Staircase forms for structured matrix pencils (and matrix polynomials)

Volker Mehrmann
TU Berlin, Inst. f. Mathematik
http://www.math.tu-berlin.de/~mehrmann/

Work supported by MATHEON, the DFG research center
Mathematics for key technologies in Berlin.
Joint work with Ralph Byers and Hongguo Xu
also results by Christian Schröder and Lena Wunderlich

Overview

- Notation.
- (Structured) staircase forms: Why bother?
- Applications.
- Structured canonical form for even/palindromic pencils.
- Structured staircase form for even/palindromic pencils.
- Computational issues.
- Conclusion and open problems.

Notation

Definition A matrix pencil of the form

$$\lambda N + M,\ M, N \in \mathbb{C}^{n,n}$$

is called

- **even** if $\lambda N + M = (-\lambda N + M)^H$, i.e. $M = M^H$ and $N = -N^H$;
- **palindromic** if $\lambda N + M = (N + \overline{\lambda} M)^H$, i.e., $M = N^H$.

Definition Let $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$. A matrix M is called

- **Hamiltonian** if $(JM) = (JM)^H$;
- **symplectic** if $M^H J M = M$.

Analogous definitions for the real and complex ^{T} case.

Observations:

Even matrix pencils generalize *Hamiltonian matrices*

Palindromic matrix pencils (almost) generalize *symplectic matrices*.

The generalized Cayley transformation $\mathcal{C}(\lambda N + M) = \mu(M - N) + (M + N)$ of an even pencil is palindromic and that of a palindromic pencil is even.

Generalization to even and palindromic matrix polynomials, Mackey^2/Mehl/M. 2006 (IWASEP V).

(Structured) staircase forms. Why bother?

- Numerical methods to compute reliable information about Jordan/Kronecker structure, sizes of blocks, eigenvalues, deflating and reducing spaces.
- Error and condition estimates.
- Deflation of singular blocks, and blocks associated with eigenvalues $\infty, 0$ in the even case, $1, -1$ in the palindromic case.

General Philosophy in Numerical Analysis

A numerical algorithm should

- be **numerically (backward) stable**, i.e. the computed solution is the exact solution of a nearby problem;
- be **as efficient as possible**, i.e. for dense eigenvalue problems the complexity should be $\mathbf{O}(n^3)$ or better;
- **reflect the structure of the physical problem and the mathematical model**, i.e. in our case preserve the even or palindromic structure;
- be as **accurate as possible even in the extreme (ill-conditioned) cases**.

Applications.

- Singularities (cracks) in anisotropic materials as functions of material or geometry parameters. Apel/M./Watkins 2002.
- Optimal control of continuous and discrete time systems, M. 1991.
- Optimal control of variable coefficient systems, Kunkel/M. 2006.
- Robust H_∞ control, Benner/Byers/M./Xu 2005.
- Optimal waveguide design, Schmidt/Friese/Zschiedrich/Deuflhard 2003.
- Resonance phenomena in tracks excited by high speed trains, Hilliges/Mehl/M. 2005.
- Passivity checking, and passivization of reduced order models in electrical field computation.
  Recent project with company CST in Darmstadt.

Optimal L_2 control for variable coefficient DAEs.

In the γ-iteration for the solution of the continuous time optimal H_∞ control problem one has to solve a boundary value problem with a pencil of matrix valued functions

$$\lambda \begin{bmatrix} 0 & E(t) & 0 \\ -E^T(t) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & A(t) & B(t) \\ (A+\dot{E})^T(t) & Q(t) & S(t) \\ B^T(t) & S^T(t) & R(t) \end{bmatrix}$$

$Q = Q^T, R = R^T$.

Goal: Determine nullspaces, and singular subspaces, index as functions of t.
Structure preserving methods, Lena Wunderlich 2006.

Robust H_∞ control.

