Chapter 8

Structured Low Rank Approximation

- Overview
- Low Rank Toeplitz Approximation
- Low Rank Circulant Approximation
- Low Rank Covariance Approximation
- Eculidean Distance Matrix Approximation
- Approximate GCD

Overview

- Given
  - ◇ A target matrix $A \in R^{n \times n}$,
  - ◇ An integer k, $1 \leq k < \mathrm{rank}(A)$,
  - ◇ A class of matrices Ω with linear structure,
  - ◇ a fixed matrix norm $\| \cdot \|$;

  Find
  - ◇ A matrix $\hat{B} \in \Omega$ of rank k, and
  - ◇

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

- Example of linear structure:
  - ◇ Toeplitz or block Toeplitz matrices.
  - ◇ Hankel or banded matrices.
- Applications:
  - ◇ Signal and image processing with Toeplitz structure.
  - ◇ Model reduction problem in speech encoding and filter design with Hankel structure.
  - ◇ Regularization of ill-posed inverse problems.

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 $\Longrightarrow$ Difficult to apply classical optimization techniques.
- Little discussion on whether lower rank matrices with specified structure actually exist.

An Example of Existence

- Physics sometimes sheds additional light.
- The Toeplitz matrix

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

  with

$$h_j := \sum_{i=1}^{k} \beta_i z_i^j, \quad j = 1, 2, \dots, 2n-1,$$

  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.

- For other types of structures, the existence question usually is a hard algebraic problem.

Another Hidden Catch

- The set of all $n \times n$ matrices with rank $\leq k$ is a *closed* set.
- The approximation problem
$$\min_{B \in \Omega, \mathrm{rank}(B) \leq k} \|A - B\|$$
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 rank $k-1$ or lower structured matrix approximation.

Low Rank Toeplitz Approximation

- Algebraic Structure of Low Rank Toeplitz Matrices.
- Constructing Low Rank Toeplitz Matrices.
  - ◇ Lift and Project Method
  - ◇ Parameterization by SVD
- Implicit Optimization
  - ◇ Engineerers' Misconception
  - ◇ Simplex Search Method
- Explicit Optimization
  - ◇ **constr** in MATLAB
  - ◇ **LANCELOT** on NEOS

A General Remark

- 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.

Algebraic Structure

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

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

  - ◇ $\mathcal{T}$ = The affine subspace of all $n \times n$ symmetric Toeplitz matrices.

- Spectral decomposition of symmetric rank k matrices:

$$M = \sum_{i=1}^{k} \alpha_i y^{(i)} {y^{(i)}}^T. \tag{2}$$

- Write $T = T([t_1, \ldots, t_n])$ in terms of (2) $\Longrightarrow$

$$\sum_{i=1}^{k} \alpha_i y_j^{(i)} y_{j+s}^{(i)} = t_{s+1},\ s = 0, 1, \ldots, n-2,\ 1 \le j \le n-s \tag{3}$$

  - ◇ Lower 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,

$$\begin{aligned} T &= \alpha_1 T([1, \ldots, 1]), \quad \text{or} \\ T &= \alpha_1 T([1, -1, 1, \ldots, (-1)^{n-1}]) \end{aligned}$$

    with arbitrary $\alpha_1 \neq 0$.

- For 4×4 symmetric Toeplitz matrices of rank 2, there are 10 unknowns in 6 equations.

$$\left\{ \begin{array}{lcl} \alpha_1 & := & \frac{\alpha_2 \left({y_1^{(2)}}^2 - {y_2^{(2)}}^2 \right)}{-{y_1^{(1)}}^2 + {y_2^{(1)}}^2}, \\ y_3^{(1)} & := & \frac{y_2^{(1)} y_1^{(2)} y_1^{(1)} + 2\, y_2^{(2)} {y_2^{(1)}}^2 - y_2^{(2)} {y_1^{(1)}}^2}{y_2^{(1)} y_1^{(2)} + y_1^{(1)} y_2^{(2)}}, \\ y_4^{(1)} & := & -\frac{{y_2^{(1)}}^3 {y_1^{(2)}}^2 - 4\, {y_2^{(1)}}^3 {y_2^{(2)}}^2 - 4\, y_1^{(1)} y_1^{(2)} y_2^{(2)} {y_2^{(1)}}^2 - 2\, y_2^{(1)} {y_1^{(1)}}^2 {y_1^{(2)}}^2 + 3\, y_2^{(1)} {y_2^{(2)}}^2 {y_1^{(1)}}^2 + 2\, y_1^{(2)} y_2^{(2)} {y_1^{(1)}}^3}{{y_2^{(1)}}^2 {y_1^{(2)}}^2 + 2\, y_2^{(1)} y_1^{(2)} y_1^{(1)} y_2^{(2)} + {y_1^{(1)}}^2 {y_2^{(2)}}^2}, \\ y_3^{(2)} & := & -\frac{y_2^{(1)} {y_1^{(2)}}^2 - 2\, y_2^{(1)} {y_2^{(2)}}^2 - y_1^{(2)} y_2^{(2)} y_1^{(1)}}{y_2^{(1)} y_1^{(2)} + y_1^{(1)} y_2^{(2)}}, \\ y_4^{(2)} & := & -\frac{3\, {y_2^{(1)}}^2 {y_1^{(2)}}^2 y_2^{(2)} - 4\, {y_2^{(1)}}^2 {y_2^{(2)}}^3 + 2\, y_2^{(1)} y_1^{(1)} {y_1^{(2)}}^3 - 4\, y_2^{(1)} y_1^{(1)} {y_2^{(2)}}^2 y_1^{(2)} - 2\, y_2^{(2)} {y_1^{(1)}}^2 {y_1^{(2)}}^2 + {y_1^{(1)}}^2 {y_2^{(2)}}^3}{{y_2^{(1)}}^2 {y_1^{(2)}}^2 + 2\, y_2^{(1)} y_1^{(2)} y_1^{(1)} y_2^{(2)} + {y_1^{(1)}}^2 {y_2^{(2)}}^2}. \end{array} \right.$$

  - ◇ Explicit description of algebraic equations for higher dimensional lower rank symmetric Toeplitz matrices becomes unbearably complicated.

