# 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 $A$; $\lambda$ 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 $0 = c_0 w + c_1 Aw + \cdots + c_n A^n w$

$$= p(A)\, w$$

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

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

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

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 $v_1$ 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 $N \to \infty$.

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

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

<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.

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

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

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

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

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 $x = \sum_{j=1}^{n} a_i v_i$ 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>: $p_A(s) = s^n - (\operatorname{tr} A) s^{n-1} + \dots + (-1)^n \det A$.

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

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

To prove our claim, compute

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

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

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

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

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 $\lambda$ is an eigenvalue of $A$, then $\lambda^N$ is an eigenvalue of $A^N$.

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

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 $\lambda$, and let $g(s) \in K[s]$.

(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 $q(A) \Leftrightarrow \det(q(A) - \mu I) = 0$.

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

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

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

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

Remark: In the case when $q(s) = p_A(s)$, 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 $\lambda_i$ is a root of $p_A(s)$.
Thus, for any $x \in X$, write $x = c_1 v_1 + \dots + c_n v_n$.

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

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

Lemma 6.6: Let $P$ and $Q$ be polynomials with matrix coefficients:

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

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

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

Proof: Exercise.

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

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

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

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

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

Examples:

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

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

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

$\Rightarrow x_1 + x_2 = 0 \Rightarrow v = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ 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 $\lambda$, and $n$ linearly independent eigenvectors, then $A = \lambda I$

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

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

$\square$

<u>Remark</u>: Every $2\times 2$ 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, 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 $\operatorname{rank}(A - \lambda I) = m-1$, 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

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

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

Def: The algebraic multiplicity of an eigenvalue 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 $\lambda_1 = \lambda_2 = 1$, $v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

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

Example: Suppose $A$ is $11 \times 11$ with an e-value $\lambda$ 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 subspaces"

$$
\begin{tikzcd}
v_5 \arrow[r, "A-\lambda I"] & v_4 \arrow[r, "A-\lambda I"] & v_3 \arrow[r, "A-\lambda I"] & v_2 \arrow[r, "A-\lambda I"] & v_1 \arrow[r, "A-\lambda I"] & 0 \\
 & & w_3 \arrow[r] & w_2 \arrow[r] & w_1 \arrow[r] & 0 \\
 & & & x_2 \arrow[r] & x_1 \arrow[r] & 0 \\
 & & & & y_1 \arrow[r] & 0
\end{tikzcd}
$$

$\uparrow$ eigenvectors.

Remarks:

$N_1 := N_{A-\lambda I} = \operatorname{span}\{v_1, w_1, x_1, y_1\}$ $\qquad \dim N_2 = 4$

$N_2 := N_{(A-\lambda I)^2} = \operatorname{Span}\{v_2, w_2, x_2, v_1, w_1, x_1, y_1\}$ $\qquad \dim N_2 = 7$

$N_3 = N_{(A-\lambda I)^3} = \operatorname{Span}\{v_3, w_3, \ldots, x, y_1\}$ $\qquad \dim N_3 = 9$

$\vdots$

Note that: $N_1 \subsetneq N_2 \subsetneq N_3 \subsetneq N_4 \subsetneq N_5 = N_6 = \cdots$

$\dim N_i = 4 < 7 < 9 < 10 < 11 = 11 = \cdots$

* It's a fundamental result that there will always be a full set of generalized eigenvectors that form a basis for $\mathbb{C}^n$. This is the Spectral theorem.

<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 $p, q \in \mathbb{C}[s]$ with no common roots. 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 <u>quotient</u>, $r$ 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 $p = md + r$ with $\deg r < \deg d$.

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

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

If not, it would have a root $\alpha$, and since $d \mid p$ & $d \mid q$, then $(s - \alpha)$ divides $p$ & $q$. $\lightning$.

Thus, $d$ is <u>constant</u>; 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 $\mathbb{C}$, & $p, q \in \mathbb{C}[s]$ with no common roots. Let $N_p$, $N_q$, $N_{pq}$ be the nullspaces 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 $a, b \in \mathbb{C}[s]$.

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

Multiply by $x \in N_{pq}$: $\quad \underbrace{a(A)p(A)x}_{\substack{\text{in } N_q \text{ because} \\ a(A)[p(A)q(A)x]=0}} + \underbrace{b(A)q(A)x}_{\substack{\text{in } N_p \text{ because} \\ b(A)[p(A)q(A)x]=0}} = x \qquad (*)$

[Here, we're using that $f(A)g(A) = g(A)f(A) \;\; \forall f, g \in \mathbb{C}[s]$.]

