Minimal Polynomial

Let $T \in \mathcal{L}(V)$ where $V$ is a finite dimensional vector space.

Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial of degree $n$.

We define $f(T) \in \mathcal{L}(V)$ by $f(T)(v) = a_n T^n v + a_{n-1} T^{n-1} v + \cdots + a_1 T v + a_0 v$.

**Proposition:** Suppose $A \in M_{nn}$. Then for any polynomial $f(x)$ and invertible $B \in M_{nn}$, then $f(B^{-1}AB) = B^{-1} f(A)B$. Furthermore, $f(A) = 0$ iff $f(B^{-1}AB) = 0$.

**Proof:** Let $f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$ be a polynomial of degree $m$. Let $A, B \in M_{nn}$ and assume $B$ is invertible. Let $v \in \mathbb{F}^n$.

$$
 f(B^{-1}AB)v = a_m (B^{-1}AB)^m v + a_{m-1} (B^{-1}AB)^{m-1} v + \cdots + a_1 B^{-1} AB v + a_0 v = 
$$

$$
 = B \left( a_m A^m + a_{m-1} A^{m-1} + \cdots + a_1 A + a_0 I_n \right) B^{-1} v = B^{-1} f(A)B v
$$

So $f(B^{-1}AB) = B^{-1} f(A)B$.

If $f(A) = 0$ then clearly $f(B^{-1}AB) = B^{-1} f(A)B = 0$.

Now suppose $f(B^{-1}AB) = 0$. Then $B^{-1} f(A)B = 0$, which implies that $f(A) = 0$ (left-mult. by $B$ and right-mult. by $B^{-1}$). □

In particular, this theorem implies that the minimal polynomial (to be defined later) for similar matrices is the same, and is hence well-defined for an arbitrary $T \in \mathcal{L}(V)$.

(ELFY) If dim $V = n$ then dim $\mathcal{L}(V) = n^2$ (as a vector space).

Hint: It suffices to show that by fixing a basis $\mathcal{B} = \{v_1, v_2, ..., v_n\}$, the set of $n^2$ linear transformations $\{T_{ij} | i = 1, 2, ..., n \text{ and } j = 1, 2, ..., n\}$ given by $T_{ij}(v_k) = \delta_{kj} v_i$ forms a basis of $\mathcal{L}(V)$. 
Consider the linear operators $I_V, T, T^2, \ldots, T^{n^2} \in \mathcal{L}(V)$. These $n^2 + 1$ linear operators cannot be linearly independent as we have that $\dim \mathcal{L}(V) = n^2$.

Hence there exists scalars $a_0, a_1, \ldots, a_{n^2} \in F$ such that $a_0I_V + a_1T + a_2T^2 + \cdots + a_{n^2}T^{n^2} = 0$.

So there exists a polynomial $f(x)$ of degree at most $n^2$ such that $f(T) = 0$.

Pick $m$ such that $m$ is the smallest positive integer such that $I_V, T, T^2, \ldots, T^m$ are linearly dependent.

Then there exists scalars $a_0, a_1, \ldots, a_m$ such that $a_m \neq 0$ (since otherwise $I_V, T, T^2, \ldots, T^{m-1}$ are linearly dependent) and $a_0I_V + a_1T + a_2T^2 + \cdots + a_mT^m = 0$.

Without loss of generality assume $a_m = 1$ (one can divide both sides of the dependence relation to achieve this). (ELFY) With $a_m = 1$, the other scalars are uniquely determined.

Then there exists a unique monic (leading coefficient of 1) polynomial $m_T(x)$ of minimal degree $m$ such that $m_T(T) = 0$. Such a polynomial is called the minimal polynomial of $T$.

**Proposition (Cayley-Hamilton Theorem):** Let $T \in \mathcal{L}(V)$. Let $c(x)$ be the characteristic polynomial of $T$. Then $c(T) = 0$.

**Proof:** Pick a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ of $V$ such that $[T]_\mathcal{B}$ is upper triangular.

Since $[T]_\mathcal{B}$ is upper triangular, the characteristic polynomial satisfies the factorization $c(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ where $\lambda_i$ are the eigenvalues (counting multiplicity) of $T$, which appear on the main diagonal of $[T]_\mathcal{B}$.

To prove that $c(T) = 0$, it suffices to show that $c(T)v_i = 0$ for each $i = 1, 2, \ldots, n$. To prove this it suffices to show that $(T - \lambda_1I_V)(T - \lambda_2I_V) \cdots (T - \lambda_jI_V)v_j = 0$ for $j = 1, 2, \ldots, n$.

We proceed by induction to prove that $(T - \lambda_1I_V)(T - \lambda_2I_V) \cdots (T - \lambda_jI_V)v_j = 0$ for $j = 1, 2, \ldots, n$.

