A Counterexample for Characterizing Invariant Subspaces of Matrices by Singularity Systems

Kathrin Schreiber
Technische Universität Berlin

schreibe@math.tu-berlin.de
http://math.tu-berlin.de/~schreibe

joint work with Hubert Schwetlick (TU Dresden)

Outline

- The eigenproblem and Newton's method
- A 2–vector iteration with optimal bordering
- Block–Newton for invariant subspaces
- Alternative bordering – optimal choice of the matrix

1. Eigenproblem and Newton's method

The linear eigenvalue problem

Consider $A \in \mathbb{C}^{n \times n}$. Find (x, λ), $0 \neq x \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$ such that

$$F(x, \lambda) = (A - \lambda I)x = 0, \quad F : \mathbb{C}^{n+1} \to \mathbb{C}^n \tag{E}$$

1. Eigenproblem and Newton's method

The linear eigenvalue problem

Consider $A \in \mathbb{C}^{n \times n}$. Find (x, λ), $0 \neq x \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$ such that

$$F(x,\lambda) = (A - \lambda I)x = 0, \quad F : \mathbb{C}^{n+1} \to \mathbb{C}^n \tag{E}$$

Let $\lambda_* \in \mathbb{C}$ be a geometrically simple eigenvalue of A, i. e., there exist eigenvectors $x_*, y_* \in \mathbb{C}^n$, $\|x_*\| = \|y_*\| = 1$ such that

$$\ker(A - \lambda_* I) = \operatorname{span}\{x_*\}, \quad \ker(A - \lambda_* I)^H = \operatorname{span}\{y_*\}$$

1. Eigenproblem and Newton's method

The linear eigenvalue problem

Consider $A \in \mathbb{C}^{n\times n}$. Find (x, λ), $0 \neq x \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$ such that

$$\boxed{F(x,\lambda) = (A - \lambda I)x = 0, \qquad F : \mathbb{C}^{n+1} \to \mathbb{C}^n} \tag{E}$$

Let $\lambda_* \in \mathbb{C}$ be a geometrically simple eigenvalue of A, i. e.,
there exist eigenvectors $x_*, y_* \in \mathbb{C}^n$, $\|x_*\| = \|y_*\| = 1$ such that

$$\boxed{\ker(A - \lambda_* I) = \operatorname{span}\{x_*\}, \qquad \ker(A - \lambda_* I)^H = \operatorname{span}\{y_*\}}$$

$\Rightarrow$

$$\boxed{(A - \lambda_* I)\, x_* = 0, \qquad y_*^H (A - \lambda_* I) = 0}$$

x_* ... right eigenvector, y_* ... left eigenvector

Newton's method on eigenproblems: Unger [50], cf. Zurmühl [53], Peters/Wilkinson [79], Osborne, Rall, Lancaster...

Add scalar linear normalizing equation to (E) $N_w(x) = w^H x - 1 = 0, \ \|w\| = 1, \ w^H x_* \neq 0$
and consider

$$\mathbf{F_w}(x, \lambda) = \begin{bmatrix} F(x, \lambda) \\ N_w(x) \end{bmatrix} = \begin{bmatrix} Ax - x\lambda \\ w^H x - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{(EE)}$$

Newton's method on eigenproblems: Unger [50], cf. Zurmühl [53], Peters/Wilkinson [79], Osborne, Rall, Lancaster...

Add scalar linear normalizing equation to (E) $N_w(x) = w^H x - 1 = 0, \; \|w\| = 1, \; w^H x_* \neq 0$
and consider

$$\mathbf{F_w}(x, \lambda) = \begin{bmatrix} F(x,\lambda) \\ N_w(x) \end{bmatrix} = \begin{bmatrix} Ax - x\lambda \\ w^H x - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{(EE)}$$

$\Rightarrow \; \mathbf{F_w}(x_*/w^H x_*, \lambda_*) = 0$, and

$$\partial\,\mathbf{F_w}(x_*, \lambda_*) = \left[\begin{array}{c|c} A - \lambda_* I & -x_* \\ \hline w^H & 0 \end{array}\right] \text{ nonsingular} \;\Leftrightarrow\; \lambda_* \text{ algebraically simple} \;\Leftrightarrow\; y_*^H x_* \neq 0$$

Newton's method on eigenproblems: Unger [50], cf. Zurmühl [53], Peters/Wilkinson [79], Osborne, Rall, Lancaster...

