Semidefinite Programming
and
Matrix Completions

Tutte Colloquium
Department of
Combinatorics and Optimization
University of Waterloo
Friday, June 16, 2000

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. Short intro. to SDP
2. SDP and positive definite matrix completions
3. Euclidean distance matrix completions
4. New characterization for EDM;
   solving large sparse problems

SDP BACKGROUND and NOTATION

Semidefinite Programming
looks just like
Linear Programming

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

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

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

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

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

for the primal

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

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

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

$AX = b$ primal feasibility

$ZX = 0$ complementary slackness

$ZX = \mu I$ perturbed

Forms the basis for:

primal simplex method
dual simplex method
interior point methods

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:

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

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

Approximate Positive Semidefinite Completions

(with Charlie Johnson
and Brenda Kroschel)

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:

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

Dual problem:

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

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

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

■

Discard linear operator K and replace with appropriate weights in H.

Use *square* perturbed optimality conditions.

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

Linearization of 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)

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

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

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, i.e. this is a square system; order is number of nonzeros/2 in H.

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

(with Abdo Alfakih)

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 (squared) **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 predistance 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

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

Define the two linear operators

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

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

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

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.

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

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

Define the composite operators

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

and

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

LEMMA

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

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)

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

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.

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

■

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

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

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

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

$$F's = -F,$$

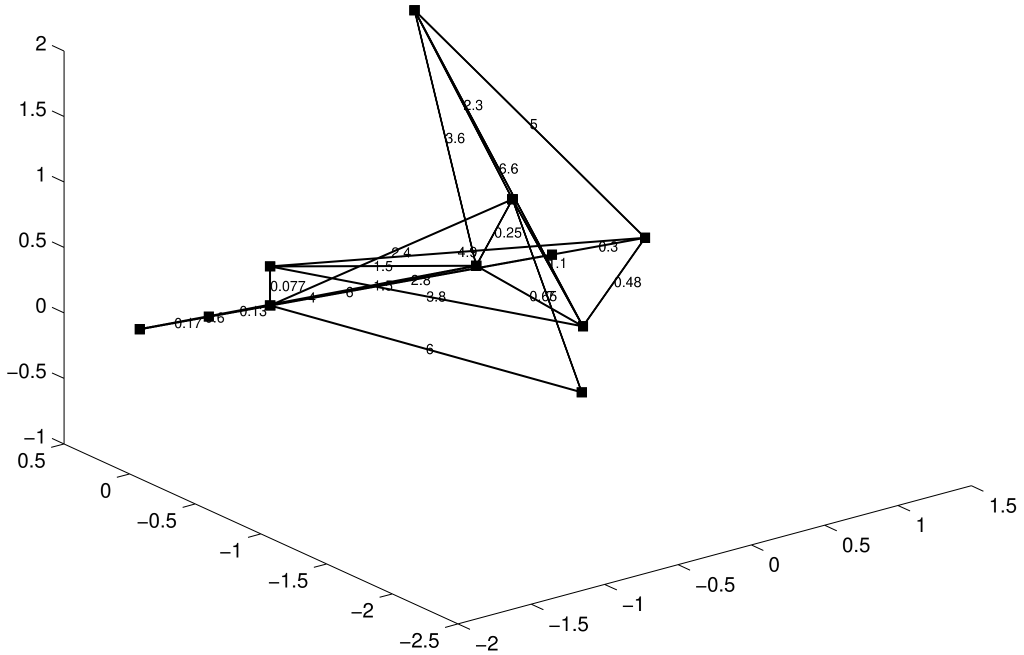
i.e.

