Note Title 2025-03-06

Def: M diagonalizable $M \in \mathbb{C}^{n \times n}$,

$$M = VDV^{-1} \qquad D = \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\text{eigenvalues}} \qquad V = \underbrace{\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}}_{\text{eigenvectors}}$$

The <u>Invariant subspace</u> associated to $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ is $\operatorname{span}(v_1, v_2, \dots v_k)$

<u>Or</u>: the invariant subspace associated to $S \subseteq \mathbb{C}$ is the one associated to $S \cap \Lambda(M) = \{\lambda_1, \dots \lambda_k\}$

$$\mathcal{V} = \operatorname{span}(v_1, v_2, \dots v_k)$$

⚠ To make this definition well-defined, we have to take <u>all</u> copies of λ if there are multiple eigenvalues, for instance, if $\lambda_1 = \lambda_2$, we need to include both v_1 and v_2, or neither.

v_i are not well-defined as a function of M, only

$$\mathcal{V}_\lambda = \ker(M - \lambda I)$$

More generally, if M is not diagonalizable, we can take

$$M = VJV^{-1} \qquad J = \operatorname{blkdiag}(J_1, J_2, \dots J_s)$$

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

The columns of V form so-called <u>Jordan chains</u>:
if we take all columns associated to blocks with

a given $\lambda \in \mathbb{C}$, their span is the <u>generalized eigenspace</u> associated to λ:

$$\mathcal{V}_\lambda = \{ v : (M - \lambda I)^k v = 0 \text{ , for some } k = 1,2,3,\dots \}$$

$= \text{span}(v_j \text{ corresponding to blocks of eigenvalue } \lambda)$

<u>Def</u>: the <u>invariant subspace</u> associated to $\lambda_1, \dots \lambda_k$ is the sum $\mathcal{V}_{\lambda_1} + \mathcal{V}_{\lambda_2} + \dots + \mathcal{V}_{\lambda_k}$

(sum of generalized eigenspaces)

$= \text{span}\{ \text{all } v_j\text{'s corresponding to blocks with eigenvalues in } \{\lambda_1, \dots \lambda_k\} \}$.

<u>Ex</u>: the <u>stable invariant subspace</u> of a matrix M is defined as

$$\mathcal{V}_s = \{ v \in \mathbb{C}^n : \lim_{k \to \infty} M^k v = 0 \}$$

For instance, in Jacobi's method, the error $e_k x_k - x$ satisfies $e_{k+1} = H e_k$; $e_k \to 0$ if and only if $e_0 \in \mathcal{V}$

We will prove that $\mathcal{V}_s$ is the invariant subspace associated to $\mathbb{D} = \{ z \in \mathbb{C} : |z| < 1 \}$.

Let us reblock the Jordan decomposition as

$$M = \left[V_1 \middle| V_2 \right] \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \left[V_1 \middle| V_2 \right]^{-1}$$

$J_1 \in \mathbb{C}^{n_1 \times n_1}$ contains all blocks with $|\lambda| < 1$

$J_2 \in \mathbb{C}^{n_2 \times n_2}$ contains all those with eigenvalues $|\lambda| \geq 1$

A vector w belongs to the invariant subspace

$\mathcal{V}$ = span (columns of V_1)

if and only if $\quad w = V_1 \cdot v_1 = \begin{bmatrix} V_1 \mid V_2 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$

$$M^k w = \begin{bmatrix} V_1 \mid V_2 \end{bmatrix} \begin{bmatrix} J_1^k & 0 \\ 0 & J_2^k \end{bmatrix} \cancel{\begin{bmatrix} V_1 \mid V_2 \end{bmatrix}^{-1} \begin{bmatrix} V_1 \mid V_2 \end{bmatrix}} \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} V_1 \mid V_2 \end{bmatrix} \begin{bmatrix} J_1^k v_1 \\ 0 \end{bmatrix} \to 0$$

Since J_1 contains blocks with eigenvalues in D, $J_1^k \to 0$

We also need to prove that if $w \notin \mathcal{V}$, then $\lim M^k w \neq 0$.

$$w = \begin{bmatrix} V_1 \mid V_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{with } v_2 \neq 0$$

$$M^k w = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} J_1^k v_1 \\ J_2^k v_2 \end{bmatrix}$$

$J_1^k v_1 \to 0$, and we would like to prove that $J_2^k v_2$ does not converge to 0.

