Section 4: Spectral theory

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
`http://www.math.clemson.edu/~macaule/`

Math 8530, Advanced Linear Algebra

Assumptions and definitions

This section is all about eigenvalues and eigenvectors of a linear map.

In most introductory courses, students learn that repeated eigenvalues often lead to "missing eigenvectors."

However, that's only half of the story – we'll see how there's always a basis of generalized eigenvectors.

This basis leads to the Jordan canonical form, and we'll see how this arises in linear differential equations.

Throughout, we will assume that A is an $n \times n$ matrix over K. Thus, it represents an endomorphism of a vector space $X \cong K^n$.

We will assume that K is algebraically closed, which means that every non-constant polynomial has a root in K.

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

Definition

If $Av = \lambda v$ for some nonzero vector v and scalar $\lambda \in K$, then v is an eigenvector and λ is an eigenvalue.

Existence of eigenvectors

Proposition 4.1

A has an eigenvector.

An example

Remark

$A - \lambda I$ is noninvertible iff $\det(A - \lambda I) = 0$. That is, λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$, and the corresponding eigenvector is any $v \neq 0$ in $N_{A-\lambda I}$.

Let's compute the eigenvalues and eigenvectors of $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$.

Linear independence of eigenvectors

Proposition 4.2

Eigenvectors of A corresponding to distinct eigenvalues are linearly independent.

Diagonalizability

Proposition 4.3

If X has a basis of eigenvectors of A, then A is similar to a diagonal matrix. We say that A is diagonalizable.

The characteristic polynomial

Throughout, $A\colon X \to X$ will be an $n \times n$ matrix over an algebraically closed field K.

Definition

The characteristic polynomial of A is

$$p_A(t) = \det(tI - A).$$

$$\det(tI-A) = \begin{vmatrix} t-a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1(n-1)} & -a_{1n} \\ -a_{21} & t-a_{22} & -a_{23} & \cdots & -a_{2(n-1)} & -a_{2n} \\ -a_{31} & -a_{32} & t-a_{33} & \cdots & -a_{3(n-1)} & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -a_{(n-1)1} & -a_{(n-1)2} & -a_{(n-1)3} & \cdots & t-a_{(n-1)(n-1)} & -a_{(n-1)n} \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & -a_{n(n-1)} & t-a_{nn} \end{vmatrix}$$

Remarks

- Recall that $\det M = \displaystyle\sum_{\pi \in S_n} \operatorname{sgn}(\pi) m_{\pi(1),1} m_{\pi(2),2} \cdots m_{\pi(n),n}$.
- The characteristic polynomial has degree n, and its roots are the eigenvalues of A.

Determinant and trace, revisited

Proposition 4.4

If the eigenvalues of A are $\lambda_1, \dots, \lambda_n$, then

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i, \qquad \det A = \prod_{i=1}^{n} \lambda_i$$

This follows from the following two observations:

$$\det(tI - A) = \begin{vmatrix} t - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1(n-1)} & -a_{1n} \\ -a_{21} & t - a_{22} & -a_{23} & \cdots & -a_{2(n-1)} & -a_{2n} \\ -a_{31} & -a_{32} & t - a_{33} & \cdots & -a_{3(n-1)} & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -a_{(n-1)1} & -a_{(n-1)2} & -a_{(n-1)3} & \cdots & t - a_{(n-1)(n-1)} & -a_{(n-1)n} \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & -a_{n(n-1)} & t - a_{nn} \end{vmatrix}$$

$$\det M = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) m_{\pi(1),1} m_{\pi(2),2} \cdots m_{\pi(n),n}.$$

Polynomials of matrices

Remark

If $Av = \lambda v$, then $A^k v = \lambda^k v$ for all $k \in \mathbb{N}$.

Actually, much more is true:

Spectral mapping theorem

If λ is an eigenvalue of A, then for any polynomial $q(t)$,

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

(b) conversely, every eigenvalue of $q(A)$ has this form.

Corollary 4.5

Every eigenvalue of $p_A(A)$ is zero.

