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 \( 0 \neq w \in \mathbb{K}^n \), consider the following:
\[ w, Aw, A^2 w, \ldots, A^n w. \]

Since \( \dim \mathbb{K}^n = n \), these vectors are linearly dependent.

Thus, we can write
\[ 0 = c_0 w + c_1 Aw + \ldots + c_n A^n w. \]

\[ = p(A)w \]

where \( p(x) = c_0 + c_1 x + \ldots + c_n x^n \in \mathbb{K}[x] \).

Since \( \mathbb{K} \) is closed, \( p(x) \) is a product of linear factors, i.e.,
\[ p(x) = c \prod_{j=1}^{n} (x - \lambda_j), \quad c \neq 0 \]

and so
\[ p(A)w = c \prod_{j=1}^{n} (A - \lambda_j I)w = 0. \]

Now, one of \( A - \lambda_j I \) must be non-invertible. (Because
p(A) is non-invertible. Suppose \( A - \lambda I \) is non-invertible, and pick \( \mathbf{v} \neq \mathbf{0} \) in the nullspace of \( A - \lambda I \).

Then, \((A - \lambda I)\mathbf{v} = \mathbf{0} \Rightarrow A\mathbf{v} = \lambda \mathbf{v} \).

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

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

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

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

Now, let's find the eigenvectors:

\( \lambda_1 = 2 \): Find \( \mathbf{v} \), such that \((A - 2I)\mathbf{v} = \mathbf{0} \).

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

Thus, \( \mathbf{v}_1 = \begin{pmatrix} -2c \\ c \end{pmatrix} \) is an eigenvector for any \( c \).

\( \lambda_2 = 5 \): Find \( \mathbf{v} \), such that \((A - 5I)\mathbf{v} = \mathbf{0} \).

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

Thus, \( \mathbf{v}_2 = \begin{pmatrix} c \\ c \end{pmatrix} \) is an eigenvector for any \( c \).

We'll say \( A \) has eigenvalues \( \lambda_1 = 2, \lambda_2 = 5 \) eigenvectors \( \mathbf{v}_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \), \( \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \).
Here, \( v_1 \) and \( v_2 \) are linearly independent. Then, 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 \( p_A(t) = \det(tI - A) \).

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

**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 \( n \) is minimal, non-zero.

(So clearly, \( c_j \neq 0 \).)
Apply $A: \ c_1 v_1 + \ldots + c_m v_m = 0$

$\Rightarrow c_1 A v_1 + \ldots + c_m A v_m = 0$

$\Rightarrow c_1 \lambda_1 v_1 + \ldots + c_m \lambda_m v_m = 0$

We now have $\sum_{j=1}^{n} c_j v_j = 0$ and $\sum_{j=1}^{n} c_j \lambda_j v_j = 0$.

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

This contradicts minimality of $m$.

Thus, $v_1, \ldots, v_m$ must be linearly independent. \[\Box\]

**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 $x = \sum_{j=1}^{n} a_j v_j$, 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$ and $\prod_{i=1}^{n} \lambda_i = \text{det} A$.

**Proof:** Claim: $P_A(t) = t^n - (\text{tr} A) t^{n-1} + \ldots + (-1)^n \text{det} A$.

Write $P_A(t) = \prod_{i=1}^{n} (t - \lambda_i)$.

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

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

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

Thus, \( \det(tI - A) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^{n} (t - a_{\pi(i),i}) \).

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

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 \( g(t) \in K[t] \) be any polynomial, say \( g(t) = \sum_{i=1}^{n} a_it^i \).

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

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

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

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

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

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

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

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

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 \( \implies \) some \( r_i \) is an eigenvalue of \( A. \)

Since \( r_i \) is a root of \( g(t) - \mu, \quad g(r_i) = \mu. \)

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

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

Proof: Case I: 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 \( p_A(t) \)

Thus, for any \( x \in \mathbb{C} \), write \( x = c_1 v_1 + \cdots + c_n v_n. \)

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

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

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

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 $Q(t) = tI - A$, $P(t) = \begin{pmatrix} p_j(t) \end{pmatrix}$, $p_{ij}(t) = (-1)^{i+j} D_{ij}(t)$

where $D_{ij}(t)$ is determinant of $ij$-th minor of $Q(t)$.

