Robust Search Directions
for
Large Sparse Semidefinite Programming (SDP)

Henry Wolkowicz

ISMP 2003

Department of Combinatorics & Optimization
University of Waterloo

Contents

1 OUTLINE

- Part I
  - Background on SDP; Notation and Motivation
  - Robust, ('non-interior') path-following algorithm for SDP (outline of Gauss-Newton (GN), PCG, method using LP)
- Part II
  - Strengthened Existence and Uniqueness Conditions for Search Directions in Semidefinite Programming

2 What is SDP? Notation and Motivation

$$
(\text{SDP}) \qquad \begin{array}{ll} \min & f(X) \\ \text{subject to} & \mathcal{A}X = b \\ & X \succeq 0, \end{array}
$$

where: $f : \mathcal{S}^n \to \mathbb{R}$ convex function

$\mathcal{S}^n$ $n \times n$ real symmetric matrices

$\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$ linear operator,

$X(\succeq) \succ 0$ denotes positive (semi)definite

$\left(\quad (\mathcal{A}X)_i = \langle A_i, X \rangle = \operatorname{trace} A_i X, \quad A_i = A_i^T, i = 1 \ldots n) \quad \right)$

2.1 Linear Primal-Dual Pair of SDPs

(looks/behaves like Linear Program, LP)

$$
\text{(PSDP)} \qquad \begin{array}{rl} \min & \langle C, X \rangle = \operatorname{trace} CX \\ \text{subject to} & \mathcal{A}X = b \\ & X \succeq 0, \end{array}
$$

$$
\text{(SDP)} \qquad \begin{array}{rl} \max & b^T y \\ \text{subject to} & \mathcal{A}^* y + Z = C \\ & Z \succeq 0, \end{array}
$$

adjoint operator: $\mathcal{A}^* y = \sum_{i=1}^m y_i A_i$

2.2 (Perturbed) Optimality Conditions

For *barrier parameter* $\mu > 0$:

$$F_\mu(X, y, Z) := \begin{pmatrix} \mathcal{A}^* y + Z - C \\ \mathcal{A}X - b \\ ZX - \mu I \end{pmatrix} = 0 \qquad \begin{pmatrix} \text{dual feasibility} \\ \text{primal feasibility} \\ \text{pert. compl. slack.} \end{pmatrix}$$

For SDP:

$F_\mu : \mathcal{S}^n \times \Re^m \times \mathcal{S}^n \rightarrow \mathcal{S}^n \times \Re^m \times \mathcal{M}^n$

i.e. overdetermined nonlinear system

3 (Non) Interior Path-Following

3.1 Illustration/Motivation on LP Case

$$
\text{(LP)} \qquad \begin{array}{rll} p^* := & \min & c^T x \quad (\text{or } \langle c, x \rangle) \\ & \text{s.t.} & Ax = b \\ & & x \geq 0 \quad (\text{or } x \succeq 0) \end{array}
$$

$$
\text{(DLP)} \qquad \begin{array}{rll} d^* := & \max & b^T y \\ & \text{s.t.} & A^T y + z = c \\ & & z \geq 0 \quad (\text{or } z \succeq 0) \end{array}
$$

$A \in \Re^{m \times n}$ full rank (onto); LP and DLP strictly feasible

dual log-barrier problem with parameter $\mu > 0$ is

$$
\text{(Dlogbarrier)} \qquad
\begin{array}{rll}
d^*_\mu := & \max & b^T y + \mu \sum_{j=1}^n \log z_j \quad (+\mu \log\det(z)) \\
& \text{s.t.} & A^T y + z = c \\
& & z > 0 \qquad (z \succ 0).
\end{array}
$$

stationary point of the Lagrangian / optimality conditions

$$
F_\mu(x,y,z) = \begin{pmatrix} A^T y + z - c \\ Ax - b \\ X - \mu Z^{-1} \end{pmatrix} = 0, \qquad
\begin{array}{c} x, z > 0, \quad (\succ 0) \\ X = \mathrm{Diag}\,(x), \\ Z = \mathrm{Diag}\,(z) \end{array}
$$

central path: set of these solutions $(x_\mu, y_\mu, z_\mu), \mu > 0$

As $\mu \to 0$, Jacobian $F'_\mu(x,y,z)$ grows ill-conditioned near central path

Cure/Fix: Make nonlinear equations *less nonlinear*, i.e. preconditioning for Newton type methods; premultiply by block-diag matrix with blocks (I, I, Z):

$$F_\mu(x,y,z) \leftarrow \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & Z \end{pmatrix} F_\mu(x,y,z) \quad = \quad \begin{pmatrix} A^T y + z - c \\ Ax - b \\ ZX - \mu I \end{pmatrix}$$

$$=: \begin{pmatrix} R_d \\ r_p \\ R_{ZX} \end{pmatrix}$$

Special structure of linearized system can be exploited;

linearization for the Newton direction $\Delta s = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$ is

$$F'_\mu(x,y,z)\Delta s = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} \Delta s = -F_\mu(x,y,z).$$

overdetermined system in SDP case:

$$\mathcal{S}^n \times \Re^m \times \mathcal{S}^n \to \mathcal{S}^n \times \Re^m \times \mathcal{M}^n$$

apply symmetrization 'undoes preconditioning'

$$\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \mathcal{S} \end{pmatrix} \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix}$$

e.g. last equation is linearization of e.g.:
$ZX + XZ - 2\mu I = 0$ (AHO search direction)

