Chapter 4

Structured Inverse Eigenvalue Problems

- Jacobi Inverse Eigenvalue Problems
- Toeplitz Inverse Eigenvalue Problem
- Nonnegative Inverse Eigenvalue Problem
- Stochastic Inverse Eigenvalue Problem
- Unitary Inverse Eigenvalue Problem
- Inverse Eigenvalue Problem with Prescribed Entries
- Inverse Singular Value Problems
- Inverse Singular/Eigenvalue Problem

Jacobi Inverse Eigenvalue Problems

- Overview.
- Subvariations.
- Existence Theory.
- Sensitivity Issues.
- Numerical Methods.

Overview

- Jacobi structure, i.e.,

$$J = \begin{bmatrix} a_1 & b_1 & 0 & & & 0 \\ b_1 & a_2 & b_2 & & & 0 \\ 0 & b_2 & a_3 & & & 0 \\ \vdots & & & \ddots & & \\ & & & & a_{n-1} & b_{n-1} \\ 0 & & & & b_{n-1} & a_n \end{bmatrix}, \quad b_i > 0,$$

  appears in many areas of applications.

  ◇ Oscillatory mass-spring systems.

  ◇ Composite pendulum.

  ◇ Sturm-Liouville problems.

- Jacobi IEP often can be solved by direct methods in finitely many steps.
- For symmetric tridiagonal matrices, there are $2n+1$ unknown entries to be determined. Thus there is in need of $2n+1$ pieces of information.
- For convenience, denote the leading principal submatrix of M by $\bar{M}$.

Subvariations

- (SIEP6a) [41, 98, 153, 164, 175, 193, 197]:
  - ◇ Given
    - ▷ Real scalars $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_1^*, \ldots, \mu_{n-1}^*\}$,
    - ▷ Interlacing property:

$$\lambda_i^* \leq \mu_i^* \leq \lambda_{i+1}^*, \quad i = 1, \ldots, n-1,$$

  - ◇ Find a Jacobi matrix J such that

$$\begin{aligned} \sigma(J) &= \{\lambda_k^*\}_{k=1}^n \\ \sigma(\bar{J}) &= \{\mu_1^*, \ldots, \mu_{n-1}^*\}. \end{aligned}$$

- (SIEP2) [40, 98, 193].
  - ◇ Given
    - ▷ Real scalars $\{\lambda_k^*\}_{k=1}^n$,
  - ◇ Find a persymmetric Jacobi matrix J such that

$$\begin{aligned} \sigma(J) &= \{\lambda_k^*\}_{k=1}^n \\ a_i &= a_{n+1-i} \\ b_i &= b_{n+2-i}. \end{aligned}$$

- (SIEP6b) [289]:
  - ◇ Given
    - ▷ Complex and distinct scalars $\{\lambda_1^*, \ldots, \lambda_{2n}^*\}$ and $\{\mu_1^*, \ldots, \mu_{2n-2}^*\} \in \mathbb{C}$,
    - ▷ Closed with complex conjugation.
  - ◇ Find tridiagonal symmetric matrices C and K for the λ-matrix $Q(\lambda) = \lambda^2 I + \lambda C + K$ so that

$$\begin{aligned} \sigma(Q) &= \{\lambda_1^*, \ldots, \lambda_{2n}^*\}, \\ \sigma(\bar{Q}) &= \{\mu_1^*, \ldots, \mu_{2n-2}^*\}. \end{aligned}$$

    - ▷ Arising from damped oscillatory systems.
    - ▷ Open Question: A practical solution requires additional conditions, i.e., positive diagonal entries, negative off-diagonal entries, and are weakly diagonally dominant.

- (SIEP7) [40, 41, 129]:
  - ◇ Given
    - ▷ Real scalars $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_1^*, \dots, \mu_{n-1}^*\}$,
    - ▷ Satisfy the interlacing property,
    - ▷ A positive number β,
  - ◇ Find a periodic Jacobi matrix J of the form

$$J = \begin{bmatrix} a_1 & b_1 & & & & b_n \\ b_1 & a_2 & b_2 & & & 0 \\ 0 & b_2 & a_3 & & & 0 \\ \vdots & & & \ddots & & \\ & & & & a_{n-1} & b_{n-1} \\ b_n & & & & b_{n-1} & a_n \end{bmatrix}$$

such that

$$\begin{aligned} \sigma(J) &= \{\lambda_k^*\}_{k=1}^n, \\ \sigma(\bar{J}) &= \{\mu_1^*, \dots, \mu_{n-1}^*\}, \\ \prod_1^n b_i &= \beta. \end{aligned}$$

- (SIEP8) [98]:
  - ◇ Given
    - ▷ Real scalars $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_1^*, \ldots, \mu_n^*\}$
    - ▷ Satisfy the interlacing property

$$\lambda_i^* \le \mu_i^* \le \lambda_{i+1}^*, \quad i = 1, \ldots, n,$$

      with $\lambda_{n+1}^* = \infty$,
  - ◇ Find Jacobi matrices J and $\tilde{J}$ so that

$$\begin{aligned} \sigma(J) &= \{\lambda_k^*\}_{k=1}^n, \\ \sigma(\tilde{J}) &= \{\mu_1^*, \ldots, \mu_n^*\}, \\ J - \tilde{J} &\ne 0, \quad \text{only at the } (n,n) \text{ position.} \end{aligned}$$

- (SIEP9):
  - ◇ Given
    - ▷ Distinct real scalars $\{\lambda_1^*, \ldots, \lambda_{2n}^*\}$,
    - ▷ A Jacobi matrix $J_n \in \mathbb{R}^{n\times n}$,
  - ◇ Find a Jacobi matrix $J_{2n} \in \mathbb{R}^{2n\times 2n}$ so that

$$\begin{aligned} \sigma(J_{2n}) &= \{\lambda_1^*, \ldots, \lambda_{2n}^*\}, \\ J_{2n}(1:n, 1:n) &= J_n. \end{aligned}$$

Physical Interpretations

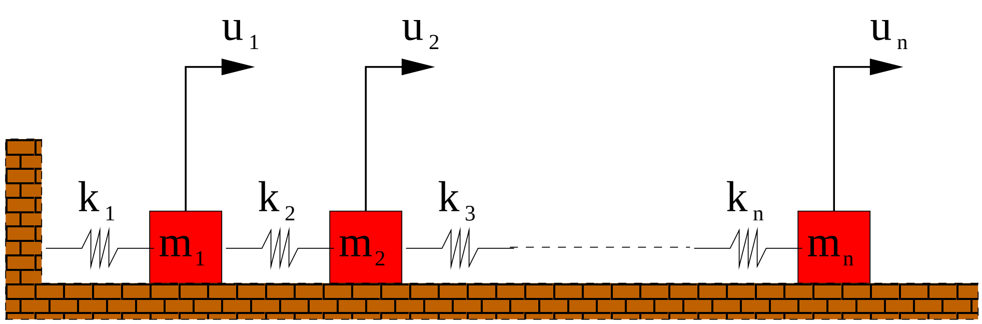


Figure 1: Mass-spring system

- Consider a serially linked mass-spring system with n particles.
  - ◇ m_i = mass of the i-th particle.
  - ◇ k_i = spring constant of the i-th spring.
  - ◇ $u_i(t)$ = displacement of the i-th particle at time t.
- Equation of motion:

$$\begin{aligned} m_1 \frac{d^2 u_1}{dt} &= -k_1 u_1 + k_2(u_2 - u_1), \\ m_i \frac{d^2 u_i}{dt} &= -k_i(u_i - u_{i-1}) + k_{i+1}(u_{i+1} - u_i), \\ & \quad i = 2, \ldots, n-1, \\ m_n \frac{d^2 u_n}{dt} &= -k_n(u_n - u_{n-1}). \end{aligned}$$

- In matrix form:

$$M\frac{d^2\mathbf{u}}{dt} = K\mathbf{u}.$$

  - ◇ $\mathbf{u} = [u_1, \ldots, u_n]^T$.
  - ◇ $M = \mathrm{diag}(m_1, \ldots, m_n)$.
  - ◇ K is the Jacobi matrix given by

$$K = \begin{bmatrix} -(k_1+k_2) & k_2 & 0 & \ldots\ 0 & 0 \\ k_2 & -(k_2+k_3) & k_3 & & 0 \\ 0 & k_3 & -(k_3+k_4) & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & k_n \\ 0 & & & k_n & -k_n \end{bmatrix}$$

- Fundamental solutions are of the form $\mathbf{u}(t) = e^{i\omega t}\mathbf{x}$.
  - ◇ Natural frequency/mode equation is governed by

$$K\mathbf{x} = -\omega^2 M\mathbf{x}.$$

  - ◇ Define $J = M^{-1/2}KM^{-1}$ and $\lambda = -\omega^2$. Then

$$J\mathbf{x} = \lambda\mathbf{x}.$$

- Knowing m_i and k_k, we can predict the natural frequencies and modes of the system.
  - ◇ The inverse problem means that we would like to calculate values such as $\frac{-k_i-k_{i+1}}{m_i}$ and $\frac{k_{i+1}}{\sqrt{m_i m_{i+1}}}$ from the spectral data.

- SIEP6a $\Longleftrightarrow$ Identifying the system from its spectrum and the spectrum of the reduced system where the last mass is held to have no motion.
- SIEP2 $\Longleftrightarrow$ Identifying the system from its spectrum if the system is symmetric with respect to its center.
- SIEP6b $\Longleftrightarrow$ Identifying the damped system, including its damper configurations, from its spectrum and the spectrum of the reduced system where the last mass is held immboile.S
- SIEP7 $\Longleftrightarrow$ Same as SIEP6a except that m_1 and m_2 are connected by another spring mechanism to form a loop.
- SIEP8 $\Longleftrightarrow$ Identifying the system from its spectrum and the spectrum of the new system whereas only the last spring constant is changed.
- SIEP9 $\Longleftrightarrow$ Identifying the system from its spectrum and physical paramters m_i, k_i of the first half particles.
- Sometimes it is impossible to gather the entire spectrum information. Partial information of some eigenvalues and some eigenvectors can also be used to determine a Jacobi matrix. See Chapter 6.

Existence Theory

- Very rich and nearly complete theory available.
- Strictly interlacing property, i.e.,

$$\lambda_i^* < \mu_i^* < \lambda_{i+1}^*, \quad i = 1, \ldots, n-1,$$

  is a necessary condition unless some subdiagonal (superdiagona) entries are zero.

  - ◇ Jacobi matrices are assumed to have positive b_i for all $i = 1, \ldots, n-1$.
  - ◇ Eigenvalues of a Jacobi matrix are necessarily real and distinct.
  - ◇ Eigenvalues of $\bar{J}$ necessarily separate those of J.

- Most existence proofs are based on the recurrence relationship between characteristic polynomials for Jacobi matrices.

- Assume that the given eigenvalues satisfy the strictly interlacing property. Then
  - ◇ The SIEP6a has a unique solution [175].
  - ◇ The SIEP8 has a unique solution.
- If $\{\lambda_k^*\}_{k=1}^n$ are distinct. Then the SIEP2 has a unique solution.
- Over the complex field $\mathbb{C}$,
  - ◇ If the given eigenvalues are distinct, then the SIEP6b is solvable and has at most $2^n(2n-3)!/(n-2)!$ different solutions [289].
  - ◇ If some eigenvalues are common, then there are infinitely many solutions for the SIEP6b.

- Assume that $\{\mu_1^*, \ldots, \mu_{n-1}^*\}$ are distinct. Then the SIEP7 is solvable if and only if

$$\prod_{k=1}^{n} |\mu_j^* - \lambda_k^*| \geq 2\beta(1 + (-1)^{n-j+1}),$$

for all $j = 1, \ldots, n-1$ [360].

  ◇ No uniqueness can be assumed.

  ◇ The eigenvalues of a periodic Jacobi matrix are not necessarily distinct.

  ◇ The eigenvalues of $\bar{J}$ need not separate those of a periodic Jacobi matrix J.

- Assume that $\{\lambda_1^*, \ldots, \lambda_{2n}^*\}$ are distinct.

  ◇ Define

$$\Delta_k = \det\left(\begin{bmatrix} 1 & \ldots & 1 & 1 & \ldots & 1 \\ \lambda_1 & \ldots & e_1^T J_n e_1 & \lambda_{k+1} & \ldots & \lambda_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{2n-1} & \ldots & e_1^T J_n^{2n-1} e_1 & \lambda_{k+1}^{2n-1} & \ldots & \lambda_{2n}^{2n-1} \end{bmatrix}\right)$$

  ◇ Then the SIEP9 has a unique solution if and only if

$$\Delta_k > 0$$

  for all $k = 1, \ldots, 2n$ [360].

Sensitivity Issues

- The function $F : \mathbb{R}^{2n-1} \longrightarrow \mathbb{R}^{2n-1}$ where
$$F(a_1, \ldots, a_n, b_1, \ldots, b_{n-1}) = (\sigma(J), \sigma(\bar{J}))$$
is differentiable, if $b_i > 0$.

- The solution J to the SIEP6a depends continuously on the given data $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_1^*, \ldots, \mu_{n-1}^*\}$ [195].

- Let J and $\tilde{J}$ be solutions to the SIEP6a with data
$$\lambda_1^* < \mu_1^* < \lambda_2^* < \ldots < \mu_{n-1}^* < \lambda_n^*,$$
$$\tilde{\lambda}_1^* < \tilde{\mu}_1^* < \tilde{\lambda}_2^* < \ldots < \tilde{\mu}_{n-1}^* < \tilde{\lambda}_n^*$$
Then there exists a constant K such that
$$\|J - \tilde{J}\|_F \leq K \left(\sum_{i=1}^{n} |\lambda_i^* - \tilde{\lambda}_i^*|^2 + \sum_{i=1}^{n-1} |\mu_i^* - \tilde{\mu}_i^*|^2 \right)^{1/2}.$$
  - ◇ K depends on the separation of the given data measured by
$$d = \max\{\lambda_n^*, \tilde{\lambda}_n^*\} - \min\{\lambda_1^*, \tilde{\lambda}_1^*\},$$
$$\epsilon_0 = \frac{\min_{j,k}\{|\lambda_j^* - \mu_k^*|, |\tilde{\lambda}_j^* - \tilde{\mu}_j^*|\}}{d},$$
$$\delta_0 = \frac{\min_{j \neq k}\{|\lambda_j^* - \lambda_k^*|, |\mu_j^* - \mu_k^*|, |\tilde{\lambda}_j^* - \tilde{\lambda}_k^*|, |\tilde{\mu}_j^* - \tilde{\mu}_k^*|, \}}{2d}.$$

Numerical Methods

- Lanczos method.
  - ◇ Given any matrix A, if $Q^T AQ = H$ where Q is orthogonal and H is upper Hessenberg with positive subdiagonal entries, then Q and H are completely determined by A and the first column of Q.
  - ◇ In our application, $J = Q^T \Lambda Q$ in symmetric diagonal.
    1. $a_1 := \mathbf{q}_1^T A\mathbf{q}_1$.
    2. $b_1 := \|\Lambda\mathbf{q}_1 - a_1\mathbf{q}_1\|$.
    3. $\mathbf{q}_2 = (\Lambda\mathbf{q}_1 - a_1\mathbf{q}_1)/b_1$.
    4. For $i = 2, \ldots, n-1$,
       (a) $a_i := \mathbf{q}_i^T A\mathbf{q}_i$.
       (b) $b_i := \|\Lambda\mathbf{q}_i - a_i\mathbf{q}_i - b_{i-1}\mathbf{q}_{i-1}\|$.
       (c) $q_{i+1} := (\Lambda\mathbf{q}_i - a_i\mathbf{q}_i - b_{i-1}\mathbf{q}_{i-1})/b_i$.
    5. $a_n := \mathbf{q}_n^T A\mathbf{q}_n$.
- Orthogonal reduction method.
  - ◇ Orthogonal tridiagonalization of a bordered diagonal matrix.

Lanczos Method (for SIEP6a)

- Basic facts:
  - ◇ Given any symmetric matrix A with orthonormal eigenpairs $(\lambda_i, \mathbf{x}_i)$, then
$$\operatorname{adj}(\lambda_i I - A) = \prod_{\substack{k=1\\k\neq i}}^{n} (\lambda_i - \lambda_k)\mathbf{x}_i\mathbf{x}_i^T.$$
  - ◇ Evaluate the above equation at the $(1,1)$ position for each $\mathbf{x}_i$ to obtain
$$x_{1i}^2 = \frac{\prod_{k=1}^{n-1}(\lambda_i - \mu_k)}{\prod_{\substack{k=1\\k\neq i}}^{n}(\lambda_i - \lambda_k)}.$$
- For SIEP6a,
  - ◇ The first column of Q for J can be expressed from the spectral data, i.e., $q_{i1} = x_{1i}$.
  - ◇ The Lanczos algorithm kicks in.
  - ◇ Need reorthogonalization along the way.

Orthogonal Reduction Method (for SIEP6a)

- Construct a bordered diagonal matrix A of the form

$$A = \begin{bmatrix} \alpha & \beta_1 & \dots & \beta_{n-1} \\ \beta_1 & \mu_1^* & & 0 \\ \vdots & & \ddots & \\ \beta_{n-1} & 0 & \dots & \mu_{n-1}^* \end{bmatrix}$$

with specified eigenvalues $\sigma(A) = \{\lambda_k^*\}_{k=1}^n$.

◇ α is trivially determined.

$$\alpha = \sum_{i=1}^{n} \lambda_i^* - \sum_{i=1}^{n-1} \mu_i^*.$$

◇ Characteristic polynomial of A is known.

$$\det(\lambda I - A) = (\lambda - \alpha) \prod_{k=1}^{n-1} (\lambda - \mu_k^*) - \sum_{i=1}^{n-1} \beta_i^2 \left(\prod_{\substack{k=1 \\ k \neq i}}^{n-1} (\lambda - \mu_k^*) \right).$$

- ◇ Border elements β_i can be calculated:

$$\beta_i^2 = -\frac{\prod_{k=1}^{n}(\mu_i^* - \lambda_k^*)}{\prod_{\substack{k=1\\k\neq i}}^{n-1}(\mu_i^* - \mu_k^*)}.$$

- Derive orthogonal matrix Q efficiently so that

$$\begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^T \end{bmatrix} A \begin{bmatrix} 1 & \mathbf{0}^T \\ 0 & Q \end{bmatrix} = \begin{bmatrix} \alpha & b_1\mathbf{e}_1^T \\ b_1\mathbf{e}_1 & \bar{J} \end{bmatrix} = J.$$

  - ◇ $b_1 = \|\beta\|$.
  - ◇ $Q^T \mathrm{diag}(\mu_1^*, \ldots, \mu_{n-1}^*)Q = \bar{J}$.
  - ◇ The first column of Q is given by β/b_1.
  - ◇ The Lanczos method can be employed.