In the γ-iteration for the solution of the continuous time optimal H_∞ control problem one has to solve the parametric even eigenvalue problem

$$\lambda \begin{bmatrix} 0 & E & 0 \\ -E^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ A^T & Q & S \\ B^T & S^T & R(\gamma) \end{bmatrix}$$

for a sequence of parameters γ. (Blocks are real.)

Goal: Determine $\gamma \in [0, \infty]$, where there is a multiple eigenvalue on the imaginary axis, the multiplicity at ∞ changes or the pencil becomes singular.

Passivity enforcing. Project with company CST (Darmstadt)

Rational matrix valued function $Y(s)$ (admittance matrix), arising from reduced order modelling of network equations including electrical field computation.

A network with admittance matrix $Y_p(s)$, depending on several parameters p is called passive if $Y_p(\overline{s}) = Y_p^H(s)$ and if $z^H(Y_p(s)^H + Y_p(s))z \geq 0$ for all $s \in \mathbb{C}$ and all $z \in \mathbb{C}^n$.

To check passivity, one performs a minimal realization

$$Y_p(s) = C(p)(sI - A(p))^{-1}B(p) + D(p)$$

and one computes the purely imaginary eigenvalues, infinite eigenvalues and singular blocks of the even pencil

$$\lambda \begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & A(p) & B(p) \\ A^T(p) & 0 & C^T(p) \\ B^T(p) & C(p) & D(p) + D(p)^T \end{bmatrix}$$

for sets of parameters.

Properties of even (palindromic) pencils

Proposition Consider a real even pencil $\lambda N + M$.
Then $(\lambda N + M)x = 0$ if and only if $x^T(-\lambda N + M) = 0$,
i.e., the eigenvalues occur in pairs λ, $-\lambda$ (Hamiltonian spectrum.)

Consider a real palindromic pencil $\lambda A^T + A$.
Then $(\lambda A^T + A)x = 0$ if and only if $x^T(A^T + \lambda A) = 0$,
i.e., the eigenvalues occur in pairs λ, $1/\lambda$ (symplectic spectrum.)

- In many applications the current methods fail near the extreme cases.
- These are the eigenvalues on or near the imaginary axis (unit circle) and in particular the eigenvalues 0 and ∞, (1 and -1).
- The (structured) perturbation theory near the extreme cases is mostly **open**, Results of Ran/Rodman 1988, Godunov/Sadkane 1997 imply that non-structured methods will usually fail.

Structure preserving equivalence transformations

To preserve the even(palindromic) structure, we use congruence transformations

$$\begin{aligned} \lambda\tilde{N}+\tilde{M} &= \lambda U^T N U + U^T M U, \\ \lambda\tilde{A}^T+\tilde{A} &= \lambda U^T A^T U + U^T A U, \end{aligned}$$

with nonsingular (unitary) U.

What is the structured canonical form under this transformation?

What is the structured condensed form under unitary transformations?

Can we compute it numerically?

A structured Kronecker form for even pencils

Theorem: Thompson 1991 If $N, M \in \mathbb{R}^{n,n}$ with $N = -N^T, M = M^T$, then there exists a nonsingular matrix $X \in \mathbb{C}^{n,n}$ such that

$$X^T(\lambda N + M)X = \mathrm{diag}(\mathcal{B}_{\mathcal{S}}, \mathcal{B}_{\mathcal{I}}, \mathcal{B}_{\mathcal{Z}}, \mathcal{B}_{\mathcal{F}}),$$

is in structured Kronecker form, where

$$\begin{array}{rcl} \mathcal{B}_{\mathcal{S}} & = & \mathrm{diag}(\mathcal{O}_\eta, \mathcal{S}_{\xi_1}, \dots, \mathcal{S}_{\xi_k}), \\ \mathcal{B}_{\mathcal{I}} & = & \mathrm{diag}\left(\mathcal{I}_{2\varepsilon_1+1}, \dots, \mathcal{I}_{2\varepsilon_l+1}, \mathcal{I}_{2\delta_1}, \dots, \mathcal{I}_{2\delta_m}\right), \\ \mathcal{B}_{\mathcal{Z}} & = & \mathrm{diag}\left(\mathcal{Z}_{2\sigma_1+1}, \dots, \mathcal{Z}_{2\sigma_r+1}, \mathcal{Z}_{2\rho_1}, \dots, \mathcal{Z}_{2\rho_s}\right), \\ \mathcal{B}_{\mathcal{F}} & = & \mathrm{diag}(\mathcal{R}_{\phi_1}, \dots, \mathcal{R}_{\phi_t}, \mathcal{C}_{\psi_1}, \dots, \mathcal{C}_{\psi_u}) \end{array}$$

