Numerical methods for the palindromic eigenvalue problem

Christian Mehl
Institut für Mathematik
Technische Universität Berlin

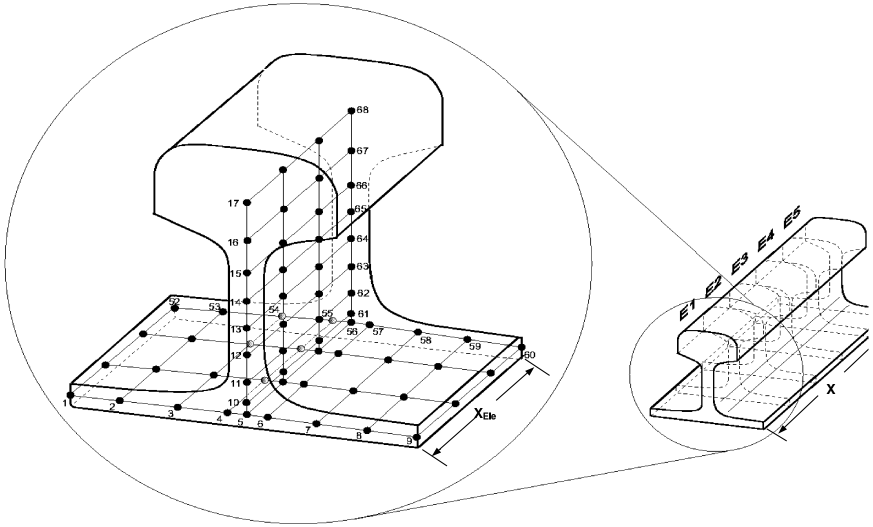
2nd ALAMA Courses on Matrix Polynomials

Castro Urdiales, May 23–24, 2013

The application

Palindromic eigenvalue problems

Application: vibration analysis of rail tracks excited by high speed trains



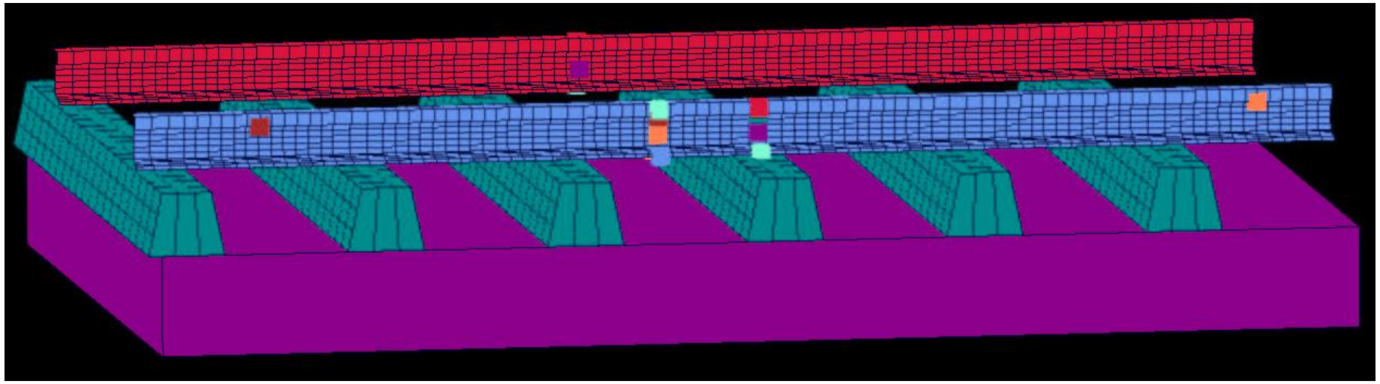
Finite element discretization of rail leads to the eigenvalue problem

$$\big(\lambda^2 A_1(\omega)^T + \lambda A_0(\omega) + A_1(\omega)\big) y = 0,$$

where $A_0, A_1 \in \mathbb{C}^{n\times n}$ and where ω is the excitation frequency.

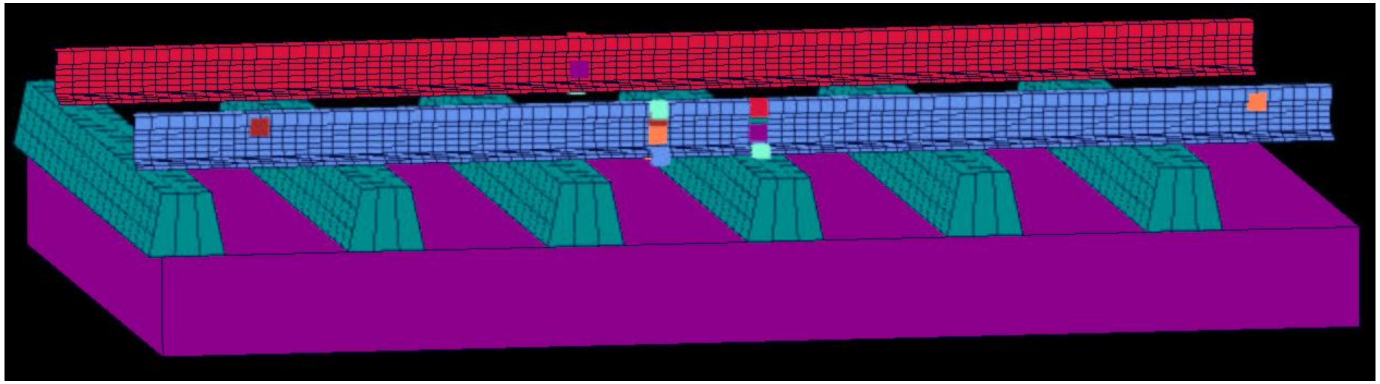
Palindromic eigenvalue problems in applications

Example: vibration analysis of rail tracks under excitation arising from fast trains



Palindromic eigenvalue problems in applications

Example: vibration analysis of rail tracks under excitation arising from fast trains

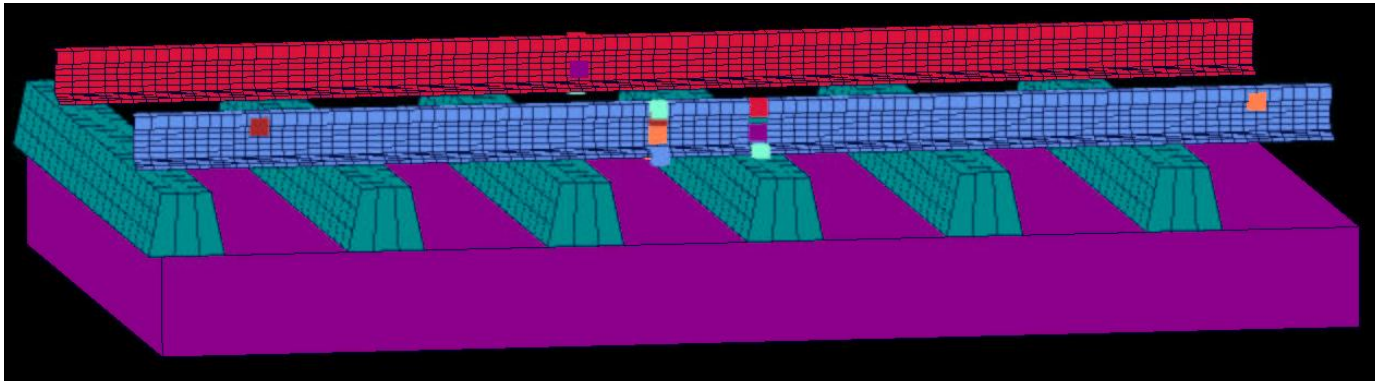


Assumptions:

1. the rail is infinite;

Palindromic eigenvalue problems in applications

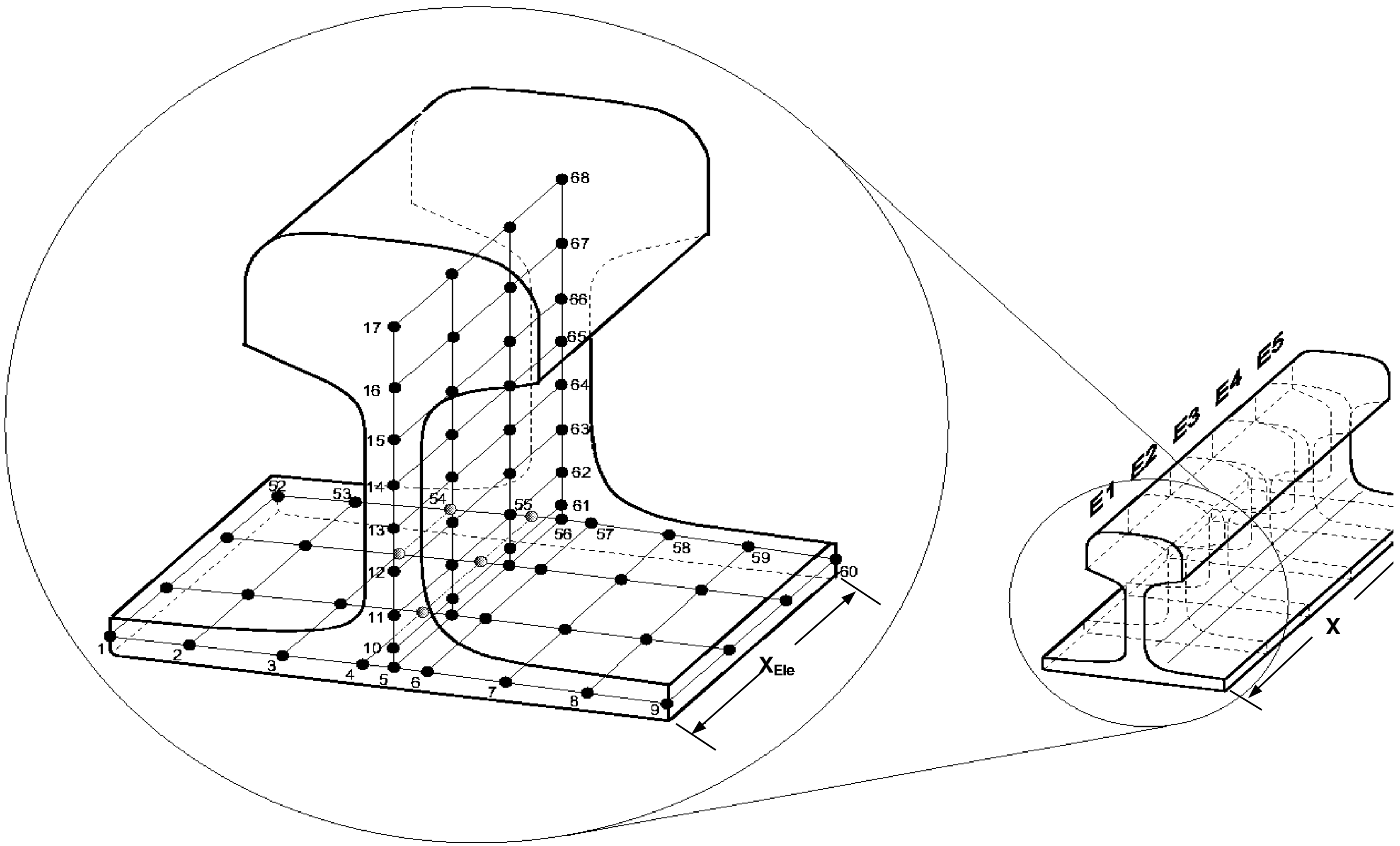
Example: vibration analysis of rail tracks under excitation arising from fast trains



Assumptions:

1. the rail is infinite;
2. identical form of track between two ties $\Rightarrow$ the system is periodic;

FE discretization of the rail



FE discretization of the rail

This leads to an infinite second order system

$$M\ddot{x} + D\dot{x} + Sx = F,$$

where

$$M = \begin{bmatrix} \ddots & \ddots & 0 & \dots & 0 \\ \ddots & M_{j-1,0} & M_{j,1} & 0 & \dots \\ 0 & M_{j,1}^T & M_{j,0} & M_{j+1,1} & 0 \\ \vdots & \ddots & M_{j+1,1}^T & M_{j+1,0} & \ddots \\ 0 & \dots & 0 & \ddots & \ddots \end{bmatrix}, \quad x = \begin{bmatrix} \vdots \\ x_{j-1} \\ x_j \\ x_{j+1} \\ \vdots \end{bmatrix}, \quad F = \begin{bmatrix} \vdots \\ F_{j-1} \\ F_j \\ F_{j+1} \\ \vdots \end{bmatrix}$$

The matrices D and S have a similar structure.