- One may also explore the bordered diagonal structure by using Householder transformation, Givens rotations, the Rutishauser method, and so on [41].

Toeplitz Inverse Eigenvalue Problem

- Overview
- Existence Theory
- Numerical Methods

Overview

- Symmetric Toeplitz Matrix $T(r) := t_{ij} = r_{|i-j|+1}$, i.e.,

$$T(r) := \begin{bmatrix} r_1 & r_2 & \dots & r_{n-1} & r_n \\ r_2 & r_1 & & r_{n-2} & r_{n-1} \\ \vdots & \ddots & \ddots & & \vdots \\ r_{n-1} & & & r_1 & r_2 \\ r_n & r_{n-1} & & r_2 & r_1 \end{bmatrix}.$$

- Is a special case of centrosymmetric matrices

$$\mathcal{C}(n) := \{M | M = M^T, M = \Xi M \Xi\}.$$

  - $\diamond$ $\Xi = [\xi_{ij}]$ = unit perdiagonal matrix,

$$\xi_{ij} = \delta_{i,n-j+1}.$$

    - $\triangleright$ Symmetric vector, if $\Xi v = v$.
    - $\triangleright$ Skew-symmetric vector, if $\Xi v = -v$.

- (ToIEP) Find $r \in R^n$ such that $T(r)$ has a prescribed set of real numbers $\{\lambda_k^*\}_{k=1}^n$ as its spectrum.

Spectral Properties of Centrosymmetric Matrices

- Any $M \in \mathcal{C}(n)$ can be decomposed as follows [55]:

n	even	odd
M	$\begin{bmatrix} A & C^T \\ C & \Xi A \Xi \end{bmatrix}$	$\begin{bmatrix} A & x & C^T \\ x^T & q & x^T \Xi \\ C & \Xi x & \Xi A \Xi \end{bmatrix}$
$\sqrt{2}K$	$\begin{bmatrix} I & -\Xi \\ I & \Xi \end{bmatrix}$	$\begin{bmatrix} I & 0 & -\Xi \\ 0 & \sqrt{2} & 0 \\ I & 0 & \Xi \end{bmatrix}$
KMK^T	$\begin{bmatrix} A - \Xi C & 0 \\ 0 & A + \Xi C \end{bmatrix}$	$\begin{bmatrix} A - \Xi C & 0 & 0 \\ 0 & q & \sqrt{2} x^T \\ 0 & \sqrt{2} x & A + \Xi C \end{bmatrix}$

◇ $A, C, \Xi \in R^{\lfloor \frac{n}{2} \times \frac{n}{2} \rfloor}$.

◇ $x \in R^{\lfloor \frac{n}{2} \rfloor}$.

◇ $q \in R$.

◇ $A = A^T$.

- Orthonormal eigenvectors $Q = K^T Z$ M can be split into two groups based on diagonal block Z.

$$Z = \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}.$$

  ◇ $Z_1 :=$ Eigenvectors of $A - \Xi C$.

  ◇ $Z_2 :=$ Eigenvectors of $\begin{bmatrix} q & \sqrt{2}x^T \\ \sqrt{2}x & A + \Xi C \end{bmatrix}$ or $A + \Xi C$.

- Eigenvectors of M enjoy special parity properties:

  ◇ $K^T \begin{bmatrix} Z_1 \\ 0 \end{bmatrix} = \lfloor \frac{n}{2} \rfloor$ skew-symmetric eigenvectors $\Rightarrow$ "Odd" eigenvalues.

  ◇ $K^T \begin{bmatrix} 0 \\ Z_2 \end{bmatrix} = \lceil \frac{n}{2} \rceil$ symmetric eigenvectors $\Rightarrow$ "Even" eigenvalues.

- Open Question: For an ToIEP to be solvable, each given eigenvalue must carry a specific parity. Can this parity be arbitrarily assigned?

A 3×3 Example

- $M \in \mathcal{C}(3)$:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ \times & m_{22} & \times \\ \times & \times & \times \end{bmatrix}.$$

  ◇ $\mathrm{trace}(M) = 0 \Longrightarrow$ Three free parameters in $\mathcal{C}(3)$.

- Isospectral subset $\mathcal{M}_{\mathcal{C}}(\lambda_1, \lambda_2, \lambda_3)$:

$$\begin{aligned} (m_{11} - \frac{\lambda_{\sigma_1}}{4})^2 + \frac{1}{2}m_{12}^2 &= \frac{(\lambda_{\sigma_2} - \lambda_{\sigma_3})^2}{16}, \\ m_{13} &= m_{11} - \lambda_{\sigma_1}. \end{aligned}$$

  ◇ $\sigma =$ A permutation of integers $\{1, 2, 3\}$.

- $\mathcal{M}_{\mathcal{C}} =$ Three ellipses.

  ◇ One circumscribes the other two.

- Check # of m_{12}-intercepts $\Rightarrow$

$$\#\text{ of solutions} = \begin{cases} 4, & \text{if distinct eigenvalues;} \\ 2, & \text{if multiplicity 2.} \end{cases}$$

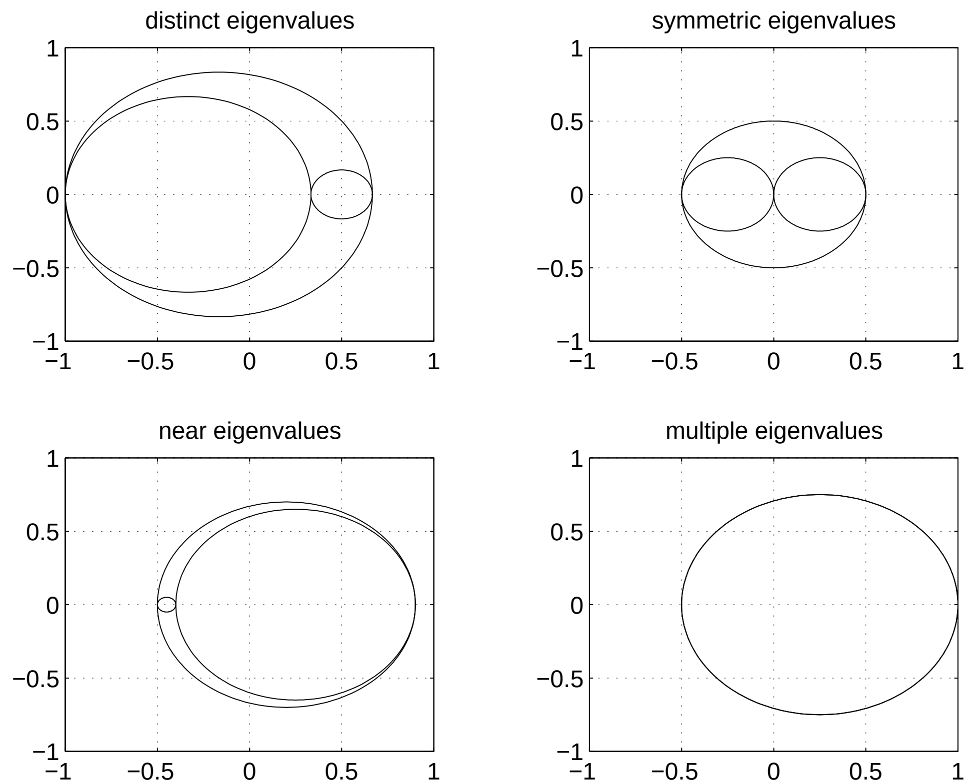


Figure 2: Plots of m_{11} versus m_{12} for $\mathcal{M}_{\mathcal{C}}$ in $\mathcal{C}(3)$.

- Each ellipse = One parity assignment among eigenvalues.
- Wrong assignment $\Rightarrow$ No Toeplitz.
- Magnitude of eigenvalues $\Rightarrow$ Solvability.
- Ordered eigenvalues alternate in parity $\stackrel{?}{\Rightarrow}$ Safeguard.

Inverse Problem for Centrosymmetric Matrices

- Close form solution:
  - ◇ Given *arbitrary*
    - ▷ Diagonal matrix $\Lambda := \mathrm{diag}\{\{\lambda_k^*\}_{k=1}^n\}$,
    - ▷ Orthogonal matrix $Z_1 \in R^{\lfloor \frac{n}{2} \rfloor \times \lfloor \frac{n}{2} \rfloor}$,
    - ▷ Orthogonal matrix $Z_2 \in R^{\lceil \frac{n}{2} \rceil \times \lceil \frac{n}{2} \rceil}$,
  - ◇ Then the matrix

$$M := K^T \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \Lambda \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}^T K.$$

    - ▷ Is centrosymmetric.
    - ▷ $\{\lambda_1^*, \ldots, \lambda_{\lfloor \frac{n}{2} \rfloor}^*\}$ = Odd eigenvalues of M.
    - ▷ $\{\lambda_{\lfloor \frac{n}{2} \rfloor + 1}^*, \ldots, \lambda_n^*\}$ = Even eigenvalues of M.
    - ▷ M may not be Toeplitz.
- Open Question: Search for a Toeplitz matrix on the isospectral surface $\mathcal{M}_{\mathcal{C}}(\{\lambda_k^*\}_{k=1}^n)$:

$$\mathcal{M}_{\mathcal{C}} := \{M \in \mathcal{C}(n) | \text{ eigenvalues} = \lambda_1^*, \ldots, \lambda_n^*\}.$$

Existence

- Solvability has been a challenge.
  - ⋄ n equations in n unknowns.
  - ⋄ $n \geq 5$ is analytically intractable.
  - ⋄ Symmetric Toeplitz matrices can have *arbitrary* real spectra [226].
    - ▷ Thus far, it is a nonconstructive proof by topological degree argument.
    - ▷ Open Question: Any algebraic proof of existence?
  - ⋄ Eigenvalues *cannot* have arbitrary parity.

Idea in Landau's Proof

- A matrix $T(c_1, \ldots, c_n)$ is *regular* if every principal submatrix $T(c_1, \ldots, c_k)$, $1 \le k \le n$ has the properties:
  - ◇ Distinct eigenvalues.
  - ◇ Alternate parity with the largest one having even parity.
- Assume the given eigenvalues $\lambda_1 \le \ldots \le \lambda_n$ are centered, i.e., $\sum_{i=1}^{n} \lambda_i = 0$.
  - ◇ Suffices to solve the ToIEP for matrices of the form $T(0, 1, t_3, \ldots, t_n)$.
  - ◇ Necessarily $\lambda_1 < 0$.
- Consider the map

$$\phi(t_3, \ldots, t_n) = (y_2, \ldots, y_{n-1})$$

  - ◇ $y_i = -\frac{\lambda_i}{\lambda_1}$, $i = 2, \ldots, n-1$.
  - ◇ $\sigma(T(0, 1, t_3, \ldots, t_n)) = \{\lambda_1, \ldots, \lambda_n\}$.

- The range of ϕ is the simplex

$$\Delta := \left\{ (y_2, \ldots, y_{n-1}) | \begin{array}{c} -1 \leq y_2 \leq \ldots \leq y_{n-1} \\ y_2 + \ldots + y_{n-2} + 2y_{n-1} \leq 1 \end{array} \right\}.$$

- Landau's approach:
  - ◇ The set $\mathcal{F}$ of regular Toeplitz matrices of the form $T(0, 1, t_3, \ldots, t_n)$ is not empty.
  - ◇ The map ϕ restricted to those $(t_3, \ldots, t_n) \in \mathbb{R}^{n-2}$ such that $T(0, 1, t_3, \ldots, t_n) \in \mathcal{F}$ is a surjective map onto Δ.
  - ◇ Any $y_1 \leq \ldots \leq y_n$ can be *shifted* and *scaled* to a unique point in Δ.

Numerical Methods

- Mostly done in $\mathcal{S}(n)$.
  - ◇ Laurie's Algorithm [227]
  - ◇ Trench's Algorithm [337]
  - ◇ Continuous method
- The calculation could be limited to the smaller space $\mathcal{C}(n)$.
  - ◇ Cayley Transform [119]
  - ◇ Newton's Refinement to Centrosymmetric Structure

Continuous Method

Refined Newton to Centrosymmetric Structure

- A tangent step
- Lift by approximation
- Lift by global ordering
- Lift by local ordering

A Classical Newton Method

- A function:

$$f : R \longrightarrow R.$$

- The scheme:

$$x^{(\nu+1)} = x^{(\nu)} - (f'(x^{(\nu)}))^{-1} f(x^{(\nu)}).$$

- The intercept:
  - ◇ $x^{(\nu+1)}$ = The x-intercept of the tangent line of the graph of f from $(x^{(\nu)}, f(x^{(\nu)}))$.
- The lifting:
  - ◇ $(x^{(\nu+1)}, f(x^{(\nu+1)}))$ = The natural "lift" of the intercept along the y-axis to the the graph of f.

An Analogy of the Newton Method

- Think of
  - ◇ $\mathcal{M}_{\mathcal{C}}(\Lambda)$ = The graph of f.
  - ◇ $\mathcal{T}(n) :=$ {Toeplitz matrices} = The x-axis.
  - ◇ Limit the iteration to $\mathcal{C}(n)$.
- Manifold $\mathcal{M}_{\mathcal{C}}(\Lambda)$:
  - ◇ Parametrization:

$$\begin{aligned} M &= Q\Lambda Q^T, \\ Q &= K^T Z, \\ Z &\in \mathcal{O}(\lfloor \frac{n}{2} \rfloor) \times \mathcal{O}(\lceil \frac{n}{2} \rceil). \end{aligned}$$

  - ◇ Tangent vector:

$$\begin{aligned} T_M(\mathcal{M}_{\mathcal{C}}) &= \tilde{S}M - M\tilde{S}, \\ \tilde{S} &:= Q \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} Q^T. \end{aligned}$$

▷ S_1 = skew-symmetric in $R^{\lfloor \frac{n}{2} \rfloor \times \lfloor \frac{n}{2} \rfloor}$.
S_2 = skew-symmetric in $R^{\lceil \frac{n}{2} \rceil \times \lceil \frac{n}{2} \rceil}$

A Tangent Step

- Given $M^{(\nu)} \in \mathcal{M}_{\mathcal{C}}(\Lambda)$ (*Parity fixed*),
  - ◇ Find $\tilde{S}^{(\nu)}$ and $r^{(\nu+1)}$ for

$$M^{(\nu)} + \tilde{S}^{(\nu)} M^{(\nu)} - M^{(\nu)} \tilde{S}^{(\nu)} = T(r^{(\nu+1)}).$$

- Equivalently:

$$\begin{aligned} \Lambda + S^{(\nu)}\Lambda - \Lambda S^{(\nu)} &= {Q^{(\nu)}}^T T(r^{(\nu+1)}) Q^{(\nu)} \\ &= {Z^{(\nu)}}^T K T(r^{(\nu+1)}) K^T Z^{(\nu)}. \end{aligned}$$

  - ◇ Spectral decomposition:

$$\begin{aligned} {Q^{(\nu)}}^T M^{(\nu)} Q^{(\nu)} &= \Lambda, \\ Q^{(\nu)} &= K^T Z^{(\nu)}. \end{aligned}$$

- Key observation:

$$K T(r^{(\nu+1)}) K^T = \begin{bmatrix} T_1^{(\nu+1)} & 0 \\ 0 & T_2^{(\nu+1)} \end{bmatrix}.$$

$\Rightarrow$ The system is split in half.

Find the Intercept

- The right-hand side of the system is linear in $r^{(\nu+1)}$.
- Diagonal elements in the system $\Rightarrow$ A linear system for $r^{(\nu+1)}$ without reference to $S^{(\nu)}$:

$$J^{(\nu)} r^{(\nu+1)} = \lambda.$$

  - ◇ $\lambda := [\phi_1, \ldots, \phi_{\lfloor \frac{n}{2} \rfloor}, \psi_1, \ldots \psi_{\lceil \frac{n}{2} \rceil}]^T$ (*Fixed parity*).
  - ◇

$$J_{ij}^{(\nu)} := \begin{cases} (Z_1^{(\nu)})_{*i}^T E_1^{[j]} (Z_1^{(\nu)})_{*i}, & \text{if } 1 \leq i \leq \lfloor \frac{n}{2} \rfloor; \\ (Z_2^{(\nu)})_{*i}^T E_2^{[j]} (Z_2^{(\nu)})_{*i}, & \text{if } \lfloor \frac{n}{2} \rfloor < i \leq n. \end{cases}$$

    - ▷

$$\begin{bmatrix} E_1^{[j]} & 0 \\ 0 & E_2^{[j]} \end{bmatrix} = K T(e_j) K^T.$$

    - ▷ $(Z_k^{(\nu)})_{*i}$:= the i^{th} column of the matrix $Z_k^{(\nu)}$.

- Only length of $\approx \frac{n}{2}$ in all vector-matrix-vector multiplications.

Compute $S^{(\nu)}$

- Once $T(r^{(\nu+1)})$ is determined, off-diagonal elments in the system $\Rightarrow S^{(\nu)}$:

$$(S_1^{(\nu)})_{ij} = \frac{(Z_1^{(\nu)})_{*i}^T T_1^{(\nu+1)} (Z_1^{(\nu)})_{*j}}{\phi_i - \phi_j},\ 1 \leq i < j \leq \lfloor \frac{n}{2} \rfloor,$$

$$(S_2^{(\nu)})_{ij} = \frac{(Z_2^{(\nu)})_{*i}^T T_2^{(\nu+1)} (Z_2^{(\nu)})_{*j}}{\psi_i - \psi_j},\ 1 \leq i < j \leq \lceil \frac{n}{2} \rceil.$$

  - ◇ Eigenvalues within each parity group must be distinct.
  - ◇ $\lambda_1, \ldots, \lambda_n$ need not be totally distinct.
- In case of multiple eigenvalues
  - ◇ Basis of eigenspace splits as evenly as possible between symmetric and skew-symmetric eigenvectors [113].
  - ◇ Multiplicity of each eigenvalue ≤ 2 can be formulated.

Find the Lift

- Coordinate-free lift (Friedland, '87; Chu, '92):

$$M^{(\nu+1)} := Q^{(\nu)} {R^{(\nu)}}^T {Q^{(\nu)}}^T M^{(\nu)} Q^{(\nu)} R^{(\nu)} {Q^{(\nu)}}^T.$$

  ◇ Lift by approximation:

$$R^{(\nu)} := \left(I + \frac{S^{(\nu)}}{2}\right)\left(I - \frac{S^{(\nu)}}{2}\right)^{-1}.$$

- In calculation, only need

$$Z^{(\nu+1)} := Z^{(\nu)} {R^{(\nu)}}^T.$$

  ◇ All matrices involved are 2-block diagonal.

