Semidefinite Programing
and
Matrix Completions for Partial Hermitian Matrices

Henry Wolkowicz

Department of Combinatorics & Optimization
University of Waterloo

1.2 Basic Properties and Notation

Basic linear **Semidefinite Programming** looks just like Linear Programming

$$
\textbf{(PSDP)} \qquad
\begin{array}{rcll}
p^* = & \max & \operatorname{trace} CX & (\langle C, X \rangle) \\
 & \text{s.t.} & \mathcal{A}X = b & \text{(linear)} \\
 & & X \succeq 0, \quad (X \in \mathcal{P}) & \text{(nonneg)}
\end{array}
$$

$\mathcal{S}^n :=$ space of $n \times n$ symmetric matrices

For $A \in \mathcal{S}^n$:
A is positive semidefinite (positive definite), (denoted $A \succeq 0$ $(A \succ 0)$,)
if $x^T A x \geq 0 (> 0), \ \forall x \neq 0$.

$\preceq$ denotes the Löwner partial order, [19]
$A \preceq B$ if $B - A \succeq 0$ (positive semidefinite)

TFAE:

1. $A \succeq 0$
2. the eigenvalues $\lambda(A) \geq 0 \quad (\lambda(A) > 0)$

The linear operator $\mathcal{A}$ and its adjoint:

$$\mathcal{A} : \mathcal{S}^n \rightarrow \Re^m, \qquad \mathcal{A}^* : \Re^m \rightarrow \mathcal{S}^n$$

$$\text{for given } A_i \in \mathcal{S}^n, i = 1, \ldots m :$$

$$(\mathcal{A}X)_i = \operatorname{trace}(A_i X), \quad \mathcal{A}^* y = \textstyle\sum_{i=1}^m y_i A_i$$

where

$$\langle \mathcal{A}X, y \rangle = \langle X, \mathcal{A}^* y \rangle, \quad \forall X, y$$

$\mathcal{P}$ - cone of positive semidefinite matrices

replaces

$\Re^n_+$ - nonnegative orthant

$\operatorname{trace} CX = \langle C, X \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij}$

SDP is equivalent to:

$$
\textbf{(PSDP)} \qquad \begin{array}{rcll} p^* = & \max & \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} & \\ & \text{s.t.} & \langle A_i, X \rangle = b_i, & i = 1, \ldots m \\ & & X \succeq 0, \quad (X \in \mathcal{P}) & \end{array}
$$

1.2.1 Duality

payoff function, player Y to player X (Lagrangian)

$$L(X, y) := \operatorname{trace}(CX) + y^T(b - \mathcal{A}X)$$

Optimal (worst case) strategy for player X:

$$p^* = \max_{X \succeq 0} \min_{y} L(X, y)$$

Using the *hidden constraint* $b - \mathcal{A}X = 0$, recovers primal problem (PSDP).

$$
\begin{aligned}
L(X,y) &= \operatorname{trace}(CX) + y^T(b - \mathcal{A}X) \\
&= b^T y + \operatorname{trace}\,(C - \mathcal{A}^* y)\, X
\end{aligned}
$$

adjoint operator, $\quad \mathcal{A}^* y = \sum_i y_i A_i$

$$
\langle \mathcal{A}^* y, X \rangle = \langle y, \mathcal{A}X \rangle, \quad \forall X, y
$$

$$
p^* = \max_{X \succeq 0} \min_y L(X,y) \leq d^* := \min_y \max_{X \succeq 0} L(X,y)
$$

For dual, use *hidden constraint* $g(y) := C - \mathcal{A}^* y \preceq 0$

$$p^* = \max_{X \succeq 0} \min_y L(X, y) \leq d^* := \min_y \max_{X \succeq 0} L(X, y)$$

dual obtained from optimal strategy of competing player Y;
use *hidden constraint* $g(y) = C - \mathcal{A}^* y \preceq 0$

$$(\mathbf{DSDP}) \qquad \begin{array}{rcc} d^* = & \min & b^T y \\ & \text{s.t.} & \mathcal{A}^* y \succeq C \end{array}$$

for the primal

$$(\mathbf{PSDP}) \qquad \begin{array}{rcc} p^* = & \max & \operatorname{trace} CX \\ & \text{s.t.} & \mathcal{A}X = b \\ & & X \succeq 0 \end{array}$$

1.2.2 Weak Duality - Optimality

Proposition 1.2.2.1 *(Weak Duality) If X is feasible in (PSDP) and y is feasible in (DSDP), then*

$$\operatorname{trace} CX - b^T y = \operatorname{trace} XZ \leq 0,$$

where $Z = C - \mathcal{A}^* y \succeq 0$ is the slack variable.

Proof. (Direct)

$$\begin{aligned} \operatorname{trace} CX - b^T y &= \operatorname{trace} (\mathcal{A}^* y - Z)X - b^T y \\ &= \operatorname{trace} y^T \mathcal{A} X - \operatorname{trace} ZX - b^T y \\ &= \operatorname{trace} y^T (\mathcal{A} X - b) - \operatorname{trace} ZX = -\operatorname{trace} ZX, \end{aligned}$$

since X is feasible. The result follows since
$\operatorname{trace} ZX = \operatorname{trace} X^{1/2} X^{1/2} Z = \operatorname{trace} X^{1/2} Z X^{1/2} \geq 0.$ ■

Characterization of optimality for the dual pair X, y (slack $Z \succeq 0$)

$$\begin{aligned} \mathcal{A}^* y - Z &= C && \text{dual feasibility} \\ \mathcal{A} X &= b && \text{primal feasibility} \\ Z X &= 0 && \text{complementary slackness} \end{aligned}$$

$$ZX = \mu I \qquad \text{perturbed}$$

Forms the basis for:

interior point methods
(primal simplex method, dual simplex method)

1.2.3 Minimizing the Maximum Eigenvalue

Arises in e.g. stability of differential equations

- The mathematical problem:
  - ◇ given $A(x)$ depending linearly on vector x
  - ◇ Find x to minimize the maximum eigenvalue of $A(x)$
- SDP Model:
  - ◇ the largest eigenvalue $\lambda_{\max}(A(x)) \leq \alpha$ <u>iff</u> $\lambda_{\max}(A(x) - \alpha I) \leq 0$ <u>iff</u> $A(x) - \alpha I \preceq 0$.
  - ◇ The DSDP (in dual form) is:

$$\max -\alpha \qquad \text{s.t.} \quad A(x) - \alpha I \preceq 0.$$

1.2.4 Pseudoconvex (Nonlinear) Optimization Problem

$$
\textbf{(PCP)} \qquad \begin{array}{rll} d^* = & \min & \frac{(c^T x)^2}{d^T x} \\ & \text{s.t.} & Ax + b \geq 0, \end{array}
$$

(where $Ax + b \geq 0 \Rightarrow d^T x > 0$) is equivalent to

$$
\begin{array}{rll} d^* = & \min & t \\ & \text{s.t.} & \begin{bmatrix} \operatorname{Diag}(Ax+b) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succeq 0 \end{array}
$$

1.2.5 Further Properties

1. For matrices P, Q compatible for multiplication,

$$\operatorname{trace} PQ = \operatorname{trace} QP.$$

2. Every symmetric matrix S has an orthogonal decomposition $S = Q\Lambda Q^T$, where Λ is diagonal with the eigenvalues of S on the diagonal, and Q has orthonormal columns consisting of eigenvectors of S.

3. For S symmetric, TFAE:
   (a) $S \succeq 0 \quad (\succ 0)$
   (b) $v^T S v \geq 0, \forall v \in \Re^n \quad (v^T S v > 0, \forall 0 \neq v \in \Re^n)$
   (c) $\lambda(S) \geq 0 \quad (> 0)$
   (d) $S = P^T P$ for some P (for some nonsingular P)

Linear Matrix Inequality $\quad g(y) = C - \mathcal{A}^* y \preceq 0$

adjoint operator $\quad \mathcal{A}^* y = \sum_i y_i A_i, \quad A_i = A_i^T$

$g(y)$ is a convex constraint, i.e.

$$g(\lambda x + (1-\lambda)y) \preceq \lambda g(x) + (1-\lambda) g(y), \quad \forall 0 \leq \lambda \leq 1, \forall x, y \in \Re^m$$

equivalently, for every $P \succeq 0$, the real valued function $f_P(x) := \operatorname{trace} P g(x)$ is convex, i.e.

$$f_P(\lambda x + (1-\lambda)y) \leq \lambda f_P(x) + (1-\lambda) f_P(y), \quad \forall 0 \leq \lambda \leq 1, \forall x, y \in \Re^m$$

feasible set $\{x \in \Re^m : g(x) \preceq 0\}$ is convex set

optimum point is on the boundary ($g(x)$ is singular)

boundary of feasible set is *not* smooth (piecewise smooth, algebraic surface)

1.3 Motivation/Examples/Applications

- Quadratic constrained quadratic programs
- Lovász theta function
- Statistics
- minimizing the L_2-operator norm of a matrix
- linear programming
- robust mathematical programming
- engineering, e.g. control theory
- combinatorial problems, e.g. the max-cut problem

1.3.1 Quadratic Constrained Quadratic Programs

What is SEMIDEFINITE PROGRAMMING?
Why use it?