Let's See It!

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

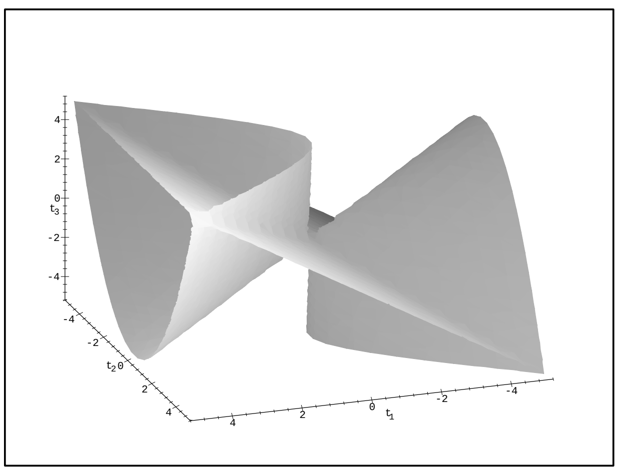


Figure 1: Lower rank, symmetric, Toeplitz matrices of dimension 3 identified in R^3.

- The number of *local* solutions to the structured lower rank approximation problem is not 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 $= \mathcal{R}(k)\bigcap\mathcal{T}$.
- Two approaches:
  - ◇ Parameterization by SVD:
    - ▷ Identify $M \in \mathcal{R}(k)$ by the triplet (U, Σ, V) of its singular value decomposition $M = U\Sigma V^T$.
      - · U and V are orthogonal matrices, and
      - · $\Sigma = \mathrm{diag}\{s_1, \ldots, s_k, 0, \ldots, 0\}$ 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 $B^{(\nu)} \in \mathcal{R}(k)$ nearest to $A^{(\nu)}$:
    - ▷ From $A^{(\nu)} \in \mathcal{T}$, first compute its SVD

      $$A^{(\nu)} = U^{(\nu)}\Sigma^{(\nu)}{V^{(\nu)}}^T.$$

    - ▷ Replace $\Sigma^{(\nu)}$ by $\operatorname{diag}\{s_1^{(\nu)}, \ldots, s_k^{(\nu)}, 0, \ldots, 0\}$ and define

      $$B^{(\nu)} := U^{(\nu)}\Sigma^{(\nu)}{V^{(\nu)}}^T.$$

  - ◇ **PROJECT**. Compute $A^{(\nu+1)} \in \mathcal{T}$ nearest to $B^{(\nu)}$:
    - ▷ From $B^{(\nu)}$, choose $A^{(\nu+1)}$ to be the matrix formed by replacing the diagonals of $B^{(\nu)}$ 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.

Geometric Sketch

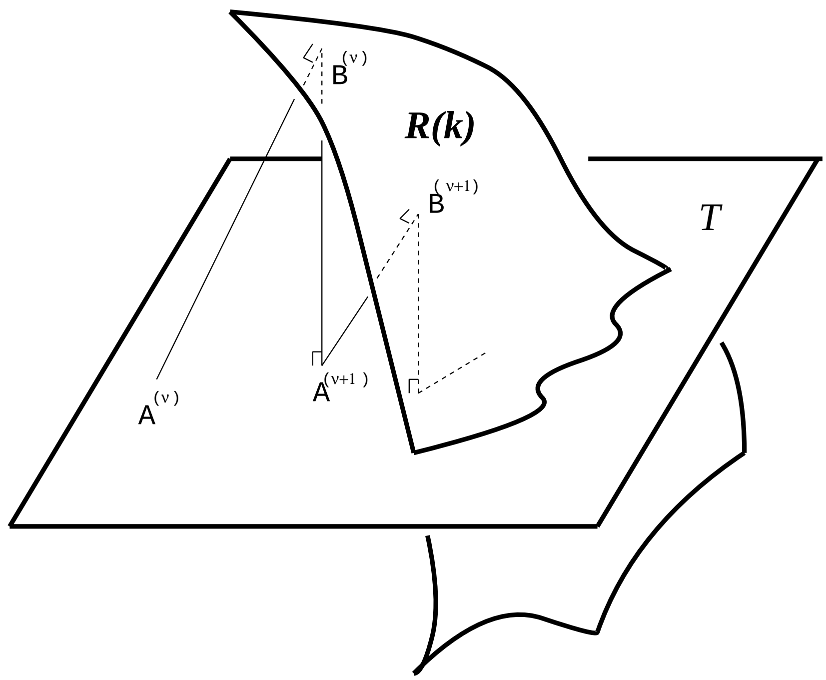


Figure 2: Algorithm 1 with intersection of lower rank matrices and Toeplitz matrices

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*

  $$P : \mathcal{T} \longrightarrow \mathcal{T} \bigcap \mathcal{R}(k).$$

  ◇ The $P(T)$ is *presumably* piecewise continuous since all projections are continuous.

