Sylvester equations:

$$AX - XB = C \qquad A \in \mathbb{C}^{m\times m} \quad B \in \mathbb{C}^{n\times n}$$

$$X, C \in \mathbb{C}^{m\times n}$$

$$\Updownarrow$$

$$M = 1 \otimes A - B^T \otimes 1 \qquad M \cdot \mathrm{vec}(X) = \mathrm{vec}(C)$$

1) Equation uniquely solvable $\Leftrightarrow$ $\Lambda(A) \cap \Lambda(B) \neq \emptyset$

2) $O(m^3+n^3)$ algorithm (Bartels-Stewart)

$$\begin{pmatrix} 1 & -X \\ 0 & 1 \end{pmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} \overset{!}{=} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

Transformation block triangular $\rightarrow$ block diagonal

$\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix}^{-1}$ so it is a similitude

$$\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & AX-XB+C \\ 0 & B \end{bmatrix}$$

if $X$ solves $AX-XB=-C$, the $(1,2)$ block becomes $0$

__Ex__: Jordan forms: $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ is not similar to $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

Indeed, if you try to apply that method, you have the Sylvester equation $\lambda \cdot x - x \cdot \lambda = -1$, which is not solvable

On the other hand, $\begin{bmatrix} \lambda & 1 \\ 0 & \mu \end{bmatrix}$ __is__ similar to $\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$

if $\lambda \neq \mu$. Indeed, $\lambda \cdot x - x \cdot \mu = -1$ is solvable.

---

$\mathcal{U} \subseteq \mathbb{C}^n$ invariant subspace for $M \in \mathbb{C}^{n \times n}$

if $M \cdot \mathcal{U} \subseteq \mathcal{U}$ $\quad$ $Mv \in \mathcal{U}$ for all $v \in \mathcal{U}$

$U_1 \in \mathbb{C}^{n \times m}$ basis matrix for $\mathcal{U}$ $(m = \dim \mathcal{U})$

Each column of $M \cdot U_1$ is inside $\mathcal{U}$, so it can be written as $U_1 \cdot v$ for a certain $v \in \mathbb{C}^m$

$\Rightarrow$ $M \cdot U_1 = U_1 \cdot A$ for a $A \in \mathbb{C}^{m \times m}$

$$\boxed{M}\,\boxed{U_1} = \boxed{U_1}\,\boxed{A}.$$

Conversely, if $MU_1 = U_1 A$ for a certain $A \in \mathbb{C}^{m \times m}$, $U_1$ with full col. rank, then $\operatorname{Im} U_1$ is an inv. subspace of $M$.

The matrix $A$ is the matrix of the linear map $u \mapsto Mu$ restricted to the subspace $\mathcal{U}$ (in the basis $U_1$)

If we complete $U_1$ to $U = [U_1 \; U_2]$ basis of $\mathbb{C}^n$ (i.e. so that $U$ is invertible), then the relation

$$MU = M[U_1 \; U_2] = [U_1 \; U_2] \left[\begin{array}{c:c} A & C \\ 0 & B \end{array}\right]$$

holds (for certain $A, B, C$ of appropriate size)

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

$U_1$ span an invariant subspace for $M$

$\Leftrightarrow$ completing it to a basis $U = [U_1 \ U_2]$, we have

$$U^{-1} M U = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}.$$

Examples..

→ Trivially, $\mathbb{C}^n$ and $\{0\}$ are invariant subspaces for any $M$

→ if $v$ is an eigenvector of $M$, $\text{span}(v)$ is invariant for $M$

$$M \cdot (\alpha v) = \alpha \lambda v \in \text{span}(v)$$

→ if $v_1, v_2, \dots v_m$ are eigenvectors of $M$, $\text{span}(v_1, v_2, \dots v_m)$ is invariant for $M$:

$$M(\alpha_1 v_1 + \dots + \alpha_m v_m) = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_m \lambda_m v_m$$
$$\in \text{span}(v_1, \dots v_m)$$

→ Given $M = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ & & & & \lambda \end{bmatrix}$ (Jordan block),

$\text{span}(e_1, \dots e_m)$ is invariant for $M$ for all $m \in \{1, 2, \dots n\}$

→ More generally, if you take $v_{1,1}, \dots v_{1,m_1}, v_{2,1}, \dots v_{2,m_2}, \dots$

$\dots v_{s,1}, \dots v_{s,m_s}$

a set of "beginnings of Jordan chains" of $M$, their span is an invariant subspace.

→ Given $M$, the set $\{v : \lim_{k \to \infty} M^k v = 0\} = S$ ("stable invariant subspace") is an invariant subspace of $M$.

Indeed, if $v \in S$, $\lim_{k \to \infty} M^k v = 0$, then $\lim_{k \to \infty} M^k (Mv) = 0$

hence $Mv \in S$ satisfies the same property.

