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:

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

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

Thus, we can write

$$0 = c_0 w + c_1 Aw + \cdots + c_n A^n w = p(A) w$$

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

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

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 $N \to \infty$.

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 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 $\lambda_1, \dots, \lambda_k$ be pairwise distinct eigenvalues, with eigenvectors $v_1, \dots, v_k$ (all non-zero).

Suppose $\sum_{j=1}^{m} c_j v_j = 0$, where m is minimal, non-zero.

(so clearly, $c_j \neq 0$.)

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

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

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

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.

Thus, $v_1, \dots, v_m$ must be linearly independent. $\square$

<u>Corollary 6.2</u>: If A has n distinct eigenvalues, then it has n linearly independent eigenvectors.

In this case, the eigenvectors form a <u>basis</u> for X, and it is easy to compute $A^N x$, for any $x \in X$:

Write $x = \sum_{j=1}^{n} a_i v_i$ $\quad$ eigenvectors $v_1, \dots, v_n$.

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

<u>Theorem 6.3</u>: If the eigenvalues of A are $\lambda_1, \dots, \lambda_n$, then

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr} A \quad \text{and} \quad \prod_{i=1}^{n} \lambda_i = \det A.$$

<u>Proof</u>: <u>Claim</u>: $p_A(t) = t^n - (\operatorname{tr} A) t^{n-1} + \dots + (-1)^n \det A$.

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

<u>Note</u>: t^{n-1} coefficient $= -\sum_{i=1}^{n} \lambda_i$, $\quad$ constant term $= (-1)^n \prod_{i=1}^{n} \lambda_i$

To prove our claim, compute

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

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

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

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

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

Remark: 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.

Let's take this further: let $g(t) \in K[t]$ be any polynomial, say $g(t) = \sum_{i=1}^{n} a_i t^i$.

If $Av = \lambda v$, then $A^i v = \lambda^i v$

$$\Rightarrow g(A)v = \sum_{i=1}^{n} a_i A^i v = \sum_{i=1}^{n} a_i \lambda^i v = g(\lambda) v.$$

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

Theorem 6.4: ("Spectral mapping theorem"). Let A have eigenvalue λ, and let $g(t) \in K[t]$.

(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 μ be an eigenvalue of $g(A) \Leftrightarrow \det(g(A)-\mu I)=0$.

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

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

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

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

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 1: All eigenvalues are distinct.

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

Thus, for any $x\in X$, write $x=c_1v_1+\dots+c_nv_n$.

$$p_A(A)x = \sum_{i=1}^{n} p_A(A)c_iv_i = \sum_{i=1}^{n} p_A(\lambda_i)c_iv_i = \sum_{i=1}^{n} 0 = 0. \quad ✓$$

For the general case (non-distinct eigenvalues), we need an additional lemma:

<u>Lemma 6.6</u>: Let P and Q be polynomials with matrix coefficients:

$$P(t) = \sum P_j t^j, \qquad Q(t) = \sum Q_k t^k, \quad \text{and let } R = PQ.$$

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

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

<u>Proof</u>: Exercise.

Now, let $Q(t) = tI - A$, $P(t) = \left(P_{ij}(t)\right)$, $P_{ij}(t) = (-1)^{i+j} D_{ji}(t)$

where $D_{ji}(t)$ = 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 $\left(Q(t)\right)^{-1} = \frac{1}{\det P(t)} P(t)$.

Put $R(t) := P(t)Q(t) = \left(\det Q(t)\right) I = p_A(t) I$

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

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

<u>Examples</u>:

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

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

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

Thus, <u>every</u> vector is an eigenvector of A.

8

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

$p_A(t) = t^2 - 2t + 1 = (t-1)^2$, so $d_1 = d_2 = 1$.

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

$\Rightarrow x_1 + x_2 = 0 \Rightarrow v = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is an eigenvector (and every multiple is too). However, this is the <u>only</u> eigenvector.

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

<u>Proof</u>: Pick $x \in X$, and write $x = a_1 x_1 + \dots + a_n x_n$.

$$Ax = a_1 A x_1 + \dots + a_n A x_n = a_1 \lambda x_1 + \dots + a_n \lambda x_n = \lambda (a_1 x_1 + \dots + a_n x_n) = \lambda x. \qquad \square$$

<u>Remark</u>: Every 2×2 matrix with $\operatorname{tr} A = 2$, $\det A = 1$, 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 λ is an eigenvalue with multiplicity m, but only one eigenvector, v_1.

