The Cayley-Hamilton Theorem

Here we will introduce the concepts of T-invariant subspaces, T-cyclic subspaces, companion matrices, and then prove the Cayley-Hamilton Theorem. The matrix version of the Cayley-Hamilton Theorem says plugging a square matrix into its characteristic polynomial will result in the zero matrix. This looks like it should work: Let $f(t) = \det(tI - A)$ for some square matrix A. Then $f(A)$ should be $\det(AI - A) = \det(A - A) = \det(0) = 0$. So what's the problem? Well, the issue at hand is whether it is ok equate: "plug in $t = A$ and take the determinant" and "take the determinant and then plug in $t = A$". It is not obvious that these should result in the same answer. It is possible to repair the above attempt at a proof,[1] but we will follow a different path.

In all that follows, we let V be a finite dimensional vector space (over some field $\mathbb{F}$) and $T : V \to V$ a linear operator on V. Also, recall that $T^k(\mathbf{v}) = \underbrace{T(T(\cdots T(\mathbf{v})\cdots))}_{k-\text{times}}$ for $k \in \mathbb{Z}_{>0}$ and $T^0(\mathbf{v}) = I(\mathbf{v}) = \mathbf{v}$ is just the identity. Also, given a polynomial $f(t) = a_m t^m + \cdots + a_0$, we have $f(T)(\mathbf{v}) = a_m T^m(\mathbf{v}) + \cdots + a_1 T(\mathbf{v}) + a_0\mathbf{v}$. Also, recall that $\det(tI - T)$ is the characteristic polynomial of T and that this can be calculated using a coordinate matrix $[T]^\beta_\beta$ for any basis β of V.

Definition: Let W be a subspace of V. We say W is **T-invariant** if $T(W) \subseteq W$ (i.e., for every $\mathbf{w} \in W$, we have $T(\mathbf{w}) \in W$). In words, T maps W back into itself. In such a case, we can consider $T\Big|_W : W \to W$ (i.e., T restricted to the domain W) as a linear operator on W.

Let W be some T-invariant subspace with basis α. We can extend this basis to a basis for V, say $\beta = \alpha \dot\cup \alpha'$ where $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and $\alpha' = \{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$. Notice for any $1 \le j \le k$, we have $T(\mathbf{v}_j) \in W$ since $\mathbf{v}_j \in \alpha \subseteq W$. Thus $T(\mathbf{v}_j) = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ for some $c_1, \ldots, c_k \in \mathbb{F}$. These are precisely the coodinates of $\left[T\Big|_W(\mathbf{v}_j)\right]_\alpha$. In other words, the β-coordinates of $T(\mathbf{v}_j)$ are the α-coordinates of $T\Big|_W(\mathbf{v}_j)$ padded out with $(n-k)$ zeros. We have:

$$[T]^\beta_\beta = \begin{bmatrix} \left[T\Big|_W\right]^\alpha_\alpha & \text{stuff} \\ 0_{(n-k)\times k} & \text{stuff} \end{bmatrix}$$

Thus T's coordinate matrix has an upper block matrix shape to it. Using similar reasoning, if $V = W_1 \oplus \cdots \oplus W_\ell$ where each W_i is T-invariant with basis α_i, then $\beta = \alpha_1 \dot\cup \cdots \dot\cup \alpha_\ell$ is a basis for V and $[T]^\beta_\beta$ is a block diagonal matrix whose diagonal blocks are $\left[T\Big|_{W_i}\right]^{\alpha_i}_{\alpha_i}$ for $i = 1, \ldots, \ell$.

Again, letting α be a basis for an T-invariant subspace W and β an extension of α to a basis for V, if $B = [T]^\beta_\beta$ and $A = [T|_W]^\alpha_\alpha$, then $B = \begin{bmatrix} A & C \\ 0 & D \end{bmatrix}$ for some matrices C, D. Therefore, $\det(tI - B) = \det(tI - A)\det(tI - D)$. Consequently, we have:

Proposition: Let W be a T-invariant subspace.

Then the characteristic polynomial of $T\Big|_W$ divides the characteristic polynomial of T.

Definition: Let $\mathbf{v} \in V$ and $W = \text{span}\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\}$ is the **T-cyclic subspace** generated by $\mathbf{v}$.

