Lecture 15: primary decomposition theorem

Why will this theorem be useful?

Lemma 0.1. *If $T : V \to V$ is linear and $\lambda \in \mathbb{F}$ then $\hat{E}_\lambda$ is T-invariant. That is, if $v \in \hat{E}_\lambda$ then $T(v) \in \hat{E}_\lambda$.*

Proof. Let $v \in \hat{E}_\lambda$. Then there is some k such that $(\lambda I - T)^k v = \vec{0}$. Now

$$(\lambda I - T)^k (T(v)) = T((\lambda I - T)^k (v)) = T(\vec{0}) = \vec{0} \ .$$

Here we have used that $(\lambda I - T)^k$ and T commute. This is because the first operator is just a combination of operators of the form T^m, all of which commute with T. Therefore $T(v) \in \hat{E}_\lambda$. □

From this lemma, we see that once we prove the primary decomposition theorem, we will be able to write V as a T-invariant direct sum of generalized eigenspaces. So letting B_i be a basis of $\hat{E}_{\lambda_i}$ and $B = \cup_{i=1}^k B_i$, B is a basis for V and $[T]_B^B$ is a block diagonal matrix. The reason is that if we take $T(v)$ for some vector $v \in B_i$ then the result lies in $\hat{E}_{\lambda_i}$ and therefore we only need to use the vectors in B_i to represent it. All entries in its expansion in terms of B corresponding to vectors in B_j (for $j \neq i$) will be zero, giving a block diagonal matrix.

Proof of the primary decomposition theorem. The proof will follow several steps.

Step 1. *Peeling off a generalized eigenspace.* The point of this step is to show that for some eigenvalue λ_1 and some subspace W_1, we can write

$$V = \hat{E}_{\lambda_1} \oplus W_1 \ .$$

Then we will restrict T to W_1 and argue by induction.

Since the characteristic polynomial c_T has coefficients from an algebraically closed field $\mathbb{F}$, it has a root $\lambda_1 \in \mathbb{F}$. Then the generalized eigenspace $\hat{E}_{\lambda_1}$ has nonzero dimension.

We claim that there is some k_1 such that $\hat{E}_{\lambda_1} = N(\lambda_1 I - T)^{k_1}$. To show this, let $v_1, \ldots, v_{t_1}$ be a basis for $\hat{E}_{\lambda_1}$. Then for each $j = 1, \ldots, t_1$ there exists $p_j \geq 1$ such that $(\lambda_1 I - T)^{p_j}(v_j) = \vec{0}$. Let $k_1 = \max\{p_1, \ldots, p_{t_1}\}$. Clearly $N(\lambda_1 I - T)^{k_1} \subset \hat{E}_{\lambda_1}$, so we need only show the other inclusion. If $v \in \hat{E}_{\lambda_1}$ we can write

$$v = a_1 v_1 + \cdots + a_{t_1} v_{t_1} \ ,$$

so

$$(\lambda_1 I - T)^{k_1}(v) = a_1 (\lambda_1 I - T)^{k_1}(v_1) + \cdots + (\lambda_1 I - T)^{k_1}(v_{t_1}) \ .$$

However $k_1 \geq p_j$ for all j so this is zero. Therefore $v \in N(\lambda_1 I - T)^{k_1}$ and we have proven the claim.

Next we show that

$$V = N(\lambda_1 I - T)^{k_1} \oplus R(\lambda_1 I - T)^{k_1} \ .$$

(As pointed out in class, this claim can be proved by just noticing that the operator $U = (\lambda_1 I - T)^{k_1}$ satisfies $N(U) = N(U^2)$ and therefore, by a homework problem, $V = N(U) \oplus R(U)$.) The rank-nullity theorem implies that their dimensions sum to the dimension of V, so we need only show that their intersection is the zero subspace. (Use the two subspace dimension theorem.) So assume that v is in the intersection, meaning that there exists $w \in V$ such that $(\lambda_1 I - T)^{k_1}(w) = v$ and $(\lambda_1 I - T)^{k_1}(v) = \vec{0}$. But then we get $(\lambda_1 I - T)^{2k_1}(w) = \vec{0}$ and therefore $w \in \hat{E}_{\lambda_1}$. This means actually

$$\vec{0} = (\lambda_1 I - T)^{k_1}(w) = v \ .$$

We now set $W_1 = R(\lambda_1 I - T)^{k_1}$ and we are done with this step.

Step 2. *T-invariance of the direct sum.* We saw above that $\hat{E}_{\lambda_1}$ is T-invariant. We claim that W_1 is as well, so that the direct sum is T-invariant, and we have obtained our first "block."

If $v \in W_1 = R(\lambda_1 I - T)^{k_1}$ then there exists $w \in V$ such that $(\lambda_1 I - T)^{k_1}(w) = v$. Then

$$T(v) = T((\lambda_1 I - T)^{k_1}(w)) = (\lambda_1 I - T)^{k_1}(T(w))$$

because T and $\lambda_1 I - T$ commute. Therefore $T(v) \in W_1$ and W_1 is T-invariant.

