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)

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X \in \mathcal{C}, \end{array}$$

$X \in \mathbf{R}^{m \times n}$ 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 $X = \mathbf{diag}(x)$

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

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

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

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, . . .)

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank} \begin{bmatrix} h_1 & h_2 & \cdots & h_n \\ h_2 & h_3 & \cdots & h_{n+1} \\ \vdots & \vdots & & \vdots \\ h_n & h_{n+1} & \cdots & h_{2n-1} \end{bmatrix} \\ \text{subject to} & F(h_1, \ldots, h_n) \preceq g \end{array}$$

variables are impulse response $h_1, \ldots, h_{2n-1}$;
linear inequality constraints involve only $h_1, \ldots, h_n$

Euclidean distance matrix (EDM)

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

r is called **embedding dimension**

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

- $D_{ii} = 0$,
- $VDV \preceq 0$, where $V = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$,
- $\mathbf{Rank}\, VDV \leq r$

Minimum embedding dimension

find EDM D with minimum embedding dimension satisfying distance bounds

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, VDV \\ \text{subject to} & D_{ii} = 0 \qquad i = 1, \ldots, n \\ & VDV \preceq 0 \\ & L_{ij} \leq D_{ij} \leq U_{ij} \qquad i, j = 1, \ldots, n. \end{array}$$

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		Trace heuristic	
minimize	$\mathbf{Rank}\, X$	minimize	$\mathbf{Tr}\, X$
subject to	$X \in \mathcal{C}$	subject to	$X \in \mathcal{C}$

- **convex** problem, hence efficiently solved

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

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, WX \\ \text{subject to} & X \in \mathcal{C} \end{array}$$

Log-det heuristic for PSD matrices

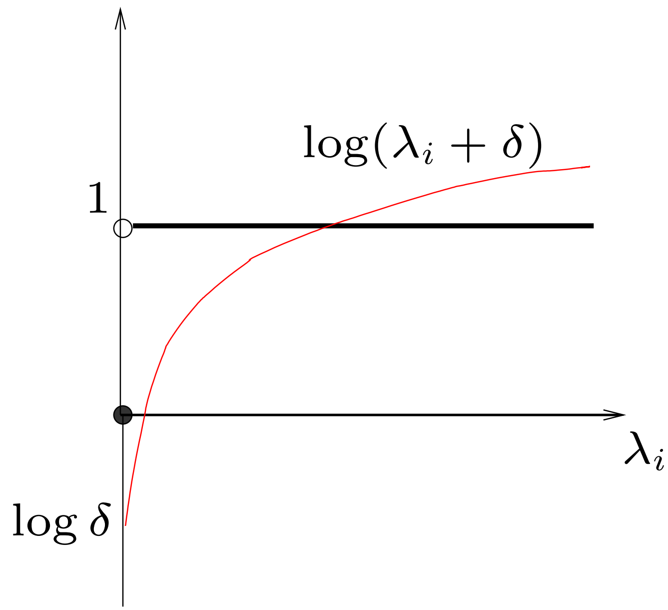
suggested heuristic: for $X = X^T \succeq 0$, (locally) minimize $\log\det(X + \delta I)$
($\delta > 0$ is small constant for regularization)

RMP	Log-det heuristic
minimize $\mathbf{Rank}\, X$ subject to $X \in \mathcal{C}$	minimize $\log\det(X + \delta I)$ subject to $X \in \mathcal{C}$

- objective is **nonconvex** (in fact, concave)
- can use any local optimization method

Idea behind log-det heuristic

$$\mathbf{Rank}\, X = \sum_{i=1}^{n} 1(\lambda_i > 0) \qquad \log\det(X + \delta I) = \sum_{i=1}^{n} \log(\lambda_i + \delta)$$



Iterative linearization method

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

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

minimize linearized objective (a convex problem):

$$X_{k+1} = \underset{X \in \mathcal{C}}{\operatorname{argmin}} \, \mathbf{Tr}(X_k + \delta I)^{-1} X$$

$i.e.$, 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 $X \in \mathbf{R}^{m \times n}$ then $\mathbf{Rank}\, X \leq r$ iff there are $Y = Y^T \in \mathbf{R}^{m \times m}$, $Z = Z^T \in \mathbf{R}^{n \times n}$, s.t.

