Syllabus

• Objectives:
  ◦ To provide some preliminaries.
  ◦ To treat some mathematics.
  ◦ To point out some applications.
  ◦ To describe some algorithms.

• Topics:
  ◦ Lecture I: Introduction
  ◦ Lecture II: General Approach
  ◦ Lecture III: Distance Geometry and Protein Structure
  ◦ Lecture IV: Singular Value Assignment with Low Rank Matrices
  ◦ Lecture V: Nonnegative Matrix Factorization

• Assignments:
  ◦ Quite a few open questions to be answered.
  ◦ Try out various existing optimization codes on large scale low rank approximation problems.
Lecture II:

General Approach

Joint work with Robert Funderlic and Robert Plemmons
Outline

• Problem Description

• Algebraic Structure:
  ◦ Algebraic Varieties
  ◦ Rank Deficient 3 × 3 Toeplitz Matrices

• Constructing Lower Rank Structured Matrices:
  ◦ Lift and Project Method
  ◦ Parameterization by SVD

• Implicit Optimization
  ◦ Engineers’ Misconception
  ◦ Simplex Search Method

• Explicit Optimization
  ◦ fmincon in MATLAB
  ◦ LANCELOT on NEOS
Structure Preserving Rank Reduction Problem

• Given
  ◦ A target matrix $A \in \mathbb{R}^{n \times n}$,
  ◦ An integer $k$, $1 \leq k < \text{rank}(A)$,
  ◦ A class of matrices $\Omega$ with a specified structure,
  ◦ a fixed matrix norm $\| \cdot \|$;

Find
  ◦ A matrix $\hat{B} \in \Omega$ of rank $k$, such that

$$\|A - \hat{B}\| = \min_{B \in \Omega, \text{rank}(B) = k} \|A - B\|.$$  \hfill (1)
Difficulties

- No easy way to characterize, either algebraically or analytically, a given class of structured lower rank matrices.
- Lack of explicit description of the feasible set → Difficult to apply classical optimization techniques.
- Little discussion on whether lower rank matrices with specified structure actually exist.
Feasibility and Approximations

• The Toeplitz matrix

\[
H := \begin{bmatrix}
h_n & h_{n+1} & \ldots & h_{2n-1} \\
\vdots & \vdots & & \\
h_2 & h_3 & \ldots & h_{n+1} \\
h_1 & h_2 & \ldots & h_n
\end{bmatrix}
\]

with

\[
h_j := \sum_{i=1}^{k} \beta_i z_i^j, \quad j = 1, 2, \ldots, 2n - 1,
\]

where \(\{\beta_i\}\) and \(\{z_i\}\) are two sequences of arbitrary nonzero numbers satisfying \(z_i \neq z_j\) whenever \(i \neq j\) and \(k \leq n\), is a Toeplitz matrix of rank \(k\).

• The general Toeplitz structure preserving rank reduction problem as described in (1) remains open.
  
  ◦ Existence of lower rank matrices of specified structure does not guarantee closest such matrices.
  ◦ No \(x > 0\) for which \(1/x\) is minimum.
Other Structures?

- For other types of structures, the existence question usually is a hard algebraic problem.
- Given real general matrices $B_0, B_1, \ldots, B_n \in \mathbb{R}^{m \times n}$, $m \geq n$, and an integer $k < n$,
  - Open Question: Can values of $c := (c_1, \ldots, c_n)^\top \in \mathbb{R}^n$ be found such that
    $$B(c) := B_0 + c_1B_1 + \ldots + c_nB_n$$
    is of rank $k$ precisely?
  - Or, $B(c)$ has a prescribed set of singular values $\{\sigma_1, \ldots, \sigma_n\}$. 
Another Hidden Catch

- The set of all $n \times n$ matrices with rank $\leq k$ is a closed set.

- The approximation problem

$$\min_{B \in \Omega, \text{rank}(B) \leq k} \|A - B\|$$

is always solvable, so long as the feasible set is non-empty.

- The rank condition is to be less than or equal to $k$, but not necessarily exactly equal to $k$.

