Section 6.2

Diagonalization

Objective

- What is a diagonalizable matrix?
- How to determine if a matrix is diagonalizable?
- How to diagonalize a matrix?
- How to compute powers of matrix using diagonalization?
- How to solve linear recurrence relation using diagonalization?

A 2x2 diagonalizable matrix

Example 6.2.2.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \qquad \boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}$$

diagonalizable   diagonalizes $\boldsymbol{A}$

$$\boxed{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}$$

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}} \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}}_{\boldsymbol{P}} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}$$

$\boldsymbol{P} \rightarrow$ diagonal

What is a diagonalizable matrix?

Definition 6.2.1

A square matrix $\boldsymbol{A}$ is called diagonalizable if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix.

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

diagonal

diagonalizable

We say: the matrix $\boldsymbol{P}$ diagonalizes $\boldsymbol{A}$

A 3x3 diagonalizable matrix

Example 6.2.2.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad \boldsymbol{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}$$

diagonalizable

diagonalizes $\boldsymbol{B}$

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{\boldsymbol{B}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}}_{\boldsymbol{P}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

A non-diagonalizable matrix

Example 6.2.2.3

We will introduce a systematic way to determine whether a matrix is diagonalizable

$$\boldsymbol{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$ not diagonalizable

Cannot find a matrix $\boldsymbol{P}$ that diagonalizes $\boldsymbol{M}$.

Prove by contradiction

Suppose there exist an invertible $\boldsymbol{P}$ such that

$$\boldsymbol{P}^{-1}\boldsymbol{M}\boldsymbol{P} = \text{Diagonal matrix.}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

Derive that: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ * & * \end{pmatrix}$ contradicts that $\boldsymbol{P}$ is invertible

How to tell whether a matrix is diagonalizable?

Example 6.2.2

$$A = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

diagonalizable

two eigenvalues : 1 and 0.95

two eigenvectors : $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

linearly independent

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

diagonalizable

two eigenvalues : 3 and 0

three eigenvectors : $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

linearly independent

$$M = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$

not diagonalizable

one eigenvalue : 2

only one eigenvector : $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

linearly independent

How to tell whether a matrix is diagonalizable?

Theorem 6.2.3

Let $\boldsymbol{A}$ be a square matrix of order n.

$\boldsymbol{A}$ is diagonalizable

if and only if

$\boldsymbol{A}$ has $\underbrace{n \text{ linearly independent eigenvectors}}_{\text{may be associated to the same eigenvalues}}$

Two observations

$$\boldsymbol{AB} = \boldsymbol{A}(\boldsymbol{b}_1 \;\; \boldsymbol{b}_2 \;\; \cdots \;\; \boldsymbol{b}_n) = (\boldsymbol{Ab}_1 \;\; \boldsymbol{Ab}_2 \;\; \cdots \;\; \boldsymbol{Ab}_n)$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{array}{c} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1 \;\; \boldsymbol{b}_2 \;\; \boldsymbol{b}_3 \end{array} = \begin{array}{c} \begin{pmatrix} 2 & 3 & 4 \\ 2 & 5 & 8 \\ 2 & 4 & 6 \end{pmatrix} \\ \boldsymbol{Ab}_1 \; \boldsymbol{Ab}_2 \; \boldsymbol{Ab}_3 \end{array}$$

$$\boldsymbol{BD} = (\boldsymbol{b}_1 \;\; \boldsymbol{b}_2 \;\; \cdots \;\; \boldsymbol{b}_n)\,\boldsymbol{D} = (d_1\boldsymbol{b}_1 \;\; d_2\boldsymbol{b}_2 \;\; \cdots \;\; d_n\boldsymbol{b}_n)$$

diagonal matrix with
diagonal entries d_i

$$\begin{array}{c} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1 \;\; \boldsymbol{b}_2 \;\; \boldsymbol{b}_3 \end{array} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{array}{c} \begin{pmatrix} 2 & 3 & 4 \\ 2 & 6 & 12 \\ 2 & 9 & 20 \end{pmatrix} \\ 2\boldsymbol{b}_1 \; 3\boldsymbol{b}_2 \; 4\boldsymbol{b}_3 \end{array}$$

The proof

$\boldsymbol{A}$ diagonalizable

$\boldsymbol{A}$ has n linearly independent eigenvectors

Theorem 6.2.3 ($\Leftarrow$)

Suppose $\boldsymbol{A}$ has n linearly independent eigenvectors.