The graph of $P(T)$

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

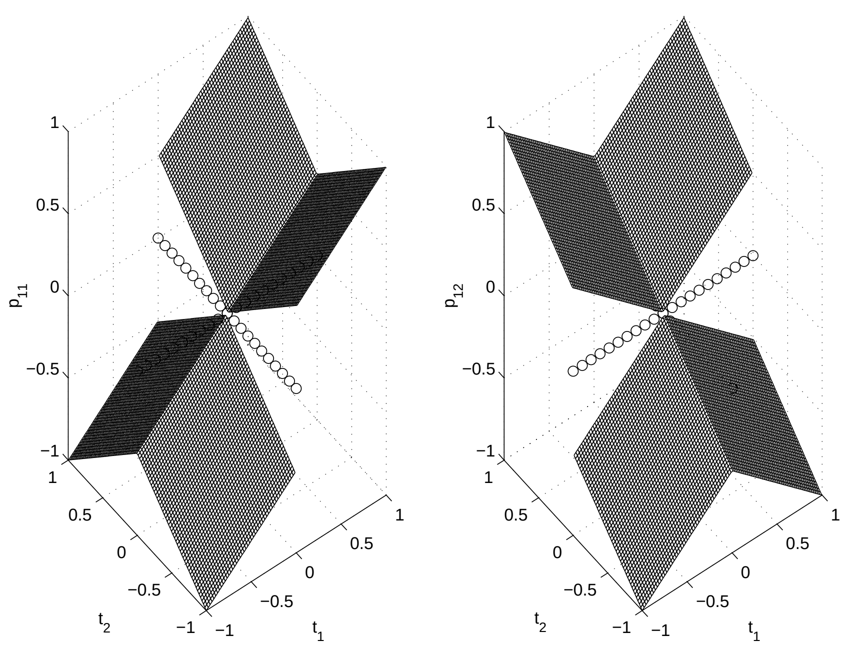


Figure 3: Graph of $P(T)$ for 2-dimensional symmetric Toeplitz T.

- 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:

$$\min_{T=\text{toeplitz}(t_1,\dots,t_n)} \|T_0 - P(T)\|. \tag{4}$$

  - ◇ 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 (4) 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 **fmins** 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) $\Longrightarrow T_k^*$ 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.

rank k	5	4	3	2	1
# of iterations	110	81	46	36	17
# of SVD calls	1881	4782	2585	2294	558
optimal solution	$\begin{bmatrix}1.1046\\1.8880\\3.1045\\3.9106\\5.0635\\5.9697\end{bmatrix}$	$\begin{bmatrix}1.2408\\1.8030\\3.0352\\4.1132\\4.8553\\6.0759\end{bmatrix}$	$\begin{bmatrix}1.4128\\1.7980\\2.8171\\4.1089\\5.2156\\5.7450\end{bmatrix}$	$\begin{bmatrix}1.9591\\2.1059\\2.5683\\3.4157\\4.7749\\6.8497\end{bmatrix}$	$\begin{bmatrix}2.9444\\2.9444\\2.9444\\2.9444\\2.9444\\2.9444\end{bmatrix}$
$\|T_0 - T_k^*\|$	0.5868	0.9851	1.4440	3.2890	8.5959
singular values	$\begin{bmatrix}17.9851\\7.4557\\2.2866\\0.9989\\0.6164\\3.4638e{-}15\end{bmatrix}$	$\begin{bmatrix}17.9980\\7.4321\\2.2836\\0.8376\\2.2454e{-}14\\2.0130e{-}14\end{bmatrix}$	$\begin{bmatrix}18.0125\\7.4135\\2.1222\\1.9865e{-}14\\9.0753e{-}15\\6.5255e{-}15\end{bmatrix}$	$\begin{bmatrix}18.2486\\6.4939\\2.0884e{-}14\\7.5607e{-}15\\3.8479e{-}15\\2.5896e{-}15\end{bmatrix}$	$\begin{bmatrix}17.6667\\2.0828e{-}14\\9.8954e{-}15\\6.0286e{-}15\\2.6494e{-}15\\2.1171e{-}15\end{bmatrix}$

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

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

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

$$\begin{array}{ll} \min & \|T_0 - M(\alpha_1, \ldots, \alpha_k, y^{(1)}, \ldots, y^{(k)})\| \quad (5) \\ \text{subject to} & m_{j,j+s-1} = m_{1,s}, \quad (6) \\ & s = 1, \ldots n-1, \\ & j = 2, \ldots, n-s+1, \end{array}$$

  if $M = [m_{ij}]$.

  ◇ Objective function in (5) 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 (6) 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 (5) can be arbitrary but is fixed.

Redundant Constraints

- Symmetric centro-symmetric matrices have special spectral properties:
  - ◇ $\lceil n/2 \rceil$ of the eigenvectors are symmetric; and
  - ◇ $\lfloor n/2 \rfloor$ are skew-symmetric.
    - ▷ $v = [v_i] \in R^n$ 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 (5) does not take this spectral structure into account in the eigenvectors $y^{(i)}$.
  - ◇ 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 Existing Optimization Codes

- Using **constr** in MATLAB
  - ◇ Routine **constr** 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.

- Using **LANCELOT** on NEOS
  - ◇ Reasons of failure of MATLAB are not clear.
    - ▷ Constraints might no longer be linearly independent.
    - ▷ Termination criteria in **constr** 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.
    - ▷ Break down the functions into sums of *element functions* to introduce sparse Hessian matrix.
    - ▷ Huge code. See

      http://www.rl.ac.uk/departments/ccd/numerical/lancelot/sif/sifhtml.html.

    - ▷ Available on the NEOS Server through a socket-based interface.
    - ▷ Uses the **ADIFOR** automatic differentiation tool.

◇ **LANCELOT** works.