- Quadratic convergence.
- Multiplicity $> 2 \Rightarrow$ No $S^{(\nu)} \Rightarrow$ No lift.
- *Can we by-pass* $S^{(\nu)}$ *to perform a lift?*

Lift by Global Ordering

- Idea:
  - ◇ Look for matrix $M^{(\nu+1)} \in \mathcal{M}_{\mathcal{C}}$ that is nearest to $T(r^{(\nu+1)})$.

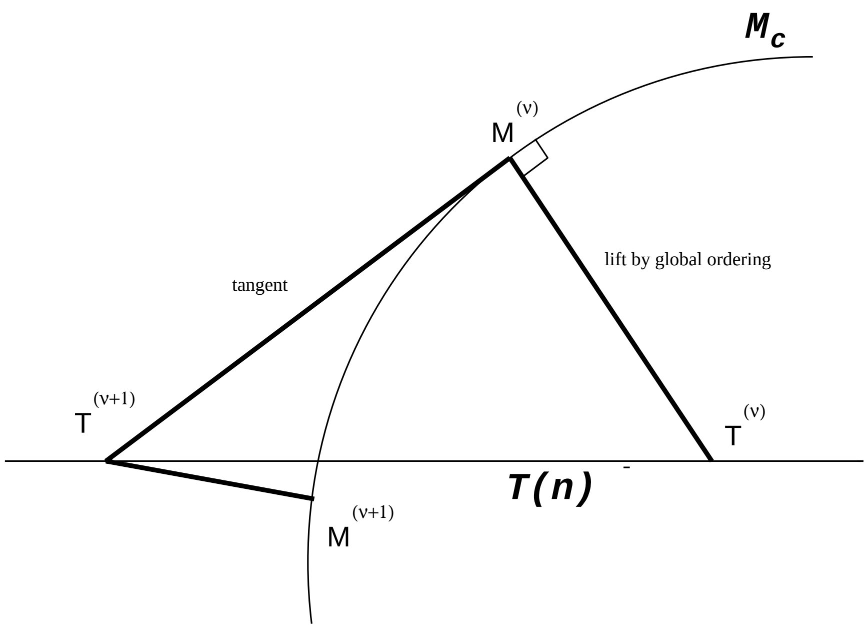


Figure 3: Geometry of Lift by Global Ordering.

- Answer: Wielandt-Hoffman theorem.
  - ◇ Spectral decomposition of $T(r^{(\nu+1)})$ is easy:

$$\overline{Z}^{(\nu+1)^T} K T(r^{(\nu+1)}) K^T \overline{Z}^{(\nu+1)} = \begin{bmatrix} \overline{\Lambda}_1^{(\nu+1)} & 0 \\ 0 & \overline{\Lambda}_2^{(\nu+1)} \end{bmatrix}.$$

  - ◇ Rearrange $\{\lambda_1, \ldots, \lambda_n\}$ in the same ordering as in $\overline{\Lambda}_1^{(\nu+1)}$ and $\overline{\Lambda}_2^{(\nu+1)}$ to obtain $\tilde{\Lambda}_1^{(\nu+1)}$ and $\tilde{\Lambda}_2^{(\nu+1)}$.
  - ◇ Define:

$$M^{(\nu+1)} := K^T \overline{Z}^{(\nu+1)} \begin{bmatrix} \tilde{\Lambda}_1^{(\nu+1)} & 0 \\ 0 & \tilde{\Lambda}_2^{(\nu+1)} \end{bmatrix} \overline{Z}^{(\nu+1)^T} K.$$

- New starting point:

$$\begin{aligned} \Lambda = \Lambda^{(\nu+1)} &:= \begin{bmatrix} \tilde{\Lambda}_1^{(\nu+1)} & 0 \\ 0 & \tilde{\Lambda}_2^{(\nu+1)} \end{bmatrix}, \\ Z^{(\nu+1)} &:= \overline{Z}^{(\nu+1)}. \end{aligned}$$

- Significance:
  - ◇ Parity assignment may be changed.
  - ◇ No $S^{(\nu)}$ is needed.
  - ◇ Multiple eigenvalues with same parity can be handled.

Lift by Local Ordering

- Would like to avoid computing $S^{(\nu)}$ as well as parity switching?
- Idea:
  - ◇ Λ is kept fixed.
  - ◇ Reorganize columns of $\overline{Z}_1^{(\nu+1)}$ and $\overline{Z}_2^{(\nu+1)}$.
- Calculation:
  - ◇ Elements in $\overline{\Lambda}_1^{(\nu+1)}, \overline{\Lambda}_2^{(\nu+1)}$ are in the same ordering as those in Λ_1 and Λ_2.
- New starting point:

$$Z^{(\nu+1)} := \text{The reorganized } \overline{Z}^{(\nu+1)}.$$

- Global ordering = Local ordering, when reaching convergence.
- Quadratic convergence:
  - ◇ Order of convergence = (order projection)*(order tangent step). (Traub)

Numerical Experiment

- Example 1: Wrong parity
- Example 2: Quadratic convergence
- Example 3: Multiplicity = 2
- Example 4: Multiplicity = 3
- Example 5: High order case

Example 1: Wrong Parity

- Test data (*Wrong parity*):

$$\left.\begin{array}{rcr} \lambda_1 &=& -2.4128 \times 10^{+0}(E) \\ \lambda_2 &=& -2.6407 \times 10^{-1}(E) \\ \lambda_3 &=& 2.6769 \times 10^{+0}(O) \end{array}\right\} \textit{Wrong parity}$$

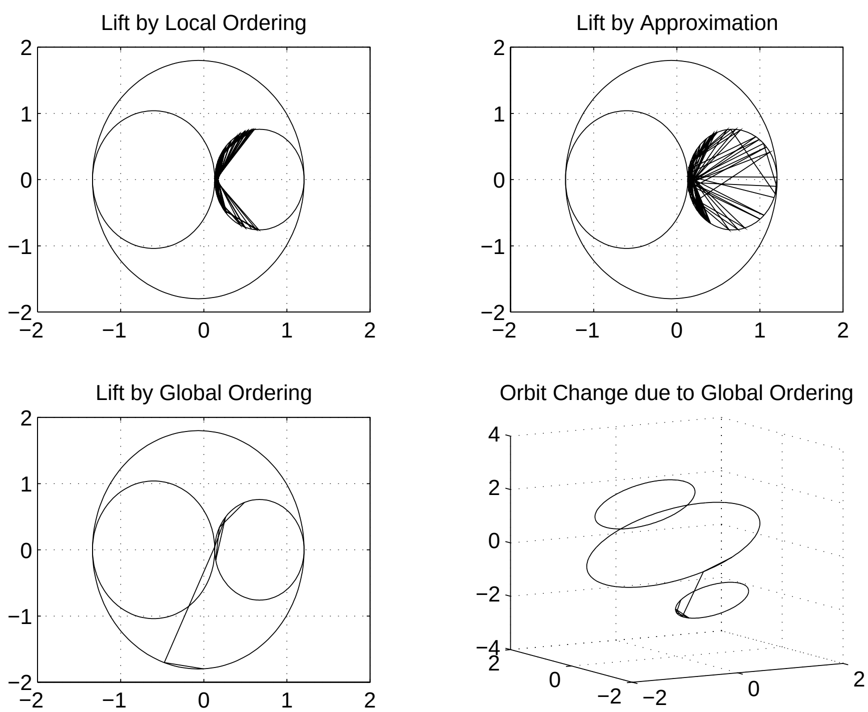


Figure 4: Behaviors of Algorithms When Starting with the Wrong Orbit.

- Lift by approximation $\Rightarrow$ Staying on the *wrong* orbit.
- Local ordering $\Rightarrow$ Wrong orbit, clustering.
- Global ordering $\Rightarrow$ Change orbit, convergence.

Example 2: Quadratic Convergence

- Limit point $r^{(*)}$ may be away from original $r^{(\#)}$, even if $r^{(0)} \approx r^{(\#)}$.
- Limit points may be different among methods, even with the same starting $r^{(0)}$.
- Eigenvalues of $T(r^{(*)}) =$ those of $T(r^{\#})$, but parity may change in the global ordering case.

	Case (a)	Case (b)	Case (c)
$r^{(\#)}$	0	0	0
	-2.0413$\times 10^{-3}$	-9.2349$\times 10^{-1}$	-3.3671$\times 10^{-1}$
	1.6065$\times 10^{+0}$	-7.0499$\times 10^{-2}$	4.1523$\times 10^{-1}$
Original Value	8.4765$\times 10^{-1}$	1.4789$\times 10^{-1}$	1.5578$\times 10^{+0}$
	2.6810$\times 10^{-1}$	-5.5709$\times 10^{-1}$	-2.4443$\times 10^{+0}$
$r^{(0)}$	0	0	0
	-2.8351$\times 10^{-1}$	-1.8024$\times 10^{+0}$	6.3658$\times 10^{-1}$
	9.3953$\times 10^{-1}$	7.3881$\times 10^{-1}$	4.0318$\times 10^{-1}$
Initial Value	8.2068$\times 10^{-1}$	1.5694$\times 10^{-1}$	1.0901$\times 10^{+0}$
	1.0634$\times 10^{+0}$	-5.2451$\times 10^{-1}$	-3.2628$\times 10^{+0}$
$r^{(*)}$	2.0426$\times 10^{-16}$	2.2204$\times 10^{-16}$	7.4940$\times 10^{-16}$
	-2.0413$\times 10^{-3}$	-9.2349$\times 10^{-1}$	-3.5391$\times 10^{-1}$
	1.6065$\times 10^{+0}$	-7.0499$\times 10^{-2}$	4.3645$\times 10^{-1}$
Local Ordering	8.4765$\times 10^{-1}$	1.4789$\times 10^{-1}$	1.5244$\times 10^{+0}$
	2.6810$\times 10^{-1}$	-5.5709$\times 10^{-1}$	-2.4655$\times 10^{+0}$
$r^{(*)}$	8.6831$\times 10^{-16}$	0	4.7184$\times 10^{-16}$
	-2.0413$\times 10^{-3}$	-9.2349$\times 10^{-1}$	-3.3671$\times 10^{-1}$
	1.6065$\times 10^{+0}$	-7.0499$\times 10^{-2}$	4.1523$\times 10^{-1}$
Approximation	8.4765$\times 10^{-1}$	1.4789$\times 10^{-1}$	1.5578$\times 10^{+0}$
	2.6810$\times 10^{-1}$	-5.5709$\times 10^{-1}$	-2.4443$\times 10^{+0}$
$r^{(*)}$	2.4113$\times 10^{-16}$	-1.1102$\times 10^{-16}$	6.1062$\times 10^{-16}$
	-9.3778$\times 10^{-2}$	-9.2646$\times 10^{-1}$	3.5391$\times 10^{-1}$
	1.5174$\times 10^{+0}$	-6.1419$\times 10^{-2}$	4.3645$\times 10^{-1}$
Global Ordering	9.9597$\times 10^{-1}$	1.3518$\times 10^{-1}$	-1.5244$\times 10^{+0}$
	5.7042$\times 10^{-1}$	-5.4694$\times 10^{-1}$	-2.4655$\times 10^{+0}$

Table 1: Initial and Final Values of $r^{(\nu)}$ for Example 2.

Iterations	Local Ordering	Approximation	Global Ordering
0	$1.3847\times10^{+0}$	$1.3847\times10^{+0}$	$1.2194\times10^{+0}$
1	7.1545×10^{-1}	7.1545×10^{-1}	4.2739×10^{-1}
2	2.1982×10^{-2}	6.3866×10^{-2}	1.4179×10^{-2}
3	5.1223×10^{-5}	2.0606×10^{-4}	4.3624×10^{-5}
4	4.4931×10^{-10}	7.1037×10^{-9}	4.7985×10^{-10}
5	1.4729×10^{-15}	2.9671×10^{-15}	1.7659×10^{-15}

Table 2: Errors of Eigenvalues for Case (a) in Example 2.

Example 3: Multiplicity = 2

- Test data (Random number):

$$\begin{array}{rc} -5.8942 \times 10^{-1} & (E) \\ \left\{ \begin{array}{r} -1.8565 \times 10^{-1} \\ -1.8565 \times 10^{-1} \end{array} \right. & \begin{array}{c} (O) \\ (E) \end{array} \\ 3.7508 \times 10^{-1} & (O) \\ 5.8564 \times 10^{-1} & (E) \end{array}$$

- Parity unknown.
  - ⋄ Assume the possibly safest assignment.
- Multiply eigenvalues split between parities.
- Quadratic convergence.

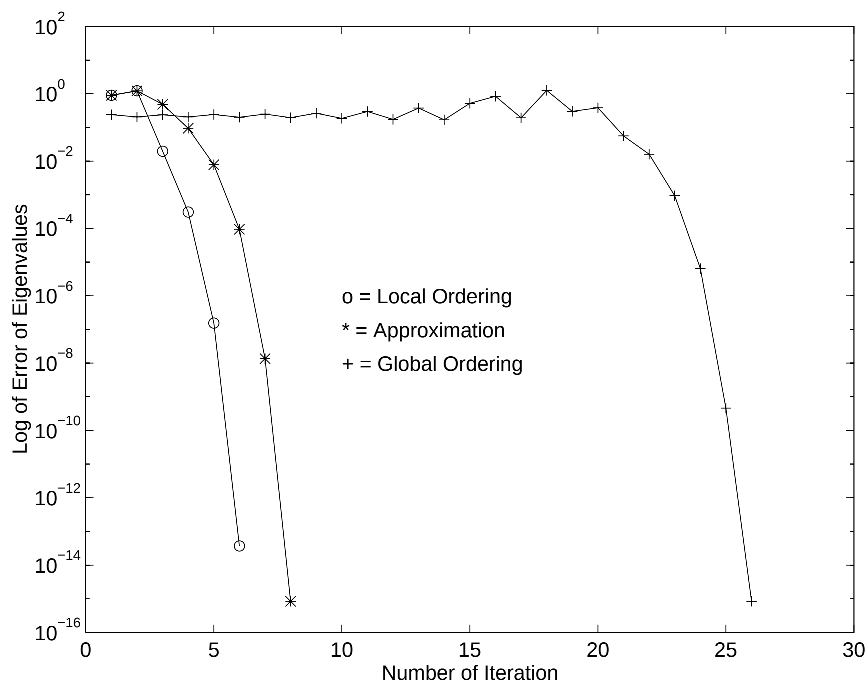


Figure 5: Number of Iteration versus Logarithmic Scale of Errors in Example 3.

Example 4: Multiplicity = 3

- Test data:

$$\begin{array}{ll} -8.4328 \times 10^{-1} & (E) \\ \left\{ \begin{array}{l} -1.2863 \times 10^{-1} \quad (O) \\ -1.2863 \times 10^{-1} \quad (E) \\ -1.2863 \times 10^{-1} \quad (O) \end{array} \right. & \\ 1.2292 \times 10^{+0} & (E) \end{array}$$

- Lift by approximation fails.
- Methods by local and global ordering converge to

$$[.2204\times10^{-16}, 4.2222\times10^{-1}, 1.2863\times10^{-1}, 4.2222\times10^{-1}, 1.2863\times10^{-1}]$$

with error history

$$\begin{array}{lll} 2.0327 & \times & 10^{+0} \\ 4.0355 & \times & 10^{-2} \\ 1.3903 & \times & 10^{-4} \\ 3.5477 & \times & 10^{-9} \\ 7.8896 & \times & 10^{-16}. \end{array}$$

Example 5: $n = 20$

- Test data:

$-1.0242\times10^{+1}$	$-9.6736\times10^{+0}$	$-5.5608\times10^{+0}$	$-2.2651\times10^{+0}$
5.5692×10^{-1}	$2.1786\times10^{+0}$	$3.3867\times10^{+0}$	$4.0016\times10^{+0}$
$6.3594\times10^{+0}$	$8.7090\times10^{+0}$		

$-1.0416\times10^{+1}$	$-9.4352\times10^{+0}$	$-4.7955\times10^{+0}$	-7.7180×10^{-1}
6.3996×10^{-1}	$2.6374\times10^{+0}$	$4.4879\times10^{+0}$	$4.7572\times10^{+0}$
$6.2222\times10^{+0}$	$9.2230\times10^{+0}$		

  - ◇ Not the safest possible parity assignment, first ten odd, last ten even.

- Method of approximation fails after 100 iterations.
- Method of global ordering performs best.

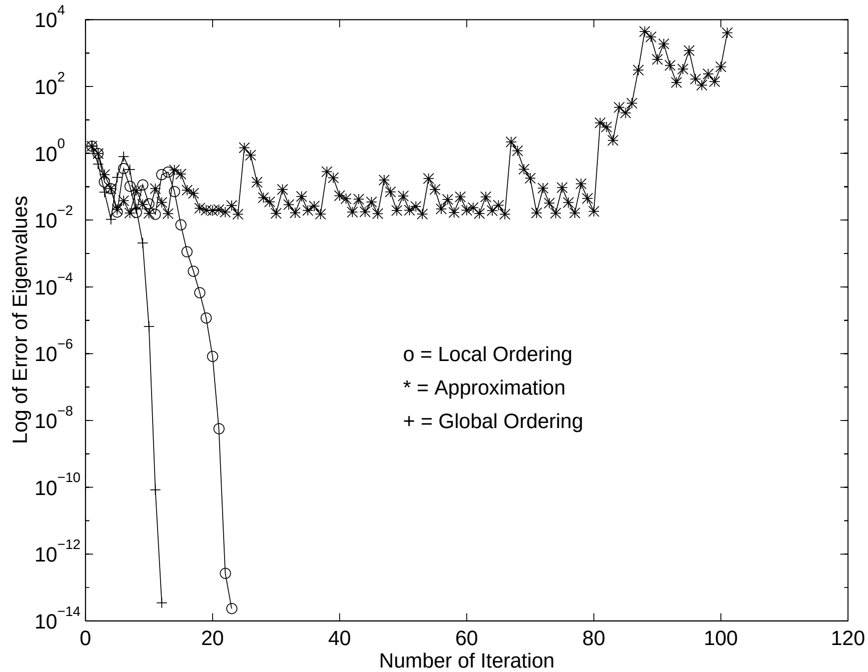


Figure 6: Number of Iteration versus Logarithmic Scale of Errors in Example 5.