This structured Kronecker canonical form is unique up to permutation of the blocks, i.e., the kind, size and number of the blocks as well as the sign characteristics are characteristic of the pencil $\lambda N + M$.

The blocks have the following properties.

1. $\mathcal{O}_\eta = \lambda 0_\eta + 0_\eta$;

2. Each $\mathcal{S}_{\xi_j}$ is a $(2\xi_j+1)\times(2\xi_j+1)$ block that combines a right singular block and a left singular block, both of minimal index ξ_j. It has the form

$$\lambda \left[\begin{array}{ccc|ccc} & & & & 1 & 0 \\ & & & \iddots & \iddots & \\ & & & 1 & 0 & \\ \hline & & -1 & & & \\ & \iddots & 0 & & & \\ -1 & \iddots & & & & \\ 0 & & & & & \end{array}\right] + \left[\begin{array}{ccc|ccc} & & & & 0 & 1 \\ & & & \iddots & \iddots & \\ & & & 0 & 1 & \\ \hline & & 0 & & & \\ & \iddots & 1 & & & \\ 0 & \iddots & & & & \\ 1 & & & & & \end{array}\right];$$

3. Each $\mathcal{I}_{2\varepsilon_j+1}$ is a $(2\varepsilon_j+1)\times(2\varepsilon_j+1)$ block that contains a single block corresponding to the eigenvalue ∞ with index $2\varepsilon_j+1$. It has the form

$$\lambda \left[\begin{array}{ccc|ccc} & & & & 1 & 0 \\ & & & \iddots & \iddots & \\ & & & 1 & 0 & \\ \hline & & -1 & 0 & & \\ & \iddots & 0 & & & \\ -1 & \iddots & & & & \\ 0 & & & & & \end{array}\right] + \left[\begin{array}{ccc|ccc} & & & & 0 & 1 \\ & & & \iddots & \iddots & \\ & & & 0 & 1 & \\ \hline & & 0 & s & & \\ & \iddots & 1 & & & \\ 0 & \iddots & & & & \\ 1 & & & & & \end{array}\right],$$

where $s \in \{1, -1\}$ is the *sign-index* or *sign-characteristic* of the block;

4. Each $\mathcal{I}_{2\delta_j}$ is a $4\delta_j \times 4\delta_j$ block that combines two $2\delta_j \times 2\delta_j$ infinite eigenvalue blocks of index δ_j. It has the form

$$\lambda \left[\begin{array}{cccc|cccc} & & & & & & 1 & 0 \\ & & & & & \iddots & \iddots & \\ & & & & 1 & \iddots & & \\ & & & & 0 & & & \\ \hline & & -1 & 0 & & & & \\ & \iddots & \iddots & & & & & \\ -1 & \iddots & & & & & & \\ 0 & & & & & & & \end{array}\right] + \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ \hline & & 1 & & & \\ & \iddots & & & & \\ 1 & & & & & \end{array}\right];$$

5. Each $\mathcal{Z}_{2\sigma_j+1}$ is a $(4\sigma_j+2) \times (4\sigma_j+2)$ block that combines two $(2\sigma_j+1) \times (2\sigma_j+1)$ Jordan blocks corresponding to the eigenvalue 0. It has the form

$$\lambda \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ \hline & & -1 & & & \\ & \iddots & & & & \\ -1 & & & & & \end{array}\right] + \left[\begin{array}{cccc|cccc} & & & & & & 1 & 0 \\ & & & & & \iddots & \iddots & \\ & & & & 1 & \iddots & & \\ & & & & 0 & & & \\ \hline & & 1 & 0 & & & & \\ & \iddots & \iddots & & & & & \\ 1 & \iddots & & & & & & \\ 0 & & & & & & & \end{array}\right];$$

