6. Spectral theory:

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

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

Prop: A has an eigenvector

Proof: Pick any $0 \neq w \in X$, consider the following:

$$w, \ Aw, \ A^2 w, \ \ldots, \ A^n w.$$

Since $\dim X = n$, these vectors are linearly dependent.

Thus, we can write
$$0 = c_0 w + c_1 A w + \cdots + c_n A^n w$$
$$= p(A)\, w$$

where $p(x) = c_0 + c_1 x + \cdots + c_n x^n \in K[x]$.

Since K is closed, $\quad p(x) = c \prod_{j=1}^{n} (x - \lambda_j), \qquad c \neq 0$

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 $v \neq 0$ in the nullspace of $A - \lambda I$.

Then, $(A - \lambda I)v = 0 \Rightarrow Av - \lambda v = 0 \Rightarrow Av = \lambda v.$ $\square$

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

<u>Example</u>: $A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$.

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

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

Now, let's find the eigenvectors.

<u>$\lambda_1 = 2$</u>: 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.

<u>$\lambda_2 = 5$</u>: 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 $N \to \infty$.

<u>Note:</u> The entries in A^N grow asympotically as $\sim 5^N$, the largest eigenvalue.

<u>Def:</u> The <u>characteristic polynomial</u> of an $n \times n$ matrix A is $p_A(s) \doteq \det(sI - A)$.

<u>Remarks:</u> $p_A(s)$ 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.

<u>Theorem 6.1:</u> Eigenvectors of A corresponding to distinct eigenvalues are linearly independent.

Suppose $\sum_{j=1}^{m} c_j v_j = 0$, where m is <u>minimal</u>, non-zero (so clearly, $c_j \neq 0$).

Apply A: $\quad c_1 v_1 + \cdots + c_m v_m = 0$

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

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

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

This contradicts minimality of m. $\square$

<u>Def</u>: If $A : X \to X$ and X has a basis of eigenvectors of A, then A is <u>diagonalizable</u>, because we can write $A = PDP^{-1}$, or $D = P^{-1}AP$.

$$\begin{tikzcd} X \arrow[r, "D"] \arrow[d, "P"'] & X \arrow[d, "P"] \\ X \arrow[r, "A"] & X \end{tikzcd}$$

Here, $D = [\lambda_1 e_1 \ \cdots \ \lambda_n e_n]$

$P = [v_1 \ \cdots \ v_n]$

$$\begin{aligned} AP = A[v_1 \cdots v_n] = [Av_1 \cdots Av_n] &= [\lambda_1 v_1 \cdots \lambda_n v_n] \\ &= [\lambda_1 Pe_1 \cdots \lambda_n Pe_n] \\ &= P[\lambda_1 e_1 \cdots \lambda_n e_n] = PD. \end{aligned}$$

Corollary 6.2: If A has distinct eigenvalues, it's diagonalizable.

In this case, it's easy to compute $A^N x$ for any $x \in X$:

write $x = \sum_{j=1}^{n} a_j v_j$, $\quad A^N x = \sum_{j=1}^{n} a_j A^N v_j = \sum_{j=1}^{n} a_j \lambda_j^N v_j$.

Theorem 6.3: If the eigenvalues of A are $\lambda_1, \dots, \lambda_n$, then
$\sum_{i=1}^{n} \lambda_i = \operatorname{tr} A$ and $\prod_{i=1}^{n} \lambda_i = \det A$.

Pf: $$p_A(s) = \prod_{i=1}^{n} (s - \lambda_i) = s^n - \underbrace{(\lambda_1 + \cdots + \lambda_n)} s^{n-1} + \cdots + (-1)^n \underbrace{\lambda_1 \cdots \lambda_n}$$

Idea: Use $p_A(s) = \det(sI - A)$ to show $\operatorname{tr} A =$ and $\det A =$

$$= \begin{vmatrix} s - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & s - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & s - a_{nn} \end{vmatrix} = \det\left(s\delta_{ij} - a_{ij}\right)_{n \times n}$$

ij-coordinate

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

$$\det(sI - A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} \left(s\,\delta_{\pi(i),i} - a_{\pi(i),i}\right)$$

Clearly, the s^{n-1} coeff. is $-\sum_{i=1}^{n} a_{ii} = \operatorname{tr} A$ ✓

and the constant term is [plug in $s=0$] $\det(-A) = (-1)^n \det A$.

