Nov 4  Annihilating Polynomials

$T : V \rightarrow V$ a (linear operator over $F$). $f(x) \in F[x]$ a polynomial.

If $f(T) = 0 : V \rightarrow V$, we say $f$ annihilates $T$.

Theorem (Cayley–Hamilton)

Let $T$ be a linear operator on a finite dimensional vector space $V$. If $f$ is the characteristic polynomial for $T$, then $f(T) = 0$; in other words, the characteristic polynomial annihilates $T$. 
Proof: The proof is a nontrivial application of the determinant.

We need a new ring

$$R = \{ \text{polynomials in } T \}$$

$R$ is a commutative ring with identity $= 1$.

Choose a basis $B = \{ e_1, \ldots, e_n \}$ of $V$ and let $A$ be the matrix of $T$ relative to the basis $B$.

So we have

$$T e_i = \sum_{j=1}^{n} A_{ji} e_j \quad 1 \leq i \leq n$$

Then we have $n$ equations.
\[ \sum_{j=1}^{n} (\delta_{ij} T - A_{ji} I) \alpha_j = 0 \quad 1 \leq i \leq n \]

Here \( \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \)

is called the Kronecker delta.

You should regard these as linear equations with coefficients in the ring \( R \).

We denote the coefficients as

\[ B_{ij} = \delta_{ij} T - A_{ji} I \]

\[ 1 \leq i \leq n \quad 1 \leq j \leq n \]

\( n \times n \) entries \( \Rightarrow \) a matrix over \( R \)

\[ B = (B_{ij}) \]
More precisely, $a_1, \ldots, a_n$ should be regarded as generators of a $R$-module $V$.) And we have $n$ relations \[ \sum_{j=1}^{n} B_{ij} a_j = 0 \quad 1 \leq i \leq n \]

Now $\det B = f(T)$

This is because the characteristic polynomial

\[ f(x) = \det (x I - A) \]

and \((x I - A)_{ij} = \delta_{ij} x - A_{ji} \)

Now we want to show

\[ f(T) a_k = \det B a_k = 0 \quad \text{for each} \quad 1 \leq k \leq n \]
Consider the adjoint matrix

\[ \widetilde{B} = \text{adj} \ B \]

a matrix over \( R \).

We know that

\[ \widetilde{B} \cdot B = \det B \cdot I \]

Therefore

\[
\begin{pmatrix}
\det B \cdot a_1 \\
\vdots \\
\det a_n
\end{pmatrix}
= \begin{pmatrix}
\det B & 0 \\
0 & \det B
\end{pmatrix}
\begin{pmatrix}
a_1 \\
\vdots \\
a_n
\end{pmatrix}
\]

\[ = \left( \widetilde{B} \cdot B \right) \begin{pmatrix}
a_1 \\
\vdots \\
a_n
\end{pmatrix}
\]

\[ = \widetilde{B} \left( B \begin{pmatrix}
a_1 \\
\vdots \\
a_n
\end{pmatrix}\right) \]
\[ = \hat{B} \begin{pmatrix} \sum_{j=1}^{n} B_{1j} \alpha_j \\ \sum_{j=1}^{n} B_{2j} \alpha_j \\ \vdots \\ \sum_{j=1}^{n} B_{nj} \alpha_j \end{pmatrix} \]

\[ = \hat{B} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \]

\[ = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \]

Therefore \( f(T) \alpha_k = \det \alpha_k = 0 \)
for any \( 1 \leq k \leq n \)

But \( \{ \alpha_1, \ldots, \alpha_n \} \) is a basis of \( V \).
\( f(T) \alpha = 0 \) for any \( \alpha \in \mathbb{V} \).

That means \( f(T) = 0 \).