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$

associating eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

Define the invertible matrix $\boldsymbol{P} = (\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n)$

$$\boldsymbol{AP} = (\boldsymbol{Au}_1 \ \ \boldsymbol{Au}_2 \ \ \cdots \ \ \boldsymbol{Au}_n)$$

$$= (\lambda_1\boldsymbol{u}_1 \ \lambda_2\boldsymbol{u}_2 \ \ \cdots \ \lambda_n\boldsymbol{u}_n)$$

$$= (\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n)\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

So $\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$ $\quad$ $\boldsymbol{A}$ is diagonalizable.

The proof

$\boldsymbol{A}$ diagonalizable

$\boldsymbol{A}$ has n linearly independent eigenvectors

Theorem 6.2.3 ($\Rightarrow$)

$\boldsymbol{A}$ is diagonalizable $\Rightarrow \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$

Let $\boldsymbol{P} = (\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n)$

$$\boxed{\boldsymbol{AP}} = \boxed{\boldsymbol{PD}}$$

$$\boldsymbol{A}(\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n) \qquad (\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n)\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

$$(\boldsymbol{Au}_1 \ \ \boldsymbol{Au}_2 \ \ \cdots \ \ \boldsymbol{Au}_n) \qquad (\lambda_1\boldsymbol{u}_1 \ \ \lambda_2\boldsymbol{u}_2 \ \ \cdots \ \ \lambda_n\boldsymbol{u}_n)$$

Compare each column on LHS and RHS

linearly independent

So $\boldsymbol{Au}_i = \lambda_i \boldsymbol{u}_i$ for all i $\Rightarrow \boldsymbol{u}_i$ are eigenvectors of $\boldsymbol{A}$ with eigenvalues λ_i

Algorithm 6.2.4 (Diagonalization)

Step 1: Solve the characteristic equation

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$$

to find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Step 2: For each λ_i, find a basis S_{λ_i} for the eigenspace E_{λ_i}.

Step 3: Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots \cup S_{\lambda_k}$.

(a) If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.

(b) If $|S| = n$, then $\boldsymbol{A}$ is diagonalizable.

Say, $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$, then the square matrix $\boldsymbol{P} = (\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n})$ diagonalizes $\boldsymbol{A}$.

How to diagonalize a matrix?

Example 6.2.6.1

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: By solving characteristic polynomial, the eigenvalues are 3 and 0.

Step 2: For $\lambda = 3$, solve $(3\boldsymbol{I} - \boldsymbol{B})\,\boldsymbol{x} = 0$

For $\lambda = 0$, solve $(0\boldsymbol{I} - \boldsymbol{B})\,\boldsymbol{x} = 0$

$S_3 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ a basis for E_3 $\quad S_0 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ a basis for E_0

Step 3: $|S| = |S_3| + |S_0| = 1 + 2 =$ order of $\boldsymbol{B}$

So $\boldsymbol{B}$ is diagonalizable

How to diagonalize a matrix?

Example 6.2.6.1

Step 3:

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ \boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3 \end{array} \quad \text{Then} \quad \underbrace{\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}}_{} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

you do not need to multiply this out!!!

P is not unique

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} 2 & -7 & 1 \\ 2 & 7 & 0 \\ 2 & 0 & -1 \end{pmatrix} \\ 2\boldsymbol{u}_1 \quad 7\boldsymbol{u}_2 \quad -\boldsymbol{u}_3 \end{array}$$

$$\mathbf{Q} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{Then} \quad \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

How to show a matrix is not diagonalizable?

Example 6.2.6.3

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$$

Step 1: The eigenvalues are 1 and 2.