6. Each $\mathcal{Z}_{2\rho_j}$ is a $2\rho_j \times 2\rho_j$ block that contains a single Jordan block corresponding to the eigenvalue 0. It has the form

$$\lambda \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ \hline & & -1 & & & \\ & \iddots & & & & \\ -1 & & & & & \end{array}\right] + \left[\begin{array}{cccc|cccc} & & & & & & 1 & 0 \\ & & & & & \iddots & \iddots & \\ & & & & 1 & \iddots & & \\ & & & s & 0 & & & \\ \hline & & 1 & 0 & & & & \\ & \iddots & \iddots & & & & & \\ 1 & \iddots & & & & & & \\ 0 & & & & & & & \end{array}\right],$$

where $s \in \{1, -1\}$ is the sign characteristic of this block;

7. Each $\mathcal{R}_{\phi_j}$ is a $2\phi_j \times 2\phi_j$ block that combines two $\phi_j \times \phi_j$ Jordan blocks corresponding to nonzero real eigenvalues a_j and $-a_j$. It has the form

$$\lambda \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ \hline & & -1 & & & \\ & \iddots & & & & \\ -1 & & & & & \end{array}\right] + \left[\begin{array}{cccc|cccc} & & & & & & 1 & a_j \\ & & & & & \iddots & \iddots & \\ & & & & 1 & \iddots & & \\ & & & & a_j & & & \\ \hline & & 1 & a_j & & & & \\ & \iddots & \iddots & & & & & \\ 1 & \iddots & & & & & & \\ a_j & & & & & & & \end{array}\right].$$

8. a. Either $\mathcal{C}_{\psi_j}$ is a $2\psi_j \times 2\psi_j$ block combining two $\psi_j \times \psi_j$ Jordan blocks with purely imaginary eigenvalues $ib_j, -ib_j$ $(b_j > 0)$. It has the form

$$\lambda\left[\begin{array}{ccc|ccc} &&&&&1\\ &&&&\iddots&\\ &&&1&&\\ \hline &&-1&&&\\ &\iddots&&&&\\ -1&&&&&\end{array}\right] + s\left[\begin{array}{cccc|cccc} &&1&b_j&&&&\\ &\iddots&\iddots&&&&&\\ 1&\iddots&&&&&&\\ b_j&&&&&&&\\ \hline &&&&&&1&b_j\\ &&&&&\iddots&\iddots&\\ &&&&1&\iddots&&\\ &&&&b_j&&&\end{array}\right],$$

where $s \in \{1, -1\}$ is the sign characteristic.

8. b. or $\mathcal{C}_{\psi_j}$ is a $4\psi_j \times 4\psi_j$ block combining $\psi_j \times \psi_j$ Jordan blocks for each of the complex eigenvalues $a_j + ib_j, a_j - ib_j, -a_j + ib_j, -a_j - ib_j$ (with $a_j \neq 0$ and $b_j \neq 0$). In this case it has form

$$\lambda\left[\begin{array}{ccc|ccc} &&&&&\Omega\\ &&&&\iddots&\\ &&&\Omega&&\\ \hline &&-\Omega&&&\\ &\iddots&&&&\\ -\Omega&&&&&\end{array}\right] + \left[\begin{array}{cccc|cccc} &&&&&&\Omega&\Lambda_j\\ &&&&&\iddots&\iddots&\\ &&&&\Omega&\iddots&&\\ &&&&\Lambda_j&&&\\ \hline &&\Omega&\Lambda_j&&&&\\ &\iddots&\iddots&&&&&\\ \Omega&\iddots&&&&&&\\ \Lambda_j&&&&&&&\end{array}\right]$$

with $\Omega = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\Lambda_j = \begin{bmatrix} -b_j & a_j \\ a_j & b_j \end{bmatrix}$.

