Lecture 39: Dealing with nondiagonalizable matrices.

(§6.7 and 7.1)

Previously... $A \in M_{n\times n}(\mathbb{F})$ is diagonalizable if and only if a) the char poly of A splits completely over $\mathbb{F}$.

b) (geom mult) = (alg mult) for all eigenvalues of A.

[Two ways to deal with non diagonalizable matrices: Jordan Canonical Form and the Singular Value Decomposition.]

Jordan Canonical Form: Over $\mathbb{C}$, any square matrix is similar to one in the following block diagonal form:

$$\begin{pmatrix} \boxed{B_1} & & & & \\ & \boxed{B_2} & & 0 & \\ & & \boxed{B_3} & & \\ & 0 & & \ddots & \\ & & & & \boxed{B_4} \end{pmatrix}$$

where each block has the form $\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & 1 & \\ 0 & & \ddots & \lambda_i \; 1 \\ & & & \lambda_i \end{pmatrix}$

That is given any $A \in M_{n\times n}(\mathbb{C})$ there exists an invertible C such that $C^{-1}AC$ has the above form.

Ex:

$$\begin{pmatrix} \boxed{\begin{matrix}2&1&0\\0&2&1\\0&0&2\end{matrix}} & & & & O \\ & \boxed{2} & & & \\ & & \boxed{\begin{matrix}3&1\\0&3\end{matrix}} & & \\ O & & & \boxed{5} & \\ & & & & \boxed{\begin{matrix}0&1\\0&0\end{matrix}} \end{pmatrix}$$

$E_2 = \text{span}(\{e_1, e_4\})$

$E_3 = \text{span}(\{e_5\})$

$E_5 = \text{span}(\{e_7\})$

$E_0 = \text{span}(\{e_8\})$

Char poly $= (2-t)^4(3-t)^2(5-t)t^2$

Fun facts about A that is Jordan Can. Form:

1) The λ_i are eigenvalues of A. Moreover, the algebraic mult of $\lambda_i = \sum$ (sizes of λ_i blocks)

[The char poly of A is easy to compute as $A - tI$ is upper triangular.]

2) For each block, the e_i cor to the first column is an eigenvector for A. In fact, $\dim(E_{\lambda_i})$ is the # of blocks assoc to λ_i.

3) Can still understand powers: (See Section 7.2, Problem 19)

$$A^n = \begin{pmatrix} \begin{matrix} 2^n & n2^{n-1} & \frac{n(n-1)}{2}2^{n-2} \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{matrix} & & & & O \\ & \boxed{2^n} & & & \\ & & \begin{matrix} 3^n & n3^{n-1} \\ 0 & 3^n \end{matrix} & & \\ O & & & \boxed{5^n} & \\ & & & & \boxed{\begin{matrix}0&0\\0&0\end{matrix}} \end{pmatrix} \quad \text{for } n > 1$$

<u>How to find</u>: Generalized eigenvectors: v where $(A-\lambda I)^p v = 0$ for some $p \geq 1$.

<u>Applications</u>:

1) Proof that any regular Markov Chain converges.

   <u>Point</u>: Blocks B with $|\lambda| < 1$ have $\lim_{n\to\infty} B^n = 0$.

2) Solving systems of linear differential equations.

<u>Limitations</u>: In general, must work over $\mathbb{C}$ even if the original matrix has entries in $\mathbb{R}$.

Singular Value Decomposition: Work over $\mathbb{R}$ for concreteness, variant works over $\mathbb{C}$.

Story so far: $T : V \xrightarrow[\text{linear}]{} V \longleftrightarrow A = [T]_\beta \in M_{n \times n}(\mathbb{R})$

Goal: Choose β so that $[T]_\beta$ is as simple as possible $\longleftrightarrow B = Q^{-1}AQ$

$\nwarrow$ diagonal, J.C.F.

Question: What about $T : V \longrightarrow W$? Can't make sense of eigenvectors here! Note we need to choose bases β for V and γ for W to write down $[T]_\beta^\gamma$. [This gives extra flexibility even if $V = W$ as $[T]_\beta$ is really $[T]_\beta^\beta$.]

Thm [Probably skip!] Suppose V and W are finite dim'l and $T : V \to W$ is linear. Then there are bases such that

$$[T]_\beta^\gamma = \left(\begin{array}{c|c} \begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix} & \overbrace{0}^{\text{nullity}(T)} \\ \hline \underbrace{0}_{\text{rank}(T)} & 0 \end{array}\right)$$

Proof: Using the Rep'l Thm, pick a basis $\beta = \{v_1, \dots, v_n\}$ whose last k vectors are a basis for $\mathcal{N}(T)$. Again via replacement, extend $\{T(v_1), \dots, T(v_{n-k})\}$ (which are lin. indep) to a basis γ for W. Then $[T]_\beta^\gamma$ has the desired form. ▨

Singular Value Decomposition: Suppose $T : V \to W$ a linear trans. between finite dim'l inner product spaces. Then there exist orthonormal bases β and γ such that

$$[T]_\beta^\gamma = \left(\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ 0 & & \ddots & & \\ & & & \sigma_r & \\ \hline & 0 & & & 0 \end{array}\right)$$

where $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$ are in $\mathbb{R}$.

