Anchored Graph Realization and Sensor Localization

Nathan Krislock, Veronica Piccialli, and Henry Wolkowicz

University of Rome and University of Waterloo

Outline

Problem Formulation, $\boldsymbol{EDMC}$, (Hard Numerical Problem)

Matrix Reformulation, $\boldsymbol{EDMC}$

- Semidefinite Programming connection

SDP Relaxation of Hard Constraint $\bar{Y} = PP^T$

Facial Reduction - Reduced Problem Model

MATLAB demonstration

————————————————————————————-

Adjoints/Duality for $\boldsymbol{EDMC - R}$

Primal-Dual Bilinear Optimality Conditions (overdetermined)

Robust Interior-Point algorithm

- Gauss-Newton Direction, crossover, exact p-d feasibility, preconditioning

Concluding Remarks

Problem

- Ad hoc wireless sensor network
- A few anchors (e.g. with GPS/bulky) have fixed, known locations
- sensors within a given range have some known distance measurements (approximate)
- Problem: Determine positions of all sensors
- Parameters: Radio range, # of anchors, noise level
- Semidefinite Relaxations/Robust Algorithm

Problem Applications

- health, military, home
- natural habitat monitoring, earthquake detection, weather/current monitoring, disaster relief operations
- random deployment in inaccessible terrains

Problem Example - with Radio Range

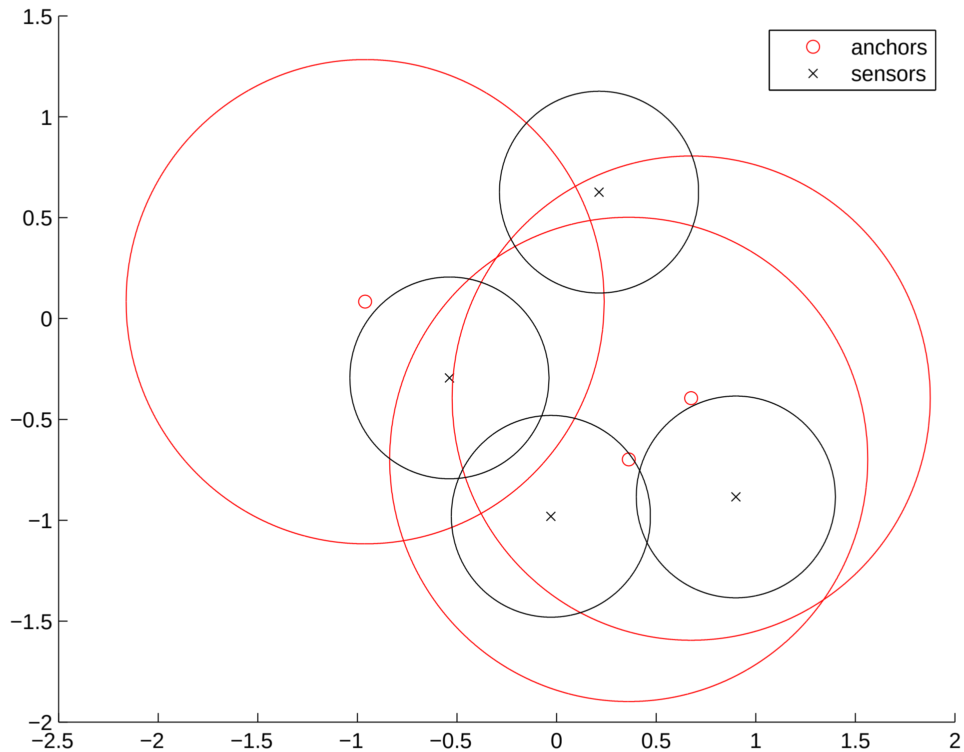


Figure 1: Connected Graph

Problem Formulation

- $p^1, \ldots, p^n \in \Re^r$ unknown (sensor) points
  $a^1, \ldots, a^m \in \Re^r$ known (anchor) points
  r embedding dimension (usually 2 or 3)
- $A^T := [a^1, a^2, \ldots, a^m]$ $X^T := [p^1, p^2, \ldots, p^n]$

$$P^T := (p^1, p^2, \ldots, p^n, a^1, a^2, \ldots, a^m)$$

$$P = \begin{pmatrix} X \\ A \end{pmatrix}$$ rows are sensor/anchor points

Sensor X and Anchor A Matrices

$$P = \begin{pmatrix} X \\ \hline A \end{pmatrix} = \begin{pmatrix} (p^1)^T \\ \ldots \\ (p^n)^T \\ \hline (a^1)^T \\ \ldots \\ (a^m)^T \end{pmatrix} = \begin{pmatrix} (1 \ -3.1) \\ \ldots \\ (-1.2 \ \ 4.2) \\ \hline (8 \ \ -2.7) \\ \ldots \\ (2 \ \ 6.1) \end{pmatrix}$$

Assumption (avoids some trivialities)

- The number of sensors and anchors, and the embedding dimension satisfy

$$n > m > r, \; A^T e = 0 \text{ and } A \text{ is full rank.}$$

Definitions

- index sets of existing values of distances d_{ij} between pairs of sensors, $\{p^i\}_1^n$:
  $\mathcal{N}_e$ distance values
  $\mathcal{N}_u$ upper bounds on distances
  $\mathcal{N}_l$ lower bounds on distances
  Similarly, the index sets $(\mathcal{M}_e, \mathcal{M}_u, \mathcal{M}_l)$ are for pairs from $\{p^i\}_1^n$ (sensors) and $\{a^k\}_1^m$ (anchors)
- partial EDM matrix E of squared distances

$$E_{ij} := \begin{cases} d_{ij}^2 & \text{if} \quad ij \in \mathcal{N}_e \cup \mathcal{M}_e \\ 0 & \text{otherwise.} \end{cases}$$

Definitions

Similarly, we define

- the (partial) matrix of (squared distances) upper bounds

  U, using $ij \in \mathcal{N}_u \cup \mathcal{M}_u$

- and the (partial) matrix of (squared distances) lower bounds

  L, using $ij \in \mathcal{N}_l \cup \mathcal{M}_l$

Weighted Least Squares Error

In the case E_{ij} have errors:

Let W_p, W_a be weight matrices. We minimize the weighted least squares error. $(\boldsymbol{EDMC})$

$$
\begin{aligned}
f_1(P) \ := \ & \sum_{(i,j)\in\mathcal{N}_e} (W_p)_{ij}(\|p^i - p^j\|^2 - E_{ij})^2 \\
& + \sum_{(i,k)\in\mathcal{M}_e} (W_a)_{ik}(\|p^i - a^k\|^2 - E_{ik})^2
\end{aligned}
$$

HARD (nonconvex) Constrained LS

$\boldsymbol{EDMC}$ Problem:

$$
\begin{array}{lll}
\text{min} & f_1(P) & \text{(weighted least squares)} \\
\text{s.t.} & \|p^i - p^j\|^2 \le U_{ij} \ \ \forall (i,j) \in \mathcal{N}_u & \left(n_u = \frac{|\mathcal{N}_u|}{2}\right) \\
 & \|p^i - a^k\|^2 \le U_{ik} \ \ \forall (i,k) \in \mathcal{M}_u & \left(m_u = \frac{|\mathcal{M}_u|}{2}\right) \\
 & \|p^i - p^j\|^2 \ge L_{ij} \ \ \forall (i,j) \in \mathcal{N}_l & \left(n_l = \frac{|\mathcal{N}_l|}{2}\right) \\
 & \|p^i - a^k\|^2 \ge L_{ik} \ \ \forall (i,k) \in \mathcal{M}_l & \left(m_l = \frac{|\mathcal{M}_l|}{2}\right)
\end{array}
$$

Trace Inner-Product

Trace (matrix) inner-product

$$\langle M, N \rangle = \operatorname{trace} M^T N = \sum_{i,j} M_{ij} N_{ij}$$

and Frobenius norm $\boxed{\|M\| = \sqrt{\operatorname{trace} M^T M}}$.

