Semidefinite Programming and Matrix Completions

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

Waterloo, Ontario N2L 3G1, Canada
hwolkowi@orion.math.uwaterloo.ca
http://orion.math.uwaterloo.ca/~hwolkowi

OUTLINE

1. Outline of SDP, history

2. SDP and combinatorics

3. SDP and matrix completions

4. Euclidean distance matrix completions

5. graph embedding

SDP: BACKGROUND and HISTORY

Semidefinite Programming
looks just like
Linear Programming

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

$\preceq$ denotes the Löwner partial order
$A \preceq B$ if $B - A \succeq 0$

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

$$\mathcal{A} : \mathcal{S}^n \to \Re^m$$

$$(\mathcal{A}X)_i = \text{trace}\,(A_i X), \text{ for given } A_i \in \mathcal{S}^n$$

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

replaces

$\Re^n_+$ - nonnegative orthant

SDP is arguably hottest area in Opt. during last 10 years.

ISMP00 in Atlanta - at least 13 minisymposia on SDP. Many conferences, papers, books, theses have been published recently, see e.g.

Handbook of Semidefinite Programming: Theory, Algorithms, and Applications
Kluwer Academic Publs. 2000.

Reasons for high interest:

many interesting **applications** in many diverse areas

SDPs are convex problems that can be **solved very efficiently**, both in theory (with a polynomial worst-case complexity) and in practice
(Nesterov and Nemirovski 1988, $\log \det A$ is a self-concordant barrier for the positive semidefinite cone, apply p-d i-p methods)

APPLICATIONS

Engineering (LMIs): Lyapunov stability of ODEs, 1890's; Yakubovitch 1960's; Boyd et al 1980's, Ricatti equations, ...

matrix completion problems 1980s, more below.

combinatorial optimization 1980s; Lovász *theta function* (*Lovász number of a graph*); strong approximation results for the max-cut problem by Goemans-Williamson; boolean quadratic approximations

OTHER APPLICATIONS

e.g. Complementarity Problems; min-max eigenvalue problems; Financial Applications; Nonlinear Programming; Statistics.

MORE SIMILARITIES WITH LP

DUALITY

payoff function, player Y to player X (Lagrangian)

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

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.

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

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

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

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

The *hidden constraint* $C - \mathcal{A}^* y \preceq 0$

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

The dual is obtained from the optimal strategy of the competing player, Y.
The *hidden constraint* $C - \mathcal{A}^* y \preceq 0$
yields the dual

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

for the primal

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

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

$$\mathcal{A}^* y - Z = C \qquad \text{dual feasibility}$$

$$AX = b \qquad \text{primal feasibility}$$

$$ZX = 0 \qquad \text{complementary slackness}$$

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

Forms the basis for:

primal simplex method
dual simplex method
interior point methods

Direct SDP Relaxation

Max-Cut Problem

$$\max \ \tfrac{1}{2}\sum_{i<j} w_{ij}(1 - x_i x_j), \qquad x \in \{\pm 1\}^n.$$

Equate $x_i = 1$ with $i \in \mathcal{I}$ and -1 otherwise.
Let

$$q(x) := x^t Q x,$$

where Q is an $n \times n$ symmetric matrix. An equivalent problem is the homogeneous (± 1)-*quadratic programming problem*

$$\mu^* := \max \ q(x), \qquad x \in \{\pm 1\}^n.$$

Replace $x \in \{\pm 1\}^n$ with constraints $x_i^2 = 1$.

Note that for

$$X = xx^t,$$

$$X \succeq 0, \ \mathrm{diag}\,(X) = e, \ q(x) = \mathrm{trace}\, XQ.$$

Relax the rank-1 condition on X to get SDP.

