Discrete Mathematics II | Spring 2022

Section 1.1: Solutions and Elementary Operations

Lecturer: Mr. Hall | *Math 270B*

Motivation

Find all solutions of the (linear) equation in one variable:

$$ax = b$$

Solution

- If $a \neq 0$, there is a unique solution $x = b/a$.
- Else if $a = 0$ and
  1) $b \neq 0$, there is no solution.
  2) $b = 0$, there are infinitely many solutions, in fact any $x \in \mathbb{R}$ is a solution.

This is a complete description of all possible solutions of $ax = b$.

Can we do the same for linear equations in more variables?

Definitions

A **linear equation** is an expression

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $n \geq 1$, $a_1, \ldots, a_n$ are real numbers, *not all of them equal to zero*, and b is a real number.

A **system of linear equations** is a set of $m \geq 1$ linear equations. It is not required that $m = n$.

A **solution** to a system of m equations in n variables is an n-tuple of numbers that satisfy each of the equations.

Solve a system means 'find *all* solutions to the system.'

Systems of Linear Equations

A **system** of **linear equations**:

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{array}$$

- variables: x_1, x_2, x_3.
- coefficients:

$$\begin{array}{rcrcrcr} 1x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -1x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{array}$$

- constant terms:

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{array}$$

$x_1 = -3$, $x_2 = -1$, $x_3 = 0$ is a solution to the system

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{array}$$

because

$$\begin{array}{rcrcrcr} (-3) & - & 2(-1) & - & 7 \cdot 0 & = & -1 \\ -(-3) & + & 3(-1) & + & 6 \cdot 0 & = & 0. \end{array}$$

Another solution to the system is $x_1 = 6$, $x_2 = 0$, $x_3 = 1$.

However, $x_1 = -1$, $x_2 = 0$, $x_3 = 0$ is not a solution to the system, because

$$\begin{array}{rcrcrcrl} (-1) & - & 2 \cdot 0 & - & 7 \cdot 0 & = & -1 & \\ -(-1) & + & 3 \cdot 0 & + & 6 \cdot 0 & = & 1 & \neq 0 \end{array}$$

The system above is **consistent**, meaning that the system has **at least one** solution.

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & + & x_3 & = & -8 \end{array}$$

is an example of an **inconsistent** system, meaning that it has no solutions.

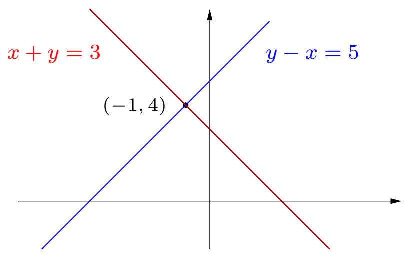
Graphical Solutions

Consider the system of linear equations in two variables

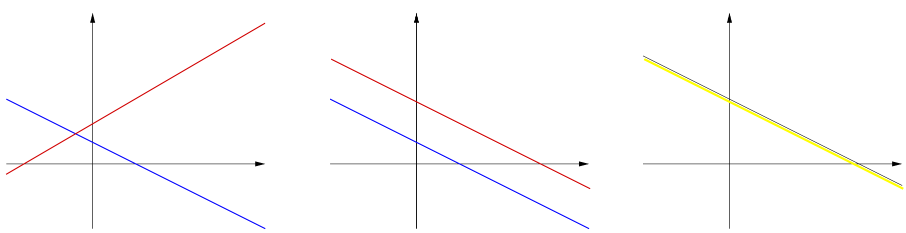
$$\begin{aligned} x + y &= 3 \\ y - x &= 5 \end{aligned}$$

A solution to this system is a pair (x, y) satisfying both equations.

Since each equation corresponds to a line, a solution to the system corresponds to a point that lies on both lines, so the solutions to the system can be found by graphing the two lines and determining where they intersect.



Given a system of two equations in two variables, graphed on the xy-coordinate plane, there are three possibilities, as illustrated below.



intersect in one point	parallel but different	lines are the same
consistent	inconsistent	consistent
(unique solution)	(no solutions)	(infinitely many solutions)

We can see that for a system of linear equations in **two variables**, exactly one of the following holds:

1. the system is **inconsistent**;
2. the system has a **unique** solution, i.e., exactly one solution;
3. the system has **infinitely many** solutions.

(We will see in what follows that this generalizes to systems of linear equations in more than two variables.)

The system of linear equations in three variables that we saw earlier

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 0, \end{array}$$

has solutions $x_1 = -3 + 9t$, $x_2 = -1 + t$, $x_3 = t$ where t is any real number (written $t \in \mathbb{R}$).

Verify this by substituting the expressions for x_1, x_2, and x_3 into the two equations.

t is called a **parameter**, and the expression

$$x_1 = -3 + 9t, x_2 = -1 + t, x_3 = t, \text{ where } t \in \mathbb{R}$$

is called the **general solution** in parametric form.

Definition

Two systems of linear equations are **equivalent** if they have exactly the same solutions.

The two systems of linear equations

$$\begin{array}{rcrcr} 2x & + & y & = & 2 \\ 3x & & & = & 3 \end{array} \quad \text{and} \quad \begin{array}{rcrcr} x & + & y & = & 1 \\ & & y & = & 0 \end{array}$$

are **equivalent** because both systems have the unique solution $x = 1$, $y = 0$.

Elementary Operations
We solve a system of linear equations by using *Elementary Operations* to transform the system into an equivalent but simpler system from which the solution can be easily obtained. Performing a sequence of elementary operations on a system of linear equations results in an equivalent system of linear equations, with the exact same solutions.

Three types of Elementary Operations

- **Type I:** Interchange two equations, $r_i \leftrightarrow r_j$.
- **Type II:** Multiply an equation by a nonzero number, kr_i.
- **Type III:** Add a multiple of one equation to a different equation, $kr_i + r_j$.

Example

Consider the system of linear equations
$$\begin{array}{rcrcrcr} 3x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 1 \\ 2x_1 & & & - & x_3 & = & 3 \end{array}$$

- Interchange first two equations (Type I elementary operation):

$$r_1 \leftrightarrow r_2 \qquad \begin{array}{rcrcrcr} -x_1 & + & 3x_2 & + & 6x_3 & = & 1 \\ 3x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ 2x_1 & & & - & x_3 & = & 3 \end{array}$$

- Multiply first equation by -2 (Type II elementary operation):

$$-2r_1 \qquad \begin{array}{rcrcrcr} -6x_1 & + & 4x_2 & + & 14x_3 & = & 2 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 1 \\ 2x_1 & & & - & x_3 & = & 3 \end{array}$$

- Add 3 times the second equation to the first equation (Type III elementary operation):

$$3r_2 + r_1 \qquad \begin{array}{rcrcrcr} & & 7x_2 & + & 11x_3 & = & 2 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 1 \\ 2x_1 & & & - & x_3 & = & 3 \end{array}$$

The Augmented Matrix

Represent a system of linear equations with its augmented matrix.

Example

The system of linear equations

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 7x_3 & = & -1 \\ -x_1 & + & 3x_2 & + & 6x_3 & = & 0 \end{array}$$

is represented by the **augmented matrix**

$$\left[\begin{array}{rrr|r} 1 & -2 & -7 & -1 \\ -1 & 3 & 6 & 0 \end{array}\right]$$

(A **matrix** is a rectangular array of numbers.)

Note: Two other **matrices** associated with a system of linear equations are the coefficient matrix and the constant matrix.

$$\left[\begin{array}{rrr} 1 & -2 & -7 \\ -1 & 3 & 6 \end{array}\right], \left[\begin{array}{r} -1 \\ 0 \end{array}\right]$$

For convenience, instead of performing **elementary operations** on a system of linear equations, perform corresponding elementary row operations on the corresponding **augmented matrix**.

Example

Type I: Interchange two rows 1 and 3.

$$\left[\begin{array}{rrrr|r} 2 & -1 & 0 & 5 & -3 \\ -2 & 0 & 3 & 3 & -1 \\ 0 & 5 & -6 & 1 & 0 \\ 1 & -4 & 2 & 2 & 2 \end{array}\right] \rightarrow^{r_1 \leftrightarrow r_3} \left[\begin{array}{rrrr|r} 0 & 5 & -6 & 1 & 0 \\ -2 & 0 & 3 & 3 & -1 \\ 2 & -1 & 0 & 5 & -3 \\ 1 & -4 & 2 & 2 & 2 \end{array}\right]$$

Type II: Multiply row 4 by 2.

$$\left[\begin{array}{rrrr|r} 2 & -1 & 0 & 5 & -3 \\ -2 & 0 & 3 & 3 & -1 \\ 0 & 5 & -6 & 1 & 0 \\ 1 & -4 & 2 & 2 & 2 \end{array}\right] \rightarrow^{2r_4} \left[\begin{array}{rrrr|r} 2 & -1 & 0 & 5 & -3 \\ -2 & 0 & 3 & 3 & -1 \\ 0 & 5 & -6 & 1 & 0 \\ 2 & -8 & 4 & 4 & 4 \end{array}\right]$$

Type III: Add 2 times row 4 to row 2.

$$\left[\begin{array}{rrrr|r} 2 & -1 & 0 & 5 & -3 \\ -2 & 0 & 3 & 3 & -1 \\ 0 & 5 & -6 & 1 & 0 \\ 1 & -4 & 2 & 2 & 2 \end{array}\right] \rightarrow^{2r_4 + r_2} \left[\begin{array}{rrrr|r} 2 & -1 & 0 & 5 & -3 \\ 0 & -8 & 7 & 7 & 3 \\ 0 & 5 & -6 & 1 & 0 \\ 1 & -4 & 2 & 2 & 2 \end{array}\right]$$

Definition

Two matrices A and B are **row equivalent** (or simply equivalent) if one can be obtained from the other by a sequence of elementary row operations.

Problem

Solve the system using an augmented matrix.

$$\begin{array}{rcrcl} x & + & 2y & = & 1 \\ 3x & + & 4y & = & 5 \end{array}$$

Solution

Discrete Mathematics II | Spring 2022

Section 1.2: Gaussian Elimination

Lecturer: Mr. Hall | *Math 270B*

Row-Echelon Matrix

- All rows consisting entirely of zeros are at the bottom.
- The first nonzero entry in each nonzero row is a 1 (called the **leading 1** for that row).
- Each leading 1 is to the right of all leading 1's in rows above it.

Example

$$\begin{bmatrix} 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where $*$ can be any number.

A matrix is said to be in *row-echelon form* (REF) if it is a row-echelon matrix.

Reduced Row-Echelon Matrix

- Row-echelon matrix.
- Each leading 1 is the only nonzero entry in its column.

Example

$$\begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where $*$ can be any number.

A matrix is said to be in *reduced row-echelon form* (RREF) if it is a reduced row-echelon matrix.

Example

Which of the following matrices are in REF?

Which matrices are in RREF?

(a) $\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix}$, (e) $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$, (f) $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Example

Suppose that the following matrix is the augmented matrix of a system of linear equations. We see from this matrix that the system of linear equations has four equations and seven variables.

$$\left[\begin{array}{rrrrrrr|r} 1 & -3 & 4 & -2 & 5 & -7 & 0 & 4 \\ 0 & 0 & 1 & 8 & 0 & 3 & -7 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

Note that the matrix is a row-echelon matrix.

- Each column of the matrix corresponds to a variable, and the leading variables are the variables that correspond to columns containing leading ones (in this case, columns 1, 3, 4, and 7).
- The remaining variables (corresponding to columns 2, 5 and 6) are called non-leading variables.

We will use elementary row operations to transform a matrix to row-echelon form (REF) or reduced row-echelon form (RREF).

Solving Systems of Linear Equations

Gaussian Elimination

To solve a system of linear equations, proceed as follows:

1. Carry the augmented matrix to a reduced row-echelon matrix using elementary row operations.
2. If a row of the form $[0\ 0\ \cdots 0 \mid 1]$ occurs, the system is inconsistent.
3. Otherwise assign the nonleading variables (if any) parameters and use the equations corresponding to the reduced row-echelon matrix to solve for the leading variables in terms of the parameters.

Problem

Solve the system

$$\begin{array}{rcrcrcr} 2x & + & y & + & 3z & = & 1 \\ x & + & 2y & - & z & = & 0 \\ x & - & 4y & + & 9z & = & 2 \end{array}$$

Solution

Given the reduced row-echelon matrix

$$\left[\begin{array}{ccr|r} 1 & 0 & \frac{7}{3} & \frac{2}{3} \\ 0 & 1 & -\frac{5}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

x and y are **leading variables**; z is a **non-leading variable** and so assign a parameter to z.

Thus the solution to the original system is given by

$$\left.\begin{array}{rcrcr} x & = & \frac{2}{3} & - & \frac{7}{3}t \\ y & = & -\frac{1}{3} & + & \frac{5}{3}t \\ z & = & t & & \end{array}\right\} \text{ where } t \in \mathbb{R}.$$

Practice

Solve the system

$$\begin{array}{rcrcrcr} x & + & y & + & 2z & = & -1 \\ 2x & + & y & + & 3z & = & 0 \\ & - & 2y & + & z & = & 2 \end{array}$$

Solution

Problem

Solve the system

$$\begin{array}{rcrcrcr} -3x_1 & - & 9x_2 & + & x_3 & = & -9 \\ 2x_1 & + & 6x_2 & - & x_3 & = & 6 \\ x_1 & + & 3x_2 & - & x_3 & = & 2 \end{array}$$

Solution

Problem

Find all values of a, b and c (or conditions on a, b and c) so that the system

$$\begin{array}{rcrcrcl} 2x & + & 3y & + & az & = & b \\ & - & y & + & 2z & = & c \\ x & + & 3y & - & 2z & = & 1 \end{array}$$

has (i) a unique solution, (ii) no solutions, and (iii) infinitely many solutions. In (i) and (iii), find the solution(s).

Solution

Definition

The **rank** of a matrix A, denoted rank A, is the number of leading 1's in any row-echelon matrix obtained from A by performing elementary row operations.

What does the rank of an augmented matrix tell us?

Suppose A is the augmented matrix of a **consistent** system of m linear equations in n variables, and rank $A = r$.

$$m\left\{\underbrace{\left[\begin{array}{cccc|c} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{array}\right]}_{n}\right. \rightarrow \underbrace{\left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]}_{r\ leading\ 1's}$$

Then the set of solutions to the system has $n - r$ parameters, so

- if $r < n$, there is at least one parameter, and the system has infinitely many solutions;
- if $r = n$, there are no parameters, and the system has a unique solution.

Example

Find the rank of $A = \left[\begin{array}{rrr} a & b & 5 \\ 1 & -2 & 1 \end{array}\right]$

Solution

Solutions to a System of Linear Equations

For **any** system of linear equations, exactly one of the following holds:

1. the system is **inconsistent**;
2. the system has a **unique** solution, i.e., exactly one solution;
3. the system has **infinitely many** solutions.

One can see what case applies by looking at the RREF matrix equivalent to the augmented matrix of the system and distinguishing three cases:

1. The last nonzero row ends with $\ldots 0\ 1]$: no solution.
2. The last nonzero row does not end with $\ldots 0\ 1]$ and all variables are leading: unique solution.
3. The last nonzero row does not end with $\ldots 0\ 1]$ and there are non-leading variables: infinitely many solutions.

Practice

Solve the system

$$\begin{array}{rcrcrcrcrcr} x_1 & - & 2x_2 & + & 2x_3 & + & 2x_4 & - & 5x_5 & = & 1 \\ -3x_1 & + & 6x_2 & - & 4x_3 & - & 9x_4 & + & 3x_5 & = & -1 \\ -x_1 & + & 2x_2 & - & 2x_3 & - & 4x_4 & - & 3x_5 & = & 3 \\ x_1 & - & 2x_2 & + & x_3 & + & 3x_4 & - & x_5 & = & 1 \end{array}$$

Solution

Solution (continued)

Discrete Mathematics II	Spring 2022

Section 1.3: Homogeneous Equations

Lecturer: Mr. Hall	*Math 270B*

Definition

A **homogeneous linear equation** is one whose constant term is equal to zero. A system of linear equations is called **homogeneous** if each equation in the system is homogeneous. A **homogeneous system** has the form

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{gathered}$$

where a_{ij} are scalars and x_i are variables, $1 \leq i \leq m$, $1 \leq j \leq n$.

Notice that $x_1 = 0, x_2 = 0, \cdots, x_n = 0$ is always a solution to a homogeneous system of equations. We call this the **trivial solution**.

We are interested in finding, if possible, **nontrivial solutions** (solutions with at least one variable not equal to zero) to homogeneous systems.

Example

Solve the system $\begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & x_3 & + & 3x_4 & = & 0 \\ -x_1 & + & 4x_2 & + & 5x_3 & - & 2x_4 & = & 0 \\ x_1 & + & 6x_2 & + & 3x_3 & + & 4x_4 & = & 0 \end{array}$

Solution

Definition

If $X_1, X_2, \ldots, X_p$ are columns with the same number of entries, and if $a_1, a_2, \ldots a_p \in \mathbb{R}$ (are scalars) then $a_1X_1 + a_2X_2 + \cdots + a_pX_p$ is a **linear combination** of columns $X_1, X_2, \ldots, X_p$.

In the previous example,

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{9}{5}s - \frac{14}{5}t \\ -\frac{4}{5}s - \frac{1}{5}t \\ s \\ t \end{bmatrix} &= \begin{bmatrix} \frac{9}{5}s \\ -\frac{4}{5}s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{14}{5}t \\ -\frac{1}{5}t \\ 0 \\ t \end{bmatrix} \\
&= s\begin{bmatrix} \frac{9}{5} \\ -\frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{14}{5} \\ -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix}
\end{aligned}$$

This gives us

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s\begin{bmatrix} \frac{9}{5} \\ -\frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{14}{5} \\ -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix} = sX_1 + tX_2,$$

where $X_1 = \begin{bmatrix} \frac{9}{5} \\ -\frac{4}{5} \\ 1 \\ 0 \end{bmatrix}$ and $X_2 = \begin{bmatrix} -\frac{14}{5} \\ -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix}$.

The columns X_1 and X_2 are called **basic solutions** to the original homogeneous system.

Notice that

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s\begin{bmatrix} \frac{9}{5} \\ -\frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{14}{5} \\ -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix} &= \frac{s}{5}\begin{bmatrix} 9 \\ -4 \\ 5 \\ 0 \end{bmatrix} + \frac{t}{5}\begin{bmatrix} -14 \\ -1 \\ 0 \\ 5 \end{bmatrix} \\
&= r\begin{bmatrix} 9 \\ -4 \\ 5 \\ 0 \end{bmatrix} + q\begin{bmatrix} -14 \\ -1 \\ 0 \\ 5 \end{bmatrix}
\end{aligned}$$

where $r, q \in \mathbb{R}$.

The columns $X_1 = \begin{bmatrix} 9 \\ -4 \\ 5 \\ 0 \end{bmatrix}$ and $X_2 = \begin{bmatrix} -14 \\ -1 \\ 0 \\ 5 \end{bmatrix}$ are also basic solutions to the original homogeneous system.

In general, any nonzero multiple of a basic solution (to a homogeneous system of linear equations) is also a basic solution.

What does the rank tell us in the homogeneous case?

Suppose A is the augmented matrix of a **homogeneous** system of m linear equations in n variables, and rank $A = r$.

$$m\left\{\underbrace{\left[\begin{array}{cccc|c} * & * & * & * & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \end{array}\right]}_{n}\right. \rightarrow \underbrace{\left[\begin{array}{cccc|c} 1 & * & * & * & 0 \\ 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]}_{r\ leading\ 1's}$$

There is always a solution, and the set of solutions to the system has $n - r$ parameters, so

- if $r < n$, there is at least one parameter, and the system has infinitely many solutions;
- if $r = n$, there are no parameters, and the system has a unique solution, the trivial solution.

Let A be an $m \times n$ matrix of rank r, and consider the homogeneous system in n variables with A as coefficient matrix. Then:

1. The system has exactly $n - r$ basic solutions, one for each parameter.
2. Every solution is a **linear combination** of these **basic solutions**.

Example

Find all values of a for which the system

$$\begin{array}{ccccccc} x & + & y & & & = & 0 \\ & & ay & + & z & = & 0 \\ x & + & y & + & az & = & 0 \end{array}$$

has nontrivial solutions, and determine the solutions.

Solution

Example

Let $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$. Determine whether $\mathbf{v} = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}$ can be written as a linear combination of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$.

Solution

Discrete Mathematics II — Spring 2022

Section 1.4: An Application to Network Flow

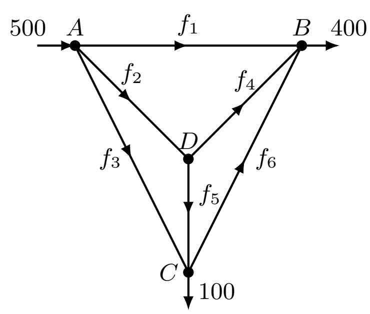
Lecturer: Mr. Hall — *Math 270B*

There are many types of problems that concern a network of conductors along which some sort of flow is observed. Examples of these include an irrigation network and a network of streets or freeways. There are often points in the system at which a net flow either enters or leaves the system. The basic principle behind the analysis of such systems is that the total flow into the system must equal the total flow out. In fact, we apply this principle at every junction in the system.

Junction Rule: At each of the junctions in the network, the total flow into that junction must equal the total flow out.

Example

A network of one-way streets is shown in the accompanying diagram. The rate of flow of cars into intersection A is 500 cars per hour, and 400 and 100 cars per hour emerge from B and C, respectively. Find the possible flows along each street.



Solution

Suppose the flows along the streets are f_1, f_2, f_3, f_4, f_5, and f_6 cars per hour in the directions shown.

Then, equating the flow in with the flow out at each intersection, we get

$$\begin{array}{ll} \text{Intersection } A & 500 = f_1 + f_2 + f_3 \\ \text{Intersection } B & f_1 + f_4 + f_6 = 400 \\ \text{Intersection } C & f_3 + f_5 = f_6 + 100 \\ \text{Intersection } D & f_2 = f_4 + f_5 \end{array}$$

These give four equations in the six variables $f_1, f_2, \ldots, f_6$.

$$\begin{aligned} f_1 + f_2 + f_3 \qquad\qquad\qquad\quad &= 500 \\ f_1 \qquad\qquad + f_4 \qquad + f_6 &= 400 \\ f_3 \qquad + f_5 - f_6 &= 100 \\ f_2 \qquad - f_4 - f_5 \qquad &= 0 \end{aligned}$$

Solution (continued)

The reduction of the augmented matrix is

$$\left[\begin{array}{cccccc|c} 1 & 1 & 1 & 0 & 0 & 0 & 500 \\ 1 & 0 & 0 & 1 & 0 & 1 & 400 \\ 0 & 0 & 1 & 0 & 1 & -1 & 100 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 1 & 400 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 100 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence, when we use f_4, f_5, and f_6 as parameters, the general solution is

$$f_1 = 400 - f_4 - f_6 \qquad f_2 = f_4 + f_5 \qquad f_3 = 100 - f_5 + f_6$$

This gives all solutions to the system of equations and hence all the possible flows.

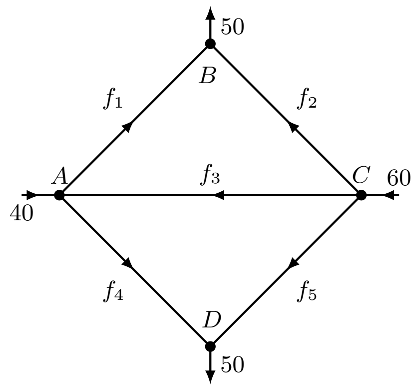
Of course, not all these solutions may be acceptable in the real situation. For example, the flows $f_1, f_2, \dots, f_6$ are all *positive* in the present context (if one came out negative, it would mean traffic flowed in the opposite direction). This imposes constraints on the flows: $f_1 \geq 0$ and $f_3 \geq 0$ become

$$f_4 + f_6 \leq 400 \qquad f_5 - f_6 \leq 100$$

Further constraints might be imposed by insisting on maximum values on the flow in each street.

Example

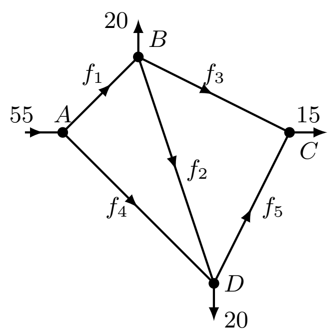
Find the possible flows in the network.



Solution

Practice

a) Find the possible flows in the network.
b) If BC is closed, what range of flow on AD must be maintained so that no edge carries a flow of more than 30?



Solution

Discrete Mathematics II Spring 2022

Section 1.5: An Application to Electrical Networks

Lecturer: Mr. Hall *Math 270B*

In an electrical network it is often necessary to find the current in amperes (A) flowing in various parts of the network. These networks usually contain resistors that retard the current. The resistors are indicated by a symbol ( ), and the resistance is measured in ohms (Ω). Also, the current is increased at various points by voltage sources (for example, a battery). The voltage of these sources is measured in volts (V), and they are represented by the symbol ( ). We assume these voltage sources have no resistance. The flow of current is governed by the following principles.

Ohm's Law

The current I and the voltage drop V across a resistance R are related by the equation $V = RI$.

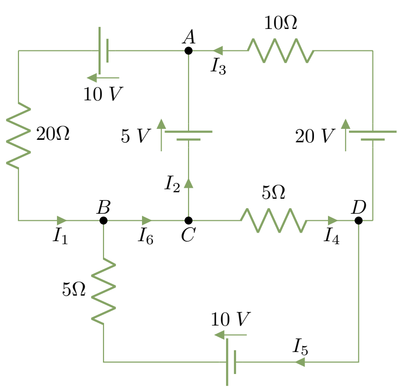
Kirchhoff's Laws

1. (Junction Rule) The current flow into a junction equals the current flow out of that junction.
2. (Circuit Rule) The algebraic sum of the voltage drops (due to resistances) around any closed circuit of the network must equal the sum of the voltage increases around the circuit.