Conclusion

- Solving the ToIEP within the subspace $\mathcal{C}(n)$ is possible.
  - ◇ Problem size and cost are halved.
  - ◇ Multiple eigenvalue case can be handled.
- Coordinate-free Newton-like methods are available.
  - ◇ Quadratic convergence is observed.
- Parity assignment of eigenvalues plays an important role in whether an ToIEP is solvable.
- Both local and global ordering, based on the Wielandt-Hoffman theorem, permit a new way of lifting.
  - ◇ Higher multiplicity eigenvalue case can now be handled.

Nonnegative Inverse Eigenvalue Problem

- Overview
- Some Existence Results
- Symmetric Nonnegative Inverse Eigenvalue Problem
- Numerical methods

Overview

- Many discussions in the literature on the subject [26, 30, 45, 130, 140, 143, 245, 262, 109, 318].
- Most discussions center around establishing a sufficient or necessary condition to qualify whether a given set of values is the spectrum of a nonnegative matrix.
- Open Question: Which sets of n real numbers occur as the spectrum of a nonnegative matrix?
- Open Question: Which sets of n real numbers occur as the spectrum of a symmetric nonnegative matrix?
- Open Question: Very few numerical algorithms.

Some Existence Results

- Suppose $\{\lambda_k^*\}_{k=1}^n$ are eigenvalues of an $n \times n$ nonnegative matrix. The the moments

$$s_k = \sum_{i=1}^{n} (\lambda_i^*)^k$$

  must satisfy

$$s_k^m \leq n^{m-1} s_{km}$$

  for all $k, m = 1, 2, \ldots$ [245].

- The set $\{\lambda_k^*\}_{k=1}^n \subset \mathbb{C}$ is the nonzero spectrum of a strictly positive matrix of size $m \geq n$ if and only if [45]
  - ◇ $\lambda_1^* > |\lambda_i^*|$ for all $i > 1$,
  - ◇ $s_k > 0$ for all $k = 1, 2, \ldots,$ and
  - ◇ The polynomial $\prod_{i=1}^{n}(t - \lambda_i^*)$ has real coefficients.

Symmetric Nonnegative Inverse Eigenvalue Problem

- There exist real numbers $\{\lambda_k^*\}_{k=1}^n$ that occur as the spectrum of a nonnegative $n \times n$ matrix, but do not occur as the spectrum of a symmetric nonnegative $n \times n$ matrix [212].

- The symmetric nonnegative inverse eigenvalue problem can be formulated as a constrained optimization problem of *minimizing* the objective function

$$F(Q, R) := \frac{1}{2}\|Q^T \Lambda Q - R \circ R\|^2,$$

subject to

$$(Q, R) \in \mathcal{O}(n) \times \mathcal{S}(n).$$

- For nonsymmetric nonnegative inverse eigenvalue problems, see the discussion for the stochastic inverse eigenvalue problems.

Numerical Method

- A dynamical system resulting from projected gradient flow can be formulated as [74]:

$$\frac{dX}{dt} = [X, [X, Y]],$$
$$\frac{dY}{dt} = 4Y \circ (X - Y).$$

  ◇ $X(t) = Q(t)^T \Lambda Q(t)$ is an isospectral matrix.

  ◇ $Y(t) = R(t) \circ R(t)$ is a symmetric nonnegative matrix.

Stochastic Inverse Eigenvalue Problem

- General View
- Karpelevič's Theorem
- Relationship to Nonnegative Matrices
- Basic Formulation
- Steepest Descent Flow
- ASVD Flow
- Convergence
- Numerical Experiment
- Conclusion

General View

- Construct a stochastic matrix with prescribed spectrum.
  - ◇ Stochastic structure.
  - ◇ No strings of symmetry.
  - ◇ Eigenvalues can appear in complex conjugate pairs.
- A hard problem [215, 262].
  - ◇ The set Θ_n of points in the complex plane that are eigenvalues of stochastic $n \times n$ matrices is completely characterized.
  - ◇ The Karpelevič theorem characterizes only one complex value a time and does not provide further insights into when two or more points in Θ_n are eigenvalues of the *same* stochastic matrix.

Karpelevič's Theorem

- A number λ is an eigenvalue for a stochastic matrix if and only if it belongs to a region Θ_n.
  - ◇ Region is symmetric about the real axis.
  - ◇ The points on the unit circles are given by $e^{2\pi i a/b}$ where a and b range over all integers such that $0 \leq a < b \leq n$.
  - ◇ The boundary of Θ_n consists of curvilinear arcs connecting these points in circular order. These arcs are characterized by specific parametric equations

$$\begin{aligned} \lambda^q(\lambda^p - t)^r &= (1-t)^r, \\ (\lambda^b - t)^d &= (1-t)^d \lambda^q, \end{aligned}$$

    where $0 \leq t \leq 1$, and b, d, p, q, r are natural integers determined certain specific rules (explicitly given in [215, 262]).

- The region Θ_4 is shown in Figure 7.

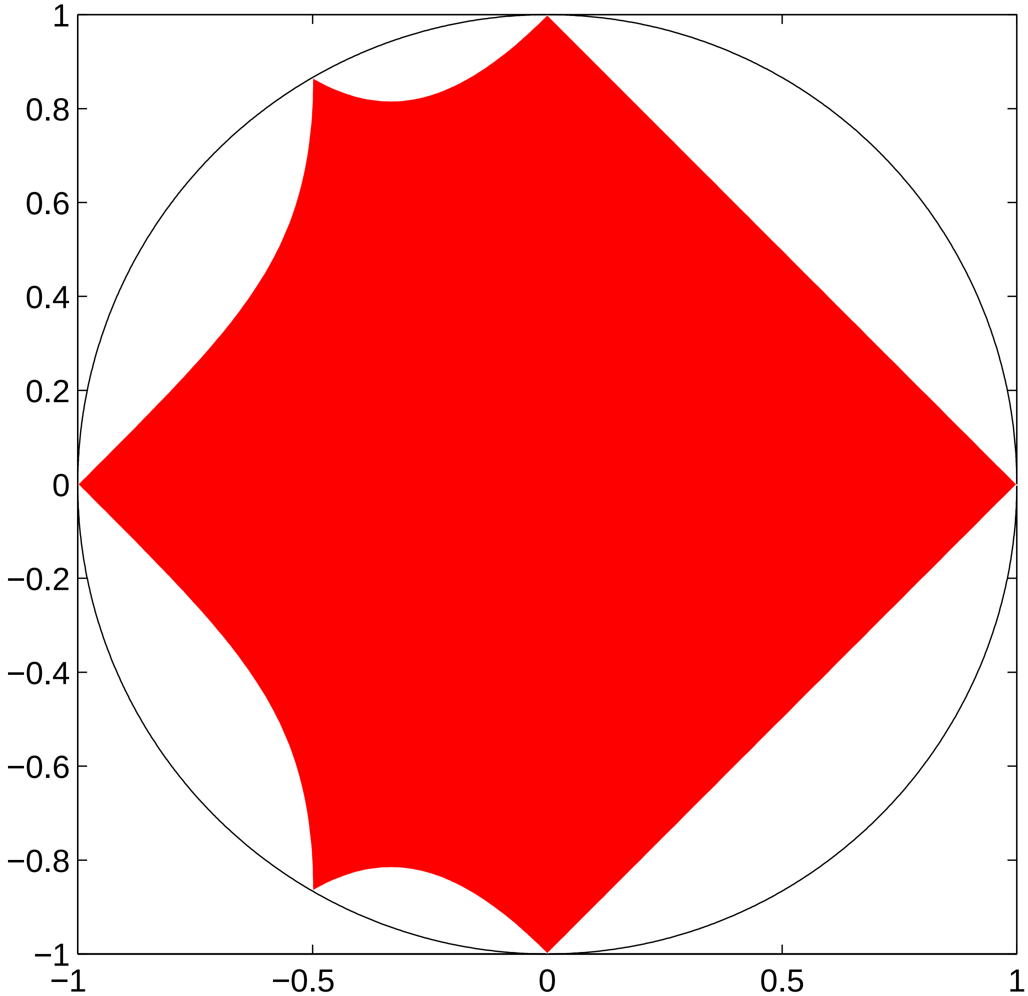


Figure 7: Θ_4 by the Karpelevič Theorem.

Relation to Nonnegative Matrices

- A complex nonzero number α is an eigenvalue of a nonnegative matrix with a positive maximal eigenvalue r if and only if α/r is an eigenvalue of a stochastic matrix.
- Key transformation:
  - ◇ Suppose A is a nonnegative matrix with positive maximal eigenvalue r and a positive maximal eigenvector x.
  - ◇ Then $D^{-1}r^{-1}AD$ is a stochastic matrix where $D := \mathrm{diag}\{x_1, \ldots, x_n\}$.
- The nonnegative inverse eigenvalue probelm (NIEP) has been discussed earlier.
  - ◇ Some necessary and a few sufficient conditions for the NIEP are available [30].
  - ◇ A continuous method for the symmetric NIEP can be formulated [74].
  - ◇ Open Question: Need a numerical algorithm for general NIEP.

Basic Formulation

- Notation:

$$\begin{aligned} \mathcal{M}(\Lambda) &:= \{P\Lambda P^{-1} | P \in R^{n\times n} \text{ is nonsingular}\} \\ \pi(R^n_+) &:= \{B \circ B | B \in R^{n\times n}\} \end{aligned}$$

  - ◇ Λ = real-valued matrix carrying the spectrum information.
  - ◇ $\circ$ = Hadamard product.

- Idea:
  - ◇ Find the intersection of $\mathcal{M}(\Lambda)$ and $\pi(R^n+)$.
  - ◇ The intersection, if exists, results in a nonnegative matrix isospectral to Λ.
  - ◇ Reduce the nonnegative matrix, if its maximal eigenvector is positive, to a stochastic matrix by diagonal similarity transformation.

Reformulation

$$\begin{array}{ll} \text{Minimize} & F(P,R) := \frac{1}{2}||PJP^{-1} - R \circ R||^2 \\ \text{Subject to} & P \in Gl(n),\ R \in gl(n) \end{array}$$

- P and R are used as coordinates to maneuver elements in $\mathcal{M}(\Lambda)$ and $\pi(R_+^n)$ to reduce the objective value.
- Feasible domains are open sets.
- A minimum may not exist.

Gradient of F

- Inner product in the product topology:

$$\langle (X_1, Y_1), (X_2, Y_2) \rangle := \langle X_1, X_2 \rangle + \langle Y_1, Y_2 \rangle.$$

- With respect to the product topology:

$$\nabla F(P, R) = \\ \big((\Delta(P,R) M(P)^T - M(P)^T \Delta(P,R)) P^{-T}, \\ -2\Delta(P,R) \circ R \big) .$$

◇ Abbreviation:

$$\begin{aligned} M(P) &:= PJP^{-1} \\ \Delta(P,R) &:= M(P) - R \circ R. \end{aligned}$$

Steepest Descent Flow

- Steepest descent flow:

$$\begin{aligned}\frac{dP}{dt} &= [M(P)^T, \Delta(P,R)]P^{-T}\\ \frac{dR}{dt} &= 2\Delta(P,R) \circ R.\end{aligned}$$

- Advantages:
  - ◇ No longer need the projection of $\nabla F(P,R)$ as does in the symmetric case.
  - ◇ The zero structure in the original matrix $R(0)$ is preserved throughout the integration — may be used to explore the possibility of constructing a Markov chain with prescribed linkages and spectrum.
- Disadvantage:
  - ◇ The solution flow $P(t)$ is susceptible to becoming unbounded — a possible frailty.
  - ◇ The involvement of P^{-1} is somewhat worrisome.

ASVD flow

- An analytic singular value decomposition of the path of matrices $P(t)$ is an analytic path of factorizations

$$P(t) = X(t)S(t)Y(t)^T$$

  where $X(t)$ and $Y(t)$ are orthogonal and $S(t)$ is diagonal.

- An ASVD exists if $P(t)$ is analytic [48, 345].
- The fact that $P(t)$ defined by the differential system is analytic follows from the Cauchy-Kovalevskaya theorem since the coefficients of the vector field are analytic.

New Coordinate System

- The two matrices P and R are used, respectively, as *coordinates* to describe the isospectral matrices and nonnegative matrices.
  - ◇ May have used more dimensions of variables than necessary — does no harm.
  - ◇ When flows $P(t)$ and $R(t)$ are introduced, in a sense a flow in $\mathcal{M}(\Lambda)$ and a flow in $\pi(R_+^n)$ are also introduced.
- The motion of the coordinate P is further described by three other variables X, S, and Y according to the ASVD.
- To produce the steepest descent flow, a coordinate system $(X(t), S(t), Y(t), R(t))$ is eventually imposed on matrices in $\mathcal{M}(\Lambda) \times \pi(R_+^n)$.

Calculating the ASVD

- Differentiate $P(t) = X(t)S(t)Y(t)^T$: (Wright '92):

$$\begin{aligned} \dot{P} &= \dot{X}SY^T + X\dot{S}Y^T + XS\dot{Y}^T \\ X^T\dot{P}Y &= \underbrace{X^T\dot{X}}_{Z}S + \dot{S} + S\underbrace{\dot{Y}^TY}_{W} \end{aligned}$$

  - ◇ Z, W are skew-symmetric matrices.

- Define $Q := X^T\dot{P}Y$.
  - ◇ Q is known since $\dot{P}$ is already specified.
  - ◇ The inverse of $P(t)$ is calculated from
  $$P^{-1} = YS^{-1}X^T.$$
  - ◇ The diagonal entries of $S = \mathrm{diag}\{s_1, \ldots, s_n\}$ provide us with information about the proximity of $P(t)$ to singularity.

- Flow for $S(t)$:

$$\frac{dS}{dt} = \mathrm{diag}(Q).$$

- Obtain $W(t)$ and $Z(t)$:

$$\begin{aligned} q_{jk} &= z_{jk}s_k + s_j w_{jk}, \\ -q_{kj} &= z_{jk}s_j + s_k w_{jk}. \end{aligned}$$

  ◇ If $s_k^2 \neq s_j^2$, then

$$z_{jk} = \frac{s_k q_{jk} + s_j q_{kj}}{s_k^2 - s_j^2},$$
$$w_{jk} = \frac{s_j q_{jk} + s_k q_{kj}}{s_j^2 - s_k^2}$$

  for all $j > k$.

- Flow for $X(t)$ and $Y(t)$:

$$\begin{aligned} \frac{dX}{dt} &= XZ. \\ \frac{dY}{dt} &= YW. \end{aligned}$$

- The flow is now ready to be integrated by any IVP solvers.

Convergence

- The approach fails only when:
  - ◇ $P(t)$ becomes singular in finite time — requires a restart.
  - ◇ $F(P(t), R(t))$ converges to a nonzero constant — a LS local solution is found.
- Gradient flows enjoy global convergence:
  - ◇ $G(t) := F(P(t), R(t))$ enjoys the property:

    $$\frac{dG}{dt} = -\|\nabla F(P(t), R(t))\|^2 \leq 0$$

    along any solution curve $(P(t), R(t))$.
  - ◇ Suppose $P(t)$ remains nonsingular. Then $G(t)$ converges.

Numerical Experiment

- Integrator: MATLAB ODE SUITE
  - ◇ **ode113** = ABM, PECE, non-stiff system.
  - ◇ **ode15s** = Klopfenstein-Shampine, quasi-constant step size, stiff system.
- Stopping criteria:
  - ◇ ABSERR = RELERR = 10^{-12}.
  - ◇ $\|\Delta(P, R)\| \leq 10^{-9} \Rightarrow$ a stochastic matrix has been found.
  - ◇ Relative improvement of $\Delta(P, R)$ between two consecutive output points $\leq 10^{-9} \Rightarrow$ a LS solution is found.

Example 1

- Spectrum:

$$\{1.0000, -0.2403, 0.1186 \pm 0.1805i, -0.1018\}$$

- Initial values:

$$P_0 = \begin{bmatrix} 0.2002 & 0.4213 & 0.9229 & 0.7243 & 0.4548 \\ 0.6964 & 0.0752 & 0.9361 & 0.2235 & 0.0981 \\ 0.7538 & 0.3620 & 0.2157 & 0.5272 & 0.2637 \\ 0.4366 & 0.3220 & 0.8688 & 0.1729 & 0.8697 \\ 0.8897 & 0.1436 & 0.7097 & 0.5343 & 0.7837 \end{bmatrix}$$

$$R_0 = .8328\mathbf{1}$$

- Limit point:

$$B = \begin{bmatrix} 0.1679 & 0.0522 & 0.4721 & 0.0000 & 0.3078 \\ 0.1436 & 0.1779 & 0.4186 & 0.1901 & 0.0698 \\ 0.0000 & 0.1377 & 0.5291 & 0.3034 & 0.0299 \\ 0.0560 & 0.4690 & 0.2404 & 0.0038 & 0.2309 \\ 0.1931 & 0.1011 & 0.5339 & 0.1553 & 0.0165 \end{bmatrix}.$$

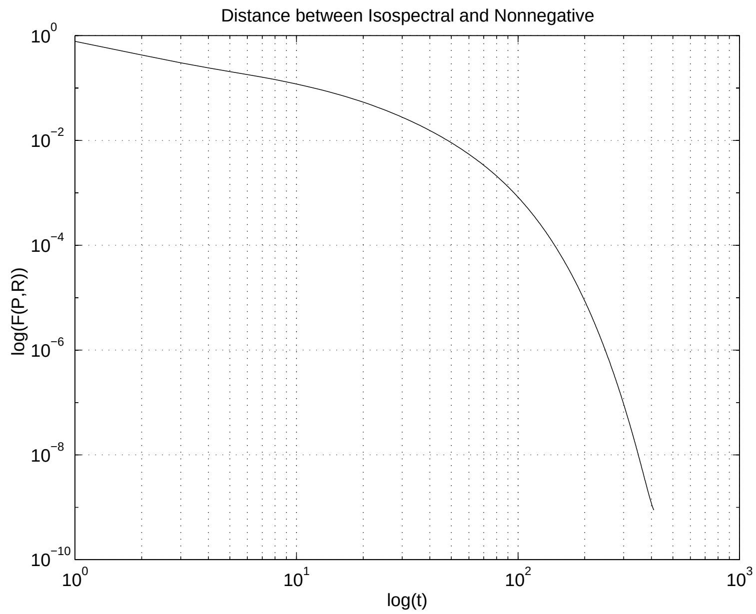


Figure 8: A log-log plot of $F(P(t), R(t))$ versus t for Example 1.

