Introduction to Linear Algebra

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

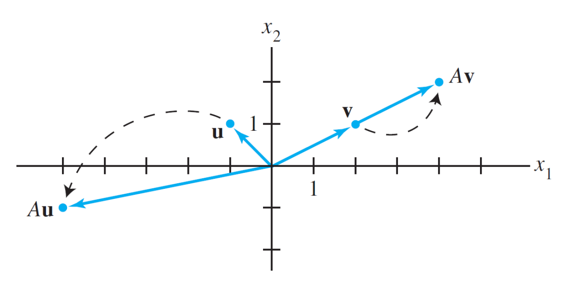
LSSU Math 305

Subsection 1

Eigenvectors and Eigenvalues

Example

- Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
- The images of $\boldsymbol{u}$ and $\boldsymbol{v}$ under multiplication by A are shown in the figure



- In fact, $A\boldsymbol{v}$ is just $2\boldsymbol{v}$.
- So A only "stretches", or dilates, $\boldsymbol{v}$.

Eigenvectors and Eigenvalues

Definition

An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda \boldsymbol{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \lambda \boldsymbol{x}$; such an $\boldsymbol{x}$ is called an **eigenvector corresponding to** λ.

Example: Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\boldsymbol{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Are $\boldsymbol{u}$ and $\boldsymbol{v}$ eigenvectors of A?

$$\begin{aligned} A\boldsymbol{u} &= \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\boldsymbol{u}, \\ A\boldsymbol{v} &= \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix}. \end{aligned}$$

Thus $\boldsymbol{u}$ is an eigenvector corresponding to an eigenvalue -4, but $\boldsymbol{v}$ is not an eigenvector of A.

Example

- Show that 7 is an eigenvalue of matrix $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ and find the corresponding eigenvectors.
- The scalar 7 is an eigenvalue of A if and only if the equation $A\boldsymbol{x} = 7\boldsymbol{x}$ has a nontrivial solution.

  This is equivalent to $A\boldsymbol{x} - 7\boldsymbol{x} = \mathbf{0}$, or $(A - 7I)\boldsymbol{x} = \mathbf{0}$.

  To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}.$$

  The columns of $A - 7I$ are obviously linearly dependent, so the original equation has nontrivial solutions.

  Thus 7 is an eigenvalue of A.

Example (Cont'd)

- To find the corresponding eigenvectors, use row operations:

$$\left[\begin{array}{rr|r} -6 & 6 & 0 \\ 5 & -5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The general solution has the form $\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} x_2 \\ x_2 \end{array}\right] = x_2 \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$.

Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

Eigenspace Corresponding to an Eigenvalue

- Note that in general $A\boldsymbol{x} = \lambda\boldsymbol{x}$ is equivalent to $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$.
- Thus λ is an eigenvalue of an $n \times n$ matrix A if and only if the equation

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

  has a nontrivial solution.
- The set of all solutions is just the null space of the matrix $A - \lambda I$.
- So this set is a subspace of $\mathbb{R}^n$ and is called the **eigenspace** of A corresponding to λ.
- The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

Example

- Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is 2. Find a basis for the corresponding eigenspace.
- Form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}.$$

Row reduce the augmented matrix for $(A - 2I)\boldsymbol{x} = \mathbf{0}$:

$$\left[\begin{array}{ccc|c} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A - 2I)\boldsymbol{x} = \mathbf{0}$ has free variables.

Example (Cont'd)

- We must have $2x_1 - x_2 + 6x_3 = 0$.
  This gives $x_1 = \frac{1}{2}x_2 - 3x_3$, with x_2 and x_3 free.
  So the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \ x_2 \text{ and } x_3 \text{ free.}$$

  The eigenspace is a two-dimensional subspace of $\mathbb{R}^3$.
  A basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Eigenvalues of Triangular Matrices

Theorem

The eigenvalues of a triangular matrix are the entries on its main diagonal.

- For simplicity, consider the 3×3 case.
  If A is upper triangular, then $A - \lambda I$ has the form

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}. \end{aligned}$$

  The scalar λ is an eigenvalue of A if and only if the equation $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution, that is, if and only if the equation has a free variable. Because of the zero entries in $A - \lambda I$, $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a free variable if and only if at least one of the entries on the diagonal of $A - \lambda I$ is zero. This happens if and only if λ equals one of the entries a_{11}, a_{22}, a_{33} in A.

