Structure Preserving Eigensolvers

GOALS

- Solve complete eigenproblem $Ax = \lambda x$
  - A is real and doubly-structured:

symmetric or *skew-symmetric*
and
persymmetric or *perskew-symmetric*

- Algorithms must
  - preserve both structures
  - have good convergence behavior
  - be parallelizable
  - be stable, accurate

I start by looking at a 2 x 2 matrix.
Sometimes I look at a 4 x 4 matrix.
That's when things get out of control and too hard.
Paul Halmos

JACOBI METHOD (1846)

Goal

$$\begin{bmatrix} \textit{real} \\ \textit{symmetric} \end{bmatrix}_{n\times n} \rightsquigarrow \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}_{n\times n}$$

Idea

Repeat until convergence
1. Choose 2×2 structured subproblem
2. Diagonalize it

End Repeat

Structured Subproblem
Choose $\{c, s \mid c^2 + s^2 = 1\}$ such that

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c \cdots s & & \\ & & \vdots \ddots \vdots & & \\ & & -s \cdots c & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} * & & & & \\ & \ddots & & & \\ & & \bullet \cdots \times & & \\ & & \vdots \ddots \vdots & & \\ & & \times \cdots \bullet & & \\ & & & \ddots & \\ & & & & * \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c \cdots -s & & \\ & & \vdots \ddots \vdots & & \\ & & s \cdots c & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \begin{bmatrix} * & & & & \\ & \ddots & & & \\ & & \bullet \cdots 0 & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots \bullet & & \\ & & & \ddots & \\ & & & & * \end{bmatrix}$$

OUR METHOD

2×2 Subproblem Insufficient
- $\mathcal{PO}(2) = \{I, -I, R, -R\}$

Choose 4×4 Subproblem
- Smaller problem must have same double structure
- Drive 4×4 matrix to canonical form

Exploit Isomorphism between $\mathbb{R}^{4\times 4}$ and $\mathbb{H} \otimes \mathbb{H}$
- *Quaternions:*
  $\mathbb{H} = \{q = q_0 + q_1 i + q_2 j + q_3 k : \ q_0, q_1, q_2, q_3 \in \mathbb{R}\}$,
  where $i^2 = j^2 = k^2 = ijk = -1$

Translate Problem to $\mathbb{H} \otimes \mathbb{H}$ and Solve
- Problem is much simpler in $\mathbb{H} \otimes \mathbb{H}$

Translate Back to $\mathbb{R}^{4\times 4}$
- Embed 4×4 solution into I_n
- Perform similarity transformation

Continue Until Convergence
- Will converge if subproblems chosen properly

STRUCTURED MATRICES

$$R = \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}_{n\times n}$$

Orthogonal $\mathcal{O}(n) = \{Q \in \mathbb{R}^{n\times n} : \ Q^TQ = QQ^T = I\}$	$\frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$
Perplectic $\mathcal{P}(n) = \{P \in \mathbb{R}^{n\times n} : \ P^TRP = R\}$	$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$
Perplectic Orthogonal $\mathcal{PO}(n) = \{P \in \mathcal{O}(n) : \ RP = PR\}$	$\frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}$
Persymmetric $p\,\mathcal{S}(n) = \{S \in \mathbb{R}^{4\times 4} : (RS)^T = RS\}$	$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 4 & 3 \\ 7 & 4 & 6 & 2 \\ 4 & 7 & 5 & 1 \end{bmatrix}$
Perskew-symmetric $p\,\mathcal{K}(n) = \{S \in \mathbb{R}^{4\times 4} : (RS)^T = -RS\}$	$\begin{bmatrix} 1 & -2 & 3 & 0 \\ -5 & 6 & 0 & -3 \\ 7 & 0 & -6 & 2 \\ 0 & -7 & 5 & -1 \end{bmatrix}$

PROPERTIES

Preserves Symmetric Structure
- Off-diagonal elements become 0

Quadratic Convergence
- P. Henrici (1958)
- H.P.M van Kempen (1966)

Rich Inherent Parallelism
- A. Sameh (1971)
- P.J. Eberlien (1987)