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

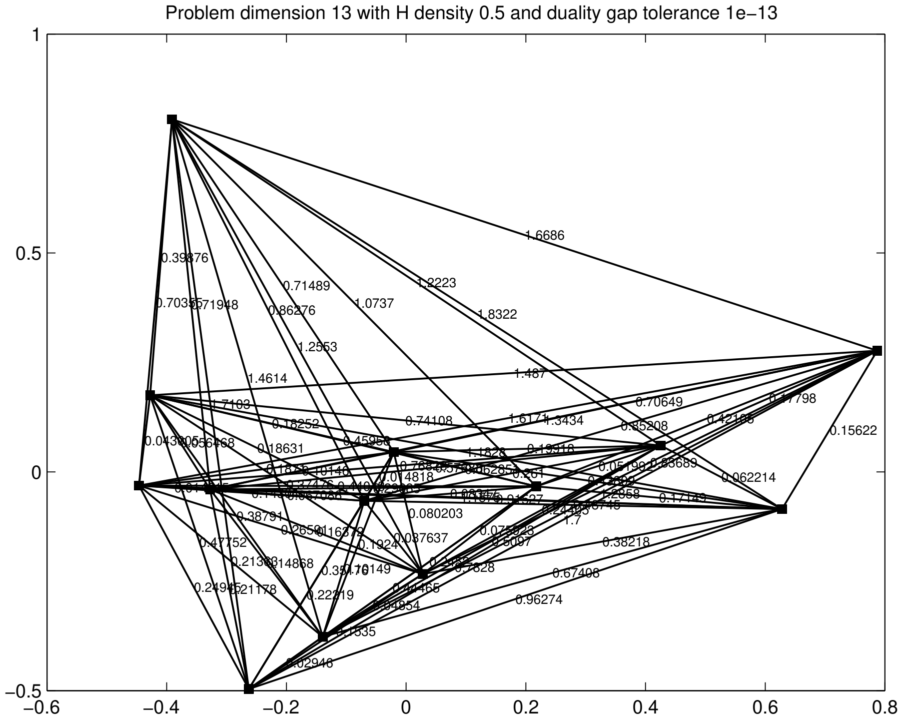
$t(n) = \frac{(n+1)n}{2}$ be dimension of $\mathcal{S}^n$.

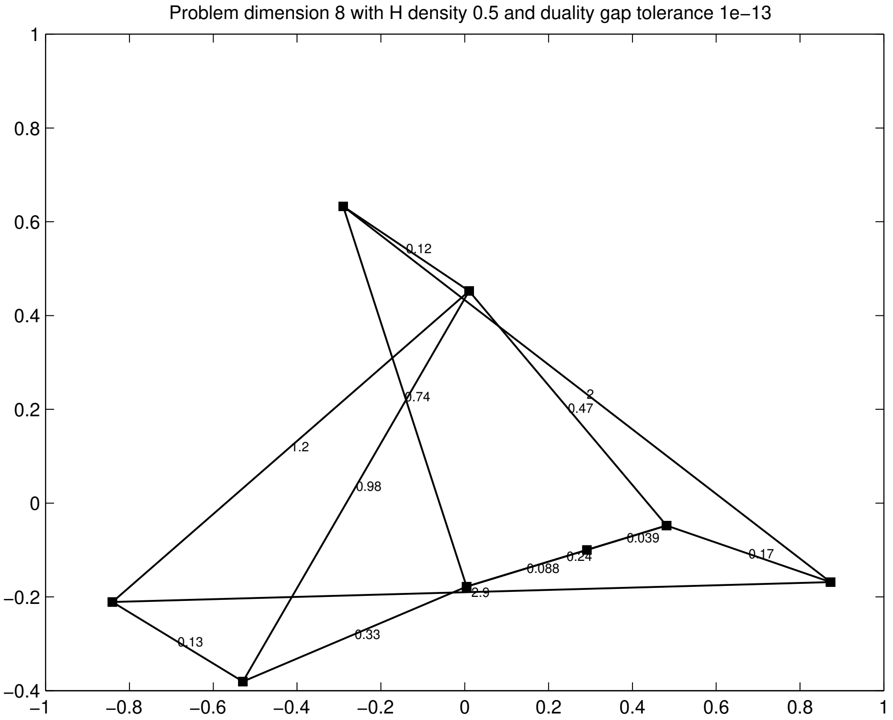
$$F's = \left[\begin{array}{cc} F'_{u1} & F'_{u2} \\ F'_{l1} & F'_{l2} \end{array}\right] \begin{pmatrix} h \\ l \end{pmatrix} = rhs = \begin{pmatrix} rhs_1 \\ rhs_2 \end{pmatrix}.$$

