6. Spectral theory:

Def: Let $A$ be an $n \times n$ matrix. A vector $v$ satisfying $Av = \lambda v$ for some $\lambda \in \mathbb{K}$, is called an eigenvector of $A$; $\lambda$ is called an eigenvalue of $A$.

Throughout, we'll assume that our field $\mathbb{K}$ is algebraically closed, i.e., every polynomial in $\mathbb{K}[x]$ has a root in $\mathbb{K}$.

The most common algebraically closed field is $\mathbb{K} = \mathbb{C}$.

Prop: $A$ has an eigenvector.

Proof: Pick any $0 \neq w \in \mathbb{C}^n$, consider the following:
$$w, \, Aw, \, A^2w, \ldots, \, A^nw.$$ Since $\dim \mathbb{C}^n = n$, these vectors are linearly dependent. Thus, we can write $0 = c_0w + c_1Aw + \ldots + c_nA^nw$
$$= p(A)w$$
where $p(x) = c_0 + c_1x + \ldots + c_nx^n \in \mathbb{K}[x]$.

Since $\mathbb{K}$ is closed, $p(x) = c \prod_{j=1}^{n} (x - \lambda_j), \quad c \neq 0$
and so $p(A)w = c \prod_{j=1}^{n} (A - \lambda_j I)w = 0$.

Now, one of $A - \lambda_jI$ must be non-invertible. (Because
\( p(A) \) is non-invertible. Suppose \( A - dI \) is non-invertible, and pick \( v \neq 0 \) in the nullspace of \( A - dI \).

Then, \((A - dI)v = 0 \Rightarrow Av - d v = 0 \Rightarrow Av = dv. \)

Remark: By Corollary to Theorem 5.7, \( A - dI \) is non-invertible iff \( \det(A - dI) = 0 \). Thus, \( \lambda \) is an eigenvalue of \( A \) iff \( \det(A - dI) = 0 \), and this is how we find all eigenvalues of \( A \).

Example: \( A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \).

\[ \det(A - dI) = \det \begin{pmatrix} 3 - d & 2 \\ 1 & 4 - d \end{pmatrix} = (3 - d)(4 - d) - 2 \\ = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5). \]

Thus, \( A \) has two eigenvalues: \( \lambda_1 = 2, \lambda_2 = 5 \).

Now, let's find the eigenvectors:

\( \lambda_1 = 2 \): Find \( v_1 \) such that \((A - 2I)v_1 = 0\).

\((A - 2I)v = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_1 + 2x_2 = 0 \Rightarrow x_1 = -2x_2 \)

Thus, \( v_1 = \begin{pmatrix} -2c \\ c \end{pmatrix} \) is an eigenvector for any \( c \).

\( \lambda_2 = 5 \): Find \( v_2 \) such that \((A - 5I)v_2 = 0\).

\((A - 5I)v = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow -2x_1 + 2x_2 = 0 \Rightarrow x_1 = x_2. \)

Thus, \( v_2 = \begin{pmatrix} -c \\ c \end{pmatrix} \) is an eigenvector for any \( c \).

We'll say \( A \) has eigenvalues \( \lambda_1 = 2, \lambda_2 = 5 \), eigenvectors \( v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \).
Here, \( v_1 \) and \( v_2 \) are linearly independent. Thus, for any \( x \in \mathbb{R}^2 \),
we can write \( x = a_1 v_1 + a_2 v_2 \).

Consider \( A^N \) for large \( N \).
\[
A^N x = A^N (a_1 v_1 + a_2 v_2) = a_1 A^N v_1 + a_2 A^N v_2 = a_1 \lambda_1^N v_1 + a_2 \lambda_2^N v_2 = 2^N a_1 v_1 + 5^N a_2 v_2.
\]
Since \( 2^N \) and \( 5^N \to \infty \) as \( N \to \infty \), it makes sense to say that \( A^N x \to \infty \) as \( N \to \infty \).

Note: The entries in \( A^N \) grow asymptotically as \( \sim 5^N \), the largest eigenvalue.

Def: The characteristic polynomial of an \( n \times n \) matrix \( A \) is
\[
\rho_A(s) = \det(sI - A).
\]

Remarks: \( \rho_A(s) \) has degree \( n \), and its roots are the eigenvalues of \( A \). Moreover, if \( K \) is closed (e.g. \( K = \mathbb{C} \)), then all \( n \) roots lie in \( K \).

Theorem 6.1: Eigenvectors of \( A \) corresponding to distinct eigenvalues are linearly independent.

Proof: Let \( \lambda_1, \ldots, \lambda_k \) be pairwise distinct eigenvalues, with
eigenvectors \( v_1, \ldots, v_k \) (all non-zero).

Suppose \( \sum_{j=1}^{n} c_j v_j = 0 \), where \( m \) is minimal, non-zero.
(\( m \) clearly, \( c_j \neq 0 \)).
Apply \( A \): \[ c_1 v_1 + \cdots + c_m v_m = 0 \]
\[ A(v_1) + \cdots + A(v_m) = 0 \]
\[ \Rightarrow c_1 A v_1 + \cdots + c_m A v_m = 0 \]
\[ \Rightarrow c_1 \lambda v_1 + \cdots + c_m \lambda v_m = 0 \]

We now have \( \sum_{j=1}^{m} c_j v_j = 0 \) and \( \sum_{j=1}^{m} c_j \lambda_j v_j = 0 \).