Fun Facts:

1) $r = \text{rank}(T)$

2) The σ_i are unique

3) Matrix form: Given $A \in M_{m \times n}(\mathbb{R})$ there exist orthogonal $P \in M_{m \times m}(\mathbb{R})$ and $Q \in M_{n \times n}(\mathbb{R})$ such that PAQ has the above form. (Like diagonalizing a symmetric matrix, except we do not insist that $P = Q^t = Q^{-1}$.)

4) Proof sketch when $W = V$ and T is invertible (<u>Skip!</u>): $T^* \circ T$ is self-adjoint and so can be diagonalized by an orthonormal $\beta = \{v_1, \dots, v_n\}$ where λ_i is the eigenvalue of $T^* \circ T$. Note that

$$\langle T(v_i), T(v_j) \rangle = \langle v_i, T^*(T(v_j)) \rangle$$

$$= \langle v_i, \lambda_j v_j \rangle = \lambda_i \langle v_i, v_j \rangle \qquad ✪$$

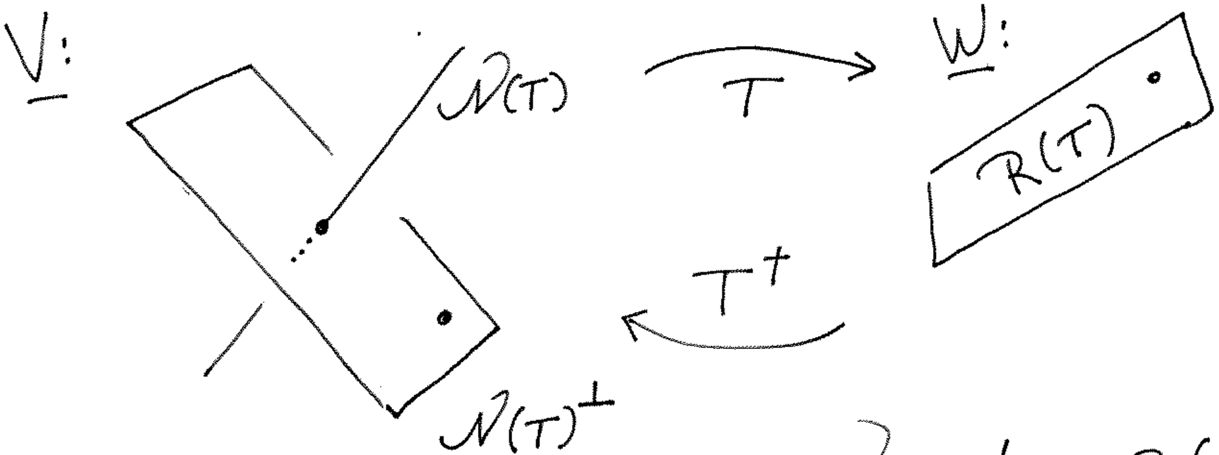
Taking $i = j$ and using that $T(v_i) \neq 0$, we learn $\lambda_i > 0$. Reindex the v_i so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. If we set $w_i = \dfrac{T(v_i)}{\|T(v_i)\|}$ then $\gamma = \{w_1, \dots, w_n\}$ is orthonormal by ✪ and you can check that $[T]_\beta^\gamma$ is $\begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}$.

Applications:

1) Computing determinants. When A is square and invertible $\det(A) = \pm \sigma_1 \sigma_2 \cdots \sigma_n$.

Pf: By HW, $\det(\text{orthogonal}) = \pm 1$.

2) Pseudo-inverse: Given $T: V \to W$ there is $T^+: W \to V$ where $T^+(W) = \mathcal{N}(T)^\perp$ the linear op $T \circ T^+$ is orthgonal projection onto $R(T)$, and $T^+ \circ T$ is the identity on $\mathcal{N}(T)^\perp$.



Point: The restriction
$$\begin{aligned} L : \mathcal{N}(T)^\perp &\to R(T) \\ x &\mapsto T(x) \end{aligned}$$
is an isomorphism and so set $T^+ = L^{-1} \circ \operatorname{proj}_{R(T)}$

3) In fact, the SVD has been a key tool in scientific computing for 50+ years.