Mathematics 160: Lecture 14 Diagonalization

Dan Sloughter

Furman University

September 30, 2011

• As in a previous example, let

$$A = \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix}.$$

We saw that

$$X_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_1=-3$ and

$$X_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.



• As in a previous example, let

$$A = \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix}.$$

We saw that

$$X_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_1=-3$ and

$$X_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.



Now let

$$P = \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}.$$

- That is, P is the matrix whose columns are the eigenvectors of A
- Then

$$P^{-1}AP = -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}$$
$$= -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 8 \\ -3 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}.$$

Now let

$$P = \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}.$$

- That is, P is the matrix whose columns are the eigenvectors of A.
- Then

$$P^{-1}AP = -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}$$
$$= -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 8 \\ -3 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}.$$

Now let

$$P = \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}.$$

- That is, P is the matrix whose columns are the eigenvectors of A.
- Then

$$P^{-1}AP = -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & 1 \end{bmatrix}$$
$$= -\frac{1}{5} \begin{bmatrix} 1 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 8 \\ -3 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}.$$

Definition

• We say an $n \times n$ matrix A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

- Suppose D is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $P^{-1}AP = D$.
- Then AP = PD
- If $X_1, X_2, ..., X_n$ are the columns of P, then we have

$$A\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

- Suppose D is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $P^{-1}AP = D$.
- Then AP = PD.
- If X_1, X_2, \ldots, X_n are the columns of P, then we have

$$A\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

- Suppose D is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $P^{-1}AP = D$.
- Then AP = PD.
- If X_1, X_2, \ldots, X_n are the columns of P, then we have

$$A\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

- Suppose D is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $P^{-1}AP = D$.
- Then AP = PD.
- If X_1, X_2, \ldots, X_n are the columns of P, then we have

$$A\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Theorem

- An $n \times n$ matrix A is diagonalizable if and only if A has n eigenvectors X_1, X_2, \ldots, X_n with the property that the matrix P whose columns are X_1, X_2, \ldots, X_n is invertible.
- If A is diagonalizable, then, with P as above,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where λ_i is the eigenvalue corresponding to the eigenvector X_i .

Theorem

- An $n \times n$ matrix A is diagonalizable if and only if A has n eigenvectors X_1, X_2, \ldots, X_n with the property that the matrix P whose columns are X_1, X_2, \ldots, X_n is invertible.
- If A is diagonalizable, then, with P as above,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where λ_i is the eigenvalue corresponding to the eigenvector X_i .

Definitions

- We say an eigenvalue λ of an $n \times n$ matrix A has multiplicity m if $x \lambda$ occurs exactly m times in the factorization of $c_A(x)$.
- If λ is an eigenvalue of an $n \times n$ matrix A, we call a set of basic solutions of the equation $(\lambda I_n A)X = O$ basic eigenvectors corresponding to λ .
- Note: an eigenvalue of multiplicity m will have at most m basic eigenvectors.

Definitions

- We say an eigenvalue λ of an $n \times n$ matrix A has multiplicity m if $x \lambda$ occurs exactly m times in the factorization of $c_A(x)$.
- If λ is an eigenvalue of an $n \times n$ matrix A, we call a set of basic solutions of the equation $(\lambda I_n A)X = O$ basic eigenvectors corresponding to λ .
- Note: an eigenvalue of multiplicity m will have at most m basic eigenvectors.

Definitions

- We say an eigenvalue λ of an $n \times n$ matrix A has multiplicity m if $x \lambda$ occurs exactly m times in the factorization of $c_A(x)$.
- If λ is an eigenvalue of an $n \times n$ matrix A, we call a set of basic solutions of the equation $(\lambda I_n A)X = O$ basic eigenvectors corresponding to λ .
- Note: an eigenvalue of multiplicity *m* will have at most *m* basic eigenvectors.