Step 2: For $\lambda = 1$, solve $(\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = 0$

For $\lambda = 2$, solve $(2\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = 0$

$S_1 = \left\{\begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}\right\}$ a basis for E_1 $\qquad$ $S_2 = \left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ a basis for E_2

Step 3: $|S| = |S_1| + |S_2| = 1 + 1 <$ order of $\boldsymbol{A}$

Only have two linearly independent eigenvectors,
so $\boldsymbol{A}$ is not diagonalizable.

Matrix with no eigenvalue

Remark 6.2.5.1

Not in scope!

The characteristic polynomial $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ may have complex roots.

$$\boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 + 1$$

roots: $\lambda = \pm\, i$

i.e. the matrix has eigenvalues that are not real numbers but complex numbers.

We can still use the algorithm to diagonalize the matrix.

However, to discuss the theory, we need the concept of vector space over complex numbers.

Upper bound of dimension of eigenspace

Remark 6.2.5.2

$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)^1 (\lambda - 2)^3 (\lambda - 4)^2$

Characteristic polynomial

multiplicity

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

Then $\dim(E_{\lambda_i}) \leq r_i$

$\dim(E_2) \leq 3$

$\dim(E_4) \leq 2$

$\dim(E_1) = 1$

The number of basis vectors in each eigenspace cannot be more than the multiplicity of the eigenvalue in the characteristic polynomial.

A is diagonalizable
if and only if
$\dim(E_{\lambda_i}) = r_i$ for all λ_i

Union of bases of eigenspaces

A = {(1,1,1), (1,2,3)}

B = {(2,2,2), (1,2,3)}

A ∪ B is linearly dependent

Remark 6.2.5.3

The set S is always linearly independent. Ex 6 Q22

$$\overbrace{\underbrace{S_{\lambda_1}}_{\text{linearly independent}} \cup \underbrace{S_{\lambda_2}}_{\text{linearly independent}} \cup \cdots \cup \underbrace{S_{\lambda_k}}_{\text{linearly independent}}}^{} = S$$

$$S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots \cup S_{\lambda_k}$$

bases $\uparrow$ of E_{λ_1}, E_{λ_2}, E_{λ_k}

In particular

If $\mathbf{u}_1 \in E_{\lambda_1}, \mathbf{u}_2 \in E_{\lambda_2}, \cdots, \mathbf{u}_k \in E_{\lambda_k}$

then $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is linearly independent

Matrix with maximum number of eigenvalues

Theorem 6.2.7

Let $\boldsymbol{A}$ be a square matrix of order n.

If $\boldsymbol{A}$ has n distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$

then $\boldsymbol{A}$ is diagonalizable. $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}$

linearly independent

Proof

We can find one eigenvector for each eigenvalue.

Hence we have n eigenvectors.

By Remark 6.2.5.3, these eigenvectors are linearly independent.

By Theorem 6.2.3, $\boldsymbol{A}$ is diagonalizable.

Matrix with maximum number of eigenvalues

Example 6.2.8

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$\boldsymbol{A}$ has 4 distinct eigenvalues 1, 2, 3, 4.

So $\boldsymbol{A}$ is diagonalizable.

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

diagonal matrices are diagonalizable

$\boldsymbol{B}$ has only 2 distinct eigenvalues 1, 2.

And $\boldsymbol{B}$ is also diagonalizable.

Matrix with maximum number of eigenvalues

Remark 6.2.9

The converse of Theorem 6.2.7 is not true.

If ***A*** is an n x n diagonalizable matrix,

A need not have n distinct eigenvalues.

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

B has only 2 distinct eigenvalues 1, 2.

And ***B*** is also diagonalizable.

How to find powers of a matrix?

Discussion 6.2.10

Let $\boldsymbol{A}$ be a diagonalizable matrix of order n

$\boldsymbol{P}$ an invertible matrix such that

$$(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^m = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}^m = \begin{pmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{pmatrix}$$

$$\boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}$$

Then $$\boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{pmatrix} \boldsymbol{P}^{-1}$$

How to find powers of a matrix?