Example

- Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$.
- The eigenvalues of A are 3, 0, and 2.
- The eigenvalues of B are 4 and 1.

Eigenvalues and Invertibility

- Suppose a matrix A has an eigenvalue of 0.
- This happens if and only if the equation

$$A\boldsymbol{x} = 0\boldsymbol{x}$$

  has a nontrivial solution.
- But this is equivalent to $A\boldsymbol{x} = \mathbf{0}$, which has a nontrivial solution if and only if A is not invertible.
- Thus 0 is an eigenvalue of A if and only if A is not invertible.
- This fact may be added to the Invertible Matrix Theorem.

Eigenvectors Corresponding to Distinct Eigenvalues

Theorem

If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

- Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly dependent.
  Since $\mathbf{v}_1$ is nonzero, we conclude by a previous theorem that one of the vectors in the set is a linear combination of the preceding vectors. Let p be the least index such that $\mathbf{v}_{p+1}$ is a linear combination of the preceding (linearly independent) vectors. Then there exist scalars $c_1, \ldots, c_p$ such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1}.$$

  Multiplying both sides by λ_{p+1}, we get

$$c_1\lambda_{p+1}\mathbf{v}_1 + \cdots + c_p\lambda_{p+1}\mathbf{v}_{p+1} = \lambda_{p+1}\mathbf{v}_{p+1}.$$

Proof (Cont'd)

- Multiplying $c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p = \boldsymbol{v}_{p+1}$ by A and using the fact that $A\boldsymbol{v}_k = \lambda_k\boldsymbol{v}_k$ for each k, we obtain

$$\begin{aligned} c_1A\boldsymbol{v}_1 + \cdots + c_pA\boldsymbol{v}_p &= A\boldsymbol{v}_{p+1}, \\ c_1\lambda_1\boldsymbol{v}_1 + \cdots + c_p\lambda_p\boldsymbol{v}_p &= \lambda_{p+1}\boldsymbol{v}_{p+1}. \end{aligned}$$

Subtracting $c_1\lambda_{p+1}\boldsymbol{v}_1 + \cdots + c_p\lambda_{p+1}\boldsymbol{v}_{p+1} = \lambda_{p+1}\boldsymbol{v}_{p+1}$ from the equation above, we have

$$c_1(\lambda_1 - \lambda_{p+1})\boldsymbol{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\boldsymbol{v}_p = \boldsymbol{0}.$$

Since $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ is linearly independent, the weights in the equation above are all zero. But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct. Hence $c_i = 0$ for $i = 1, \ldots, p$. But then we get $\boldsymbol{v}_{p+1} = \boldsymbol{0}$, which is impossible.

Hence $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ cannot be linearly dependent and therefore must be linearly independent.

Subsection 2

The Characteristic Equation

Example

- Find the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.
- We must find all scalars such that the matrix equation $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution.

  By the Invertible Matrix Theorem, this problem is equivalent to finding all λ such that the matrix $A - \lambda I$ is not invertible, where

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}.$$

  By a previous result, this matrix fails to be invertible precisely when its determinant is zero.

  So the eigenvalues of A are the solutions of the equation

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0.$$

Example (Cont'd)

- Recall that

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

So

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix} \\
&= (2-\lambda)(-6-\lambda) - 3 \cdot 3 \\
&= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\
&= \lambda^2 + 4\lambda - 21 \\
&= (\lambda - 3)(\lambda + 7).
\end{aligned}$$

If $\det(A - \lambda I) = 0$, then $\lambda = 3$ or $\lambda = -7$.

So the eigenvalues of A are 3 and -7.