Add scalar linear normalizing equation to (E) $N_w(x) = w^H x - 1 = 0,\ \|w\| = 1,\ w^H x_* \neq 0$
and consider

$$\mathbf{F_w}(x,\lambda) = \begin{bmatrix} F(x,\lambda) \\ N_w(x) \end{bmatrix} = \begin{bmatrix} Ax - x\lambda \\ w^H x - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{(EE)}$$

$\Rightarrow\ \mathbf{F_w}(x_*/w^H x_*, \lambda_*) = 0$, and

$$\partial\,\mathbf{F_w}(x_*,\lambda_*) = \left[\begin{array}{c|c} A - \lambda_* I & -x_* \\ \hline w^H & 0 \end{array}\right] \quad \text{nonsingular} \quad \begin{array}{c} \Leftrightarrow \\ \Leftrightarrow \end{array} \quad \begin{array}{c} \lambda_* \text{ algebraically simple} \\ y_*^H x_* \neq 0 \end{array}$$

Given: $(u,\theta) \approx (x_*,\lambda_*),\ \|u\| = 1$, look for: $(u_+,\theta_+) = (u+s, \theta+\mu)$, i.e., perform Newton step for (EE) with $w = u \iff$

$$\begin{bmatrix} A - \theta I & u \\ u^H & 0 \end{bmatrix} \begin{bmatrix} s \\ -\mu \end{bmatrix} = - \begin{bmatrix} (A-\theta I)u \\ 0 \end{bmatrix} \Leftrightarrow \begin{array}{rcl} (A-\theta I)u_+ & = & u\,\mu \\ u^H u_+ & = & 1 \end{array}$$

Notice: At the solution (x_*, λ_*) we have

$$\|\partial\,\mathbf{F}(x_*, \lambda_*)^{-1}\| = \left\| \begin{bmatrix} (A - \lambda_* I) & -x_* \\ x_*^H & 0 \end{bmatrix}^{-1} \right\| \geq \frac{1}{|\, y_*^H x_* |}$$

Notice: At the solution (x_*, λ_*) we have

$$\|\partial\,\mathbf{F}(x_*, \lambda_*)^{-1}\| = \left\| \begin{bmatrix} (A - \lambda_* I) & -x_* \\ x_*^H & 0 \end{bmatrix}^{-1} \right\| \geq \frac{1}{|\,y_*^H x_*|}$$

$\Rightarrow$ expect problems if y_*, x_* almost orthogonal!

Notice: At the solution (x_*, λ_*) we have

$$\|\partial\, \mathbf{F}(x_*, \lambda_*)^{-1}\| = \left\| \begin{bmatrix} (A - \lambda_* I) & -x_* \\ x_*^H & 0 \end{bmatrix}^{-1} \right\| \geq \frac{1}{|\, y_*^H x_* |}$$

$\Rightarrow$ expect problems if y_*, x_* almost orthogonal!

- Possible if A highly nonnormal
- Then x_* bad choice for bordering $(A - \lambda_* I \mid -x_*)$, whereas $y_* \in \ker(A - \lambda_* I)^H = \mathrm{im}(A - \lambda_* I)^\perp$ would be optimal

Notice: At the solution (x_*, λ_*) we have

$$\|\partial\,\mathbf{F}(x_*,\lambda_*)^{-1}\| = \left\| \begin{bmatrix} (A-\lambda_* I) & -x_* \\ x_*^H & 0 \end{bmatrix}^{-1} \right\| \geq \frac{1}{|\,y_*^H x_*|}$$

$\Rightarrow$ expect problems if y_*, x_* almost orthogonal!

- Possible if A highly nonnormal
- Then x_* bad choice for bordering $(A - \lambda_* I \mid -x_*)$, whereas $y_* \in \ker(A-\lambda_* I)^H = \mathrm{im}(A-\lambda_* I)^\perp$ would be optimal

$\Rightarrow$ **Look for methods that work with matrices of the following kind**

$$\left[\begin{array}{c|c} A-\theta I & v \\ \hline u^H & 0 \end{array}\right] \quad \text{instead of} \quad \left[\begin{array}{c|c} A-\theta I & u \\ \hline u^H & 0 \end{array}\right]$$

where $v \approx y_*,\ \|v\| = 1$

Idea

Consider $(\hat{x}, \hat{\lambda}) = (0, \lambda_*)$ as simple bifurcation point of $F(x, \lambda) = (A - \lambda I)x = 0$ and apply bifurcation point algorithm Allgower/Schwetlick [ZAMM 1997].