When applying rule 2, select a direction (clockwise or counterclockwise) around the closed circuit and then consider all voltages and currents positive when in this direction and negative when in the opposite direction. This is why the term *algebraic sum* is used in rule 2. Here is an example.

Example

Find the various currents in the circuit shown.



First apply the junction rule at junctions A, B, C, and D to obtain

$$\begin{array}{ll} \text{Junction } A & I_1 = I_2 + I_3 \\ \text{Junction } B & I_6 = I_1 + I_5 \\ \text{Junction } C & I_2 + I_4 = I_6 \\ \text{Junction } D & I_3 + I_5 = I_4 \end{array}$$

Note that these equations are not independent (in fact, the third is an easy consequence of the other three).

Next, the circuit rule insists that the sum of the voltage increases (due to the sources) around a closed circuit must equal the sum of the voltage drops (due to resistances). By Ohm's law, the voltage loss across a resistance R (in the direction of the current I) is RI. Going counterclockwise around three closed circuits yields

$$\begin{array}{lr} \text{Upper left} & 10 + \quad 5 = 20I_1 \\ \text{Upper right} & -5 + \quad 20 = 10I_3 + 5I_4 \\ \text{Lower} & -10 = -5I_5 - 5I_4 \end{array}$$

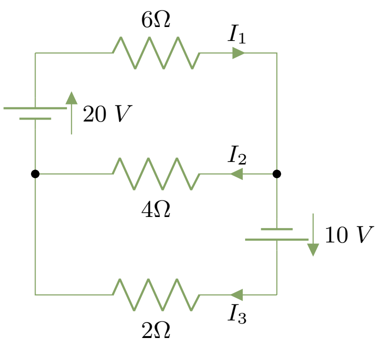
Hence, disregarding the redundant equation obtained at junction C, we have six equations in the six unknowns $I_1, \ldots, I_6$. The solution is

$$\begin{array}{ll} I_1 = \frac{15}{20} & I_4 = \frac{28}{20} \\ I_2 = \frac{-1}{20} & I_5 = \frac{12}{20} \\ I_3 = \frac{16}{20} & I_6 = \frac{27}{20} \end{array}$$

The fact that I_2 is negative means, of course, that this current is in the opposite direction, with a magnitude of $\frac{1}{20}$ amperes.

Example

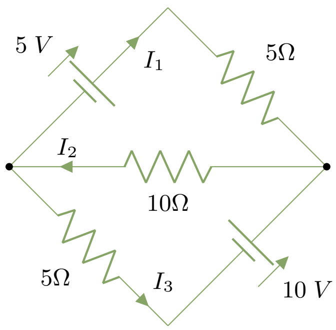
Find the currents in the circuit.



Solution

Practice

Find the currents in the circuit.



Solution

Discrete Mathematics II — Spring 2022

Section 2.1: Matrix Addition, Scalar Multiplication, and Transposition

Lecturer: Mr. Hall — *Math 270B*

Definitions

Let m and n be positive integers.

- An $m \times n$ **matrix** is a rectangular array of numbers having m rows and n columns. Such a matrix is said to have **size** $m \times n$.
- A **row matrix (or row vector)** is a $1 \times n$ matrix, and a **column matrix (or column vector)** is an $m \times 1$ matrix.
- A **square matrix** is an $n \times n$ matrix.
- The (i, j)**-entry of a matrix** is the entry in row i and column j. For a matrix A, the (i, j)-entry of A is often written as a_{ij}.

General notation for an $m \times n$ matrix, A:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]$$

Matrices – Properties and Operations

1. **Equality:** two matrices are equal if and only if they have the same size and the corresponding entries are equal.
2. **Zero Matrix:** an $m \times n$ matrix with all entries equal to zero.
3. **Addition:** matrices must have the same size; add corresponding entries.
4. **Scalar Multiplication:** multiply each entry of the matrix by the scalar.
5. **Negative of a Matrix:** for an $m \times n$ matrix A, its negative is denoted $-A$ and $-A = (-1)A$.
6. **Subtraction:** for $m \times n$ matrices A and B, $A - B = A + (-1)B$.

Matrix Addition

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices. Then $A + B = C$ where C is the $m \times n$ matrix $C = [c_{ij}]$ defined by

$$c_{ij} = a_{ij} + b_{ij}$$

Example

Add the matrices $A = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -2 \\ 6 & 1 \end{bmatrix}$.

Solution

Properties of Matrix Addition

Let A, B and C be $m \times n$ matrices. Then the following properties hold.

1. $A + B = B + A$ (matrix addition is commutative).

2. $(A + B) + C = A + (B + C)$ (matrix addition is associative).

3. There exists an $m \times n$ zero matrix, 0, such that $A + 0 = A$.
   (existence of an additive identity).

4. There exists an $m \times n$ matrix $-A$ such that $A + (-A) = 0$.
   (existence of an additive inverse).

Scalar Multiplication

Let $A = [a_{ij}]$ be an $m \times n$ matrix and let k be a scalar. Then $kA = [ka_{ij}]$.

Example

Let $A = \begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & -2 \\ 0 & 4 & 5 \end{bmatrix}$. Find $3A$.

Solution

Properties of Scalar Multiplication

Let A and B be $m \times n$ matrices and let $k, p \in \mathbb{R}$ (scalars). Then the following properties hold.

1. $k(A + B) = kA + kB$.
   (scalar multiplication distributes over matrix addition).

2. $(k + p)A = kA + pA$.
   (addition distributes over scalar multiplication).

3. $k(pA) = (kp)A$. (scalar multiplication is associative).

4. $1A = A$. (existence of a multiplicative identity).

Example

Find the sum:

$$2\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} + 4\begin{bmatrix} -2 & 1 \\ 3 & 0 \end{bmatrix} - \begin{bmatrix} 6 & 8 \\ 1 & -1 \end{bmatrix}$$

Solution

Example

Let A and B be $m \times n$ matrices. Simplify the expression

$$2[9(A-B)+7(2B-A)]-2[3(2B+A)-2(A+3B)-5(A+B)]$$

Solution

The Transpose of a Matrix

If A is an $m \times n$ matrix, then its **transpose**, denoted A^T, is the $n \times m$ matrix whose i^{th} row is the j^{th} column of A, $1 \leq j \leq n$; i.e., if $A = [a_{ij}]$, then

$$A^T = [a_{ij}]^T = [a_{ji}]$$

i.e., the (i, j)-entry of A^T is the (j, i)-entry of A.

Properties of the Transpose of a Matrix

Let A and B be $m \times n$ matrices, C be an $n \times p$ matrix, and $r \in \mathbb{R}$ a scalar. Then

1. $(A^T)^T = A$
2. $(rA)^T = rA^T$
3. $(A + B)^T = A^T + B^T$
4. $(AC)^T = C^T A^T$

Example

Find the matrix A if $\left(A + 3\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix}\right)^T = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix}$.

Solution

Symmetric Matrices

Let $A = [a_{ij}]$ be an $m \times n$ matrix. The entries $a_{11}, a_{22}, a_{33}, \ldots$ are called the **main diagonal** of A.

The matrix A is called **symmetric** if and only if $A^T = A$. Note that this immediately implies that A is a **square** matrix.

Examples

$$\begin{bmatrix} 2 & -3 \\ -3 & 17 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 5 \\ 0 & 2 & 11 \\ 5 & 11 & -3 \end{bmatrix}, \begin{bmatrix} 0 & 2 & 5 & -1 \\ 2 & 1 & -3 & 0 \\ 5 & -3 & 2 & -7 \\ -1 & 0 & -7 & 4 \end{bmatrix}$$

are symmetric matrices, and each is symmetric about its main diagonal.

Example

Show that if A and B are symmetric matrices, then $A^T + 2B$ is symmetric.

Proof

Skew-Symmetric Matrices

An $n \times n$ matrix A is said to be **skew-symmetric** if $A^T = -A$.

Example

$$\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 9 & 4 \\ -9 & 0 & -3 \\ -4 & 3 & 0 \end{bmatrix}$$

are skew-symmetric matrices.

Example

Show that if A is a square matrix, then $A - A^T$ is skew-symmetric.

Proof

Discrete Mathematics II	Spring 2022

Section 2.2: Equations, Matrices, and Transformations

Lecturer: Mr. Hall | *Math 270B*

Definitions

A row matrix or column matrix is often called a **vector**, and such matrices are referred to as **row vectors** and **column vectors**, respectively. If $\mathbf{x}$ is a row vector of size $1 \times n$, and $\mathbf{y}$ is a column vector of size $m \times 1$, then we write

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Vector form of a system of linear equations

Consider the system of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Such a system can be expressed in **vector form** or as a **vector equation** by using linear combinations of column vectors:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Example

Express the following system of linear equations in vector form.

$$\begin{array}{ccccccc} 2x_1 & + & 4x_2 & - & 3x_3 & = & -6 \\ & - & x_2 & + & 5x_3 & = & 0 \\ x_1 & + & x_2 & + & 4x_3 & = & 1 \end{array}$$

Solution

Matrix Vector Multiplication

Let $A = [a_{ij}]$ be an $m \times n$ matrix with columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, written $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$, and let $\mathbf{x}$ be an $n \times 1$ column vector,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then the product of the matrix A and (column) vector $\mathbf{x}$ is the $m \times 1$ column vector given by

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \sum_{j=1}^{n} x_j\mathbf{a}_j$$

that is, $A\mathbf{x}$ is a linear combination of the columns of A.

Example

Compute the product $A\mathbf{x}$ for

$$A = \begin{bmatrix} 1 & 4 \\ 5 & 0 \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Solution

Example

Compute $A\mathbf{x}$ for

$$A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & -1 & 0 & 1 \\ 3 & 1 & 3 & 1 \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 4 \end{bmatrix}$$

Solution

Matrix form of a system of linear equations

Consider the system of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Such a system can be expressed in **matrix form** using matrix vector multiplication,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Thus a system of linear equations can be expressed as a **matrix equation** $A\mathbf{x} = \mathbf{b}$, where A is the **coefficient matrix**, $\mathbf{b}$ is the **constant matrix**, and $\mathbf{x}$ is the **matrix of variables.**

Example

Express the following system of linear equations in matrix form.

$$\begin{array}{ccccccc} 2x_1 & + & 4x_2 & - & 3x_3 & = & -6 \\ & - & x_2 & + & 5x_3 & = & 0 \\ x_1 & + & x_2 & + & 4x_3 & = & 1 \end{array}$$

Solution

Matrix and Vector Equations

1. Every system of m linear equations in n variables can be written in the form $A\mathbf{x} = \mathbf{b}$ where A is the coefficient matrix, $\mathbf{x}$ is the matrix of variables, and $\mathbf{b}$ is the constant matrix.

2. The system $A\mathbf{x} = \mathbf{b}$ is consistent (i.e., has at least one solution) if and only if $\mathbf{b}$ is a linear combination of the columns of A.

3. The vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a solution to the system $A\mathbf{x} = \mathbf{b}$ if and only if $x_1, x_2, \ldots, x_n$ is a solution to the vector equation
$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots x_n\mathbf{a}_n = \mathbf{b}$$
where $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are the columns of A.

Example

Let
$$A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & -1 & 0 & 1 \\ 3 & 1 & 3 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Express $\mathbf{b}$ as a linear combination of the columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ of A, or show that this is impossible.

Solution

Solve the system $A\mathbf{x} = \mathbf{b}$ where $\mathbf{x}$ is a column vector with four entries.

Do so by putting the **augmented matrix** $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ in reduced row-echelon form.

$$\left[\begin{array}{cccc|c} 1 & 0 & 2 & -1 & 1 \\ 2 & -1 & 0 & 1 & 1 \\ 3 & 1 & 3 & 1 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & \frac{1}{7} \\ 0 & 1 & 0 & 1 & -\frac{5}{7} \\ 0 & 0 & 1 & -1 & \frac{3}{7} \end{array}\right]$$

Since there are infinitely many solutions (x_4 is assigned a parameter), choose any value for x_4.

Choosing $x_4 = 0$ (which is the simplest thing to do) gives us
$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{5}{7}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \frac{3}{7}\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} = \frac{1}{7}\mathbf{a}_1 - \frac{5}{7}\mathbf{a}_2 + \frac{3}{7}\mathbf{a}_3 + 0\mathbf{a}_4.$$

Algebraic Properties of Matrix-Vector Multiplication

Let A and B be $m \times n$ matrices, and let $\mathbf{x}$ and $\mathbf{y}$ be n-vectors in $\mathbb{R}^n$. Then:

1. $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}$.
2. $A(a\mathbf{x}) = a(A\mathbf{x}) = (aA)\mathbf{x}$ for all scalars a.
3. $(A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$.

This provides a useful way to describe the solutions to a system $A\mathbf{x} = \mathbf{b}$.

There is a related system

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

called the **associated homogeneous system**, obtained from the original system $A\mathbf{x} = \mathbf{b}$ by replacing all the constants by zeros. Suppose $\mathbf{x}_1$ is a solution to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x}_0$ is a solution to $A\mathbf{x} = \mathbf{0}$ (that is $A\mathbf{x}_1 = \mathbf{b}$ and $A\mathbf{x}_0 = \mathbf{0}$). Then $\mathbf{x}_1 + \mathbf{x}_0$ is another solution to $A\mathbf{x} = \mathbf{b}$. Indeed,

$$A(\mathbf{x}_1 + \mathbf{x}_0) = A\mathbf{x}_1 + A\mathbf{x}_0 = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

This observation has a useful converse.

Suppose $\mathbf{x}_1$ is any particular solution to the system $A\mathbf{x} = \mathbf{b}$ of linear equations. Then every solution $\mathbf{x}_2$ to $A\mathbf{x} = \mathbf{b}$ has the form

$$\mathbf{x}_2 = \mathbf{x}_0 + \mathbf{x}_1$$

for some solution $\mathbf{x}_0$ of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

Suppose $\mathbf{x}_2$ is also a solution to $A\mathbf{x} = \mathbf{b}$, so that $A\mathbf{x}_2 = \mathbf{b}$. Write $\mathbf{x}_0 = \mathbf{x}_2 - \mathbf{x}_1$. Then $\mathbf{x}_2 = \mathbf{x}_0 + \mathbf{x}_1$ and we compute

$$A\mathbf{x}_0 = A(\mathbf{x}_2 - \mathbf{x}_1) = A\mathbf{x}_2 - A\mathbf{x}_1 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Hence $\mathbf{x}_0$ is a solution to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

Example

Express every solution of the system as a sum of a specific solution plus a solution of the associated homogeneous system.

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 2 \\ 2x_1 & + & x_2 & & & = & 3 \\ x_1 & - & x_2 & - & 3x_3 & = & 0 \end{array}$$

Solution

The Dot Product

If $(a_1, a_2, \dots, a_n)$ and $(b_1, b_2, \dots, b_n)$ are two ordered n-tuples, their **dot product** is defined to be the number

$$a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

obtained by multiplying corresponding entries and adding the results.

This is very much related to the matrix produxt $A\mathbf{x}$.

Let A be an $m \times n$ matrix and let $\mathbf{x}$ be an n-vector. Then each entry of the vector $A\mathbf{x}$ is the dot product of the corresponding row of A with $\mathbf{x}$.

Example

If $A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & -1 & 0 & 1 \\ 3 & 1 & 3 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 4 \end{bmatrix}$, compute $A\mathbf{x}$.

Solution

The Identity Matrix

For each $n \geq 2$, the **identity matrix** I_n is the $n \times n$ matrix with 1's on the main diagonal (upper left to lower right), and zeros elsewhere.

The first few identity matrices are

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \ldots$$

For each $n \geq 2$ we have $I_n\mathbf{x} = \mathbf{x}$ for each n-vector $\mathbf{x}$ in $\mathbb{R}^n$.

We verify the case $n = 4$. Given the 4-vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ the dot product rule gives

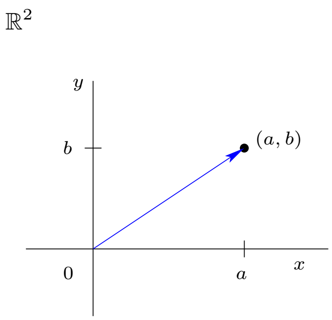
$$I_4\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 0 + 0 + 0 \\ 0 + x_2 + 0 + 0 \\ 0 + 0 + x_3 + 0 \\ 0 + 0 + 0 + x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{x}.$$

In general, $I_n\mathbf{x} = \mathbf{x}$ because entry k of $I_n\mathbf{x}$ is the dot product of row k of I_n with $\mathbf{x}$, and row k of I_n has 1 in position k and zeros elsewhere.

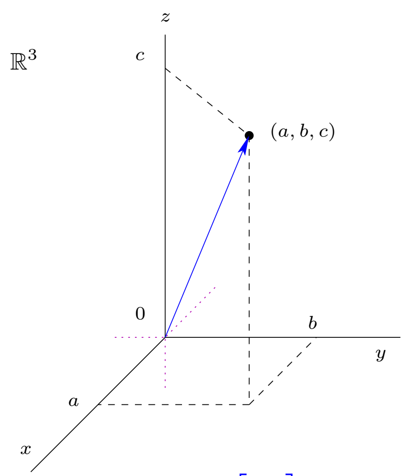
Transformations

- We have already used $\mathbb{R}$ to denote the set of real numbers.
- We use $\mathbb{R}^2$ to the denote the set of all column vectors of length two, and we use $\mathbb{R}^3$ to the denote the set of all column vectors of length three (the length of a vector is the number of entries it contains).
- In general, we write $\mathbb{R}^n$ for the set of all column vectors of length n.

Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ have convenient geometric interpretations as **position vectors** of points in the 2-dimensional (Cartesian) plane and in 3-dimensional space, respectively.



The vector $\begin{bmatrix} a \\ b \end{bmatrix}$.



The vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

A **transformation** is a function $T : \mathbb{R}^n \to \mathbb{R}^m$, sometimes written

$$\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m,$$

and is called a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

If $m = n$, then we say T is a transformation of $\mathbb{R}^n$.

What do we mean by a function?

Informally, a function $T : \mathbb{R}^n \to \mathbb{R}^m$ is a rule that assigns exactly one vector of $\mathbb{R}^m$ to each vector of $\mathbb{R}^n$.
We use the notation $T(\mathbf{x})$ to mean the transformation T applied to the vector $\mathbf{x}$.

If T acts by matrix multiplication of a matrix A, we call T a **matrix transformation**, and write $T_A(\mathbf{x}) = A\mathbf{x}$.

Example

$T : \mathbb{R}^3 \to \mathbb{R}^4$ defined by

$$T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b \\ b+c \\ a-c \\ c-b \end{bmatrix}$$

is a transformation

that transforms the vector $\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}$ in $\mathbb{R}^3$ into the vector

$$T\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 1+4 \\ 4+7 \\ 1-7 \\ 7-4 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ -6 \\ 3 \end{bmatrix}.$$

Example

Consider the matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix}$. By matrix multiplication, A transforms vectors in $\mathbb{R}^3$ into vectors in $\mathbb{R}^2$.

Consider the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Transforming this vector by A looks like:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y \\ 2x+y \end{bmatrix}$$

For example:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Rotations in $\mathbb{R}^2$

Let A be an $m \times n$ matrix. The transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T(\mathbf{x}) = A\mathbf{x} \text{ for each } \mathbf{x} \in \mathbb{R}^n$$

is called the **matrix transformation induced by** A.

The transformation

$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$

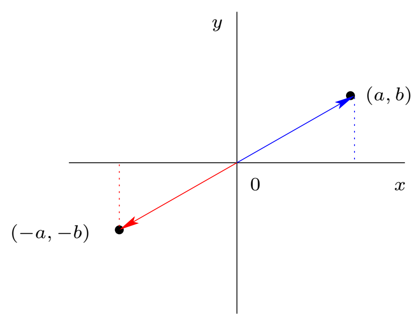
denotes counterclockwise rotation about the origin through an angle of θ.

Example

We denote by

$$R_\pi : \mathbb{R}^2 \to \mathbb{R}^2$$

counterclockwise rotation about the origin through an angle of π.



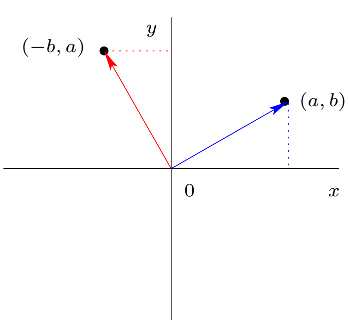
We see that $R_\pi \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a \\ -b \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$, so R_π is a matrix transformation.

Example

We denote by

$$R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$$

counterclockwise rotation about the origin through an angle of $\pi/2$.



We see that $R_{\pi/2} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$, so $R_{\pi/2}$ is a matrix transformation.

Discrete Mathematics II	Spring 2022

Section 2.3: Matrix Multiplication

Lecturer: Mr. Hall	*Math 270B*

Definition

Let A be an $m \times n$ matrix and let $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix}$ be an $n \times p$ matrix, whose columns are $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$. The **product** of A and B is the matrix

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

i.e., the first column of AB is $A\mathbf{b}_1$, the second column of AB is $A\mathbf{b}_2$, etc. Note that AB has size $m \times p$.

Example

Find the product AB, where

$$A = \begin{bmatrix} -1 & 0 & 3 \\ 2 & -1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 1 & 2 \\ 0 & -2 & 4 \\ 1 & 0 & 0 \end{bmatrix}$$

Solution

Compatibility for Matrix Multiplication

Let A and B be matrices, and suppose that A is $m \times n$.

- In order for the product AB to exist, the number of rows in B must be equal to the number of columns in A, implying that B is an $n \times p$ matrix for some p.
- When defined, AB is an $m \times p$ matrix.

If the product is defined, then A and B are said to be **compatible** for (matrix) multiplication.

As we saw in the previous problem

$$\overset{2\times 3}{\begin{bmatrix} -1 & 0 & 3 \\ 2 & -1 & 1 \end{bmatrix}} \overset{3\times 3}{\begin{bmatrix} -1 & 1 & 2 \\ 0 & -2 & 4 \\ 1 & 0 & 0 \end{bmatrix}} = \overset{2\times 3}{\begin{bmatrix} 4 & -1 & -2 \\ -1 & 4 & 0 \end{bmatrix}}$$

Note that the product

$$\overset{3\times 3}{\begin{bmatrix} -1 & 1 & 2 \\ 0 & -2 & 4 \\ 1 & 0 & 0 \end{bmatrix}} \overset{2\times 3}{\begin{bmatrix} -1 & 0 & 3 \\ 2 & -1 & 1 \end{bmatrix}}$$

does not exist.

The (i,j)-entry of a Matrix Product

Let $A = [a_{ij}]$ be an $m \times n$ matrix and $B = [b_{ij}]$ be an $n \times p$ matrix. Then the (i, j)-entry of AB is given by the dot product of row i of A and column j of B:

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Using the above definition, the $(2, 3)$-entry of the product

$$\begin{bmatrix} -1 & 0 & 3 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 \\ 0 & -2 & 4 \\ 1 & 0 & 0 \end{bmatrix}$$

is computed by the dot product of the second row of the first matrix and the third column of the second matrix:

$$2 \times 2 + (-1) \times 4 + 1 \times 0 = 4 - 4 + 0 = 0.$$

Example

Let

$$A = \begin{bmatrix} 1 & 2 \\ -3 & 0 \\ 1 & -4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & -1 & 2 & 0 \\ 3 & -2 & 1 & -3 \end{bmatrix}$$

- Does AB exist? If so, compute it.
- Does BA exist? If so, compute it.

Solution

Example

Let

$$G = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } H = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

- Does GH exist? If so, compute it.
- Does HG exist? If so, compute it.

Solution

Example

Let

$$P = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} -1 & 1 \\ 0 & 3 \end{bmatrix}$$

- Does PQ exist? If so, compute it.
- Does QP exist? If so, compute it.

Solution

Fact

The three preceding problems illustrate an important property of matrix multiplication.

> In general, matrix multiplication is **not** commutative, i.e., the order of the matrices in the product is important.

In other words, in general $AB \neq BA$.

Properties of Matrix Multiplication

Let A, B, and C be matrices of the appropriate sizes, and let $r \in \mathbb{R}$ be a scalar. Then the following properties hold.

1. $A(B + C) = AB + AC$.
   (matrix multiplication distributes over matrix addition).
2. $(B + C)A = BA + CA$.
   (matrix multiplication distributes over matrix addition).
3. $A\,(BC) = (AB)\,C$. (matrix multiplication is associative).
4. $r(AB) = (rA)B = A(rB)$.

Elementary Proofs

Example

Let A and B be $m \times n$ matrices, and let C be an $n \times p$ matrix. Show that if A and B commute with C, i.e., $AC = CA$ and $BC = CB$, then $A + B$ commutes with C.

Proof

Example

Let A, B and C be $n \times n$ matrices, and suppose that both A and B commute with C, i.e., $AC = CA$ and $BC = CB$. Show that AB commutes with C.

Proof

Section 2.4: Matrix Inverses

Definition

For each $n \geq 2$, the $n \times n$ **identity matrix**, denoted I_n, is the matrix having ones on its main diagonal and zeros elsewhere, and is defined for all $n \geq 2$.

Example

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition

Let $n \geq 2$. For each j, $1 \leq j \leq n$, we denote by $\mathbf{e}_j$ the j^{th} column of I_n.

Example

When $n = 3$, $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Theorem

Let A be an $m \times n$ matrix Then $AI_n = A$ and $I_m A = A$.

Proof

The (i, j)-entry of AI_n is the product of the i^{th} row of $A = [a_{ij}]$, namely $\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \end{bmatrix}$ with the j^{th} column of I_n, namely $\mathbf{e}_j$. Since $\mathbf{e}_j$ has a one in row j and zeros elsewhere,

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \end{bmatrix} \mathbf{e}_j = a_{ij}$$

Since this is true for all $i \leq m$ and all $j \leq n$, $AI_n = A$.