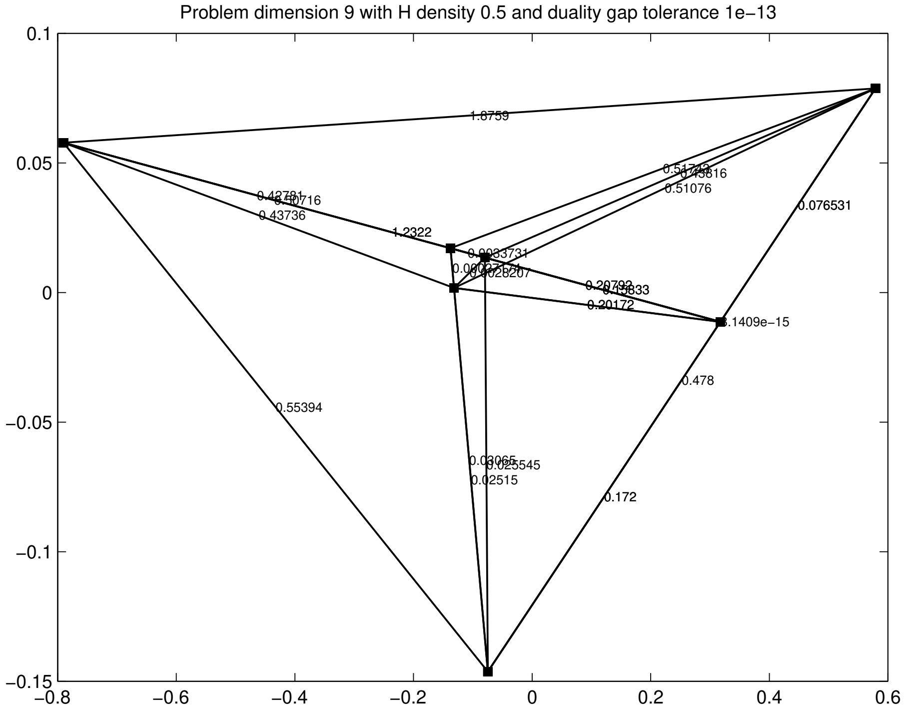
Operator F' maps $\Re^{2(t(n-1))}$ to $\Re^{t(n-1)+(n-1)^2}$.



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
0.39876
0.71489
1.2223
1.0737
1.8322
0.86276
1.2553
1.4614
1.487
0.70649
0.47798
0.15622
0.42166
0.74108
1.3434
0.85208
0.18252
0.45958
0.18631
0.080203
0.087637
0.062214
0.17149
0.38791
0.47752
0.38218
0.67408
0.96274
0.24945
0.221178
0.22219
0.1535
0.02946
0.1924
0.5097
1.7
0.261


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


Problem dimension 9 with H density 0.5 and duality gap tolerance 1e−13
1.8759
0.42781
0.50716
0.43736
1.2322
0.51742
0.45816
0.51076
0.076531
0.0033731
0.0028207
0.20792
0.15833
0.20172
3.1409e−15
0.478
0.55394
0.03065
0.025545
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

Larger Models

often result in improved numerical complexity. Instead of projecting and reducing the dimension to get Slater's condition, add a variable and increase the dimension.

Lemma 1 *Let*

$$\begin{array}{rcll} \mathcal{F} & := & \{X \in \mathcal{S}^n : v^T e = 0 & \Rightarrow \quad v^T X v \leq 0\}, \\ \mathcal{F}_0 & := & \{X \in \mathcal{S}^n : X - \alpha e e^t \preceq 0, & \text{for some } \alpha \geq 0\}, \\ \mathcal{F}_1 & := & \{X \in \mathcal{S}^n : X - \alpha e e^t \preceq 0, & \forall\ \alpha \geq \bar{\alpha}, \\ & & \qquad \text{for some } \bar{\alpha} \geq 0\}. & \end{array}$$

Then

$$\operatorname{ri}(\mathcal{F}) \subset \mathcal{F}_0 = \mathcal{F}_1 \subset \mathcal{F} \subset \overline{\mathcal{F}_0}. \qquad (1)$$

Proof. Suppose that $\bar{X} \in \mathrm{ri}\,(\mathcal{F})$ (i.e. $v^T e = 0, v \neq 0 \Rightarrow v^T \bar{X} v < 0$) but $\bar{X} \notin \mathcal{F}_0$. 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, $\mathrm{ri}\,(\mathcal{F}) \subset \mathcal{F}_0$. That $\mathcal{F}_0 = \mathcal{F}_1$ 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. $\mathcal{F}_0 \subset \mathcal{F}$. The final inclusion comes from the first and the fact that $\mathcal{F}$ is closed. ∎

Corollary 1 *Let*

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

Then

$$\mathcal{E} = \mathcal{E}_0 = \mathcal{E}_1. \tag{2}$$

Proof. The proof is similar to that in the above Lemma 1. We only include the details about the closure.