Theorem

- An n × n matrix A is not diagonalizable if and only if A has an eigenvalue of multiplicity m which has fewer than m basic eigenvectors.
- In particular, if an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Theorem

- An n × n matrix A is not diagonalizable if and only if A has an eigenvalue of multiplicity m which has fewer than m basic eigenvectors.
- In particular, if an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Suppose

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

• From our previous work,

$$X_1 = \begin{bmatrix} -2\\1\\0 \end{bmatrix}$$
 and $X_2 = \begin{bmatrix} -1\\0\\1 \end{bmatrix}$

are basic eigenvectors for the eigenvalue $\lambda_1=1$

Suppose

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

From our previous work,

$$X_1 = \begin{bmatrix} -2\\1\\0 \end{bmatrix}$$
 and $X_2 = \begin{bmatrix} -1\\0\\1 \end{bmatrix}$

are basic eigenvectors for the eigenvalue $\lambda_1 = 1$.

And

$$X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

is a basic eigenvector for the eigenvalue $\lambda_3 = 5$.

Then

$$P = \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

You may check that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$



And

$$X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

is a basic eigenvector for the eigenvalue $\lambda_3 = 5$.

Then

$$P = \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

You may check that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$



And

$$X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

is a basic eigenvector for the eigenvalue $\lambda_3 = 5$.

Then

$$P = \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

• You may check that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$



Suppose

$$A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 2 & 3 \\ 2 & -3 & -4 \end{bmatrix}.$$

The characteristic polynomial is

$$c_{A}(\lambda) = \det \begin{bmatrix} \lambda - 1 & 2 & 2 \\ 2 & \lambda - 2 & -3 \\ -2 & 3 & \lambda + 4 \end{bmatrix}$$

$$= (\lambda - 1)((\lambda - 2)(\lambda + 4) + 9) - 2(2(\lambda + 4) - 6)$$

$$+ 2(6 + 2(\lambda - 2))$$

$$= (\lambda - 1)(\lambda^{2} + 2\lambda + 1) - (4\lambda + 4) + (4\lambda + 4)$$

$$= (\lambda - 1)(\lambda + 1)^{2}.$$

• So the eigenvalues are $\lambda_1=1$ and $\lambda_2=\lambda_3=-1$.



Suppose

$$A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 2 & 3 \\ 2 & -3 & -4 \end{bmatrix}.$$

• The characteristic polynomial is

$$c_{A}(\lambda) = \det \begin{bmatrix} \lambda - 1 & 2 & 2 \\ 2 & \lambda - 2 & -3 \\ -2 & 3 & \lambda + 4 \end{bmatrix}$$

$$= (\lambda - 1)((\lambda - 2)(\lambda + 4) + 9) - 2(2(\lambda + 4) - 6)$$

$$+ 2(6 + 2(\lambda - 2))$$

$$= (\lambda - 1)(\lambda^{2} + 2\lambda + 1) - (4\lambda + 4) + (4\lambda + 4)$$

$$= (\lambda - 1)(\lambda + 1)^{2}.$$

• So the eigenvalues are $\lambda_1=1$ and $\lambda_2=\lambda_3=-1$.

Suppose

$$A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 2 & 3 \\ 2 & -3 & -4 \end{bmatrix}.$$

The characteristic polynomial is

$$c_{A}(\lambda) = \det \begin{bmatrix} \lambda - 1 & 2 & 2 \\ 2 & \lambda - 2 & -3 \\ -2 & 3 & \lambda + 4 \end{bmatrix}$$

$$= (\lambda - 1)((\lambda - 2)(\lambda + 4) + 9) - 2(2(\lambda + 4) - 6)$$

$$+ 2(6 + 2(\lambda - 2))$$

$$= (\lambda - 1)(\lambda^{2} + 2\lambda + 1) - (4\lambda + 4) + (4\lambda + 4)$$

$$= (\lambda - 1)(\lambda + 1)^{2}.$$

• So the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = -1$.

• For $\lambda_1 = 1$, we have

$$\begin{bmatrix} 0 & 2 & 2 & 0 \\ 2 & -1 & -3 & 0 \\ -2 & 3 & 5 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & -1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ -2 & 3 & 5 & 0 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 2 & -1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{3}{2} & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

• So if we let $x_3 = t$, then $x_2 = -t$ and $x_1 = t$.