▷ Find optimal solutions of problem (5) 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.

rank k	5	4	3	2	1
# of variables	35	28	21	14	7
# of f/c calls	108	56	47	43	19
total time	12.99	4.850	3.120	1.280	.4300

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

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.

Low Rank Circulant Approximation

- Basic Properties
- (Inverse) Eigenvalue Problem
- Conjugate Evenness
- Low Rank Approximation
- Tree Structure
- Numerical Experiment

Basic Properties

- A circulant matrix $C = Circul(c)$

$$C = \begin{bmatrix} c_0 & c_1 & & \dots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \dots & c_{n-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_1 & c_2 & & c_{n-1} & c_0 \end{bmatrix}$$

  is uniquely determined by the first row c.

  ◇ Define

$$\Pi := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & & \dots & 0 \end{bmatrix}.$$

  Then

$$Circul(c) = \sum_{k=0}^{n-1} c_k \Pi^k = P_c(\Pi)$$

  with characteristic polynomial

$$P_c(x) = \sum_{k=0}^{n-1} c_k x^k.$$

Elementary Spectral Properties

- Define

$$\Omega := \mathrm{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}), \quad \omega := \exp(\frac{2\pi i}{n}).$$

- Define the Fourier matrix F where

$$F^* := \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \ldots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2n-2} \\ \vdots & & & & \vdots \\ 1 & \omega^{n-1} & \omega^{n-2} & \ldots & \omega \end{bmatrix}.$$

  - ◇ F is unitary.

- The forward shift matrix Π is unitarily diagonalizable.

$$\Pi = F^*\Omega F.$$

- The circulant matrix $Circul(c)$ with any given row vector c has a spectral decomposition

$$Circul(c) = F^* P_c(\Omega) F.$$

- Spectral properties:
  - ◇ Closely related to the discrete Fourier transform.
  - ◇ Explicit solution for eigenvalue and inverse eigenvalue problems.
  - ◇ FFT calculation.

(Inverse) Eigenvalue Problem

- Forward problem:
  - ◇ Eigenvalues of $Circul(c)$:
  $$\lambda = [P_c(1), \ldots P_c(\omega^{n-1})].$$
  - ◇ Can be computed from
  $$\lambda^T = \sqrt{n} F^* c^T.$$
- Inverse problem:
  - ◇ Given any vector $\lambda := [\lambda_0, \ldots, \lambda_{n-1}] \in C^n$, define
  $$c^T = \frac{1}{\sqrt{n}} F \lambda^T.$$
  - ◇ $Circul(c)$ has eigenvalues in vector λ.
- Both matrix-vector products involved done in $O(n \log_2 n)$ flops.
- If all the eigenvalues are distinct, there are precisely $n!$ distinct circulant matrices with the prescribed spectrum.

Conjugate Evenness

- $c^T = \frac{1}{\sqrt{n}} F \lambda^T$ is real if and only if $\lambda^T = \sqrt{n} F^* c^T$ is conjugate-even.
  - ◇ If $n = 2m$,

    $$\lambda = [\lambda_0, \lambda_1, \dots, \lambda_{m-1}, \lambda_m, \overline{\lambda_{m-1}}, \dots, \overline{\lambda_1}].$$

    - ▷ $\lambda_0, \lambda_m \in \mathbb{R}$. (Absolutely real, others real or complex.)
  - ◇ If $n = 2m + 1$,

    $$\lambda := [\lambda_0, \lambda_1, \dots, \lambda_m, \overline{\lambda_m}, \dots, \overline{\lambda_1}].$$

    - ▷ $\lambda_0 \in \mathbb{R}$. (Absolutely real.)
- Singular value decomposition of $Circul(c)$:

  $$Circul(c) = (F^* P_c(\Omega) |P_c(\Omega)|^{-1}) |P_c(\Omega)| F$$

  - ◇ Singular values are $|P_c(\omega^k)|, k = 0, 1, \cdots n - 1$.
  - ◇ At most $\lceil \frac{n+1}{2} \rceil$ distinct singular values.

Low Rank Approximation

- Given $A \in R^{n\times n}$, its nearest circulant matrix approximation $Circul(c)$ is given by the projection (T. Chan)

$$c_k := \frac{1}{n}\langle A, \Pi^k\rangle, \quad k = 0, \ldots, n-1,$$

  - ◇ $Circul(c)$ is generally of **full rank** even if A has lower rank.

- How to reduce the rank of $Circul(c)$?
  - ◇ The truncated singular value decomposition (TSVD) gives rise to the nearest low rank approximation in Frobenius norm.
  - ◇ The TSVD of $Circul(c)$ is automatically circulant.
  - ◇ But: the TSVD can lead to a **complex** circulant approximation.

Trivial $O(n \log n)$ TSVD Algorithm

- Given a real matrix A and a fixed rank $\ell \leq n$,
  1. Use the projection to find the nearest real circulant matrix approximation $Circul(c)$ of A.
  2. Use the FFT to calculate the spectrum λ of the matrix $Circul(c)$.
  3. Arrange all elements of $|\lambda|$ in descending order, including those with equal modulus.
  4. Let $\hat{\lambda}$ be the vector consisting of elements of λ, but those corresponding to the last $n-\ell$ singular values in the descending order are set to zero.
  5. Apply the inverse FFT to $\hat{\lambda}$ to determine a nearest circulant matrix $Circul(\hat{c})$ of rank ℓ to A.
- The resulting matrix $Circul(\hat{c})$ is complex-valued in general.
  - ◇ Need to preserve the conjugate-even structure.
  - ◇ Need to modify the TSVD strategy.