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 $(Q(t))^{-1} = \frac{1}{\det P(t)} P(t)$.

Put $R(t) := P(t)Q(t) = (\det Q(t)) I = P_A(t) I$

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

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

Examples:

1. $A = I$, then $P_A(t) = \det(tI - I) = (t-1)^n$

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

2. $A - I = 0$, 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
\[ p_A(t) = t^2 - 2t + 1 = (t-1)^2, \]
so \( d_1 = d_2 = 1 \).

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

**Proof:** 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 \( x = a_1 x_1 + \ldots + a_n x_n \). \( A x = a_1 A x_1 + \ldots + a_n A x_n = a_1 \lambda x_1 + \ldots + a_n \lambda x_n = \lambda (a_1 x_1 + \ldots + a_n x_n) = \lambda x \).

Remark: Every 2x2 matrix with \( \text{tr} \ A = 2 \), \( \det A = 1 \), has \( \lambda = 1 \) as a double root of \( p_A(t) \). 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 \((A-\lambda I)v_2 = v_1 \), \( \implies (A-\lambda I)^2 v_2 = 0 \).
Similarly, we can find $v_3$ such that

$$(A-\lambda I)v_3 = v_2 \Rightarrow (A-\lambda I)^2v_2 \neq 0 \text{ but } (A-\lambda I)^3v_3 = 0.$$ 

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

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

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

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

For convenience, pick $c = 0$ to get $v_2 = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix}$.

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 (generic or generalized) of $A$. 

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

Proof: Let \( I = \{ ap + bq : a, b \in \mathbb{C}[t] \} \), and pick \( d \in I \) with minimal degree.

Claim 1: \( d \) divides \( p \) and \( q \). (i.e., it is a "greatest common divisor").

Suppose it did not. Using the division algorithm, we could write: \( r = p - md, \) \( \deg r < \deg d \).

Since \( p, d \in I \), then \( p - md \in I \). But \( d \) had minimal degree in \( I \). \( \Box \)

Claim 2: \( \deg d = 0 \).

If not, it would have a root \( \alpha \), and since \( d \mid p \) and \( d \mid q \), \( \alpha \) is a root of \( p \) and \( q \). \( \Box \)

Thus, \( d \) is a constant; we may assume \( 1 \), since we're over \( \mathbb{C} \).

Lemma 6.9: Let \( A \) be an \( n \times n \) matrix over \( \mathbb{C} \), and let \( p, q \in \mathbb{C}[t] \) with no common roots. Let \( N_p, N_q, N_{pq} \) be the null spaces of \( p(A), q(A), \) and \( p(A)q(A) \), respectively. Then \( N_{pq} = N_p \oplus N_q \).

Proof: Write \( ap + bq = 1 \) for \( a, b \in \mathbb{C}[t] \).
Plug in $A$: \[ a(A)p(A) + b(A)q(A) = I. \]

Take $x \in N_{p_0}$: \[ a(A)p(A)x + b(A)q(A)x = x. \]  \hspace{1cm} (8)

Note: \[ q(A)[a(A)p(A)x] = a(A)p(A)q(A)x = 0 \] \hspace{1cm} (since $x \in N_{p_0}$)

Thus, \[ a(A)p(A)x \in N_q. \]

Similarly, \[ p(A)[b(A)q(A)x] = b(A)p(A)q(A)x = 0 \] \hspace{1cm} (since $x \in N_{p_0}$)

Thus, \[ b(A)q(A)x \in N_p. \]

The expression (8) is \[ x = x_p + x_q, \]

\[ = b(A)q(A)x + a(A)p(A)x \in N_p \cap N_q. \]

To show $N_{p_0} = N_p \oplus N_q$, we must show this decomposition is unique.

Suppose \[ x = x_p + x_q = x'_p + x'_q. \]

Put \[ y = x_p - x'_p = x_q - x'_q \in N_p \cap N_q. \]

\[ 0 = a(A)p(A)y + b(A)q(A)y = Iy = y \implies y = 0. \]

**Corollary 6.10:** Let \( p_1, \ldots, p_k \in \mathbb{C}[t] \) be pairwise coprime (no common roots). Let \( N_{p_1} \cdots p_k \) be the nullspace of \( p_1(A) \cdots p_k(A) \).

