Math 221: LINEAR ALGEBRA

Chapter 5. Vector Space $\mathbb{R}^n$
§5-5. Similarity and Diagonalization

Le Chen[1]

Emory University, 2020 Fall

(last updated on 10/27/2020)

Similar Matrices

Diagonalization Revisited

Algebraic and Geometric Multiplicities

Characterizing Diagonalizable Matrices

Complex Eigenvalues

Eigenvalues of Real Symmetric Matrices

Similar Matrices

Definition (Similar Matrices)

Let A and B be $n \times n$ matrices. A is similar to B, written $A \sim B$, if there exists an invertible matrix P such that $B = P^{-1}AP$.

Lemma

Similarity is an equivalence relation, i.e., for $n \times n$ matrices A, B and C

1. $A \sim A$ (reflexive);
2. if $A \sim B$, then $B \sim A$ (symmetric);
3. if $A \sim B$ and $B \sim C$, then $A \sim C$ (transitive).

Proof.

1. Since $A = I_n A I_n$ and $I_n^{-1} = I_n$, $A = I_n^{-1} A I_n$. Therefore, $A \sim A$.
2. Suppose $A \sim B$. Then there exists an invertible $n \times n$ matrix P such that $B = P^{-1}AP$. Multiplying both sides on the left by P, on the right by P^{-1}, and simplifying gives us $PBP^{-1} = A$. Therefore, $A = (P^{-1})^{-1}A(P^{-1})$, so $A \sim B$.

Proof. (continued)

3. Since $A \sim B$ and $B \sim C$, there exist invertible $n \times n$ matrices P and Q such that
$$B = P^{-1}AP \quad \text{and} \quad C = Q^{-1}BQ.$$

Thus
$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (Q^{-1}P^{-1})A(PQ) = (PQ)^{-1}A(PQ),$$

where PQ is invertible, and hence $A \sim C$. ■

Definition

If $A = [a_{ij}]$ is an $n \times n$ matrix, then the trace of A is

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Lemma (Properties of trace)

For $n \times n$ matrices A and B, and any $k \in \mathbb{R}$,

1. $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$;
2. $\text{tr}(kA) = k \cdot \text{tr}(A)$;
3. $\text{tr}(AB) = \text{tr}(BA)$.

Proof.

The proofs of (1) and (2) are trivial. As for (3), ...

Recall that for any n × n matrix A, the characteristic polynomial of A is

$$c_A(x) = \det(xI - A),$$

and is a polynomial of degree n.

Theorem (Properties of Similar Matrices)

If A and B are n × n matrices and A ∼ B, then

1. $\det(A) = \det(B)$;
2. rank (A) = rank (B);
3. $\operatorname{tr}(A) = \operatorname{tr}(B)$;
4. $c_A(x) = c_B(x)$;
5. A and B have the same eigenvalues.

Since $A \sim B$, there exists an $n \times n$ invertible matrix P so that $B = P^{-1}AP$.

1. $\det(B) = \det(P^{-1}AP) = \det(P^{-1}) \cdot \det(A) \cdot \det(P)$.

   Since P is invertible, $\det(P^{-1}) = \frac{1}{\det(P)}$, so

   $$\det(B) = \frac{1}{\det(P)} \cdot \det(A) \cdot \det(P) = \frac{1}{\det(P)} \cdot \det(P) \cdot \det(A) = \det(A).$$

   Therefore, $\det(B) = \det(A)$.

2. rank (B) = rank $(P^{-1}AP)$.

   Since P is invertible, rank $(P^{-1}AP)$ = rank $(P^{-1}A)$,

   since P^{-1} is invertible, rank $(P^{-1}A)$ = rank (A).

   Therefore, rank (B) = rank (A).

3. $\operatorname{tr}(B) = \operatorname{tr}[(P^{-1}A)P] = \operatorname{tr}[P(P^{-1}A)] = \operatorname{tr}[(PP^{-1})A] = \operatorname{tr}(IA) = \operatorname{tr}(A)$.