3.2 Reduction/Block Elimination for the Normal Equations

Step 1 (Eliminate Δz):

$$\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -X & 0 & I \end{pmatrix} \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & -XA^T & 0 \end{pmatrix}.$$

We let

$$P_Z = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -X & 0 & I \end{pmatrix}, \quad K = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & -XA^T & 0 \end{pmatrix}.$$

with right-hand side

$$-\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -X & 0 & I \end{pmatrix} \begin{pmatrix} R_d \\ r_p \\ R_{ZX} - \mu e \end{pmatrix} = \begin{pmatrix} -R_d \\ -r_p \\ XR_d - R_{ZX} \end{pmatrix}$$

Step 2 (Eliminate Δx):

$$F_n := P_n K \quad := \quad \begin{pmatrix} I & 0 & 0 \\ 0 & I & -AZ^{-1} \\ 0 & 0 & Z^{-1} \end{pmatrix} \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & -XA^T & 0 \end{pmatrix}$$

$$= \quad \begin{pmatrix} 0 & A^T & I_n \\ 0 & AZ^{-1}XA^T & 0 \\ I_n & -Z^{-1}XA^T & 0 \end{pmatrix}$$

$AZ^{-1}XA^T$ can have:

- uniformly bounded condition number, e.g. Güler et al 1993
- structured singularity, e.g. M. Wright 1997

But $\mathrm{cond}(F_n) \to \infty$.

The right-hand side becomes

$$-P_n P_Z \begin{pmatrix} R_d \\ r_p \\ R_{ZX} \end{pmatrix} =$$

$$\begin{pmatrix} -R_d \\ -r_p + A(x - Z^{-1}XR_d - \mu Z^{-1}e) \\ Z^{-1}XR_d - x + \mu Z^{-1}e \end{pmatrix}$$

Proposition The condition number of $F_n^T F_n$ diverges to infinity if $x(\mu)_i / z(\mu)_i$ diverges to infinity, for some i, as μ converges to 0. The condition number of $(F'_\mu)^T F'_\mu$ is uniformly bounded if there exists a unique primal-dual solution.

PROOF: Note that

$$F_n^T F_n = \begin{pmatrix} I_n & -Z^{-1}XA^T & 0 \\ -AXZ^{-1} & (AA^T + (AZ^{-1}XA^T)^2 + AZ^{-2}X^2A^T) & A \\ 0 & A^T & I_n \end{pmatrix}.$$

By interlacing of eigenvalues, ... ■

EXAMPLE: *getting too close to boundary* (worse for SDP)

$A = \begin{pmatrix} 1 & 1 \end{pmatrix}, c = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, b = 1.$

$x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, y^* = -1, z^* = \begin{pmatrix} 0 \\ 2 \end{pmatrix};$

initial points: $x = \begin{pmatrix} 9.183012e-001 \\ 1.356397e-008 \end{pmatrix}, z = \begin{pmatrix} 2.193642e-008 \\ 1.836603e+000 \end{pmatrix},$
$y = -1.163398e+000.$

residuals and duality gap:

$\|r_b\| = 0.081699, \quad \|R_d\| = 0.36537, \quad \mu = x^T z/n = 2.2528e-008$

5 decimals rounding before/after arithmetic

centering with $\sigma = .1$

search directions found using full matrix F'_μ and backsolve matrix F_n:

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} 8.17000e-02 \\ -1.35440e-08 \\ 1.63400e-01 \\ -2.14340e-08 \\ 1.63400e-01 \end{pmatrix}; \quad = \begin{pmatrix} -6.06210e+01 \\ -1.35440e-08 \\ 1.63400e-01 \\ 0.00000e+00 \\ 1.63400e-01 \end{pmatrix}$$

error in Δy is small; error after backsubstitution for $(\Delta x)_1$ is large.

$$\begin{pmatrix} AZ^{-1}XA^T \\ -Z^{-1}XA^T \end{pmatrix} = \begin{pmatrix} 4.18630e+07 \\ -4.18630e+07 \\ -7.38540e-09 \end{pmatrix}$$

3.3 Alternate Second Step; Stable Reduction

Assuming! $A = [I_m \; E]$.

Partition diagonal matrix Z, X using vectors
$z = \begin{pmatrix} z_m \\ z_v \end{pmatrix}, x = \begin{pmatrix} x_m \\ x_v \end{pmatrix}$, $XA^T = \begin{pmatrix} X_m \\ X_v E^T \end{pmatrix}$

$$
\begin{aligned}
F_s : &= P_s K = \begin{pmatrix} I_n & 0 & 0 & 0 \\ 0 & I_m & 0 & 0 \\ 0 & -Z_m & I_m & 0 \\ 0 & 0 & 0 & I_v \end{pmatrix} \begin{pmatrix} 0 & 0 & A^T & I_n \\ I_m & E & 0 & 0 \\ Z_m & 0 & -X_m & 0 \\ 0 & Z_v & -X_v E^T & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & A^T & I_n \\ I_m & E & 0 & 0 \\ 0 & -Z_m E & -X_m & 0 \\ 0 & Z_v & -X_v E^T & 0 \end{pmatrix}.
\end{aligned}
$$

The right-hand side becomes

$$
\begin{aligned}
-P_s P_Z \begin{pmatrix} A^T y + z - c \\ Ax - b \\ ZXe - \mu e \end{pmatrix} &= -P_s \begin{pmatrix} R_d \\ r_p \\ -XR_d + ZXe - \mu e \end{pmatrix} \\
&= \begin{pmatrix} -R_d \\ -r_p \\ -Z_m r_p - X_m (R_d)_m + Z_m X_m e - \mu e \\ -X_v (R_d)_v + Z_v X_v e - \mu e \end{pmatrix}
\end{aligned}
$$

4 Applications

Exploit structure in:

1. Finding the Nearest Correlation Matrix
2. SDP Relaxations of:
   Max-Cut Problem; Theta Function; Quadratic Assignment Problem; Graph Partitioning Problem.