- It is possible that a given target matrix $A$ does not have a nearest rank $k$ structured matrix approximation, but does have a nearest structured matrix approximation of rank $k - 1$ or lower.
Our Approach

• Introduce two procedures to tackle the structure preserving rank reduction problem numerically.
• The procedures can be applied to problems of any norm, any linear structure, and any matrix norm.
• Use the symmetric Toeplitz structure with Frobenius matrix norm to illustrate the ideas.
Some other approaches

• (van der Veen’96) Given \( A \in \mathbb{R}^{m \times n} \) which is known to have \( k \) singular values less than \( \epsilon \), find all rank-\( k \) matrices \( B \in \mathbb{R}^{m \times n} \) such that

\[
\| A - B \|_2 < \epsilon.
\]

△ Not seeking the best approximation, only the one in the \( \epsilon \)-neighborhood of \( A \).
△ No structure involved.
△ **Open Question:** Can it be done this way for structured matrices?

• (Manton, Mahony, and Hua’03) Consider the weighted low rank approximation

\[
\min_{B \in \mathbb{R}^{m \times n}, \text{rank}(B) \leq k} \| A - B \|_Q^2,
\]

where

\[
\| X \|_Q^2 = \text{vec}(X)^\top Q \text{vec}(X)
\]

and \( Q \in \mathbb{R}^{mn \times mn} \) is a SPD matrix.

△ Reformulate the minimization as

\[
\min_{N \in \mathbb{R}^{n \times (n-k)}, N^\top N = I} \left( \min_{B \in \mathbb{R}^{m \times n}, BN = 0} \| A - B \|_Q^2 \right).
\]

A quadratic programming problem

△ (Schuermans, Lemmerling, Van Huffel’03) Using a modified operator \( \text{vec}_2 \) to dictate the underlying linear structure.

• (Frieze, Kannan, and Vempala’98) Monte-Carlo algorithm for finding a matrix \( B^* \) of rank at most \( k \) so that

\[
\| A - B^* \|_F \leq \min_{B \in \mathbb{R}^{m \times n}, \text{rank}(B) \leq k} \| A - B \|_F + \epsilon \| A \|_F
\]

holds with probability \( 1 - \delta \).

△ The algorithm takes time polynomial in \( k, 1/\epsilon \) and \( \log(1/\delta) \) only, and is independent of \( m, n \).
△ **Open question:** Can a structure be built in? With what probability?
Representing Low Rank Toeplitz Matrices

- Identify a *symmetric* Toeplitz matrix by its first row,

$$T = T([t_1, \ldots, t_n]) = \begin{bmatrix}
    t_1 & t_2 & \cdots & t_n \\
    t_2 & t_1 & \ddots & t_{n-1} \\
    \vdots & \ddots & \ddots & \ddots \\
    t_{n-1} & t_{n-2} & \cdots & t_2 \\
    t_n & t_{n-1} & \cdots & t_2 & t_1
\end{bmatrix}.$$  

○ $T$ = The affine subspace of all $n \times n$ symmetric Toeplitz matrices.

- Spectral decomposition of symmetric rank $k$ matrices:

$$M = \sum_{i=1}^{k} \alpha_i y^{(i)} y^{(i)\top}. \quad (3)$$

- Write $T = T([t_1, \ldots, t_n])$ in terms of (3) $\implies$

$$\sum_{i=1}^{k} \alpha_i y_j^{(i)} y_{j+s}^{(i)} = t_{s+1}, \; s = 0, 1, \ldots, n - 1, \; 1 \leq j \leq n - s \quad (4)$$

○ Low rank matrices form an *algebraic variety*, i.e, solutions of polynomial systems.
Some Examples

- The case $k = 1$ is trivial.
  - Rank-one Toeplitz matrices form two simple one-parameter families,
    \[
    T = \alpha_1 T([1, \ldots, 1]), \quad \text{or} \quad T = \alpha_1 T([1, -1, 1, \ldots, (-1)^{n-1}])
    \]
    with arbitrary $\alpha_1 \neq 0$.