Thus, \( (\lambda_m \sum_{j=1}^{m} c_j v_j) - (\sum_{j=1}^{m} c_j \lambda_j v_j) = \sum_{j=1}^{m-1} (c_j \lambda_m - c_j \lambda_j) v_j = 0 \).

This contradicts minimality of \( m \).

Thus, \( v_1, \ldots, v_m \) must be linearly independent. \( \square \)

**Corollary 6.2**: If \( A \) has \( n \) distinct eigenvalues, then it has \( n \) linearly independent eigenvectors.

In this case, the eigenvectors form a basis for \( X \), and it is easy to compute \( A^n x \), for any \( x \in X \):

Write \( x = \sum_{j=1}^{n} a_j v_j \) \( \Rightarrow \) eigenvectors \( v_1, \ldots, v_n \).

\[ A^n x = \sum_{j=1}^{n} a_j A^n v_j = \sum_{j=1}^{n} a_j \lambda_j^n v_j \]

**Theorem 6.3**: If the eigenvalues of \( A \) are \( \lambda_1, \ldots, \lambda_n \), then
\[ \sum_{i=1}^{n} \lambda_i = \text{tr} A \quad \text{and} \quad \prod_{i=1}^{n} \lambda_i = \det A \]

**Proof**: Claim: \( p_A(s) = s^n - (\text{tr} A) s^{n-1} + \cdots + (-1)^n \det A \).

Write \( p_A(s) = \prod_{i=1}^{n} (s - \lambda_i) \).

Note: \( s^{n-1} \) coefficient = \( -\sum_{i=1}^{n} \lambda_i \), constant term = \( (-1)^n \prod_{i=1}^{n} \lambda_i \).
To prove our claim, compute

\[ p_n(t) = \det(sI - A) = \det \begin{pmatrix} s - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & s - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & s - a_{nn} \end{pmatrix} \]

Recall that \[ \det A = \sum_{\pi \in S_n} \text{sgn}(\pi) a_{\pi(1),1} \cdots a_{\pi(n),n} \]

Thus, \[ \det(sI - A) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^n (s \delta_{\pi(i),i} - a_{\pi(i),i}) \]

Clearly, the \((n-1)\)-coefficient is \[ -\sum_{i=1}^n a_{ii} = tr(A) \quad \checkmark \]

and the constant term is \[ \det(-A) = (-1)^n \det(A). \]

**Remark:** If \( Av = \lambda v \), then \( A^2v = \lambda^2v \). Thus, if \( \lambda \) is an eigenvalue of \( A \), then \( \lambda^n \) is an eigenvalue of \( A^n \).

Let's take this further. Let \( q(s) \in K[s] \) be any polynomial, say \( q(s) = \sum_{i=1}^n a_is^i \).

If \( Av = \lambda v \), then \( A^iv = \lambda^iv \)

\[ \Rightarrow q(A)v = \sum_{i=1}^n a_i A^iv = \sum_{i=1}^n a_i \lambda^iv = q(\lambda)v. \]

Thus, \( q(\lambda) \) is an eigenvalue of \( q(A) \). In fact, the converse holds too:

**Theorem 6.9:** ("Spectral mapping theorem"). Let \( A \) have eigenvalue \( \lambda \), and let \( q(s) \in K[s] \).

(a) \( q(\lambda) \) is an eigenvalue of \( q(A) \).

(b) Conversely, every eigenvalue of \( q(A) \) is of the form \( q(\lambda) \).
Proof: (a) We just did this.

(b) Let $\mu$ be an eigenvalue of $g(A)$ \iff $\det(g(A) - \mu I) = 0$.

Consider $g(s) - \mu = c \prod_{i=1}^{n} (s - r_i), \quad r_i \in \mathbb{C}$.

and $g(A) - \mu I = c \prod_{i=1}^{n} (A - r_i I)$

Since $g(A) - \mu I$ is not invertible, one of $A - r_i I$ is not invertible \iff some $r_i$ is an eigenvalue of $A$.

Since $r_i$ is a root of $g(s) - \mu$, $g(r_i) = \mu$.

\[ \square \]

Remark: In the case when $g(s) = p_A(s)$, we conclude that all eigenvalues of $p_A(A)$ are zero. Actually, even more is true.

\textbf{Theorem 6.5 (Cayley-Hamilton theorem).} Every matrix satisfies its characteristic polynomial: $p_A(A) = 0$.

\textbf{Proof:} Case 1: All eigenvalues are distinct.

By Theorem 6.2, $A$ has $n$ linearly independent eigenvectors $v_1, \ldots, v_n$. Each eigenvalue $\lambda_i$ is a root of $p_A(s)$.

