Log-det heuristic for matrix rank minimization
(with applications to Hankel and Euclidean distance matrices)

Maryam Fazel
California Institute of Technology

Haitham Hindi  Stephen Boyd
Stanford University

June 5, 2003
Outline

- Rank Minimization Problem (RMP)
- Examples
- Solution methods
  - trace & log-det heuristics for PSD matrices
  - semidefinite embedding technique
  - trace & log-det heuristics for general matrices
Rank Minimization Problem (RMP)

\[
\begin{align*}
\text{minimize} & \quad \text{Rank } X \\
\text{subject to} & \quad X \in \mathcal{C},
\end{align*}
\]

\(X \in \mathbb{R}^{m \times n}\) is the optimization variable; \(\mathcal{C}\) is convex set

- RMP is difficult nonconvex problem (NP-hard)
- RMP arises in many application areas
- usual meaning: find simplest or minimum order system; model with fewest parameters . . . (Occam's razor)
Maximum sparsity problem

important special case of RMP: variable $X = \text{diag}(x)$

then $\text{Rank } X = \text{card}(x)$, number of nonzero $x_i$

RMP reduces to finding the \textbf{sparsest} vector in convex set $\mathcal{C}$:

\[
\begin{align*}
\text{minimize} & \quad \text{card}(x) \\
\text{subject to} & \quad x \in \mathcal{C}
\end{align*}
\]

meaning: find simplest model, design with fewest components, sparse signal representation, . . .
Minimum order system realization

find minimum order system that satisfies time-domain specs (rise-time, slew-rate, overshoot, settling-time, ...)

\[
\begin{align*}
\text{minimize} & \quad \text{Rank}
\begin{bmatrix}
h_1 & h_2 & \cdots & h_n \\
h_2 & h_3 & \cdots & h_{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
h_n & h_{n+1} & \cdots & h_{2n-1}
\end{bmatrix} \\
\text{subject to} & \quad F(h_1, \ldots, h_n) \leq g
\end{align*}
\]

variables are impulse response \( h_1, \ldots, h_{2n-1} \);
linear inequality constraints involve only \( h_1, \ldots, h_n \)
Euclidean distance matrix (EDM)

\[ D \in \mathbb{R}^{n \times n} \text{ is EDM if there are } x_1, \ldots, x_n \in \mathbb{R}^r \text{ s.t. } D_{ij} = \|x_i - x_j\|^2 \]

\( r \) is called embedding dimension

[Schoenberg '35] \( D = D^T \in \mathbb{R}^{n \times n} \) is EDM with embedding dimension \( r \) iff

- \( D_{ii} = 0 \),
- \( VDV \preceq 0 \), where \( V = I - \frac{1}{n} \mathbf{1} \mathbf{1}^T \),
- \( \text{Rank} \ VDV \leq r \)
Minimum embedding dimension

find EDM $D$ with minimum embedding dimension satisfying distance bounds

\[
\text{minimize} \quad \text{Rank } VDV \\
\text{subject to} \quad D_{ii} = 0 \quad i = 1, \ldots, n \\
VDV \preceq 0 \\
L_{ij} \leq D_{ij} \leq U_{ij} \quad i, j = 1, \ldots, n.
\]

applications:

- statistics/psychometrics (multidimensional scaling)
- chemistry (molecular conformation)
Solution methods

- analytical solutions for special cases (e.g., using SVD)
- global optimization (e.g., branch & bound) for small problems
- heuristic methods:
  - alternating projections [Grigoriadis & Beran '00]
  - factorization methods, e.g., alternating LMIs [Iwasaki '99]
  - analytic anti-centering [David '94]

an especially simple yet effective heuristic: trace heuristic
Trace heuristic for PSD matrices

observation: for $X = X^T \succeq 0$, minimizing trace tends to give low rank solution [Mesbahi '97, Pare '00, ...]

RMP

minimize $\text{Rank } X$
subject to $X \in C$

Trace heuristic

minimize $\text{Tr } X$
subject to $X \in C$

- convex problem, hence efficiently solved

variation: weighted trace minimization ($W = W^T > 0$)

minimize $\text{Tr } WX$
subject to $X \in C$
Log-det heuristic for PSD matrices

suggested heuristic: for $X = X^T \succeq 0$, (locally) minimize $\log \det (X + \delta I)$
($\delta > 0$ is small constant for regularization)

RMP

\begin{align*}
\text{minimize} & \quad \text{Rank } X \\
\text{subject to} & \quad X \in C
\end{align*}

Log-det heuristic

\begin{align*}
\text{minimize} & \quad \log \det (X + \delta I) \\
\text{subject to} & \quad X \in C
\end{align*}