□

<u>Remark:</u> If $Av = \lambda v$, then $A^2 v = \lambda^2 v$. Thus, if λ is an eigenvalue of A, then λ^N is an eigenvalue of A^N.

Take this further: let $q(s) \in K[s]$, say $q(s) = \sum_{i=0}^{n} a_i s^i$

Then $q(A)v = \sum_{i=0}^{n} a_i A^i v = \sum_{i=0}^{n} a_i \lambda^i v = q(\lambda) v$.

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

<u>Theorem 6.4</u>: ("Spectral mapping theorem"). Let A have e-value λ, and let $q(s) \in K[s]$.

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

(b) Conversely, <u>every</u> eigenvalue of $q(A)$ has the form $q(\lambda)$.

<u>Proof</u>: (a) We just did this.

(b) Let μ be an eigenvalue of $q(A) \iff \det\big(\overbrace{q(A) - \mu I}^{\text{not invertible!}}\big) = 0$.

Consider $q(s) - \mu = c(s - r_1) \cdots (s - r_n)$

Plug in A: $q(A) - \mu I = c(A - r_1 I) \cdots (A - r_n I)$ is <u>singular</u>.

Thus, some $A - r_j I$ must be non-invertible.

$\Rightarrow r_j$ is an e-value of A.

Since r_j is a root of $q(s) - \mu$, $q(r_j) = \mu$ $\square$

Cor: All eigenvalues of $p_A(A)$ are zero.

Actually, much more is true:

Theorem 6.5 (Cayley-Hamilton): $p_A(A) = 0$.

Proof: Case 1 (easy): A is diagonalizable.

$$p_A(A)x = \sum_{j=1}^{n} p_A(A) c_j v_j = \sum_{j=1}^{n} p_A(\lambda_j) c_j v_j = \sum_{j=1}^{n} 0 \, c_j v_j = 0 \checkmark$$

write x as $c_1 v_1 + \dots + c_n v_n$.

For the general case, we need a lemma.

Lemma 6.6: Let P & Q be polynomials with matrix coeffs:

$$P(s) = \sum_j P_j s^j, \qquad Q(s) = \sum Q_k s^k. \qquad \text{Let } R = PQ.$$

Then
$$R(s) = P(s)Q(s) = (P_n s^n + \dots + P_1 s + P_0)(Q_m s^m + \dots + Q_1 s + Q_0)$$
$$= R_{n+m} s^{n+m} + \dots + R_1 s + R_0.$$
$$\text{where } R_\ell = \sum_{j+k=\ell} P_j Q_k.$$

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

Proof: Exercise.

Now, let $\boxed{Q(s) = sI - A}$, so $\det Q(s) = p_A(s)$.

Recall Cramer's theorem: $(Q^{-1})_{ij} = (-1)^{i+j} \dfrac{\det Q_{ji}}{\det Q}$, if $\det Q \neq 0$.

Even if $\det Q = 0$, we can multiply through by $(\det Q)$ & Q:

$$I \cdot \det Q = \begin{bmatrix} \det Q & & \\ & \ddots & \\ & & \det Q \end{bmatrix} = \left((-1)^{i+j} \det Q_{ji}\right) \cdot Q$$

Now, let $P(s) := \left((-1)^{i+j} Q_{ji}(s)\right)_{n \times n}$.

$R(s) := P(s)\,Q(s) = (\det Q)\, I = p_A(s)$.

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

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

$\square$

Remark: $p_A(s)$ is a degree-n monic polynomial s.t. $p_A(A) = 0$.

Let $I = \{ p(s) \in \mathbb{C}[s] : p(A) = 0 \}$. This is an *ideal* of $\mathbb{C}[s]$ since it's closed under $+$, $-$, scalar multiplication.