• Quadratic approximations are better than linear approximations. And, we can solve relaxations of quadratic approximations efficiently using semidefinite programming.

HOW DOES SDP arise from quadratic approximations?

$$q_i(y) = \frac{1}{2} y^T Q_i y + y^T b_i + c_i, \;\; y \in \Re^n$$

$$(\mathbf{QQP}) \qquad \begin{cases} q^* = & \min & q_0(y) \\ & \text{s.t.} & q_i(y) = 0 \\ & & i = 1, \ldots m \end{cases}$$

- Lagrangian (quadratic, linear, constant in y terms):

$$L(y, x) = \frac{1}{2} y^T (Q_0 - \sum_{i=1}^m x_i Q_i) y + y^T (b_0 - \sum_{i=1}^m x_i b_i) + (c_0 - \sum_{i=1}^m x_i c_i)$$

- Primal-Dual pair: $\quad q^* = \min_y \max_x L(y, x) \geq d^* = \max_x \min_y L(y, x)$

- homogenize (add y_0): $\quad y_0 y^T (b_0 - \sum_{i=1}^m x_i b_i), \;\; y_0^2 = 1.$

$$
\begin{aligned}
d^* &= \max_x \min_y && L(y,x) \\
&= \max_x \min_{y_0^2=1} && \tfrac{1}{2} y^T (Q_0 - \textstyle\sum_{i=1}^m x_i Q_i) y \quad (+t y_0^2) \\
& && + y_0 y^T (b_0 - \textstyle\sum_{i=1}^m x_i b_i) \\
& && + (c_0 - \textstyle\sum_{i=1}^m x_i c_i) \quad (-t)
\end{aligned}
$$

hidden semidefinite constraint yields SDP constraint

$$
B - \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} \succeq 0.
$$

($\mathcal{A} : \Re^{m+1} \to \mathcal{S}_{n+1}$)

$$
B = \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix}, \quad \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{bmatrix} -t & \sum_{i=1}^m x_i b_i^T \\ \sum_{i=1}^m x_i b_i & \sum_{i=1}^m x_i Q_i \end{bmatrix}
$$

The dual program is equivalent to the SDP (with $c_0 = 0$)

$$
(\mathbf{D}) \qquad \begin{array}{rll} d^* = & \sup & -\sum_{i=1}^m x_i c_i - t \\ & \text{s.t.} & \mathcal{A}\begin{pmatrix} t \\ x \end{pmatrix} \preceq B \\ & & x \in \Re^m, t \in \Re \end{array}
$$

As in linear programming, the dual is obtained from the optimal strategy of the competing player:

$$
(\mathbf{DD}) \qquad \begin{array}{rll} d^* = & \inf & \operatorname{trace} BU \\ & \text{s.t.} & \mathcal{A}^* U = \begin{pmatrix} -1 \\ -c \end{pmatrix} \\ & & U \succeq 0. \end{array}
$$

4.5 Matrix Completions

4.5.1 Outline

- background
- Positive Semidefinite matrix completions
- Euclidean Distance matrix completions
- exploiting sparsity
- A New Characterization of EDMs using Gale Matrices

4.5.2 Positive Definite Completions of Partial Hermitian Matrices

- $\mathcal{G}(V, E)$ *finite undirected graph*, (node set V, edge set E)

- $A(\mathcal{G})$ is a $\mathcal{G}$*-partial matrix*, (a_{ij} defined *iff* $\{i, j\} \in E$)

- $A(\mathcal{G})$ is a $\mathcal{G}$*-partial positive matrix* if $a_{ij} = \overline{a_{ji}}, \forall \{i, j\} \in E$ and all existing principal minors are positive.

- with $\mathcal{J} = (V, \bar{E}), E \subset \bar{E}$ *a* $\mathcal{J}$*-partial matrix* $B(\mathcal{J})$ *extends* the $\mathcal{G}$-partial matrix $A(\mathcal{G})$ if $b_{ij} = a_{ij}, \forall \{i, j\} \in E$

- $\mathcal{G}$ is *positive completable* if every $\mathcal{G}$-partial positive matrix can be extended to a positive definite matrix.

$\mathcal{G}$ is **chordal** if there are no minimal cycles of length ≥ 4.
(i.e. every cycle of length ≥ 4 has a chord)

Theorem 4.5.2.1 *(Grone, Johnson, Sa, Wolkowicz)*
$\mathcal{G}$ is positive completable **iff** *$\mathcal{G}$ is chordal.* ■

Equivalently, positive completability can be expressed as strict feasibility in SDP:

$$\begin{aligned} \operatorname{trace} E_{ij} P &= a_{ij}, \quad \forall \{i,j\} \in E \\ P &\succ 0 \end{aligned}$$

where $E_{ij} = e_i e_j^T + e_j e_k^T$

4.5.3 Approximate Positive Semidefinite Completions

GIVEN:

- $H = H^T \geq 0$ a real, nonnegative (elementwise) **symmetric matrix of weights**
  (with positive diagonal elements $H_{ii} > 0, \ \forall i$)
- $A = A^*$, the **given partial Hermitian matrix**
  i.e. some elements approximately fixed, others free (for notational purposes, assume free elements set to 0 if not specified.)

$||A||_F = \sqrt{\mathrm{trace}\, A^* A}$ *Frobenius norm*, $\circ$ denotes *Hadamard product.*

$$f(P) := ||H \circ (A - P)||_F^2$$

the weighted, best approximate, completion problem

$$(\mathrm{AC}) \qquad \begin{array}{rcc} \mu^* := & \min & f(P) \\ & \text{subject to} & \mathcal{K}P = b \\ & & P \succeq 0, \end{array}$$

where $\mathcal{K} : \mathcal{H}^n \to \mathcal{C}^m$ linear operator

Lagrangian : $\quad L(P,y,\Lambda) = f(P) + \langle y, b - \mathcal{K}P\rangle - \operatorname{trace}\Lambda P$

Dual problem:

$$
\text{(DAC)} \qquad \begin{array}{cc} \max & f(P) + \langle y, b - \mathcal{K}P\rangle - \operatorname{trace}\Lambda P \\ \text{subject to} & \nabla f(P) - \mathcal{K}^* y - \Lambda = 0 \\ & \Lambda \succeq 0. \end{array}
$$

Theorem 4.5.3.1 *The matrix* $\bar{P} \succeq 0$ *and vector-matrix* $\bar{y}, \bar{\Lambda} \succeq 0$ *solve AC and DAC resp.* **iff**

$$
\begin{array}{cc} \mathcal{K}\bar{P} = b & \textit{primal feas.} \\ 2H^{(2)} \circ (\bar{P} - A) - \mathcal{K}^* \bar{y} - \bar{\Lambda} = 0 & \textit{dual feas.} \\ \operatorname{trace}\bar{\Lambda}\bar{P} = 0 & \textit{compl. slack.} \end{array}
$$

■

For simplicity, replace $\mathcal{K}$ with appropriate weights in H.
Use *square* perturbed optimality conditions.
(For $P, \Lambda \succeq 0$: $\operatorname{trace} P\Lambda = 0 \iff P\Lambda = 0$.)

$$\begin{aligned} 2H^{(2)} \circ (P - A) - \Lambda = 0 &\quad \text{dual feasibility} \\ -P + \mu\Lambda^{-1} = 0 &\quad \text{perturbed C.S.} \end{aligned}$$

Linearize second equation and solve for h and l

$$h = \mu\Lambda^{-1} - \mu\Lambda^{-1} l \Lambda^{-1} - P$$

$$l = \frac{1}{\mu}\{-\Lambda(P + h)\Lambda\} + \Lambda$$

Dual Step First: (if many elements of P are free; similar approach for primal step first if many elements are fixed)

We can eliminate the primal step h and solve for the dual step l.

$$
\begin{aligned}
l &= 2H^{(2)} \circ h + (2H^{(2)} \circ (P - A) - \Lambda) \\
&= 2H^{(2)} \circ (\mu\Lambda^{-1} - \mu\Lambda^{-1} \\
&\qquad l\Lambda^{-1} - P) + (2H^{(2)} \circ (P - A) - \Lambda).
\end{aligned}
$$

Equivalently, we get the Newton equation

$$
2H^{(2)} \circ (\mu\Lambda^{-1} l \Lambda^{-1}) + l = 2H^{(2)} \circ (\mu\Lambda^{-1} - A) - \Lambda.
$$

This shows that l, Λ have the same sparsity pattern as H (order is number of nonzeros in $H/2$.)