Since the x-part $\hat{x}$ of the bifurcation point is known to be 0, the bifurcation equations reduce to scalar equations $\mu(\lambda) = 0$, where $x = x(\lambda), \mu = \mu(\lambda)$ are implicitly defined by the

singularity system

$$\begin{aligned} (A - \lambda I)x + v\mu &= 0 \\ u^H x &= 1 \end{aligned}$$

i.e.,

$$\underbrace{\begin{bmatrix} A - \lambda I & v \\ u^H & 0 \end{bmatrix}}_{C(\lambda, u, v)} \underbrace{\begin{bmatrix} x \\ \mu \end{bmatrix}}_{\hat{x}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e_{n+1}$$

where $u, v,\ \|u\| = \|v\| = 1$ are the current approximations to x_*, y_*, resp., and we suppose

$$x_*^H u \neq 0, \qquad y_*^H v \neq 0$$

2. Generalized RQI

Schwetlick/Lösche [00]

Given: Approximations u, v, θ , $\|u\| = \|v\| = 1.$ Step $(u, v, \theta) \mapsto (u_+, v_+, \theta_+) :$

2. Generalized RQI

Schwetlick/Lösche [00]

Given: Approximations u, v, θ , $\|u\| = \|v\| = 1.$ Step $(u, v, \theta) \mapsto (u_+, v_+, \theta_+)$:

S1: Set

$$C = C(u, v, \theta) = \begin{bmatrix} (A - \theta I) & v \\ u^H & 0 \end{bmatrix}$$

S2: Solve

$$C \begin{bmatrix} \tilde{u} \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C^H \begin{bmatrix} \tilde{v} \\ \nu \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

S3: Set

$$u_+ = \frac{\tilde{u}}{\|\tilde{u}\|}, \quad v_+ = \frac{\tilde{v}}{\|\tilde{v}\|}$$

S4: Set

$$\theta_+ = \frac{\tilde{v}^H A \tilde{u}}{\tilde{v}^H \tilde{u}} = \frac{v_+^H A u_+}{v_+^H u_+}$$

Now: If $\theta \notin \lambda(A)$ there holds

$$u_+ \sim (A - \theta I)^{-1} v, \quad v_+ \sim (A - \theta I)^{-H} u$$

cf. Parlett's alternating RQI [74]

Now: If $\theta \notin \lambda(A)$ there holds

$$u_+ \sim (A - \theta I)^{-1} v, \quad v_+ \sim (A - \theta I)^{-H} u$$

cf. Parlett's alternating RQI [74]

Convergence: Let λ_* be simple and $u_0^H x_* \neq 0$, $v_0^H y_* \neq 0$. Then there exists $\varepsilon_0 > 0$, such that GRQI is well defined for all θ_0 with $|\theta_0 - \lambda_*| \leq \varepsilon_0$ and there holds

(i) $$\|C(u_k, v_k, \theta_k)^{-1}\| \leq K_0$$

(ii) $$\lim_{k\to\infty} \|C(u_k, v_k, \theta_k)^{-1}\| = \max\{\|(A - \lambda_* I)^\dagger\|, 1\}$$

(iii) $$|\theta_{k+1} - \lambda_*| \leq K_1 |\theta_k - \lambda_*|^2$$

(iv) $$\sin \xi_{k+1} \;\leq\; K_2 \sin \xi_k \sin \eta_k, \quad \sin \eta_{k+1} \;\leq\; K_2 \sin \eta_k \sin \xi_k$$

where $\xi_k = \measuredangle(\text{span}\{u_k\}, \text{span}\{x_*\})$, $\eta_k = \measuredangle(\text{span}\{v_k\}, \text{span}\{y_*\})$, and $M^\dagger$ denotes the Moore–Penrose Pseudo-Inverse of M

3. Block–Newton for invariant subspaces

Aim: computation of multiple or clustered eigenvalues when coefficient matrices singular or bad conditioned in solution (Krylov methods may fail)

$A = A^T \in \mathbb{R}^{n \times n}$: Lösche/Schwetlick/Timmermann [LAA 98]

$A \in \mathbb{R}^{n \times n}$ arbitrary: Beyn/Kleß/Thümmler [01]

3. Block–Newton for invariant subspaces

Aim: computation of multiple or clustered eigenvalues when coefficient matrices singular or bad conditioned in solution (Krylov methods may fail)

$A = A^T \in \mathbb{R}^{n\times n}$: Lösche/Schwetlick/Timmermann [LAA 98]
$A \in \mathbb{R}^{n\times n}$ arbitrary: Beyn/Kleß/Thümmler [01]

Invariant subspace: $\mathcal{X} = \mathrm{im}(X_1),\ X_1 = [x_1, \ldots, x_p] \in \mathbb{C}^{n\times p},\ \mathrm{rank}(X_1) = p$ invariant subspace of dimension p , if

$$x \in \mathcal{X} \ \Rightarrow\ Ax \in \mathcal{X} \quad \Longleftrightarrow \quad AX_1 = X_1 L_1 \quad \text{with } L_1 \in \mathbb{C}^{p\times p}$$