The Determinant Reviewed

- Let A be an $n \times n$ matrix;
- Let U be any echelon form obtained from A by row replacements and row interchanges (without scaling).
- Let r be the number of such row interchanges.
- Then the **determinant** of A, written as $\det A$, is $(-1)^r$ times the product of the diagonal entries $u_{11}, \ldots, u_{nn}$ in U.
- If A is invertible, then $u_{11}, \ldots, u_{nn}$ are all pivots (because $A \sim I_n$ and the u_{ii} have not been scaled to 1's).
- Otherwise, at least u_{nn} is zero, and the product $u_{11} \cdots u_{nn}$ is zero.
- Thus

$$\det A = \begin{cases} (-1)^r \cdot (\text{product of pivots of } U), & \text{if } A \text{ is invertible,} \\ 0, & \text{if } A \text{ is not invertible.} \end{cases}$$

Example

- Compute $\det A$ for $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.
- The following row reduction uses one row interchange:

$$A \longrightarrow \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U.$$

So $\det A$ equals $(-1)^1 \cdot (1 \cdot (-2) \cdot (-1)) = -2$.

The Invertible Matrix Theorem (Revisited)

Theorem (The Invertible Matrix Theorem Cont'd)

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

(s) The number 0 is not an eigenvalue of A.

(t) The determinant of A is not zero.

- When A is a 3×3 matrix, $|\det A|$ turns out to be the volume of the parallelepiped determined by the columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ of A.

  This volume is nonzero if and only if the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ are linearly independent, in which case the matrix A is invertible.

Properties of Determinants

Theorem (Properties of Determinants)

Let A and B be $n \times n$ matrices.

(a) A is invertible if and only if $\det A \neq 0$.

(b) $\det AB = (\det A)(\det B)$.

(c) $\det A^T = \det A$.

(d) If A is triangular, then $\det A$ is the product of the entries on the main diagonal of A.

(e) A row replacement operation on A does not change the determinant.

A row interchange changes the sign of the determinant.

A row scaling also scales the determinant by the same scalar factor.

The Characteristic Equation

- Part (a) of the preceding theorem shows how to determine when a matrix of the form $A - \lambda I$ is not invertible.
- The scalar equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** of A.
- By previous work we know the following:

A scalar is an eigenvalue of an $n \times n$ matrix A if and only if satisfies the characteristic equation

$$\det(A - \lambda I) = 0.$$

Example

- Find the characteristic equation of $A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.
- We have

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\ &= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda). \end{aligned}$$

The characteristic equation is

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0.$$

Expanding the product, we can also write

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0.$$

The Characteristic Polynomial

- It can be shown that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n called the **characteristic polynomial** of A.
- An eigenvalue a of A is said to have **multiplicity** k if $(\lambda - a)$ occurs k times as a factor of the characteristic polynomial.
- In general, the (**algebraic**) **multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic equation.

Example

- The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.
- Factor the polynomial

$$\begin{aligned} \lambda^6 - 4\lambda^5 - 12\lambda^4 &= \lambda^4(\lambda^2 - 4\lambda - 12) \\ &= \lambda^4(\lambda - 6)(\lambda + 2). \end{aligned}$$

The eigenvalues are 0 (multiplicity 4), 6 (multiplicity 1) and -2 (multiplicity 1).

Similarity

- If A and B are $n \times n$ matrices, then A is **similar to** B if there is an invertible matrix P such that
$$P^{-1}AP = B, \quad \text{or, equivalently,} \quad A = PBP^{-1}.$$
- Writing Q for P^{-1}, we have $Q^{-1}BQ = A$.
- So B is also similar to A.
- We say simply that A and B are **similar**.
- Changing A into $P^{-1}AP$ is called a **similarity transformation**.

Similarity and Characteristic Polynomials

Theorem

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

- If $B = P^{-1}AP$, then

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P.$$

Using the multiplicative property of the determinant, we compute

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P). \end{aligned}$$

But $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$.

So we get $\det(B - \lambda I) = \det(A - \lambda I)$.

Subsection 3

Diagonalization

Example: Powers of Diagonal Matrices

- Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$.
- Then $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$.
- Also $D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$
- In general, $D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}$.

Example

- Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$.
  Find a formula for A^k, given that $A = PDP^{-1}$, where
  $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$.
- The standard formula for the inverse of a 2×2 matrix yields
  $P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$.
- Then, by associativity of matrix multiplication,
$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDDP^{-1} \\ &= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}. \end{aligned}$$
- Again
  $A^3 = (PDP^{-1})A^2 = (PDP^{-1})(PD^2P^{-1}) = PDD^2P^{-1} = PD^3P^{-1}$.