More Accurate Than QR Algorithm
- J.W. Demmel and K. Veseliè (1992)
- R. Mathias (1995)

Stopping Criteria $\sim$ *log n* Sweeps
- R.P. Brent and F.T. Luk (1895)

CANONICAL FORMS

Structure	4×4	$n \times n$
Symmetric Persymmetric	$\begin{bmatrix} a_1 & 0 & 0 & b_1 \\ 0 & a_2 & b_2 & 0 \\ 0 & b_2 & a_2 & 0 \\ b_1 & 0 & 0 & a_1 \end{bmatrix}$	$\begin{bmatrix} a_1 & & 0 & & b_1 \\ & a_2 & & b_2 & \\ 0 & & \times & & 0 \\ & b_2 & & a_2 & \\ b_1 & & 0 & & a_1 \end{bmatrix}$
Skew-symmetric Persymmetric	$\begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & 0 \\ 0 & -b & 0 & 0 \\ -a & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & & -b & -a \\ & & / & \\ a & b & & 0 \end{bmatrix}$
Symmetric Perskew-symmetric	$\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & -b & 0 \\ 0 & 0 & 0 & -a \end{bmatrix}$	$\begin{bmatrix} a & b & & 0 \\ & & \diagdown & \\ 0 & & -b & -a \end{bmatrix}$
Skew-symmetric Perskew-symmetric	$\begin{bmatrix} 0 & a & b & 0 \\ -a & 0 & 0 & -b \\ -b & 0 & 0 & -a \\ 0 & b & a & 0 \end{bmatrix}$	$\begin{bmatrix} A_1 & & 0 & & B_1 \\ & A_2 & & B_2 & \\ 0 & & z & & 0 \\ & -B_2 & & -A_2 & \\ -B_1 & & 0 & & -A_1 \end{bmatrix}$

STRUCTURED MATRICES

Theorem

1. *For any symmetric persymmetric $A \in \mathbb{R}^{n\times n}$ there exists a perplectic orthogonal $P \in \mathcal{PO}(n)$ such that P^TAP is in canonical form.*
2. *For any skew-symmetric persymmetric $A \in \mathbb{R}^{n\times n}$ there exists a perplectic orthogonal $P \in \mathcal{PO}(n)$ such that P^TAP is in canonical form.*
3. *For any symmetric perskew-symmetric $A \in \mathbb{R}^{n\times n}$ there exists a perplectic orthogonal $P \in \mathcal{PO}(n)$ such that P^TAP is in canonical form.*
4. *For any skew-symmetric perskew-symmetric $A \in \mathbb{R}^{n\times n}$ there exists a perplectic orthogonal $P \in \mathcal{PO}(n)$ such that P^TAP is in canonical form.*

Daniel M. Dunlavy *
Western Michigan University

D. Steven Mackey
Kalamazoo College

Niloufer Mackey
Western Michigan University

SUBPROBLEM ALGORITHM

Input: *skew-symmetric persymmetric* matrix $A \in \mathbb{R}^{4\times4}$

Output: *perplectic orthogonal* matrix $P \in \mathbb{R}^{4\times4}$ such that

$$PAP^T = \begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & 0 \\ 0 & -b & 0 & 0 \\ -a & 0 & 0 & 0 \end{bmatrix}$$

Note: Eigenvalues of A are ai, $-ai$, bi, $-bi$

Algorithm:

$$u = \begin{bmatrix} -A_{12} & -\frac{1}{2}(A_{14}+A_{23}) \end{bmatrix}$$

$$v = \begin{bmatrix} A_{13} & \frac{1}{2}(A_{14}-A_{23}) \end{bmatrix}$$

$$d_1 = \|u\|_2 + u_2$$

$$d_2 = \|v\|_2 + v_2$$

$$\alpha = 2\sqrt{\|u\|_2\,\|v\|_2\,d_1\,d_2}$$

$$P = \frac{1}{\alpha}\begin{bmatrix} d_1 & 0 & u_1 & 0 \\ 0 & d_1 & 0 & -u_1 \\ -u_1 & 0 & d_1 & 0 \\ 0 & u_1 & 0 & d_1 \end{bmatrix}\begin{bmatrix} d_2 & v_1 & 0 & 0 \\ -v_1 & d_2 & 0 & 0 \\ 0 & 0 & d_2 & -v_1 \\ 0 & 0 & v_1 & d_2 \end{bmatrix}$$

