Section 5.3

Diagonalization

Motivation: Difference equations

Now do multiply matrices

Many real-word (linear algebra problems):

- Start with a *given situation* (v_0) and
- want to know **what happens after some time** (iterate a transformation):

$$v_n = Av_{n-1} = \ldots = A^n v_0.$$

- Ultimate question: *what happens in the long run* (find v_n as $n \to \infty$)

Old Example

Recall our example about *rabbit populations*: using eigenvectors was easier than matrix multiplications, but ...

- Taking *powers of diagonal* matrices is easy!
- Working with *diagonalizable matrices* is also easy.
- We need to use the eigenvalues and eigenvectors of the dynamics.

Powers of Diagonal Matrices

If D is diagonal

Then D^n is also diagonal, the diagonal entries of D^n are the *nth powers of the diagonal* entries of D

Example

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \qquad M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix},$$

$$D^2 = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \qquad M^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix},$$

$$\vdots \qquad \vdots$$

$$D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \qquad M^n = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & 0 & \frac{1}{3^n} \end{pmatrix}.$$

Powers of Matrices that are Similar to Diagonal Ones

When is A is not diagonal?

Example

Let $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$. Compute A^n. Using that

$$A = PDP^{-1} \quad \text{where} \quad P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

From the first expression:

$$\begin{aligned}
A^2 &= (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PD^2P^{-1} \\
A^3 &= (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1} = PDID^2P^{-1} = PD^3P^{-1} \\
&\vdots \\
A^n &= PD^nP^{-1}
\end{aligned}$$

Closed formula in terms of n: easy to compute

Plug in P and D:

$$A^n = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2^{n+1} - 3^n & -2^{n+1} + 2 \cdot 3^n \\ 2^n - 3^n & -2^n + 2 \cdot 3^n \end{pmatrix}.$$

Diagonalizable Matrices

Definition

An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix:

$$A = PDP^{-1} \qquad \text{for } D \text{ diagonal.}$$

Important

$$\text{If } A = PDP^{-1} \text{ for } D = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix} \text{ then}$$

$$A^k = PD^K P^{-1} = P \begin{pmatrix} d_{11}^k & 0 & \cdots & 0 \\ 0 & d_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn}^k \end{pmatrix} P^{-1}.$$

So diagonalizable matrices are *easy to raise to any power*.

Diagonalization

The Diagonalization Theorem

An $n \times n$ matrix A is **diagonalizable** if and only if A has n *linearly independent eigenvectors*.

In this case, $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the *corresponding eigenvalues* (in the same order).

Important

- If A has *n distinct eigenvalues* then A **is diagonalizable**.
  Fact 2 in 5.1 lecture notes: eigenvectors with distinct eigenvalues are always linearly independent.
- If A is diagonalizable matrix it *need not have n distinct* eigenvalues though.

Problem: Diagonalize $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$.

The *characteristic polynomial* is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\,\lambda + \det(A) = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

Therefore the *eigenvalues are* 2 *and* 3. Let's compute some eigenvectors:

$$(A - 2I)x = 0 \iff \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = 2y$, so $v_1 = \binom{2}{1}$ is an eigenvector with eigenvalue 2.

$$(A - 3I)x = 0 \iff \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = y$, so $v_2 = \binom{1}{1}$ is an eigenvector with eigenvalue 3.

The eigenvectors v_1, v_2 are *linearly independent*, so the Diagonalization Theorem says

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Problem: Diagonalize $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

The *characteristic polynomial* is

$$f(\lambda) = \det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 1)^2(\lambda - 2).$$

Therefore the *eigenvalues are* 1 *and* 2, with respective multiplicities 2 and 1.

First compute the 1-eigenspace:

$$(A - I)x = 0 \iff \begin{pmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric vector form is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Hence a basis for the 1-eigenspace is

$$\mathcal{B}_1 = \{v_1, v_2\} \quad \text{where} \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now let's compute the 2-eigenspace:

$$(A-2I)x=0 \iff \begin{pmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & -1 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = 3z, y = 2z$, so an eigenvector with eigenvalue 2 is

$$v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Note that v_1, v_2 form a basis for the 1-eigenspace, and v_3 has a distinct eigenvalue. Thus, the eigenvectors v_1, v_2, v_3 are linearly independent and the Diagonalization Theorem says

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

In this case: there are 3 linearly independent eigenvectors and only 2 distinct eigenvalues.