Instead of AI_n and I_mA we often write AI and IA, respectively, since the size of the identity matrix is clear from the context: the sizes of A and I must be compatible for matrix multiplication.

Thus

$$AI = A \text{ and } IA = A$$

which is why I is called an **identity** matrix – it is an identity for matrix multiplication.

Definition

Let A be an $n \times n$ matrix. Then B is an **inverse** of A if and only if $AB = I_n$ and $BA = I_n$.

Note that since A and I_n are both $n \times n$, B must also be an $n \times n$ matrix.

Example

Verify that B is an inverse of A.

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

Solution

Does the matrix

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

have an inverse?

No! To see this, suppose

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is an inverse of A.

Then

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ c & d \end{bmatrix}$$

which is never equal to I_2.

The Uniqueness of an Inverse

If A is a square matrix and B and C are inverses of A, then $B = C$.

Proof

Since B and C are inverses of A, $AB = I = BA$ and $AC = I = CA$. Then

$$C = CI = C(AB) = CAB$$

and

$$B = IB = (CA)B = CAB$$

so $B = C$.

Example

For $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$, we saw that

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The preceding theorem tells us that B is the inverse of A, rather than just an inverse of A.

Definition

Let A be a square matrix, i.e., an $n \times n$ matrix.

- The **inverse** of A, if it exists, is denoted A^{-1}, and

$$AA^{-1} = I = A^{-1}A$$

- If A has an inverse, then we say that A is **invertible**.

Finding the inverse of a 2×2 matrix

Suppose that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

If $ad - bc \neq 0$, then there is a formula for A^{-1}:

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

This can easily be verified by computing the products AA^{-1} and $A^{-1}A$.

$$\begin{aligned} AA^{-1} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \frac{1}{ad-bc} \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \right) \\ &= \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & -bc+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Problem

Suppose that A is any $n \times n$ matrix.

- How do we know whether or not A^{-1} exists?
- If A^{-1} exists, how do we find it?

Answer

The matrix inversion algorithm.

Although the formula for the inverse of a 2×2 matrix is quicker and easier to use than the matrix inversion algorithm, the general formula for the inverse of an $n \times n$ matrix, $n \geq 3$ (which we will see later), is more complicated and difficult to use than the matrix inversion algorithm. To find inverses of square matrices that are not 2×2, the matrix inversion algorithm is the most efficient method to use.

The Matrix Inversion Algorithm

Let A be an $n \times n$ matrix. To find A^{-1}, if it exists,

- take the $n \times 2n$ matrix
$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$
obtained by augmenting A with the $n \times n$ identity matrix, I_n.
- Perform elementary row operations to transform $\left[\begin{array}{c|c} A & I_n \end{array}\right]$ into a reduced row-echelon matrix.

Let A be an $n \times n$ matrix. Then the following conditions are equivalent.

1. A is invertible.
2. the reduced row-echelon form of A is I.
3. $\left[\begin{array}{c|c} A & I_n \end{array}\right]$ can be transformed into $\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$ using the Matrix Inversion Algorithm.

Example

Find, if possible, the inverse of $\begin{bmatrix} 1 & 0 & -1 \\ -2 & 1 & 3 \\ -1 & 1 & 2 \end{bmatrix}$.

Solution

Practice

Let $A = \begin{bmatrix} 3 & 1 & 2 \\ 1 & -1 & 3 \\ 1 & 2 & 4 \end{bmatrix}$. Find the inverse of A, if it exists.

Solution

Systems of Linear Equations and Inverses

Suppose that a system of n linear equations in n variables is written in matrix form as $A\mathbf{x} = \mathbf{b}$, and suppose that A is invertible.

The system of linear equations

$$\begin{array}{rcl} 2x - 7y & = & 3 \\ 5x - 18y & = & 8 \end{array}$$

can be written in matrix form as $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 2 & -7 \\ 5 & -18 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix}$$

You can check that $A^{-1} = \begin{bmatrix} 18 & -7 \\ 5 & -2 \end{bmatrix}$.

Since A^{-1} exists and has the property that $A^{-1}A = I$, we obtain the following.

$$\begin{array}{rcl} A\mathbf{x} & = & \mathbf{b} \\ A^{-1}(A\mathbf{x}) & = & A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{x} & = & A^{-1}\mathbf{b} \\ I\mathbf{x} & = & A^{-1}\mathbf{b} \\ \mathbf{x} & = & A^{-1}\mathbf{b} \end{array}$$

i.e., $A\mathbf{x} = \mathbf{b}$ has the unique solution given by $\mathbf{x} = A^{-1}\mathbf{b}$.

Therefore,

$$\mathbf{x} = A^{-1} \begin{bmatrix} 3 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 & -7 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 8 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

Inverses of Transposes and Products

Example

Suppose A is an invertible matrix. Then

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$$

and

$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$$

This means that $(A^T)^{-1} = (A^{-1})^T$.

Example

Suppose A and B are invertible $n \times n$ matrices. Then

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

This means that $(AB)^{-1} = B^{-1}A^{-1}$.

Inverses of Transposes and Products

1. If A is an invertible matrix, then $(A^T)^{-1} = (A^{-1})^T$.
2. If A and B are invertible matrices, then AB is invertible and
$$(AB)^{-1} = B^{-1}A^{-1}$$
3. If $A_1, A_2, \ldots, A_k$ are invertible, then $A_1A_2\cdots A_k$ is invertible and
$$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1}\cdots A_2^{-1}A_1^{-1}$$

Properties of Inverses

1. I is invertible, and $I^{-1} = I$.
2. If A is invertible, so is A^{-1}, and $(A^{-1})^{-1} = A$.
3. If A is invertible, so is A^k, and $(A^k)^{-1} = (A^{-1})^k$.
   (A^k means A multiplied by itself k times)
4. If A is invertible and $p \in \mathbb{R}$ is nonzero, then pA is invertible, and $(pA)^{-1} = \frac{1}{p}A^{-1}$.

Example

Given $(3I - A^T)^{-1} = 2\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$, find the matrix A.

Solution

A Fundamental Result

Let A be an $n \times n$ matrix, and let $\mathbf{x}$, $\mathbf{b}$ be $n \times 1$ vectors. The following conditions are equivalent.

1. A is invertible.
2. The rank of A is n.
3. The reduced row echelon form of A is I_n.
4. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$.
5. A can be transformed to I_n by elementary row operations.
6. The system $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$ for any choice of $\mathbf{b}$.
7. There exists an $n \times n$ matrix C with the property that $CA = I_n$.
8. There exists an $n \times n$ matrix C with the property that $AC = I_n$.

The following example illustrates why "an inverse" of a non-square matrix doesn't make sense. If A is $m \times n$ and B is $n \times m$, where $m \neq n$, then even if $AB = I$, it will never be the case that $BA = I$.

Example

Let $A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 4 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

and

$$BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 5 & 1 \end{bmatrix} \neq I_3$$

The Inverse of a Linear Transformation

Let $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be a transformation given by

$$T\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} x+y \\ y \end{array}\right]$$

Then T is a linear transformation induced by $A = \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]$.

Notice that the matrix A is invertible. Therefore the transformation T has an inverse, T^{-1}, induced by

$$A^{-1} = \left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \end{array}\right]$$

Consider the action of T and T^{-1}.

$$T\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} x+y \\ y \end{array}\right]$$

$$T^{-1}\left[\begin{array}{c} x+y \\ y \end{array}\right] = \left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} x+y \\ y \end{array}\right] = \left[\begin{array}{c} x \\ y \end{array}\right]$$

Therefore

$$T^{-1}\left(T\left[\begin{array}{c} x \\ y \end{array}\right]\right) = \left[\begin{array}{c} x \\ y \end{array}\right]$$

Discrete Mathematics II	Spring 2022

Section 2.5: Elementary Matrices

Lecturer: Mr. Hall	*Math 270B*

Definition

An **elementary matrix** is a matrix obtained from an identity matrix by performing *a single* elementary row operation.

The **type** of an elementary matrix is given by the type of row operation used to obtain the elementary matrix. Recall the elementary row operations.

- **Type I:** Interchange two rows.
- **Type II:** Multiply a row by a nonzero number.
- **Type III:** Add a (nonzero) multiple of one row to a different row.

Example

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

are examples of elementary matrices of types I, II and III, respectively.

Let

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

We are interested in the effect that (left) multiplication of A by E, F and G has on the matrix A.

Compute EA, FA, and GA.

Solution

Inverses of Elementary Matrices

Every elementary matrix E is invertible, and E^{-1} is also an elementary matrix (of the same type). Moreover, E^{-1} corresponds to the inverse of the row operation that produces E.

The following table gives the inverse of each type of elementary row operation:

Type	**Operation**	**Inverse Operation**
I	Interchange rows p and q	Interchange rows p and q
II	Multiply row p by $k \neq 0$	Multiply row p by $1/k$
III	Add k times row p to row $q \neq p$	Subtract k times row p from row q

Example

Without using the matrix inversion algorithm, find the inverse of the elementary matrix

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

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

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution

The Form $B = UA$

Suppose A is an $m \times n$ matrix and that B can be obtained from A by a sequence of k elementary row operations.

Then there exist elementary matrices $E_1, E_2, \ldots E_k$ such that

$$B = E_k(E_{k-1}(\cdots(E_2(E_1A))\cdots))$$

Since matrix multiplication is associative, we have

$$B = (E_kE_{k-1}\cdots E_2E_1)A$$

or, more concisely, $B = UA$ where $U = E_kE_{k-1}\cdots E_2E_1$.

To find U so that $B = UA$, we could find $E_1, E_2, \ldots, E_k$ and multiply these together (in the correct order), but there is an easier method for finding U.

Definition

Let A be an $m \times n$ matrix. We write

$$A \to B$$

if B can be obtained from A by a sequence of elementary row operations.

Suppose A is an $m \times n$ matrix and that $A \to B$. Then

1. There exists an invertible $m \times m$ matrix U such that $B = UA$;
2. U can be computed by performing elementary row operations on $\left[\begin{array}{c|c} A & I_m \end{array}\right]$ to transform it into $\left[\begin{array}{c|c} B & U \end{array}\right]$;
3. $U = E_kE_{k-1}\cdots E_2E_1$, where $E_1, E_2, \ldots, E_k$ are elementary matrices corresponding, in order, to the elementary row operations used to obtain B from A.

Example

Let $A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}$, and let R be the reduced row-echelon form of A.

Find a matrix U so that $R = UA$.

Solution

A Matrix as a Product of Elementary Matrices

Let

$$A = \begin{bmatrix} 1 & 2 & -4 \\ -3 & -6 & 13 \\ 0 & -1 & 2 \end{bmatrix}$$

Suppose we use row operations to put A in reduced row-echelon form, and write down the corresponding elementary matrices.

$$\begin{bmatrix} 1 & 2 & -4 \\ -3 & -6 & 13 \\ 0 & -1 & 2 \end{bmatrix} \underset{E_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & -4 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{bmatrix} \underset{E_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \underset{E_3}{\longrightarrow}$$
$$\begin{bmatrix} 1 & 2 & -4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \underset{E_4}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \underset{E_5}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Notice that the reduced row-echelon form of A equals I_3. Now find the matrices E_1, E_2, E_3, E_4 and E_5.

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$E_4 = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

It follows that

$$\begin{aligned} (E_5(E_4(E_3(E_2(E_1A))))) &= I \\ (E_5E_4E_3E_2E_1)A &= I \end{aligned}$$

and therefore

$$A^{-1} = E_5E_4E_3E_2E_1$$

Since $A^{-1} = E_5E_4E_3E_2E_1$,

$$\begin{aligned} A^{-1} &= E_5E_4E_3E_2E_1 \\ (A^{-1})^{-1} &= (E_5E_4E_3E_2E_1)^{-1} \\ A &= E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1} \end{aligned}$$

This example illustrates the following result.

Let A be an $n \times n$ matrix. Then, A^{-1} exists if and only if A can be written as the product of elementary matrices.

Example

Express $A = \begin{bmatrix} 4 & 1 \\ -3 & 2 \end{bmatrix}$ as a product of elementary matrices.

Solution

Let A be an $m \times n$ matrix of rank r. There exist invertible matrices U and V of size $m \times m$ and $n \times n$, respectively, such that

$$UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m \times n}$$

Moreover, if R is the reduced row-echelon form of A, then:

1. U can be computed by $\begin{bmatrix} A & I_m \end{bmatrix} \to \begin{bmatrix} R & U \end{bmatrix}$;
2. V can be computed by $\begin{bmatrix} R^T & I_n \end{bmatrix} \to \begin{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n \times m} & V^T \end{bmatrix}$.

Definition

If A is an $m \times n$ matrix of rank r, the matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ is called the **Smith normal form** of A. Whereas the reduced row-echelon form of A is the "nicest" matrix to which A can be carried by row operations, the Smith normal form is the "nicest" matrix to which A can be carried by *row and column* operations. This is because doing row operations to R^T amounts to doing *column* operations to R and then transposing.

Example

Given $A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 0 & 3 \\ 0 & 1 & -4 & 1 \end{bmatrix}$, find invertible matrices U and V such that $UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where $r = \text{rank } A$.

Solution

Section 2.6: Linear Transformations

Definition

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation** if it satisfies the following two properties for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and all (scalars) $a \in \mathbb{R}$.

1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ (preservation of addition)
2. $T(a\mathbf{x}) = aT(\mathbf{x})$ (preservation of scalar multiplication)

Properties of Linear Transformations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\mathbf{x} \in \mathbb{R}^n$.

Since T preserves scalar multiplication,

1. $T(0\mathbf{x}) = 0T(\mathbf{x})$ implying $T(\mathbf{0}) = \mathbf{0}$, so T preserves the zero vector.
2. $T((-1)\mathbf{x}) = (-1)T(\mathbf{x})$, implying $T(-\mathbf{x}) = -T(\mathbf{x})$, so T preserves the negative of a vector.

Suppose $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are vectors in $\mathbb{R}^n$ and

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k$$

for some $a_1, a_2, \ldots, a_k \in \mathbb{R}$.

Then

3. $$\begin{aligned} T(\mathbf{y}) &= T(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k) \\ &= a_1T(\mathbf{x}_1) + a_2T(\mathbf{x}_2) + \cdots + a_kT(\mathbf{x}_k), \end{aligned}$$

   so T preserves linear combinations.

Example

Let $T : \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation such that

$$T\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 0 \\ -2 \end{bmatrix} \text{ and } T\begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ -1 \\ 5 \end{bmatrix}. \text{ Find } T\begin{bmatrix} -7 \\ 3 \\ -9 \end{bmatrix}.$$

Solution

Practice

Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be a linear transformation such that

$$T\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} \text{ and } T\begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}. \text{ Find } T\begin{bmatrix} 1 \\ 3 \\ -2 \\ -4 \end{bmatrix}.$$

Solution

Matrix Transformations

Theorem

Every matrix transformation is a linear transformation.

Proof

Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation induced by the $m \times n$ matrix A,

i.e., $T(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x} \in \mathbb{R}^n$.

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and let $a \in \mathbb{R}$.

Then

$$T(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y}),$$

proving that T preserves addition.

Also,

$$T(a\mathbf{x}) = A(a\mathbf{x}) = a(A\mathbf{x}) = aT(\mathbf{x}),$$

proving that T preserves scalar multiplication.

Since T preserves addition and scalar multiplication T is a linear transformation.

Example

Recall $T : \mathbb{R}^3 \to \mathbb{R}^4$ defined by

$$T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b \\ b+c \\ a-c \\ c-b \end{bmatrix}$$

Is T a matrix transformation?

Consider $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix}$, then

$$A\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b \\ b+c \\ a-c \\ c-b \end{bmatrix} = T\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

So in this case T is a matrix transformation!

Not all transformations are matrix transformations

Consider $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(\mathbf{x}) = \mathbf{x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ for all } \mathbf{x} \in \mathbb{R}^2.$$

Why is T not a matrix transformation?

We have $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(\mathbf{x}) = \mathbf{x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ for all } \mathbf{x} \in \mathbb{R}^2.$$

Since every matrix transformation is a linear transformation,

we consider $T(0)$, where 0 is the zero vector of $\mathbb{R}^2$.

$$T \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

violating one of the properties of a linear transformation.

Therefore, T is not a linear transformation, and hence is not a matrix transformation.

Matrix Transformations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then we can find an $m \times n$ matrix A such that

$$T(\mathbf{x}) = A\mathbf{x}$$

In this case, we say that T is induced, or determined, by A and we write

$$T_A(\mathbf{x}) = A\mathbf{x}$$

Definition

The set of columns $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of I_n is called the **standard basis** of $\mathbb{R}^n$.

The Matrix of T

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

Then T is a matrix transformation.

Furthermore, T is induced by the unique matrix

$$A = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array} \right],$$

where $\mathbf{e}_j$ is the j^{th} column of I_n, and $T(\mathbf{e}_j)$ is the j^{th} column of A.

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation defined by

$$T \left[\begin{array}{c} x \\ y \end{array} \right] = \left[\begin{array}{c} x + 2y \\ x - y \end{array} \right]$$

for each $\mathbf{x} \in \mathbb{R}^2$. Find the matrix, A, of T.

Solution

Example

Sometimes T is not defined so nicely for us. Suppose T is given as

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \; T\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Find the matrix A of T.

Solution

Determining if a Transformation is Linear

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a transformation defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ y \\ -x+2y \end{bmatrix}$.

One way to show that T is a linear transformation is to show that it preserves addition and scalar multiplication.

However, now that we know that linear transformations are matrix transformations, we can use this to our advantage.

If T were a linear transformation, then T would be induced by the matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} T\begin{bmatrix} 1 \\ 0 \end{bmatrix} & T\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ -1 & 2 \end{bmatrix}.$$

Since

$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ y \\ -x+2y \end{bmatrix} = T\begin{bmatrix} x \\ y \end{bmatrix},$$

T is a matrix transformation, and therefore a linear transformation.

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a transformation defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ x+y \end{bmatrix}$.

If T were a linear transformation, then T would be induced by the matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} T\begin{bmatrix} 1 \\ 0 \end{bmatrix} & T\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

However,

$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ x+y \end{bmatrix}.$$

We see from this that if $x = 0$ or $y = 0$, then $xy = 0$, so $A\begin{bmatrix} x \\ y \end{bmatrix} = T\begin{bmatrix} x \\ y \end{bmatrix}$.

But if we take $x = y = 1$, then

$$A\begin{bmatrix} x \\ y \end{bmatrix} = A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \text{ while } T\begin{bmatrix} x \\ y \end{bmatrix} = T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

i.e., $A\begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq T\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Therefore, T is **not** a linear transformation.

Definition

Suppose $T : \mathbb{R}^k \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations.

The **composite** (or composition) of S and T is

$$S \circ T : \mathbb{R}^k \to \mathbb{R}^m,$$

is defined by

$$(S \circ T)(\mathbf{x}) = S(T(\mathbf{x})) \text{ for all } \mathbf{x} \in \mathbb{R}^k.$$

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, bend left, "T"] \arrow[rr, bend right, "S \circ T"'] & \mathbb{R}^n \arrow[r, bend left, "S"] & \mathbb{R}^m \end{tikzcd}$$

Be careful with the order of the transformations! We write $S \circ T$, but it is the transformation T that is applied first, followed by the transformation S.

Let $\mathbb{R}^k \xrightarrow{T} \mathbb{R}^n \xrightarrow{S} \mathbb{R}^m$ be linear transformations, and suppose that S is induced by matrix A, and T is induced by matrix B. Then $S \circ T$ is a linear transformation, and $S \circ T$ is induced by the matrix AB.

Example

Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformations defined by

$$S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix} \text{ and } T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} \text{ for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2.$$

Then S and T are induced by matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

respectively.

The composite of S and T is the transformation $S \circ T : \mathbb{R}^2 \to \mathbb{R}^2$

defined by

$$(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix} = S\left(T\begin{bmatrix} x \\ y \end{bmatrix}\right),$$

and has matrix (or is induced by the matrix)

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

Therefore the composite of S and T is the linear transformation

$$(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix} = AB\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ -x \end{bmatrix},$$

for each $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

Compare this with the composite of T and S which is the linear transformation

$$(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix},$$

for each $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

Rotations in $\mathbb{R}^2$

The transformation

$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$

denotes a counterclockwise rotation about the origin through an angle of θ.

Rotation through an angle of θ preserves scalar multiplication.

Rotation through an angle of θ preserves vector addition.

Since R_θ preserves addition and scalar multiplication, R_θ is a linear transformation, and hence a matrix transformation.

The Matrix for R_θ

The rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, and is induced by the matrix

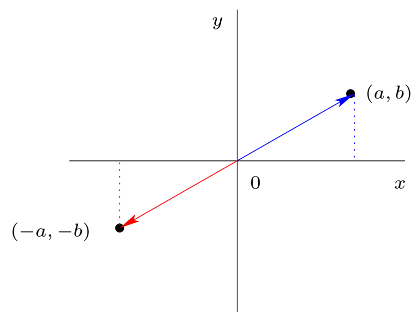
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Example (Rotation through π)

We denote by

$$R_\pi : \mathbb{R}^2 \to \mathbb{R}^2$$

a counterclockwise rotation about the origin through an angle of π.



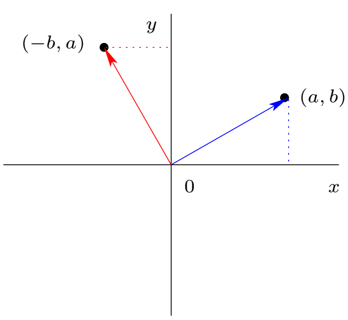
We see that $R_\pi \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a \\ -b \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$, so R_π is a matrix transformation.

Example (Rotation through $\pi/2$)

We denote by

$$R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$$

a counterclockwise rotation about the origin through an angle of $\pi/2$.



We see that $R_{\pi/2} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$, so $R_{\pi/2}$ is a matrix transformation.

Reflections in $\mathbb{R}^2$

Example

In $\mathbb{R}^2$, reflection in the x-axis, which transforms $\begin{bmatrix} a \\ b \end{bmatrix}$ to $\begin{bmatrix} a \\ -b \end{bmatrix}$, is a matrix transformation because

$$\begin{bmatrix} a \\ -b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

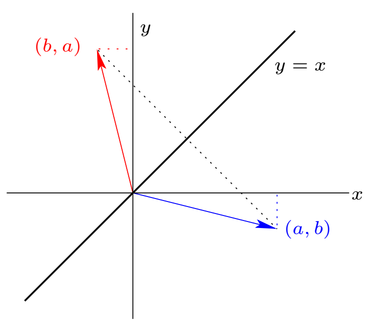
Example

In $\mathbb{R}^2$, reflection in the y-axis transforms $\begin{bmatrix} a \\ b \end{bmatrix}$ to $\begin{bmatrix} -a \\ b \end{bmatrix}$. This is a matrix transformation because

$$\begin{bmatrix} -a \\ b \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

Example

Reflection in the line $y = x$ transforms $\begin{bmatrix} a \\ b \end{bmatrix}$ to $\begin{bmatrix} b \\ a \end{bmatrix}$.

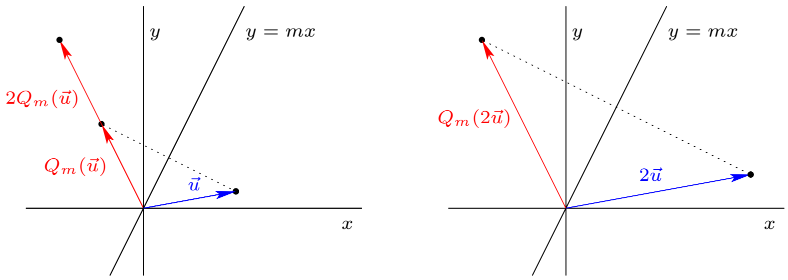


This is a matrix transformation because

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

Reflection in $y = mx$ preserves scalar multiplication

Let $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection in the line $y = mx$, and let $\vec{u} \in \mathbb{R}^2$.



The figure indicates that $Q_m(2\vec{u}) = 2Q_m(\vec{u})$.

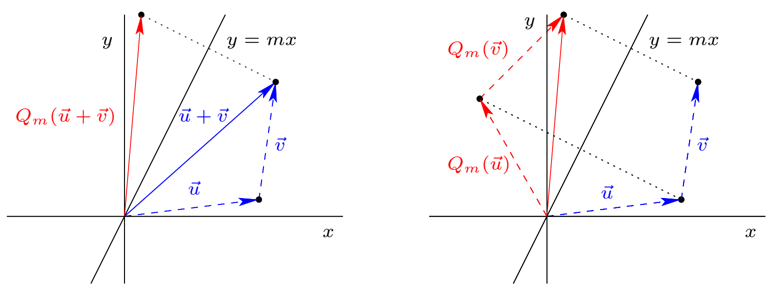
In general, for any scalar k,

$$Q_m(kX) = kQ_m(X),$$

i.e., Q_m preserves scalar multiplication.

Reflection in $y = mx$ preserves vector addition

Let $\vec{u}, \vec{v} \in \mathbb{R}^2$.



The figure indicates that

$$Q_m(\vec{u}) + Q_m(\vec{v}) = Q_m(\vec{u} + \vec{v}),$$

i.e., Q_m preserves vector addition.

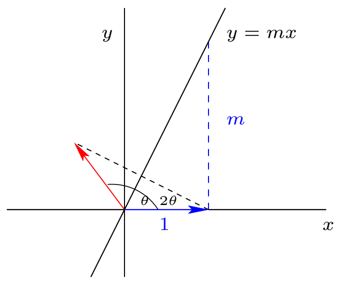
Q_m is a linear transformation

Since Q_m preserves addition and scalar multiplication, Q_m is a linear transformation, and hence a matrix transformation.