Data Matching Problem

- The low rank "real" circulant approximation problem is equivalent to a data matching problem:

  (DMP) *Given a conjugate-even vector $\lambda \in C^n$, find its nearest conjugate-even approximation $\hat{\lambda} \in C^n$ subject to the constraint that $\hat{\lambda}$ has exactly $n - \ell$ zeros.*

- How to solve the DMP?
  - ◇ Write $\hat{\lambda} = [\hat{\lambda}_1, 0] \in C^n$ with $\hat{\lambda}_1 \in C^\ell$ being arbitrary.
  - ◇ Consider the problem of minimizing

    $$F(P, \hat{\lambda}) = \|P\hat{\lambda}^T - \lambda^T\|^2$$

    with a permutation matrix P.
    - ▷ P is used to search for the match.
  - ◇ Write $P = [P_1, P_2]$ with $P_1 \in R^{n \times \ell}$.

- ◇ A least squares problem:

$$F(P, \hat{\lambda}) = \|P_1 \hat{\lambda}_1^T - \lambda^T\|^2$$

- ◇ The optimal solution is

$$\hat{\lambda}_1 = \lambda P_1.$$

  - ▷ The entries of $\hat{\lambda}_1$ must be a portion of λ.

- ◇ The objective function becomes

$$F(P, \hat{\lambda}) = \|(P_1 P_1^T - I)\lambda\|^2.$$

  - ▷ $P_1 P_1^T - I$ is just a projection.
  - ▷ The optimal permutation P should be such that $P_1 P_1^T$ projects λ to its first ℓ components with largest modulus.

- Without the conjugate-even constraint, the answer to the data matching problem corresponds precisely to the usual TSVD selection criterion.

- With the conjugate-even constraint, the above criterion remains effective subject to the conjugate-even structure inside λ.

An Example using a Tree Structure

- Consider the case $n = 6$.
- Suppose λ_1, λ_2 are complex and distinct.
- Six possible conjugate-even structures.
- Tree graph:
  - ◇ Each node in the tree represents an element of λ.
  - ◇ Arrange the nodes from top to bottom in descending order of their moduli.
  - ◇ In case of a tie,
    - ▷ Complex conjugate nodes stay at the same level.
    - ▷ Real node is below the complex nodes.
- If $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$, then the tree is given by:

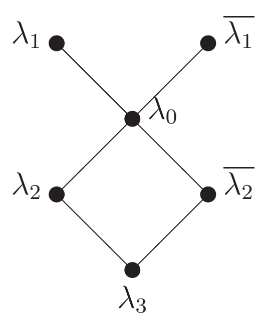
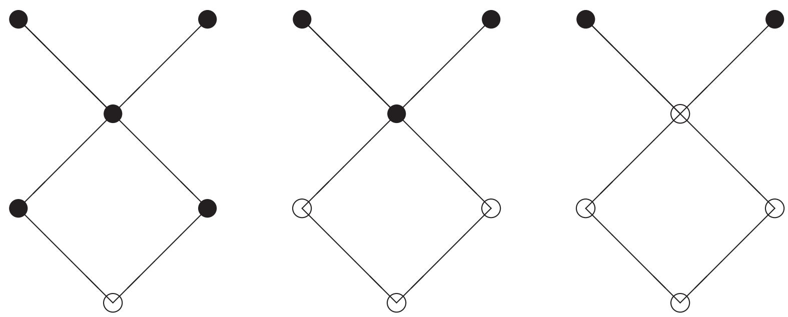


Figure 4: Tree graph of $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$ with $|\lambda_1| \geq |\lambda_0| > |\lambda_2| \geq |\lambda_3|$.



rank = 5 rank = 3 rank = 2

Figure 5: Tree graphs of $\hat{\lambda}$ with rank 5, 3, and 2.

$2|\lambda_2|^2 < |\lambda_0|^2 + |\lambda_3|^2$ $2|\lambda_2|^2 > |\lambda_0|^2 + |\lambda_3|^2$

Figure 6: Tree graphs of $\hat{\lambda}$ with rank 4.

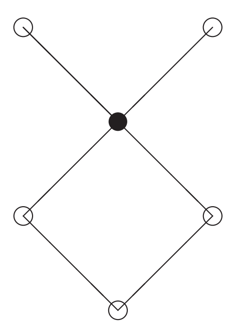


Figure 7: Tree graph of $\hat{\lambda}$ with rank 1.

