1 Semidefinite Programming

Primal:

$$\sup_{Y \in \mathcal{R}^g,\ S \in \mathcal{SR}^q} \quad \langle B, Y \rangle : \qquad A(Y) + S = C, \quad S \succeq 0$$

Dual:

$$\inf_{X \in \mathcal{SR}^q} \quad \langle C, X \rangle : \quad A^*(X) = B, \quad X \succeq 0$$

1.1 Self-Dual Embedding

The discussion in this section is needed in case there is not a feasible initial point on the central available to initialize the interior-point algorithms that will be discussed later. When such a point is available we would have the matrices $\bar{B}$ and $\bar{C}$ (defined below) reduce to zero with significant simplifications to the overal computations. These simplifications parallel the talk given by Mauricio in June 2011.

Define

$$\begin{aligned} \bar{B} &:= B - A^*(I), & \bar{z} &:= 1 + \langle C, I \rangle - \langle B, Y_0 \rangle, \\ \bar{C} &:= C - A(Y_0) - I, & \beta &:= n + 1 \end{aligned}$$

The following semidefinite program

$$\begin{array}{rrrrrrcr} & & & \inf \theta\beta & & & & \\ & A^*(X) & -\tau B & +\theta \bar{B} & & & = & 0 \\ -A(Y) & & +\tau C & -\theta \bar{C} & -S & & = & 0 \\ \langle B, Y \rangle & -\langle C, X \rangle & & +\theta \bar{z} & & -\rho & = & 0 \\ -\left\langle \bar{B}, Y \right\rangle & +\left\langle \bar{C}, X \right\rangle & -\tau \bar{z} & & & & = & -\beta \end{array}$$

where

$$\begin{aligned} X &\succeq 0, & \tau &\geq 0, \\ S &\succeq 0, & \rho &\geq 0, \end{aligned}$$

is self-dual. Furthermore one can verify that

$$Y = Y_0, \qquad X = S = I, \qquad \tau = \theta = \rho = 1$$

is feasible because

$$\begin{aligned} A^*(I) - B + \bar{B} = 0, \\ -A(Y_0) + C - \bar{C} - I = 0, \\ \langle B, Y_0 \rangle - \langle C, I \rangle + \bar{z} - 1 = 0, \end{aligned}$$

and

$$\begin{aligned} \beta &= \left\langle \bar{B}, Y_0 \right\rangle - \left\langle \bar{C}, I \right\rangle + \bar{z}, \\ &= \langle B - A^*(I), Y_0 \rangle - \langle C - A(Y_0) - I, I \rangle + 1 + \langle C, I \rangle - \langle B, Y_0 \rangle, \\ &= -\langle A^*(I), Y_0 \rangle + \langle A(Y_0) + I, I \rangle + 1, \\ &= \langle I, I \rangle + 1 = n + 1. \end{aligned}$$

1.2 Central Path

As the embedded primal and dual programs are the same we can limit our attention to either the primal or dual central path. For instance, the embedded primal central path is the curve

$$\begin{array}{rrrrrrcl}
 & A^*(X) & -\tau B & +\theta\bar{B} & & & = & 0 \\
-A(Y) & & +\tau C & -\theta\bar{C} & -S & & = & 0 \\
\langle B, Y\rangle & -\langle C, X\rangle & & +\theta\bar{z} & & -\rho & = & 0 \\
-\left\langle \bar{B}, Y\right\rangle & +\left\langle \bar{C}, X\right\rangle & -\tau\bar{z} & & & & = & -\beta
\end{array}$$

$$XS = \mu I, \quad X \succ 0, \quad S \succ 0$$
$$\tau\rho = \mu, \qquad \tau > 0, \quad \rho > 0$$

TODO: ADD COMMENT ON WHAT IS DISPLAYED IN THE COLUMNS OF THE CURRENT IMPLEMENTATION OF THE ALGORITHM.

Let

$$\mathcal{S}_P(U) := \frac{1}{2}\left(PUP^{-1} + P^{-T}U^TP^T\right). \tag{1.2.1}$$

The key set of equations is the linearized central-path, namely

$$\begin{array}{rrrrrrcl}
 & A^*(\Delta_X) & -\delta_\tau B & +\delta_\theta\bar{B} & & & = & 0 \\
-A(\Delta_Y) & & +\delta_\tau C & -\delta_\theta\bar{C} & -\Delta_S & & = & 0 \\
\langle B, \Delta_Y\rangle & -\langle C, \Delta_X\rangle & & +\delta_\theta\bar{z} & & -\delta_\rho & = & 0 \\
-\left\langle \bar{B}, \Delta_Y\right\rangle & +\left\langle \bar{C}, \Delta_X\right\rangle & -\delta_\tau\bar{z} & & & & = & 0
\end{array} \tag{1.2.2}$$

$$\begin{gathered}
\mathcal{S}_P(\Delta_X S_k + X_k\Delta_S) = \sigma_k\mu_k I - \mathcal{S}_P(X_kS_k), \\
\rho_k\delta_\tau + \tau_k\delta_\rho = \sigma_k\mu_k - \tau_k\rho_k
\end{gathered} \tag{1.2.3}$$

where $\sigma_k \in [0, 1]$, and most of this note is devoted on how to solve it. In case of an available initial feasible solution on the central path we can have $\delta_\tau = \delta_\theta = 0$ and the third and fourth equations in (1.2.2) and the second equation in (1.2.3) can be dropped out.

When working with predictor-corrector methods we will modify equation (1.2.3) to include an extra term

$$\mathcal{S}_P(\Delta_X S_k + X_k\Delta_S) = \sigma_k\mu_k I - \mathcal{S}_P(X_kS_k + \tilde{\Delta}_X\tilde{\Delta}_S), \tag{1.2.4}$$
$$\rho_k\delta_\tau + \tau_k\delta_\rho = \sigma_k\mu_k - \tau_k\rho_k - \tilde{\delta}_\tau\tilde{\delta}_\rho \tag{1.2.5}$$

where $\tilde{\Delta}_X$, $\tilde{\Delta}_S$, $\tilde{\delta}_\tau$ and $\tilde{\delta}_\rho$ will be given matrices. The standard central path is recovered by setting these matrices to zero.

Note that equation (1.2.4) can be spelled out as

$$\begin{aligned}
&P(\Delta_X S_k + X_k\Delta_S)P^{-1} + P^{-T}(S_k\Delta_X + \Delta_S X_k)P^T = \\
&\qquad 2\sigma_k\mu_k I - PX_kS_kP^{-1} - P^{-T}S_kX_kP^T - P\tilde{\Delta}_X\tilde{\Delta}_SP^{-1} - P^{-T}\tilde{\Delta}_S\tilde{\Delta}_XP^T.
\end{aligned}$$

Equivalently, if we multiply by P^T on the left and by P^{-T} on the right we obtain

$$\begin{aligned}
&Q(\Delta_X S_k + X_k\Delta_S)Q^{-1} + (S_k\Delta_X + \Delta_S X_k) = \\
&\qquad 2\sigma_k\mu_k I - QX_kS_kQ^{-1} - S_kX_k - Q\tilde{\Delta}_X\tilde{\Delta}_SQ^{-1} - \tilde{\Delta}_S\tilde{\Delta}_X,
\end{aligned} \tag{1.2.6}$$

which depends exclusively on the symmetric matrix $Q = P^TP$.

1.2.1 HRVM/KSH/M Primal Scaling

Choose $P = S_k^{1/2}$, $Q = S_k$. Then (1.2.6) becomes

$$2S_k\Delta_X + S_kX_k\Delta_S S_k^{-1} + \Delta_S X_k = 2\sigma_k\mu_k I - 2S_kX_k - S_k\tilde{\Delta}_X\tilde{\Delta}_S S_k^{-1} - \tilde{\Delta}_S\tilde{\Delta}_X.$$

After multiplication by S_k^{-1} on the left we obtain

$$\Delta_X + \frac{1}{2}\left(X_k\Delta_S S_k^{-1} + S_k^{-1}\Delta_S X_k\right) = \sigma_k\mu_k S_k^{-1} - X_k - \frac{1}{2}\left(\tilde{\Delta}_X\tilde{\Delta}_S S_k^{-1} + S_k^{-1}\tilde{\Delta}_S\tilde{\Delta}_X\right). \qquad (1.2.7)$$

1.2.2 HRVM/KSH/M Dual Scaling

Choose $P = X_k^{-1/2}$, $Q = X_k^{-1}$. Then (1.2.6) becomes

$$X_k^{-1}(\Delta_X S_k + X_k\Delta_S)X_k + (S_k\Delta_X + \Delta_S X_k) = 2\sigma_k\mu_k I - 2S_kX_k - X_k^{-1}\tilde{\Delta}_X\tilde{\Delta}_S X_k - \tilde{\Delta}_S\tilde{\Delta}_X.$$

After multiplication by X_k^{-1} on the left we obtain

$$\Delta_S + \frac{1}{2}\left(X_k^{-1}\Delta_X S_k + S_k\Delta_X X_k^{-1}\right) = \sigma_k\mu_k X_k^{-1} - S_k - \frac{1}{2}\left(X_k^{-1}\tilde{\Delta}_X\tilde{\Delta}_S + \tilde{\Delta}_S\tilde{\Delta}_X X_k^{-1}\right). \qquad (1.2.8)$$