Example 6.2.11.1

invertible

$$\boldsymbol{A} = \begin{pmatrix} -4 & 0 & -6 \\ 2 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix}$$

Use Algorithm 6.2.4 to find the eigenvalues and eigenvectors

We have

$$\boldsymbol{P} = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

obtain this diagonal matrix from eigenvalues, not matrix multiplication!

$$\mathbf{A}^{m} = \mathbf{P} \begin{pmatrix} (-1)^{m} & 0 & 0 \\ 0 & 1^{m} & 0 \\ 0 & 0 & 2^{m} \end{pmatrix} \mathbf{P}^{-1}$$

$$\mathbf{A}^{-1} = \mathbf{P} \begin{pmatrix} (-1)^{-1} & 0 & 0 \\ 0 & 1^{-1} & 0 \\ 0 & 0 & 2^{-1} \end{pmatrix} \mathbf{P}^{-1}$$

Some applications

- Weather forecast (Markov chain)
- Population growth
- Cards shuffling
- Genetics
- Linear recurrence relation

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \to \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \to \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \cdots \to \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

stage 0 (initial)   stage 1   stage 2   stage n

$\mathbf{x_0}$ $\mathbf{x_1} = \mathbf{A}\mathbf{x_0}$ $\mathbf{x_2} = \mathbf{A}\mathbf{x_1}$ $\mathbf{x_n} = \mathbf{A}\mathbf{x_{n-1}}$

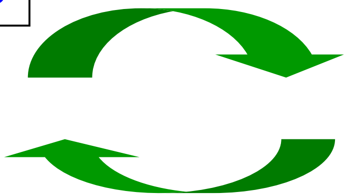
Application to modeling

Example 6.1.1 (Population)

Population after n years

a_n rural population

4%



1%

Urban population b_n

Long term effect ?

Ans: ~ 20% rural population, ~ 80% urban population

$$a_n = 0.96a_{n-1} + 0.01b_{n-1}$$
$$b_n = 0.04a_{n-1} + 0.99b_{n-1}$$

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}$$

Application to modeling

Example 6.1.1

$$\underbrace{\begin{pmatrix} a_n \\ b_n \end{pmatrix}}_{\boldsymbol{x_n}} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}}_{\boldsymbol{x_{n-1}}} \quad \Rightarrow \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}^n}_{\mathbf{A}^n} \underbrace{\begin{pmatrix} a_0 \\ b_0 \end{pmatrix}}_{\mathbf{x}_0}$$

$$\boldsymbol{x_n} = \boldsymbol{A}\boldsymbol{x_{n-1}} = \boldsymbol{A}^2\boldsymbol{x_{n-2}} = \boldsymbol{A}^3\boldsymbol{x_{n-3}} = \cdots = \boldsymbol{A}^n\boldsymbol{x_0}$$

current population

long term effect $\longrightarrow$ a_n and b_n for large n

$\longrightarrow$ $\boldsymbol{x}_n$ for large n

$\longrightarrow$ $\boldsymbol{A}^n$ for large n

$$\mathbf{A}^{(\text{big } n)} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \boxed{0.95^{(\text{big } n)}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}$$

$\rightarrow 0$

$$\begin{pmatrix} a_{(\text{big } n)} \\ b_{(\text{big } n)} \end{pmatrix} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 0.2(a_0 + b_0) \\ 0.8(a_0 + b_0) \end{pmatrix}$$

Fibonacci Numbers

0 1 1 2 3 5 8 13 21 34

What's the 100th number?

How to solve recurrence relation?

Example 6.2.11.2

Denote the Fibonacci numbers by $a_0, a_1, a_2, \ldots$

$a_0 = 0$ $a_1 = 1$ $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

initial conditions recurrence relation

What is the value of a_n ?

$$\boxed{a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n}$$

due to eigenvalues

Example: $$a_{100} = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{100} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{100}$$

354224848179261915075

How to find recurrence matrix?