The expression $(*)$ is $\quad x = \underset{b(A)q(A)x}{x_p} + \underset{a(A)p(A)x}{x_q}$

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

Suppose $x = x_p + x_q = x_p' + x_q'$. Put $y := x_p - x_p' = x_q - x_q' \in N_p \cap N_q$

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

$\Rightarrow y = 0$.

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

<u>Corollary 6.10</u>: Let $p_1, \dots, p_k \in \mathbb{C}[s]$ 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 $N_{p_1 \cdots p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k}$.

<u>Proof</u>: Exercise. (Induct on $k$.)

<u>Proof of Spectral theorem</u>: 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 $m(A)x = \prod_{j=1}^{R} (A - \lambda_j I)^{m_j} x = 0$, each $\lambda_j$ is e-value.

$$= \prod_{j=1}^{R} \underbrace{(A - \lambda_j I)^{m_j}}_{p_j(A)} x = 0$$

<u>Remarks</u>:
- $p_j(s) = (s - \lambda_j)^{m_j}$ and $\lambda_i \neq \lambda_j$
- The $x$ above is in $N_{p_1 \cdots p_R} = N_{p_1} \oplus \cdots \oplus N_{p_R}$.
- If $x = x_{p_1} + \cdots + x_{p_R}$, with $x_{p_i} \in N_{p_i}$, then each $x_{p_i}$ is a <u>generalized eigenvector</u>: $(A - \lambda_i I)^{m_i} x = 0$.

$\square$

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

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

<u>Lemma</u>: $I$ 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 <u>principal ideal</u> of $\mathbb{C}[s]$.)

<u>Proof</u>: Let $m \in I$ have minimal degree.

<u>Uniqueness</u>: Clear. [If there were 2, subtract them.]

<u>Existence</u>: Suppose $p \in I$ were not a multiple of $m$.

By division algorithm, write $p = q m + r$, $\deg r < \deg m$.

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

<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

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

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

$N_{d-1} \subsetneq N_d = N_{d+1}$, called the <u>index</u> of the eigenvalue $\lambda$.

<u>Theorem 6.11</u>: If $A$ is $n \times n$ & has distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ with indices $d_1, \ldots, d_k$, the its minimal polynomial is

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

<u>Proof</u>: Exercise.

Denote $N_{d_j}(\lambda_j)$ by $N^{(j)}$. The spectral theorem can be stated as follows: $\mathbb{C}^n = N^{(1)} \oplus N^{(2)} \oplus \cdots \oplus N^{(k)}$.

<u>Remark</u>: $\dim N^{(j)}$ 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 <u>invariant</u> under $A$.

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

<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 for each eigenvalue $\lambda_j$, $\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$, where

$N_m(\lambda_j) = \text{nullspace of } (A-\lambda_j I)^m, \quad M_m(\lambda_j) = \text{nullspace of } (B-\lambda_j I)^m.$

<u>Proof</u>: "$\Rightarrow$" 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 $N_i := \text{nullspace}(A-\lambda I)^i$.

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

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

<u>Lemma</u>: The map $A-\lambda I$ carries over to a well-defined map on the quotient spaces:

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

Moreover, it is injective.

<u>Proof</u>: 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 $N_d$ in "batches."

Let $\bar{x}_1, \dots, \bar{x}_{\ell_0}$ be a basis for $N_d/N_{d-1}$ (so $x_1, \dots, x_{\ell_0}$ lin. indep. in $N_d$).

By lemma, $\underbrace{(A-\lambda I)\bar{x}_1}_{\bar{x}_1'}, \dots, \underbrace{(A-\lambda I)\bar{x}_{\ell_0}}_{\bar{x}_{\ell_0}'}$ are linearly independent in $N_{d-1}/N_{d-2}$.

Extend to a basis $\bar{x}_1', \dots, \bar{x}_{\ell_0}', \bar{x}_{\ell_0+1}', \dots, \bar{x}_{\ell_1}'$ of $N_{d-1}/N_{d-2}$.

