Scribe: Zhe Yu

$$(P) \quad \begin{array}{rrcl} \min_X & C \bullet X & & \\ & A_i \bullet X & = & b_i, \quad i = 1, \dots, m, \\ & X & \succeq & 0, \end{array}$$

$$(D) \quad \begin{array}{rrcccl} \max_{y,S} & b^T y & & & & \\ & \sum_i y_i A_i & + & S & = & C, \\ & & & S & \succeq & 0. \end{array}$$

Theorem 1 *If $\mathcal{F}(P)$,$\mathcal{F}^\circ(D)$ are nonempty, then*
(a) The set of optimal solutions to (P) is nonempty and compact.
(b) There is no duality gap.

To reverse this result, we show how to write the primal as the dual and vice-versa. Firstly, we get rid of some trivial cases:

We can suppose there is some $H \in \mathbb{M}^n$ with $\mathcal{A}H = b$. Otherwise, (P) is infeasible. Also, there is some y with $\mathcal{A}^* y = 0$ but $b^T y \neq 0$, w.l.o.g., $b^T y > 0$. So, if (D) is feasible, it is unbounded.

Also, if the A_i's are linearly dependent, then we can remove redundant constraints in (P). In addition, with $(\exists H : \mathcal{A}H = b)$, $\mathcal{A}^* y = 0 \Rightarrow b^T y = 0$. So we can correspondingly eliminate "redundant" A_i's and the correspondingly y_i's and b_i's from (D).

Let $\mathfrak{A} = \text{span}\{A_1, \dots, A_m\} \subseteq \mathbb{M}^n$. Then $A_i \bullet X = b_i$, all i, if and only if $X - H \in \mathfrak{A}^\perp$. We let $F_1, F_2, \dots, F_l$ be a basis for $\mathfrak{A}^\perp$. (If the A_i's are linearly independent, $l = n(n+1)/2 - m$.)

Then we can rewrite (P) as following

$$\begin{array}{rl} (P) & \min\{C \bullet X : A_i \bullet X = b_i, i = 1, \dots, m, X \succeq 0\} \\ \equiv & \min_{z,X}\{C \bullet (H - \sum_{k=1}^l z_k F_k) : X = H - \sum_{k=1}^l z_k F_k, X \succeq 0\} \\ \equiv & C \bullet H - \max_{z,X}\{\sum_k g_k z_k : \sum z_k F_k + X = H, X \succeq 0\} \end{array}$$

with $g_k = C \bullet F_k, k = 1 \dots, l$, so (P) is equivalent to a problem is standard dual form.

Similarly:

$$\begin{array}{rl} (D) & \max_{y,S}\{b^T y : \sum y_i A_i + S = C, S \succeq 0\} \\ \equiv & \max_{y,S}\{(\mathcal{A}H)^T y : S = C - \mathcal{A}^* y, S \succeq 0\} \\ \equiv & \max_{y,S}\{H \bullet \mathcal{A}^* y : S = C - \mathcal{A}^* y, S \succeq 0\} \\ \equiv & \max_{y,S}\{H \bullet (C - S) : S - C \in \mathfrak{A}, S \succeq 0\} \\ \equiv & \max_{y,S}\{H \bullet (C - S) : F_k \bullet (S - C) = 0, k = 1, \dots, l, S \succeq 0\} \\ \equiv & H \bullet C - \min_S\{H \bullet S : F_k \bullet S = g_k, k = 1, \dots, l, S \succeq 0\}, \end{array}$$

a problem in standard primal form.

Corollary 1 *If $\mathcal{F}^\circ(P)$, $\mathcal{F}(D)$ are nonempty, then*
(a) (D) has a nonempty set of optimal solutions and the set of optimal S's is compact; and
(b) there is no duality gap.

Corollary 2 *If both $\mathcal{F}^\circ(P)$, $\mathcal{F}^\circ(D)$ are nonempty, then both problems have optimal solutions and there is no duality gap.*

If this is the case, X and (y, S) are optimal if and only if

$$\begin{array}{rll} \mathcal{A}^* y + S & = C, & S \succeq 0 \\ \mathcal{A} \bullet X \qquad & = b, & X \succeq 0 \\ X \bullet S & = 0 & \end{array}$$

or the last equation can be replaced by $XS = 0$.

Let's assume we have optimal solutions $X_*, (y_*, S_*)$ and there is no duality gap. (We'll say strong duality holds.)

What can we say about the ranks r_* and s_* of X_* and S_*?

Proposition 1 $r_* + s_* \leq n$.