Example (Cont'd)

- In general, for $k \geq 1$,

$$
\begin{aligned}
A^k &= PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 5^k & 3^k \\ -5^k & -2 \cdot 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ -2 \cdot 5^k + 2 \cdot 3^k & -5^k + 2 \cdot 3^k \end{bmatrix}.
\end{aligned}
$$

Diagonalizability

- A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D.

Theorem (The Diagonalization Theorem)

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

- In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$.

  We call such a basis an **eigenvector basis** of $\mathbb{R}^n$.

Proof of the Diagonalization Theorem

- First, observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and if D is any diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then

$$AP = A[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n] = [A\mathbf{v}_1 \ A\mathbf{v}_2 \ \cdots \ A\mathbf{v}_n].$$

- Also

$$PD = P \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1 \mathbf{v}_1 \ \lambda_2 \mathbf{v}_2 \ \cdots \ \lambda_n \mathbf{v}_n].$$

Now suppose A is diagonalizable and $A = PDP^{-1}$. Then right-multiplying this relation by P, we have $AP = PD$. In this case, we get that

$$[A\mathbf{v}_1 \ A\mathbf{v}_2 \ \cdots \ A\mathbf{v}_n] = [\lambda_1 \mathbf{v}_1 \ \lambda_2 \mathbf{v}_2 \ \cdots \lambda_n \mathbf{v}_n].$$

Proof of the Diagonalization Theorem (Cont'd)

- Equating columns, we find that $A\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1$, $A\boldsymbol{v}_2 = \lambda_2 \boldsymbol{v}_2$, ..., $A\boldsymbol{v}_n = \lambda_n \boldsymbol{v}_n$. Since P is invertible, its columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ must be linearly independent. Also, since these columns are nonzero, the last equations show that $\lambda_1, \ldots, \lambda_n$ are eigenvalues and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are corresponding eigenvectors. This argument proves the "only if" parts of the first and second statements, along with the third statement, of the theorem.

  Finally, given any n eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, use them to construct the columns of P and use corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ to construct D. Then, we get

$$AP = A[\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_n] = [A\boldsymbol{v}_1 \ \cdots A\boldsymbol{v}_n] = [\lambda_1 \boldsymbol{v}_1 \ \cdots \ \lambda_n \boldsymbol{v}_n] = PD.$$

  This is true without any condition on the eigenvectors. If, in fact, the eigenvectors are linearly independent, then P is invertible (by the Invertible Matrix Theorem), and $AP = PD$ implies that $A = PDP^{-1}$.

The Diagonalization Process

- Diagonalize the the matrix $A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$, if possible.

  That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

  Step 1 (Find the eigenvalues of A): We have

$$\begin{aligned} 0 &= \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} \\ &= (1-\lambda)(-5-\lambda)(1-\lambda) - 27 - 27 \\ &\quad - 9(-5-\lambda) + 9(1-\lambda) + 9(1-\lambda) \\ &= -\lambda^3 + 2\lambda^2 - \lambda - 5\lambda^2 + 10\lambda - 5 - 27 - 27 \\ &\quad + 45 + 9\lambda + 9 - 9\lambda + 9 - 9\lambda \\ &= -\lambda^3 - 3\lambda^2 + 4 \end{aligned}$$

The Diagonalization Process (Step 1 Cont'd)

$$\begin{aligned} -\lambda^3 - 3\lambda^2 + 4 &= -\lambda^3 + \lambda^2 - 4\lambda^2 + 4 \\ &= -\lambda^2(\lambda - 1) - 4(\lambda - 1)(\lambda + 1) \\ &= (\lambda - 1)(-\lambda^2 - 4\lambda - 4) \\ &= -(\lambda - 1)(\lambda + 2)^2. \end{aligned}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$.

The Diagonalization Process (Step 2)

Step 2 (Find three linearly independent eigenvectors of A): Three vectors are needed because A is a 3×3 matrix. This is the critical step. If it fails, then the theorem says that A cannot be diagonalized.