Since $\mathbb{C}[s]$ is a *principal ideal domain* (PID), $I = \langle m_A \rangle$ for some monic $m_A(s)$ of minimal degree called the *minimal polynomial*. All polynomials s.t. $p(A) = 0$ are scalars of $m_A(s)$.

Here's a direct proof (unnecessary for those with a ring theory background)

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

Existence: Division algorithm. Suppose $p \in I$ were not a multiple of $m = m_A$.

Write $p = qm + r$, $\deg r < \deg m$.

Then $r = p - qm \in I$. $\lightning$ $\square$

Fact (we'll prove later): $p_A(s)$ and $m_A(s)$ have exactly the same set of roots.

Examples: (1) $A = I$. $p_A(s) = \det(sI - I) = (s-1)^n$

$\Rightarrow \lambda = 1$ 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.

Note that $m_A(s) = s - 1 \neq p_A(s)$

(2) $A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$ $\operatorname{tr} A = 2$, $\det A = 1$,

So $p_A(s) = s^2 - 2s + 1 = (s-1)^2$, so $\lambda_1 = \lambda_2 = 1$.

Since $A - I \neq 0$, $m_A(s) = (s-1)^2 = p_A(s)$.

Eigenvectors: $(A-I)v = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\Rightarrow x_1 + x_2 = 0 \Rightarrow v = \begin{pmatrix} c \\ -c \end{pmatrix}$ is the only eigenvector.

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

Proof: Pick $x \in X$, write $x = c_1 x_1 + \cdots + c_n x_n$.

$$Ax = A(c_1 x_1 + \cdots + c_n x_n) = c_1 \lambda x_1 + \cdots + c_n \lambda x_n = \lambda(c_1 x_1 + \cdots + c_n x_n)$$
$$= \lambda x. \qquad \square$$

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

Ex: $A = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$ $\qquad A = \begin{bmatrix} \boxed{\begin{matrix} 1 & 1 \\ & 1 \end{matrix}} & \\ & \boxed{1} \end{bmatrix}$ $\qquad A = \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$

$m_A(s) = s - 1$ $\qquad m_A(s) = (s-1)^2$ $\qquad m_A(s) = (s-1)^3$

3 e-vectors $\qquad$ 2 e-vectors $\qquad$ 1 e-vector

Suppose λ is an eigenvalue with multiplicity m, but only one eigenvector, v_1.

This means: $(A - \lambda I)v_1 = 0$, $\dim N_{A-\lambda I} = 1$, $\operatorname{rank}(A - \lambda I) = m - 1$.

<u>Big idea</u>: we can find some v_2 s.t. $(A-\lambda I)v_2 = v_1$

$$\Rightarrow (A-\lambda I)^2 v_2 = 0$$

Similarly, we can find v_3 s.t.

$$(A-\lambda I)v_3 = v_2 \Rightarrow (A-\lambda I)^2 = v_1 \neq 0 \quad \text{but} \quad (A-\lambda I)^3 v_3 = 0.$$

Picture of this: $v_m \xrightarrow{A-\lambda I} \cdots \longrightarrow v_3 \longrightarrow v_2 \longrightarrow v_1 \xrightarrow{A-\lambda I} 0$

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

<u>Example</u>: $A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$, $\lambda_1 = \lambda_2 = 1$, $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

To find a generalized eigenvector v_2, we need to solve

$$(A-\lambda I)v_2 = v_1 \Rightarrow \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\Rightarrow 2x_1 + 2x_2 = 1 \Rightarrow x_2 = \tfrac{1}{2} - x_1$$

$$\Rightarrow v = \begin{pmatrix} c \\ \frac{1}{2} - c \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} + \begin{pmatrix} c \\ c \end{pmatrix} \quad \text{for any } c.$$

Pick $c = 0$. We have $\begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \xmapsto{A-\lambda I} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xmapsto{\quad} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

<u>Def</u>: The <u>algebraic multiplicity</u> of λ is the largest m s.t. $(s-\lambda)^m$ appears as a factor of $p_A(s)$.

The <u>geometric multiplicity</u> of λ is the number of linearly independent eigenvectors, i.e., $\dim N_{A-\lambda I}$

<u>Preview</u>: The <u>spectral theorem</u> guarantees that there will always be a basis for $\mathbb{C}^n$ consisting of generalized eigenvectors of A that "behave nicely."