Then \[ N_{p_1} \cdots p_k = N_{p_1} \oplus \cdots \oplus N_{p_k}. \]

**Proof:** Exercise. (Induct on \( k \).)
Proof of Spectral Theorem: Pick \( x \in \mathbb{C}^n \). The vectors 
\( x, Ax, A^2 x, \ldots, A^n x \) are linearly dependent, thus there is a polynomial of degree \( \leq n \) such that \( p(A)x = 0 \).

Factor \( p: p(A)x = \prod_{j=1}^{k} (A - \lambda_j I)^{m_j} x = 0 \), so the roots of \( p \) are \( \lambda_j \), with multiplicity \( m_j \).

Moreover, if \( \lambda_j \) is not an eigenvalue, then \( A - \lambda_j I \) is invertible and can be removed. Thus, assume each \( \lambda_j \) is an eigenvalue of \( A \).

Write \( p_j(t) = (t - \lambda_j)^{m_j} \), so we have \( \prod_{j=1}^{k} p_j(A)x = 0 \), and \( x \in N_{p_1} \cdots N_{p_k} \).

Clearly, none of \( p_1, \ldots, p_k \) have (pairwise common roots), so \( N_{p_1} \cdots N_{p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k} \).

\[ x = x_{p_1} + \ldots + x_{p_k}, \quad x_{p_i} \in N_{p_i} \]

Note that each \( x_{p_i} \in N_{p_i} \) is a generalized eigenvector.

Let \( I = I_A \) be the set of polynomials \( p(t) \in \mathbb{C}[t] \) for which \( p(A) = 0 \).

Note that \( I \) is closed under addition and multiplication (of not just scalars, but polynomials too).
Lemma: I contains a unique monic polynomial $m = m_\lambda$ of minimal degree, and all other polynomials in $I$ are scalar multiples of $m$ (i.e., $I$ is a principal ideal of $\mathbb{C}[t]$).

Proof: Let $m \in I$ have minimal degree. Clearly, $m$ is unique, because if there were another, their difference would have strictly smaller degree.

Now, suppose $p \in I$ were not a multiple of $m$. Using the division (Euclidean) algorithm, write $p = qm + r$ with $\deg r < \deg m$, contradicting minimality. \(\blacksquare\)

Def: The minimal polynomial of a matrix $A$, denoted $m_\lambda$, is the unique monic polynomial of minimal degree for which $m_\lambda(A) = 0$. Let $N_\lambda = N_{m_\lambda}(\lambda)$ be the nullspace of $(A - \lambda I)^m$.

Note that $N_\lambda$ consists of generalized eigenvectors, and

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

for some index $d$. Denote $d = d(\lambda)$ be the minimal index such that $N_{d-1} \subseteq N_d = N_{d+1}$, called the index of the eigenvalue $\lambda$. 
Theorem 6.11: Let $A$ be an $n \times n$ matrix, with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, with indices $d_1, \ldots, d_k$. Then, the minimal polynomial of $A$ is $m_A(t) = \prod_{i=1}^{k} (t - \lambda_i)^{d_i}$.

Proof: Exercise.

Denote $N_{d_j}(\lambda_j)$ by $N^{(i)}$. The Spectral Theorem (Thm 6.7) can now be stated as follows:

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

Remark: $\dim N^{(i)}$ is the algebraic multiplicity of $\lambda_i$.

(This will be proved later.)

Note that $A$ maps $N^{(i)}$ 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 $N_1, \ldots, N_k$.

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 $A$ and $B$ share the same eigenvalues $\lambda_1, \ldots, \lambda_k$, and $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$ for each $j = 1, \ldots, k$, where $N_m(\lambda_j) = \text{nullspace of } (A - \lambda_j I)^n$ and $M_m(\lambda_j) = \text{nullspace of } (B - \lambda_j I)^n$.
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.

It follows that \( \lambda \) is an eigenvalue of \( A \) iff it is an eigenvalue of \( B \).

\( \Leftarrow \) Let \( \lambda = \alpha_j \) be an eigenvalue of \( A \), with \( N_i = \text{nullspace}(A-\alpha_j I)^i \).