Thus, for any $x \epsilon \mathbb{R}^n$, write $x = c_1 v_1 + \cdots + c_n v_n$.

$$p_A(A)x = \sum_{i=1}^{n} p_A(A) c_i v_i = \sum_{i=1}^{n} p_A(\lambda_i) c_i v_i = \sum_{i=1}^{n} 0 = 0.$$  \[ \checkmark \]

For the general case (non-distinct eigenvalues), we need an additional lemma:
Lemma 6.6: Let $P$ and $Q$ be polynomials with matrix coefficients:

$$P(t) = \sum P_j t^j, \quad Q(s) = \sum Q_k s^k,$$

and let $R = PQ$.

Then, $R(t) = \sum R_k t^k$ with $R_k = \sum_{j+k=k} P_j Q_k$.

Moreover, if $A$ commutes with the $Q_k$'s, then $P(A)Q(A) = R(A)$.

Proof: Exercise.

Now, let $Q(s) = s I - A$, $P(s) = (P_{ij}(s))$, $P_{ij}(s) = (-1)^{ij} D_{ij}(s)$

where $D_{ij}(s)$ is the determinant of the $ij$th minor of $Q(s)$.

Recall Theorem 5.12, the formula for a matrix inverse:

$$(Q^{-1})_{ki} = (-1)^{i+k} \frac{\det Q_{ik}}{\det Q}$$

In our context, this means that $(Q(s))^{-1} = \frac{1}{\det Q(s)} P(s)$.

Put $R(s) = P(s)Q(s) = (\det Q(s)) I = P_A(s) I$.

Clearly, $A$ commutes with the coefficients of $Q(s)$, and $Q(A)=0$.

By Lemma 6.6, $R(A) = P(A)Q(A) = P_A(A) I = 0 \Rightarrow P_A(A)=0$.

Example:

(i) $A = I$, then $P_A(s) = \det (sI-I) = (s-1)^n$

$\Rightarrow \lambda = 1$ is an eigenvalue with multiplicity $n$.

$A-I=0$, so $(A-I)v=0$ for all $v$.

Thus, every vector is an eigenvector of $A$. 

(2) \( A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix} \). \( \text{tr } A = 2 \), \( \det A = 1 \), so
\[ p_A(s) = s^2 - 2s + 1 = (s - 1)^2, \] so \( \lambda_1 = \lambda_2 = 1 \).

To find the eigenvectors: \((A - I)v = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}v = 0 \Rightarrow x_1 + x_2 = 0 \Rightarrow v = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \) is an eigenvector (and every multiple is too). However, this is the only eigenvector.

Prop: If \( A \) has only one eigenvalue \( \lambda \), and \( n \) linearly independent eigenvectors, then \( A = \lambda I \).

Proof: Pick \( x \in \mathbb{R}^n \), and write \( x = a_1 x_1 + \ldots + a_n x_n \).

\[ Ax = a_1 Ax_1 + \ldots + a_n Ax_n = a_1 \lambda x_1 + \ldots + a_n \lambda x_n = \lambda (a_1 x_1 + \ldots + a_n x_n) = \lambda x. \]

Remark: Every 2x2 matrix with \( \text{tr } A = 2 \), \( \det A = 1 \), has \( \lambda = 1 \) as a double root of \( p_A(s) \). These matrices form a 2-parameter family, and only \( A = I \) has 2 linearly independent eigenvectors.

In cases like these, we have a notion of "generalized eigenvectors."

Suppose \( \lambda \) is an eigenvalue with multiplicity \( m \), but only one eigenvector, \( v_1 \).

Then \((A - \lambda I)v_1 = 0\).

Since \( \text{rank } (A - \lambda I) = m - 1 \), there is some \( v_2 \) such that
\[ (A - \lambda I)v_2 = v_1 \Rightarrow (A - \lambda I)v_2 = 0. \]
Similarly, we can find \( v_3 \) such that
\[(A-\lambda I)v_3 = v_2 \Rightarrow (A-\lambda I)^2v_3 \neq 0 \text{ but } (A-\lambda I)^3v_3 = 0.\]

Picture of this: \[V_m \rightarrow A-\lambda I \rightarrow \ldots \rightarrow V_3 \rightarrow A-\lambda I \rightarrow V_2 \rightarrow A-\lambda I \rightarrow V_1 \rightarrow A-\lambda I \rightarrow 0\]

Def: The **algebraic multiplicity** of an eigenvalue \( \lambda \) is the largest \( m \) such that \((s-\lambda)^m\) appears as a factor of \( p_A(s)\).

The **geometric multiplicity** of \( \lambda \) is the number of linearly independent eigenvectors it has, or equivalently, the rank of the nullspace of \( A-\lambda I \).

Def: A vector \( v \) is a **generalized eigenvector** of \( A \) with eigenvalue \( \lambda \) if \((A-\lambda I)^m v = 0\) for some \( m \in \mathbb{N}\).

Example: \( A = \begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix} \), which has \( \lambda_1 = \lambda_2 = 1 \), \( v = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \).