A structured Kronecker form for palindromic pencils

Theorem: Schröder, Horn/Sergejchuk 2006

If $A \in \mathbb{R}^{n,n}$, then there exists a nonsingular matrix $X \in \mathbb{R}^{n,n}$ such that

$$\lambda X^T A^T X + X^T A X = \mathrm{diag}(\lambda A_1 + A_1^T, \ldots, \lambda A_\ell^T + A_\ell)$$

is in structured Kronecker form.

This structured Kronecker canonical form is unique up to permutation of the blocks, i.e., the kind, size and number of the blocks as well as the sign characteristics are characteristic of the pencil $\lambda A^T + A$.

The blocks have the following forms:

$$
S_p = \left[\begin{array}{cccc|cccc}
 & & & & & & & 0 \\
 & 0_{p+1} & & & & & \iddots & 1 \\
 & & & & 0 & \iddots & & \\
 & & & & 1 & & & \\
\hline
 & & 1 & 0 & & & & \\
 & \iddots & \iddots & & & 0_p & & \\
1 & 0 & & & & & &
\end{array}\right] \in \mathbb{R}^{2p+1,2p+1},\ p \in \mathbb{N}_0;
$$

$$
L_{1,p}(\lambda) = \left[\begin{array}{cccc|cccc}
 & & & & & & & \lambda \\
 & 0_p & & & & & \iddots & 1 \\
 & & & & & \iddots & \iddots & \\
 & & & & \lambda & 1 & & \\
\hline
 & & & 1 & & & & \\
 & & \iddots & & & & & \\
 & \iddots & & & & & 0_p & \\
1 & & & & & & &
\end{array}\right] \in \mathbb{R}^{2p,2p},
$$

where $p \in \mathbb{N}, \lambda \in \mathbb{R}, |\lambda| < 1;$

$$L_{2,p}(\alpha,\beta) = \left[\begin{array}{c|c} 0_{2p} & \begin{array}{cccc} & & & \Lambda \\ & & \iddots & I_2 \\ & \iddots & \iddots & \\ \Lambda & I_2 & & \end{array} \\ \hline \begin{array}{ccc} & & I_2 \\ & \iddots & \\ I_2 & & \end{array} & 0_{2p} \end{array}\right] \in \mathbb{R}^{4p,4p},$$

where $p \in \mathbb{N}, \Lambda = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}, \alpha, \beta \in \mathbb{R} \setminus \{0\}, \beta < 0, |\alpha + i\beta| < 1;$

$$\sigma U_{1,p} = \sigma \left[\begin{array}{c|c|c} 0_{\lfloor \frac{p}{2} \rfloor} & & \begin{array}{ccc} & & 1 \\ & \iddots & 1 \\ 1 & \iddots & \end{array} \\ \hline & 1 & 1 \\ \hline \begin{array}{ccc} & & 1 \\ & \iddots & \\ 1 & & \end{array} & & 0_{\lfloor \frac{p}{2} \rfloor} \end{array}\right] \in \mathbb{R}^{p,p}, \text{ where } p \in \mathbb{N} \text{ is odd}, \sigma \in \{1,-1\};$$

$$U_{2,p} = L_p(1) = \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & 0_p & & & \iddots & 1 \\ & & & \iddots & \iddots & \\ & & & 1 & 1 & \\ \hline & & 1 & & & \\ & \iddots & & & & \\ \iddots & & & & 0_p & \\ 1 & & & & & \end{array}\right] \in \mathbb{R}^{2p,2p},$$

where $p \in \mathbb{N}$ is even;

$$U_{3,p} = L_p(-1) = \left[\begin{array}{ccc|ccc} & & & & & -1 \\ & 0_p & & & \iddots & 1 \\ & & & \iddots & \iddots & \\ & & & -1 & 1 & \\ \hline & & 1 & & & \\ & \iddots & & & & \\ \iddots & & & & 0_p & \\ 1 & & & & & \end{array}\right] \in \mathbb{R}^{2p,2p},$$ where $p \in \mathbb{N}$ is odd;