Let $\mathbf{0} \neq \mathbf{v} \in V$ and let $W = \text{span}\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\}$ be the T-cyclic subspace generated by $\mathbf{v}$. Now $\dim(W) \le \dim(V) < \infty$. Thus $\{\mathbf{v}, T(\mathbf{v}), \ldots\}$ (an infinite set) must be linearly dependent. Let k be the smallest positive integer such that $T^k(\mathbf{v}) \in \text{span}\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{k-1}(\mathbf{v})\}$ (call this space W'). Thus $T^k(\mathbf{v}) = c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_{k-1}T^{k-1}(\mathbf{v})$ for some $c_0, c_1, \ldots, c_{k-1} \in \mathbb{F}$. Let $g(t) = t^k - c_{k-1}t^{k-1} + \cdots - c_1t - c_0$. Then $g(T)(\mathbf{v}) = T^k(\mathbf{v}) - c_{k-1}T^{k-1}(\mathbf{v}) - \cdots - c_0\mathbf{v} = \mathbf{0}$. Consider some $\ell > 0$. Divide t^ℓ by $g(t)$ and get $t^\ell = q(t)g(t) + r(t)$ for some polynomials $q(t), r(t) \in \mathbb{F}[t]$ with $r(t) = 0$ or $\deg(r(t)) < \deg(g(t)) = k$. Thus $r(t) = b_{k-1}t^{k-1} + \cdots + b_1t + b_0$ for some $b_0, \ldots, b_{k-1} \in \mathbb{F}$. We have $T^\ell(\mathbf{v}) = q(T)g(T)(\mathbf{v}) + r(T)(\mathbf{v}) = q(T)(\mathbf{0}) + r(T)(\mathbf{v}) = b_{k-1}T^{k-1}(\mathbf{v}) + \cdots + b_1T(\mathbf{v}) + b_0\mathbf{v} \in W'$. This implies $W = W' = \text{span}\{\mathbf{v}, \ldots, T^{k-1}(\mathbf{v})\}$. Notice $\beta = \{\mathbf{v}, \ldots, T^{k-1}(\mathbf{v})\}$ must be linear independent since otherwise we would have some $T^\ell(\mathbf{v})$ as a linear combination of $\{\mathbf{v}, \ldots, T^{\ell-1}(\mathbf{v})\}$ for $\ell < k$ (contradicting k being the smallest such positive integer). We say $g(t)$ is W's associated polynoial and have:

Proposition: Let $W = \text{span}\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\}$ be the T-cyclic subspace generated by $\mathbf{v}$.

Then $\beta = \{\mathbf{v}, T(\mathbf{v}), \ldots, T^{k-1}(\mathbf{v})\}$ is a basis for W where $k = \dim(W)$.

[1] See Neal McCoy's *Rings and Ideals* where he carefully develops the theory of polynomials with non-commutative coefficients.

Consider a k-dimensional T-cyclic subpace W generated by $\mathbf{v} \neq \mathbf{0}$. Again, $\beta = \{\mathbf{v}, \ldots, T^{k-1}(\mathbf{v})\}$ is a basis for W, and say $T^k(\mathbf{v}) = c_0\mathbf{v} + \cdots + c_{k-1}T^{k-1}(\mathbf{v})$. Then we have:

$$\left[T\Big|_W\right]_\beta^\beta = \Big[[T(\mathbf{v})]_\beta\ [T(T(\mathbf{v}))]_\beta\ \cdots\ [T(T^{k-1}(\mathbf{v})]_\beta\Big] = \Big[[T(\mathbf{v})]_\beta\ [T^2(\mathbf{v})]_\beta\ \cdots\ [T^{k-1}(\mathbf{v})]_\beta\ [c_0\mathbf{v} + \cdots + c_{k-1}T^{k-1}(\mathbf{v})]_\beta\Big]$$

$$= \begin{bmatrix} 0 & 0 & \cdots & 0 & c_0 \\ 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & & 0 & c_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & c_{k-1} \end{bmatrix} \quad \text{(for convenience, call this matrix } C\text{)}.$$

Definition: This is called the **companion matrix** associated with the polynomial $t^k - c_{k-1}t^{k-1} - \cdots - c_1t - c_0$.

Proposition: The characteristic polynomial of a companion matrix is its associated polynomial.

Proof: Proceed by induction. The base case is silly: $C = [c_0]$ then $\det(tI - C) = \det([t - c_0]) = t - c_0$. Assume this is true for $(k-1) \times (k-1)$ companion matrices and consider a $k \times k$ companion matrix C. To compute our characteristic polynomial, we expand the determinant along the first row:

$$\det(tI - C) = \det\begin{bmatrix} t & & & -c_0 \\ -1 & \ddots & & \vdots \\ & \ddots & t & -c_{k-2} \\ & & -1 & t - c_{k-1} \end{bmatrix} = t \cdot \det\begin{bmatrix} t & & & -c_1 \\ -1 & \ddots & & \vdots \\ & \ddots & t & -c_{k-2} \\ & & -1 & t - c_{k-1} \end{bmatrix} + (-1)^{k+1}(-c_0)\det\begin{bmatrix} -1 & t & & \\ & \ddots & \ddots & \\ & & & t \\ & & & -1 \end{bmatrix}$$

$$= t \cdot (t^{k-1} - c_{k-1}t^{k-2} - \cdots c_2t - c_1) + (-1)^{k+1}(-c_0)(-1)^{k-1} = t^k - c_{k-1}t^{k-1} - \cdots - c_2t^2 - c_1t - c_0$$

where we used our inductive hypothesis on the first $(k-1)\times(k-1)$ determinant and the fact that other $(k-1)\times(k-1)$ determinant is just the product of its diagonals: $(-1)(-1)\cdots(-1) = (-1)^{k-1}$. Therefore, if the statement holds for $k-1$, it also holds for k. Thus our proposition follows by induction. ■