Actually, even much more is true:

Cayley-Hamilton theorem

Every matrix satisfies its characteristic polynomial. That is, $p_A(A) = 0$.

Lemma 4.6 (exercise)

Let P and Q be polynomials with matrix coefficients:

$$P(t) = P_n t^n + \cdots + P_1 t + P_0, \qquad Q(t) = Q_m t^m + \cdots + Q_1 t + Q_0.$$

Their product is a polynomial

$$\begin{aligned} R(t) = P(t)Q(t) &= (P_n t^n + \cdots + P_1 t + P_0)(Q_m t^m + \cdots + Q_1 t + Q_0) \\ &= R_{n+m} t^{n+m} + \cdots + R_1 t + R_0, \end{aligned}$$

where $R_k = \sum_{i+j=k} P_i Q_j$. Moreover, if A commutes with the Q_i's, then $P(A)Q(A) = R(A)$.

We will apply this to the polynomial $Q(t) = tI - A$, and so $\det Q(t) = p_A(t)$.

Let C_{ji} be the (j,i) cofactor of $Q(t)$. By Cramer's theorem, $\det Q(t) I = (C_{ji})Q(t)$.

If we let $P(t) = (C_{ji})$, then

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

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

$$R(A) = P(A)Q(A) = \det Q(A) I = p_A(A) = 0.$$

The minimal polynomial

Throughout, $A\colon X \to X$ will be an $n \times n$ matrix over an algebraically closed field K.

Let I be the set of polynomials

$$I = \{p(t) \in K[t] \mid p(A) = 0\}.$$

This is an ideal of $K[t]$ since it's closed under addition, subtraction, and multiplication.

Since $K[t]$ is a principal ideal domain (PID), I is generated by a single element.

That is, $I = \langle m_A(t) \rangle$, for some monic polynomial $m_A(t)$, called the minimal polynomial of A.

All polynomials $p(t)$ such that $p(A) = 0$ are multiples of $m_A(t)$.

Let's verify existence and uniqueness of $m_A(t)$ without using ring theoretic ideas.

2×2 examples

Examples

1. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
2. $A = \begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix}$

Remark

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

3×3 examples

Suppose A is a 3×3 matrix and $p_A(t) = (t-1)^3$. Since $m_A(t)$ divides $p_A(t)$, there are three possibilities:

1. $m_A(t) = t - 1$
2. $m_A(t) = (t-1)^2$
3. $m_A(t) = (t-1)^3$.

Generalized eigenvectors

Suppose λ is an eigenvalue with multiplicity m, but only one eigenvector, $v_1 \in X$. Then

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

Big idea

We can *always* find some $v_2 \in X$ such that

$$(A-\lambda I)v_2 = v_1, \qquad \Longrightarrow \qquad (A-\lambda I)^2 v_2 = 0.$$

Similarly, we can find $v_3 \in X$ such that

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

Definition

A vector $v \in X$ is a generalized eigenvector of A with eigenvalue λ if $(A-\lambda I)^m v = 0$ for some $m \geq 1$. The "genuine" eigenvectors are when $m = 1$.

2×2 examples, revisited

Examples

1. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
2. $A = \begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix}$

Invariant subspaces and block diagonal matrices

Throughout, X is an n-dimensional vector space over an algebraically closed field K.

Definition

An invariant subspace of $A\colon X \to X$ is any $Y \leq X$ for which $A(Y) \subseteq Y$.

Suppose $X = Y \oplus Z$, both A-invariant.

If $y_1, \dots, y_k$ and $z_{k+1}, \dots, z_n$ are bases for Y and Z, then the matrix of A with respect to

$$y_1, \dots, y_k, z_{k+1}, \dots, z_n$$

is block-diagonal. It is easy to see how this extends to a sum of A-invariant subspaces,

$$X = Y_1 \oplus \cdots \oplus Y_\ell.$$

Suppose we have a collection $v_1, \dots, v_m$ of generalized eigenvectors:

$$v_{m-1} = (A-\lambda I)v_m, \quad v_{m-2} = (A-\lambda I)^2 v_m, \quad \dots, \quad v_2 = (A-\lambda I)^{m-2} v_m, \quad v_1 = (A-\lambda I)^{m-1} v_m.$$

Notice that $Y = \mathrm{Span}(v_1, \dots, v_m)$ is invariant under both $(A - \lambda I)$ and A.