- Both solvers work reasonably.
  - ◇ **ode15s** advances with larger step sizes at the cost of solving implicit algebraic equations.
  - ◇ Jacobians are calculated by finite difference. Function calls could be reduced by fewer output points.
- Different initial values lead to different stochastic matrices.

Example 2

- Spectrum:

$$\{1.0000, -0.2608, 0.5046, 0.6438, -0.4483\}$$

- Looking for a Markov chain with ring linkage, i.e., each state is linked at most to its two immediate neighbors.

- Initial values:

$$P_0 = \begin{bmatrix} 0.1825 & 0.7922 & 0.2567 & 0.9260 & 0.9063 \\ 0.1967 & 0.5737 & 0.7206 & 0.5153 & 0.0186 \\ 0.5281 & 0.2994 & 0.9550 & 0.6994 & 0.1383 \\ 0.7948 & 0.6379 & 0.5787 & 0.1005 & 0.9024 \\ 0.5094 & 0.8956 & 0.3954 & 0.6125 & 0.4410 \end{bmatrix}$$

$$R_0 = 0.9210 \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

- Limit point:

$$D = \begin{bmatrix} 0.0000 & 0.3094 & 0 & 0 & 0.6906 \\ 0.0040 & 0.5063 & 0.4896 & 0 & 0 \\ 0 & 0.0000 & 0.5134 & 0.4866 & 0 \\ 0 & 0 & 0.7733 & 0.2246 & 0.0021 \\ 0.4149 & 0 & 0 & 0.3900 & 0.1951 \end{bmatrix}$$

Example 3

- Spectrum

$$\{1.0000, -0.2403, 0.3090 \pm 0.5000i, -0.1018\}$$

- Initial values: same as Example 1 (or modify R_0).
- Slow convergence:

$$E = \begin{bmatrix} 0.3818 & 0.0000 & 0.4568 & 0.0000 & 0.1614 \\ 0.5082 & 0.3314 & 0.0871 & 0.0049 & 0.0684 \\ 0.0000 & 0.0000 & 0.5288 & 0.4712 & 0.0000 \\ 0.0266 & 0.7634 & 0.0292 & 0.0310 & 0.1498 \\ 0.5416 & 0.0524 & 0.3835 & 0.0196 & 0.0029 \end{bmatrix}$$

$$F = \begin{bmatrix} 0.3237 & 0 & 0.4684 & 0 & 0.2079 \\ 0.4742 & 0.3184 & 0.1303 & 0.0007 & 0.0764 \\ 0 & 0.0000 & 0.5231 & 0.4769 & 0 \\ 0.0066 & 0.7536 & 0.0372 & 0.0958 & 0.1068 \\ 0.5441 & 0.0429 & 0.3959 & 0.0022 & 0.0149 \end{bmatrix}$$

Conclusion

- The theory of solvability on the StIEP or the NIEP is yet to be developed.
- An ODE approach capable of solving the StIEP or the NIEP numerically, if the prescribed spectrum is feasible, is proposed.
- The method is easy to implement by existing ODE solvers.
- The method can also be used to approximate least squares solutions or linearly structured matrices.

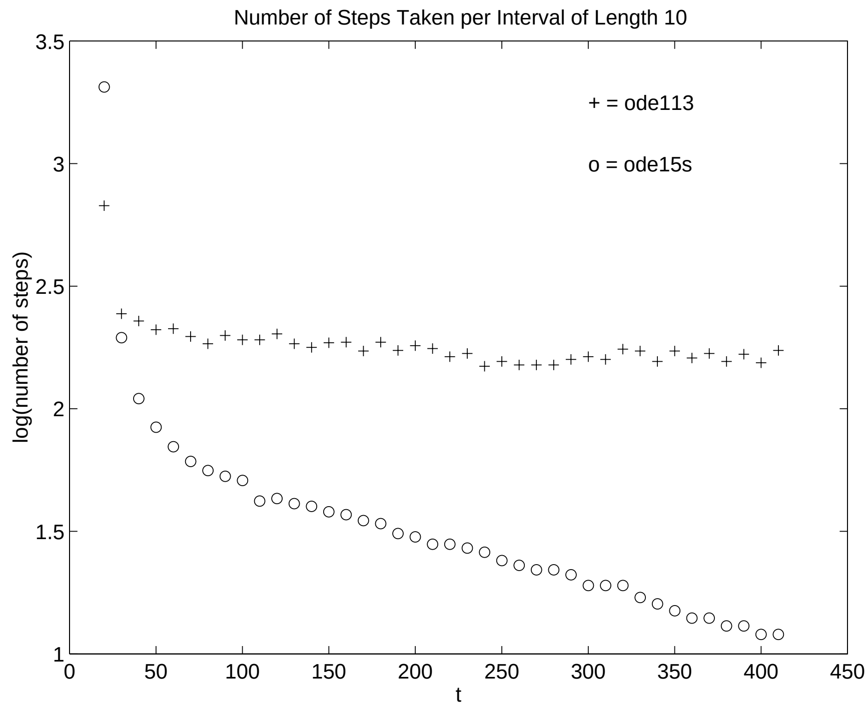


Figure 9: A comparison of steps taken by **ode113** and **ode15s** for Example 1.

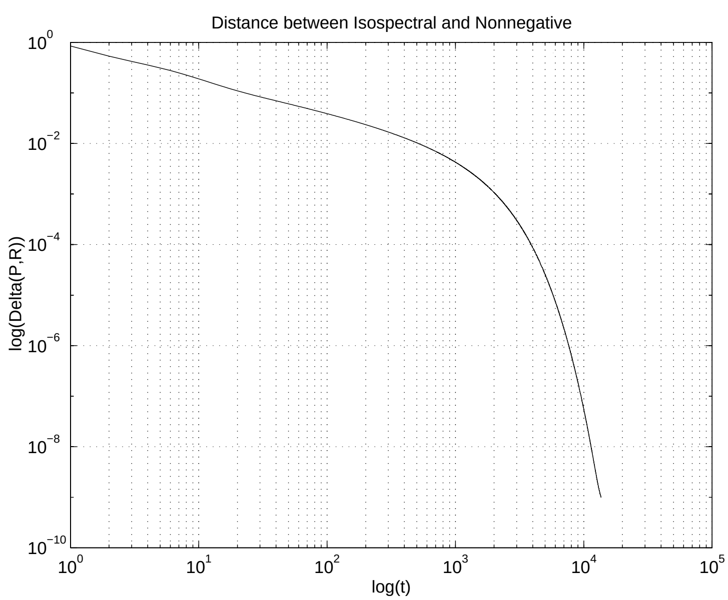


Figure 10: A log-log plot of $F(P(t), R(t))$ versus t for Example 3.

Unitary Inverse Eigenvalue Problem

- Overview
- Formulation
- Existence Theory

Overview

- Eigenvalue of unitary matrices are on the unit circle.
- Suffices to concentrate on unitary Hessenberg matrices.
- Any upper Hessenberg unitary matrix H with positive subdiagonal entries can uniquely expressed as the product

$$H = G_1(\eta_1) \dots G_{n-1}(\eta_{n-1})\tilde{G}_n(\eta_n),$$

  - ◇ $\eta_k \in \mathrm{C}$ with $|\eta_k| < 1$ for $1 \leq k < n$ and $|\eta_n| = 1$.
  - ◇ Each $G_k(\eta_k)$, $k = 1, \dots, n-1$ is a Givens rotation,

$$G_k(\eta_k) = \begin{bmatrix} I_{k-1} & & & \\ & -\eta_k & \zeta_k & \\ & \zeta_k & \bar{\eta}_k & \\ & & & I_{n-k+1} \end{bmatrix}$$

    with $\zeta_k := \sqrt{1 - |\eta_k|^2}$.
  - ◇ $\tilde{G}_n(\eta_n) = \mathrm{diag}[I_{n-1}, -\eta_n]$.
- Each upper Hessenbergunitary matrix is determined by $2n - 1$ *real* parameters.
  - ◇ $\{\eta_k\}_{j=1}^{n}$ are called the Schur parameters.
  - ◇ $H = H(\eta_1, \dots, \eta_n)$.

Formulation

- The Schur parametrization of an upper Hesserberg unitary matrix requies $2n-1$ pieces of information.
  - ◇ The complementary parameters $\{\zeta_k\}_{k=1}^{n-1}$ are the subdiagonal elements of H and cannot be independently given.
- Upper Hessenberg unitary matrices with positive subdiagonal entries are related to orthogonal polynomials on the unit circle.
  - ◇ Jacobi matrices are related to orthogonal polynomials on an interval.
  - ◇ There should considerably similarity between unitary inverse eigenvalue problems the Jacobi inverse eigenvalue problems.
- Need a concept for modified principal submatrices of H.

Existence Theory

- Analogue of SIEP8:
  - ◇ Given
    - ▷ Two sets $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_k^*\}_{k=1}^n$ strictly interlaced on the unit circle,
  - ◇ Then there exist a unique $H = H(\eta_1, \ldots, \eta_n)$ and a unique $\alpha \in \mathbb{C}$ of unit modulus such that
    - ▷ $\sigma(H) = \{\lambda_k^*\}_{k=1}^n$.
    - ▷ $\sigma(H(\alpha\eta_1, \ldots, \alpha\eta_n)) = \{\mu_k^*\}_{k=1}^n$.
- Analogue of SIEP6a [4]:
  - ◇ Given
    - ▷ Two sets $\{\lambda_k^*\}_{k=1}^n$ and $\{\mu_k^*\}_{k=0}^{n-1}$ strictly interlaced on the unit circle,
  - ◇ Then there exist a unique $H = H(\eta_1, \ldots, \eta_n)$ such that
    - ▷ $\sigma(H) = \{\lambda_k^*\}_{k=1}^n$.
    - ▷ $\sigma(H(\eta_1, \ldots, \eta_{n-2}, \rho_{n-1})) = \{\mu_k^*\}_{k=1}^{n-1}$ with

$$\rho_{n-1} = \frac{\eta_{n-1}\bar{\mu}_0\eta_n}{1 + \bar{\mu}_0\bar{\eta}_{n-1}\eta_n}.$$

Inverse Eigenvalue Problems with Prescribed Entries

- Overview
- Prescribed Entries Along the Diagonal
- Prescribed Entries at Arbitrary Location
- Numerical Methods

Overview

- The PEIEP is a special kind of matrix completion problem [217]:
  - ◇ Given
    - ▷ A certain subset $\mathcal{K} = \{(i_t, j_t\}_{t=1}^k$ of pairs of subscripts,
    - ▷ A certain set of values $\{a_1, \ldots, a_k\} \subset \mathbf{F}$,
    - ▷ Another set of n values $\{\lambda_k^*\}_{k=1}^n$,
  - ◇ Find a matrix $X \in \mathbf{F}^{n \times n}$ such that
    - ▷ $\sigma(X) = \{\lambda_k^*\}_{k=1}^n$.
    - ▷ $X_{i_t, j_t} = a_t$ for $t = 1, \ldots, k$.
- Positions that do not belong to $\mathcal{K}$ are *free*, whose $n^2 - k$ entries are to be determined.
  - ◇ Jacobi structure is a special case.
  - ◇ Sometimes only need to fill $\mathcal{K}$ positions with prescribed values, but not in any specific order.
- What is the minimal/maximal count of k for the problem to make sense?

Prescribed Entries along the Diagonal

- Schur-Horn Theorem (on Hermitain matrices).
- Mirsky Theorem (on general matrices).
- Sing-Thompson Theorem (on singular values).
- de Oliveira Theorem (on general diagonals).

Schur-Horn Theorem

- Concerns with the relationship between diagonal entries and eigenvalues of a Hermitian matrix.
- The vector $a \in \mathbb{R}^n$ is said to majorize $\lambda \in \mathbb{R}^n$ if, assuming the ordering

$$\begin{aligned} a_{j_1} &\leq \ldots \leq a_{j_n}, \\ \lambda_{m_1} &\leq \ldots \leq \lambda_{m_n}, \end{aligned}$$

  the following relationships hold:

$$\begin{aligned} \sum_{i=1}^{k} \lambda_{m_i} &\leq \sum_{i=1}^{k} a_{j_i}, \quad \text{for } k = 1, \ldots n, \\ \sum_{i=1}^{n} \lambda_{m_i} &= \sum_{i=1}^{n} a_{j_i}. \end{aligned}$$

- A Hermitian matrix H with eigenvalues λ and diagonal entries a exists if and only if a *majorizes* λ.
- The proof for the sufficient part is the hard part.
  - ◇ (SHIEP) Construct such a Hermitian matrix with given diagonals a and eigenvalues λ, if a majorizes λ [78, 379].

Mirsky Theorem

- Is there any similar connection between eigenvalues and diagonal entries of a general matrix?
- A matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and main diagonal elements $a_1, \ldots, a_n$ exists if and only if

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \lambda_i.$$

  - ⋄ Not an interesting inverse eigenvalue problems.

Sing-Thompson Theorem

- Concerns with the relationship singular values and diagonal entries of a general matrix.
- Given vectors $d, s \in \mathcal{R}^n$,
  - ◇ Assume
$$\begin{aligned} s_1 &\geq s_2 &\geq \ldots s_n, \\ |d_1| &\geq |d_2| &\geq \ldots |d_n|. \end{aligned}$$
  - ◇ A real matrix with singular values s and main diagonal entries d (possibly in different order) exists if and only if
$$\begin{aligned} \sum_{i=1}^{k} |d_i| &\leq \sum_{i=1}^{k} s_i, \quad \text{for } k = 1, \ldots, n, \\ \left(\sum_{i=1}^{n-1} |d_i|\right) - |d_n| &\leq \left(\sum_{i=1}^{n-1} s_i\right) - s_n. \end{aligned}$$
- (STISVP) Construct such a square matrix with given diagonals and singular values [71].

de Oliveira Theorem

- Corresponding to a given permutation ρ, the set $\mathcal{D} = \{(i, \rho(i))\}_1^n$ is called a ρ-diagonal.
  - ◇ Let $\rho = \rho_1 \dots \rho_s$ be the representation of ρ as the product of disjoint cycles ρ_k.
- A generalizaton of the Mirsky Theorem [105, 106, 107]:
  - ◇ Given
    - ▷ Arbitrary $\{\lambda_k^*\}_{k=1}^n \subset \mathbf{F}$,
    - ▷ Arbitrary numbers $\{a_1, \dots, a_n\} \subset \mathbf{F}$,
    - ▷ Suppose that at least one of the cycles $\rho_1, \dots, \rho_s$ has length > 2.
  - ◇ Then there exists a matrix $X \in \mathbf{F}^{n\times n}$ such that
    - ▷ $\sigma(X) = \{\lambda_k^*\}_{k=1}^n$.
    - ▷ $X_{i,\rho(i)} = a_i$ for $i = 1, \dots n$.

Prescribed Entries at Arbitrary Locations

- London-Minc Theorem [246, 105]:
  - ◇ Given
    - ▷ Arbitrary $\{\lambda_k^*\}_{k=1}^n \subset \mathbf{F}$,
    - ▷ Arbitrary values $a_1, \ldots, a_{n-1}$,
    - ▷ Arbitrary but distinct positions $\{(i_t, j_t)\}_{t=1}^{n-1}$,
  - ◇ There exists a matrix $X \in \mathbf{F}^{n\times n}$ such that
    - ▷ $\sigma(X) = \{\lambda_k^*\}_{k=1}^n$.
    - ▷ $X_{i_t,j_t} = a_t$ for $t = 1, \ldots, n-1$.
- Can matrices have arbitrary $n-1$ prescribed entries and prescribed characteristic polynomials? (See [116, 217].)
- Open Question: How many more entries of a matrix can be specified with prescribed eigenvalues?

Cardinality and Locations

- Specific locations:
  - ◇ Both the SHIEP and the STISVP have n prescribed entries that are located at the diagonal.
    - ▷ Certain inequalities involving the prescribed eigenvalues and diagonal entries must be satisfied.
  - ◇ The AIEP has $n^2 - n$ prescribed entries that are located at the off-diagonal.
    - ▷ The AIEP is always solvable over an algebraically closed field and there at most $n!$ solutions.
- Arbitrary locations with $|\mathcal{K}| = n$ [217]:
  - ◇ Suppose that
    - ▷ $\mathbf{F}$ is algebraically closed.
    - ▷ The Mirsky condition is satisfied, if $\mathcal{K} = \{(i,i)\}_{i=1}^{n}$.
    - ▷ $a_i = \lambda_j$ for some j, if $\mathcal{K} = \{(i, j_t)\}_{t=1}^{n}$ and $a_t = 0$ for all $j_t \neq i$.
  - ◇ Then the PEIPE is solvable via rational algorithm over $\mathbf{F}$.

- Arbitrary location with $|\mathcal{K}| = 2n - 3$ [191]:
  - ◇ Suppose that
    - ▷ $\mathbf{F}$ is algebraically closed.
    - ▷ The Mirsky condition is satisfied, if $\mathcal{K} \supseteq \{(i,i)\}_{i=1}^{n}$.
    - ▷ $a_i = \lambda_j$ for some j, if $\mathcal{K} \supseteq \{(i, j_t)\}_{t=1}^{n}$ and $a_t = 0$ for all $j_t \neq i$.
  - ◇ Then the PEIEP is solvable in $\mathbf{F}$.

Numerical Methods

- Projected gradient method can be applied [78].
- An induction proof can be implemented as a recursive algorithm, provided the computer permits a subprogram to invoke itself recursively.
  - ◇ Fast recursive algorithms have been proposed for inverse problem associated with the SHIEP and the STISVP Theorem.
  - ◇ Details are similar to discussion for the inverse singular/eigenvalue problem.
- Open Question: Has not seen the numerical implementation of either the de Oliveira Theorem or the London-Minc Theorem, though this could be done in finitely many steps.
- Open Question: Need an algorithm to implement the Hershkowitz results.

Inverse Singular Value Problems

- IEP versus ISVP.
- Existence Question.
- A Continuous Approach.
- An Iterative Method for IEP.
- An Iterative Approach for ISVP.

IEP versus ISVP

- Inverse Eigenvalue Problem (IEP):
  - ◇ Given
    - ▷ Symmetric matrices $A_0, A_1, \ldots, A_n \in R^{n\times n}$;
    - ▷ Real numbers $\lambda_1^* \geq \ldots \geq \lambda_n^*$,
  - ◇ Find
    - ▷ Values of $c := (c_1, \ldots, c_n)^T \in R^n$
    - ▷ Eigenvalues of the matrix

$$A(c) := A_0 + c_1 A_1 + \ldots + c_n A_n$$

are precisely $\lambda_1^*, \ldots, \lambda_n^*$.