3. Block–Newton for invariant subspaces

Aim: computation of multiple or clustered eigenvalues when coefficient matrices singular or bad conditioned in solution (Krylov methods may fail)

$A = A^T \in \mathbb{R}^{n\times n}$: Lösche/Schwetlick/Timmermann [LAA 98]
$A \in \mathbb{R}^{n\times n}$ arbitrary: Beyn/Kleß/Thümmler [01]

Invariant subspace: $\mathcal{X} = \mathrm{im}(X_1),\ X_1 = [x_1, \ldots, x_p] \in \mathbb{C}^{n\times p},\ \mathrm{rank}(X_1) = p$ invariant subspace of dimension p , if

$$x \in \mathcal{X} \ \Rightarrow\ Ax \in \mathcal{X} \quad \Longleftrightarrow \quad AX_1 = X_1 L_1 \quad \text{with } L_1 \in \mathbb{C}^{p\times p}$$

Assume: $A \in \mathbb{C}^{n\times n}$ can be reduced to block triangular form by a unitary similarity transformation

$$A = [X_1\ Y_2] \begin{bmatrix} L_1 & H \\ 0 & L_2 \end{bmatrix} [X_1\ Y_2]^H, \qquad [\underbrace{X_1}_{p}\ \underbrace{Y_2}_{q}] \text{ unitary,}$$

$L_1 \in \mathbb{C}^{p\times p},\ L_2 \in \mathbb{C}^{q\times q},\ p+q = n,\ \lambda(L_1)\cup\lambda(L_2) = \lambda(A)$ $\quad \Rightarrow\ AX_1 = X_1L_1,\ X_1 \in \mathbb{C}^{n\times p}$

Consider the extended bilinear system

$$\mathbf{F_W}(X,\Lambda) = \begin{bmatrix} F(X,\Lambda) \\ N_W(X) \end{bmatrix} = \begin{bmatrix} AX - X\Lambda \\ W^H X - I_p \end{bmatrix} = \begin{bmatrix} 0_{n\times q} \\ 0_{q\times q} \end{bmatrix}$$

where $W \in \mathbb{C}^{n\times p}$, $W^H W = I_p$, provided that

$$X_1^H W \in \mathbb{C}^{p\times p} \quad \text{nonsingular} \quad \Longleftrightarrow \quad \psi = \measuredangle(\operatorname{im} X_1, \operatorname{im} W) < \pi/2$$

Consider the extended bilinear system

$$\mathbf{F_W}(X,\Lambda) = \begin{bmatrix} F(X,\Lambda) \\ N_W(X) \end{bmatrix} = \begin{bmatrix} AX - X\Lambda \\ W^H X - I_p \end{bmatrix} = \begin{bmatrix} 0_{n\times q} \\ 0_{q\times q} \end{bmatrix}$$

where $W \in \mathbb{C}^{n\times p}$, $W^H W = I_p$, provided that

$$X_1^H W \in \mathbb{C}^{p\times p} \quad \text{nonsingular} \quad \Longleftrightarrow \quad \psi = \measuredangle(\operatorname{im} X_1, \operatorname{im} W) < \pi/2$$

Now: $\mathbf{F_W}(X_*,\Lambda_*) = 0$ where

$$X_* = X_1 S,\ \Lambda_* = S^{-1} L_1 S \quad \text{and } S = (X_1^H W)^{-H}$$

Consider the extended bilinear system

$$\mathbf{F_W}(X,\Lambda) = \begin{bmatrix} F(X,\Lambda) \\ N_W(X) \end{bmatrix} = \begin{bmatrix} AX - X\Lambda \\ W^H X - I_p \end{bmatrix} = \begin{bmatrix} 0_{n\times q} \\ 0_{q\times q} \end{bmatrix}$$

where $W \in \mathbb{C}^{n\times p}$, $W^H W = I_p$, provided that

$$X_1^H W \in \mathbb{C}^{p\times p} \quad \text{nonsingular} \quad \Longleftrightarrow \quad \psi = \measuredangle(\operatorname{im} X_1, \operatorname{im} W) < \pi/2$$

Now: $\mathbf{F_W}(X_*,\Lambda_*) = 0$ where

$$X_* = X_1 S,\ \Lambda_* = S^{-1} L_1 S \quad \text{and } S = (X_1^H W)^{-H}$$

Derivative:

$$\partial \mathbf{F_W}(X,\Lambda)[S,M] = \begin{bmatrix} AS - S\Lambda - XM \\ W^H S \end{bmatrix}$$

$$\partial \mathbf{F_W}(X_*,\Lambda_*) \text{ nonsingular} \quad \Longleftrightarrow \quad X_1 \text{ simple i.e., } \lambda(L_1) \cap \lambda(L_2) = \emptyset$$