Proof: : We know $X_* S_* = 0$. So X_* and S_* commute, and they can be simultaneously diagonalized.

$$X_* = Q \Lambda_* Q^T, \qquad S_* = Q \Omega_* Q^T$$

with $\Lambda_* = \begin{bmatrix} \hat{\Lambda} & 0 \\ 0 & 0 \end{bmatrix}$, with $\hat{\Lambda} = \mathbb{M}^{r_*}_{++}$. Then

$$X_* S_* = Q \Lambda_* (Q^T Q) \Omega_* Q^T = Q \begin{bmatrix} \lambda_1 \omega_1 & & \\ & \ddots & \\ & & \lambda_n \omega_n \end{bmatrix} Q^T.$$

Since $X_* S_* = 0$, $\omega_1 = \ldots = \omega_{r_*} = 0$, and so $s_* \leq n - r_*$. $\square$

Note: in LP, we have strict complementarity, i.e., there are an optimal x_* and an optimal slack s_* with either $x_{*j} > 0$ or $s_{*j} > 0$ for all j.

However there are SDPs with $r_* + s_* < n$ for all optimal solutions.

Proposition 2 *Suppose X is an extreme point of (P) with rank$(X) = r$. Then $r(r+1)/2 \leq m$.*

Proof: : Suppose $S \in \mathcal{F}(P)$ has rank r with $r(r+1)/2 > m$.

Write $X = Q \Lambda Q^T$, where

$$\Lambda = \begin{bmatrix} \hat{\Lambda} & 0 \\ 0 & 0 \end{bmatrix},$$

$\hat{\Lambda} \in \mathbb{M}^r_{++}$.

Recall that (P) is equivalent to $(\bar{P})$, defined by

$$\bar{A}_i = Q^T A_i Q, i = 1, \ldots, m, \qquad \bar{C} = Q^T C Q.$$

$\bar{X}$ is feasible in $(\bar{P})$ if and only if $X = Q\bar{X}Q^T$ is feasible in (P), and so Λ is feasible in $\bar{P}$.

Suppose

$$\bar{A}_i = \begin{bmatrix} \hat{A}_i & B_i^T \\ B_i & D_i \end{bmatrix}$$

with $\hat{A}_i \in \mathbb{M}^r$.

Then $\hat{A}_i \bullet \hat{\Lambda} = b_i, i = 1, \ldots, m$.

So by counting dimensions, there is $\hat{H} \in \mathbb{M}^r \setminus \{0\}$ with $\hat{A} \bullet \hat{H} = 0, i = 1, \ldots, m$.

Then $\begin{bmatrix} \hat{\Lambda} \pm \varepsilon H & 0 \\ 0 & 0 \end{bmatrix}$ is feasible in $(\bar{P})$ for all sufficient small ε.

Hence, $Q \begin{bmatrix} \hat{\Lambda} \pm \varepsilon H & 0 \\ 0 & 0 \end{bmatrix} Q^T$ is feasible in (P) for all sufficient small ε.

So X is not an extreme point. □

Corollary 3 *Suppose (P) has an optimal solution; then it has one with rank at most r, with r the largest integer s.t. $r(r+1)/2 \leq m$.*

Proof: : Let X be an optimal solution of minimum rank, say $\bar{r}$ with $\bar{r}(\bar{r}+1)/2 > m$. Then, as in the proof above, we find $H \in \mathbb{M}^n, H \neq 0$ with $X \pm \varepsilon H$ also feasible, and hence also optimal.

By choosing an appropriate sign and then ε as large as possible, we obtain an optimal solution of smaller rank. □

For example, in the SDP relaxation for Max Cut, we have n equality constraints so there is an optimal solution with rank r, $r(r+1)/2 \leq n$, so $r \leq \sqrt{2n}$.

Note that such low-rank solutions may not respect the block-diagonal structure of the problem. (E.g., if C and all A_i's are diagonal, the only diagonal solution may have rank n.)

We can sometimes get conditions by applying these results blockwise.

The Log Barrier Function:

Define

$$F(X) = -\text{lndet}(X) := \begin{cases} -\ln\det X & \text{if } X \succ 0, \\ +\infty & \text{o.w.} \end{cases}$$

Note that $-\text{lndet}(X) = -\sum_j \ln \lambda_j(X)$, and that F "respects the structure" of X.

Proposition 3 *(a) If $X = \text{Diag}\,(X_1, \ldots, X_p)$, $F(X) = \sum_{i=1}^p F(X_i)$.*
(b) If $X = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$, then $F(X) = F(A) + F(C - BA^{-1}B^T) = F(C) + F(A - B^T C^{-1} B)$.