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

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

RMP in embedded PSD form

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

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X \in \mathcal{C} \end{array}$$

equivalent to PSD RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank} \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\ & X \in \mathcal{C}, \end{array}$$

with variables $X \in \mathbf{R}^{m \times n}$, $Y = Y^T \in \mathbf{R}^{m \times m}$, $Z = Z^T \in \mathbf{R}^{n \times n}$

can now apply any method for symmetric PSD RMP

Trace heuristic for general matrices

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr} \begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\ & X \in \mathcal{C} \end{array}$$

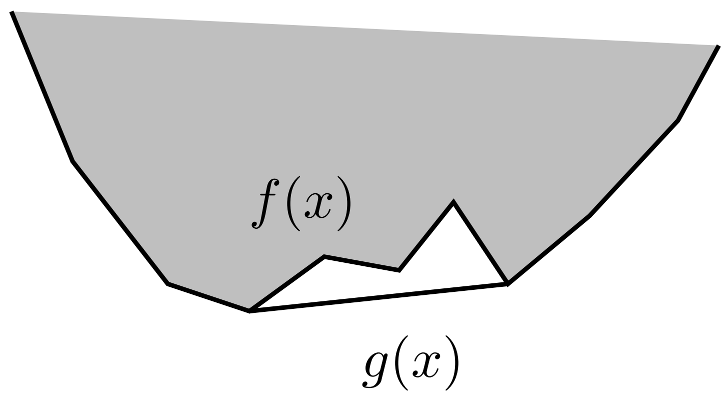
can show this is equivalent to

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & X \in \mathcal{C}, \end{array}$$

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 $f : C \to \mathbf{R}$ is largest convex function g s.t. $g(x) \le f(x)$ 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 $\mathbf{Rank}\, X$ on $\{X \in \mathbf{R}^{m\times n} \mid \|X\| \leq 1\}$.
[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

$$\begin{array}{ll} \text{minimize} & \log\det\left(\begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix} + \delta I\right) \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\ & X \in \mathcal{C} \end{array}$$

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

- each iteration minimizes a *weighted trace* cost function

Iterative ℓ_1 heuristic for maximum sparsity problem

log-det heuristic for maximum sparsity problem yields

$$\begin{array}{ll} \text{minimize} & \sum_i \log(|x_i| + \delta) \\ \text{subject to} & x \in \mathcal{C}. \end{array}$$

iterative linearization/minimization yields

$$x^{(k+1)} = \underset{x \in \mathcal{C}}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^{(k)} |x_i|, \qquad w_i^{(k)} = \frac{1}{|x_i^{(k)}| + \delta}$$

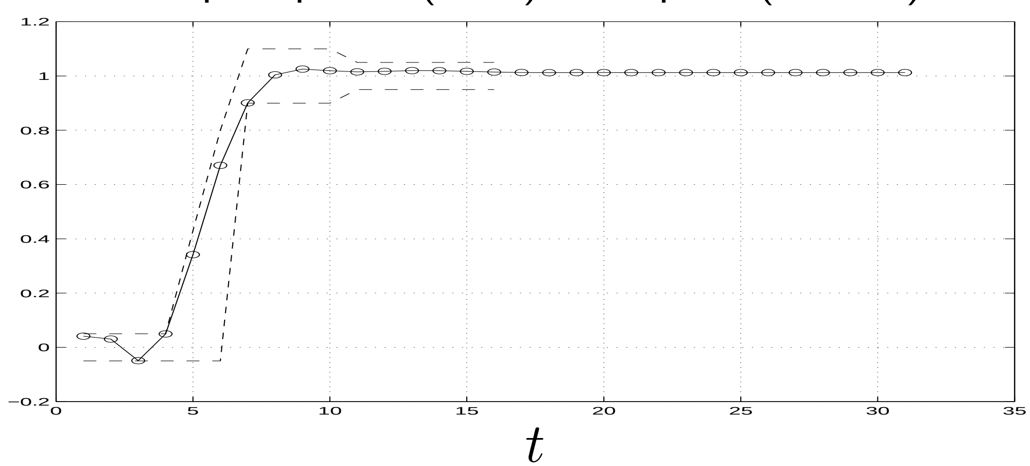
- each step is **weighted ℓ_1 norm** minimization
- when $x_i^{(k)}$ small, weight in next step is large; hence, small entries in x are pushed towards zero (subject to $x \in \mathcal{C}$)