In $\mathcal{S}^n$, $n \times n$ symmetric matrices,

$A = A^T, B = B^T$:

$$\langle A, B \rangle = \operatorname{trace} AB = \sum_{i,j} A_{ij} B_{ij}$$

Löwner Partial Order

$B \in \mathcal{S}^n$ is positive semidefinite if
$x^T B x \geq 0,\ \forall x \in \Re^n$ ($\iff \lambda_i(B) \geq 0,\ \forall i$)

$$B \succeq 0 \quad \text{(is positive semidefinite)}$$

$$\iff$$

$$\exists P \text{ with } B = PP^T,\ \operatorname{rank}(B) = \operatorname{rank}(P)$$

the positive semidefinite (Löwner) partial order is:

$$A \succeq B\ (A \succ B) \text{ if } A - B \succeq 0\ (A - B \succ 0)$$

Daniel Fylstra

Introducing Convex and Conic Optimization for the Quantitative Finance Professional

Few people are aware of a quiet revolution that has taken place in optimization methods over the last decade

Optimization has played an important role in quantitative finance ever since Markowitz published his original paper on portfolio selection in 1952[1]. Most “quants” have some knowledge of linear programming as used in bond duration matching, quadratic programming as used in equity portfolio optimization, and nonlinear optimization used with portfolios of derivatives. The finance industry and technical literature is home to many different, competing approaches that apply these standard optimization methods to different models and risk measures, in an effort to achieve better investment results.

But fewer people are aware of a quiet revolution that has taken place in the optimization methods themselves over the last decade. A better understanding of the properties of optimization models, and new algorithms – notably interior point or barrier methods – have led to a changed view of the whole field of optimization. It is not an exaggeration to say that linear and quadratic programming are being replaced – or more properly subsumed – by more powerful and general methods of convex and conic optimization.

Semidefinite Matrix

$$0 \preceq B \;=\; PP^T = \begin{pmatrix} \dots & \dots & \dots & \dots \\ \dots & (p_i)^T p_i & \dots & (p_i)^T p_j \\ \dots & \dots & \dots & \dots \\ \dots & (p_j)^T p_i & \dots & (p_j)^T p_j \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

AND:

$$\|p^i - p^j\|^2 = (p^i)^T p^i + (p^j)^T p^j - 2(p^i)^T p^j$$

$\mathcal{K}(\boldsymbol{SDP}) = \boldsymbol{EDM}$

$B = PP^T$ **($\boldsymbol{SDP}$)**. $B_{ii} = (p^i)^T p^i$; $B_{ij} = (p^i)^T p^j$

The squared distance

$$
\begin{array}{rcccccc}
D_{ij} & = & \|p^i - p^j\|^2 & & \textbf{(}\boldsymbol{EDM}\textbf{)} & & \\
 & = & (p^i)^T p^i & + & (p^j)^T p^j & - & 2(p^i)^T p^j \\
 & = & \updownarrow & & \updownarrow & & \updownarrow \\
 & = & (\operatorname{diag}(B) e^T & + & e \operatorname{diag}(B)^T & - & 2B)_{ij} \\
 & =: & (\mathcal{K}(B))_{ij} & & & &
\end{array}
$$

$D = \mathcal{K}(B)$ change $\boldsymbol{EDM}$ $D \leftrightarrow$ $\boldsymbol{SDP}$ B

Matrix Reformulation of $EDMC$

Let $\bar{Y} := PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & AA^T \end{pmatrix}$

We get the equivalent $\boldsymbol{EDMC}$

$$\begin{array}{ll} \min & f_2(\bar{Y}) := \frac{1}{2}\|W \circ (\mathcal{K}(\bar{Y}) - E)\|_F^2 \\ \text{subject to} & g_u(\bar{Y}) := H_u \circ (\mathcal{K}(\bar{Y}) - U) \le 0 \\ & g_l(\bar{Y}) := H_l \circ (\mathcal{K}(\bar{Y}) - L) \ge 0 \\ & \text{hard constraint } \boxed{\bar{Y} - PP^T = 0} \end{array}$$

SDP Relaxation of Hard Constraint

$$\bar{Y} = PP^T = \begin{pmatrix} XX^T & XA^T \\ AX^T & \boxed{AA^T} \end{pmatrix} \quad \text{holds}$$

$$\Longleftrightarrow$$

$$\bar{Y}_{11} = XX^T \text{ and } \bar{Y}_{21} = AX^T, \boxed{\bar{Y}_{22} = AA^T}.$$

Relax $\bar{Y} = PP^T$ to (Löwner partial order)

$\boxed{\bar{Y}_{22} = AA^T}$ $\quad PP^T - \bar{Y} \preceq 0 \quad$ quadr convex constr

(But .... why this relaxation?)

Convex wrt Löwner Partial Order

The constraint $g(P,Y) = PP^T - Y \preceq 0$ is $\succeq$-convex, since each function

$$\phi_Q(P,Y) = \operatorname{trace} Qg(P,Y) \quad \text{is convex } \forall Q \succeq 0.$$

Note

$$\begin{aligned} \operatorname{trace} QPP^T &= \operatorname{trace} QPIP^T \\ &= \operatorname{vec}(P)^T (I \otimes Q) \operatorname{vec}(P) \end{aligned}$$

Hessian is $I \otimes Q \succeq 0$;
and the cone $\boldsymbol{SDP}$ is self-polar.

Linearization of SDP Relaxation

$$PP^T - \bar{Y} \preceq 0, \qquad P = \begin{pmatrix} X \\ A \end{pmatrix}, \boxed{\bar{Y}_{22} = AA^T}$$

$\Longleftrightarrow$ (by Schur complement)

$$Z = \begin{pmatrix} I & P^T \\ P & \bar{Y} \end{pmatrix} \succeq 0, \quad P = \begin{pmatrix} X \\ A \end{pmatrix}, \boxed{\bar{Y}_{22} = AA^T}$$

$\Longleftrightarrow$ (ignore $\bar{\cdot}$)

$$Z = \begin{pmatrix} I & X^T & A^T \\ X & Y & Y_{21}^T \\ A & Y_{21} & \boxed{AA^T} \end{pmatrix} \succeq 0, \qquad \left(\begin{array}{c} \text{NOT!} \ \succ 0 \\ \Rightarrow Y_{21} = AX^T \end{array} \right)$$

Facial Reduction

$$Z_s := \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix}$$

(NE 2×2 block)

(Lin.Tr. but NOT onto)