Each test appears on one line and includes 20 test problems.

dim	toler	Hdens./infty	Apsd	cond(A)	Hpd	min/max	iters
15	10^{-5}	.0751/.02	yes	222.5	no	8/17	10.3
15	10^{-6}	.1/.95	yes	19.6	no	10/23	15.6
15	10^{-6}	.01/.95	yes	21.	no	10/20	13.2
19	10^{-6}	.005/.1	yes	14.8	no	10/18	12.9
21	10^{-6}	.005/.1	yes	20.3	no	8/24	15
38	10^{-6}	1/.99	yes	49.6	yes	14/24	16.
45	10^{-6}	1/.99	yes	46.8	yes	15/22	17.
55	10^{-6}	1/.99	yes	37.2	yes	15/30	17.5
85	10^{-5}	.0219/.02	yes	1374.5	no	16/23	18.9
95	10^{-5}	.0206/.02	yes	2.7	no	8/14	11.1
95	10^{-6}	1/.999	yes	196.	yes	14/18	16.8
145	10^{-6}	.01/.997	yes	658.5	yes	13/17	14.9

data for primal-step-first (20 problems per test):
dimension; tolerance for duality gap;
density of nonzeros in H/ density of infinite values in H;
positive semidefiniteness of A; condition number of A; positive definiteness of H;
min and max number of iterations; average number of iterations.

dim	toler	Hdens./infty	Apsd	cond(A)	Hpd	min/max	iters
60	10^{-6}	.01/.001	yes	79.7	no	15/23	16.8
65	10^{-6}	.015/.001	yes	49.9	yes	18/24	21.3
83	10^{-6}	.007/.001	no	235.1	no	24/29	25.5
85	10^{-5}	.008/.001	yes	94.7	no	11/17	13.1
85	10^{-6}	.0075/.001	no	299.9	no	23/27	25.2
87	10^{-6}	.006/.001	yes	74.2	yes	14/19	16.9
89	10^{-6}	.006/.001	no	179.3	no	23/28	15.2
110	10^{-6}	.007/.001	yes	172.3	yes	15/20	17.8
155	10^{-6}	.01/0	yes	643.9	yes	14/18	15.3
655	10^{-6}	.017/0	yes	1.4	no	14/14	14.
755	10^{-6}	.002/0	yes	1.5	no	15/15	15.

data for dual-step-first (20 problems per test): dimension; tolerance for duality gap; density of nonzeros in H/ density of infinite values in H; positive semidefiniteness of A; condition number of A; positive definiteness of H;

(only one test for: 655,755)

4.5.4 Euclidean Distance Matrix Completion Problem

What are EDMs?

- A **pre-distance matrix** (or dissimilarity matrix) is: an $n \times n$ symmetric matrix $D = (d_{ij})$ with nonnegative elements and zero diagonal
- A (squared) **Euclidean distance matrix** (EDM) is: a pre-distance matrix such that there exists points $x^1, x^2, \ldots, x^n$ in $\Re^r$ such that

$$d_{ij} = \|x^i - x^j\|^2, \quad i, j = 1, 2, \ldots, n.$$

- The smallest value of r is called **the embedding dimension** of D. (r is always $\leq n-1$)

EDM problem:
Given a partial symmetric matrix A with certain elements specified, the Euclidean distance matrix completion problem (EDMCP) consists in finding the unspecified elements of A that make A a EDM.

WHY?

e.g.:

- The shape of an enzyme determines it chemical function. Once the shape is known, then the proper drug can be designed.

- distance geometry on molecules: Atoms are points in space with pairwise distances; find a set of points which yield those distances.

• **approximate EDMCP**, let: A be a pre-distance matrix, H be an $n \times n$ symmetric matrix with nonnegative elements, $\|A\|_F = \sqrt{\operatorname{trace} A^T A}$ denote the *Frobenius norm* of A.

• **objective function**

$$f(D) := \|H \circ (A - D)\|_F^2,$$

where $\circ$ denotes *Hadamard product.*

• **weighted, closest Euclidean distance matrix problem** is

$$(CDM_0) \quad \begin{array}{rcc} \mu^* := & \min & f(D) \\ & \text{subject to} & D \in \mathcal{D}, \end{array}$$

where $\mathcal{D}$ denotes the cone of EDMs.

DISTANCE GEOMETRY

A pre-distance matrix D is a EDM if and only if D is negative semidefinite on

$$M := \left\{ x \in \Re^n : x^T e = 0 \right\},$$

where e is the vector of all ones.

Define V $n \times (n-1)$, full column rank such that $V^T e = 0$. Then

$$J := VV^\dagger = I - \frac{ee^T}{n}$$

is the orthogonal projection onto M, where $V^\dagger$ denotes *Moore-Penrose generalized inverse.*

Define the **centered** and **hollow** subspaces

$$\begin{aligned}\mathcal{S}_C &:= \{B \in \mathcal{S}^n : Be = 0\},\\ \mathcal{S}_H &:= \{D \in \mathcal{S}^n : \operatorname{diag}(D) = 0\}.\end{aligned}$$

Define the two linear operators

$$\begin{aligned}\mathcal{K}(B) &:= \operatorname{diag}(B)\, e^T + e \operatorname{diag}(B)^T - 2B,\\ \mathcal{T}(D) &:= -\tfrac{1}{2} JDJ.\end{aligned}$$

The operator $-2\mathcal{T}$ is an orthogonal projection onto $\mathcal{S}_C$.

Theorem 4.5.4.1

$$\begin{aligned}\mathcal{K}(\mathcal{S}_C) &= \mathcal{S}_H,\\ \mathcal{T}(\mathcal{S}_H) &= \mathcal{S}_C,\end{aligned}$$

and $\mathcal{K}_{|\mathcal{S}_C}$ *and* $\mathcal{T}_{|\mathcal{S}_H}$ *are inverses of each other.* ∎

D (hollow matrix) is EDM $\iff$ $B = \mathcal{T}(D) \succeq 0$

D is EDM $\iff$ $D = \mathcal{K}(B)$, for some B with $Be = 0$ and $B \succeq 0$.

In this case the embedding dimension r is given by the rank of B. Moreover if $B = XX^T$, then the coordinates of the points $x^1, x^2, \ldots, x^n$ that generate D are given by the rows of X and, since $Be = 0$, it follows that the origin coincides with the centroid of these points.

The cone of EDMs, $\mathcal{D}$, has empty interior.
This can cause problems for interior-point methods.

$$V \cdot V : \mathcal{S}_{n-1} \to \mathcal{S}_n$$

$$V \cdot V : \mathcal{P}_{n-1} \to \mathcal{P}_n$$

Define the composite operators

$$\mathcal{K}_V(X) \quad := \quad \mathcal{K}(VXV^T),$$

and

$$\mathcal{T}_V(D) \quad := \quad V^\dagger \mathcal{T}(D)(V^\dagger)^T = -\tfrac{1}{2} V^\dagger D (V^\dagger)^T.$$

LEMMA

$$\mathcal{K}_V(\mathcal{S}_{n-1}) = \mathcal{S}_H,$$
$$\mathcal{T}_V(\mathcal{S}_H) = \mathcal{S}_{n-1},$$

and $\mathcal{K}_V$ and $\mathcal{T}_V$ are inverses of each other on these two spaces. ■

COROLLARY

$$\begin{aligned} \mathcal{K}_V(\mathcal{P}) &= \mathcal{D}, \\ \mathcal{T}_V(\mathcal{E}) &= \mathcal{P}. \end{aligned}$$

■

Summary

(Re)Define the closest EDM problem:

$$\begin{aligned} f(X) &:= \|H \circ (A - \mathcal{K}_V(X))\|_F^2 \\ &= \|H \circ \mathcal{K}_V(B - X)\|_F^2, \end{aligned}$$

where $B = \mathcal{T}_V(A)$.
($\mathcal{K}_V$ and $\mathcal{T}_V$ are both linear operators)

$$\begin{array}{llll} & \mu^* := & \min & f(X) \\ (CDM) & & \textbf{subject to} & \mathcal{A}X = b \\ & & & X \succeq 0. \end{array}$$

The additional constraint using $\mathcal{A} : \mathcal{S}_{n-1} \longrightarrow \Re^m$, could represent some of the fixed elements in the given matrix A.