To find a generalized eigenvector \( v_2 \), we need to solve
\[(A-\lambda I)v_2 = v_1 \Rightarrow \begin{pmatrix} 2 & 2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \]

\[\Rightarrow \begin{cases} 2x_1 + 2x_2 = -1 \\ -2x_1 - 2x_2 = 1 \end{cases} \Rightarrow 2x_1 + 2x_2 = -1 \Rightarrow x_2 = \frac{-1 - x_1}{2}

So, \( v = \begin{pmatrix} c \\ -\frac{1}{2} - c \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix} + \begin{pmatrix} c \\ c \end{pmatrix} \) is a generalized eigenvector.

For convenience, pick \( c = 0 \). We have: \( \begin{pmatrix} 0 \\ -1/2 \end{pmatrix} \xrightarrow{A-\lambda I} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \xrightarrow{A-\lambda I} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \).
Example: Suppose $A$ is $11 \times 11$ with an eigenvalue $\lambda$ of algebraic multiplicity 11, and geometric multiplicity 4. [So $\dim (N - \lambda I) = 4$].

The following is one possibility for the generalized eigenvectors:

These rows are called "invariant subspace":

\[
\begin{align*}
V_5 & \rightarrow A - \lambda I \rightarrow V_4 \rightarrow A - \lambda I \rightarrow V_3 \rightarrow A - \lambda I \rightarrow V_2 \rightarrow A - \lambda I \rightarrow V_1 \rightarrow A - \lambda I \rightarrow 0 \\
W_3 & \rightarrow W_2 \rightarrow W_1 \rightarrow 0 \\
x_2 & \rightarrow x_1 \rightarrow 0 \\
y_1 & \rightarrow 0
\end{align*}
\]

Remarks:

$N_1 := N_{A - \lambda I} = \text{Span} \{ V_5, W_1, x_1, y_1 \}$  \hspace{1cm} \text{dim } N_2 = 4

$N_2 := (A - \lambda I)^2 = \text{Span} \{ V_5, W_2, x_2, V_1, W_1, x_1, y_1 \}$  \hspace{1cm} \text{dim } N_2 = 7

$N_3 := (A - \lambda I)^3 = \text{Span} \{ V_5, W_3, \ldots, x, y \}$  \hspace{1cm} \text{dim } N_3 = 9

$\vdots$

Note that: $N_1 \subseteq N_2 \subseteq N_3 \subseteq N_4 \subseteq N_5 = N_6 = \ldots$

$\dim N_1 = 4 < 7 < 9 < 10 < 11 = 11 = \ldots$

It's a fundamental result that there will always be a full set of generalized eigenvectors that form a basis for $C^n$. This is the Spectral Theorem.
Theorem 6.7: (Spectral Theorem) Let $A$ be an $n \times n$ matrix over $\mathbb{C}$.

Then $\mathbb{C}^n$ has a basis of eigenvectors (genuine or generalized) of $A$.

To prove Theorem 6.7, we need some algebraic results first.

Lemma 6.8: Let $p, q \in \mathbb{C}[x]$ with no common roots. Then we can write $ap + bq = 1$ for some other $a, b \in \mathbb{C}[x]$.

Remark: This is by the division algorithm. If these are integers, then we can write, $m = q \cdot r + r$, $r < n$. [e.g., $49 = 9 \cdot 5 + 4$]

$q$ is the quotient, $r$ is the remainder.

Proof: Let $I = \{ap + bq : a, b \in \mathbb{C}[x]\}$, the ideal generated by $p$ and $q$.

Pick $d \in I$ with minimal degree.

Claim 1: $d \mid p$ and $d \mid q$.

Suppose it did not; say $d \nmid p$.

By division algorithm, write $p = md + r$ with $\deg r < \deg d$.

Since $p, d \in I$, $r = p - md \in I$. But $d$ had minimal degree.

Claim 2: $\deg d = 0$.

If not, it would have a root $x$, and since $d \mid p$ and $d \mid q$,

then $(x - d)$ divides $p$ and $q$.

Thus, $d$ is constant; we may assume $1$ since we're over $\mathbb{C}$. □
Lemma 6.9: Let $A$ be an $n \times n$ matrix over $\mathbb{C}$, $\mathfrak{p}, \mathfrak{q} \in \mathcal{C}[s]$ with no common roots. Let $N_{\mathfrak{p}}, N_{\mathfrak{q}}, N_{\mathfrak{pq}}$ be the nullspaces of $\mathfrak{p}(A), \mathfrak{q}(A),$ and $\mathfrak{p}(A)\mathfrak{q}(A)$, respectively. Then $N_{\mathfrak{pq}} = N_{\mathfrak{p}} \oplus N_{\mathfrak{q}}$.

Proof: Write $ap + bq = 1$ for $a, b \in \mathcal{C}[s]$.

Plug in $A$: $a(\mathfrak{p})(\mathfrak{p}(A)) + b(\mathfrak{q})(\mathfrak{q}(A)) = I$.

Multiply by $x \in N_{\mathfrak{pq}}$: $\underbrace{a(\mathfrak{p})(\mathfrak{p}(A))x + b(\mathfrak{q})(\mathfrak{q}(A))x} = x$.