The SDP Relaxation:

$$\begin{array}{ll} \max & \text{trace}\, XQ \\ \text{subject to} & \text{diag}\,(X) = e \\ & X \succeq 0 \end{array}$$

If the weights are nonnegative:
Goemans-williamson randomized approximation algorithm with performance guarantee of 0.87856, i.e.

$$\text{opt MC} \geq \text{rand. value} \geq 0.87856(\text{opt SDP})$$

How does SDP arise from general quadratic approximations?

Let

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

$$\textbf{(QQP)} \quad \begin{array}{rll} q^* = & \min & q_0(y) \\ & \text{s.t.} & q_i(y) \leq 0 \\ & & i = 1, \ldots m \end{array}$$

Lagrangian:

$$L(y, x) \;\; = \;\; q_0(y) + \sum_{i=1}^m x_i q_i(y)$$

or equivalently

$$\begin{array}{rcl} L(y, x) & = & \frac{1}{2} y^t (Q_0 + \sum_{i=1}^m x_i Q_i) y \\ & & + y^t (b_0 + \sum_{i=1}^m x_i b_i) \\ & & + (c_0 + \sum_{i=1}^m x_i c_i) \end{array}$$

Weak duality follows from "hidden constraint"

$$d^* = \max_{x \geq 0} \min_y L(y, x) \leq q^* = \min_y \max_{x \geq 0} L(y, x).$$

Now homogenize; multiply linear term by new variable y_0

$$y_0 y^t (b_0 + \sum_{i=1}^m x_i b_i), \quad y_0^2 = 1.$$

$$
\begin{aligned}
d^* &= \\
&= \max_{x\geq 0} \min{}_y \quad L(y,x) \\
&= \max_{x\geq 0} \min_{y_0^2=1} \quad \tfrac{1}{2} y^t (Q_0 + \textstyle\sum_{i=1}^m x_i Q_i) y + t y_0^2 \\
&\qquad + y_0 y^t (b_0 + \textstyle\sum_{i=1}^m x_i b_i) \\
&\qquad + (c_0 + \textstyle\sum_{i=1}^m x_i c_i) - t \\
&= \max_{x\geq 0, t} \min_{y} \quad \tfrac{1}{2} y^t (Q_0 + \textstyle\sum_{i=1}^m x_i Q_i) y + t y_0^2 \\
&\qquad + y_0 y^t (b_0 + \textstyle\sum_{i=1}^m x_i b_i) \\
&\qquad + (c_0 + \textstyle\sum_{i=1}^m x_i c_i) - t
\end{aligned}
$$

The hidden semidefinite constraint yields the semidefinite program.

We get

$$B := \begin{pmatrix} 0 & b_0^t \\ b_0 & Q_0 \end{pmatrix},$$

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

$$A \begin{pmatrix} t \\ x \end{pmatrix} := - \begin{bmatrix} t & \sum_{i=1}^m x_i b_i^t \\ \sum_{i=1}^m x_i b_i & \sum_{i=1}^m x_i Q_i \end{bmatrix}$$

and the SDP constraint

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

NOTE There is NO hidden constraint on the Q_i if all q_i are convex. There is no need for SDP in the convex case. Better algorithms exist, e.g. proximal methods, using quadratic cones, ...

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

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

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

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

Positive Definite Completions of Partial Hermitian Matrices

$\mathcal{G}(V, E)$ finite undirected graph

$A(\mathcal{G})$ 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. (every cycle of length ≥ 4 has a chord)

THEOREM (Grone, Johnson, Sa, Wolkowicz)
$\mathcal{G}$ is positive completable iff $\mathcal{G}$ is chordal. ∎

equivalently - strict feasibility for SDP:

$$\begin{array}{rl} \text{trace}\, E_{ij} P = a_{ij}, & \forall \{i,j\} \in E \\ P \succ 0 & \end{array}$$

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

Approximate Positive Semidefinite Completions

given:
$H = H^t \geq 0$ real, nonnegative (elementwise) **symmetric matrix of weights** with positive diagonal elements $H_{ii} > 0, \ \forall i$;

and $A = A^*$ 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.)

