Matrix pencils

Definition: Matrix pencil

$A + xB$, with $A, B \in \mathbb{C}^{m \times n}$, x indeterminate.

A pencil is called regular if $n = m$ and $\det(A + xB)$ does not vanish identically, i.e., if there is $\lambda \in \mathbb{C}$ for which it is square invertible.

An eigenvalue λ is a value for which $\det(A + \lambda B) = 0$.
Eigenvector, Jordan chains...

If $\det(A + xB)$ has degree less than n, the 'missing' eigenvalues are said to be "at infinity".

Example

$$\begin{bmatrix} x+1 & x \\ x & x+1 \end{bmatrix} \quad \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$$

Eigenvalues of singular pencils

Can still be defined via 'unusual rank drop'. For instance:

$$A + xB = \begin{bmatrix} 2 & x & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ x & x & 0 \end{bmatrix}$$

has typical rank 2. More formally, $\operatorname{rank}_{\mathbb{C}(x)}(A + xB) = 2$.
But $A + 2B$ has rank 1.

Generalized Schur factorization

Theorem

For any pair of square $A, B \in \mathbb{C}^{m\times m}$, one can find orthogonal Q, Z such that $QAZ = T_A, QBZ = T_B$ are upper triangular (at the same time).

Regular unless $(T_A)_{ii} = (T_B)_{ii} = 0$ for some i.

Eigenvalues $= \frac{(T_A)_{ii}}{(T_B)_{ii}}$ (incl. ∞).

Canonical form

Equivalence relation $\sim$: for each two square $P \in \mathbb{C}^{m \times m}, Q \in \mathbb{C}^{n \times n}$ square invertible, $A + xB$ and $P(A + xB)Q$ are said to be equivalent.

Equivalent $\implies$ same eigenvalues, singularity. . .

Is there a canonical form for this equivalence relation?

If B is square nonsingular, there is little new in this theory:
$A + xB \sim J - xI$, where J is the Jordan canonical form of $-B^{-1}A$ (or $-AB^{-1}$).
Computing eigenvalues of $A + xB \iff$ computing eigenvalues of $-B^{-1}A$

Theorem (Weierstrass canonical form)

For a regular matrix pencil $A + xB \in \mathbb{C}[x]^{n\times n}$, there are nonsingular $P, Q \in \mathbb{C}^{n\times n}$ such that $P(A + xB)Q$ is the direct sum (`blkdiag`) of blocks of the forms

$$J_\lambda(x) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} - xI, \quad J_\infty(x) = I - x \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

WCF — Proof

Proof (sketch):
Underlying idea: make a projective transform $x \mapsto \frac{1}{x-c}$ such that the new point at ∞ is a certain c s.t. $A + cB$ is nonsingular; then it's just a Jordan form.

- Take c such that $A + cB$ is invertible;
- $A + xB \sim I + (x - c)(A + cB)^{-1}B$;
- $A + xB \sim I + (x - c)\,\mathrm{blkdiag}(J_1, \dots, J_s)$,
- Consider separately each $I + (x - c)J_i = I + (x - c)(\lambda I + N)$.
- If $\lambda = 0$, block $\sim I - xM$, where $M = \mathrm{toeplitztriu}(0, 1, \dots)$.
- If $\lambda \neq 0$, block $\sim M - xI$, where $M = \mathrm{toeplitztriu}(\frac{c\lambda-1}{\lambda}, \frac{1}{\lambda^2}, \dots)$.

Jordan chains

Jordan chains can be constructed:
$AQ = P^{-1}\,\mathrm{blkdiag}(\cdot)$, $BQ = P^{-1}\,\mathrm{blkdiag}(\cdot)$.

Call q_j the columns of Q, p_j those of P^{-1}.

- at λ: $-Aq_1 = p_1 = \lambda Bq_1$, $-Aq_2 = p_2 = Bq_1 + \lambda Bq_2, \dots$.
- at ∞: $-Bq_1 = 0$, $-Bq_2 = Aq_1, \dots$.