In $N_{\mathfrak{p}}$ because $b(\mathfrak{q})(\mathfrak{q}(A))x = 0$.

In $N_{\mathfrak{q}}$ because $a(\mathfrak{p})(\mathfrak{p}(A))x = 0$.

[Here, we're using that $f(\mathfrak{p})g(\mathfrak{q}) = g(\mathfrak{q})f(\mathfrak{p}) \neq \pm f, g \in \mathcal{C}[s].$]

The expression $(\star)$ is $x = x_{\mathfrak{p}} + x_{\mathfrak{q}}$.

$b(\mathfrak{q})(\mathfrak{q}(A))x + a(\mathfrak{p})(\mathfrak{p}(A))x$.

This shows $N_{\mathfrak{pq}} = N_{\mathfrak{p}} \oplus N_{\mathfrak{q}}$. To show $\oplus$, we need uniqueness.

Suppose $x = x_{\mathfrak{p}} + x_{\mathfrak{q}} = x_{\mathfrak{p}}' + x_{\mathfrak{q}}'$. Put $y := x_{\mathfrak{p}} - x_{\mathfrak{p}}' = x_{\mathfrak{q}} - x_{\mathfrak{q}}' \in N_{\mathfrak{p}} \cap N_{\mathfrak{q}}$.

Clearly, $y \in N_{\mathfrak{pq}}$, so $y = Iy = \left[a(\mathfrak{p})(\mathfrak{p}(A)) + b(\mathfrak{q})(\mathfrak{q}(A))\right]y = 0$.

$\Rightarrow y = 0$.

Thus, $N_{\mathfrak{pq}} = N_{\mathfrak{p}} \oplus N_{\mathfrak{q}}$. $\square$
Corollary 6.10: Let \( P_{1}, \ldots, P_{k} \in \mathbb{C}(x) \) be pairwise coprime (no common roots). Let \( N_{P_{1}} \ldots P_{k} \) be the nullspace of \( P_{1}(A) \ldots P_{k}(A) \).

Then \( N_{P_{1}} \ldots P_{k} = N_{P_{1}} \oplus \cdots \oplus N_{P_{k}} \).

Proof: Exercise. (Induct on \( k \)).

Proof of Spectral Theorem: Pick \( x \in \mathbb{C}^{n} \).

Write \( P_{A}(A)x = \prod_{j=1}^{R} (A - \lambda_{j}I)^{n_{j}}x = 0 \).

Remove all factors \( A - \lambda_{j}I \) that are invertible, so we're left with a polynomial \( m(A)x = \prod_{j=1}^{R} (A - \lambda_{j}I)^{n_{j}}x = 0 \), each \( \lambda_{j} \) is e-value.

\[
= \prod_{j=1}^{R} (A - \lambda_{j}I)^{n_{j}}x = 0 \quad \text{with } \quad P_{j}(A)
\]

Remarks: 
- \( P_{j}(s) = (s - \lambda_{j})^{n_{j}} \) and \( \lambda_{i} \neq \lambda_{j} \)
- The \( x \) above is in \( N_{P_{1}} \ldots P_{k} = N_{P_{1}} \oplus \cdots \oplus N_{P_{k}} \).
- If \( x = x_{P_{1}} + \cdots + x_{P_{k}} \), with \( x_{P_{i}} \in N_{P_{i}} \), then each \( x_{P_{i}} \) is a generalized eigenvector: \( (A - \lambda_{i}I)^{n_{j}}x = 0 \). \( \square \)
Let \( I = \text{Im} \) be the set of polynomials \( p(x) \in \mathbb{C}[x] \) s.t. \( p(A) = 0 \).

Note that \( I \) is closed under addition & multiplication (of not just scalars, but polynomials too.)

**Lemma:** \( I \) contains a unique monic polynomial \( m = m_A \) of minimal degree, and all other polynomials in \( I \) are scalar multiples of \( m_A \) (i.e., \( I = \langle m_A \rangle \) is a principal ideal of \( \mathbb{C}[x] \).)

**Proof:** Let \( m \in I \) have minimal degree.

**Uniqueness:** Clear. [If there were 2, subtract them.]

**Existence:** Suppose \( p \in I \) were not a multiple of \( m \).

By division algorithm, write \( p = qm + r \), \( \deg r < \deg m \).

Then \( r = p - qm \in I \). \( \square \)

**Def:** The **minimal polynomial** of a matrix \( A \), denoted \( m_A \), is the unique monic polynomial of minimal degree for which \( m_A(A) = 0 \).

Let \( N_m = N_{m_A}(\lambda) \) be the nullspace of \( (A - \lambda I)^m \).
Note that $N_m$ consists of generalized eigenvectors, and

$$N_1 \subset N_2 \subset \ldots \subset N_d = N_{d+1} = \ldots$$

for some index $d$. Let $d = d(\lambda)$ be the minimal index such that

$$N_{d-1} \subset N_d = N_{d+1},$$

called the index of the eigenvalue $\lambda$.

