lyryx with Open Texts

A First Course in

LINEAR ALGEBRA

Lecture Notes

by Karen Seyffarth

Spectral Theory: Diagonalization

advancing learning

Champions of Access to Knowledge

OPEN TEXT

ONLINE ASSESSMENT

SUPPORT

INSTRUCTOR SUPPLEMENTS

Contact Lyryx Today!

info@lyryx.com

A First Course in Linear Algebra

Lecture Notes

Current Lecture Notes Revision: Version 2017 — Revision A

These lecture notes were originally developed by Karen Seyffarth of the University of Calgary. Edits, additions, and revisions have been made to these notes by the editorial team at Lyryx Learning to accompany their text A First Course in Linear Algebra based on K. Kuttler's original text.

In addition we recognize the following contributors. All new content contributed is released under the same license as noted below.

- **Ilijas Farah, York University**
- **Ken Kuttler, Brigham Young University**
- **Asia Weiss, York University**

BE A CHAMPION OF OER!

Contribute suggestions for improvements, new content, or errata:
A new topic
A new example or problem
A new or better proof to an existing theorem
Any other suggestions to improve the material

Contact Lyryx at info@lyryx.com with your ideas.

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 that similarity is transitive.

Since $A \sim B$ and $B \sim C$, there exist invertible $n \times n$ matrices P and Q such that

$$B = P^{-1}AP \text{ and } C = Q^{-1}BQ.$$

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

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

Lemma (Properties of trace)

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

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

The proofs of these are exercises in manipulating sums.

Reminder (Characteristic Polynomial)

For any $n \times n$ 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 $n \times n$ matrices and $A \sim B$, then

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

Proof.

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. $\text{rank}\ (B) = \text{rank}\ (P^{-1}AP)$.
   Since P is invertible, $\text{rank}\ (P^{-1}AP) = \text{rank}\ (P^{-1}A)$, and since P^{-1} is invertible, $\text{rank}\ (P^{-1}A) = \text{rank}\ (A)$. Therefore, $\text{rank}\ (B) = \text{rank}\ (A)$.
3. $\text{tr}(B) = \text{tr}[(P^{-1}A)P] = \text{tr}[P(P^{-1}A)] = \text{tr}[(PP^{-1})A] = \text{tr}(IA) = \text{tr}(A)$.

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.

$\square$

Diagonal Matrices

Definition

An $n \times n$ matrix A is diagonal if it is both upper triangular and lower triangular. Equivalently, all entries except those on the main diagonal are zeros.

Notation

An $n \times n$ diagonal matrix

$$D = \begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & a_n \end{bmatrix}$$

is written $D = \text{diag}(a_1, a_2, a_3, \ldots, a_{n-1}, a_n)$.

Diagonalizability

Definition

An $n \times n$ matrix A is said to be diagonalizable if there exists an invertible $n \times n$ matrix P such that $A = PDP^{-1}$.

Definition

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

Diagonalizing a Matrix

Let A be an $n \times n$ matrix. The process of finding an invertible matrix P and a diagonal matrix D so that $A = PDP^{-1}$ is referred to as diagonalizing the matrix A, and P is called the diagonalizing matrix for A.

The key to diagonalizing a matrix (finding the matrices P and D) lies in the eigenvectors and eigenvalues of the matrix A.

Reminder

Let A be an $n \times n$ matrix and λ a real number. If λ is an eigenvalue of A, then

$$AX = \lambda X$$

for some **nonzero** vector X in $\mathbb{R}^n$. Such a vector X is called a λ-eigenvector of A or an eigenvector of A corresponding to λ.

Theorem

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

1. *A is diagonalizable if and only if it has eigenvectors $X_1, X_2, \ldots, X_n$ so that $P = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$ is invertible.*
2. *If P is invertible, then*

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

where λ_i is the eigenvalue of A corresponding to the eigenvector X_i, i.e., $AX_i = \lambda_i X_i$.

Example

Diagonalize, if possible, the matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$.

Solution

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

Therefore, A has eigenvalues $\lambda_1 = 1$ of multiplicity two and $\lambda_2 = -3$ of multiplicity one.

Eigenvectors for $\lambda_1 = 1$: solve $(I - A)X = 0$.

$$\left[\begin{array}{ccc|c} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ so } X = \begin{bmatrix} s \\ t \\ 0 \end{bmatrix}, s, t \in \mathbb{R}.$$

Solution (continued)

Therefore, basic eigenvectors corresponding to $\lambda_1 = 1$ are $\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right]$ and $\left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right]$.

