - \frac{1}{2}PXSP^{-1} - \frac{1}{2}P^{-T}SXP^{T},$$

  or,

$$(P \odot P^{-T}S)(\Delta X) + (PX \odot P^{-T})(\Delta S) = \nu I - \frac{1}{2}PXSP^{-1} - \frac{1}{2}P^{-T}SXP^{T}.$$

  Or, pre-multiplying by P^T and post-multiplying by P:

$$(M \odot S)(\Delta X) + (MX \odot I)(\Delta S) = \nu M - \frac{1}{2}MXS - \frac{1}{2}SXM,$$

  where $M := P^TP \succ 0$.

  We observe that ΔX and ΔS depend on P only through $M := P^TP$. So, we can assume that $P = M^{1/2} \succ 0$.

  Note that

  - If $M = I,\ P = I$, then we get the AHO direction.
  - If $M = X^{-1},\ P = X^{-1/2}$, then we get the dual HRVW-KSH direction.
  - If $M = S,\ P = S^{1/2}$, then we get (after pre and post-multiplying by S^{-1}) the HRVW-KSH direction.

This way of deriving the HRVW-KSH directions is due to Monteiro. So, we'll use "HKM" and "dual HKM" to refer to these directions from now on.

Zhang developed the approach for general P, so any such direction is in the MZ (Monteiro-Zhang) family.

Note: If we scale the problems as follows:

$$\begin{array}{rcl}\tilde{X} & \rightarrow & \hat{X} = P\tilde{X}P^{T},\\ \tilde{S} & \rightarrow & \hat{S} = P^{-T}\tilde{S}P^{-1},\end{array}$$

then $\tilde{X}\tilde{S}$ is transformed to $\hat{X}\hat{S} = P\tilde{X}\tilde{S}P^{-1}$. Hence, the MZ approach can be viewed as scaling $\tilde{X}$ to $\hat{X} = P\tilde{X}P^T$ and $\tilde{S}$ to $\hat{S} = P^{-T}\tilde{S}P^{-1}$, applying the AHO formula in the scaled space, and then scaling the directions back to get

$$\begin{aligned} \Delta X &= P^{-1}\widehat{\Delta X}P^{-T}, \\ \Delta S &= P^{-1}\widehat{\Delta S}P. \end{aligned}$$

Remark 1 *The AHO scaling leaves X and S unchanged. The HKM scaling sends X to $S^{1/2}XS^{1/2}$ and S to I. The dual HKM scaling sends X to I and S to $X^{1/2}SX^{1/2}$. Also note that the last two directions make the scaled iterates commute.*

Another viewpoint on scaling

Think of X as the matrix representation of a self-adjoint, positive definite linear operator $\chi : V \to V^*$ from an n-dimensional real vector space V into its dual, V^*. Then, let $\langle \cdot, \cdot \rangle \to \mathbb{R}$ be the pairing of V^* and V, so

$$\langle \chi v, \bar{v} \rangle = \langle \chi \bar{v}, v \rangle$$

and

$$\langle \chi v, v \rangle > 0, \ \forall v \neq 0,$$

since χ is self-adjoint and positive definite.

Choose a basis $(b_1, \ldots, b_n)$ in V and let X be the matrix with entries:

$$x_{ij} := \langle \chi b_j, b_i \rangle$$

for all i, j.

If instead we use the basis $(c_1, \ldots, c_n)$ with

$$c_i = \sum_{k=1}^{n} p_{ik} b_k,$$

where $P = (p_{ik})$ is an invertible matrix, then the new representation turns out to be PXP^T.

Similarly, view S as the matrix representation of $\sigma : V^* \to V$, also self-adjoint and positive definite, with

$$s_{ij} = \langle b_i^*, \sigma b_j^* \rangle,$$

where $(b_1^*, \ldots, b_n^*)$ is a basis for V^*. In particular, we choose $b_1^*, \ldots, b_n^*$ as the dual basis, with:

$$\langle b_i^*, b_j \rangle = \delta_{ij} = \left\{ \begin{array}{ll} 0, & i \neq j \\ 1, & i = j \end{array} \right. .$$

Then, under the corresponding change of dual bases, S transforms to $P^{-T}SP^{-1}$.

Also, XS is a matrix representation of $\chi\sigma$, where

$$\chi\sigma : V^* \to V^*, \ \sigma\chi : V \to V.$$

So, trace $(\chi\sigma)$ = trace $(\sigma\chi)$ makes sense, but $\chi\sigma + \sigma\chi$ does not!