$$\sigma U_{4,p} = \sigma \left[\begin{array}{cccc|cccc} & & & & & & & -1 \\ & 0_{\frac{p}{2}} & & & & & \iddots & 1 \\ & & & & & \iddots & \iddots & \\ & & & & -1 & 1 & & \\ \hline & & & 1 & 1 & & & \\ & & \iddots & & & & & \\ & \iddots & & & & & & \\ 1 & & & & & & & \end{array}\right] \in \mathbb{R}^{p,p},$$

where $p \in \mathbb{N}$ is even, $\sigma \in \{1, -1\}$;

$$\sigma U_{5,p}(\alpha,\beta) = \sigma \left[\begin{array}{c|c|c} 0_{2\lfloor \frac{p}{2} \rfloor} & & \begin{array}{ccc} & & \Lambda \\ & \iddots & I_2 \\ \Lambda & \iddots & \end{array} \\ \hline & \Lambda^{\frac{1}{2}} & I_2 \\ \hline \begin{array}{ccc} & & I_2 \\ & \iddots & \\ I_2 & & \end{array} & & 0_{2\lfloor \frac{p}{2} \rfloor} \end{array}\right] \in \mathbb{R}^{2p,2p},$$

where $p \in \mathbb{N}$ is odd, $|\alpha + i\beta| = 1, \beta < 0, \Lambda = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$,

$\Lambda^{\frac{1}{2}}$ is defined as rotation matrix with rotation angle $\frac{\phi}{2} \in (0,\pi)$,

where $\phi = \arctan(\frac{\beta}{\alpha})$ is the rotation angle of the rotation matrix Λ, $\sigma \in \{1,-1\}$;

$$
\sigma U_{6,p}(\alpha,\beta) = \sigma \left[\begin{array}{cccc|cccc}
 & & & & & & & \Lambda \\
 & 0_{2\frac{p}{2}} & & & & & \iddots & I_2 \\
 & & & & & \iddots & \iddots & \\
 & & & & \Lambda & I_2 & & \\
\hline
 & & & I_2 & I_2 & & & \\
 & & \iddots & & & & & \\
 & \iddots & & & & & & \\
I_2 & & & & & & &
\end{array}\right] \in \mathbb{R}^{2p,2p},
$$

where $p \in \mathbb{N}$ is even, $|\alpha + i\beta| = 1, \beta < 0, \Lambda = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}, \sigma \in \{1, -1\}$.

Similar results als in complex * and complex ^{T} case.

Corollary: Exact characterization when a symplectic matrix S has a factorization $S = A^{-T}A$ in terms of dimensions of blocks associated with eigenvalues 1 and -1. Schröder 2006.

Consequences of even (palindromic) Kronecker form

- The even (palindromic) Kronecker forms are nice theoretical results.
- The transformation matrix X may be arbitrarily ill-conditioned.
- The even (palindromic) Kronecker cannot be computed well with finite precision algorithms.
- The information given in the even (palindromic) Kronecker form is essential for the understanding of the computational problems.
- We need alternatives, from which we can derive the information, that allows the deflation of singular blocks and blocks associated with $0, \infty$ (1 and -1).

We need structured staircase forms!

Why not just use standard equivalence $\lambda QNU + QMU$?

Example Consider a 3×3 even pencil with matrices

$$N = Q \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T, \qquad M = Q \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} Q^T,$$

where Q is a random real orthogonal matrix. The pencil is congruent to

$$\lambda \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It has a triple eigenvalue ∞ with geom. multiplicity 1 and algebr. multiplicity 3.

For different randomly generated orthogonal matrices Q the QZ algorithm in MATLAB produced all variations of eigenvalues that are possible in a general 3×3 pencil.