To conclude this step, we define T_1 to be T restricted to W_1. That is, we view W_1 as a vector space of its own and $T_1 : W_1 \to W_1$ as a linear transformation defined by $T_1(w) = T(w)$ for $w \in W_1$.

Step 3. $\hat{E}_{\lambda_2}, \ldots, \hat{E}_{\lambda_k}$ *are in* W_1. We now show that if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T then $\hat{E}_{\lambda_2}, \ldots, \hat{E}_{\lambda_k}$ are contained in W_1.

So first let $v \in \hat{E}_{\lambda_j}$ for some $j = 2, \ldots, k$. By definition of the generalized eigenspace we can find t such that $v \in N(\lambda_j I - T)^t$. We will now use a lemma that follows from the homework.

Lemma 0.2. *There exist polynomials $p, q \in \mathbb{F}[x]$ such that*

$$(\lambda_j - x)^t p + (\lambda_1 - x)^{k_1} q = 1 \ .$$

Proof. Since $(\lambda_j - x)^t$ and $(\lambda_1 - x)^{k_1}$ do not have a common root, you proved in the homework that their greatest common divisor is 1. Then the lemma follows from the result on the homework: if $r, s \in \mathbb{F}[x]$ have greatest common divisor d then there exist $p, q \in \mathbb{F}[x]$ such that $rp + sq = d$. $\square$

We will use the lemma but in its transformation form. For any polynomial $a \in \mathbb{F}[x]$ of the form $a(x) = a_n x^n + \cdots + a_0$ we define

$$a(T) = a_n T^n + \cdots + a_1 T + a_0 I \ .$$

Therefore

$$I = (\lambda_j I - T)^t p(T) + (\lambda_1 I - T)^{k_1} q(T) \ .$$

Applying this to v, we get

$$v = (\lambda_j I - T)^t p(T)(v) + (\lambda_1 I - T)^{k_1} q(T)(v) \ .$$

But all these polynomial transformations commute, so we can use $v \in N(\lambda_j I - T)^t$ to find

$$v = (\lambda_1 I - T)^{k_1}(q(T)(v)) \in R(\lambda_1 I - T)^{k_1} = W_1 \ .$$

Therefore $v \in W_1$ and so $\hat{E}_{\lambda_j} \subset W_1$ for $j = 2, \ldots, k$.

Step 4. $\hat{E}_{\lambda_2}, \ldots, \hat{E}_{\lambda_k}$ *are the generalized eigenspaces of* T_1. Let $w \in W_1$ be a vector in a generalized eigenspace of T_1 with eigenvalue λ. Then for some t, $(\lambda I - T_1)^t(w) = \vec{0}$ and since T_1 acts the same as T, we find w is a generalized eigenvector of T. This means that $w \in \hat{E}_{\lambda_j}$ for some j and thus $\lambda = \lambda_j$. If $j = 1$ then we would have $w \in W_1 \cap \hat{E}_{\lambda_1}$, giving $w = \vec{0}$. Therefore either $j > 1$ or $w = \vec{0}$, meaning in either case that $w \in \hat{E}_{\lambda_2} \cup \cdots \cup \hat{E}_{\lambda_k}$.

Conversely if $w \in \hat{E}_{\lambda_j}$ for some $j = 2, \ldots, k$ then $w \in W_1$. Then for some t, $(\lambda_j I - T)^t(w) = \vec{0}$. But T acts the same as T_1 on W_1 so $(\lambda_j I - T_1)(w) = \vec{0}$ and w is a generalized eigenvector of T_1.

Step 5. *The inductive step.* We will argue for the theorem by induction on the number of distinct eigenvalues of T. Let $e(T)$ be this number. If $e(T) = 1$ then we have seen that $V = \hat{E}_{\lambda_1} \oplus W_1$ but that T_1 has no eigenvalues. This means that W_1 must have dimension zero and $V = \hat{E}_{\lambda_1}$. Therefore $V = \hat{E}_{\lambda_1}$ and we are done.

Now assume that the theorem holds for all linear $U : V \to V$ such that $e(U) \leq k$ (for some $k \geq 1$) and let $T : V \to V$ be linear with $e(T) = k + 1$. Then let λ_1 be an eigenvalue of T and decompose $V = \hat{E}_{\lambda_1} \oplus W_1$. The transformation T_1 has $e(T_1) \leq k$ so we can write W_1 as a direct sum of its generalized eigenspaces. These are just $\hat{E}_{\lambda_2}, \ldots, \hat{E}_{\lambda_k}$, the other generalized eigenspaces of T, so

$$W_1 = \hat{E}_{\lambda_2} \oplus \cdots \oplus \hat{E}_{\lambda_k} \ .$$

Therefore $V = \hat{E}_{\lambda_1} \oplus \cdots \oplus \hat{E}_{\lambda_k}$ and we are done. $\square$