1.2.3 Nesterov-Todd Search Scaling

The Nesterov-Todd direction is obtained with $P = W_k^{-1/2}$, $Q = W_k^{-1}$, where

$$W_k := X_k^{1/2}(X_k^{1/2}S_kX_k^{1/2})^{-1/2}X_k^{1/2} \succ 0$$

is such that

$$W_kS_kW_k = X_k, \qquad\qquad W_kX_k^{-1}W_k = S_k^{-1}.$$

In this case, equation (1.2.6) reduces to

$$\begin{aligned}W_k^{-1}(\Delta_X S_k + X_k\Delta_S)W_k + (S_k\Delta_X + \Delta_S X_k) &= \\ &2\sigma_k\mu_k I - W_k^{-1}X_kS_kW_k - S_kX_k - W_k^{-1}\tilde{\Delta}_X\tilde{\Delta}_S W_k - \tilde{\Delta}_S\tilde{\Delta}_X.\end{aligned}$$

After multiplication by X_k^{-1} on the right, the above equation becomes

$$\begin{aligned}&\left(\Delta_S + W_k^{-1}\Delta_X W_k^{-1} - \sigma_k\mu_k X_k^{-1} + S_k + W_k^{-1}KW_k^{-1}\right) + \\ &\qquad S_k\left(\Delta_X + W_k\Delta_S W_k - \sigma_k\mu_k S_k^{-1} + X_k + K\right)X_k^{-1} = 0,\end{aligned}$$

where $K = K^T$ is the solution to the Lyapunov equation

$$KS_kW_k + W_kS_kK = \tilde{\Delta}_X\tilde{\Delta}_S W_k + W_k\tilde{\Delta}_S\tilde{\Delta}_X. \qquad (1.2.9)$$

Indeed K is such that

$$(W_k^{-1}\tilde{\Delta}_X\tilde{\Delta}_S W_k + \tilde{\Delta}_S\tilde{\Delta}_X)X_k^{-1} = W_k^{-1}KW_k^{-1} + S_kKX_k^{-1}.$$

It is therefore enough to solve one of the equations

$$\Delta_X + W_k\Delta_S W_k = \sigma_k\mu_k S_k^{-1} - X_k - K, \qquad (1.2.10)$$

or

$$\Delta_S + W_k^{-1}\Delta_X W_k^{-1} = \sigma_k\mu_k X_k^{-1} - S_k - W_k^{-1}KW_k^{-1} \qquad (1.2.11)$$

which in this very special case are equivalent.

JUNE 27th 2011: The Nesterov-Todd search scaling is the current default choice in the current code.

1.3 Search Direction

For the NT and the KSH Primal direction we can solve for Δ_X and Δ_S explicitely to arrive at a reduced system of equations. We obtain

$$\Delta_X = -P - \frac{1}{2}\left(E\Delta_S F + F\Delta_S E\right) \tag{1.3.1}$$

where

$$E = X_k, \qquad F = S_k^{-1}, \qquad P = X_k - \sigma_k\mu_k S_k^{-1} + \frac{1}{2}\left(\tilde{\Delta}_X\tilde{\Delta}_S S_k^{-1} + S_k^{-1}\tilde{\Delta}_S\tilde{\Delta}_X\right), \tag{1.3.2}$$

for the KSH Primal direction and

$$E = W_k, \qquad F = W_k, \qquad P = X_k - \sigma_k\mu_k S_k^{-1} + K, \tag{1.3.3}$$

for the NT direction.

The notation E and F allows the solver to handle multiple choices of the search directions as discussed in Section 1.2. Henceforth we can use the letters E and F and not have to worry about the particular search direction in hand.

The rest of this section discusses a strategy for solving equations (1.2.2) and (1.2.3).

Substituting (1.3.1) in the linearized central path equations (1.2.2) and (1.2.3) we obtain the reduced equations

$$-\frac{1}{2}A^*\left(E\Delta_S F + F\Delta_S E\right) = A^*(P) + \delta_\tau B - \delta_\theta \bar{B}, \tag{1.3.4}$$

$$A(\Delta_Y) \qquad + \Delta_S \qquad = \qquad \delta_\tau C - \delta_\theta \bar{C}, \tag{1.3.5}$$

These equations can be reduced even further by solving for Δ_S as in

$$\Delta_S = \delta_\tau C - \delta_\theta \bar{C} - A(\Delta_Y). \tag{1.3.6}$$

This produces

$$\frac{1}{2}A^*\left(EA(\Delta_Y)F + FA(\Delta_Y)E\right) = A^*(P) + [A^*(P_\tau) + B]\delta_\tau - [A^*(P_\theta) + \bar{B}]\delta_\theta, \tag{1.3.7}$$

where

$$P_\tau := \frac{1}{2}\left(ECF + FCE\right), \qquad P_\theta := \frac{1}{2}\left(E\bar{C}F + F\bar{C}E\right). \tag{1.3.8}$$

In order to solve these equations and the remaining two scalar equations in δ_τ and δ_θ we first solve for three right-hand sides. That is, we solve

$$\frac{1}{2}A^*\left(EA(\bar{\Delta}_Y)F + FA(\bar{\Delta}_Y)E\right) = A^*(P), \tag{1.3.9}$$

$$\frac{1}{2}A^*\left(EA(\Delta_Y^\tau)F + FA(\Delta_Y^\tau)E\right) = A^*(P_\tau) + B \tag{1.3.10}$$

$$\frac{1}{2}A^*\left(EA(\Delta_Y^\theta)F + FA(\Delta_Y^\theta)E\right) = -(A^*(P_\theta) + \bar{B}), \tag{1.3.11}$$

Note that the left hand side of all above equations is the same. However the right hand side is different, which may affect the accuracy of the computed solution. In particular we have observed that the numerical residual on the last two equations can be orders of magnitude higher than the residual obtained in the first equation. This might require additional iterative refinements of the computed solutions. One of the reasons might be that the right hand side of the first equation is on the range space of A^*.

We then compute the corresponding Δ_X's

$$\bar{\Delta}_X = \frac{1}{2}\left(EA(\bar{\Delta}_Y)F + FA(\bar{\Delta}_Y)E\right) - P, \tag{1.3.12}$$

$$\Delta_X^\tau = \frac{1}{2}\left(EA(\Delta_Y^\tau)F + FA(\Delta_Y^\tau)E\right) - P_\tau, \tag{1.3.13}$$

$$\Delta_X^\theta = \frac{1}{2}\left(EA(\Delta_Y^\theta)F + FA(\Delta_Y^\theta E\right) + P_\theta, \tag{1.3.14}$$

from (1.3.1) and

$$\bar{\Delta}_S = -A(\bar{\Delta}_Y), \qquad \Delta_S^\tau = C - A(\Delta_Y^\tau), \qquad \Delta_S^\theta = \bar{C} - A(\Delta_Y^\theta). \tag{1.3.15}$$

Note that these come from solving (1.3.4)-(1.3.5) considering three independent right hand sides as discussed above.

After computing the Δ_Y's and Δ_X's we compute

$$\delta_\rho = -p - (\rho_k/\tau_k)\delta_\tau \tag{1.3.16}$$

where

$$p = \rho_k - \sigma_k\mu_k\tau_k^{-1} + \tilde{\delta}_\tau\tilde{\delta}_\rho\tau_k^{-1}, \tag{1.3.17}$$

for both KSH and NT directions. Then we solve for δ_τ and δ_θ

$$\begin{bmatrix} \langle C, \Delta_X^\tau\rangle - \langle B, \Delta_Y^\tau\rangle - (\rho_k/\tau_k) & \left\langle C, \Delta_X^\theta\right\rangle - \left\langle B, \Delta_Y^\theta\right\rangle - \bar{z} \\ \left\langle \bar{C}, \Delta_X^\tau\right\rangle - \left\langle \bar{B}, \Delta_Y^\tau\right\rangle - \bar{z} & \left\langle \bar{C}, \Delta_X^\theta\right\rangle - \left\langle \bar{B}, \Delta_Y^\theta\right\rangle \end{bmatrix} \begin{pmatrix} \delta_\tau \\ \delta_\theta \end{pmatrix} = \begin{pmatrix} p + \left\langle B, \bar{\Delta}_Y\right\rangle - \left\langle C, \bar{\Delta}_X\right\rangle \\ \left\langle \bar{B}, \bar{\Delta}_Y\right\rangle - \left\langle \bar{C}, \bar{\Delta}_X\right\rangle \end{pmatrix},$$

The resulting search direction is

$$\begin{aligned} \Delta_Y &= \bar{\Delta}_Y + \Delta_Y^\tau\delta_\tau + \Delta_Y^\theta\delta_\theta \\ \Delta_X &= \bar{\Delta}_X + \Delta_X^\tau\delta_\tau + \Delta_X^\theta\delta_\theta, \\ \Delta_S &= \delta_\tau C - \delta_\theta\bar{C} - A(\Delta_Y), \\ \delta_\rho &= \langle B, \Delta_Y\rangle - \langle C, \Delta_X\rangle + \delta_\theta\bar{z}. \end{aligned}$$

1.4 Variations of the linear algebra subproblem related to strategies in the algorithms

The next sections discuss minor variations on the right hand side of the linear algebra subproblem discussed above at the various phases of a predictor-correctot type algorithm.

1.4.1 Predictor

The affine-scaling (predictor) direction $(\Delta_Y^{\text{aff}}, \Delta_X^{\text{aff}}, \Delta_S^{\text{aff}}, \delta_\theta^{\text{aff}}, \delta_\tau^{\text{aff}}, \delta_\rho^{\text{aff}})$ is obtained as in § 1.3 with

$$\sigma_k = 0, \qquad \tilde{\Delta}_X = \tilde{\Delta}_S = 0, \qquad P = P_{\text{aff}} = X_k, \qquad p = r_{\text{aff}} = \rho_k, \tag{1.4.1}$$

for both KSH and NT directions.