A structured staircase form for even pencils, Byers/M./Xu 2005

For $\lambda N + M$ with $N = -N^T, M = M^T \in \mathbb{R}^{n,n}$, there exists a real orthogonal matrix $U \in \mathbb{R}^{n,n}$, such that

$$U^T N U =$$

$$\left[\begin{array}{cccc|c|cccc}
N_{11} & \cdots & \cdots & N_{1,m} & N_{1,m+1} & N_{1,m+2} & \cdots & N_{1,2m} & 0 \\
\vdots & \ddots & & \vdots & \vdots & \vdots & ⋰ & ⋰ & \\
\vdots & & \ddots & \vdots & \vdots & N_{m-1,m+2} & ⋰ & & \\
-N_{1,m}^T & \cdots & \cdots & N_{m,m} & N_{m,m+1} & 0 & & & \\
\hline
-N_{1,m+1}^T & \cdots & \cdots & -N_{m,m+1}^T & N_{m+1,m+1} & & & & \\
\hline
-N_{1,m+2}^T & \cdots & -N_{m-1,m+2}^T & 0 & & & & & \\
\vdots & ⋰ & ⋰ & & & & & & \\
-N_{1,2m}^T & ⋰ & & & & & & & \\
0 & & & & & & & &
\end{array}\right]
\begin{array}{c}
n_1 \\ \vdots \\ \vdots \\ n_m \\ \hline l \\ \hline q_m \\ \vdots \\ q_2 \\ q_1
\end{array}$$

$$U^T M U =$$

$$\left[\begin{array}{cccc|c|cccc}
M_{11} & \cdots & \cdots & M_{1,m} & M_{1,m+1} & M_{1,m+2} & \cdots & \cdots & M_{1,2m+1} \\
\vdots & \ddots & & \vdots & \vdots & \vdots & & \iddots & \\
\vdots & & \ddots & \vdots & \vdots & \vdots & \iddots & & \\
M_{1,m}^T & \cdots & \cdots & M_{m,m} & M_{m,m+1} & M_{m,m+2} & & & \\
\hline
M_{1,m+1}^T & \cdots & \cdots & M_{m,m+1}^T & M_{m+1,m+1} & & & & \\
\hline
M_{1,m+2}^T & \cdots & \cdots & M_{m,m+2}^T & & & & & \\
\vdots & & \iddots & & & & & & \\
\vdots & \iddots & & & & & & & \\
M_{1,2m+1}^T & & & & & & & &
\end{array}\right]
\begin{array}{c}
n_1 \\ \vdots \\ \vdots \\ n_m \\ \hline l \\ \hline q_m \\ \vdots \\ \vdots \\ q_1
\end{array},$$

where $q_1 \geq n_1 \geq q_2 \geq n_2 \geq \ldots \geq q_m \geq n_m$,

$$\begin{aligned}
N_{j,2m+1-j} &\in \mathbb{R}^{n_j,q_{j+1}}, \qquad 1 \leq j \leq m-1, \\
N_{m+1,m+1} &= \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta = -\Delta^T \in \mathbb{R}^{2p,2p}, \\
M_{j,2m+2-j} &= \begin{bmatrix} \Gamma_j & 0 \end{bmatrix} \in \mathbb{R}^{n_j,q_j}, \quad \Gamma_j \in \mathbb{R}^{n_j,n_j}, \qquad 1 \leq j \leq m, \\
M_{m+1,m+1} &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}, \quad \Sigma_{11} = \Sigma_{11}^T \in \mathbb{R}^{2p,2p}, \quad \Sigma_{22} = \Sigma_{22}^T \in \mathbb{R}^{l-2p,l-2p},
\end{aligned}$$

and the blocks Σ_{22} and Δ and Γ_j, $j = 1, \ldots, m$ are nonsingular.

The middle block

The middle block

$$\lambda N_{m+1,m+1} + M_{m+1,m+1} = \lambda \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix},$$

contains all the blocks associated with finite eigenvalues and 1×1 blocks associated with the eigenvalue ∞.

With this, all the dynamics and constraints can be identified.