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

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

weighted, best approximate, completion problem

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

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

Lagrangian:

$$L(P, y, \Lambda) = f(P) + \langle y, b - KP \rangle - \operatorname{trace} \Lambda P$$

Dual problem:

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

THEOREM The matrix $\bar{P} \succeq 0$ and vector-matrix $\bar{y}, \bar{\Lambda} \succeq 0$ solve AC and DAC if and only if

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

■

Solve the perturbed nonlinear system,

$$\text{trace}\,\bar{\Lambda}\bar{P} = \mu I,$$

using Newton's method i.e. take one Newton step then update/decrease the barrier parameter μ

Two algorithms: dual-step-first; primal-step-first

Exploit sparsity, i.e. depends on whether there are *many* elements fixed or *many* elements free.

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)

Euclidean Distance Matrix Completion Problem

What are EDMs?

A **pre-distance matrix** (or dissimilarity matrix):
• an $n \times n$ symmetric matrix $D = (d_{ij})$ with nonnegative elements and zero diagonal

---

A **Euclidean distance matrix** (EDM):
• 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.

For the approximate EDMCP, let A be a pre-distance matrix, H be an $n \times n$ symmetric matrix with nonnegative elements, and let $\|A\|_F = \sqrt{\text{trace}\, A^t A}$ denote the *Frobenius norm* of A. Consider the objective function

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

where $\circ$ denotes *Hadamard product*. The **weighted, closest Euclidean distance matrix problem** is

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

where $\mathcal{E}$ 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

$$\mathcal{K}(B) := \operatorname{diag}(B)\, e^t + e \operatorname{diag}(B)^t - 2B,$$

and

$$\mathcal{T}(D) := -\tfrac{1}{2} JDJ.$$

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

THEOREM The linear operators satisfy

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

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

■

A hollow matrix D is EDM
if and only if
$B = \mathcal{T}(D) \succeq 0$ (positive semidefinite)

D is EDM
if and only if
$D = \mathcal{K}(B)$, 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, $\mathcal{E}$, has empty interior. This can cause problems for interior-point methods.

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

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

Define the composite operators

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

and

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

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

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

■

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)

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

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:

Step 1 derive a dual program

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

Step 3 find a search direction for solving the perturbed optimality conditions

Step 4 take a step and backtrack to stay strictly feasible (positive definite)

Step 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)

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

or

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

the duality gap,
$f(X) - (f(X) + \langle y, b - \mathcal{A}(X)\rangle - \text{trace}\,\Lambda X)$,
in the case of primal and dual feasibility, is given by the complementary slackness condition:

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

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 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.

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

■

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

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

(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 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).

For the EDM:

search direction (Gauss-Newton direction) is the Frobenius norm lss of

$$F's = -F,$$

i.e.