1.4.2 Centering parameter

After computing the predictor direction we compute the scalar

$$\alpha_{\rm aff} = \arg\max\{\alpha \in [0,1] : \quad X_k + \alpha\Delta_X^{\rm aff} \succeq 0, \quad \tau_k + \alpha\delta_\tau^{\rm aff} \geq 0, \quad S_k + \alpha\Delta_S^{\rm aff} \succeq 0, \quad \rho_k + \alpha\delta_\rho^{\rm aff} \geq 0\},$$

and

$$\mu_{\rm aff} = \frac{1}{n+1}\left[\left\langle X_k + \alpha_{\rm aff}\Delta_X^{\rm aff}, S_k + \alpha_{\rm aff}\Delta_S^{\rm aff}\right\rangle + (\tau_k + \alpha_{\rm aff}\delta_\tau^{\rm aff})(\rho_k + \alpha_{\rm aff}\delta_\rho^{\rm aff})\right]$$

to produce the centering parameter

$$\sigma_k = (\mu_{\rm aff}/\mu_k)^3.$$

1.4.3 Corrector

We now compute the corrector direction $(\Delta_Y^{\rm cc}, \Delta_X^{\rm cc}, \Delta_S^{\rm cc}, \delta_\theta^{\rm cc}, \delta_\tau^{\rm cc}, \delta_\rho^{\rm cc})$ as in § 1.3 with

$$\sigma_k = (\mu_{\rm aff}/\mu_k)^3, \qquad \tilde{\Delta}_X = \Delta_X^{\rm aff}, \qquad \tilde{\Delta}_S = \Delta_S^{\rm aff}.$$

1.5 Line Search

Compute the bound

$$\alpha_{\rm cc} = \arg\max\{\alpha \in [0,1] : \quad X_k + \alpha\Delta_X^{\rm cc} \succeq 0, \quad \tau_k + \alpha\delta_\tau^{\rm cc} \geq 0, \quad S_k + \alpha\Delta_S^{\rm cc} \succeq 0, \quad \rho_k + \alpha\delta_\rho^{\rm cc} \geq 0\},$$

and

$$\alpha = \min(1, \kappa\,\alpha_{\rm cc}), \qquad \kappa \in [0,1).$$

and update

$$\begin{array}{llll} Y_{k+1} = Y_k + \alpha\Delta_Y, & \theta_{k+1} = \theta_k + \alpha\delta_\theta, & & \\ X_{k+1} = X_k + \alpha\Delta_X, & \tau_{k+1} = \tau_k + \alpha\delta_\tau, & X_{k+1} \succ 0, & \tau_{k+1} > 0 \\ S_{k+1} = S_k + \alpha\Delta_S, & \rho_{k+1} = \rho_k + \alpha\delta_\rho & S_{k+1} \succ 0, & \rho_{k+1} > 0. \end{array}$$

In the algorithm, $\kappa = 0.99$.

1.6 Computing the Nesterov-Todd Direction

This is mostly from [**?**]. Compute the Cholesky factors

$$X_k = L^T L, \qquad S_k = R^T R,$$

and the SVD decomposition

$$UDV^T = RL^T, \qquad U^T U = I, \qquad V^T V = I, \qquad D = \mathrm{diag}(d), \qquad d = (d_1, \ldots, d_n).$$

Then $Q = L^{-T} X^{1/2}$ is orthogonal [**?**, Lemma 3.3] and

$$X^{1/2} S X^{1/2} = Q^T (LR^T)(RL^T) Q = Q^T V D^2 V^T Q$$

and

$$(X^{1/2}SX^{1/2})^{-1/2} = Q^T V D^{-1} V^T Q$$

and finally

$$W_k = L^T V D^{-1} V^T L = GG^T, \qquad G := L^T V D^{-1/2}$$

Note that

$$G^T S_k G = G^{-1} X_k G^{-T} = D.$$

We now turn to the computation of the matrix K satisfying the Lyapunov equation (1.2.9), which we rewrite as

$$G^{-1} K S_k G + G^T S_k K G^{-T} = G^{-1}(\tilde{\Delta}_X \tilde{\Delta}_S W_k + W_k \tilde{\Delta}_S \tilde{\Delta}_X) G^{-T}.$$

Let $K = G\tilde{K}G^T$ so that

$$\tilde{K} D + D\tilde{K} = G^{-1} \tilde{\Delta}_X \tilde{\Delta}_S G + G^T \tilde{\Delta}_S \tilde{\Delta}_X G^{-T}.$$

Because D is diagonal, we have

$$\tilde{K} = (G^{-1} \tilde{\Delta}_X \tilde{\Delta}_S G + G^T \tilde{\Delta}_S \tilde{\Delta}_X G^{-T}) \, ./ \, (de^T + ed^T),$$

where e is a vector with all entries equal to one and ./ indicates entrywise division.

2 Iterative Solution

Taking a step back, the linear algebra problem we need to solve is the saddle-point problem:

$$\begin{array}{ccccccc} & \frac{1}{2}A^*(E\Delta_S F + F\Delta_S E) & +\delta_\tau B & -\delta_\theta \bar{B} & = & -A^*(P), \\ A(\Delta_Y) & +\Delta_S & -\delta_\tau C & +\delta_\theta \bar{C} & = & 0, \\ \langle B, \Delta_Y \rangle & +\langle FCE, \Delta_S \rangle & +\delta_\tau(\rho_k/\tau_k) & +\delta_\theta \bar{z} & = & -\langle C, P\rangle - p \\ \left\langle \bar{B}, \Delta_Y \right\rangle & +\left\langle F\bar{C}E, \Delta_S \right\rangle & +\delta_\tau \bar{z} & & = & -\left\langle \bar{C}, P \right\rangle \end{array} \tag{2.0.1}$$

Using vectorized notation we obtain

$$\begin{bmatrix} 0 & \tilde{A}^* & B & -\bar{B} \\ A & I & -C & \bar{C} \\ B^T & \tilde{C}^T & \rho_k/\tau_k & \bar{z} \\ \bar{B}^T & \tilde{\bar{C}}^T & \bar{z} & 0 \end{bmatrix} \begin{pmatrix} \Delta_Y \\ \Delta_S \\ \delta_\tau \\ \delta_\theta \end{pmatrix} = \begin{pmatrix} -A^*(P) \\ 0 \\ -\langle C, P \rangle - p \\ -\left\langle \bar{C}, P \right\rangle \end{pmatrix} \tag{2.0.2}$$

where

$$\tilde{A}^*(\Delta_S) = \frac{1}{2} A^*(E\Delta_S F + F\Delta_S E) \qquad \tilde{C} = FCE, \qquad \tilde{\bar{C}} = F\bar{C}E. \tag{2.0.3}$$

A first iterative algorithm is one that *refines* the solution to the above linear system. In this case, we can assume we have already computed the factors of the positive definite matrix

$$H := \tilde{A}^* A. \tag{2.0.4}$$

Take, for instance, a *Block Gauss-Seidel* type algorithm were we split (2.0.2) in the form

$$(L+U)\,x = b \tag{2.0.5}$$

and iterate

$$x_{k+1} = L^{-1}(b - Ux_k). \tag{2.0.6}$$

Let

$$L = \begin{bmatrix} 0 & \tilde{A}^* & 0 & 0 \\ A & I & 0 & 0 \\ B^T & \tilde{C}^T & \rho_k/\tau_k & \bar{z} \\ \bar{B}^T & \tilde{\bar{C}}^T & \bar{z} & 0 \end{bmatrix}, \qquad U = \begin{bmatrix} 0 & 0 & B & -\bar{B} \\ 0 & 0 & -C & \bar{C} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.0.7}$$

so that

$$b - Ux_k = \begin{pmatrix} -A^*(P) - B\delta_\tau^k + \bar{B}\delta_\theta^k \\ C\delta_\tau^k - \bar{C}\delta_\theta^k \\ -\langle C, P\rangle - p \\ -\left\langle \bar{C}, P\right\rangle \end{pmatrix} \tag{2.0.8}$$

and partition

$$L = L_\Delta L_\delta, \qquad L_\Delta := \begin{bmatrix} 0 & \tilde{A}^* & 0 & 0 \\ A & I & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad L_\delta := \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ B^T & \tilde{C}^T & \rho_k/\tau_k & \bar{z} \\ \bar{B}^T & \tilde{\bar{C}}^T & \bar{z} & 0 \end{bmatrix} \tag{2.0.9}$$

to obtain

$$x_{k+1/2} := L_\Delta^{-1}(b - Ux_k) \tag{2.0.10}$$

$$= \begin{bmatrix} -H^{-1} & H^{-1}\tilde{A}^* & 0 & 0 \\ AH^{-1} & I - AH^{-1}\tilde{A}^* & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} -A^*(P) - B\delta_\tau^k + \bar{B}\delta_\theta^k \\ C\delta_\tau^k - \bar{C}\delta_\theta^k \\ -\langle C, P\rangle - p \\ -\left\langle \bar{C}, P\right\rangle \end{pmatrix} \tag{2.0.11}$$

$$= \begin{pmatrix} r_k \\ C\delta_\tau^k - \bar{C}\delta_\theta^k - A\,r_k \\ -\langle C, P\rangle - p \\ -\left\langle \bar{C}, P\right\rangle \end{pmatrix}, \qquad r_k := H^{-1}\left[A^*(P) + \tilde{A}^*(C\delta_\tau^k - \bar{C}\delta_\theta^k) + B\delta_\tau^k - \bar{B}\delta_\theta^k\right] \tag{2.0.12}$$

and

$$x_{k+1} = L_\delta^{-1} x_{k+1/2} \tag{2.0.13}$$

$$= \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ -\bar{z}^{-1}\bar{B}^T & -\bar{z}^{-1}\tilde{\bar{C}}^T & 0 & \bar{z}^{-1} \\ -\bar{z}^{-1}B^T + \rho_k/(\bar{z}^2\tau_k)\bar{B}^T & -\bar{z}^{-1}\bar{C}^T + \rho_k/(\bar{z}^2\tau_k)\tilde{\bar{C}}^T & \bar{z}^{-1} & -\rho_k/(\bar{z}^2\tau_k) \end{bmatrix} \begin{pmatrix} r_k \\ C\delta_\tau^k - \bar{C}\delta_\theta^k - Ar_k \\ -\langle C, P\rangle - p \\ -\left\langle \bar{C}, P\right\rangle \end{pmatrix}. \tag{2.0.14}$$