Next, we will explore what happens when we have multiple genuine eigenvectors, and the invariant subspaces that arise.

Our 11×11 running example

Suppose $A\colon X \to X$ has characteristic polynomial $p_A(t) = (t-\lambda)^{11}$, and $\dim N_{A-\lambda I} = 4$.

Here is one such possibility for the generalized eigenvectors:

$$v_5 \xmapsto{A-\lambda I} v_4 \xmapsto{A-\lambda I} v_3 \xmapsto{A-\lambda I} v_2 \xmapsto{A-\lambda I} v_1 \xmapsto{A-\lambda I} 0$$

$$w_3 \xmapsto{A-\lambda I} w_2 \xmapsto{A-\lambda I} w_1 \xmapsto{A-\lambda I} 0$$

$$x_2 \xmapsto{A-\lambda I} x_1 \xmapsto{A-\lambda I} 0$$

$$y_1 \xmapsto{A-\lambda I} 0$$

What invariant subspaces do you see?

Let $N_j := N_{(A-\lambda I)^j}$. Notice that

$$\cdots = N_6 = N_5 \supsetneq N_4 \supsetneq N_3 \supsetneq N_2 \supsetneq N_1 \supsetneq 0.$$

The anatomy of an eigenvalue

Key idea

For any $A\colon X \to X$, there is always a basis of generalized eigenvectors of A.

Definition & preview

The algebraic multiplicity of λ is:

- the largest k such that $(t-\lambda)^k$ is a factor of $p_A(t)$
- the maximum number of linearly independent generalized λ-eigenvectors of A
- the number of diagonal entries of λ in the Jordan canonical form.

The geometric multiplicity of λ is:

- $\dim N_{A-\lambda I}$
- the maximum number of linearly independent genuine λ-eigenvectors of A
- the number of Jordan blocks corresponding to λ.

The index of λ is:

- the smallest d such that $N_d = N_{d+1}$
- the "length of the longest chain" of generalized eigenvectors
- the largest m such that $(t-\lambda)^m$ is a factor of $m_A(t)$
- the size of the largest Jordan block corresponding to λ.

A key technical lemma

Lemma 4.7 (HW exercise)

The map $A-\lambda I$ is a well-defined injective map on quotient spaces:

$$A-\lambda I\colon N_{j+1}/N_j \longrightarrow N_j/N_{j-1}, \qquad\qquad A-\lambda I\colon \bar{x} \longmapsto \overline{(A-\lambda I)x}.$$

Therefore, $\dim(N_{j+1}/N_j) \leq \dim(N_j/N_{j-1})$.

$$v_5 \xmapsto{A-\lambda I} v_4 \xmapsto{A-\lambda I} v_3 \xmapsto{A-\lambda I} v_2 \xmapsto{A-\lambda I} v_1 \xmapsto{A-\lambda I} 0$$

$$w_3 \xmapsto{A-\lambda I} w_2 \xmapsto{A-\lambda I} w_1 \xmapsto{A-\lambda I} 0$$

$$x_2 \xmapsto{A-\lambda I} x_1 \xmapsto{A-\lambda I} 0$$

$$y_1 \xmapsto{A-\lambda I} 0$$

$$\cdots \;=\; N_6 \;=\; N_5 \;\supsetneq\; N_4 \;\supsetneq\; N_3 \;\supsetneq\; N_2 \;\supsetneq\; N_1 \;\supsetneq\; 0.$$

The idea of the spectral theorem

Throughout, assume K is algebraically closed, and $\dim X = n$. A generalized eigenvector of A is any $v \in X$ such that $(A-\lambda I)^m v = 0$ for some $m \geq 1$.

Spectral theorem

Let $A \colon X \to X$ be linear. Then X has a basis of generalized eigenvectors of A.