- For $4 \times 4$ symmetric Toeplitz matrices of rank 2, there are 10 unknowns in 6 equations (by dropping the references to $t_1, \ldots, t_4$).

  \[
  \begin{align*}
  \alpha_1 & := \frac{\alpha_2 (y_1^{(2)^2} - y_2^{(2)^2})}{-\alpha_2 (y_1^{(2)^2} + y_2^{(2)^2})}, \\
  y_3^{(1)} & := \frac{y_2^{(2)} y_1^{(2)} y_1^{(1)} y_2^{(1)} - y_2^{(1)} y_1^{(1)} y_2^{(2)} - y_2^{(1)} y_1^{(1)^2} y_2^{(2)} + \alpha_1 y_2^{(1)^2} y_1^{(1)} y_2^{(2)} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(2)} + 3 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} + 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2}}, \\
  y_4^{(1)} & := \frac{y_2^{(1)} y_1^{(2)} y_1^{(2)} y_2^{(1)} - y_2^{(1)} y_1^{(1)^2} y_2^{(2)} - y_2^{(1)} y_1^{(1)} y_2^{(2)} - y_2^{(1)} y_1^{(1)^2} y_2^{(2)} + \alpha_1 y_2^{(1)^2} y_1^{(1)} y_2^{(2)} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(2)} + 3 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} + 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2} - 2 y_2^{(1)} y_1^{(1)^2} y_2^{(1)^2}}, \\
  y_3^{(2)} & := \frac{y_2^{(2)^2} y_1^{(2)} y_1^{(2)} y_2^{(2)} - y_2^{(2)^2} y_1^{(2)} y_1^{(2)^2} y_2^{(2)} - y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + \alpha_1 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + 3 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)}}, \\
  y_4^{(2)} & := \frac{y_2^{(2)^2} y_1^{(2)^2} y_1^{(2)^2} y_2^{(2)} - y_2^{(2)^2} y_1^{(2)^2} y_1^{(2)^2} y_2^{(2)} - y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + \alpha_1 y_2^{(2)^2} y_2^{(2)^2} y_2^{(2)^2} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + 3 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} + 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)} - 2 y_2^{(2)^2} y_1^{(2)^2} y_2^{(2)}},
  \end{align*}
  \]

- The eigstructure of symmetric and centrosymmetric matrices has a special parity property, but that has not been taken into account.
- Explicit description of algebraic equations for higher dimensional low rank symmetric Toeplitz matrices becomes unbearably complicated.
About Uniqueness

- Consider rank deficient $T([t_1, t_2, t_3])$
  - $\det(T) = (t_1 - t_3)(t_1^2 + t_1t_3 - 2t_2^2) = 0$.
  - A union of two algebraic varieties.

![Figure 1: Low rank, symmetric, Toeplitz matrices of dimension 3 identified in $\mathbb{R}^3$.](image)

- The number of *local* solutions to the structured lower rank approximation problem is not unique.
Dimensionality

- (Adamjan, Arov and Krein’71) Suppose the underlying matrices are of infinite dimension. Then the closest approximation to a Hankel matrix by a low rank Hankel matrix always exists and is unique.
Constructing Lower Rank Toeplitz Matrices

• Idea:
  ◊ Rank $k$ matrices in $R^{n\times n}$ form a surface $\mathcal{R}(k)$.
  ◊ Rank $k$ Toeplitz matrices = $\mathcal{R}(k) \cap \mathcal{T}$.

• Two approaches:
  ◊ Parameterization by SVD:
    ▶ Identify $M \in \mathcal{R}(k)$ by the triplet $(U, \Sigma, V)$ of its singular value decomposition $M = U\Sigma V^\top$.
    · $U$ and $V$ are orthogonal matrices, and
    · $\Sigma = \text{diag}\{s_1, \ldots, s_k, 0, \ldots, 0\}$ with $s_1 \geq \ldots \geq s_k > 0$.
    ▶ Enforce the structure.
  ◊ Alternate projections between $\mathcal{R}(k)$ and $\mathcal{T}$ to find intersections. (Cheney & Goldstein’59, Catzow’88)
Lift and Project Algorithm