THEOREM:

$$Z = \begin{pmatrix} I & X^T & A^T \\ X & Y & Y_{21}^T \\ A & Y_{21} & \boxed{AA^T} \end{pmatrix} \succeq 0$$

$$\Longleftrightarrow$$

$Z_s \succeq 0$ and $Y_{21} = AX^T$

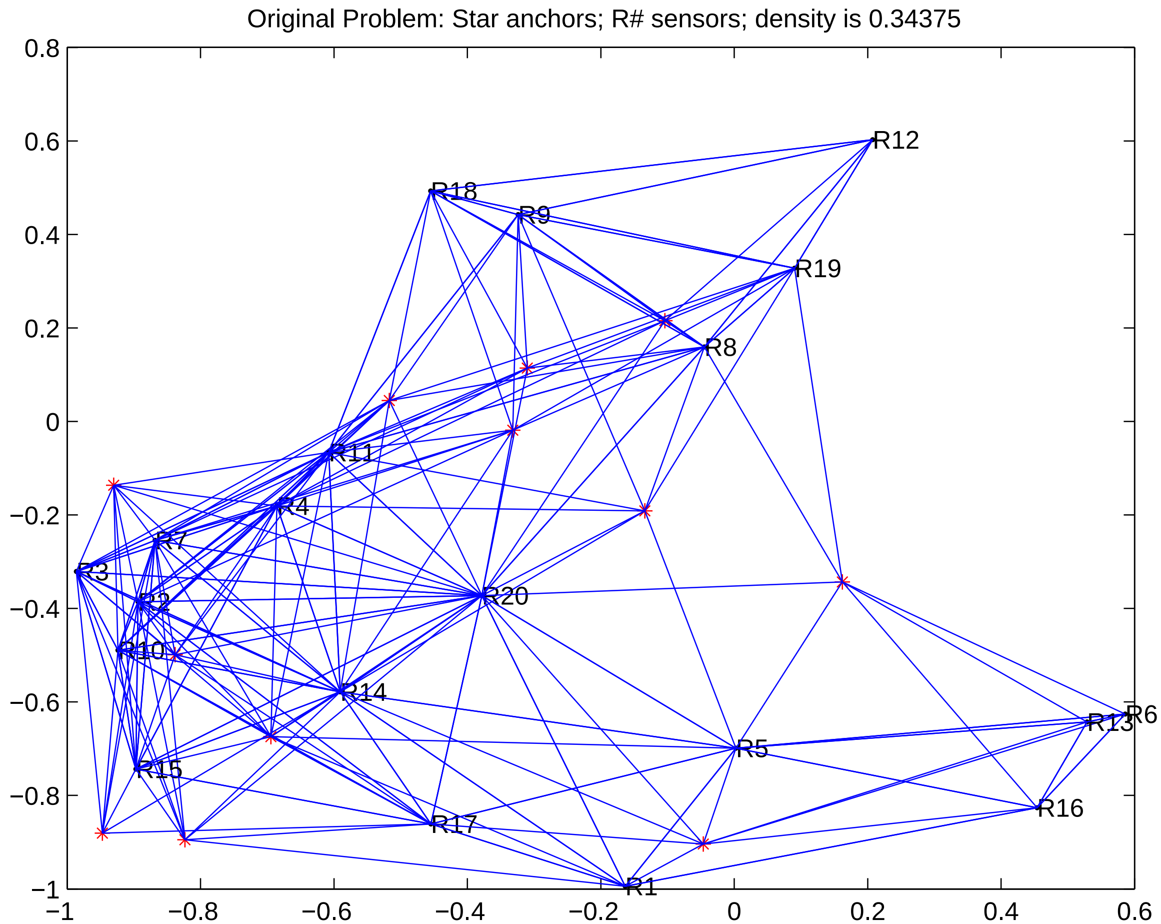
Equivalent Reduced Problem Model

$(\boldsymbol{EDMC-R})$

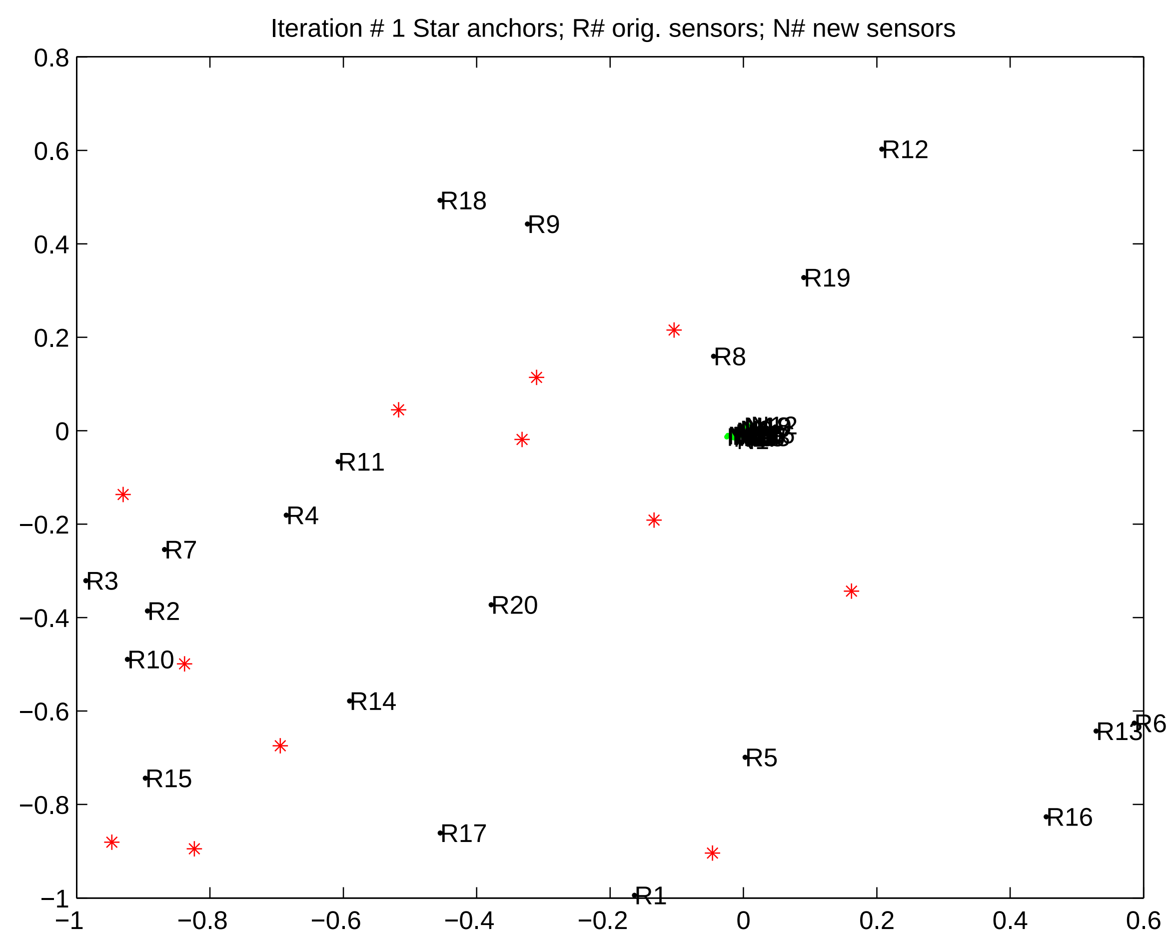
$$
\begin{array}{lrcl}
\min & f_3(x,y) := \frac{1}{2}\|W \circ (\mathcal{K}(\mathcal{Y}(x,y))) - \bar{E}\|_F^2 & & \\
\textbf{s.t.} & g_u(x,y) := H_u \circ \mathcal{K}(\mathcal{Y}(x,y)) - \bar{U} & \leq & 0 \\
 & g_l(x,y) := \bar{L} - H_l \circ \mathcal{K}(\mathcal{Y}(x,y)) & \leq & 0 \\
 & \mathrm{sBlk}_1(I) + \mathcal{Z}_s(x,y) & \succeq & 0
\end{array}
$$

(objective is ℓ_2 rather than ℓ_1 in the literature, e.g. H. Jin(05), A. So, Y. Ye(05), P. Biswas, T. Liang, K. Toh, T. Wang, Y. Ye(06).)