<u>Example</u>: Suppose $A: X \to X$, $p_A(s) = (s-\lambda)^{11}$, $\dim N_{A-\lambda I} = 4$.

The following is one possibility for the generalized eigenvectors:

$$
\begin{array}{ccccccccccc}
v_5 & \xrightarrow{A-\lambda I} & v_4 & \longrightarrow & v_3 & \longrightarrow & v_2 & \longrightarrow & v_1 & \xrightarrow{A-\lambda I} & 0 \\
 & & & & w_3 & \longrightarrow & w_2 & \longrightarrow & w_1 & \longrightarrow & 0 \\
 & & & & & & x_2 & \longrightarrow & x_1 & \longrightarrow & 0 \\
 & & & & & & & & y_1 & \longrightarrow & 0
\end{array}
$$

<u>Remarks</u>:

$$N_1 := N_{A-\lambda I} = \operatorname{Span}\{v_1, w_1, x_1, y_1\}$$

$$N_2 := N_{(A-\lambda I)^2} = \operatorname{Span}\{v_1, w_1, x_1, y_1, v_2, w_2, x_2\}$$

$$N_3 := N_{(A-\lambda I)^3} = \operatorname{Span}\{v_1, w_1, x_1, y_1, v_2, w_2, x_2, v_3, x_3\}$$

$$\vdots$$

and
$$\cdots = N_6 = N_5 \supsetneq N_4 \supsetneq N_3 \supsetneq N_2 \supsetneq N_1 \supsetneq 0$$
$$\dim N_i = 11 = 11 > 10 > 9 > 7 > 4 > 0$$

<u>Def</u>: The <u>index</u> of λ = $\min d$ s.t. $N_d = N_{d+1}$.

<u>Def</u>: If $A: X \to X$, the a subspace $Y \leq X$ is <u>invariant under A</u> if $A(Y) \subseteq Y$.

<u>More remarks</u>:

- If you believe that these generalized eigenvectors form a basis of $\mathbb{C}^n$, then $m_A(s) = (s-\lambda)^{\text{index}} = (s-\lambda)^5$.

- Each "row" is an invariant subspace of $A - \lambda I$, and of A. [What other invariant subspaces are there?]

- If $X = Y \oplus Z$, both invariant under A, and $y_1, \dots, y_k, z_1, \dots, z_\ell$ is a basis, then the matrix form of A wrt this basis is <u>block-diagonal</u>:

- The matrix of A restricted to $\text{Span}\{v_1, \dots, v_5\}$ is

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

$$\begin{aligned} A v_1 &= \lambda v_1 \\ (A - \lambda I) v_2 &= v_1 \implies A v_2 = v_1 + \lambda v_2 \\ & \qquad\qquad\quad\; A v_3 = v_2 + \lambda v_3 \\ & \qquad\qquad\qquad\quad \vdots \end{aligned}$$

Clearly, $J - \lambda I = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$ satisfies $(J - \lambda I)^5 = 0$, $(J - \lambda I)^4 \neq 0$.

This is called a <u>Jordan block</u> for λ.

<u>Theorem 6.7</u> (Spectral theorem): Let A be an $n \times n$ matrix over $\mathbb{C}$. Then $\mathbb{C}^n$ has a basis of generalized eigenvectors of A.

We need some algebraic results first.

<u>Lemma 6.8</u>: Let $p, q \in \mathbb{C}[s]$, co-prime. Then we can write

$ap + bq = 1$ for some other $a, b \in \mathbb{C}[s]$.

<u>Remark</u>: This is by the division algorithm. If these are integers, then we can write $m = qn + r$, $r < n$

[e.g., $49 = \underline{9} \cdot 5 + \underline{4}$, q is the quotient, r is the remainder]

<u>Proof</u>: Let $I = \{ap + bq : a, b \in \mathbb{C}[s]\}$; the <u>ideal</u> generated by p & q. Pick $d \in I$ with minimal degree.

<u>Claim 1</u>: $d \mid p$ and $d \mid q$.