A related difference equation

Ansatz: $F_j = \hat{F}_j e^{i\omega t}$, $x_j = \hat{x}_j e^{i\omega t}$, where ω is the excitation frequency

$$\rightsquigarrow A_{j-1,j}^T \hat{x}_{j-1} + A_{jj}\hat{x}_j + A_{j,j+1}\hat{x}_{j+1} = \hat{F}_j,$$

where

$$\begin{aligned} A_{j,j+1} &= -\omega^2 M_{j1} + i\omega D_{j1} + S_{j1} \\ A_{jj} &= -\omega^2 M_{j0} + i\omega D_{j0} + S_{j0} \end{aligned}$$

Observation: the rail track has identical form between two ties

$\rightsquigarrow$ the system matrices vary periodically

A related difference equation

Combine the parts of the rail between to ties into one vector:

$$y_j = \left[\begin{array}{cccc} \hat{x}_j^T & \hat{x}_{j+1}^T & \ldots & \hat{x}_{j+l}^T \end{array} \right]^T$$

$\rightsquigarrow$ constant coefficient difference equation:

$$A_1^T y_{j-1} + A_0 y_j + A_1 y_{j+1} = \tilde{F}_j,$$

where

$$A_0 = \begin{bmatrix} A_{j,j} & A_{j,j+1} & & 0 \\ A_{j,j+1}^T & A_{j+1,j+1} & \ddots & \\ & \ddots & \ddots & A_{j+l-1,j+l} \\ 0 & & A_{j+l-1,j+l}^T & A_{j+l,j+l} \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 \\ A_{j+l,j+l+1} & 0 & \ldots & 0 \end{bmatrix}.$$

A quadratic eigenvalue problem

Ansatz: $y_{j+1} = \lambda y_j \quad \rightsquigarrow (A_1^T + \lambda A_0 + \lambda^2 A_1)y = 0$

- A_0 is complex symmetric ($A_0 = A_0^T$)
- A_1 is complex and singular
- coefficient sequence A_1^T–A_0–A_1 can also be read backwards

A quadratic eigenvalue problem

Ansatz: $y_{j+1} = \lambda y_j \quad \rightsquigarrow (A_1^T + \lambda A_0 + \lambda^2 A_1)y = 0$

- A_0 is complex symmetric ($A_0 = A_0^T$)
- A_1 is complex and singular
- coefficient sequence A_1^T–A_0–A_1 can also be read backwards
- This is a quadratic **palindromic eigenvalue problem**!

Palindromic matrix polynomials

Reminder: A matrix polynomial $P(\lambda) = A_0 + \lambda A_1 + \cdots + \lambda^k A_k$ is called T-palindromic (in short: **palindromic**) if

$$P(\lambda) = \sum_{j=0}^{k} \lambda^{k-j} A_j^T.$$

Examples:

- $P(\lambda) = A + \lambda B + \lambda^2 B^T + \lambda^3 A^T$;
- $P(\lambda) = A_2^T + \lambda A_1^T + \lambda^2 A_0 + \lambda^3 A_1 + \lambda^4 A_2$, where A_0 is symmetric;
- palindromic pencils $\lambda Z + Z^T$.

Formal resemblance with linguistic palindroms like "**¡La moral, claro, mal!**".

Properties of palindromic matrix polynomials

General assumption: all matrix polynomials under consideration are regular, i.e., $\det P(\lambda) \not\equiv 0$.

Spectral symmetry: Palindromic matrix polynomials have a **symplectic spectrum**.

- if λ_0 is an eigenvalue of $P(\lambda)$, then so is λ_0^{-1};
- pairing occurs also in algebraic, geometric, and partial multiplicities;
- symmetry degenerates for $\lambda_0 = 1$ and $\lambda_0 = -1$;

The linearization

How to solve structured eigenvalue problems

Given: structured eigenvalue problem $(\lambda^m A_m + \cdots + \lambda A_1 + A_0)x = 0.$

Standard approach: linearize the problem; use the companion form

$$\lambda \begin{bmatrix} A_m & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix} \begin{bmatrix} \lambda^{m-1}x \\ \vdots \\ \lambda x \\ x \end{bmatrix} = \begin{bmatrix} -A_{m-1} & \dots & \dots & -A_0 \\ I & & & \\ & \ddots & & \\ & & I & 0 \end{bmatrix} \begin{bmatrix} \lambda^{m-1}x \\ \vdots \\ \lambda x \\ x \end{bmatrix}$$

Task: find a linearization having the same symmetry structure as the original matrix polynomial that still has the nice properties of the companion form

An ansatz space for linearizations

Given: structured eigenvalue problem $(\lambda^m A_m + \cdots + \lambda A_1 + A_0)x = 0$.

Standard approach: linearize the problem; use the companion form

$$\lambda \underbrace{\begin{bmatrix} A_m & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix}}_{\mathcal{A}} \underbrace{\begin{bmatrix} \lambda^{m-1} \\ \vdots \\ \lambda \\ 1 \end{bmatrix}}_{=:\Lambda} x = \underbrace{\begin{bmatrix} -A_{m-1} & \dots & -A_1 & -A_0 \\ I & & & \\ & \ddots & & \\ & & I & 0 \end{bmatrix}}_{\mathcal{B}} \underbrace{\begin{bmatrix} \lambda^{m-1} \\ \vdots \\ \lambda \\ 1 \end{bmatrix}}_{=\Lambda} x$$

The companion form $C(\lambda) = \lambda\mathcal{A} - \mathcal{B}$ satisfies

$$C(\lambda)\Lambda x = \begin{bmatrix} P(\lambda) \\ 0 \\ \vdots \\ 0 \end{bmatrix} x = e_1 \otimes P(\lambda)x$$

An ansatz space for linearizations

The companion form $C(\lambda) = \lambda\mathcal{A} - \mathcal{B}$ satisfies

$$C(\lambda)\Lambda x = \begin{bmatrix} P(\lambda) \\ 0 \\ \vdots \\ 0 \end{bmatrix} x = e_1 \otimes P(\lambda)x$$

Idea: Allow linearizations $L(\lambda)$ satisfying

$$L(\lambda)\Lambda x = \begin{bmatrix} v_1 P(\lambda) \\ v_2 P(\lambda) \\ \vdots \\ v_m P(\lambda) \end{bmatrix} x = v \otimes P(\lambda)x$$

Then $L(\lambda)$ still has the eigenvector recovery property as $C(\lambda)$.

We will call v an **ansatz vector**.

An ansatz space for linearizations

This can be said differently. Consider again $\mathcal{A}$ and $\mathcal{B}$:

$$\mathcal{A} = \begin{bmatrix} A_m & 0 & \cdots & 0 \\ 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \end{bmatrix}, \quad -\mathcal{B} = \begin{bmatrix} A_{m-1} & \dots & A_1 & A_0 \\ -I & & & 0 \\ & \ddots & & \vdots \\ & & -I & 0 \end{bmatrix}$$

Introduce the so-called **shifted sum**:

$$\begin{matrix} & n & n(m-1) \\ nm & [\,\mathcal{A}_1 & \mathcal{A}_2 \,] \end{matrix} \boxplus \begin{matrix} n(m-1) & n \\ [\,\mathcal{B}_1 & \mathcal{B}_2\,] \end{matrix} := \begin{matrix} n & n(m-1) & n \\ [\,\mathcal{A}_1 & \mathcal{A}_2 + \mathcal{B}_1 & \mathcal{B}_2\,] \end{matrix}$$

An ansatz space for linearizations

This can be said differently. Consider again $\mathcal{A}$ and $\mathcal{B}$:

$$\mathcal{A} = \begin{bmatrix} A_m & 0 & \cdots & 0 \\ 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \end{bmatrix}, \quad -\mathcal{B} = \begin{bmatrix} A_{m-1} & \dots & A_1 & A_0 \\ -I & & & 0 \\ & \ddots & & \vdots \\ & & -I & 0 \end{bmatrix}$$

Introduce the so-called **shifted sum**:

$$\begin{array}{c} \\ nm \end{array} \overset{\begin{array}{cc} n & n(m-1) \end{array}}{\begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 \end{bmatrix}} \boxplus \overset{\begin{array}{cc} n(m-1) & n \end{array}}{\begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \end{bmatrix}} := \overset{\begin{array}{ccc} n & n(m-1) & n \end{array}}{\begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 + \mathcal{B}_1 & \mathcal{B}_2 \end{bmatrix}}$$

An ansatz space for linearizations

This can be said differently. Consider again $\mathcal{A}$ and $\mathcal{B}$:

$$\mathcal{A} = \begin{bmatrix} A_m & 0 & \cdots & 0 \\ 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \end{bmatrix}, \quad -\mathcal{B} = \begin{bmatrix} A_{m-1} & \dots & A_1 & A_0 \\ -I & & & 0 \\ & \ddots & & \vdots \\ & & -I & 0 \end{bmatrix}$$

We obtain:

$$\mathcal{A} \boxplus (-\mathcal{B}) = \begin{bmatrix} A_m & A_{m-1} & \dots & A_1 & A_0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

The first block row consists of the coefficients of the polynomial $P(\lambda)$.

An ansatz space for linearizations

Ansatz: Look for linearizations $L(\lambda)$ in the vector space

$$\mathbb{L}_1(P) = \left\{\lambda\mathcal{A} - \mathcal{B} \;:\; (\lambda\mathcal{A} - \mathcal{B})\Lambda = v \otimes P(\lambda), v = (v_1, \ldots, v_m)^T \in \mathbb{C}^m\right\}$$

$$= \left\{\lambda\mathcal{A} - \mathcal{B} \;:\; \mathcal{A} \boxplus (-\mathcal{B}) = \begin{bmatrix} v_1 A_m & \cdots & v_1 A_1 & v_1 A_0 \\ \vdots & \ldots & \vdots & \vdots \\ v_m A_m & \cdots & v_m A_1 & v_m A_0 \end{bmatrix}, v_i \in \mathbb{C}\right\}$$

An ansatz space for linearizations

Ansatz: Look for linearizations $L(\lambda)$ in the vector space

$$\mathbb{L}_1(P) = \left\{\lambda\mathcal{A} - \mathcal{B} \;:\; (\lambda\mathcal{A} - \mathcal{B})\Lambda = v \otimes P(\lambda), v = (v_1, \ldots, v_m)^T \in \mathbb{C}^m\right\}$$

$$= \left\{ \lambda\mathcal{A} - \mathcal{B} \;:\; \mathcal{A} \boxplus (-\mathcal{B}) = \begin{bmatrix} v_1 A_m & \cdots & v_1 A_1 & v_1 A_0 \\ \vdots & \ldots & \vdots & \vdots \\ v_m A_m & \cdots & v_m A_1 & v_m A_0 \end{bmatrix}, v_i \in \mathbb{C} \right\}$$

Theorem (Mackey, Mackey, M., Mehrmann, 2004):
$L(\lambda) \in \mathbb{L}_1(P)$ is a linearization of the regular matrix polynomial $P(\lambda)$ if and only if $L(\lambda)$ is regular, i.e., $\det L(\lambda) \not\equiv 0$.