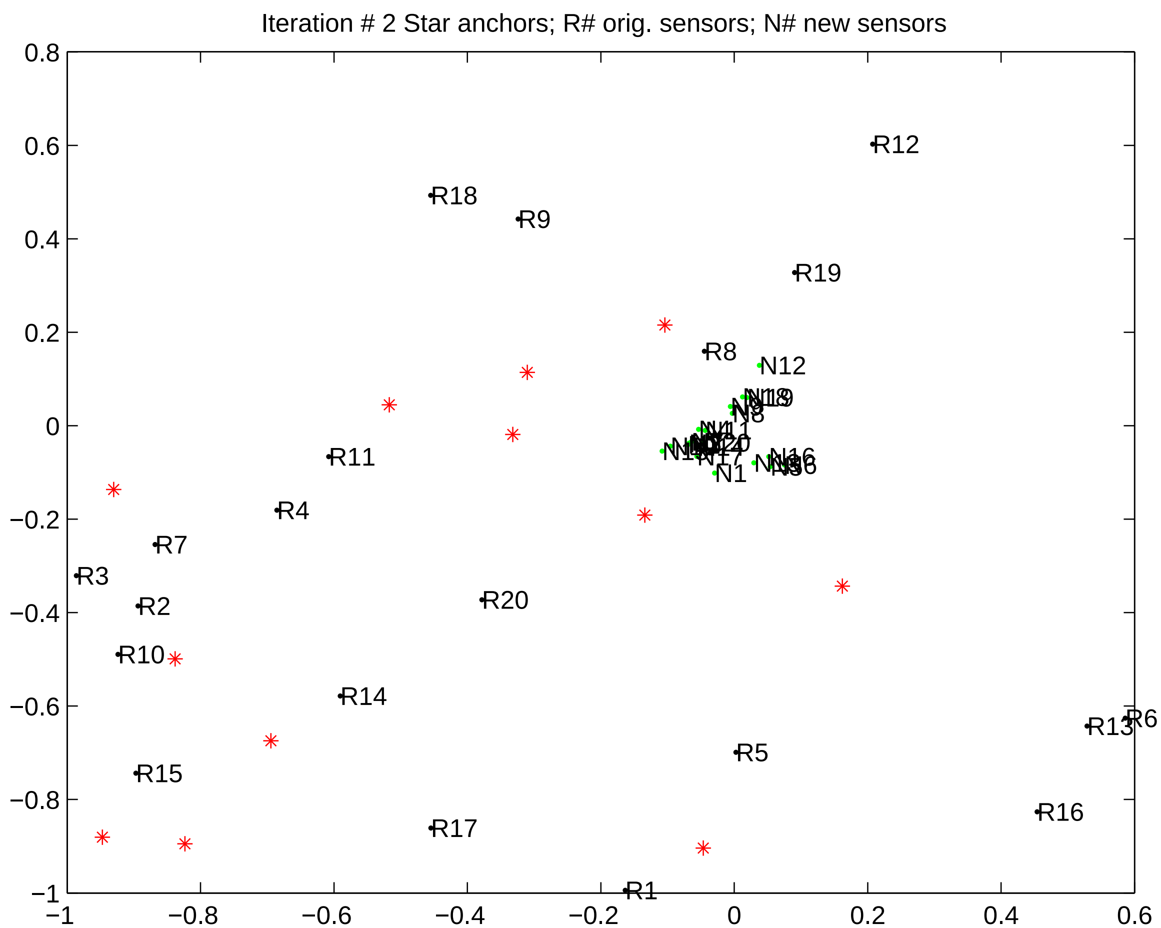
MATLAB Demo - Original Data



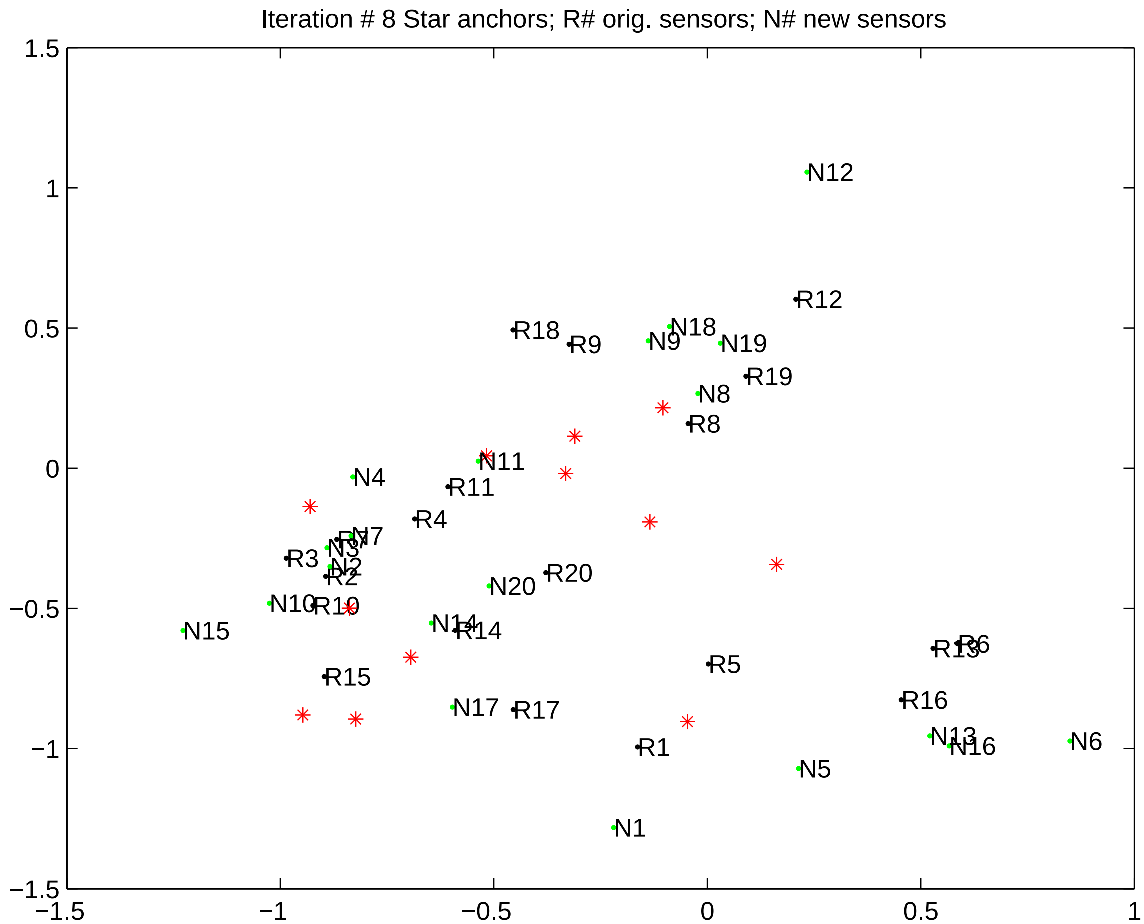
MATLAB Demo - After 1 Iteration



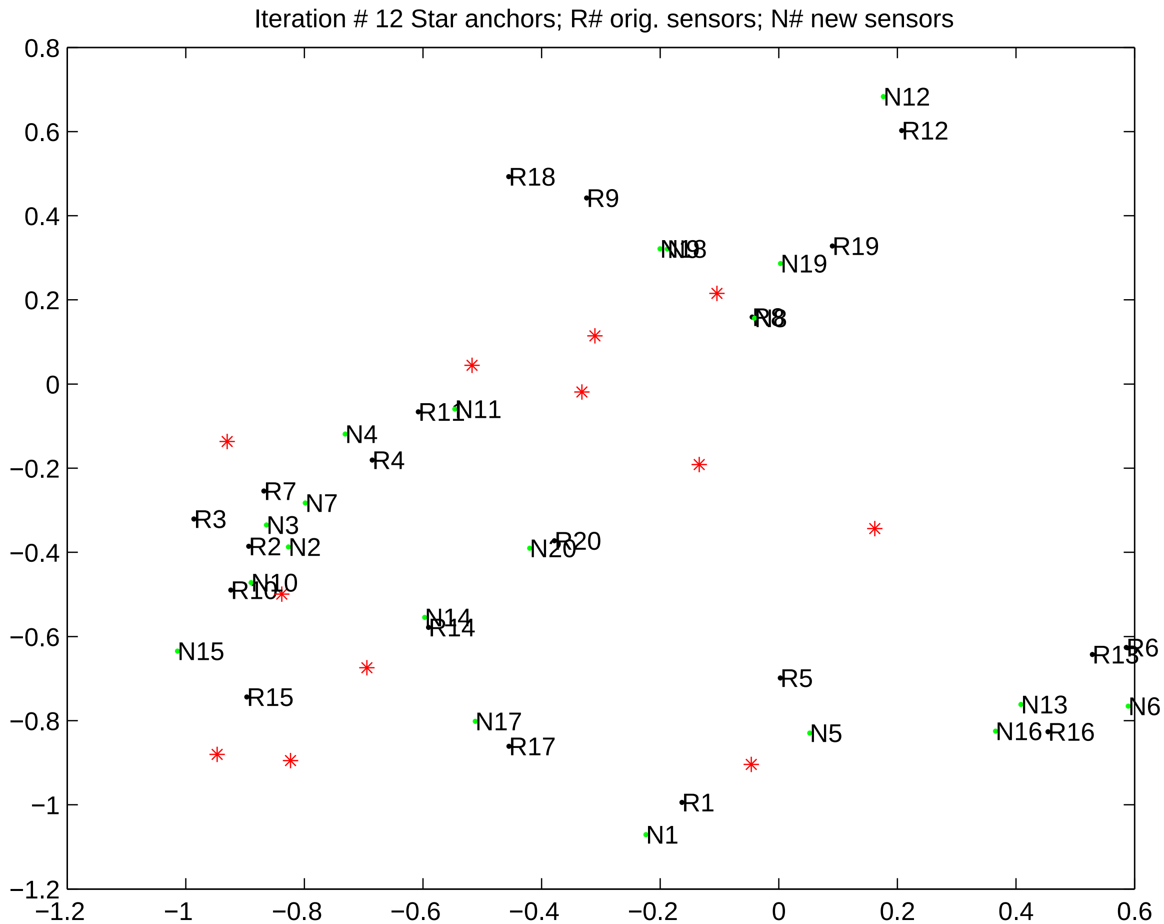
MATLAB Demo - After 2 Iteration



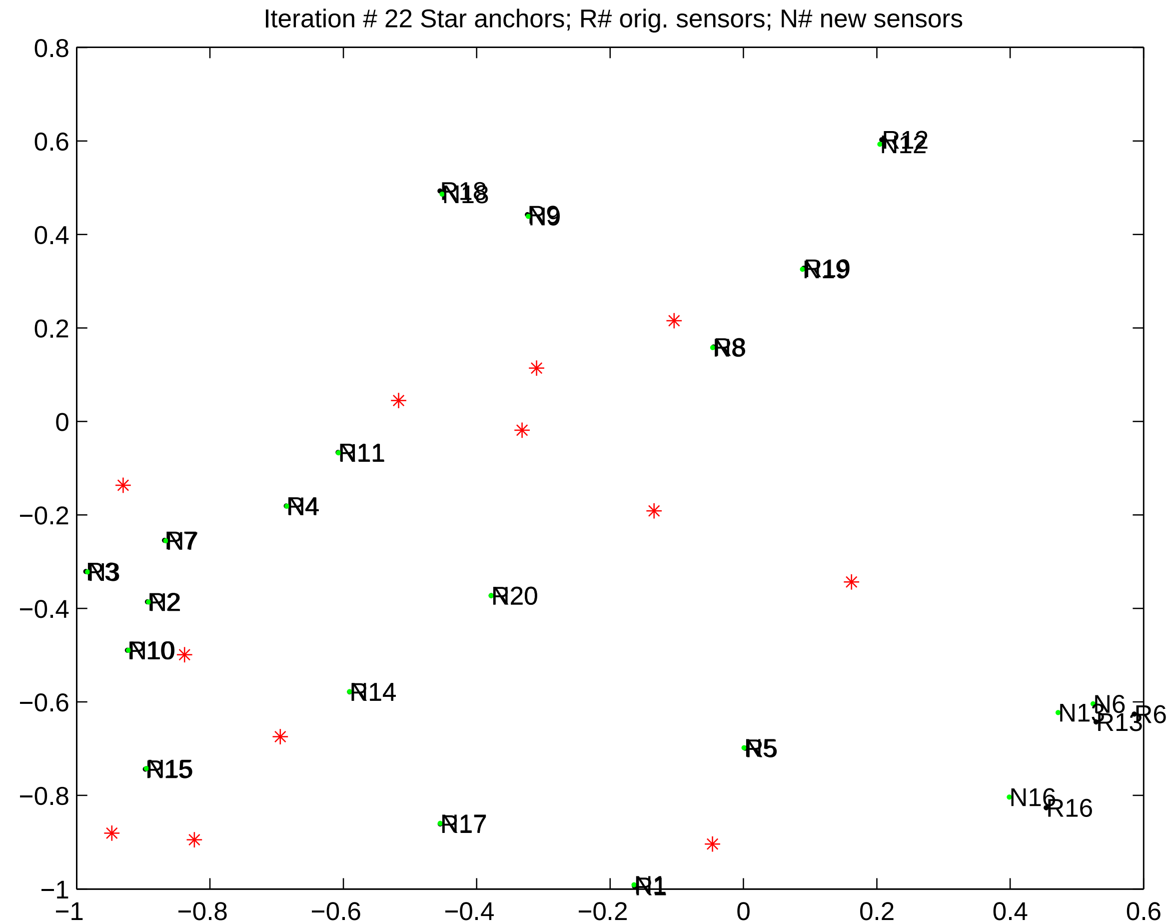
MATLAB Demo - After 8 Iteration



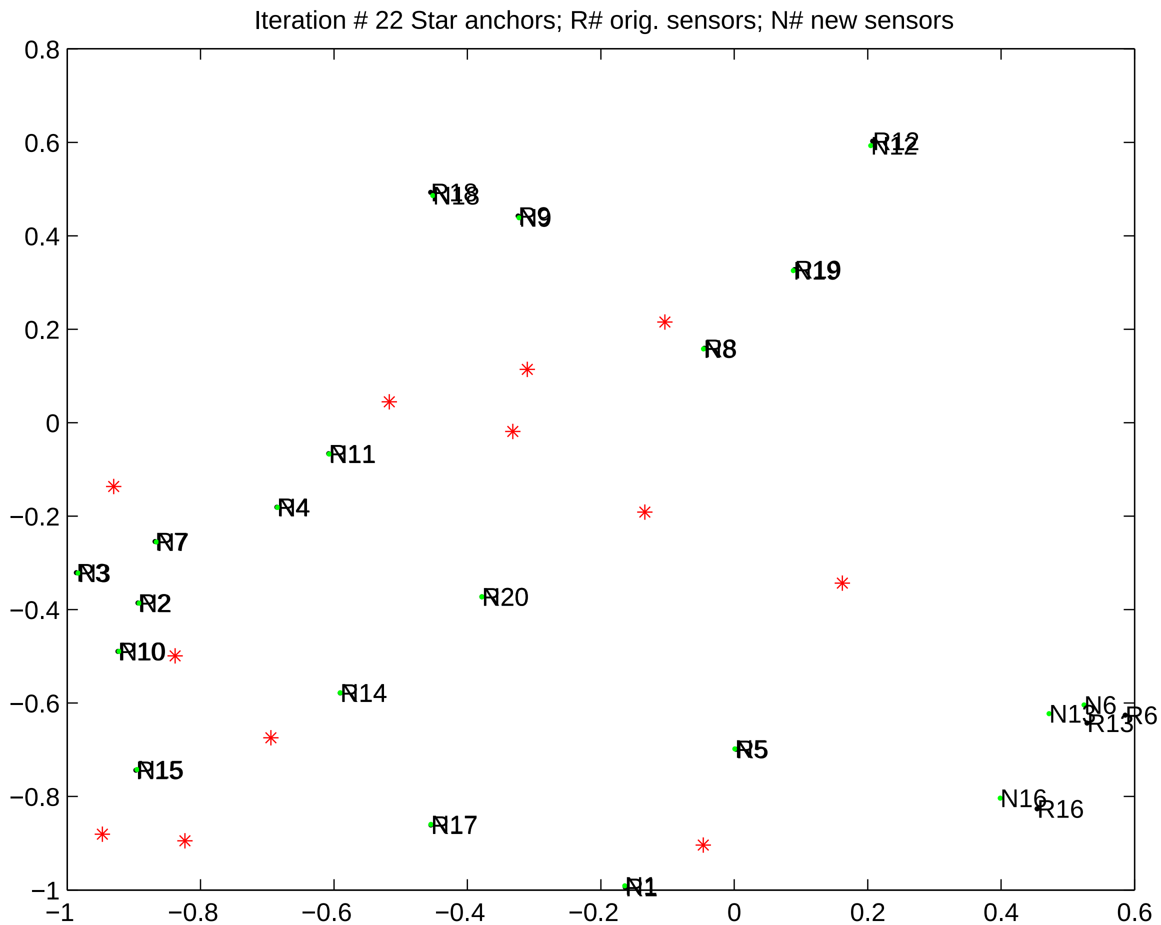
MATLAB Demo - After 12 Iteration



MATLAB Demo - After 22 Iteration



MATLAB Demo


Iteration # 22 Star anchors; R# orig. sensors; N# new sensors
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
R12 N12
R18 N18
R9 N9
R19 N19
R8 N8
R11 N11
R4 N4
R7 N7
R3 N3
R2 N2
R20 N20
R10 N10
R14 N14
N6 R6
N13 R13
R5 N5
R15 N15
N16 R16
R17 N17
R1 N1

Facial Reduction - Proof Outline

(compact) singular value decomposition

$\boxed{A = U\Sigma V^T}$ $\;U m \times r, V r \times r$

$$Z = Z_1 := \begin{pmatrix} I & X^T & V\Sigma U^T \\ X & Y & Y_{21}^T \\ U\Sigma V^T & Y_{21} & U\Sigma^2 U^T \end{pmatrix} \succeq 0$$

choose $\bar{U}$ so that $\begin{pmatrix} U & \bar{U} \end{pmatrix}$ is orthogonal;

Facial Reduction - Proof Outline cont...

Nonsingular congruence (apply Sylvester Lemma on inertia)

$$0 \preceq Z_2 := T^T Z T =$$

$$= \begin{pmatrix} V^T & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \begin{pmatrix} U & \bar{U} \end{pmatrix}^T \end{pmatrix} Z \begin{pmatrix} V & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \begin{pmatrix} U & \bar{U} \end{pmatrix} \end{pmatrix}$$

Facial Reduction - Congruence cont...