- **Basis for** $\lambda = 1$:

$$(A - I)\boldsymbol{v} = \boldsymbol{0} \Rightarrow \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \boldsymbol{0}$$

$$\left[\begin{array}{ccc|c} 0 & 3 & 3 & 0 \\ -3 & -6 & -3 & 0 \\ 3 & 3 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left\{ \begin{array}{l} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{array} \right. \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

So $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ is a basis.

The Diagonalization Process (Step 2 Cont'd)

Basis for $\lambda = -2$:

$$(A+2I)\boldsymbol{v} = \mathbf{0} \;\Rightarrow\; \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{ccc|c} 3 & 3 & 3 & 0 \\ -3 & -3 & -3 & 0 \\ 3 & 3 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$x_1 = -x_2 - x_3 \;\Rightarrow\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

So $\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ form a basis.

We can check that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a linearly independent set.

The Diagonalization Process (Steps 3 and 4)

Step 3 (Construct P from the vectors in Step 2): The order of the vectors is unimportant. Using the order chosen in Step 2, form

$$P = [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \boldsymbol{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Step 4 (Construct D from the corresponding eigenvalues): In this step, it is essential that the order of the eigenvalues matches the order chosen for the columns of P. Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

The Diagonalization Process (Final Check)

- We check that P and D really work.
- To avoid computing P^{-1}, simply verify that $AP = PD$.

$$\begin{aligned} AP &= \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}; \\ PD &= \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}. \end{aligned}$$

Example

- Diagonalize the matrix $A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$, if possible.
- The characteristic equation of A is

$$\begin{array}{rcl} 0 & = & \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} \\ & = & (2-\lambda)(-6-\lambda)(1-\lambda) - 36 - 36 \\ & & +9(6+\lambda) + 9(2-\lambda) + 16(1-\lambda) \\ & = & -\lambda^3 - 4\lambda^2 + 12\lambda + \lambda^2 + 4\lambda - 12 - 36 - 36 \\ & & +54 + 9\lambda + 18 - 9\lambda + 16 - 16\lambda \\ & = & -\lambda^3 - 3\lambda^2 + 4 \\ & \stackrel{\text{previous}}{=} & \cdots \\ & = & -(\lambda-1)(\lambda+2)^2. \end{array}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$.

Example (Cont'd)

- Consider $\lambda = 1$:

  We have

$$(A - I)\boldsymbol{v} = \mathbf{0} \;\Rightarrow\; \begin{bmatrix} 1 & 4 & 3 \\ -4 & -7 & -3 \\ 3 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ -4 & -7 & -3 & 0 \\ 3 & 3 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 0 & 9 & 9 & 0 \\ 0 & -9 & -9 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left\{ \begin{array}{rcl} x_1 + 4x_2 + 3x_3 & = & 0 \\ x_2 + x_3 & = & 0 \end{array} \right. \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

So a basis for $\lambda = 1$ consists of $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

Example (Cont'd)

- Consider $\lambda = -2$:
  We have

$$(A+2I)\mathbf{v} = \mathbf{0} \;\Rightarrow\; \begin{bmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{rrr|r} 4 & 4 & 3 & 0 \\ -4 & -4 & -3 & 0 \\ 3 & 3 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ -4 & -4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left\{ \begin{array}{rcl} x_1 + x_2 + x_3 & = & 0 \\ -x_3 & = & 0 \end{array} \right. \;\Rightarrow\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

So a basis for $\lambda = 1$ consists of $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

Example (Cont'd)

- There are no other eigenvalues, and every eigenvector of A is a multiple of either $\boldsymbol{v}_1$ or $\boldsymbol{v}_2$.

  Hence it is impossible to construct a basis of $\boldsymbol{R}^3$ using eigenvectors of A.

  So by the theorem, A is not diagonalizable.

Sufficient Condition for Diagonalizability

Theorem

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

- Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of a matrix A.

  Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is linearly independent, by a previous theorem.

  Hence A is diagonalizable.
- It is not necessary for an $n \times n$ matrix to have n distinct eigenvalues in order to be diagonalizable.

  The 3×3 matrix in a previous example was diagonalizable even though it has only two distinct eigenvalues.