4. Alternative bordering

Idea: Analog to idea of GRQI, define $X = X(\Lambda),\ M = M(\Lambda)$ by

$$G(\Lambda)[X, M] = \begin{bmatrix} AX - X\Lambda + VM \\ U^H X \end{bmatrix} = \begin{bmatrix} 0_{n\times p} \\ I_{p\times p} \end{bmatrix}$$

and apply Newton to $M(\Lambda) = 0$ with starting point $\Lambda = \Theta$, i.e.
$\Theta_+ = \Theta - \partial M(\Theta)^{-1}[M(\Theta)]$.

4. Alternative bordering

Idea: Analog to idea of GRQI, define $X = X(\Lambda),\ M = M(\Lambda)$ by

$$G(\Lambda)[X, M] = \begin{bmatrix} AX - X\Lambda + VM \\ U^H X \end{bmatrix} = \begin{bmatrix} 0_{n\times p} \\ I_{p\times p} \end{bmatrix}$$

and apply Newton to $M(\Lambda) = 0$ with starting point $\Lambda = \Theta$, i.e.
$\Theta_+ = \Theta - \partial M(\Theta)^{-1}[M(\Theta)]$.
In vector notation this reads as

$$\begin{bmatrix} \mathrm{vec}(0) \\ \mathrm{vec}(I) \end{bmatrix} = \begin{bmatrix} \mathcal{R}\ \mathrm{vec}(X) + \mathcal{V}\mathrm{vec}(M) \\ \mathcal{U}^H \mathrm{vec}(X) \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & 0 \end{bmatrix} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(M) \end{bmatrix} =: \mathcal{G}(\Lambda) \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(M) \end{bmatrix}$$

where

$$\mathcal{R} = \mathcal{R}(\Lambda) = I_p \otimes A - \Lambda^T \otimes I_n, \quad \mathcal{U} = I_p \otimes U, \quad \mathcal{V} = I_p \otimes V$$

and one has

$$G(\Lambda_*) \text{ nonsingular} \iff \mathcal{G}(\Lambda_*) \text{ nonsingular}$$

In order to apply Newton's method we need to show, that $\mathcal{G}$ is nonsingular at the solution point Λ_*.

Unfortunately: $\mathcal{G}$ can be singular as the following example shows

$$A = \left[\begin{array}{cc|ccc} \frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} L_1 & H \\ \hline 0 & L_2 \end{array}\right] = (X_1|Y_2)^H A (X_1|Y_2)$$

where $(X_1|\, Y_2) = I_5 = [e_1\, e_2 \mid e_3\, e_4\, e_5]$, i.e. $AX_1 = X_1 L_1$

Unfortunately: $\mathcal{G}$ can be singular as the following example shows

$$A = \left[\begin{array}{cc|ccc} \frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} L_1 & H \\ \hline 0 & L_2 \end{array}\right] = (X_1|Y_2)^H A (X_1|Y_2)$$

where $(X_1|\,Y_2) = I_5 = [e_1\, e_2 \mid e_3\, e_4\, e_5]$, i.e. $AX_1 = X_1 L_1$

$$\mathcal{G}(L_1) = \begin{bmatrix} \mathcal{R}(L_1) & \mathcal{Y}_1 \\ \mathcal{X}_1 & 0 \end{bmatrix} =$$

$$\left[\begin{array}{c|c} I \otimes A - L_1^T \otimes I & I \otimes Y_1 \\ \hline I \otimes X_1^T & 0 \end{array}\right] = \left[\begin{array}{cccccccccc|cccc}
 & 1 & -0.5 & & & & & & & & 1 & & & \\
 & -1 & & & & & & & & & & 1 & & \\
 & & -0.5 & 1 & & & & & & & -1 & & & \\
 & & & -0.5 & 1 & & & & & & & & & \\
 & & & & -0.5 & & & & & & & & & \\
-1 & & & & & 1 & 1 & -0.5 & & & & & 1 & \\
 & -1 & & & & & & & & & & & & 1 \\
 & & -1 & & & & & 0.5 & 1 & & & & -1 & \\
 & & & -1 & & & & & 0.5 & 1 & & & & \\
 & & & & -1 & & & & & 0.5 & & & & \\
\hline
1 & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & \\
 & & & & & 1 & & & & & & & & \\
 & & & & & & 1 & & & & & & &
\end{array}\right]$$