An ansatz space for linearizations

Ansatz: Look for linearizations $L(\lambda)$ in the vector space

$$\mathbb{L}_1(P) = \{\lambda\mathcal{A} - \mathcal{B} \,:\, (\lambda\mathcal{A} - \mathcal{B})\Lambda = v \otimes P(\lambda), v = (v_1, \ldots, v_m)^T \in \mathbb{C}^m\}$$

$$= \left\{\lambda\mathcal{A} - \mathcal{B} \,:\, \mathcal{A} \boxplus (-\mathcal{B}) = \begin{bmatrix} v_1 A_m & \cdots & v_1 A_1 & v_1 A_0 \\ \vdots & \ldots & \vdots & \vdots \\ v_m A_m & \cdots & v_m A_1 & v_m A_0 \end{bmatrix}, v_i \in \mathbb{C}\right\}$$

Note: These linearizations share nice properties with the companion form:

- eigenvectors have the form $\Lambda = \begin{bmatrix} \lambda^{m-1} x \\ \vdots \\ x \end{bmatrix}$, where $P(\lambda)x = 0$;
- information on eigenvalue ∞ is properly reflected (strong linearization)

Palindromic linearizations

Question: Can we find a palindromic $L(\lambda) \in \mathbb{L}_1(P)$?

Palindromic linearizations

Question: Can we find a palindromic $L(\lambda) \in \mathbb{L}$?

Answer: Yes! Use a *palindromic* v, i.e., $v \in \mathbb{R}^m$ satisfying

$$v = \left[\, v_1 \;\; v_2 \;\; \ldots \;\; v_{m-1} \;\; v_m \,\right] = \left[\, v_m \;\; v_{m-1} \;\; \ldots \;\; v_2 \;\; v_1 \,\right].$$

Palindromic linearizations

Question: Can we find a palindromic $L(\lambda) \in \mathbb{L}$?

Answer: Yes! Use a *palindromic* v, i.e., $v \in \mathbb{R}^m$ satisfying

$$v = \begin{bmatrix} v_1 & v_2 & \dots & v_{m-1} & v_m \end{bmatrix} = \begin{bmatrix} v_m & v_{m-1} & \dots & v_2 & v_1 \end{bmatrix}.$$

Theorem (Mackey, Mackey, M., Mehrmann, 2004):
Let $v \in \mathbb{R}^m$ be palindromic. Then there exists a unique palindromic pencil $L(\lambda) = \lambda Z + Z^T$ such that

$$Z \boxplus Z^T = \begin{bmatrix} v_1 A_m & \cdots & v_1 A_1 & v_1 A_0 \\ \vdots & \dots & \vdots & \vdots \\ v_m A_m & \cdots & v_m A_1 & v_m A_0 \end{bmatrix} \in \mathbb{L}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & \mathbf{?} & \mathbf{?} \\ \mathbf{0} & \mathbf{?} & \mathbf{?} \\ A & \mathbf{?} & \mathbf{?} \end{bmatrix}, \quad Z^T = \begin{bmatrix} \mathbf{?} & \mathbf{?} & A^T \\ \mathbf{?} & \mathbf{?} & \mathbf{0} \\ \mathbf{?} & \mathbf{?} & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & \mathbf{?} & \mathbf{?} \\ \mathbf{0} & \mathbf{?} & \mathbf{?} \\ A & \mathbf{0} & A \end{bmatrix}, \quad Z^T = \begin{bmatrix} A^T & \mathbf{0} & A^T \\ \mathbf{?} & \mathbf{?} & \mathbf{0} \\ \mathbf{?} & \mathbf{?} & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & B - A^T & B^T \\ \mathbf{0} & \mathbf{?} & \mathbf{?} \\ A & \mathbf{0} & A \end{bmatrix}, \quad Z^T = \begin{bmatrix} A^T & \mathbf{0} & A^T \\ \mathbf{?} & \mathbf{?} & \mathbf{0} \\ B & B^T - A & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & B - A^T & B^T \\ \mathbf{0} & \mathbf{?} & B - A^T \\ A & \mathbf{0} & A \end{bmatrix}, \quad Z^T = \begin{bmatrix} A^T & \mathbf{0} & A^T \\ B^T - A & \mathbf{?} & \mathbf{0} \\ B & B^T - A & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & B - A^T & B^T \\ \mathbf{0} & A - B^T & B - A^T \\ A & \mathbf{0} & A \end{bmatrix}, \quad Z^T = \begin{bmatrix} A^T & \mathbf{0} & A^T \\ B^T - A & A^T - B & \mathbf{0} \\ B & B^T - A & A^T \end{bmatrix}$$

Construction of palindromic linearizations

Given: Palindromic matrix polynomial $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$.
We construct a palindromic pencil corresponding to $v = (1, 0, 1)^T \in \mathbb{R}^3$.
We know:

$$Z \boxplus Z^T = \begin{bmatrix} A & B & B^T & A^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A & B & B^T & A^T \end{bmatrix}$$

We obtain:

$$Z = \begin{bmatrix} A & B - A^T & B^T \\ \mathbf{0} & A - B^T & B - A^T \\ A & \mathbf{0} & A \end{bmatrix}, \quad Z^T = \begin{bmatrix} A^T & \mathbf{0} & A^T \\ B^T - A & A^T - B & \mathbf{0} \\ B & B^T - A & A^T \end{bmatrix}$$

This is a palindromic linearization of $P(\lambda)$ if $\det(\lambda Z + Z^T) \not\equiv 0$.

Question: How do we check regularity of $\lambda Z + Z^T \in \mathbb{L}_1(P)$?

The "v-polynomial theorem"

Given: palindromic matrix polynomial $P(\lambda) = \lambda^m A_m + \cdots + \lambda A_1 + A_0$

The "v-polynomial theorem"

Given: palindromic matrix polynomial $P(\lambda) = \lambda^m A_m + \cdots + \lambda A_1 + A_0$

Theorem (Mackey, Mackey, M., Mehrmann, 2004):
Let $v = (v_1, \ldots, v_m)$ be palindromic and let $\lambda Z - Z^T \in \mathbb{L}_1(P)$ be associated with v. Consider the *v-polynomial* $\lambda^{m-1} v_1 + \cdots + \lambda v_{m-1} + v_m$. Then $\lambda Z - Z^T$ is regular if and only if the roots of the v-polynomial are disjoint from the eigenvalues of $P(\lambda)$.

The "v-polynomial theorem"

Given: palindromic matrix polynomial $P(\lambda) = \lambda^m A_m + \cdots + \lambda A_1 + A_0$

Theorem (Mackey, Mackey, M., Mehrmann, 2004):
Let $v = (v_1, \ldots, v_m)$ be palindromic and let $\lambda Z - Z^T \in \mathbb{L}_1(P)$ be associated with v. Consider the *v-polynomial* $\lambda^{m-1}v_1 + \cdots + \lambda v_{m-1} + v_m$. Then $\lambda Z - Z^T$ is regular if and only if the roots of the v-polynomial are disjoint from the eigenvalues of $P(\lambda)$.

Example: The v-polynomial of $v = (1, 0, 1)^T \in \mathbb{R}^3$ is $\lambda^2 + 1$. Thus

$$\lambda \begin{bmatrix} A & B - A^T & B^T \\ 0 & A - B^T & B - A^T \\ A & 0 & A \end{bmatrix} + \begin{bmatrix} A^T & 0 & A^T \\ B^T - A & A^T - B & 0 \\ B & B^T - A & A^T \end{bmatrix}$$

is a linearization of $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda B^T + A^T$ if and only if $\pm i$ are no eigenvalues of $P(\lambda)$.

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.
- The corresponding v-polynomial is $\lambda + 1$.

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.
- The corresponding v-polynomial is $\lambda + 1$.
- We obtain the following possible linearization:

$$\lambda \begin{bmatrix} A_1 & ? \\ A_1 & ? \end{bmatrix} + \begin{bmatrix} ? & A_1^T \\ ? & A_1^T \end{bmatrix}$$

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.
- The corresponding v-polynomial is $\lambda + 1$.
- We obtain the following possible linearization:

$$\lambda \begin{bmatrix} A_1 & \textbf{?} \\ A_1 & A_1 \end{bmatrix} + \begin{bmatrix} A_1^T & A_1^T \\ \textbf{?} & A_1^T \end{bmatrix}$$

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.
- The corresponding v-polynomial is $\lambda + 1$.
- We obtain the following possible linearization:

$$\lambda \begin{bmatrix} A_1 & A_0 - A_1^T \\ A_1 & A_1 \end{bmatrix} + \begin{bmatrix} A_1^T & A_1^T \\ A_0 - A_1 & A_1^T \end{bmatrix}$$

The railway eigenvalue problem

Aim: Find a palindromic linearization for $P(\lambda) = \lambda^2 A_1 + \lambda A_0 + A_1^T$, where A_0 is complex symmetric.

- There is only one palindromic $v \in \mathbb{R}^2$ up to scaling: $v = (1, 1)$.
- The corresponding v-polynomial is $\lambda + 1$.
- We obtain the following possible linearization:

$$\lambda \begin{bmatrix} A_1 & A_0 - A_1^T \\ A_1 & A_1 \end{bmatrix} + \begin{bmatrix} A_1^T & A_1^T \\ A_0 - A_1 & A_1^T \end{bmatrix}$$

- This is a linearization if and only if -1 is not an eigenvalue of $P(\lambda)$, i.e., if and only if $A_1 - A_0 + A_1^T$ is nonsingular.

The anti-triangular Schur form

How to solve linear palindromic eigenvalue problems

Task: Solve the generalized eigenvalue problem for $\lambda Z + Z^T$.

- T-congruence transformations preserve the structure:

$$(\lambda Z + Z^T) \mapsto P^T(\lambda Z + Z^T)P, \quad P \text{ invertible}$$

- Numerical stability: Choose $P = U$ unitary if possible.

- Look for condensed forms under simultaneous unitary consimilarity:

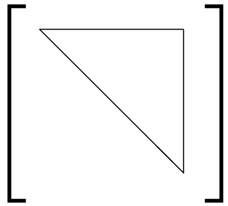
$$(\lambda Z + Z^T) \mapsto \overline{U}^{-1}(\lambda Z + Z^T)U, \quad U \text{ unitary}$$

Advantage: We have to store and work on Z only.