Example

- Determine if the matrix $A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$ is diagonalizable.
- Since the matrix is triangular, its eigenvalues are $5, 0$ and -2.

  Since A is a 3×3 matrix with three distinct eigenvalues, A is diagonalizable.

Matrices Whose Eigenvalues Are Not Distinct

Theorem

Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.

(a) For $1 \le k \le p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

(b) The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if:

  (i) the characteristic polynomial factors completely into linear factors;

  (ii) the dimension of the eigenspace for each k equals the multiplicity of k.

(c) If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Example

- Diagonalize the matrix $A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}$, if possible.
- Since A is a triangular matrix, the eigenvalues are 5 and -3, each with multiplicity 2.

  We find a basis for each eigenspace.

Example (Cont'd)

- For $\lambda = 5$:

$$(A - 5I)\mathbf{v} = \mathbf{0} \Rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 \\ 1 & -2 & 0 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 & 0 \\ 1 & -2 & 0 & -8 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 0 & 8 & 0 \\ 0 & 1 & -4 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -8x_3 - 16x_4 \\ 4x_3 + 4x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}.$$

Example (Cont'd)

- For $\lambda = -3$:

$$(A+3I)\boldsymbol{v} = \mathbf{0} \Rightarrow \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ -1 & -2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{cccc|c} 8 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ -1 & -2 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Example (Cont'd)

- The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\} = \left\{ \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is linearly independent, by the preceding theorem.

  So the matrix $P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \mathbf{v}_4]$ is invertible, and $A = PDP^{-1}$, where

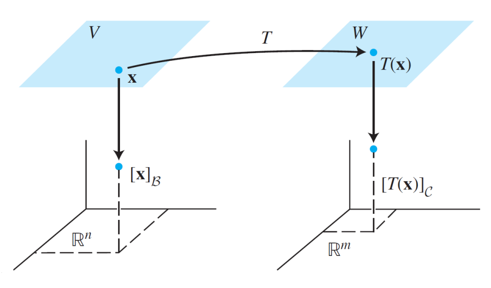
$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

Subsection 4

Eigenvectors and Linear Transformations

Linear Transformations and Bases

- Let V be an n-dimensional vector space.
- Let W be an m-dimensional vector space.
- Let T be any linear transformation from V to W.
- To associate a matrix with T, choose (ordered) bases $\mathcal{B}$ and $\mathcal{C}$ for V and W, respectively.
- Given any $\boldsymbol{x}$ in V, the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ is in $\mathbb{R}^n$ and the coordinate vector of its image, $[T(\boldsymbol{x})]_{\mathcal{C}}$, is in $\mathbb{R}^m$:



The Matrix of a Linear Transformation

- We establish a connection between $[\boldsymbol{x}]_{\mathcal{B}}$ and $[T(\boldsymbol{x})]_{\mathcal{C}}$.
- Let $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be the basis $\mathcal{B}$ for V.
- If $\boldsymbol{x} = r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n$, then $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$.
- Moreover, since T is linear,

$$T(\boldsymbol{x}) = T(r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n) = r_1T(\boldsymbol{b}_1) + \cdots + r_nT(\boldsymbol{b}_n).$$

- Since the coordinate mapping from W to $\mathbb{R}^m$ is linear we get

$$[T(\boldsymbol{x})]_{\mathcal{C}} = r_1[T(\boldsymbol{b}_1)]_{\mathcal{C}} + \cdots + r_n[T(\boldsymbol{b}_n)]_{\boldsymbol{C}}.$$

The Matrix of a Linear Transformation (Cont'd)

- Since $\mathcal{C}$-coordinate vectors are in $\mathbb{R}^m$, the last vector equation can be written as a matrix equation, namely,
$$[T(\boldsymbol{x})]_{\mathcal{C}} = M[\boldsymbol{x}]_{\mathcal{B}},$$
where
$$M = \begin{bmatrix} [T(\boldsymbol{b}_1)]_{\mathcal{C}} & [T(\boldsymbol{b}_2)]_{\mathcal{C}} & \cdots & [T(\boldsymbol{b}_n)]_{\mathcal{C}} \end{bmatrix}.$$
- The matrix M is a matrix representation of T, called the **matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$**.