SUBPROBLEM EXAMPLE

Input:

$$A = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$

Algorithm:

$$u = \begin{bmatrix} -1 & 0 \end{bmatrix}$$

$$v = \begin{bmatrix} -1 & 0 \end{bmatrix}$$

$$d_1 = 1$$

$$d_2 = 1$$

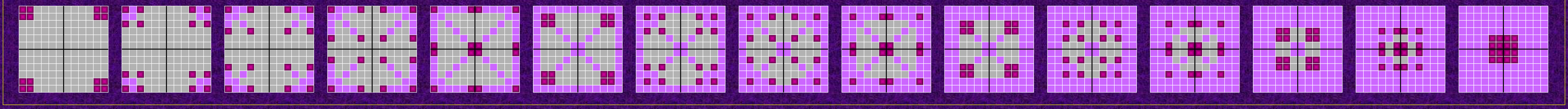
$$\alpha = 2$$

$$P = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \Longrightarrow PAP^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

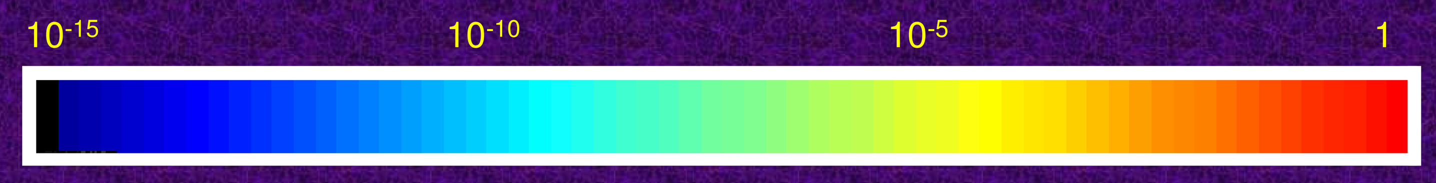
MATLAB: `eig(A)` =

```
-1.45249e-016 + 1.99999e+000i
-1.45249e-016 - 1.99999e+000i
-4.86657e-016
            0
```

