Lecture 10 - Eigenbases

Saw how to associate to a linear operator $L: V \to V$ a polynomial $p_A(\lambda)$ by choosing a basis and forming

$$\det(A - \lambda I).$$

This polynomial has the form $(-1)^n \lambda^n + \ldots$ & we'll often instead consider $(-1)^n p_A(\lambda) = \lambda^n + \ldots$

For a 2x2 matrix, we have

$$p_A(\lambda) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - \underbrace{(a+d)}_{\substack{\text{Trace}(A) \\ = tr(A)}}\lambda + \underbrace{(ad-bc)}_{|A|}$$

This gives us checks on possible eigenvalues: do they sum to the trace? $\lambda_1 + \lambda_2 = tr(A)$

multiply to the det? $\lambda_1 \cdot \lambda_2 = |A|$

We want to simplify our problem.

Def A subspace $W \subseteq V$ is L-invariant if

$$L(W) \subseteq W$$

This obviously applies equally to matrices, so we'll use those terms

Prop If W is A-invariant, then W is invariant under any polynomial in A.

Here if $q(x) = a_n x^n + \ldots + a_0$, then $q(A) = a_n A^n + \ldots + a_0 I$

Pf Let $\bar{w} \in W$. Then by assumption, $A(\bar{w}) \in W$, so by induction, $A^n(\bar{w}) = A^{n-1}(A\bar{w}) \in W$. Thus

$$(a_n A^n + \ldots + a_0 I)\bar{w} = a_n A^n(\bar{w}) + \ldots + a_0 \bar{w} \in W,$$ since W is a subspace. ☒

Prop If $V = U \oplus W$, U and W A invariant, then by choosing a basis for U and one for W, we can put A into a block form:

$$A \sim \begin{array}{cc} \begin{array}{cc} \overbrace{\quad\;}^{\dim U} & \overbrace{\quad\quad}^{\dim W} \end{array} & \\ \left[\begin{array}{c|c} A' & 0 \\ \hline 0 & A'' \end{array}\right] & \begin{array}{l} \}\,\dim U \\ \}\,\dim W \end{array} \end{array}$$

There is an important invariant subspace:

$\ker(A)$:

$\bar{w} \in \ker(A)$, then $A\bar{w} = \bar{0} \in \ker(A)$.

This is then invariant under $A - \lambda_0 I$ for any $\lambda_0 \in \mathbb{F}$.

Cor The eigenspaces, $\ker(A - \lambda I)$, are A-invariant.

Combining this with the previous proposition gives us a very easy set-up.

Prop If $V = \ker(A - \lambda_1 I) \oplus \ker(A - \lambda_2 I) \oplus \ldots \oplus \ker(A - \lambda_r I)$, where $\dim \ker(A - \lambda_s I) = n_s$, then V has a basis in which A is diagonal:

$$\begin{array}{cc} & \begin{array}{cccc} n_1 & n_2 & \cdots & n_r \end{array} \\ \begin{array}{c} n_1 \\ \\ n_2 \\ \\ \vdots \end{array} & \left[\begin{array}{c|c|c|c} \begin{matrix} \lambda_1 & 0 & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix} & 0 & \cdots & 0 \\ \hline 0 & \begin{matrix} \lambda_2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_2 \end{matrix} & \cdots & 0 \\ \hline \vdots & & \ddots & \\ \hline 0 & 0 & & \begin{matrix} \lambda_r & 0 \\ 0 & \lambda_r \end{matrix} \end{array}\right] \end{array}$$

Pf Take any basis of $\ker(A - \lambda_i I)$ for each i. The union is a basis for V. In this basis, A has the desired form. $\boxtimes$

Def A matrix A is **diagonalizable** if in some basis, A is diagonal (so A is similar to a diagonal matrix).

Thus if we have enough linearly independent eigenvectors, then A is diagonalizable. The linear independence is actually automatic.

Prop If $\overline{v}_1, \dots, \overline{v}_r$ are non-zero eigenvectors with distinct eigenvalues $\lambda_1, \dots, \lambda_r$, respectively, then

$$\{\overline{v}_1, \dots, \overline{v}_r\} \text{ is linearly independent.}$$

Remark Our method of proof applies more generally: if we have a set of eigenvectors such that each set of vectors with the same eigenvalue is linearly independent then the entire set is linearly independent.