- Inverse Singular Value Problem ISVP:
  - ◇ Given
    - ▷ General matrices $B_0, B_1, \ldots, B_n \in R^{m\times n}$, $m \geq n$;
    - ▷ Nonnegative real numbers $\sigma_1^* \geq \ldots \geq \sigma_n^*$,
  - ◇ Find
    - ▷ Values of $c := (c_1, \ldots, c_n)^T \in R^n$
    - ▷ Singular values of the matrix

$$B(c) := B_0 + c_1 B_1 + \ldots + c_n B_n$$

are precisely $\sigma_1^*, \ldots, \sigma_n^*$.

Existence Question

- Not always does the IEP have a solution.
- Inverse Toeplitz Eigenvalue Problem (ITEP)
  - ◇ A special case of the (IEP) where $A_0 = 0$ and $A_k := (A_{ij}^{(k)})$ with

$$A_{ij}^{(k)} := \begin{cases} 1, & \text{if } |i-j| = k-1; \\ 0, & \text{otherwise.} \end{cases}$$

  - ◇ Symmetric Toeplitz matrices can have *arbitrary* real spectra [226].
- Not aware of any result concerning the existence question for ISVP.

Notation

- $\mathcal{O}(n) :=$ All orthogonal matrices in $R^{n\times n}$;
- $\Sigma = (\Sigma_{ij}) :=$ A "diagonal" matrix in $R^{m\times n}$

$$\Sigma_{ij} := \begin{cases} \sigma_i^*, & \text{if } 1 \leq i = j \leq n; \\ 0, & \text{otherwise.} \end{cases}$$

- $\mathcal{M}_s(\Sigma) := \{U\Sigma V^T | U \in \mathcal{O}(m), V \in \mathcal{O}(n)\}$
  - ◇ Contains all matrices in $R^{m\times n}$ whose singular values are precisely $\sigma_1^*, \ldots, \sigma_n^*$.
- $\mathcal{B} := \{B(c) | c \in R^n\}$.
- Solving the ISVP $\equiv$ Finding an intersection of the two sets $\mathcal{M}_s(\Sigma)$ and $\mathcal{B}$.

A Continuous Approach

- Assume
  - ◇ $\langle B_i, B_j \rangle = \delta_{ij}$ for $1 \leq i \leq j \leq n$.
  - ◇ $\langle B_0, B_k \rangle = 0$ for $1 \leq k \leq n$.
- The projection of X onto the linear subspace spanned by $B_1, \ldots, B_n$:

$$P(X) = \sum_{k=1}^{n} \langle X, B_k \rangle B_k.$$

- The distance from X to the affine subspace $\mathcal{B}$:

$$dist(X, \mathcal{B}) = \|X - (B_0 + P(X))\|.$$

- Define, for any $U \in R^{m \times m}$ and $V \in R^{n \times n}$, a residual matrix:

$$R(U, V) := U\Sigma V^T - (B_0 + P(U\Sigma V^T)).$$

- Consider the optimization problem:

$$\begin{array}{rl} \text{Minimize} & F(U, V) := \dfrac{1}{2}\|R(U, V)\|^2 \\ \text{Subject to} & (U, V) \in \mathcal{O}(m) \times \mathcal{O}(n). \end{array}$$

Compute the Projected Gradient

- Frobenius inner product on $R^{m\times m} \times R^{n\times n}$:

$$\langle (A_1, B_1), (A_2, B_2)\rangle := \langle A_1, A_2\rangle + \langle B_1, B_2\rangle.$$

- The gradient ∇F may be interpreted as the pair of matrices:

$$\nabla F(U, V) = (R(U, V)V\Sigma^T, R(U, V)^T U\Sigma).$$

- Tangent space can be split:

$$\mathcal{T}_{(U,V)}(\mathcal{O}(m) \times \mathcal{O}(n)) = \mathcal{T}_U\mathcal{O}(m) \times \mathcal{T}_V\mathcal{O}(n).$$

- Projection is easy because:

$$\begin{aligned} R^{n\times n} &= T_V\mathcal{O}(n) \oplus T_V\mathcal{O}(n)^\perp \\ &= V\mathcal{S}(n)^\perp \oplus V\mathcal{S}(n) \end{aligned}$$

- Project the gradient $\nabla F(U,V)$ onto the tangent space $\mathcal{T}_{(U,V)}(\mathcal{O}(m)\times\mathcal{O}(n))$:

$$g(U,V) = \left(\frac{R(U,V)V\Sigma^T U^T - U\Sigma V^T R(U,V)^T}{2}U, \frac{R(U,V)^T U\Sigma V^T - V\Sigma^T U^T R(U,V)}{2}V\right).$$

- Descent vector field:

$$\frac{d(U,V)}{dt} = -g(U,V)$$

defines a steepest descent flow on the manifold $\mathcal{O}(m)\times\mathcal{O}(n)$ for the objective function $F(U,V)$.

The Differential Equation on $\mathcal{M}_s(\Sigma)$

- Define

$$X(t) := U(t)\Sigma V(t)^T.$$

- $X(t)$ satisfies the differential system:

$$\begin{aligned}\frac{dX}{dt} &= X\frac{X^T(B_0+P(X)) - (B_0+P(X))^T X}{2} \\ &\quad -\frac{X(B_0+P(X))^T - (B_0+P(X))^T X}{2}X.\end{aligned}$$

- $X(t)$ moves on the iso-singular-value surface $\mathcal{M}_s(\Sigma)$ in the steepest descent direction to minimize $dist(X(t), \mathcal{B})$.
- This is a continuous method for the ISVP.

Remarks

- No assumption on the multiplicity of singular values is needed.
- Any tangent vector $T(X)$ to $\mathcal{M}_s(\Sigma)$ at a point $X \in \mathcal{M}_s(\Sigma)$ about which a local chart can be defined must be of the form

$$T(X) = XK - HX$$

for some skew symmetric matrices $H \in R^{m\times m}$ and $K \in R^{n\times n}$.

An Iterative Approach for ISVP

- An analogous Newton iteration for PIEP has been discussed.
- Assume
  - ◇ Matrices $B_0, B_1, \ldots, B_n$ are arbitrary.
  - ◇ All singular values $\sigma_1^*, \ldots, \sigma_n^*$ are positive and distinct.
- Given $X^{(\nu)} \in \mathcal{M}_s(\Sigma)$
  - ◇ There exist $U^{(\nu)} \in \mathcal{O}(m)$ and $V^{(\nu)} \in \mathcal{O}(n)$ such that

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

  - ◇ Seek a $\mathcal{B}$-intercept $B(c^{(\nu+1)})$ of a line that is tangent to the manifold $\mathcal{M}_s(\Sigma)$ at $X^{(\nu)}$.
  - ◇ Lift the matrix $B(c^{(\nu+1)}) \in \mathcal{B}$ to a point $X^{(\nu+1)} \in \mathcal{M}_s(\Sigma)$.

Find the Intercept

- Find
  - ◇ Skew-symmetric matrices $H^{(\nu)} \in R^{m\times m}$ and $K^{(\nu)} \in R^{n\times n}$, and
  - ◇ A vector $c^{(\nu+1)} \in R^n$,

  such that

$$X^{(\nu)} + X^{(\nu)}K^{(\nu)} - H^{(\nu)}X^{(\nu)} = B(c^{(\nu+1)})$$

- Equivalently,

$$\Sigma + \Sigma\tilde{K}^{(\nu)} - \tilde{H}^{(\nu)}\Sigma = {U^{(\nu)}}^T B(c^{(\nu+1)})V^{(\nu)}$$

  - ◇ Underdetermined skew-symmetric matrices:

$$\begin{aligned}\tilde{H}^{(\nu)} &:= {U^{(\nu)}}^T H^{(\nu)}U^{(\nu)},\\ \tilde{K}^{(\nu)} &:= {V^{(\nu)}}^T K^{(\nu)}V^{(\nu)}.\end{aligned}$$

- Can determine $c^{(\nu+1)}$, $H^{(\nu)}$ and $K^{(\nu)}$ separately.

- Totally $\frac{m(m-1)}{2} + \frac{n(n-1)}{2} + n$ unknowns — the vector $c^{(\nu+1)}$ and the skew matrices $\tilde{H}^{(\nu)}$ and $\tilde{K}^{(\nu)}$.
- Only mn equations.
- Observe: $\tilde{H}_{ij}^{(\nu)}$, $n+1 \le i \ne j \le m$,
  - ◇ $\frac{(m-n)(m-n-1)}{2}$ unknowns.
  - ◇ Locate at the lower right corner of $\tilde{H}^{(\nu)}$.
  - ◇ Are not bound to any equations at all.
  - ◇ Set
$$\tilde{H}_{ij}^{(\nu)} = 0 \text{ for } n+1 \le i \ne j \le m.$$
- Denote
$$W^{(\nu)} := {U^{(\nu)}}^T B(c^{(\nu+1)}) V^{(\nu)}.$$

  Then
$$W_{ij}^{(\nu)} = \Sigma_{ij} + \Sigma_{ii}\tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)}\Sigma_{jj},$$

Determine $c^{(\nu+1)}$

- For $1 \leq i = j \leq n$,

$$J^{(\nu)} c^{(\nu+1)} = \sigma^* - b^{(\nu)}$$

  ◇ Know quantities:

$$\begin{aligned} J_{st}^{(\nu)} &:= {u_s^{(\nu)}}^T B_t v_s^{(\nu)}, \text{ for } 1 \leq s, t \leq n, \\ \sigma^* &:= (\sigma_1^*, \ldots, \sigma_n^*)^T, \\ b_s^{(\nu)} &:= {u_s^{(\nu)}}^T B_0 v_s^{(\nu)}, \text{ for } 1 \leq s \leq n. \\ u_s^{(\nu)} &= \text{column vectors of } U^{(\nu)}, \\ v_s^{(\nu)} &= \text{column vectors of } V^{(\nu)}. \end{aligned}$$

- The vector $c^{(\nu+1)}$ is obtained.
- $c^{(\nu+1)} \Rightarrow W^{(\nu)}$.

Determine $H^{(\nu)}$ and $K^{(\nu)}$

- For $n+1 \leq i \leq m$ and $1 \leq j \leq n$,

$$\tilde{H}_{ij}^{(\nu)} = -\tilde{H}_{ji}^{(\nu)} = -\frac{W_{ij}^{(\nu)}}{\sigma_j^*}.$$

- For $1 \leq i < j \leq n$,

$$\begin{aligned} W_{ij}^{(\nu)} &= \Sigma_{ii}\tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)}\Sigma_{jj}, \\ W_{ji}^{(\nu)} &= \Sigma_{jj}\tilde{K}_{ji}^{(\nu)} - \tilde{H}_{ji}^{(\nu)}\Sigma_{ii} \\ &= -\Sigma_{jj}\tilde{K}_{ij}^{(\nu)} + \tilde{H}_{ij}^{(\nu)}\Sigma_{ii}. \end{aligned}$$

  Solving for $\tilde{H}_{ij}^{(\nu)}$ and $\tilde{K}_{ij}^{(\nu)}$ $\Rightarrow$

$$\begin{aligned} \tilde{H}_{ij}^{(\nu)} = -\tilde{H}_{ji}^{(\nu)} &= \frac{\sigma_i^* W_{ji}^{(\nu)} + \sigma_j^* W_{ij}^{(\nu)}}{(\sigma_i^*)^2 - (\sigma_j^*)^2}, \\ \tilde{K}_{ij}^{(\nu)} = -\tilde{K}_{ji}^{(\nu)} &= \frac{\sigma_i^* W_{ij}^{(\nu)} + \sigma_j^* W_{ji}^{(\nu)}}{(\sigma_i^*)^2 - (\sigma_j^*)^2}. \end{aligned}$$

- The intercept is now completely found.

Find the Lift-Up

- Define orthogonal matrices

$$R := (I + \frac{H^{(\nu)}}{2})(I - \frac{H^{(\nu)}}{2})^{-1},$$
$$S := (I + \frac{K^{(\nu)}}{2})((I - \frac{K^{(\nu)}}{2})^{-1}.$$

- Define the lifted matrix on $\mathcal{M}_s(\Sigma)$:

$$X^{(\nu+1)} := R^T X^{(\nu)} S.$$

- Observe

$$X^{(\nu+1)} \approx R^T(e^{H^{(\nu)}} B(c^{(\nu+1)}) e^{-K^{(\nu)}})S$$

and

$$R^T e^{H^{(\nu)}} \approx I_m$$
$$e^{-K^{(\nu)}} S \approx I_n,$$

if $\|H^{(\nu)}\|$ and $\|K^{(\nu)}\|$ are small.

- For computation,
  - ◇ Only need orthogonal matrices

$$\begin{aligned} U^{(\nu+1)} &:= R^T U^{(\nu)} \\ V^{(\nu+1)} &:= S^T V^{(\nu)}. \end{aligned}$$

  - ◇ Does not need to form $X^{(\nu+1)}$ explicitly.

Quadratic Convergence

- Measure the discrepancy between $(U^{(\nu)}, V^{(\nu)}) \in R^{m\times m} \times R^{n\times n}$ in the induced Frobenius norm.
- Observe:
  - ◇ Suppose:
    - ▷ The ISVP has an exact solution at c^*.
    - ▷ SVD of $B(c^*) = \hat{U}\Sigma\hat{V}^T$.
  - ◇ Define error matrix::

$$E := (E_1, E_2) := (U - \hat{U}, V - \hat{V}).$$

  - ◇ If $U\hat{U}^T = e^H$ and $V\hat{V}^T = e^K$, then

$$\begin{aligned} U\hat{U}^T &= (E_1 + \hat{U})\hat{U}^T \\ &= I_m + E_1\hat{U}^T \\ &= e^H = I_m + H + O(||H||^2). \end{aligned}$$

    and a similar expression for $V\hat{V}^T$.
  - ◇ Thus,

$$||(H, K)|| = O(||E||).$$

- At the ν-th stage, define

$$E^{(\nu)} := (E_1^{(\nu)}, E_2^{(\nu)}) = (U^{(\nu)} - \hat{U}, V^{(\nu)} - \hat{V}).$$

- How far is ${U^{(\nu)}}^T B(c^*) V^{(\nu)}$ away from Σ?

  ◇ Write

$$\begin{aligned} {U^{(\nu)}}^T B(c^*) V^{(\nu)} &:= e^{-\hat{H}^{(\nu)}} \Sigma e^{\hat{K}^{(\nu)}} \\ &:= ({U^{(\nu)}}^T e^{-H_*^{(\nu)}} U^{(\nu)}) \Sigma ({V^{(\nu)}}^T e^{K_*^{(\nu)}} V^{(\nu)}) \end{aligned}$$

  with

$$\begin{aligned} H_*^{(\nu)} &:= U^{(\nu)} \hat{H}^{(\nu)} {U^{(\nu)}}^T, \\ K_*^{(\nu)} &:= V^{(\nu)} \hat{K}^{(\nu)} {V^{(\nu)}}^T, \end{aligned}$$

  ◇ Then

$$\begin{aligned} e^{H_*^{(\nu)}} &= U^{(\nu)} \hat{U}^T, \\ e^{K_*^{(\nu)}} &= V^{(\nu)} \hat{V}^T. \end{aligned}$$

  ◇ So

$$||(H_*^{(\nu)}, K_*^{(\nu)})|| = O(||E^{(\nu)}||).$$

  ◇ Norm invariance under orthogonal transformations $\Rightarrow$

$$||(\hat{H}^{(\nu)}, \hat{K}^{(\nu)})|| = O(||E^{(\nu)}||).$$

- Rewrite
$$U^{(\nu)^T} B(c^*) V^{(\nu)} = \Sigma + \Sigma \hat{K}^{(\nu)} - \hat{H}^{(\nu)} \Sigma + O(\|E^{(\nu)}\|^2).$$

- Compare:
$$\begin{aligned} & U^{(\nu)^T} (B(c^*) - B(c^{(\nu+1)})) V^{(\nu)} \\ = \ & \Sigma(\hat{K}^{(\nu)} - \tilde{K}^{(\nu)}) - (\hat{H}^{(\nu)} - \tilde{H}^{(\nu)})\Sigma \\ + \ & O(||E^{(\nu)}||^2). \end{aligned}$$

- Diagonal elements $\Rightarrow$
$$J^{(\nu)}(c^* - c^{(\nu+1)}) = O(||E^{(\nu)}||^2).$$
  - ⋄ Thus
$$||c^* - c^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

- Off-diagonal elements $\Rightarrow$
$$\begin{aligned} ||\hat{H}^{(\nu)} - \tilde{H}^{(\nu)}|| &= O(||E^{(\nu)}||^2), \\ ||\hat{K}^{(\nu)} - \tilde{K}^{(\nu)}|| &= O(||E^{(\nu)}||^2). \end{aligned}$$
  - ⋄ Therefore,
$$||(\tilde{H}^{(\nu)}, \tilde{K}^{(\nu)})|| = O(||E^{(\nu)}||).$$

- Together,
$$\begin{aligned} ||H^{(\nu)} - H_*^{(\nu)}|| &= O(||E^{(\nu)}||^2), \\ ||K^{(\nu)} - K_*^{(\nu)}|| &= O(||E^{(\nu)}||^2). \end{aligned}$$

- Observe:

$$\begin{aligned} E_1^{(\nu+1)} &:= U^{(\nu+1)} - \hat{U} = R^T U^{(\nu)} - e^{-H_*^{(\nu)}} U^{(\nu)} \\ &= \left[(I - \frac{H^{(\nu)}}{2}) - (I - H_*^{(\nu)} + O(||H_*^{(\nu)}||^2) \right. \\ &\quad \left. (I + \frac{H^{(\nu)}}{2}) \right] (I + \frac{H^{(\nu)}}{2})^{-1} U^{(\nu)} \\ &= \left[H_*^{(\nu)} - H^{(\nu)} + O(||H_*^{(\nu)} H^{(\nu)}|| \right. \\ &\quad \left. + ||H^{(\nu)}||^2) \right] (I + \frac{H^{(\nu)}}{2})^{-1} U^{(\nu)}. \end{aligned}$$

◇ It is clear now that

$$||E_1^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

- A similar argument works for $E_2^{(\nu+1)}$.
- We have proved that

$$||E^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

Multiple Singular Values

- Previous definition in finding the $\mathcal{B}$-intercept of a tangent line of $\mathcal{M}_s(\Sigma)$ allows
  - ◇ No zero singular values.
  - ◇ No multiple singular values.
- Now assume
  - ◇ All singular values are positive.
  - ◇ Only the first singular value σ_1^* is multiple, with multiplicity p.