12 x 12 Perplectic Sweep

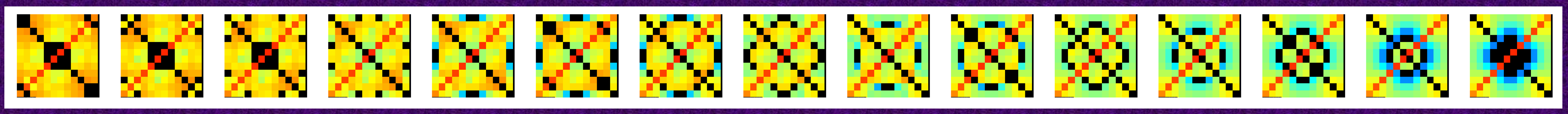


Scale



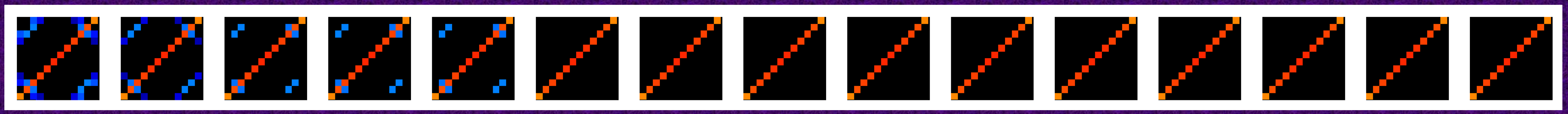
Sweep 1

Sweep 4



Sweep 5

Sweep 6



NUMERICAL RESULTS

Experiment: 100 matrices, entries from $N(0,1)$

n	Sweeps	RelOff	$\lVert P^TRP-R\rVert_F$	$\lVert P^TP-I\rVert_F$	Block	RelEig
50	7.27	6.57×10^{-16}	6.70×10^{-14}	6.70×10^{-14}	7.75×10^{-15}	4.81×10^{-14}
100	8.02	3.28×10^{-15}	2.21×10^{-13}	2.21×10^{-13}	1.63×10^{-14}	1.57×10^{-13}
150	8.36	2.42×10^{-15}	4.32×10^{-13}	4.32×10^{-13}	2.51×10^{-14}	2.71×10^{-13}
200	8.85	1.75×10^{-15}	7.00×10^{-13}	7.00×10^{-13}	3.40×10^{-14}	2.72×10^{-13}
1000	11	3.49×10^{-15}	1.79×10^{-12}	1.79×10^{-12}	1.84×10^{-13}	1.51×10^{-13}

Table 1: $n\times n$ symmetric persymmetric matrices

n	Sweeps	RelOff	$\lVert P^TRP-R\rVert_F$	$\lVert P^TP-I\rVert_F$	Block	RelEig
50	7.89	9.16×10^{-16}	3.11×10^{-14}	3.12×10^{-14}	8.19×10^{-15}	2.12×10^{-14}
100	8.72	2.01×10^{-15}	7.57×10^{-14}	7.58×10^{-14}	1.71×10^{-14}	7.41×10^{-14}
150	9.02	2.17×10^{-15}	1.33×10^{-13}	1.33×10^{-13}	2.62×10^{-14}	5.91×10^{-14}
200	9.31	5.28×10^{-15}	2.10×10^{-13}	2.10×10^{-13}	3.54×10^{-14}	1.01×10^{-13}
1000	10	4.13×10^{-15}	2.73×10^{-12}	2.72×10^{-12}	1.74×10^{-13}	2.10×10^{-13}

Table 2: $n\times n$ skew-symmetric persymmetric matrices

n	Sweeps	RelOff	$\lVert P^TRP-R\rVert_F$	$\lVert P^TP-I\rVert_F$	Block	RelEig
50	7.20	1.49×10^{-15}	3.48×10^{-14}	3.48×10^{-14}	7.72×10^{-15}	6.53×10^{-14}
100	8.00	1.22×10^{-15}	8.98×10^{-14}	8.98×10^{-14}	1.62×10^{-14}	5.76×10^{-14}
150	8.08	3.05×10^{-15}	1.79×10^{-13}	1.79×10^{-13}	2.48×10^{-14}	8.88×10^{-14}
200	8.58	5.71×10^{-15}	2.60×10^{-13}	2.60×10^{-13}	3.36×10^{-14}	1.50×10^{-13}
1000	10	7.86×10^{-17}	1.06×10^{-11}	1.06×10^{-11}	1.82×10^{-13}	5.84×10^{-13}

Table 3: $n\times n$ symmetric perskew-symmetric matrices

RelOff $= \sqrt{\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}} \Big/ \|A\|_F$

$\|P^TRP-R\|_F$ = check of perplectic structure of P

$\|P^TP-I\|_F$ = check of orthogonal structure of P

Block $= \|\{P_{1:n,1:n} - RP_{n+1:2n,n+1:2n}R\}\| + \|\{P_{1:n,n+1:2n} - RP_{n+1:2n,1:n}R\}\|$ (check of perplectic orthogonal structure of P)

RelEig $= \max_j|\lambda_j^{eig} - \lambda_j^{jac}| \,/\, |\lambda_j^{eig}|$ (check against MATLAB's `eig` function)

NUMERICAL RESULTS

Convergence:
- methods always converged
- reloff always decreased monotonically
- rate is asymptotically quadratic

