10 Eigenbases and the Jordan Form

Change of basis is a powerful tool, and often, we would like to work in as natural a basis as possible.

> **Guiding Question**
> Given a linear operator, how can we find a basis in which the matrix is as nice as possible?

10.1 Review

Last time, we learned about eigenvectors and eigenvalues of linear operators, or more concretely, matrices, on vector spaces. An **eigenvector** is a (nonzero) vector sent to itself, up to scaling, under the linear operator, and its associated **eigenvalue** is the scaling factor; that is, if $A\vec{v} = \lambda\vec{v}$ for some scalar λ, $\vec{v}$ is an eigenvector with associated eigenvalue λ.

If there exists an eigenbasis[30], then in that basis, the linear operator $P^{-1}AP$[31] will simply become $\begin{pmatrix} \lambda_1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda_n \end{pmatrix}$,

since each basis vector $\vec{v}_i$ is sent to $\lambda_i \vec{v}_i$. So having an eigenbasis is equivalent to the matrix A being similar to a diagonal matrix.

In order to concretely find the eigenvectors, it is easier to first find the eigenvalues, which are the roots of the **characteristic polynomial** $p_A(t) = \det(tI_n - A)$. Each root λ of p_A has at least one corresponding eigenvector. The eigenvectors for λ are precisely the nonzero vectors in $\ker(\lambda I_n - A)$.

10.2 The Characteristic Polynomial

Let's start with an example.

> **Example 10.1**
> If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $p_A(t) = t^2 - (a+d)t + (ad - bc)$.

In general, for an $n \times n$ matrix,

$$p_A(t) = t^n - (a_{11} + \cdots + a_{nn})t^{n-1} + (-1)^n \det(A).$$

The coefficient of t^{n-1} is the sum of the entries on the diagonal, called the **trace** of A.

Because the characteristic polynomial can be written as a determinant, it can be defined for general linear operators without specifying a basis, and so each of the coefficients are basis-independent. In particular, we get that

$$\text{Tr}(P^{-1}AP) = \text{Tr}(A).$$

> **Guiding Question**
> What can go wrong when hunting for an eigenbasis? How can we fix this?

> **Example 10.2**
> Over the real numbers $\mathbb{R}$, take $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. The characteristic polynomial is $p_A(t) = t^2 - 2\cos\theta + 1$,
> which has no real roots (unless $\theta = \pi$.) Geometrically, that makes sense, because under a rotation by $\theta \neq k\pi$, every vector will end up pointing an a different direction than it initially was, so there should be no real eigenvectors.

Over a general field F, it is certainly possible for the characteristic polynomial not to have any roots at all; in order to fix this issue, we work over a field like $\mathbb{C}$,[32] where every degree n polynomial always has n roots (with

---

[30]A basis consisting of eigenvectors
[31]This comes from the change of basis formula
[32]Fields where every non-constant polynomial has roots are called **algebraically closed**.

multiplicity). So $p_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$, where the λ_i can repeat. For the rest of the lecture, we will only consider linear operators on vector spaces over $\mathbb{C}$, which takes care of the first obstacle of finding eigenvalues.

However, even over $\mathbb{C}$, not every linear operator has an eigenbasis.

Example 10.3

Consider $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. The characteristic polynomial is $p_A(t) = t^2$, so if A were similar to some diagonal matrix, it would be similar to the zero matrix; this would mean that A would be the zero matrix, and thus A cannot be diagonalizable.

In other words, $p_A(t)$ only has one root, 0, so any eigenvector would be in $\ker(0I_2 - A) = \text{Span}(\vec{e}_1)$. So there is only a one-dimensional space of eigenvectors, which is not enough to create an eigenbasis.

In some sense, which we will make precise later on in this lecture, this is the most important counterexample for why linear operators can be nondiagonalizable.

Proposition 10.4

Given an $n \times n$ matrix A, eigenvectors $\vec{v}_1, \cdots, \vec{v}_k$, and distinct eigenvalues $\lambda_1, \cdots, \lambda_k$, the vectors $\vec{v}_i$ are all linearly independent.

Proof. Let's prove this by induction.

Base Case. If $k = 1$, by the definition of an eigenvector, $\vec{v}_k \neq 0$ so $\{\vec{v}_i\}$ is linearly independent.

Inductive Hypothesis. Suppose the proposition is true for $k - 1$.

Inductive Step. Now, suppose the proposition is not true for k. Then there exist coefficients a_i such that

$$\sum a_i \vec{v}_i = 0.$$

Applying A to both sides, we get

$$\sum a_i \lambda_i \vec{v}_i = 0,$$

which is another linear relation between the k vectors. Subtracting k times the first relation from the second one results in the linear relation

$$\sum a_i(\lambda_i - \lambda_k) = \sum_{i=0}^{i=k-1} a_i(\lambda_i - \lambda_k) = 0.$$

Since in the last term $\lambda_k - \lambda_k = 0$, while $\lambda_i - \lambda_k \neq 0$ for $i \neq k$ since the λ_i are distinct, we obtain a linear relation between $\vec{v}_1, \cdots, \vec{v}_{k-1}$, which is a contradiction of the inductive hypothesis. Thus, $\{v_i\}_{i=0,\cdots,k}$ is linearly independent. $\square$

Corollary 10.5

Consider a matrix A. If the characteristic polynomial is

$$p_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$$

where each λ_i is distinct, A will have an eigenbasis and will thus be diagonalizable.