- Given $A^{(0)} = A$, repeat projections until convergence:
  - **LIFT.** Compute $B^{(ν)} \in \mathcal{R}(k)$ nearest to $A^{(ν)}$:
    - From $A^{(ν)} \in \mathcal{T}$, first compute its SVD
      \[ A^{(ν)} = U^{(ν)} \Sigma^{(ν)} V^{(ν)\top}. \]
    - Replace $\Sigma^{(ν)}$ by $\text{diag}\{ s_1^{(ν)}, \ldots, s_k^{(ν)}, 0, \ldots, 0 \}$ and define
      \[ B^{(ν)} := U^{(ν)} \Sigma^{(ν)} V^{(ν)\top}. \]
  - **PROJECT.** Compute $A^{(ν+1)} \in \mathcal{T}$ nearest to $B^{(ν)}$:
    - From $B^{(ν)}$, choose $A^{(ν+1)}$ to be the matrix formed by replacing the diagonals of $B^{(ν)}$ by the averages of their entries.

- The general approach remains applicable to any other linear structure, and symmetry can be enforced.
  - The only thing that needs to be modified is the projection in the projection (second) step.
Figure 2: Lift-and-project algorithm with intersection of low rank matrices and Toeplitz matrices
Black-box Function

• Descent property:

\[ \|A^{(\nu+1)} - B^{(\nu+1)}\|_F \leq \|A^{(\nu+1)} - B^{(\nu)}\|_F \leq \|A^{(\nu)} - B^{(\nu)}\|_F. \]

○ Descent with respect to the Frobenius norm which is not necessarily the norm used in the structure preserving rank reduction problem.

• If all \( A^{(\nu)} \) are distinct then the iteration converges to a Toeplitz matrix of rank \( k \).

○ In principle, the iteration could be trapped in an impasse where \( A^{(\nu)} \) and \( B^{(\nu)} \) would not improve any more, but not experienced in practice.

• The lift and project iteration provides a means to define a black-box function

\[ P : T \rightarrow T \cap R(k). \]

○ The \( P(T) \) is presumably piecewise continuous since all projections are continuous.
Consider $P : \mathbb{R}^2 \rightarrow \mathbb{R}^2$:

- Use the $xy$-plane to represent the domain of $P$ for $2 \times 2$ symmetric Toeplitz matrices $T(t_1, t_2)$.
- Use the $z$-axis to represent the image $p_{11}(T)$ and $p_{12}(T)$, respectively.

![Graph of $P(T)$ for 2-dimensional symmetric Toeplitz $T$.](image_url)

Figure 3: Graph of $P(T)$ for 2-dimensional symmetric Toeplitz $T$.

- Toeplitz matrices of the form $T(t_1, 0)$ or $T(0, t_2)$, corresponding to points on axes, converge to the zero matrix.
Implicit Optimization

- Implicit formulation:

\[
\min_{T=\text{toeplitz}(t_1, \ldots, t_n)} \| T_0 - P(T) \|.
\] (5)

△ $T_0$ is the given target matrix.

△ $P(T)$, regarded as a black box function evaluation, provides a handle to manipulate the objective function $f(T) := \| T_0 - P(T) \|$.

△ The norm used in (5) can be any matrix norm.

- Engineers’ misconception:

△ $P(T)$ is not necessarily the closest rank $k$ Toeplitz matrix to $T$.

△ In practice, $P(T_0)$ has been used “as a cleansing process whereby any corrupting noise, measurement distortion or theoretical mismatch present in the given data set (namely, $T_0$) is removed.”

△ More needs to be done in order to find the closest lower rank Toeplitz approximation to the given $T_0$ as $P(T_0)$ is merely known to be in the feasible set.
Numerical Experiment

• An ad hoc optimization technique:
  ◊ The simplex search method by Nelder and Mead requires only function evaluations.
  ◊ Routine \texttt{fminsearch} in MATLAB, employing the simplex search method, is ready for use in our application.