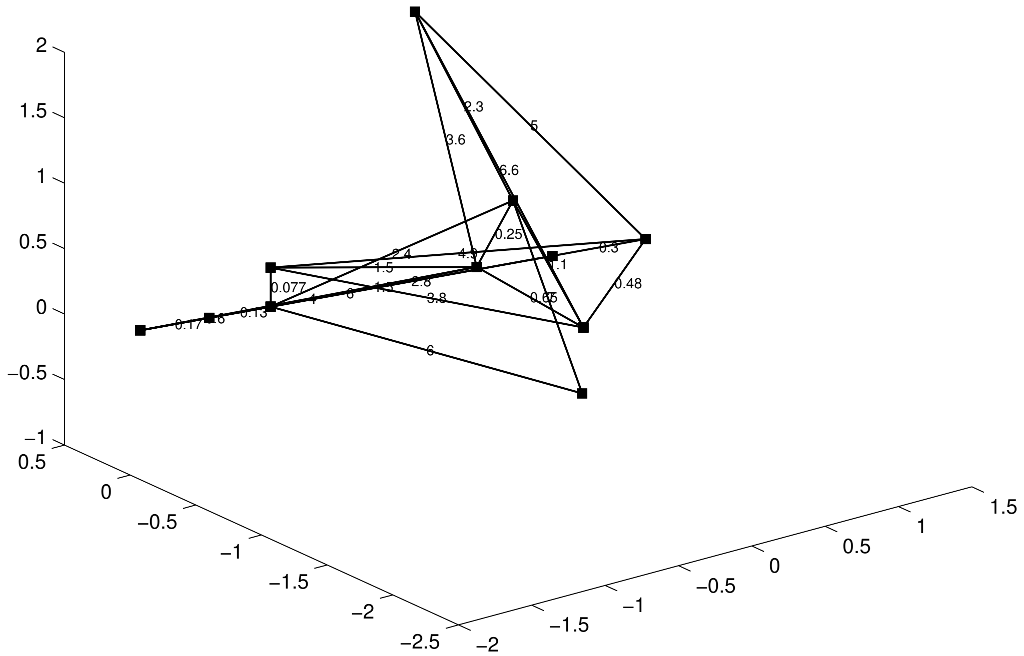
$$\begin{array}{rcl} 2\mathcal{K}_V^*\left(H^{(2)} \circ \mathcal{K}_V(h)\right) - l & = & -F_d \\ \Lambda h + lX & = & -F_c. \end{array}$$

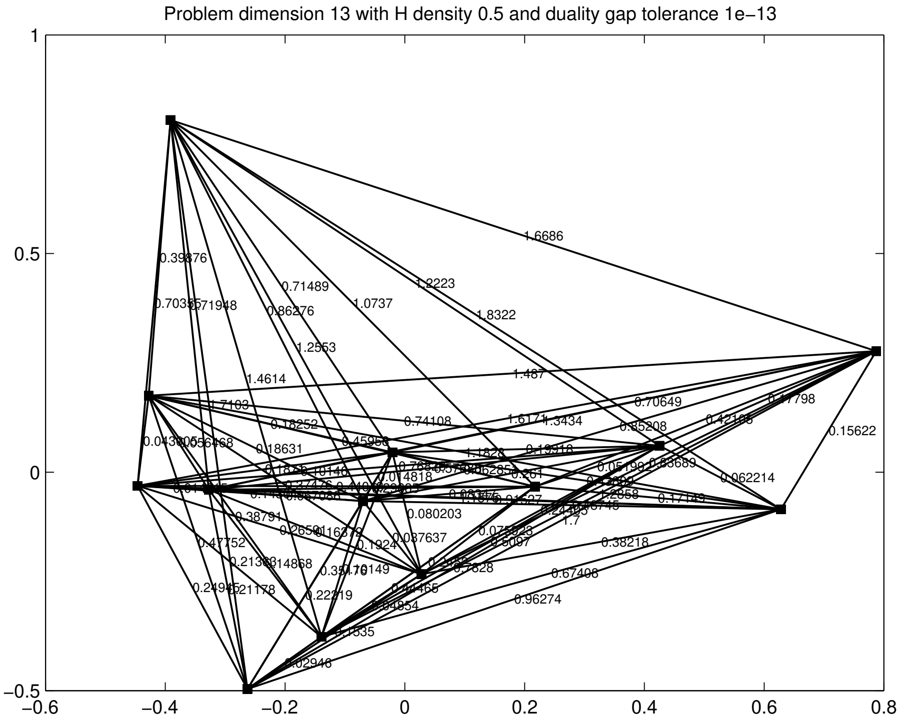
$t(n) = \frac{(n+1)n}{2}$ be dimension of $\mathcal{S}^n$.