The matrix that induces Q_m can be found by computing $Q_m(\mathbf{e}_1)$ and $Q_m(\mathbf{e}_2)$, where

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

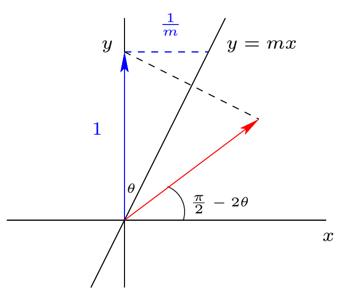
Computing $Q_m(\mathbf{e}_1)$



$$\cos\theta = \frac{1}{\sqrt{1+m^2}} \quad \text{and} \quad \sin\theta = \frac{m}{\sqrt{1+m^2}}$$

$$Q_m(\mathbf{e}_1) = \begin{bmatrix} \cos(2\theta) \\ \sin(2\theta) \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta \\ 2\sin\theta\cos\theta \end{bmatrix} = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 \\ 2m \end{bmatrix}$$

Computing $Q_m(\mathbf{e}_2)$



$$\cos\theta = \frac{m}{\sqrt{1+m^2}} \quad \text{and} \quad \sin\theta = \frac{1}{\sqrt{1+m^2}}$$

$$\begin{aligned} Q_m(\mathbf{e}_2) &= \begin{bmatrix} \cos(\frac{\pi}{2} - 2\theta) \\ \sin(\frac{\pi}{2} - 2\theta) \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{2}\cos(2\theta) + \sin\frac{\pi}{2}\sin(2\theta) \\ \sin\frac{\pi}{2}\cos(2\theta) - \cos\frac{\pi}{2}\sin(2\theta) \end{bmatrix} \\ &= \begin{bmatrix} \sin(2\theta) \\ \cos(2\theta) \end{bmatrix} = \begin{bmatrix} 2\sin\theta\cos\theta \\ \cos^2\theta - \sin^2\theta \end{bmatrix} = \frac{1}{1+m^2}\begin{bmatrix} 2m \\ m^2-1 \end{bmatrix} \end{aligned}$$

The Matrix for Reflection in $y = mx$

The transformation $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$, reflection in the line $y = mx$, is a linear transformation and is induced by the matrix

$$\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.$$

Multiple Actions

Reflection followed by Rotation

Example

Find the rotation or reflection that equals reflection in the x-axis followed by rotation through an angle of $\frac{\pi}{2}$.

Solution

Let Q_0 denote the reflection in the x-axis, and $R_{\frac{\pi}{2}}$ denote the rotation through an angle of $\frac{\pi}{2}$. We want to find the matrix for the transformation $R_{\frac{\pi}{2}} \circ Q_0$.

Q_0 is induced by $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and $R_{\frac{\pi}{2}}$ is induced by

$$B = \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Hence $R_{\frac{\pi}{2}} \circ Q_0$ is induced by

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

Notice that $BA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a reflection matrix.

Compare BA to

$$Q_m = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$$

Now, since $1 - m^2 = 0$, we know that $m = 1$ or $m = -1$. But $\frac{2m}{1+m^2} = 1 > 0$, so $m > 0$, implying $m = 1$.

Therefore,

$$R_{\frac{\pi}{2}} \circ Q_0 = Q_1,$$

reflection in the line $y = x$.

Reflection followed by Reflection

Example

Find the rotation or reflection that equals reflection in the line $y = -x$ followed by reflection in the y-axis.

Solution

We must find the matrix for the transformation $Q_Y \circ Q_{-1}$.

Q_{-1} is induced by

$$A = \frac{1}{2}\begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix},$$

and Q_Y is induced by

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

Therefore, $Q_Y \circ Q_{-1}$ is induced by BA.

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

What transformation does BA induce?

Rotation through an angle θ such that

$$\cos\theta = 0 \text{ and } \sin\theta = -1.$$

Therefore, $Q_Y \circ Q_{-1} = R_{-\frac{\pi}{2}} = R_{\frac{3\pi}{2}}$.

Summary

In general,

- The composite of two rotations is a rotation

$$R_\theta \circ R_\eta = R_{\theta+\eta}$$

- The composite of two reflections is a rotation.

$$Q_m \circ Q_n = R_\theta$$

  where θ is $2\times$ the angle between lines $y = mx$ and $y = nx$.

- The composite of a reflection and a rotation is a reflection.

$$R_\theta \circ Q_n = Q_m \circ Q_n \circ Q_n = Q_m$$

Discrete Mathematics II | Spring 2022

Section 2.7: LU-Factorization

Lecturer: Mr. Hall | *Math 270B*

Triangular Matrices

A matrix $A = [a_{ij}]$ is called **upper triangular** if $a_{ij} = 0$ whenever $i > j$. Thus the entries below the main diagonal equal 0.

$$\begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & \vdots \\ \vdots & \vdots & \ddots & * \\ 0 & \cdots & 0 & * \end{bmatrix}$$

where $*$ refers to any number.

A lower triangular matrix is defined similarly, as a matrix for which all entries **above** the main diagonal are equal to zero.

Definition

An **LU-Factorization** of a matrix A is written

$$A = LU$$

where L is a lower triangular matrix and U is an upper triangular matrix.

The LU-Factorization often helps to quickly solve equations of the form $AX = B$.

An LU-Factorization can be found for a matrix A provided that the row-echelon form of A can be calculated without interchanging rows.

Example

Find an LU-factorization of $A = \begin{bmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{bmatrix}$.

Solution

We lower reduce A to row-echelon form as follows:

$$A = \begin{bmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 6 & 12 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = U$$

The circled columns are determined as follows: The first is the leading column of A, and is used (by lower reduction) to create the first leading 1 and create zeros below it. This completes the work on row 1, and we repeat the procedure on the matrix consisting of the remaining rows. Thus the second circled column is the leading column of this smaller matrix, which we use to create the second leading 1 and the zeros below it. As the remaining row is zero here, we are finished. Then $A = LU$ where

$$L = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & 6 & 1 \end{bmatrix}$$

This matrix L is obtained from I_3 by replacing the bottom of the first two columns by the circled columns in the reduction. Note that the rank of A is 2 here, and this is the number of circled columns.

LU-Algorithm

Let A be an $m \times n$ matrix of rank r, and suppose that A can be lower reduced to a row-echelon matrix U. Then $A = LU$ where the lower triangular, invertible matrix L is constructed as follows:

1. If $A = 0$, take $L = I_m$ and $U = 0$.
2. If $A \neq 0$, write $A_1 = A$ and let $\mathbf{c}_1$ be the leading column of A_1. Use $\mathbf{c}_1$ to create the first leading 1 and create zeros below it (using lower reduction). When this is completed, let A_2 denote the matrix consisting of rows 2 to m of the matrix just created.
3. If $A_2 \neq 0$, let $\mathbf{c}_2$ be the leading column of A_2 and repeat Step 2 on A_2 to create A_3.
4. Continue in this way until U is reached, where all rows below the last leading 1 consist of zeros. This will happen after r steps.
5. Create L by placing $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r$ at the bottom of the first r columns of I_m.

Example

Find an LU-factorization for $A = \begin{bmatrix} 5 & -5 & 10 & 0 & 5 \\ -3 & 3 & 2 & 2 & 1 \\ -2 & 2 & 0 & -1 & 0 \\ 1 & -1 & 10 & 2 & 5 \end{bmatrix}$.

Solution

The next example deals with a case where no row of zeros is present in U (in fact, A is invertible).

Example

Find an LU-factorization for $A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix}$.

Solution

The reduction to row-echelon form is

$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = U$$

Hence $A = LU$ where $L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 2 & 5 \end{bmatrix}$.

Solving Systems using LU-Factorization

Suppose we wish to find all solutions $\mathbf{x}$ to the system $A\mathbf{x} = \mathbf{b}$. The LU factorization of A can assist in this process.

Consider the following reduction:

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ (LU)\mathbf{x} &= \mathbf{b} \\ L(U\mathbf{x}) &= \mathbf{b} \\ L\mathbf{y} &= \mathbf{b} \end{aligned}$$

Therefore if we can solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, then all that remains is to solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$.

Example

Find all solutions to

$$\begin{bmatrix} 1 & 3 & 2 & 0 \\ 3 & 10 & 5 & 1 \\ 0 & -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

Solution

Section 2.9: An Application to Markov Chains

Lecturer: Mr. Hall *Math 270B*

Markov Chains are used to model systems (or processes) that evolve through a series of stages. At each stage, the system is in one of a finite number of states.

Example

Three states: sunny (S), cloudy (C), rainy (R).

Stages: days.

The state that the system occupies at any stage is determined by a set of probabilities.

Important fact: probabilities are always real numbers between zero and one, inclusive.

- If it is sunny one day, then there is a 40% chance it will be sunny the next day, and a 40% chance that it will be cloudy the next day (and a 20% chance it will be rainy the next day).

The values 40%, 40% and 20% are **transition probabilities**, and are assumed to be known.

- If it is cloudy one day, then there is a 40% chance it will be rainy the next day, and a 25% chance that it will be sunny the next day.
- If it is rainy one day, then there is a 30% chance it will be rainy the next day, and a 50% chance that it will be cloudy the next day.

We put the transition probabilities into a **transition matrix**,

$$P = \begin{bmatrix} 0.4 & 0.25 & 0.2 \\ 0.4 & 0.35 & 0.5 \\ 0.2 & 0.4 & 0.3 \end{bmatrix}$$

Note. Transition matrices are **stochastic**, meaning that the sum of the entries in each column is equal to one.

Suppose that it is rainy on Thursday. What is the probability that it will be sunny on Sunday?

The **initial state** vector, $\mathbf{s}_0$, corresponds to the state of the weather on Thursday, so

$$\mathbf{s}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

What is the state vector for Friday?

$$\mathbf{s}_1 = \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.25 & 0.2 \\ 0.4 & 0.35 & 0.5 \\ 0.2 & 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = P\mathbf{s}_0.$$

To find the state vector for Saturday:

$$\mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} 0.4 & 0.25 & 0.2 \\ 0.4 & 0.35 & 0.5 \\ 0.2 & 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.265 \\ 0.405 \\ 0.33 \end{bmatrix}$$

Finally, the state vector for Sunday is

$$\mathbf{s}_3 = P\mathbf{s}_2 = \begin{bmatrix} 0.4 & 0.25 & 0.2 \\ 0.4 & 0.35 & 0.5 \\ 0.2 & 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} 0.265 \\ 0.405 \\ 0.33 \end{bmatrix} = \begin{bmatrix} 0.27325 \\ 0.41275 \\ 0.314 \end{bmatrix}$$

The probability that it will be sunny on Sunday is 27.325%.

Important fact: the sum of the entries of a state vector is always one.

Definition

Let P be the transition matrix for an n-state Markov chain. If $\mathbf{s}_m$ is the state vector at stage m, then

$$\mathbf{s}_{m+1} = P\mathbf{s}_m \text{ for } m = 0, 1, 2, \ldots$$

Example

- A customer always eats lunch either at restaurant A or restaurant B.
- The customer never eats at A two days in a row.
- If the customer eats at B one day, then the next day she is three times as likely to eat at B as at A.

Give the probability transition matrix.

Example

A wolf pack always hunts in one of three regions, R_1, R_2, and R_3.

- If it hunts in some region one day, it is as likely as not to hunt there again the next day.
- If it hunts in R_1, it never hunts in R_2 the next day.
- If it hunts in R_2 or R_3, it is equally likely to hunt in each of the other two regions the next day.

If the pack hunts in R_1 on Monday, find the probability that it will hunt in R_3 on Friday.

Solution

Sometimes, state vectors converge to a particular vector, called the **steady state** vector.

How do we know if a Markov chain has a steady state vector? If the Markov chain has a steady state vector, how do we find it?

One condition ensuring that a steady state vector exists is that the transition matrix P be **regular**, meaning that for some integer $k > 0$, all entries of P^k are positive (i.e., greater than zero).

In the restaurant example,

$P = \begin{bmatrix} 0 & \frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix}$ is regular because

$$P^2 = \begin{bmatrix} 0 & \frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{3}{16} \\ \frac{3}{4} & \frac{3}{16} \end{bmatrix}$$

has all entries greater than zero.

Definition

If P is the transition matrix of a Markov chain and P is regular, then the steady state vector can be found by solving the system

$$\mathbf{s} = P\mathbf{s}$$

for $\mathbf{s}$, and then ensuring that the entries of $\mathbf{s}$ sum to one.

Notice that if $\mathbf{s} = P\mathbf{s}$, then

$$\begin{aligned} \mathbf{s} - P\mathbf{s} &= \mathbf{0} \\ I\mathbf{s} - P\mathbf{s} &= \mathbf{0} \\ (I - P)\mathbf{s} &= \mathbf{0} \end{aligned}$$

- This last line represents a system of linear equations that is homogeneous.
- The structure of P ensures that $I - P$ is not invertible, and so the system has infinitely many solutions.
- Choose the value of the parameter so that the entries of $\mathbf{s}$ sum to one.

In the restaurant example,

$$P = \begin{bmatrix} 0 & \frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix},$$

and we've already verified that P is regular.

Now solve the system $(I - P)\mathbf{s} = \mathbf{0}$.

$$I - P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{4} \\ -1 & \frac{1}{4} \end{bmatrix}$$

Solving $(I - P)\mathbf{s} = \mathbf{0}$:

$$\left[\begin{array}{rr|r} 1 & -\frac{1}{4} & 0 \\ -1 & \frac{1}{4} & 0 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & -\frac{1}{4} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The general solution in parametric form is

$$s_1 = \frac{1}{4}t, s_2 = t \text{ for } t \in \mathbb{R}.$$

Since $s_1 + s_2 = 1$,

$$\begin{aligned} \frac{1}{4}t + t &= 1 \\ \frac{5}{4}t &= 1 \\ t &= \frac{4}{5} \end{aligned}$$

Therefore, the steady state vector is

$$S = \left[\begin{array}{c} \frac{1}{5} \\ \frac{4}{5} \end{array}\right] = \left[\begin{array}{c} 0.2 \\ 0.8 \end{array}\right]$$

Practice

Referring to the wolf pack problem,

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

Is there a steady state vector? If so, find it.

Solution

Discrete Mathematics II | Spring 2022

Section 3.1: The Cofactor Expansion

Lecturer: Mr. Hall | *Math 270B*

The Cofactor Expansion

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then the **determinant** of A is defined as

$$\det A = ad - bc,$$

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

Notation. For $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we often write $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, i.e., use vertical bars instead of square brackets.

More generally, the determinant of an $n \times n$ matrix is computed using determinants of $(n-1) \times (n-1)$ submatrices.

Let $A = [a_{ij}]$ be an $n \times n$ matrix.

- The sign of the (i, j) position is $(-1)^{i+j}$.
  Thus the sign is 1 if $(i + j)$ is even, and -1 if $(i + j)$ is odd.

Let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j.

- The (i, j)-cofactor of A is
  $$c_{ij}(A) = (-1)^{i+j} \det(A_{ij}).$$

Finally,

- $\det A = a_{11}c_{11}(A) + a_{12}c_{12}(A) + a_{13}c_{13}(A) + \cdots + a_{1n}c_{1n}(A)$,
  and is called the cofactor expansion of A along row 1.

Example

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Find $\det A$ by using cofactor expansion along row 1.

Solution

Example

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Find $\det A$ by using cofactor expansion along column 2.

Solution

Fact

The determinant of an $n \times n$ matrix A can be computed using the cofactor expansion along any row or column of A.

That is, $\det A$ can be computed by multiplying each entry of the row or column by the corresponding cofactor and adding the results.

The next example illustrates the importance of this fact.

Example

Let $A = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 5 & 0 & 0 & 7 \\ 0 & 1 & -1 & 0 \\ 3 & 0 & 0 & 2 \end{bmatrix}$. Find $\det A$.

Solution

Example

Find $\det A$ for $A = \begin{bmatrix} -8 & 1 & 0 & -4 \\ 5 & 7 & 0 & -7 \\ 12 & -3 & 0 & 8 \\ -3 & 11 & 0 & 2 \end{bmatrix}$.

Solution

Elementary Row Operations and Determinants

Let $A = \begin{bmatrix} 2 & 0 & -3 \\ 0 & 4 & 0 \\ 1 & 0 & -2 \end{bmatrix}$. Then

$$\det A = 4(-1)^4 \begin{vmatrix} 2 & -3 \\ 1 & -2 \end{vmatrix} = 4(-1) = -4.$$

Let B_1, B_2, and B_3 be obtained from A by performing a type 1, 2 and 3 elementary row operation, respectively, i.e.,

$$B_1 = \begin{bmatrix} 2 & 0 & -3 \\ 1 & 0 & -2 \\ 0 & 4 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 2 & 0 & -3 \\ 0 & 4 & 0 \\ -3 & 0 & 6 \end{bmatrix}, B_3 = \begin{bmatrix} 2 & 0 & -3 \\ 0 & 4 & 0 \\ 5 & 0 & -8 \end{bmatrix}.$$

Compute $\det B_1$, $\det B_2$, and $\det B_3$.

Definition

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

1. If A has a row or column of zeros, then $\det A = 0$.
2. If B is obtained from A by exchanging two different rows (or columns) of A, then $\det B = -\det A$.
3. If B is obtained from A by multiplying a row (or column) of A by a scalar $k \in \mathbb{R}$, then $\det B = k \det A$.
4. If B is obtained from A by adding k times one row of A to a different row of A (or adding k times one column of A to a different column of A) then $\det B = \det A$.
5. If two different rows (or columns) of A are identical, then $\det A = 0$.

Example

Find $\det A$ for $A = \begin{bmatrix} 3 & 1 & 2 & 4 \\ -1 & -3 & 8 & 0 \\ 1 & -1 & 5 & 5 \\ 1 & 1 & 2 & -1 \end{bmatrix}$

Solution

Example

If $\det \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} = -1$, find $\det \begin{bmatrix} -x & -y & -z \\ 3p+a & 3q+b & 3r+c \\ 2p & 2q & 2r \end{bmatrix}$.

Solution

Example

Suppose A is a 3×3 matrix with $\det A = -1$. What is $\det(-3A)$?

Write $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Then $-3A = \begin{bmatrix} -3a_{11} & -3a_{12} & -3a_{13} \\ -3a_{21} & -3a_{22} & -3a_{23} \\ -3a_{31} & -3a_{32} & -3a_{33} \end{bmatrix}$.

$$\det(-3A) = \begin{vmatrix} -3a_{11} & -3a_{12} & -3a_{13} \\ -3a_{21} & -3a_{22} & -3a_{23} \\ -3a_{31} & -3a_{32} & -3a_{33} \end{vmatrix} = (-3) \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ -3a_{21} & -3a_{22} & -3a_{23} \\ -3a_{31} & -3a_{32} & -3a_{33} \end{vmatrix}$$

$$= (-3)(-3) \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -3a_{31} & -3a_{32} & -3a_{33} \end{vmatrix} = (-3)(-3)(-3) \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$= (-3)^3 \det A = (-27) \times (-1) = 27.$

Determinant of a Scalar Multiple of a Matrix

If A is an $n \times n$ matrix and $k \in \mathbb{R}$ is a scalar, then

$$\det(kA) = k^n \det A.$$

Example

Let

$$A = \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} \text{ and } B = \begin{bmatrix} 2a+p & 2b+q & 2c+r \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{bmatrix}$$

Show that $\det B = 9 \det A$.

Solution

Definitions

1. An $n \times n$ matrix A is called **upper triangular** if and only if all entries **below** the main diagonal are zero.
2. An $n \times n$ matrix A is called **lower triangular** if and only if all entries **above** the main diagonal are zero.
3. An $n \times n$ matrix A is called **triangular** if and only if it is upper triangular or lower triangular.

The Determinant of a Triangular Matrix

If $A = \left[\ a_{ij}\ \right]$ is an $n \times n$ triangular matrix, then

$$\det A = a_{11}a_{22}a_{33}\cdots a_{nn},$$

i.e., $\det A$ is the product of the entries of the main diagonal of A.

Example

Find $\det A$ for $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{bmatrix}$

Solution

Determinants of Block Matrices

Consider the matrices

$$\begin{bmatrix} A & X \\ 0 & B \end{bmatrix} \text{ and } \begin{bmatrix} A & 0 \\ Y & B \end{bmatrix}$$

where A and B are square matrices. Then

$$\det \begin{bmatrix} A & X \\ 0 & B \end{bmatrix} = \det A \det B \text{ and } \det \begin{bmatrix} A & 0 \\ Y & B \end{bmatrix} = \det A \det B.$$

Example

$$\det \begin{bmatrix} 1 & -1 & 2 & 0 & -2 \\ 0 & 1 & 0 & 4 & 1 \\ 1 & 1 & 5 & 0 & 0 \\ 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} = \det \left[\begin{array}{ccc|cc} 1 & -1 & 2 & 0 & -2 \\ 0 & 1 & 0 & 4 & 1 \\ 1 & 1 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

$$= \det \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & 5 \end{bmatrix} \det \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & 2 \\ 1 & 5 \end{bmatrix} \det \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}$$

$$= 3 \times (-2) = -6.$$

Section 3.2: Determinants and Matrix Inverses

Determinant of a Matrix Product

If A and B are $n \times n$ matrices, then
$$\det(AB) = \det A \det B.$$

Determinant of a Matrix Transpose

If A is an $n \times n$ matrix, then $\det(A^T) = \det A$.

Determinant of a Matrix Inverse

An $n \times n$ matrix A is invertible if and only if $\det A \neq 0$. In this case,
$$\det(A^{-1}) = \frac{1}{\det A}.$$

Example

Find all values of c for which $A = \begin{bmatrix} c & 1 & 0 \\ 0 & 2 & c \\ -1 & c & 5 \end{bmatrix}$ is invertible.

Solution

Example

Suppose A is a 3×3 matrix. Find $\det A$ and $\det B$ if

$$\det(2A^{-1}) = -4 = \det(A^3 (B^{-1})^T).$$

Solution

Example

Suppose A, B and C are 4×4 matrices with

$$\det A = -1, \det B = 2, \text{ and } \det C = 1.$$

Find $\det(2A^2(B^{-1})(C^T)^3 B(A^{-1}))$.

Solution

Example

A square matrix A is **orthogonal** if and only if $A^T = A^{-1}$. What are the possible values of $\det A$ if A is orthogonal?

Solution

Since $A^T = A^{-1}$,

$$\begin{aligned} \det A^T &= \det(A^{-1}) \\ \det A &= \frac{1}{\det A} \\ (\det A)^2 &= 1 \end{aligned}$$

Assuming A is a real matrix, this implies that $\det A = \pm 1$, i.e., $\det A = 1$ or $\det A = -1$.

Adjugates

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the **adjugate** of A is defined as

$$\text{adj}A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

and observe that

$$\begin{aligned} A(\text{adj}A) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} \\ &= (\det A)I_2 \end{aligned}$$

Furthermore, if $\det A \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det A}\text{adj}A.$$

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

$$\text{adj}A = \begin{bmatrix} c_{ij}(A) \end{bmatrix}^T,$$

where $c_{ij}(A)$ is the (i,j)-cofactor of A, i.e., adj$A$ is the transpose of the matrix of cofactors.

Reminder. $c_{ij}(A) = (-1)^{i+j}\det(A_{ij})$.

Example

Find adj$A$ when $A = \begin{bmatrix} 2 & 1 & 3 \\ 5 & -7 & 1 \\ 3 & 0 & -6 \end{bmatrix}$.

Solution

The Adjugate Formula

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

$$A(\mathrm{adj}A) = (\det A)I = (\mathrm{adj}A)A.$$

Furthermore, if $\det A \neq 0$, then

$$A^{-1} = \frac{1}{\det A}\mathrm{adj}A.$$

Note. Except in the case of a 2×2 matrix, the adjugate formula is a very inefficient method for computing the inverse of a matrix; the matrix inversion algorithm is much more practical. However, the adjugate formula is of theoretical significance.

Example

For an $n \times n$ matrix A, show that $\det(\mathrm{adj}A) = (\det A)^{n-1}$.

Proof

Using the adjugate formula,

$$\begin{array}{rcl} A(\mathrm{adj}A) & = & (\det A)I \\ \det(A(\mathrm{adj}A)) & = & \det((\det A)I) \\ (\det A) \times \det(\mathrm{adj}A) & = & (\det A)^n(\det I) \\ (\det A) \times \det(\mathrm{adj}A) & = & (\det A)^n \end{array}$$

If $\det A \neq 0$, then divide both sides of the last equation by $\det A$:

$$\det(\mathrm{adj}A) = (\det A)^{n-1}.$$

If $\det A = 0$, then

$$A(\mathrm{adj}A) = (\det A)I = (0)I = 0,$$

i.e., $A(\mathrm{adj}A)$ is the zero matrix.

In this case, if $\det(\mathrm{adj}A)$ were **not** equal to zero, then $\mathrm{adj}A$ would be invertible, and $A(\mathrm{adj}A) = 0$ would imply $A = 0$.

However, if $A = 0$, then $\mathrm{adj}A = 0$ and is **not** invertible, and thus has determinant equal to zero, i.e., $\det(\mathrm{adj}A) = 0$.

Therefore, if $\det A = 0$, then

$$\det(\mathrm{adj}A) = 0 = 0^{n-1} = (\det A)^{n-1}.$$

Example

Let A and B be $n \times n$ matrices. Show that $\det(A + B^T) = \det(A^T + B)$.

Solution

Cramer's Rule

If A is an $n \times n$ invertible matrix, then the solution to $A\mathbf{x} = \mathbf{b}$ can be given in terms of determinants of matrices.

Let A be an $n \times n$ invertible matrix, the solution to the system $A\mathbf{x} = \mathbf{b}$ of n equations in the variables x_1, $x_2 \cdots x_n$ is given by

$$x_1 = \frac{\det A_1}{\det A}, x_2 = \frac{\det A_2}{\det A}, \cdots, x_n = \frac{\det A_n}{\det A}$$

where, for each k, the matrix A_k is obtained from A by replacing column k with $\mathbf{b}$.