Anti-triangular forms

Naive approach: reduce Z to triangular form:

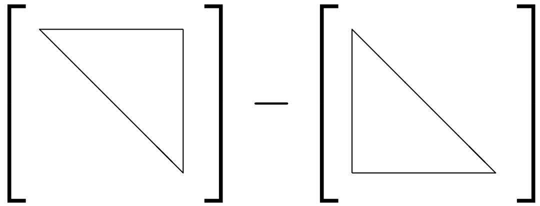
$$U^T Z U \quad \hat{=}$$



Anti-triangular forms

Naive approach: reduce Z to triangular form:

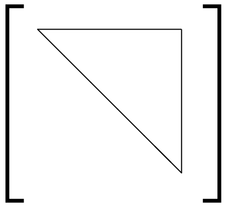
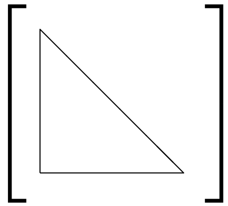
$$\lambda U^T Z U - U^T Z^T U \hat{=} \lambda$$



Drawback: the eigenvalues of $\lambda Z + Z^T$ cannot be read off;

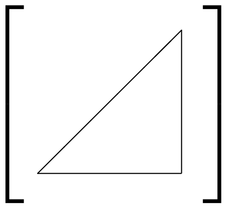
Anti-triangular forms

Naive approach: reduce Z to triangular form:

$$\lambda U^T Z U - U^T Z^T U \mathrel{\hat{=}} \lambda$$  $-$ 

Drawback: the eigenvalues of $\lambda Z + Z^T$ cannot be read off;

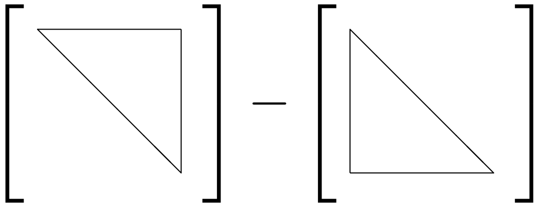
Idea: reduce Z to **anti-triangular form**;

$$U^T Z U \quad \mathrel{\hat{=}}$$ 

Anti-triangular forms

Naive approach: reduce Z to triangular form:

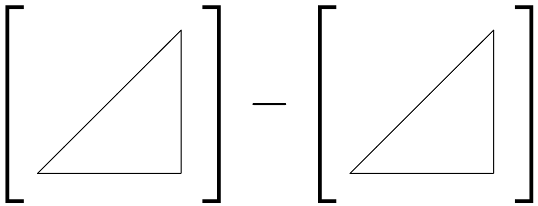
$$\lambda U^T Z U - U^T Z^T U \hat{=} \lambda$$



Drawback: the eigenvalues of $\lambda Z + Z^T$ cannot be read off;

Idea: reduce Z to **anti-triangular form**;

$$\lambda U^T Z U - U^T Z^T U \hat{=} \lambda$$



Then the eigenvalues of $\lambda Z + Z^T$ can be easily read off.

The anti-triangular Schur form

Theorem: Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^T Z U = \begin{bmatrix} 0 & \dots & 0 & z_{1n} \\ \vdots & \iddots & z_{2,n-1} & \vdots \\ 0 & \iddots & \iddots & \vdots \\ z_{n1} & \dots & \dots & z_{nn} \end{bmatrix}$$

is in **anti-triangular Schur form**.

We now have

$$\lambda U^T Z U + U^T Z U = \lambda \begin{bmatrix} 0 & \dots & 0 & z_{1n} \\ \vdots & \iddots & z_{2,n-1} & \vdots \\ 0 & \iddots & \iddots & \vdots \\ z_{n1} & \dots & \dots & z_{nn} \end{bmatrix} + \begin{bmatrix} 0 & \dots & 0 & z_{n1} \\ \vdots & \iddots & z_{n-1,2} & \vdots \\ 0 & \iddots & \iddots & \vdots \\ z_{1n} & \dots & \dots & z_{nn} \end{bmatrix}$$

The anti-triangular Schur form

Theorem: Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^T Z U = \begin{bmatrix} 0 & \dots & 0 & z_{1n} \\ \vdots & \iddots & z_{2,n-1} & \vdots \\ 0 & \iddots & \iddots & \vdots \\ z_{n1} & \dots & \dots & z_{nn} \end{bmatrix}$$

is in **anti-triangular Schur form**.

Consequence: If $\det(\lambda Z + Z^T) \not\equiv 0$ then the eigenvalues of $\lambda Z + Z^T$ are

$$-\frac{z_{n1}}{z_{1n}}, \dots, -\frac{z_{1n}}{z_{n1}}, \qquad \text{(where } \frac{z}{0} := \infty\text{)}.$$

The anti-triangular Schur form

Theorem: Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^T Z U = \begin{bmatrix} 0 & \dots & 0 & z_{1n} \\ \vdots & \iddots & z_{2,n-1} & \vdots \\ 0 & \iddots & \iddots & \vdots \\ z_{n1} & \dots & \dots & z_{nn} \end{bmatrix}$$

is in **anti-triangular Schur form**.

Observation: If x is an eigenvector of $\lambda Z + Z^T$, then:

$$\lambda x^T Z x = -x^T Z^T x = -x^T Z x$$

Consequence: $x^T Z x = 0$, i.e., x is Z-**neutral**, provided that $\lambda \neq -1$.

The anti-triangular Schur form

Proof (for the generic case $\lambda Z + Z^T$ is regular and -1 is not an eigenvalue):

$$Z = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

- If x is a normalized eigenvector,
- then x is Z-neutral, i.e. $x^T Z x = 0$, i.e., x and $\overline{Zx}$ are orthogonal.
- Let $Q = [x, q_2, \ldots, q_n]$ be unitary with $q_n = \frac{\overline{Zx}}{\|Zx\|}$ (or $\frac{\overline{Z^T x}}{\|Z^T x\|}$, when $Zx = 0$).

The anti-triangular Schur form

Proof (for the generic case $\lambda Z + Z^T$ is regular and -1 is not an eigenvalue):

$$Z = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

- If x is a normalized eigenvector,
- then x is Z-neutral, i.e. $x^T Z x = 0$, i.e., x and $\overline{Zx}$ are orthogonal.
- Let $Q = [x, q_2, \ldots, q_n]$ be unitary with $q_n = \frac{\overline{Zx}}{\|Zx\|}$ (or $\frac{\overline{Z^T x}}{\|Z^T x\|}$, when $Zx = 0$).
- Compute $\widetilde{Z} = Q^T Z Q$.

The anti-triangular Schur form

Proof (for the generic case $\lambda Z + Z^T$ is regular and -1 is not an eigenvalue):

$$\widetilde{Z} = \begin{bmatrix} 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

- If x is a normalized eigenvector,
- then x is Z-neutral, i.e. $x^T Z x = 0$, i.e., x and $\overline{Zx}$ are orthogonal.
- Let $Q = [x, q_2, \ldots, q_n]$ be unitary with $q_n = \frac{\overline{Zx}}{\|Zx\|}$ (or $\frac{\overline{Z^T x}}{\|Z^T x\|}$, when $Zx = 0$).
- Compute $\widetilde{Z} = Q^T Z Q$. **Then** $q_j^T Z q_1 = q_j^T Z x = 0$ **for** $j = 1, \ldots, n-1$.

The anti-triangular Schur form

Proof (for the generic case $\lambda Z + Z^T$ is regular and -1 is not an eigenvalue):

$$\widetilde{Z} = \begin{bmatrix} 0 & 0 & 0 & 0 & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

- If x is a normalized eigenvector,
- then x is Z-neutral, i.e. $x^T Z x = 0$, i.e., x and $\overline{Zx}$ are orthogonal.
- Let $Q = [x, q_2, \ldots, q_n]$ be unitary with $q_n = \frac{\overline{Zx}}{\|Zx\|}$ (or $\frac{\overline{Z^T x}}{\|Z^T x\|}$, when $Zx = 0$).
- Compute $\widetilde{Z} = Q^T Z Q$. **We also have $q_j^T Z^T q_1 = q_j^T Z^T x = 0, j = 1, \ldots, n-1$, because Zx and $Z^T x$ are linearly dependent (x is eigenvector).**

The anti-triangular Schur form

Proof (for the generic case $\lambda Z + Z^T$ is regular and -1 is not an eigenvalue):

$$\widetilde{Z} = \begin{bmatrix} 0 & 0 & 0 & 0 & * \\ 0 & \color{red}{*} & \color{red}{*} & \color{red}{*} & * \\ 0 & \color{red}{*} & \color{red}{*} & \color{red}{*} & * \\ 0 & \color{red}{*} & \color{red}{*} & \color{red}{*} & * \\ * & * & * & * & * \end{bmatrix}$$

- If x is a normalized eigenvector,
- then x is Z-neutral, i.e. $x^T Z x = 0$, i.e., x and $\overline{Zx}$ are orthogonal.
- Let $Q = [x, q_2, \ldots, q_n]$ be unitary with $q_n = \frac{\overline{Zx}}{\|Zx\|}$ (or $\frac{\overline{Z^T x}}{\|Z^T x\|}$, when $Zx = 0$).
- Compute $\widetilde{Z} = Q^T Z Q$. **Continue with the red submatrix**.

Method No. 1: Structured Deflation

Method 1: Structured Deflation

Reminder: construction of anti-triangular Schur form

- start: choose an eigenvector x of $\lambda Z + Z^T$.

- x is Z-neutral, i.e. $x^T Z x = 0$,
- if the corresponding eigenvalue is not -1.

- This allows the construction of a unitary U with x and $\overline{Zx}$ (after normalization) as columns.
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: choose an eigenvector x of $\lambda Z + Z^T$.
- x is Z-neutral, i.e. $x^T Z x = 0$,
- if the corresponding eigenvalue is not -1.

- This allows the construction of a unitary U with x and $\overline{Zx}$ (after normalization) as columns.
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.

- x is Z-neutral, i.e. $x^T Z x = 0$,
- if the corresponding eigenvalue is not -1.

- This allows the construction of a unitary U with x and $\overline{Zx}$ (after normalization) as columns.
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.
- Q is Z-neutral, i.e. $Q^T Z Q = 0$,
- if the corresponding eigenvalue is not -1.

- This allows the construction of a unitary U with x and $\overline{Zx}$ (after normalization) as columns.
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.
- Q is Z-neutral, i.e. $Q^T Z Q = 0$,
- if the set Λ of associated eigenvalues is reciprocal-free:

$$\mu \in \Lambda \Rightarrow \frac{1}{\mu} \notin \Lambda;$$

- This allows the construction of a unitary U with x and $\overline{Zx}$ (after normalization) as columns.
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.

- Q is Z-neutral, i.e. $Q^T Z Q = 0$,
- if the set Λ of associated eigenvalues is reciprocal-free:

$$\mu \in \Lambda \Rightarrow \frac{1}{\mu} \notin \Lambda;$$

- This allows the construction of a unitary matrix U with Q and $\overline{ZQ}$ as "columns" (after orthogonalization and normalization).
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.
- Q is Z-neutral, i.e. $Q^T Z Q = 0$,
- if the set Λ of associated eigenvalues is reciprocal-free:

$$\mu \in \Lambda \Rightarrow \frac{1}{\mu} \notin \Lambda;$$

- This allows the construction of a unitary matrix U with Q and $\overline{ZQ}$ as "columns" (after orthogonalization and normalization).
- Congruence trafo with U gets Z "closer" to anti-triangular Schur form.

Method 1: Structured Deflation

Idea: replace x with a suitable deflating subspace

- start: let Q span a deflating subspace of $\lambda Z + Z^T$.

- Q is Z-neutral, i.e. $Q^T Z Q = 0$,
- if the set Λ of associated eigenvalues is reciprocal-free:

$$\mu \in \Lambda \Rightarrow \frac{1}{\mu} \notin \Lambda;$$

- This allows the construction of a unitary matrix U with Q and $\overline{ZQ}$ as "columns" (after orthogonalization and normalization).
- Congruence trafo with U gets Z to anti-triangular Schur form, if the deflating subspace has dimension $\frac{n}{2}$ or $\frac{n-1}{2}$.