If $n = 1$, then $Tv_1 = \lambda_1v_1$, and $c(x) = (\lambda_1 - x)$. Clearly $(\lambda_1I_V - T)v_1 = 0$. So the base case is done.
Assume

\[(\lambda_1 I_V - T)\mathbf{v}_1 = 0,\]

\[(\lambda_1 I_V - T)(\lambda_2 I_V - T)\mathbf{v}_2 = 0,\]

\[\vdots\]

\[\vdots\]

\[(T - \lambda_1 I_V)(T - \lambda_2 I_V) \cdots (T - \lambda_{j-1} I_V)\mathbf{v}_{j-1} = 0\]

Since \([T]_B\) is upper triangular it follows that \((T - \lambda_j I_V)\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j-1}\}\) as the \(j\)th through \(n\)th components of \(([T]_B - \lambda_j I_n)e_j\) are 0.

Thus

\[(T - \lambda_1 I_V)(T - \lambda_2 I_V) \cdots (T - \lambda_j I_V)\mathbf{v}_j = 0.\]

So the proof is complete \(\square\)

(ELFY) Let \(T \in \mathcal{L}(V)\). If \(f(x)\) is a polynomial and \(f(T) = 0\) then the minimal polynomial \(m_T(x)\) is a factor of \(f(x)\).

Important consequence: The min. polynomial of \(T\) has maximum degree \(n = \dim V\).
**Proposition:** Let \( T \in \mathcal{L}(V) \). \( \lambda \) is a root of the minimal polynomial \( m_T(x) \) iff \( \lambda \) is an eigenvalue of \( T \).

**Proof:** Let the minimum polynomial be \( m_T(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0 \).

Let \( \lambda \) be a root of \( m_T(x) \). Then \( m_T(x) = (x - \lambda)g(x) \), where \( g(x) \) is a monic polynomial of degree \( m - 1 \).

Because the degree of \( g(x) \) is less than \( m \), there exists a non-zero vector \( \mathbf{v} \in V \) such that \( g(T)\mathbf{v} \neq \mathbf{0} \). However, \( (T - \lambda I_V)g(T)\mathbf{v} = \mathbf{0} \to \lambda \) is an eigenvalue of \( T \).

Now for the reverse direction assume \( \lambda \) is an eigenvalue of \( T \). Then there is a non-zero \( \mathbf{v} \in V \) such that \( T\mathbf{v} = \lambda \mathbf{v} \).

Then \( T^j\mathbf{v} = \lambda^j \mathbf{v} \) for every non-negative integer \( j \).

\[
0 = m_T(T)\mathbf{v} = (T^m + a_{m-1}T^{m-1} + \cdots + a_1T + a_0I_V) \mathbf{v} \to \\
(\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0) \mathbf{v} = m_T(\lambda)\mathbf{v} = \mathbf{0} \to m_T(\lambda) = 0.
\]

So \( \lambda \) is a root of \( m_T(x) \).

Important: The min. poly. must divide the char. poly. and have the same roots.
Examples:

\[
A = \begin{pmatrix}
1 & 1 & 4 \\
0 & 1 & 3 \\
0 & 0 & 2 \\
\end{pmatrix}.
\]

Clearly the characteristic polynomial is \(c(x) = -(x-1)^2(x-2)\).

Then \(c(A) = 0\) as...

\[c(A) = -(A - I_3)^2(A - 2I_3) = -\begin{pmatrix}
0 & 1 & 4 \\
0 & 0 & 3 \\
0 & 0 & 1 \\
\end{pmatrix}^2 \begin{pmatrix}
-1 & 1 & 4 \\
0 & -1 & 3 \\
0 & 0 & 0 \\
\end{pmatrix} = ... = 0.\]

Possible minimal polynomials: \(m_A = (x-1)(x-2)\) or \(m_A = (x-1)^2(x-2)\).

(ELFY) Show the latter is the min poly.

\[
B = \begin{pmatrix}
1 & 0 & 4 \\
0 & 1 & 3 \\
0 & 0 & 2 \\
\end{pmatrix}.
\]

Clearly the characteristic polynomial is \(c(x) = -(x-1)^2(x-2)\).

Then \(c(B) = 0\) as...

\[c(B) = -(B - I_3)^2(B - 2I_3) = -\begin{pmatrix}
0 & 0 & 4 \\
0 & 0 & 3 \\
0 & 0 & 1 \\
\end{pmatrix}^2 \begin{pmatrix}
-1 & 0 & 4 \\
0 & -1 & 3 \\
0 & 0 & 0 \\
\end{pmatrix} = ... = 0.\]

Possible minimal polynomials: \(m_B = (x-1)(x-2)\) or \(m_B = (x-1)^2(x-2)\).

(ELFY) Show the former is min. poly.