$$= \begin{pmatrix} I & V^TX^T & \begin{pmatrix} \Sigma & 0 \end{pmatrix} \\ XV & Y & \begin{pmatrix} Y_{21}^TU & Y_{21}^T\bar{U} \end{pmatrix} \\ \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} & \begin{pmatrix} U^TY_{21} \\ \bar{U}^TY_{21} \end{pmatrix} & \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \perp \begin{pmatrix} 0 & 0 \\ 0 & \boxed{I} \end{pmatrix}$$

$$\Rightarrow Z \perp \left[T \begin{pmatrix} 0 & 0 \\ 0 & \boxed{I} \end{pmatrix} T^T \right] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \boxed{\bar{U}\bar{U}^T} \end{pmatrix} := Q$$

Minimal and Conjugate Faces

Conjugate face to feasible set $\mathcal{F}_Z$ is

$$\boldsymbol{SDP} \cap \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{U}\bar{U}^T \end{pmatrix} \right\}^\perp$$

Minimal face of $\boldsymbol{SDP}$ containing $\mathcal{F}_Z = \text{cone}\,\mathcal{F}_Z$ is

$$\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix} \mathcal{S}_+^{2r+n} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix}^T$$

Minimal Face

each given feasible $Z = \begin{pmatrix} I & X^T & A^T \\ X & Y & Y_{21}^T \\ A & Y_{21} & AA^T \end{pmatrix} \succeq 0,$

can be expressed as (using $A = U\Sigma V^T$)

$$= \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix} \begin{pmatrix} I & X^T & V\Sigma \\ X & Y & XV\Sigma \\ \Sigma V^T & \Sigma V^T X^T & \Sigma^2 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U \end{pmatrix}$$

Matrix $\leftrightarrow$ Vector Notation I

vector $\boxed{v = \operatorname{vec} V}$ is matrix V taken columnwise

$$\boxed{\operatorname{sblk}_{21}\begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix} = \sqrt{2}X}$$ pulls out the 21 block -