Method 1: Structured Deflation

Theorem: (generalizes a trick by A. Laub for the computation of the Hamiltonian Schur form) Let $\lambda Z + Z^T \in \mathbb{C}^{2n\times 2n}$ be regular and let

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} \\ 0 & X_{22} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} \\ 0 & Y_{22} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}.$$

be its generalized Schur decomposition, where $X_{11}, Y_{11} \in \mathbb{C}^{n\times n}$. If

$$\mu \in \sigma(\lambda X_{11} + Y_{11}) \implies \frac{1}{\mu} \notin \sigma(\lambda X_{11} + Y_{11})$$

then

$$U = \begin{bmatrix} W_{11} & Q_{11}^T R_n \\ W_{21} & Q_{12}^T R_n \end{bmatrix}, \qquad \left(R_n := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \right)$$

is unitary and

$$U^T Z U = \begin{bmatrix} 0 & Y_{11}^T R_n \\ R_n X_{11} & * \end{bmatrix}.$$

is in anti-triangular form.

Method 1: Structured Deflation

Algorithm: (for regular $\lambda Z + Z^T$ not having eigenvalues with modulus 1)

1. Compute the generalized Schur decomposition
$$\lambda Z + Z^T = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} \\ 0 & X_{22} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} \\ 0 & Y_{22} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}.$$
2. Reorder the eigenvalues such that $\lambda X_{11} + Y_{11}$ contains all eigenvalues with $|\lambda| > 1$.
3. Set $U = \begin{bmatrix} W_{11} & Q_{11}^T R_n \\ W_{21} & Q_{12}^T R_n \end{bmatrix}$.
4. Compute $Z_{22} = \begin{bmatrix} R_n Q_{11} & R_n Q_{12} \end{bmatrix} Z \begin{bmatrix} Q_{11}^T R_n \\ Q_{12}^T R_n \end{bmatrix}$.
5. Set $\widetilde{Z} := \begin{bmatrix} 0 & Y_{11}^T R_n \\ R_n X_{11} & Z_{22} \end{bmatrix}$.

Method 1: Structured Deflation

Properties:

+ cost is essentially the cost of QZ with reordering;

− only applicable if Z has even dimension and if $\lambda Z + Z^T$ does not have eigenvalues with modulus 1;

− problems if there are eigenvalues with modulus close to ± 1; $\leadsto$ QZ might detect more or less than n eigenvalues λ with $|\lambda| > 1$.

Questions: Are there other methods?

Method No. 2: A Jacobi-like Method

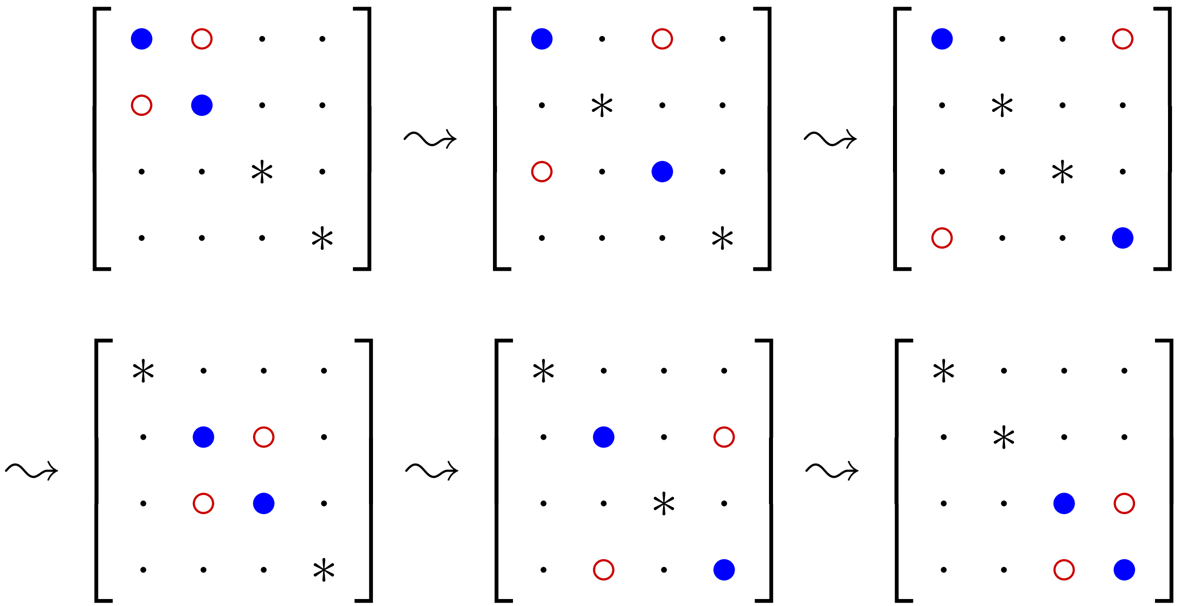
Jacobi(-like) methods

- 1846 Jacobi (symmetric matrices)
- ...
- 1955 Greenstadt (general complex matrices)
- 1985 Stewart (general complex matrices)
- 1990 Byers (Hamiltonian matrices)
- 1997 Bunse-Gerstner/Faßbender (Hamiltonian matrices)
- 2004 Hilliges, M., Mehrmann (palindromic pencils

Jacobi's method for symmetric matrices

Main theme: Elimination of one (or two) pivot elements by diagonalization of a 2×2 problem.

Sweep (cyclic-by-column): $n(n-1)/2$ steps with column-wise choise of pivots from top-to-bottom



Jacobi's method for symmetric matrices

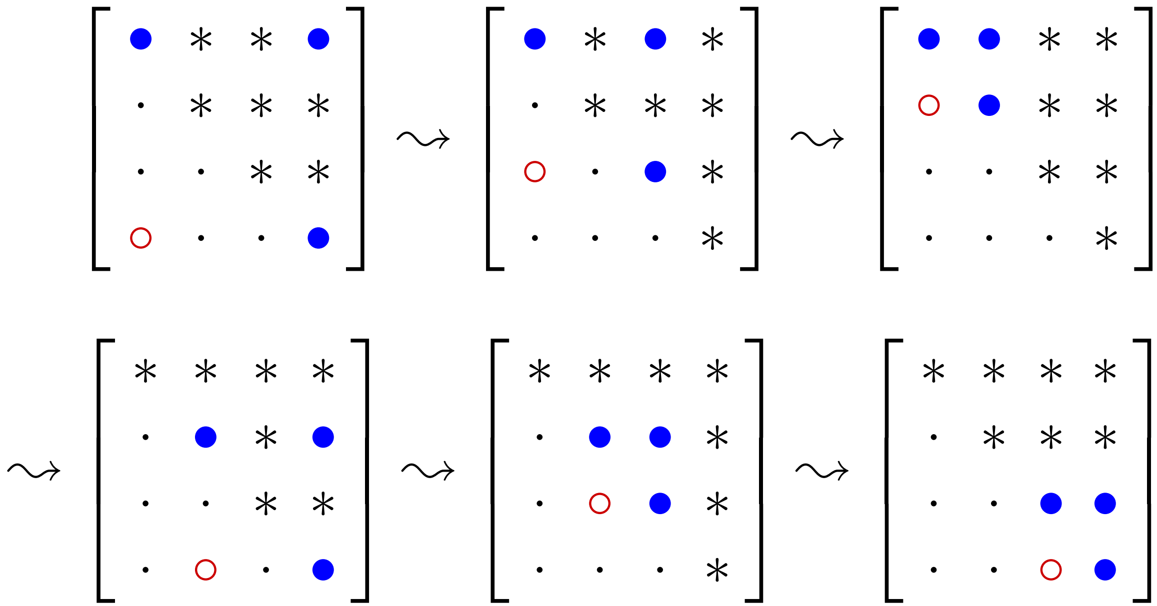
Properties:

- „off-Norm" off$(A) = \sqrt{\sum_{j \neq i} a_{ij}^2}$ is reduced in each step
- global convergence
- asymptotic quadratic convergence
- stable, parallelizable, more accurate than QR, fast if matrix is close to diagonal
- in general much slower than QR (not competitive)

Nonsymmetric Jacobi method

Main theme: Elimination of one (or two) pivot elements by trigonalization of a 2×2 problem.

Sweep (cyclic-by-column): $n(n-1)/2$ steps with column-wise choise of pivots from bottom-to-top

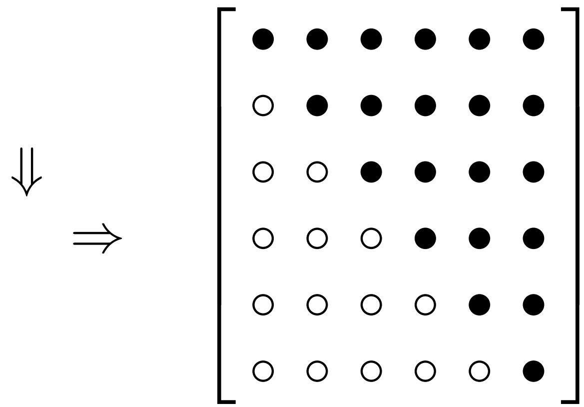


Surprise: sweep $\neq$ sweep

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \circ & \bullet \end{bmatrix}$$

Consider two “cyclic-by-column” sweeps:

Surprise: sweep $\neq$ sweep



Consider two "cyclic-by-column" sweeps:

i) "top-to-bottom": $(2,1),\ldots,(n,1),(3,2),\ldots,(n,2),\ldots,(n,n-1)$

Surprise: sweep $\neq$ sweep

$$\Uparrow \quad \Rightarrow \quad \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \circ & \bullet \end{bmatrix}$$

Consider two "cyclic-by-column" sweeps:

i) "top-to-bottom": $(2,1), \ldots, (n,1), (3,2), \ldots, (n,2), \ldots, (n, n-1)$

ii) "bottom-to-top": $(n,1), \ldots, (2,1), (n,2), \ldots, (3,2), \ldots, (n, n-1)$

Surprise: sweep $\neq$ sweep

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \circ & \bullet \end{bmatrix}$$

Consider two “cyclic-by-column” sweeps:

i) “top-to-bottom”: $(2,1),\ldots,(n,1),(3,2),\ldots,(n,2),\ldots,(n,n-1)$

ii) “bottom-to-top”: $(n,1),\ldots,(2,1),(n,2),\ldots,(3,2),\ldots,(n,n-1)$

symmetric case:
quadratic asymptotic convergence for both sweep selections;

Surprise: sweep $\neq$ sweep

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \bullet & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \bullet & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \bullet & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \bullet & \bullet \\ \circ & \circ & \circ & \circ & \circ & \bullet \end{bmatrix}$$

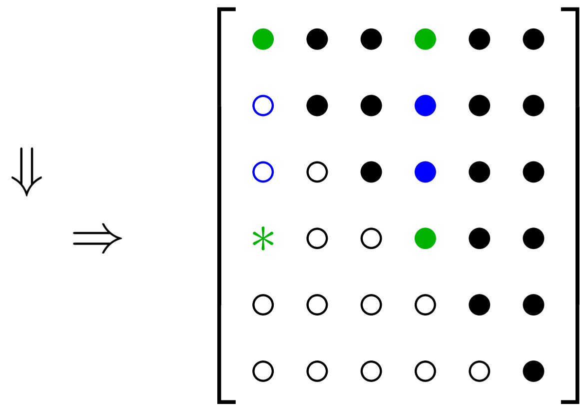
Consider two “cyclic-by-column” sweeps:

i) “top-to-bottom”: $(2,1),\ldots,(n,1),(3,2),\ldots,(n,2),\ldots,(n,n-1)$

ii) “bottom-to-top”: $(n,1),\ldots,(2,1),(n,2),\ldots,(3,2),\ldots,(n,n-1)$

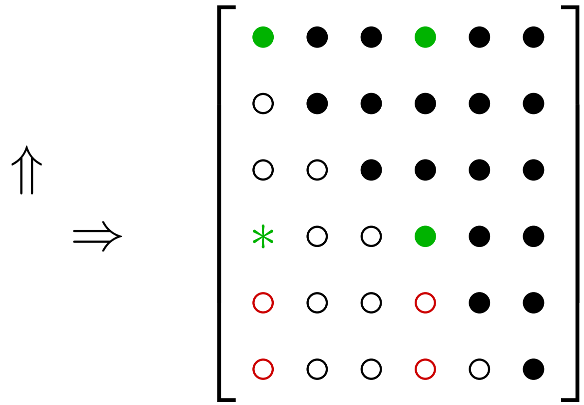
nonsymmetric case:
quadratic asymptotic convergence for “bottom-to-top”-sweeps,
but only linear asymptotic convergence for “top-to-bottom”-sweeps;

