Lecture 33: Generalized eigenvectors and nilpotent operators

Last time: Polar decomposition and singular value decomposition

This time: Return to arbitrary finite-dimensional complex vector spaces.

Let $T \in \mathcal{L}_{\mathbb{C}}(V)$, then if $\lambda \in \mathbb{C}$ is an eigenvalue of T, the eigenspace $\mathrm{Eig}(\lambda, T)$ consists of the zero vector and all eigenvectors of T corresponding to λ.

Let $\{\lambda_1, \dots, \lambda_m\}$ be the distinct eigenvalues of T, then

(*) $V = \mathrm{Eig}(\lambda_1, T) \oplus \dots \oplus \mathrm{Eig}(\lambda_m, T)$ if and only if T is diagonalizable
if and only if V has a basis
consisting of eigenvectors of T.

Goal for the next lectures: Find similar decomposition of V as in (*) that holds for all operators $T \in \mathcal{L}_{\mathbb{C}}(V)$.

Recall: v is an eigenvector of T with eigenvalue λ, then

$$v \in \mathrm{Eig}(\lambda, T) = \ker(T - \lambda I_V)$$

Def 8.9, 8.10 Let $T \in \mathcal{L}(V)$ and λ be an eigenvalue of T.

- $v \in V$ is called a generalized eigenvector of T corr. to λ, if $v \neq 0$ and $(T - \lambda I_V)^j(v) = 0$ for some $j \in \mathbb{N}$.
- The generalized eigenspace of T corr. to the eigenvalue λ is denoted by $G(\lambda, T)$, and consists of 0 and all generalized eigenvectors of T corr. to λ.

Remark:
- $Eig(\lambda, T) \subseteq G(\lambda, T)$ by definition.
- $G(\lambda, T) = \ker (T - \lambda I_V)^n$, where $n = \dim V$, as we now prove.

Preparation: kernels of powers of operators

Prop 8.2 Let $T \in \mathcal{L}(V)$. Then $\{0\} = \ker T^0 \subseteq \ker T \subseteq \ker T^2 \subseteq \ker T^3$

$(k \in \mathbb{N})$ $\qquad \ldots \subseteq \ker T^k \subseteq \ker T^{k+1} \subseteq \ldots$

Proof: Let $k \in \mathbb{N}$ be arbitrary and $v \in \ker T^k$, i.e., $T^k(v) = 0$.

Then also $0 = T(0) = T(T^k(v)) = T^{k+1}(v) \Rightarrow v \in \ker T^{k+1}$. $\square$

Prop 8.3 Let $T \in \mathcal{L}(V)$. If $\ker T^m = \ker T^{m+1}$ for some $m \in \mathbb{N}$,

then $\ker T^m = \ker T^{m+1} = \ker T^{m+k}$ for all $k \in \mathbb{N}$.

Proof: Let $h \geq 1$, then we want to show that $\ker T^{m+h} = \ker T^{m+h+1}$.

($\subseteq$) $\ker T^{m+h} \subseteq \ker T^{m+h+1}$ by Prop 8.2.

($\supseteq$) Let $v \in \ker T^{m+h+1}$, then $0 = T^{m+h+1}(v) = T^{m+1}(T^h(v))$

$\Rightarrow T^h(v) \in \ker T^{m+1} = \ker T^m$, i.e., $T^m(T^h(v)) = T^{m+h}(v) = 0$

$\Rightarrow v \in \ker T^{m+h} \Rightarrow \ker T^{m+h+1} \subseteq \ker T^{m+h}$. $\square$

This stabilization eventually happens, since V is finite-dim.:

$\boxed{\text{Prop 8.4}}$ Let $T \in \mathcal{L}(V)$ and $n = \dim V$.

Then $\ker T^n = \ker T^{n+1} = \ker T^{n+h}$ $\forall h \in \mathbb{N}$.

Proof: assume $\ker T^n \subsetneq \ker T^{n+1}$, then

$$\{0\} = \ker T^0 \subsetneq \ker T \subsetneq \ker T^2 \subsetneq \ldots \subsetneq \ker T^n \subsetneq \ker T^{n+1}$$

Since each inclusion is strict, the dimension grows by at least 1 in each step. But then $\dim \ker T^{n+1} \geq n+1$, which is impossible since $\ker T^{n+1} \subseteq V$ and $n = \dim V$. $\lightning$.