(work with: Anjos, Higham, Kruk, Rendl, Sotirov, Takouda, Wei)

5 Conclusion - Gauss-Newton direction

Advantages

Robust, warm starts are simple, longer steps, no pred.-corr.

exact primal and dual feasibility at each iteration

Can apply CG-type approaches

q-quadratic convergence

scale-invariant on the right

Disadvantages/Future:

Still too expensive

Need large sparse QR efficient as Cholesky

6 Part II: Strengthened Existence and Uniqueness Conditions for Search Directions in Semidefinite Programming

Levent Tunçel and Henry Wolkowicz

Department of Combinatorics & Optimization
University of Waterloo

(Acknowledgment: Kathrin Schäcke)

7 Preliminaries

7.1 OUTLINE

- Preview: summary of known and strengthened results
- Kronecker and Symmetric Kronecker Products
- quadratic constrained quadratic programs
- main results: **-strengthened sufficient conditions**
  - **-E&U generic**

7.2 Primal-Dual Pair of SDPs

$$
\text{(P)} \qquad \begin{array}{rll}
\min & \operatorname{trace} CX & (\langle C, X \rangle) \\
\text{subject to} & \operatorname{trace} A_i X = b_i,\ \ i = 1, \ldots, m, & (\mathcal{A}X = b) \\
& X \in \mathcal{S}^n_+ & (X \succeq 0)
\end{array}
$$

$$
\text{(D)} \qquad \begin{array}{rll}
\max & b^T y & (\langle b, y \rangle) \\
\text{subject to} & \sum_{i=1}^m y_i A_i + S = C, & (\mathcal{A}^* y + S = C) \\
& S \in \mathcal{S}^n_+ & (S \succeq 0).
\end{array}
$$

7.3 Perturbed Optimality Conditions

Barrier parameter: $\mu > 0$

$$\begin{array}{rcll} \mathcal{A}^* y + S - C & = & 0 & \text{dual feasibility} \\ \mathcal{A} X - b & = & 0 & \text{primal feasibility} \\ SX - \mu I & = & 0 & \text{perturbed complementary slackness} \\ X \succ 0, S \succ 0 & & & \end{array}$$

overdetermined nonlinear system of matrix-vector equations; apply symmetrization and Newton's Method.

7.4 AHO Search Direction

Alizadeh-Haeberly-Overton (AHO) $(\Delta X, \Delta y, \Delta S)$,
at a point $(X, y, S) \in \mathcal{S}^n_+ \times \mathbb{R}^m \times \mathcal{S}^n_+$ with centrality parameter $\sigma \in (0, 1]$ and barrier parameter $0 < \mu = \operatorname{trace}(XS)/n$:

$$\begin{array}{rcl} \mathcal{A}^*\Delta y + \Delta S & = & C - \mathcal{A}^* y - S \\ \mathcal{A}\Delta X & = & b - \mathcal{A}X \\ X(\Delta S) + (\Delta S)X + S(\Delta X) + (\Delta X)S & = & 2\sigma\mu I - XS - SX \end{array}$$

i.e. linearization of symmetrized optimality conditions, AHOlin:

$$\mathrm{AHO}'_\mu(\Delta X, \Delta y, \Delta S) = -\text{residuals}$$

7.4.1 Example:

$$X = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix}, \; S = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix}, \; \mathcal{A}X = \operatorname{trace} \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} X = 3$$

So AHOlin (with appropriate C) is:

$$\begin{array}{rcll} \Delta y \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} + \Delta S & = & 0, & \left(\Delta S = -\Delta y \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} \right) \\ \operatorname{trace} \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} \Delta X & = & 0 & \left(\Delta X = \begin{pmatrix} 2\sqrt{2}\Delta v_1 & \Delta v_1 \\ \Delta v_1 & \Delta v_2 \end{pmatrix} \right) \end{array}$$

So substitute and the third block

$$X(\Delta S) + (\Delta S)X + S(\Delta X) + (\Delta X)S = 2\sigma\mu I - XS - SX$$

becomes:

$$
\begin{aligned}
-\Delta y\left\{\begin{pmatrix}1 & \sqrt{2}\\ \sqrt{2} & 3\end{pmatrix}\begin{pmatrix}-1 & \sqrt{2}\\ \sqrt{2} & 0\end{pmatrix}+\begin{pmatrix}-1 & \sqrt{2}\\ \sqrt{2} & 0\end{pmatrix}\begin{pmatrix}1 & \sqrt{2}\\ \sqrt{2} & 3\end{pmatrix}\right\} \quad &+ \\
\left\{\begin{pmatrix}1 & 0\\ 0 & 11\end{pmatrix}\begin{pmatrix}2\sqrt{2}\Delta v_1 & \Delta v_1\\ \Delta v_1 & \Delta v_2\end{pmatrix}+\begin{pmatrix}2\sqrt{2}\Delta v_1 & \Delta v_1\\ \Delta v_1 & \Delta v_2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 11\end{pmatrix}\right\} \quad &= \\
= \begin{pmatrix}34\sigma - 2 & -12\sqrt{2}\\ -12\sqrt{2} & 34\sigma - 66\end{pmatrix} &
\end{aligned}
$$