Proof. Each eigenvalue must have at least one eigenvector. Taking $\vec{v}_1, \cdots, \vec{v}_n$ to be eigenvectors for $\lambda_1, \cdots, \lambda_n$. Since there are n eigenvectors, which is the same as the dimension of the vector space, and by Proposition 10.4 they are linearly independent, they form an eigenbasis and A is diagonalizable. $\square$

If there are repeated roots, then there will not necessarily be enough eigenvectors to form a basis. Luckily for us, it is usually true that a matrix will be diagonalizable.[33]

---

[33]More concretely, the space of $n \times n$ square matrices can be thought of as a metric space, and the non-diagonalizable matrices will be a set of measure zero. In particular, the diagonalizable matrices are **dense** in the space of all square matrices. Intuitively, given a non-diagonalizable matrix, perturbing the entries by a little bit will perturb the roots a little bit, making them non-distinct.

In general, $p_A(t) = (t - \lambda_1)^{e_1} \cdots (t - \lambda_k)^{e_k}$, where the λ_i are distinct. Let $V_{\lambda_i} = \ker(\lambda_i I - A)$. Any vector $v \in V_{\lambda_i}$ is an eigenvector with eigenvalue λ_i. We know that for each i, $\dim V_{\lambda_i} \geq 1$. Using our proposition, given a basis for each subspace V_{λ_i}, if there are enough to get n total vectors, combining all the bases would give an eigenbasis for A, since they would all be linearly independent.[34]

10.3   Jordan Form

Keeping in mind the matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we have the following question.

> **Guiding Question**
> If a matrix is *not* diagonalizable, what is nicest form it can take on under a change of basis?

Let's see a class of matrices that always have the issue of repeated eigenvalues.

> **Definition 10.6**
> Given $a \geq 1$ and $\lambda \in F$, let the Jordan block be an $a \times a$ matrix
>
> $$J_a(\lambda) = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ 0 & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda \end{pmatrix}$$
>
> with λ on the diagonal and 1s above each λ.[a]
>
> ---
> [a]The notation here is a little different from the textbook.

> **Example 10.7**
> For $\lambda = 1, 2, 3$, we get
>
> $$(\lambda), \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

For $J_a(\lambda)$, the characteristic polynomial is $(t - \lambda)^a$, and when $a > 1$, the only eigenvalue is $\vec{e}_1$ so it will not be diagonalizable. Although this these Jordan blocks are very specific matrices, in some sense they are exactly the sources of all the problems.

> **Example 10.8** $(J_4(0))$
> The matrix $J_4(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Applying it to the basis vectors gives a "chain of vectors"
>
> $$\vec{e}_4 \mapsto \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto 0,$$
>
> where each basis vector is mapped to the next.

This leads us to the main theorem.

---
[34]Computationally, it is simple to find basis vectors, and in a more computational class we would go further in-depth on finding these.

> **Theorem 10.9** (Jordan Decomposition Theorem)
> Given a linear operator $T : V \longrightarrow V$, where $\dim V = n$, there exists a basis $\vec{v}_1, \cdots, \vec{v}_n$, there exist pairs $(a_1, \lambda_1), \cdots, (a_r, \lambda_r)$ such that the matrix of T in this basis is a **block diagonal** matrix
>
> $$\begin{pmatrix} J_{a_1}(\lambda_1) & & & \\ & J_{a_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{a_r}(\lambda_r) \end{pmatrix},$$
>
> where all other entries are 0.

While it is not possible to diagonalize every linear operator, it is possible to write them as a block diagonal matrix with Jordan blocks on the diagonal. Additioanlly, these Jordan blocks are unique up to rearrangement. This block diagonal matrix is called the Jordan decomposition of the linear operator T.

Student Question. *Do the a_i correspond to the exponents of the roots?*

Answer. *Not quite, as we will see promptly.*

Let's continue with some examples.

> **Example 10.10** ($n = 4$)
>
> If we have $a_1 = 4$, then the Jordan decomposition will look like $\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$. If $a_1 = 3$ and $a_2 = 1$,
>
> then it will look like $\begin{pmatrix} \lambda_1 & 1 & 0 & \\ 0 & \lambda_1 & 1 & \\ 0 & 0 & \lambda_1 & \\ & & & \lambda_2 \end{pmatrix}$. For $a_1 = 2$ and $a_2 = 2$, we have $\begin{pmatrix} \lambda_1 & 1 & & \\ 0 & \lambda_1 & & \\ & & \lambda_2 & 1 \\ & & 0 & \lambda_2 \end{pmatrix}$. Where
>
> $a_1 = 2, a_2 = 1$, and $a_3 = 1$, we would have $\begin{pmatrix} \lambda_1 & 1 & & \\ 0 & \lambda_1 & & \\ & & \lambda_2 & \\ & & & \lambda_3 \end{pmatrix}$, and for $a_1, a_2, a_3, a_4 = 1$, we would just
>
> get $\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{pmatrix}$, which is really just a diagonal matrix.

Essentially, every matrix is similar to some Jordan decomposition matrix, where it is diagonalizable if and only if each $a_i = 1$.

The characteristic polynomial of T will be $(t - \lambda_1)^{a_1} \cdots (t - \lambda_r)^{a_r}$. These exponents a_i are not quite the same as the exponents e_i from before, since the λ_i in the characteristic polynomial of T can repeat. However, for eigenvalues equal to λ_j, the sum of all the exponents will in fact be e_j.

From the characteristic polynomial of a matrix, it is not possible to precisely figure out the Jordan decomposition, but it does provide some amount of information. Next class, we will continue seeing what information we get from the Jordan Decomposition Theorem.

MIT OpenCourseWare
https://ocw.mit.edu

Resource: Algebra I Student Notes
Fall 2021
Instructor: Davesh Maulik
Notes taken by Jakin Ng, Sanjana Das, and Ethan Yang