Note that the most expensive operation is the computation of

$$r_k = H^{-1}\left[A^*(P) + \tilde{A}^*(C\delta_\tau^k - \bar{C}\delta_\theta^k) + B\delta_\tau^k - \bar{B}\delta_\theta^k\right], \tag{2.0.15}$$

which can be effectively handled by pre-computing and storing the Cholesky factors of the positive definite matrix H.

3 Sylvester

Let

$$Y = \{Y_1, \cdots, Y_n\}, \qquad\qquad A(Y) = \{A_1(Y), \cdots, A_m(Y)\} \tag{3.0.16}$$

where each

$$A_\iota(Y) = \frac{1}{2}\left(\sum_{i=1}^{n}\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota} Y_i R_{ik_\iota} + R_{ik_\iota}^T Y_i^T L_{ik_\iota}^T\right) \tag{3.0.17}$$

The variables Y_i's can be symmetric or not. Now consider the dual mapping obtained through

$$\sum_{\iota=1}^{m} \langle A_\iota(Y), X_\iota\rangle = \sum_{\iota=1}^{m}\left(\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{\ell_{i_\iota}} \langle L_{ik_\iota} Y_i R_{ik_\iota} + R_{ik_\iota}^T Y_i^T L_{ik_\iota}^T, X_\iota\rangle\right) \tag{3.0.18}$$

$$= \sum_{i=1}^{n}\left(\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} \langle L_{ik_\iota}^T X_\iota R_{ik_\iota}^T, Y_i\rangle\right) \tag{3.0.19}$$

from where

$$A^*(X) = \{A_1^*(X), \cdots, A_n^*(X)\} \tag{3.0.20}$$

where

$$A_i^*(X) = \sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota}^T X_\iota R_{ik_\iota}^T. \tag{3.0.21}$$

if Y_i is not symmetric. Note that the variables X_ι are always symmetric. If Y_i is symmetric then

$$A_i^*(X) = \frac{1}{2}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} (L_{ik_\iota}^T X_\iota R_{ik_\iota}^T + R_{ik_\iota} X_\iota L_{ik_\iota}) \tag{3.0.22}$$

Note that

$$\frac{1}{2} A^*\left(EA(Y)F + FA(Y)E\right) = \{S_1(Y), \cdots, S_n(Y)\} \tag{3.0.23}$$

where each component is

$$\begin{aligned} S_i(Y) &= \frac{1}{2}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota}^T\left(E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota\right) R_{ik_\iota}^T \\ &= \frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota R_{ik_\iota}^T) + \\ &\qquad \frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota R_{ik_\iota}^T) \end{aligned}$$

When $E_\iota = F_\iota = W_\iota$

$$S_i(Y) = \frac{1}{2}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L_{ik_\iota}^T W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota R_{ik_\iota}^T + L_{ik_\iota}^T W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota R_{ik_\iota}^T)$$

which is also a Sylvester map.

When Y_i is symmetric then we have

$$\begin{aligned} S_i(Y) = &\frac{1}{4}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota}^T \left(E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota\right) R_{ik_\iota}^T + \\ &\frac{1}{4}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota} \left(E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota\right) L_{ik_\iota} \\ = &\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L_{ik_\iota}^T E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota R_{ik_\iota}^T) + \\ &\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L_{ik_\iota}^T E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota R_{ik_\iota}^T) + \\ &\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota} E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota L_{ik_\iota} + R_{ik_\iota} F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota L_{ik_\iota}) + \\ &\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota} E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota L_{ik_\iota} + R_{ik_\iota} F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota L_{ik_\iota}). \end{aligned}$$

When $E_\iota = F_\iota = W_\iota$

$$\begin{aligned} S_i(Y) = &\frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L_{ik_\iota}^T W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota R_{ik_\iota}^T + L_{ik_\iota}^T W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota R_{ik_\iota}^T) + \\ &\frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota} W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota L_{ik_\iota} + R_{ik_\iota} W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota L_{ik_\iota}). \end{aligned}$$

Further simplifications will be possible when vectorizing such maps

3.1 Vectorized Sylvester Mappings

Vectorized mappings can be obtained directly from the formulas developed in the previous sections by using Kronecker products.

We have that

$$\operatorname{vec} A(Y) = \mathcal{A} \operatorname{vec} Y, \qquad \operatorname{vec} Y = \begin{pmatrix} \operatorname{vec} Y_1 \\ \vdots \\ \operatorname{vec} Y_n \end{pmatrix}, \qquad \mathcal{A} = [\mathcal{A}_{\iota i}], \tag{3.1.1}$$

where each

$$\begin{aligned}
\mathcal{A}_{\iota i} &= \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}} \left(R_{ik_\iota}^T \otimes L_{ik_\iota} + (L_{ik_\iota} \otimes R_{ik_\iota}^T)K_{m_i,n_i}\right), \\
&= \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}} \left(R_{ik_\iota}^T \otimes L_{ik_\iota} + K_{p_\iota,p_\iota}(R_{ik_\iota}^T \otimes L_{ik_\iota})\right), \\
&= \frac{1}{2}(I + K_{p_\iota,p_\iota}) \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota}\right).
\end{aligned}$$

Now

$$\operatorname{vec} A^*(X) = \mathcal{A}^* \operatorname{vec} X, \qquad \operatorname{vec} X = \begin{pmatrix} \operatorname{vec} X_1 \\ \vdots \\ \operatorname{vec} X_m \end{pmatrix}, \qquad \mathcal{A}^* = [\mathcal{A}_{i\iota}], \tag{3.1.2}$$

where

$$\mathcal{A}_{i\iota}^* = \sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T.$$

If Y_i is symmetric then

$$\begin{aligned}
\mathcal{A}_{i\iota}^* &= \frac{1}{2}\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T + L_{ik_\iota}^T \otimes R_{ik_\iota} \\
&= \frac{1}{2}\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T + K_{m_i,n_i}(R_{ik_\iota} \otimes L_{ik_\iota}^T)K_{p_\iota,p_\iota}
\end{aligned}$$

3.1.1 Compound scaled mapping (Alternative)

$$\begin{aligned}
\frac{1}{2}\operatorname{vec} A^*(EA(Y)F + FA(Y)E) &= \frac{1}{2}\mathcal{A}^* \operatorname{vec}(EA(Y)F + FA(Y)E) \qquad (3.1.3) \\
&= \frac{1}{2}\mathcal{A}^*((F \otimes E) + (E \otimes F))\mathcal{A} \operatorname{vec} Y \qquad (3.1.4)
\end{aligned}$$

$$\operatorname{vec} A^*(WA(Y)W) = \mathcal{A}^*(W \otimes W)\mathcal{A} \operatorname{vec} Y \tag{3.1.5}$$

$$\operatorname{vec} A^*(WA(Y)W) = \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}(R_{ik_\iota}\otimes L_{ik_\iota}^T)(W_\iota\otimes W_\iota)\left(R_{i\kappa_\iota}^T\otimes L_{i\kappa_\iota}+(L_{i\kappa_\iota}\otimes R_{i\kappa_\iota}^T)K_{m_i,n_i}\right), \tag{3.1.6}$$

$$= \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}(R_{ik_\iota}W_\iota\otimes L_{ik_\iota}^T W_\iota)\left(R_{i\kappa_\iota}^T\otimes L_{i\kappa_\iota}+(L_{i\kappa_\iota}\otimes R_{i\kappa_\iota}^T)K_{m_i,n_i}\right), \tag{3.1.7}$$

$$= \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}(R_{ik_\iota}W_\iota R_{i\kappa_\iota}^T\otimes L_{ik_\iota}^T W_\iota L_{i\kappa_\iota}+(R_{ik_\iota}W_\iota L_{i\kappa_\iota}\otimes L_{ik_\iota}^T W_\iota R_{i\kappa_\iota}^T)K_{m_i,n_i}, \tag{3.1.8}$$

$$\tag{3.1.9}$$

If Y_i is symmetric

$$\operatorname{vec} A^*(WA(Y)W) = \frac{1}{4}\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}(R_{ik_\iota}\otimes L_{ik_\iota}^T + L_{ik_\iota}^T\otimes R_{ik_\iota})(W_\iota\otimes W_\iota)\left(R_{i\kappa_\iota}^T\otimes L_{i\kappa_\iota}+(L_{i\kappa_\iota}\otimes R_{i\kappa_\iota}^T)K_{m_i,n_i}\right), \tag{3.1.10}$$

$$= \frac{1}{4}(I+K_{m_i,n_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}R_{ik_\iota}W_\iota R_{i\kappa_\iota}^T\otimes L_{ik_\iota}^T W_\iota L_{i\kappa_\iota} \tag{3.1.11}$$

$$\frac{1}{4}(I+K_{m_i,n_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{\kappa=1}^{\ell_{i_\iota}}(R_{ik_\iota}W_\iota L_{i\kappa_\iota}\otimes L_{ik_\iota}^T W_\iota R_{i\kappa_\iota}^T)K_{m_i,n_i} \tag{3.1.12}$$