The left-hand-side coefficient matrix is:

$$
\begin{pmatrix}2 & 4\sqrt{2} & 0\\ 3\sqrt{2} & 12 & 0\\ 4 & 0 & 22\end{pmatrix}\begin{pmatrix}-\Delta y\\ \Delta v_1\\ \Delta v_2\end{pmatrix} = \begin{pmatrix}34\sigma - 2\\ -12\sqrt{2}\\ 34\sigma - 66\end{pmatrix}
$$

SINGULAR - inconsistent for $\sigma \in [0,1]$

7.5 Conditioning of Search Directions

Most popular directions: AHO, HrvwKshM, NT

Cost: AHO - High; NT - low; HKM - lowest

-

---

well-defined $\cong$ well-conditioned

Under standard nondegeneracy assumptions:

1. **well-defined at** $X, S \succ 0$**:**

   HKM and NT *is* always; but, AHO *not* always

2. **well-defined at optimal** X, S**:**

   HKM and NT *not*; but, AHO *is*

8 Preview Sufficient Conditions

8.1 Shida-Shindoh-Kojima (SSK) Type Sufficient Conditions

AHO search direction is well-defined (existence and uniqueness), whenever

$$\text{(SSKcond)} \qquad X, S \succ 0 \text{ and } XS + SX \succeq 0.$$

8.1.1 Strengthened (SSKcond) Condition

$$\left\{ \begin{array}{c} \text{AHO is well-defined whenever:} \\ \boxed{2\lambda_{\min}(S)X} + (\text{SX+XS}) \ \text{ is definite.} \end{array} \right.$$

-

Note: if and only if

$$K := \frac{1}{2}\left[I \otimes (SX + XS)\right] + \left[X \otimes S\right] \text{ is definite,}$$

where $\otimes$ denotes the Kronecker product. Or the stronger:

$$K + TKT + TK + KT \text{ is semidefinite with rank } n(n+1)/2,$$

where T is the matrix representation of the transpose operator.

8.2 Monteiro-Zanjácomo (MZ) Type Sufficient Conditions

centrality type (spectral norm):

$$\text{(MZcond)} \quad X, S \succ 0 \text{ and } \|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| \leq \frac{1}{2}\nu, \quad \text{for some } \nu > 0$$

Central Path:

$$X, S \succ 0 \text{ and } SX - \nu I = 0, \text{ for some } \nu > 0$$

The condition is equivalent to the bound on the condition number ($\gamma(S^{\frac{1}{2}}XS^{\frac{1}{2}})$ = largest eigenvalue/smallest eigenvalue)

$$\text{(MZequivalent)} \qquad X, S \succ 0 \text{ and } \gamma(S^{\frac{1}{2}}XS^{\frac{1}{2}}) \leq 3,$$

-

8.2.1 STRENGTHENED MZ

- AHO is well-defined whenever $\frac{1}{2}\nu$ above is replaced by $\frac{3}{5}\nu$
- equivalent condition number bound increases from 3 to 4.

-

Or is it strengthened??

8.3 Monteiro-Todd (MT) Sufficient Conditions

centrality type (spectral norm):

$$\text{(MTcond)} \quad X, S \succ 0 \text{ and } \|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| < \frac{1}{\sqrt{2}}\nu, \quad \text{for some } \nu > 0$$

$$\text{(MTequivalent)} \qquad X, S \succ 0 \text{ and } \gamma(S^{\frac{1}{2}} X S^{\frac{1}{2}}) < (3 + 2\sqrt{2}),$$

8.3.1 STRENGTHENED MT

$1/\sqrt{2}(\cong .7071) \rightarrow (\sqrt{3} - 1)(\cong .7321) \qquad < \alpha = .9193$ (example)

$(3 + 2\sqrt{2})(\cong 5.8284) \rightarrow (3 + 2\sqrt{3})(\cong 6.4641)$

9 Kronecker and Symmetric Kronecker Algebra

for compatible matrices $K, N, M \in \mathcal{M}^n$:

$$\operatorname{vec}(NKM^T) = (M \otimes N)\operatorname{vec}(K),$$

$\operatorname{vec}(K)$ is: vector formed from the columns of K

$$M \otimes N = \{M_{ij}N\} = \begin{pmatrix} M_{11}N & \dots & M_{1n}N \\ \dots & \dots & \dots \\ M_{n1}N & \dots & M_{nn}N \end{pmatrix} \quad \textit{Kronecker product.}$$

quadratic form is:

$$\left\langle L, NKM^T \right\rangle = \operatorname{trace} L^T NKM^T = \operatorname{vec}(L)^T (M \otimes N)\operatorname{vec}(K)$$

9.1 Some Properties of $\otimes$ and $\overset{s}{\otimes}$

Let λ, v and μ, w be eigenpairs for M, N, respectively. An eigenpair for $\otimes$ is given by:

$$(M \otimes N)(v \otimes w) = \lambda\mu(v \otimes w)$$

For compatible matrices

$$(A \otimes B)(C \otimes D) = (AC \otimes BD)$$

and

$$(A \otimes B)^T = (A^T \otimes B^T).$$

Special Quadratic Form:

Lemma: Let A, B, C be $n \times n$ matrices. Then

$$2\operatorname{trace}(ABC) = \left[\operatorname{vec}(B)^T | \operatorname{vec}(C)^T\right] \left[\begin{array}{c|c} 0 & I \otimes A^T \\ \hline I \otimes A & 0 \end{array}\right] \left[\begin{array}{c} \operatorname{vec}(B) \\ \hline \operatorname{vec}(C) \end{array}\right]$$

$\blacksquare$

Let $t(n) = n(n+1)/2$;
$\mathrm{s2vec}\,(X) \in \mathbb{R}^{t(n)}$ denote isometry $\mathcal{S}^n \longleftrightarrow \mathbb{R}^{t(n)}$ (upper triang. X)
$\mathrm{trace}\, XY = \mathrm{s2vec}\,(X)^T \mathrm{s2vec}\,(Y)$

-

---

symmetric Kronecker product for $U \in \mathcal{S}^n$

$$(M \overset{s}{\otimes} N)\mathrm{s2vec}\,(U) := \mathrm{s2vec}\,\left(\frac{1}{2}(NUM^T + MUN^T)\right).$$

The quadratic form with $U \in \mathcal{S}^n$ is

$$\begin{aligned}
\mathrm{trace}\,\tfrac{1}{2}(NUM^T + MUN^T)U &= \mathrm{s2vec}\,(U)^T (M \overset{s}{\otimes} N)\mathrm{s2vec}\,(U) \\
&= \mathrm{s2vec}\,(U)^T (M \overset{s}{\otimes} N)^T \mathrm{s2vec}\,(U) \\
&= \mathrm{s2vec}\,(U)^T (M^T \overset{s}{\otimes} N^T)\mathrm{s2vec}\,(U).
\end{aligned}$$

Define $Q \in \mathbb{R}^{t(n) \times n^2}$ so that: rows consist of $\operatorname{vec}(E_{kl})$, where $\{E_{kl} : 1 \leq k \leq l \leq n\}$ is the set of $t(n)$ orthonormal basis matrices of $\mathcal{S}^n$.

$$Q_{(i,j),(k,l)} := \begin{cases} 1 & \text{if } i = j = k = l \\ \frac{1}{\sqrt{2}} & \text{if } i = k \neq j = l \text{ or } i = l \neq j = k \\ 0 & \text{otherwise} \end{cases}$$

Then $QQ^T = I$, $Q^T Q$ is matrix representation of orthogonal projection of $\mathcal{M}^n$ (as $\mathbb{R}^{n^2}$) onto $\mathcal{S}^n$ (as a subspace of $\mathbb{R}^{n^2}$) and, for $A, B \in \mathcal{S}^n$,

$$A \overset{s}{\otimes} B = \frac{1}{2} Q \left(A \otimes B + B \otimes A \right) Q^T \in \mathcal{S}^{t(n)}$$

9.2 Eigenvalue Relationships

Similar eigenpair result for *commuting matrix pair*

Theorem: Let $A, B \in \mathcal{S}^n$ such that A and B commute. Let λ_i, μ_i, v_i denote the eigenvalues and the corresponding (common) eigenvectors of A and B, respectively. Then, for $1 \leq i \leq j \leq n$, $\frac{1}{2}(\lambda_i \mu_j + \lambda_j \mu_i)$ are the eigenvalues and s2vec $\left(\frac{1}{2}\left[v_i v_j^T + v_j v_i^T\right]\right)$ are the corresponding eigenvectors of $(A \overset{s}{\otimes} B)$. ■

Proposition: If $Av = 0$, then vv^T is an eigenvector for $A \overset{s}{\otimes} B$. ■

Theorem: (Surprising result) Let $A, B \in \mathcal{S}^n$. Then
$\left(A \overset{s}{\otimes} B\right) \succeq 0$ (resp. $\succ 0$) $\iff$ $(A \otimes B) \succeq 0$ (resp. $\succ 0$).
Proof Outline. (necessity for $\succeq$) Let
$\lambda := \min\{0, \lambda_{\min}(A)\}$, $\mu := \min\{0, \lambda_{\min}(B)\}$. Suppose contrapositive, i.e. $\lambda_{\min}(A \otimes B) < 0$. Wlog: $\lambda < 0$. Now show: $\lambda_{\min}(A \overset{s}{\otimes} B) < 0$ (using eigenvector $\mathrm{s2vec}\,(vv^T)$ for $A \overset{s}{\otimes} B$ from 0 eigenvalue/vector v of A or B)

$$
\begin{aligned}
2(A \overset{s}{\otimes} B) &= Q(A \otimes B + B \otimes A)Q^T \\
&= Q((A - \lambda I) \otimes (B - \mu I) + (B - \mu I) \otimes (A - \lambda I))Q^T + \\
&\quad \mu Q((A - \lambda I) \otimes I + I \otimes (A - \lambda I))Q^T + \\
&\quad \lambda Q(I \otimes (B - \mu I) + (B - \mu I) \otimes I)Q^T
\end{aligned}
$$

■

Theorem: (More Surprise)

Let $A, B \in \mathcal{M}^n$. For $u \in \mathbb{R}^{t(n)}$, we have the eigenpair relationship

$$\left(A \overset{s}{\otimes} B\right) u = \lambda u \quad \Longrightarrow \quad \frac{1}{2}\left(A \otimes B + B \otimes A\right)\left(Q^T u\right) = \lambda(Q^T u).$$