Proof. (continued)

4.

$$
\begin{aligned}
c_B(x) = \det(xI - B) &= \det(xI - P^{-1}AP) \\
&= \det(xP^{-1}P - P^{-1}AP) \\
&= \det(P^{-1}xP - P^{-1}AP) \\
&= \det[P^{-1}(xI - A)P] \\
&= \det(P^{-1}) \cdot \det(xI - A) \cdot \det(P) \\
&= \det(P^{-1}) \cdot \det(P) \cdot \det(xI - A)
\end{aligned}
$$

Since P is invertible, $\det(P^{-1}) = \frac{1}{\det(P)}$, so

$$
c_B(x) = \frac{1}{\det(P)} \cdot \det(P) \cdot \det(xI - A) = \det(xI - A) = c_A(x).
$$

5. Since the eigenvalues of a matrix are the roots of the characteristic polynomial, $c_B(x) = c_A(x)$ implies that A and B have the same eigenvalues. ∎

Diagonalization Revisited

Recall that if λ is an eigenvalue of A, then $A\vec{x} = \lambda\vec{x}$ for some nonzero vector $\vec{x}$ in $\mathbb{R}^n$. Such a vector $\vec{x}$ is called a λ-eigenvector of A or an eigenvector of A corresponding to λ.

Definition (Diagonalizable – rephrased)

An $n \times n$ matrix A is diagonalizable if $A \sim D$ for some diagonal matrix D.

Remark (Diagonalizability)

Determining whether or not a square matrix A is diagonalizable is done by checking whether

<div align="center">

the number of linearly independent eigenvectors
– geometric multiplicity

$||?$

the multiplicity of each eigenvalue
– algebraic multiplicity

</div>

Example

Let $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $\lambda = -1$ is an eigenvalue of A, and $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a (-1)-eigenvector of A since

$$A\vec{x} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Theorem

Suppose A is an $n \times n$ matrix.

1. The eigenvalues of A are the roots of $c_A(x)$.
2. The λ-eigenvectors of A are all the nonzero solutions to $(\lambda I - A)\vec{x} = \vec{0}_n$.

Problem

Determine all eigenvalues of $A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 3 & 6 & 0 & 0 \\ -1 & 0 & 6 & 0 \\ 4 & 2 & -1 & 1 \end{bmatrix}$.

Solution

$$\det(xI - A) = \begin{vmatrix} x+2 & 0 & 0 & 0 \\ -3 & x-6 & 0 & 0 \\ 1 & 0 & x-6 & 0 \\ -4 & -2 & 1 & x-1 \end{vmatrix} = (x+2)(x-6)(x-6)(x-1).$$

Thus, the eigenvalues of A are $-2, 6, 6$ and 1, precisely the elements on the main diagonal of A. ■

Remark

In general, the eigenvalues of any triangular matrix are the entries on its main diagonal.

Theorem

Let A be an $n \times n$ matrix.

1. A is diagonalizable if and only if $\mathbb{R}^n$ has a basis $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ of eigenvectors of A.

2. If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ are eigenvectors of A and form a basis of $\mathbb{R}^n$, then

$$P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix}$$

is an invertible matrix such that

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where λ_i is the eigenvalue of A corresponding to $\vec{x}_i$.

This result was covered earlier, but without the use of term basis.

Theorem

Let A be an n × n matrix, and suppose that A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. For each i, let $\vec{x}_i$ be a λ_i-eigenvector of A. Then $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is linearly independent.

Proof.

We need to show that $t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k = \vec{0}$ only has trivial solution $t_1 = \cdots = t_k = 0$. Notice that

$$t_1 A\vec{x}_1 + t_2 A\vec{x}_2 + \cdots + t_k A\vec{x}_k = t_1\lambda_1\vec{x}_1 + t_2\lambda_2\vec{x}_2 + \cdots + t_k\lambda_k\vec{x}_k = \vec{0}$$

$$t_1 A^2\vec{x}_1 + t_2 A^2\vec{x}_2 + \cdots + t_k A^2\vec{x}_k = t_1\lambda_1^2\vec{x}_1 + t_2\lambda_2^2\vec{x}_2 + \cdots + t_k\lambda_k^2\vec{x}_k = \vec{0}$$

$$\vdots \qquad\qquad \vdots$$

$$t_1 A^{k-1}\vec{x}_1 + \cdots\cdots + t_k A^{k-1}\vec{x}_k = t_1\lambda_1^{k-1}\vec{x}_1 + \cdots\cdots + t_k\lambda_k^{k-1}\vec{x}_k = \vec{0}$$

Proof.