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

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

Since $p, d \in I$, $r = p - md \in I$. But $\deg d$ is min'l. $\lightning$.

<u>Claim 2</u>: $\deg d = 0$.

If not, then $(s-\alpha) \mid d \Rightarrow s-\alpha \mid p$ and q $\lightning$.

Thus d is <u>constant</u>, assume $d = 1$ since we're over $\mathbb{C}$. $\square$

Lemma 6.9: Let $A : \mathbb{C}^n \to \mathbb{C}^n$, $p, q \in \mathbb{C}[s]$, co-prime.

Let N_p, N_q, N_{pq} be the nullspaces of $p(A)$, $q(A)$, $p(A)q(A)$.

Then $N_{pq} = N_p \oplus N_q$.

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

Plug in A: $\quad a(A)p(A) + b(A)q(A) = I$

Multiply by $x \in N_{pq}$: $\quad \underbrace{a(A)p(A)x}_{\text{in } N_q \text{ because}} + \underbrace{b(A)q(A)}_{\text{in } N_p \text{ because}} x = x. \quad (*)$

$a(A)[p(A)q(A)x] = 0 \qquad b(A)[p(A)q(A)x] = 0.$

Note: $f(A)g(A) = g(A)f(A) \quad \forall f, g \in \mathbb{C}[s]$.

The expression $(*)$ is $\quad x = x_p + x_q$

$$= b(A)q(A)x + a(A)p(A)x$$

This shows $N_{pq} = N_p + N_q$. To show $\oplus$, we need uniqueness.

Suppose $x = x_p + x_q = x_p' + x_q'$. Collect terms:

$$y := x_p - x_p' = x_q' - x_q \in N_p \cap N_q.$$

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

$\Rightarrow y = 0$. Thus, $N_{pq} = N_p \oplus N_q$. $\square$

Corollary 6.10: Let $p_1, \ldots, p_k \in \mathbb{C}[s]$, pairwise co-prime, &

$N_{p_1 \cdots p_k}$ = 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 of spectral theorem: Write $p_A(s) = (s-\lambda_1)^{n_1}(s-\lambda_2)^{n_2} \cdots (s-\lambda_k)^{n_k}$,

$(\lambda_i \neq \lambda_j)$ $\quad := p_1(s) \cdot p_2(s) \cdots p_k(s)$.

Since $p_A(A) = 0$, $\quad \mathbb{C}^n = N_{p(A)} = N_{p_1 \cdots p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k}$.

Thus, for any $x \in \mathbb{C}^n$, $\quad x = x_1 + \cdots + x_k$, $\quad x_j \in N_{p_j}$

Note: $x_j \in N_{p_j} \iff (A - \lambda_j I)^{n_j} x_j = 0$

$\iff x_j$ is a generalized e-vector for λ_j.

$\square$

Remark: Take any basis $\mathcal{B}_j$ of N_{p_j}. Then

$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is a basis of $\mathbb{C}^n$ consisting of

generalized eigenvectors of A.

Spectral theorem (refinement): $\mathbb{C}^n = N^{(1)} \oplus \cdots \oplus N^{(k)}$,

where $N^{(j)} = N_{(A-\lambda_j I)^{d_j}}$; $\quad d_j = \text{index}(\lambda_j)$

and $\dim N^{(j)}$ = algebraic multiplicity of λ_j.

Corollary: $M_A(s) = \prod_{j=1}^{k} (s-\lambda_j)^{d_j}$

It turns out that A (up to choice of basis) is completely determined by the dimensions of $N_1(\lambda), \dots, N_{d_\lambda}(\lambda)$ for each λ.

Theorem 6.12: Two matrices A, B are similar iff they have the same eigenvalues, and the dimensions of the corresponding eigenspaces are the same. That is, if for each λ_j, $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$,

where $N_m(\lambda_j) = N_{(A-\lambda_j I)^m}$, $M_m(\lambda_j) = N_{(B-\lambda_j I)^m}$.

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

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

"$\Leftarrow$" Let $\lambda = \lambda_j$ be an eigenvalue of A, $N_i := N_i(\lambda)$.