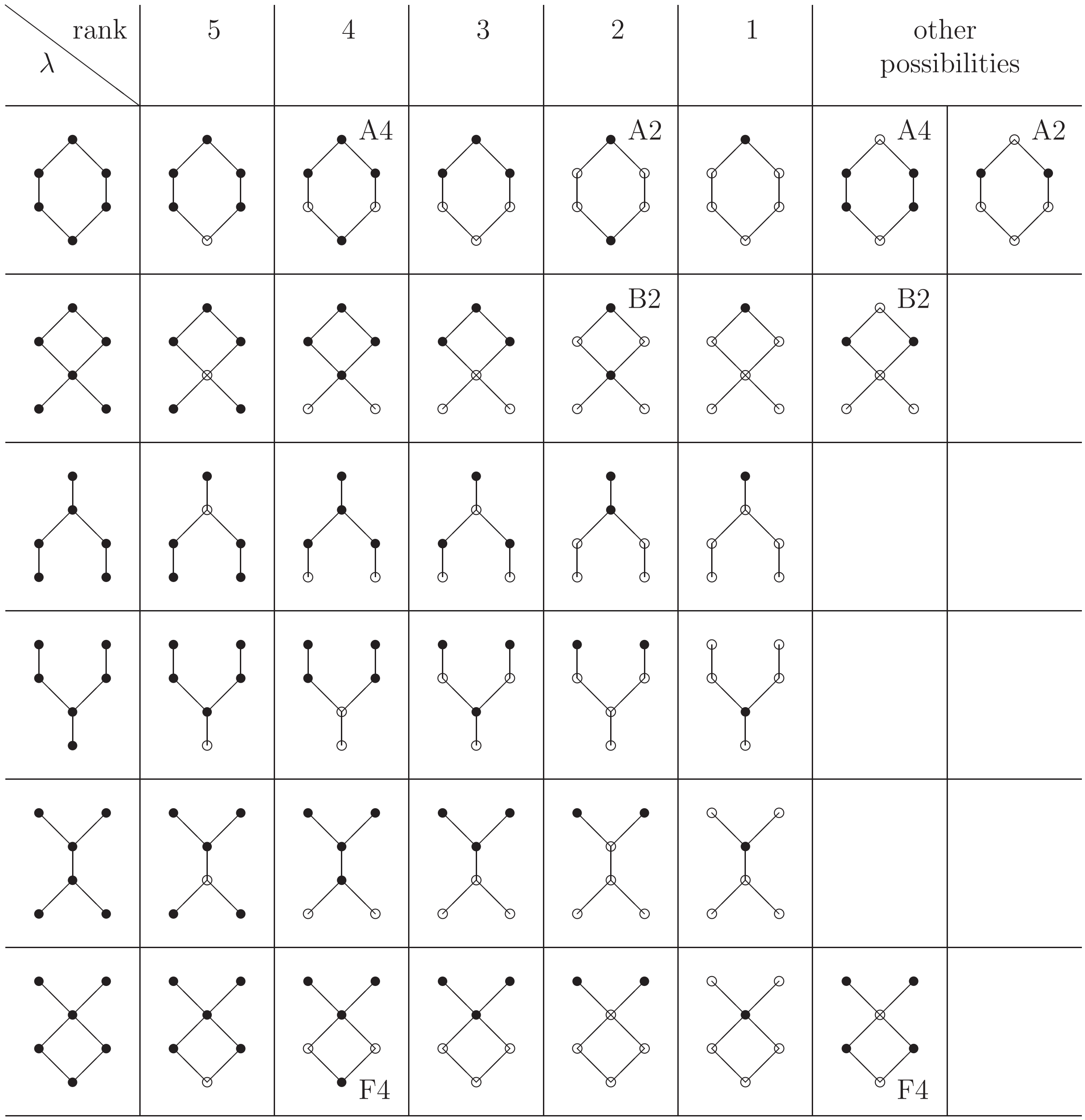


Figure 8: Possible solutions to the DMP when $n = 6$.

Numerical Experiment

- The low rank approximation algorithm needs to be smart enough to explore the conjugate-even structure, to truncate, and to reassemble the conjugate-even structure.
- Numerical complexity is $O(n \log n)$ flops.

Example 1: Symmetric Illustration

Consider the 8×8 symmetric $Circul(c)$:

$$c = \left[0.5404,\ 0.2794,\ 0.1801,\ \text{-}0.0253,\ \text{-}0.2178,\ \text{-}0.0253,\ 0.1801,\ 0.2794\right].$$

- Eigenvalues (in descending order):

$$\left[1.1909,\ 1.1892,\ 1.1892,\ 0.3273,\ 0.3273,\ \mathbf{0.1746},\ \text{-}0.0376,\ \text{-}0.0376\right]$$

- For rank 7 approximation, the usual TSVD would set -0.0376 to zero, resulting in a complex matrix.
- Use the conjugate-even eigenvalues

$$\hat{\lambda} = \left[1.1909,\ 1.1892,\ 0.3273,\ \text{-}0.0376,\ \mathbf{0}\ \text{-}0.0376,\ 0.3273,\ 1.1892\right],$$

  to obtain the best real-valued, rank 7, approximation $Circul(\hat{c})$ via the FFT:

$$\hat{c} = \left[0.5186,\ 0.3657,\ 0.0670,\ \text{-}0.0680,\ \text{-}0.0572,\ \text{-}0.0680,\ 0.0670,\ 0.3657\right].$$

- To obtain the best real-value, rank 4, circulant approximation, use eigenvalues $\hat{\lambda}$

$$\hat{\lambda} = \left[1.1909,\ 1.1892,\ 0,\ 0,\ 0.3273,\ 0,\ 0,\ 1.1892\right].$$

Example 2: Complex Illustration

Consider the 9×9 $Circul(c)$ with

$$c = \big[1.6864, 1.7775, 1.9324, 2.9399, 1.9871, 1.7367, 4.0563, 1.2848, 2.5989\big].$$

- Eigenvalues: structure given by

  $\big[$**20.0000**,

  -2.8130 + 1.9106i, -2.8130 - 1.9106i, 3.0239 - 1.0554i, 3.0239 + 1.0554i,

  -1.3997 + 0.7715i, -1.3997 - 0.7715i, -1.2223 - 0.2185i, -1.2223 + 0.2185i$\big]$.

- To obtain a real-valued circulant approximation of rank 8, we have no choice but to select the set the *largest* eigenvalue (singular value) of $Circul(c)$ to zero.
  - ◇ Setting the largest singular value to zero to obtain the nearest real low rank approximation is quite counter-intuitive to the usual sense of TSVD.
  - ◇ Apply algorithm to reduce the rank further, to 7.