For example, if \( n = 2 \), the proof says this:

Choose a basis \( \{ \alpha_1, \alpha_2 \} \)

\[
T \alpha_1 = A_{11} \alpha_1 + A_{21} \alpha_2
\]

\[
T \alpha_2 = A_{12} \alpha_1 + A_{22} \alpha_2
\]

\[
(T - A_{11}) \alpha_1 - A_{21} \alpha_2 = 0
\]

\[
-A_{12} \alpha_1 + (T - A_{22}) \alpha_2 = 0
\]

\[
B = \begin{pmatrix}
T - A_{11}I & -A_{21}I \\
-A_{12}I & T - A_{22}I
\end{pmatrix}
\]
\[
B \begin{pmatrix}
\alpha_1 \\
\alpha_2
\end{pmatrix}
= \begin{pmatrix}
0 \\
0
\end{pmatrix}
\]

\[
\det B = (T - A_{11} I)(T - A_{22} I)
- A_{12} A_{21} I
\]

\[
= T^2 - (A_{11} + A_{22}) T
+ (A_{11} A_{22} - A_{12} A_{21}) I
\]

\[
= f(T)
\]

\[
\tilde{B} = \text{adj } B = \begin{pmatrix}
T - A_{22} I & A_{21} I \\
A_{12} I & T - A_{11} I
\end{pmatrix}
\]

and \[
\tilde{B} \cdot B = \begin{pmatrix}
\det B & 0 \\
0 & \det B
\end{pmatrix}
\begin{pmatrix}
\det B \alpha_1 \\
\det B \alpha_2
\end{pmatrix}
= \begin{pmatrix}
\det B & 0 \\
0 & \det B
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\alpha_2
\end{pmatrix}
\]
\[
\begin{align*}
\vec{B} B \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} &= \vec{B} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\vec{B} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{align*}
\]

That means

\[\det B \alpha_1 = 0\]

\[\det B \alpha_2 = 0\]

Fix a linear operator \(T\).

Consider the set of all polynomials annihilate \(T\)

\[I = \{ f(x) \text{ such that } f(T) = 0 \}\]

Cayley-Hamilton: the characteristic polynomial is in \(I\).
Now we pick a nonzero polynomial with the minimal degree \( p(x) \) from \( I \).

We can make \( p(x) \) monic by rescaling.

a) \( \deg p \leq n = \dim V \)

Since the characteristic polynomial has degree \( n \), \( p(x) \) has a degree \( \leq n \).

b) Such \( p(x) \) is unique.

If there is another \( p'(x) \) monic with minimal degree

\[
\deg p(x) = \deg p'(x)
\]

\( p - p' \) will have a lower degree.
and \( p - p'(T) = p(T) - p'(T) \)
\[
= 0 - 0
\]
\[
= 0
\]
\( p - p' \in I \)

If we want \( \deg p(x) \) to be minimal among nonzero polynomials from \( I \), then \( p - p' = 0 \)

c) For each nonzero polynomial \( f(x) \in I \), \( p(x) \) divides \( f(x) \).

In general
\[
f(x) = p(x)g(x) + r(x)
\]
\( r(x) \) is the remainder.

such that if \( r(x) \neq 0 \), \( \deg r(x) < \deg p(x) \)
Now $f(T) = 0$
$p(T) = 0$
$\Rightarrow r(T) = 0$

Therefore if $r(x) \neq 0$, it would contradict with the assumption that $p(x)$ is minimal in $I$.

Definition. We call this unique monic polynomial with the minimal degree in $I$ the minimal polynomial of $T$.

In particular, the minimal polynomial divides the characteristic polynomial.
We can similarly define the minimal polynomial for a square matrix.

Theorem: Let $T$ be a linear operator. The characteristic polynomial and the minimal polynomial for $T$ have the same roots, except for multiplicities.

Proof: Let $p(x)$ be the minimal polynomial. We want to show a scalar $\lambda \in F$ is a root of $p$ if and only if $\lambda$ is an eigenvalue of $T$. 
Suppose \( p(\lambda) = 0 \)

Then \( p(x) = (x - \lambda) q(x) \)
for \( q(x) \) with lower degree

So \( q(T) \neq 0 \)

Then we have \( d = q(T) \beta \neq 0 \)

\[
0 = p(T) \beta = (T - \lambda) q(T) \beta
= (T - \lambda) \alpha
\]

\( Td = \lambda \alpha \quad d \neq 0 \)

\( \lambda \) is an eigenvalue

\( \Leftarrow \) Suppose \( \lambda \) is an eigenvalue

There exists \( d \neq 0 \)

\( Td = \lambda \alpha \)

\[
0 = p(T) \alpha = p(\lambda) d
\]

Since \( \alpha \neq 0 \) \( p(\lambda) = 0 \) \( \square \)