Theorem 1 Suppose $E$, $F$ and $I$ are linear operators mapping $S^R_{n \times n}$ to itself, with $E$ invertible and $I$ the identity mapping. Also assume that $A_1, A_2, \ldots, A_m$ are linearly independent. Then if $E^{-1}F$ is positive definite (not necessarily self-adjoint), there is a unique solution to

\[
\begin{pmatrix}
0 & A^* & I \\
A & 0 & 0 \\
E & 0 & F
\end{pmatrix}
\begin{pmatrix}
U \\
v \\
W
\end{pmatrix}
= 
\begin{pmatrix}
P \\
q \\
R
\end{pmatrix}
\quad (\star)
\]

given by

\[v = (AE^{-1}FA^*)^{-1}(q - AE^{-1}(R - FP)),\]
\[W = P - A^*v,\]
\[U = E^{-1}(R - FW).\]

**Proof:** If $(\star)$ has a solution, it must satisfy $W = P - A^*v$, and $U = E^{-1}(R - FW) = E^{-1}(R - FP + FA^*v)$ and using the second set of equations

\[A(E^{-1}(R - FP) + E^{-1}FA^*v) = q\]

or

\[AE^{-1}FA^*v = q - AE^{-1}(R - FP).\]

But if $AE^{-1}FA^*y = 0$ then we must have $y^TAE^{-1}FA^*y = 0$, i.e. $(A^*y) \cdot (E^{-1}FA^*y) = 0$ which implies $A^*y = 0$, so $y = 0$ (since $E^{-1}F$ is pd). Hence $AE^{-1}FA^*$ is invertible. So the solution, if it exists, is unique. We can check that this solution indeed solves $(\star)$. □

Corollary 1 If $F^o(P)$ and $F^o(D)$ are non-empty and the $A_i$’s are linearly independent, then \{$(X(\mu), y(\mu), S(\mu))$\} forms a smooth path lying in $F^o := F^o(P) \times F^o(D)$.

**Proof:** The proof uses the implicit function theorem. Since we can choose $E(H) := \mu X^{-1}HX^{-1}$ and $F(H) := H$, we have $E^{-1}(H) = \mu^{-1}XHX$, and $V \cdot E^{-1}FV = \mu^{-1}V \cdot XVX = \mu^{-1}\|X^{-1/2}VX^{-1/2}\|^2_F > 0$ if $V \neq 0$. □

We want to follow this path as $\mu$ approaches 0.
We want to measure the “proximity” of our iterates to the central path. Since we’ll concentrate on feasible algorithms, we want to measure the deviation of $XS$ from $\mu I$ for some positive $\mu > 0$. Given $pd X, S$ define $\mu := \mu(X, S) = \frac{S \cdot X}{n}$ and

$$d_F(X, S) = \|X^{1/2}SX^{1/2} - \mu I\|_F = (\text{trace}(X^{1/2}SX^{1/2} - \mu X^{1/2}SX^{1/2} + \mu^2 I))^{1/2} = (\text{trace}((S - \mu X^{-1})X(S - \mu X^{-1})X))^{1/2} = \|S - \mu X^{-1}\|_X^* = \|S + \mu F'(X)\|_S,$$

also

$$d_F(X, S) = (\text{trace}(X^{1/2}S^{1/2}SX^{1/2} - \mu X^{1/2}S^{1/2}S^{1/2}X^{1/2} + \mu^2 I))^{1/2} = (\text{trace}(S^{1/2}XS^{1/2} - \mu S^{1/2}XS^{1/2} + \mu^2 I))^{1/2} = \|S^{1/2}XS^{1/2} - \mu I\|_F = \|X + \mu F'(S)\|_S.$$ 

So this is symmetric in $X$ and $S$.

For any $\beta > 0$, we can define

$$\mathcal{N}_F(\beta) := \{(X, y, S) \in \mathcal{F}^0 : d_F(X, S) \leq \beta \mu(X, S)\}$$

which is the same as requiring $\|\lambda(X^{1/2}SX^{1/2}) - \mu e\|_2 \leq \mu \beta$. We could also require

$$\mathcal{N}_\infty(\beta) := \{(X, y, S) \in \mathcal{F}^0 : \|\lambda(X^{1/2}SX^{1/2}) - \mu e\|_\infty \leq \mu \beta\}$$

($\beta < 1$) which would be a weaker assumption and hence define a larger neighborhood or even use

$$\mathcal{N}_{-\infty}(\beta) := \{(X, y, S) \in \mathcal{F}^0 : \max_j \mu - \lambda_j(X^{1/2}SX^{1/2}) \leq \mu \beta\}$$

which would cover everything as $\beta$ goes to 1.