$$
\begin{array}{ccccccccc}
t_1\lambda_1\vec{x}_1 & + & t_2\lambda_2\vec{x}_2 & + & \cdots & + & t_k\lambda_k\vec{x}_k & = & \vec{0}\\
t_1\lambda_1^2\vec{x}_1 & + & t_2\lambda_2^2\vec{x}_2 & + & \cdots & + & t_k\lambda_k^2\vec{x}_k & = & \vec{0}\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
t_1\lambda_1^{k-1}\vec{x}_1 & + & t_2\lambda_2^{k-1}\vec{x}_2 & + & \cdots & + & t_k\lambda_k^{k-1}\vec{x}_k & = & \vec{0}
\end{array}
$$

$$\Updownarrow$$

$$
\begin{pmatrix}\vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_k\end{pmatrix}
\begin{bmatrix}
t_1 & 0 & 0 & 0\\
0 & t_2 & 0 & 0\\
0 & 0 & \ddots & 0\\
0 & 0 & 0 & t_k
\end{bmatrix}
\begin{pmatrix}
\lambda_1^0 & \lambda_1^1 & \cdots & \lambda_1^{k-1}\\
\lambda_2^0 & \lambda_2^1 & \cdots & \lambda_2^{k-1}\\
\vdots & \vdots & \vdots & \vdots\\
\lambda_k^0 & \lambda_k^1 & \cdots & \lambda_k^{k-1}
\end{pmatrix}
= O_{k\times k}.
$$

Proof.

Since λ_i are distinct, the Vandermonde matrix is invertible, hence,

$$\begin{pmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_k \end{pmatrix} \begin{bmatrix} t_1 & 0 & 0 & 0 \\ 0 & t_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & t_k \end{bmatrix} = O_{k \times k}.$$

$$\Updownarrow$$

$$t_i \vec{x}_i = 0 \quad \text{for all } i = 1, \cdots, k$$

$$\Downarrow$$

$$t_i = 0 \quad \text{for all } i = 1, \cdots, k$$

Only trivial solution is found. Hence, $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is independent. ∎

Proof. (Another proof left for you to study)

The proof is by induction on k, the number of distinct eigenvalues.

Basis. If $k = 1$, then $\{\vec{x}_1\}$ is an independent set because $\vec{x}_1 \neq \vec{0}_n$.

Suppose that for some $k \geq 1$, $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is independent, where $\vec{x}_i$ is an eigenvector of A corresponding to λ_i, $1 \leq i \leq k$, and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct. (This is the Inductive Hypothesis.) Now suppose $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ are distinct eigenvalues of A that have corresponding eigenvectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{k+1}$, respectively. Consider

$$t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_{k+1}\vec{x}_{k+1} = \vec{0}_n, \text{ for } t_1, t_2, \ldots, t_{k+1} \in \mathbb{R}. \quad (1)$$

Multiplying equation (1) by A (on the left) gives us

Proof. (continued)

$$t_1 A\vec{x}_1 + t_2 A\vec{x}_2 + \cdots + t_{k+1} A\vec{x}_{k+1} = \vec{0}_n,$$

$$\Downarrow$$

$$t_1 \lambda_1 \vec{x}_1 + t_2 \lambda_2 \vec{x}_2 + \cdots + t_{k+1} \lambda_{k+1} \vec{x}_{k+1} = \vec{0}_n. \qquad (2)$$

Also, multiplying (1) by λ_{k+1} gives us

$$t_1 \lambda_{k+1} \vec{x}_1 + t_2 \lambda_{k+1} \vec{x}_2 + \cdots + t_{k+1} \lambda_{k+1} \vec{x}_{k+1} = \vec{0}_n, \qquad (3)$$

and subtracting (3) from (2) results in

$$t_1 (\lambda_1 - \lambda_{k+1}) \vec{x}_1 + t_2 (\lambda_2 - \lambda_{k+1}) \vec{x}_2 + \cdots + t_k (\lambda_k - \lambda_{k+1}) \vec{x}_k = \vec{0}_n.$$

Proof. (continued)

By the inductive hypothesis, $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is independent, so

$$t_i(\lambda_i - \lambda_{k+1}) = 0 \text{ for } i = 1, 2, \ldots k.$$

Since $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct, $(\lambda_i - \lambda_{k+1}) \neq 0$ for $i = 1, 2, \ldots, k$, and thus $t_i = 0$ for $i = 1, 2, \ldots, k$. Substituting these values into (1) yields

$$t_{k+1}\vec{x}_{k+1} = \vec{0}_n,$$

implying that $t_{k+1} = 0$, since $\vec{x}_{k+1} \neq \vec{0}_n$.

Therefore, $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{k+1}\}$ is an independent set, and the result follows by induction. ∎

The next result is an easy consequence of the previous Theorem.

Theorem (Covered earlier, but now with a proof)

If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

Proof.

Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denote the n (distinct) eigenvalues of A, and let $\vec{x}_i$ be an eigenvector of A corresponding to λ_i, $1 \leq i \leq n$. By the previous Theorem, $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ is an independent set. A subset of n linearly independent vectors of $\mathbb{R}^n$ also spans $\mathbb{R}^n$, and thus $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ is a basis of $\mathbb{R}^n$. Thus A is diagonalizable. ∎

Problem

Is the matrix

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 8 & 6 & -2 \\ 0 & 0 & -3 \end{bmatrix}$$

diagonalizable?

Solution

Because A has characteristic polynomial

$$c_A(x) = (x+3)(x-2)(x-4),$$

A has distinct eigenvalues $-3, 2$ and 4.

Since A has three distinct eigenvalues, A is diagonalizable. ■

Is $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ diagonalizable? Explain.

Solution

First, $c_A(x) = (x - 2)(x + 1)^2$, so the eigenvalues of A are $\lambda_1 = 2, \lambda_2 = -1$, and $\lambda_3 = -1$. Since the eigenvalues are not distinct, it isn't immediately obvious that A is diagonalizable. The general solution to $(-I - A)\vec{x} = \vec{0}_3$:

$$\begin{bmatrix} -1 & -1 & -1 & | & 0 \\ -1 & -1 & -1 & | & 0 \\ -1 & -1 & -1 & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

is $x_1 = -s - t, x_2 = s$, and $x_3 = t$ for $s, t \in \mathbb{R}$, leading to basic solutions

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

that are linearly independent. Therefore, there is a basis of $\mathbb{R}^3$ consisting of eigenvectors of A, so A is diagonalizable. ■

Lemma (Technical but useful)

Let A be an $n \times n$ matrix, with independent eigenvectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$. Extend $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ to a basis $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, \ldots, \vec{x}_n\}$ of $\mathbb{R}^n$, and let $P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix}$. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the (not necessarily distinct) eigenvalues corresponding to $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$, then

$$P^{-1}AP = \begin{bmatrix} \text{diag}(\lambda_1, \ldots, \lambda_k) & B \\ 0_{(n-k) \times k} & A_1 \end{bmatrix},$$

where B is an $k \times (n-k)$ matrix and A_1 is an $(n-k) \times (n-k)$ matrix.

Proof.

$$\begin{bmatrix} A\vec{x}_1 & | & \cdots & | & A\vec{x}_k & | & A\vec{x}_{k+1} & | & \cdots & | & A\vec{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{x}_1 & | & \cdots & | & \lambda_k\vec{x}_k & | & A\vec{x}_{k+1} & | & \cdots & | & A\vec{x}_n \end{bmatrix}$$

$$\|$$

$$A \begin{bmatrix} \vec{x}_1 & | & \vec{x}_2 & | & \cdots & | & \vec{x}_n \end{bmatrix} \qquad\qquad \|$$

$$\begin{bmatrix} \vec{x}_1 & | & \cdots & | & \vec{x}_k & | & \vec{x}_{k+1} & | & \cdots & | & \vec{x}_n \end{bmatrix}
\begin{bmatrix}
\lambda_1 & & & & a_{1,k+1} & \cdots & a_{1,k+1} \\
& \ddots & & & \vdots & \vdots & \vdots \\
& & \lambda_k & & a_{k,k+1} & \cdots & a_{k,k+1} \\
\hline
& & & & a_{k+1,k+1} & \cdots & a_{k+1,k+1} \\
& 0 & & & \vdots & \vdots & \vdots \\
& & & & a_{n,k+1} & \cdots & a_{n,k+1}
\end{bmatrix}$$

$$\uparrow \qquad \cdots \qquad \uparrow$$
$$P^{-1}A\vec{x}_{k+1} \quad \cdots \quad P^{-1}A\vec{x}_n$$

$$\Updownarrow$$

$$AP = P \begin{bmatrix} \operatorname{diag}(\lambda_1,\ldots,\lambda_k) & B \\ 0_{(n-k)\times k} & A_1 \end{bmatrix}$$

$$\Updownarrow$$

$$P^{-1}AP = \begin{bmatrix} \operatorname{diag}(\lambda_1,\ldots,\lambda_k) & B \\ 0_{(n-k)\times k} & A_1 \end{bmatrix}$$

Proof. (Another proof)