Example 3: Perturbed Case

- Let $C_\kappa \in R^{n\times n}$ be a given circulant matrix of rank κ. Random noise added to C_κ will destroy the circulant structure as well as the rank condition.
- Let $E \in R^{n\times n}$ denote a random but fixed circulant matrix with unit Frobenius norm, and let

$$W_j = C_\kappa + 10^{-j}E, \quad j = 1, \ldots, 12.$$

- W_j will almost certainly be of full rank. Note that $\|W_j - C_\kappa\| = 10^{-j}$. It will be interesting to see if W_j has any closer circulant matrix approximation of rank κ other than C_κ, especially when j is large.
- Test case with $n = 100$, $\kappa = 73$, and a predetermined matrix C_{73}.
- Using our algorithm to find the best circulant approximation Z_j to W_j, we find that it is always the case that

$$\|W_j - Z_j\| < \|W_j - C_\kappa\|$$

for all j, i.e., our real circulant approximation is closest.

Conclusion

- For any given real data matrix, its nearest real circulant approximation can simply be determined from the average of its diagonals.
- The nearest low rank (possibly complex) approximation to the circulant matrix can be determined effectively from the TSVD and the FFT.
- To construct a real circulant matrix with specified spectrum, the eigenvalues must appear in conjugate-even form. So, the truncation criteria for a nearest low rank, real, circulant matrix approximation must be modified.
- We have proposed a fast algorithm with $O(n \log n)$ complexity to accomplish all of these objectives.

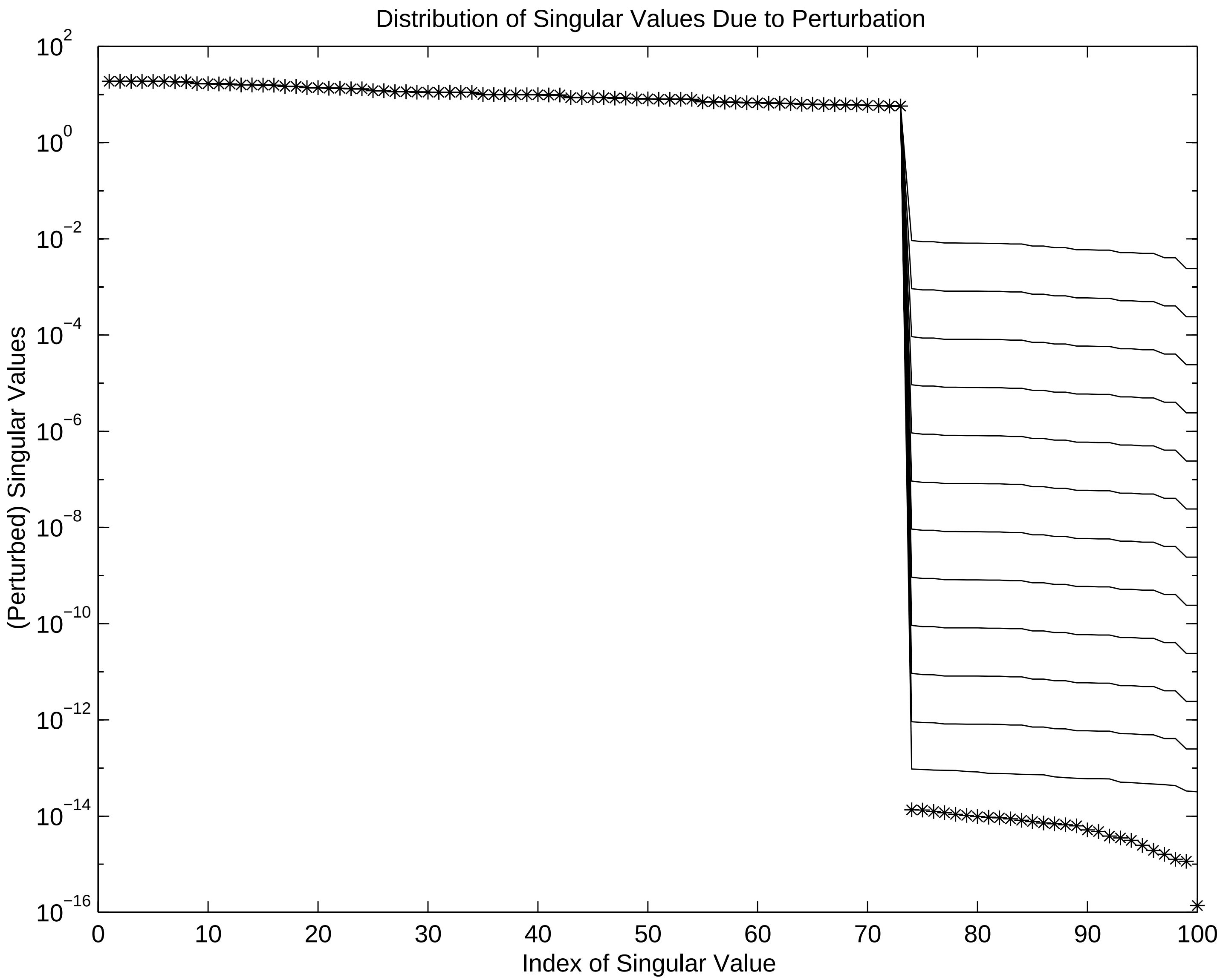


Figure 9: Distribution of Singular Values.

