Cyclic subspaces for linear operators

Let V be a finite dimensional vector space and $T : V \to V$ be a linear operator. One way to create T-invariant subspaces is as follows. Choose a non-zero vector $\mathbf{v} \in V$, and let $k \in \mathbf{N}$ be the smallest integer such that $\{\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots, T^k\mathbf{v}\}$ is a dependent set. Let

$$H_{\mathbf{v}} = \operatorname{span}\{\mathbf{v}, T\mathbf{v}, \ldots, T^{k-1}\mathbf{v}\}.$$

Then $H_{\mathbf{v}}$ is called the *cyclic subspace generated by* $\mathbf{v}$. By our choice of k, we have a non-trivial linear combination of $\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}$ that vanishes. Moreover, the coefficient of $T^k\mathbf{v}$ in this combination must be non-zero, because the vectors $\mathbf{v}, \ldots, T^{k-1}\mathbf{v}$ are independent. Hence after dividing the combination by the coefficient of $T^k\mathbf{v}$, we arrive at

$$T^k\mathbf{v} + c_{k-1}T^{k-1}\mathbf{v} + \cdots + c_0\mathbf{v} = \mathbf{0},$$

for some scalars $c_0, \ldots, c_{k-1}$. We associate to the linear combination on the right a polynomial

$$p_{\mathbf{v}}(x) := x^k + c_{k-1}x^{k-1} + \cdots + c_0.$$

Theorem 0.1. *$H_{\mathbf{v}}$ is the smallest T-invariant subspaces that contains $T^\ell\mathbf{v}$ for every $\ell \in \mathbf{N}$. Relative to the basis $\{\mathbf{v}, T\mathbf{v}, \ldots, T^{k-1}\mathbf{v}\}$, the restricited transformation $T : H_{\mathbf{v}} \to H_{\mathbf{v}}$ has matrix*

$$A = \begin{bmatrix} 0 & 0 & \ldots & 0 & c_0 \\ 1 & 0 & \ldots & 0 & c_1 \\ 0 & 1 & \ldots & 0 & c_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \ldots & 1 & c_{k-1} \end{bmatrix}.$$

Hence the characteristic polynomial of $T|_H$ is $p_{\mathbf{v}}$.

Proof. First I show by induction that $T^\ell\mathbf{v} \in H_{\mathbf{v}}$ for every $\ell \in \mathbf{N}$. When $\ell = 0$, this is true by definition of $H_{\mathbf{v}}$. Suppose it's true for $\ell = m - 1$. That is,

$$T^{m-1}\mathbf{v} = a_0\mathbf{v} + \cdots + a_{k-1}T^{k-1}\mathbf{v} \in H_{\mathbf{v}}.$$

Then

$$T^m\mathbf{v} = T(T^{m-1}\mathbf{v}) = a_0T\mathbf{v} + \cdots + a_{k-2}T^{k-1}\mathbf{v} + a_{k-1}T^k\mathbf{v}.$$

All terms on the right, except the last one, belong to $H_{\mathbf{v}}$ by definition. The final term belongs to $H_{\mathbf{v}}$, because as we saw above, $T^k\mathbf{v}$ is a equal to a linear combination of $\mathbf{v}, \ldots, T^{k-1}\mathbf{v}$. Hence $T^m\mathbf{v} \in H_{\mathbf{v}}$, which completes the induction step and the proof that $T^\ell\mathbf{v} \in H_{\mathbf{v}}$ for every $\ell \in \mathbf{N}$.

Note that this fact implies that T maps each of the basis vectors $\mathbf{v}, \ldots, T^{k-1}$ for $H_{\mathbf{v}}$ back into $H_{\mathbf{v}}$. Hence $H_{\mathbf{v}}$ is T-invariant. Note further that if H is any subspace (let alone an invariant one) containing $T^\ell(\mathbf{v})$ for every ℓ, then in particular, H contains the basis $\{\mathbf{v}, \ldots, T^{k-1}\mathbf{v}\}$ for $H_{\mathbf{v}}$. Therefore H contains $H_{\mathbf{v}}$. That is, $H_{\mathbf{v}}$ is the smallest subspace of V that contains $T^\ell\mathbf{v}$ for all $\ell \in \mathbf{N}$.

Now the matrix of T relative to $\mathcal{B} = \{\mathbf{v}, \dots, T^{k-1}\mathbf{v}\}$ is

$$\begin{aligned} A &= [[T(\mathbf{v})]_{\mathcal{B}}\,[T(T\mathbf{v})]_{\mathcal{B}}\,\dots\,[T(T^{k-1}\mathbf{v})]_{\mathcal{B}}] = [[T\mathbf{v}]_{\mathcal{B}}\,[T^2\mathbf{v}]_{\mathcal{B}}\,\dots\,[T^k\mathbf{v}]_{\mathcal{B}}] \\ &= \begin{bmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -c_{k-1} \end{bmatrix}, \end{aligned}$$

where, in the last column, I have used the formula for $T^k(\mathbf{v})$ that precedes the statement of the theorem.

Thus I evaluate the characteristic polynomial of T by cofactor expansion about the last column in

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & \dots & 0 & c_0 \\ -1 & \lambda & \dots & 0 & c_1 \\ 0 & -1 & \dots & 0 & c_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \dots & -1 & \lambda + c_{k-1} \end{vmatrix} = (-1)^{2k}(\lambda + c_{k-1}) \det A_{kk} + \sum_{j=1}^{k-1} (-1)^{j+k}(c_{j-1}) \det A_{jk}.$$

Here A_{jk} is the jk-minor of $\lambda I - A$. It has block diagonal form

$$A_{jk} = \begin{bmatrix} B_j & 0 \\ 0 & C_j \end{bmatrix},$$

where the $(j-1) \times (j-1)$ matrix B_j and the $(k-j) \times (k-j)$ matrix C_j are given by

$$B_j = \begin{bmatrix} \lambda & 0 & \dots & 0 & 0 \\ -1 & \lambda & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ & & \ddots & & \vdots \\ 0 & 0 & \dots & -1 & \lambda \end{bmatrix}, \qquad C_j = \begin{bmatrix} -1 & \lambda & 0 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda \\ 0 & 0 & 0 & \dots & -1 \end{bmatrix},$$

Thus $\det A_{jk} = (\det B_j)(\det C_j) = (\lambda)^{j-1}(-1)^{k-j}$. Plugging back into the cofactor formula, we find that the characteristic polynomial of T is

$$(-1)^{2k}(\lambda + c_{k-1})\lambda^{k-1} + \sum_{j=1}^{k-1} (-1)^{2k} c_{j-1} \lambda^{j-1} = \lambda^k + c_{k-1}\lambda^{k-1} + c_{k-2}\lambda^{k-2} + \dots + c_0$$

as asserted. □