The upper/left part of $\mathcal{G}$ has full row rank $n \cdot p = 10$, but $\mathcal{G}(L_1)$ has rank drop 1

Lemma 1 *Let* $B = \left[\begin{array}{c|c} D & E \\ \hline F^H & G^H \end{array}\right]$, *with* $D \in \mathbb{C}^{n\times n}, E, F \in \mathbb{C}^{n\times p}, G \in \mathbb{C}^{p\times p}$ *and let* $\begin{bmatrix} X \\ M \end{bmatrix} \in \mathbb{C}^{(n+p)\times p}$ *s.t.* $\ker[\, D \,|\, E \,] = \operatorname{im} \begin{bmatrix} X \\ M \end{bmatrix}$. *Then:*

$$B \ \textit{nonsingular} \quad \Longleftrightarrow \quad \operatorname{rank}\,[\, D \,|\, E \,] = n \quad \textit{and} \quad \begin{bmatrix} F \\ G \end{bmatrix}^H \begin{bmatrix} X \\ M \end{bmatrix} \quad \textit{nonsingular}.$$

Lemma 1 *Let* $B = \left[\begin{array}{c|c} D & E \\ \hline F^H & G^H \end{array}\right]$, *with* $D \in \mathbb{C}^{n\times n}, E, F \in \mathbb{C}^{n\times p}, G \in \mathbb{C}^{p\times p}$ *and let* $\begin{bmatrix} X \\ M \end{bmatrix} \in \mathbb{C}^{(n+p)\times p}$ *s.t.* $\ker[\, D \,|\, E \,] = \operatorname{im} \begin{bmatrix} X \\ M \end{bmatrix}$. *Then:*

$$B \;\; \textit{nonsingular} \quad \Longleftrightarrow \quad \operatorname{rank}\,[\, D \,|\, E \,] = n \quad \textit{and} \quad \begin{bmatrix} F \\ G \end{bmatrix}^H \begin{bmatrix} X \\ M \end{bmatrix} \quad \textit{nonsingular}.$$

Remarks:

- $p > \dim\ker(D)$ is feasible
- The second condition is equivalent to $\measuredangle(\,\operatorname{im}\begin{bmatrix} F \\ G \end{bmatrix}, \operatorname{im}\begin{bmatrix} X \\ M \end{bmatrix}) < \frac{\Pi}{2}$.

4.1. Nullspace of $\mathbf{X} \mapsto \mathbf{AX} - \mathbf{X\Lambda} =: \mathbf{R}(\mathbf{\Lambda})[\mathbf{X}]$

- Again: $\Lambda = S^{-1}L_1S$ where S arbitrary, nonsingular $\Rightarrow$

$$R = AX - XS^{-1}L_1S = (A(XS^{-1}) - (XS^{-1})L_1)S$$

4.1. Nullspace of $\mathbf{X} \mapsto \mathbf{AX} - \mathbf{X\Lambda} =: \mathbf{R}(\mathbf{\Lambda})[\mathbf{X}]$

- Again: $\Lambda = S^{-1}L_1S$ where S arbitrary, nonsingular $\Rightarrow$

$$R = AX - XS^{-1}L_1S = (A(XS^{-1}) - (XS^{-1})L_1)S$$

- Set $\mathbf{Z_1} = \mathbf{Y_1^H X S^{-1}}$, $\mathbf{Z_2} = \mathbf{Y_2^H X S^{-1}}$, then

$$XS^{-1} = (X_1 X_2)(X_1 X_2)^{-1} XS^{-1} = X_1 Z_1 + X_2 Z_2$$

4.1. Nullspace of $\mathbf{X} \mapsto \mathbf{AX} - \mathbf{X\Lambda} =: \mathbf{R(\Lambda)[X]}$

- Again: $\Lambda = S^{-1} L_1 S$ where S arbitrary, nonsingular $\Rightarrow$

$$R = AX - XS^{-1}L_1S = (A(XS^{-1}) - (XS^{-1})L_1)S$$

- Set $\mathbf{Z_1} = \mathbf{Y_1^H X S^{-1}}$, $\mathbf{Z_2} = \mathbf{Y_2^H X S^{-1}}$, then

$$XS^{-1} = (X_1 X_2)(X_1 X_2)^{-1} XS^{-1} = X_1 Z_1 + X_2 Z_2$$

- Some algebra yields

$$R = 0 \;\Leftrightarrow\; L_1 Z_1 - Z_1 L_1 = 0,\; T(L_2, L_1)[Z_2] = L_2 Z_2 - Z_2 L_1 = 0$$

$\Rightarrow Z_2 = 0$, since $\lambda(L_1) \cap \lambda(L_2) = \emptyset$ and hence