Example

Use Cramer's Rule to solve the system for x_3

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & -1 \\ 5x_1 & + & 2x_2 & & & = & 2 \\ x_1 & + & x_2 & - & x_3 & = & 1 \end{array}$$

Solution

Polynomial Interpolation

Given data points $(0,1)$, $(1,2)$, $(2,5)$ and $(3,10)$, find an interpolating polynomial $p(x)$ of degree at most three, and then estimate the value of y corresponding to $x = \frac{3}{2}$.

We want to find the coefficients r_0, r_1, r_2 and r_3 of

$$p(x) = r_0 + r_1 x + r_2 x^2 + r_3 x^3$$

so that $p(0) = 1$, $p(1) = 2$, $p(2) = 5$, and $p(3) = 10$.

$$\begin{array}{rcl}
p(0) & = & r_0 = 1 \\
p(1) & = & r_0 + r_1 + r_2 + r_3 = 2 \\
p(2) & = & r_0 + 2r_1 + 4r_2 + 8r_3 = 5 \\
p(3) & = & r_0 + 3r_1 + 9r_2 + 27r_3 = 10
\end{array}$$

Solve this system of four equations in the four variables r_0, r_1, r_2 and r_3.

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 4 & 8 & 5 \\ 1 & 3 & 9 & 27 & 10 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Therefore $r_0 = 1$, $r_1 = 0$, $r_2 = 1$, $r_3 = 0$, and so

$$p(x) = 1 + x^2.$$

The estimate is

$$y = p\left(\frac{3}{2}\right) = 1 + \left(\frac{3}{2}\right)^2 = \frac{13}{4}.$$

Definition

Given n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ with the x_i **distinct**, there is a unique polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}$$

such that $p(x_i) = y_i$ for $i = 1, 2, \ldots, n$.

The polynomial $p(x)$ is called the **interpolating polynomial** for the data.

To find $p(x)$, set up a system of n linear equations in the n variables $r_0, r_1, r_2, \ldots, r_{n-1}$.
$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}$:

$$\begin{array}{rcl} r_0 + r_1 x_1 + r_2 x_1^2 + \cdots + r_{n-1} x_1^{n-1} & = & y_1 \\ r_0 + r_1 x_2 + r_2 x_2^2 + \cdots + r_{n-1} x_2^{n-1} & = & y_2 \\ r_0 + r_1 x_3 + r_2 x_3^2 + \cdots + r_{n-1} x_3^{n-1} & = & y_3 \\ \vdots & \vdots & \vdots \\ r_0 + r_1 x_n + r_2 x_n^2 + \cdots + r_{n-1} x_n^{n-1} & = & y_n \end{array}$$

The coefficient matrix for this system is

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

The determinant of a matrix of this form is called a **Vandermonde** determinant.

The Vandermonde Determinant

Let $a_1, a_2, \ldots, a_n$ be real numbers, $n \geq 2$. Then the corresponding Vandermonde determinant is given by

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \prod_{1 \leq j < i \leq n} (a_i - a_j)$$

In our earlier example with the data points $(0,1)$, $(1,2)$, $(2,5)$ and $(3,10)$, we have

$$a_1 = 0, a_2 = 1, a_3 = 2, a_4 = 3$$

giving us the Vandermonde determinant

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{vmatrix}$$

This determinant is equal to

$$\begin{aligned} & (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)(a_4 - a_1)(a_4 - a_2)(a_4 - a_3) \\ = \; & (1-0)(2-0)(2-1)(3-0)(3-1)(3-2) = 2 \times 3 \times 2 \\ = \; & 12. \end{aligned}$$

The Vandermonde determinant is nonzero if $a_1, a_2, \ldots, a_n$ are **distinct**.

This means that given n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ with **distinct** x_i, then there is a unique interpolating polynomial

$$p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}.$$

Example

Let $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$. Find A^{100}.

How can we do this efficiently?

Consider the matrix $P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$. Observe that P is invertible, and that

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

Furthermore,

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

where D is a **diagonal** matrix.

This is significant, because

$$\begin{aligned} P^{-1}AP &= D \\ P(P^{-1}AP)P^{-1} &= PDP^{-1} \\ (PP^{-1})A(PP^{-1}) &= PDP^{-1} \\ IAI &= PDP^{-1} \\ A &= PDP^{-1}, \end{aligned}$$

and so

$$\begin{aligned} A^{100} &= (PDP^{-1})^{100} \\ &= (PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)D(P^{-1}\cdots P)DP^{-1} \\ &= PDIDIDI\cdots IDP^{-1} \\ &= PD^{100}P^{-1}. \end{aligned}$$

Now,

$$D^{100} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}^{100} = \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix}.$$

Therefore,

$$\begin{aligned} A^{100} &= PD^{100}P^{-1} \\ &= \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix} \left(\frac{1}{3}\right) \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2^{100} + 2 \cdot 5^{100} & 2 \cdot 2^{100} - 2 \cdot 5^{100} \\ 2^{100} - 5^{100} & 2 \cdot 2^{100} + 5^{100} \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2^{100} + 2 \cdot 5^{100} & 2^{101} - 2 \cdot 5^{100} \\ 2^{100} - 5^{100} & 2^{101} + 5^{100} \end{bmatrix} \end{aligned}$$

Diagonalization and Matrix Powers

If $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$ for each $k = 1, 2, 3, \ldots$

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.

Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix, λ a real number, and $\mathbf{x} \neq \mathbf{0}$ an n-vector. If $A\mathbf{x} = \lambda\mathbf{x}$, then λ is an **eigenvalue** of A, and $\mathbf{x}$ is an **eigenvector** of A corresponding to λ, or a **λ-eigenvector**.

Example

Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

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

This means that 3 is an eigenvalue of A, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to 3 (or a 3-eigenvector of A).

Finding all Eigenvalues and Eigenvectors of a Matrix

Suppose that A is an $n \times n$ matrix, $\mathbf{x} \neq \mathbf{0}$ an n-vector, $\lambda \in \mathbb{R}$, and that $A\mathbf{x} = \lambda\mathbf{x}$.

Then

$$\begin{aligned} \lambda\mathbf{x} - A\mathbf{x} &= \mathbf{0} \\ \lambda I\mathbf{x} - A\mathbf{x} &= \mathbf{0} \\ (\lambda I - A)\mathbf{x} &= \mathbf{0} \end{aligned}$$

Since $\mathbf{x} \neq \mathbf{0}$, the matrix $\lambda I - A$ has no inverse, and thus

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

Definition

The **characteristic polynomial** of an $n \times n$ matrix A is

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

Example

Find the characteristic polynomial of $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$

Solution

Definition

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

1. The eigenvalues of A are the roots of $c_A(\lambda)$.
2. The λ-eigenvectors, $\mathbf{x}$, are the nontrivial solutions to $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

Example

For $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$, we have

$$c_A(\lambda) = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5),$$

so A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$.

To find the 2-eigenvectors of A, solve $(2I - A)\mathbf{x} = \mathbf{0}$:

$$\left[\begin{array}{rr|r} -2 & 2 & 0 \\ 1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & -1 & 0 \\ -2 & 2 & 0 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The general solution, in parametric form, is

$$\mathbf{x} = \begin{bmatrix} t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ where } t \in \mathbb{R}.$$

To find the 5-eigenvectors of A, solve $(5I - A)\mathbf{x} = \mathbf{0}$:

$$\left[\begin{array}{rr|r} 1 & 2 & 0 \\ 1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The general solution, in parametric form, is

$$\mathbf{x} = \begin{bmatrix} -2t \\ t \end{bmatrix} = t\begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ where } t \in \mathbb{R}.$$

Definition

A **basic eigenvector** of an $n \times n$ matrix A is any nonzero multiple of a basic solution to $(\lambda I - A)\mathbf{x} = \mathbf{0}$, where λ is an eigenvalue of A.

From the previous example, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ are basic eigenvectors of the matrix

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

corresponding to the eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$, respectively.

Example

For $A = \begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$, find $c_A(\lambda)$, the eigenvalues of A, and the corresponding basic eigenvectors.

Solution

Solution (continued)

Geometric Interpretation of Eigenvalues and Eigenvectors

Let A be a 2×2 matrix. Then A can be interpreted as a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.

How does the linear transformation affect the eigenvectors of the matrix?

Let $\mathbf{x}$ be a nonzero vector in $\mathbb{R}^2$. We denote by $L_{\mathbf{x}}$ the unique line in $\mathbb{R}^2$ that contains $\mathbf{x}$ and the origin.

Let $\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}$ be a nonzero vector in $\mathbb{R}^2$. Then $L_{\mathbf{x}}$ is the set of all scalar multiples of $\mathbf{x}$, i.e.,

$$L_{\mathbf{x}} = \mathbb{R}\mathbf{x} = \{t\mathbf{x} \mid t \in \mathbb{R}\}.$$

Definition

Let A be a 2×2 matrix and L a line in $\mathbb{R}^2$ through the origin. Then L is said to be A**-invariant** if the vector $A\mathbf{x}$ lies in L whenever $\mathbf{x}$ lies in L,

i.e., $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$,

i.e., $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar $\lambda \in \mathbb{R}$,

i.e., $\mathbf{x}$ is an eigenvector of A.

Diagonalization

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 $\mathrm{diag}(a_1, a_2, a_3, \ldots, a_{n-1}, a_n)$.

Recall that if A is an $n \times n$ matrix and P is an invertible $n \times n$ matrix so that $P^{-1}AP$ is diagonal, then P is called a **diagonalizing matrix** of A, and A is **diagonalizable**.

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

1. A is diagonalizable if and only if it has eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ so that
$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$$
is invertible.
2. If P is invertible, then
$$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$
where λ_i is the eigenvalue of A corresponding to the eigenvector $\mathbf{x}_i$, i.e., $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$.

From our earlier example

$A = \begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$ has eigenvalues and corresponding basic eigenvectors

$$\lambda_1 = 1 \text{ and } \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \lambda_2 = 2 \text{ and } \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}; \lambda_3 = 3 \text{ and } \mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Let $P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Then P is invertible, so by the above Theorem,

$$P^{-1}AP = \mathrm{diag}(1, 2, 3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Note

It is not always possible to find n eigenvectors so that P is invertible.

Example

Let $A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 6 & -6 \\ 1 & 2 & -1 \end{bmatrix}$. Then

$$c_A(\lambda) = \begin{vmatrix} \lambda - 1 & 2 & -3 \\ -2 & \lambda - 6 & 6 \\ -1 & -2 & \lambda + 1 \end{vmatrix} = \cdots = (\lambda - 2)^3.$$

A has only one eigenvalue, $\lambda_1 = 2$, with multiplicity three.

To find the 2-eigenvectors of A, solve the system $(2I - A)\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ -2 & -4 & 6 & 0 \\ -1 & -2 & 3 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The general solution in parametric form is

$$\mathbf{x} = \begin{bmatrix} -2s + 3t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, s, t \in \mathbb{R}.$$

Since the system has only **two** basic solutions, there are only two basic eigenvectors, implying that the matrix A is **not diagonalizable**.

Example

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

Solution

Matrix Diagonalization Test

A square matrix A is diagonalizable if and only if every eigenvalue λ of multiplicity m yields exactly m basic eigenvectors, i.e., the solution to $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has m parameters.

A special case of this is:

Distinct Eigenvalues and Diagonalization

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

Example

Show that $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is not diagonalizable.

Solution

Linear Dynamical Systems

A **linear dynamical system** consists of

- an $n \times n$ matrix A and an n-vector $\mathbf{v}_0$;
- a matrix recursion defining $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$ by $\mathbf{v}_{k+1} = A\mathbf{v}_k$; i.e.,

$$\begin{array}{rcl} \mathbf{v}_1 & = & A\mathbf{v}_0 \\ \mathbf{v}_2 & = & A\mathbf{v}_1 = A(A\mathbf{v}_0) = A^2\mathbf{v}_0 \\ \mathbf{v}_3 & = & A\mathbf{v}_2 = A(A^2\mathbf{v}_0) = A^3\mathbf{v}_0 \\ \vdots & \vdots & \vdots \\ \mathbf{v}_k & = & A^k\mathbf{v}_0. \end{array}$$

Linear dynamical systems are used, for example, to model the evolution of populations over time.

If A is diagonalizable, then
$$P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$
where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of A.

Thus $A = PDP^{-1}$, and $A^k = PD^kP^{-1}$. Therefore,
$$\mathbf{v}_k = A^k\mathbf{v}_0 = PD^kP^{-1}\mathbf{v}_0.$$

Example

Consider the linear dynamical system $\mathbf{v}_{k+1} = A\mathbf{v}_k$ with
$$A = \left[\begin{array}{rr} 2 & 0 \\ 3 & -1 \end{array}\right], \text{ and } \mathbf{v}_0 = \left[\begin{array}{r} 1 \\ -1 \end{array}\right].$$

Find a formula for $\mathbf{v}_k$.

Solution

First, $c_A(\lambda) = (\lambda - 2)(\lambda + 1)$, so A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -1$, and thus is diagonalizable.

Solve $(2I - A)\mathbf{x} = \mathbf{0}$:
$$\left[\begin{array}{rr|r} 0 & 0 & 0 \\ -3 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$
has general solution $\mathbf{x} = \left[\begin{array}{c} t \\ t \end{array}\right]$, $t \in \mathbb{R}$, and basic solution $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$.

Solve $(-I-A)\mathbf{x}=\mathbf{0}$:

$$\left[\begin{array}{rr|r} -3 & 0 & 0 \\ -3 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

has general solution $\mathbf{x}=\left[\begin{array}{c} 0 \\ t \end{array}\right]$, $t\in\mathbb{R}$, and basic solution $\mathbf{x}_2=\left[\begin{array}{c} 0 \\ 1 \end{array}\right]$.

Thus, $P=\left[\begin{array}{rr} 1 & 0 \\ 1 & 1 \end{array}\right]$ is a diagonalizing matrix for A,

$$P^{-1}=\left[\begin{array}{rr} 1 & 0 \\ -1 & 1 \end{array}\right], \text{ and } P^{-1}AP=\left[\begin{array}{rr} 2 & 0 \\ 0 & -1 \end{array}\right].$$

Therefore,

$$\begin{array}{rcl}
\mathbf{v}_k & = & A^k\mathbf{v}_0 \\
 & = & PD^kP^{-1}\mathbf{v}_0 \\
 & = & \left[\begin{array}{rr} 1 & 0 \\ 1 & 1 \end{array}\right]\left[\begin{array}{rr} 2 & 0 \\ 0 & -1 \end{array}\right]^k\left[\begin{array}{rr} 1 & 0 \\ -1 & 1 \end{array}\right]\left[\begin{array}{r} 1 \\ -1 \end{array}\right] \\
 & = & \left[\begin{array}{rr} 1 & 0 \\ 1 & 1 \end{array}\right]\left[\begin{array}{cc} 2^k & 0 \\ 0 & (-1)^k \end{array}\right]\left[\begin{array}{r} 1 \\ -2 \end{array}\right] \\
 & = & \left[\begin{array}{cc} 2^k & 0 \\ 2^k & (-1)^k \end{array}\right]\left[\begin{array}{r} 1 \\ -2 \end{array}\right] \\
 & = & \left[\begin{array}{c} 2^k \\ 2^k-2(-1)^k \end{array}\right]
\end{array}$$

Remark

Often, instead of finding an exact formula for $\mathbf{v}_k$, it suffices to estimate $\mathbf{v}_k$ as k gets large.

This can easily be done if A has a **dominant eigenvalue with multiplicity one**: an eigenvalue λ_1 with the property that

$$|\lambda_1|>|\lambda_j| \text{ for } j=2,3,\ldots,n.$$

Suppose that

$$\mathbf{v}_k=PD^kP^{-1}\mathbf{v}_0,$$

and assume that A has a dominant eigenvalue, λ_1, with corresponding basic eigenvector $\mathbf{x}_1$ as the first column of P.

For convenience, write $P^{-1}\mathbf{v}_0=\left[\begin{array}{cccc} b_1 & b_2 & \cdots & b_n \end{array}\right]^T$.

Then

$$\begin{aligned}
\mathbf{v}_k &= PD^kP^{-1}\mathbf{v}_0 \\
&= \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array}\right] \left[\begin{array}{cccc} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{array}\right] \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array}\right] \\
&= b_1\lambda_1^k\mathbf{x}_1 + b_2\lambda_2^k\mathbf{x}_2 + \cdots + b_n\lambda_n^k\mathbf{x}_n \\
&= \lambda_1^k\left(b_1\mathbf{x}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{x}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{x}_n\right)
\end{aligned}$$

Now, $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$ for $j = 2, 3, \ldots n$, and thus $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$.

Therefore, for large values of k, $\mathbf{v}_k \approx \lambda_1^k b_1 \mathbf{x}_1$.

Example

If

$$A = \left[\begin{array}{rr} 2 & 0 \\ 3 & -1 \end{array}\right], \text{ and } \mathbf{v}_0 = \left[\begin{array}{r} 1 \\ -1 \end{array}\right],$$

estimate $\mathbf{v}_k$ for large values of k.

Solution

In our previous example, we found that A has eigenvalues 2 and -1. This means that $\lambda_1 = 2$ is a dominant eigenvalue; let $\lambda_2 = -1$.

As before $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$ is a basic eigenvector for $\lambda_1 = 2$, and $\mathbf{x}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ is a basic eigenvector for $\lambda_2 = -1$, giving us

$$P = \left[\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array}\right], \text{ and } P^{-1} = \left[\begin{array}{rr} 1 & 0 \\ -1 & 1 \end{array}\right].$$

$$P^{-1}\mathbf{v}_0 = \left[\begin{array}{rr} 1 & 0 \\ -1 & 1 \end{array}\right]\left[\begin{array}{r} 1 \\ -1 \end{array}\right] = \left[\begin{array}{r} 1 \\ -2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right]$$

For large values of k,

$$\mathbf{v}_k \approx \lambda_1^k b_1 \mathbf{x}_1 = 2^k(1)\left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{c} 2^k \\ 2^k \end{array}\right]$$

Let's compare this to the formula for V_k that we obtained earlier:

$$\mathbf{v}_k = \left[\begin{array}{c} 2^k \\ 2^k - 2(-1)^k \end{array}\right]$$

Section 3.4: An Application to Linear Recurrences

Lecturer: Mr. Hall

Definitions

A sequence of numbers $x_0, x_1, x_2, x_3, \dots$ is defined **recursively** if each number in the sequence is determined by the numbers that occur before it in the sequence.

A **linear recurrence** of length k has the form

$$x_{n+k} = a_1 x_{n+k-1} + a_2 x_{n+k-2} + \cdots + a_k x_n, n \geq 0,$$

for some real numbers $a_1, a_2, \dots, a_k$.

The simplest linear recurrence has length one, so has the form

$$x_{n+1} = a x_n \text{ for } n \geq 0,$$

with $a \in \mathbb{R}$ and some initial value x_0.

In this case,

$$\begin{array}{rcl} x_1 & = & a x_0 \\ x_2 & = & a x_1 = a^2 x_0 \\ x_3 & = & a x_2 = a^3 x_0 \\ \vdots & \vdots & \vdots \\ x_n & = & a x_{n-1} = a^n x_0 \end{array}$$

Therefore, $x_n = a^n x_0$.

Example

Solve the linear recurrence relation

$$x_{k+2} = 2x_{k+1} + 3x_k \text{ for } k \geq 0,$$

with $x_0 = 0$ and $x_1 = 1$.

Solution

Practice

Solve the linear recurrence relation

$$x_{k+2} = 5x_{k+1} - 6x_k, k \geq 0$$

with $x_0 = 0$ and $x_1 = 1$.

Solution

Discrete Mathematics II | Spring 2022

Section 4.1: Vectors and Lines

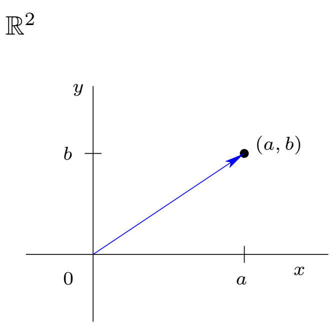
Lecturer: Mr. Hall | *Math 270B*

Scalar quantities versus vector quantities

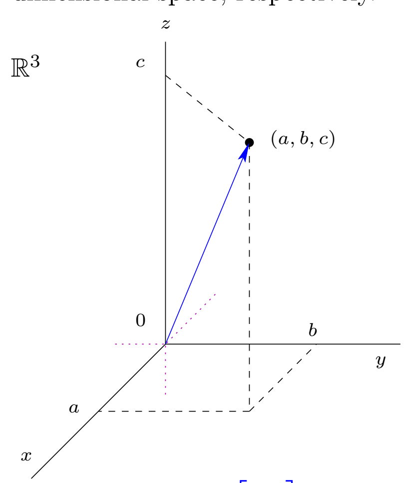
- A **scalar** quantity has only magnitude; e.g. time, temperature.
- A **vector** quantity has both magnitude and direction; e.g. displacement, force, wind velocity.

Whereas two scalar quantities are equal if they are represented by the same value, two vector quantities are equal if and only if they have the same **magnitude** and **direction**.

Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ have convenient geometric representations as **position vectors** of points in the 2-dimensional (Cartesian) plane and in 3-dimensional space, respectively.



The vector $\begin{bmatrix} a \\ b \end{bmatrix}$.



The vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Notation

- If P is a point in $\mathbb{R}^3$ with coordinates (x, y, z) we denote this by $P = (x, y, z)$.
- If $P = (x, y, z)$ is a point in $\mathbb{R}^3$, then
$$\overrightarrow{0P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
is often used to denote the position vector of the point.
- Instead of using a capital letter to denote the vector (as we generally do with matrices), we emphasize the importance of the geometry and the direction with an arrow over the name of the vector.

- The notation $\overrightarrow{0P}$ emphasizes that this vector goes from the origin 0 to the point P. We can also use lower case letters for names of vectors. In this case, we write $\overrightarrow{0P} = \mathbf{p}$.
- Any vector
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ in } \mathbb{R}^3$$
is associated with the point (x_1, x_2, x_3).
- Often, there is no distinction made between the vector $\mathbf{x}$ and the point (x_1, x_2, x_3), and we say that both $(x_1, x_2, x_3) \in \mathbb{R}^3$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$.

Facts

Let $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ be vectors in $\mathbb{R}^3$. Then

1. $\mathbf{v} = \mathbf{w}$ if and only if $x = x_1$, $y = y_1$, and $z = z_1$.
2. $||\mathbf{v}|| = \sqrt{x^2 + y^2 + z^2}$.
3. $\mathbf{v} = \mathbf{0}$ if and only if $||\mathbf{v}|| = 0$.
4. For any scalar a, $||a\mathbf{v}|| = |a| \cdot ||\mathbf{v}||$.

Analogous results hold for $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$, i.e.,
$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}, \mathbf{w} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}.$$

In this case, $||\mathbf{v}|| = \sqrt{x^2 + y^2}$.

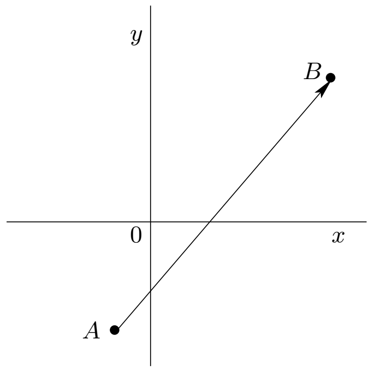
Example

Let $\mathbf{p} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$, and $-2\mathbf{q} = \begin{bmatrix} -6 \\ 2 \\ 4 \end{bmatrix}$

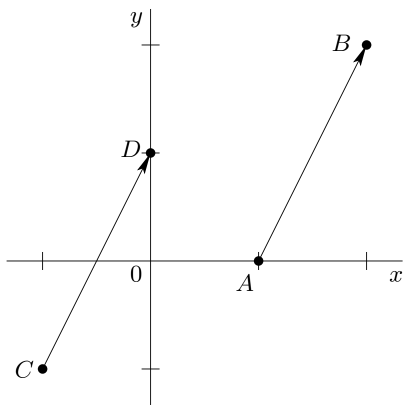
Compute $||\mathbf{p}||$, $||\mathbf{q}||$, and $|| - 2\mathbf{q}||$.

Solution

Geometric Vectors



- $\overrightarrow{AB}$ is the geometric vector from A to B.
- A is the tail of $\overrightarrow{AB}$.
- B is the tip of $\overrightarrow{AB}$.
- the magnitude of $\overrightarrow{AB}$ is its length, and is denoted $||\overrightarrow{AB}||$.



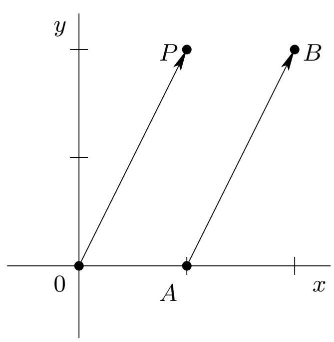
- $\overrightarrow{AB}$ is the vector from $A(1,0)$ to $B(2,2)$.
- $\overrightarrow{CD}$ is the vector from $C(-1,-1)$ to $D(0,1)$.

- $\overrightarrow{AB} = \overrightarrow{CD}$ because the vectors have the same length and direction.

Definition

A vector is in **standard position** if its tail is at the origin.

We co-ordinatize vectors by putting them in standard position, and then identify them with their tips.



Thus $\overrightarrow{AB} = \overrightarrow{0P}$ where $P = P(1,2)$, and we write $\overrightarrow{0P} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \overrightarrow{AB}$.

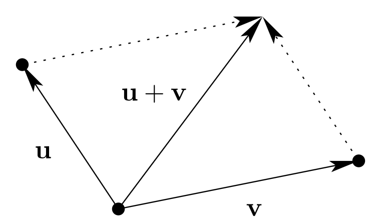
$\overrightarrow{0P}$ is the **position vector** for $P(1,2)$.

More generally, if $P(x, y, z)$ is a point in $\mathbb{R}^3$, then $\overrightarrow{0P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is the position vector for P.