Theorem 6.11: If $A$ is $n \times n$ and has distinct eigenvalues $\lambda_1, \ldots, \lambda_d$ with indices $d_1, \ldots, d_k$, then its minimal polynomial is

$$m_A(t) = \prod_{i=1}^{d} (t - \lambda_i)^{d_i}.$$ 

Proof: Exercise.

Denote $N_j(\lambda_j)$ by $N^{(j)}$. The spectral theorem can be stated as follows:

$$\mathbb{C}^n = N^{(1)} \oplus N^{(2)} \oplus \ldots \oplus N^{(k)}.$$ 

Remark: $\dim N^{(j)}$ is the algebraic multiplicity of $\lambda_j$ (this will be proved later).

Note that $A$ maps $N^{(j)}$ into itself. We call such a subspace invariant under $A$.

It turns out that $A$ (up to choice of basis) is completely determined by the dimensions of $N_1(\lambda), \ldots, N_d(\lambda)$ for each $\lambda$. 
**Theorem 6.12:** Two matrices $A, B$ are similar if and only if they have the same eigenvalues, and the dimensions of the corresponding eigenspaces are the same. That is, if for each eigenvalue $\lambda_j$, $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$, where $N_m(\lambda_j) = \text{nullspace of } (A - \lambda_j I)^m$, $M_m(\lambda_j) = \text{nullspace of } (B - \lambda_j I)^m$.

**Proof:** $\Rightarrow$ If $A = S^{-1}BS$, then $(A - \lambda I)^m = S^{-1}(B - \lambda I)^m S$.

Therefore, $(A - \lambda I)^m$ and $(B - \lambda I)^m$ have the same nullity.

$\Leftarrow$ Let $\lambda = \lambda_j$ be an eigenvalue of $A$, and $N_i := \text{nullspace of } (A - \lambda I)^i$.

**Goal:** Construct a basis for $N_d$ under which $A - \lambda I$ admits a nice matrix form (the "Jordan Canonical Form").

**Recall:** $N_{d+1} \supseteq N_d \supseteq N_{d-1} \supseteq \cdots \supseteq N_2 \supseteq N_1 \supseteq N_0 = 0$.

**Lemma:** The map $A - \lambda I$ carries over to a well-defined map on the quotient spaces: $A - \lambda I : \frac{N_{i+1}}{N_i} \longrightarrow \frac{N_i}{N_{i-1}}$.

Moreover, it is injective.

**Proof:** Exercise (HW).
By lemma 6.13, \( \dim(N_{i+1}/N_i) \leq \dim(N_i/N_{i-1}) \).

We will construct our basis for \( Nd \) in "batches."

Let \( \overline{x}_{1}, \ldots, \overline{x}_{k} \) be a basis for \( N_d/N_{d-1} \) (so \( x_1, \ldots, x_k \) lin. indep. in \( N_d \)).

By lemma, \( (A-\lambda I)\overline{x}_1, \ldots, (A-\lambda I)\overline{x}_k \) are linearly independent in \( N_{d-1}/N_{d-2} \).

Extend to a basis \( \overline{x}_1', \ldots, \overline{x}_k', \overline{x}_{k+1}', \ldots, \overline{x}_l' \) of \( N_{d-1}/N_{d-2} \).

Repeat this process:

\( (A-\lambda I)\overline{x}_1', \ldots, (A-\lambda I)\overline{x}_k' \) are linearly independent in \( N_{d-1}/N_{d-2} \).

Picture of this:

\[
\begin{align*}
X_1 & \xrightarrow{A-\lambda I} X_1' \\
\vdots & \qquad \vdots \\
X_{k} & \xrightarrow{A-\lambda I} X_k' \\
X_{k+1} & \xrightarrow{A-\lambda I} X_{k+1}' \\
& \quad \vdots \\
X_l & \xrightarrow{A-\lambda I} X_l'
\end{align*}
\]

\[
\begin{align*}
A-\lambda I & \quad X_1^{(d)} \quad A-\lambda I \quad 0 \\
\vdots & \quad \vdots \\
A-\lambda I & \quad X_{k+1}^{(d)} \quad 0 \\
A-\lambda I & \quad X_{k+1}^{(d)} \quad 0 \\
\vdots & \quad \vdots \\
A-\lambda I & \quad X_{l}^{(d)} \quad 0 \\
\vdots & \quad \vdots \\
A-\lambda I & \quad X_{l}^{(d)} \quad 0 \\
\vdots & \quad \vdots \\
A-\lambda I & \quad X_{l}^{(d)} \quad 0
\end{align*}
\]
Remarks:

- $N_d(\lambda) =$ space spanned by all of the vectors
  = set of generalized eigenvectors for $\lambda$.
- Algebraic multiplicity of $\lambda = \dim N_d(\lambda) =$ total # vectors spanned.
- Geometric multiplicity of $\lambda = \dim N_1(\lambda) =$ # of rows
  = # of linearly independent eigenvectors for $\lambda$.
- Index of $\lambda =$ length of longest row.
- Each "row" of vectors spans an invariant subspace of $A - \lambda I$.