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

Since $\operatorname{rank}(A - \lambda I) = m - 1$, there is some v_2 such that

$(A - \lambda I) v_2 = v_1 \Rightarrow (A - \lambda I)^2 v_2 = 0$.

Similarly, we can find v_3 such that

$$(A-\lambda I)v_3 = v_2 \implies (A-\lambda I)^2 v_3 \neq 0 \text{ but } (A-\lambda I)^3 \vec{v}_3 = 0.$$

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

The <u>geometric multiplicity</u> of λ 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 <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>: $A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$, which has $\lambda_{1,2} = 1$, $v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

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

$$\implies \begin{cases} 2x_1 + 2x_2 = -1 \\ -2x_1 - 2x_2 = 1 \end{cases} \implies 2x_1 + 2x_2 = -1. \implies x_2 = -\tfrac{1}{2} - x_1$$

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

<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 eigenvectors (genuine or generalized) of A.

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

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

<u>Proof</u>: Let $I = \{ap + bq : a, b \in \mathbb{C}[t]\}$, and pick $d \in I$ with minimal degree.

<u>Claim 1</u>: 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 $p - md = r \in I$. But d had minimal degree in I. $\lightning$.

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

If not, it would have a root α, and since $d \mid p$ and $d \mid q$, α is a root of p and q. $\lightning$

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

<u>Lemma 6.9</u>: 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

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

<u>Proof</u>: Write $ap + bq = 1$ for $a, b \in \mathbb{C}[t]$.

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

Take $x \in N_{pq}$: $\quad a(A)p(A)x + b(A)q(A)x = x. \quad (*)$

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

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

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

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

The expression $(*)$ is

$$\begin{aligned} x &= x_p + x_q \\ &= b(A)q(A)x + a(A)p(A)x \in N_p + N_q. \end{aligned}$$

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

Suppose $x = x_p + x_q = x_p' + x_q'$.

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

$0 = a(A)p(A)y + b(A)q(A)y = Iy = y \quad \Rightarrow y = 0.$ $\square$

Corollary 6.10: Let $p_1, \dots, p_k \in \mathbb{C}[t]$ 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)

<u>Proof of Spectral theorem</u>: Pick $x \in \mathbb{C}^n$. The vectors $x, Ax, A^2x, \ldots, A^n x$ are linearly dependent, thus there is a polynomial of degree $\leq n$ such that $p(A)x = 0$.

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

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

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

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

$$x = x_{p_1} + \ldots + x_{p_R}, \qquad x_{p_i} \in N_{p_i}.$$

Note that each $x_{p_i} \in N_{p_i}$ is a generalized eigenvector. $\square$

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

Note that $\mathcal{I}$ is closed under addition and multiplication (of not just scalars, but polynomials too).

<u>Lemma</u>: $\mathcal{I}$ contains a unique monic polynomial $m = m_A$ of minimal degree, and all other polynomials in $\mathcal{I}$ are scalar multiples of m_A (i.e., $\mathcal{I}$ is a <u>principal ideal</u> of $\mathbb{C}[t]$).

<u>Proof</u>: Let $m \in \mathcal{I}$ have minimal degree. Clearly, m is unique, because if there were another, their difference would have strictly smaller degree.

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

<u>Def</u>: The <u>minimal polynomial</u> of a matrix A, denoted m_A, is the unique monic polynomial of minimal degree for which $m_A(A) = 0$.

Let $N_m = N_m(\lambda)$ be the nullspace of $(A - \lambda I)^m$.

Note that N_m consists of generalized eigenvectors, and

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

for some index d. Denote $d = d(\lambda)$ be the minimal index such that $N_{d-1} \subsetneq N_d = N_{d+1}$, called the <u>index</u> of the eigenvalue λ.

<u>Theorem 6.11</u>: Let A be an $n \times n$ matrix, with distinct eigenvalues $\lambda_1, \dots, \lambda_k$, with indices $d_1, \dots, d_k$. Then, the minimal polynomial of A is $m_A(t) = \prod_{i=1}^{k} (t - \lambda_i)^{d_i}$.

<u>Proof</u>: Exercise.

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

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

<u>Remark</u>: $\dim N^{(j)}$ is the algebraic multiplicity of λ_j. (This will be proved later.)

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

It turns out that A (up to choice of basis) is completely determined by the dimensions of $N_1, \dots, N_d$.