Recall our running example, a linear map with $p_A(t) = (t-\lambda)^{11}$, and $\dim N_{A-\lambda I} = 4$:

$$v_5 \xmapsto{A-\lambda I} v_4 \xmapsto{A-\lambda I} v_3 \xmapsto{A-\lambda I} v_2 \xmapsto{A-\lambda I} v_1 \xmapsto{A-\lambda I} 0$$

$$w_3 \xmapsto{A-\lambda I} w_2 \xmapsto{A-\lambda I} w_1 \xmapsto{A-\lambda I} 0$$

$$x_2 \xmapsto{A-\lambda I} x_1 \xmapsto{A-\lambda I} 0$$

$$y_1 \xmapsto{A-\lambda I} 0$$

If $N_j := N_{(A-\lambda I)^j}$, then

$$\cdots = N_6 = N_5 \supsetneq N_4 \supsetneq N_3 \supsetneq N_2 \supsetneq N_1 \supsetneq 0.$$

Supporting lemmas

Lemma 4.8

Let $p, q \in K[t]$ be co-prime. Then we can write $ap + bq = 1$ for some $a, b \in K[t]$.

Lemma 4.9

Let $A \colon X \to X$, and $p, q \in K[t]$ be co-prime. If N_p, N_q, N_{pq} are the nullspaces of $p(A)$, $q(A)$, and $p(A)q(A)$, then

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

Corollary 4.10

If $p_1, \dots, p_k \in K[t]$ are pairwise co-prime, and $N_{p_1 \cdots p_k}$ is the nullspace of $p_1(A) \cdots p_k(A)$, then

$$N_{p_1 \cdots p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k}.$$

Generalized eigenspaces

Definition

Let λ be an eigenvalue of $A\colon X \to X$ with index $d_\lambda = \text{index}(\lambda)$. The generalized eigenspace of λ is

$$E_\lambda := N_{(A-\lambda I)^{d_\lambda}} = \bigcup_{j=1}^{\infty} N_{(A-\lambda I)^j}.$$

Spectral theorem (stronger)

Let $A\colon X \to X$ be linear, with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Then

$$X = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Goals

Assume K is algebraically closed, and $\dim X = n$. Last time, we proved the following:

Spectral theorem

Let $A \colon X \to X$ be linear. Then

$$X = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k},$$

where $E_{\lambda_j} = \bigcup_{m=1}^{\infty} N_{(A-\lambda_j I)^m}$ is the generalized eigenspace of λ_j.

We motivated it with a running example, a map with $p_A(t) = (t-\lambda)^{11}$, and $\dim N_{A-\lambda I} = 4$:

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

However, we haven't actually proven that the generalized eigenvectors have this structure. Now, we will show how to explicitly construct such a basis.

We'll also see why the generalized eigenspace structure determines the similarity class of A.

Generalized eigenspaces characterize similarity

Let $A\colon X \to X$ have eigenvalue λ of degree d_λ. For each $m = 1, 2, \dots$, define

$$N_m(\lambda) = N_{(A-\lambda I)^m}, \qquad \text{and note that} \quad E_\lambda = \bigcup_{m=1}^{\infty} N_m(\lambda).$$

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

For another $B\colon X \to X$ with eigenvalue λ, denote its eigen-subspaces by $M_m(\lambda) = N_{(B-\lambda I)^m}$.

Theorem 4.11

The linear maps A and B are similar if and only if for each eigenvalue λ,

$$\dim N_m(\lambda) = \dim M_m(\lambda), \qquad \text{for all } m = 1, 2, \dots.$$

The "$\Rightarrow$" implication is easy. Let $A = PBP^{-1}$.

Then $(A - \lambda I)^m = P(B - \lambda I)^m P^{-1}$, and similar maps have the same nullity.

For the "$\Leftarrow$" implication, we need to construct a basis for E_λ under which $A - \lambda I$ (and hence $B - \lambda I$) admits a nice matrix form.

This is the Jordan canonical form.