Hence, $\ker T^n = \ker T^{n+1}$, and the rest follows from Prop 8.3. $\square$

Now, $G(\lambda, T) = \ker (T - \lambda I_V)^n$ where $n = \dim V$ follows from Prop. 8.4.

$$G(\lambda,T) = \{v \in V : (T - \lambda I_V)^j(v) = 0 \text{ for some } j \in \mathbb{N}\}$$

$\boxed{\text{Prop } 8.13}$ Let $T \in \mathcal{L}_{\mathbb{C}}(V)$ with distinct eigenvalues $\{\lambda_1, \dots, \lambda_m\}$ and corresponding generalized eigenvectors $v_1, \dots, v_m$. Then $\{v_1, \dots, v_m\}$ are linearly independent.

Proof: Let $a_1 v_1 + \dots + a_m v_m = 0 \ ^{(*)}$ for some $a_i \in \mathbb{C}$. To show: $a_i = 0 \ \forall i$.

Let $h \in \mathbb{N}$ be the largest integer s.t. $(T - \lambda_1 I)^h (v_1) = w \neq 0$.

Then $(T - \lambda_1 I_V)(w) = (T - \lambda_1 I_V)^{h+1}(v) = 0$, so

w is an eigenvector of T corr. to the eigenvalue λ_1.

$\Rightarrow T(w) = \lambda_1 w$, and hence for $\lambda \in \mathbb{C}$ we have

$$(T - \lambda I_V)(w) = (\lambda_1 - \lambda)w$$

and $$(T - \lambda I_V)^n (w) = (\lambda_1 - \lambda)^n w. \qquad (**)$$

Note also that $\underline{(T - \lambda_j I_V)^n (v_j) = 0 \text{ for } j \geq 2}$. $\quad (***)$

Applying the operator $(T - \lambda_1 I_V)^h (T - \lambda_2 I_V)^n \dots (T - \lambda_m I_V)^n$ to $(*)$:

$= w$

$$0 = a_1 (T - \lambda_1 I_V)^h (T - \lambda_2 I_V)^n \dots (T - \lambda_m I_V)^n (v_1)$$
$$+ a_2 (T - \lambda_1 I_V)^h \underbrace{(T - \lambda_2 I_V)^n \dots (T - \lambda_m I_V)^n (v_2)}$$

$= 0$ by $(***)$

$$\vdots$$

$$= a_1 (T - \lambda_2 I_V)^n \dots (T - \lambda_m I_V)^n (w)$$
$$= a_1 (\underbrace{\lambda_1 - \lambda_2}_{\neq 0})^n \dots (\underbrace{\lambda_1 - \lambda_m}_{\neq 0})^n w \qquad \Rightarrow a_1 = 0 \ (w \neq 0)$$

Apply the same steps for each a_j, showing $a_j = 0 \ \forall j$

$\Rightarrow \{v_1, \dots, v_m\}$ are linearly indep. $\square$

Nilpotent operators

$\boxed{\text{Def 8.16}}$ An operator $T \in \mathcal{L}(V)$ is called nilpotent if $T^k = 0$ for some $k \in \mathbb{N}$.

Ex.: ·) $\partial : P_d(\mathbb{F}) \to P_d(\mathbb{F})$ is nilpotent since $\partial^{d+1} = 0$.

·) $T : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} y \\ 0 \end{pmatrix}$ is nilpotent, since $T^2 = 0$.

$\boxed{\text{Prop 8.18}}$ Suppose $N \in \mathcal{L}(V)$ is nilpotent, then $N^{\dim V} = 0$.

Proof: Let k be s.t. $N^k = 0$, then $\ker N^k = V$

If $k \le \dim V$, then clearly also $N^{\dim V} = 0$

If $k \ge \dim V$, then by Prop 8.4, $\ker N^{\dim V} = \ker N^k = V$,

so $N^{\dim V} = 0$. $\square$

Prop 8.19 Let $N \in \mathcal{L}(V)$ be nilpotent. Then there exists a basis $\mathcal{B}$ of V s.t. $\mathcal{M}(N)_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$.

Proof: Recall that $\{0\} = \ker N^0 \subseteq \ker N \subseteq \ker N^2 \subseteq \ldots \subseteq \ker N^{\dim V} = V$ by Prop 8.2 and Prop 8.18.

Now choose a basis for $\ker N$, extend this to a basis for $\ker N^2$, and so forth until you reach $\ker N^{\dim V} = V$.

Let $2 \leq j \leq \dim V$, and consider a basis $\{v_1, \ldots, v_m\}$ for $\ker N^j$.

Then $N(v_k) \in \ker N^{j-1}$ $\left(0 = N^j(v_k) = N^{j-1}(N(v_k))\right)$ and hence $N(v_k)$ can be expressed as a linear combination of basis vectors for $\ker N^{j-1}$, which proves the claimed matrix representation. $\square$