4.1. Nullspace of $\mathbf{X} \mapsto \mathbf{AX} - \mathbf{X\Lambda} =: \mathbf{R}(\mathbf{\Lambda})[\mathbf{X}]$

- Again: $\Lambda = S^{-1}L_1S$ where S arbitrary, nonsingular $\Rightarrow$

$$R = AX - XS^{-1}L_1S = (A(XS^{-1}) - (XS^{-1})L_1)S$$

- Set $\mathbf{Z_1} = \mathbf{Y_1^H XS^{-1}}$, $\mathbf{Z_2} = \mathbf{Y_2^H XS^{-1}}$, then

$$XS^{-1} = (X_1X_2)(X_1X_2)^{-1}XS^{-1} = X_1Z_1 + X_2Z_2$$

- Some algebra yields

$$R = 0 \;\Leftrightarrow\; L_1Z_1 - Z_1L_1 = 0,\; T(L_2, L_1)[Z_2] = L_2Z_2 - Z_2L_1 = 0$$

$\Rightarrow Z_2 = 0$, since $\lambda(L_1) \cap \lambda(L_2) = \emptyset$ and hence

$$\ker R(\Lambda) = \{X : \; AX - X\Lambda = 0\} = \{XS^{-1} = X_1Z_1,\; L_1Z_1 - Z_1L_1 = 0\}$$

4.1. Nullspace of $\mathbf{X} \mapsto \mathbf{A}\mathbf{X} - \mathbf{X}\boldsymbol{\Lambda} =: \mathbf{R}(\boldsymbol{\Lambda})[\mathbf{X}]$

- Again: $\Lambda = S^{-1}L_1 S$ where S arbitrary, nonsingular $\Rightarrow$

$$R = AX - XS^{-1}L_1S = (A(XS^{-1}) - (XS^{-1})L_1)S$$

- Set $\mathbf{Z_1} = \mathbf{Y_1^H X S^{-1}}$, $\mathbf{Z_2} = \mathbf{Y_2^H X S^{-1}}$, then

$$XS^{-1} = (X_1 X_2)(X_1 X_2)^{-1} XS^{-1} = X_1 Z_1 + X_2 Z_2$$

- Some algebra yields

$$R = 0 \ \Leftrightarrow \ L_1 Z_1 - Z_1 L_1 = 0, \ T(L_2, L_1)[Z_2] = L_2 Z_2 - Z_2 L_1 = 0$$

$\Rightarrow Z_2 = 0$, since $\lambda(L_1) \cap \lambda(L_2) = \emptyset$ and hence

$$\ker R(\Lambda) = \{X : \ AX - X\Lambda = 0\} = \{XS^{-1} = X_1 Z_1, \ L_1 Z_1 - Z_1 L_1 = 0\}$$

Problem: which dimension $\hat{p}$ has $\ker R(\Lambda_*)$?

Clearly: $$p \leq \hat{p} := \dim \ker R(\Lambda_*) \leq p^2$$

exact value depends on structure of Jordan form, see [Gantmacher 1958]

5. Feasible choice of the block borderings

Consider the modified matrix

$$\mathcal{G}(\Lambda) := \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & \mathcal{W} \end{bmatrix}$$

5. Feasible choice of the block borderings

Consider the modified matrix

$$\mathcal{G}(\Lambda) := \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & \mathcal{W} \end{bmatrix}$$

It was: $R(\Lambda)[X] = AX - X\Lambda = 0 \quad \Longleftrightarrow \quad X = X_1 Z_1 S,\ Z_1\ s.t.\ L_1 Z_1 = Z_1 L_1,\ L_1, Z_1 \in \mathbb{C}^{p\times p}$ where $\dim\ker(R(\Lambda)) =: \hat{p},\ p \le \hat{p} \le p^2,$ i.e., there exist linearly independent $\{B_i\}_{i=1}^{\hat{p}}$, such that

$$X^{(i)} = X_1 B_i S \quad \text{ON–System} \quad w.r.t.\ \langle X, Y\rangle = \mathrm{vec}(X)^H \mathrm{vec}(Y) = \mathrm{trace}(X^H Y)$$

and in vector notation

$x^{(i)} = \mathrm{vec}(X^{(i)}) = (I \otimes X_1)\mathrm{vec}(B_i S),$ where $x^{(i)H} x^{(j)} = \delta_{ij}$

5. Feasible choice of the block borderings

Consider the modified matrix

$$\mathcal{G}(\Lambda) := \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & \mathcal{W} \end{bmatrix}$$