Let m be the ___last___ nonzero entry of v_2

$$J_2^k v_2 = \begin{array}{c} \\ \\ \\ m \\ \\ \\ \end{array} \begin{bmatrix} \lambda_1^k & & & & & \\ & \lambda_1^k & & & & \\ & & \ddots \; \lambda_i^k & & & \\ & 0 \cdots 0 & & \ddots & & \\ & 0 & & & \lambda_s^k & \\ & & & & & \ddots \; \lambda_s^k \end{bmatrix} \begin{bmatrix} \vdots \\ \\ \\ (v_2)_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \lambda_i^k (v_2)_m \\ \\ \\ \\ \end{bmatrix} \leftarrow m\text{-th entry}$$

The m-th entry is $(\lambda_i)^k (v_2)_m$, where $(v_2)_m \neq 0$, and $|\lambda_i| \geq 1$, so it does ___not___ converge to 0.

This proves that $\mathcal{V}_s$ is the invariant subspace of M

associated to the unit disc. □

Teo: suppose we have a decomposition

$$M = \begin{bmatrix} \underset{n_1}{U_1} & \underset{n_2}{U_2} \end{bmatrix} \begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \end{array} \begin{array}{c} n_1 \\ n_2 \end{array} \begin{bmatrix} \underset{n_1}{U_1} & \underset{n_2}{U_2} \end{bmatrix}^{-1}.$$

with $\Lambda(A) \cap \Lambda(B) \neq 0$.

Then, $\operatorname{Im} U_1$ is the invariant subspaces associated to $\Lambda(A)$.

Proof: we know that there is X such that

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1}$$

(it's the solution of a certain Sylvester equation)

Now take Jordan forms

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} V_A & 0 \\ 0 & V_B \end{bmatrix} \begin{bmatrix} J_A & 0 \\ 0 & J_B \end{bmatrix} \begin{bmatrix} V_A^{-1} & 0 \\ 0 & V_B^{-1} \end{bmatrix}.$$

$$M = \underbrace{\begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} V_A & 0 \\ 0 & V_B \end{bmatrix}}_{V} \begin{bmatrix} J_A & 0 \\ 0 & J_B \end{bmatrix} \underbrace{\begin{bmatrix} V_A^{-1} & \\ & V_B^{-1} \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}^{-1}}_{V^{-1}}$$

is a Jordan form of M.

The leftmost columns of V span the invariant subspace associated to $\Lambda(A)$

$$V = \begin{bmatrix} U_1 V_A & \boxed{(U_1 X + U_2) V_B} \end{bmatrix}.$$

$\mathcal{V} = \mathrm{Im}\,(U_1 V_A) = \mathrm{Im}\,(U_1)$

Note that if we take $w \in \mathcal{V}$, then $w = [U_1 | U_2]\begin{bmatrix} v_1 \\ 0 \end{bmatrix}$

$$Mw = [U_1 | U_2]\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\cancel{[U_1\ U_2]^{-1}\,[U_1 | U_2]}\begin{bmatrix} v_1 \\ 0 \end{bmatrix}$$

$$= [U_1\ \ U_2]\begin{bmatrix} Av_1 \\ 0 \end{bmatrix} = U_1(Av_1)$$

(*) This shows $w \in \mathcal{V} \Rightarrow Mw \in \mathcal{V}$, and that

$$f: w \to Mw \qquad f: \mathcal{V} \to \mathcal{V}$$

has associated matrix A_1 in the basis given by the columns of U_1

<u>Remark</u>: • $M\mathcal{V} \subseteq \mathcal{V}$ can be a strict inclusion if $\Lambda(M) \ni 0$