Recall that $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ is the standard basis of R_n. Since $I_n = P^{-1}P$,

$$\begin{aligned} \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \cdots & \vec{e}_n \end{bmatrix} = P^{-1}P &= P^{-1}\begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix} \\ &= \begin{bmatrix} P^{-1}\vec{x}_1 & P^{-1}\vec{x}_2 & \cdots & P^{-1}\vec{x}_n \end{bmatrix} \end{aligned}$$

Thus for each j, $1 \leq j \leq n$, $P^{-1}\vec{x}_j = \vec{e}_j$. Also,

$$\begin{aligned} P^{-1}AP &= P^{-1}A\begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix} \\ &= \begin{bmatrix} P^{-1}A\vec{x}_1 & P^{-1}A\vec{x}_2 & \cdots & P^{-1}A\vec{x}_n \end{bmatrix}, \end{aligned}$$

so the j$^{\text{th}}$ column of $P^{-1}AP$, $1 \leq j \leq k$, is equal to

$$P^{-1}(A\vec{x}_j) = P^{-1}(\lambda_j\vec{x}_j) = \lambda_j(P^{-1}\vec{x}_j) = \lambda_j\vec{e}_j.$$

This gives us the first k columns of $P^{-1}AP$, and the result follows. ∎

Definition

Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. The eigenspace of A corresponding to λ is the set
$$E_\lambda(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda\vec{x}\}.$$

Remark

1. The eigenspace $E_\lambda(A)$ is indeed a subspace of $\mathbb{R}^n$ because

$$E_\lambda(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda\vec{x}\} = \{\vec{x} \in \mathbb{R}^n \mid (\lambda I - A)\vec{x} = \vec{0}_n\} = \text{null}(\lambda I - A).$$

2. If λ is not an eigenvalue of A, then $E_\lambda(A) = \{0\}$.

Definition

1. If A is an $n \times n$ matrix and λ is an eigenvalue of A, then the (algebraic) multiplicity of λ is the largest value of m for which

$$c_A(x) = (x - \lambda)^m g(x)$$

for some polynomial $g(x)$, i.e., the multiplicity of λ is the number of times that λ occurs as a root of $c_A(x)$.

2. $\dim(E_\lambda(A))$ is called the geometric multiplicity of λ.

Lemma

If A is an $n \times n$ matrix, and λ is an eigenvalue of A of multiplicity m, then

$$\dim(E_\lambda(A)) \leq m,$$

that is,

$$\text{Geometric multiplicity} \leq \text{Algebraic multiplicity}.$$

Proof.

Let $d = \dim(E_\lambda(A))$, and let $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_d\}$ be a basis of $E_\lambda(A)$. As a consequence, we know that there exists an invertible $n \times n$ matrix P so that

$$P^{-1}AP = \left[\begin{array}{cc} \text{diag}(\lambda, \ldots, \lambda) & B \\ 0_{(n-d) \times d} & A_1 \end{array} \right] = \left[\begin{array}{cc} \lambda I_d & B \\ 0_{(n-d) \times d} & A_1 \end{array} \right]$$

where B is $d \times (n-d)$ and A_1 is $(n-d) \times (n-d)$.

Define $A' = P^{-1}AP$. Then $A \sim A'$, so A and A' have the same characteristic polynomial. Thus

$$\begin{array}{rcl} c_A(x) = c_{A'}(x) = \det(xI - A') & = & \det \left[\begin{array}{cc} (x - \lambda)I_d & -B \\ 0_{(n-d) \times d} & xI_{n-d} - A_1 \end{array} \right] \\ & = & \det[(x - \lambda)I_d] \det(xI_{n-d} - A_1) \\ & = & (x - \lambda)^d c_{A_1}(x) \\ & = & (x - \lambda)^d g(x). \end{array}$$

Since λ has multiplicity m, $d \leq m$, and therefore $\dim(E_\lambda(A)) \leq m$ as required. ■

Characterizing Diagonalizable Matrices

The crucial consequence of this Lemma is the characterization of matrices that are diagonalizable.

Theorem (Covered earlier, here with new terminology)

For an $n \times n$ matrix A, the following two conditions are equivalent.

1. A is diagonalizable.
2. For each eigenvalue λ of A, $\dim(E_\lambda(A))$ is equal to the multiplicity of λ, i.e.,

Diagonalizable

$\Updownarrow$

Geometric multiplicity $=$ Algebraic multiplicity, for all λ.

Problem (Covered earlier, here with new terminology)

If possible, diagonalize the matrix $A = \begin{bmatrix} 3 & 1 & 6 \\ 2 & 1 & 0 \\ -1 & 0 & -3 \end{bmatrix}$. Otherwise, explain why A is not diagonalizable.