• An example:
  ◊ Suppose \( T_0 = T(1, 2, 3, 4, 5, 6) \).
  ◊ Start with \( T^{(0)} = T_0 \), and set worst case precision to \( 10^{-6} \).
  ◊ Able to calculate all lower rank matrices while maintaining the symmetric Toeplitz structure. Always so?
  ◊ Nearly machine-zero of smallest calculated singular value(s) \( \implies T^*_k \) is computationally of rank \( k \).
  ◊ \( T^*_k \) is only a local solution.
  ◊ \( \|T^*_k - T_0\| < \|P(T_0) - T_0\| \) which, though represents only a slight improvement, clearly indicates that \( P(T_0) \) alone does not give rise to an optimal solution.
<table>
<thead>
<tr>
<th>rank $k$</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td># of iterations</td>
<td>110</td>
<td>81</td>
<td>46</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td># of SVD calls</td>
<td>1881</td>
<td>4782</td>
<td>2585</td>
<td>2294</td>
<td>558</td>
</tr>
<tr>
<td>optimal solution</td>
<td>$\begin{bmatrix} 1.1046 \ 1.8880 \ 3.1045 \ 3.9106 \ 5.0635 \ 5.9697 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 1.2408 \ 1.8030 \ 3.0352 \ 4.1132 \ 4.8553 \ 6.0759 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 1.4128 \ 1.7980 \ 2.8171 \ 4.1089 \ 5.2156 \ 5.7450 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 1.9591 \ 2.1059 \ 2.5683 \ 3.4157 \ 4.7749 \ 6.8497 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2.9444 \ 2.9444 \ 2.9444 \ 2.9444 \ 2.9444 \ 2.9444 \end{bmatrix}$</td>
</tr>
<tr>
<td>$|T_0 - T_k^*|$</td>
<td>0.5868</td>
<td>0.9851</td>
<td>1.4440</td>
<td>3.2890</td>
<td>8.5959</td>
</tr>
<tr>
<td>singular values</td>
<td>$\begin{bmatrix} 17.9851 \ 7.4557 \ 2.2866 \ 0.9989 \ 0.6164 \ 3.4638e-15 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 17.9980 \ 7.4321 \ 2.2836 \ 0.8376 \ 2.2454e-14 \ 2.0130e-14 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 18.0125 \ 7.4135 \ 2.1222 \ 1.9865e-14 \ 9.0753e-15 \ 6.5255e-15 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 18.2486 \ 6.4939 \ 2.0884e-14 \ 7.5607e-15 \ 3.8479e-15 \ 2.5896e-15 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 17.6667 \ 2.0828e-14 \ 9.8954e-15 \ 6.0286e-15 \ 2.6494e-15 \ 2.1171e-15 \end{bmatrix}$</td>
</tr>
</tbody>
</table>

Table 1: Test results for a case of $n = 6$ symmetric Toeplitz structure
Explicit Optimization

- Difficult to compute the gradient of $P(T)$.
- Other ways to parameterize structured lower rank matrices:
  - Use eigenvalues and eigenvectors for symmetric matrices;
  - Use singular values and singular vectors for general matrices.
  - Robust, but might have *overdetermined* the problem.
An Illustration

- Define

\[ M(\alpha_1, \ldots, \alpha_k, y^{(1)}, \ldots, y^{(k)}) := \sum_{i=1}^{k} \alpha_i y^{(i)} y^{(i)\top}. \]

- Reformulate the symmetric Toeplitz structure preserving rank reduction problem explicitly as

\[
\begin{align*}
\min & \quad \|T_0 - M(\alpha_1, \ldots, \alpha_k, y^{(1)}, \ldots, y^{(k)})\| \\
\text{subject to} & \quad m_{j, j+s-1} = m_{1, s}, \\
& \quad s = 1, \ldots, n-1, \\
& \quad j = 2, \ldots, n-s+1,
\end{align*}
\]

if \( M = [m_{ij}] \).