- Observe:
  - ◇ All formulas work, except
    - ▷ For $1 \leq i < j \leq p$, only know

$$W_{ij}^{(\nu)} + W_{ji}^{(\nu)} = 0.$$

    - ▷ No values for $\tilde{H}_{ij}^{(\nu)}$ and $\tilde{K}_{ij}^{(\nu)}$ can be determined.
    - ▷ Additional $q := \frac{p(p-1)}{2}$ equations for the vector $c^{(\nu+1)}$ arise.
- Multiple singular values gives rise to an overdetermined system for $c^{(\nu+1)}$.
  - ◇ Tangent lines from $\mathcal{M}_s(\Sigma)$ may not intercept the affine subspace $\mathcal{B}$ at all.
  - ◇ The ISVP needs to be modified.

Modified ISVP

- Given
  - ◇ Positive values $\sigma_1^* = \ldots = \sigma_p^* > \sigma_{p+1}^* > \ldots > \sigma_{n-q}^*$,
- Find
  - ◇ Real values of $c_1, \ldots, c_n$,
  - ◇ The $n-q$ largest singular values of the matrix matrix $B(c)$ are $\sigma_1^*, \ldots, \sigma_{n-q}^*$.

Find the Intercept

- Use the equation

$$\hat{\Sigma} + \hat{\Sigma}\tilde{K}^{(\nu)} - \tilde{H}^{(\nu)}\hat{\Sigma} = {U^{(\nu)}}^{T} B(c^{(\nu+1)}) V^{(\nu)}$$

  to find the $\mathcal{B}$-intercept where

  ◇ The diagonal matrix

$$\hat{\Sigma} := diag\{\sigma_1^*, \ldots, \sigma_{n-q}^*, \hat{\sigma}_{n-q+1}, \ldots, \hat{\sigma}_n\}$$

  ◇ Additional singular values $\hat{\sigma}_{n-q+1}, \ldots, \hat{\sigma}_n$ are free parameters.

The Algorithm

Given $U^{(\nu)} \in \mathcal{O}(m)$ and $V^{(\nu)} \in \mathcal{O}(n)$,

- Solve for $c^{(\nu+1)}$ from the system of equations:

$$\sum_{k=1}^{n} \left(u_i^{(\nu)^T} B_k v_i^{(\nu)} \right) c_k^{(\nu+1} = \sigma_i^* - u_i^{(\nu)^T} B_0 v_i^{(\nu)},$$

for $i = 1, \ldots, n-q$

$$\sum_{k=1}^{n} \left(u_s^{(\nu)^T} B_k v_t^{(\nu)} + u_t^{(\nu)^T} B_k v_s^{(\nu)} \right) c_k^{(\nu+1)} =$$

$$-u_s^{(\nu)^T} B_0 v_t^{(\nu)} - u_t^{(\nu)^T} B_0 v_s^{(\nu)},$$

for $1 \le s < t \le p$.

- Define $\hat{\sigma}_k^{(\nu)}$ by

$$\hat{\sigma}_k^{(\nu)} := \begin{cases} \sigma_k^*, & \text{if } 1 \le k \le n-q; \\ u_k^{(\nu)^T} B(c^{(\nu+1)}) v_k^{(\nu)}, & \text{if } n-q < k \le n \end{cases}$$

- Once $c^{(\nu+1)}$ is determined, calculate $W^{(\nu)}$.

- Define skew symmetric matrices $\tilde{K}^{(\nu)}$ and $\tilde{H}^{(\nu)}$:
  - ◇ For $1 \le i < j \le p$, the equation to be satisfied is

$$W_{ij}^{(\nu)} = \hat{\sigma}_i^{(\nu)} \tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)} \hat{\sigma}_j^{(\nu)}.$$

    - ▷ Many ways to define $\tilde{K}_{ij}^{(\nu)}$ and $\tilde{H}_{ij}^{(\nu)}$.
    - ▷ Set $\tilde{K}_{ij}^{(\nu)} \equiv 0$ for $1 \le i < j \le p$.
  - ◇ $\tilde{K}^{(\nu)}$ is defined by

$$\tilde{K}_{ij}^{(\nu)} := \begin{cases} \frac{\hat{\sigma}_i^{(\nu)} W_{ij}^{(\nu)} + \hat{\sigma}_j^{(\nu)} W_{ji}^{(\nu)}}{(\hat{\sigma}_i^{(\nu)})^2 - (\hat{\sigma}_j^{(\nu)})^2}, & \text{if } 1 \le i < j \le n;\ p < j; \\ 0, & \text{if } 1 \le i < j \le p \end{cases}$$

  - ◇ $\tilde{H}^{(\nu)}$ is defined by

$$\tilde{H}_{ij}^{(\nu)} := \begin{cases} -\frac{W_{ij}^{(\nu)}}{\hat{\sigma}_j^{(\nu)}}, & \text{if } 1 \le i < j \le p; \\ -\frac{W_{ij}^{(\nu)}}{\hat{\sigma}_j^{(\nu)}}, & \text{if } n+1 \le i \le m;\ 1 \le j \le n; \\ \frac{\hat{\sigma}_i^{(\nu)} W_{ji}^{(\nu)} + \hat{\sigma}_j^{(\nu)} W_{ij}^{(\nu)}}{(\hat{\sigma}_i^{(\nu)})^2 - (\hat{\sigma}_j^{(\nu)})^2}, & \text{if } 1 \le i < j \le n;\ p < j; \\ 0, & \text{if } n+1 \le i \ne j \le m. \end{cases}$$

- Once $\tilde{H}^{(\nu)}$ and $\tilde{K}^{(\nu)}$ are determined, proceed the lifting in the same way as for the ISVP.

Remarks

- No longer on a fixed manifold $\mathcal{M}_s(\Sigma)$ since $\hat{\Sigma}$ is changed per step.
- The algorithm for multiple singular value case converges quadratically.

Zero Singular Value

- Zero singular value $\Rightarrow$ rank deficiency.
- Finding a lower rank matrix in a generic affine subspace $\mathcal{B}$ is intuitively a more difficult problem.
- More likely the ISVP does not have a solution.
- Consider the simplest case where $\sigma_1^* > \ldots > \sigma_{n-1}^* > \sigma_n^* = 0$.
  - ◇ Except for $\tilde{H}_{in}$ (and $\tilde{H}_{ni}$), $i = n+1, \ldots, m$, all other quantities including $c^{(\nu+1)}$ are well-defined.
  - ◇ It is necessary that
  $$W_{in}^{(\nu)} = 0 \text{ for } i = n+1, \ldots, m.$$
  - ◇ If the necessary condition fails, then no tangent line of $\mathcal{M}_s(\Sigma)$ from the current iterate $X^{(\nu)}$ will intersect the affine subspace $\mathcal{B}$.

Example of the Continuous Approach

- Integrator — Subroutine ODE (Shampine et al, '75).
  - ◇ ABSERR and RELERR $= 10^{-12}$.
  - ◇ Output values examined at interval of 10.
- Two consecutive output points differ by less than 10^{-10} $\Rightarrow$ Convergence.
- Stable equilibrium point is not necessarily a solution to the ISVP.
- Change to different initial value $X(0)$ if necessary.

Example of the Iterative Approach

- Easy implementation by MATLAB.
  - ◇ Consider the case when $m = 5$ and $n = 4$.
  - ◇ Randomly generated basis matrices by the Gaussian distribution.
- Numerical experiment meant solely to examine the behavior of quadratic convergence.
  - ◇ Randomly generate a vector $c^{\#} \in R^4$.
  - ◇ Singular values of $B(c^{\#})$ used as the prescribed singular values.
  - ◇ Perturb each entry of $c^{\#}$ by a uniform distribution between -1 and 1.
  - ◇ Use the perturbed vector as the initial guess.

Observations

- The limit point c^* is not necessary the same as the original vector $c^{\#}$.
- Singular values of $B(c^*)$ do agree with those of $B(c^{\#})$.
- Differences between singular values of $B(c^{(\nu)})$ and $B(c^*)$ are measured in the 2-norm.
- Quadratic convergence is observed.

Example of Multiple Singular Values

- Construction of an example is not trivial.
  - ◇ Same basis matrices as before.
  - ◇ Assume $p = 2$.
  - ◇ Prescribed singular values $\sigma^* = (5, 5, 2)^T$.
  - ◇ Initial guess of $c^{(0)}$ is searched by trials
- The order of singular values could be altered.
  - ◇ The value 5 is no longer the largest singular value.
  - ◇ Unless the initial guess $c^{(0)}$ is close enough to an exact solution c^*, no reason to expect that the algorithm will preserve the ordering.
  - ◇ Once convergence occurs, then σ^* must be part of the singular values of the final matrix.
- At the initial stage the convergence is slow, but eventually the rate is picked up and becomes quadratic.

Inverse Singualar/Eigenvalue Problem

- Overview
- A Recursive Algorithm
- The Matrix Structure
- Numerical Experiment

Overview

- The Schur-Horn Theorem gives the connection between diagonal entries and eigenvalues of a Hermitian matrix.
- The Mirsky Theorem gives a connection between diagonal entries and eigenvalues of a general matrix.
- The Sing-Thompson Theorem gives the connection between diagonal entries and singular values of a general matrix.
- What is the connection between singular values and eigenvalues of a matrix?
  - ◇ singular value = |eigenvalue|, if Hermitian matrices.
  - ◇ How about general matrices?
- Can we create matrices with prescribed singular values and eigenvalues?
  - ◇ Desirable for test matrices.

Weyl-Horn Theorem

- Given vectors $\lambda \in C^n$ and $\alpha \in \mathbb{R}^n$,
  - ◇ Assume
$$\begin{aligned} |\lambda_1| &\geq \ldots \geq |\lambda_n|, \\ \alpha_1 &\geq \ldots \geq \alpha_n. \end{aligned}$$
  - ◇ Then a matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and singular values $\alpha_1, \ldots, \alpha_n$ exists if and only if
$$\begin{aligned} \prod_{j=1}^{k} |\lambda_j| &\leq \prod_{j=1}^{k} \alpha_j, \quad k = 1, \ldots, n-1, \\ \prod_{j=1}^{n} |\lambda_j| &= \prod_{j=1}^{n} \alpha_j. \end{aligned}$$
    - ▷ If $|\lambda_n| > 0$, then $\log \alpha$ majorizes $\log |\lambda|$.
- How to solve the inverse singular eigenvalue problem numerically?

A Recursive Algorithm

- The Building Block — 2×2 Case
- The Original Proof by Induction
- An Innocent Mistake
- A Recursive Clause in Programming

The 2×2 Case

- The Weyl-Horn Condition:

$$\begin{cases} |\lambda_1| & \leq \ \alpha_1, \\ |\lambda_1||\lambda_2| & = \ \alpha_1\alpha_2. \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \alpha_2 \leq |\lambda_2| \leq |\lambda_1| \leq \alpha_1 \\ |\lambda_1|^2 + |\lambda_2|^2 \leq \alpha_1^2 + \alpha_2^2. \end{cases}$$

- The building block — A triangular matrix

$$A = \begin{bmatrix} \lambda_1 & \mu \\ 0 & \lambda_2 \end{bmatrix}$$

has singular value $\{\alpha_1, \alpha_2\}$ if and only if

$$\mu = \sqrt{\alpha_1^2 + \alpha_2^2 - |\lambda_1|^2 - |\lambda_2|^2}.$$

  - ◇ A is complex-valued when eigenvalues are complex.
  - ◇ A stable way of computing μ:

$$\mu = \begin{cases} 0, & \text{if } |(\alpha_1-\alpha_2)^2-(|\lambda_1|-|\lambda_2|)^2| \leq \epsilon \\ \sqrt{|(\alpha_1-\alpha_2)^2-(|\lambda_1|-|\lambda_2|)^2|}, & \text{otherwise.} \end{cases}$$

Ideas in Horn's Proof

- Reduce the original inverse problem to two problems of *smaller sizes.*
- Problems of smaller sizes are guaranteed to be solvable by the *induction hypothesis.*
- The subproblems are *affixed* together by working on a suitable 2×2 *corner*.
- The 2×2 problem has an explicit solution.

Key to the Algorithmic Success

- The eigenvalues and singular values of each of the two subproblems can be derived *explicitly.*
- Each of the two subproblems can further be down-sized.
- The original problem is *divided* into subproblems of size 2×2 or 1×1.
- The smaller problems can be *conquered* to build up the original size.
- In an environment that allows a subprogram to invoke itself recursively, only one-step of the divide-and-conquer procedure will be enough.
- Very similar to the radix-2 FFT $\Longrightarrow$ fast algorithm.

Outline of Proof

- Suppose $\alpha_i > 0$ for all $i = 1, \ldots, n$. So $\lambda_i \neq 0$ for all i.
  - ◇ The case of zero singular values can be handled in a similar way.
- Define
$$\begin{cases} \sigma_1 := \alpha_1, \\ \sigma_i := \sigma_{i-1}\frac{\alpha_i}{|\lambda_i|}, \quad \text{for } i = 2, \ldots, n-1. \end{cases}$$
  - ◇ Assume $\sigma := \min_{1 \leq i \leq n-1} \sigma_i$ occurs at the index j.
- Define
$$\rho := \frac{|\lambda_1 \lambda_n|}{\sigma}.$$
- The following three sets of inequalities are true. The numbers satisfy the Weyl-Horn conditions.
$$\begin{cases} |\lambda_1| \geq |\lambda_n|, \\ \sigma \geq \rho. \end{cases}$$
$$\begin{cases} \sigma \geq |\lambda_2| \geq \ldots \geq |\lambda_j|, \\ \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_j. \end{cases}$$
$$\begin{cases} |\lambda_{j+1}| \geq \ldots \geq |\lambda_{n-1}| \geq \rho, \\ \alpha_{j+1} \geq \ldots \geq \alpha_{n-1} \geq \alpha_n. \end{cases}$$

- By induction hypothesis,
  - ◇ There exist unitary matrices U_1, $V_1 \in C^{j\times j}$ and *triangular* matrices A_1 such that

$$U_1 \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 \\ \vdots & \ddots & & \\ 0 & 0 & \dots & \alpha_j \end{bmatrix} V_1^* = A_1 = \begin{bmatrix} \sigma & \times & \times & \dots & \times \\ 0 & \lambda_2 & & & \times \\ & & & & \\ \vdots & & & \ddots & \\ 0 & 0 & & & \lambda_j \end{bmatrix}.$$

  - ◇ There exist unitary matrices U_2, $V_2 \in C^{(n-j)\times(n-j)}$, and *triangular* matrix A_2 such that

$$U_2 \begin{bmatrix} \alpha_{j+1} & 0 & \dots 0 \\ 0 & \alpha_{j+2} & 0 \\ \vdots & & \ddots \\ 0 & 0 & \dots \alpha_n \end{bmatrix} V_2^* = A_2 = \begin{bmatrix} \lambda_{j+1} & \times & \dots & \times & \times \\ 0 & \lambda_{j+2} & & & \times \\ \vdots & & \ddots & & \vdots \\ & & & \lambda_{n-1} & \times \\ 0 & 0 & \dots & 0 & \rho \end{bmatrix}.$$

- Horn's claim: The block matrix

$$\begin{bmatrix} A_1 & \bigcirc \\ \bigcirc & A_2 \end{bmatrix}$$

can be *permuted* to the triangular matrix

$$\begin{bmatrix}
\lambda_2 & \times & \dots & \times & \times & & & & & \\
0 & & & & \times & & & & & \\
\vdots & & \ddots & & \vdots & & & \bigcirc & & \\
 & & & \lambda_j & \times & & & & & \\
0 & & \dots & 0 & \sigma & 0 & & & & \\
0 & 0 & \dots & 0 & 0 & \rho & \times & \times & \dots & \times \\
 & & & & & & \lambda_{j+1} & & & \times \\
 & & \bigcirc & & & & & & & \\
 & & & & & \vdots & & & \ddots & \\
 & & \dots & & & 0 & 0 & & & \lambda_{n-1}
\end{bmatrix}$$

- The 2×2 corner can now be glued together by

$$U_0 \begin{bmatrix} \sigma & 0 \\ 0 & \rho \end{bmatrix} V_0^* = A_0 = \begin{bmatrix} \lambda_1 & \mu \\ 0 & \lambda_n \end{bmatrix}.$$

- How to do the permutation, or is it a mistake?
  - ◇ It takes more than permutation to rearrange the diagonals of a triangular matrix.

A MATLAB Program

```
function [A]=svd_eig(alpha,lambda);
n = length(alpha);
if n == 1                                    % The 1 by 1 case
   A = [lambda(1)];
elseif n == 2                                % The 2 by 2 case
   [U,V,A] = two_by_two(alpha,lambda);
else                                         % Check zero singular values
   tol = n*alpha(1)*eps;
   k = sum(alpha > tol); m = sum(abs(lambda) > tol);
   if k == n                                 % Nonzero singular values
      j = 1; s = alpha(1); temp = s;
      for i = 2:n-1
         temp = temp*alpha(i)/abs(lambda(i));
         if temp < s, j = i; s = temp; end
      end
      rho = abs(lambda(1)*lambda(n))/s;
      [U0,V0,A0] = two_by_two([s;rho],[lambda(1);lambda(n)]);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
      [A1] = svd_eig(alpha(1:j),[s;lambda(2:j)]);         % RECURSIVE %
      [A2] = svd_eig(alpha(j+1:n),[lambda(j+1:n-1);rho]); %  CALLING  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
      A = [A1,zeros(j,n-j);zeros(n-j,j),A2];
      Temp = A;
         A(1,:)=U0(1,1)*Temp(1,:)+U0(1,2)*Temp(n,:);
         A(n,:)=U0(2,1)*Temp(1,:)+U0(2,2)*Temp(n,:);
      Temp = A;
         A(:,1)=V0(1,1)*Temp(:,1)+V0(1,2)*Temp(:,n);
         A(:,n)=V0(2,1)*Temp(:,1)+V0(2,2)*Temp(:,n);
   else                                      % Zero singular values
      beta = prod(abs(lambda(1:m)))/prod(alpha(1:m-1));
      [U3,V3,A3] = svd_eig([alpha(1:m-1);beta],lambda(1:m));
      A = zeros(n); A(1:m,1:m) = V3'*A3*V3;
      for i = m+1:k, A(i,i+1) = alpha(i); end
      A(m,m+1) = sqrt(abs(alpha(m)^2-beta^2));
   end
end
```

Matrix Structure

- A Modified Proof
- A Symbolic Example

Correct the "Mistake"

- Horn's requirement:
  - ◇ Both intermediate matrices A_1 and A_2 are upper triangular matrices.
  - ◇ Diagonal entries are arranged in a certain order.
    - ▷ Valid from the Schur decomposition theorem.
    - ▷ More than permutation, not easy for computation.
    - ▷ To rearrange diagonal entries via unitary similarity transformations while maintaining the upper triangular structure is expensive.
- Our contribution:
  - ◇ The triangular structure is entirely unnecessary.
  - ◇ The matrix A produced from our algorithm is generally not triangular.
  - ◇ Do not need to rearrange the diagonal entries
  - ◇ Modify the first and the last rows and columns of the block diagonal matrix $\begin{bmatrix} A_1 & \bigcirc \\ \bigcirc & A_2 \end{bmatrix}$, *as if nothing happened*, is enough.