Surprise: sweep $\neq$ sweep



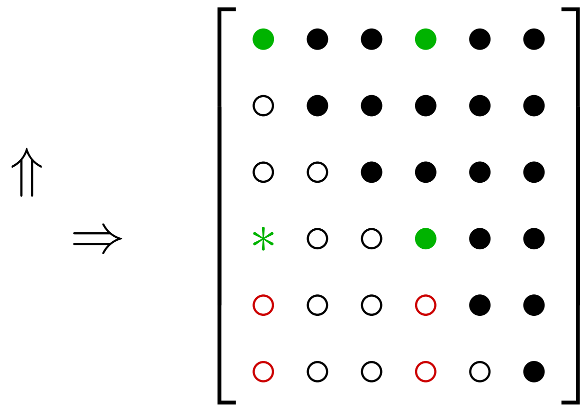
Heuristic explanation: in "top-to-bottom"-sweeps **elements** that have already been annihilated in the current sweep are recombined with potentially large elements from the upper triangular part.

Surprise: sweep $\neq$ sweep



Heuristic explanation: in "bottom-to-top"-sweeps such **elements** are recombined with usually small elements from the lower triangular part.

Surprise: sweep $\neq$ sweep



Theorem (M., 2008): The nonsymmetric Jacobi algorithm is asympotically quadratically convergent **if north-east-directed sweeps are used**.

north-east-directed sweep: the sequence of indices $((i_1, j_1), ..., (i_N, j_N))$, $N = n(n-1)/2$ in which order the elements are annihilated satisfies

$$\nu < \mu \quad \Rightarrow \quad (i_\nu > i_\mu \text{ or } j_\nu < j_\mu)$$

Nonsymmetric Jacobi method

Properties:

- „off-Norm“ $\text{off}(M) = \sqrt{\sum_{j<i} m_{ij}^2}$ is *not* reduces in each step
- experimentally: global convergence;
  open problem: proof of convergence
- experimentally: asymptotic linear or quadratic convergence;
  proof of local convergence for north-east-directed sweeps;
- fast if matrix close to Schur form
- in general much slower than QR (not competitive)

Method 2: A Jacobi-like method

Idea: Annihilate one diagonal or two off diagonal pivot elements in the strict upper anti-triangular part of Z in each Jacobi-step:

$$
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & {\color{red}\bullet} & \cdot & \cdot & \cdot & \cdot & * & * \\
\cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & * & * & * & * & * & * \\
\cdot & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
\qquad
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & * & * \\
\cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & * & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & * & * & * & * & * & * \\
\cdot & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
$$

This can always be achieved via a unitary consimilarity transformation.

Method 2: A Jacobi-like method

Diagonal pivots:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & {\color{red}\bullet} & \cdot & \cdot & \cdot & \cdot & {\color{blue}\bullet} & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & \cdot & \cdot & * & * & * & * \\ \cdot & \cdot & \cdot & * & * & * & * & * \\ \cdot & \cdot & * & * & * & * & * & * \\ \cdot & {\color{blue}\bullet} & * & * & * & * & {\color{blue}\bullet} & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

Consider the colored 2×2 subproblem:

$$\begin{bmatrix} {\color{red}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}$$

Method 2: A Jacobi-like method

Diagonal pivots:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & {\color{red}\bullet} & \cdot & \cdot & \cdot & \cdot & {\color{blue}\bullet} & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & \cdot & \cdot & * & * & * & * \\ \cdot & \cdot & \cdot & * & * & * & * & * \\ \cdot & \cdot & * & * & * & * & * & * \\ \cdot & {\color{blue}\bullet} & * & * & * & * & {\color{blue}\bullet} & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

Compute the anti-triangular form of the 2×2 problem:

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \begin{bmatrix} {\color{red}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} {\color{red}\circ} & * \\ * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Diagonal pivots:

$$
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & & u_{21} & \\
 & & 1 & & & & & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & & & & & 1 & & \\
 & u_{12} & & & & & u_{22} & \\
 & & & & & & & 1
\end{bmatrix}
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & \color{red}{\bullet} & \cdot & \cdot & \cdot & \cdot & \color{blue}{\bullet} & * \\
\cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & * & * & * & * & * & * \\
\cdot & \color{blue}{\bullet} & * & * & * & * & \color{blue}{\bullet} & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & & u_{12} & \\
 & & 1 & & & & & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & & & & & 1 & & \\
 & u_{21} & & & & & u_{22} & \\
 & & & & & & & 1
\end{bmatrix}
$$

Then update the $n \times n$ matrix.

$$
\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}
\begin{bmatrix} \color{red}{\bullet} & \color{blue}{\bullet} \\ \color{blue}{\bullet} & \color{blue}{\bullet} \end{bmatrix}
\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}
=
\begin{bmatrix} \color{red}{\circ} & * \\ * & * \end{bmatrix}
$$

Method 2: A Jacobi-like method

Diagonal pivots:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \circ & \cdot & \cdot & \cdot & \cdot & * & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & \cdot & \cdot & * & * & * & * \\ \cdot & \cdot & \cdot & * & * & * & * & * \\ \cdot & \cdot & * & * & * & * & * & * \\ \cdot & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

Then update the $n \times n$ matrix.

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} \circ & * \\ * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Off-diagonal pivots:

$$
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & \cdot & \bullet & \cdot & \cdot & \cdot & * & * \\
\cdot & \bullet & \cdot & \cdot & \cdot & * & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & * & * & * & * & * & * \\
\cdot & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
$$

Question: Why consider two pivots?

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & {\color{blue}\bullet} & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & \cdot & \cdot & * & * & * & * \\ \cdot & \cdot & \cdot & * & * & * & * & * \\ \cdot & \cdot & {\color{blue}\bullet} & * & * & * & {\color{blue}\bullet} & * \\ \cdot & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

Consider the colored 2×2 problem:

$$\begin{bmatrix} {\color{green}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}$$

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \cdot & \color{green}{\bullet} & \cdot & \cdot & \cdot & \color{blue}{\bullet} & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & \cdot & \cdot & * & * & * & * \\ \cdot & \cdot & \cdot & * & * & * & * & * \\ \cdot & \cdot & \color{blue}{\bullet} & * & * & * & \color{blue}{\bullet} & * \\ \cdot & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

Compute the anti-triangular form of the 2×2 problem:

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \begin{bmatrix} \color{green}{\bullet} & \color{blue}{\bullet} \\ \color{blue}{\bullet} & \color{blue}{\bullet} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} \color{green}{\circ} & * \\ * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ast \\ \cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & {\color{blue}\bullet} & \ast \\ \cdot & \cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast \\ \cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast \\ \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast & \ast \\ \cdot & \cdot & {\color{blue}\bullet} & \ast & \ast & \ast & {\color{blue}\bullet} & \ast \\ \cdot & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\ \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \end{bmatrix}$$

We may use different unitary transformation from the left and the right,

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \begin{bmatrix} {\color{green}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} {\color{green}\circ} & \ast \\ \ast & \ast \end{bmatrix}$$

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{21} & & \\
 & & v_{11} & & & & v_{21} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{12} & & & & u_{22} & & \\
 & & v_{12} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ast \\
\cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & {\color{blue}\bullet} & \ast \\
\cdot & \cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast \\
\cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast \\
\cdot & \cdot & \cdot & \ast & \ast & \ast & \ast & \ast \\
\cdot & \cdot & {\color{blue}\bullet} & \ast & \ast & \ast & {\color{blue}\bullet} & \ast \\
\cdot & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast
\end{bmatrix}
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{12} & & \\
 & & v_{11} & & & & v_{12} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{21} & & & & u_{22} & & \\
 & & v_{21} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
$$

because of the position of the subproblem. $\leadsto$ more freedom in the parameters

$$
\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}
\begin{bmatrix} {\color{green}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}
\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}
=
\begin{bmatrix} {\color{green}\circ} & \ast \\ \ast & \ast \end{bmatrix}
$$

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{21} & & \\
 & & v_{11} & & & & v_{21} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{12} & & & & u_{22} & & \\
 & & v_{12} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & {\color{blue}\bullet} & * \\
\cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & \bullet & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & {\color{blue}\bullet} & * & * & * & {\color{blue}\bullet} & * \\
\cdot & \bullet & * & * & * & \bullet & * & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{12} & & \\
 & & v_{11} & & & & v_{12} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{21} & & & & u_{22} & & \\
 & & v_{21} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
$$

Simultaneously, a second 2×2 system – marked by $\bullet$ – will be transformed.

$$
\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}
\begin{bmatrix} {\color{green}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}
\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}
=
\begin{bmatrix} {\color{green}\circ} & * \\ * & * \end{bmatrix}
$$

$$
\begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}
\begin{bmatrix} {\color{green}\bullet} & \bullet \\ \bullet & \bullet \end{bmatrix}
\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}
=
\begin{bmatrix} {\color{green}*} & * \\ * & * \end{bmatrix}
$$

Method 2: A Jacobi-like method

Off-diagonal pivots: assume, we only consider one off-diagonal pivot;

$$
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{21} & & \\
 & & v_{11} & & & & v_{21} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{12} & & & & u_{22} & & \\
 & & v_{12} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ast \\
\cdot & \cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & {\color{blue}\bullet} & \ast \\
\cdot & {\color{green}\bullet} & \cdot & \cdot & \cdot & \bullet & \ast & \ast \\
\cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast \\
\cdot & \cdot & \cdot & \ast & \ast & \ast & \ast & \ast \\
\cdot & \cdot & {\color{blue}\bullet} & \ast & \ast & \ast & {\color{blue}\bullet} & \ast \\
\cdot & \bullet & \ast & \ast & \ast & \bullet & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast
\end{bmatrix}
\begin{bmatrix}
1 & & & & & & & \\
 & u_{11} & & & & u_{12} & & \\
 & & v_{11} & & & & v_{12} & \\
 & & & 1 & & & & \\
 & & & & 1 & & & \\
 & u_{21} & & & & u_{22} & & \\
 & & v_{21} & & & & v_{22} & \\
 & & & & & & & 1
\end{bmatrix}
$$

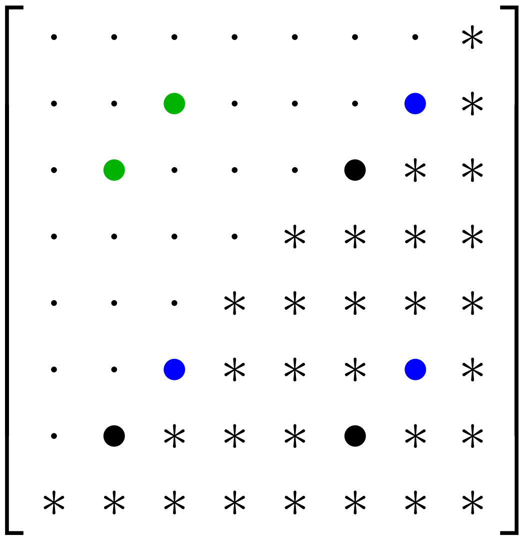
Use the freedom in the parameters to anti-triangularize the second system as well.