If we aren't concerned with the locations of the tail and tip, we simply write $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Description of Vectors

- **vector equality**: same length and direction.
- **0**: the vector with length zero and no direction.
- **scalar multiplication**: if $\mathbf{v} \neq \mathbf{0}$ and $a \in \mathbb{R}$, $a \neq 0$, then $a\mathbf{v}$ has length $|a| \cdot ||\mathbf{v}||$ and
  - the same direction as $\mathbf{v}$ if $a > 0$;
  - direction opposite to $\mathbf{v}$ if $a < 0$.
- **addition**: $\mathbf{u} + \mathbf{v}$ is the diagonal of the parallelogram defined by $\mathbf{u}$ and $\mathbf{v}$, and having the same tail as $\mathbf{u}$ and $\mathbf{v}$ (parallelogram law).



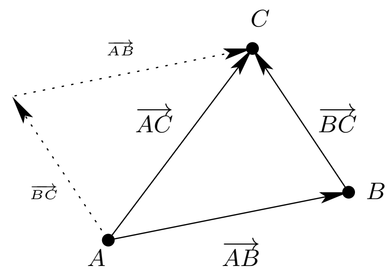
If we have a coordinate system, then

- **vector equality**: $\mathbf{u} = \mathbf{v}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are equal as matrices.
- **0**: has all coordinates equal to zero.
- **scalar multiplication**: $a\mathbf{v}$ is obtained from $\mathbf{v}$ by multiplying each entry of $\mathbf{v}$ by a (matrix scalar multiplication).
- **addition**: $\mathbf{u} + \mathbf{v}$ is represented by the matrix sum of the columns $\mathbf{u}$ and $\mathbf{v}$.

Tip-to-Tail Method for Vector Addition

For points A, B and C,

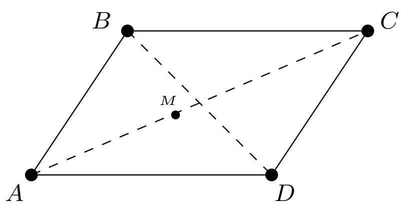
$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$



Example

Show that the diagonals of any parallelogram bisect each other.

Denote the parallelogram by its vertices, $ABCD$.

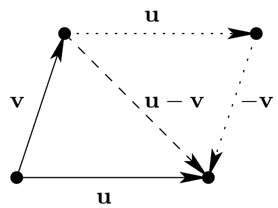


- Let M denote the midpoint of $\overrightarrow{AC}$. Then $\overrightarrow{AM} = \overrightarrow{MC}$.
- It now suffices to show that $\overrightarrow{BM} = \overrightarrow{MD}$.

Solution

Vector Subtraction

- If we have a coordinate system, then subtract the vectors as you would subtract matrices.
- For the intrinsic description:



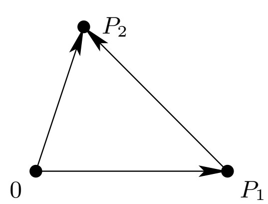
$\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v})$ and is the diagonal from the tip of $\mathbf{v}$ to the tip of $\mathbf{u}$ in the parallelogram defined by $\mathbf{u}$ and $\mathbf{v}$.

Distance between two points in $\mathbb{R}^3$.

Let $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ be two points. Then

1. $$\overrightarrow{P_1P_2}=\begin{bmatrix} x_2-x_1 \\ y_2-y_1 \\ z_2-z_1 \end{bmatrix}.$$

2. The distance between P_1 and P_2 is
$$\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$



$\overrightarrow{0P_1}+\overrightarrow{P_1P_2}=\overrightarrow{0P_2}$, so $\overrightarrow{P_1P_2}=\overrightarrow{0P_2}-\overrightarrow{0P_1}$,
and the distance between P_1 and P_2 is $||\overrightarrow{P_1P_2}||$.

Example

For $P(1,-1,3)$ and $Q(3,1,0)$ Find $\overrightarrow{PQ}$ and $||\overrightarrow{PQ}||$.

Solution

Definition

A **unit vector** is a vector of length one.

Example

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, are examples of unit vectors.

$\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$ is not a unit vector, since $||\mathbf{v}|| = \sqrt{14}$. However,

$$\mathbf{u} = \frac{1}{\sqrt{14}}\mathbf{v} = \begin{bmatrix} \frac{-1}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \end{bmatrix}$$

is a unit vector in the same direction as $\mathbf{v}$, i.e.,

$$||\mathbf{u}|| = \frac{1}{\sqrt{14}}||\mathbf{v}|| = \frac{1}{\sqrt{14}}\sqrt{14} = 1.$$

In general,

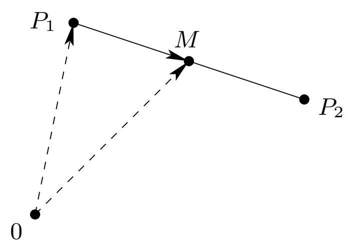
If $\mathbf{v} \neq \mathbf{0}$, then

$$\frac{1}{||\mathbf{v}||}\mathbf{v}$$

is a unit vector in the same direction as $\mathbf{v}$.

Example

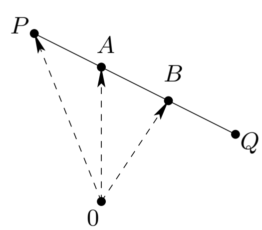
Find the point, M, that is midway between $P_1(-1,-4,3)$ and $P_2(5,0,-3)$.



Solution

Example

Find the two points trisecting the segment between $P(2,3,5)$ and $Q(8,-6,2)$.



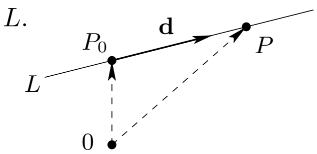
Definition

Two nonzero vectors $\mathbf{v}$ and $\mathbf{w}$ are **parallel** if and only if one is a scalar multiple of the other.

In particular, if $\mathbf{v}$ and $\mathbf{w}$ are nonzero and have the same direction, then $\mathbf{v} = \frac{||\mathbf{v}||}{||\mathbf{w}||}\mathbf{w}$; if $\mathbf{v}$ and $\mathbf{w}$ have opposite directions, then $\mathbf{v} = -\frac{||\mathbf{v}||}{||\mathbf{w}||}\mathbf{w}$.

Equations of Lines

Let L be a line, $P_0(x_0, y_0, z_0)$ a fixed point on L, $P(x, y, z)$ an arbitrary point on L, and $\mathbf{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ a **direction vector** for L, i.e., a vector parallel to L.



Then $\overrightarrow{0P} = \overrightarrow{0P_0} + \overrightarrow{P_0P}$, and $\overrightarrow{P_0P}$ is parallel to $\mathbf{d}$, so $\overrightarrow{P_0P} = t\mathbf{d}$ for some $t \in \mathbb{R}$.

Vector Equation of a Line

$$\overrightarrow{0P} = \overrightarrow{0P_0} + t\mathbf{d}, t \in \mathbb{R}.$$

Notation in the text: $\mathbf{p} = \overrightarrow{0P}$, $\mathbf{p}_0 = \overrightarrow{0P_0}$, so $\mathbf{p} = \mathbf{p}_0 + t\mathbf{d}$.

In component form, this is written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \begin{bmatrix} a \\ b \\ c \end{bmatrix}, t \in \mathbb{R}.$$

Parametric Equations of a Line

$$\begin{array}{rcl} x & = & x_0 + ta \\ y & = & y_0 + tb \\ z & = & z_0 + tc \end{array}, \quad t \in \mathbb{R}.$$

Example

Find an equation for the line through two points $P(2,-1,7)$ and $Q(-3,4,5)$.

Solution

Example

Find an equation for the line through $Q(4,-7,1)$ and parallel to the line

$$L: \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ -5 \\ 3 \end{bmatrix}.$$

Solution

Example

Given two lines L_1 and L_2, find the point of intersection, if it exists.

$$L_1 : \begin{array}{rcl} x &=& 3+t \\ y &=& 1-2t \\ z &=& 3+3t \end{array} \qquad\qquad L_2 : \begin{array}{rcl} x &=& 4+2s \\ y &=& 6+3s \\ z &=& 1+s \end{array}$$

Solution

Discrete Mathematics II | Spring 2022

Section 4.2: Projections and Planes

Lecturer: Mr. Hall | *Math 270B*

Definition

Let $\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ be vectors in $\mathbb{R}^3$. The **dot product** of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} \cdot \mathbf{v} = x_1 x_2 + y_1 y_2 + z_1 z_2,$$

i.e., $\mathbf{u} \cdot \mathbf{v}$ is a scalar.

Note. Another way to think about the dot product is as the 1×1 matrix

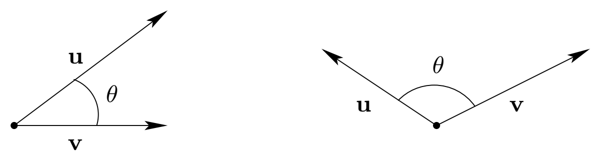
$$\mathbf{u}^T \mathbf{v} = \begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 + y_1 y_2 + z_1 z_2 \end{bmatrix}.$$

Properties of the Dot Product

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors in $\mathbb{R}^3$ (or $\mathbb{R}^2$) and let $k \in \mathbb{R}$.

1. $\mathbf{u} \cdot \mathbf{v}$ is a real number.
2. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. (commutative property)
3. $\mathbf{u} \cdot \mathbf{0} = 0$.
4. $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$.
5. $(k\mathbf{u}) \cdot \mathbf{v} = k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$. (associative property)
6. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$. (distributive properties)
   $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$.

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in $\mathbb{R}^3$ (or $\mathbb{R}^2$). There is a unique angle θ between $\mathbf{u}$ and $\mathbf{v}$ with $0 \le \theta \le \pi$.



Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors, and let θ denote the angle between $\mathbf{u}$ and $\mathbf{v}$. Then

$$\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\ ||\mathbf{v}|| \cos\theta.$$

- If $0 \le \theta < \frac{\pi}{2}$, then $\cos\theta > 0$.
- If $\theta = \frac{\pi}{2}$, then $\cos\theta = 0$.
- If $\frac{\pi}{2} < \theta \le \pi$, then $\cos\theta < 0$.

Therefore, for nonzero vectors $\mathbf{u}$ and $\mathbf{v}$,

- $\mathbf{u} \cdot \mathbf{v} > 0$ if and only if $0 \le \theta < \frac{\pi}{2}$.
- $\mathbf{u} \cdot \mathbf{v} = 0$ if and only if $\theta = \frac{\pi}{2}$.
- $\mathbf{u} \cdot \mathbf{v} < 0$ if and only if $\frac{\pi}{2} < \theta \le \pi$.

Definition

Vectors $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if and only if $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$ or $\theta = \frac{\pi}{2}$.

Thus, vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

Example

Find the angle between $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Solution

Practice

Find the angle between $\mathbf{u} = \begin{bmatrix} 7 \\ -1 \\ 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$.

Solution

Example

Find all vectors $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ orthogonal to both

$$\mathbf{u} = \begin{bmatrix} -1 \\ -3 \\ 2 \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Solution

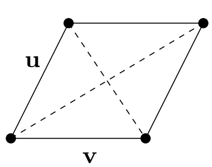
Example

Are $A(4,-7,9)$, $B(6,4,4)$ and $C(7,10,-6)$ the vertices of a right angle triangle?

Solution

Example

A rhombus is a parallelogram with sides of equal length. Prove that the diagonals of a rhombus are perpendicular.



Define the parallelogram (rhombus) by vectors $\mathbf{u}$ and $\mathbf{v}$.

Then the diagonals are $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-\mathbf{v}$.

Show that $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-\mathbf{v}$ are perpendicular.

Solution

Projections

Given nonzero vectors $\mathbf{u}$ and $\mathbf{d}$, express $\mathbf{u}$ as a sum $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, where $\mathbf{u}_1$ is parallel to $\mathbf{d}$ and $\mathbf{u}_2$ is orthogonal to $\mathbf{d}$.

$\mathbf{u}_1$ is **the projection of u onto d**, written $\mathbf{u}_1 = \mathrm{proj}_{\mathbf{d}}\mathbf{u}$.

Since $\mathbf{u}_1$ is parallel to $\mathbf{d}$, $\mathbf{u}_1 = t\mathbf{d}$ for some $t \in \mathbb{R}$.

Furthermore, if $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, then $\mathbf{u}_2 = \mathbf{u} - \mathbf{u}_1$. Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal,

$$\begin{aligned}
\mathbf{u}_2 \cdot \mathbf{u}_1 &= 0 \\
\mathbf{u}_2 \cdot (t\mathbf{d}) &= 0 \\
t(\mathbf{u}_2 \cdot \mathbf{d}) &= 0 \\
\mathbf{u}_2 \cdot \mathbf{d} &= 0 \\
(\mathbf{u} - \mathbf{u_1}) \cdot \mathbf{d} &= 0 \\
\mathbf{u} \cdot \mathbf{d} - \mathbf{u_1} \cdot \mathbf{d} &= 0 \\
\mathbf{u} \cdot \mathbf{d} - (t\mathbf{d}) \cdot \mathbf{d} &= 0 \\
\mathbf{u} \cdot \mathbf{d} - t(\mathbf{d} \cdot \mathbf{d}) &= 0 \\
\mathbf{u} \cdot \mathbf{d} - t||\mathbf{d}||^2 &= 0 \\
\mathbf{u} \cdot \mathbf{d} &= t||\mathbf{d}||^2
\end{aligned}$$

Since $\mathbf{d} \neq \mathbf{0}$, we get

$$t = \frac{\mathbf{u} \cdot \mathbf{d}}{||\mathbf{d}||^2},$$

and therefore

$$\mathbf{u}_1 = \mathrm{proj}_{\mathbf{d}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{d}}{||\mathbf{d}||^2}\mathbf{d}.$$

Summary

Let $\mathbf{u}$ and $\mathbf{d}$ be vectors with $\mathbf{d} \neq \mathbf{0}$.

1.
$$\mathbf{u_1} = \mathrm{proj}_{\mathbf{d}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{d}}{||\mathbf{d}||^2}\mathbf{d}$$
is parallel to $\mathbf{d}$.

2.
$$\mathbf{u_2} = \mathbf{u} - \mathrm{proj}_{\mathbf{d}}\mathbf{u}$$
is orthogonal to $\mathbf{d}$.

Example

Let $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$. Find vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ so that $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, with $\mathbf{u}_1$ parallel to $\mathbf{d}$ and $\mathbf{u}_2$ orthogonal to $\mathbf{d}$.

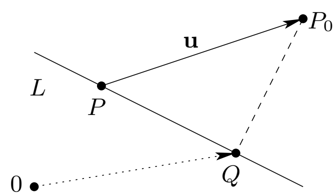
Solution

Example

Let $P_0(3, 2, -1)$ be a point in $\mathbb{R}^3$ and L a line with equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}.$$

Find the shortest distance from P_0 to L, and find the point Q on L that is closest to P_0.



Equations of Planes

Given a point P_0 and a nonzero vector $\mathbf{n}$, there is a unique plane containing P_0 and orthogonal to $\mathbf{n}$.

A nonzero vector $\mathbf{n}$ is a **normal vector** to a plane if and only if $\mathbf{n} \cdot \mathbf{v} = 0$ for every vector $\mathbf{v}$ in the plane.

Consider a plane containing a point P_0 and orthogonal to vector $\mathbf{n}$, and let P be an arbitrary point on this plane.

Then
$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0,$$
or, equivalently,
$$\mathbf{n} \cdot (\overrightarrow{0P} - \overrightarrow{0P_0}) = 0,$$
and is a **vector equation** of the plane.

The vector equation
$$\mathbf{n} \cdot (\overrightarrow{0P} - \overrightarrow{0P_0}) = 0$$
can also be written as
$$\mathbf{n} \cdot \overrightarrow{0P} = \mathbf{n} \cdot \overrightarrow{0P_0}.$$

Now suppose $P_0 = P_0(x_0, y_0, z_0)$, $P = P(x, y, z)$, and $\mathbf{n} = \begin{bmatrix} a & b & c \end{bmatrix}^T$.

Then the previous equation becomes
$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix},$$
so
$$ax + by + cz = ax_0 + by_0 + cz_0,$$
which is a **scalar equation** of a plane.

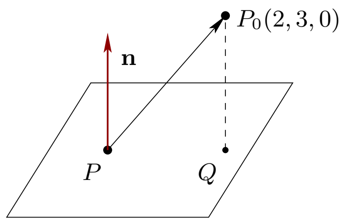
Example

Find an equation of the plane containing $P_0(1, -1, 0)$ and orthogonal to $\mathbf{n} = \begin{bmatrix} -3 & 5 & 2 \end{bmatrix}^T$.

Solution

Example

Find the shortest distance from the point $P_0(2,3,0)$ to the plane with equation $5x + y + z = -1$, and find the point Q on the plane that is closest to P_0.



Solution

The Cross Product

Let $\mathbf{u} = \begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix}^T$ and $\mathbf{v} = \begin{bmatrix} x_2 & y_2 & z_2 \end{bmatrix}^T$. Then

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} y_1 z_2 - z_1 y_2 \\ -(x_1 z_2 - z_1 x_2) \\ x_1 y_2 - y_1 x_2 \end{bmatrix}.$$

Note. $\mathbf{u} \times \mathbf{v}$ is a vector that is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

Cofactor expansion (down column 1):

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & x_1 & x_2 \\ \mathbf{j} & y_1 & y_2 \\ \mathbf{k} & z_1 & z_2 \end{vmatrix}, \text{ where } \mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Facts

Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$.

1. $\mathbf{v} \times \mathbf{w}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.
2. If $\mathbf{v}$ and $\mathbf{w}$ are both nonzero, then $\mathbf{u} \times \mathbf{w} = \mathbf{0}$ if and only if $\mathbf{v}$ and $\mathbf{w}$ are parallel.

Example

Find all vectors orthogonal to both $\mathbf{u} = \begin{bmatrix} -1 & -3 & 2 \end{bmatrix}^T$ and $\mathbf{v} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$.

(We previously solved this using the dot product.)

Solution

Discrete Mathematics II	Spring 2022

Section 4.3: More on the Cross Product

Lecturer: Mr. Hall	*Math 270B*

Theorem

If $\mathbf{u} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$. Then

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ z_0 & z_1 & z_2 \end{bmatrix}.$$

Shorthand: $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix}$.

Proof

Let $\mathbf{u} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$.

Then

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \cdot \begin{bmatrix} y_1 z_2 - z_1 y_2 \\ -(x_1 z_2 - z_1 x_2) \\ x_1 y_2 - y_1 x_2 \end{bmatrix} \\ &= x_0(y_1 z_2 - z_1 y_2) - y_0(x_1 z_2 - z_1 x_2) + z_0(x_1 y_2 - y_1 x_2) \\ &= x_0 \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} - y_0 \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} + z_0 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \\ &= \begin{vmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ z_0 & z_1 & z_2 \end{vmatrix}. \end{aligned}$$

Properties of the Cross Product

Let $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ be in $\mathbb{R}^3$.

1. $\mathbf{u} \times \mathbf{v}$ is a vector.
2. $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.
3. $\mathbf{u} \times \mathbf{0} = \mathbf{0}$ and $\mathbf{0} \times \mathbf{u} = \mathbf{0}$.
4. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.
5. $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$.
6. $(k\mathbf{u}) \times \mathbf{v} = k(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \times (k\mathbf{v})$ for any scalar k.
7. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$.
8. $(\mathbf{v} + \mathbf{w}) \times \mathbf{u} = \mathbf{v} \times \mathbf{u} + \mathbf{w} \times \mathbf{u}$.

The Lagrange Identity

If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, then

$$||\mathbf{u} \times \mathbf{v}||^2 = ||\mathbf{u}||^2||\mathbf{v}||^2 - (\mathbf{u} \cdot \mathbf{v})^2.$$

As a consequence of the Lagrange Identity and the fact that

$$\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\ ||\mathbf{v}|| \cos\theta,$$

we have

$$\begin{aligned} ||\mathbf{u} \times \mathbf{v}||^2 &= ||\mathbf{u}||^2||\mathbf{v}||^2 - (\mathbf{u} \cdot \mathbf{v})^2 \\ &= ||\mathbf{u}||^2||\mathbf{v}||^2 - ||\mathbf{u}||^2||\mathbf{v}||^2 \cos^2\theta \\ &= ||\mathbf{u}||^2||\mathbf{v}||^2(1 - \cos^2\theta) \\ &= ||\mathbf{u}||^2||\mathbf{v}||^2 \sin^2\theta. \end{aligned}$$

Taking square roots of both sides yields,

$$||\mathbf{u} \times \mathbf{v}|| = ||\mathbf{u}||\ ||\mathbf{v}|| \sin\theta.$$

Note that since $0 \le \theta \le \pi$, $\sin\theta \ge 0$.

If $\theta = 0$ or $\theta = \pi$, then $\sin\theta = 0$, and $||\mathbf{u} \times \mathbf{v}|| = 0$. This is consistent with our earlier observation that if $\mathbf{u}$ and $\mathbf{v}$ are parallel, then $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.

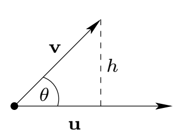
Theorem

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^3$, and let θ denote the angle between $\mathbf{u}$ and $\mathbf{v}$.

1. $||\mathbf{u} \times \mathbf{v}|| = ||\mathbf{u}||\ ||\mathbf{v}|| \sin\theta$, and is the area of the parallelogram defined by $\mathbf{u}$ and $\mathbf{v}$.
2. $\mathbf{u}$ and $\mathbf{v}$ are parallel if and only if $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.

Proof

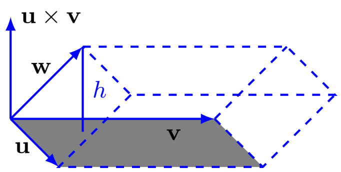
The area of the parallelogram defined by $\mathbf{u}$ and $\mathbf{v}$ is $||\mathbf{u}||h$, where h is the height of the parallelogram.



$\sin\theta = \frac{h}{||\mathbf{v}||}$, implying that $h = ||\mathbf{v}|| \sin\theta$. Therefore, the area is

$$||\mathbf{u}||\ ||\mathbf{v}|| \sin\theta.$$

Volume of a Parallelepiped



If three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are given, they determine a "squashed" rectangular solid called a **parallelepiped**, and it is often useful to be able to find the volume of such a solid. The base of the solid is the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$, so it has area $A = \|\mathbf{u} \times \mathbf{v}\|$. The height of the solid is the length h of the projection of $\mathbf{w}$ on $\mathbf{u} \times \mathbf{v}$. Hence

$$h = \left|\frac{\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})}{\|\mathbf{u} \times \mathbf{v}\|^2}\right| \|\mathbf{u} \times \mathbf{v}\| = \frac{|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|}{\|\mathbf{u} \times \mathbf{v}\|} = \frac{|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|}{A}$$

Thus the volume of the parallelepiped is $hA = |\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|$.

This proves that the volume of the parallelepiped determined by three vectors $\mathbf{w}$, $\mathbf{u}$, and $\mathbf{v}$ is given by $|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|$.

Example

Find the area of the triangle having vertices $A(3, -1, 2)$, $B(1, 1, 0)$ and $C(1, 2, -1)$.

Solution

Example

Find the volume of the parallelepiped determined by the vectors $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$.

Solution

Example

If $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are the coordinate vectors, verify that $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

Solution

Example

Show that $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ need not equal $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$ by calculating both when

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution

Example

Find two unit vectors orthogonal to both $\mathbf{u}$ and $\mathbf{v}$ if: $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$

Solution

Discrete Mathematics II | Spring 2022

Section 4.4: Linear Operators on $\mathbb{R}^3$

Lecturer: Mr. Hall | *Math 270B*

Definition

Recall that a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called *linear* if $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(a\mathbf{x}) = aT(\mathbf{x})$ holds for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and all scalars a. In this case, we showed that there exists an $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$, and we say that T is the **matrix transformation induced** by A.

Definition

A linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

is called a **linear operator** on $\mathbb{R}^n$.

In this section we investigate : Rotations about a line through the origin, reflections in a plane through the origin, and projections onto a plane or line through the origin in $\mathbb{R}^3$. In every case we show that the operator is linear, and we find the matrices of all the reflections and projections. To do this we must prove that these reflections, projections, and rotations are actually *linear* operators on $\mathbb{R}^3$. In the case of reflections and rotations, it is convenient to examine a more general situation. A transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be **distance preserving** if the distance between $T(\mathbf{v})$ and $T(\mathbf{w})$ is the same as the distance between $\mathbf{v}$ and $\mathbf{w}$ for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$; that is,

$$\|T(\mathbf{v}) - T(\mathbf{w})\| = \|\mathbf{v} - \mathbf{w}\| \text{ for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } \mathbb{R}^3 \tag{1}$$

Clearly reflections and rotations are distance preserving, and both carry $\mathbf{0}$ to $\mathbf{0}$, so the following theorem shows that they are both linear.

Theorem

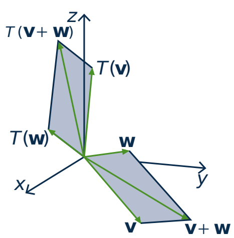
If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is distance preserving, and if $T(\mathbf{0}) = \mathbf{0}$, then T is linear.

Proof

Since $T(\mathbf{0}) = \mathbf{0}$, taking $\mathbf{w} = \mathbf{0}$ in (1) shows that $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v}$ in $\mathbb{R}^3$, that is T preserves length.
Also, $\|T(\mathbf{v}) - T(\mathbf{w})\|^2 = \|\mathbf{v} - \mathbf{w}\|^2$ by (1).
Since $\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 - 2\mathbf{v} \cdot \mathbf{w} + \|\mathbf{w}\|^2$ always holds, it follows that $T(\mathbf{v}) \cdot T(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{v}$ and $\mathbf{w}$.
Hence, the angle between $T(\mathbf{v})$ and $T(\mathbf{w})$ is the same as the angle between $\mathbf{v}$ and $\mathbf{w}$ for all (nonzero) vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$.