- Algorithm:
  - ◇ Denote $U_0 = [u_{0,st}]$ and $V_0 = [v_{0,st}]$.
  - ◇ Then

$$\begin{bmatrix} u_{0,11} & 0 & u_{0,12} \\ 0 & I_{n-1} & 0 \\ u_{0,21} & 0 & u_{0,22} \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 \\ \vdots & & \ddots & \\ 0 & & & \alpha_n \end{bmatrix} \begin{bmatrix} V_1^* & 0 \\ 0 & V_2^* \end{bmatrix} \begin{bmatrix} v_{0,11} & 0 & v_{0,12} \\ 0 & I_{n-1} & 0 \\ v_{0,21} & 0 & v_{0,22} \end{bmatrix}^*$$

is the desired matrix.

- A has the structure

$$A = \begin{bmatrix} \lambda_1 & \otimes & \dots & \otimes & \otimes & * & * & & \mu \\ \otimes & \lambda_2 & & & \times & 0 & 0 & & * \\ \vdots & & \ddots & & \vdots & & & \bigcirc & \\ & & & \lambda_{j-1} & \times & & & & \\ \otimes & \times & \dots & \times & \lambda_j & & & & * \\ * & 0 & \dots & 0 & 0 & \lambda_{j+1} & \times & \times & \dots & \otimes \\ * & & & & 0 & \times & \lambda_{j+2} & & \otimes \\ & & \bigcirc & & & & & & \\ & & & & & \vdots & & \ddots & \\ 0 & * & \dots & * & * & \otimes & \otimes & & \lambda_n \end{bmatrix}.$$

  - ◇ $\times$ = unchanged, original entries from A_1 or A_2.
  - ◇ $\otimes$ = entries of A_1 or A_2 that are modified by scalar multiplications.
  - ◇ $*$ = possible new entries that were originally zero.

A Variation of Horn's Proof

- Does the algorithm really works?
  - ◇ Clearly, A has singular values $\{\alpha_1, \ldots, \alpha_n\}$.
  - ◇ Need to show that A has eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$.
- What has been changed?

(P1) Diagonal entries of A_1 and A_2 are in fixed orders, $\sigma, \lambda_2, \ldots, \lambda_j$ and $\lambda_{j+1}, \ldots, \lambda_{n-1}, \rho$, respectively.

(P2) Each A_i is similar through permutations, which need not to be known, to a lower triangular matrix whose diagonal entries constitute the same set as the diagonal entries of A_i. (Thus, each A_i has precisely its own diagonal entries as its eigenvalues.)

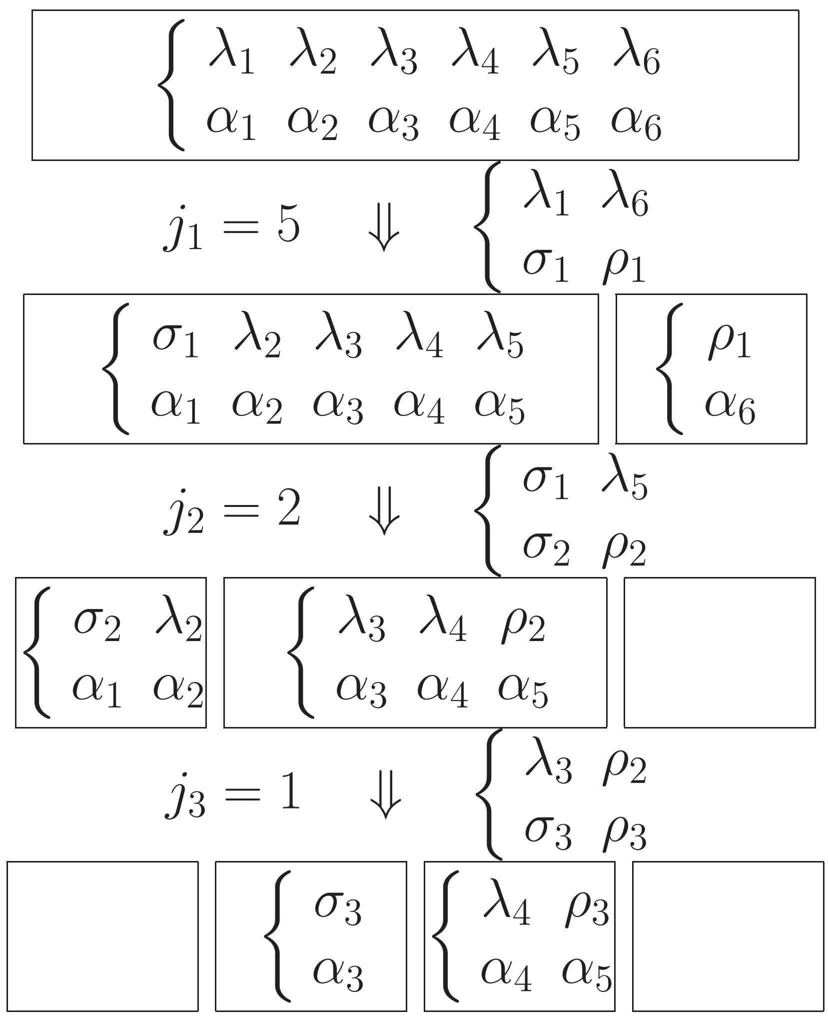
(P3) The first row and the last row have the same zero pattern except that the lower-left corner is always zero.

(P4) The first column and the last column have the same zero pattern except that the lower-left corner is always zero.

- Use graph theory to show that the affixed matrix A has exactly the same properties.

A Symbolic Example

- Dividing process:



- Conquering process:

$$\begin{bmatrix} \lambda_1 & \otimes & 0 & 0 & \otimes & * \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ \otimes & 0 & \lambda_3 & 0 & \times & * \\ \otimes & 0 & \times & \lambda_4 & \times & * \\ 0 & \times & 0 & 0 & \lambda_5 & 0 \\ 0 & * & 0 & 0 & * & \lambda_6 \end{bmatrix}$$

$$j_1 = 5 \quad \Uparrow \quad \begin{bmatrix} \sigma_1 & \times & 0 & 0 & \times & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ \times & 0 & \lambda_3 & 0 & \times & 0 \\ \times & 0 & \times & \lambda_4 & \times & 0 \\ 0 & \times & 0 & 0 & \lambda_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_1 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_1 & \otimes & 0 & 0 & * \\ 0 & \lambda_2 & 0 & 0 & 0 \\ * & 0 & \lambda_3 & 0 & \otimes \\ * & 0 & \times & \lambda_4 & \otimes \\ 0 & * & 0 & 0 & \lambda_5 \end{bmatrix} \qquad \begin{bmatrix} \rho_1 \end{bmatrix}$$

$$j_2 = 2 \quad \Uparrow \quad \begin{bmatrix} \sigma_2 & \times & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & \times \\ 0 & 0 & \times & \lambda_4 & \times \\ 0 & 0 & 0 & 0 & \rho_2 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_2 & * \\ 0 & \lambda_2 \end{bmatrix} \qquad \begin{bmatrix} \lambda_3 & 0 & * \\ * & \lambda_4 & \otimes \\ 0 & 0 & \rho_2 \end{bmatrix}$$

$$j_3 = 1 \quad \Uparrow \quad \begin{bmatrix} \sigma_3 & 0 & 0 \\ 0 & \lambda_4 & \times \\ 0 & 0 & \rho_3 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_3 \end{bmatrix} \qquad \begin{bmatrix} \lambda_4 & * \\ 0 & \rho_3 \end{bmatrix}$$

Numerical Experiment

- The divide-and-conquer feature brings on fast computation.
- The overall cost is estimated at the order of $O(n^2)$.
- A numerical simulation:

Rosser Test

- Rosser matrix R:

$$R = \begin{bmatrix} 611 & 196 & -192 & 407 & -8 & -52 & -49 & 29 \\ 196 & 899 & 113 & -192 & -71 & -43 & -8 & -44 \\ -192 & 113 & 899 & 196 & 61 & 49 & 8 & 52 \\ 407 & -192 & 196 & 611 & 8 & 44 & 59 & -23 \\ -8 & -71 & 61 & 8 & 411 & -599 & 208 & 208 \\ -52 & -43 & 49 & 44 & -599 & 411 & 208 & 208 \\ -49 & -8 & 8 & 59 & 208 & 208 & 99 & -911 \\ 29 & -44 & 52 & -23 & 208 & 208 & -911 & 99 \end{bmatrix}.$$

  ◇ Has one double eigenvalue, three nearly equal eigenvalues, one zero eigenvalue, two dominant eigenvalues of opposite sign and one small nonzero eigenvalue.

  ◇ The *computed* eigenvalues and singular values of R are

$$\lambda = \begin{bmatrix} -1.020049018429997e{+}03 \\ 1.020049018429997e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359278e{+}03 \\ 1.000000000000001e{+}03 \\ 9.999999999999998e{+}02 \\ 9.804864072152601e{-}02 \\ 4.851119506099622e{-}13 \end{bmatrix}, \alpha = \begin{bmatrix} 1.020049018429997e{+}03 \\ 1.020049018429996e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359279e{+}03 \\ 1.000000000000000e{+}03 \\ 9.999999999999998e{+}02 \\ 9.804864072162672e{-}02 \\ 1.054603342667098e{-}14 \end{bmatrix}.$$

- Using the above λ and α,
  - ◇ A nonsymmetric matrix is produced:

$$\begin{bmatrix} 1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0199e{+}03 & 0 & 0 & 1.4668e{-}09 & 0 \\ 0 & 0 & 0 & 0 & 1.0000e{+}03 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0000e{+}03 & 0 & 0 \\ 0 & 0 & 0 & -1.5257e{-}05 & 0 & 0 & 9.8049e{-}02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.4045e{-}07 & 0 \end{bmatrix}.$$

  - ◇ The re-computed eigenvalues and singular values of A are

$$\hat{\lambda} = \begin{bmatrix} -1.020049018429997e{+}03 \\ 1.020049018429997e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359278e{+}03 \\ 1.000000000000001e{+}03 \\ 9.999999999999998e{+}02 \\ 9.80486407215721e{-}02 \\ 0 \end{bmatrix}, \hat{\alpha} = \begin{bmatrix} 1.020049018429997e{+}03 \\ 1.020049018429997e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359279e{+}03 \\ 1.000000000000001e{+}03 \\ 9.999999999999998e{+}02 \\ 9.804864072162672e{-}02 \\ 0 \end{bmatrix}.$$

  - ◇ The re-computed eigenvalues and singular values agree with those of R up to the machine accuracy.

Wilkinson Test

- Wilkinson's matrices:
  - ◇ All are symmetric and tridiagonal.
  - ◇ Have nearly, but not exactly, equal eigenvalue pairs.
- Using these data:
  - ◇ Discrepancy in eigenvalues and singular values between our constructed matrices and Wilkinson's matrices.
  - ◇ Matrices constructed are nearly but not symmetric.

Conclusion

- Weyl-Horn Theorem completely characterizes the relationship between eigenvalues and singular values of a general matrix.
- The original proof has been modified.
- With the aid of programming languages that allow a subprogram to invoke itself recursively, an induction proof can be implemented as a recursive algorithm.
- The resulting algorithm is fast. The cost of construction is approximately $O(n^2)$.
- The matrix being constructed usually is not symmetric and is complex-valued, if complex eigenvalues are present.
- Numerical experiment on some very challenging problems suggests that our method is quite robust.

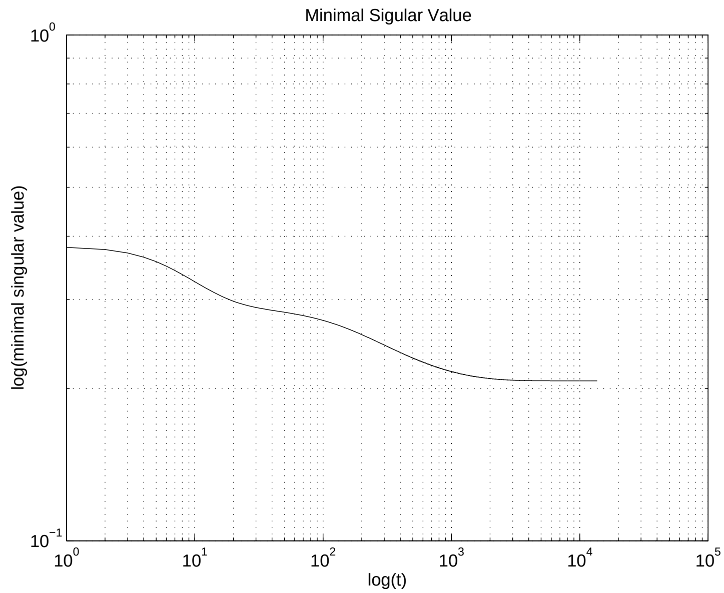


Figure 11: History of the smallest singular value for Example 3.

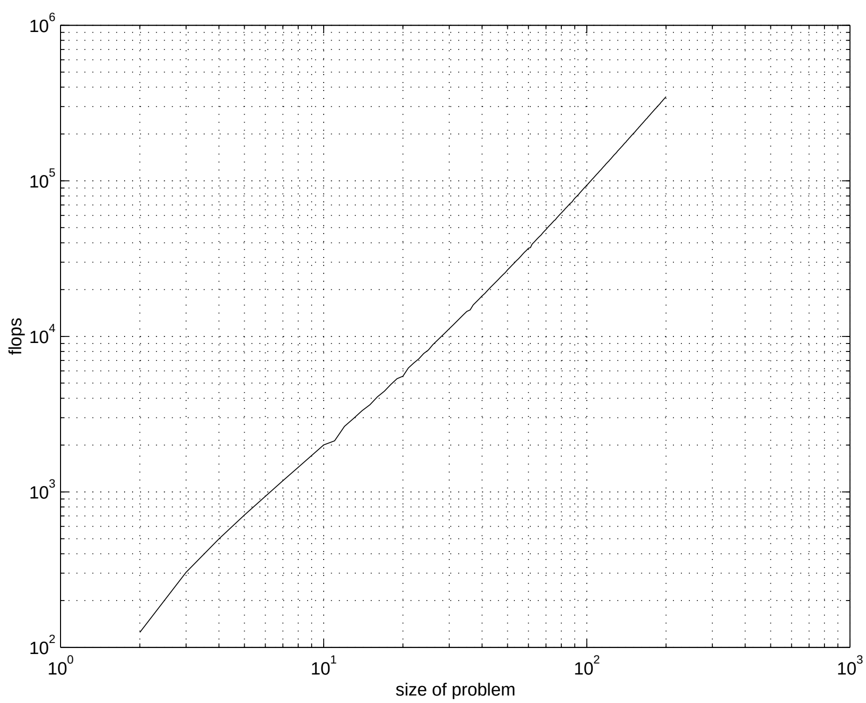


Figure 12: log-log plot of computational flops versus problem sizes

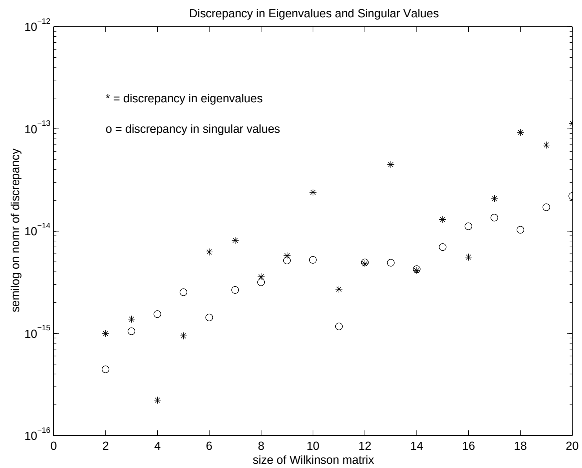


Figure 13: L_2 norm of discrepancy in eigenvalues and singular values.

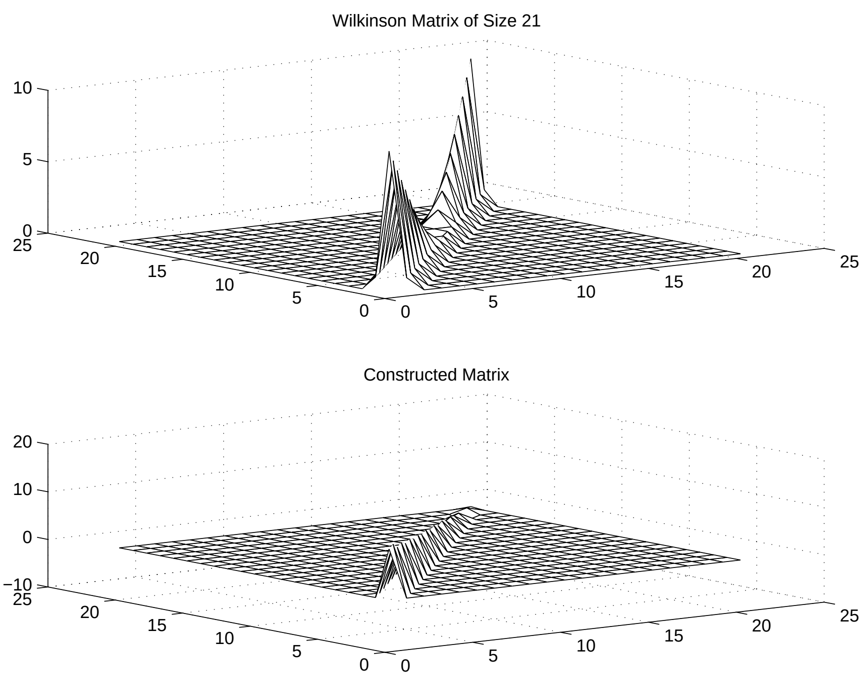


Figure 14: 3-D mesh representation of 21×21 matrices