Goal: Construct a basis for \( N_d \) under which \( A-\alpha_j I \) admits a nice matrix form.

Remark: \( 0 = N_0 \subset N_1 \subset N_2 \subset \cdots \subset N_{d-1} \subset N_d = N_{d+1} \) , where \( d \) is the index of \( \lambda \).

Lemma 6.13: The map \( A-\alpha_j I \) carries over to a well-defined map on the quotient spaces: \( A-\alpha_j I : N_{i+1}/N_i \longrightarrow N_i/N_{i-1} \).

Moreover, it is injective. \( \{x\} \rightarrow \{(A-\alpha_j I)x\} \)

Proof: Exercise. (HW) \( \Box \)

By Lemma 6.13, \( \dim (N_{i+1}/N_i) \leq \dim (N_i/N_{i-1}) \).

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

First, let \( \bar{X}_1, \ldots, \bar{X}_{\alpha} \) be a basis for \( N_d/N_{d-1} \).

By Lemma, \((A-\alpha_j I)\bar{X}_1, \ldots, (A-\alpha_j I)\bar{X}_{\alpha} \) are linearly independent in \( N_{d-1}/N_{d-2} \).

Extend to a basis by adding \( \bar{X}_{\alpha+1}, \ldots, \bar{X}_d \in N_{d-1}/N_{d-2} \).
Repeat this process: 

\[(A-\lambda I)^2 \bar{x}_{i}, (A-\lambda I)^2 \bar{x}_{i+1}, \ldots, (A-\lambda I)^2 \bar{x}_{i+k}, \ldots, (A-\lambda I) \bar{x}_{i}, \text{ are linearly independent in } N_{d-2}/N_{d-3}.\]

Extend to a basis by adding \(\bar{x}_{i+1}, \ldots, \bar{x}_{i+k} \in N_{d-2}/N_{d-3}\).

Eventually, this terminates because \(N_{0} = 0\).

\[\text{Remark: Each "row" of vectors spans a subspace that is invariant under } A-\lambda I.\]
The matrix of $A - \lambda I$ restricted to the vectors in the first row has the form,

$$
\begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \cdots & 0
\end{bmatrix}_{d \times d}
$$

Thus, the matrix of $A$ restricted to this invariant subspace is

$$
\begin{bmatrix}
\lambda & 1 & 0 & 0 \\
\lambda & \lambda & 1 & 0 \\
0 & \lambda & \ddots & 1 \\
0 & 0 & \cdots & \lambda
\end{bmatrix}_{d \times d}
$$

Such a matrix is called a Jordan block.

Similarly, each "row" of the previous diagram corresponds to a Jordan block.

When put together, over all "rows," and then all distinct eigenvalues, the matrix $A$ can be written in block-diagonal form. The full $n \times n$ matrix for $A$ with respect to such a basis is called the Jordan canonical form of $A$. Since it only depends on the eigenvalues and dimensions of the eigenspaces,
If $A$, $B$ have the same eigenvalues and eigenspace dimensions, then they are similar to the same matrix, their common Jordan canonical form. Hence they are similar matrices, proving Theorem 6.17.

The following is a generalization of the spectral mapping theorem.

**Theorem 6.14**: Let $A, B: X \to X$ 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 $X = N^{(1)} \oplus \ldots \oplus N^{(k)}$, where each summand is a generalized eigenspace of $A$. (i.e., $N^{(j)} = \text{nullspace of } (A - \lambda_j I)^{d_j}$.)

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

To show this, consider: $B(A - \lambda I)^d x = (A - \lambda I)^d B x$. $(\star)$

If $\lambda$ is an eigenvalue of $A$, then the LHS of $(\star)$ is zero.

Thus, $B x \in N^{(i)}$.

Now apply the spectral theorem to $B$, restricted to each $N^{(i)}$ separately. $\square$
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 a linear map, 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.

Theorem 6.17: Let $X$ be a finite-dimensional space over $\mathbb{C}$, and $A : X \to X$ a linear map. Let $\lambda \neq \lambda'$ be eigenvalues of $A$ (and thus also of $A'$). If $v \in X$ is an eigenvector for $\lambda$ and $\ell \in X'$ an eigenvector for $\lambda'$, then $(\ell, x) = 0$.