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

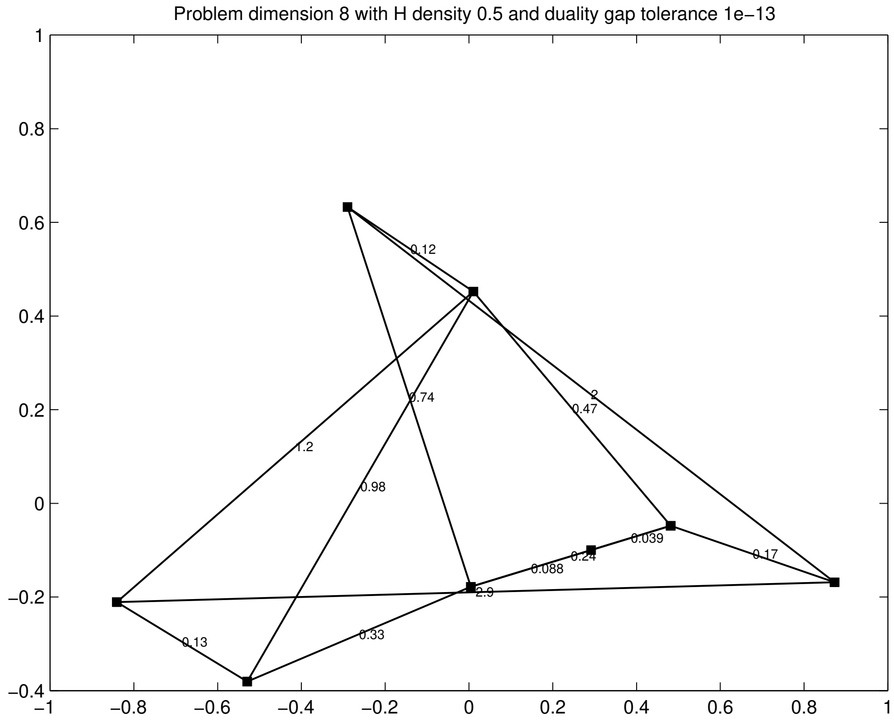
Operator F' maps $\Re^{2(t(n-1))}$ to $\Re^{t(n-1)+(n-1)^2}$.