Actually $S$ is the span of all Jordan chains of $M$ associated to eigenvalues $|\lambda| < 1$:

$M = V \cdot J V^{-1}$, $\quad J = \left[\begin{array}{cccc|cccc} J_1 & & & & & & & \\ & J_2 & & & & & 0 & \\ & & \ddots & & & & & \\ & & & J_r & & & & \\ \hline & & & & J_{r+1} & & & \\ & & 0 & & & \ddots & & \\ & & & & & & & J_s \end{array}\right]$ $\quad J_1 \cdots J_r$ have eigenvalues with abs. value $< 1$

We need to show that $\lim J^k v = 0$

if and only if $v = \begin{bmatrix} v_1 \\ \vdots \\ v_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Indeed, if $|\lambda_i| < 1$, then $\lim J_i^k \to 0$

and if $|\lambda_i| \geq 1$, then $\lim J_i^k v$ is <u>not</u> zero

for every vector $v$:

$$J_i^k v = \begin{bmatrix} \lambda_i^k & & & \\ 0 & \lambda_i^k & & \\ 0 & 0 & \lambda_i^k & \\ \vdots & & \ddots & \\ 0 & 0 & \cdots\ 0 & \lambda_i^k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_s \end{bmatrix} \begin{matrix} v_t \\ 0 \\ 8 \end{matrix}$$

If $v_t$ is the last non-zero entry of $v$, then the $t$-th entry of $J_i^k v$ is $\lambda_i^k \cdot v_t$, and this does not converge to $\emptyset$.

---

Schur decomposition: $M = QTQ^*$

$Q$ unitary

$T$ upper triangular

$t_{ii}$ eigenvalues of $M$

Problem: given a Schur form $(Q, T)$ of $M$, return a different one $(\hat{Q}, \hat{T})$ such that the eigenvalues come in a different order.

$$Q^{-1} M Q = \begin{bmatrix} \diagdown & \\ 0 & \end{bmatrix}$$

The first block $Q_1$ of $Q = [Q_1 \; Q_2]$ spans an invariant subspace of $M$

So if $(\hat{Q}, \hat{T})$ is a Schur form of $M$ in which the eigenvalues in the first block are inside the unit disk (or, more generally, in a certain region).

$$T = \begin{bmatrix} T_{11} & * \\ 0 & T_{22} \end{bmatrix} \qquad \begin{array}{l} T_{11}: |t_{ii}| < 1 \\ T_{22}: |t_{ii}| \geq 1 \end{array}$$

Then $\underbrace{Q_1 \text{ spans the stable invariant subspace}}$ $S$.

How does the procedure (ordschur) work?

Idea: it is enough to know how to swap two blocks of eigenvalues

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \longrightarrow \begin{bmatrix} \hat{B} & \hat{C} \\ 0 & \hat{A} \end{bmatrix} \quad \text{with } \begin{array}{l} \Lambda(A) = \Lambda(\hat{A}) \\ \Lambda(B) = \Lambda(\hat{B}) \end{array}$$

If $X$ solves $AX - XB + C = 0$, then

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

Multiply left and right by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

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

But this is not an orthogonal transformation...

$$\begin{bmatrix} X & 1 \\ 1 & 0 \end{bmatrix} = QR$$

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

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

where $\hat{B} = R_{11} B R_{11}^{-1}$, $\hat{A} = R_{22} B R_{22}^{-1}$ have the same eigenvalues as $B$ and $A$ respectively

In practice, construct

$$Q = qr\left(\begin{bmatrix} X & 1 \\ 1 & 0 \end{bmatrix}\right),$$

and use it to form $Q^* \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} Q$, which will be upper triangular.

---

Setting: a matrix $M$ with inv. subspace $U_1$

perturbed matrix $M + \delta_M$; does it have an invariant subspace $U_1 + \delta U_1$?

We can assume with a change of basis (orthogonal!) that

$$M = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \qquad U_1 = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

__Recall:__

$$\operatorname{sep}(A,B) := \sigma_{\min}\left(I \otimes A - B^T \otimes I\right) = \min \frac{\|AZ - ZB\|_F}{\|Z\|_F}$$

$$M + \delta_M = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} + \begin{bmatrix} \delta_A & \delta_C \\ \delta_D & \delta_B \end{bmatrix}, \quad \|\delta_A\| = a \text{ and so on}$$

__Thm:__ if $a, b, c, d$ small enough, there exists

$$X \text{ with } \|X\|_F \leq 2 \frac{d}{\operatorname{sep}(A,B) - a - b}$$

such that $\begin{bmatrix} I \\ X \end{bmatrix}$ spans an inv. sub. of $M + \delta_M$.