$$
\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}
\begin{bmatrix} {\color{green}\bullet} & {\color{blue}\bullet} \\ {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}
\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}
=
\begin{bmatrix} {\color{green}\circ} & \ast \\ \ast & \ast \end{bmatrix}
$$

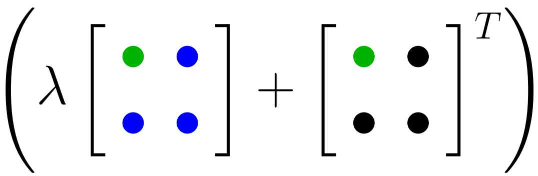
$$
\begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}
\begin{bmatrix} {\color{green}\bullet} & \bullet \\ \bullet & \bullet \end{bmatrix}
\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}
=
\begin{bmatrix} {\color{green}\ast} & \ast \\ \ast & \ast \end{bmatrix}
$$

Method 2: A Jacobi-like method

Off-diagonal pivots:

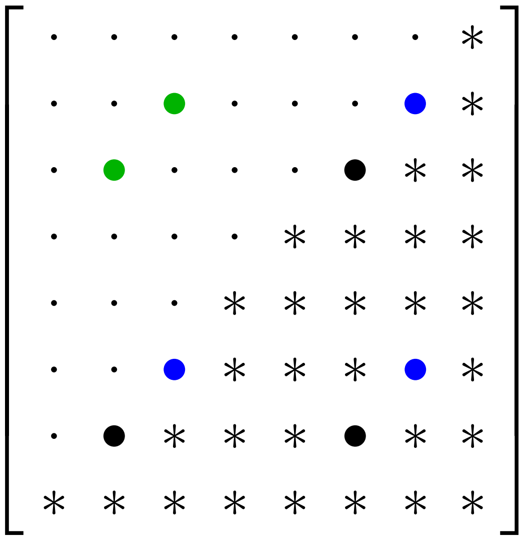


Anti-triangularize the colored/black generalized 2×2 problem:



Method 2: A Jacobi-like method

Off-diagonal pivots:



Anti-triangularize the colored/black generalized 2×2 problem:

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \left(\lambda \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix} + \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}^T \right) \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \lambda \begin{bmatrix} \circ & * \\ * & * \end{bmatrix} + \begin{bmatrix} \circ & * \\ * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Off-diagonal pivots:

$$
\begin{bmatrix}
1 &&&&&&& \\
& u_{11} &&&& u_{21} && \\
&& v_{11} &&&& v_{21} & \\
&&& 1 &&&& \\
&&&& 1 &&& \\
& u_{12} &&&& u_{22} && \\
&& v_{12} &&&& v_{22} & \\
&&&&&&& 1
\end{bmatrix}
\begin{bmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * \\
\cdot & \cdot & \color{green}{\bullet} & \cdot & \cdot & \cdot & \color{blue}{\bullet} & * \\
\cdot & \color{green}{\bullet} & \cdot & \cdot & \cdot & \bullet & * & * \\
\cdot & \cdot & \cdot & \cdot & * & * & * & * \\
\cdot & \cdot & \cdot & * & * & * & * & * \\
\cdot & \cdot & \color{blue}{\bullet} & * & * & * & \color{blue}{\bullet} & * \\
\cdot & \bullet & * & * & * & \bullet & * & * \\
* & * & * & * & * & * & * & *
\end{bmatrix}
\begin{bmatrix}
1 &&&&&&& \\
& u_{11} &&&& u_{12} && \\
&& v_{11} &&&& v_{12} & \\
&&& 1 &&&& \\
&&&& 1 &&& \\
& u_{21} &&&& u_{22} && \\
&& v_{21} &&&& v_{22} & \\
&&&&&&& 1
\end{bmatrix}
$$

Update the $n \times n$ matrix.

$$
\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}
\left(\lambda \begin{bmatrix} \color{green}{\bullet} & \color{blue}{\bullet} \\ \color{blue}{\bullet} & \color{blue}{\bullet} \end{bmatrix} + \begin{bmatrix} \color{green}{\bullet} & \bullet \\ \bullet & \bullet \end{bmatrix}^T \right)
\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}
= \lambda \begin{bmatrix} \color{green}{\circ} & * \\ * & * \end{bmatrix} + \begin{bmatrix} \color{green}{\circ} & * \\ * & * \end{bmatrix}
$$

Method 2: A Jacobi-like method

Off-diagonal pivots:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ast \\ \cdot & \cdot & \circ & \cdot & \cdot & \cdot & \ast & \ast \\ \cdot & \circ & \cdot & \cdot & \cdot & \ast & \ast & \ast \\ \cdot & \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast \\ \cdot & \cdot & \cdot & \ast & \ast & \ast & \ast & \ast \\ \cdot & \cdot & \ast & \ast & \ast & \ast & \ast & \ast \\ \cdot & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\ \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \end{bmatrix}$$

Update the $n \times n$ matrix.

$$\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix} \left(\lambda \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix} + \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}^T \right) \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \lambda \begin{bmatrix} \circ & \ast \\ \ast & \ast \end{bmatrix} + \begin{bmatrix} \circ & \ast \\ \ast & \ast \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \circ & \cdot & \cdot & \cdot & \cdot & \bullet \\ \cdot & \cdot & \cdot & \cdot & * & * \\ \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & * & * & * & * \\ \cdot & * & * & * & * & * \\ \bullet & * & * & * & * & \bullet \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \circ & \cdot & \cdot & \cdot & \bullet \\ \circ & \cdot & \cdot & \cdot & \bullet & * \\ \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & * & * & * & * \\ \cdot & \bullet & * & * & * & \bullet \\ \bullet & * & * & * & \bullet & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

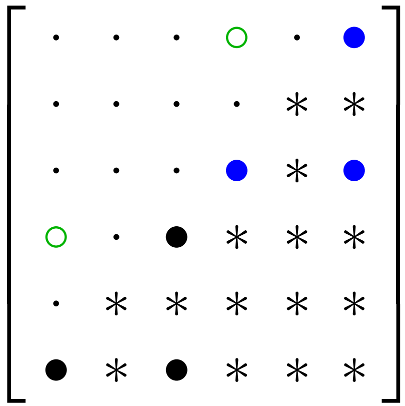
E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \cdot & \circ & \cdot & \cdot & \bullet \\ \cdot & \cdot & \cdot & \cdot & * & * \\ \circ & \cdot & \cdot & \bullet & * & * \\ \cdot & \cdot & \bullet & * & * & \bullet \\ \cdot & * & * & * & * & * \\ \bullet & * & * & \bullet & * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:



Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \circ & \bullet \\ \cdot & \cdot & \cdot & \cdot & \bullet & \bullet \\ \cdot & \cdot & \cdot & \ast & \ast & \ast \\ \cdot & \cdot & \ast & \ast & \ast & \ast \\ \circ & \bullet & \ast & \ast & \ast & \ast \\ \bullet & \bullet & \ast & \ast & \ast & \ast \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \circ & \cdot & \cdot & \bullet & * \\ \cdot & \cdot & \cdot & * & * & * \\ \cdot & \cdot & * & * & * & * \\ \cdot & \bullet & * & * & \bullet & * \\ * & * & * & * & * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \cdot & \circ & \cdot & \bullet & * \\ \cdot & \circ & \cdot & \bullet & * & * \\ \cdot & \cdot & \bullet & * & \bullet & * \\ \cdot & \bullet & * & \bullet & * & * \\ * & * & * & * & * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

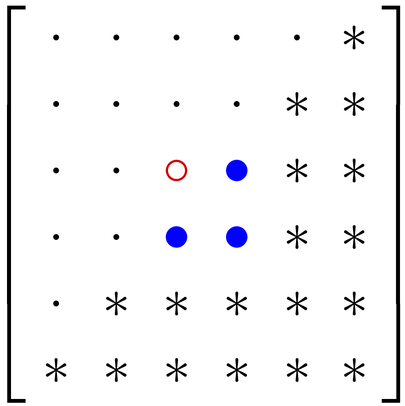
E.g., cyclic-by-row-sweep:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & * \\ \cdot & \cdot & \cdot & \circ & \bullet & * \\ \cdot & \cdot & \cdot & \bullet & \bullet & * \\ \cdot & \circ & \bullet & * & * & * \\ \cdot & \bullet & \bullet & * & * & * \\ * & * & * & * & * & * \end{bmatrix}$$

Method 2: A Jacobi-like method

Sweep: Annihilate each pivot element at least once.

E.g., cyclic-by-row-sweep:



Method 2: A Jacobi-like method

Properties of the algorithm:

+ locally and asymptotically quadratically convergent;

+ globally convergent in experiments;

+ converges fast for matrices Z close to anti-triangular form

− expensive in general (cost of 3 sweeps $\hat{=}$ cost of `QZ`)

− convergence problems for badly scaled problems

− convergence problems for large n

Method No. 3: A Hybrid Method

Method 3: A hybrid method

Laub-trick:

$+$ works for moderate sizes of n;

$+$ essentially cost of `QZ`;

$-$ problems for eigenvalues with modulus near one;

Jacobi:

$+$ works nicely if problem is small and eigenvalues do not differ too much in modulus;

Idea: Combine the positive properties of these two algorithms. Use the Laub-trick for getting all eigenvalues sufficiently far away from the unit circle and use Jacobi for the eigenvalues near the unit circle.

Method 3: A hybrid method

$$\lambda Z + Z^T = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ 0 & X_{22} & X_{23} \\ 0 & 0 & X_{33} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ 0 & Y_{22} & Y_{23} \\ 0 & 0 & Y_{33} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

Step 1: Given a tolerance $\alpha > 1$ and a regular $\lambda Z + Z^T \in \mathbb{C}^{2n \times 2n}$, compute its generalized Schur decomposition, where the eigenvalues are ordered in such a way that

$$\begin{aligned} \sigma(\lambda X_{11} + Y_{11}) &\subseteq \{\lambda \in \mathbb{C} : |\lambda| \geq \alpha\}, \\ \sigma(\lambda X_{22} + Y_{22}) &\subseteq \{\lambda \in \mathbb{C} : \alpha > |\lambda| > \frac{1}{\alpha}\}, \\ \sigma(\lambda X_{33} + Y_{33}) &\subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq \frac{1}{\alpha}\}. \end{aligned}$$

Method 3: A hybrid method

$$\lambda Z + Z^T = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ 0 & X_{22} & X_{23} \\ 0 & 0 & X_{33} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ 0 & Y_{22} & Y_{23} \\ 0 & 0 & Y_{33} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

Step 2: By the Laub trick, the matrix

$$\begin{bmatrix} W_{11} & Q_{11}^T R_m \\ W_{21} & Q_{12}^T R_m \\ W_{31} & Q_{13}^T R_m \end{bmatrix}$$

has orthonormal columns. Extend this matrix to a unitary matrix

$$U := \begin{bmatrix} W_{11} & U_{12} & Q_{11}^T R_m \\ W_{21} & U_{22} & Q_{12}^T R_m \\ W_{31} & U_{32} & Q_{13}^T R_m \end{bmatrix}.$$

Method 3: A hybrid method

$$\lambda Z + Z^T = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ 0 & X_{22} & X_{23} \\ 0 & 0 & X_{33} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ 0 & Y_{22} & Y_{23} \\ 0 & 0 & Y_{33} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

Step 3: Compute

$$U^T Z U = \begin{bmatrix} 0 & 0 & Y_{11}^T R_m \\ 0 & Z_{22} & Z_{23} \\ R_m X_{11} & Z_{32} & Z_{33} \end{bmatrix},$$

where $Y_{11}^T R_m \in \mathbb{C}^{m \times m}$ and $R_m X_{11} \in \mathbb{C}^{m \times m}$ are in anti-triangular form and $Z_{22} \in \mathbb{C}^{(n-2m)\times(n-2m)}$ has only eigenvalues in $\{\lambda \in \mathbb{C} : \alpha > |\lambda| > \frac{1}{\alpha}\}$.