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

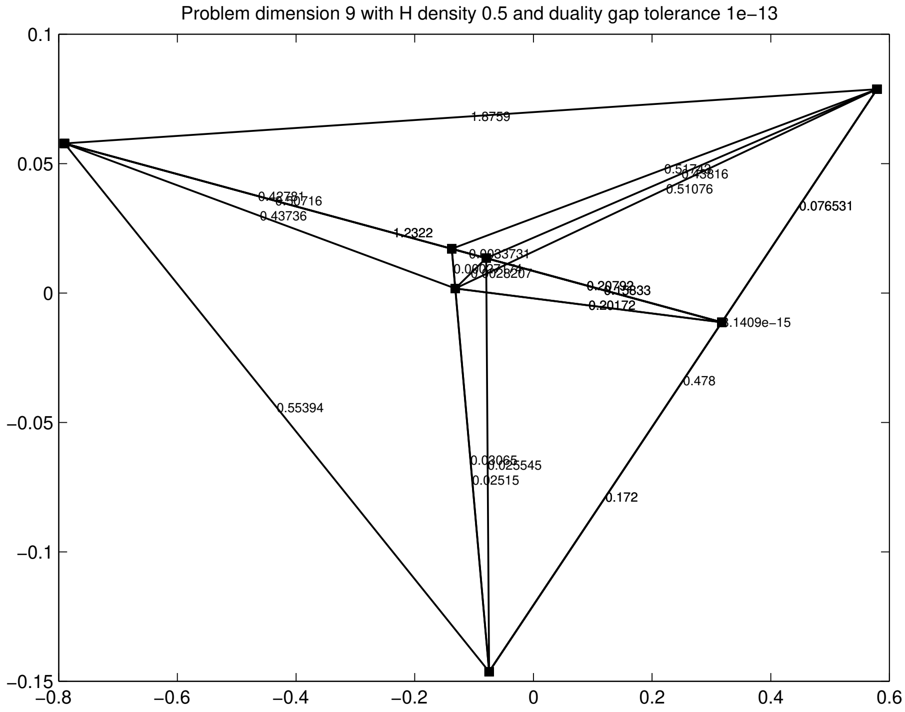




Problem dimension 13 with H density 0.5 and duality gap tolerance 1e−13



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
0.24
0.17
0.088
2.9
0.13
0.33
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
0.51076
0.076531
1.2322
0.43736
0.20172
0.478
0.55394
0.02515
0.172
0.1
0.05
0
−0.05
−0.1
−0.15
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6

REMARKS:

An initial stepsize of 1.5 was taken; tests were also done with 2 and improved convergence was noted.

A centering parameter σ changes adaptively depending on the length of the stepsize, i.e. it decreases if the stepsize is large and increases to 1 if the stepsize is small.

The barrier parameter μ also changes adaptively; there is an aggressive decrease $\mu/4n$ if the stepsize is large; this decrease is smaller when small stepsizes are chosen.

Backtracking was initially done with the value .95 to be conservative, as is usually the case with SDP. However, tests were done with the aggressive .997 as is done in LP. This did not result in faster convergence. However, robustness was maintained.

The upper part of F', i.e. $[F'_{u1} \ F'_{u2}]$, need only be evaluated once at the beginning; while the lower (simpler) part of the operator needs to be evaluated at each iteration. Both evaluations take very little time compared to the solution of the least squares problem.

No predictor-corrector approach was used. Rather, centering was used to speed up convergence. However, the step lengths were consistently larger than 1 which suggests that no corrector step is needed, i.e. centering keeps the steps very close to the central path.

Though the initial starting points were feasible, dual feasibility was lost after the first iteration, i.e. unlike Newton's method, the Gauss-Newton direction does not maintain feasibility. However, complementarity and dual feasibility converged together to 0, i.e. at a similar rate. This might account for the strong convergence obtained with this direction.

The search direction was very robust, i.e. machine accuracy was attainable whenever asked for. This was the case even when the aggressive initial stepsize and reduction of μ were used.

Problem dimension was restricted to $n \leq 42$. This was because of the high cost of the least squares solution in MATLAB. (Done using qls, Version 1.6, from Pontus Matstoms, Linkoping University.) However, if one solves the least squares problem using, for example, a Lanczos approach, then one need never form the matrices which represent F' and thus can exploit sparsity and solve much larger problems. These tests are ongoing.

We have assumed, without loss of generality, that the graph of the matrix H was connected, i.e. if this was not the case then we can solve two smaller problems independently and more simply. In fact, one should always solve the smaller problems independently. Note that Slater's constraint qualification is always satisfied for the primal and so there is no duality gap. However, if Slater's constraint qualification fails for the dual (which is equivalent to the graph being disconnected), then the set of optimal solutions for the primal is unbounded and the algorithm has great difficulty converging. Here is an example of a class of problems where strong duality holds but interior-point methods have difficulty.

A Heuristic for Low Dimensional Realizations of Weighted Graphs

with Abdo Alfakih

$G = (V, E, \omega)$ edge-weighted connected graph,
n nodes, m edges
with points $p^1, p^2, \dots, p^n \in \Re^r$

$$P := \begin{bmatrix} p^1 & p^2 & \dots & p^n \end{bmatrix}^t, \quad n \times r \text{ matrix}$$

is called a *realization* of G in $\Re^r$ if

$$\|p^i - p^j\| = \omega_{ij} \quad \text{for all } (v_i, v_j) \in E,$$

We assume:

G is r-*realizable*, i.e. there exists a realization of G in $\Re^r$.