• $\mathrm{Im}\, U_2$ is <u>not</u> invariant

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \cdot \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

but $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \notin \mathrm{span}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$

• these are <u>not</u> the only M-invariant subspaces, when M has multiple eigenvalues! E.g.,

$$M = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

has only one generalized eigenspace $\mathcal{V}_\lambda$, so the dim invariant subspace associated to a part of the spectrum is $V_\lambda = \mathbb{R}^3$

However, $\operatorname{span}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right)$ and $\operatorname{span}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right)$ are also M-invariant, because

$$M = I \cdot \left[\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ \hline 0 & 0 & \lambda \end{array}\right] \cdot I^{-1}$$

and one can repeat the argument. (*)

How to compute invariant subspaces?

- compute a Schur canonical form

$$M = Q \begin{array}{c} \\ n_1 \\ n_2 \end{array}\begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} \end{array} Q^*$$

such that $\Lambda(U_{11})$ are the eigenvalues associated to the subspace

- then, the first n_1 columns of Q span the invariant subspace.

Problem: the QR/Francis algorithm does not allow us to specify the order of the eigenvalues on the diagonal!

Solution: algorithm to reorder the Schur form:

given $M = QUQ^*$, we want to compute another Schur form $M = \hat{Q}\hat{U}\hat{Q}^*$ with the eigenvalues in a different order on the diagonal.

The main building block is being able to swap two diagonal blocks:

$$M = Q \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} Q^* \to M = \hat{Q} \begin{bmatrix} \hat{B} & \hat{C} \\ 0 & \hat{A} \end{bmatrix} \hat{Q}^*$$

with $\Lambda(B)=\Lambda(\hat{B})$, $\Lambda(A)=\Lambda(\hat{A})$

Combining many of these transformations gives arbitrary orders.

$$\begin{bmatrix} 1 & & \\ & \square & \\ & & 1 \end{bmatrix}\begin{bmatrix} U_{11} & \times & \times & \times \\ 0 & U_{22} & \times & \times \\ 0 & 0 & U_{33} & \times \\ 0 & 0 & 0 & U_{44} \end{bmatrix}\begin{bmatrix} 1 & & \\ & \square^* & \\ & & 1 \end{bmatrix} = \begin{bmatrix} U_{11} & \times & \times & \times \\ 0 & \hat{U}_{33} & \times & \times \\ 0 & 0 & \hat{U}_{22} & \times \\ 0 & 0 & 0 & U_{44} \end{bmatrix}$$

If X solves $AX-XB=-C$, then

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

Multiply left and right by $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$:

$$\begin{bmatrix} 0 & I \\ I & -X \end{bmatrix}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\begin{bmatrix} X & I \\ I & 0 \end{bmatrix} = \begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}$$

one the inverse of the other, but not orthogonal.

This does not work, because $\begin{bmatrix} X & I \\ I & 0 \end{bmatrix}$ is not orthogonal.

However, it works if we take $QR = \begin{bmatrix} X & I \\ I & 0 \end{bmatrix}$:

$$R^{-1}Q^*\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}QR = \begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}$$

$$Q^*\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}Q = R\begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}R^{-1}$$

all are upper triangular!

$$= \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}\begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}\begin{bmatrix} R_{11}^{-1} & * \\ 0 & R_{22}^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} R_{11}BR_{11}^{-1} & * \\ 0 & R_{22}AR_{22}^{-1} \end{bmatrix}$$ → upper triangular

$$\Lambda(R_{11} B R_{11}^{-1}) = \Lambda(B), \qquad \Lambda(R_{22} A R_{22}^{-1}) = \Lambda(A).$$

<u>Remark</u>: we only need the "back-substitution" part of the Bartels-Stewart algorithm, since A, B are already upper triangular.

[Matlab examples in the video or notes]