Basis construction (algebraic description)

Lemma 4.7 (HW)

The map $A-\lambda I$ is a well-defined injective map on quotient spaces, i.e.,

$$A-\lambda I\colon N_{j+1}/N_j \hookrightarrow N_j/N_{j-1}, \qquad\qquad A-\lambda I\colon \bar{x} \longmapsto \overline{(A-\lambda I)x}.$$

Therefore, $\dim(N_{j+1}/N_j) \leq \dim(N_j/N_{j-1})$.

We will construct our basis in batches, from "left-to-right", starting with $N_d = E_\lambda$.

Let $\bar{x}_1, \dots, \bar{x}_{\ell_0}$ be a basis for N_d/N_{d-1}.

Apply $A-\lambda I$, to get $(A-\lambda I)\bar{x}_j \mapsto \bar{x}'_j$.

The vectors $\bar{x}'_1, \dots, \bar{x}'_{\ell_0}$ are linearly independent in N_{d-1}/N_{d-2}. Extend to a basis $\bar{x}'_1, \dots, \bar{x}'_{\ell_1}$.

Apply $A-\lambda I$, to get $(A-\lambda I)\bar{x}'_j \mapsto \bar{x}''_j$.

The vectors $\bar{x}''_1, \dots, \bar{x}''_{\ell_1}$ are linearly independent in N_{d-2}/N_{d-3}. Extend to a basis $\bar{x}''_1, \dots, \bar{x}''_{\ell_2}$.

Repeat this process, until we reach the genuine eigenvectors. The collection of representatives we've constructed is a basis for E_λ.

Basis construction (visualization)

Key points

$$A - \lambda I \colon N_{j+1}/N_j \hookrightarrow N_j/N_{j-1} \qquad \Longrightarrow \qquad \dim(N_{j+1}/N_j) \leq \dim(N_j/N_{j-1}).$$

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

Jordan blocks

Spectral theorem

Let $A\colon X \to X$ be linear. Then

$$X = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k},$$

where $E_{\lambda_j} = \displaystyle\bigcup_{m=1}^{\infty} N_{(A-\lambda_j I)^m}$ is the generalized eigenspace of λ_j.

Moreover, each E_{λ_j} is a direct sum of subspaces invariant under both A and $(A-\lambda_j I)$.

Let's recall an old example where λ has algebraic multiplicity $\dim E_\lambda = 11$ and geometric multiplicity $\dim N_{A-\lambda I} = 4$.

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

The matrix of A with respect to this is block-diagonal, consisting of Jordan blocks.

Jordan canonical form

A Jordan block is a matrix of the form

$$J_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$$

Every matrix A is similar to a Jordan matrix — a block-diagonal matrix of Jordan blocks:

$$J = \begin{bmatrix} J_{\lambda_1,1} & & & & & & \\ & \ddots & & & & & \\ & & J_{\lambda_1,n_1} & & & & \\ & & & \ddots & & & \\ & & & & J_{\lambda_k,1} & & \\ & & & & & \ddots & \\ & & & & & & J_{\lambda_k,n_k} \end{bmatrix}$$

This is called the Jordan normal form, or Jordan canonical form (JCF) of A.

Summary of key spectral concepts

Two linear maps $A, B\colon X \to X$ are similar iff they have the same Jordan canonical form.

For each eigenvalue λ, the algebraic multiplicity of λ is the:

- degree of $(t-\lambda)$ in $p_A(t)$
- maximum number of linearly independent generalized λ-eigenvectors of A
- number of diagonal entries of λ in the Jordan canonical form.

The geometric multiplicity of λ is the:

- $\dim N_{A-\lambda I}$
- maximum number of linearly independent genuine λ-eigenvectors of A
- number of Jordan blocks corresponding to λ.

The index of λ is the:

- smallest d such that $N_d = N_{d+1}$ (length of the largest "chain")
- degree of $(t-\lambda)$ in $m_A(t)$
- size of the largest Jordan block corresponding to λ.

A is diagonalizable if:

- X has a basis of genuine eigenvectors
- $m_A(t)$ has no repeated roots
- the Jordan canonical form is a diagonal matrix.