Goal: Construct a basis for N_d under which $A-\lambda I$ admits a nice matrix form ("Jordan canonical form")

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

Key lemma: The map $A-\lambda I$ carries over to a well-defined 1-1 map on quotient spaces:

$$A-\lambda I : N_{i+1}/N_i \longrightarrow N_i/N_{i-1}$$
$$\bar{x} \longmapsto \overline{(A-\lambda I)x}.$$

Proof: HW.

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

We'll construct our basis for N_d in "batches."

Let $\bar{x}_1, \ldots, \bar{x}_{\lambda_0}$ be a basis for N_d/N_{d-1} (so $x_1, \ldots, x_{\lambda_0}$ lin ind. in N_d).

By lemma, $(A-\lambda I)\bar{x}_1, \ldots, (A-\lambda I)\bar{x}_{\lambda_0}$ are lin. indep. in N_{d-1}/N_{d-2}.

Extend to a basis $\bar{x}_1', \ldots, \bar{x}_{\lambda_0}', \bar{x}_{\lambda_0+1}', \ldots, \bar{x}_{\ell_1}'$ of N_{d-1}/N_{d-2}.

Repeat this process: $(A-\lambda I)\bar{x}_1', \ldots, (A-\lambda I)\bar{x}_{\ell_1}'$ are lin. indep. in N_{d-1}/N_{d-2}.

$$\begin{array}{cccccccccccc}
x_1 & \xrightarrow{A-\lambda I} & x_1' & \longrightarrow & x_1'' & \longrightarrow & \cdots & \longrightarrow & x_1^{(d)} & \xrightarrow{A-\lambda I} & 0 \\
\vdots & & & & & & & & & & \\
x_{\lambda_0} & \longrightarrow & x_{\lambda_0}' & \longrightarrow & x_{\lambda_0}'' & \longrightarrow & \cdots & \longrightarrow & x_{\lambda_0}^{(d)} & \longrightarrow & 0 \\
 & & x_{\lambda_0+1}' & \longrightarrow & x_{\lambda_0}'' & \longrightarrow & \cdots & \longrightarrow & x_{\lambda_0+1}^{(d)} & \longrightarrow & 0 \\
 & & \vdots & & \vdots & & & & \vdots & & \\
 & & x_{\ell_1}' & \longrightarrow & x_{\ell_1}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_1}^{(d)} & \longrightarrow & 0 \\
 & & & & x_{\ell_1+1}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_1+1}^{(d)} & \longrightarrow & 0 \\
 & & & & \vdots & & & & \vdots & & \\
 & & & & x_{\ell_2}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_2}^{(d)} & \longrightarrow & 0 \\
 & & & & & & & & \vdots & & \\
 & & & & & & & & x_{\ell_d}^{(d)} & \longrightarrow & 0
\end{array}$$

Note that by this construction, A and B are similar to the same matrix $\Rightarrow$ A & B are similar. $\square$

The following is a generalization of the spectral mapping theorem.

<u>Theorem 6.14</u>: Let $A, B : X \to X$ be commuting maps, $\dim X < \infty$. Then there is a basis for X consisting of generalized eigenvectors of A & B.

<u>Proof</u>: Write $X = N^{(1)} \oplus \cdots \oplus N^{(k)}$, where each summand is a generalized eigenspace $N^{(j)} = N_{d_j}(\lambda_j) = \text{nullspace}\,(A - \lambda_j I)^{d_j}$.

<u>Claim</u>: $N^{(j)}$ is B-invariant.

Let $d = d_j$, $\lambda = \lambda_j$, $x \in N^{(j)}$. Then $(A - \lambda I)^d x = 0$.