Suppose that $0 \neq X_k \in \mathcal{E}_0$, i.e. $\operatorname{diag}(X_k) = 0, X_k \preceq \alpha_k E$, for some α_k; and, suppose that $X_k \to \bar{X}$. Since X_k is hollow it has exactly one positive eigenvalue and this must be smaller than α_k. However, since X_k converges to $\bar{X}$, we conclude that $\bar{X} \leq \lambda_{\max}(\bar{X})E$, where $\lambda_{\max}(\bar{X})$ is the largest eigenvalue of $\bar{X}$. ■

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

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

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

(perturbed) Optimality Conditions are:

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

h denotes the step for P
w denotes the step for α
l denotes the step for Λ
s denotes the step for y.

maintain

$$\text{diag}\,(h) = \text{diag}\,(P) = 0$$

linearization of complementary slackness

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

solve for h

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

(or solve for l)

linearization dual feasibility

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

substitute for h, s

Newton equation is

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

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

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,

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

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:

$$
\begin{aligned}
&k \neq l, i \neq j \text{ LHS} = \\
&= \operatorname{trace} E_{kl} \left\{ 2H^{(2)} \circ \left(\mu \Lambda^{-1} E_{ij} \Lambda^{-1} \right) - \operatorname{Diag}\operatorname{diag}\left(E_{ij}\right) \right. \\
&\qquad\qquad\qquad \left. + E_{ij} \right\} \\
&= \mu \operatorname{trace}\left(e_k e_l^t + e_l e_k^t\right) \left(H^{(2)} \circ \Lambda^{-1} (e_i e_j^t + e_j e_i^t) \Lambda^{-1} \right) \\
&\qquad\qquad\qquad + \delta_{(ij)(kl)} \\
&= \mu \left[2e_l^t \left(H^{(2)} \circ \Lambda^{-1}_{:,i} \Lambda^{-1}_{j:} \right) e_k + 2e_k^t \left(H^{(2)} \circ \Lambda^{-1}_{:,i} \Lambda^{-1}_{j:} \right) e_l \right] \\
&\qquad\qquad + \delta_{(ij)(kl)}; \\
&k \neq l, i \neq j \text{ LHS} = \\
&= 2\mu H^{(2)}_{kl} \left(\Lambda^{-1}_{li} \Lambda^{-1}_{jk} + \Lambda^{-1}_{ki} \Lambda^{-1}_{jl} \right) + \delta_{(ij)(kl)}; \\
&k \neq l, i = j \text{ LHS} = \\
&= \operatorname{trace} E_{kl} \left\{ 2\mu H^{(2)} \circ \left[\Lambda^{-1} E_{jj} \Lambda^{-1} \right] - \operatorname{Diag}\operatorname{diag}\left(E_{jj}\right) \right. \\
&= 2\sqrt{2} \mu \operatorname{trace} e_k e_l^t \left(H^{(2)} \circ \Lambda^{-1} e_j e_j^t \Lambda^{-1} \right) \\
&= 2\sqrt{2} \mu H^{(2)}_{kl} \left(\Lambda^{-1}_{lj} \Lambda^{-1}_{jk} \right); \\
&k = l, i \neq j \text{ LHS} = \\
&\quad = \sqrt{2} \mu \Lambda^{-1}_{ki} \Lambda^{-1}_{jk}, \qquad k = 1, \ldots n; \\
&k = l, i = j \text{ LHS} = \\
&= \mu \Lambda^{-1}_{ki} \Lambda^{-1}_{ik}, \qquad k = 1, \ldots n.
\end{aligned}
$$

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

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

last row of LHS:

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

Newton system is:

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

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

Solve for svec (l):

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

$$L_{pq} =$$

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

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

$RHS_p =$

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