It was: $R(\Lambda)[X] = AX - X\Lambda = 0 \iff X = X_1 Z_1 S,\ Z_1\ s.t.\ L_1 Z_1 = Z_1 L_1,\ L_1, Z_1 \in \mathbb{C}^{p \times p}$ where $\dim\ker(R(\Lambda)) =: \hat{p},\ p \leq \hat{p} \leq p^2,$ i.e., there exist linearly independent $\{B_i\}_{i=1}^{\hat{p}}$, such that

$$X^{(i)} = X_1 B_i S \quad \text{ON–System} \quad w.r.t.\ \langle X, Y\rangle = \mathrm{vec}(X)^H \mathrm{vec}(Y) = \mathrm{trace}(X^H Y)$$

and in vector notation

$x^{(i)} = \mathrm{vec}(X^{(i)}) = (I \otimes X_1)\mathrm{vec}(B_i S),$ where $x^{(i)^H} x^{(j)} = \delta_{ij}$

$$\tilde{X} = \left[x^{(1)} \dots x^{(\hat{p})} \mid x^{(\hat{p}+1)} \dots x^{(p^2)} \mid x^{(p^2+1)} \dots x^{(np)}\right] = \left[\tilde{X}_{11} \mid \tilde{X}_{12} \mid \tilde{X}_2\right] \quad \textbf{unitary}$$

$$\ker \mathcal{R}(\Lambda) = \mathrm{im}(\tilde{X}_{11}), \quad \text{i.e.} \quad \mathcal{R}\tilde{X}_{11} = 0$$

Looking at the upper block of the adjoint matrix $\mathcal{G}^H(\Lambda) = \begin{bmatrix} \mathcal{R}^H & \mathcal{U} \\ \mathcal{V}^H & \mathcal{W}^H \end{bmatrix}$ and considering its transposed $\begin{bmatrix} \mathcal{R} \\ \mathcal{U}^H \end{bmatrix}$ we get

$$\mathcal{R}x = 0 \quad \Longleftrightarrow \quad x = \tilde{X}_{11}\,\xi, \qquad \xi \in \mathbb{C}^{\hat{p}}, \quad \tilde{X}_{11}^H \tilde{X}_{11} = I_{\hat{p}}$$

Looking at the upper block of the adjoint matrix $\mathcal{G}^H(\Lambda) = \begin{bmatrix} \mathcal{R}^H & \mathcal{U} \\ \mathcal{V}^H & \mathcal{W}^H \end{bmatrix}$ and considering its transposed $\begin{bmatrix} \mathcal{R} \\ \mathcal{U}^H \end{bmatrix}$ we get

$$\mathcal{R}x = 0 \quad \Longleftrightarrow \quad x = \tilde{X}_{11}\,\xi, \qquad \xi \in \mathbb{C}^{\hat{p}}, \quad \tilde{X}_{11}^H \tilde{X}_{11} = I_{\hat{p}}$$

$$\mathcal{U}^H x = 0 \quad \Longleftrightarrow \quad \underbrace{\mathcal{U}^H \tilde{X}_{11}}_{=:\mathcal{U}_{11}}\,\xi = 0 \;\Rightarrow\; \xi = 0 \quad \Longleftrightarrow \quad \operatorname{rank} \mathcal{U}_{11} = \hat{p}, \quad (\mathcal{U}_{11} \in \mathbb{C}^{p^2 \times \hat{p}})$$

If we choose $U = X_1$ the matrix $\mathcal{U}_{11}$ is orthonormal, moreover

$\mathcal{U}^H \tilde{X} = [\, \underbrace{\mathcal{U}_{11}}_{\hat{p}} \mid \underbrace{\mathcal{U}_{12}}_{p^2-\hat{p}} \mid 0 \,]$, where $[\, \mathcal{U}_{11} \mid \mathcal{U}_{12} \,]$ is unitary.

With Lemma 1 the upper block bordering is feasible in the sense of having full rank. Analogously, the left block is proven to have full rank if $V = Y_1$ where Y_1 is a basis for the left invariant subspace.

Open questions

(i) **Can one choose W in a practicable way (using available information) such that $\mathcal{W} = I \otimes W$ gives a regular**

$$\mathcal{G}(\Lambda) := \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & \mathcal{W} \end{bmatrix}$$

(ii) **What is the best structured choice (in terms of the matrix formulation) for the bordering blocks?**

Open questions

(i) **Can one choose W in a practicable way (using available information) such that $\mathcal{W} = I \otimes W$ gives a regular**

$$\mathcal{G}(\Lambda) := \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & \mathcal{W} \end{bmatrix}$$

(ii) **What is the best structured choice (in terms of the matrix formulation) for the bordering blocks?**

Thank you for your attention