- objective is nonconvex (in fact, concave)
- can use any local optimization method
Idea behind log-det heuristic

\[
\text{Rank } X = \sum_{i=1}^{n} 1(\lambda_i > 0) \quad \log \det(X + \delta I) = \sum_{i=1}^{n} \log(\lambda_i + \delta)
\]
Iterative linearization method

linearize (concave) objective at $X_k \succ 0$:

$$\log \det(X + \delta I) \approx \log \det(X_k + \delta I) + \text{Tr}(X_k + \delta I)^{-1}(X - X_k)$$

minimize linearized objective (a convex problem):

$$X_{k+1} = \arg\min_{X \in \mathcal{C}} \text{Tr}(X_k + \delta I)^{-1}X$$

i.e., iterative weighted trace minimization

- with $X_0 = I$, first iteration same as trace heuristic
- in practice only a few iterations needed (about 5 or 6)
Semidefinite embedding

**question:** can we extend trace & log-det heuristics to general (nonsquare, non PSD) matrices?

**Lemma:** let $X \in \mathbb{R}^{m \times n}$ then $\text{Rank } X \leq r$ iff there are $Y = Y^T \in \mathbb{R}^{m \times m}$, $Z = Z^T \in \mathbb{R}^{n \times n}$, s.t.

$$\text{Rank} \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} \leq 2r, \quad \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0.$$ 

thus, can embed general (non PSD) RMP in a (larger) PSD RMP
**RMP in embedded PSD form**

recall general RMP \( (X \in \mathbb{R}^{m \times n}) \)

\[
\begin{align*}
\text{minimize} & \quad \text{Rank } X \\
\text{subject to} & \quad X \in \mathcal{C}
\end{align*}
\]

equivalent to PSD RMP

\[
\begin{align*}
\text{minimize} & \quad \text{Rank } \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} \\
\text{subject to} & \quad \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\
& \quad X \in \mathcal{C},
\end{align*}
\]

with variables \( X \in \mathbb{R}^{m \times n}, Y = Y^T \in \mathbb{R}^{m \times m}, Z = Z^T \in \mathbb{R}^{n \times n} \)

can now apply any method for symmetric PSD RMP

ACC 2003
Trace heuristic for general matrices

\[
\begin{align*}
\text{minimize} & \quad \text{Tr} \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} \\
\text{subject to} & \quad \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\
& \quad X \in C
\end{align*}
\]

can show this is equivalent to

\[
\begin{align*}
\text{minimize} & \quad \|X\|_* \\
\text{subject to} & \quad X \in C,
\end{align*}
\]

where \( \|X\|_* = \sum_{i=1}^n \sigma_i(X) \), called nuclear norm of \( X \), is dual of spectral (maximum singular value) norm.
Convex envelope

**convex envelope** of $f : C \rightarrow \mathbb{R}$ is largest convex function $g$ s.t.
$g(x) \leq f(x)$ for all $x \in C$

- ‘best’ convex lower approximation
- epigraph of $g$ is convex hull of epigraph of $f$
Convex envelope of rank

**Theorem:** $\|X\|_*$ is cvx envelope of $\text{Rank} \ X$ on $\{X \in \mathbb{R}^{m \times n} \mid \|X\| \leq 1\}$.

[Fazel, Hindi, Boyd ’01]

**Conclusions:**

- trace heuristic minimizes convex envelope of rank (i.e., the best convex approximation to rank) over unit ball in matrix norm
- hence, heuristic gives lower bound on objective
- provides theoretical support for use of trace/nuclear norm heuristic
Log-det heuristic for general matrices

minimize \[ \log \det \left( \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} + \delta I \right) \]

subject to \[ \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \preceq 0 \]

\[ X \in \mathcal{C} \]

can linearize as before to obtain iterations in \( X, Y, Z \)

- each iteration minimizes a weighted trace cost function
Iterative $\ell_1$ heuristic for maximum sparsity problem

log-det heuristic for maximum sparsity problem yields

\[
\begin{align*}
\text{minimize} & \quad \sum_i \log(|x_i| + \delta) \\
\text{subject to} & \quad x \in C.
\end{align*}
\]

iterative linearization/minimization yields

\[
x^{(k+1)} = \arg\min_{x \in C} \sum_{i=1}^n w_i^{(k)} |x_i|, \quad w_i^{(k)} = \frac{1}{|x_i^{(k)}| + \delta}
\]

- each step is weighted $\ell_1$ norm minimization
- when $x_i^{(k)}$ small, weight in next step is large; hence, small entries in $x$ are pushed towards zero (subject to $x \in C$)
Example: minimum order system realization with step response constraints

find minimum order system that satisfies $l_i \leq s_i \leq u_i$, $i = 1, \ldots, 16$

- $n = 16$
- trace/nuclear norm heuristic yields rank 5
- log-det heuristic converges in 5 steps, yields rank 4

step response (solid) and specs (dashed)
Conclusions

- RMP is difficult nonconvex problem with many applications; maximum sparsity problem is special case

- for PSD matrices, log-det heuristic is an improvement over trace heuristic

- using semidefinite embedding, general RMP can be cast as a PSD RMP (thus, trace and log-det heuristics for PSD RMP are extended to general case)

- generalization of trace heuristic to general matrices is **nuclear norm**, which is convex envelope of rank