$a_0 = 0,\ a_1 = 1,$
$a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Example 6.2.11.2

Form the vector: $\boldsymbol{x_n} = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ $\quad \boldsymbol{x_{n-1}} = \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$

$$\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \mathbf{x}_3 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \cdots$$

The recurrence matrix $\boldsymbol{A}$: $\quad \boldsymbol{x_n} = \boldsymbol{A}\boldsymbol{x_{n-1}}$ for all n

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$$

Compare coefficients $\quad a_n = \boxed{0}\, a_{n-1} + \boxed{1}\, a_n$

Recurrence relation $\quad a_{n+1} = \boxed{1}\, a_{n-1} + \boxed{1}\, a_n$

$a_{n+1} = a_n + a_{n-1}$

$$\begin{aligned} a_n &= \boxed{}\, a_{n-1} + \boxed{}\, a_n \\ a_{n+1} &= \boxed{}\, a_{n-1} + \boxed{}\, a_n \end{aligned}$$

Example (Additional)

$$a_0 = 1,\ a_1 = 3,$$
$$a_n = 3a_{n-1} + 5a_{n-2} \ \text{ for } n \geq 2$$

What is the recurrence matrix ?

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 5 & 3 \end{pmatrix}$$

In general,

$$a_0 = s,\ a_1 = t,$$
$$a_n = pa_{n-1} + qa_{n-2} \ \text{ for } n \geq 2$$

The recurrence matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$

How to find the explicit formula?

$a_0 = 0,\ a_1 = 1,$
$a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Example 6.2.11.2

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$ has two eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ $\lambda_2 = \frac{1-\sqrt{5}}{2}$

So $\boldsymbol{A}$ is diagonalizable

Diagonalized by $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$

$$\boldsymbol{x_n} = \boldsymbol{A}^n \boldsymbol{x_0}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \mathbf{P} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \mathbf{P}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n \\ \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{pmatrix}$$

$$\boxed{a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n}$$

$a_0 = 0,\ a_1 = 1,\ a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Solving linear recurrence relation

$a_0 = u \qquad a_1 = v \qquad a_n = pa_{n-1} + qa_{n-2}$ for $n \geq 2$

Form the recurrence matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Find the eigenvalues of $\boldsymbol{A}$

$$\lambda_1 = \tfrac{1+\sqrt{5}}{2} \qquad \lambda_2 = \tfrac{1-\sqrt{5}}{2}$$

If $\boldsymbol{A}$ is diagonalizable, find the matrix $\boldsymbol{P}$ that diagonalizes $\boldsymbol{A}$

$$\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \tfrac{1+\sqrt{5}}{2} & \tfrac{1-\sqrt{5}}{2} \end{pmatrix}$$

Set up $\boldsymbol{x_n} = \boldsymbol{A}^n \boldsymbol{x}_0$ and diagonalize $\boldsymbol{A}^n$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix} \boldsymbol{P}^{-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

Multiply out the RHS and equate the first component

$$a_n = s\left(\lambda_1\right)^n + t\left(\lambda_2\right)^n$$

Section 6.3

Orthogonal Diagonalization

Objective

- What is orthogonal diagonalization?
- When is a matrix orthogonally diagonalizable?
- How to orthogonally diagonalize a symmetric matrix?

What is an orthogonally diagonalizable matrix

Definition 6.3.2

Recall: Section 6.2

A square matrix **A** is called
diagonalizable
if there exists an invertible matrix **P** such that
$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a diagonal matrix.
We say the matrix **P** diagonalizes **A**.

A square matrix **A** is called
orthogonally diagonalizable
if there exists an orthogonal matrix **P** such that
$\mathbf{P}^{T}\mathbf{A}\mathbf{P}$ is a diagonal matrix.
We say the matrix **P** orthogonally diagonalizes **A**.