the $\sqrt{2}$ is for isometry in Frobenius norm

$$x := \operatorname{vec}\left(\operatorname{sblk}_{21}\begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix}\right) = \sqrt{2}\operatorname{vec}(X)$$

$\boxed{y := \operatorname{svec}(Y)}$ $(Y = Y^T,$ isometry$)$

Matrix $\leftrightarrow$ Vector Notation II

(adjoints: $\mathrm{sblk}_{21}^{*}(X) =$)

$$\mathrm{sBlk}_{21}(X) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix}$$

$$\mathrm{svec}^{-1}(\cdot) = \mathrm{svec}^{*}(\cdot) = \mathrm{sMat}(\cdot)$$

$$\mathcal{Z}_s^x(x) := \mathrm{sBlk}_{21}(\mathrm{Mat}(x)), \quad \mathcal{Z}_s^y(y) := \mathrm{sBlk}_{2}(\mathrm{sMat}($$

$$\mathcal{Z}_s(x,y) := \mathcal{Z}_s^x(x) + \mathcal{Z}_s^y(y), \quad Z_s := \mathrm{sBlk}_1(I) + \mathcal{Z}_s(x$$

to build $\quad Z_s = \begin{pmatrix} I & X^T \\ X & Y \end{pmatrix} \in \mathcal{S}^{r+n}$

Matrix $\leftrightarrow$ Vector Notation III

$$\mathcal{Y}^x(x) = \mathrm{sBlk}_{21}(A\mathrm{Mat}\,(x)^T),\ \mathcal{Y}^y(y) = \mathrm{sBlk}_1(\mathrm{sMat}$$

$$\mathcal{Y}(x,y) = \mathcal{Y}^x(x) + \mathcal{Y}^y(y),\ \boxed{\bar{Y} = \mathrm{sBlk}_2(AA^T) + \mathcal{Y}(x}$$

$$\begin{aligned}\bar{E} &:= W \circ \left[E - \mathcal{K}(\mathrm{sBlk}_2(AA^T))\right]\\ \bar{U} &:= H_u \circ \left[\mathcal{K}(\mathrm{sBlk}_2(AA^T)) - U\right]\\ \bar{L} &:= H_l \circ \left[L - \mathcal{K}(\mathrm{sBlk}_2(AA^T))\right]\end{aligned}$$

The unknown matrix $\bar{Y}$ is equal to $\mathcal{Y}(x,y)$ (with additional constant $2,2$ block), i.e. unknowns are the vectors x, y.