• For $\lambda_1 = 1$, we have

$$\begin{bmatrix} 0 & 2 & 2 & 0 \\ 2 & -1 & -3 & 0 \\ -2 & 3 & 5 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & -1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ -2 & 3 & 5 & 0 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 2 & -1 & -3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{3}{2} & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

• So if we let $x_3 = t$, then $x_2 = -t$ and $x_1 = t$.

• That is,

$$X = t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

• In particular,

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_1=1.$

• That is,

$$X=t\begin{bmatrix}1\\-1\\1\end{bmatrix}.$$

• In particular,

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$.

• For the eigenvalue $\lambda_2 = \lambda_3 = -1$, we have

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

• If we let $x_3 = t$, then $x_2 = -t$ and $x_1 = 0$.

• For the eigenvalue $\lambda_2 = \lambda_3 = -1$, we have

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

• If we let $x_3 = t$, then $x_2 = -t$ and $x_1 = 0$.

That is,

$$X = t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

ullet Hence we have only one basic eigenvector corresponding to $\lambda_2=-1$

$$X_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

• Hence A is not diagonalizable. We say A is defective

That is,

$$X=t\begin{bmatrix}0\\-1\\1\end{bmatrix}.$$

• Hence we have only one basic eigenvector corresponding to $\lambda_2=-1$:

$$X_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

• Hence A is not diagonalizable. We say A is defective

• That is,

$$X=t\begin{bmatrix}0\\-1\\1\end{bmatrix}.$$

• Hence we have only one basic eigenvector corresponding to $\lambda_2=-1$:

$$X_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

• Hence A is not diagonalizable. We say A is defective

- Given an $n \times n$ matrix A, the Octave command eig(A) returns the eigenvalues of A.
- The command [P, D] = eig(A) returns P as the matrix of eigenvectors and D as the corresponding diagonal matrix of eigenvalues.
- P(:,j) will return the jth column of P, that is, the jth eigenvector of A.
- Note: Octave chooses eigenvectors normalized to have length one.
- The command poly(A) returns the coefficients of the characteristic polynomial of A.

- Given an $n \times n$ matrix A, the Octave command eig(A) returns the eigenvalues of A.
- The command [P, D] = eig(A) returns P as the matrix of eigenvectors and D as the corresponding diagonal matrix of eigenvalues.
- P(:,j) will return the jth column of P, that is, the jth eigenvector of
 A.
- Note: Octave chooses eigenvectors normalized to have length one.
- The command poly(A) returns the coefficients of the characteristic polynomial of A.

- Given an $n \times n$ matrix A, the Octave command eig(A) returns the eigenvalues of A.
- The command [P, D] = eig(A) returns P as the matrix of eigenvectors and D as the corresponding diagonal matrix of eigenvalues.
- P(:,j) will return the jth column of P, that is, the jth eigenvector of A.
- Note: Octave chooses eigenvectors normalized to have length one.
- The command poly(A) returns the coefficients of the characteristic polynomial of A.

- Given an $n \times n$ matrix A, the Octave command eig(A) returns the eigenvalues of A.
- The command [P, D] = eig(A) returns P as the matrix of eigenvectors and D as the corresponding diagonal matrix of eigenvalues.
- P(:,j) will return the jth column of P, that is, the jth eigenvector of A.
- Note: Octave chooses eigenvectors normalized to have length one.
- The command poly(A) returns the coefficients of the characteristic polynomial of A.

- Given an $n \times n$ matrix A, the Octave command eig(A) returns the eigenvalues of A.
- The command [P, D] = eig(A) returns P as the matrix of eigenvectors and D as the corresponding diagonal matrix of eigenvalues.
- P(:,j) will return the jth column of P, that is, the jth eigenvector of A.
- Note: Octave chooses eigenvectors normalized to have length one.
- The command poly(A) returns the coefficients of the characteristic polynomial of A.