Commuting maps

Lemma 4.12

Let $A, B\colon X \to X$ be commuting linear maps, and $E_\lambda = \bigcup_{j=1}^{\infty} N_{(A-\lambda I)^j}$, the generalized λ-eigenspace of A. Then E_λ is B-invariant.

Theorem 4.13

Let $A, B\colon X \to X$ be commuting linear maps. There is a basis for X consisting of generalized eigenvectors of A and B.

Corollary 4.14

Let $A, B\colon X \to X$ be commuting diagonalizable linear maps. Then they are simultaneously diagonalizable. That is, for some invertible $P\colon X \to X$,

$$A = PD_AP^{-1} \qquad \text{and} \qquad B = PD_BP^{-1}.$$

Application: ODEs with repeated roots

Recall how to solve the differential equation $y'' - 3y' + 2y = 0$:

- Look for a solution of the form $y(t) = e^{rt}$.
- Plug back in to get $e^{rt}(r^2 - 3r + 2) = 0$, and so $r = 1$ or $r = 2$.
- The general solution is thus $y(t) = C_1 e^t + C_2 e^{2t}$.

A "problem case" occurs when the "characteristic equation" has repeated roots.

For example, consider $y'' - 2\lambda y' + \lambda^2 y = 0$.

The same process gives $r_1 = r_2 = \lambda$, so we only get one solution, $y_1(t) = e^{\lambda t}$.

However, the solution space is two-dimensional. It turns out that $y_2(t) = te^{\lambda t}$ is also a solution.

Now, we'll see how this arises as a generalized eigenfunction of a differential operator.

The derivative operator

Clearly, $y_1(t) = e^{\lambda t}$ is an eigenfunction of $D = \frac{d}{dt}$.

Equivalently, it is in $N_{D-\lambda I}$, and solves the ODE

$$(D - \lambda I)y = 0 \qquad \Leftrightarrow \qquad \big(\tfrac{d}{dt} - \lambda\big)y = 0 \qquad \Leftrightarrow \qquad y' - \lambda y = 0.$$

Generalized eigenfunctions in $N_{(D-\lambda I)^2}$ are solutions to the second order ODE

$$(D - \lambda I)^2 y = 0, \qquad \Leftrightarrow \qquad \big(\tfrac{d}{dt} - \lambda\big)^2 y = 0, \qquad \Leftrightarrow \qquad y'' - 2\lambda y' + \lambda^2 y = 0$$

It is easy to see that $y_2(t) = te^{\lambda t}$ is in $N_{(D-\lambda I)^2}$, because

$$D(y_2) = D\big(te^{\lambda t}\big) = e^{\lambda t} + \lambda te^{\lambda t} = y_1 + \lambda y_2.$$

Similarly, $y_3(t) = \frac{1}{2!}t^2 e^{\lambda t}$ is in $N_{(D-\lambda I)^3}$, because

$$D(y_3) = D\big(\tfrac{1}{2!}t^2 e^{\lambda t}\big) = te^{\lambda t} + \lambda \tfrac{1}{2!}t^2 e^{\lambda t} = y_2 + \lambda y_3.$$

Repeating in this manner, we see that the generalized eigevectors for D are:

$$\cdots \xmapsto{D-\lambda I} \tfrac{1}{4!}t^4 e^{\lambda t} \xmapsto{D-\lambda I} \tfrac{1}{3!}t^3 e^{\lambda t} \xmapsto{D-\lambda I} \tfrac{1}{2!}t^2 e^{\lambda t} \xmapsto{D-\lambda I} te^{\lambda t} \xmapsto{D-\lambda I} e^{\lambda t} \xmapsto{D-\lambda I} 0$$

The generalized eigenspace of D for eigenvalue λ is thus

$$E_\lambda = \big\{p(t)e^{\lambda t} \mid p \in K[t]\big\}.$$

Systems of linear differential equations

Consider the linear system $x' = Ax$:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

It is easy to check that if $Av = \lambda v$, then $x(t) = e^{\lambda t} v$ is a solution.