- Objective function in (6) is described in terms of the non-zero eigenvalues \( \alpha_1, \ldots, \alpha_k \) and the corresponding eigenvectors \( y^{(1)}, \ldots, y^{(k)} \) of \( M \).
- Constraints in (7) are used to ensure that \( M \) is symmetric and Toeplitz.

- For other types of structures, we only need modify the constraint statement accordingly.

- The norm used in (6) can be arbitrary but is fixed.
Redundant Constraints

- Symmetric centro-symmetric matrices have special spectral properties (Cantoni and Butler’76):
  - $\lceil n/2 \rceil$ of the eigenvectors are symmetric; and
  - $\lfloor n/2 \rfloor$ are skew-symmetric.
  - $\uparrow \mathbf{v} = [v_i] \in \mathbb{R}^n$ is symmetric (or skew-symmetric) if $v_i = v_{n-i}$ (or $v_i = -v_{n-i}$).
- Symmetric Toeplitz matrices are symmetric and centro-symmetric.
- The formulation in (6) does not take this spectral structure of eigenvectors $y^{(i)}$ into account.
  - More variables than needed have been introduced.
  - May have overlooked any internal relationship among the $\frac{n(n-1)}{2}$ equality constraints.
  - May have caused, inadvertently, additional computation complexity.
Using fmincon in MATLAB

- Routine **fmincon** in MATLAB:
  - Uses a sequential quadratic programming method.
  - Solve the Kuhn-Tucker equations by a quasi-Newton updating procedure.
  - Can estimate derivative information by finite difference approximations.
  - Readily available in Optimization Toolbox.

- Our experiments:
  - Use the same data as in the implicit formulation.
  - Case $k = 5$ is computationally the same as before.
  - Have trouble in cases $k = 4$ or $k = 3$,
    - Iterations will not improve approximations at all.
    - MATLAB reports that the optimization is terminated successfully.
<table>
<thead>
<tr>
<th>f-Count</th>
<th>Function</th>
<th>MAX{g}</th>
<th>STEP</th>
<th>Procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>0.958964</td>
<td>8.65974e-15</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>0.958964</td>
<td>2.66454e-14</td>
<td>1.91e-06</td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>0.958964</td>
<td>2.70894e-14</td>
<td>2.98e-08</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>185</td>
<td>0.958964</td>
<td>2.70894e-14</td>
<td>2.98e-08</td>
<td></td>
</tr>
<tr>
<td>239</td>
<td>0.958964</td>
<td>2.73115e-14</td>
<td>2.98e-08</td>
<td></td>
</tr>
<tr>
<td>289</td>
<td>0.958964</td>
<td>2.77556e-14</td>
<td>4.77e-07</td>
<td></td>
</tr>
<tr>
<td>337</td>
<td>0.958964</td>
<td>2.77556e-14</td>
<td>1.91e-06</td>
<td></td>
</tr>
<tr>
<td>393</td>
<td>0.958964</td>
<td>2.77556e-14</td>
<td>7.45e-09</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>445</td>
<td>0.958964</td>
<td>5.28466e-14</td>
<td>1.19e-07</td>
<td></td>
</tr>
<tr>
<td>501</td>
<td>0.958964</td>
<td>5.68434e-14</td>
<td>7.45e-09</td>
<td></td>
</tr>
<tr>
<td>557</td>
<td>0.958964</td>
<td>5.70655e-14</td>
<td>7.45e-09</td>
<td>Hessian not updated</td>
</tr>
<tr>
<td>613</td>
<td>0.958964</td>
<td>5.66214e-14</td>
<td>7.45e-09</td>
<td></td>
</tr>
<tr>
<td>667</td>
<td>0.958964</td>
<td>5.55112e-14</td>
<td>2.98e-08</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>713</td>
<td>0.958964</td>
<td>3.17302e-13</td>
<td>7.63e-06</td>
<td></td>
</tr>
<tr>
<td>761</td>
<td>0.958964</td>
<td>2.61569e-13</td>
<td>1.91e-06</td>
<td></td>
</tr>
<tr>
<td>812</td>
<td>0.958964</td>
<td>2.60014e-13</td>
<td>-2.38e-07</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>856</td>
<td>0.958964</td>
<td>2.57794e-13</td>
<td>3.05e-05</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>900</td>
<td>0.958964</td>
<td>2.56462e-13</td>
<td>3.05e-05</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>948</td>
<td>0.958964</td>
<td>2.57128e-13</td>
<td>1.91e-06</td>
<td></td>
</tr>
<tr>
<td>994</td>
<td>0.958964</td>
<td>2.56684e-13</td>
<td>7.63e-06</td>
<td></td>
</tr>
<tr>
<td>1038</td>
<td>0.958964</td>
<td>3.42837e-13</td>
<td>3.05e-05</td>
<td></td>
</tr>
<tr>
<td>1083</td>
<td>0.958964</td>
<td>3.41727e-13</td>
<td>-1.53e-05</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1124</td>
<td>0.958964</td>
<td>3.92575e-13</td>
<td>0.000244</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1161</td>
<td>0.958964</td>
<td>5.04485e-13</td>
<td>0.00391</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1200</td>
<td>0.958964</td>
<td>5.12923e-13</td>
<td>0.000977</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1233</td>
<td>0.958964</td>
<td>5.61551e-13</td>
<td>0.0625</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1272</td>
<td>0.958964</td>
<td>5.86642e-13</td>
<td>0.000977</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1308</td>
<td>0.958964</td>
<td>4.84279e-13</td>
<td>0.00781</td>
<td>Hessian modified twice</td>
</tr>
<tr>
<td>1309</td>
<td>0.958964</td>
<td>4.84723e-13</td>
<td>1</td>
<td>Hessian modified twice</td>
</tr>
</tbody>
</table>