How to **diagonalize a matrix** A:

1. *Find the eigenvalues* of A using the characteristic polynomial.
2. *Compute a basis* $\mathcal{B}_\lambda$ for each λ-eigenspace of A.
3. If there are fewer than n total vectors in the union of all of the eigenspace bases $\mathcal{B}_\lambda$, then the matrix is not diagonalizable.
4. *Otherwise*, the n vectors $v_1, v_2, \ldots, v_n$ in your eigenspace bases are linearly independent, and $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_i is the eigenvalue for v_i.

Problem: Show that $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is **not diagonalizable**.

The *characteristic polynomial* is

$$f(\lambda) = \det(A - \lambda I) = (\lambda - 1)^2.$$

Let's compute the 1-eigenspace:

$$(A - I)x = 0 \iff \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x = 0.$$

A basis for the 1-eigenspace is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$; solution has only one free variable!

Conclusion:

- All eigenvectors of A are multiples of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- So A has only one linearly independent eigenvector
- If A was diagonalizable, there would be *two linearly independent eigenvectors*!

Poll

Poll

Which of the following matrices *are diagonalizable*, and why?

A. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ B. $\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$ C. $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ D. $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

Matrix D is *already diagonal*!

Matrix B is *diagonalizable* because it has two distinct eigenvalues.

Matrices A and C are *not diagonalizable*: Same argument as previous slide:

All eigenvectors are multiples of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Non-Distinct Eigenvalues

Definition

Let λ be an eigenvalue of a square matrix A. The **geometric multiplicity** of λ is the *dimension of the λ-eigenspace*.

Theorem

Let λ be an eigenvalue of a square matrix A. Then

$$1 \leq (\text{the geometric multiplicity of } \lambda) \leq (\text{the algebraic multiplicity of } \lambda).$$

- Note: If λ is an eigenvalue, then the λ-eigenspace has dimension at least 1.
- ...but it might be smaller than what the characteristic polynomial suggests. The intuition/visualisation is *beyond the scope of this course*.

Multiplicities all one

If there are n eigenvalues all with algebraic multiplicity 1 (so does the geometric multiplicities), then their corresponding *eigenvectors are linearly independent*. Therefore A is diagonalizable.

Non-Distinct Eigenvalues

(Good) examples

From *previous exercises* we know:

Example

The matrix $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ has characteristic polynomial

$$f(\lambda) = -(\lambda - 1)^2(\lambda - 2).$$

The matrix $B = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$ has characteristic polynomial

$$f(\lambda) = (1 - \lambda)(4 - \lambda) + 2 = (\lambda - 2)(\lambda - 3).$$

Matrix A	Geom. M.	Alg. M.
$\lambda = 1$	2	2
$\lambda = 2$	1	1

Matrix B	Geom. M.	Alg. M.
$\lambda = 2$	1	1
$\lambda = 3$	1	1

Thus, *both matrices are diagonalizable*.

Example

The matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has characteristic polynomial $f(\lambda) = (\lambda - 1)^2$.
We showed before that the *1-eigenspace has dimension* 1 and A was **not diagonalizable**. The geometric multiplicity is smaller than the algebraic.

Eigenvalue	Geometric	Algebraic
$\lambda = 1$	1	2

The Diagonalization Theorem (Alternate Form)

Let A be an $n \times n$ matrix. The following are equivalent:

1. A is **diagonalizable**.
2. The *sum of the geometric multiplicities* of the eigenvalues of A equals n.
3. The sum of all algebraic multiplicities is n. And for each eigenvalue, the *geometric and algebraic* multiplicity are equal.

Applications to Difference Equations

Let $D = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$.

Start with a vector v_0, and let $v_1 = Dv_0$, $v_2 = Dv_1, \ldots, v_n = D^n v_0$.

Question: What happens to the v_i*'s* for different starting vectors v_0?