Primal-Dual Interior-Point Framework:

1. derive a dual program
2. state optimality conditions for log-barrier problem (perturbed primal-dual optimality conditions)
3. find a search direction for solving the perturbed optimality conditions
4. take a step and backtrack to stay strictly feasible (positive definite)
5. Update and go to Step 3 (adaptive update of log-barrier parameter)

Step 1. derive a dual program:

$\Lambda \in \mathcal{S}_{n-1}, \Lambda \succeq 0$ and $y \in R^m$,
Lagrangian is

$$L(X, y, \Lambda) = f(X) + \langle y, b - \mathcal{A}(X) \rangle - \langle \Lambda, X \rangle$$

primal program (CDM) is

$$= \min_X \max_{\substack{y \\ \Lambda \succeq 0}} L(X, y, \Lambda).$$

dual program is:

$$= \max_{\substack{y \\ \Lambda \succeq 0}} \min_X L(X, y, \Lambda),$$

The inner minimization of the convex, in X, Lagrangian is unconstrained so we add the hidden constraint which makes the minimization redundant.

dual program (DCDM)

$$\max_{\substack{\nabla f(X) - \mathcal{A}^* y = \Lambda \\ \Lambda \succeq 0}} f(X) + \langle y, b - \mathcal{A}(X) \rangle - \operatorname{trace} \Lambda X.$$

or

$$\begin{array}{cc} \max & f(X) + \langle y, b - \mathcal{A}(X) \rangle - \operatorname{trace} \Lambda X \\ \text{subject to} & \nabla f(X) - \mathcal{A}^* y - \Lambda = 0 \\ & \Lambda \succeq 0, (X \succeq 0). \end{array}$$

the duality gap,

$f(X) - (f(X) + \langle y, b - \mathcal{A}(X)\rangle - \operatorname{trace} \Lambda X)$,

in the case of primal and dual feasibility, is given by the complementary slackness condition:

$$\operatorname{trace} X(\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(X - B)) - \mathcal{A}^* y) = 0,$$

or equivalently

$$X(\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(X - B)) - \mathcal{A}^* y) = 0,$$

where $H^{(2)} = H \circ H$.

Theorem 4.5.4.2 *Suppose that Slater's condition holds. Then $\bar{X} \succeq 0$, and $\bar{y}$, $\bar{\Lambda} \succeq 0$ solve (CDM) and (DCDM), respectively, if and only if the following three equations hold.*

$$\begin{array}{rr} \mathcal{A}(\bar{X}) = b & \textit{prim. feas.} \\ 2\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(\bar{X} - B)) - \mathcal{A}^*\bar{y} - \bar{\Lambda} = 0 & \textit{dual feas.} \\ \operatorname{trace} \bar{\Lambda}\bar{X} = 0 & \textit{C.S.} \end{array}$$

$\blacksquare$

Lemma 4.5.4.1 *Let H be an $n \times n$ symmetric matrix with nonnegative elements and 0 diagonal such that the graph of H is connected. Then*

$$\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(I)) \succ 0,$$

where $I \in \mathcal{S}_{n-1}$ is the identity matrix. ■

i.e. when $\mathcal{A} = 0$, we have a Slater point for the dual (and primal)

Step 2. state optimality conditions for log-barrier problem (perturbed primal-dual optimality conditions):

The log-barrier problem for (CDM) is

$$\min_{X \succ 0} B_\mu(X) := f(X) - \mu \log \det(X),$$

where $\mu \downarrow 0$.

For each $\mu > 0$ we take one Newton step for solving the stationarity condition

$$\nabla B_\mu(X) = 2\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(X - B)) - \mu X^{-1} = 0.$$

Let

$$C := 2\mathcal{K}_V^*(H^{(2)} \circ \mathcal{K}_V(B)) = 2\mathcal{K}_V^*(H^{(2)} \circ A).$$

Then the stationarity condition is equivalent to

$$\nabla B_\mu(X) = 2\mathcal{K}_V^* \left(H^{(2)} \circ \mathcal{K}_V(X)\right) - C - \mu X^{-1} = 0.$$

equating $\Lambda = \mu X^{-1}$ and multiplying through by X

optimality conditions, $F := \begin{pmatrix} F_d \\ F_c \end{pmatrix} = 0,$

$$\begin{array}{rcll} 2\mathcal{K}_V^* \left(H^{(2)} \circ \mathcal{K}_V(X)\right) - C - \Lambda & = & 0 & \text{dual feas.} \\ \Lambda X - \mu I & = & 0 & \text{pert. C.S.,} \end{array}$$

(an OVERDETERMINED nonlinear system since ΛX not symmetric)

estimate of the barrier parameter

$$\mu = \frac{1}{n-1} \text{ trace } \Lambda X$$

σ_k centering parameter
$\mathcal{F}^0$ set of strictly feasible primal-dual points
F' derivative of F

Algorithm 4.5.4.1 *(p-d i-p framework:)*
Given $(X^0, \Lambda^0) \in \mathcal{F}^0$
for $k = 0, 1, 2 \ldots$
solve *for the search direction*

$$F'(X^k, \Lambda^k) \begin{pmatrix} \delta X^k \\ \delta \Lambda^k \end{pmatrix} = \begin{pmatrix} -F_d \\ -\Lambda^k X^k + \sigma_k \mu_k I \end{pmatrix}$$

where σ_k *centering,* $\mu_k = \frac{\operatorname{trace} X^k \Lambda^k}{(n-1)}$

$$(X^{k+1}, \Lambda^{k+1}) = (X^k, \Lambda^k) + \alpha_k (\delta X^k, \delta \Lambda^k)$$

so that $(X^{k+1}, \Lambda^{k+1}) \succ 0$
end (for).

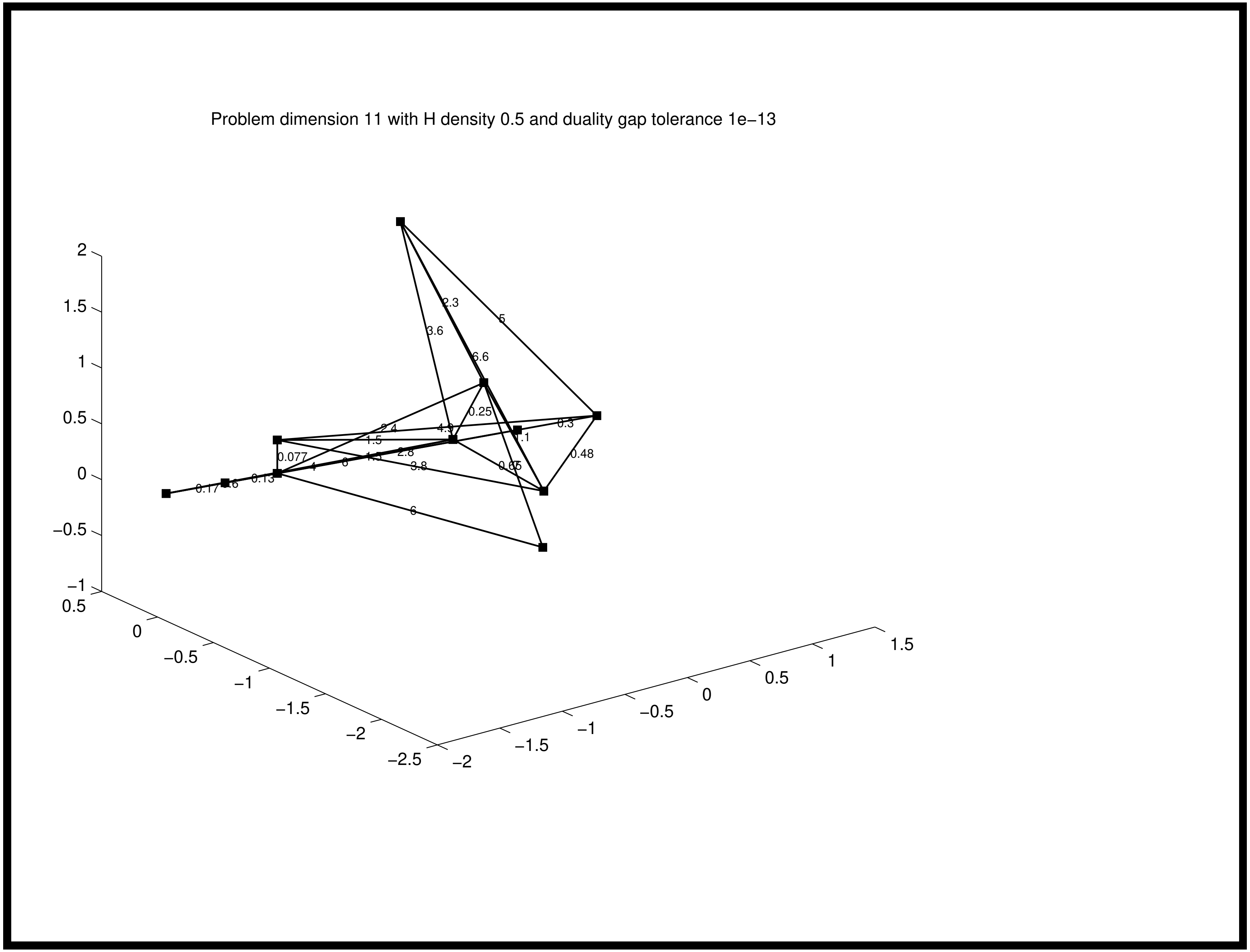
search direction (Gauss-Newton direction) - the Frobenius norm lss of $F's = -F$, i.e.