Pf This will be by induction on r. If $r=1$, then it's obvious. Assume true for $r-1$. Now consider

$$a_1\overline{v}_1 + \dots + a_r\overline{v}_r = \overline{0}.$$

Applying $(A-\lambda_1 I)$ gives

$$a_2(\lambda_2-\lambda_1)\overline{v}_2 + \dots + a_r(\lambda_r-\lambda_1)\overline{v}_r = \overline{0}.$$

By the inductive hypothesis,

$$\begin{matrix} a_2(\lambda_2-\lambda_1)=0 \\ \vdots \\ a_r(\lambda_r-\lambda_1)=0 \end{matrix},$$

and since $\lambda_i \neq \lambda_j$ for $i \neq j$, we can divide both sides by $\lambda_i - \lambda_1$ for each $i \Rightarrow$

$$\begin{matrix} a_2 = 0 \\ \vdots \\ a_r = 0 \end{matrix}.$$

Thus our original equation is

$$a_1\overline{v}_1 = 0 \Rightarrow a_1 = 0. \qquad \boxtimes$$

Thus we need only have enough eigenvectors (a spanning set of them) to get diagonalizability.

Ex Let $P_n(\mathbb{R})$ be the vector space of real polynomials of degree at most n. The standard monomial basis shows this to be $(n+1)$-dimensional. Let $L = \frac{d}{dx}$.

In fact $\frac{d}{dx}$ factors:

$$\begin{tikzcd} d/dx: P_n(\mathbb{R}) \arrow[r] \arrow[dr] & P_{n-1}(\mathbb{R}) \arrow[d] \\ & P_n(\mathbb{R}) \end{tikzcd}$$

For a more specific example, let $n=2$. Then

$$\left[\frac{d}{dx}\right] = \begin{array}{c|ccc|} & 1 & x & x^2 \\ \hline 1 & 0 & 1 & 0 \\ x & 0 & 0 & 2 \\ x^2 & 0 & 0 & 0 \end{array}$$

The characteristic polynomial is $-\lambda^3$, so the only eigenvalue is 0. On the other hand, iterating $\frac{d}{dx}$ yields 0.

Def A is nilpotent if there is some m such that $A^m = 0$.

We now consider a chain of subspaces

$$\ker(A) \subseteq \ker(A^2) \subseteq \cdots \subseteq V$$

Prop 1) If $\ker(A^m) = \ker(A^{m+1})$, then $\ker(A^m) = \ker(A^{m+r})$ $\forall r$

2) For some $m > 0$, $\ker(A^m) = \ker(A^{m+1})$

Pf 1) We show that $\ker(A^m) = \ker(A^{m+1}) \Rightarrow \ker(A^{m+1}) = \ker(A^{m+2})$.

Assume $\bar{y} \in \ker(A^{m+2})$. Then $A\bar{y} \in \ker(A^{m+1})$. Since $\ker(A^m) = \ker(A^{m+1})$, $A\bar{y} \in \ker(A^m) \Rightarrow \bar{y} \in \ker(A^{m+1})$.

2) Consider the corresponding sequence of dimensions

$$\dim(\ker(A)) \leq \dim(\ker(A^2)) \leq \cdots \leq n = \dim V$$

Since these are all integers and the sequence is bounded above, for some m, $\dim(\ker A^m) = \dim(\ker A^{m+1})$. Thus $\ker(A^m) = \ker(A^{m+1})$. ∎

Thm $V = E_0 \oplus U$, where E_0 and U are A-invariant, A is nilpotent on E_0, and A is injective on U.

Pf $E_0 = \ker(A^m)$ for the smallest m where it stabilizes. Let $U = \operatorname{Im}(A^m)$. Then if $\bar{w} \in U$, then $\bar{w} = A^m \bar{v}$, and so $A\bar{w} = A^{m+1}\bar{v} = A^m(A\bar{v}) \in U$. $\Rightarrow$ U A-invariant

Consider $\bar{w} \in \ker(A^m) \cap \operatorname{Im}(A^m)$. Then $\bar{w} = A^m \bar{u}$, and $A^m \bar{w} = \bar{0}$ $\Rightarrow A^{m+m} \bar{u} = \bar{0}$. By assumption, $\ker(A^m) = \ker(A^{2m})$, so $\bar{w} = A^m \bar{u} = \bar{0}$. A similar argument shows that $\ker(A) \cap \operatorname{Im}(A^m) = \bar{0}$.

$\Rightarrow$ A is injective on U.

Now by rank-nullity,

$$\dim V = \dim(\ker A^m) + \dim(\operatorname{Im}(A^m))$$

So $\ker(A^m) \cap \operatorname{Im}(A^m) = \{\bar{0}\} \Rightarrow V = \ker(A^m) \oplus \operatorname{Im}(A^m)$. ☒