With this we can show that T is linear.
Given nonzero vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$, the vector $\mathbf{v} + \mathbf{w}$ is the diagonal of the parallelogram determined by $\mathbf{v}$ and $\mathbf{w}$. By the preceding paragraph, the effect of T is to carry this *entire parallelogram* to the parallelogram determined by $T(\mathbf{v})$ and $T(\mathbf{w})$, with diagonal $T(\mathbf{v}+\mathbf{w})$. But this diagonal is $T(\mathbf{v})+T(\mathbf{w})$ by the parallelogram law (see the Figure).

In other words, $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$. A similar argument shows that $T(a\mathbf{v}) = aT(\mathbf{v})$ for all scalars a, proving that T is indeed linear.

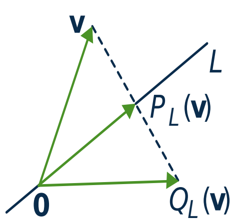


Projetions

Let L denote a line through the origin in $\mathbb{R}^3$.
Given a vector $\mathbf{v}$ in $\mathbb{R}^3$, the reflection $Q_L(\mathbf{v})$ of $\mathbf{v}$ in L and the projection $P_L(\mathbf{v})$ of $\mathbf{v}$ on L are defined in the Figure. In the same Figure, we see that

$$P_L(\mathbf{v}) = \mathbf{v} + \frac{1}{2}[Q_L(\mathbf{v}) - \mathbf{v}] = \frac{1}{2}[Q_L(\mathbf{v}) + \mathbf{v}] \qquad (2)$$

so the fact that Q_L is linear shows that P_L is also linear.



We saw in Section 4.2 how to get the matrix of P_L directly.

In fact, if $\mathbf{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \mathbf{0}$ is a direction vector for L, and we write $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, then

$$\begin{aligned} P_L(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\mathbf{d} &= \frac{ax+by+cz}{a^2+b^2+c^2}\begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= \frac{1}{a^2+b^2+c^2}\begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

Note that this shows directly that P_L is a matrix transformation and so gives another proof that it is linear.

Theorem

Let L denote the line through the origin in $\mathbb{R}^3$ with direction vector $\mathbf{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \mathbf{0}$. Then P_L and Q_L are both linear and

$$P_L \text{ has matrix } \frac{1}{a^2+b^2+c^2}\begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$$

$$Q_L \text{ has matrix } \frac{1}{a^2+b^2+c^2}\begin{bmatrix} a^2-b^2-c^2 & 2ab & 2ac \\ 2ab & b^2-a^2-c^2 & 2bc \\ 2ac & 2bc & c^2-a^2-b^2 \end{bmatrix}$$

Proof

It remains to find the matrix of Q_L. But (2) implies that $Q_L(\mathbf{v}) = 2P_L(\mathbf{v}) - \mathbf{v}$ for each $\mathbf{v}$ in $\mathbb{R}^3$, so if $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ we obtain (with some matrix arithmetic):

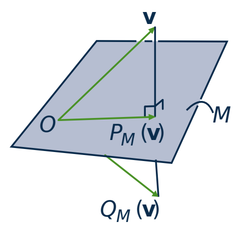
$$\begin{aligned} Q_L(\mathbf{v}) &= \left\{\frac{2}{a^2+b^2+c^2}\begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \frac{1}{a^2+b^2+c^2}\begin{bmatrix} a^2-b^2-c^2 & 2ab & 2ac \\ 2ab & b^2-a^2-c^2 & 2bc \\ 2ac & 2bc & c^2-a^2-b^2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

as required.

Reflections

In $\mathbb{R}^3$ we can reflect in planes as well as lines. Let M denote a plane through the origin in $\mathbb{R}^3$. Given a vector $\mathbf{v}$ in $\mathbb{R}^3$, the reflection $Q_M(\mathbf{v})$ of $\mathbf{v}$ in M and the projection $P_M(\mathbf{v})$ of $\mathbf{v}$ on M are defined in the Figure. As before, we have

$$P_M(\mathbf{v}) = \mathbf{v} + \frac{1}{2}[Q_M(\mathbf{v}) - \mathbf{v}] = \frac{1}{2}[Q_M(\mathbf{v}) + \mathbf{v}]$$



so the fact that Q_M is linear shows that P_M is also linear.

Again we can obtain the matrix directly. If $\mathbf{n}$ is a normal for the plane M, then the Figure shows that

$$P_M(\mathbf{v}) = \mathbf{v} - \text{proj}_{\mathbf{n}}\mathbf{v} = \mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n} \text{ for all vectors } \mathbf{v}.$$

If $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \mathbf{0}$ and $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, a computation like the above gives

$$\begin{aligned} P_M(\mathbf{v}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \frac{ax+by+cz}{a^2+b^2+c^2}\begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= \frac{1}{a^2+b^2+c^2}\begin{bmatrix} b^2+c^2 & -ab & -ac \\ -ab & a^2+c^2 & -bc \\ -ac & -bc & b^2+c^2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

Theorem

Let M denote the plane through the origin in $\mathbb{R}^3$ with normal $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \mathbf{0}$. Then P_M and Q_M are both linear and

$$P_M \text{ has matrix } \frac{1}{a^2+b^2+c^2}\begin{bmatrix} b^2+c^2 & -ab & -ac \\ -ab & a^2+c^2 & -bc \\ -ac & -bc & a^2+b^2 \end{bmatrix}$$

$$Q_M \text{ has matrix } \frac{1}{a^2+b^2+c^2}\begin{bmatrix} b^2+c^2-a^2 & -2ab & -2ac \\ -2ab & a^2+c^2-b^2 & -2bc \\ -2ac & -2bc & a^2+b^2-c^2 \end{bmatrix}$$

Since $Q_M(\mathbf{v}) = 2P_M(\mathbf{v}) - \mathbf{v}$ for each $\mathbf{v} \in \mathbb{R}^3$, the computation is similar to that for Q_L.

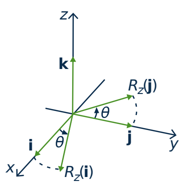
Rotations

Let $R_{z,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ denote rotation of $\mathbb{R}^3$ about the z axis through an angle θ from the positive x axis toward the positive y axis. Show that $R_{z,\theta}$ is linear and find its matrix.

First, R is distance preserving and so is linear. Hence from Section 2.6 we obtain the matrix of $R_{z,\theta}$.

Let $\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ denote the standard basis of $\mathbb{R}^3$; we must find $R_{z,\theta}(\mathbf{i})$, $R_{z,\theta}(\mathbf{j})$, and $R_{z,\theta}(\mathbf{k})$. Clearly $R_{z,\theta}(\mathbf{k}) = \mathbf{k}$. The effect of $R_{z,\theta}$ on the x-y plane is to rotate it counterclockwise through the angle θ. Hence the Figure gives

$$R_{z,\theta}(\mathbf{i}) = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \quad R_{z,\theta}(\mathbf{j}) = \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix}$$



so, by Section 2.6, $R_{z,\theta}$ has matrix

$$\begin{bmatrix} R_{z,\theta}(\mathbf{i}) & R_{z,\theta}(\mathbf{j}) & R_{z,\theta}(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In each case solve the problem by finding the matrix of the operator.

Example

Find the projection of $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}$ on the line with equation $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}$.

Solution

Example

Find the reflection of $\mathbf{v} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$ in the line with equation $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$.

Solution

Example

Find the projection of $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ on the plane with equation $3x - 5y + 2z = 0$.

Solution

Example

Find the reflection of $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ in the plane with equation $x - y + 3z = 0$.

Solution

Section 5.1: Subspaces and Spanning

Definitions

1. $\mathbb{R}$ denotes the set of **real** numbers, and is an example of a set of **scalars**.
2. $\mathbb{R}^n$ is the set of all n-tuples of real numbers, i.e.,
$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, 1 \leq i \leq n\}.$$
3. The **vector space** $\mathbb{R}^n$ consists of the set $\mathbb{R}^n$ written as **column matrices**, along with the (matrix) operations of addition and scalar multiplication. Unless stated otherwise, $\mathbb{R}^n$ means the vector space $\mathbb{R}^n$.

We are interested in *nice* subsets of $\mathbb{R}^n$, defined as follows.

Definition

A subset U of $\mathbb{R}^n$ is a **subspace** of $\mathbb{R}^n$ if

S1. The zero vector of $\mathbb{R}^n$, $\mathbf{0}_n$, is in U;

S2. U is closed under addition, i.e., for all $\mathbf{u}, \mathbf{w} \in U$, $\mathbf{u} + \mathbf{w} \in U$;

S3. U is closed under scalar multiplication, i.e., for all $\mathbf{u} \in U$ and $k \in \mathbb{R}$, $k\mathbf{u} \in U$.

The subset $U = \{\mathbf{0}_n\}$ is a subspace of $\mathbb{R}^n$, as is the set $\mathbb{R}^n$ itself. Any other subspace of $\mathbb{R}^n$ is a **proper subspace** of $\mathbb{R}^n$.

If U is a subset of $\mathbb{R}^n$, we write $U \subseteq \mathbb{R}^n$.

Example

In $\mathbb{R}^3$, the line L through the origin that is parallel to the vector $\mathbf{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \mathbf{0}$ has (vector) equation $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $t \in \mathbb{R}$, so

$$L = \{t\mathbf{d} \mid t \in \mathbb{R}\}.$$

Prove that L is a subspace of $\mathbb{R}^3$.

Proof

Example

In $\mathbb{R}^3$, let M denote the plane through the origin having equation $ax + by + cz = 0$; then M has normal vector $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. If $\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, then

$$M = \left\{ \mathbf{u} \in \mathbb{R}^3 \mid \mathbf{n} \bullet \mathbf{u} = 0 \right\},$$

where $\mathbf{n} \bullet \mathbf{u}$ is the dot product of vectors $\mathbf{n}$ and $\mathbf{u}$.

Prove that M is a subspace of $\mathbb{R}^3$.

Proof

Example

Let $U = \left\{ \left[\begin{array}{c} a \\ b \\ c \\ d \end{array} \right] \;\middle|\; a, b, c, d \in \mathbb{R} \text{ and } 2a - b = c + 2d \right\}$

Prove that U is a subspace of $\mathbb{R}^4$.

Proof

Example

Is $U = \left\{ \begin{bmatrix} 1 \\ s \\ t \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$ a subspace of $\mathbb{R}^3$? Justify your answer.

Solution

Example

Is $U = \left\{ \begin{bmatrix} r \\ 0 \\ s \end{bmatrix} \;\middle|\; r, s \in \mathbb{R} \text{ and } r^2 + s^2 = 0 \right\}$ a subspace of $\mathbb{R}^3$? Justify your answer.

Solution

Definitions

Let A be an $m \times n$ matrix. The **null space** of A is defined as

$$\text{null}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}_m\},$$

and the **image space** of A is defined as

$$\text{im}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}.$$

In the language of Chapter 2, null(A) consists of all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$, and im(A) is the set of all vectors $\mathbf{y}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{y}$ *has* a solution $\mathbf{x}$. Note that $\mathbf{x}$ is in null(A) if it satisfies the *condition* $A\mathbf{x} = \mathbf{0}$, while im(A) consists of vectors of the *form* $A\mathbf{x}$ for some $\mathbf{x}$ in $\mathbb{R}^n$.

Example

Prove that if A is an $m \times n$ matrix, then null(A) is a subspace of $\mathbb{R}^n$.

Proof

Example

Prove that if A is an $m \times n$ matrix, then $\text{im}(A)$ is a subspace of $\mathbb{R}^m$.

Proof

Definition

Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. The **eigenspace of A corresponding to λ** is the set

$$E_\lambda(A) = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

Note that

$$\begin{aligned} E_\lambda(A) &= \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}, \\ &= \{\mathbf{v} \in \mathbb{R}^n \mid \lambda\mathbf{v} - A\mathbf{v} = \mathbf{0}_n\} \\ &= \{\mathbf{v} \in \mathbb{R}^n \mid (\lambda I - A)\mathbf{v} = \mathbf{0}_n\} \end{aligned}$$

showing that

$$E_\lambda(A) = \text{null}(\lambda I - A).$$

It follows that

- if λ is an eigenvalue of A, then $E_\lambda(A) \neq \{\mathbf{0}_n\}$;
- the nonzero vectors of $E_\lambda(A)$ are the eigenvectors of A corresponding to λ;
- the eigenspace of A corresponding to λ is a subspace of $\mathbb{R}^n$.

Definition

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ and $t_1, t_2, \ldots, t_k \in \mathbb{R}$. Then the vector

$$\mathbf{x} = t_1\mathbf{x_1} + t_2\mathbf{x_2} + \cdots + t_k\mathbf{x_k}$$

is called a **linear combination** of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$; the (scalars) $t_1, t_2, \ldots, t_k \in \mathbb{R}$ are the **coefficients**.

The set of *all* linear combinations of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}$ is called **the span of** $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, and is written

$$\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}\} = \{t_1\mathbf{x_1} + t_2\mathbf{x_2} + \cdots + t_k\mathbf{x_k} \mid t_1, t_2, \ldots, t_k \in \mathbb{R}\}.$$

Additional Terminology. If $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}\}$, then

- U **is spanned by** the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}$.
- the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}$ **span** U.
- the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_k}\}$ is a **spanning set** for U.

Example

Let $\mathbf{x} \in \mathbb{R}^3$ be a nonzero vector. Then $\text{span}\{\mathbf{x}\} = \{k\mathbf{x} \mid k \in \mathbb{R}\}$ is a line through the origin having direction vector $\mathbf{x}$.

Example

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ be nonzero vectors that are not parallel. Then

$$\text{span}\{\mathbf{x}, \mathbf{y}\} = \{k\mathbf{x} + t\mathbf{y} \mid k, t \in \mathbb{R}\}$$

is a plane through the origin containing $\mathbf{x}$ and $\mathbf{y}$.

Example

Let $\mathbf{x} = \begin{bmatrix} 8 \\ 3 \\ -13 \\ 20 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ -3 \\ 5 \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} -1 \\ 0 \\ 2 \\ -3 \end{bmatrix}$. Is $\mathbf{x} \in \text{span}\{\mathbf{y}, \mathbf{z}\}$?

Solution

Practice

Let $\mathbf{w} = \begin{bmatrix} 8 \\ 3 \\ -13 \\ 21 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ -3 \\ 5 \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} -1 \\ 0 \\ 2 \\ -3 \end{bmatrix}$. Is $\mathbf{w} \in \text{span}\{\mathbf{y}, \mathbf{z}\}$?

Solution

Theorem

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ and let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$.

Then

1. U is a subspace of $\mathbb{R}^n$ containing each $\mathbf{x_i}$, $1 \leq i \leq k$;
2. if W is a subspace of $\mathbb{R}^n$ and $\mathbf{x_1}, \mathbf{x}_2, \ldots, \mathbf{x}_k \in W$, then $U \subseteq W$.
   (This is saying that U is the "smallest" subspace of $\mathbb{R}^n$ that contains $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$.)

Proof

1. Since $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ and $0\mathbf{x_1} + 0\mathbf{x_2} + \cdots + 0\mathbf{x_k} = \mathbf{0}_n$, $\mathbf{0}_n \in U$.

   Suppose $\mathbf{x}, \mathbf{y} \in U$. Then $\mathbf{x} = s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k}$ and $\mathbf{y} = t_1\mathbf{x_1} + t_2\mathbf{x_2} + \cdots + t_k\mathbf{x_k}$ for some $s_i, t_i \in \mathbb{R}$, $1 \leq i \leq k$. Thus
   $$\begin{aligned} \mathbf{x} + \mathbf{y} &= (s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k}) + (t_1\mathbf{x_1} + t_2\mathbf{x_2} + \cdots + t_k\mathbf{x_k}) \\ &= (s_1 + t_1)\mathbf{x_1} + (s_2 + t_2)\mathbf{x_2} + \cdots + (s_k + t_k)\mathbf{x_k}. \end{aligned}$$
   Since $s_i + t_i \in \mathbb{R}$ for all $1 \leq i \leq k$, $\mathbf{x} + \mathbf{y} \in U$, i.e., U is closed under addition.

   Suppose $\mathbf{x} \in U$ and $a \in \mathbb{R}$. Then $\mathbf{x} = s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k}$ for some $s_i \in \mathbb{R}$, $1 \leq i \leq k$. Thus
   $$\begin{aligned} a\mathbf{x} &= a(s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k}) \\ &= (as_1)\mathbf{x_1} + (as_2)\mathbf{x_2} + \cdots + (as_k)\mathbf{x_k}. \end{aligned}$$
   Since $as_i \in \mathbb{R}$ for all $1 \leq i \leq k$, $a\mathbf{x} \in U$ and U is closed under scalar multiplication.
   Therefore, U is a subspace of $\mathbb{R}^n$. Furthermore, since
   $$\mathbf{x_i} = \sum_{j=1}^{i-1} 0\mathbf{x_j} + 1\mathbf{x_i} + \sum_{j=i+1}^{k} 0\mathbf{x_j},$$
   it follows that $\mathbf{x_i} \in U$ for all i, $1 \leq i \leq k$.
2. To prove that $U \subseteq W$, prove that if $\mathbf{x} \in U$, then $\mathbf{x} \in W$.
   Suppose $\mathbf{x} \in U$. Then $\mathbf{x} = s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k}$ for some $s_i \in \mathbb{R}$, $1 \leq i \leq k$. Since W contains each $\mathbf{x_i}$ and W is closed under scalar multiplication, it follows that $s_i\mathbf{x_i} \in W$ for each i, $1 \leq i \leq k$. Furthermore, since W is closed under addition, $\mathbf{x} = s_1\mathbf{x_1} + s_2\mathbf{x_2} + \cdots + s_k\mathbf{x_k} \in W$. Therefore, $U \subseteq W$.

Example

Is $U = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \;\middle|\; a, b, c, d \in \mathbb{R} \text{ and } 2a - b = c + 2d \right\}$ a subspace of $\mathbb{R}^4$? Justify your answer.

Solution

Example

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $U_1 = \operatorname{span}\{\mathbf{x}, \mathbf{y}\}$, and $U_2 = \operatorname{span}\{2\mathbf{x} - \mathbf{y}, 2\mathbf{y} + \mathbf{x}\}$. Prove that $U_1 = U_2$.

Proof

Definition

Let $\mathbf{e}_j$ denote the j^{th} column of I_n, the $n \times n$ identity matrix; $\mathbf{e}_j$ is called the j^{th} **coordinate vector** of $\mathbb{R}^n$.

Claim

$\mathbb{R}^n = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

Proof

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$. Then $\mathbf{x} = x_1\mathbf{e_1} + x_2\mathbf{e_2} + \cdots + x_n\mathbf{e_n}$, where $x_1, x_2, \ldots, x_n \in \mathbb{R}$. Therefore, $\mathbf{x} \in \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, and thus $\mathbb{R}^n \subseteq \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

Conversely, since $\mathbf{e}_i \in \mathbb{R}^n$ for each i, $1 \leq i \leq n$ (and $\mathbb{R}^n$ is a vector space), it follows that $\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \subseteq \mathbb{R}^n$. The equality now follows.

Example

Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

Does $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ span $\mathbb{R}^4$? (Equivalently, is $\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\} = \mathbb{R}^4$?)

Solution

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

Show that $\operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} \neq \mathbb{R}^4$.

Solution

If A is an $m \times n$ matrix A, the next two examples show that it is a routine matter to find spanning sets for null$A$ and im$A$.

Example

Given an $m \times n$ matrix A, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ denote the basic solutions to the system $A\mathbf{x} = \mathbf{0}$ given by the gaussian algorithm. Then

$$\text{null } A = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$$

Solution

If $\mathbf{x}$ is in null$A$, then $A\mathbf{x} = \mathbf{0}$ so $\mathbf{x}$ is a linear combination of the basic solutions;
that is, null$A \subseteq \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$. On the other hand, if $\mathbf{x}$ is in span$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$,
then $\mathbf{x} = t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k$ for scalars t_i, so

$$A\mathbf{x} = t_1A\mathbf{x}_1 + t_2A\mathbf{x}_2 + \cdots + t_kA\mathbf{x}_k = t_1\mathbf{0} + t_2\mathbf{0} + \cdots + t_k\mathbf{0} = \mathbf{0}$$

This shows that $\mathbf{x}$ is in null$A$, and hence that span$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subseteq$ null$A$. Thus we have equality.

Example

Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ denote the columns of the $m \times n$ matrix A. Then

$$\text{im } A = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$$

Solution

If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, observe that

$$\begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \cdots & A\mathbf{e}_n \end{bmatrix} = A\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix} = AI_n = A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots \mathbf{c}_n \end{bmatrix}.$$

Hence $\mathbf{c}_i = A\mathbf{e}_i$ is in im$A$ for each i, so span$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \subseteq$ im$A$.

Conversely, let $\mathbf{y}$ be in im$A$, say $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x}$ in $\mathbb{R}^n$. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then by definition

$$\mathbf{y} = A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n \text{ is in span}\{\mathbf{c}_1, \mathbf{c}_2, \ \ldots, \mathbf{c}_n\}$$

This shows that im$A \subseteq \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$, and the result follows.

Discrete Mathematics II	Spring 2022

Section 5.2: Independence and Dimension

Lecturer: Mr. Hall | *Math 270B*

Definition

Let $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be a subset of $\mathbb{R}^n$. The set S is **linearly independent** (or simply independent) if the following condition is satisfied:

if $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}_n$, then $t_1 = t_2 = \cdots = t_k = 0$,

i.e., the only linear combination of vectors of S that vanishes (is equal to the zero vector) is the trivial one (all coefficients equal to zero).

A set that is **not** linearly independent is called **dependent**.

Example

Is $S = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\}$ linearly independent?

Solution

Example

Consider the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \subseteq \mathbb{R}^n$, and suppose $t_1, t_2, \ldots, t_n \in \mathbb{R}$ are such that

$$t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \cdots t_n\mathbf{e}_n = \mathbf{0}_n.$$

Since

$$t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \cdots t_n\mathbf{e}_n = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix},$$

the only linear combination that vanishes is the trivial one, i.e., the one with $t_1 = t_2 = \cdots = t_n = 0$. Therefore, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent.

Example

Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be an independent subset of $\mathbb{R}^n$. Is $\{\mathbf{u} + \mathbf{v}, 2\mathbf{u} + \mathbf{w}, \mathbf{v} - 5\mathbf{w}\}$ linearly independent?

Solution

Example

Let $X \subseteq \mathbb{R}^n$ and suppose that $\mathbf{0}_n \in X$. Show that X is linearly dependent.

Solution

Let $X = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}\}$ for some $k \geq 1$, and suppose $\mathbf{x}_1 = \mathbf{0_n}$. Then

$$1\mathbf{x}_1 + 0\mathbf{x}_2 + \cdots + 0\mathbf{x}_k = 1 \cdot \mathbf{0} + 0\mathbf{x}_2 + \cdots + 0\mathbf{x}_k = \mathbf{0},$$

i.e., we have found a nontrivial linear combination of the vectors of X that vanishes. Therefore, X is dependent. This means that the zero vector of $\mathbb{R}^n$ cannot belong to any independent set.

Theorem

Let $U = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subseteq \mathbb{R}^n$ be an independent set. Then any vector $\mathbf{x} \in \text{span}(U)$ has a unique representation as a linear combination of vectors of U.

Proof

Suppose that there is a vector $\mathbf{x} \in \text{span}(U)$ such that

$$\begin{aligned} \mathbf{x} &= s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_k\mathbf{v}_k, \text{ for some } s_1, s_2, \ldots, s_k \in \mathbb{R}, \text{ and} \\ \mathbf{x} &= t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k, \text{ for some } t_1, t_2, \ldots, t_k \in \mathbb{R}. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{0}_n &= \mathbf{x} - \mathbf{x} \\ &= (s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_k\mathbf{v}_k) - (t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k) \\ &= (s_1 - t_1)\mathbf{v}_1 + (s_2 - t_2)\mathbf{v}_2 + \cdots + (s_k - t_k)\mathbf{v}_k. \end{aligned}$$

Since U is independent, the only linear combination that vanishes is the trivial one, so $s_i - t_i = 0$ for all i, $1 \leq i \leq k$.

Therefore, $s_i = t_i$ for all i, $1 \leq i \leq k$, and the representation is unique.

Example

Suppose that $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$. Prove that $\{\mathbf{u}, \mathbf{v}\}$ is dependent if and only if $\mathbf{u}$ and $\mathbf{v}$ are parallel.

Proof

Example

Suppose that $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^3$, and that $\{\mathbf{v}, \mathbf{w}\}$ is independent. Prove that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent if and only if $\mathbf{u} \notin \operatorname{span}\{\mathbf{v}, \mathbf{w}\}$.

Proof

Facts

Suppose A is an $m \times n$ matrix with columns $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n \in \mathbb{R}^m$. Then

1. $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is linearly independent if and only if $A\mathbf{x} = \mathbf{0}_m$ with $\mathbf{x} \in \mathbb{R}^n$ implies $\mathbf{x} = \mathbf{0}_n$.
2. $\mathbb{R}^m = \operatorname{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ if and only if $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^m$.

These are useful facts because:

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$.

1. Are $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ linearly independent?
2. Do $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ span $\mathbb{R}^n$?

To answer both questions, simply let A be a matrix whose columns are the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$. Find R, a row-echelon form of A.

- The answer to the first question is "yes" if and only if each column of R has a leading one.
- The answer to the second question is "yes" if and only if each row of R has a leading one.

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

Show that $\operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} \neq \mathbb{R}^4$.

Solution

The Invertible Matrix Theorem

Let A be an $n \times n$ matrix. The following are equivalent.

1. A is invertible.
2. The columns of A are independent.
3. The columns of A span $\mathbb{R}^n$.
4. The rows of A are independent, i.e., the columns of A^T are independent.
5. The rows of A span the set of all $1 \times n$ rows, i.e., the columns of A^T span $\mathbb{R}^n$.

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

Show that $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} \neq \mathbb{R}^4$.

Solution

Practice

Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix}.$$

Is $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ independent?