$$\begin{array}{rcl} 2\mathcal{K}_V^*\left(H^{(2)} \circ \mathcal{K}_V(h)\right) - l & = & -F_d \\ \Lambda h + lX & = & -F_c. \end{array}$$

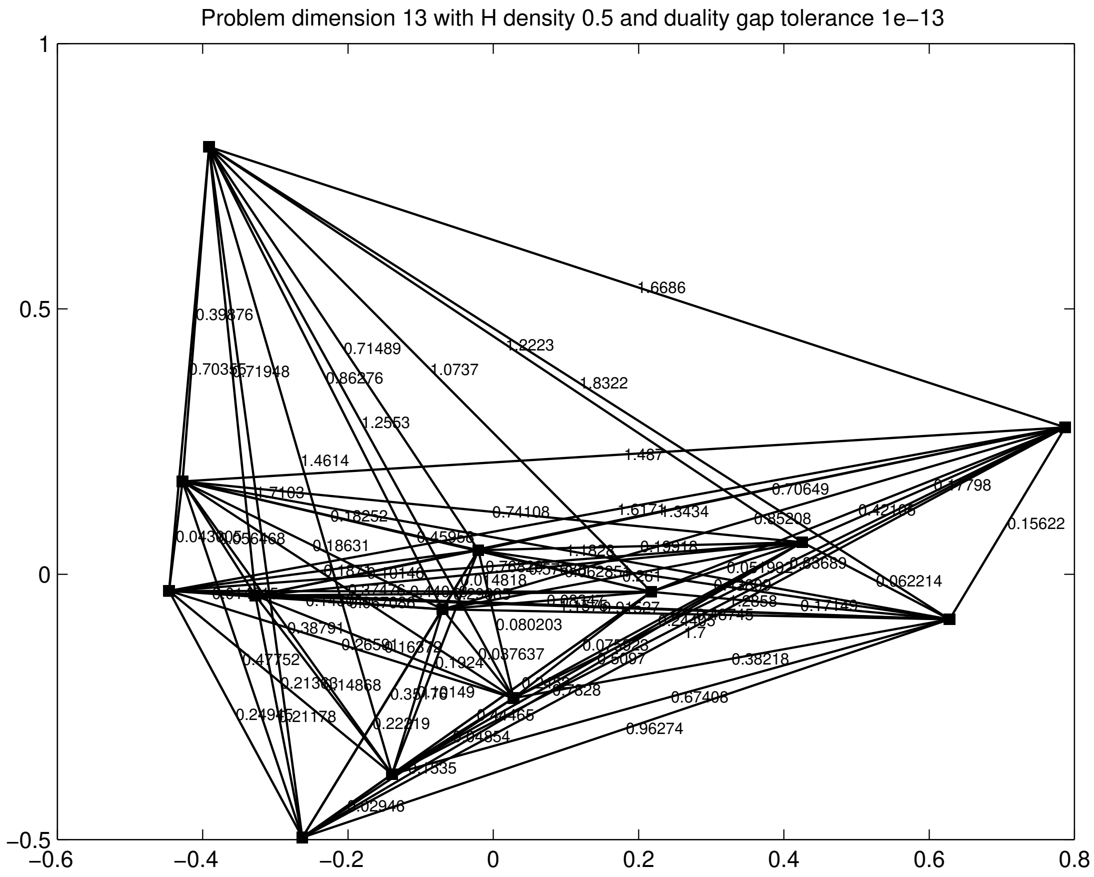
$t(n) = \frac{(n+1)n}{2}$ be dimension of $\mathcal{S}^n$.

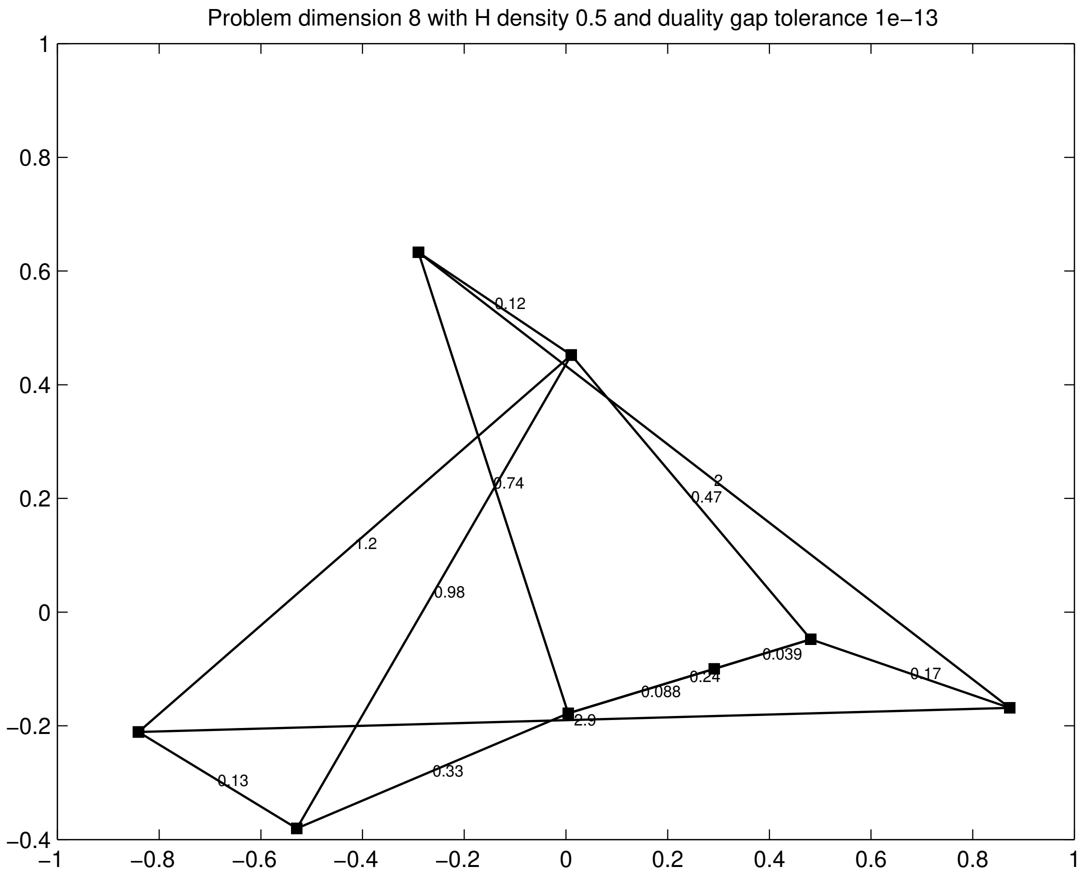
$$F's = \begin{bmatrix} F'_{u1} & F'_{u2} \\ F'_{l1} & F'_{l2} \end{bmatrix} \begin{pmatrix} h \\ l \end{pmatrix} = rhs = \begin{pmatrix} rhs_1 \\ rhs_2 \end{pmatrix}.$$