$\blacksquare$

10 E&U Conditions

10.1 E&U Conditions Using an Indefinite Trust Region Subproblem, TRS

Lemma: Suppose that $(\Delta X, \Delta y, \Delta S)$ solves AHO linearization (AHOlin) for a given pair $X \succ 0$, $S \succ 0$. Then the following holds:

1. $\operatorname{trace}(\Delta X \Delta S) = 0$ (from: $\Delta X \in \mathcal{N}(\mathcal{A})$, $\Delta S \in \mathcal{R}(\mathcal{A}^*)$)

2. The matrix
$$M := X(\Delta S) + (\Delta X)S$$
is *skew-symmetric*, i.e. $M = -M^T$
(from: last *symmetrized* block in AHOlin)

3.

$$M = X(\Delta S) + (\Delta X)S = 0$$

if and only if

$$\text{AHOlin} \Rightarrow \Delta X = \Delta S = 0, \quad \Delta y = 0.$$

(from:

$$\begin{aligned} 0 = \operatorname{trace} \Delta X \Delta S &= \operatorname{trace}(\Delta X)\left[-X^{-1}(\Delta X)S\right] \\ &= -\operatorname{trace} S^{\frac{1}{2}}(\Delta X)X^{-\frac{1}{2}}X^{-\frac{1}{2}}(\Delta X)S^{\frac{1}{2}} \\ &= -\|X^{-\frac{1}{2}}(\Delta X)S^{\frac{1}{2}}\|_F^2 \quad) \end{aligned}$$

Corollary:

$$\left\{\begin{array}{rcl} 0 = & \min & q_0(\Delta X, \Delta S) := \operatorname{trace}\left(X(\Delta S) + (\Delta X)S\right)^2 \\ & \text{subject to} & q_1(\Delta X, \Delta S) := \operatorname{trace} \Delta S \Delta X = 0 \\ & & q_2(\Delta X, \Delta S) := \|L_S(\Delta X) + L_X(\Delta S)\|^2 = 0 \end{array}\right.$$

$$(L_S(\Delta X) := S(\Delta X) + (\Delta X)S \rightarrow \text{ Lyapunov operator})$$

if and only if

$$\text{AHOlin} \Rightarrow \Delta X = \Delta S = 0 \quad (\text{ indep. of } \mathcal{A})$$

■

Optimality Conditions=? Use to find sufficient conditions for E&U?

10.1.1 Strengthened MZ

Theorem: Suppose that $S, X \succ 0$. Then the equivalent (centrality type) conditions, in Items 1 and 2, imply that $(X, S) \in AHO(\mathcal{L})$.

1. $$\|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| \leq \frac{3}{5}\nu, \quad \text{for some scalar } \nu > 0.$$

2. The condition number
$$\gamma(S^{\frac{1}{2}} X S^{\frac{1}{2}}) \leq 4.$$

PROOF: Set $D^2 := S^{-1/2}\left(S^{1/2}XS^{1/2}\right)^{1/2}S^{-1/2}$ and $V := DSD = D^{-1}XD^{-1}$. Note eigenvalues of V^2 and $S^{1/2}XS^{1/2}$ are the same. Define $\overline{\Delta X} := D^{-1}(\Delta X)D^{-1}$, and $\overline{\Delta S} := D(\Delta S)D$. Note that

$$\text{trace}\left[(\overline{\Delta X})(\overline{\Delta S})\right] = \text{trace}\left[(\Delta X)(\Delta S)\right] = 0.$$

By above Corollary, we need only show that

$$\text{TRS}_M \qquad 0 \leq \min\{\text{trace}\left(M^2\right) : \text{trace}\,(\overline{\Delta X})(\overline{\Delta S}) = 0\}.$$

$$
\begin{aligned}
\operatorname{trace}\left(M^2\right) &= \operatorname{trace}\left\{\left[X(\Delta S)+(\Delta X)S\right]^2\right\} \\
&= \operatorname{trace}\left[V(\overline{\Delta S})V(\overline{\Delta S})\right]+\operatorname{trace}\left[V(\overline{\Delta X})V(\overline{\Delta X})\right]+ \\
&\qquad 2\operatorname{trace}\left[(\overline{\Delta X})V^2(\overline{\Delta S})\right] \\
&= \operatorname{trace}\left[V(\overline{\Delta S})V(\overline{\Delta S})\right]+\operatorname{trace}\left[V(\overline{\Delta X})V(\overline{\Delta X})\right] \\
&\qquad +2\operatorname{trace}\left[(V^2-\mu I)(\overline{\Delta X})(\overline{\Delta S})\right].
\end{aligned}
$$

The μ plays the role of a Lagrange multiplier for TRS_M problem. Let $s = \operatorname{s2vec}\left(\overline{\Delta S}\right), x = \operatorname{s2vec}\left(\overline{\Delta X}\right)$. Then using above Lemma yields

$$
\operatorname{trace}\left(M^2\right) = \begin{pmatrix} x \\ s \end{pmatrix}^T \left\{ \begin{pmatrix} V \overset{s}{\otimes} V & 0 \\ 0 & V \overset{s}{\otimes} V \end{pmatrix} + \begin{pmatrix} 0 & I \overset{s}{\otimes} \left(V^2-\mu I\right) \\ I \overset{s}{\otimes} \left(V^2-\mu I\right) & 0 \end{pmatrix} \right\} \begin{pmatrix} x \\ s \end{pmatrix}
$$