Equivalent Reduced Problem Model

$(\boldsymbol{EDMC-R})$

$$
\begin{array}{lrcl}
\min & f_3(x,y) := \frac{1}{2}\|W \circ (\mathcal{K}(\mathcal{Y}(x,y))) - \bar{E}\|_F^2 & & \\
\text{s.t.} & g_u(x,y) := H_u \circ \mathcal{K}(\mathcal{Y}(x,y)) - \bar{U} & \leq & 0 \\
 & g_l(x,y) := \bar{L} - H_l \circ \mathcal{K}(\mathcal{Y}(x,y)) & \leq & 0 \\
 & \mathrm{sBlk}_1(I) + \mathcal{Z}_s(x,y) & \succeq & 0
\end{array}
$$

(objective is ℓ_2 rather than ℓ_1 in the literature, e.g. H. Jin(05), A. So, Y. Ye(05), P. Biswas, T. Liang, K. Toh, T. Wang, Y. Ye(06).)

Problems with Relaxation

1. $\{(\bar{Y}, P) : \bar{Y} = PP^T\} \subset \{(\bar{Y}, P) : \bar{P}P^T - \bar{Y} \preceq 0\}$
   (But, is Lagrangian relaxation stronger?)
2. linearization (using Schur complement) results in a constraint that is *NOT* onto, i.e. two relaxations *NOT* numerically equivalent
3. Least squares problem is (usually) underdetermined.

Lagrangian of $EDMC-R$

$$
\begin{aligned}
L(x,y,\Lambda_u,\Lambda_l,\Lambda) = {} & \\
& \tfrac{1}{2}\|W \circ \mathcal{K}(\mathcal{Y}(x,y) - \bar{E}\|_F^2 \\
& \quad + \langle \Lambda_u, H_u \circ \mathcal{K}(\mathcal{Y}(x,y)) - \bar{U} \rangle \\
& \qquad + \langle \Lambda_l, \bar{L} - H_l \circ \mathcal{K}(\mathcal{Y}(x,y)) \rangle \\
& \qquad\quad - \langle \Lambda, \mathrm{sBlk}\,_1(I) + \mathcal{Z}_s(x,y) \rangle \,,
\end{aligned}
$$

where $0 \leq \Lambda_u, 0 \leq \Lambda_l \in \mathcal{S}^{m+n}, \quad 0 \preceq \Lambda \in \mathcal{S}^{m+n}$

$$
\Lambda = \begin{pmatrix} \Lambda_1 & \Lambda_{21}^T \\ \Lambda_{21} & \Lambda_2 \end{pmatrix}, \qquad \langle A, B \rangle = \operatorname{trace} A^T B.
$$

Matrix $\leftrightarrow$ Vector Dual Variable Notation

$$
\begin{array}{cc}
\lambda_u := \operatorname{svec}(\Lambda_u), & \lambda_l := \operatorname{svec}(\Lambda_l), \\
h_u := \operatorname{svec}(H_u), & h_l := \operatorname{svec}(H_l), \\
\lambda := \operatorname{svec}(\Lambda), & \lambda_1 := \operatorname{svec}(\Lambda_1), \\
\lambda_2 := \operatorname{svec}(\Lambda_2), & \lambda_{21} := \operatorname{vec}\operatorname{sblk}_{21}(\Lambda).
\end{array}
$$

Adjoints

To differentiate the Lagrangian, we need the adjoints of the various linear transformations, e.g. part of $\mathcal{K}$:

$$
\begin{aligned}
\mathcal{D}_e(B) &= \operatorname{diag}(B)\, e^T + e \operatorname{diag}(B)^T \\
\mathcal{D}_e^*(D) &= 2\operatorname{Diag}(De) \\
\langle \mathcal{D}_e(B), D \rangle &= \operatorname{trace}\left(\operatorname{diag}(B) e^T D + e \operatorname{diag}(B)^T D\right) \\
&= \operatorname{trace}\left(De(\operatorname{diag} B)^T + De(\operatorname{diag} B)^T\right) \\
&= 2\operatorname{trace}\ (\operatorname{diag} B)^T (De) \\
&= \langle B, \mathcal{D}_e^*(D) \rangle, \forall D, B
\end{aligned}
$$

Primal-Dual Optimal. Conditions 1

THEOREM: The primal-dual variables $x, y, \Lambda, \lambda_u, \lambda_l$ are optimal for $\boldsymbol{EDMC - R}$ if and only if:

1. Primal Feasibility:

The slack variables satisfy

$$S_u = \bar{U} - H_u \circ (\mathcal{K}(\mathcal{Y}(x,y))), \; s_u = \operatorname{svec} S_u \geq 0$$

$$S_l = H_l \circ (\mathcal{K}(\mathcal{Y}(x,y))) - \bar{L}, \quad s_l = \operatorname{svec} S_l \geq 0$$

and

$$\begin{aligned} Z_s &= \operatorname{sBlk}_1(I) + \operatorname{sBlk}_2 \operatorname{sMat}(y) + \operatorname{sBlk}_{21} \operatorname{Mat} \\ &\succeq 0 \end{aligned}$$

Primal-Dual Optimal. Conditions 2a

2a. Dual Feasibility:

The stationarity equations ($\Rightarrow$ exact p-d feas.)

$$
\begin{aligned}
(\mathcal{Z}_s^x)^*(\Lambda) &= \lambda_{21} \quad \text{(eliminated)}\\
&= [W \circ (\mathcal{K}\mathcal{Y}^x)]^* \left(W \circ \mathcal{K}(\mathcal{Y}(x,y)) - \right.\\
&\quad\quad + [H_u \circ (\mathcal{K}\mathcal{Y}^x)]^* (\Lambda_u)\\
&\quad\quad\quad - [H_l \circ (\mathcal{K}\mathcal{Y}^x)]^* (\Lambda_l)\\
(\mathcal{Z}_s^y)^*(\Lambda) &= \lambda_{2} \quad \text{(eliminated)}\\
&= [W \circ (\mathcal{K}\mathcal{Y}^y)]^* \left(W \circ \mathcal{K}(\mathcal{Y}(x,y)) - \right.\\
&\quad\quad + [H_u \circ (\mathcal{K}\mathcal{Y}^y)]^* (\Lambda_u)\\
&\quad\quad\quad - [H_l \circ (\mathcal{K}\mathcal{Y}^y)]^* (\Lambda_l)
\end{aligned}
$$

Primal-Dual Optimal. Conditions 2b

2b. Dual Feasibility:

Nonnegativity

$$\Lambda = \mathrm{sBlk}_1\mathrm{sMat}(\lambda_1) + \mathrm{sBlk}_2\mathrm{sMat}(\lambda_2) + \mathrm{sBlk}_{21}\mathrm{Mat}(\lambda_{21}) \succeq 0;$$

$$\lambda_u \geq 0; \lambda_l \geq 0$$

$$\Lambda = \Lambda(\lambda_1, x, y, \lambda_u, \lambda_l)$$ (from stationarity)

Primal-Dual Optimal. Conditions 3 (C.S.)