- The matrix $A - \lambda I$ restricted to this
  subspace has the form:

$$
\begin{bmatrix}
0 & \cdots & 0 \\
0 & \ddots & \vdots \\
0 & \cdots & 0
\end{bmatrix}
$$

- The matrix $A$ restricted to this
  subspace has the form, called a Jordan block.

Reason: If $X_d \xrightarrow{A - \lambda I} X_{d-1} \xrightarrow{A - \lambda I} \cdots X_2 \xrightarrow{A - \lambda I} X_1 \xrightarrow{A - \lambda I} 0$
then write basis $X_1, \ldots, X_d$,

$$(A - \lambda I) X_j = X_{j-1} \Rightarrow A X_j = \lambda X_j + X_{j-1} \Rightarrow \text{row } j \text{ is}
\begin{bmatrix}
\lambda \\
1
\end{bmatrix}$$

If we use a basis of generalized eigenvectors for $C^\perp$, then
the matrix for $A$ is block-diagonal, consisting of Jordan blocks.
Such a matrix is called the Jordan canonical form of $A$. $J = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_k \end{bmatrix}$.

Since it depends only on the eigenvalues and eigenspace dimensions, if two matrices $A$ and $B$ have the same eigenvalues and $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$ as in Theorem 6.12, then they must be similar to the same "Jordan matrix."

The following is a generalization of the spectral mapping theorem:

**Theorem 6.14**: Let $A, B : X \rightarrow X$ be commuting maps, $\dim X < \infty$.

Then there is a basis for $X$ consisting of eigenvectors and generalized eigenvectors of $A$ and $B$.

**Proof**: Write $X = N^{(1)} \oplus \cdots \oplus N^{(k)}$, where each summand is a generalized eigenspace $N^{(i)} = N_{d_j}(\lambda_j) = \text{null space } (A - \lambda_j I)^{d_j}$.

**Claim**: $B$ maps $N^{(i)}$ into $N^{(i)}$.

To show this, let $d = d_j$ and $\lambda = \lambda_j$. For a gen. eigenvector $x$,

$$0 = (A - \lambda I)^d x = B(A - \lambda I)^d x = (A - \lambda I)^d B x \Rightarrow B x \in N^{(i)}.$$ 

Now apply the spectral theorem to $B$, restricted to each $N^{(i)}$ separately.
Conclusion: \( B \mid_{N(i)} : N(i) \to N(i) \) and by the spectral theorem, \( N(i) \) has a basis of generalized eigenvectors of \( B \). But there are also generalized eigenvectors of \( A \) for \( \lambda \).

\[ \square \]

**Corollary 6.15:** Theorem 6.14 remains true for any number (even infinite) of pairwise commuting maps.

**Proof:** Exercise.

**Theorem 6.16:** Every square matrix \( A \) is similar to its transpose.

**Proof:** Let \( A : X \to X \) be linear and \( A' : X' \to X' \) its transpose.

Note that \( (A-\lambda I)' = A' - \lambda I' \).

Thus, \( A \) and \( A' \) have the same eigenvalues, and the eigenspaces have the same dimension.

The transpose of \( (A-\lambda I)^j \) is \( (A'-\lambda I)^j \), thus their nullspaces have the same dimension.

Theorem 6.12 now implies that \( A \) and \( A' \) are similar.

\[ \square \]

**Theorem 6.17:** Let \( X \) be a finite-dimensional space over \( \mathbb{C} \), and \( A : X \to X \) linear. Let \( \lambda \neq \lambda' \) be eigenvalues of \( A \) (and thus also of \( A' \)). If \( Av = \lambda v \) and \( A' \ell = \lambda' \ell \), then \((\ell, x) = 0\).
Proof: By assumption, $A v = \lambda v$ and $A' e_i = \lambda_i e_i$

$\Rightarrow \lambda (e_i, v) = (e_i, \lambda v) = (e_i, A v) = (e_i, A' e_i, v) = \lambda_i (e_i, v)$

Since $\lambda \neq \lambda_i$, $(e_i, v) = 0$.

Application of Theorem 6.17:

Theorem 6.18: Suppose $A$ has distinct eigenvalues $\lambda_1, ..., \lambda_n$ and corresponding eigenvectors $v_1, ..., v_n \in \mathbb{C}^n$ and let $e_1, ..., e_n$ be the corresponding eigenvectors in $A'$.

Then: (a) $(e_i, v_i) \neq 0$ for each $i$

(b) If $x = \sum_{i=1}^{n} a_i v_i$, then $a_i = \frac{(e_i, x)}{(e_i, v_i)}$.

Def: When $A$ has linearly independent eigenvectors $v_1, ..., v_n$, we say that $A$ is diagonalizable, because its Jordan canonical form is a diagonal matrix $D$. In this case, we can write $A = P^{-1}DP$, or equivalently, $D = PAP^{-1}$.