Theorem: (Cayley-Hamilton Theorem) Let T be a linear operator on a finite dimensional vector space V with characteristic polynomial $f(t) = \det(tI - T)$. Then $f(T) = 0$.

Proof: Obviously, $f(T)(\mathbf{0}) = \mathbf{0}$. Let $\mathbf{0} \neq \mathbf{v} \in V$ and $W = \text{span}\{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\}$ be the T-cyclic subspace generated by $\mathbf{v}$. We know that if $\dim(W) = k$, then $\beta = \{\mathbf{v}, \ldots, T^{k-1}(\mathbf{v})\}$ is a basis for W and $T^k(\mathbf{v}) = c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_{k-1}T^{k-1}(\mathbf{v})$ for some $c_0, \ldots, c_{k-1} \in \mathbb{F}$. Lettting $g(t) = t^k - c_{k-1}t^{k-1} - \cdots - c_1t - c_0$, we have that $C = [T|_W]_\beta^\beta$ is the companion matrix associated with $g(t)$. Also, we have shown that the charateristic polynomial of C is just $g(t)$ itself. We also know when W is T-invariant, the characteristic polynomial of $T\Big|_W$ (i.e., $g(t)$) must divide the characteristic polynomial of T (i.e., $f(t)$). Thus there exists some polynomial $q(t) \in \mathbb{F}[t]$ such that $f(t) = q(t)g(t)$. We have $f(T)(\mathbf{v}) = q(T)(g(T)(\mathbf{v})) = q(T)(T^k(\mathbf{v}) - c_{k-1}T^{k-1}(\mathbf{v}) - \cdots - c_1T(\mathbf{v}) - c_0\mathbf{v}) = q(T)(\mathbf{0}) = \mathbf{0}$. We have shown that $f(T)(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v} \in V$. Thus $f(T) = 0$ (i.e., the zero operator). ■

Although we won't prove it here, it turns out that one can always find a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ such that V is a direct sum of the T-cyclic subspaces generated by these vectors. The existence of such generating vectors follows from the classification of finitely generated modules working over a principal ideal domain (in particular, $R = \mathbb{F}[t]$). The same classification applied to $R = \mathbb{Z}$ gives a classification of finitely generated Abelian groups.

Say we have such a set of vectors and let $W_{g_i(t)}$ be the T-cyclic subspace generated by $\mathbf{v}_i$ with associated polynomial is $g_i(t)$. Then one has $V = W_{g_1(t)} \oplus \cdots \oplus W_{g_\ell(t)}$. Let $\beta_i = \{\mathbf{v}_i, T(\mathbf{v}_i), \ldots, T^{k_i-1}(\mathbf{v}_i)\}$ be our usual basis for $W_{g_i(t)}$. Combine these: $\beta = \beta_1 \dot\cup \cdots \dot\cup \beta_\ell$ to get a basis for V. Then $[T]_\beta^\beta$ will be block diagonal and have companion matrices as its blocks on the diagonal.

Moreover, it is possible to find cyclic subspaces such that our associated polynomials successively divide one another: $g_1(t)$ divides $g_2(t)$, $g_2(t)$ divides $g_3(t)$, etc. These polynomials are *uniquely determined* by T and are called the **invariant factors** of T. Thus we can find a basis such that the coordinate matrix of T is block diagonal with companion matrices (associated with T's invariant factors) blocks. This is T's **rational canonical form**. It is a kind of standardized coordinate matrix for T. One can also find vectors such that each associated polynomial $g_i(t)$, now called **elementary divisors**, is a power of an irreducible polynomial. In this case, the corresponding coordinate matrix is called a **primary rational canonical form** for T.

These canonical forms are kind of unwieldy, but theoretically useful since they *always exist*. On the other hand, when T's characteristic polynomial splits (i.e., factors into linear factors) over our field of scalars, we can do better. Here the T-invariant subspaces associated with elementary divisors turn out to be spaces of **generalized eigenvectors**. By choosing a slightly different basis than the cyclic ones we've been using thus far, one can find a coordinate matrix for T in **Jordan form**. Jordan form is very useful and we will discuss it later.