And, wlog that $P^T e = 0$, ($e \in \Re^n$ vector of ones) rank$P = r$

The

Euclidean distance matrix (EDM) associated with a realization P

is the $n \times n$ matrix

$$D_P = (d_{ij}) = \|p^i - p^j\|^2.$$

(D_P is symmetric with zero diagonal and $d_{ij} = \omega_{ij}^2$ for all $(v_i, v_j) \in E$)

again, with V $n \times (n-1)$ matrix

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

we get three equivalent representations for the realization:

THEOREM Given a realization P of a weighted graph $G = (V, E, \omega)$ such that $P^T e = 0$, then P can be equivalently represented by the matrices B or X where $B := PP^T$ and $X := V^T B V$.

■

we use the $(n-1) \times (n-1)$ positive semidefinite matrix X as the representation of a given realization of a graph G since:

it allows a much simpler characterization of Ω, the set of all realizations of graph G in all dimensional Euclidean spaces;

the set Ω has the desired property of being full dimensional, whenever the given graph G has a realization in $\Re^{(n-1)}$.

$H = (h_{ij})$ adjacency matrix of graph G

$$\mathcal{A}(X) := H \circ \mathcal{K}_V(X), \quad \text{linear operator}$$

($\circ$ Hadamard product)

$$\mathcal{K}_V(X) \;:=\; \text{diag}(VXV^T)\, e^T + e\, \text{diag}(VXV^T)^T - 2VXV^T$$

$\bar{m}$ number of missing edges of G
$\bar{m} = n(n-1)/2 - m$

$E^{ij} \in \mathcal{S}^n$ matrix with ones in (i,j)th and (j,i)th entries

if $h_{ij} = 0$ and $i < j$, let

$$M^k := -\frac{1}{2} V^T E^{ij} V, \;\; k = 1, \ldots, \bar{m}$$

Then $\{ M^k : k = 1, \ldots, \bar{m} \}$ is linearly independent set; forms a basis for $\mathcal{N}(\mathcal{A})$

$$\mathcal{N}(\mathcal{A}) = \{B \in \mathcal{S}_{n-1} : B = \sum_{k=1}^{\bar{m}} y_k M^k, \text{some } y \in \Re^{\bar{m}}\}.$$

Recall that $X^* = V^T B^* V = V^T P^* P^{*T} V$ and that $X^* \in \mathcal{S}_{n-1}$ is positive semidefinite with rank r.

Let Ω_r be the subset of $\Re^{\bar{m}}$ defined by

$$\begin{aligned}\Omega_r \;:=\; \{y \in \Re^{\bar{m}} : \;& X(y) = X^* + \textstyle\sum_{k=1}^{\bar{m}} y_k M^k \succeq 0, \\ & \text{and rank } X(y) = r\}.\end{aligned}$$

THEOREM Given a framework (G, X^*) in $\Re^r$, then $\{X(y) : y \in \Omega_r\}$ is the set of all realizations of G in $\Re^r$. ■

By relaxing the rank condition on $X(y)$, we get the following closed, convex, and generally non-polyhedral set.

$$\Omega := \{y \in \Re^{\bar{m}} \ : \ X(y) = X^* + \sum_{k=1}^{\bar{m}} y_k M^k \ \succeq 0\}.$$

It readily follows that $\{X(y) : y \in \Omega\}$ is the set of all realizations of G in all finite-dimensional Euclidean spaces. Since $X^* \succeq 0$, the origin is always contained in Ω. Furthermore, Ω is bounded if and only if the given graph G is connected.

Heuristic

Step 1

Given $G = (V, E, \omega)$, let

$$D = (d_{ij}) := \left\{ \begin{array}{cl} \omega_{ij}^2 & \text{if } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{array} \right\}.$$

Then solve the following SDP:

$$\begin{array}{rll} \mu^* := & \min & \| H \circ (D - \mathcal{K}_V(X) \|_F^2 \\ & & X \succeq 0. \end{array}$$

$\mu^* = 0$ iff G is realizable. Assume G is realizable and let X^* be the optimal solution of the above SDP.

