Dealing with no known feasible solution in interior-point methods

1) Just take a "Newton" step to approximate a point on the central path, even if the current iterate is not feasible. I.e., the rhs of the first two sets of eqs become nonzero:

$$C - \mathcal{A}^* y - S$$
$$b - \mathcal{A}X$$

These extensions work well in practice. One difficulty in the analysis: we don't have $\Delta X \bullet \Delta S = 0$ anymore, but there are polynomial bounds under suitable assumptions.

2) Reduce to the feasible case using an extension of the homogeneous self-dual approach of Ye, Todd, Mizuno for LP. Consider the (homogeneous, self-dualP problem:

$$\begin{array}{llllllll}
\min & & & & (n+1)\theta & & \\
S = & & -\mathcal{A}^* y & +C\tau & -\bar{C}\theta & \succeq & 0 \\
 & \mathcal{A}X & & -b\tau & +\bar{b}\theta & = & 0 \\
\kappa = & -C \bullet X & +b^T y & & +\bar{z}\theta & \geq & 0 \\
 & \bar{C} \bullet X & -\bar{b}^T y & -\bar{z}\tau & & = & -(n+1) \\
 & X \succeq 0, \tau \geq 0 & & & , & &
\end{array} \tag{HSDP}$$

where $\bar{C} = C - I, \bar{b} = b - \mathcal{A}I, \bar{z} = C \bullet I + 1$, so $(X = I, y = 0, \tau = 1, \theta = 1)$ is strictly feas with $S = I, \kappa = 1$.

Theorem 1 *a)* (HSDP) *is equivalent to its dual, has a strictly feasible solution and has an optimal solution with optimal value 0.*

b) If (HSDP) *has an optimal solution* $(X^*, y^*, S^*, \tau^*, \kappa^*, 0)$ *with* $\tau^* > 0$*, then* $\kappa^* = 0$*, and* $x^*/\tau^*, (y^*/\tau^*, S^*/\tau^*)$ *are optimal in (P) and (D) respectively.*

c) If (HSDP) *has an optimal solution* $(X^*, y^*, S^*, \tau^*, \kappa^*, 0)$ *with* $\kappa^* > 0$*, then* $\tau^* = 0$ *and either* (y^*, S^*) *shows (P) infeasible or* X^* *shows (D) infeasible, or both.*

Key fact: Weak path following restrictions imply that, if there is an optimal solution with $\tau(\kappa)$ positive, then the iterations converge to such a solution.

Since the constraint "matrix" is skew-symmetric,

$$-X \bullet S - \tau\kappa = -(n+1)\theta$$

and combining the equations gives

$$I \bullet X + I \bullet S + \tau + \kappa = (n+1)\theta.$$

Applying a path-following algorithm to (HSDP) is slightly complicated. We want to keep the primal and dual iterates the same, and the directions the same, and then we can use a linear system only a little larger than that used for (P) and (D).

The spectral bundle method (Helmberg&Rendl)

Consider

$$\begin{array}{lll} \max & C \bullet X & \\ \text{s.t.} & \mathcal{A}X & = 0 \\ & I \bullet X & = 1 \\ & X & \succeq 0 \end{array} \tag{P}$$

and its dual

$$\begin{array}{ll} \min & \lambda \\ \text{s.t.} & \lambda I \succeq C - \mathcal{A}^* y \end{array} \tag{D}$$

or

$$\min_y \quad \underbrace{\lambda_{\max}(C - \mathcal{A}^* y).}_{:=f(y),\text{convex and nonsmooth}}$$

Note that (P) appears a rather special SDP problem; but HW4 shows that it is in fact quite general, only requiring that the feasible region of the primal problem be bounded. HW4 also shows that, if q is a unit eigenvector corresponding to the largest eigenvalue of $C - \mathcal{A}^* y$, then $-\mathcal{A}(qq^T)$ is a subgradient of f at y, and if q is a unit vector with $q^T(C - \mathcal{A}^* y)q \geq f(y) - \epsilon$, then $-\mathcal{A}(qq^T)$ is an ϵ-subgradient of f at y.

$$\begin{aligned} (D) &\equiv \min_y \max_{qq^T=1} (C - \mathcal{A}^* y) \bullet qq^T \\ &\equiv \min_y \max_{X, I \bullet X = 1, X \succeq 0} (C - \mathcal{A}^* y) \bullet X. \end{aligned}$$

We proceed as follows:

1) Restrict the set of X's:

$$X = \lambda W + PVP^T$$

where $P \in \mathbb{R}^{n \times r}$ has orthonormal columns, $W \in \mathbb{M}^n_+$ has $I \bullet W = 1$, $\lambda \geq 0$, and $V \in \mathbb{M}^r_+$, where $\lambda + I \bullet V = 1$. This gives

$$\min_y \max_{\lambda \geq 0, V \succeq 0, \lambda + I \bullet V = 1} (C - \mathcal{A}^* y) \bullet (\lambda W + PVP^T).$$

If V is further restricted to be diagonal, this max is

$$\max\{(C - \mathcal{A}^* y) \bullet W, \max_i (C - \mathcal{A}^* y) \bullet p_i p_i^T\},$$

a piecewise-linear underestimate of f.

2) Add a proximal term:
If our current y is z, then consider

$$\min_y \max_{\lambda, V} (C - \mathcal{A}^* y) \bullet (\lambda W + PVP^T) + \frac{\sigma}{2}\|y - z\|^2.$$

Under standard assumptions this is equivalent to

$$\max_{\lambda, V} \min_y (C - \mathcal{A}^* y) \bullet (\lambda W + PVP^T) + \frac{\sigma}{2}\|y - z\|^2,$$

giving $y = z + \frac{1}{\sigma}\mathcal{A}(\lambda W + PVP^T)$. This gives

$$\max_{\lambda \geq 0, V \succeq 0, \lambda + I \bullet V = 1} (C - \mathcal{A}^* y) \bullet (\lambda W + PVP^T) - \frac{1}{2\sigma}\|\mathcal{A}(\lambda W + PVP^T)\|^2, \tag{1}$$

which is a quadratic SDP with "order" $r + 1$.

Algorithm 1 Spectral Bundle Method

Given $C, A_1, \ldots, A_m, z = z_0 \in \mathbb{R}^m$
Evaluate $f(z)$ and get a corresponding eigenvector v
Set $W = I/n$, say, and $P = [v]$.
while termination criteria not satisfied **do**
  Solve (1) with the current z, W, P to get λ, V, y
  Evaluate $f(y)$, also getting a corresponding eigenvector v
  Update W, P, z:

$W := \lambda W + PVP^T$
$P :=$ columns of PQ_1, where Q_1 has as columns the top few eigenvectors of V, and v
$z := y$ if $f(y) < f(z)$

end while