Lecture 39: Dealing with nondiagonalizble matrices.

(§6.7 and §7.1)

Last time: For a linear op T on a finite dim'l inner product space V, the following are equivalent:

1) T is an isometry, i.e. $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for all $x, y \in V$.
2) $T \circ T^* = T^* \circ T = I_V$
3-4) For every (some) orthonormal basis β for V, the image $T(\beta)$ is an orthonormal basis.
5) $\|T(x)\| = \|x\|$ for all $x \in V$.

Def: A square matrix is unitary when $A^*A = I$ and orthogonal when $A^tA = I$.

[With respect to an orthonormal basis, an isometry has a unitary (and orthogonal if $\mathbb{F} = \mathbb{R}$) matrix.]

Thm: Suppose $A \in M_{n\times n}(\mathbb{R})$ is orthogonal. Then L_A is an isometry of $(\mathbb{R}^n, \text{dot})$.

Pf: Set $a_i = Ae_i = i^{th}$ col of A. Then $I =$

$$A^tA = \begin{pmatrix} — a_1 — \\ \vdots \\ — a_n — \end{pmatrix}\begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix} = G \text{ where } G_{ij} = \langle a_i, a_j\rangle$$

In particular $\{L_A(e_i)\}$ is an orthonormal basis of $\mathbb{R}^n$. ∎

<u>Cor</u>: $A \in M_{n\times n}(\mathbb{R})$. TFAE:

i) A is orthogonal

ii) $A^t = A^{-1}$

iii) the columns of A are an orthonormal basis for $\mathbb{R}^n$.

iv) — rows ———— " ————

v) L_A is an isometry of $(\mathbb{R}^n, \text{dot})$.

<u>Pf</u>: Exercise.

<u>Restated Thm</u>: Suppose $A \in M_{n\times n}(\mathbb{R})$ is symmetric. Then there is an orthogonal Q with $Q^tAQ = Q^{-1}AQ$ diagonal.

——— ∘ ———

<u>What to do when $A \in M_{n\times n}(\mathbb{R})$ is not diagonalizable?</u>

<u>Jordan Canonical Form</u>: Over $\mathbb{C}$, any square matrix is similar to one of the following form:

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

where each block has the form

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & 1 & \\ & & \ddots & \lambda_i \; 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

<u>Ex:</u>

$$A = \begin{pmatrix} \boxed{\begin{matrix}2&1&0\\0&2&1\\0&0&2\end{matrix}} & & & & \huge 0 \\ & \boxed{2} & & & \\ & & \boxed{\begin{matrix}3&1\\0&3\end{matrix}} & & \\ & & & \boxed{5} & \\ \huge 0 & & & & \boxed{\begin{matrix}0&1\\0&0\end{matrix}} \end{pmatrix}$$

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

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

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

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

<u>Remarks:</u> Suppose A is in Jordan Can. Form.

1) The λ_i are the eigenvalues of A.

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

3) Can still understand powers: (see page 519 for details)

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

for $n > 1$.

Applications:

1) Proof that any regular Markov chain converges.

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

2) Solving systems of linear differential equations.

Singular Value Decomposition: Work over $\mathbb{R}$ for concreteness.

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

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

↖ diagonal, J.C.F., etc.

Problems: 1) To use J.C.F. need to work over $\mathbb{C}$.

2) What about $T: V \to W$?

In (2), there's not even any notion of an eigenvector! However we do get to choose two bases, β for V and γ for W, in order to write down $[T]_\beta^\gamma$. [This gives more flexibility even if $W = V$ since $[T]_\beta = [T]_\beta^\beta$.]

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

$$[T]_\beta^\gamma = \left(\begin{array}{c:c} \begin{matrix} 1 & 0 & \\ 0 & 1 & \\ & & \ddots \; 1 \end{matrix} & 0 \\ \hdashline 0 & 0 \\ \underbrace{\qquad}_{\operatorname{rank}(T)} & \underbrace{\quad}_{\leftarrow \operatorname{nullity}(T)} \end{array}\right).$$

Pf (Skip!): Using the rep'l. theorem, 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 is linearly independent) to a basis γ for W. Then $[T]_\beta^\gamma$ has the desired form. ▨

Singular Value Decomposition: Suppose T is a linear map between real inner product spaces V and W. Then there exist orthonormal basis β and γ so that

$$[T]_\beta^\gamma = \left(\begin{array}{c:c} \begin{matrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_r \end{matrix} & 0 \\ \hdashline 0 & 0 \end{array}\right)$$

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

Notes: 1) Here $r = \text{rank}(T)$.

2) It turns out that the singular values σ_i are unique.

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

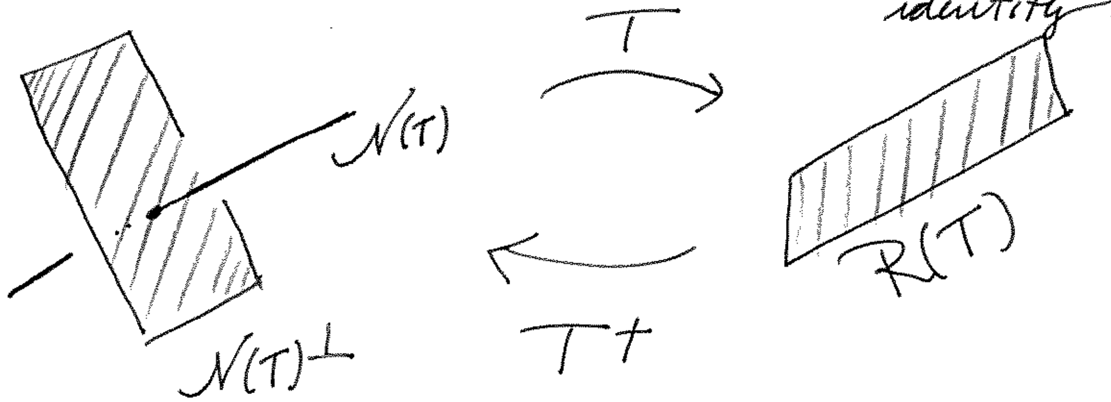
Proof idea for SVD: Diagonalize $T^* \circ T$.

[Only talked about T^* for linear ops, not general linear transformations, so just think about the case when $W = V$.]

Applications:

1) Pseudo-inverse: Given $T: V \to W$ there is a $T^\dagger: W \to V$ where

$T^{+}(W) = \mathcal{N}(T)^{\perp}$, the linear op $T \circ T^{+}$ is orthogonal projection onto $R(T)$, and $T^{+} \circ T$ is the identity on $\mathcal{N}(T)^{\perp}$.



Point: The restriction

$$\begin{aligned} L : \mathcal{N}(T)^{\perp} &\longrightarrow R(T) \\ x &\longrightarrow T(x) \end{aligned}$$

is an isomorphism, and $T^{+} = L^{-1} \circ \mathrm{proj}_{R(T)}$

2) Computing determinants. When A is square and invertible, $\det(A) = \pm \sigma_1 \cdots \sigma_r$

Proof: By HW, det(orthogonal) = ±1.

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