Solution

$c_A(x) = (x-3)(x+1)^2$, so A has eigenvalues $\lambda_1 = 3$, $\lambda_2 = \lambda_3 = -1$. Find the dimension of $E_{-1}(A)$ by solving the linear system $(-I - A)\vec{x} = \vec{0}_3$.

$$\begin{bmatrix} 4 & -1 & -6 & | & 0 \\ -2 & -2 & 0 & | & 0 \\ 1 & 0 & 2 & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & | & 0 \\ 0 & 1 & -2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

From this, we see that $\dim(E_{-1}(A)) = 1$. Since -1 is an eigenvalue of multiplicity two, A is not diagonalizable. ■

Problem (Covered earlier, here with new terminology)

Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Show that A is diagonalizable, and that B is not diagonalizable.

Solution

Both A and B are triangular matrices, so we immediately see that A and B have the same eigenvalues: $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$. Thus for each matrix, 1 is an eigenvalue of multiplicity two.

Solving the system $(I - A)\vec{x} = \vec{0}_3$:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

we see that there are two parameters in the general solution, so $\dim(E_1(A)) = 2$. Therefore, A is diagonalizable.

Solution (continued)

Solving the system $(I - B)\vec{x} = \vec{0}_3$:

$$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

we see that the general solution has only one parameter, so $\dim(E_1(B)) = 1$. However, the algebraic multiplicity of $\lambda = 1$ is 2. Therefore, B is not diagonalizable. ■

Complex Eigenvalues

If a matrix has eigenvalues that have imaginary parts (and aren't simply real numbers), we can still find eigenvectors and possibly diagonalize the matrix.

Problem

Diagonalize, if possible, the matrix $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

Solution

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x - 1 & -1 \\ 1 & x - 1 \end{vmatrix} = x^2 - 2x + 2.$$

The roots of $c_A(x)$ are distinct complex numbers: $\lambda_1 = 1 + i$ and $\lambda_2 = 1 - i$, so A is diagonalizable. Corresponding eigenvectors are

$$\vec{x}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{x}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix},$$

respectively.

A diagonalizing matrix for A is

$$P = \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix},$$

and

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

■

Remark

Notice that A is a real matrix, but has complex eigenvalues (and eigenvectors).

Eigenvalues of Real Symmetric Matrices

Theorem

The eigenvalues of any **real symmetric** matrix are **real**.

Proof.

Let A be an $n \times n$ real symmetric matrix, and let λ be an eigenvalue of A. To prove that λ is real, it is enough to prove that $\overline{\lambda} = \lambda$, i.e., λ is equal to its (complex) conjugate.

We use $\overline{A}$ to denote the matrix obtained from A by replacing each entry by its conjugate. Since A is real, $\overline{A} = A$.

Suppose

$$\vec{x} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

is a λ-eigenvector of A. Then $A\vec{x} = \lambda\vec{x}$.

Proof. (continued)

Let $c = \vec{x}^T\overline{\vec{x}} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix} \begin{bmatrix} \overline{z}_1 \\ \overline{z}_2 \\ \vdots \\ \overline{z}_n \end{bmatrix}$.

Then $c = z_1\overline{z}_1 + z_2\overline{z}_2 + \cdots + z_n\overline{z}_n = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2$; since $\vec{x} \neq \vec{0}$, c is a positive real number. Now

$$
\begin{aligned}
\lambda c &= \lambda(\vec{x}^T\overline{\vec{x}}) = (\lambda\vec{x}^T)\overline{\vec{x}} = (\lambda\vec{x})^T\overline{\vec{x}} \\
&= (A\vec{x})^T\overline{\vec{x}} = \vec{x}^T A^T\overline{\vec{x}} \\
&= \vec{x}^T A\overline{\vec{x}} \qquad \text{(since A is symmetric)} \\
&= \vec{x}^T \overline{A}\ \overline{\vec{x}} \qquad \text{(since A is real)} \\
&= \vec{x}^T(\overline{A\vec{x}}) = \vec{x}^T(\overline{\lambda\vec{x}}) = \vec{x}^T\ \overline{\lambda}\ \overline{\vec{x}} \\
&= \overline{\lambda}(\vec{x}^T\overline{\vec{x}}) \\
&= \overline{\lambda}c.
\end{aligned}
$$

Thus, $\lambda c = \overline{\lambda}c$. Since $c \neq 0$, it follows that $\lambda = \overline{\lambda}$, and therefore λ is real. ∎