Eigenvectors for $\lambda_2 = -3$: solve $(-3I - A)X = 0$.

$$\left[\begin{array}{rrr|r} -4 & 0 & -1 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ so } X = \left[\begin{array}{c} -\frac{1}{4}t \\ 0 \\ t \end{array}\right], t \in \mathbb{R}.$$

Therefore, a basic eigenvector corresponding to $\lambda_2 = -3$ is $\left[\begin{array}{r} -1 \\ 0 \\ 4 \end{array}\right]$

Let

$$P = \left[\begin{array}{rrr} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & 0 & 0 \end{array}\right].$$

Then P is invertible (easily checked by computing $\det P$).

Solution (continued)

Furthermore,

$$P^{-1}AP = D = \mathrm{diag}(-3, 1, 1) = \left[\begin{array}{rrr} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

The eigenvalues of A in D (from left to right) occur in the same order as their corresponding eigenvectors as columns of P.

Example

Diagonalize the matrix $A = \begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$

Solution

You can check that A has eigenvalues and corresponding basic eigenvectors:

$$\lambda_1 = 3 \text{ and } X_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}; \lambda_2 = 2 \text{ and } X_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}; \lambda_3 = 1 \text{ and } X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Let $P = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$. Then P is invertible (check this!), so by the previous theorem,

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

Eigenvalues, Eigenvectors, and Diagonalization

Theorem

Let A be an $n \times n$ matrix, and suppose that A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. For each i, let X_i be a λ_i-eigenvector of A. Then $\{X_1, X_2, \ldots, X_m\}$ is linearly independent.

Diagonalizability

Determining whether or not a square matrix A is diagonalizable can be done using eigenvalues and eigenvectors of the matrix A.

Theorem

Let A be an $n \times n$ matrix and suppose it has n distinct eigenvalues. Then it follows that A is diagonalizable.

Proof.

Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denote the n (distinct) eigenvalues of A, and let X_i be an eigenvector of A corresponding to λ_i, $1 \leq i \leq n$. Then $\{X_1, X_2, \ldots, X_n\}$ is an independent set.

A subset of n linearly independent vectors of $\mathbb{R}^n$ also spans $\mathbb{R}^n$, and thus $\{X_1, X_2, \ldots, X_n\}$ is a basis of $\mathbb{R}^n$. Therefore A is diagonalizable. $\square$

Example

Show that the matrix

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

is diagonalizable.

Solution

A has characteristic polynomial

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

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

Since A is 3×3 and has three distinct eigenvalues, A is diagonalizable.

Definition

Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. The eigenspace of A corresponding to λ, written $E_\lambda(A)$ is the set of all eigenvectors corresponding to λ.

In other words, the eigenspace $E_\lambda(A)$ is all X such that $AX = \lambda X$. Notice that this set can be written $E_\lambda(A) = \text{null}(\lambda I - A)$, showing that $E_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

Definition

Let A be an $n \times n$ matrix with characteristic polynomial given by $\det(\lambda I - A)$. Then, the multiplicity of an eigenvalue λ of A is the number of times λ occurs as a root of that characteristic polynomial.

Lemma

If A is an $n \times n$ matrix, then

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

where λ is an eigenvalue of A of multiplicity m.

This result tells us that if λ is an eigenvalue of A, then the number of linearly independent λ-eigenvectors is never more than the multiplicity of λ.

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

Theorem

Let A be an $n \times n$ matrix A. Then A is diagonalizable if and only if for each eigenvalue λ of A, $\dim(E_\lambda(A))$ is equal to the multiplicity of λ.

Example

Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Determine if A is diagonalizable

Solution

Through the usual procedure, we find that the eigenvalues of A are $\lambda_1 = 1, \lambda_2 = 1$. Solving as usual, we find that the eigenvectors are given by

$$t \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and the basic eigenvector is } X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

This means that $\dim(E_\lambda(A)) = 1$, but the multiplicity of $\lambda = 1$ is 2. Therefore this matrix 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

$$X_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix} \text{ and } X_2 = \begin{bmatrix} i \\ 1 \end{bmatrix},$$

respectively.

Solution (continued)

A diagonalizing matrix for A is

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

and

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

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

Eigenvalues of a Real Symmetric Matrix

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

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

is a λ-eigenvector of A. Then $AX = \lambda X$.

Proof (continued).

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

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

$$\begin{aligned}
\lambda y &= \lambda(X^T\overline{X}) = (\lambda X^T)\overline{X} = (\lambda X)^T\overline{X} \\
&= (AX)^T\overline{X} = X^T A^T \overline{X} \\
&= X^T A \overline{X} \quad \text{(since } A \text{ is symmetric)} \\
&= X^T\ \overline{AX} \quad \text{(since } A \text{ is real)} \\
&= X^T(\overline{AX}) = X^T(\overline{\lambda X}) = X^T\ \overline{\lambda X} \\
&= \overline{\lambda}(X^T\overline{X}) \\
&= \overline{\lambda} y.
\end{aligned}$$

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