Solution

Bases and Dimension

Let U be a subspace of $\mathbb{R}^n$ that is spanned by m vectors. If U contains a subset of k linearly independent vectors, then $k \leq m$.

Definition

Let U be a subspace of $\mathbb{R}^n$. A set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is a **basis** of U if

1. $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent;
2. $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$.

As a consequence of all this, if $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is a basis of a subspace U, then every $\mathbf{u} \in U$ has a unique representation as a linear combination of the vectors $\mathbf{x}_i$, $1 \leq i \leq m$.

Example

The subset $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis of $\mathbb{R}^n$, called the **standard basis** of $\mathbb{R}^n$. (We've already seen that $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent and that $\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.)

Definition

The **dimension** of a subspace U of $\mathbb{R}^n$ is the number of vectors in any basis of U, and is denoted $\dim(U)$.

Example

In $\mathbb{R}^n$, we define the dimension of the subspace $\{\mathbf{0}_n\}$ to be 0. This says that $\{\mathbf{0}_n\}$ has a basis containing no vectors. This should make sense because as we showed before, the zero vector of $\mathbb{R}^n$ cannot belong to any independent set.

Example

Since $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis of $\mathbb{R}^n$, $\mathbb{R}^n$ has dimension n.

This is why the Cartesian plane, $\mathbb{R}^2$, is called 2-dimensional, and $\mathbb{R}^3$ is called 3-dimensional.

Example

Let

$$U = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; a - b = d - c \right\}.$$

Show that U is a subspace of $\mathbb{R}^4$, find a basis of U, and find $\dim(U)$.

Solution

Properties of Bases

Let U be a subspace of $\mathbb{R}^n$. Then

1. U has a basis, and $\dim(U) \leq n$.
2. Any independent set of U can be extended (by adding vectors) to a basis of U.
3. Any spanning set of U can be cut down (by deleting vectors) to a basis of U.

Example

Previously, we showed that

$$U = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; a - b = d - c \right\}$$

is a subspace of $\mathbb{R}^4$, and that $\dim(U) = 3$.

Also, it is easy to verify that

$$S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix} \right\},$$

is an independent subset of U.

S can be extended to a basis of U. To do so, find a vector in U that is **not** in $\text{span}(S)$.

$$\begin{bmatrix} 1 & 2 & ? \\ 1 & 3 & ? \\ 1 & 3 & ? \\ 1 & 2 & ? \end{bmatrix}$$

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

Therefore, S can be extended to the basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\} \text{ of } U.$$

Example

Let

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 0 \\ 3 \\ -1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 4 \\ 4 \\ 11 \\ -3 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 3 \\ -2 \\ 2 \\ -1 \end{bmatrix},$$

and let $U = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Find a basis of U that is a subset of $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$, and find $\dim(U)$.

Theorem

Let U be a subspace of $\mathbb{R}^n$ with $\dim(U) = m$, and let $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ be a subset of U. Then B is linearly independent if and only if B spans U.

Proof

($\Rightarrow$) Suppose B is independent. If $\text{span}(B) \neq U$, then extend B to a basis B' of U by adding appropriate vectors from U. Then B' is a basis of size more than $m = \dim(U)$, which is impossible. Therefore, $\text{span}(B) = U$, and hence B is a basis of U.

($\Leftarrow$) Conversely, suppose $\text{span}(B) = U$. If B is not independent, then cut B down to a basis B' of U by deleting appropriate vectors. But then B' is a basis of size less than $m = \dim(U)$, which is impossible. Therefore, B is independent, and hence B is a basis of U.

What is the significance of this result?

Answer

Let U be a subspace of $\mathbb{R}^n$ and suppose $B \subseteq U$.

- If B spans U and $|B| = \dim(U)$, then B is also independent, and hence B is a basis of U.
- If B is independent and $|B| = \dim(U)$, then B also spans U, and hence B is a basis of U.

Therefore if $|B| = \dim(U)$, it is sufficient to prove that B is either independent or spans U in order to prove it is a basis.

Discrete Mathematics II	Spring 2022

Section 5.3: Orthogonality

Lecturer: Mr. Hall	*Math 270B*

Definitions

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1. The **dot product** of $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots x_ny_n = \mathbf{x}^T\mathbf{y}.$$

   Note: $\mathbf{x} \cdot \mathbf{y}$ is a scalar, but we also treat it as a 1×1 matrix.

2. The **length** of $\mathbf{x}$, denoted $||\mathbf{x}||$ is

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 \cdots + x_n^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

3. $\mathbf{x}$ is called a **unit vector** if $||\mathbf{x}|| = 1$.

Properties

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ (the dot product is commutative)
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ (the dot product distributes over addition)
3. $(a\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (a\mathbf{y})$
4. $||\mathbf{x}||^2 = \mathbf{x} \cdot \mathbf{x}$.
5. $||\mathbf{x}|| \geq 0$ with equality if and only if $\mathbf{x} = \mathbf{0}_n$.
6. $||a\mathbf{x}|| = |a|\ ||\mathbf{x}||$.

Example

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$. Furthermore, suppose that there exists a vector $\mathbf{x} \in \mathbb{R}^n$ for which $\mathbf{x} \cdot \mathbf{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\mathbf{x} = \mathbf{0}_n$.

Proof

Write $\mathbf{x} = t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_k\mathbf{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k$ span $\mathbb{R}^n$).

Then

$$\begin{aligned} ||\mathbf{x}||^2 &= \mathbf{x} \cdot \mathbf{x} \\ &= \mathbf{x} \cdot (t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_k\mathbf{f}_k) \\ &= \mathbf{x} \cdot (t_1\mathbf{f}_1) + \mathbf{x} \cdot (t_2\mathbf{f}_2) + \cdots + \mathbf{x} \cdot (t_k\mathbf{f}_k) \\ &= t_1(\mathbf{x} \cdot \mathbf{f}_1) + t_2(\mathbf{x} \cdot \mathbf{f}_2) + \cdots + t_k(\mathbf{x} \cdot \mathbf{f}_k) \\ &= t_1(0) + t_2(0) + \cdots + t_k(0) = 0. \end{aligned}$$

Since $||\mathbf{x}||^2 = 0$, $||\mathbf{x}|| = 0$. Now, $||\mathbf{x}|| = 0$ if and only if $\mathbf{x} = \mathbf{0}_n$. Therefore, $\mathbf{x} = \mathbf{0}_n$.

Thus, the only vector orthogonal to every vector of a spanning set of $\mathbb{R}^n$ is the zero vector.

The Cauchy Inequality

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}||\ ||\mathbf{y}||$ with equality if and only if $\{\mathbf{x}, \mathbf{y}\}$ is linearly dependent.

Proof

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 \leq ||t\mathbf{x} + \mathbf{y}||^2 &= (t\mathbf{x} + \mathbf{y}) \cdot (t\mathbf{x} + \mathbf{y}) \\ &= t^2\mathbf{x} \cdot \mathbf{x} + 2t\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \\ &= t^2||\mathbf{x}||^2 + 2t(\mathbf{x} \cdot \mathbf{y}) + ||\mathbf{y}||^2. \end{aligned}$$

The quadratic $t^2||\mathbf{x}||^2 + 2t(\mathbf{x} \cdot \mathbf{y}) + ||\mathbf{y}||^2$ in t is always nonnegative, so it does not have distinct real roots. Thus, if we use the quadratic formula to solve for t, the discriminant must be nonpositive, i.e., $(2\mathbf{x} \cdot \mathbf{y})^2 - 4||\mathbf{x}||^2||\mathbf{y}||^2 \leq 0$. Therefore, $(2\mathbf{x} \cdot \mathbf{y})^2 \leq 4||\mathbf{x}||^2||\mathbf{y}||^2$ Since both sides of the inequality are nonnegative, we can take (positive) square roots of both sides:

$$|2\mathbf{x} \cdot \mathbf{y}| \leq 2||\mathbf{x}||\ ||\mathbf{y}||$$

Therefore, $|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}||\ ||\mathbf{y}||$. What remains is to show that $|\mathbf{x} \cdot \mathbf{y}| = ||\mathbf{x}||\ ||\mathbf{y}||$ if and only if $\{\mathbf{x}, \mathbf{y}\}$ is linearly dependent.

First suppose that $\{\mathbf{x}, \mathbf{y}\}$ is dependent. Then by symmetry (of $\mathbf{x}$ and $\mathbf{y}$), $\mathbf{x} = k\mathbf{y}$ for some $k \in \mathbb{R}$. Hence

$$|\mathbf{x} \cdot \mathbf{y}| = |(k\mathbf{y}) \cdot \mathbf{y}| = |k|\ |\mathbf{y} \cdot \mathbf{y}| = |k|\ ||\mathbf{y}||^2, \text{ and } ||\mathbf{x}||\ ||\mathbf{y}|| = ||k\mathbf{y}||\ ||\mathbf{y}|| = |k|\ ||\mathbf{y}||^2,$$

so $|\mathbf{x} \cdot \mathbf{y}| = ||\mathbf{x}||\ ||\mathbf{y}||$.

Conversely, suppose $\{\mathbf{x}, \mathbf{y}\}$ is independent; then $t\mathbf{x} + \mathbf{y} \neq \mathbf{0}_n$ for all $t \in \mathbb{R}$, so $||t\mathbf{x} + \mathbf{y}||^2 > 0$ for all $t \in \mathbb{R}$. Thus the quadratic

$$t^2||\mathbf{x}||^2 + 2t(\mathbf{x} \cdot \mathbf{y}) + ||\mathbf{y}||^2 > 0$$

so has no real roots. It follows that the the discriminant is negative, i.e.,

$$(2\mathbf{x} \cdot \mathbf{y})^2 - 4||\mathbf{x}||^2||\mathbf{y}||^2 < 0.$$

Therefore, $(2\mathbf{x}\cdot\mathbf{y})^2 < 4||\mathbf{x}||^2||\mathbf{y}||^2$; taking square roots of both sides (they are both nonnegative) and dividing by two gives us

$$|\mathbf{x}\cdot\mathbf{y}| < ||\mathbf{x}||\ ||\mathbf{y}||,$$

showing that equality is impossible.

The Triangle Inequality

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$.

Proof

$$\begin{aligned}
||\mathbf{x}+\mathbf{y}||^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) \\
&= \mathbf{x}\cdot\mathbf{x} + 2\mathbf{x}\cdot\mathbf{y} + \mathbf{y}\cdot\mathbf{y} \\
&= ||\mathbf{x}||^2 + 2\mathbf{x}\cdot\mathbf{y} + ||\mathbf{y}||^2 \\
&\leq ||\mathbf{x}||^2 + 2||\mathbf{x}||\ ||\mathbf{y}|| + ||\mathbf{y}||^2 \text{ by the Cauchy Inequality} \\
&= (||\mathbf{x}|| + ||\mathbf{y}||)^2.
\end{aligned}$$

Since both sides of the inequality are nonnegative, we take (positive) square roots of both sides:

$$||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||.$$

Definition

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **distance** between $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||.$$

Properties of the distance function

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Then

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$.
2. $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.
4. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$.

Orthogonality

- Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We say that $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$.
- More generally, $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ is an **orthogonal set** if each $\mathbf{x_i}$ is nonzero, and every pair of distinct vectors of X is orthogonal, i.e., $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq k$.
- A set $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ is an **orthonormal** set if X is an orthogonal set of unit vectors, i.e., $||\mathbf{x_i}|| = 1$ for all i, $1 \leq i \leq k$.

Examples

1. The standard basis of $\mathbb{R}^n$ is an orthonormal set (and hence an orthogonal set).
2. $$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$
   is an orthogonal (but not orthonormal) subset of $\mathbb{R}^4$.
3. If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$ and $p \neq 0$, then $\{p\mathbf{x}_1, p\mathbf{x}_2, \ldots, p\mathbf{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$.
4. $$\left\{ \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$
   is an orthonormal subset of $\mathbb{R}^4$.

Definition

Normalizing an orthogonal set is the process of turning an orthogonal (but not orthonormal) set into an orthonormal set. If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$, then

$$\left\{ \frac{1}{||\mathbf{x}_1||}\mathbf{x_1}, \frac{1}{||\mathbf{x}_2||}\mathbf{x_2}, \ldots, \frac{1}{||\mathbf{x}_k||}\mathbf{x_k} \right\}$$

is an orthonormal set.

Example

Verify that

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} \right\}$$

is an orthogonal set, and normalize this set.

Solution

Pythagoras' Theorem

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k||^2 = ||\mathbf{x}_1||^2 + ||\mathbf{x}_2||^2 + \cdots + ||\mathbf{x}_k||^2.$$

Proof

Start with

$$\begin{aligned} ||\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k||^2 &= (\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k) \cdot (\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k) \\ &= (\mathbf{x}_1 \cdot \mathbf{x}_1 + \mathbf{x}_1 \cdot \mathbf{x}_2 + \cdots + \mathbf{x}_1 \cdot \mathbf{x}_k) \\ &\quad +(\mathbf{x}_2 \cdot \mathbf{x}_1 + \mathbf{x}_2 \cdot \mathbf{x}_2 + \cdots + \mathbf{x}_2 \cdot \mathbf{x}_k) \\ &\quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ &\quad +(\mathbf{x}_k \cdot \mathbf{x}_1 + \mathbf{x}_k \cdot \mathbf{x}_2 + \cdots + \mathbf{x}_k \cdot \mathbf{x}_k) \\ &= \mathbf{x}_1 \cdot \mathbf{x}_1 + \mathbf{x}_2 \cdot \mathbf{x}_2 + \cdots + \mathbf{x}_k \cdot \mathbf{x}_k \\ &= ||\mathbf{x}_1||^2 + ||\mathbf{x}_2||^2 + \cdots + ||\mathbf{x}_k||^2. \end{aligned}$$

The second to last equality follows from the fact that the set is orthogonal, so for all i and j, $i \neq j$ and $1 \leq i, j \leq k$, $\mathbf{x}_i \cdot \mathbf{x}_j = 0$. Thus, the only nonzero terms are those of the form $\mathbf{x}_i \cdot \mathbf{x}_i$, $1 \leq i \leq k$.

Theorem

Every orthogonal set in $\mathbb{R}^n$ is linearly independent.

Proof

Let $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$ be an orthogonal set in $\mathbb{R}^n$ and suppose a linear combination vanishes, say: $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}$.

Then $0 = \mathbf{x}_1 \cdot \mathbf{0} = \mathbf{x}_1 \cdot (t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k) = t_1(\mathbf{x}_1 \cdot \mathbf{x}_1) + t_2(\mathbf{x}_1 \cdot \mathbf{x}_2) + \cdots + t_k(\mathbf{x}_1 \cdot \mathbf{x}_k) = t_1||\mathbf{x}_1||^2 + t_2(0) + \cdots + t_k(0) = t_1||\mathbf{x}_1||^2$

Since $||\mathbf{x}_1||^2 \neq 0$, this impliles that $t_1 = 0$. Similarly $t_i = 0$ for each $1 \leq i \leq k$. Therefore, $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$ is linearly independent.

The Fourier Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal basis of a subspace U of $\mathbb{R}^n$. Then for any $\mathbf{x} \in U$,

$$\mathbf{x} = \left(\frac{\mathbf{x}\cdot\mathbf{f}_1}{||\mathbf{f}_1||^2}\right)\mathbf{f}_1 + \left(\frac{\mathbf{x}\cdot\mathbf{f}_2}{||\mathbf{f}_2||^2}\right)\mathbf{f}_2 + \cdots + \left(\frac{\mathbf{x}\cdot\mathbf{f}_m}{||\mathbf{f}_m||^2}\right)\mathbf{f}_m.$$

This expression is called the **Fourier expansion** of $\mathbf{x}$, and

$$\frac{\mathbf{x}\cdot\mathbf{f}_j}{||\mathbf{f}_j||^2},$$

$j = 1, 2, \ldots, m$ are the **Fourier coefficients.**

Proof

Let $\mathbf{x} \in U$. Since $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is a basis of U, $\mathbf{x} = t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_m\mathbf{f}_m$ for some $t_1, t_2, \ldots, t_m \in \mathbb{R}$.

Notice that for any i, $1 \leq i \leq m$,

$$\begin{array}{rcl} \mathbf{x}\cdot\mathbf{f}_i & = & (t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_m\mathbf{f}_m)\cdot\mathbf{f}_i \\ & = & t_i\mathbf{f}_i\cdot\mathbf{f}_i \quad \text{since } \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\} \text{ is orthogonal} \\ & = & t_i||\mathbf{f}_i||^2. \end{array}$$

Since $\mathbf{f}_i$ is nonzero, we obtain

$$t_i = \frac{\mathbf{x}\cdot\mathbf{f}_i}{||\mathbf{f}_i||^2}.$$

The result now follows.

Note

If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthonormal basis, then the Fourier coefficients are simply $t_i = \mathbf{x}\cdot\mathbf{f}_i$, $i = 1, 2, \ldots, m$.

Example

Let $\mathbf{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\mathbf{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Express $\mathbf{x}$ as a linear combination of $\mathbf{f}_1, \mathbf{f}_2,$ and $\mathbf{f}_3$.

Solution

Example

Let

$$\mathbf{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{f}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \mathbf{f}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Show that $B = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is an orthogonal basis of $\mathbb{R}^4$, and express $\mathbf{x} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T$ as a linear combination of $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ and $\mathbf{f}_4$.

Solution

Example

Show that B is an orthogonal basis of $\mathbb{R}^3$ and use the Fourier Expansion Theorem to expand $\mathbf{x} = (a, b, c)$ as a linear combination of the basis vectors. $B = \{(1, -1, 3), (-2, 1, 1), (4, 7, 1)\}$

Solution

Practice

Show that B is an orthogonal basis of $\mathbb{R}^3$ and use the Fourier Expansion Theorem to expand $\mathbf{x} = (a, b, c)$ as a linear combination of the basis vectors. $B = \{(1, 2, 3), (-1, -1, 1), (5, -4, 1)\}$

Solution

Example

Find all $(a, b, c, d) \in \mathbb{R}^4$ such that the given set is orthogonal. $\{(1, 2, 1, 0), (1, -1, 1, 3), (2, -1, 0, -1), (a, b, c, d)\}$

Solution

Discrete Mathematics II	Spring 2022
Section 5.4: Rank of a Matrix	
Lecturer: Mr. Hall	*Math 270B*

Definitions

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

- The **column space of** A, denoted $\mathrm{col}(A)$ is the subspace of $\mathbb{R}^m$ spanned by the columns of A.
- The **row space of** A, denoted $\mathrm{row}(A)$ is the subspace of $\mathbb{R}^n$ spanned by the rows of A (or the columns of A^T).

Note

From the work done in 5.1, we see that $\mathrm{col}(A) = \mathrm{im}(A)$.

Lemma

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

1. If $A \to B$ by elementary row operations, then $\mathrm{row}(A) = \mathrm{row}(B)$.
2. If $A \to B$ by elementary column operations, then $\mathrm{col}(A) = \mathrm{col}(B)$.

Proof

It suffices to prove only part one (the proof of part two is analogous), and when $A \to B$ for a single row operation.

Thus let $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ denote the rows of A.

- If B is obtained from A by interchanging two rows of A, then A and B have exactly the same rows, so $\mathrm{row}(B) = \mathrm{row}(A)$.

- Suppose $p \neq 0$, and suppose that for some j, $1 \leq j \leq m$, B is obtained from A by multiplying row j by p. Then
$$\mathrm{row}(B) = \mathrm{span}\{\mathbf{r}_1, \ldots, p\mathbf{r}_j, \ldots, \mathbf{r}_m\}.$$
Since
$$\{\mathbf{r}_1, \ldots, p\mathbf{r}_j, \ldots, \mathbf{r}_m\} \subseteq \mathrm{row}(A),$$
it follows that $\mathrm{row}(B) \subseteq \mathrm{row}(A)$.

  Conversely, since
$$\{\mathbf{r}_1, \ldots, \mathbf{r}_m\} \subseteq \mathrm{row}(B),$$
it follows that $\mathrm{row}(A) \subseteq \mathrm{row}(B)$. Therefore, $\mathrm{row}(B) = \mathrm{row}(A)$.

- Suppose $p \neq 0$, and suppose that for some i and j, $1 \leq i, j \leq m$, B is obtained from A by adding p times row j to row i. Without loss of generality, we may assume $i < j$.

  Then
  $$\text{row}(B) = \text{span}\{\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i + p\mathbf{r}_j, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_m\}.$$
  Since
  $$\{\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i + p\mathbf{r}_j, \ldots, \mathbf{r}_m\} \subseteq \text{row}(A),$$
  it follows that $\text{row}(B) \subseteq \text{row}(A)$.

  Conversely, since
  $$\{\mathbf{r}_1, \ldots, \mathbf{r}_m\} \subseteq \text{row}(B),$$
  it follows that $\text{row}(A) \subseteq \text{row}(B)$. Therefore, $\text{row}(B) = \text{row}(A)$.

Corollary

Let A be an $m \times n$ matrix, U an invertible $m \times m$ matrix, and V an invertible $n \times n$ matrix. Then $\text{row}(UA) = \text{row}(A)$ and $\text{col}(AV) = \text{col}(A)$.

Proof

Since U is invertible, U is a product of elementary matrices, implying that $A \to UA$ by a sequence of elementary row operations. By Lemma 2, $\text{row}(UA) = \text{row}(A)$.

Now consider AV: $\text{col}(AV) = \text{row}((AV)^T) = \text{row}(V^T A^T)$ and V^T is invertible (a matrix is invertible if and only if its transpose is invertible). It follows from the first part of this Corollary that
$$\text{row}(V^T A^T) = \text{row}(A^T).$$
But $\text{row}(A^T) = \text{col}(A)$, and therefore $\text{col}(AV) = \text{col}(A)$.

Lemma

If R is a row-echelon matrix then

1. the nonzero rows of R are a basis of $\text{row}(R)$;
2. the columns of R containing the leading ones are a basis of $\text{col}(R)$.

Example

Let
$$R = \begin{bmatrix} 1 & 2 & 2 & -2 & 0 & 0 \\ 0 & 1 & 3 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

1. Since the nonzero rows of R are linearly independent, they form a basis of $\text{row}(R)$.
2. Let $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\} \subseteq \mathbb{R}^5$. Then B is linearly independent and spans $\text{col}(R)$, and thus is a basis of $\text{col}(R)$. This tells us that $\dim(\text{col}(R)) = 4$. Now let X denote the set of columns of R that contain the leading ones.

   Then X is a linearly independent subset of $\text{col}(R)$ with $4 = \dim(\text{col}(R))$ vectors. It follows that X spans $\text{col}(R)$, and therefore is a basis of $\text{col}(R)$.

Example

Find a basis of $U = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 5 \\ 7 \end{bmatrix} \right\}$ and find $\dim(U)$.

Solution

Definition

For any matrix A, the **rank of** A is defined as rank $(A) = \dim(\text{row}(A))$.

The Rank Theorem

Let $A = \left[\begin{array}{cccc} \mathbf{c_1} & \mathbf{c_2} & \cdots & \mathbf{c_n} \end{array}\right]$ be an $m \times n$ matrix with columns $\{\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_n}\}$, and suppose that rank $(A) = r$. Then

$$\dim(\text{row}(A)) = \dim(\text{col}(A)) = r.$$

Furthermore, if R is a row-echelon form of A then

1. the r nonzero rows of R are a basis of row(A);
2. if $S = \{\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \ldots, \mathbf{c}_{j_r}\}$ are the r columns of A corresponding to the columns of R containing leading ones, then S is a basis of col(A).

Example

For the following matrix A, find rank (A) and find bases for row(A) and col(A).

$$A = \left[\begin{array}{rrrr} 2 & -4 & 6 & 8 \\ 2 & -1 & 3 & 2 \\ 4 & -5 & 9 & 10 \\ 0 & -1 & 1 & 2 \end{array}\right].$$

Solution

Example (revisited)

Find a basis of $U = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 5 \\ 7 \end{bmatrix} \right\}$ and find $\dim(U)$.

Solution

Example

For the following matrix A, find bases for $\text{null}(A)$ and $\text{im}(A)$, and find their dimensions.

$$A = \begin{bmatrix} 2 & -4 & 6 & 8 \\ 2 & -1 & 3 & 2 \\ 4 & -5 & 9 & 10 \\ 0 & -1 & 1 & 2 \end{bmatrix}.$$

Solution

Example

Can a 5×6 matrix have independent columns? independent rows? Justify your answer.

Solution

Discrete Mathematics II | Spring 2022

Section 5.5: Similarity and Diagonalization

Lecturer: Mr. Hall | *Math 270B*

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

Definition

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

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

Properties of the trace

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

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

Reminder (The Characteristic Polynomial)

For any $n \times n$ matrix A, the characteristic polynomial of A is

$$c_A(\lambda) = \det(\lambda I - A),$$

and is a polynomial of degree n.

Properties of Similar Matrices

If A and B are $n \times n$ matrices and $A \sim B$, then

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

Proof (continued)

4.

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

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

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

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

Definition

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

Determining whether or not a square matrix A is diagonalizable can be done using eigenvalues and eigenvectors of the matrix A. If λ is an eigenvalue of A, then

$$A\mathbf{v} = \lambda\mathbf{v}$$

for some **nonzero** vector $\mathbf{v}$ in $\mathbb{R}^n$. Such a vector $\mathbf{v}$ 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 $\mathbb{R}^n$ has a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of eigenvectors of A.
2. If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are eigenvectors of A and form a basis of $\mathbb{R}^n$, then

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$$

is an invertible matrix such that

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

where λ_i is the eigenvalue of A corresponding to $\mathbf{v}_i$.

Example

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

Solution

Definition

Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. The **eigenspace of A corresponding to λ** is the set

$$E_\lambda(A) = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

Note

$$\{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\} = \{\mathbf{v} \in \mathbb{R}^n \mid (\lambda I - A)\mathbf{v} = \mathbf{0}_n\} = \text{null}(\lambda I - A),$$

showing that $E_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

Lemma

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

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

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.

Diagonalizable Matrices

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.

Example

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

Example

Let

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

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