3.1.2 Compound scaled mapping

Proceeding with

$$\frac{1}{2}\operatorname{vec} A^*(EA(Y)F+FA(Y)E) = \mathcal{S}\operatorname{vec} Y, \qquad \mathcal{S} = \sum_{\iota=1}^{m}[\mathcal{S}_{ij}^\iota] \tag{3.1.13}$$

where

$$\mathcal{S}_{ij}^\iota := \frac{1}{4}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota R_{jk_\iota}^T\otimes L_{ik_\iota}^T E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R_{jk_\iota}^T\otimes L_{ik_\iota}^T F_\iota L_{jk_\iota})+$$
$$\frac{1}{4}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota L_{jk_\iota}\otimes L_{ik_\iota}^T E_\iota R_{jk_\iota}^T + R_{ik_\iota}E_\iota L_{jk_\iota}\otimes L_{ik_\iota}^T F_\iota R_{jk_\iota}^T)K_{m_j,n_j}$$

When $E_\iota = F_\iota = W_\iota$

$$\mathcal{S}_{ij}^\iota := \frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}\left((R_{ik_\iota}W_\iota R_{jk_\iota}^T\otimes L_{ik_\iota}^T W_\iota L_{jk_\iota}+(R_{ik_\iota}W_\iota L_{jk_\iota}\otimes L_{ik_\iota}^T W_\iota R_{jk_\iota}^T)K_{m_j,n_j}\right)$$

If Y_i is symmetric then

$$\begin{aligned}
S^\iota_{ij} = & \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota L_{jk_\iota} \otimes L^T_{ik_\iota}E_\iota R^T_{jk_\iota} + R_{ik_\iota}E_\iota L_{jk_\iota} \otimes L^T_{ik_\iota}F_\iota R^T_{jk_\iota})K_{m_j,n_j}+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L^T_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}E_\iota L_{jk_\iota} + L^T_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L^T_{ik_\iota}F_\iota L_{jk_\iota} \otimes R_{ik_\iota}E_\iota R^T_{jk_\iota} + L^T_{ik_\iota}E_\iota L_{jk_\iota} \otimes R_{ik_\iota}F_\iota R^T_{jk_\iota})K_{m_j,n_j} \\
= & \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}K_{m_i,m_i}(L^T_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}E_\iota L_{jk_\iota} + L^T_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L^T_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}E_\iota L_{jk_\iota} + L^T_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}K_{m_i,m_i}(R_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}F_\iota L_{jk_\iota}) \\
= & \frac{1}{8}(I+K_{m_i,m_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}F_\iota L_{jk_\iota})+ \\
& \frac{1}{8}(I+K_{m_i,m_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L^T_{ik_\iota}F_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}E_\iota L_{jk_\iota} + L^T_{ik_\iota}E_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}F_\iota L_{jk_\iota})
\end{aligned}$$

and when $E_\iota = F_\iota = W_\iota$

$$S^\iota_{ij} = \frac{1}{4}(I+K_{m_i,m_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}W_\iota R^T_{jk_\iota} \otimes L^T_{ik_\iota}W_\iota L_{jk_\iota} + L^T_{ik_\iota}W_\iota R^T_{jk_\iota} \otimes R_{ik_\iota}W_\iota L_{jk_\iota})$$

3.1.3 Reduced mappings

It is possible to reduce the above mappings when some or all of variables have structured, e.g. they are symmetric. In this case, for a symmetric matrix X, we can define a projection matrix P_m such that

$$\operatorname{svec} X = P_m \operatorname{vec} X, \qquad \frac{1}{2}P_m(I+K_{m,m}) = P_m K_{m,m} = P_m. \tag{3.1.14}$$

Note that, in general

$$\operatorname{vec} X = Q_m \operatorname{svec} X, \qquad Q_m = P^T_m(P_m P^T_m)^{-1}, \tag{3.1.15}$$

which implies that

$$\frac{1}{2}(I + K_{m,m})Q_m = K_{m,m}Q_m = Q_m. \tag{3.1.16}$$

For example, for $m = 2$ we have

$$K_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad P_2 = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$\operatorname{vec}\begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \\ x_3 \end{pmatrix} = Q_2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \operatorname{svec}\begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P_2\begin{pmatrix} x_1 \\ x_2 \\ x_2 \\ x_3 \end{pmatrix}, \quad Q_2^T \operatorname{vec}\begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} = \begin{pmatrix} x_1 \\ 2\,x_2 \\ x_3 \end{pmatrix}$$

Consider the one equation in a single symmetric variable Y

$$A^*(A(Y)) = b$$

Applying svec we have

$$\operatorname{svec} A^*(A(Y)) = P_m\,\mathcal{A}^*\mathcal{A}\,Q_m \operatorname{svec} Y, \qquad \operatorname{svec} b = P_m \operatorname{vec} b,$$

where we have used the fact that

$$\operatorname{vec} A^*(X) = \mathcal{A}^* \operatorname{vec} X, \qquad \operatorname{vec} A(Y) = \mathcal{A} \operatorname{vec} Y = \mathcal{A}\,Q_m \operatorname{svec} Y.$$

In order to preserve symmetry we multiply the above equations by $(P_m P_m^T)^{-1}$ on the left to obtain

$$(P_m P_m^T)^{-1} \operatorname{svec} A^*(A(Y)) = Q_m^T\,\mathcal{A}^*\mathcal{A}\,Q_m \operatorname{svec} Y, \qquad (P_m P_m^T)^{-1} \operatorname{svec} b = Q_m^T \operatorname{vec} b,$$

If the variable Y_i is symmetric then the correspoding entry on the matrix $\mathcal{A}$ becomes

$$\mathcal{A}_{\iota i} = \frac{1}{2}(I + K_{p_\iota,p_\iota})\left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota}\right) Q_{m_i}$$

and the reduced dual mapping

$$\begin{aligned} \mathcal{A}_{i\iota}^* &= \frac{1}{2}P_{m_i}\left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T + L_{ik_\iota}^T \otimes R_{ik_\iota}\right) \\ &= \frac{1}{2}P_{m_i}\left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T + K_{n_i,m_i}(R_{ik_\iota} \otimes L_{ik_\iota}^T)K_{p_\iota,p_\iota}\right) \\ &= \frac{1}{2}P_{m_i}\left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T\right)(I + K_{p_\iota,p_\iota}) \end{aligned}$$

Note that the matrices $\mathcal{A}$ and $\mathcal{A}^*$ are now transposes of each other. This will happen if we note that the $A_\iota' s$ maps onto a space of symmetric matrices and therefore that the dual variables X_ι's are symmetric. So taking the projections on these symmetric subspaces we obtain

$$\mathcal{A}_{\iota i} = \frac{1}{2} P_\iota (I + K_{p_\iota, p_\iota}) \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right) = P_\iota \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right),$$

when Y_j is not symmetric and

$$\mathcal{A}_{\iota i} = P_\iota \left(\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota} \otimes R_{ik_\iota}^T \right) Q_{m_i} = P_\iota \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right) Q_{m_i}$$

when Y_j is symmetric.

Likewise, the dual mapping becomes

$$\mathcal{A}_{i\iota}^* = \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) Q_{p_\iota}$$

when Y_i is not symmetric and

$$\mathcal{A}_{i\iota}^* = \frac{1}{2} P_{m_i} \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) (I + K_{p_\iota, p_\iota}) Q_{p_\iota} = P_{m_i} \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) Q_{p_\iota}$$

when Y_i is symmetric. Note that finally we have $A_{\iota i}^* = A_{\iota i}^T$.