Note: TRS_M is an indefinite trust region subproblem and the optimum can be characterized using positive semidefiniteness of the Hessian of the Lagrangian, i.e. the optimal value is finite (so nonnegative) if and only if there exits a μ that makes the Hessian positive semidefinite.

$$\begin{pmatrix} V \overset{s}{\otimes} V & I \overset{s}{\otimes} \left(V^2 - \mu I\right) \\ I \overset{s}{\otimes} \left(V^2 - \mu I\right) & V \overset{s}{\otimes} V \end{pmatrix} \succeq 0, \text{ for some } \mu. \qquad (*)$$

Take Schur complement in (*) to get the equivalent condition

$$V \overset{s}{\otimes} V - \left(I \overset{s}{\otimes} \left(V^2 - \mu I\right)\right) \left(V \overset{s}{\otimes} V\right)^{-1} \left(I \overset{s}{\otimes} \left(V^2 - \mu I\right)\right) \succeq 0. \qquad (**)$$

Note that the matrices $V, (V^2 - \mu I)$ and I all commute. We can diagonalize all four terms to yield the following result for the eigenvalues λ_i of V,

$$\lambda_i^2 \lambda_j^2 - (\lambda_i^2 - \mu)^2 \geq 0, \quad \forall i, j$$

$$\text{Equivalently}: \qquad \lambda_i \lambda_j \geq |\lambda_i^2 - \mu|, \quad \forall i, j.$$

$$\text{Equivalently}: \qquad \lambda_{\max}^2 - \lambda_{\max}\lambda_{\min} \leq \mu \leq 2\lambda_{\min}^2.$$

divide by $\lambda_{\min}$; let $\gamma = \frac{\lambda_{\max}}{\lambda_{\min}}$,
then condition number satisfies $\gamma^2 - \gamma - 2 \leq 0$, i.e. $\gamma \leq 2$ or $\gamma^2 \leq 4$.

Equivalently $|\lambda_i^2 - \nu| \leq \frac{3}{5}\nu, \forall i$, i.e. $\lambda_j^2 - \frac{3}{5}\nu \leq \nu \leq \lambda_i^2 + \frac{3}{5}\nu$
equivalently: $\frac{5}{8}\lambda_j^2 \leq \nu \leq \frac{5}{2}\lambda_i^2$
equivalently condition number has bound of 4 $\blacksquare$

10.2 Conditions Using a Subspace

Definition Let $\mathcal{L} \subset \mathcal{S}^n$ be a subspace. Define

$$AHO(\mathcal{L}) := \{(X,S) \in \mathcal{S}^n \times \mathcal{S}^n : \quad \text{AHOlin, determined by } (X,S) \;\&\; \mathcal{L}, \text{ has unique solution } \}$$

$$\overline{AHO}(\mathcal{L}) := \{(X,S) \in \mathcal{S}^n \times \mathcal{S}^n : (X,S) \notin AHO(\mathcal{L})\}$$

For $\mathcal{L} = \mathcal{N}(\mathcal{A})$:

$$AHO(\mathcal{A}) = AHO(\mathcal{L}), \; \overline{AHO}(\mathcal{A}) = \overline{AHO}(\mathcal{L}).$$

Proposition: Let $(X, S) \in \mathcal{S}^n \times \mathcal{S}^n$ and a subspace $\mathcal{L} \subset \mathcal{S}^n$ be given. Then $(X, S) \in \overline{AHO}(\mathcal{L})$ if and only if

$$L_X(\mathcal{L}^\perp) \cap L_S(\mathcal{L}) \neq \{0\} \qquad (*)$$

Moreover, if L_X is invertible (similar if L_S invertible), then (*) holds if and only if

$$L_X^{-1} L_S\,(\Delta X) \in \mathcal{L}^\perp, \quad \text{for some } 0 \neq \Delta X \in \mathcal{L}.$$

Equivalently, for $\mathcal{L} = \mathcal{N}(\mathcal{A})$, if and only if $\exists\ \Delta X \in \mathcal{N}(\mathcal{A}) \setminus \{0\}$ such that $\Delta S = L_X^{-1} L_S(\Delta X) \in \mathbb{R}(\mathcal{A}^*)$.

Corollary: If L_X is invertible (similar result if L_S is invertible), then

$$\left(L_X^{-1} L_S + (L_X^{-1} L_S)^*\right) \text{ is definite } \Rightarrow (X, S) \in AHO(\mathcal{A}).$$

PROOF: Note that

$$L_X^{-1} L_S\,(\Delta X) \in \mathcal{L}^\perp, \quad \text{for some } 0 \neq \Delta X \in \mathcal{L}.$$

holds only if there exists $0 \neq \Delta X \in \mathcal{L}$ such that

$$\left\langle \Delta X, L_X^{-1} L_S\,(\Delta X)\right\rangle = \operatorname{trace} \Delta X L_X^{-1} L_S\,(\Delta X) = 0.$$

This cannot happen if the quadratic form is definite. ∎

10.2.1 Strengthened SSK

The four (equivalent) conditions 1.a-d:

-

(1a) • $K = \frac{1}{2}\left[I \otimes (SX + XS)\right] + \left[X \otimes S\right]$ is definite