$$\begin{tikzcd} \boldsymbol{x} \arrow[r, "T"] \arrow[d] & T(\boldsymbol{x}) \arrow[d] \\ {[\boldsymbol{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } M"'] & {[T(\boldsymbol{x})]_{\mathcal{C}}} \end{tikzcd}$$

Example

- Suppose $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ is a basis for V and $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$ is a basis for W.
- Let $T : V \to W$ be a linear transformation with the property that

$$T(\boldsymbol{b}_1) = 3\boldsymbol{c}_1 - 2\boldsymbol{c}_2 + 5\boldsymbol{c}_3 \quad \text{and} \quad T(\boldsymbol{b}_2) = 4\boldsymbol{c}_1 + 7\boldsymbol{c}_2 - \boldsymbol{c}_3.$$

- Find the matrix M for T relative to $\mathcal{B}$ and $\mathcal{C}$.
- The $\mathcal{C}$-coordinate vectors of the images of $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are

$$[T(\boldsymbol{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \text{and} \quad [T(\boldsymbol{b}_2)]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}.$$

- Hence

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}.$$

Linear Transformations from V into V

- In case W is the same as V and the basis $\mathcal{C}$ is the same as $\mathcal{B}$, the matrix M is called the **matrix for T relative to $\mathcal{B}$**, or simply the **$\mathcal{B}$-matrix for T**, and is denoted by $[T]_{\mathcal{B}}$:

$$\begin{array}{ccc} \boldsymbol{x} & \xrightarrow{\quad T \quad} & T(\boldsymbol{x}) \\ \downarrow & & \downarrow \\ [\boldsymbol{x}]_{\mathcal{B}} & \xrightarrow[\text{by } [T]_{\mathcal{B}}]{\text{Multiplication}} & [T(\boldsymbol{x})]_{\mathcal{B}} \end{array}$$

- The $\mathcal{B}$-matrix for $T : V \to V$ satisfies

$$[T(\boldsymbol{x})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}, \text{ for all } \boldsymbol{x} \text{ in } V.$$

Example

- The mapping $T : \mathbb{P}_2 \to \mathbb{P}_2$ defined by
$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$
is a linear transformation.
  (a) Find the $\mathcal{B}$-matrix for T, where $\mathcal{B}$ is the basis $\{1, t, t^2\}$.
  (b) Verify that $[T(\boldsymbol{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}$ for each $\boldsymbol{p}$ in $\mathbb{P}_2$.

(a) Compute the images of the basis vectors:
$$T(1) = 0, \quad T(t) = 1, \quad T(t^2) = 2t.$$
Then write the $\mathcal{B}$-coordinate vectors of $T(1)$, $T(t)$ and $T(t^2)$ (which are found by inspection in this example): and place them together as the $\mathcal{B}$-matrix for T:
$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \left[T(t^2)\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

Example (Cont'd)

We place the coordinate vectors together to form $[T]_{\mathcal{B}} = [[T(1)]_{\mathcal{B}}\ [T(t)]_{\mathcal{B}}\ [T(t^2)]_{\mathcal{B}}]$:

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

(b) For a general $\boldsymbol{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$\begin{aligned} [T(\boldsymbol{p})]_{\mathcal{B}} &= [a_1 + 2a_2 t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}. \end{aligned}$$

Diagonal Matrix Representation

Theorem (Diagonal Matrix Representation)

Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$ matrix. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

- Denote the columns of P by $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$, so that $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ and $P = [\boldsymbol{b}_1 \ \cdots \ \boldsymbol{b}_n]$.