Proceeding with the scaled mapping we have that when Y_j is symmetric but Y_i is not we have

$$\mathcal{S}_{ij}^\iota = \frac{1}{4} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} F_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T E_\iota L_{jk_\iota} + R_{ik_\iota} E_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T F_\iota L_{jk_\iota}) P_{m_j}^T +$$
$$\frac{1}{4} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} F_\iota L_{jk_\iota} \otimes L_{ik_\iota}^T E_\iota R_{jk_\iota}^T + R_{ik_\iota} E_\iota L_{jk_\iota} \otimes L_{ik_\iota}^T F_\iota R_{jk_\iota}^T) P_{m_j}^T$$

and when $E_\iota = F_\iota = W_\iota$

$$\mathcal{S}_{ij}^\iota := \frac{1}{2} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \left(R_{ik_\iota} W_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T W_\iota L_{jk_\iota} + R_{ik_\iota} W_\iota L_{jk_\iota} \otimes L_{ik_\iota}^T W_\iota R_{jk_\iota}^T \right) P_{m_j}^T.$$

If Y_i is symmetric but Y_j is not

$$S_{ij}^\iota(X) = \frac{1}{4} P_{m_i} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} F_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T E_\iota L_{jk_\iota} + R_{ik_\iota} E_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T F_\iota L_{jk_\iota}) +$$
$$\frac{1}{4} P_{m_i} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T F_\iota R_{jk_\iota}^T \otimes R_{ik_\iota} E_\iota L_{jk_\iota} + L_{ik_\iota}^T E_\iota R_{jk_\iota}^T \otimes R_{ik_\iota} F_\iota L_{jk_\iota})$$

and when $E_\iota = F_\iota = W_\iota$

$$S_{ij}^\iota(X) = \frac{1}{2} P_{m_i} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} W_\iota R_{jk_\iota}^T \otimes L_{ik_\iota}^T W_\iota L_{jk_\iota} + L_{ik_\iota}^T W_\iota R_{jk_\iota}^T \otimes R_{ik_\iota} W_\iota L_{jk_\iota})$$

If Y_i and Y_j are both symmetric

$$S^{\iota}_{ij}(X) = \frac{1}{4}P_{m_i}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}F_\iota R^T_{jk_\iota}\otimes L^T_{ik_\iota}E_\iota L_{jk_\iota} + R_{ik_\iota}E_\iota R^T_{jk_\iota}\otimes L^T_{ik_\iota}F_\iota L_{jk_\iota})P^T_{m_j} + \frac{1}{4}P_{m_i}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(L^T_{ik_\iota}F_\iota R^T_{jk_\iota}\otimes R_{ik_\iota}E_\iota L_{jk_\iota} + L^T_{ik_\iota}E_\iota R^T_{jk_\iota}\otimes R_{ik_\iota}F_\iota L_{jk_\iota})P^T_{m_j}$$

and when $E_\iota = F_\iota = W_\iota$

$$S^{\iota}_{ij} = \frac{1}{2}P_{m_i}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(R_{ik_\iota}W_\iota R^T_{jk_\iota}\otimes L^T_{ik_\iota}W_\iota L_{jk_\iota} + L^T_{ik_\iota}W_\iota R^T_{jk_\iota}\otimes R_{ik_\iota}W_\iota L_{jk_\iota})P^T_{m_j}.$$

3.2 Summary of Vectorized Sylvester Mappings

3.2.1 Not Reduced

$$\mathcal{A}_{\iota i} = \frac{1}{2}(I + K_{p_\iota,p_\iota})\left(\sum_{k=1}^{\ell_{i_\iota}} R^T_{ik_\iota}\otimes L_{ik_\iota}\right).$$

$$\mathcal{A}^*_{i\iota} = \begin{cases} \displaystyle\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota}\otimes L^T_{ik_\iota}, & Y_i \quad \text{not symmetric} \\ \displaystyle\frac{1}{2}\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota}\otimes L^T_{ik_\iota} + K_{p_\iota,p_\iota}\left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota}\otimes L^T_{ik_\iota}\right)K_{m_i,m_i}, & Y_i \quad \text{symmetric} \end{cases}$$

$$\frac{1}{2}\operatorname{vec} A^*(EA(Y)F + FA(Y)E) = \mathcal{S}\operatorname{vec} Y, \qquad \mathcal{S} = \sum_{\iota=1}^{m}[\mathcal{S}^\iota_{ij}] \tag{3.2.1}$$

where

$$\mathcal{S}^\iota_{ij} = \begin{cases} \displaystyle\frac{1}{2}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}\mathcal{U}^\iota_{ik_\iota,jk_\iota} + \frac{1}{2}\left(\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}\mathcal{V}^\iota_{ik_\iota,jk_\iota}\right)K_{m_j,n_j}, & Y_i \quad \text{not symmetric,} \\ \displaystyle\frac{1}{4}(I + K_{m_i,m_i})\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}}(\mathcal{U}^\iota_{ik_\iota,jk_\iota} + \mathcal{V}^\iota_{ik_\iota,jk_\iota}), & Y_i \quad \text{symmetric.} \end{cases}$$

$$\mathcal{U}^\iota_{ik_\iota,jk_\iota} := R_{ik_\iota}W_\iota R^T_{jk_\iota}\otimes L^T_{ik_\iota}W_\iota L_{jk_\iota} \qquad \mathcal{V}^\iota_{ik_\iota,jk_\iota} := R_{ik_\iota}W_\iota L_{jk_\iota}\otimes L^T_{ik_\iota}W_\iota R^T_{jk_\iota}. \tag{3.2.2}$$

3.2.2 Reduced Y

$$\mathcal{A}_{\iota i} = \begin{cases} \frac{1}{2}(I + K_{p_\iota,p_\iota}) \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right), & Y_i \quad \text{not symmetric} \\ \frac{1}{2}(I + K_{p_\iota,p_\iota}) \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right) Q_{m_i}, & Y_i \quad \text{symmetric} \end{cases}$$

$$\mathcal{A}_{i\iota}^* = \begin{cases} \sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T, & Y_i \quad \text{not symmetric} \\ \frac{1}{2} P_{m_i} \left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) (I + K_{p_\iota,p_\iota}), & Y_i \quad \text{symmetric} \end{cases}$$

$$\mathcal{S}_{ij}^\iota = \begin{cases} \frac{1}{2} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \mathcal{U}_{ik_\iota,jk_\iota}^\iota + \frac{1}{2} \left(\sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \mathcal{V}_{ik_\iota,jk_\iota}^\iota \right) K_{m_j,n_j}, & Y_i \quad \text{not symmetric}, \quad Y_j \quad \text{not symmetric}, \\ \frac{1}{2} Q_{m_i}^T \left(\sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \mathcal{U}_{ik_\iota,jk_\iota}^\iota + \mathcal{V}_{ik_\iota,jk_\iota}^\iota \right), & Y_i \quad \text{symmetric}, \quad Y_j \quad \text{not symmetric}, \\ \frac{1}{2} \left(\sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \mathcal{U}_{ik_\iota,jk_\iota}^\iota + \mathcal{V}_{ik_\iota,jk_\iota}^\iota \right) Q_{m_j}, & Y_i \quad \text{not symmetric}, \quad Y_j \quad \text{symmetric}, \\ \frac{1}{2} Q_{m_i}^T \left(\sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} \mathcal{U}_{ik_\iota,jk_\iota}^\iota + \mathcal{V}_{ik_\iota,jk_\iota}^\iota \right) Q_{m_j}, & Y_i \quad \text{symmetric}, \quad Y_j \quad \text{symmetric}. \end{cases}$$

3.2.3 Reduced Y and X

$$\mathcal{A}_{\iota i} = \begin{cases} P_{p_\iota} \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right), & Y_i \quad \text{not symmetric} \\ P_{p_\iota} \left(\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota}^T \otimes L_{ik_\iota} \right) Q_{m_i}, & Y_i \quad \text{symmetric} \end{cases}$$

$$\mathcal{A}_{i\iota}^* = \begin{cases} \left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) Q_{p_\iota}, & Y_i \quad \text{not symmetric} \\ P_{m_i} \left(\sum_{k=1}^{\ell_{i\iota}} R_{ik_\iota} \otimes L_{ik_\iota}^T \right) Q_{p_\iota}, & Y_i \quad \text{symmetric} \end{cases}$$

$\mathcal{S}_{ij}^\iota$ is the same

4 Tensor Product LMIs

Let

$$Y = \{Y_1, \cdots, Y_n\}, \qquad A(Y) = \{A_1(Y), \cdots, A_m(Y)\} \tag{4.0.3}$$

where each

$$A_\iota(Y) = \sum_{i=1}^{n} A_{i_\iota} \otimes Y_i \tag{4.0.4}$$

The coefficient matrices $A_{i\iota}$ are all real symmetric matrices of dimension q_ι and the variables Y_i's are all real symmetric matrices of dimension r.

Dual variables X_ι are always real symmetric matrices of dimension $q_\iota r$, and the dual mapping is obtained through

$$\begin{aligned}
\sum_{\iota=1}^{m} \langle A_\iota(Y), X_\iota \rangle &= \sum_{\iota=1}^{m} \sum_{i=1}^{n} \langle A_{i_\iota} \otimes Y_i, X_\iota \rangle && (4.0.5)\\
&= \sum_{\iota=1}^{m} \sum_{i=1}^{n} \sum_{k=1}^{q_\iota} \sum_{\ell=1}^{q_\iota} \langle (A_{i_\iota})_{k\ell} Y_i, (X_\iota)_{k\ell} \rangle && (4.0.6)\\
&= \sum_{i=1}^{n} \left\langle \sum_{\iota=1}^{m} \sum_{k=1}^{q_\iota} \sum_{\ell=1}^{q_\iota} (A_{i_\iota})_{k\ell} (X_\iota)_{k\ell}, Y_i \right\rangle && (4.0.7)\\
& && (4.0.8)
\end{aligned}$$

from where

$$A^*(X) = \{A_1^*(X), \cdots, A_n^*(X)\} \tag{4.0.9}$$

and

$$A_i^*(X) = \sum_{\iota=1}^{m} A_{i_\iota} \bullet X_\iota, \tag{4.0.10}$$

after defining

$$A \bullet X = \left\{ \mathbb{SR}^{qr} \to \mathbb{SR}^{r} : \quad A \bullet X = \sum_{k=1}^{q} \sum_{\ell=1}^{q} (A)_{k\ell} (X_\iota)_{k\ell} \right\} \tag{4.0.11}$$