__Proof:__ Is there a matrix $X$ s.t.

$$\begin{bmatrix} 1 & 0 \\ -X & 1 \end{bmatrix}\begin{bmatrix} A+\delta_A & C+\delta_C \\ \delta_D & B+\delta_B \end{bmatrix}\begin{bmatrix} 1 & 0 \\ X & 1 \end{bmatrix} = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} ?$$

If there is one, then $\begin{bmatrix} 1 \\ X \end{bmatrix}$ is an inv. subspace of $M+\delta_M$.

$$\text{LHS} = \begin{bmatrix} A+\delta_A & C+\delta C \\ -X(A+\delta_A)+\delta_D & -X(C+\delta_C)+B+\delta_B \end{bmatrix}\begin{bmatrix} 1 & 0 \\ X & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} * & * \\ \boxed{-X(A+\delta_A)+\delta_D-X(C+\delta_C)X+(B+\delta_B)X} = 0 & * \end{bmatrix}$$

(Riccati matrix equation)

$$\underbrace{X(A+\delta_A)-(B+\delta_B)X} = \delta_D - X(C+\delta_C)X$$

$$\hat{T} \cdot \operatorname{vec} X = \operatorname{vec}\left(\delta_D - X\left(C + \delta_C\right)X\right)$$

$$\hat{T} = -1 \otimes \left(B + \delta_B\right) + \left(A + \delta_A\right)^T \otimes 1$$

$$\operatorname{vec} X = \hat{T}^{-1}\left(\delta_D - X\left(C + \delta_C\right)X\right)$$

$$x = \operatorname{vec} X$$

Has the form $x = \varphi(x)$

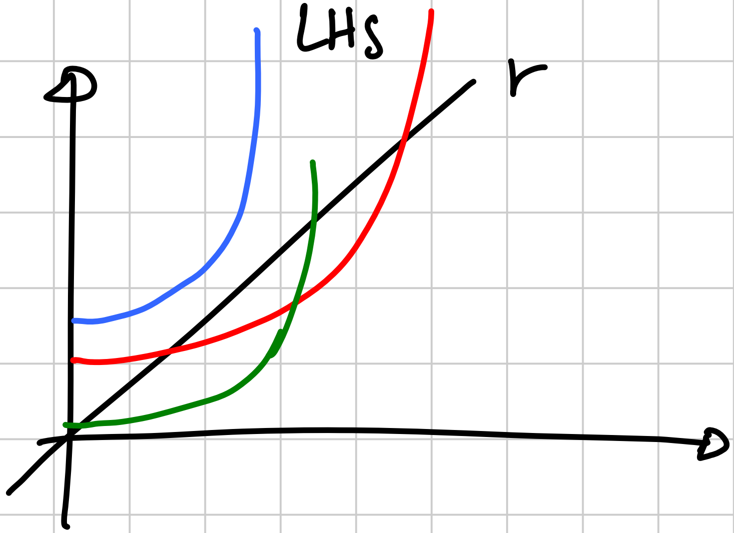
We would like to show that $\varphi\left(B(0,r)\right) \subseteq B(0,r)$ for a ball of sufficiently small radius $r$ centered in $0$

If $\|x\| \leq r$, then

$$\|\varphi(x)\| \leq \|\hat{T}^{-1}\| \left( \|\mathrm{vec}\, \delta_D\| + \|x\| \cdot \|\mathrm{vec}(C+\delta_C)\| \cdot \|x\| \right)$$

$$\leq \underbrace{\|\hat{T}^{-1}\| \left( d + r^2 \left( \|C\|_F + c \right) \right)}_{\text{LHS}} \overset{!}{=} r \qquad r_{\pm} = \frac{1 \pm \sqrt{\phantom{x}}}{2(\|C\|_F + c\|}$$

We look for an $r$ such that this is an exact equality!




If the two plots meet, (red case), then there is a solution $r$

For sufficiently small perturbations, the constant term of the LHS tends to zero, so the two lines will meet

$$\left[ \|\hat{T}^{-1}\| = \left\| \left( -1 \otimes (B + \delta_B) + (A+\delta_A)^T \otimes 1 \right)^{-1} \right\| \right.$$

$$= \sigma_{min}\left(-1 \otimes (B+\delta_B) + (A+\delta_A)^T \otimes I\right)^{-1}$$

$$\sigma_{min}\left(-1 \otimes (B+\delta_B) + (A+\delta_A)^T \otimes 1\right) \geq$$ (triangle inequality)

$$\underbrace{\sigma_{min}\left(-1 \otimes B + A^T \otimes 1\right)}_{sep(A,B)} \underbrace{-}_{-} \underbrace{\|1 \otimes \delta_B\|}_{b} \underbrace{-}_{-} \underbrace{\|\delta_A^T \otimes 1\|}_{a}$$

$$\|T^{-1}\| \leq \frac{1}{sep(A,B) - a - b}$$

If the discriminant is $\geq 0 \Rightarrow$ lines meet $\Rightarrow$

$\varphi$ sends $B(0,r)$ to itself $\Rightarrow$ By Brouwer's fixed-point theorem, there is $x$ s.t. $x = \varphi(x)$ (and $\|x\| \leq r$)

$\Rightarrow X$ solves the Riccati equation

$\Rightarrow \begin{bmatrix} I \\ X \end{bmatrix}$ is an inv. sub. of $M + \delta_M$

$$\|X\|_F \leq r \leq \frac{2d}{\mathrm{sep}(A,B) - a - b}$$