Thus, the general solution is

$$x(t) = C_1 e^{3t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + C_2 e^{-t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} C_1 e^{3t} + C_2 e^{-t} \\ 2C_1 e^{3t} - 2C_2 e^{-t} \end{bmatrix}.$$

Now, consider an example that has only one eigenvector:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad x_1(t) = e^{\lambda t} v_1 = e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

In an ODE course, one is taught to look for a solution of the form

$$x_2(t) = te^{-2t} v + e^{-2t} w,$$

and solve for v and w.

We'll see that what we're really doing is finding generalized eigenvectors of A.

Solving $\mathbf{x}' = \mathbf{A}\mathbf{x}$ with repeated eigenvalues

Suppose that $Av = \lambda v$, and so $x_1(t) = e^{\lambda t} v$ is a solution. Consider

$$x_2(t) = te^{\lambda t} v + e^{\lambda t} w,$$

and plug this back into $x' = Ax$:

- $Ax_2 = te^{\lambda t} Av + e^{\lambda t} Aw$.
- $x_2' = (e^{\lambda t} v + \lambda te^{\lambda t} v) + \lambda e^{\lambda t} w$.

Equate like terms and divide by $e^{\lambda t}$:

- $te^{\lambda t}$: $Av = \lambda v$
- $e^{\lambda t}$: $Aw = v + \lambda w$.

In other words, $v = v_1$ is the eigenvector, and $w = v_2$ a generalized eigenvector. The general solution is

$$x(t) = C_1 x_1(t) + C_2 x_2(t) = C_1 e^{\lambda t} v_1 + C_2 e^{\lambda t}(tv_1 + v_2).$$

In summary, if the generalized eigenvectors of A are

$$v_2 \xmapsto{A-\lambda I} v_1 \xmapsto{A-\lambda I} 0$$

then the generalized eigenvectors of $A - \frac{d}{dt}$ are

$$\cdots \xmapsto{A-\frac{d}{dt}} e^{\lambda t}\left(\tfrac{t^2}{2!} v_1 + tv_2 + v_3\right) \xmapsto{A-\frac{d}{dt}} e^{\lambda t}(tv_1 + v_2) \xmapsto{A-\frac{d}{dt}} e^{\lambda t} v_1 \xmapsto{A-\frac{d}{dt}} 0$$

A Jordan matrix perspective

Formally, suppose we have the system $x' = Ax$, and $A = PJP^{-1}$.

$$(P^{-1}x)' = J(P^{-1}x), \qquad \text{let } z = P^{-1}x \quad \Leftrightarrow \quad x = Pz.$$

Now, we just have to analyze $z' = Jz$ for a Jordan matrix.

The solution is

$$z = e^{\lambda t} \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ & 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k-2}}{(k-2)!} \\ & & 1 & t & \cdots & \frac{t^{k-3}}{(k-3)!} \\ & & & \ddots & \ddots & \vdots \\ & & & & 1 & t \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_{n-1} \\ C_n \end{bmatrix} = e^{Jt}c.$$

It is easy to extend this to one where J has multiple Jordan blocks.

Finishing our example

Let's return to our example of $x' = Ax$, with only one eigenvector:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad x_1(t) = e^{\lambda t} v_1 = e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The Jordan canonical form $A = PJP^{-1}$ is

$$\begin{bmatrix} -1 & -1 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

The solution is $x = Pz$, where $z = e^{\lambda t} e^{Jt} c$:

$$x(t) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} e^{-2t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = e^{-2t} \begin{bmatrix} 1 & t+1 \\ 1 & t \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} C_1 e^{-2t} + C_2 e^{-2t}(t+1) \\ C_1 e^{-2t} + C_2 t e^{-2t} \end{bmatrix}.$$

Notice that we can rearrange terms to get this into a familiar form:

$$x(t) = C_1 e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + C_2 e^{-2t} \left(t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = C_1 e^{-2t} v_1 + C_2 e^{-2t} (t v_1 + v_2).$$

In other words, the generalized eigenvectors are:

$$e^{-2t} \left(t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \xmapsto{A - \frac{d}{dt}} e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \xmapsto{A - \frac{d}{dt}} \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$