Theorem (Kronecker canonical form)

For a ~~regular~~ matrix pencil $A + xB \in \mathbb{C}[x]^{m\times n}$, there are nonsingular $P \in \mathbb{C}^{m\times m}, Q \in \mathbb{C}^{n\times n}$ such that $P(A + xB)Q$ is the direct sum (`blkdiag`) of blocks of the form $J_\lambda(x), J_\infty(x)$, and

$$\begin{bmatrix} 1 & x & & & \\ & 1 & x & & \\ & & \ddots & \ddots & \\ & & & 1 & x \end{bmatrix} \in \mathbb{C}[x]^{k\times(k+1)}, \quad \begin{bmatrix} 1 & & & \\ x & 1 & & \\ & x & \ddots & \\ & & \ddots & 1 \\ & & & x \end{bmatrix} \in \mathbb{C}[x]^{(k+1)\times k},$$

(This includes 1×0 and 0×1 empty blocks).

Examples

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & x \\ 1 & x \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & x \\ 1 & 0 & 0 \\ x & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \ldots$$

Kernel in $\mathbb{C}(x)$

The singular blocks in the KCF are related to kernel in $\mathbb{C}(x)$.

The $(k \times (k+1))$ Kronecker blocks have kernel $\begin{bmatrix}(-1)^k x^k & (-1)^{k-1} x^{k-1} & \dots & -x & 1\end{bmatrix}^T$, the other blocks have full column rank.

This can be used to characterize $\ker_{\mathbb{C}(x)}(A + Bx)$, using the fact that $\ker \operatorname{blkdiag}(C, D) = \operatorname{blkdiag}(\ker C, \ker D)$.

Remark: it can be proved that this procedure gives a minimal basis, i.e., all other polynomial bases of $\ker_{\mathbb{C}(x)}(A + Bx)$ have higher degrees.

Proof (sketch): [Gantmacher book '59]

- Suppose $(A + xB)v(x) = 0$ for some $v \in \mathbb{C}(x)^n$
- We may assume $v = v_0 + v_1 x + \cdots + v_d x^d \in \mathbb{C}[x]^n$, clearing denominators.
- This implies singularity of $(d+1) \times d$ $\begin{bmatrix} A & & & \\ B & A & & \\ & \ddots & \ddots & \\ & & B & A \\ & & & B \end{bmatrix}$.
- Assume d minimal.
- We wish to show that the v_i are linearly independent. Suppose they are not so; then one can choose $\alpha(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_e x^e$ (of minimal degree $e \leq d$) such that $w(x) = \alpha(x)v(x)$ has a zero coefficient w_e. But then $Aw_0 = 0$, $Aw_1 + Bw_0 = 0$, $\ldots$, $Bw_{e-1} = 0$, which contradicts minimality of d.

(cont.)

- Take bases $V = [v_0, v_1, \dots, v_d, *]$, and $W = [Av_1, \dots, Av_d, *]$. Then,
$$W^{-1}(A + Bx)V = \begin{bmatrix} K(x) & L(x) \\ 0 & M(x) \end{bmatrix},$$
where $K(x)$ is a $d \times (d+1)$ Kronecker block.
- Repeat on $M(x)$ until all the kernel is gone, obtaining a block triangular pencil.
- Repeat on $M^T(x)$ until all the left kernel is gone, obtaining a block triangular pencil $W^{-1}(A + Bx)V = S(x) =$
$$\begin{bmatrix} S_{1,1}(x) & S_{1,2}(x) & \dots & S_{1,k}(x) \\ 0 & S_{2,2}(x) & \ddots & \vdots \\ \vdots & \ddots & S_{k-1,k-1}(x) & S_{k-1,k}(x) \\ 0 & \dots & 0 & S_{k,k}(x) \end{bmatrix}.$$
- Then, one can 'decouple' each pair of blocks with $\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} S(x) \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$, and X, Y solve certain paired Sylvester equations…(some work needed; we do not show details).

Application: differential-algebraic equations

Linear systems of ODEs in the form

$$B\dot{x} = Ax + f(t),$$

where A, B are constant matrices, appear in several applications; for instance, finite element methods often return equations with a 'mass matrix' B in front of the differential term.

When B is invertible, all is good. When B is not, several 'pathological' behaviors may arise. (Note that the Cauchy existence/uniqueness theorem does not apply to this form!)

With a change of variables, one may assume (A, B) in KCF. What equations correspond to the various blocks? Examples

(Similar theory for difference equations $Bx_{k+1} = Ax_k + f_k$.)