Optimization Converged Successfully
Using LANCELOT on NEOS

• Reasons of failure of MATLAB are not clear.
  ◦ Constraints might no longer be linearly independent.
  ◦ Termination criteria in `fmincon` might not be adequate.
  ◦ Difficult geometry means hard-to-satisfy constraints.

• Using more sophisticated optimization packages, such as LANCELOT.
  ◦ A standard Fortran 77 package for solving large-scale nonlinearly constrained optimization problems (Conn, Could, and Toint’92).
  ◦ Break down the functions into sums of element functions to introduce sparse Hessian matrix.
  ◦ Huge code. See
    [http://www.rl.ac.uk/departments/ced/numerical/lancelot/sif/sifhtml.html](http://www.rl.ac.uk/departments/ced/numerical/lancelot/sif/sifhtml.html)
  ◦ Available on the NEOS Server through a socket-based interface.
  ◦ Uses the ADIFOR automatic differentiation tool.
• **LANCELOT** works, so far.
  
  ◦ Find optimal solutions of problem (6) for all values of $k$.
  ◦ Results from **LANCELOT** agree, up to the required accuracy $10^{-6}$, with those from **fmins**.
  ◦ Rank affects the computational cost nonlinearly.

<table>
<thead>
<tr>
<th>rank $k$</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td># of variables</td>
<td>35</td>
<td>28</td>
<td>21</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td># of f/c calls</td>
<td>108</td>
<td>56</td>
<td>47</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>total time</td>
<td>12.99</td>
<td>4.850</td>
<td>3.120</td>
<td>1.280</td>
<td>.4300</td>
</tr>
</tbody>
</table>

Table 3: Cost overhead in using **LANCELOT** for $n = 6$.

• It is not clear whether the **LANCELOT** would run into the same trouble as **fmincon** when applied to larger size problems.
• There are many other algorithms available in NEOS.
Conclusions

- Structure preserving rank reduction problems arise in many important applications, particularly in the broad areas of signal and image processing.
- Constructing the nearest approximation of a given matrix by one with any rank and any linear structure is difficult in general.
- We have proposed two ways to formulate the problems as standard optimization computations.
- It is now possible to tackle the problems numerically via utilizing standard optimization packages.
- The ideas were illustrated by considering Toeplitz structure with Frobenius norm.
- Our approach can be readily generalized to consider rank reduction problems for any given linear structure and of any given matrix norm.