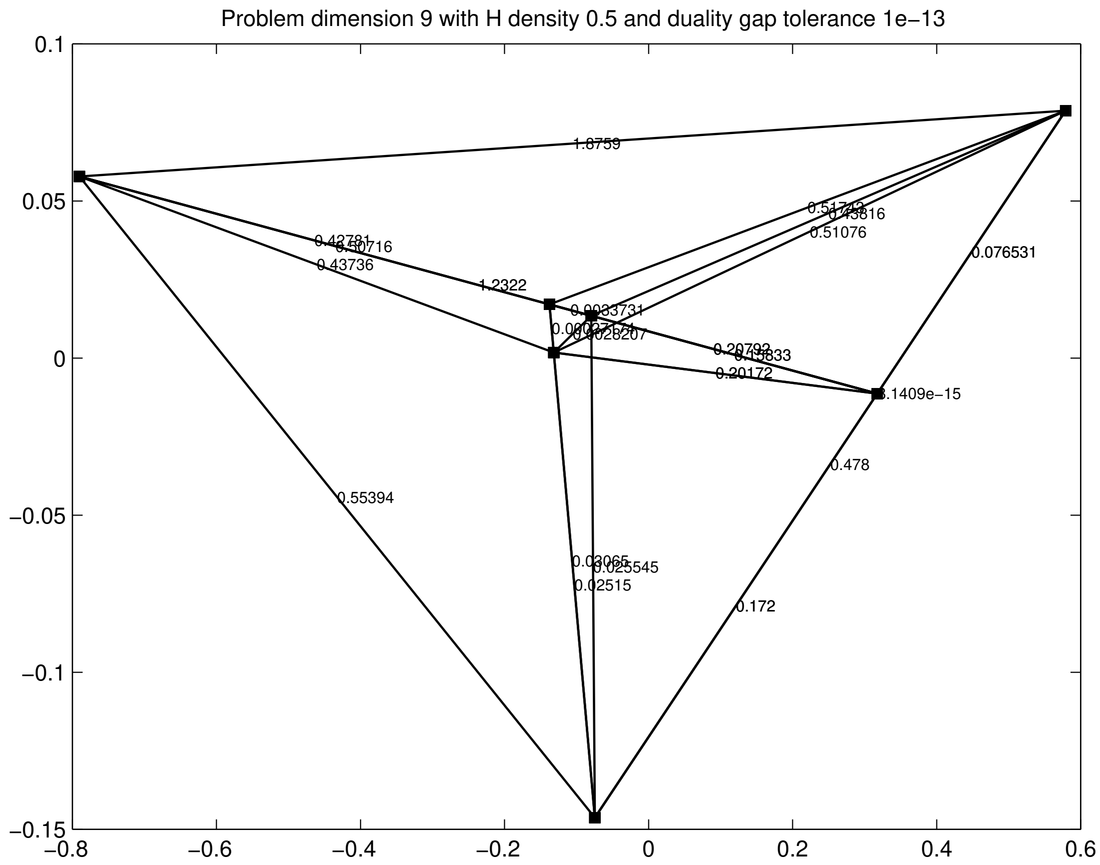
Operator F' maps $\Re^{2(t(n-1))}$ to $\Re^{t(n-1)+(n-1)^2}$.



Problem dimension 11 with H density 0.5 and duality gap tolerance 1e−13


Problem dimension 13 with H density 0.5 and duality gap tolerance 1e−13
1
0.5
0
−0.5
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1.6686
1.2223
1.8322
0.71489
1.0737
0.86276
1.2553
1.4614
1.487
0.70649
0.17798
0.15622
0.42165
0.85208
1.3434
0.74108
0.18252
0.18631
0.45958
0.080203
0.087637
0.38218
0.67406
0.96274
0.062214
0.17149
0.38791
0.47752
0.22219
0.02946
0.1535
0.21178
0.14868


Problem dimension 8 with H density 0.5 and duality gap tolerance 1e−13
0.12
0.74
2
0.47
1.2
0.98
0.039
0.17
0.24
0.088
2.9
0.33
0.13
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1



Problem dimension 9 with H density 0.5 and duality gap tolerance 1e−13
1.8759
0.51743
0.43816
0.51076
0.076531
0.42781
0.30716
0.43736
1.2322
0.0033731
0.0002174
0.0028207
0.20792
0.15633
0.20172
3.1409e−15
0.478
0.55394
0.03065
0.025545
0.02515
0.172

4.5.5 Large Sparse Problems

Instead of projecting and reducing the dimension to get Slater's condition, add a variable and increase the dimension.

Lemma 4.5.5.1 *Let*

$$
\begin{aligned}
\mathcal{D} &:= \{X \in \mathcal{S}_H : v^T e = 0 \quad \Rightarrow \quad v^T X v \leq 0\}, \\
\mathcal{D}_0 &:= \{X \in \mathcal{S}_H : X - \alpha e e^T \preceq 0, \quad \textit{for some } \alpha\}, \\
\mathcal{D}_1 &:= \{X \in \mathcal{S}_H : X - \alpha e e^T \preceq 0, \quad \forall\ \alpha \geq \bar{\alpha},\ \textit{for some } \bar{\alpha}\}.
\end{aligned}
$$

Then

$$
\operatorname{ri}(\mathcal{D}) \subset \mathcal{D}_0 = \mathcal{D}_1 \subset \mathcal{D} \subset \overline{\mathcal{D}_0}. \tag{56}
$$

Proof. Suppose that $\bar{X} \in \mathrm{ri}\,(\mathcal{D})$ (i.e. $v^T e = 0, v \neq 0 \Rightarrow v^T \bar{X} v < 0$) but $\bar{X} \notin \mathcal{D}_0$. Then, for each $\alpha \geq 0$, there exists w_α with $||w_\alpha|| = 1$, such that $w_\alpha \to \bar{w}$, as $\alpha \to \infty$ and

$$w_\alpha^T (\bar{X} - \alpha e e^T) w_\alpha > 0, \quad \forall\ \alpha \geq 0,$$

i.e.

$$w_\alpha^T \bar{X} w_\alpha > \alpha w_\alpha^T e e^T w_\alpha, \quad \forall\ \alpha \geq 0.$$

Since w_α converges and the left-hand-side of the above inequality must be finite, this implies that $e^T \bar{w} = \bar{w}^T \bar{X} \bar{w} = 0$, a contradiction. Therefore, $\mathrm{ri}\,(\mathcal{D}) \subset \mathcal{D}_0$. That $\mathcal{D}_0 = \mathcal{D}_1$ is clear.

Now suppose that $\bar{X} - \alpha e e^T \preceq 0,\ \alpha \geq 0$. Let $v^T e = 0$. Then $0 \geq v^T (\bar{X} - \alpha e e^T) v = v^T \bar{X} v$, i.e. $\mathcal{D}_0 \subset \mathcal{D}$. The final inclusion comes from the first and the fact that $\mathcal{D}$ is closed. ■

let: $E = ee^T$; $f(P) := \|H \circ (A - P)\|_F^2$;
$\mathcal{K}$ lin. operator with constraint $\operatorname{diag}(P) = 0$.
primal problem is:

$$
\text{(CDM)} \qquad
\begin{array}{rcc}
\mu^* := & \min & f(P) \\
& \text{subject to} & \mathcal{K}P = b \\
& & \alpha E - P \succeq 0
\end{array}
$$

and **dual problem (DCDM)** is

$$
\begin{array}{rc}
\mu^* = \nu^* := \max & f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda(\alpha E - P) \\
\text{s.t.} & \nabla_P f(P) - \mathcal{K}^* y + \Lambda = 0 \\
& -\operatorname{trace} \Lambda E = 0 \\
& \Lambda \succeq 0.
\end{array}
$$

(Slater's holds for primal but fails for dual.)

(perturbed) Optimality Conditions are:

$$
\begin{array}{rcl}
\operatorname{diag}(P) = 0 & & \text{primal feas.} \\
2H^{(2)} \circ (P - A) - \operatorname{Diag}(y) + \Lambda = 0, & & \\
-\operatorname{trace} \Lambda E = 0 & & \text{dual feas.} \\
-(\alpha E - P) + \mu \Lambda^{-1} = 0, & & \text{pert. C.S.}
\end{array}
$$

$$
\begin{array}{rl}
h & \text{denotes the step for} \quad P \\
w & \text{denotes the step for} \quad \alpha \\
l & \text{denotes the step for} \quad \Lambda \\
s & \text{denotes the step for} \quad y.
\end{array}
$$

maintain

$$\operatorname{diag}(h) = \operatorname{diag}(P) = 0$$

linearization of complementary slackness

$$-(\alpha + w)E + (P + h) + \mu\Lambda^{-1} - \mu\Lambda^{-1} l \Lambda^{-1} \quad = \quad 0,$$

solve for h

$$h \quad = \quad -\mu\Lambda^{-1} + \mu\Lambda^{-1} l \Lambda^{-1} - P + (\alpha + w)E.$$

(or solve for l)

linearization dual feasibility

$$\begin{aligned} 2H^{(2)} \circ h - \operatorname{Diag}(s) + l \quad &= \quad -(2H^{(2)} \circ (P - A) \\ &\qquad - \operatorname{Diag}(y) + \Lambda) \\ -\operatorname{trace} lE \quad &= \quad \operatorname{trace} \Lambda E \end{aligned}$$

substitute for h, s
Newton equation is

$$\begin{aligned} 2H^{(2)} \circ \left(wE + \mu\Lambda^{-1} l \Lambda^{-1}\right) - \operatorname{Diag} \operatorname{diag}(l) + l \\ = 2H^{(2)} \circ \left\{\mu\Lambda^{-1} + A - \alpha E\right\} + \operatorname{Diag}(y) - \Lambda \\ \operatorname{diag}\ \left(\mu\Lambda^{-1} l \Lambda^{-1}\right) + we \\ = \operatorname{diag}\ \left(\mu\Lambda^{-1}\right) - \alpha e \\ \operatorname{trace}(lE) = -\operatorname{trace}(\Lambda E). \end{aligned}$$

square system, order $1 + nnz$ where nnz are the number of nonzeros in the upper triangular part of H, $(\operatorname{diag}(H) = 0)$.