Repeat this process:

$\underbrace{(A-\lambda I)\bar{x}_1'}_{\bar{x}_1''}, \dots, \underbrace{(A-\lambda I)\bar{x}_{\ell_1}'}_{\bar{x}_{\ell_1}''}$ are linearly independent in $N_{d-1}/N_{d-2}$

<u>Picture of this:</u>

$$
\begin{array}{ccccccccccc}
x_1 & \xrightarrow{A-\lambda I} & x_1' & \xrightarrow{A-\lambda I} & x_1'' & \xrightarrow{A-\lambda I} & \cdots & \xrightarrow{A-\lambda I} & x_1^{(d)} & \xrightarrow{A-\lambda I} & 0 \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \\
x_{\ell_0} & \xrightarrow{A-\lambda I} & x_{\ell_0}' & \xrightarrow{A-\lambda I} & x_{\ell_0}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_0}^{(d)} & \longrightarrow & 0 \\
 & & x_{\ell_0+1}' & \longrightarrow & x_{\ell_0+1}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_0+1}^{(d)} & \longrightarrow & 0 \\
 & & \vdots & & \vdots & & & & \vdots & & \\
 & & x_{\ell_1}' & \longrightarrow & x_{\ell_1}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_1}^{(d)} & \longrightarrow & 0 \\
 & & & & x_{\ell_1+1}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_1+1}^{(d)} & \longrightarrow & 0 \\
 & & & & \vdots & & & & \vdots & & \\
 & & & & x_{\ell_2}'' & \longrightarrow & \cdots & \longrightarrow & x_{\ell_2}^{(d)} & \longrightarrow & 0 \\
 & & & & & & & & \vdots & & \\
 & & & & & & & & x_{\ell_d}^{(d)} & \longrightarrow & 0
\end{array}
$$

Remarks: • $N_d(\lambda)$ = space spanned by all of these vectors

= set of generalized eigenvectors for $\lambda$.

• Algebraic multiplicity of $\lambda$ = $\dim N_d(\lambda)$ = total # vectors shown.

• 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 <u>invariant subspace</u> of $A-\lambda I$.

• The <u>matrix</u> $A-\lambda I$ restricted to this subspace has the form:

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \bigcirc \\ & & 0 & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$

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

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \bigcirc \\ & & \lambda & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}$$

<u>Reason</u>: 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 wrt basis $x_1, \dots, x_d$,

$(A-\lambda I)x_j = x_{j-1} \Rightarrow Ax_j = \lambda x_j + x_{j-1} \Rightarrow$ row $j$ is $\begin{bmatrix} \lambda \\ 1 \end{bmatrix}$

If we use a basis of generalized eigenvectors for $\mathbb{C}^n$, then the matrix for $A$ is <u>block-diagonal</u>, consisting of Jordan blocks.

Such a matrix is called the

<u>Jordan canonical form</u> of A.

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}_{n\times n}$$

Since it depends only on the eigenvalues and eigenspace dimensions, if two matrices A & 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:

<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 & generalized eigenvectors of A and B.

<u>Proof</u>: Write $X = N^{(1)} \oplus \dots \oplus N^{(k)}$, 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>: B maps $N^{(j)}$ into $N^{(j)}$.

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

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

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

Conclusion: $B\big|_{N^{(j)}} : N^{(j)} \longrightarrow N^{(j)}$ and by the spectral theorem, $N^{(j)}$ has a basis of generalized eigenvectors of $B$. But these 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 $A': X' \longrightarrow X'$ 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 $A: X \longrightarrow X$ 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, $Av = \lambda v$ and $A'l = \lambda' l$

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

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

Application of Theorem 6.17:

Theorem 6.18: Suppose $A$ has distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $v_1, \ldots, v_n \in A$ and let $l_1, \ldots, l_n$ 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 $a_i = \dfrac{(l_i, x)}{(l_i, v_i)}$.

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 $A = P^{-1} D P$, 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., $D = (\lambda_1 e_1, \ldots, \lambda_n e_n)$, $P = (v_1, \ldots, v_n)$.