Also let

$$\Omega := \{y \in \Re^{\bar{m}} \ : \ X(y) = X^* + \sum_{k=1}^{\bar{m}} y_k M^k \succeq 0\}.$$

Then

$$\{X(y) : y \in \Omega\}$$

is the set of all realizations of G in all finite-dimensional Euclidean spaces. Furthermore, Ω is a nonpolyhedral convex closed set and it is bounded iff G is connected.

Note that the origin belongs to Ω since $X^* \succeq 0$.

Now it is easy to show that if $\hat{y}$ is an extreme point of Ω, then rank $X(\hat{y}) \leq r^*$. r^* is Barvinok bound. Thus by moving to an extreme point of Ω we are guaranteed a realization of G in a Euclidean space of dimension $\leq r^*$. Obviously our goal is not to move to just any extreme point of Ω but rather to an extreme point y such that rank of $X(y)$ is as small as possible. That is done in step 2.

Step 2

Now chose a vector c randomly and solve the following SDP.

$$\begin{array}{cc} \text{max} & c^T y \\ \text{subject to} & y \in \Omega. \end{array}$$

It can be proved that if $\hat{y}$ is a vertex of Ω. i.e., the normal cone $N(\Omega, \hat{y})$ is full dimensional, then rank of $X(\hat{y})$ is *small*. Then since c is random, we have more chances of obtaining such a point $\hat{y}$ as a solution of the above SDP.

Following is a derivation of optimality conditions for the SDP pair:

$$\begin{array}{cl} \max & c^T y \\ \text{subject to} & X^* + \sum_{k=1}^{\bar{m}} y_k M^k \succeq 0 \end{array}$$

and

$$\begin{array}{cl} \min & \langle X^* Y \rangle \\ \text{subject to} & \left\langle Y M^k \right\rangle = -c_k \quad \text{for } k = 1, \ldots, \bar{m}, \\ & Y \succeq 0. \end{array}$$

the perturbed optimality conditions are :

$$\begin{array}{rcll} Z & = & X^* + \sum_{k=1}^{\bar{m}} y_k \; M^k & \\ c_k & = & - \left\langle Y M^k \right\rangle & \text{for } k = 1, \ldots, \bar{m} \\ ZY & = & \mu I \;, \quad Z \succeq 0, \; Y \succeq 0 & \end{array}$$

Linearizing yields:

$$\begin{array}{rcll} \delta Z - \sum_{k=1}^{\bar{m}} \delta y_k \; M^k & = & -F_d & \\ \left\langle \delta Y M^k \right\rangle & = & -F_p^k & \text{for } k = 1, \ldots, \bar{m} \\ Z \; \delta Y + \delta Z \; Y & = & -ZY + \mu I & \end{array}$$

where

$$\begin{array}{rcll} F_d & = & -X^* + Z - \sum_{k=1}^{\bar{m}} y_k \; M^k & \\ F_p^k & = & c_k + \left\langle Y M^k \right\rangle & \text{for } k = 1, \ldots, \bar{m} \end{array}$$

Solving for the search direction yields

$$\begin{array}{rcl}\delta y & = & A^{-1} b \\ \delta Z & = & -F_d + \sum_{k=1}^{\bar{m}} \delta y_k \, M^k \\ \delta Y & = & -Y + \mu Z^{-1} + Z^{-1} F_d Y - \sum_{k=1}^{\bar{m}} \delta y_k Z^{-1} M^k Y \end{array}$$

where

$$\begin{array}{rcl} A = (a_{ij}) & = & \left(\left\langle Z^{-1} M^j Y M^i \right\rangle \right) \\ b = (b_i) & = & \left(F_p^i - \left\langle Y M^i \right\rangle + \right. \\ & & \left. \mu \left\langle Z^{-1} M^i \right\rangle + \left\langle Z^{-1} F_d Y M^i \right\rangle \right) \end{array}$$