Answer: Note that D is diagonal, so

$$D^n \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1^n & 0 \\ 0 & 1/2^n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b/2^n \end{pmatrix}.$$

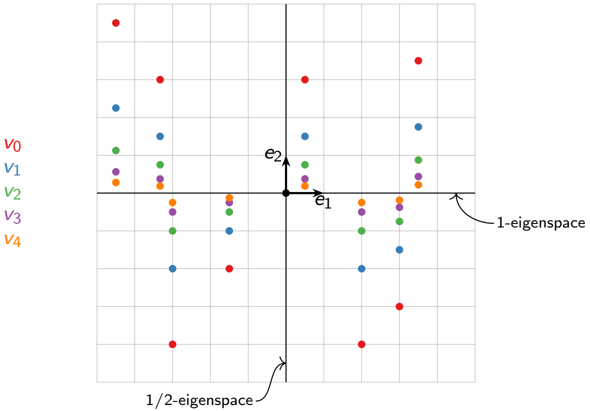
If we start with $v_0 = \begin{pmatrix} a \\ b \end{pmatrix}$, then

- the x-coordinate equals the initial coordinate,
- the y-coordinate gets halved every time.

Applications to Difference Equations

Picture

$$D\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b/2 \end{pmatrix}$$



So all vectors get *"collapsed into the x-axis"*, which is the 1-eigenspace.

Let $A = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$.

Start with a vector v_0, and let $v_1 = Av_0$, $v_2 = Av_1, \ldots, v_n = A^n v_0$.

Question: What happens to the v_i*'s* for different starting vectors v_0?

Matrix Powers: This is a diagonalization question. Bottom line: $A = PDP^{-1}$ for

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

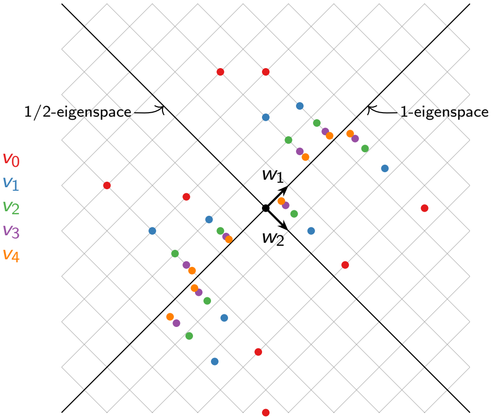
Hence $v_n = PD^n P^{-1} v_0$.
Details: The characteristic polynomial is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\,\lambda + \det(A) = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right).$$

We compute eigenvectors with eigenvalues 1 and 1/2 to be, respectively,

$$w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad w_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

$A^n = PD^nP^{-1}$ acts on the usual coordinates of v_0 in the same way that D^n acts on the $\mathcal{B}$-coordinates, where $\mathcal{B} = \{w_1, w_2\}$.



So all vectors get *"collapsed into the 1-eigenspace"*.

Extra: Proof Diagonalization Theorem

Why is the Diagonalization Theorem true?

A diagonalizable implies A has n linearly independent eigenvectors: Suppose $A = PDP^{-1}$, where D is diagonal with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $v_1, v_2, \ldots, v_n$ be the columns of P. They are linearly independent because P is invertible. So $Pe_i = v_i$, hence $P^{-1}v_i = e_i$.

$$Av_i = PDP^{-1}v_i = PDe_i = P(\lambda_i e_i) = \lambda_i Pe_i = \lambda_i v_i.$$

Hence v_i is an eigenvector of A with eigenvalue λ_i. So the columns of P form n linearly independent eigenvectors of A, and the diagonal entries of D are the eigenvalues.

A has n linearly independent eigenvectors implies A is diagonalizable: Suppose A has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let P be the invertible matrix with columns $v_1, v_2, \ldots, v_n$. Let $D = P^{-1}AP$.

$$De_i = P^{-1}APe_i = P^{-1}Av_i = P^{-1}(\lambda_i v_i) = \lambda_i P^{-1}v_i = \lambda_i e_i.$$

Hence D is diagonal, with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Solving $D = P^{-1}AP$ for A gives $A = PDP^{-1}$.