To see this, note that

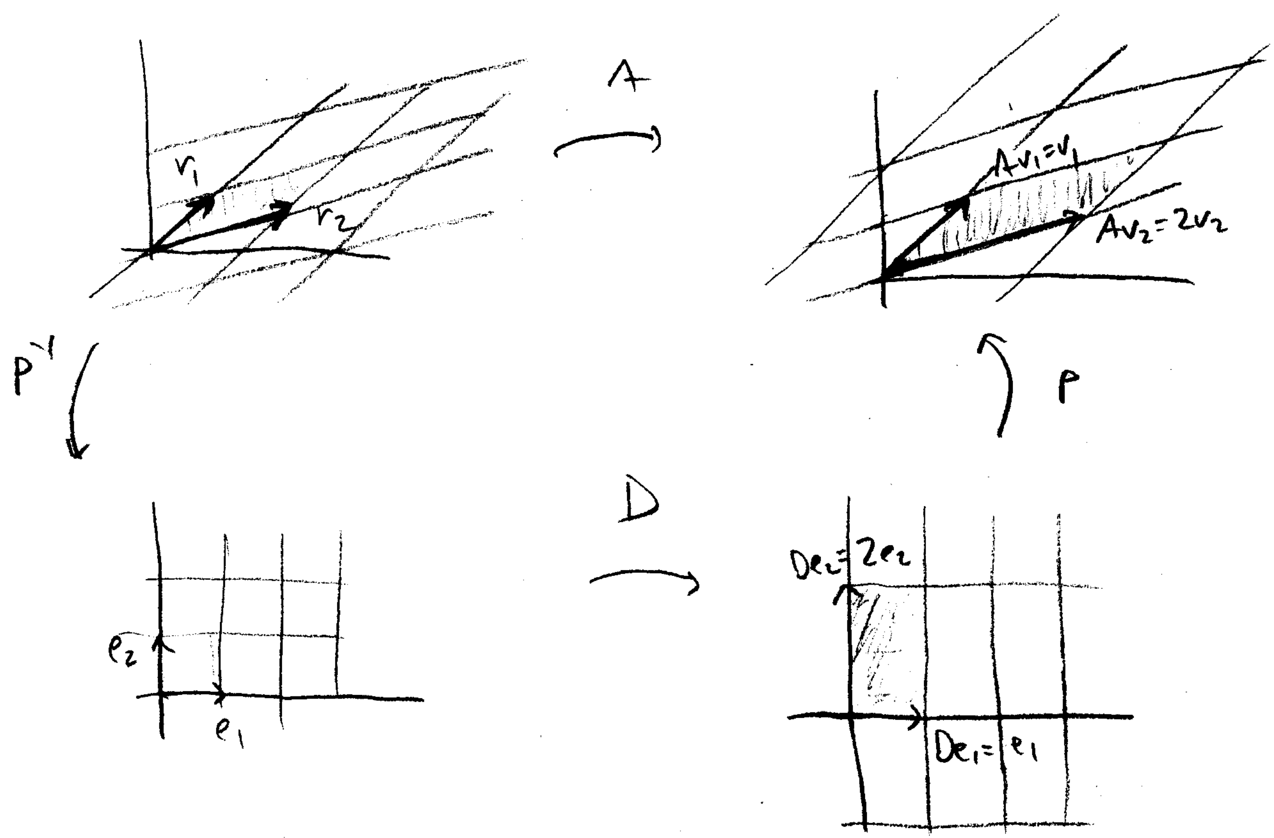
$$
\begin{aligned}
AP = A(v_1,\dots,v_n) = (Av_1,\dots,Av_n) &= (\lambda_1 v_1,\dots,\lambda_n v_n) \\
&= (\lambda_1 Pe_1,\dots,\lambda_n Pe_n) \\
&= P(\lambda_1 e_1,\dots,\lambda_n e_n) = PD.
\end{aligned}
$$

Diagram:

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

Example: $\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \Rightarrow \begin{matrix} \lambda_1 = 1 & v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \lambda_2 = 2 & v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \end{matrix}$

$\quad\quad\quad\quad A \quad = \quad P \quad\quad D \quad\quad P^{-1}$



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 $n$ linearly independent eigenvectors $v_1, \dots, v_n$.

Note: $\vec{x}_i(t) = e^{\lambda_i}\vec{v}_i$ 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's well-known that the nullspace is $n$-dimensional.

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

In matrix form, this is $\begin{bmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} = e^{Dt}\vec{x}_0$

Here, $\vec{x}_0 = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}$ and we're using basis $\vec{v}_1, \dots, \vec{v}_n$.

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

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

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 $\vec{x}(t) = C e^{-2t}\vec{v}_1 + C_2 e^{-2t}(t\vec{v}_1 + \vec{v}_2)$.

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