(1b) • $\frac{1}{2}(\bar{S}D_X + D_X\bar{S}) + \lambda_{\min}(S)D_X$ is definite

(1c) • $2\lambda_{\min}(S)X + (SX + XS)$ is definite

(1d) • If a symmetric nonsingular square root of X exists :

$\frac{1}{2}\lambda_{\min}(R + R^T) >$ (or $<$) $- \lambda_{\min}(R)$, with $R = X^{-\frac{1}{2}}SX^{\frac{1}{2}}$

IMPLY the four (equivalent) conditions (2a)-(2d) $\Downarrow \Downarrow$

(2a) • $\frac{1}{2}\left[I \overset{s}{\otimes} (SX + XS)\right] + \left[X \overset{s}{\otimes} S\right]$ is definite

(2b) • $L_X^{-1} L_S + L_S L_X^{-1}$ is definite

(2c) • $L_S L_X + L_X L_S$ is definite

(2d) • $K + TKT + TK + KT$ is semidefinite with rank $t(n)$

IMPLY

$$\Downarrow \quad \Downarrow$$

$$(X, S) \in AHO(\mathcal{L}), \text{ for all subspaces } \mathcal{L} \in \mathcal{S}^n \, .$$

Proof Outline. Again, expand quadratic form:

$$
\begin{aligned}
& \langle U, (L_S L_X + L_X L_S)(U) \rangle = \\
= \ & \operatorname{trace}\left\{U\left[L_S(XU + UX) + L_X(SU + US)\right]\right\} \\
= \ & 2\operatorname{trace}\left(SUXU + XUSU\right) \\
& +\operatorname{trace}\left[(XS + SX)UIU + IU(SX + XS)U\right] \\
= \ & 4\,\mathrm{s2vec}\,(U)^T \left(X \overset{s}{\otimes} S\right) \mathrm{s2vec}\,(U) \\
& +2\,\mathrm{s2vec}\,(U)^T \left(I \overset{s}{\otimes} (XS + SX)\right) \mathrm{s2vec}\,(U).
\end{aligned}
$$

$\blacksquare$

11 Generic Properties

Theorem: For every subspace $\mathcal{L}$ of $\mathcal{S}^n$, we have $AHO(\mathcal{L})$ is an open set and $\overline{AHO}(\mathcal{L})$ is a set of measure zero.

PROOF: Recall that $(X, S) \in AHO(\mathcal{L})$ <u>iff</u> $\mathcal{K} = \begin{pmatrix} 0 & \mathcal{A}^* & I \\ \mathcal{A} & 0 & 0 \\ L_S & 0 & L_X \end{pmatrix}$ nonsingular. Since $\det(\mathcal{K}(X, S))$ analytic in X, S, the result follows from Baire category theorem. (zero set of analytic function is set of measure zero and complement is open by continuity)
(Alternatively: $\overline{AHO}(\mathcal{L})$ and therefore $AHO(\mathcal{L})$ are both semi-algebraic sets; and semi-algebraic sets without interior are of measure zero.) ∎

12 Extended QQP/Hessian

Using Corollary on Transparency 49; with

$$s := \Delta S, x := \Delta X, \overline{s} := \overline{\Delta S}, \overline{x} := \overline{\Delta X} \in \mathcal{S}^n .$$

Lagrangian is:

$$\begin{aligned} L(x, s, \mu, \lambda, \alpha) &= q_0(x, s) + \mu q_1(x, s) + \lambda q_2(x, s) \\ &= \text{trace } \{XsXs + SxSx + 2XsxS + \\ &\qquad\qquad \mu\,[xs] + \\ &\qquad\qquad \lambda\,[\; XsXs + SxSx + 2XsxS + \\ &\qquad\qquad\quad X^2ss + S^2xx + 2SxXs \;]\,\} . \end{aligned}$$

Rewrite using definitions of D, V on Transp. 51:

$$
\begin{aligned}
&L(\overline{x}, \overline{s}, \mu, \lambda, \alpha) = \\
&\quad = \operatorname{trace} \Big\{ \left[V\overline{x}V\overline{x} + V\overline{s}V\overline{s} + 2\overline{x}V^2\overline{s}^T \right] + \\
&\qquad\qquad \mu \left[\overline{xs}^T \right] + \\
&\qquad\qquad \lambda \Big[V\overline{x}V\overline{x} + V\overline{s}V\overline{s} + 2\overline{x}V^2\overline{s}^T \; + \\
&\qquad\qquad\quad VD^{-2}V\overline{x}D^2\overline{x}^T + VD^2V\overline{s}D^{-2}\overline{s}^T + 2D^{-2}V\overline{x}D^2V\overline{s} \; \Big] \Big\}
\end{aligned}
$$

Hessian of Lagrangian

$$\nabla^2 L \;=\; \begin{pmatrix} V \overset{s}{\otimes} V & V^2 \overset{s}{\otimes} I \\ V^2 \overset{s}{\otimes} I & V \overset{s}{\otimes} V \end{pmatrix} + \mu \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$+\lambda \left[\begin{pmatrix} V \overset{s}{\otimes} V & V^2 \overset{s}{\otimes} I \\ V^2 \overset{s}{\otimes} I & V \overset{s}{\otimes} V \end{pmatrix} + \begin{pmatrix} D^2 \overset{s}{\otimes} V D^{-2} V & V D^2 \overset{s}{\otimes} D^{-2} V \\ D^2 V \overset{s}{\otimes} V D^{-2} & V D^2 V \overset{s}{\otimes} D^{-2} \end{pmatrix} \right]$$