Note that

$$E(A \otimes Y)F = [(M)_{k\ell}] \tag{4.0.12}$$

where

$$(M)_{k\ell} = \sum_{\kappa} (E)_{k\kappa} (N)_{\kappa\ell}, \qquad (N)_{\kappa\ell} = \sum_{j} (A)_{\kappa j} Y (F)_{j\ell} \tag{4.0.13}$$

which implies

$$(M)_{k\ell} = \sum_{\kappa} \sum_{j} (A)_{\kappa j} (E)_{k\kappa} Y (F)_{j\ell} \tag{4.0.14}$$

and

$$A \bullet E(A \otimes Y)F = \sum_{k=1}^{q} \sum_{\ell=1}^{q} (A)_{k\ell} (M)_{k\ell} \tag{4.0.15}$$

$$= \sum_{k=1}^{q} \sum_{\ell=1}^{q} (A)_{k\ell} \sum_{\kappa} \sum_{j} (A)_{\kappa j} (E)_{k\kappa} Y (F)_{j\ell} \tag{4.0.16}$$

$$\frac{1}{2} A^* \left(EA(Y)F + FA(Y)E \right) = \{S_1(Y), \cdots, S_n(Y)\} \tag{4.0.17}$$

where each component is

$$\begin{aligned}
S_i(Y) &= \sum_{\iota=1}^{m} A_{i_\iota} \bullet (E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota) \\
A_\iota(Y) &= \sum_{i=1}^{n} A_{i_\iota} \otimes Y_i \\
&= \frac{1}{2} \sum_{\iota=1}^{m} \sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota}^T R_{ik_\iota}^T \\
&= \frac{1}{4} \sum_{j=1}^{n} \sum_{\iota=1}^{m} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota R_{ik_\iota}^T) + \\
&\qquad \frac{1}{4} \sum_{j=1}^{n} \sum_{\iota=1}^{m} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota R_{ik_\iota}^T)
\end{aligned}$$

When $E_\iota = F_\iota = W_\iota$

$$S_i(Y) = \frac{1}{2} \sum_{j=1}^{n} \sum_{\iota=1}^{m} \sum_{k=1}^{\ell_{i_\iota}} \sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota R_{ik_\iota}^T + L_{ik_\iota}^T W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota R_{ik_\iota}^T)$$

which is also a Sylvester map.

When Y_i is symmetric then we have

$$\begin{aligned}
S_i(Y) = \frac{1}{4}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} L_{ik_\iota}^T \left(E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota\right) R_{ik_\iota}^T + \\
\frac{1}{4}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}} R_{ik_\iota} \left(E_\iota A_\iota(Y) F_\iota + F_\iota A_\iota(Y) E_\iota\right) L_{ik_\iota} \\
= \frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota R_{ik_\iota}^T) + \\
\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota R_{ik_\iota}^T + L_{ik_\iota}^T F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota R_{ik_\iota}^T) + \\
\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} E_\iota L_{jk_\iota} Y_j R_{jk_\iota} F_\iota L_{ik_\iota} + R_{ik_\iota} F_\iota L_{jk_\iota} Y_j R_{jk_\iota} E_\iota L_{ik_\iota}) + \\
\frac{1}{8}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} E_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T F_\iota L_{ik_\iota} + R_{ik_\iota} F_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T E_\iota L_{ik_\iota}).
\end{aligned}$$

When $E_\iota = F_\iota = W_\iota$

$$\begin{aligned}
S_i(Y) = \frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (L_{ik_\iota}^T W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota R_{ik_\iota}^T + L_{ik_\iota}^T W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota R_{ik_\iota}^T) + \\
\frac{1}{4}\sum_{j=1}^{n}\sum_{\iota=1}^{m}\sum_{k=1}^{\ell_{i_\iota}}\sum_{k=1}^{\ell_{j_\iota}} (R_{ik_\iota} W_\iota L_{jk_\iota} Y_j R_{jk_\iota} W_\iota L_{ik_\iota} + R_{ik_\iota} W_\iota R_{jk_\iota}^T Y_j^T L_{jk_\iota}^T W_\iota L_{ik_\iota}).
\end{aligned}$$

Further simplifications will be possible when vectorizing such maps

5 AHSS

Let us foccus on the saddle-point problem

$$\begin{bmatrix} W_k^{-1} \otimes W_k^{-1} & -A \\ A^* & 0 \end{bmatrix} \begin{pmatrix} \bar{\Delta}_X & \Delta_X^\tau & \Delta_X^\theta \\ \bar{\Delta}_Y & \Delta_Y^\tau & \Delta_Y^\theta \end{pmatrix} = \begin{bmatrix} \sigma_k \mu_k X_k^{-1} - S_k & -C & \bar{C} \\ 0 & B & -\bar{B} \end{bmatrix}$$

Or for easy of notation

$$\begin{bmatrix} W_k^{-1} \otimes W_k^{-1} & -A \\ A^* & 0 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

AHSS is

1. Compute the residuals

$$r^{(k)} = f - [(W_k^{-1} \otimes W_k^{-1})x^{(k)} - Ay^{(k)}] \qquad s^{(k)} = g - A^* x^{(k)}$$

2. Compute the auxiliary vectors

$$u^{(k)} = \frac{2}{\alpha+1} r^{(k)} \qquad v^{(k)} = 2s^{(k)} - A^*(W_k \otimes W_k)u^{(k)}$$

3. Compute the update vectors

$$\left(\beta C + \frac{1}{\alpha} A^*(W_k \otimes W_k)A\right) w^{(k)} = v^{(k)} \qquad t^{(k)} = (W_k \otimes W_k)[u^{(k)} + Aw^{(k)}]$$

4. Compute next iterate

$$x^{(k+1)} = x^{(k)} + t^{(k)}, \qquad y^{(k+1)} = y^{(k)} + w^{(k)}.$$

After some manipulations the computation of the update vector is

$$\begin{aligned}
\left(\beta C + \frac{1}{\alpha} A^*(W_k \otimes W_k)A\right) w^{(k)} &= 2g - 2A^* x^{(k)} - A^*(W_k \otimes W_k)u^{(k)} \\
&= 2g - 2A^* x^{(k)} - \frac{2}{\alpha+1} A^*(W_k \otimes W_k) r^{(k)} \\
&= 2g - 2A^* \left[x^{(k)} - \frac{1}{\alpha+1}(W_k \otimes W_k) r^{(k)}\right] \\
&= 2g - 2A^* \left[x^{(k)} + \frac{1}{\alpha+1} x^{(k)} - \frac{1}{\alpha+1}(W_k \otimes W_k)[f + Ay^{(k)}]\right] \\
&= 2g - 2A^* \left[\frac{\alpha+2}{\alpha+1} x^{(k)} - \frac{1}{\alpha+1}(W_k \otimes W_k)[f + Ay^{(k)}]\right]
\end{aligned}$$

or

$$\begin{aligned}
A^*\left[\beta(Z_k \otimes Z_k) + \frac{1}{\alpha}(W_k \otimes W_k)\right] Aw^{(k)} &= 2g - A^*\left[2x^{(k)} - (W_k \otimes W_k)u^{(k)}\right] \\
&= 2g - A^*\left[2x^{(k)} - (W_k \otimes W_k)[f - Ay^{(k)}] + x^{(k)}\right]
\end{aligned}$$

AHSS is

1. Compute the residuals

$$r^{(k)} = f - [(W_k^{-1} \otimes W_k^{-1})x^{(k)} - Ay^{(k)}] \qquad s^{(k)} = g - A^* x^{(k)}$$

2. Compute the auxiliary vectors

$$u^{(k)} = \frac{2}{\alpha + 1} r^{(k)} \qquad v^{(k)} = 2s^{(k)} - A^*(W_k \otimes W_k)u^{(k)}$$

3. Compute the update vectors

$$\left(\beta C + \frac{1}{\alpha} A^*(W_k \otimes W_k)A\right) w^{(k)} = v^{(k)} \qquad t^{(k)} = (W_k \otimes W_k)[u^{(k)} + Aw^{(k)}]$$

4. Compute next iterate

$$x^{(k+1)} = x^{(k)} + t^{(k)}, \qquad y^{(k+1)} = y^{(k)} + w^{(k)}.$$

5.1 Approximate Solution

If the above linear algebra problem is solved for Δ_Y only approximately then the iterates Δ_X may not satisfy the

Given Δ_Y and δ_θ solve

$$\begin{aligned} A^*(\Delta_X) - \delta_\tau B + \delta_\theta \bar{B} &= 0 \\ -\left\langle \bar{B}, \Delta_Y \right\rangle + \left\langle \bar{C}, \Delta_X \right\rangle - \delta_\tau \bar{z} &= 0 \end{aligned}$$

for Δ_X and δ_τ. One option is to solve for Δ_X in

$$A^*(\Delta_X) - B \left\langle \bar{C}/\bar{z}, \Delta_X \right\rangle = B \left\langle \bar{B}/\bar{z}, \Delta_Y \right\rangle - \delta_\theta \bar{B}$$

then set

$$\begin{aligned} \delta_\tau &= \left\langle \bar{C}/\bar{z}, \Delta_X \right\rangle - \left\langle \bar{B}/\bar{z}, \Delta_Y \right\rangle \\ \Delta_S &= -A(\Delta_Y) + \delta_\tau C - \delta_\theta \bar{C} \\ \delta_\rho &= \langle B, \Delta_Y \rangle - \langle C, \Delta_X \rangle + \delta_\theta \bar{z} \end{aligned}$$

Note that

$$\left\langle A^*(\Delta_X) - \left\langle \bar{C}/\bar{z}, \Delta_X \right\rangle B, Z \right\rangle = \left\langle A(Z) - \langle B, Z \rangle \bar{C}/\bar{z}, \Delta_X \right\rangle$$