**Framework for a primal-dual path-following interior-point algorithm:**

Generate a sequence of iterates $\{(X_k, y_k, S_k) : k \geq 0\} \in \mathcal{F}^0$ as follows: start with $(X_0, y_0, S_0) \in \mathcal{N}_F(\beta)$ for some $0 < \beta < 1$, $\epsilon > 0$, and set $k = 0$. At iteration $k$, stop if $S_k \cdot X_k \leq \epsilon$. Otherwise, determine search directions $(\Delta X_k, \Delta y_k, \Delta S_k) \in SR^{n \times n} \times \mathbb{R}^n \times SR^{n \times n}$ and choose step length $\alpha_k > 0$ so that $(X_{k+1} := X_k + \alpha_k \Delta X_k, y_{k+1} := y_k + \alpha_k \Delta y_k, S_{k+1} := S_k + \alpha_k \Delta S_k) \in \mathcal{N}_F(\beta)$, increase $k$ and repeat.

The search directions are designed to help us get close to the central path, in particular to $(X(\nu), y(\nu), S(\nu))$, where $\nu = \sigma \mu, \mu = \mu(X_k, S_k), 0 \leq \sigma \leq 1$. We’ll use a Newton-like algorithm: Assume that the current iterate is $(X, y, S) \in \mathcal{F}^0$ and use $(\tilde{X}, \tilde{y}, \tilde{S})$ for generic points in the appropriate space. We want to take a Newton step for

$$A^* \tilde{y} + \tilde{S} = C,$$

symmetrion of $(\tilde{X} \tilde{S} = \nu I)$. 

We’ll follow the approach of Monteiro (for two special cases) and Zhang (for the general case): apply a similarity transformation to the last set of equations $\tilde{X} \tilde{S} = \nu I$. 

2
First, a simple symmetrization: what if we require

\[ \bar{S}\bar{X}/2 + \bar{X}\bar{S}/2 = \nu I? \quad (\dagger) \]

If this holds, \( \bar{X}\bar{S} = \nu I + \) a skew-symmetric matrix. But the eigenvalues of this skew-symmetric matrix are all real and all purely imaginary at the same time, hence they are equal to 0. We could apply Newton’s method to the new system with \((\dagger)\) as its last set of equations.

Let \( M(\bar{X}, \bar{S}) = \bar{X}\bar{S}/2 + \bar{S}\bar{X}/2 - \nu I \). Taking the derivatives \( E := DXM(\bar{X}, \bar{S})(X, S) \) and \( F := DS(M(\bar{X}, \bar{S}))(X, S) \), we get \( E(H) = HS/2 + SH/2 = S \odot I(H) \) and similarly \( F = X \odot I \).

Here, for matrices \( P, Q \in \mathbb{R}^{n \times n} \) we define \( P \odot Q := \mathbb{R}^{n \times n} \to S \mathbb{R}^{n \times n} \) by \( (P \odot Q)K = \frac{PKQ^T + QK^TP^T}{2} \).

We want \( E \) invertible and \( E^{-1}F \) pd. Note that we can have \( X \) and \( S \) pd but \( E^{-1}F \) not pd (see Todd, Toh, and Tutuncu). This leads to the algorithm of Alizadeh, Haeberly, and Overton. (Hence the algorithm may fail, and furthermore even if it doesn’t fail, each iteration is expensive since we need to invert \( E \!).

Here is an alternative: We can solve the Newton system ignoring symmetry and make sure the search directions are symmetric afterwards. \( \Delta S \) is automatically symmetric since it equals \(-A^*\Delta y \), so we need to symmetrize \( \Delta X \).

We have

\[ \Delta XS + X\Delta S = \nu I - XS, \text{ or} \]
\[ \Delta X + X\Delta SS^{-1} = \nu S^{-1} - X; \]

hence instead we can require

\[ \Delta X + 1/2(X\Delta SS^{-1} + S^{-1}\Delta SX) = \nu S^{-1} - X. \]

This leads to choosing \( E = I \) and \( F = X \odot S^{-1} \), hence we have \( E^{-1}F = F = X \odot S^{-1} \) and

\[ U \cdot E^{-1}FU = 1/2(U \cdot (XUS^{-1} + S^{-1}UX)) \]
\[ = \text{trace}(UXUS^{-1}) = \text{trace}(X^{1/2}US^{-1/2}S^{-1/2}UX^{1/2}) \]
\[ = \|X^{1/2}US^{-1/2}\|_F^2 > 0, \]

if \( U \neq 0 \). Note that \( E^{-1}F \) is a positive definite self-adjoint operator. This method of computing the search directions is due independently to Helmberg, Rendl, Vanderbei, and Wolkowicz and to Kojima, Shindoh, and Hara. We’ll see next time how we can view this as applying a similarity transformation and then Newton’s method.