3. Complementary Slackness:

$$\begin{aligned} \lambda_u \circ s_u &= 0 \\ \lambda_l \circ s_l &= 0 \\ \Lambda Z_s &= 0 \end{aligned}$$

(equivalently $\operatorname{trace} \Lambda Z_s = 0$)

Perturbed Compl. Slack. Conditions

$$F_\mu(x, y, \lambda_u, \lambda_l, \lambda_1) := \begin{pmatrix} \lambda_u \circ s_u - \mu_u e \\ \lambda_l \circ s_l - \mu_l e \\ \boxed{\Lambda Z_s - \mu_c I} \end{pmatrix} = 0,$$

where $s_u = s_u(x, y)$, $s_l = s_l(x, y)$,
$\Lambda = \Lambda(\lambda_1, x, y, \lambda_u, \lambda_l)$, $Z_s = Z_s(x, y)$
an overdetermined bilinear system with
$(m_u + n_u) + (m_l + n_l) + (n + r)^2$ equations
$nr + t(n) + (m_u + n_u) + (m_l + n_l) + t(r)$ variables.

Gauss-Newton Search Direction

$$\Delta s := \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \lambda_u \\ \Delta \lambda_l \\ \Delta \lambda_1 \end{pmatrix}$$

overdetermined linearized system is:

$$F'_\mu(\Delta s) \cong F'_\mu(x, y, \lambda_u, \lambda_l, \lambda_1)(\Delta s) = -F_\mu(x, y, \lambda_u, \lambda_l,$$

Notation - Compos. of Lin. Tr.

$$
\begin{aligned}
\mathcal{K}_H^x(x) &:= H \circ (\mathcal{K}(\mathcal{Y}^x(x))), \\
\mathcal{K}_H^y(y) &:= H \circ (\mathcal{K}(\mathcal{Y}^y(y))), \\
\mathcal{K}_H(x, y) &:= H \circ (\mathcal{K}(\mathcal{Y}(x, y))).
\end{aligned}
$$

GN: Three blocks of Equations

1. $\lambda_u \circ \operatorname{svec} \mathcal{K}_{H_u}(\Delta x, \Delta y) + s_u \circ \Delta\lambda_u = \mu_u e - \lambda_u \circ s_u$
2. $\lambda_l \circ \operatorname{svec} \mathcal{K}_{H_l}(\Delta x, \Delta y) + s_l \circ \Delta\lambda_l = \mu_l e - \lambda_l \circ s_l$
3.

$$
\begin{aligned}
&\Lambda \mathcal{Z}_s(\Delta x, \Delta y) + [\operatorname{sBlk}_1(\operatorname{sMat}(\Delta\lambda_1))\\
&\quad + \operatorname{sBlk}_2(\operatorname{sMat}\{(\mathcal{K}_W^y)^* \mathcal{K}_W(\Delta x, \Delta y) +\\
&\qquad (\mathcal{K}_{H_u}^y)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^y)^*(\operatorname{sMat}(\Delta\lambda_l))\})\\
&\qquad + \operatorname{sBlk}_{21}(\operatorname{Mat}\{(\mathcal{K}_W^x)^* \mathcal{K}_W(\Delta x, \Delta y)\\
&\qquad\quad + (\mathcal{K}_{H_u}^x)^*(\operatorname{sMat}(\Delta\lambda_u)) - (\mathcal{K}_{H_l}^x)^*(\operatorname{sMat}(\Delta\lambda_l))\\
&= \mu_c I - \Lambda Z_s
\end{aligned}
$$

Initial Str. Feas. Start Heuristic

If the graph is connected, we can use the stationarity equations and get a strictly feasible primal-dual starting point and *maintain exact numerical primal-dual feasibility* throughout the iterations.

Diagonal Preconditioning

Given $A \in \mathcal{M}^{m\times n}$, $m \geq n$ full rank matrix; and using condition number of $K \succ 0$:

$\omega(K) = \frac{\operatorname{trace}(K)/n}{\det(K)^{1/n}}$, the optimal diagonal scaling

$$\min_{D\succ 0} \omega\left((AD)^T(AD)\right), \quad D^* = \operatorname{Diag}\left(1/\|A_{:,i}\|\right)$$

(cite Dennis-W.) Therefore, need to evaluate columns of $F'_\mu(\cdot)$ (can be done explicitly/efficiently)

(Partial block Cholesky precondioning)

dens: W .75,L .8; n 15, m 5, r 2

nf	optvalue	relaxation	cond.number	sv($\mathcal{Z}_s$)	sv(F'_μ)
0.0000e+000	3.9909e-009	1.1248e-005	3.8547e+006	15	19
5.0000e-002	7.5156e-004	4.4637e-002	1.0244e+011	6	27
1.0000e-001	3.7103e-003	1.1286e-001	1.9989e+010	5	25
1.5000e-001	6.2623e-003	1.3125e-001	1.0065e+010	6	14
2.0000e-001	1.3735e-002	1.3073e-001	6.8833e+009	7	12
2.5000e-001	2.3426e-002	2.4828e-001	2.4823e+010	8	6
3.0000e-001	6.0509e-002	2.3677e-001	3.4795e+010	7	7
3.5000e-001	5.5367e-002	3.7260e-001	2.3340e+008	6	4
4.0000e-001	7.6703e-002	3.6343e-001	8.9745e+010	8	3
4.5000e-001	1.2493e-001	6.9625e-001	3.2590e+010	6	9
5.0000e-001	1.3913e-001	3.9052e-001	2.2870e+005	8	0
5.5000e-001	8.8552e-002	3.8742e-001	5.8879e+007	8	2
6.0000e-001	4.2425e-001	4.1399e-001	4.9251e+012	8	4

dens: W .75,L .8; n 15, m 5, r 2

nf	optvalue	relaxation	cond.number	sv($\mathcal{Z}_s$)	sv(F'_μ)
0.0000e+000	3.9909e-009	1.1248e-005	3.8547e+006	15	19
5.0000e-002	7.5156e-004	4.4637e-002	1.0244e+011	6	27
5.5000e-001	8.8552e-002	3.8742e-001	5.8879e+007	8	2
6.0000e-001	4.2425e-001	4.1399e-001	4.9251e+012	8	4
6.5000e-001	2.0414e-001	6.6054e-001	2.4221e+010	7	4
7.0000e-001	1.2028e-001	3.4328e-001	1.9402e+010	7	6
7.5000e-001	2.6590e-001	7.9316e-001	1.3643e+011	7	4
8.0000e-001	4.7155e-001	3.7822e-001	6.6910e+009	8	2
8.5000e-001	1.8951e-001	5.8652e-001	1.4185e+011	6	7
9.0000e-001	2.1741e-001	9.8757e-001	2.9077e+005	8	0
9.5000e-001	4.4698e-001	4.6648e-001	2.7013e+006	9	2

Table 1: Robust Algorithm for Ill-posed Problem