Method 3: A hybrid method

$$\lambda Z + Z^T = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \left(\lambda \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ 0 & X_{22} & X_{23} \\ 0 & 0 & X_{33} \end{bmatrix} + \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ 0 & Y_{22} & Y_{23} \\ 0 & 0 & Y_{33} \end{bmatrix} \right) \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

Step 3: Compute

$$U^T Z U = \begin{bmatrix} 0 & 0 & Y_{11}^T R_m \\ 0 & Z_{22} & Z_{23} \\ R_m X_{11} & Z_{32} & Z_{33} \end{bmatrix},$$

where $Y_{11}^T R_m \in \mathbb{C}^{m \times m}$ and $R_m X_{11} \in \mathbb{C}^{m \times m}$ are in anti-triangular form and $Z_{22} \in \mathbb{C}^{(n-2m)\times(n-2m)}$ has only eigenvalues in $\{\lambda \in \mathbb{C} : \alpha > |\lambda| > \frac{1}{\alpha}\}$.

Step 4: Anti-triangularize Z_{22} by some expensive, but accurate method (e.g., palindromic Jacobi algorithm, Schröder's palindromic QR algorithm).

Method No. 4: Palindromic QR

Method No. 4: Palindromic QR

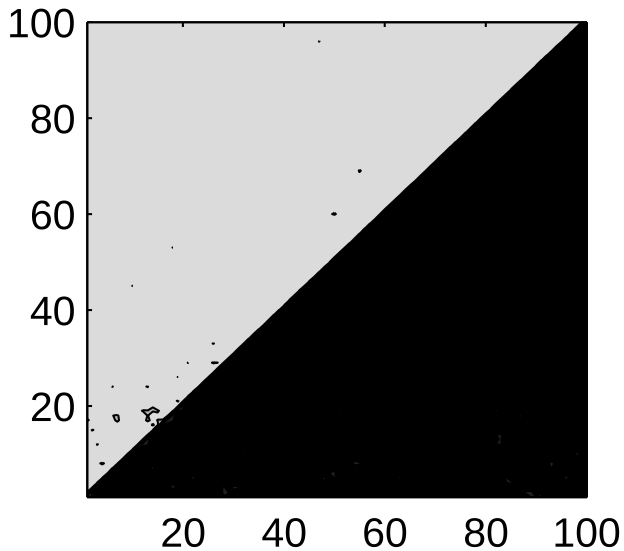
Ph.D. thesis by Christian Schröder (2008) – too much for this talk

Numerical experiments

Numerical experiments

Test: 100 random 100×100 matrices with 10 eigenvalues in an annulus in the complex plane with outer radius $1+10^{-12}$ and inner radius $1/(1+10^{-12})$.

Typical behavior: (black: elements of modulus larger than one; light grey: elements of modulus 10^{-15})

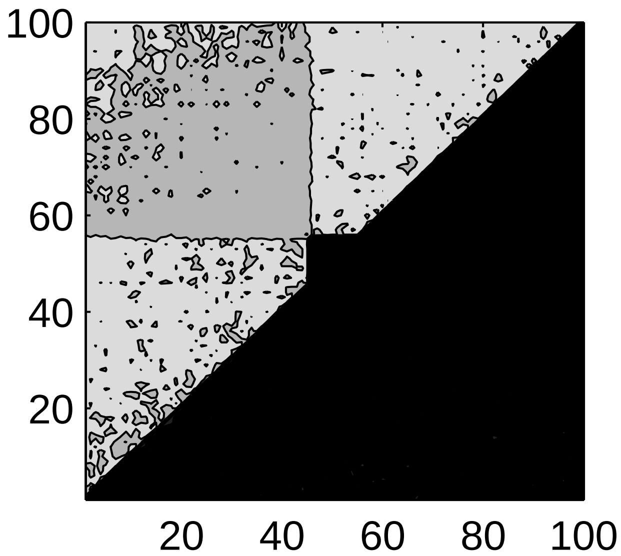


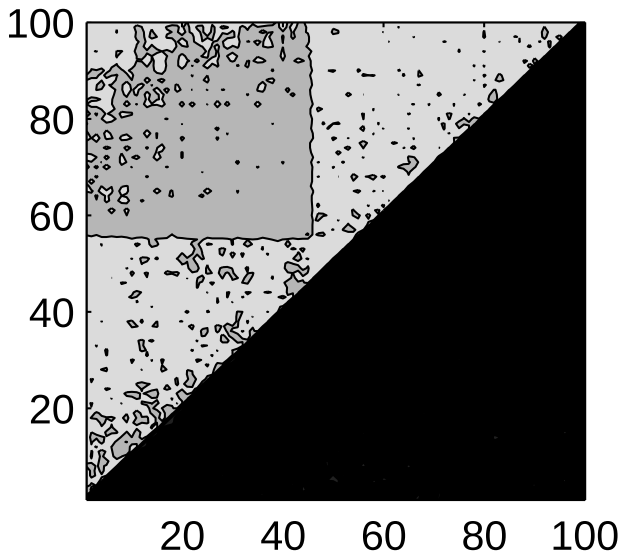
similar results when solving the small subproblem with palindromic Jacobi or the palindromic QR

Numerical experiments

Tough test: random 100×100 matrices with 10 eigenvalues in a disk with radius 10^{-12} and center 1.

Typical behavior: (black: elements of modulus larger than one; light grey: elements of modulus 10^{-15})



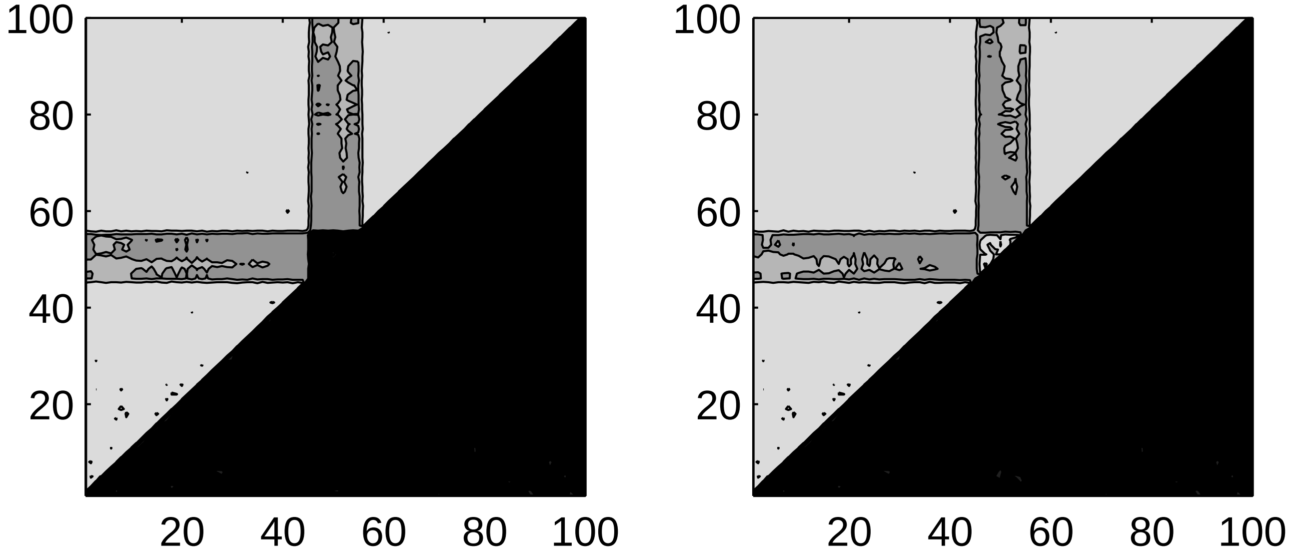


left: method 4 before applying palindromic QR to the small problem
right: method 4 before applying palindromic QR to the small problem

Numerical experiments

Tough test: random 100×100 matrices with 10 eigenvalues in a disk with radius 10^{-12} and center 1.

Typical behavior: (black: elements of modulus larger than one; light grey: elements of modulus 10^{-15})

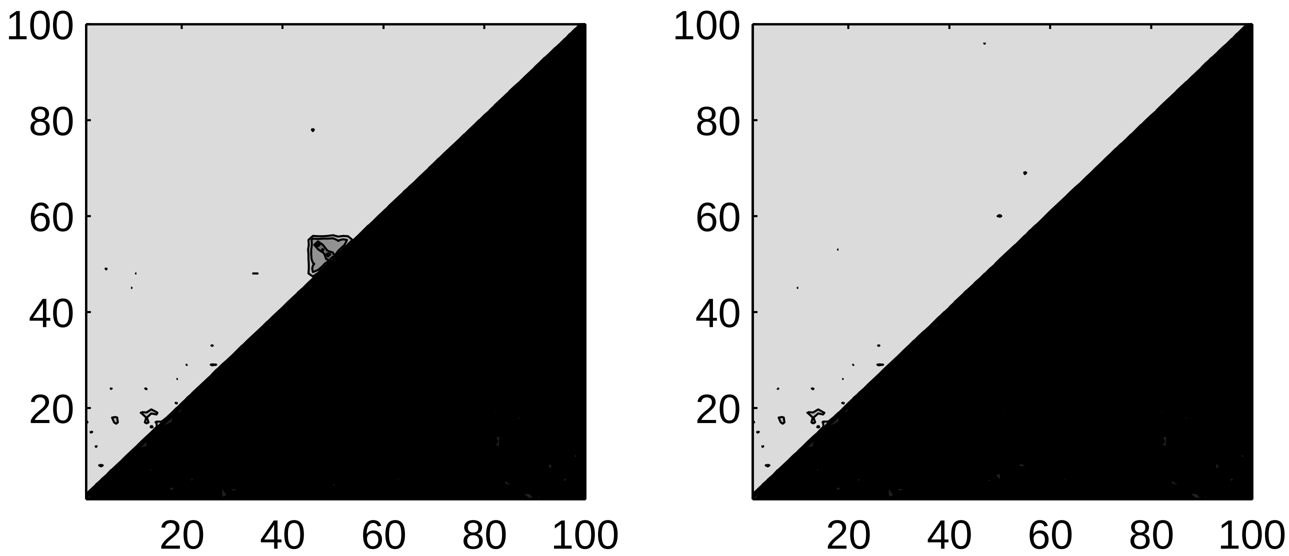


left: method 4 followed by a full sweep of Jacobi before...
right: ...and after solving the small problem by the palindromic QR

Numerical experiments

Tough test: random 100×100 matrices with 10 eigenvalues in a disk with radius 10^{-12} and center 1.

Typical behavior: (black: elements of modulus larger than one; light grey: elements of modulus 10^{-15})



left: method 4 with palindromic QR, then full Jacobi sweep before...
right: and after solving the small problem once again with palindromic QR

Thank you for your attention!