The finite spectrum in the middle block

The finite spectrum of is obtained from the even pencil

$$\lambda\Delta + \Sigma = \lambda\Delta + (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^T)$$

with Δ invertible.

The matrix Δ has a skew-Cholesky factorization $\Delta = LJL^T$, with

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

Bunch 78, Benner/et al 2000.

Thus, the spectral information can be obtained from the Hamiltonian matrix

$$\mathcal{H} = JL^{-1}\Sigma L^{-T}.$$

What can we do with the staircase form

- All the information about the invariants (Kronecker indices) can be read off. Formulas are given in Byers/M./Xu 2005.
- Singularities and high order blocks to the eigenvalue ∞ can be deflated off.
- The best treatment of infinite eigenvalue in the middle block $\lambda N_{m+1,m+1} + M_{m+1,m+1}$ is unclear.
  Is the use of skew-Cholesky better than projecting out the nullspace with unitary (symplectic) transformations?
- The Hamiltonian part associated with the finite eigenvalues in the middle block can be treated with the structured methods for Hamiltonian problems Benner/M./Xu 1998, Byers/Benner/M./Xu 2002, Chu/Liu/M. 2004.
  See Benner/Kressner's HAPACK.

Computational procedure

- The procedure consists of a recursive sequence of singular value decompositions.
- The staircase form essentially determines a least generic even pencil within the rounding error cloud surrounding $\lambda N + M$.

Computational difficulties

- Rank decisions face the usual difficulties and have to be adapted to the recursive procedure.
- Similar difficulties as in standard staircase form, GUPTRI Demmel/Kågström 1993.
- What to do in case of doubt? In applications, assume worst case, see Mattheij/Wijckmans 1998.
- Perturbation analysis is essentially open for singular and higher order blocks associated with ∞.
- Perturbation theory for Hamiltonian matrices, Ran/Rodman 1988, Godunov/Sadkane 1997.
- Extension to regular even pencils Bora/M. 2005.

Example revisited

Our MATLAB implementation of the structured staircase Algorithm determined that in the cloud of rounding-error-small perturbations of each even $\lambda N + M$, there is an even pencil with structured staircase form

$$\lambda \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

with one block I_3 with sign-characteristic 1.

The algorithm successfully located a least generic even pencil within the cloud.

Conclusions

- Palindromic/even matrix pencils appear in many applications.
- Palindromic/even matrix polynomials can be linearized to palindromic/even matrix pencils, Mackey^2/Mehl/M. 2006.
- Singular and high index blocks should be deflated in a structured way.
- There exist structure preserving staircase forms containing all the information.
- This can be used to deflate all the infinite and singular blocks.
- After deflation, the structured methods for Hamiltonian eigenvalue problems can be applied.

Open problems

- Structured perturbation theory mostly open.
- Rank decisions in process.
- Practical implemention of structured staircase algorithms. Schröder Phd
- Extension of deflation procedures to structured matrix polynomials without linearization. Schröder Phd
- Structured tuples of matrix functions. Wunderlich PhD
  Nonstructured case M./Shi 2006.

Recent References

- R. Byers, V. Mehrmann and H. Xu. *A structured staircase algorithm for skew-symmetric/symmetric pencils.* MATHEON PREPRINT *url: http://www.matheon.de/* 2005
- R. Horn and V. Sergeichuk. *Canonical forms for complex congruence and* * *congruence*. To appear in Lin. Alg. Appl. 2006.
- D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. *Palindromic Polynomial Eigenvalue Problems: Good Vibrations from Good Linearizations.* To appear in SIMAX 2006.
- V. Mehrmann und C. Shi. *Transformation of high order linear differential-algebraic systems to first order*. Numerical Algorithms 2006 to appear.
- C. Schröder. *A canonical form for palindromic pencils and palindromic factorizations*, MATHEON PREPRINT , *url: http://www.matheon.de/* 2006
- L. Wunderlich. *Structure preserving canonical forms for pairs of Hermitian matrices and matrix functions.* MATHEON PREPRINT *url: http://www.matheon.de/* 2006.