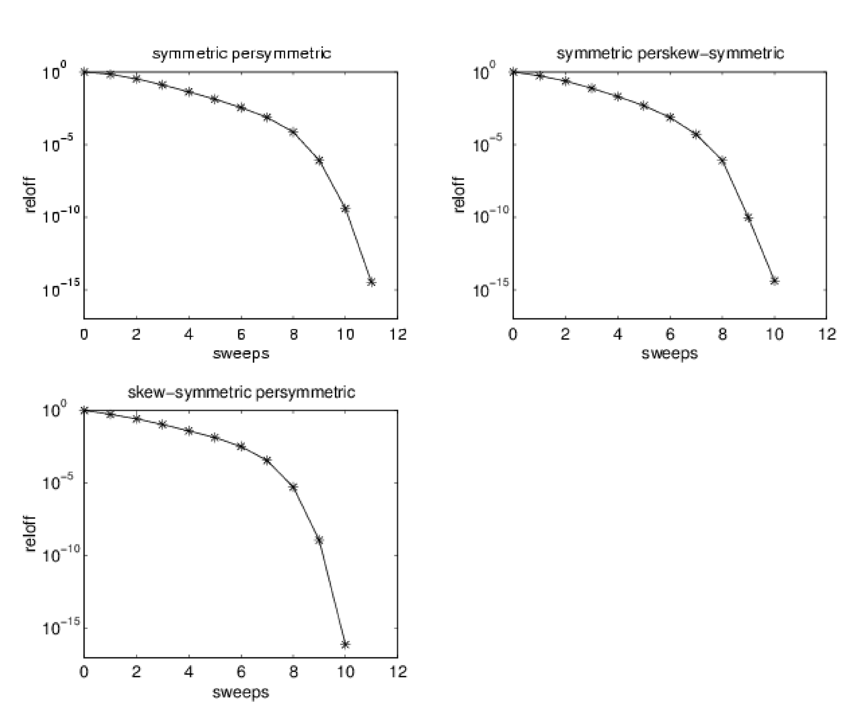


Figure 1: Typical convergence behavior of 1000×1000 matrices

NUMERICAL RESULTS

Sweeps:
- depends only on n
- increases roughly as $\log n$
- leads to *a priori* stopping criteria

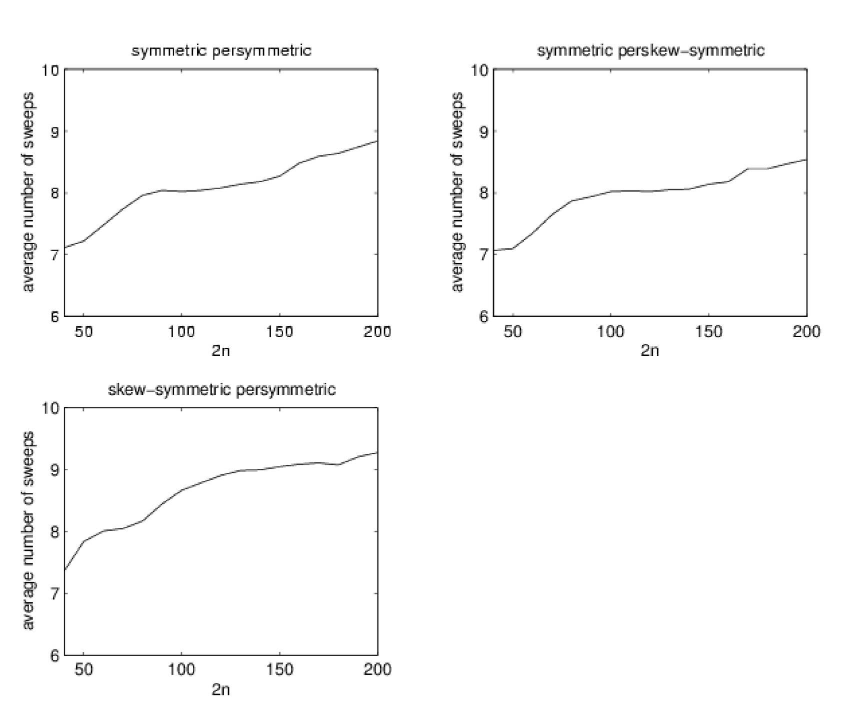


Figure 2: Average number of sweeps for convergence for 1000×1000 matrices