If need to improve set Δ_X and δ_τ fixed then solve

$$\begin{aligned} A(\Delta_Y) + \delta_\theta \bar{C} &= S_k - \sigma_k \mu_k X_k^{-1} + W_k^{-1} \Delta_X W_k^{-1} + \delta_\tau C \\ \langle B, \Delta_Y \rangle + \delta_\theta \bar{z} &= \sigma_k \mu_k \tau_k^{-1} - \rho_k + \langle C, \Delta_X \rangle - \delta_\tau (\rho_k / \tau_k) \end{aligned}$$

for Δ_Y and δ_θ. One option is to solve for Δ_Y in

$$A(\Delta_Y) - \langle B, \Delta_Y \rangle \bar{C}/\bar{z} = S_k - \sigma_k \mu_k X_k^{-1} + W_k^{-1} \Delta_X W_k^{-1} + \delta_\tau C - [\sigma_k \mu_k \tau_k^{-1} - \rho_k + \langle C, \Delta_X \rangle - \delta_\tau (\rho_k / \tau_k)] \bar{C}/\bar{z}$$

then set

$$\delta_\theta = - \langle B/\bar{z}, \Delta_Y \rangle + [\sigma_k \mu_k \tau_k^{-1} - \rho_k + \langle C, \Delta_X \rangle - \delta_\tau (\rho_k / \tau_k)]/\bar{z}$$

5.2 Self Duality

The above problem is self dual. Indeed defining the shorthand notation

$$\begin{bmatrix} 0 & -\mathcal{A}^* & -\bar{B} \\ \mathcal{A} & \mathcal{Q} & \mathcal{R} \\ \bar{B}^* & -\mathcal{R}^* & 0 \end{bmatrix} = \left[\begin{array}{c|cc|c} 0 & -A^* & B & -\bar{B} \\ \hline A & 0 & -C & \bar{C} \\ -B^* & C^* & 0 & -\bar{z} \\ \hline \bar{B}^* & -\bar{C}^* & \bar{z} & 0 \end{array}\right], \qquad \mathcal{X} = \begin{pmatrix} X \\ \tau \end{pmatrix}, \qquad \mathcal{S} = \begin{pmatrix} S \\ \rho \end{pmatrix},$$

we can rewrite the primal program as

$$\begin{aligned} &\inf_{Y,\mathcal{Y},\mathcal{S}} \beta\theta \\ &\quad -\mathcal{A}^*(\mathcal{X}) - \bar{B}\theta = 0 \\ &\quad \mathcal{A}(Y) + \mathcal{Q}(\mathcal{X}) + \mathcal{R}\theta + \mathcal{S} = 0 \\ &\quad \left\langle \bar{B}, Y \right\rangle - \mathcal{R}^*(\mathcal{X}) = \beta \\ &\quad \mathcal{X} \succeq 0, \qquad \mathcal{S} \succeq 0 \end{aligned}$$

Defining the Lagrangian

$$\begin{aligned} L(Y,\mathcal{Y}) &= \beta\theta - \left\langle Z, \mathcal{A}^*(\mathcal{X}) + \bar{B}\theta \right\rangle + \langle \mathcal{Y}, \mathcal{A}(Y) + \mathcal{Q}(\mathcal{X}) + \mathcal{R}\theta \rangle + \nu \left[\left\langle \bar{B}, Y \right\rangle - \mathcal{R}^*(\mathcal{X}) - \beta \right] \\ &= \beta\nu + \left\langle Y, \mathcal{A}^*(\mathcal{Y}) + \bar{B}\nu \right\rangle + \langle \mathcal{X}, -\mathcal{A}(Z) - \mathcal{Q}(\mathcal{Y}) - \mathcal{R}\nu \rangle + \theta \left[\beta - \left\langle \bar{B}, Z \right\rangle + \mathcal{R}^*(\mathcal{Y}) \right] \end{aligned}$$

we can compute the dual program

$$\begin{aligned} &\sup_{Z,\mathcal{Z},\mathcal{V}} -\beta\nu \\ &\quad \mathcal{A}^*(\mathcal{Y}) + \bar{B}\nu = 0 \\ &\quad -\mathcal{A}(Z) - \mathcal{Q}(\mathcal{Y}) - \mathcal{R}\nu - \mathcal{V} = 0 \\ &\quad \left\langle \bar{B}, Z \right\rangle - \mathcal{R}^*(\mathcal{Y}) = \beta \\ &\quad \mathcal{Y} \succeq 0, \qquad \mathcal{V} \succeq 0 \end{aligned}$$

which is equivalent to primal.

6 HSS

Rearrange the linear problem (**??**) into the form

$$\begin{bmatrix} W_k^{-1} \otimes W_k^{-1} & C & -A & -\bar{C} \\ -C^* & (\rho_k/\tau_k) & B^* & \bar{z} \\ A^* & -B & 0 & \bar{B} \\ \bar{C}^* & -\bar{z} & -\bar{B}^* & 0 \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \\ \Delta_Y \\ \delta_\theta \end{pmatrix} = \begin{pmatrix} \sigma_k\mu_k X_k^{-1} - S_k \\ \sigma_k\mu_k\tau_k^{-1} - \rho_k \\ 0 \\ 0 \end{pmatrix}.$$

Let $\alpha, \beta > 0$ be given. HSS is:
Step 1: compute the partial vectors

$$\begin{bmatrix} \alpha I + W_k^{-1} \otimes W_k^{-1} & C \\ -C^* & \alpha I + (\rho_k/\tau_k) \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell+\frac{1}{2})} = \alpha \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell)} + \begin{bmatrix} A & \bar{C} \\ -B^* & -\bar{z} \end{bmatrix} \begin{pmatrix} \Delta_Y \\ \delta_\theta \end{pmatrix}^{(\ell)} + \begin{pmatrix} \sigma_k\mu_k X_k^{-1} - S_k \\ \sigma_k\mu_k\tau_k^{-1} - \rho_k \end{pmatrix}$$

$$\begin{bmatrix} \beta I & \bar{B} \\ -\bar{B}^* & \beta I \end{bmatrix} \begin{pmatrix} \Delta_Y \\ \delta_\theta \end{pmatrix}^{(\ell+\frac{1}{2})} = \begin{bmatrix} A^* & -B \\ \bar{C}^* & -\bar{z} \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell)} + \beta \begin{pmatrix} \Delta_Y \\ \delta_\theta \end{pmatrix}^{(\ell)}.$$

Step 2: compute the partial vectors

$$\begin{bmatrix} \alpha I & 0 & -A & -\bar{C} \\ 0 & \alpha I & B^* & \bar{z} \\ A^* & -B & \beta I & 0 \\ \bar{C}^* & -\bar{z} & 0 & \beta I \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \\ \Delta_Y \\ \delta_\theta \end{pmatrix}^{(\ell+1)} = \begin{bmatrix} \begin{pmatrix} \alpha I - W_k^{-1} \otimes W_k^{-1} & -C \\ C^* & \alpha I - (\rho_k/\tau_k) \end{pmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell+\frac{1}{2})} + \begin{pmatrix} \sigma_k \mu_k X_k^{-1} - S_k \\ \sigma_k \mu_k \tau_k^{-1} - \rho_k \end{pmatrix} \\ \begin{bmatrix} \beta I & -\bar{B} \\ \bar{B}^* & \beta I \end{bmatrix} \begin{pmatrix} \Delta_Y \\ \delta_\theta \end{pmatrix}^{(\ell+\frac{1}{2})} \end{bmatrix}$$

$$\begin{bmatrix} \alpha I + W_k^{-1} \otimes W_k^{-1} & C \\ -C^* & \alpha I + (\rho_k/\tau_k) \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell+\frac{1}{2})} = \begin{pmatrix} r_X \\ r_\tau \end{pmatrix}$$

Now note that

$$\begin{bmatrix} I & 0 \\ C^*[\alpha I + W_k^{-1} \otimes W_k^{-1}]^{-1} & 1 \end{bmatrix} \begin{bmatrix} \alpha I + W_k^{-1} \otimes W_k^{-1} & C \\ -C^* & \alpha I + (\rho_k/\tau_k) \end{bmatrix} = \begin{bmatrix} \alpha I + W_k^{-1} \otimes W_k^{-1} & C \\ 0 & \alpha I + (\rho_k/\tau_k) + C^*[\alpha I + W_k^{-1} \otimes W_k^{-1}]^{-1} C \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ C^*[\alpha I + W_k^{-1} \otimes W_k^{-1}]^{-1} & 1 \end{bmatrix} \begin{bmatrix} \alpha I + W_k^{-1} \otimes W_k^{-1} & C \\ -C^* & \alpha I + (\rho_k/\tau_k) \end{bmatrix} \begin{pmatrix} \Delta_X \\ \delta_\tau \end{pmatrix}^{(\ell+\frac{1}{2})} = \begin{bmatrix} I & 0 \\ C^*[\alpha I + W_k^{-1} \otimes W_k^{-1}]^{-1} & 1 \end{bmatrix} \begin{pmatrix} r_X \\ r_\tau \end{pmatrix}$$

Using the fact that

$$[\alpha I + W_k^{-1} \otimes W_k^{-1}]^{-1} = \alpha^{-1} I - \alpha^{-1}[I + \alpha W_k \otimes W_k]^{-1}$$

$$[\alpha I + W_k^{-1} \otimes W_k^{-1}] \operatorname{vec}(X) = \operatorname{vec}(B)$$

$$\alpha X + W_k^{-1} X W_k^{-1} = B$$