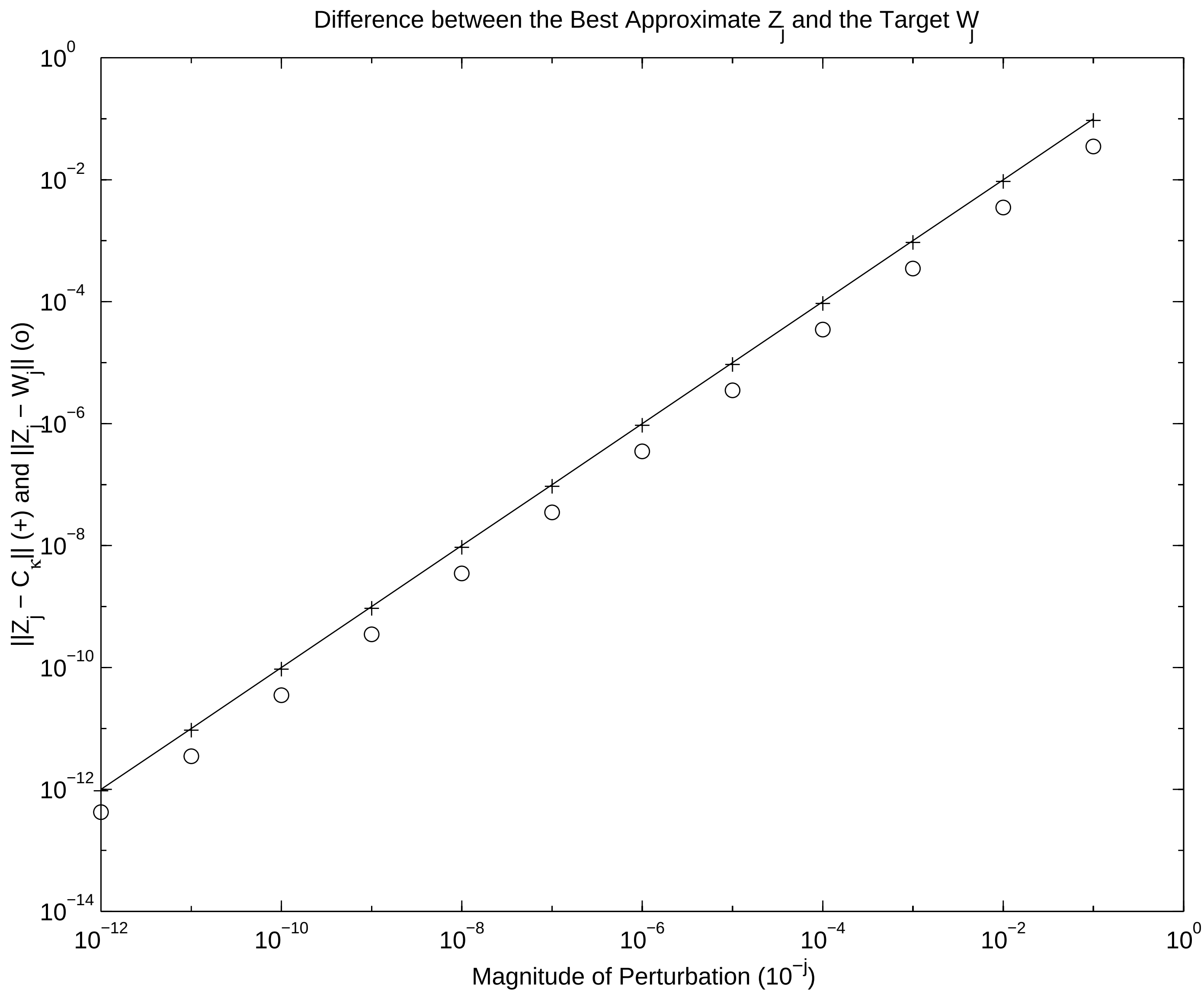


Figure 10: Errors in Approximation.

Lor Rank Covarance Approximation

Euclidian Distance Matrix Approximation

Approximate GCD

f-COUNT	FUNCTION	MAX{g}	STEP	Procedures
29	0.958964	8.65974e-15	1	
77	0.958964	2.66454e-14	1.91e-06	
131	0.958964	2.70894e-14	2.98e-08	Hessian modified twice
185	0.958964	2.70894e-14	2.98e-08	
239	0.958964	2.73115e-14	2.98e-08	
289	0.958964	2.77556e-14	4.77e-07	
337	0.958964	2.77556e-14	1.91e-06	
393	0.958964	2.77556e-14	7.45e-09	Hessian modified twice
445	0.958964	5.28466e-14	1.19e-07	
501	0.958964	5.68434e-14	7.45e-09	
557	0.958964	5.70655e-14	7.45e-09	Hessian not updated
613	0.958964	5.66214e-14	7.45e-09	
667	0.958964	5.55112e-14	2.98e-08	Hessian modified twice
713	0.958964	3.17302e-13	7.63e-06	
761	0.958964	2.61569e-13	1.91e-06	
812	0.958964	2.60014e-13	-2.38e-07	Hessian modified twice
856	0.958964	2.57794e-13	3.05e-05	Hessian modified twice
900	0.958964	2.56462e-13	3.05e-05	Hessian modified twice
948	0.958964	2.57128e-13	1.91e-06	
994	0.958964	2.56684e-13	7.63e-06	
1038	0.958964	3.42837e-13	3.05e-05	
1083	0.958964	3.41727e-13	-1.53e-05	Hessian modified twice
1124	0.958964	3.92575e-13	0.000244	Hessian modified twice
1161	0.958964	5.04485e-13	0.00391	Hessian modified twice
1200	0.958964	5.12923e-13	0.000977	Hessian modified twice
1233	0.958964	5.61551e-13	0.0625	Hessian modified twice
1272	0.958964	5.86642e-13	0.000977	Hessian modified twice
1308	0.958964	4.84279e-13	0.00781	Hessian modified twice
1309	0.958964	4.84723e-13	1	Hessian modified twice

Optimization Converged Successfully

Table 2: A typical output of intermediate results from **constr**.