$\Rightarrow (A - \lambda I)^d Bx = B(A - \lambda I)^d x = B \cdot 0 = 0 \Rightarrow Bx \in N^{(j)}$ ✓

<u>Conclusion</u>: $B|_{N^{(j)}} : N^{(j)} \to N^{(j)}$ and by the spectral theorem, $N^{(j)}$ has a basis of generalized e-vectors of B. But these are already generalized e-vectors of A. $\square$

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

$$\left(N_1^{(1)} \oplus \cdots \oplus N_{i_1}^{(1)}\right) \oplus \left(N_1^{(2)} \oplus \cdots \oplus N_{i_2}^{(2)}\right) \oplus \cdots \oplus \left(N_1^{(k)} \oplus \cdots \oplus N_{i_k}^{(k)}\right)$$

Remark: We can pick a basis $\mathcal{B}$ for $\mathbb{C}^n$ so that ${}_{\mathcal{B}}[A]_{\mathcal{B}}$ is a Jordan matrix. But we can't assume that ${}_{\mathcal{B}}[B]_{\mathcal{B}}$ is a Jordan matrix.

Here, by Jordan matrix, we mean $J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{bmatrix}$, where each J_i is a Jordan block.

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

Proof: Exercise

If we restrict Theorem 6.14 to diagonalizable maps, we can say more.

Theorem 6.16: Let $A, B : X \to X$ be diagonalizable and $AB = BA$. Then A, B are simultaneously diagonalizable, i.e., there is some matrix P such that

$A = P D_A P^{-1}$, and $B = P D_B P^{-1}$, where D_A, D_B are diagonal.

Proof: HW exercise. The approach is similar to Thm 6.14. $\square$

This means that $\mathbb{C}^n$ has a basis of common eigenvectors to both A and B.

Theorem 6.17: Every square matrix is similar to its transpose.

Proof: HW. Show that $A: X \to X$, $A': X' \to X'$ have the same eigenvalues, and $(A-\lambda I)^j$ & $[(A'-\lambda I)']^j$ have the same dimension. Apply Theorem 6.12. □

Application to differential equations:

(1) Consider a system of n linear ODEs: $\vec{x}' = A\vec{x}$.

Suppose A has eigenvalues $\lambda_1, \dots, \lambda_n$ and e-vectors $v_1, \dots, v_n$.

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

Solutions to $\vec{x}' = A\vec{x}$ 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 $\vec{x}(t) = C_1 e^{\lambda_1 t}\vec{v}_1 + \dots + C_n e^{\lambda_n t}\vec{v}_n$.

Matrix form: $\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} = e^{Dt}\vec{x}_0$.

Note: $\vec{x}_0 = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, we're using basis $\vec{v}_1, \dots, \vec{v}_n$.

With respect to basis $e_1, \dots, e_n$, $e^{Dt}\vec{x}_0$ becomes

$$e^{At}\vec{x}_0 = e^{PDP^{-1}t}\vec{x}_0 = (Pe^{D}P^{-1})\vec{x}_0.$$

Though $e^{At} = \sum_{i=0}^{\infty} \frac{1}{n!} A^n t^n$ is hard to compute, e^{Dt} and $Pe^{Dt}P^{-1}$ are easy to compute.

In summary, if A has n linearly independent e-vectors, then the general solution to $\vec{x}\,' = A\vec{x}$, $\vec{x}(0) = \vec{x}_0$ is

$$\vec{x}(t) = e^{At}\vec{x}_0 = Pe^{Dt}P^{-1}\vec{x}_0, \quad \text{where} \quad A = PDP^{-1}.$$

(2) Consider $\begin{cases} x_1' = -x_1 - x_2 \\ x_2' = x_1 - 3x_2 \end{cases}$ i.e., $A = \begin{pmatrix} -1 & -1 \\ 1 & -3 \end{pmatrix}$.

Check: $\lambda_1 = \lambda_2 = -2$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

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

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

Plug back in: $\vec{x}_2\,' = -2e^{-2t}(t\vec{v} + \vec{w}) + e^{-2t}\vec{v} = e^{-2t}(tA\vec{v} + A\vec{w})$.

Equate coeffs: te^{-2t}: $-2\vec{v} = A\vec{v} \Rightarrow (A + 2I)\vec{v} = \vec{0}$.

e^{-2t}: $\vec{v} - 2\vec{w} = A\vec{w} \Rightarrow (A + 2I)\vec{w} = \vec{v}$.

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

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

Or $\vec{x}(t) = e^{Jt}\vec{x}_0$, where $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, $\lambda = 2$.