F denotes $(nnz+n)\times 2$ matrix
row p contains indices of the p-th nonzero, upper triangular, element of $H+I$ ordered by columns,

$$\begin{aligned}&\{(F_{p1}, F_{p2})_{p=1,\ldots nnz+n}\}\\ &\qquad = \{ij : H_{ij} \neq 0, i \leq j, \text{ ordered by columns}\}.\end{aligned}$$

δ_{ij} is *Kronecker delta function*
$\delta_{(ij)(kl)}$ is 1 if $(ij)=(kl)$, 0 otherwise.
$E_{ij} = \left(e_i e_j^T + e_j^T e_i\right)/\sqrt{2}$, ij unit matrix in $\mathcal{S}^n$, where
$E_{ij} = \left(e_i e_j^T + e_j^T e_i\right)/2$ if $i=j$.
(orthonormal basis of $\mathcal{S}^n$)

operator equation:

$$
\begin{aligned}
& k \neq l, i \neq j \text{ LHS} = \\
& = \operatorname{trace} E_{kl} \left\{ 2H^{(2)} \circ \left(\mu \Lambda^{-1} E_{ij} \Lambda^{-1} \right) - \operatorname{Diag} \operatorname{diag} (E_{ij}) \right. \\
& \qquad\qquad\qquad \left. + E_{ij} \right\} \\
& = \mu \operatorname{trace} (e_k e_l^T + e_l e_k^T) \left(H^{(2)} \circ \Lambda^{-1} (e_i e_j^T + e_j e_i^T) \Lambda^{-1} \right) \\
& \qquad\qquad\qquad + \delta_{(ij)(kl)} \\
& = \mu \left[2e_l^T \left(H^{(2)} \circ \Lambda^{-1}_{:,i} \Lambda^{-1}_{j:} \right) e_k + 2e_k^T \left(H^{(2)} \circ \Lambda^{-1}_{:,i} \Lambda^{-1}_{j:} \right) e_l \right] \\
& \qquad\qquad + \delta_{(ij)(kl)};
\end{aligned}
$$

$k \neq l, i \neq j$ LHS $=$

$$= 2\mu H^{(2)}_{kl} \left(\Lambda^{-1}_{li} \Lambda^{-1}_{jk} + \Lambda^{-1}_{ki} \Lambda^{-1}_{jl} \right) + \delta_{(ij)(kl)};$$

$k \neq l, i = j$ LHS $=$

$$= \operatorname{trace} E_{kl} \left\{ 2\mu H^{(2)} \circ \left[\Lambda^{-1} E_{jj} \Lambda^{-1} \right] - \operatorname{Diag} \operatorname{diag} \left(E_{jj} \right) + E_{jj} \right\}$$

$$= 2\sqrt{2}\mu \operatorname{trace} e_k e_l^T \left(H^{(2)} \circ \Lambda^{-1} e_j e_j^T \Lambda^{-1} \right)$$

$$= 2\sqrt{2}\mu H^{(2)}_{kl} \left(\Lambda^{-1}_{lj} \Lambda^{-1}_{jk} \right);$$

$k = l, i \neq j$ LHS $=$

$$= \sqrt{2}\mu \Lambda^{-1}_{ki} \Lambda^{-1}_{jk}, \quad k = 1, \ldots n;$$

$k = l, i = j$ LHS $=$

$$= \mu \Lambda^{-1}_{ki} \Lambda^{-1}_{ik}, \quad k = 1, \ldots n.$$

last column of LHS, matrix $l = 0$ and $w = 1$:

$$\begin{aligned} w = 1, k \neq l \text{ LHS} &= \operatorname{trace}\left(E_{kl}(2H^{(2)} \circ E)\right); \\ w = 1, k = l \text{ LHS} &= 1. \end{aligned}$$

last row of LHS:

$$\begin{aligned} i \neq j \text{ LHS} &= \operatorname{trace}\left(E_{ij}E\right)) = \sqrt{2}; \\ i = j \text{ LHS} &= 1. \end{aligned}$$

Newton system is:

$$\mathrm{sMat}\,[L(\mathrm{svec}\,(l))] = \mathrm{sMat}\,[\mathrm{svec}\,(RHS)]\,,$$

$\mathrm{svec}\,(S)$ vector formed from nonzero elements of columns of upper triangular part, where strict upper triangular part is multiplied by $\sqrt{2}$. (trace $XY = \mathrm{svec}\,(X)^T \mathrm{svec}\,(Y)$, i.e. isometry) sMat is inverse

Solve for $\mathrm{svec}\,(l)$:

$$L(\mathrm{svec}\,(l)) = \mathrm{svec}\,(RHS).$$

$$L_{pq} =$$

$$\begin{cases} 2\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_1}} + \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_2}}\right) \\ \qquad \text{if } p \neq q,\ k \neq l,\ i \neq j; \\ 2\sqrt{2}\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_2}} \Lambda^{-1}_{F_{q_2},F_{p_1}}\right) \\ \qquad \text{if } p \neq q,\ k \neq l,\ i = j; \\ 2\sqrt{2}\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_2}} \Lambda^{-1}_{F_{q_2},F_{p_1}}\right) \\ \qquad \text{if } p = q,\ k \neq l,\ i = j; \\ 2\mu H^{(2)}_{F_{p_2},F_{p_1}} \left(\Lambda^{-1}_{F_{p_2},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_1}} + \Lambda^{-1}_{F_{p_1},F_{q_1}} \Lambda^{-1}_{F_{q_2},F_{p_2}}\right) + 1 \\ \qquad \text{if } p = q, k \neq l,\ i \neq j; \end{cases}$$

$$
\begin{cases}
\sqrt{2}\mu\Lambda^{-1}_{F_{p_1},F_{q_1}}\Lambda^{-1}_{F_{q_2},F_{p_1}} \\
\qquad\qquad\qquad\qquad \text{if } k = l,\ i \neq j; \\
\mu\Lambda^{-1}_{F_{p_1},F_{q_1}}\Lambda^{-1}_{F_{q_1},F_{p_1}} & \text{if } k = l,\ i = j \\
2\sqrt{(}2)H^{(2)}_{F_{p_2},F_{p_1}} & \text{if } w = 1,\ k \neq l \\
1 & \text{if } w = 1,\ k = l.
\end{cases}
$$

The p-th row calculated using Hadamard product of pairs of columns of Λ^{-1},

$$
\Lambda^{-1}_{F_{p_2},F_{:,1}} \circ \Lambda^{-1}_{F_{p_1},F_{:,2}}.
$$

complete vectorization

$p = kl, k \leq l$, and last row, component of the right-hand-side of the system is
$RHS_p =$

$$\begin{cases} \sqrt{2}\left(2H_p^{(2)} \circ \{\mu\Lambda_p^{-1} + A_p - \alpha\} - \Lambda_p\right), & \text{if } k \neq l \\ \mu\Lambda_{kk}^{-1} - \alpha & \text{if } k = l \\ -\text{trace}\,(\Lambda E) & \text{last row} \end{cases}$$

4.5.6 On a New Characterization of EDMs

The above application for EDM uses

$$D = \lambda E - P,\ P \succeq 0,\ \operatorname{diag} P = \lambda e,\ \lambda \geq 0;$$

this raises the **question**:

> **which** *Euclidean distance matrices D can be expressed as $D = \lambda(E - C)$ for some nonnegative scalar λ and some correlation matrix C, where E is the matrix of all ones.*

We **show (and characterize)** that the cones

$$\operatorname{cone}(E - \mathcal{E}_n) \subsetneq \overline{\operatorname{cone}(E - \mathcal{E}_n)} = \mathcal{D}_n,$$

where $\mathcal{E}_n$ is the elliptope (set of correlation matrices) and $\mathcal{D}_n$ is the (closed convex) cone of Euclidean distance matrices.

The characterization is given using the Gale transform of the points generating D. We also show that given points $p^1, p^2, \ldots, p^n \in \Re^r$, for any scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$\sum_{j=1}^{n} \lambda_j \, p^j = 0, \qquad \sum_{j=1}^{n} \lambda_j = 0,$$

we have

$$\sum_{j=1}^{n} \lambda_j \, \|p^i - p^j\|^2 = \alpha \text{ for all } i = 1, \ldots, n,$$

for some scalar α independent of i.