  In this case, P is the change-of-coordinates matrix $P_{\mathcal{B}}$ where

$$P[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x} \quad \text{and} \quad [\boldsymbol{x}]_{\mathcal{B}} = P^{-1}\boldsymbol{x}.$$

Diagonal Matrix Representation (Cont'd)

- If $T(\boldsymbol{x}) = A\boldsymbol{x}$ for $\boldsymbol{x}$ in $\mathbb{R}^n$, then

$$\begin{array}{rcl} [T]_{\mathcal{B}} & = & [[T(\boldsymbol{b}_1)]_{\mathcal{B}} \ \cdots \ [T(\boldsymbol{b}_n)]_{\mathcal{B}}] \\ & = & [[A\boldsymbol{b}_1]_{\mathcal{B}} \ \cdots \ [A\boldsymbol{b}_n]_{\mathcal{B}}] \\ & = & [P^{-1}A\boldsymbol{b}_1 \ \cdots \ P^{-1}A\boldsymbol{b}_n] \\ & = & P^{-1}A[\boldsymbol{b}_1 \ \cdots \ \boldsymbol{b}_n] \\ & = & P^{-1}AP. \end{array}$$

Since $A = PDP^{-1}$, we have $[T]_{\mathcal{B}} = P^{-1}AP = D$.

Example

- Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\boldsymbol{x}) = A\boldsymbol{x}$, where $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$.

  Find a basis $\mathcal{B}$ for $\mathbb{R}^2$ with the property that the $\mathcal{B}$-matrix for T is a diagonal matrix.
- We diagonalize A by finding its eigenvalues and the corresponding eigenvectors.

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 7-\lambda & 2 \\ -4 & 1-\lambda \end{vmatrix} = (7-\lambda)(1-\lambda) + 8 \\ &= \lambda^2 - 8\lambda + 15 = (\lambda - 3)(\lambda - 5). \end{aligned}$$

- For $\lambda = 3$:

$$(A - 3I)\boldsymbol{v} = \boldsymbol{0} \quad \Rightarrow \quad \begin{bmatrix} 4 & 2 \\ -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boldsymbol{0}$$

$$\left[\begin{array}{cc|c} 4 & 2 & 0 \\ -4 & -2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \quad \Rightarrow \quad \boldsymbol{v}_1 = \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}.$$

Example

- For $\lambda = 5$:

$$(A - 5I)\boldsymbol{v} = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} 2 & 2 \\ -4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

$$\left[\begin{array}{cc|c} 2 & 2 & 0 \\ -4 & -4 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad \Rightarrow \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

So we get $A = PDP^{-1}$, with $P = \begin{bmatrix} -\frac{1}{2} & -1 \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$.

Since the columns of P are eigenvectors of A, by the preceding theorem, D is the $\mathcal{B}$-matrix for T when $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$.

The mappings $\boldsymbol{x} \mapsto A\boldsymbol{x}$ and $\boldsymbol{u} \mapsto D\boldsymbol{u}$ describe the same linear transformation, relative to different bases.

Similarity of Matrix Representations

- If A is similar to a matrix C, with $A = PCP^{-1}$, then $C = P^{-1}AP$ is the $\mathcal{B}$-matrix for the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ when the basis $\mathcal{B}$ is formed from the columns of P.

$$\begin{tikzcd}
\boldsymbol{x} \arrow[r, "\text{Multiplication}", "\text{by } A"'] \arrow[d] & A\boldsymbol{x} \\
\boldsymbol{x}_{\mathcal{B}} \arrow[r, "\text{Multiplication}", "\text{by } C"'] & {[A\boldsymbol{x}]_{\mathcal{B}}} \arrow[u]
\end{tikzcd}$$

- Conversely, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $T(\boldsymbol{x}) = A\boldsymbol{x}$, and if $\mathcal{B}$ is any basis for $\mathbb{R}^n$, then the $\mathcal{B}$-matrix for T is similar to A.
- Thus, the set of all matrices similar to a matrix A coincides with the set of all matrix representations of the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

Example

- Let $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, $\boldsymbol{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

  The characteristic polynomial of A is $(\lambda + 2)^2$, but the eigenspace for the eigenvalue -2 is only one-dimensional; so A is not diagonalizable.

  However, the basis $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ has the property that the $\mathcal{B}$-matrix C for the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is a triangular matrix called the **Jordan form** of A. Find this $\mathcal{B}$-matrix C.

- If $P = [\boldsymbol{b}_1 \ \boldsymbol{b}_2]$, then the $\mathcal{B}$-matrix is $C = P^{-1}AP$.

  Compute

$$\begin{aligned} AP &= \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} \\ C &= P^{-1}AP = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}. \end{aligned}$$