**Proof:** By assumption, $Av = \lambda v$ and $A'l = \lambda' l$.

\[
\lambda (\ell, v) = (\ell, \lambda v) = (\ell, Av) = (A'l, v) = (\lambda'l, v) = \lambda'(\ell, v).
\]

Since $\lambda \neq \lambda'$, $(\ell, v) = 0$. \qed
Theorem 6.17 has a useful application:

Theorem 6.18: Suppose \( A \) has distinct eigenvalues \( \lambda_1, \ldots, \lambda_n \) and corresponding eigenvectors \( v_1, \ldots, v_n \in \mathbb{A} \), and let \( l_1, \ldots, l_n \) be the corresponding eigenvectors in \( \mathbb{A}' \).

Then:
1. \( (l_i, v_i) \neq 0 \) for each \( i \)
2. If \( x = \sum \limits_{i=1}^{n} a_i v_i \), then \( a_i = \frac{(l_i, x)}{(l_i, v_i)} \).

Proof: Exercise. (HW).

Remark: Theorems 6.17 and 6.18 together tell us that the eigenvectors \( l_1, \ldots, l_n \in \mathbb{A}' \) is the dual basis to the eigenvectors \( v_1, \ldots, v_n \in \mathbb{A} \).

Def: When \( A \) has linearly independent eigenvectors \( v_1, \ldots, v_n \), 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, \( D = \text{PAP}^{-1} \). The matrix \( D \) has the eigenvalues down the diagonal, and the columns of \( P \) are the corresponding eigenvectors, i.e., \( D = (\lambda_1 e_1, \ldots, \lambda_n e_n) \), \( P = (v_1, \ldots, v_n) \).
To see this, note that
\[ A \mathbf{v} = A (v_1, ..., v_n) = (A v_1, ..., A v_n) = (\lambda_1 v_1, ..., \lambda_n v_n) = (\lambda_1 e_1, ..., \lambda_n e_n) = P D. \]

Applications to differential equations

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

Suppose that \( A \) has eigenvalues \( \lambda_1, ..., \lambda_n \) and \( n \) linearly independent eigenvectors \( v_1, ..., v_n \).

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

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

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

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

In matrix form, this is \[ \begin{bmatrix} e^{\lambda_1 t} & \cdots & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} = e^{Dt} \dot{X}_0. \]

Here, \( \dot{X}_0 = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} \) and we are using the basis \( v_1, ..., v_n \).
With respect to the basis $e_1, \ldots, e_n$, $e^{Dt}x_0$ becomes

$$e^{At}x_0 = e^{P^{-1}Dp}P^{-1}e^{Dp}P^{-1}x_0,$$

while it may seem that $e^{At} = \sum_{i=0}^{\infty} \frac{1}{i!} A^i$ is hard to compute, $e^{Dp}$, and $P^{-1}e^{Dp}P$ 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

$$\dot{x}(t) = e^{At}x_0 = P^{-1}e^{Dp}P x_0,$$

where $A = P^{-1}DP$.

(2) Consider

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

i.e., $\dot{x} = A\vec{x}$, $A = \begin{pmatrix} -1 & -1 \\ 1 & -2 \end{pmatrix}$.

It's easy to check that $\lambda_1 = \lambda_2 = -2$ is an eigenvalue of $A$, with eigenvector $\vec{v}_1 = (1, 0)$.

Thus, $\dot{x}_1(t) = e^{-2t}\vec{v}_1$ is a solution to $\dot{x'} = Ax$.

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

Plug back in: $\begin{cases} x'_1(t) = -2e^{-2t}(t\vec{v} + \vec{w}) + e^{-2t}\vec{v} = e^{-2t}(tA\vec{v} + A\vec{w}) \\ e^{-2t}: t\vec{v} - 2\vec{w} = A\vec{w} \Rightarrow (A + 2I)\vec{w} = t\vec{v} \end{cases}$

So, $\vec{v} = \vec{v}_1$, and $\vec{w} = \vec{w}_2$, a generalized eigenvector ($\vec{w}_2 = (0, 1)$ works).

Thus, the general solution is $\dot{x}(t) = C_1 e^{-2t}\vec{v}_1 + C_2 e^{-2t}(t\vec{v} + \vec{w}_2)$.

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