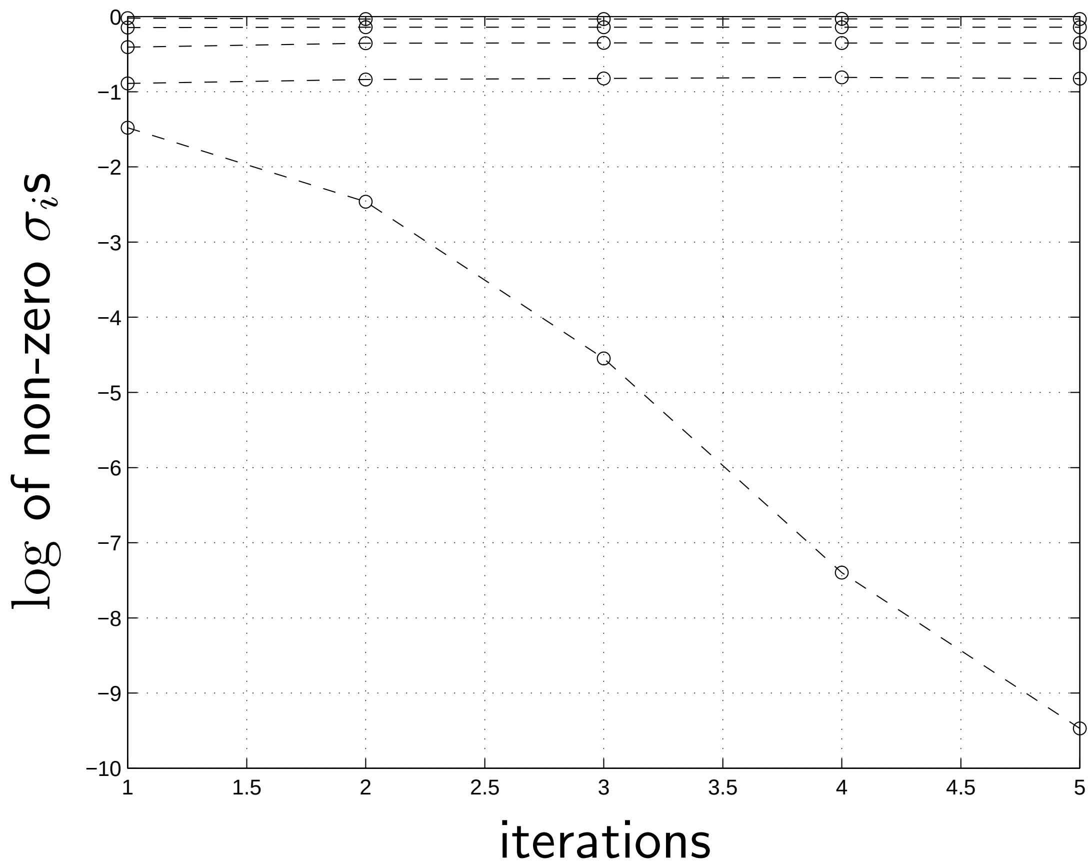
Example: minimum order system realization with step response constraints

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)





0
−1
−2
−3
−4
−5
−6
−7
−8
−9
−10
1
1.5
2
2.5
3
3.5
4
4.5
5
log of non-zero σ_is
iterations

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