Recall:

- $n \times n$ matrix $D = (d_{ij})$ is *Euclidean distance matrix (EDM)* if

$$\exists \quad p^1, p^2, \ldots, p^n \in \Re^r, \text{ such that } \|p^i - p^j\|^2 = d_{ij}, \quad \forall i, j$$

- dimension of smallest Euclidean space containing $p^1, p^2, \ldots, p^n$ is the *embedding dimension* of D.
- D (hollow) EDM $\iff$ D negative semidefinite on $e^\perp$
- the set of $n \times n$ EDM matrices is a closed convex cone, $\mathcal{D}_n$

Let $\mathcal{E}_n$ denote set of $n \times n$ *correlation matrices*, i.e., set of positive semidefinite symmetric matrices whose diagonal is equal to e.

It is also well known [8] that $\mathcal{D}_n$ is the tangent cone of $\mathcal{E}_n$ at E, the matrix of all ones, i.e.,

$$\mathcal{D}_n = \overline{\text{cone }(E - \mathcal{E}_n)} = \overline{\{\lambda(E - C) : \lambda \geq 0, C \in \mathcal{E}_n\}}, \qquad (57)$$

where $\bar{\cdot}$ denotes closure.

- Let V be $n \times (n-1)$ such that

$$V^T e = 0\,, \quad V^T V = I_{n-1}\,. \tag{58}$$

- orthogonal projection on $M = e^{\perp}$, denoted by J, is $J := VV^T = I - ee^T/n$.
- D hollow is EDM $\iff$ $B := -\frac{1}{2}\, JDJ \succeq 0$
- embedding dimension of D is equal, r, rank of B.
- the points $p^1, p^2, \ldots, p^n$ that generate D are given by the rows of the $n \times r$ matrix P where $B := PP^T$.
- since $Be = 0$, centroid of the points p^i, $i = 1, \ldots, n$ coincides with the origin

Let $p^1, p^2, \ldots, p^n$ be points in $\Re^r$ whose centroid coincides with the origin, and are not contained in a proper hyperplane. Then

$$P := \begin{bmatrix} {p^1}^T \\ {p^2}^T \\ \vdots \\ {p^n}^T \end{bmatrix}$$

is of rank r. Let $B = PP^T$. Then it easily follows that the EDM matrix D generated by p^i, $i = 1, \ldots, n$ is given by

$$D = \operatorname{diag} Be^T + e\,(\operatorname{diag} B)^T - 2B. \tag{59}$$

- Let $\bar{r} = n - 1 - r$ and

$$Z \text{ be } n \times \bar{r}, \text{ full column rank with, } P^T Z = 0, e^T Z = 0$$

- ${z^i}^T$ denotes the i-th row of Z. i.e.,

$$Z^T := \left[z^1 \mid z^2 \mid \ldots \mid z^n \right].$$

- z^i is called Gale transform of p^i; Z is called a *Gale matrix* corresponding to D.
- The columns of Z represent the *affine dependence relations* among the points $p^1, p^2, \ldots, p^n$, i.e., among the rows of P.

Main Results

Theorem 4.5.6.1 *Let D be a Euclidean distance matrix and let Z be a Gale matrix corresponding to D. Then, the columns of DZ are proportional to e.*

Equivalently

Theorem 4.5.6.2 *Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be coefficients, not all zero, of the affine dependence equation of the points $p^1, p^2, \ldots, p^n$, in $\Re^r$, i.e.,*

$$\sum_{j=1}^{n} \lambda_j p^j = 0, \quad \sum_{j=1}^{n} \lambda_j = 0.$$

Then

$$\sum_{j=1}^{n} \lambda_j \; \|p^i - p^j\|^2 = \alpha \textit{ for all } i = 1, \ldots, n,$$

for some scalar α independent of i.

Theorem 4.5.6.3 *Let D be a Euclidean distance matrix and let Z be a Gale matrix corresponding to D. Then* **the following are equivalent***:*

1.

$$D = \lambda(E - C), \tag{60}$$

for some nonnegative scalar λ and some correlation matrix C;

2.

$$DZ = 0. \tag{61}$$

Proof of the Main Results

Lemma 4.5.6.1 *Let D be a Euclidean distance matrix and let $B = -\frac{1}{2}JDJ$. Then:*

1.

$$-\frac{1}{2}V^TDV = V^TBV;$$

2.

$$\mathcal{N}(V^TDV) = \mathcal{N}(P^TV).$$

Proof. The first part follows directly from (59) and the definition of V. This yields the second part since $B = PP^T$ and $\mathcal{N}(V^TBV) = \mathcal{N}(V^TPP^TV) = \mathcal{N}(P^TV)$. ■

Lemma 4.5.6.2 *Let D be a Euclidean distance matrix and let U be the matrix whose columns form an orthonormal basis of the null space of $V^T DV$. Then VU is a Gale matrix corresponding to D.*

Proof. It follows from Lemma 4.5.6.1 that $P^T VU = V^T DVU = 0$ and from the definition of V in (58) that $e^T VU = 0$. Hence, the columns of VU form an orthonormal basis for the null space of $\begin{bmatrix} P^T \\ e^T \end{bmatrix}$. ■

Proof of Theorem 4.5.6.1. Let Z be a Gale matrix corresponding to D. Then It follows from Lemma 4.5.6.2 that $VU = ZQ$ for some nonsingular $\bar{r} \times \bar{r}$ matrix Q. Thus $V^T DZ = V^T DVUQ^{-1} = 0$. Hence, the columns of DZ are proportional to e. ∎

Proof of Theorem 4.5.6.3. $D = \lambda(E - C)$ for some nonnegative scalar λ and some correlation matrix C if and only if $E - \frac{1}{\lambda}D$ is positive semidefinite. Let $Q = [\frac{e}{\sqrt{n}}\ V]$. Then, $E - D/\lambda \succeq 0$ if and only if $Q^T\ (E - D/\lambda)\ Q \succeq 0$. But

$$Q^T\ (E - D/\lambda)\ Q = \begin{bmatrix} n - \frac{1}{\lambda\, n}\, e^T De & -\frac{1}{\lambda\,\sqrt{n}}\, e^T DV \\ -\frac{1}{\lambda\,\sqrt{n}}\, V^T De & -\frac{1}{\lambda}\, V^T DV \end{bmatrix}.$$

Recall that $V^T(-D)V \succeq 0$ follows from Lemma 4.5.6.1. Let W and U be the matrices whose columns form an orthonormal basis for the range space and null space of $V^T(-D)V$, respectively.

Hence, $V^T(-D)V = W\Lambda W^T$, where Λ is the diagonal matrix of the positive eigenvalues of $V^T(-D)V$. Let $Q' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & W & U \end{bmatrix}$. Then, $E - D/\lambda$ is positive semidefinite if and only if

$$\begin{aligned} R &= {Q'}^T Q^T (E - D/\lambda) Q Q' \\ &= \begin{bmatrix} n - \frac{1}{\lambda n} e^T D e & -\frac{1}{\lambda\sqrt{n}} e^T D V W & -\frac{1}{\lambda\sqrt{n}} e^T D V U \\ -\frac{1}{\lambda\sqrt{n}} W^T V^T D e & \frac{1}{\lambda}\Lambda & 0 \\ -\frac{1}{\lambda\sqrt{n}} U^T V^T D e & 0 & 0 \end{bmatrix} \succeq 0. \end{aligned} \tag{62}$$

Now for sufficiently large λ the submatrix

$$\begin{bmatrix} n - \frac{1}{\lambda\, n}\, e^T De & -\frac{1}{\lambda\, \sqrt{n}}\, e^T DVW \\ -\frac{1}{\lambda\, \sqrt{n}}\, W^T V^T De & \frac{1}{\lambda}\, \Lambda \end{bmatrix}$$

is positive definite.

Thus $E - D/\lambda$ is positive semidefinite if and only if $e^T DVU = e^T DZ = 0$. But it follows from Theorem 4.5.6.1 that $e^T DZ = 0$ if and only if $DZ = 0$ and the result follows. ■

Given the two Euclidean distance matrices:

$$D_1 = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 0 & 1 \\ 4 & 1 & 0 \end{bmatrix}, \qquad D_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

Gale matrices Z_1, Z_2 corresponding to D_1, D_2, resp. are:

$$Z_1^T = [1 \ -2 \ 1], \quad Z_2^T = [1 \ 0 \ -1],$$

Now $D_1 Z_1 = 2e$ and $D_2 Z_2 = 0$. Then: $D_2 = E - C_2$, where

$$C_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \succeq 0.$$

However, there exists no $\lambda \geq 0$ such that $D_1 = \lambda(E - C_1)$ for some correlation matrix C_1.