When is a matrix orthogonally diagonalizable

Theorem 6.3.4

$$\begin{pmatrix} 0 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

A square matrix is orthogonally diagonalizable
if and only if
it is symmetric.

beyond the scope of this course

A is orthogonally diagonalizable

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$$

$$\Rightarrow \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T$$

$$\Rightarrow \mathbf{A}^T = (\mathbf{P}\mathbf{D}\mathbf{P}^T)^T$$

$$\Rightarrow \mathbf{A}^T = (\mathbf{P}^T)^T\mathbf{D}^T(\mathbf{P}^T)$$

$$\Rightarrow \mathbf{A}^T = \mathbf{P}\mathbf{D}\mathbf{P}^T = \mathbf{A}$$

So **A** is symmetric

Algorithm 6.3.5 **A**: symmetric matrix

Step 1: Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Step 2: For each eigenvalue λ_i,

Step 2a: find a basis S_{λ_i} for the eigenspace E_{λ_i}

Step 2b: use the Gram-Schmidt Process (Theorem 5.2.19) to transform S_{λ_i} to an orthonormal basis T_{λ_i} .

Step 3: Let $T = T_{\lambda_1} \cup T_{\lambda_2} \cup \ldots \cup T_{\lambda_k}$

say $T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

The square matrix $\mathbf{P} = (\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n)$ is an orthogonal matrix that diagonalizes **A**.

Eigenvalues of symmetric matrix

Remark 6.3.6.1 **A**: symmetric matrix

In Step 1, the eigenvalues of a symmetric matrix are always real numbers.

Idea:

Let λ be an eigenvalue of a symmetric matrix

Write $\lambda = a + ib$ (a, b are real)

Conjugate $\bar{\lambda} = a - ib$ also an eigenvalue of the matrix

Try to show $\lambda = \bar{\lambda}$, which implies λ is real.

Dimension of Eigenspaces

In general
$\dim(E_{\lambda_i}) \le r_i$

Remark 6.3.6.2 **A**: symmetric matrix

Suppose the characteristic polynomial of **A** is

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots (\lambda - \lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of **A**.

Then for each eigenvalue λ_i,

$$\dim(E_{\lambda_i}) = r_i.$$

number of basis vectors in the eigenspace for λ_i

multiplicity of λ_i in the characteristic polynomial

$r_1 + r_2 + \cdots + r_k =$ degree of polynomial = order of **A**

$\dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_k} =$ no.lin.indep.eigenvectors

A symmetric matrix is always diagonalizable.

T is an orthonormal set

Step 2b: use the Gram-Schmidt Process to transform S_{λ_i} to an orthonormal basis T_{λ_i}

Remark 6.3.6

$$\begin{array}{ccccccccc} & & S_{\lambda_1} & & S_{\lambda_2} & & \dots & & S_{\lambda_k} \\ T & = & T_{\lambda_1} & \cup & T_{\lambda_2} & \cup & \dots & \cup & T_{\lambda_k} \end{array}$$

3. In Step 3, the set T is always orthonormal.

   Not immediate

4. Since T is always orthonormal, the square matrix **P** in Step 3 is always orthogonal.

   Immediate from Theorem 5.4.6

Ex6 Q26 Proof later

Let **A** be a symmetric matrix.

If **u** and **v** are two eigenvectors of **A** associated with eigenvalues λ and μ, resp. where $\lambda \neq \mu$,

Then $\mathbf{u} \cdot \mathbf{v} = 0$.

OD a 2x2 symmetric matrix

Example 6.3.7.1

$$\mathbf{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$$

Step 1: The eigenvalues are 1/2 and 3/2.

Step 2a: Bases for $E_{1/2}$ and $E_{3/2}$: $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$

Step 2b: Orthonormal bases: $\left\{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\right\}$

No need Gram-Schmidt

Step 3: $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$

OD a 3x3 symmetric matrix

Discussion 6.3.1

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: The eigenvalues are 3 and 0.

Step 2a: Bases for E_3 and E_0: $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

Step 2b: Orthonormal bases: $\left\{\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}\right\}$

just normalize

Gram-Schmidt

Step 3: $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$ and $\mathbf{P}^T\mathbf{B}\mathbf{P} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$