The matrix $D$ has the eigenvalues down the diagonal, and the columns of $P$ are the corresponding eigenvectors, i.e., $D = (\lambda_1 e_1, ..., \lambda_n e_n)$, $P = (v_1, ..., v_n)$.
To see this, note that

\[ A \rho = A(u_1, \ldots, u_n) = (Av_1, \ldots, Av_n) = (\lambda_1 v_1, \ldots, \lambda_n v_n) = (\lambda_1 p e_1, \ldots, \lambda_n p e_n) = \rho (\lambda_1 e_1, \ldots, \lambda_n e_n) = \rho D. \]

**Diagram:**

\[ \begin{align*}
\mathbb{R}^n & \xrightarrow{A} \mathbb{R}^n \\
\rho & \uparrow \\
\mathbb{R}^n & \xrightarrow{D} \mathbb{R}^n
\end{align*} \]

**Example:**

\[
\begin{bmatrix}
3 & -2 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
0 & 2
\end{bmatrix} \begin{bmatrix}
1 & 2 \\
1 & 1
\end{bmatrix}^{-1} \implies \lambda_1 = 1, \quad v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},
\]
\[
\lambda_2 = 2, \quad v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}
\]
Application to differential equations

1. Consider a system of n linear ODEs: \( \dot{\mathbf{x}} = A \mathbf{x} \).

   Suppose \( A \) has eigenvalues \( \lambda_1, \ldots, \lambda_n \) and \( n \) linearly independent eigenvectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \).

   Note: \( \mathbf{x}_i(t) = e^{\lambda_i t} \mathbf{v}_i \) is a solution (easy to check this).

   Solutions to \( \dot{\mathbf{x}} = A \mathbf{x} \) are vectors in the nullspace of \( \frac{d}{dt} - A \).

   It's well-known that the nullspace is \( n \)-dimensional.

   Thus, the general solution is \( \mathbf{x}(t) = C_1 e^{\lambda_1 t} \mathbf{v}_1 + \cdots + C_n e^{\lambda_n t} \mathbf{v}_n \).

   In matrix form, this is
   \[
   \begin{bmatrix}
   e^{\lambda_1 t} & 0 & \cdots & 0 \\
   0 & e^{\lambda_2 t} & \cdots & 0 \\
   \vdots & \vdots & \ddots & \vdots \\
   0 & 0 & \cdots & e^{\lambda_n t}
   \end{bmatrix}
   \begin{bmatrix}
   C_1 \\
   C_2 \\
   \vdots \\
   C_n
   \end{bmatrix}
   = e^{Dt} \mathbf{x}_0
   \]

   Here, \( \mathbf{x}_0 = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} \) and we're using basis \( \mathbf{v}_1, \ldots, \mathbf{v}_n \).

   With respect to the basis \( e_1, \ldots, e_n \), \( e^{Dt} \mathbf{x}_0 \) becomes
   \[
   e^{At} \mathbf{x}_0 = e^{P^{-1} D P t} P e_0 = (P e^{D t} P) \mathbf{x}_0,
   \]

   While it may seem that \( e^{At} = \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} \) is hard to compute,

   \( e^{Dt} \) and \( P^{-1} e^{Dt} P \) are easy to compute.
In summary, if $A$ has $n$ linearly independent eigenvectors, then the general solution to $\dot{x} = Ax$, $x(0) = x_0$ is

$$\dot{x}(t) = e^{At}x_0 = p^T e^{Dt} p x_0,$$


(2) Consider

$$\begin{cases}
\dot{x}_1 &= -x_1 - x_2 \\
\dot{x}_2 &= x_1 - 3x_2
\end{cases}$$

i.e., $\dot{x} = Ax$, $A = \begin{pmatrix} -1 & -1 \\ 1 & -3 \end{pmatrix}$.

It's easy to check that $\lambda_1 = \lambda_2 = -2$ is an eigenvalue of $A$ with eigenvector $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Thus, $\dot{x}_1(t) = e^{-2t}v_1$ is a solution to $\dot{x} = Ax$.

We need another. Try $\dot{x}_2 = e^{-2t}(t\dot{v} + \ddot{w})$, solve for $\dot{v}$, $\ddot{w}$.

Plug back in:

$$\begin{align*}
\dot{x}_2 &= -2e^{-2t}(t\dot{v} + \ddot{w}) + e^{-2t} \dot{v} = e^{-2t}(tA\dot{v} + A\ddot{w})
\end{align*}$$

Equate coeffs:

$$\begin{cases}
e^{-2t} : -2\dot{v} = A\dot{v} & \Rightarrow (A + 2I)\dot{v} = 0 \\
e^{-2t} : -2\ddot{w} = A\ddot{w} & \Rightarrow (A + 2I)\ddot{w} = 0
\end{cases}$$

So, $\dot{v} = \ddot{v}_1$ and $\ddot{w} = \ddot{v}_2$, a generalized eigenvector ($\ddot{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ works).

Thus, the general solution is $\dot{x}(t) = Ce^{-2t}\dot{v}_1 + C_2 e^{t} (t\dot{v}_1 + \ddot{v}_2)$.

Or $\dot{x}(t) = e^{Jt}x_0$, where $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (Jordan canonical form; here $\lambda = 2$).