<u>Theorem 6.12</u>: Two matrices A, B are similar if and only if they have the same eigenvalues, <u>and</u> the dimensions of the corresponding eigenspaces are the same. That is, if A and B share the same eigenvalues $\lambda_1, \dots, \lambda_k$, and $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$ for each $j = 1, \dots, k$, 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.

It follows that λ is an eigenvalue of A iff it is an eigenvalue of B.

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

Goal: Construct a basis for N_d under which $A-\lambda I$ 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 λ.

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

$$\{x\} \longmapsto \{(A-\lambda I)x\}$$

Moreover, it is injective.

Proof: Exercise. (HW) $\square$

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 $\bar{x}_1, \ldots, \bar{x}_{l_0}$ be a basis for N_d/N_{d-1}.

By lemma, $(A-\lambda I)\bar{x}_1, \ldots, (A-\lambda I)\bar{x}_{l_0}$ are linearly independent in N_{d-1}/N_{d-2}.

Extend to a basis by adding $\bar{x}_{l_0+1}, \ldots, \bar{x}_{l_1} \in N_{d-1}/N_{d-2}$

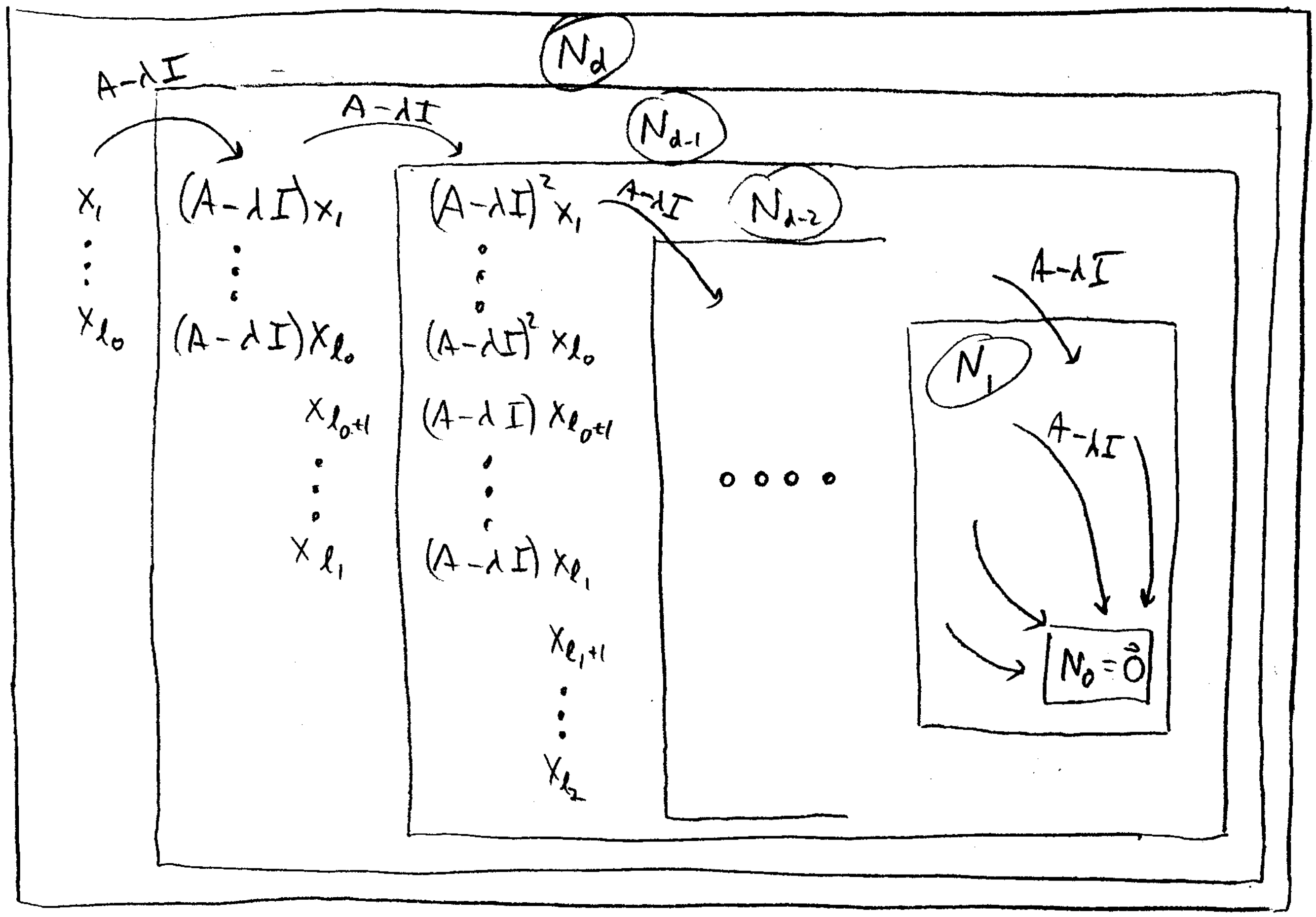
Repeat this process:

$(A-\lambda I)^2 \bar{x}_1, \ldots, (A-\lambda I)^2 \bar{x}_{\ell_0}, (A-\lambda I)\bar{x}_{\ell_0+1}, \ldots, (A-\lambda I)\bar{x}_{\ell_1}$ are linearly independent in N_{d-2}/N_{d-3}.

Extend to a basis by adding $\bar{x}_{\ell_1+1}, \ldots, \bar{x}_{\ell_2} \in N_{d-2}/N_{d-3}$

⋮

Eventually, this terminates because $N_0 = 0$.



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,

because:

$$x_1 \rightarrow (A-\lambda I)x_1 \rightarrow (A-\lambda I)^2 x_1 \rightarrow \cdots$$

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

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

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

Such a matrix is called a <u>Jordan block</u>.

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

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}.$$

<u>Jordan canonical form</u> of A. Since it <u>only</u> 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 6.12. □

The following is a generalization of the spectral mapping theorem:

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

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

<u>Claim</u>: B maps $N^{(j)}$ into $N^{(j)}$.

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

If λ is an eigenvalue of A, then the LHS of (✱) is zero.

Thus, $Bx \in N^{(j)}$. ✓

Now apply the spectral theorem to B, restricted to each $N^{(j)}$ separately. □

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 λ and $l \in X'$ an eigenvector for λ', then $(l, x) = 0$.

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

$$\lambda(l, v) = (l, \lambda v) = (l, Av) = (A'l, v) = (\lambda' l, v) = \lambda'(l, v).$$

Since $\lambda \neq \lambda'$, $(l, v) = 0$. $\square$

Theorem 6.17 has a useful application:

<u>Theorem 6.18</u>: Suppose A has distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding eigenvectors $v_1, \dots, v_n \in A$, and let $l_1, \dots, l_n$ be the corresponding eigenvectors in A'.

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

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

<u>Proof</u>: Exercise. (HW).

<u>Remark</u>: Theorems 6.17 and 6.18 together tell us that the eigenvectors $l_1, \dots, l_n \in A'$ is the <u>dual basis</u> to the eigenvectors $v_1, \dots, v_n \in A$.

<u>Def</u>: When A has linearly independent eigenvectors $v_1, \dots, v_n$, we say that A is <u>diagonalizable</u>, because it's Jordan canonical form is a diagonal matrix D. In this case, we can write $A = P^{-1} D P$, 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., $D = (\lambda_1 e_1, \dots, \lambda_n e_n)$, $P = (v_1, \dots, v_n)$.

To see this, note that

$$AP = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (\lambda_1 v_1, \dots, \lambda_n v_n)$$
$$= (\lambda_1 P e_1, \dots, \lambda_n P e_n)$$
$$= P(\lambda_1 e_1, \dots, \lambda_n e_n) = PD.$$

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] \arrow[d, "P"'] & \mathbb{R}^n \arrow[d, "P"] \\ \mathbb{R}^n \arrow[r, "D"] & \mathbb{R}^n \end{tikzcd}$$

<u>Applications to differential equations</u>

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

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

<u>Note</u>: $\vec{x}_i(t) = e^{\lambda_i} \vec{v}_i$ is a solution (easy to check this)

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

In matrix form, this is $\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = e^{Dt} \vec{x}_0$.

Here, $\vec{x}_0 = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ and we are using the basis $v_1, \dots, v_n$.

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

$$e^{At}\vec{x}_0 = e^{P^{-1}DPt}\vec{x}_0 = (P^{-1}e^{D}P)\vec{x}_0$$

While it may seem that $e^{At} = \sum_{i=0}^{\infty} \frac{1}{n!}A^n t^n$ 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 $\vec{x}' = A\vec{x}$, $\vec{x}(0) = \vec{x}_0$ is

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

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

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

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 $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ (Jordan Canonical form; here $\lambda = -2$)